MODELLING THE INTERPLAY BETWEEN HUMAN BEHAVIOUR AND THE SPREAD OF INFECTIOUS DISEASES: FROM TOY MODELS TO QUANTITATIVE APPROACHES

Benjamin Franz Josef Steinegger

Benjamin Steinegger

Modelling the interplay between human behaviour and the spread of infectious diseases

From toy models to quantitative approaches

Ph.D. Thesis
Supervised by Dr. Alex Arenas
Co-Supervised by Dr. Clara Granell

Departament d'Enginyeria Informàtica i Matemàtiques de la Seguretat

UNIVERSITAT
ROVIRA i VIRGILI

Universitat Rovira i Virgili
Tarragona, Catalonia, Spain

December 2022

Author:
Benjamin Steinegger

Supervisor & Co-Supervisor:
Dr. Alexandre Arenas Moreno & Dr. Clara Granell

Universitat Rovira i Virgili

Cover art by: Dall·E 2
Printed by: Arola Editors

UNIVERSITAT
ROVIRA i VIRGILI

We STATE that the present study, entitled *Modelling the interplay between human behaviour and the spread of infectious diseases*, presented by Benjamin Steinegger for the award of the degree of Doctor, has been carried out under our supervision at the Departament d'Enginyeria Informàtica i Matemàtiques de la Seguretat of this university.

Tarragona, 15th December, 2022

Doctoral Thesis Supervisor,

Dr. Alexandre Arenas Moreno

Doctoral Thesis Co-Supervisor,

Dr. Clara Granell

UNIVERSITAT ROVIRA I VIRGILI
MODELLING THE INTERPLAY BETWEEN HUMAN BEHAVIOUR AND THE SPREAD OF INFECTIOUS DISEASES: FROM TOY MODELS TO QUANTITATIVE APPROACHES
Benjamin Franz Josef Steinegger

To all my loved ones

UNIVERSITAT ROVIRA I VIRGILI
MODELLING THE INTERPLAY BETWEEN HUMAN BEHAVIOUR AND THE SPREAD OF INFECTIOUS DISEASES: FROM TOY MODELS TO QUANTITATIVE APPROACHES
Benjamin Franz Josef Steinegger

ACKNOWLEDGEMENTS

This thesis was completed with the help and support from a series of individuals that I deeply appreciate. I would like to use the opportunity here to express them my gratitude.

First of all, I would like to thank my supervisor Dr. Alex Arenas for the support and the freedom he gave me during these four years. It was a privilege to complete the dissertation learning from you and knowing that you would always have my back if necessary. Together with my Co-Supervisor Dr. Clara Granell, we could share many great experiences on our numerous trips to conferences in different corners of the world. And beyond research, you introduced me to the joys of gastronomy, which will stick with me for the rest of my life.

Thank you Lluís for your friendship! You welcomed me with open arms, and introduced me to many people which made my period in Tarragona so enjoyable. I will be forever in your debt. Keep alive your enthusiasm and you will be able to do anything.

I want also to thank all the other former and current members of the Alephsys lab. I especially thank Alessio, which, besides being a good friend, introduced me to the world of complex systems and supervised my first steps in the world of science. Equally, I want to thank Eugenio for the time we spent together in Tarragona and his assistance in navigating the scientific community. Thank you Joan for helping me to become a semi-decent coder. Thank you Giulio and Giacomo for the great company, the scientific collaboration and the many laughs we had throughout our time together. Thank you Sergio for the always precise handling of any administrative issues. Thank you Albert for always cheering up our lunch time. And finally also thanks to Mattia, Piergiorgio and Aleix for our short but great time together.

Besides the members of the Alephsys lab, I had the pleasure to meet too many other great scientists and individuals to mention them all. But I want like to thank Jesús for proposing the project with which this thesis started. I want to thank David for the great collaborations and the friendship. Finally, I would also thank the entire team of the Complexity72h workshop. It is

incredible how we managed to wrap up somehow a project in the midst of a global pandemic.

During my time in Tarragona, I was lucky enough to also become friends with people outside of academia. All of you made my time in Tarragona fly by, and allowed me to switch off from work from time to time. I would like to especially thank Adria, David and Enric for the friendship, the time we spent together, and the deep dive in local culture you provided me with.

This thesis could have never come to life without the skill set I developed during my undergraduate studies. For this reason, I want to thank all the people who made my time in Lausanne on the other side of the "Röstigraben" instructive and amusing in so many different ways. In particular, I owe my gratitude to Loïc, Vincent, Alexandre, Thomas, Riad, Marko, Ivan, and Noé. With tears in my eyes, I realise that your demons, Noé, prevented us from celebrating our achievements together. I hope that they do not haunt you anymore and you could find some peace. I am sorry for not having picked up the phone to call you when I should have.

I would also like to thank my friends and family in Switzerland that supported me along this years. I especially thank my brother Matthias, my godfather Christoph, my godmother Derja and my sponsor Peter. Thank you Jonas for your legendary visits in Tarragona. It was always relaxing and encouraging to spend time with all of you. I truly appreciated to be able to disconnect and talk about anything but research. Obviously, I would also like to thank my parents for the unconditional support all along my education. The decision to do the PhD in Tarragona may not have been easy to comprehend for them, but they nonetheless encouraged me in my choices. I hope that this thesis may convince them that I was actually working from time to time during my time in Spain.

Finally, I want to thank Selina. I am deeply grateful for your support and generosity during all these years. It never felt as if you were far away, even though we were living a thousand kilometres apart, in different countries. Your cheerful attitude was, and still is to this day, always able to put a smile on my face. I am looking forward to now build our lives together with Zurich as a home base. But who knows where life will take us in the future.

PUBLICATIONS

This thesis includes work by the author that has been published or submitted for publication. These publications are the own work of the author of this thesis, and the author has the permission of the publishers to reproduce the contents of these publications for academic purposes. In particular, some data, ideas, opinions and figures presented in this thesis have previously appeared or may appear shortly after the submission of this thesis as follows:

- "Pulsating campaigns of human prophylaxis driven by risk perception palliate oscillations of direct contact transmitted diseases", B. Steinegger, A. Arenas, J. Gómez-Gardeñes and C. Granell, Phys. Rev. Res. **2**, 023181 (2020). (Section 3.1)

- "Behavioural response to heterogeneous severity of COVID-19 explains temporal variation of cases among different age groups", B. Steinegger, L. Arola-Fernández, C. Granell, J. Gómez-Gardeñes and A. Arenas, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences **380**, 20210119 (2022). (Section 3.2)

- "Homophily in the adoption of digital proximity tracing apps shapes the evolution of epidemics", G. Burgio, B. Steinegger, G. Rapisardi and A. Arenas, Phys. Rev. Res. **3**, 033128 (2021). (Section 4.1.1)

- "Homophily impacts the success of vaccine roll-outs", G. Burgio, B. Steinegger and A. Arenas, Commun Phys **5**, 70 (2022). (Section 4.1.2)

- "Non-selective distribution of infectious disease prevention may outperform risk-based targeting", B. Steinegger, I. Iacopini, A.S. Teixeira, A. Bracci, P. Casanova-Ferrer, A. Antonioni and E. Valdano, Nat Commun **13**, 3028 (2022). (Section 4.2)

- "Modeling the spatiotemporal epidemic spreading of COVID-19 and the impact of mobility and social distancing interventions", A. Arenas, W. Cota, J. Gómez-Gardeñes, S. Gómez, C. Granell, J.T. Matamalas, D. Soriano-Panos and B. Steinegger, Phys. Rev. X **10**, 041055 (2021). (Section 5.1)

- "Retrospective study of the first wave of COVID-19 in Spain", B. Steinegger, C. Granell, G. Rapisardi, S. Gómez, J.T. Matamalas, D. Soriano Paños, J. Gómez-Gardeñes and Alex Arenas, in preparation (2022). (Section 5.2)
- "The Spread of the Delta Variant in Catalonia During Summer 202", B. Steinegger, G. Burgio, G. Rapisardi, R. Guimerà and Alex Arenas, in preparation (2022). (Section 5.3)

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Preventing the spread of infectious diseases is one of the greatest challenges of humanity's past, present, and foreseeable future. Many infectious diseases are transmitted upon contact, and hence the complex web of human interactions acts as a substrate for their propagation. For this reason, epidemic models always comprise, either explicitly or implicitly, a description of how humans interact. Research in mathematical epidemiology has opened many pathways to include some of the diverse aspects of human behaviour into epidemic models. However, the quest for a general theory of the interplay between human behaviour and the spread of pathogens is far from complete. Moreover, in this quest, a pronounced gap exists between quantitative models and more qualitative approaches that focus on the phenomenology.

The aim of this thesis is to contribute to the mathematical description of human behaviour in the context of infectious diseases, working with both quantitative and qualitative models. The first chapter develops two qualitative toy models to outline how dynamical risk-based prophylaxis can sustain epidemic cycles. In the second chapter, we consider specific static aspects of human behaviour – homophily and heterogeneous contact patterns – and analyse their implications on epidemic control. In contrast to previous belief, we show that homophily in the adoption of many prophylactic tools can be either beneficial, detrimental or non trivially affect their efficacy. Furthermore, we question the current paradigm of risk-based immunisation strategies and show that targeting hubs is only optimal for protection with high efficacy. The last chapter of this thesis focuses on quantitative approaches to model the spread of SARS-CoV-2, in particular, the first wave and the spread of the Delta variant. Besides the methodological advances, the results add evidence of how voluntary behavioural adaption shaped the course of the epidemic beyond the introduction of non-pharmaceutical interventions.

Overall, this thesis unveils new phenomenology, adds additional empirical evidence, and provides new tools to analyse how human behaviour and epidemics coevolve. The flexible blend of quantitative and qualitative approaches may also provide a pathway to analyse and interpret the vast amount of data currently collected during the SARS-CoV-2 pandemic.

UNIVERSITAT ROVIRA I VIRGILI
MODELLING THE INTERPLAY BETWEEN HUMAN BEHAVIOUR AND THE SPREAD OF INFECTIOUS DISEASES: FROM TOY MODELS TO QUANTITATIVE APPROACHES
Benjamin Franz Josef Steinegger

1 INTRODUCTION

"A new pandemic is one of the biggest threats for today's interconnected, globalised human society" was a sentence with which the author of this thesis started an unsuccessful grant proposal in late 2019. Probably any researcher working in the modelling of infectious diseases wrote similar sentences in their grant proposals or the introduction of their articles. While many probably truly believed this statement, it was also, to some extent, a platitude that was used, at least partially, to motivate the research in epidemic modelling. Unfortunately, shortly after, in early 2020, the platitude became a reality, and we saw SARS-CoV-2 spread around the globe. Entire countries were shut down, and in particular western society came to a halt. SARS-CoV-2 seems like a singular event to many of us, and it may be unique with respect to the velocity the virus has spread to almost every part of the world. However, human history is intrinsically linked with the spread of infectious diseases. Smallpox ravaged human society for thousands of years before being eradicated by vaccine mandates in the 20th century. The bubonic plague, also called the Black Death, killed over 20 million people in the middle age. Also, today, HIV and Tuberculosis are responsible for around a million deaths every year [1], causing indescribable destruction in mainly African countries.

Considering the immense suffering infectious diseases have caused historically, it seems only natural that humanity invested considerable efforts to develop the necessary knowledge and tools to prevent them from spreading. More generally, one can understand large parts of the technological advances as a quest of humanity to become independent from the whims of nature. No painting better illustrates how death was omnipresent in people's lives in the past than the Triumph of Death by Bruegel, which is shown in Fig. 1.1. The associated despair led people to find ways to prevent famines, floods, or earthquakes from causing total destruction. In many cases, science was at the core of these technological advances, but some were more immediate

Figure 1.1: The Triumph of Death by Pieter Bruegel, the Elder painted in the 16th-century [2]. The original painting is in the Museo del Prado in Madrid, Spain.

than others. Human creativity can move mountains, given that the obstacle is clear and visible. However, visibility was precisely the problem for infectious diseases since bacteria and viruses are not tangible. More specifically, it took until the late 19th century, and a fierce Franco-German rivalry fueled by the war of 1870, for Louis Pasteur and Robert Koch to identify that the transmission of bacteria caused anthrax and tuberculosis disease, respectively [3]. Shortly after, in 1892, Dmitri Iossifowitsch Iwanowski identified that the pathogen for tobacco disease was a virus [4]. These breakthroughs sparked the rapid development of different vaccines, which allowed the containment of many infectious diseases during the 20th-century [5].

The understanding that human contact is at the origin of the spread of many pathogens paved the way to develop mathematical models which describe the mechanics of human-to-human transmission behind the propagation of infectious diseases. Already one of the first mathematical models, developed in the 1920s by William Ogilvy Kermack and Anderson Gray McKendrick, allowed for an accurate description of an actual plague outbreak. In analogy to meteorology, the availability of a mathematical description fed the hopes to not only be able to reproduce but also to predict the evolution of epidemics. However, contrary to weather forecasts, not the chaotic nature of the underlying equations made the prediction difficult, but instead the

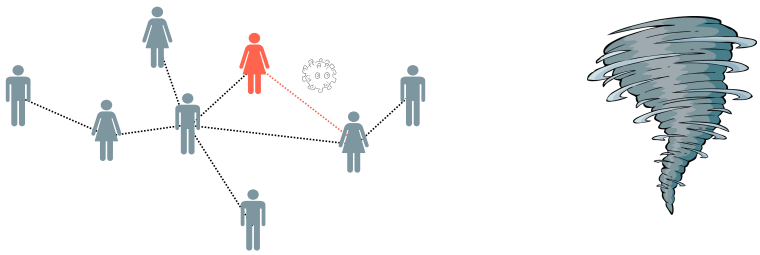


Figure 1.2: Schematic human-to-human disease transmission compared to a meteorological phenomenon for which no human-made obstacles exist.

fact that many diseases are transmitted through human contact. Topologically speaking, the disease propagates on top of a physical contact network. Accordingly, if humans adapt their behaviour, for example, they reduce their social contacts facing the danger of an expanding epidemic, the evolution of the epidemic will be fundamentally different. In contrast, for meteorological phenomena such as a Tornado, nothing humanly possible can be done to alter its course since it freely moves through physical space instead on top of a human-made topology.

The behavioural impact on the epidemic evolution leads to the possibility of disease forecasting becoming a self-defeating prophecy. If models suggest a devastating impact of an epidemic in the near future, individuals may act more cautiously, or authorities may implement measures to limit the epidemic's impact. In both cases, the prediction is self-defeating because human behaviour is intrinsically linked to the epidemic. This intrinsic link is apparent when one considers the reproduction number R, i.e. the number of infections made on average by an infected individual. The disease can spread for $R > 1$, while for $R < 1$, the disease slowly dies out. Given a contact rate k, an infectious period τ_I, a transmission probability per contact λ, and the fraction of susceptible individuals in the population S, R reads

$$R = k\tau_I\lambda S\,. \tag{1.1}$$

In the above expression, none of the quantities is merely biological. Human behaviour can contribute to the reduction of any of them in order to contain the disease. To name an example for each quantity, individuals can practise social distancing to decrease k, quickly self-isolate after the presence of symp-

toms to reduce τ_I, wear face masks to reduce λ or get vaccinated to decrease susceptibility, S, to the pathogen in the population. The strong dependence of R on quantities that are shaped by human behaviour indicates how the latter can fundamentally alter the evolution of an epidemic.

Beyond the immediate behavioural reaction to a disease, one can also see the evolution of human society and infectious diseases as one of coevolution [6]. Disease prophylaxis has been woven into the cultural fabric in many societies before the discovery of viruses and bacteria. For example, nomad tribespeople in the region of Manchuria, today in between China, Russia and Mongolia, incorporated reasonable, sound epidemiological rules into their religion to prevent bubonic infection from marmots [7]. Suppose a close marmot colony showed signs of sickness; the tribe stroke its tents and moved away to avoid bad luck. In contrast, as Chinese settlers moved into the territory in the early 20th century, they discarded the local customs, and a series of plague outbreaks were recorded. The cultural evolution of the nomad tribespeople cemented a successful behavioural adaption, which prevented disease propagation. Similarly, on an even more significant time scale, the biological evolution of pathogens and the human species are intertwined. Without viruses and bacteria, there would be no need for the human immune system to exist. Furthermore, viruses and bacteria constantly adapted themselves to find a loophole in the immune system [8]. The most recent and prominent example is the Omicron variant of SARS-CoV-2 that evades preexisting immunity from previous infection or vaccination through mutation [9]. However, behavioural adaption may be not only cultural but also partially biological. For example, different animals, such as bats [10] or mandrills [11], show social distancing. For these animals, the origin of the prophylactic behaviour is unlikely to be cultural and almost surely biological. Eventually, the interdependence between infectious diseases and humanity results in different time scales in their simultaneous evolution.

Apart from the temporal scales, different spatial scales exist along which infectious diseases can spread [12]. Locally, diseases spread through a tight-knit human contact network. On a larger scale, mobility introduces diseases to different regions, countries or continents. The technological advances of the last centuries have substantially reduced the time for a pathogen to spread globally. Previous to the existence of cars, trains and aeroplanes, it took multiple years for a pathogen to spread across countries, typically along trade routes [13]. Further, there is not only a dynamic between hosts (human-to-human transmission) but also within, at an even smaller spatial and temporal scale.

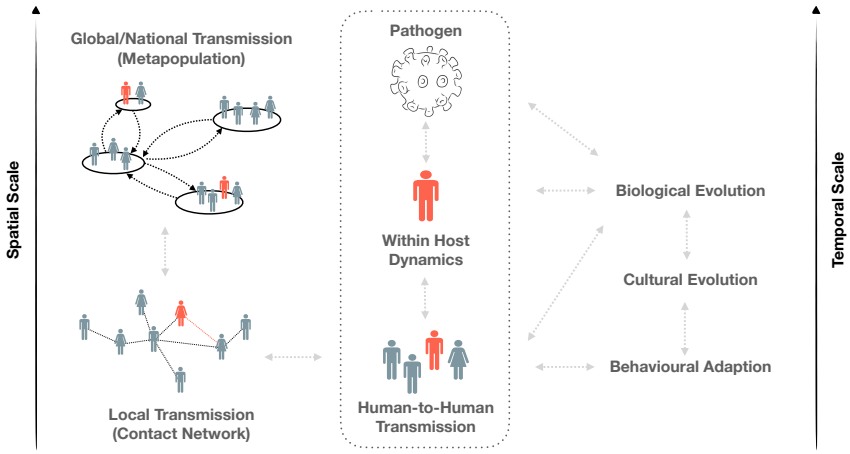


Figure 1.3: Schematic coevolution of infectious diseases and humanity, inspired by Ref. [14]. Arrows denote interdependence.

Pathogens replicate inside the host, battling the human immune system [14]. The within-host dynamics is responsible for the existence of an incubation period and time-dependent infectivity after infection.

The coevolution of pathogens and humanity is one of the most paradigmatic examples of a complex system due to the different temporal and spatial scales involved. However, it is not evident to give a precise definition of what a complex system is. It is probably easier to illustrate and define the term complex systems by its use [15]. In this sense, the term complex system is generally used to refer to systems in which many entities interact at a lower scale, while their interaction leads to an emergent phenomenology at higher spatiotemporal scales [16]. A typical example is ferromagnetic materials, where the alignment of the spins of the individual atoms leads to the emergent property of magnetism below a critical temperature at the macroscopic level. Similar emergent systems have been found in economics, sociology and biology [17]. In the context of infectious diseases, individual entities correspond to the hosts whose interactions enable the pathogen's spread and sometimes endemic state. The equivalent to the critical temperature is the condition $R > 1$ for the disease to spread.

To describe complex systems, the main challenge, which often constitutes a walk on a tightrope, is to find a balance between empirical accuracy and model complexity. For example, to model epidemic spreading, one needs to decide at which resolution level one wants to describe onward transmis-

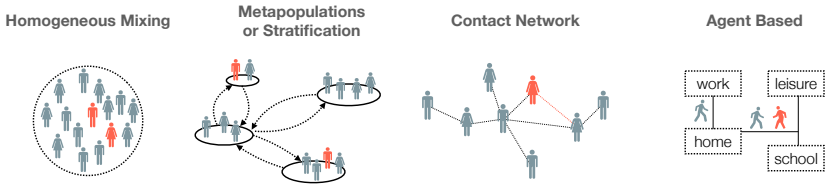


Figure 1.4: Different types of approximations to describe how humans interact.

sion. There are many approaches, from considering interactions taking place randomly, stratifying the population according to age or geography (metapopulations), using a real-world contact network, or making "The Sims"-like agent-based simulations [18]. The specific choice strongly depends on the model's purpose. For prediction or quantification, the primary goal is accuracy. In contrast, transparency is more critical if the model aims to advance understanding. Quantitative models, which are usually fitted to, or at least informed by, empirical data, do not only serve to make predictions. As previously mentioned, prediction is often very challenging due to the interplay with human behaviour, which frequently only allows making a scenario analysis. Accordingly, quantitative approaches are often much more helpful in inferring characteristics of the pathogen or human behaviour than in making actual predictions. Qualitative models that aim to advance understanding are often referred to as toy models and generally have two applications. First, toy models can consider an empirical finding and analyse its implications for the dynamics as part of a thought experiment. Second, one can also start with an empirical finding regarding the dynamics and leverage a toy model to analyse which ingredients are necessary and thus responsible for the empirically observed phenomenology. In both cases, transparency is vital to building causal relations between model ingredients and model outcomes inside the framework.

This thesis blends qualitative toy models and quantitative approaches to describe the interplay between the spread of directly transmitted infectious diseases and human behaviour. Such approaches to incorporate human behaviour more explicitly into epidemic models are manifold but also very scattered. There is no all-encompassing theory [19–22], as for complex systems in general. This thesis will also not succeed in providing this all-encompassing theory. However, we present advances on several fronts, hopefully contributing to constructing a unified framework in the future. The

blend of qualitative and quantitative models also aims to bridge the gap between the toy models primarily used in complex systems and the extensive parametrisation of more traditional quantitative models in mathematical epidemiology. Situating the content of the thesis within Fig. 1.3, we will mainly focus on the immediate behavioural response of individuals and fixed aspects of human behaviour which are not related to the presence of a disease. We will not describe the coevolution of human society and pathogens at the time scale of cultural or even biological evolution. Instead, regarding the epidemic dynamics, we will model onward transmission at different resolution levels, more precisely, all of which are highlighted in Fig. 1.4 except for agent-based modelling. While the toy models we develop, for which interpretability is critical, mainly rely on a mean-field description, the quantitative approaches utilise a data-informed stratification of the population.

The quantitative approaches presented in this thesis all focus on modelling the spread of SARS-CoV-2. Additionally, some of the more theoretically centred contributions were motivated by questions that arose during the SARS-CoV-2 pandemic. For this reason, SARS-CoV-2 appears in all chapters of the thesis. Nevertheless, the thesis should not be seen as one about SARS-CoV-2 since different projects had already been started before the pandemic outbreak. Furthermore, while the pandemic served as an excellent study case and raised many questions, the answers to these questions, particularly the more theoretical ones, often have a much broader scope than SARS-CoV-2.

The thesis is organised into six separate chapters. After this introduction, Chapter 2 provides the necessary theoretical background to follow this thesis. In addition to a short historical overview, the chapter also contextualises the contributions made by different fields to model epidemics. Subsequently, Chapters 3-5 will outline the contributions of this thesis in three thematic blocks.

The first part, Chapter 3, introduces a toy model in which risk-based individual prophylaxis leads to recurrent epidemic waves. It focuses on the dynamic interplay between human behaviour and the spread of infectious diseases. From a methodological point of view, we show that incorporating a behavioural dimension allows for the emergence of epidemic waves without needing memory in the dynamics. Initially, the purpose of the model was to explain trends in sexually transmitted infections. However, it turned out that the model may also provide a mechanistic explanation for reoccurring surges in SARS-CoV-2 infections, which are driven by a relaxation in social distancing and prophylactic practice in general. Subsequently, we extend the model

to incorporate the possibility of heterogeneity in the risk of a severe course of infection. This extension of the model shows how the age-dependent risk of severe illness after a SARS-CoV-2 infection can explain the temporal variation of the case distribution across time, observed in many countries.

After focusing on the behavioural adaption in the presence of infectious diseases, the second part of the thesis, Chapter 4, focuses on two static features of human behaviour – homophily and heterogeneous contact rates – and their effects on epidemic control. In contrast to the first part, the specific aspects of human behaviour under consideration are not a reaction to the presence of an infectious disease but also exist in its absence. The first feature we consider, homophily, refers to the tendency of humans to interact with others with similar characteristics. Contrary to previous belief, we show that homophily in adopting prophylactic tools (vaccines, contact tracing apps, face masks, etc.) is not always detrimental to epidemic control. More precisely, we identify three dynamical regimes where the population-level efficacy of the prophylactic tool can either decrease, increase, or vary non-monotonously with respect to the level of homophily. Afterwards, we revisit a paradigm in the distribution of prophylactic tools in a population with heterogeneous contact rates. Standard theory tells us that individuals with the highest contact rates should be given prophylactic tools to maximise their impact on prevalence. However, we show that, due to the increased risk of breakthrough infections, this result only holds if the efficacy of the prophylactic tool is sufficiently high with respect to the prevalence of the pathogen in the population. Otherwise, there is a specific contact rate that maximises the impact of the prophylactic tool, and even random distribution can outperform risk-based distribution in some cases.

Finally, the third part, Chapter 5, focuses on the quantitative approaches we have developed to model the spread of SARS-CoV-2 in Spain. Firstly, we outline a model that we used to describe the spatiotemporal spread of SARS-CoV-2 in Spain during the early phase of the virus in 2020. The model is relevant from a methodological perspective since it combines the description of the spatial spread of the virus through mobility with the modelling of containment measures. However, increasing data availability after the end of the first wave forced us to revisit previously drawn conclusions and analyse the happening of events during the first wave retrospectively. Through a blend of a model-based approach, direct analysis of epidemiological data, and the evaluation of large-scale mobility data, we highlight a substantial decrease in the reproduction number before the nationwide lockdown in Spain on March

15. Furthermore, counterfactual scenarios indicate that the early decrease in the reproduction number reduced the number of fatalities and infections during the first wave by around 30%. Eventually, we will analyse the spread of the Delta variant and the associated rise of infections in Catalonia, a region in Spain, during the summer of 2021. In this sense, we also tangentially touch the biological time scale since the Delta variant is a mutation of the original virus strain. Summer 2021 was a particularly interesting period as the relaxation of restrictions, the presence of a new variant and an ongoing vaccination campaign took place simultaneously. The analysis is made via an epidemiological model that incorporates the daily number of vaccines distributed in the different age strata and the presence of two variants (Alpha & Delta). It allows for a different evolution of the contact rates in each age stratum. The fitting of the model to epidemiological data indicates that the rapid explosion in cases was mainly fuelled by increasing social interactions of the younger population, which peaked during the regional holiday of Sant Joan. However, the results also indicate that contact rates were reduced before the introduction of restrictions by the local authorities and that the latter had little effect on the epidemic evolution. The decoupling between containment measures and the adoption of social distancing by the population is particularly interesting from a behavioural point of view. It suggests the presence of what is generally referred to as pandemic fatigue.

The conclusions of the thesis will follow these contributions in Chapter 6, which summarises the work but also includes examples for future lines of research. However, before looking into the future, let us first focus on the past with the theoretical background. This overview will outline how we arrived at and what consists of present-day theory in epidemic modelling.

UNIVERSITAT ROVIRA I VIRGILI
MODELLING THE INTERPLAY BETWEEN HUMAN BEHAVIOUR AND THE SPREAD OF INFECTIOUS DISEASES: FROM TOY MODELS TO QUANTITATIVE APPROACHES
Benjamin Franz Josef Steinegger

2

THEORETICAL BACKGROUND

This chapter provides the necessary theoretical background to follow the rest of this thesis. Due to the vast amount of contributions in epidemiology, this overview is non-exhaustive. Nevertheless, we will provide a short historical review of the critical contributions that lead to the modern-day theory in epidemiology. After this historical perspective, we will also introduce a series of state-of-the-art techniques used in epidemic modelling and surveillance nowadays. Finally, while we dedicate a separate section to modelling human behaviour in the context of infectious diseases, this chapter, in general, should highlight how the spread of epidemics and human behaviour are intrinsically linked.

2.1 FIRST EPIDEMIC MODELS

The first mathematical analysis of an epidemiological problem is generally attributed to Daniel Bernoulli [23], one of the many academics in the Bernoulli family. In the middle of the 18th century, Bernoulli put forward a mathematical analysis for the benefit of life expectancy if smallpox was eliminated through variolation. Variolation, the predecessor of the smallpox vaccine, was a technique developed in China/India and imported from the Ottomans to Europe that gives immunity against smallpox by inserting smallpox scabs into superficial scratches made to the skin [24]. The approach of Bernoulli did not include a description of the epidemic contagion dynamics but rather assumed a constant annual infection risk and case fatality risk across all age groups. Bernoulli estimated that with the eradication of smallpox, the median age would increase from 11.5 years to 25.5 years. It is unknown whether Bernoulli's estimations influenced the smallpox vaccine mandate introduced in England in 1853 and public health policy in general. Interestingly, the analysis of Bernoulli sparked controversy with d'Alembert, which published a

critique of Bernoulli's work six years prior to its publication [25]. The correspondence in which Bernoulli discusses his dislike for d'Alembert with Euler is very entertaining – if not enlightening – to read, given that one has a weak spot for the gossip around science [26]. Unfortunately, Bernoulli's work, published a hundred years before Koch's and Pasteur's discovery that infectious diseases were caused by living organisms, received more attention in the actuarial sciences than in epidemiology.

2.1.1 *Towards the law of mass action*

Only in the 19th century did the first theories to understand the dynamics behind infectious disease curves emerge. More precisely, there were two competing explanations: Farr's hypothesis and Snow's hypothesis [27]. Farr claimed that epidemics end since the potency of the responsible organism is reduced with every individual it passes through [28, 29]. In contrast, John Snow hypothesised that an epidemic ends as it runs out of "fuel", i.e. of susceptible individuals [30]. Starting from the discovery of Koch and Pasteur, the increasing knowledge of how infectious diseases are transmitted caused Snow's hypothesis to gain the upper hand.

William Heaton Hamer probably best understood the implication of Snow's hypothesis implied for the dynamics of an epidemic [31]. Hamer was particularly interested in the subsequent infection waves of measles that biennially occurred in England during the 19th century [27]. Accordingly, Hamer extensively studied the available information on measles cases, and put forward a series of hypotheses in 1906 during his Milroy Lectures at the Royal College of Physicians, London [32]. The central insight that Hamer presented was that "explosions" in infections occur when the accumulation of susceptible individuals is sufficient. Beyond this basic concept, Hamer also explained the existence of inflexion points in the epidemic curves and summarised his findings in a drawing shown in Fig. 2.1.

Hamer explains the different stages of the epidemic curves with respect to the pool of susceptible individuals available. First, Hamer notices that at points B and C, the inflexion points, the number of susceptible individuals removed through infection and added through birth must be equal. Since, after point B (C), the decrease (increase) in the incidence starts to slow down and hence the number of susceptible individuals added exceeds (succeeds) the number of infected individuals. Further, Hamer explains that since the de-

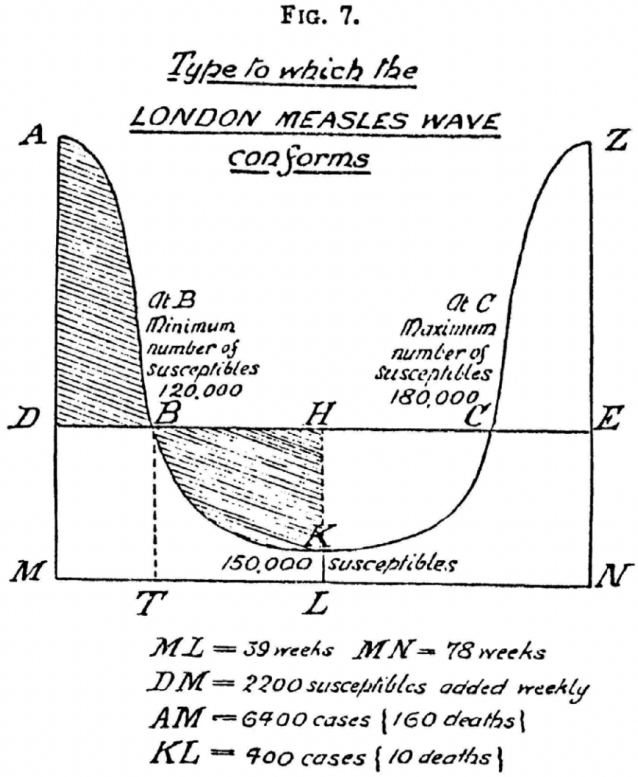


Figure 2.1: Schema – measles cases as a function of time – for the London measles wave by Hamer. The drawing is taken from Ref. [32].

crease (increase) of the incidence is maximal at B (C), the number of susceptible individuals must be minimal (maximal) at B (C). In other words, we see that Hamer developed an intuition for how the evolution of the infected individuals, i.e. the first derivative, is proportional to the number of susceptible individuals. If one wants to describe his findings in a mathematical formula [25], which he did not, we could write:

$$\text{cases of next interval} \sim \text{current cases} \times \text{current susceptibles}\,. \tag{2.1}$$

The proportionality factor would correspond to the reproduction number, which we will define later. In line with this, Hamer showed his intuition for the concept of the reproduction number and correctly assessed that in points A, K, and Z, infectious individuals infect, at least on average, precisely one susceptible individual. Overall, Hamer correctly described how

an ebbing and flowing of the susceptible population drives the subsequent measles waves. However, Hamer did not try to translate his intuition into a mathematical theory. Fortunately, only a decade later, Ross and McKendrick came forward with the first mathematical theory that described the temporal evolution of an epidemic.

Ronald Ross [33] and Anderson McKendrick [34] – both medical doctors and self-taught mathematicians – can be seen as the true founders of mathematical epidemiology. Curiously, they seem not to have been aware of the works of Hamer and did not cite his work even once. It is difficult to isolate the respective contributions of Ross and McKendrick since they worked on their theories simultaneously. Furthermore, they knew each other, and McKendrick even served in 1901 under Ross for an antimalaria campaign in Sierra Leone. Also later, they maintained regular correspondence, in which, a few years prior to the publication of their respective theories, Ross wrote to McKendrick [27]:

> We shall end by establishing a new science. But first let you and me unlock the door and then anybody can go in who likes.

The prediction could not have been more accurate. In the subsequent years, both Ross [35–37] and McKendrick [38–41] published a series of articles that laid the foundations for epidemic models as we know them today. Their main contribution is introducing the concept of mass action into epidemiology. The law of mass action mathematically describes chemical reaction kinetics and states that the rate of a chemical reaction is proportional to the product of the concentrations of the reactants [42]. Ross and McKendrick drew the analogy between molecules floating around and reacting and individuals moving in space and transmitting the disease. In other words, the law of mass action describes mathematically the intuition of Hamer that the number of newly infected individuals is proportional to the product of the number of current susceptible and infected individuals. McKendrick explicitly translated the law of mass action from chemistry to epidemiology. Ross, in contrast, saw the law of mass action as a natural extension of his theory but realised the connection solely in hindsight [27]. The different understandings may also explain the choice of their collaborators: While Ross developed the model with the mathematician Hilda Phoebe Hudson, McKendrick collaborated with the chemist William Ogilvy Kermack. Nevertheless, both arrived at the same concept, and the approaches are very similar. The simplicity of the concept of mass action, i.e. the random contact process, made it possible to analytically investigate

the dynamical behaviour of an epidemic and unveil some of the most fundamental aspects of the transmission dynamics. In particular, this mathematical approach has shown that an epidemic is not necessarily terminated by the exhaustion of the susceptible members of the community – as thought by Snow – but, instead, the interplay between population density, infectivity, recovery and death rates defines the end of an epidemic. In the following, we will outline the initial theory proposed by Kermack and McKendrick [39] since one of its special cases – the SIR model – can be found in every modern textbook on epidemic modelling. Very similar conclusions would, however, also hold for the contributions of Ross and Hudson.

2.1.2 *Original formulation of the Kermack–McKendrick model*

The seminal work by Kermack and McKendrick was published in 1927 [39]. Even though almost 100 years have passed since its publication, one can only hardly better outline the problem under consideration than the authors in the original article. Hence, we directly reproduce here their problem summary:

> One (or more) infected person is introduced into a community of individuals, more or less susceptible to the disease in question. The disease spreads from the affected to the unaffected by contact infection. Each infected person runs through the course of his sickness, and finally is removed from the number of those who are sick, by recovery or by death. The chances of recovery or death vary from day to day during the course of his illness. The chances that the affected may convey infection to the unaffected are likewise dependent upon the stage of the sickness. As the epidemic spreads, the number of unaffected members of the community becomes reduced. Since the course of an epidemic is short compared with the life of an individual, the population may be considered as remaining constant, except in as far as it is modified by deaths due to the epidemic disease itself. In the course of time the epidemic may come to an end. One of the most important problems in epidemiology is to ascertain whether this termination occurs only when no susceptible individuals are left, or whether the interplay of the various factors of infectivity, recovery and mortality, may result in termination, whilst many susceptible individuals are still present in the unaffected population.

In the above paragraph, the authors introduce a series of model assumptions that can be summarised as [43]:

1. Transmission occurs upon physical contact.
2. Infection results in either complete immunity or death.
3. Individuals are equally susceptible.
4. The population is closed, i.e. the influx of susceptible individuals is negligible at the time scale of the epidemic.

The authors then introduce the following two additional assumptions later in the article:

5. The population size is large enough such that stochastic effects can be neglected, and a deterministic description is possible.
6. Contacts are according to the law of mass action.

Starting from these assumptions, Kermack and McKendrick build a discrete-time model to determine the criteria under which an epidemic ends. They define $v_{t,\theta}$ as the number of infected individuals that have been infected for θ time steps at time t, i.e. that got infected at time $t-\theta$. Defining ψ_θ as the probability to be removed, through death or recovery, θ time steps after infection, $v_{t,\theta}$ follows the recurrence relation

$$\begin{aligned} v_{t,\theta} &= v_{t-1,\theta-1}\,(1-\psi_{\theta-1}) \\ &= v_{t-2,\theta-2}\,(1-\psi_{\theta-1})\,(1-\psi_{\theta-2}) \\ &= v_{t-\theta,0}\underbrace{\prod_{k=1}^{\theta}(1-\psi_{\theta-k})}_{B_\theta}\,. \end{aligned} \tag{2.2}$$

The quantity B_θ corresponds to the survival function of an infected individual. Having expressed the evolution of the infected individuals, we can now describe the newly infected individuals at time t. Defining s_t as the density of susceptible individuals (unaffected) at time t, the newly infected individuals $v_t = v_{t,0}$ at time t correspond to $S_t - S_{t+1}$. It is to say that $v_t = v_{t,0}$ holds for $t \neq 0$. For $t = 0$, given the initial number of infected individuals I_0, we have

$v_0 + I_0 = v_{0,0}$. With ϕ_θ being the transmission probability, i.e. infectivity, θ time steps after infection, the law of mass action allows v_t to be expressed as

$$v_t = s_t \sum_{\theta=1}^{t} \phi_\theta v_{t,\theta} \,. \tag{2.3}$$

Hence, given the recurrence relation in Eq. (2.2), we can use the above equation to express $S_{t+1} - S_t$ as

$$S_{t+1} - S_t = -S_t \left(\sum_{\theta=1}^{t} \underbrace{\phi_\theta B_\theta}_{A_\theta} v_{t-\theta} + \underbrace{\phi_t B_t}_{A_t} I_0 \right) . \tag{2.4}$$

Here, A_θ serves as an auxiliary variable to simplify notation. With the removal probability ψ_θ and the survival function B_θ, the density of removed individuals R_t follows, similarly to the susceptible ones, the equation

$$R_{t+1} - R_t = \sum_{\theta=1}^{t} \underbrace{\psi_\theta B_\theta}_{C_\theta} v_{t-\theta} + \underbrace{\psi_t B_t}_{C_t} I_0 \,. \tag{2.5}$$

The quantity $C_\theta = \psi_\theta B_\theta$ encapsulates the probability to be removed precisely after θ time steps. Due to the assumption of a closed population, the following conservation equation holds

$$N = S_t + I_t + R_t \,, \tag{2.6}$$

where I_t is number of infected individuals at time t and N the population size. Accordingly, following Eqs. (2.4) & (2.5), the number of infected individuals evolves as

$$I_{t+1} - I_t = S_t \left(\sum_{\theta=1}^{t} A_\theta v_{t-\theta} + A_t I_0 \right) - \left(\sum_{\theta=1}^{t} C_\theta v_{t-\theta} + C_t I_0 \right) . \tag{2.7}$$

The Eqs. (2.4), (2.5) and (2.7) describe the dynamics of the system in discrete time. Kermack and McKendrick then consider the continuous-time limit of this set of equations, which only requires replacing sums with integrals and

differences with derivatives. Accordingly, the system of differential equations becomes

$$\frac{\mathrm{d}S_t}{\mathrm{d}t} = -S_t \left(\int_0^t A_\theta v_{t-\theta} \mathrm{d}\theta + A_t I_0 \right) \tag{2.8}$$

$$\frac{\mathrm{d}I_t}{\mathrm{d}t} = S_t \left(\int_0^t A_\theta v_{t-\theta} \mathrm{d}\theta + A_t I_0 \right) - \int_0^t C_\theta v_{t-\theta} \mathrm{d}\theta + C_t I_0 \tag{2.9}$$

$$\frac{\mathrm{d}R_t}{\mathrm{d}t} = \int_0^t C_\theta v_{t-\theta} \mathrm{d}\theta + C_t I_0 \,, \tag{2.10}$$

where

$$v_t = -\frac{\mathrm{d}S_t}{\mathrm{d}t} \,. \tag{2.11}$$

The survival probability B_θ becomes

$$B_\theta = \exp\left(- \int_0^\theta \psi_\alpha \mathrm{d}\alpha \right) \,. \tag{2.12}$$

The discrete-time expression of B_θ in Eq. (2.2) can be recovered by replacing the integral in Eq. (2.12) by a sum and subsequently making a first-order Taylor expansion. Further, the relations $A_\theta = \phi_\theta B_\theta$ and $C_\theta = \psi_\theta B_\theta$ still hold. From there, Kermack and McKendrick make a lengthy and tedious mathematical analysis to determine the conditions under which the epidemic stops. We will skip this analysis and directly focus on the special case they present afterwards, the case of a constant removal rate and infectivity.

2.1.3 *The successful particular case - The SIR model*

Let us denote the constant removal rate with μ and the constant infectivity with β such that $\psi_\theta = \mu$ and $\phi_\theta = \beta \; \forall \; \theta \geq 0$. Accordingly, we have $A_\theta = \beta B_\theta$ and $C_\theta = \mu B_\theta$ with $B_\theta = \exp(-\mu\theta)$. With these expressions, A_θ and C_θ are proportional to each other, i.e. $A_\theta / C_\theta = \beta / \mu$, implying that in Eqs. (2.8)-(2.10) the following expression appears four times:

$$\int_0^t B_\theta v_{t-\theta} \mathrm{d}\theta + B_t I_0 \,.$$

Let us inspect this expression closer. The integrand in the first term expresses, through the survival function B_θ, the number of individuals infected at $t-\theta$ ($v_{t-\theta}$), which have not been removed until time t. Accordingly, the entire integral defines the total number of individuals infected during the epidemic, which have not been removed until time t. Similarly, the second term expresses the number of initially infected individuals at time $t=0$ that have not been removed until time t. Accordingly, the above term is nothing else than the number of infected individuals at time t, i.e. I_t, and, eventually, Eqs. (2.8)-(2.10) become

$$\frac{\mathrm{d}S_t}{\mathrm{d}t} = -\beta S_t I_t \tag{2.13}$$

$$\frac{\mathrm{d}I_t}{\mathrm{d}t} = \beta S_t I_t - \mu I_t \tag{2.14}$$

$$\frac{\mathrm{d}R_t}{\mathrm{d}t} = \mu I_t\,. \tag{2.15}$$

The above equations correspond to the standard susceptible-infected-recovered (SIR) model. We have shown that the SIR model is a special case of the general framework put forward by Kermack and McKendrick. It assumes constant infectivity over the course of an infection and a constant removal rate μ. These assumptions could not be further off from reality, but they allow for a simple description of the epidemic dynamics and tractability of the model. We will discuss in Sec. 2.3.2 how a non-trivially shaped infectivity profile affects the dynamics of an epidemic.

Making a dimensional analysis, we note that both μ and β are rates, i.e. they are in the units of 1/time. For the removal rate, this is quite intuitive since we describe removal through a constant rate (homogeneous Poisson process), which results in an exponentially distributed waiting time from infection until removal. In contrast, it is difficult to grasp the concept of an infectivity rate β. The difficulty stems from the fact that β acts, at least conceptually, as an effective parameter that combines the transmission probability λ and the contact rate k. It seems much more intuitive to understand the spreading process as individuals that enter into contact at a given rate k during which the virus is spread with a probability λ, implying $\beta = \lambda k$.

The dimensionless interaction term $S_t I_t$ corresponds to the law of mass action. As for chemical reactions, it is a density-based approach, wherefore the variables S_t, I_t, and R_t theoretically refer to densities per surface area and not to absolute numbers. However, since the dynamic occurs in a fixed

surface area, one can rescale the densities by the considered area and work with the absolute numbers. However, in this framework, the dynamic is not independent of the population size. Actually, if we consider a population of size N, the fraction of infected individuals i_t evolves according to Eqs. (2.13)-(2.15) as

$$\frac{\mathrm{d}i_t}{\mathrm{d}t} = \lambda(Nk)s_t i_t - \mu i_t\,, \tag{2.16}$$

where s_t refers to the fraction of susceptible individuals. We note that the contact rate k scales with N. In other words, the larger the population, i.e. the denser the population, the faster the epidemic spreads. Therefore, this framework is generally known as the *density dependent* approach [18, 44]. This approach makes sense for plant and animal diseases, where the closer animals are, the more often they interact. However, if the virus hosts are humans, such an assumption may not be reasonable. Someone in London (population of 8 million) has probably not over 10 times more contacts than an individual living in Zurich (population of 500'000). Social constraints in these contexts are more likely to determine disease dynamics. Whether the contact rate is dependent on density – maybe not linear – is still a topic of research, and evidence is contradictory [45–48]. For this reason, there also exists an alternative *frequency dependent* approach that assumes no dependence of the contact rate on the population size [18, 44]. In this framework, individuals interact with the contact rate k independently of the population size, which comes down to removing N in Eq. (2.16). Accordingly, in the frequency-dependent approach, the differential equations read

$$\frac{\mathrm{d}s(t)}{\mathrm{d}t} = -\lambda k s(t) i(t) \tag{2.17}$$

$$\frac{\mathrm{d}i(t)}{\mathrm{d}t} = \lambda k s(t) i(t) - \mu i(t) \tag{2.18}$$

$$\frac{\mathrm{d}r(t)}{\mathrm{d}t} = \mu i(t)\,, \tag{2.19}$$

where $s(t)$, $i(t)$ and $r(t)$ are the fraction of susceptible, infected and removed individuals, respectively. Please note that we adapted the time index to a more modern notation. The frequency-dependent approach is more commonly found in textbooks nowadays. In the following, we will present the basic phenomenology that emerges from this minimal model.

2.1.3.1 Phenomenology of the SIR model

According to Kermack and McKendrick, the most immediate question which needs to be answered is to see what terminates an epidemic. To this aim, we will consider the numerical solution of the Eqs. (2.17)-(2.19) which is presented in Fig. 2.2 (A). The temporal evolution shows that the intuition by Hamer was pretty accurate: the depletion of susceptible individuals modulates the propagation of epidemics. As previously thought, the epidemic does not stop when no susceptible individuals are available anymore, but a specific value of recovered individuals seems to terminate the epidemic.

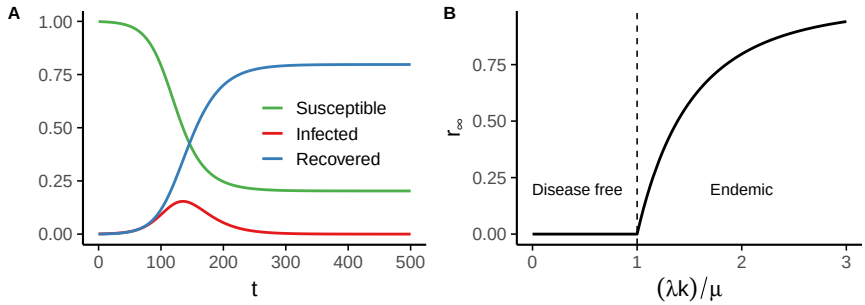


Figure 2.2: **(A)**: Exemplary temporal evolution of the SIR model. **(B)**: Final attack rate, r_∞, as a function of the parameters of the SIR model.

This specific value is generally referred to as the final attack rate $r_\infty = \lim_{t\to\infty} r(t)$ and corresponds to the total fraction of individuals that had been infected at the end of the epidemic. Unfortunately, there is no explicit analytical expression for r_∞ but only a transcendental equation. To get to this equation, let us divide Eq. (2.17) by Eq. (2.19), giving us

$$\frac{\mathrm{d}s}{\mathrm{d}r} = -\frac{\lambda k}{\mu} s\,. \tag{2.20}$$

Integrating this equation, we get

$$s(t) = s(0)\exp\left(-\frac{\lambda k}{\mu}(r(t) - r(0))\right)\,. \tag{2.21}$$

As laid out by Kermack and McKendrick, we consider the introduction of only a few infected individuals into a fully susceptible population. We can therefore assume that $s(0) \approx 1.0$ and $r(0) \approx 0$. Furthermore, at the end of the epidemic, no infected individuals remain, and $i_\infty = 0$. Due to the

conservation equation, we thus have the relation $s_\infty = 1 - r_\infty$ and Eq. (2.21) becomes

$$r_\infty = 1 - \exp\left(-\frac{\lambda k}{\mu} r_\infty\right) . \tag{2.22}$$

As previously mentioned, this equation has no explicit analytical solution besides the trivial solution $r_\infty = 0$. However, we can analyse under which condition an additional non-trivial solution exists. The variable r_∞ is bounded between 0 and 1. At the upper boundary, $r_\infty = 1$, the left-hand side (LHS) of Eq. (2.22) is bigger than the right-hand side (RHS). Since the LHS and the RHS are both 0 at the lower boundary, $r_\infty = 0$, an additional solution can thus only exist if the RHS grows faster than the LHS as r_∞ increases. In mathematical terms, this means that the derivative of the RHS with respect to r_∞ evaluated at 0 must exceed the one of the LHS. The comparison of the two derivatives yields the inequality

$$\frac{\lambda k}{\mu} \geq 1 . \tag{2.23}$$

Accordingly, if the above condition is fulfilled, Eq. (2.22) also has a non-trivial solution. However, the question remains of whether these two solutions are stable. We thus analyse the stability of the trivial solution by linearising Eq. (2.18) around the disease-free equilibrium $i \approx \epsilon$ and $s \approx 1 - \epsilon$ leading to

$$\frac{\mathrm{d}i(t)}{dt} \approx (\lambda k - \mu)\, i(t) . \tag{2.24}$$

Hence, whenever $\lambda k \geq \mu$, the coefficient in Eq. (2.24) is positive, and the disease grows exponentially, making the disease-free equilibrium $i = 0$ unstable. In this case, the condition in Eq. (2.23) is fulfilled, and the system evolves towards the non-trivial equilibrium point with $r_\infty > 0$. Fig. 2.2 (B) shows how the criteria $(\lambda k)/\mu \geq 1$ separates a disease free phase ($r_\infty = 0$) from an endemic phase ($r_\infty > 0$). Physicists, also in the context of epidemic models, often refer to this as a phase transition [49]. Analogous to the critical temperature in magnetism, one can define the critical transmission probability $\lambda_c = \mu/k$, often referred to as the epidemic threshold, above which the disease is endemic. In the context of dynamical systems, the phase separation is generally referred to as a bifurcation, i.e. a change in the stability of certain equilibrium points [50].

In the above example, we assumed that infected individuals are introduced in a fully susceptible population. However, assuming that a disease may be

reintroduced, as was the case for measles considered by Hamer, not the entire population may be susceptible. If the population is not fully susceptible, one cannot assume $s \approx 1 - \epsilon$ since s is far from one. Accordingly, in Eq. (2.24) s would also multiply the term λk and the condition Eq. (2.23) would thus read

$$\frac{\lambda k s}{\mu} \geq 1 \,. \tag{2.25}$$

Therefore, the disease can actually spread as s increases and exceeds the threshold $s_c = \mu/(\lambda k)$. This dependence on s thus completes the intuition of Hamer, how an increasing pool of susceptible individuals can lead to the emergence of new epidemic waves. However, beyond that, the mathematical analysis of Kermack and McKendrick unveiled how the interplay between a series of parameters – contact rate, removal rate, transmission probability – determines whether a disease can spread or an epidemic terminates. This threshold, which separates the disease-free state from the endemic state, has an intuitive interpretation that Kermack and McKendrick completely missed [51]. The intuitive interpretation comes from a quantity called the *basic reproduction number*, which we will detail in the next section.

2.1.4 *The basic reproduction number*

Let us first consider the threshold in Eq. (2.23), which assumes a fully susceptible population. The parameter μ is the rate at which individuals stop being infectious. The infectious period thus follows an exponential distribution with mean τ_I given by $\tau_I = 1/\mu$. Accordingly, given that k is the number of contacts per unit of time, the term $k/\mu = k\tau_I$ corresponds to the total number of contacts made by an individual while being infectious, i.e. during the infectious period. Multiplying this, the number of contacts, by the transmission probability λ, thus gives the number of people infected during the course of an individual's infectious period. The number of cases caused by an infected individual – primary case – is generally referred to as the number of secondary cases. Therefore, the threshold condition in Eq. (2.23) expresses that the disease can only propagate if the number of secondary cases is above one.

This condition is intuitive if we make the connection to demography or, even better, to ecology. The most straightforward example is to focus on hermaphrodites that do not have opposite sexes and only reproduce through

self-fertilisation. In this case, it seems evident that, on average, exemplars of the species need to reproduce at least once to guarantee survival of the species. Accordingly, defining the number of offspring as the reproduction number of the species, if it is above 1, the population will grow exponentially. In contrast, the population will decrease if it is below 1, and the species will eventually die out.

Now, in the context of a disease, it is not individuals that reproduce but hosts for the virus. Similarly, suppose an infected individual (primary host) does not infect at least one other individual (secondary host). In that case, the virus will not reproduce since the primary host will recover at some point. We define the basic reproduction number R_0 as the number of secondary cases generated by a primary case in a fully susceptible population. As mentioned before, for the SIR model, R_0 is expressed as

$$R_0 = \frac{\lambda k}{\mu} = \lambda k \tau_I \,. \tag{2.26}$$

The connection to population growth is immediate if we consider Eq. (2.24). In line with the definition of R_0, the linearised equation describes the epidemic growth in a fully susceptible population. If we rescale time by $1/\mu = \tau_I$, Eq. (2.24) can be written as

$$\frac{\mathrm{d}i(t)}{\mathrm{d}t} = (R_0 - 1)\, i(t) \,. \tag{2.27}$$

The equation describes the initial growth of an epidemic, where R_0 determines the growth factor. There is the subtraction since the infected individual will recover after a time τ_I, which is the time scale of the dynamics after rescaling. In line with the condition in Eq.(2.23), the epidemic grows exponentially if $R_0 > 1$ and decays exponentially if $R_0 < 1$.

Curiously, it took almost fifty years until the concept of the reproduction number was imported from demography to mathematical epidemiology. Alfred J. Lotka, in 1925, was the first to introduce the concept of the reproduction number [52, 53]. He came up with it in the context of demography and referred to it as the rate of natural increase per head. At the time, Lotka worked at Metropolitan Life Insurance in New York City. After Daniel Bernoulli, this was the second time epidemiology and actuarial sciences crossed paths. Furthermore, Lotka worked in ecology and is mainly known for the Lotka-Volterra equations [54]. Lotka also contributed early on to mathematical epidemiology, was well aware of the work by Ross, and pro-

Disease	R_0
Syphilis	1-2
EBOV (2014 outbreak)	1.5 - 2.5
IAV H1N1 (1918 pandemic)	2-3
HIV-1	2-5
SARS-CoV-1 (2003)	2-5
SARS-CoV-2 (2020)	2-4
Mumps	4-7
Polio	5-7
Rubella	5-7
Measles	12-17

Table 2.1: Basic reproduction number R_0 for a series of viral pathogens [8, 57].

posed the first solution to the malaria model laid out by Ross in the footnotes of his lecture notes in 1911 [55]. Nevertheless, Lotka, working mainly on malaria, could not translate the concept of the reproduction number from demography to epidemiology. Moreover, the concept of a reproduction number is not straightforward to understand for vector-borne diseases, such as malaria, since mosquitoes transmit the virus. Despite this difficulty, George MacDonald, which developed the mathematical theory to model malaria transmission together with Ross [56], was the first to use the term reproduction rate in the context of epidemics [51]. Nevertheless, it took decades for the term basic reproduction number to become widely used. In particular, Roy Anderson and Robert May, both having a background in ecology, popularised the concept of R_0 in the 80s and 90s [51].

Eventually, the basic reproduction number became an intuitive and instructive concept. It is specific to a particular disease, i.e. its transmissibility λ and infectious period τ_I. Accordingly, the basic reproduction number classified diseases according to their capability to spread in a population. Table 2.1 shows the estimated basic reproduction number for a series of diseases. The reproduction numbers show some heterogeneity, with measles having the largest basic reproduction number.

In addition to classifying different pathogens, via the basic reproduction number, one can estimate the number of individuals that need to be immunised to reach *herd immunity* and eradicate a disease [58, 59]. Typically, immunisation is achieved through vaccination, which effectively reduces the pool of susceptible individuals. Accordingly, in the presence of vaccinated individu-

als, the population is not fully susceptible, and we can consider the condition in Eq. (2.25): $R_0 s \geq 1$. After an intervention, such as the distribution of vaccines for example, the LHS in the inequality $R_0 s$ is generally referred to as the *reproduction number* instead of *basic reproduction number*; but it still represents the same quantity, i.e. the number of secondary cases. Nevertheless, since not the entire population is susceptible, not every contact of an infectious individual can yield a secondary case and thus the multiplication by s. However, we will discuss this in detail in Sec. 2.3. Assuming that a fraction v of the population is vaccinated and the vaccine is 100% effective, we can express the susceptible population as $s = 1 - v$. Inserting the relation into the transmission threshold yields a critical vaccination coverage $v_c = 1 - 1/R_0$ necessary to eradicate the disease. Accordingly, to eradicate measles, more than 90% of the population need to be vaccinated.

2.1.5 *From the SIR model to compartmental models in general*

The simple SIR model not only provides a rich phenomenology but can also quite accurately describe the evolution of actual real-world epidemics. Already Kermack and McKendrick successfully adjusted their model framework to a plague outbreak between 1905 and 1906 in the islands of Bombay, India. However, while the three health statuses *susceptible*, *infected* and *removed* may describe the plague well, diseases which do not confer life-long immunity and reinfection is, thus, possible, require a different framework. Such an example is Syphilis, where mostly asymptomatic reinfections make it very challenging to control the spread of the disease [60].

The simplest model that incorporates reinfections is the SIS model (susceptible-infected-susceptible). In this model, recovered individuals are not removed but become susceptible again. Accordingly, the following set of differential equations governs the dynamics:

$$\frac{\mathrm{d}s(t)}{\mathrm{d}t} = -\lambda k s(t) i(t) + \mu i(t) \tag{2.28}$$

$$\frac{\mathrm{d}i(t)}{\mathrm{d}t} = \lambda k s(t) i(t) - \mu i(t)\,. \tag{2.29}$$

In contrast to the SIR model, the SIS model allows for an explicit analytical expression for the non-trivial stationary state. The stationary state is found as $i^* = 1 - \mu/(\lambda k)$ from the equilibrium conditions. Accordingly, for $\lambda k < \mu$, we have $i^* < 0$. In this case, the disease-free state, i.e. the trivial equilibrium,

is stable. The disease-free state loses stability as $\lambda k > \mu$ while the non-trivial state is stable. We thus have the same critical transmission probability λ_c for the SIS model as for the SIR model.

The SIS model is just one example of all the existing epidemiological models. More generally, the different health statuses in the model – the dynamical variables of the system – are referred to as *compartments*, and the models are known as *compartmental models* [61]. On a conceptual level, the epidemic compartments effectively describe the within-host dynamics resulting from the immune response to the replicating virus. Epidemiological models can include many compartments, mainly if the clinical evolution, such as hospital admission, ICU admission or death, is modelled in detail. For the modelling of SARS-CoV-2 in Chapter 5, we will encounter examples that, besides the clinical evolution, also include different virus variants and the vaccination status. However, additional compartments generally only complicate the notation of epidemiological models. Once infected, individuals transit between compartments according to the defined probabilities and rates. Writing the model equations in these cases is straightforward. However, solely the infection term can be non-straightforward to express. Often, the infection term is expressed as $\lambda(t)s(t)$, where $\lambda(t)$ is referred to as the *force of infection* [59]. In the above examples, the force of infection corresponds to the product between the transmission probability λ, the contact rate k, and the fraction of infected individuals $i(t)$, i.e. $\lambda(t) = \lambda k i(t)$.

A standard compartment that can be found in epidemiological models is the exposed compartment E, in which individuals are infected but not yet infectious. For example, for SARS-CoV-2, this latency period until individuals become infectious is between 2 and 3 days [62]. For a disease that provides immunity after infection but has a latency period, we refer to the SEIR model (susceptible-exposed-infected-recovered). However, some diseases provide immunity after infection, but the immunity wanes over time. The waning immunity may be caused by decreasing antibody levels, such as for SARS-CoV-2 [63], or by the virus' mutation in the case of influenza [64], for example. From a modelling perspective, this means that individuals can transfer from the recovered compartment to the susceptible compartment, wherefore, assuming that the disease has a latency period, we refer to this model as the SEIRS model. In analogy to the SIS model, the possibility of reinfection violates the second assumption by Kermack and McKendrick that infection results in either complete immunity or death. Similarly to Ross [36], Kermack and McKendrick have softened this assumption already in

their second article [40]. Further, they also soften the fourth assumption – closed population – and include a birth process in the model. Nevertheless, their mathematical analysis stays quite similar to the original article. However, since it was central to their approach, Kermack and McKendrick never considered dropping assumptions 3 and 6. In combination, the two assumptions imply that all individuals interact with the same contact rate and that contacts are made randomly, i.e. according to the law of mass action. In the following, we will show how the study of more complex contact structures was the main focus in mathematical epidemiology during the second half of the 20th century [65].

2.2 INCORPORATING HETEROGENEITY INTO EPIDEMIC MODELS

To this point, we have considered models that assume a population of indistinguishable agents interacting randomly and with the same contact rate. However, this is a minimal description of the human population's diversity and the patterns that underlie our interactions. Most apparently, geography shapes how we physically interact and along which paths diseases can propagate. Similarly, not all individuals interact at the same frequency. In many circumstances, individuals are less successful in establishing and maintaining social contacts. This section will illustrate how different aspects such as age, geography, contact heterogeneity and contact networks can be incorporated into epidemic models. All the mentioned examples introduce heterogeneity into the disease dynamics and are at the core of a diverse set of empirical phenomena.

2.2.1 *Age stratified models*

From the 1950s onward, researchers in epidemiology have put much work into modelling childhood diseases such as measles, or mumps [59]. Studies have indicated that stochastic fluctuations [66], an accumulation of newly born susceptibles and parametric resonance with the seasonally varying transmission probability [67] may be the drivers behind the recurrent measles outbreaks. Naturally, researchers also devoted their attention to designing effective immunisation strategies through mass vaccination to eradicate childhood diseases and prevent these recurrent outbreaks [68, 69]. Most of these models assumed a global population of equally susceptible individuals that

interacted according to the law of mass action [70]. This assumption corresponds to assumptions 3 and 6 in the model of Kermack and McKendrick. It is generally referred to as the *homogeneous mixing* assumption under which contacts occur randomly and with equal probability [18]. The existing models struggled to conform to various datasets, wherefore researchers questioned if homogeneous mixing may be too crude of an approximation for the actual contact structure [70, 71]. The heterogeneous behaviour in different age groups has shown to be a key determinant for the spread of childhood diseases [59]. Daily contacts at school for children and work for adults cause human contact patterns not to be random but rather follow a clear structure [72]. Fig. 2.3 A shows a paradigmatic example of such a contact pattern. For the above reasons, individuals generally interact more often with others in the same group, i.e. the contact structure is modular [73]. Further, the contact structure is fork-like due to the interactions between parents and their children.

To model these structured interactions, researchers put forward frameworks that stratify the population into different age groups [70, 71]. The contact rates between the different strata are then encoded in a contact matrix C, where the matrix elements C_{ij} define the frequency at which individuals in age group i interact with the ones in age group j. The matrix is sometimes also referred to as the *who acquires infection from whom* (WAIFW) matrix [18]. If we assume an SIS-like model, the fraction of infected individuals x_i in age group i evolves as

$$\dot{x}_i(t) = -\mu x_i(t) + (1 - x_i(t))\lambda \sum_{j=1}^{M} C_{ij} x_j(t)\,, \tag{2.30}$$

where M refers to the total number of age strata. As previously mentioned, the notation of the force of infection $\lambda_i(t)$ is often taken [18]. The force of infection subsumes the infection term as

$$\lambda_i(t) = \lambda \sum_{j=1}^{M} C_{ij} x_j(t)\,, \tag{2.31}$$

allowing to rewrite Eq. (2.30) as

$$\dot{x}_i(t) = -\mu x_i(t) + \lambda_i(t)(1 - x_i(t))\,. \tag{2.32}$$

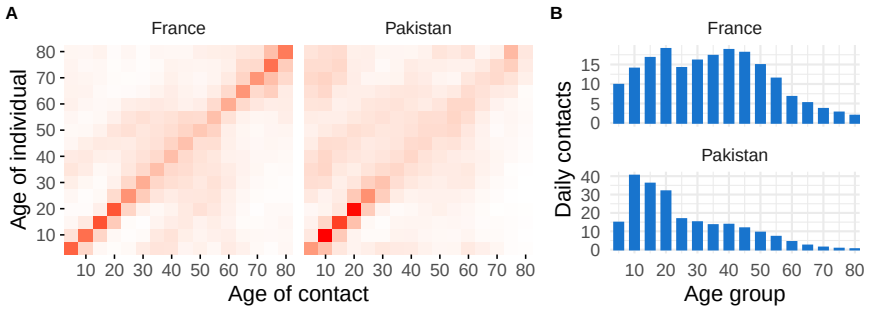


Figure 2.3: **(A)**: Relative mixing patterns across age groups in France and Pakistan [72]. **(B)**: Number of daily contacts of the different age groups in France and Pakistan [72].

The approach of the force of infection is particularly attractive since, almost independently of the model's complexity, the equations can always be brought into the above form. The potentially complicated infection term can be hidden in $\lambda_i(t)$ allowing for a clean notation. Eq. (2.30) can be written in many different but equivalent ways. For example, instead of expressing the evolution of the fraction of infected individuals, one can work with the absolute number of infected individuals X_j in the respective age groups. In this case, the term x_j does not appear in the force of infection but rather X_j/N_j. Further, instead of fixing the absolute contact rate between age groups with the contact matrix C, one can parametrise the relative distribution of contacts with the matrix C ($\sum_j C_{ij} = 1$) and encode the total contact rate of age group i as k_i. In this case, the frequency at which individuals in age group i interact with the ones belonging to age group j reads $k_i C_{ij}$. However, these are only differences in notation that do not affect the actual dynamics.

The contact matrix C is generally informed by empirical studies that estimate through surveys how individuals from different age groups interact with each other [72, 74–77]. Substantial differences between countries are found that can often be attributed to cultural peculiarities. For example, in Fig. 2.3 A, we can see that the elderly in France interact primarily with each other, while in Pakistan, the elderly quite frequently interact with children. This difference may be attributed to an absence of daycare facilities, requiring grandparents to watch the children during the day. The data does not only indicate that individuals do not interact non-random, i.e. not according to the law of mass action, but also that the number of contacts is very different across age groups. In general, younger individuals have many more contacts per

day than the elderly (Fig. 2.3 B). Accordingly, not all individuals are equally susceptible. Now, for regular contacts, the heterogeneity in the population is not very pronounced. The difference is only around one order of magnitude. However, facing the accelerating spread of sexually transmitted diseases in the 60s and 70s, researchers quickly realised that the more pronounced heterogeneity in sexual contacts is one of the key drivers behind their spread [78, 79]. Eventually, advances paved the way to model the spread of an epidemic in actual contact networks, which we will detail in the following sections.

2.2.2 *Heterogeneity in the contact rates*

There was a general worry in the 1960s that the introduction of hormonal contraception, together with the sexual revolution, could lead to an increase in sexually risky behaviour and eventually accelerate the spread of sexually transmitted infections. For example, annual Gonorrhea cases in the United States increased from 200'000 to almost 1'000'000 between the 60s and the 70s [79]. In hindsight, this was partially due to the increase mentioned above in risky behaviour, but probably also due to better surveillance [18]. Facing this rapid increase, researchers built models to unveil the driving factors behind the spread of Gonorrhea [78]. However, when parametrised with realistic values, models which assume homogeneous mixing would suggest that Gonorrhea should not be able to spread, i.e. it is below threshold. However, it turned out that solely considering the average partner acquisition rate and neglecting its standard deviation underestimates the virus' ability to spread [18]. In other words, this discrepancy required challenging assumption 3 that all individuals are equally susceptible since more frequent interaction increases the probability (susceptibility) of (to) infection.

Accordingly, models were built to distinguish different activity classes [78, 80]. In these models, the population is not stratified according to age but rather according to the different partner acquisition rates (contact rates). We thus consider M activity classes, where the contact rate of activity class i is k_i. We further express the probability that an interaction of an individual in activity class i takes place with an individual in activity class j as $P(i, j)$. This approach corresponds to the normalised mixing matrix mentioned before.

Since reinfections are possible for Gonorrhea, we will consider an SIS model for which the model equations then read

$$\dot{x}_i(t) = -\mu x_i(t) + (1 - x_i(t))\lambda k_i \sum_{j=1}^{M} P(i,j) x_j(t)\,. \tag{2.33}$$

Naturally, the question arises how to describe the relative mixing $P(k,l)$. Evidence suggests that the mixing pattern is assortative, i.e. that individuals tend to more frequently interact with others in a similar activity class [81–83]. However, including assortativity makes the mathematical analysis of the dynamics more challenging. Hence, once more, we will reintroduce assumption 6 and employ the law of mass action to simplify calculations.

The law of mass action assumes that individuals interact randomly, equivalent to molecules. Since we consider different activity classes, not all individuals interact equally often. Accordingly, in this context, randomly interacting does not mean interacting with a random individual but rather with a random available contact. In other words, the probability of interacting with an individual in activity class i is proportional to the fraction of the population p_i belonging to the activity class i and their contact rate k_i. This approach to express the interaction term is often referred to as *proportionate mixing* [18]. Therefore, we can express the probability $P(i,j)$ as

$$P(i,j) = \frac{p_j k_j}{\sum_{l=1}^{M} k_l p_l} = \frac{p_j k_j}{\langle k \rangle}\,. \tag{2.34}$$

The term $\langle k \rangle$ denotes the average contact rate in the population. Therefore, the dynamics in Eq. (2.33) can be written as

$$\dot{x}_i(t) = -\mu x_i(t) + (1 - x_i(t))\lambda \frac{k_i}{\langle k \rangle} z(t)\,, \tag{2.35}$$

with

$$z(t) = \sum_{j=1}^{M} p_j k_j x_j(t)\,. \tag{2.36}$$

The quantity $z(t)$ expresses the average contact rate of an infected individual. Looking at Eq. (2.35), we see that, at equilibrium, $x_i > 0$ for any i holds if and only if $z > 0$. Accordingly, to find the threshold condition, i.e. the criterium that separates the disease-free phase from the endemic one, we can focus on the equilibrium value of z. Similarly, as for the standard SIR model

(sec. 2.1.3.1), in order to find the epidemic threshold, we assume $x_i \approx \epsilon$. Accordingly, multiplying both sides in Eq. (2.35) by $p_j k_j$ and summing over j, we find

$$\dot{z}(t) = \left(-\mu + \lambda \frac{\langle k^2 \rangle}{\langle k \rangle} \right) z(t) \,. \tag{2.37}$$

The term $\langle k^2 \rangle$ is defined as $\sum_j p_J k_j^2$. Therefore, the disease-free equilibrium is only unstable, and the dynamics evolve towards the endemic state if

$$\frac{\lambda}{\mu} \frac{\langle k^2 \rangle}{\langle k \rangle} > 1 \,, \tag{2.38}$$

with a critical transmission probability

$$\lambda_c = \mu \frac{\langle k \rangle}{\langle k^2 \rangle} \,. \tag{2.39}$$

Accordingly, as already pointed out, in the presence of different activity classes, not only the average contact rate determines whether the disease can spread, but also the dispersion in the contact rates (second moment). For a homogeneous population, in which everyone interacts at the same contact rate k and is thus equally susceptible (assumption 3), the above threshold immediately reduces to the one in Eq. (2.23) since $\langle k^2 \rangle = k^2$.

To highlight the impact of the heterogeneity in the contact rates, let us consider a population in which a fraction of the population has a contact rate k_{low} and the rest have a contact rate $k_{high} = \alpha k_{low}$ with $\alpha > 1$. The parameter α controls the heterogeneity in the contact rates. Further, we fix the average degree in the population as $\langle k \rangle = \gamma k_{low}$ with $\alpha > \gamma > 1$. Accordingly, the fraction of the population p interacting with contact rate k_{low} is given by

$$p = \frac{\alpha - \gamma}{\alpha - 1} \,. \tag{2.40}$$

Please note that the condition $\langle k \rangle = \gamma k_{low}$ imposes γ as a lower bound on α. With the expression of p, the ratio between the second moment $\langle k^2 \rangle$ and the first moment $\langle k \rangle$ is expressed as

$$\frac{\langle k^2 \rangle}{\langle k \rangle} = \left(\frac{\alpha^2}{\alpha - 1} \frac{(\gamma - 1)}{(\alpha - 1)} + \frac{\alpha - \gamma}{\alpha - 1} \right) k_{low} \,. \tag{2.41}$$

Calculations to show that the expression in the above equations monotonously increases with α are tedious. Therefore, we consider the limit for $\alpha \gg 1$,

where we see that $\langle k^2 \rangle / \langle k \rangle$ scales linearly with α. Accordingly, variations in the contact rates promote the spread of a disease even though the average contact rate stays constant (constant γ). For this reason, Gonorrhea can still spread even though homogeneous mixing would suggest that Gonorrhea is below threshold.

In modelling sexually transmitted infections, considering different activity classes, often also called *risk classes*, is a standard approach [78, 84–88]. Nevertheless, it is often not straightforward and somewhat arbitrary to decide the number of risk classes that stratify the population [89]. In reality, every individual has a different contact rate and, more specifically, a distinct contact pattern that results in a contact network at the population level. From the mid-80s onward, in the context of HIV transmission, researchers slowly started to realise that these contact networks crucially shape disease transmission [90–97]. In the next section, we will see how the properties of these networks affect the spreading dynamics and how these phenomena are modelled.

2.2.3 *Epidemics on networks*

In today's day and age, most of us associate the term network directly with online social networks. However, social networks shaped our society and were studied long before Myspace or Facebook existed. Also, in the analogue days, individuals interacted in a very structured way that can generally be represented as a network. From the 1930s onward, when Moreno started the field of sociometry [98], which later evolved into social network analysis, sociologists extensively studied the structure of social networks [99–101]. Moreover, they analysed how the topological structure shapes group dynamics [102], the spread of innovations [103], or power dynamics [104], to name a few. As sociologists developed these concepts, they became aware that graph theory, a somewhat forgotten branch of mathematics at that time, provided the necessary tools to describe social networks mathematically [105, 106].

Starting from the famous Königsberg Bridge Problem by Euler, graph theory studied a wide range of topics but was restricted, in most cases, to more theoretical questions [107]. Probably one of the most famous problems in graph theory is the *colouring problem* since it was the first theorem to be proven with the assistance of a computer program [108, 109]. Graphs in the mathematical language were identical to what sociologists referred to as net-

works. Accordingly, sociologists could leverage the formalism to describe different properties of their empirically built networks such as centrality measures [110], clustering coefficients [111], or community structure [112]. As we previously mentioned, epidemiologists started to realise the importance of networks in the context of the AIDS epidemic. Since the AIDS epidemic mainly affected vulnerable groups such as sex workers, men who have sex with men, or drug users, many of these early works were right at the intersection between public health and sociology [90, 92, 96]. It turned out that the centrality measures built by sociologists gave powerful insight into the drivers behind the spread of HIV [90, 96]. This epiphany led epidemiologists to use the inferred contact networks to study the spread of diseases on top of them through computer simulations [113–115]. However, with few exceptions [116], the analytical description for the spread of epidemics on networks was not developed by trained epidemiologists but by statistical physicists, who worked in the then relatively young field of *complex networks* [117, 118].

Finding a definition for the term complex networks is not trivial. Generally, the attribute *complex* can be understood as describing networks that are not as regular as a lattice [119] but are more structured than random graphs, for which all connections are built randomly and with equal probability [120]. Most empirical networks live in this space between lattices and random graphs [82], which is best illustrated by one of the earliest contributions in complex networks, the *small-world model* by Duncan Watts and Steven Strogatz [121]. The small-world model starts with a lattice and randomly rewires the links with a given probability, continuously interpolating between a lattice and a random graph. The networks created from their model showed clustering (transitivity) while still exhibiting the "small-world" property [122] – generally known as *the six degrees of separation*. In contrast to most empirical networks, previous network models from graph theory could not exhibit both properties simultaneously.

As illustrated by the work of Watts and Strogatz, research in complex networks generally takes a less formal and more empirical approach than graph theory. This more hand-wavy approach can be partially explained by the vast amount of physicists that worked early on in the field of complex networks. There are at least two reasons why so many statistical physicists turned their attention towards networks in the late 90s and early 2000s. First, the study of structure is inherent to solid-state physics. It started with periodic crystalline structures [123] and evolved towards disordered systems, where network-like structures were considered early on [124]. Second, in the early days of

the field of complex networks, empirical evidence was found that the degree distribution of real-world networks followed a power law [125, 126]. This finding suggested that some universal mechanism could explain the ubiquity of power law degree distributions [127, 128]. Moreover, universality was a particularly attractive concept to physicists back then. With the discovery of the rich topological structure of empirical networks, physicists quickly jumped to analyse how this may affect different dynamical processes [129–131]. Among those processes was epidemic spreading, and, as previously mentioned, the field of complex networks was mainly responsible for developing an analytical framework to describe spreading processes on networks [65, 131, 132]. It was quite a journey between sociology, epidemiology, and graph theory that eventually led to physicists modelling the spread of epidemics on networks. In the following, we will detail some of the frameworks they developed.

2.2.3.1 Mathematical description of a network

Networks can be described by a set of entities or *nodes* that are connected through *edges* [82]. Nodes and edges can be anything from web pages, and hyperlinks [133], proteins and chemical reactions [134], power stations and transmission lines [135], or, in the case of an epidemic, individuals and physical interactions [136]. The empirical measurement of these networks is not always straightforward, and interactions are often defined with respect to a threshold, which, in some cases, can fundamentally alter the inferred structure [128]. However, in mathematical terms, the definition of a network is straightforward. Interactions are encoded in the adjacency matrix A, whose elements A_{ij} are defined as

$$A_{ij} = \begin{cases} 1 \text{ if node } i \text{ is connected to node } j \\ 0 \text{ otherwise} \end{cases} . \tag{2.42}$$

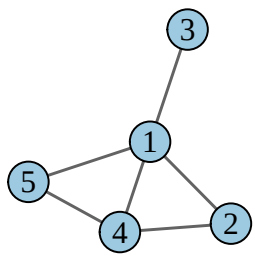
The size of A corresponds to the number of nodes in the network. Fig. 2.4 shows an example of a network and its corresponding adjacency matrix.

If interactions are bidirectional, as in Fig. 2.4, the adjacency matrix is symmetric ($A_{ij} = A_{ji}\ \forall i, j$) and the network is referred to as undirected. With few exceptions [137, 138], contact networks for epidemic models are undirected since the virus can spread in both directions. We refer to the number of connections of node i as its degree k_i, which can be expressed as $k_i = \sum_j A_{ij}$. Networks are not necessarily binary, i.e. either two nodes interact with each

other, or they do not. The connections (edges) can also take continuous values $A_{ij} = w_{ij}$ (weights), which makes the network weighted. In this case, the sum over the weights over a node is its strength $s_i = \sum_j w_{ij}$. Epidemic models can consider a weighted adjacency matrix since interactions between individuals are not all equally intense and frequent [139]. However, probably for convenience, most of the time, researchers work with binary networks. The different degrees of the nodes lead to a degree distribution p_k of the network. Further, there are a series of centrality measures for the nodes that can be directly calculated from the network's adjacency matrix. This thesis will not focus on these centrality measures, and we refer the reader to Ref. [82]. Equally, there are a series of extensions such as multilayer networks [140] or higher-order networks that are not covered in this thesis [141].

In the context of epidemics, a vast amount of empirical contact networks have been collected in the past. While networks were first inferred from surveys [94, 142], the technological advancements later allowed automatised contact collection through Bluetooth devices [143–146], and cell phones [147–149]. Many of these empirical networks also evolve in time and are thus *temporal networks*. In this case, an adjacency matrix exists for every time-step t denoted $A(t)$. We have previously mentioned that these contact networks can be used to simulate the spread of an epidemic over the network [132]. However, if one wants to focus on a concrete topological aspect, researchers often employ synthetic network models that allow for varying a specific network property. We will present two network models here, the Erdős-Rényi model [150] and the Barabási-Albert [127] model.

The Erdős-Rényi model assumes that edges are created randomly. More specifically, given a network of N nodes, edges exist with probability p among them. It is the standard recipe for creating a random graph and is intrinsically



$$A = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \end{pmatrix}$$

Figure 2.4: Exemplary network with the corresponding adjacency matrix A. The node's label refers to the ordering for the entries in A.

linked to a percolation process [120]. The regular mechanism to form edges leads to a relatively homogeneous degree distribution that follows a binomial distribution

$$p_k = \binom{N-1}{k} p^k (1-p)^{N-1-k}, \tag{2.43}$$

with mean degree $\langle k \rangle = p(N-1) \approx pN$ and second moment $\langle k^2 \rangle = \langle k \rangle (1-p) + \langle k \rangle^2$. We see that the degree distribution is not very heterogeneous. For a large average degree, we have $\langle k^2 \rangle / \langle k \rangle^2 \approx 1$. As the system size increases, the degree distribution converges to a Poisson distribution as in general for a binomial distribution.

However, empirical evidence has shown that many real-world networks are more heterogeneous than the Poisson distribution and even have a degree distribution that follows, or at least can be approximated by, a power law, i.e. $p_k \sim k^{-\gamma}$ [151]. The first evidence for such power law was found for the World-Wide Web [125] in 1999. More precisely, the distribution of the number of outgoing and incoming hyperlinks of websites was shown to be scale-free. At the same time, the Internet topology was found to follow a power law [152]. Power law distributions $f(k)$ are referred to as scale-free since they are scale-invariant, i.e. $f(\alpha k) \sim f(k)$. For this reason, Réka Albert and Albert-László Barabási developed the *preferential attachment* model, today most known as the Barabási-Albert model, as a universal mechanism to describe the emergence of degree distributions that follow a power law [127]. The mechanism assumes that nodes with a fixed number of edges are sequentially added to the system. As nodes enter the system, they connect with the nodes already present with a probability proportional to their degree. Interestingly, this is a special case of the Price model, which, published in 1976, describes cumulative advantage processes [153]. The motivation of Derek De Solla Price was to explain the heterogeneous distribution of the frequency of publication by authors in a given field, which follows approximately a power law. This empirical observation, known as Lotka's law, was first made in 1926 by Alfred J. Lotka [154], the scientist that also first defined the reproduction number (Sec. 2.1.4). Barabási-Albert showed that preferential attachment resulted in a degree distribution that follows a power law with exponent 3, i.e. $p_k \sim k^{-3}$. The resulting degree distribution implies that few nodes (hubs) participate in most links. In contrast, most nodes (leaves) have only a few connections. Small changes in the edge formation rule also allow for fine-tuning of the exponent of the power law [155].

Simultaneously, as scale-free networks were discovered at the turn of the millennium, computer viruses became a focus as a substantial threat to the digital infrastructure and the functioning of our society. In this context, it was proposed to leverage the tools developed in epidemiology to describe the spread of computer viruses [156]. So it came that two physicists, Romualdo Pastor-Satorras and Alessandro Vespignani, developed a mathematical framework to describe the impact of the scale-free degree distribution on the spread of computer viruses [157].

2.2.3.2 Annealed network formalism

Since the entirety of the World-Wide Web or the Internet was not available as a dataset (besides being computationally impossible to handle), Pastor-Satorras and Vespignani decided to work solely with the degree distribution. Similarly to the stratification with respect to the different contact rates (Sec. 2.2.2), their model separates the population into different degree classes. The variables $x_k(t)$ denote the probability that a node (individual) with degree k is infected at time t. For the contact structure between the different degree classes, Pastor-Satorras and Vespignani assumed proportionate mixing. In network science, this approach is often referred to as *annealed network* formalism, or *heterogeneous mean-field* formalism [131]. The annealed network formalism assumes that the time scale at which the network evolves is much smaller than the epidemic. Accordingly, it is a reasonable assumption that nodes interact randomly but with a frequency proportional to their degree [131]. The formalism is then called annealed in analogy to annealed disorder in physics [119]. Since the study also assumes an SIS model, the model equations are equivalent to the ones presented in Eq. (2.35) and read

$$\dot{x}_i(t) = -\mu x_i(t) + (1 - x_i(t))\lambda k \sum_l \frac{l p_l}{\langle k \rangle} x_l(t)\,. \tag{2.44}$$

The overall prevalence I in the population can then be expressed as $I = \langle x_k \rangle \sum_l p_l x_l$. One can see that the framework is equivalent to the one introduced in Sec. 2.2.2 to describe how heterogeneous activity levels affect the transmission of sexually transmitted diseases. However, it appears that Pastor-Satorras and Vespignani were not aware of the respective literature in epidemiology when they developed the heterogeneous mean-field formalism [157, 158]. More precisely, they refer to the contact-process model [159] as

the starting point from where they developed the model. Nevertheless, they arrived at the same model equations as epidemiologists some years earlier.

However, in contrast to epidemiologists, Pastor-Satorras and Vespignani took the empirical finding of a scale-free degree distribution literally and to the extreme. As we previously mentioned, the dynamics described in Eq. (2.44) leads to a critical transmission probability λ_c given by

$$\lambda_c = \mu \frac{\langle k \rangle}{\langle k^2 \rangle} . \tag{2.45}$$

Now, to calculate the moments of the power law distribution ($p(k) \sim k^{-\gamma}$), we interchange the sum into an integral with the boundaries k_{min} and k_{max}. The maximal degree in the network is $N - 1 \approx N$, where N is the number of nodes in the network. Accordingly, we fix $k_{max} = N$ and the moments of the degree distribution scale as

$$\langle k^m \rangle \sim k^{m-\gamma+1}|_{k_{min}}^{N} . \tag{2.46}$$

In other words, the second-moment scales with respect to the system size as $\langle k^2 \rangle \sim N^{3-\gamma}$. Therefore, if $\gamma < 3$ [1], the epidemic threshold λ_c vanishes in the thermodynamic limit ($\lim_{N\to\infty} \lambda_c = 0$). This divergence implies that, in such a network, a virus can survive independently of the transmission probability and the infectious period, here $1/\mu$. However, it is worth mentioning that the epidemic threshold only vanishes in the thermodynamic limit. In the context of computer networks, the thermodynamic limit is a useful concept due to the huge size of the network and the low cost of making a connection. Nevertheless, since all empirical networks are finite, they will have a non-zero epidemic threshold. Additionally, many real-world networks include some physical limits that make it impossible for the maximal degree to scale with the system size. For example, it was shown that the number of sexual interactions follows a power law [126]. Logically, the number of sexual partners cannot increase indefinitely since an individual has only a finite amount of available time to have sexual intercourse. In these cases, the distributions are not pure power laws, but upper truncated power laws or power laws with a sharp cutoff [151]. Despite the hype in the early 2000s, it is quite contested today whether scale-free networks actually are so ubiquitous in nature [160–164].

1 In the case $\gamma = 3$, the second moment also diverges but with $\ln(N)$.

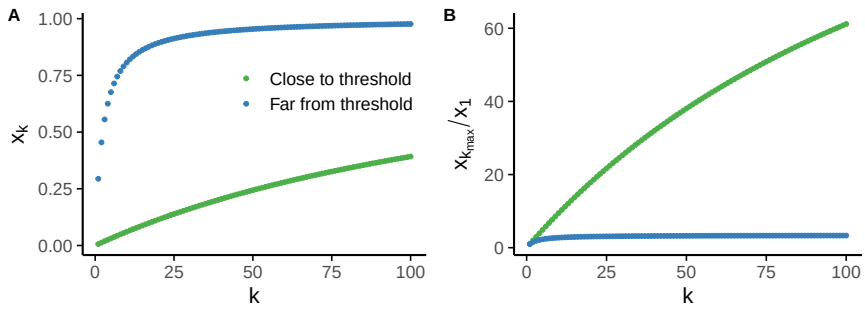


Figure 2.5: **(A)**: Dependence of the infection probability x_k on the degree of a node k. For the green dots, the system is close to the epidemic threshold ($\lambda = 0.03$). In contrast, the blue dots represent a system far from the epidemic threshold ($\lambda = 0.1$). Overall prevalence is 1% for the green vs 30% for the blue. The recovery probability was fixed as $\mu = 0.1$. The degree distribution is a power law with exponent 3. **(B)**: Same difference between green and blue dots, but this time the relative variation in x_k is shown, i.e. $x_{k_{max}}/x_1$ with $k_{max} = 100$.

Setting aside the question of whether heterogeneous contact networks are scale-free, it is true that the unequal distribution of contacts impacts the ability of a disease to propagate substantially. The network's hubs drive the spread of the disease due to their high number of contacts and corresponding increased infection risk. The increased infection risk with the number of contacts can be seen when considering the equilibrium condition of Eq. (2.44), which gives an expression of x_k as

$$x_k = \frac{\frac{\lambda}{\mu}\frac{k}{\langle k\rangle}z}{1+\frac{\lambda}{\mu}\frac{k}{\langle k\rangle}z}\,. \tag{2.47}$$

The quantity z is defined as in Sec. 2.2.2. Accordingly, from Eq. (2.47), the infection risk of a node increases with its degree as

$$\frac{\mathrm{d}x_k}{\mathrm{d}k} = \frac{\frac{\lambda}{\mu}\frac{z}{\langle k\rangle}}{\left(1+\frac{\lambda}{\mu}\frac{k}{\langle k\rangle}z\right)^2}\,. \tag{2.48}$$

These calculations mathematically show how the infection probability increases with the degree of a node. While the increase for low k is almost linear, as k increases further, the derivative saturates, and the infection risk only varies minimally. The same can be observed considering the expression

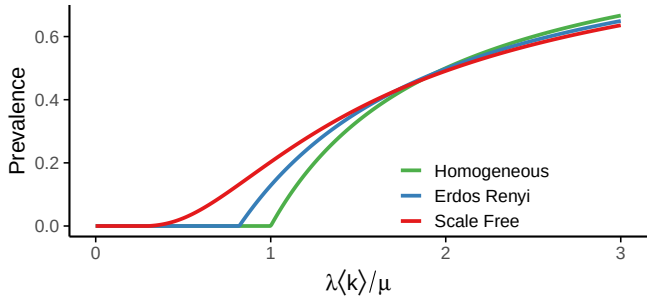


Figure 2.6: Prevalence as a function of the transmission probability λ (rescaled) for a homogeneous population, an Erdös-Rényi (random) network and a scale-free networks. All networks have the same mean degree. Prevalence was obtained by numerically solving Eqs. (2.44).

of x_k in Eq. (2.47) directly. The variation of x_k is shown in Fig. 2.5A. However, as predicted by Eq. (2.48), the variation in x_k saturates much slower close to the epidemic threshold (low z) and the difference in the infection probability between hubs (high k) and leaves (low k) is much more pronounced (Fig. 2.5B). This pronounced heterogeneity in the infection probability allows the disease to stay endemic for much lower values of the transmissibility λ.

To illustrate how the heterogeneous network structure facilitates the spread of a pathogen, we highlight, in Fig. 2.6, how the increasing heterogeneity from a homogeneous population ($p_k = \delta_{\langle k \rangle j}$), to a random graph ($p_k = 1/\lambda e^{-\lambda k}$), and to a scale-free graph ($p_k = k^{-\gamma}$) continuously decreases the epidemic threshold. In contrast, as the system is further away from the epidemic threshold, prevalence is maximal for homogeneous mixing due to the absence of leaves with a low infection probability.

Despite the insight it gives, the annealed network approach neglects many topological features of the network and focuses only on the degree distribution. The assumption of the much faster-evolving network than the epidemic and proportionate mixing reintroduces the law of mass action (assumption 3). The contrast of an annealed disorder in physics is quenched disorder, where the structure does not evolve, as, for example, a Spin Glass [124]. In the following, we will take the same approach and describe how the dynamics evolve on top of a fixed network (adjacency matrix).

2.2.3.3 Quenched mean-field dynamics

The quenched mean-field approach assumes that the epidemic evolves much faster than the contact structure [165, 166]. Accordingly, we can assume a constant adjacency matrix A_{ij} on top of which the disease propagates. Instead of stratifying the population with respect to the different degree classes, the approach describes the infection probabilities x_i at the individual level for every node i. In this case, assuming an SIS model, the dynamics are described by the following equations

$$\dot{x}_i(t) = -\mu x_i(t) + \lambda(1 - x_i(t)) \sum_{j=1}^{N} A_{ij} x_j(t) \,. \tag{2.49}$$

The second term guarantees that node i only interacts with its neighbours in the network. The prevalence in the population is given by $I = \sum_i x_i / N$. In order to find the epidemic threshold, we develop around the disease-free equilibrium $x_i \approx \epsilon$. Accordingly, upon rescaling time, Eq. (2.49) becomes

$$\dot{x}_i(t) = \sum_{j=1}^{N} \underbrace{\left(-\delta_{ij} + \frac{\lambda}{\mu} A_{ij} \right)}_{J_{ij}} x_j \,, \tag{2.50}$$

where J_{ij} is the Jacobian of the dynamics. The disease-free equilibrium loses stability when the largest eigenvalue of J becomes positive [50]. However, this is equivalent to the condition $\Lambda_{max}\lambda/\mu \geq 1$, where $\Lambda_{max}(A)$ is the leading eigenvalue of the adjacency matrix A. This condition leads to a critical transmission probability (epidemic threshold) λ_c given by

$$\lambda_c = \frac{\mu}{\Lambda_{max}(A)} \,. \tag{2.51}$$

One can show that for large, uncorrelated random networks, the largest eigenvalue of the adjacency matrix is given by $\langle k^2 \rangle / \langle k \rangle$ [167, 168]. Hence, one recovers the results of the annealed formalism (Eq. (2.45)). To build such uncorrelated networks, one fixes the degree distribution and samples a degree sequence, i.e. the degrees of all nodes i. The stubs of the nodes are then randomly connected effectively following proportionate mixing. There are some subtleties regarding self-loops and multiple edges for small graphs, but it works overall pretty well. The method is referred to as the *configuration model* [82].

The quenched mean-field equations give a relatively accurate description when compared with numeric simulations [131], which are generally referred to as Monte Carlo methods [169, 170]. However, the continuous-time evolution of the dynamics makes the comparison with the numerical simulations more complex. To simulate the continuous time process, one generally uses the Gillespie algorithm that was introduced to simulate chemical reactions [171, 172]. The algorithm calculates the distribution of what and when future events occur, samples from this distribution and correspondingly updates the system. This continuous time approach is closest to the actual nature of a spreading process. However, since the Gillespie is computationally costly, researchers often rely on a discrete-time description, which is easier to simulate. In the discrete-time simulations, for every time step, infected individuals recover with a probability μ and infections are transmitted along links between infected and susceptible individuals with probability λ. The simplicity of the approach, and the corresponding low computational cost, made the discrete-time description particularly appealing. The widespread use of computer simulations motivated researchers also to provide a corresponding analytical description for the discrete-time process [173–175], which will be presented in the next section.

2.2.3.4 Microscopic Markov Chain Approach

The first discrete-time model of an epidemic was already published in 1889 by the Russian physician Pyotr Dimitrievich En'ko [176]. Since the text was published in Russian, it did not receive much attention and was only rediscovered by Klaus Dietz in 1989. However, even though En'ko did not consider a network, his approach is similar to today's discrete-time descriptions. The first analytical framework to describe the discrete-time infection dynamics on networks was proposed by Wang et al. [173, 174]. Gómez et al. further extended the theory and solidified it under the name Microscopic Markov Chain Approach (MMCA) [175, 177]. The description is microscopic since – as in the quenched network formalism – it provides equations for the evolution of the individual infection probabilities of node i at time step t, generally denoted as $p_i(t)$ for discrete-time models. One can understand the dynamical equations as a Markov Chain for the individual probabilities $p_i(t)$. For an

adjacency matrix A of size $N \times N$, a recovery probability μ and transmission probability λ, the dynamics of the SIS model read

$$p_i(t+1) = p_i(t) - \mu p_i(t) + (1 - p_i(t))(1 - q_i(t)) \,, \tag{2.52}$$

where $q_i(t)$ is the probability that node i is not infected by any of its neighbours during time step i and is expressed as

$$q_i(t) = \prod_{j=1}^{N} (1 - \lambda A_{ij} p_j(t)) \,. \tag{2.53}$$

The overall prevalence can then be written as $I(t) = \sum_i p_i(t)/N$. It is a microscopic Markov Chain as it does not consider the possible 2^N (each node either susceptible or infected) macroscopic states of the system. Such a macroscopic description would be computationally much more costly [178, 179]. In contrast, the microscopic approach requires only $N(\Omega - 1)$ equations, where Ω is the number of compartments of the model. The reduction in complexity is possible since the MMCA approach, as well as the quenched mean-field formalism, assumes independence between the infection probability of two nodes i and j. The states of the nodes are actually Bernoulli variables $X_i(t) \in \{0, 1\}$, where $X_i(t) = 0$ ($X_i(t) = 1$) if node i is susceptible (infected) at time t. Accordingly, from a mathematical perspective [131], the independence assumption is expressed as

$$\mathbb{E}[X_i X_j] = \mathbb{E}[X_i]\mathbb{E}[X_j] = p_i(t) p_j(t) \;\forall\; i \neq j \,, \tag{2.54}$$

where $\mathbb{E}[X]$ refers to the expected value of the random variable X. The state of node i is not independent of the state of node j, particularly if the two nodes are neighbours. However, it turns out that the dynamical correlations are not very persistent and, despite the independence assumption, the MMCA provides a very good match with Monte Carlo simulations, even on empirical networks [175]. There exist also extensions of the formalism which take into account correlations along edges [180], or even among cliques of size 3 (triangles) [181, 182] that are particularly relevant in the case of higher-order dynamics [183].

Close to the epidemic threshold, assuming $p_i \approx \epsilon$, we effectively recover Eq. (2.50) by linearising Eq. (2.53) and plugging it into Eq. (2.52). The only difference is that, due to the discrete-time approach, there is not a derivative but a difference $p_i(t+1) - p_i(t)$. However, it leads to the same stability

criteria of the disease-free equilibrium, and thus also to the same epidemic threshold λ_c of Eq. (2.51). Furthermore, the quenched mean-field formalism can be recovered by considering shorter time intervals. Decreasing the time intervals effectively leads to a smaller value of the transmission probability λ as contacts are of shorter duration. Then, in the continuous time limit for small λ, one recovers Eq. (2.49). The formalism can also be easily adapted to incorporate a time-varying network. One simply replaces the constant adjacency matrix A with the time-dependent adjacency matrix $A(t)$. Only the calculation of the epidemic threshold is somewhat challenging since it requires periodicity [184]. The analysis becomes more subtle if one assumes the adjacency matrix to vary in continuous time [185].

We have provided various descriptions of the epidemic dynamics on networks. In particular, we have presented all the formalisms employed in this thesis. However, for the sake of completeness, the *pair approximation* [132, 186] and the *percolation approach* need to be mentioned. The pair approximation consists of a link-based description of the dynamics and relies on moment closure approximations [187]. It is essentially the continuous time pendant of Ref. [180]. Percolation theory, combined with a message-passing approach [124], can be used to estimate the final attack rate of an SIR model. Early research was able to map the SIR model to a percolation process [188, 189], which has later been slightly rectified [190]. This overview completes the existing description of epidemics on contact networks. In the following section, we will present approaches to incorporating mobility networks into the problem of epidemic spreading.

2.2.4 *Modelling the spread of epidemics through human mobility*

Until now, we have considered how the actual contact structure, represented by the adjacency matrix, can be incorporated into an epidemic model. However, besides very recent advances in data collection [148, 149, 191], empirical networks were generally limited to a small number of participants, and a specific locality [143–147]. Accordingly, the available networks are of good use to model the spread of a pathogen inside a small, geographically fixed population. Sadly, these empirical networks do not suffice to describe the propagation of diseases between regions, countries or continents. At this larger scale, diseases are imported through human movement.

One of the most well-known examples where a disease was imported through human movement is the introduction of smallpox to the Americas in the early 16th century [7]. Since the Aztecs completely lacked immunity to smallpox, a substantial part of the population deceased after entering into contact with the men of Hernán Cortés, and he was able to conquer the Aztec land with a relatively small army. Nevertheless, during this period, the geographic propagation of diseases was relatively slow due to the limitation of the existing means of transportation. For example, for the bubonic plague, often referred to as the Black Death, the newest evidence suggests that it took almost 10 years for the virus to make its way along the silk road to western Europe from its origin in the Tian Shan region in Centralasia [13]. In contrast, today, due to the intense connectivity of the world, pathogens can spread in only a few weeks around the globe, as we experienced during the SARS-CoV-2 pandemic [192–195]. The connections at different scales (planes, trains, ships, buses, cars) [196] effectively rearrange the geometry on which the disease propagates [197].

Due to the altered geometry, a diffusive spatial description of disease propagation is inadequate. For this reason, researchers employ the so-called *metapopulation models* that, first developed in ecology [198, 199], stratify the population into different geographic locations and describe the interactions among them [59, 200, 201]. The basic framework is very similar to the one we introduced in Sec. 2.2.1 for the age stratification. The main difference is that the index i refers to different locations, called *patches* or *subpopulations*, and the contact matrix is now a mobility matrix R that controls the interactions, i.e. mobility between patches. The subpopulations can be anything from municipalities to regions or even entire nations. These approaches often assume a homogeneous mixing inside patches, but more sophisticated solutions exist [202]. The mobility matrix is nowadays generally estimated from cell phone data [203–206]. Assuming the same number of contacts in every patch, the entry R_{ij} corresponds to the proportion of time spent in patch j by individuals residing in patch i. Many algorithms exist to infer which patch an individual resides in [207]. Further, assuming that the number of contacts of individuals from patch i in patch j is proportional to the time spent there is a rather crude approximation since individuals have different types of contacts and distinct contact frequencies depending on why they spent time in another patch. Many such subtleties can and sometimes should be considered when inferring the mobility matrix R from raw data [204, 206].

Beyond cell phone data, one can often also use census surveys to determine the daily commuting patterns in a given country, or region [204]. The census surveys generally allow us to estimate recurrent mobility as the number of individuals that reside in patch i and work in patch j. This estimation can then be encoded in a normalised mobility matrix R, whose elements R_{ij} define the fraction of individuals living in patch i and working in patch j. Gomez-Gardeñes et al. [208] extended the MMCA to describe the impact of recurrent mobility patterns on the epidemic dynamics. We will use this approach to describe the initial spread of SARS-CoV-2 in Spain, presented in Chapter 5. In this MMCA approach, variables $\rho_i(t)$ do not refer to the probability of an individual i being infected, but rather to the probability of an individual *residing in patch i* being infected. Assuming an SIS compartmental model, the recurrence equations can be written as

$$\rho_i(t+1) = \rho_i(t) - \mu\rho_i(t) + (1-\rho_i(t))\Pi_i(t) \tag{2.55}$$

where $\Pi_i(t)$ represents the probability that a susceptible individual residing in patch i gets infected during time step t. This quantity $\Pi_i(t)$ is expressed as

$$\Pi_i(t) = (1-p)P_i(t) + p\sum_j R_{ij}P_j(t)\,. \tag{2.56}$$

The mobility parameter p controls the fraction of the population that commutes every day. It was introduced more with a theoretical interest to study explicitly the role of mobility. Given a real-world empirical mobility network, one can fix $p = 1$. The variable $P_i(t)$ describes the probability of a susceptible individual getting infected when staying in its residence patch i. It is expressed as

$$P_i(t) = 1 - \left(1 - \lambda\frac{\tilde{I}_i(t)}{\tilde{n}_i(t)}\right)^{k_i}\,. \tag{2.57}$$

Here, k_i is the average number of contacts in patch i. The number of contacts k_i can depend on different demographic and socio-economical factors. The quantity $\tilde{n}_i$ ($\tilde{I}_i(t)$) is the effective (infected) population in patch i (at time t).

Given the number of inhabitants n_i in patch i, these quantities are expressed as

$$\tilde{n}_i = \sum_j n_j \left[(1-p)\delta_{ij} + pR_{ji}\right] \tag{2.58}$$

$$\tilde{I}_i(t) = \sum_j n_j \rho_j(t) \left[(1-p)\delta_{ij} + pR_{ji}\right] . \tag{2.59}$$

The framework is a movement-interaction-return model. Individuals move to another patch according to the mobility matrix R, interact and return to their origin. If they move, individuals never interact in the home patch, except for individuals working in their patch of origin (diagonal entry R_{ii}). Interaction with the home patch only takes place implicitly since everyday individuals that move from patch i to patch j are chosen randomly. Extensions of the model exist, which substantially increase model complexity, where the same individuals move every day from patch i to patch j, and interactions take place in the patch of origin as well in the patch of destination [209, 210].

Now, as many times before, we can linearise the equations around the disease-free equilibrium to find the epidemic threshold ($\rho_i \approx \epsilon$). The linearised form of the equations is particularly instructive since the impact of mobility becomes apparent. To linearise the equations, we will assume a density-dependent formulation, where in every patch, all the individuals interact with each other (all-to-all), i.e. $k_i = \tilde{n}_i$. With these assumptions, we find at equilibrium the following equation for the difference $\Delta_i = \rho_i(t+1) - \rho_i(t)$:

$$\Delta_i = \left(-\delta_{ij} + \frac{\lambda}{\mu} \sum_j \underbrace{n_j \left[(1-p)^2\delta_{ij} + p(1-p)\left(R_{ij} + R_{ji}\right) + p^2 \sum_l R_{il}R_{jl} \right]}_{M_{ij}} \right) \rho_j . \tag{2.60}$$

Please note that we have dropped the time dependence since we consider the equilibrium condition. The three summation terms of the matrix M encode the three possible interactions between individuals residing in patch i and patch j. In the first term, both individuals reside in the same patch ($i = j$), and they do not move ($(1-p)^2$). In the second term, one individual moves to the patch ($R_{ij} + R_{ji}$), in which the other resides and stays ($p(1-p)$). In the third term, both individuals move (p^2), and they meet in any patch l ($R_{il}R_{jl}$). In the most often used approaches, which use continuous time and the mobility matrix R is defined such that $p = 1$ (time spent in other patches), only the

third term appears in the model equations allowing for a compact and elegant formulation [204]. The above equation then leads to the following threshold condition $(\lambda\Lambda_{max}(M))/\mu > 1$. Previously, we have shown that, for the simple SIR model, the stability criterion of the disease-free equilibrium coincides with the expression of the basic reproduction number. In the following, we will see that the same holds even when heterogeneities are present in the model, as shown in this Sec. 2.2.

2.2.5 *The next-generation matrix approach*

The approach we have been using so far to calculate the epidemic threshold consisted in calculating the stability of the disease-free equilibrium through the Jacobian of the dynamics, as it is the standard in dynamical system theory [50]. Mathematically speaking, this corresponds to a dynamical system following a set of first-order differential equations

$$\dot{x}_i(t) = f_i(x(t))\,, \tag{2.61}$$

where the property $f(0) = 0$ guarantees the existence of the disease-free equilibrium. According to dynamical system theory, the disease-free equilibrium loses stability when the largest eigenvalue of the Jacobian matrix J evaluated at $x = 0$, $J(0)$, becomes positive. Accordingly, the disease will actually spread after introducing a primary case in the population. Previously, we have seen that the condition $\Lambda_{max}(J(0)) > 0$ is equivalent to $R_0 > 1$ in the SIR model. Now, it turns out that the same holds in heterogeneous populations. This equivalence was proven through the construction of the *next generation matrix* (NGM) in 1990 by Diekmann et al. [211–214]. First, let us realise that evaluating the Jacobian J at the disease-free equilibrium $x = 0$ is the same as linearising Eq. (2.61) around the disease-free equilibrium ($x \approx \epsilon$). We used this equivalence many times in the models presented before without explicit mention. Accordingly, in its linearised form, Eq. (2.61) can be written as

$$\dot{x}(t) = J(0)x\,, \tag{2.62}$$

where we dropped the index i to simplify the notation. Diekmann et al. proposed to separate $J(0)$ into a transmission part T and a transition part Σ such that $J(0) = T + \Sigma$. The transmission part T refers to the entries in the Jacobian that account for infections taking place. The transition part Σ refers

to the transition between compartments without new infections. To illustrate this better, let us consider the example for the standard SIR model, whose equations are given in Eqs. (2.17) & (2.18). We can drop the redundant equation for S due to the conservation equation $S + I + R = 1$. In this case [2], the infection terms and transition terms are kSI and $\pm\mu I$, respectively. With respect to the Jacobian, this means $J(0)$ can be expressed as

$$J(0) = \begin{pmatrix} \lambda k - \mu & 0 \\ 0 & \mu \end{pmatrix} = \underbrace{\begin{pmatrix} \lambda k & 0 \\ 0 & 0 \end{pmatrix}}_{T} + \underbrace{\begin{pmatrix} -\mu & 0 \\ 0 & \mu \end{pmatrix}}_{\Sigma} . \tag{2.63}$$

Diekmann et al. then propose to rescale "time" by Σ^{-1}. Effectively, this is equivalent to the rescaling of time in Sec. 2.1.4 by the infectious period $\tau_I = 1/\mu$, allowing to pass from the time-dependent description of the SIR model to a description based on subsequent generations. The NGM approach generalises in the case of a more complex distribution of infectious compartments. Eq. (2.62) then becomes

$$\dot{x}(t) = -\left(T\Sigma^{-1} + 1\right) x\,. \tag{2.64}$$

The minus sign appears here since the matrix Σ is negative signed. According to the above equations, the condition of the epidemic threshold is $\Lambda_{max}(T\Sigma^{-1} + 1) > 0$, which is equivalent to $\Lambda_{max}(-T\Sigma^{-1}) > 1$. The matrix $\mathrm{NGM} = -T\Sigma^{-1}$ is referred to as the next generation matrix, and its largest eigenvalue, i.e. its spectral radius, coincides with the basic reproduction number of the disease. For the SIR model, we have

$$\mathrm{NGM} = \begin{pmatrix} \frac{\lambda k}{\mu} & 0 \\ 0 & 0 \end{pmatrix} , \tag{2.65}$$

and we see that $\Lambda_{max}(\mathrm{NGM}) = (\lambda k)/\mu = R_0$. In this example, the expression of the NGM in Eq. (2.65) has only one non-trivial dimension. Generally, only infectious compartments need to be included in the construction of the NGM. In other words, we could have directly excluded the compartment R. Even then, sometimes dimensions appear, mostly with a zero row, that do not con-

2 Please note that the variables S, I and R correspond here to the fraction of susceptible, infected and recovered individuals. Previously, in Sec. 2.1.3, we used the distinction with capital and small letters only to make explicit the description in absolute numbers and the one working with proportions.

tribute to the largest eigenvalue. One can simply exclude the dimensions in these cases since the leading eigenvalue will not be affected. We do not prove here that $R_0 = \Lambda_{max}(\mathrm{NGM})$ holds for all compartmental models. We refer the reader for this to Ref. [214]. However, to better illustrate the inner working of the NGM approach, let us consider SIR-like dynamics with two distinct groups. The two groups interact according to a contact matrix K, with distinct recovery probabilities μ_1 and μ_2, respectively. We thus consider two separate infectious compartments I_1 and I_2, as well as susceptible compartments S_1 and S_2. Accordingly, the model equations can be written as

$$\dot{S}_1 = -\lambda k_{11} I_1 S_1 - \lambda k_{21} I_2 S_1 \tag{2.66}$$
$$\dot{S}_2 = -\lambda k_{22} I_2 S_2 - \lambda k_{12} I_1 S_2 \tag{2.67}$$
$$\dot{I}_1 = -\mu_1 I_1 + \lambda k_{11} I_1 S_1 + \lambda k_{21} I_2 S_1 \tag{2.68}$$
$$\dot{I}_2 = -\mu_2 I_2 + \lambda k_{22} I_2 S_2 + \lambda k_{12} I_1 S_2 \tag{2.69}$$
$$\dot{R} = \mu_1 I_1 + \mu_2 I_1 \,. \tag{2.70}$$

The quantities k_{ij} are the entries $(K)_{ij}$ of the contact matrix K and define the contact rate between group i and group j. To construct the NGM matrix, we focus only on the infectious compartments I_1 and I_2. Accordingly, we have

$$T = \lambda \begin{pmatrix} k_{11} & k_{21} \\ k_{12} & k_{22} \end{pmatrix} , \; \Sigma = \begin{pmatrix} -\mu_1 & 0 \\ 0 & -\mu_2 \end{pmatrix} . \tag{2.71}$$

We thus see that the matrix $(-\Sigma^{-1})_{ij} = \tau_i \delta_{ij} = 1/\mu_i \delta_{ij}$ is defined by the infectious periods of the two population groups. More generally, the entry $(-\Sigma^{-1})_{ij}$ corresponds to the expected time an individual currently in state i will spend in state j [214]. We then find the NGM matrix as

$$\mathrm{NGM} = \begin{pmatrix} \frac{\lambda k_{11}}{\mu_1} & \frac{\lambda k_{21}}{\mu_2} \\ \frac{\lambda k_{12}}{\mu_1} & \frac{\lambda k_{22}}{\mu_2} \end{pmatrix} = \begin{pmatrix} R_{11} & R_{21} \\ R_{12} & R_{22} \end{pmatrix} . \tag{2.72}$$

If we assume proportionate mixing, i.e. that individuals interact in proportion to their contact rates as detailed in Sec. 2.2.2 & 2.2.3.2, the entries k_{ij} are given by $k_{ij} = k_j k_i p_i / \langle k \rangle$, where p_i refers to the fraction of the population that belongs to group i. Making these quantities explicit and assuming $\mu = \mu_1 = \mu_2$, it is straightforward to show that one recovers the previously found epidemic threshold, i.e. $R_0 = \lambda \langle k^2 \rangle / (\mu \langle k \rangle) \geq 1$.

Further, we see that the entry $(\mathrm{NGM})_{ij}$, here denoted as R_{ji}, is nothing else than the number of secondary cases generated in group j by an infected individual, the primary case, in group i. Hence, the multiplication of the NGM with a two-dimensional vector y, whose entries define the number of infected individuals in each of the two groups, describes how generations reproduce [215]. In other words, the new infections per generation g will grow as

$$y(g+1) = \mathrm{NGM}\, y(g)\,. \tag{2.73}$$

Accordingly, any vector y will rapidly align with the normalised eigenvector $\hat{y}$ associated with the leading eigenvalue R_0, and the case distribution among the present groups will thus follow $\hat{y}$. By definition, we have the relation $R_0\hat{y} = \mathrm{NGM}\,\hat{y}$. Therefore, summing over all the entries of $R_0\hat{y}$, we have the relation

$$R_0 = \sum_i \sum_j (\mathrm{NGM})_{ij}\hat{y}_j = \sum_j \hat{y}_j \underbrace{\sum_i R_{ji}}_{R_j}\,. \tag{2.74}$$

Therefore, the basic reproduction number is nothing more than a weighted average with respect to the case distribution, $\hat{y}$, of the individual reproduction numbers R_j, which corresponds to the number of secondary cases generated by a primary case in group j. Later, in Chapter 5, we will use this relation to calculate the reproduction number when the case distribution does not follow $\hat{y}$ due to constant changes in the contact rates. But first, we will consider how one can estimate the reproduction number from case data during an evolving epidemic.

2.3 THE INSTANTANEOUS REPRODUCTION NUMBER

The basic reproduction number is defined as the number of secondary cases produced by primary cases in a fully susceptible population. However, during the course of an epidemic, the number of susceptible individuals will continuously decrease. We have already discussed in Sec. 2.1.4 that for a population which is not fully susceptible, the epidemic threshold, and thus the reproduction number reads $R = \lambda k S \tau_I$ ($\tau_I = 1/\mu$) for the SIR model. The reproduction number is thus time-dependent since the fraction of susceptibles $S(t)$ is constantly changing throughout an epidemic. Furthermore, the fraction of susceptible individuals is not even constant during the infectious period of an infected individual. In this sense, $\lambda k S(t) \tau_I$ does not reflect the

number of secondary cases produced by an individual at time t. Accordingly, an important question arises: how can one define the reproduction number in an evolving environment? Generally, there are two definitions of the time-dependent reproduction number that people employ: the *instantaneous reproduction number* [216] and the *case reproduction number* [217].

The case reproduction number, R_c, is the expected number of secondary cases produced by an individual infected at time t (primary case). The case reproduction number thus explicitly considers the changing environment during the infectious period of an infected individual. In contrast, the instantaneous reproduction number, R_t, explicitly discards the changing environment. It is defined as the expected number of secondary cases produced by an individual infected at time t (primary case) should the conditions not change. For example, the quantity R_t assumes that the fraction of susceptible individuals stays constant during the infectious period, but all the other parameters are also assumed constant. From a mathematical perspective, in a simple SIR framework, the two quantities R_c and R_t are defined as

$$R_c(t) = \int_t^\infty k(t')S(t')\lambda e^{-\frac{(t'-t)}{\tau_I}} dt' \tag{2.75}$$

$$R_t(t) = \lambda k(t)S(t)\tau_I\,, \tag{2.76}$$

where we assume that the contact rate $k(t)$ may also change over time. The mathematical description of $R_t(t)$ is immediate from its definition. One assumes that the fraction of susceptible individuals $S(t)$ and the contact rate $k(t)$ stay constant during the entire infectious period τ_I. For the case reproduction number R_c, changes in the environment are modulated by the survival function e^{-t/τ_I} of the infected individual. The survival function quantifies the probability that an individual, infected at $t = 0$, is still infectious at time t.

From a conceptual point of view, the case reproduction number R_c follows the course of infection of an individual infected at time t. The quantity informs how individuals infected at different times contribute to the spread of a disease. In contrast, the instantaneous reproduction number R_t is a more abstract concept that is useful to quantify, for example, how the circumstances of a specific day impacted the spread of the virus. For this reason, the instantaneous reproduction number became a heavily used quantity to estimate the impact of *non pharmaceutical interventions* (NPIs) during the SARS-CoV-2 pandemic. Any NPI that reduces, for example, the number of contacts $k(t)$ leads

to a sharp decrease in R_t, as reflected in Eq. (2.76). On the contrary, for the case reproduction number, the impact of an NPI will be spread out over multiple days due to the integral formulation in Eq. (2.75). For this reason, estimating R_t from the epidemiological data (case reports) was an essential tool to determine the actual impact of NPIs. The most popular framework used to infer R_t is called *EpiEstim*, which we will present in the next section.

2.3.1 *EpiEstim - Estimating R_t*

EpiEstim, a framework to infer the instantaneous reproduction number, was initially proposed and developed by Cori et al. [218–220]. The approach assumes a simple epidemiological model with a homogeneous population interacting randomly. However, the framework is not limited to a constant infectivity rate but allows for any possible distribution. The model equations are similar to the original formulation of the Kermack-McKendrick model presented in Sec. 2.1.2. In the approach by Cori et al., given an infectivity profile w_s, the daily incidence I_t (new infections on day t) evolves according to a renewal equation as

$$I_t = R_t \sum_{s=0}^{t} I_{t-s} w_s \,. \tag{2.77}$$

The infectivity profile w_s, which is often referred to as generation time distribution, is the product of the survival function and the relative variation in infectiousness. So far, we mainly assumed constant infectiousness, but most real diseases display substantial differences over the course of an infection [18]. In the following section, we will detail the interpretation of w_s. The EpiEstim framework then compares the simple model in Eq. (2.77) to the actual reported incidence data. Assuming the incidence is Poisson distributed with mean given by I_t, the time series R_t can then be estimated through Bayesian inference, which will be discussed in Sec. 2.4. For the actual expression of the likelihood or other details, such as the priors of the inference or the smoothing factor for the time series R_t, we refer directly to Refs. [218–220]. Please note that the quantity R_t here refers to the number of secondary cases produced by the average infected individual. Since the framework assumes a homogeneous population, the distinct reproduction numbers existing among different groups are aggregated and averaged, similarly as in Eq. (2.74), giv-

ing R_t [3]. EpiEstim estimates the time series R_t from the reported incidence, assuming that the infectivity profile w_s is known. However, often it is challenging to estimate w_s; we will see why in the next section.

2.3.2 *Non-Markovian infectivity profiles*

As previously highlighted, the typical compartmental models assume simple within-host dynamics, i.e. a constant transmission probability and a constant recovery rate. This assumption results in Markovian dynamics since the probability of infecting someone or recovering does not depend on how many days someone has been infectious. The Markovian assumption facilitates the mathematical modelling and simplifies the model complexity substantially. However, in most cases, neither infectiousness nor the recovery rate is constant [223]. Most diseases have a typical time for the start of infectiousness, the onset of symptoms (incubation period), or the recovery [62, 224–226]. The disease characteristics also lead to specific intervals in transmission chains between primary and secondary cases. The two most relevant intervals are the *generation time*, and the *serial interval* [227, 228]. In analogy to population dynamics, the generation time is the time that passes between infection of a primary case and the subsequent onward transmission to a secondary case. The serial interval, in contrast, is the time that passes between the symptom onset of the primary case and the symptom onset of its secondary case. Together with the incubation period, the intervals are illustrated in Fig. 2.7.

The generation time is equivalent to the infectivity profile used in EpiEstim. It describes when, after infection, an individual is most likely to transmit the disease. The generation time combines the transmission probability ϕ_θ and the recovery probability ψ_θ in the original formulation of the Kermack-McKendrick model (Sec. 2.1.2). However, instead of explicitly considering when an individual recovers, the generation time accounts for this by a decreasing probability of onward transmission. In this case, even though individuals may not have recovered, they do not contribute to the dynamics in Eq. (2.77).

The discrete-time dynamics, as in Eq. (2.77) and originally formulated by Kermack-McKendrick, had their revival in the modelling and analysis of

3 In reality, the variation of the reproduction number across the population follows a distribution, whose heterogeneity is generally captured by the overdispersion of a negative binomial [221]. The overdispersion results from a non-trivial interplay between the heterogeneity in the number of contacts and viral shedding across individuals [222].

SARS-CoV-2 [229]. Estimating the impact of NPIs, which often should show an impact on a specific day, requires using the correct inter-generation time. Using the Markovian approximation [4] could be misleading in these cases, in particular, if various NPIs have been introduced in a relatively short time. The continuous time analogy is often computationally very costly [231]. The only problem is that the generation time can often be hard to estimate. Due to the incubation period, it is generally challenging to determine when individuals got infected. Creative detours need to be found in order to infer the generation time [232]. In contrast, determining the day of symptom onset is straightforward; therefore, it is also simple to estimate the serial interval. Often, researchers use the serial interval as a proxy for the generation time [225, 233]. The two quantities share the same mean, but the variance may differ [227], which generally affects the estimations of R_t [225]. Essentially, the almost equivalence comes from the fact that the serial interval, $t_4 - t_2$, is simply $t_3 - t_1$, the generation time convoluted by the incubation period. The convolution does not affect the average but changes the variance.

4 A trick is to divide the compartments into various subcompartments. For example, instead of one infectious compartment I, take n infectious compartments, I_1, I_2, I_3, etc., while the time spent in each compartment is τ_I/n. In this case, the dynamics stay Markovian. However, the generation time will follow an Erlang distribution which is a better approximation, in general, than an exponential distribution [230].

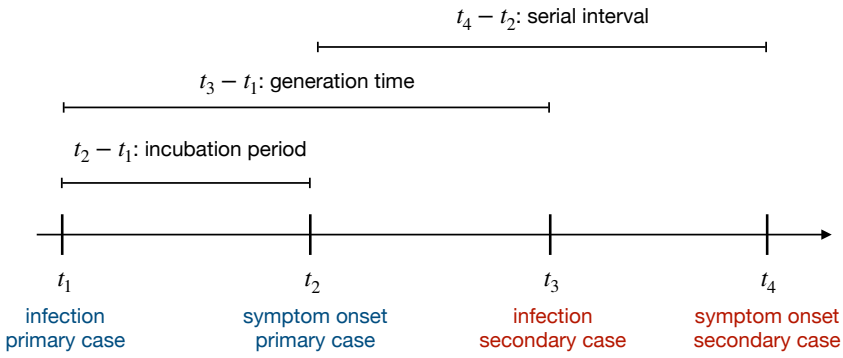


Figure 2.7: Schema for different intervals in transmission chains. It includes the incubation period, the generation time and the serial interval. Please note that the serial interval can be negative while the generation time is positive by definition.

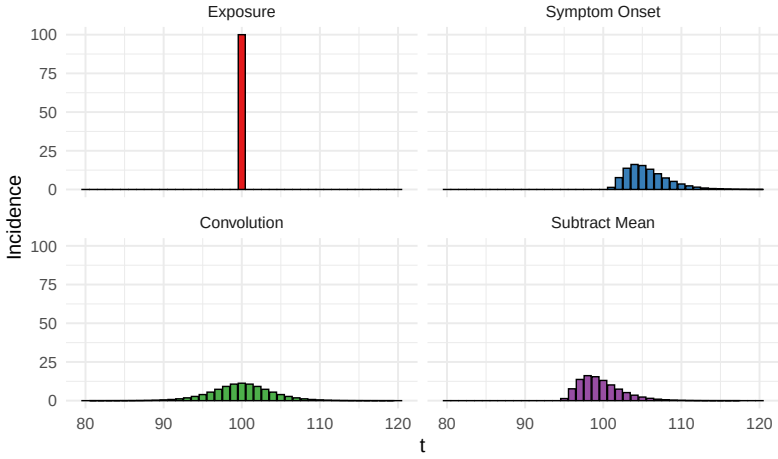


Figure 2.8: Original incidence (top-left), observed symptom onset (top-right) and two examples of how the exposure times should not be reconstructed. Both doing a convolution instead of deconvolution (bottom-left) and a simple subtraction of the mean (bottom-right). Adapted from Ref. [233]. The time series of symptom onset was obtained by convoluting the exposures with a gamma distribution with a mean of 5.2 and a standard deviation of 2.8 (incubation period).

2.3.3 *Reconstructing the exposure times*

A similar problem arises when using the reported cases as a proxy of incidence to estimate R_t. In general, cases are only diagnosed and reported once individuals have symptoms. Accordingly, the reported cases reflect a convolution of the actual incidence with the time between infection and reporting. Hence, when one uses the reported cases to infer the reproduction number, the estimated R_t is also convoluted by the incubation period. The present delay then hinders any analysis regarding the impact of NPIs. Therefore, before estimating R_t, it is necessary to deconvolute the time series and infer the actual exposure times (day of infection) [225, 233]. Luckily, there are several algorithms available to perform this deconvolution [234–236].

Fig. 2.8 illustrates why performing the deconvolution, and doing it properly, is crucial before estimating R_t. On the top-left, we show the number of newly infected individuals on $t = 100$. Assuming that the date of symptom onset is reported, we would observe the top-right, the time series convoluted by the incubation period. Additionally to the shift, as in image processing, this corresponds to a blurring of the original data. Accordingly, if one per-

formed a convolution to recover the exposure times (bottom-left), one would blur the data further. Similarly, subtracting the incubation average does not allow for recovering the exposure times (bottom-right). The direct estimation from the time series of symptom onset, performing a convolution instead of a deconvolution, and simply subtracting the mean are common mistakes that health authorities and researchers made during the SARS-CoV-2 pandemic. We hope that the above figure has illustrated why these errors should be avoided when estimating R_t.

One can also deconvolute the time series of hospital admission or daily deaths directly to reconstruct the exposure times [233]. Since R_t is defined as a fraction in Eq. (2.77), a constant underreporting rate does not influence the estimation of R_t. However, as we will see in Chapter 5, the reporting rate can rapidly decrease during a growing epidemic as test facilities get saturated. Therefore, hospital admissions or daily deaths can often be a more reliable data stream in these cases. For the same reason, hospital admissions, ICU admissions and daily deaths are generally preferred to fit epidemiological models instead of relying on the reported infections. In the next section, we will present one of the most common practises for doing the actual fitting.

2.4 FITTING EPIDEMIC MODELS TO DATA

There are many ways to fit an epidemic model to data, and many data types exist [237]. In this thesis, the data we adjusted the models to mainly consisted of reported case numbers (hospital admissions, ICU admissions, deaths and infections). For such standard data streams, it is luckily quite straightforward to express a likelihood that compares the model with the data[5]. Traditional approaches have mainly relied on maximum likelihood estimations, or least-square minimisation [239], to fit epidemic models to the data, and infer the corresponding parameters [240]. However, the increasing availability of computational power in recent years has made it possible to evaluate the entire posterior and hence do actual *Bayesian inference* [241], beyond the estimation of the maximum-likelihood.

Bayesian inference leverages Bayes theorem for conditional probabilities to express the probability, referred to as posterior distribution, that a specific set of parameters of a given model may explain the observed data [242]. More

5 If the likelihood cannot be expressed, or is computationally too costly, one can also do approximate Bayesian computation (ABC) [238], which often simply uses a least-square approximation for the likelihood.

specifically, given a set of parameters θ and the observed data X, the posterior probability distribution is expressed as

$$p(\theta|X) = \frac{p(X|\theta)p(\theta)}{p(X)} \sim p(X|\theta)p(\theta)\,. \tag{2.78}$$

The conditional probability $p(X|\theta)$ is the probability of observing the data X given the model parameters θ, which corresponds to the likelihood employed in the maximum likelihood estimation. The quantity $p(X)$ denotes the probability of making the observation X. Only God will ever be able to express $p(X)$. However, since it is only a normalising factor and does not depend on θ, it can be dropped from the expression of the posterior distribution $p(\theta|X)$. The last term $p(\theta)$ is the prior distribution. It reflects our prior knowledge about the parameters of the system. For example, previous estimations of a parameter can be incorporated into the prior distribution. Similarly, some parameters may only take positive values, which would be reflected in a sharp cutoff in $p(\theta)$. If there is no prior information on the parameters, the posterior $p(\theta|X)$ is directly proportional to the likelihood $p(X|\theta)$.

Given the expression of the posterior distribution in Eq. (2.78), one needs to explore the phase space and evaluate $p(\theta|X)$ to estimate, for example, the mean $\langle\theta\rangle$. The mean can be directly expressed as $\langle\theta\rangle = \int p(\theta|X)d\theta$. However, depending on the number of parameters of the model, the parameter space can often be very high-dimensional. In these cases, evaluating the posterior distribution in the entire parameter phase space is computationally not feasible. To contour this problem, researchers designed a set of Markov chain Monte Carlo (MCMC) methods to explore the parameter phase space more efficiently [243]. These methods build a Markov Chain with a given update rule that, after a sufficient number of steps, converges to the target distribution, the posterior $p(\theta|X)$. The most famous of these MCMC methods is the *Metropolis-Hastings algorithm* that was originally developed in statistical physics [169, 244]. Nowadays, more efficient samplers such as the *Hamiltonian Monte Carlo* [245] exist. Ref. [246] provides a good overview of the available different methods.

Assuming that the data, X, consists of case reports, the likelihood $p(X|\theta)$, i.e. the observation process, can be described by a Poisson distribution, where the mean corresponds to the model prediction. However, reported data often underlies substantial variation. For example, reporting is generally much lower on weekends than on weekdays. To allow for such variance, people

often employ a negative binomial distribution instead of a Poisson to model the count data [247]. The dispersion parameter of the negative binomial distribution is part of the inference process. Generally, it is desirable to fit the model to incidence data instead of cumulative cases. Fitting to cumulative cases can lead to substantial overconfidence in the estimations [248]. There are many excellent ready-made implementations to perform Bayesian inference. We will use Stan in this thesis [249–251], but others such as Turing [252] in Julia [253], or emcee [254] in Python offer similar advantages. The actual fit is then made by running multiple chains that take a fixed number of samples. Initial samples (warm-up) are often discarded to allow the chain to reach the vicinity of the posterior distribution. Convergence can either be checked visually, by looking at the temporal evolution of the chains (trace plots) or through the Gelman-Rubin statistics [255]. Once the fit is performed, credible intervals – the equivalent of confidence intervals in the frequentist approach – can be calculated for any quantity by propagating the posterior. An excellent introduction to the workflow for epidemic modelling with Stan is provided by Grinsztajn et al. [256].

2.5 EXPLICITLY MODELLING HUMAN BEHAVIOUR IN THE CONTEXT OF EPIDEMICS

We have presented a series of approaches to describe the spread of infectious diseases. Extensions of the original approach by Kermack-McKenrick essentially incorporated different features of human behaviour. Stratification of the population, for example, according to age and geography, or the use of real-world contact networks allowed us to describe more accurately the structure that underlies human interactions. We may also see the work by Kermack-McKenrick as an attempt to model human behaviour relying on the assumption that individuals homogeneously interact according to the law of mass action. In this sense, models of infectious diseases, which are transmitted upon human contact, always rely on a description of human behaviour. For this reason, we added the word "explicitly" to the title of this section. In general, modelling human behaviour in the context of epidemics is primarily understood as focusing on the explicit dynamical behavioural reaction in the presence of a disease [6, 19–22, 257].

The behavioural reaction in the face of a propagating disease almost always aims to reduce the individual or global infection risk. Such behavioural

changes and practices to contain the spread of a disease have been observed throughout human history. Examples are the introduction of a quarantine in Dubrovnik (Croatia) during the 14th century to prevent the plague from entering the city [258], the use of face masks during the H1N1 influenza A pandemic in 1918 [258], the increasing demand for condoms after the emergence of HIV in the 1980s [259], or the adoption of intensified hygiene practices in the face of the 2009 H1N1 pandemic [260]. Similarly, we can consider the rapid development and adoption of vaccines against SARS-CoV-2 as a collective behavioural reaction. We have highlighted in the introduction that the interaction between diseases and human behaviour can be understood as a coevolution at the biological and cultural timescale. However, mathematical models are far from modelling such coevolution. For this reason, mathematical modelling of adaptive behaviour in the face of an epidemic focuses mainly on the individual adoption of prophylaxis [19, 21, 22]. The prophylactic practice can be anything from vaccines, face masks or social distancing, to name a few. Generally, the uptake of such prophylactic measures is modelled either through game theory or the spread of individual awareness. We will detail and overview both approaches in the following.

2.5.1 *Vaccine adoption formulated as a dilemma in game theory*

The voluntary adoption of vaccines can be understood as a social dilemma. The dilemma arises since not vaccinated individuals profit from others who get vaccinated. The presence of vaccinated individuals also decreases the infection risk for not vaccinated individuals. Accordingly, individuals that do not get the vaccine can be seen as free riders taking advantage of the ones that get the vaccine. In the extreme case, where the vaccination coverage is sufficient to reach herd immunity, non-vaccinated individuals have no incentive to get vaccinated (besides solidarity) since the infection risk is 0. If the number of free riders (not vaccinated) exceeds a threshold, the population may lose herd immunity, hence the dilemma.

A natural framework to describe such a dilemma is game theory. Initially proposed by John Von Neumann in 1928 [261], and popularised with the book "Theory of Games and Economic Behaviour" by Von Neumann and Morgenstern [262], game theory rapidly became a widely used tool to analyse collective decision making. It focuses on situations where the outcome depends not only on the decision maker but also on the decision of others. Game theory

was rapidly adapted in economics to model the decision-making of rational agents [263] but also found application to describe deterrence theory in the context of nuclear warfare [264]. In the framework of game theory, individuals choose between a set of strategies, and each strategy has an associated *payoff*. The equilibrium, referred to as Nash Equilibrium, is reached when none of the individuals can increase their benefit (payoff) by changing the strategy [263]. If one allows for mixed strategies, i.e. the possibility for an individual to adopt various strategies according to different probabilities, all individuals, independent of the adopted strategy, will necessarily have the same payoff at equilibrium.

The description of voluntary vaccine uptake through game theory was probably the first approach that took explicitly into account the behavioural adaption [265, 266]. The vaccination "game" has two associated strategies, individuals either get vaccinated (V) or not (NV). The corresponding payoffs, P_{V} and P_{NV}, are determined by the vaccination cost c, and the infection cost T (we assume $c < T$) modulated by the infection probability R_∞. The infection probability is generally inferred from an SIR-like outbreak after the vaccination uptake. The infection cost c is not merely financial but may also include the fear of secondary effects. The infection cost T reflects the possible complications when contracting the disease. Accordingly, the payoffs are expressed as

$$P_{\mathrm{V}} = -c \tag{2.79}$$

$$P_{\mathrm{NV}} = -TR_\infty\,. \tag{2.80}$$

Here we assume a perfect vaccine, wherefore vaccinated individuals cannot get infected, and the vaccination cost completely determines their payoff. The product TR_∞ reflects the expected infection cost. The infection risk R_∞ is a function of the fraction of vaccinated individuals y. From the equilibrium condition $P_{\mathrm{V}} = P_{\mathrm{NV}}$, we find $R_\infty(y^*) = c/T$, where y^* is the vaccination coverage at equilibrium. Assuming that y_c is the critical vaccination coverage to reach herd immunity, the equilibrium indicates that herd immunity will not be reached, i.e. $y^* < y_c$. Otherwise, we had $R_\infty(y^*) = 0$, and the equilibrium condition would not be fulfilled. In this sense, game theory perfectly captures the essence of the vaccination dilemma.

Now, the question that arises is how such an equilibrium is reached. In the above framework, individuals are completely rational and can calculate their payoffs, i.e. they can evaluate their future infection risk R_∞. More real-

istically, reaching an equilibrium requires repeated outbreaks which allow individuals to adapt their strategy. To model the temporal evolution towards equilibrium, researchers in social sciences borrowed techniques that were developed in biology [267]. In the 1970s, game theory was suddenly discovered as the general framework to describe biological evolution. The difference was that strategies corresponded to different phenotypes and payoffs to their reproductive ability (fitness). Intuitively, equilibrium can only be reached if all phenotypes reproduce at the same rate; otherwise, the most successful phenotype would invade the others. The reproduction leads to a dynamics where the frequency of a phenotype x_i evolves as

$$\dot{x}_i = x_i \left(f_i(x) - \langle f(x) \rangle \right) , \tag{2.81}$$

which is known as the replicator equation [268]. The entry $f_i(x)$ corresponds to the fitness of phenotype i, given the frequency distribution x. Then, $\langle f(x) \rangle$ is the average fitness in the population, making apparent why equilibrium is reached when all individuals have the same fitness. Eq. (2.81) then governs whether and how the equilibrium is reached given an initial condition.

In the context of social processes, there is no apparent reproductive process as heredity in biology. However, economists introduced a temporal evolution in game theory by proposing imitation as a reproduction mechanism [269]. Suppose the game is played repetitively, and individuals imitate the most successful strategy with a probability proportional to the payoff difference. In that case, the dynamics follow the replicator equation in Eq. (2.81) [269]. The imitation dynamics have been used in the context of vaccine adoption, where individuals update their strategies after each subsequent outbreak [270–272]. Beyond vaccination, we will employ a similar approach in Chapter 3 to describe the temporal coevolution of the adoption of prophylactic tools and the spread of an epidemic. However, first, let us introduce the second approach based on individual awareness.

2.5.2 *Awareness models*

The second approach to modelling the behavioural reaction is assuming that individual awareness leads to a reduction in the infection probability. Many different mechanisms to model this reduction have been considered [19, 21, 22]. The emergence of the individual awareness is then either *prevalence-based* or *belief-based* [19]. The seminal study by Funk et al. already analysed ex-

plicitly both approaches [273]. The prevalence-based approach assumes that the awareness is fuelled by the presence of the disease and can only be sustained if the disease is endemic [273, 274]. The simplest example would be to assume that the transmission probability decreases as the global prevalence increases, I, as $\lambda(I) = \lambda(1 - I)$. In these cases, similarly to the game theoretic approach, the epidemic threshold is not affected by the possibility of awareness [273, 275–277]. At the epidemic threshold, prevalence is effectively zero. Accordingly, it seems intuitive that a process which can only be triggered and sustained for a non-zero prevalence is not present close to the disease-free state and, hence, does not affect the epidemic threshold.

In contrast, in the belief-based approach, the epidemic threshold is affected [273, 278–280]. The belief-based approach assumes that the awareness can spread, also in the absence of infected individuals (locally or globally), as a social contagion [281]. Every "infected individual" with this social contagion shows awareness and has a reduced probability of infection. Since the social contagion can spread in the absence of the disease, individual awareness slows down the spread of the pathogen even close to the disease-free equilibrium. Accordingly, the belief-based approach leads to an increase in the epidemic threshold. Now, there is a third specific category of models that we have yet to mention, those called adaptive networks [282–285]. In the context of epidemics, the framework of adaptive networks considers that individuals reshape their contacts in the presence of a propagating disease. Typically, ties between susceptible and infected individuals are cut and randomly rewired for the susceptible individual. Effectively, the process corresponds to the isolation of detected infected individuals, which impacts and increases the epidemic threshold. However, the extent to which the epidemic threshold is affected is bounded by the presymptomatic transmission rate, i.e. the transmission risk before individuals are detected.

This overview completes the necessary theoretical background to follow the different contributions of the thesis. Next, in Chapter 3, we will outline two models, which serve to understand how prophylactic practice can give rise to subsequent epidemic waves. As previously mentioned, we will use a game theoretic framework instead of the awareness-based description.

UNIVERSITAT ROVIRA I VIRGILI
MODELLING THE INTERPLAY BETWEEN HUMAN BEHAVIOUR AND THE SPREAD OF INFECTIOUS DISEASES: FROM TOY MODELS TO QUANTITATIVE APPROACHES
Benjamin Franz Josef Steinegger

3

EPIDEMIC CYCLES DRIVEN BY HUMAN BEHAVIOUR

Infectious diseases that provide short- or long-term immunity often exhibit recurrent incidence cycles [286–289]. For example, influenza outbreaks occur annually during the winter months and are fueled by the rapid mutation of the virus strains [289]. In the case of childhood diseases, these cycles are – and luckily in many regions, they are not anymore – driven by the accumulation of susceptible individuals through birth and external seasonal factors [59]. However, researchers also suggested that human behaviour may be at the core of the epidemic cycles. More precisely, Chris T. Bauch put forward a framework that explains prevalence cycles of whole-cell pertussis (whooping cough) in England and Wales during the 1970s through a risk-based vaccine adoption [270]. The whole-cell pertussis vaccine was controversial since it had rare but significant side effects, and the general public even came to associate neurological issues with the vaccine, even though they were demonstrated to be unrelated [290]. In his model, Bauch assumed that, due to the perceived side effects, parents only vaccinate their newborns against whole-cell pertussis if the infection risk is sufficiently high. Describing the parents' decision as an evolutionary game, the model effectively reproduced the phenomenology of the recurrent epidemic cycles.

Another example where human behaviour was proposed as the origin for epidemic cycles is syphilis [291]. The 8-11 year cycles in the syphilis incidence across the US were first identified and studied by Grassly et al. [292]. They hypothesised that the cycles were unforced, endogenous oscillations driven by the natural dynamics of a syphilis infection, namely the partial protective immunity. However, these findings were heavily contested later on, particularly the parameter choices made by Grassly et al. [293]. Althouse et al. then proposed that the cycles may result from the adaptive behaviour of individuals when facing detected infections among their sexual partners [291]. They

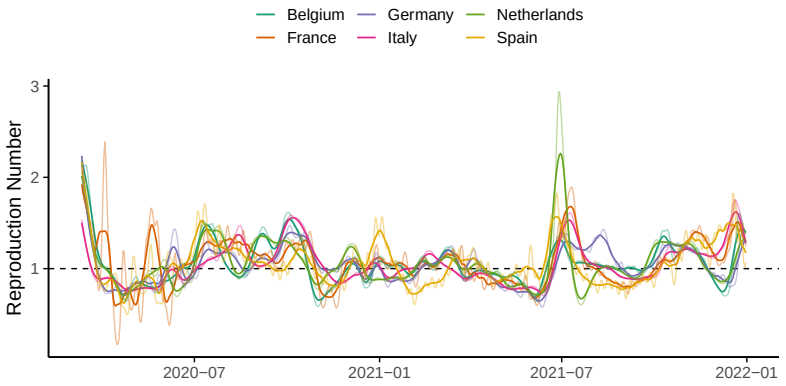


Figure 3.1: Estimation of the reproduction number from the reported cases [294] for a series of European countries. Credible intervals are not illustrated to improve visibility. The reproduction numbers were estimated using the framework and parameters provided in Ref. [236]. Solid lines represent a 14-day rolling average of the actual estimation (shaded line).

assume individuals disconnect (stop interacting) from infected sexual partners at a given rate. This framework is generally known as epidemics on adaptive networks [282–285], which exhibits sustained epidemic cycles in a specific parameter region.

With the emergence of SARS-CoV-2, we have encountered another infectious disease that shows recurrent increases and decreases in the incidence. We have experienced many surges of SARS-CoV-2 infections and subsequent decreases since early 2020. The recurrent epidemic waves result in a reproduction number that constantly "oscillates" around 1, as illustrated in Fig. 3.1. Accordingly, the disease is not eradicated but is more or less controlled. While the synchronised peak during the winter months can be attributed to seasonal factors [295, 296], increases in spring and summer are much more incoherent across countries (Fig. 3.1). In particular during the summer of 2020, when no new, more transmissible variant was introduced to Europe. These peaks can be associated with a relaxation in social distancing, and a subsequent strengthening [297–300]. Obviously, in the case of SARS-CoV-2, practices of social distancing result from a complex interplay between individual decisions and governmental restrictions. However, evidence suggests that, beyond policy-induced social distancing, voluntary adoption played a crucial role in the evolution of the pandemic [300].

The presence of epidemic cycles beyond childhood diseases sparked our interest to see whether a game theoretic approach could reproduce the observed phenomenology. In contrast to the adoption of vaccines, prophylactic practices such as using face masks or social distancing are reversible decisions. It is hence not possible to directly adapt the framework in Ref. [270]. The approach by Althouse et al., using adaptive networks, is very specific to sexually transmitted diseases. It does not seem reasonable that, for example, SARS-CoV-2 infections lead to a long-term restructuring of the contact network. An infected individual may be temporarily isolated, but social contacts are quickly reestablished afterwards. Isolation translates as a reduced effective infectious period, which can be incorporated into a standard SIR model; it does not lead to epidemic cycles. Additionally, adaptive networks generally have only a small parameter range in which epidemic cycles are sustained [282]; only the introduction of memory can extend this region [285]. Somewhat similar, an awareness kernel has been proposed that acts on the contact rates or the transmission probability [275]. The kernel is defined such that the higher the prevalence in the past, the stronger the prophylactic practice will be. However, we will introduce here a simple toy model that exhibits sustained epidemic cycles and does not require memory if the behavioural dimension is explicitly introduced. The model is presented in the following Sec. 3.1. The second part of this chapter, Sec. 3.2, will then extend the model to incorporate heterogeneity in the risk of a severe infection, which will explain the temporal variation of reported SARS-CoV-2 cases across age groups.

3.1 EPIDEMIC CYCLES IN A HOMOGENEOUS POPULATION

We outlined in Sec. 2.5.1 how the underlying dilemma of the voluntary adoption of vaccines can be understood in game theoretical terms [265, 266, 270]. The adoption of prophylactic measures such as the use of face masks or social distancing can be understood in a similar way [301–303]. However, in contrast with the assumption made in Sec. 2.5.1 where we assumed 100% efficacy, most actual prophylactic measures provide only partial protection. In other words, anyone that adopts prophylaxis can still get infected. This subtlety requires a slight reformulation of the payoffs. Furthermore, transparency is critical since the model here aims to explain how epidemic cycles emerge and are sustained through individual prophylaxis. To this aim, we will introduce a series of approximations that allow for the analytical tractability

of the model, facilitating the exploration of the phase space. After introducing the model and the phenomenology it exhibits, we will also propose an intervention policy that could efficiently prevent recurrent outbreaks.

3.1.1 *Recurrence equations*

To describe the system under consideration, we must model both the spread of the virus (epidemics) and the adoption of prophylactic tools (human behaviour). The coupling of two dynamics can rapidly increase the number of parameters and, thus, the model's complexity. However, since our interest lies in the phenomenology and not in a quantitative description, we intend to keep things as straightforward as possible while retaining the key ingredients. Therefore, for the sake of simplicity, we describe the spread of the epidemic with a standard SIS model, including a recovery probability μ and transmission probability λ that spreads on a network with an adjacency matrix A. The main advantage of taking an SIS model over an SIR model on a closed population is the existence of a non-trivial stationary state or at least a limit cycle. This way, the system is less dependent on the initial conditions, and the transient state can be discarded.

As pointed out previously, game theory is a natural choice to describe the adoption of vaccines and prophylactic tools. We thus formulate the decision problem as a two-strategy game where individuals either adopt prophylaxis and are protected (P) or do not and are unprotected (NP). In this sense, the adoption can be understood as a trade-off between an adoption cost c and an associated benefit stemming from the reduced infection risk. Please note that the adoption cost should not be understood as a mere financial cost. It also encodes the personal and social disadvantages that go along with the prophylactic tool. For example, the use of a face mask can be physically uncomfortable which also contributes to the adoption cost. The cost associated with an infection is the infection cost T. Similarly, the infection cost is not primarily financial. Depending on the disease under consideration, it includes the risk of a severe course of infection, physical pain, or mental stress stemming from the possibility of having transmitted the disease. It is a general challenge in game theory to quantify non-monetary costs and translate them into payoffs. However, since we take a qualitative approach and we will explore a wide range of parameters, such abstraction is not further concerning. Given the infection cost, the benefit of the prophylactic tool is determined by the asso-

ciated reduction of the infection risk. The decision to adopt the prophylactic tool is thus risk-based. Indeed, empirical studies for different prophylactic tools and diseases indicate that protective behaviour increases with prevalence [259, 265, 304–307]. We incorporate this into the model by considering an *expected infection cost*, which is the product between the infection risk and the infection cost. Including the infection risk in the payoffs acts as the coupling between the behavioural adaptation and the epidemic dynamics. The dynamics can thus be considered a game with environmental feedback, the epidemic [308].

In the model, we have a total of four compartments —two behavioural and two epidemiological—: susceptible and not protected S_{np}, susceptible and protected S_p, infected and not protected I_{np}, or infected and protected S_p. Accordingly, the payoffs of being protected P_p or unprotected P_{np} can be written as

$$P_p(t) = -c - T\frac{I_p(t)}{I_p(t) + S_p(t)} \tag{3.1}$$

$$P_{np}(t) = -T\frac{I_{np}(t)}{I_{np}(t) + S_{np}(t)}\,. \tag{3.2}$$

In the above expressions, the two ratios express the global infection risk depending on whether individuals are protected. Note that we assume that the individuals have knowledge of the macroscopic state of the epidemic. Once the payoffs are defined, the question arises how individuals adopt either strategy. As highlighted before, prophylactic tools can be adopted and discarded repetitively in contrast to vaccines. Therefore, the decision process should be able to evolve in time. For this reason, we adopt an evolutionary game instead of a static one [269]. However, instead of opting for the imitation dynamics, we assume that individuals adopt the most successful strategy with a probability that is proportional to the difference in payoffs (Eqs. (3.1) & (3.2)). This choice simplifies the model with respect to the imitation dynamics and makes it analytically tractable. On a macroscopic level, these transition probabilities lead to a dynamics very similar to those of the replication equation [268]. More specifically, the transition probabilities are defined as

$$\Gamma_{p\to np}(t) = \alpha\frac{\Delta P_{np-p}(t)}{T+c}\Theta[\Delta P_{np-p}(t)] \tag{3.3}$$

$$\Gamma_{np\to p}(t) = \alpha\frac{\Delta P_{p-np}(t)}{T+c}\Theta[\Delta P_{p-np}(t)]\,. \tag{3.4}$$

Θ represents the Heaviside function, where $\Theta(x) = 1$ if $x >= 0$ and $\Theta(x) = 0$ if $x < 0$. The differences in payoffs are defined as $\Delta P_{np-p}(t) = -\Delta P_{p-np}(t) = P_{np}(t) - P_p(t)$. The parameter $\alpha \in [0, 1]$ controls the time scale between the behavioural and the epidemic dynamics. The smaller α, the slower individuals adapt their strategies in the face of evolving epidemics. The denominator $T + c$ is the maximum possible difference in payoff and normalises the probabilities. Due to the normalisation, the system's dynamics is invariant under a simultaneous rescaling of the infection and protection cost. Solely the ratio between the protection and infection cost (c/T) determines the adoption of prophylaxis. Therefore, without loss of generality, we will fix $c = 1$ and only vary T in the following.

Prophylactic tools very rarely provide complete protection against infection. Face masks [309], social distancing [310], or condom use [311] all have effectiveness well below 100%. In analogy to leaky vaccines [312], we consider a leaky prophylactic tool with effectiveness $1 - \gamma$ upon each contact. For $\gamma = 0$, protection is perfect, while $\gamma = 1$ means there is no protection. We assume that the effectiveness of the prophylactic tool to prevent transmission and infection is the same. To be more precise, the probability for transmission is $\gamma\lambda$, independent of whether the infected individual or the susceptible one is protected. Furthermore, we assume that if two protected individuals meet, the transmission probability is only linearly reduced. Both of these assumptions are reasonable for condom use but, for example, do not hold for face masks. However, the assumptions help to simplify the number of parameters and are unlikely to affect the overall phenomenology. Eventually, the transition between epidemic compartments can be described by the following reactions between agents i and j

$$S^i_{np}(t) + I^j_{np}(t) \xrightarrow{\lambda} I^i_{np}(t) + I^j_{np}(t) \tag{3.5}$$

$$S^i_{np}(t) + I^j_{p}(t) \xrightarrow{\gamma\lambda} I^i_{np}(t) + I^j_{p}(t) \tag{3.6}$$

$$S^i_{p}(t) + I^j_{np}(t) \xrightarrow{\gamma\lambda} I^i_{p}(t) + I^j_{np}(t) \tag{3.7}$$

$$S^i_{p}(t) + I^j_{p}(t) \xrightarrow{\gamma\lambda} I^i_{p}(t) + I^j_{p}(t)\,. \tag{3.8}$$

To write the model equations, we employ the discrete-time Microscopic Markov Chain Approach [175, 177], which we have detailed in Sec. 2.2.3.4. In this approach, for every individual $i = 1, \dots N$, the variables $S^i_{np}(t)$, $S^i_p(t)$,

$I^i_{np}(t)$ and $I^i_p(t)$ describe the individuals' probability to be in each of the four states at time t. The model equations read

$$\begin{aligned} S^i_p(t+1) = &\left[1-\Gamma_{p\to np}(t)\right]\left\{S^i_p(t)\left[1-q^i_p(t)\right]+I^i_p(t)\mu\right\} \\ &+\Gamma_{np\to p}(t)\left\{S^i_{np}(t)\left[1-q^i_p(t)\right]+I^i_{np}(t)\mu\right\} \qquad (3.9) \\ S^i_{np}(t+1) = &\,\Gamma_{p\to np}(t)\left\{S^i_p(t)\left[1-q^i_{np}(t)\right]+I^i_p(t)\mu\right\} \\ &+\left[1-\Gamma_{np\to p}(t)\right]\left\{S^i_{np}(t)\left[1-q^i_{np}(t)\right]+I^i_{np}(t)\mu\right\} \qquad (3.10) \\ I^i_p(t+1) = &\left[1-\Gamma_{p\to np}(t)\right]\left\{S^i_p(t)q^i_p(t)+I^i_p(t)(1-\mu)\right\} \\ &+\Gamma_{np\to p}(t)\left\{S^i_{np}(t)q^i_p(t)+I^i_{np}(t)(1-\mu)\right\} \qquad (3.11) \\ I^i_{np}(t+1) = &\,\Gamma_{p\to np}(t)\left\{S^i_p(t)q^i_{np}(t)+I^i_p(t)(1-\mu)\right\} \\ &+\left[1-\Gamma_{np\to p}(t)\right]\left\{S^i_{np}(t)q^i_{np}(t)+I^i_{np}(t)(1-\mu)\right\}. \qquad (3.12) \end{aligned}$$

In the above equations, the terms in brackets describe the transitions between the epidemic compartments, while the others control the behavioural transitions. The two quantities $q^i_p(t)$ and $q^i_{np}(t)$ express the probability that agent i will get infected at time t given the adopted strategy p or np. These probabilities are given by

$$q^i_p(t) = 1-\prod_{j=1}^{N}\left\{1-A_{ij}\lambda\gamma\left[I^j_p(t)+I^j_{np}(t)\right]\right\} \qquad (3.13)$$

$$q^i_{np}(t) = 1-\prod_{j=1}^{N}\left\{1-A_{ij}\lambda\left[\gamma I^j_p(t)+I^j_{np}(t)\right]\right\}. \qquad (3.14)$$

3.1.2 *Sustained epidemic cycles*

As a first step, we numerically solve the dynamic equations to analyse the temporal evolution of the system. We observe in Fig. 3.2A that the system exhibits a limit cycle for both the prevalence and the fraction of protected individuals. Looking carefully into the temporal evolution, we see that peaks in the prevalence and the fraction of protected individuals alternate with each other. The evolution of the payoffs for both strategies, shown in Fig. 3.2B, unveils the driving force behind this temporal pattern. Focusing on the grey-shaded area, we see that high prophylaxis levels result in low prevalence, reducing infection pressure. Further, in a low prevalence environment, the

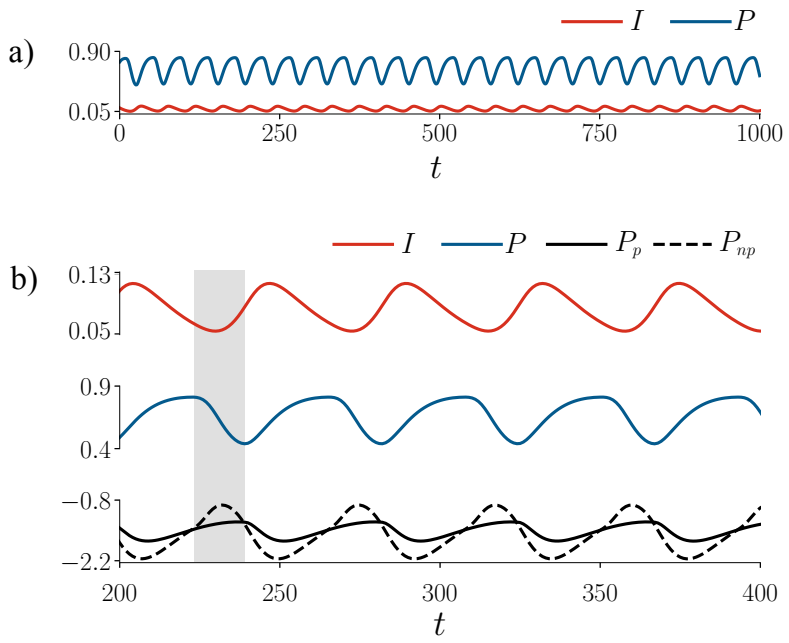


Figure 3.2: Numerical results of the risk-driven epidemic spreading model on a power-law network of size $N = 2000$ and exponent 2.5. Default parameters are $c = 1$, $\mu = 0.1$, $T = 10$, $\lambda = 0.05$, $\alpha = 0.1$, and $\gamma = 0.1$. **(A)**: Fraction of protected ($P = S_p + I_p$) (blue upper line) and infected ($I = I_p + I_{np}$) (red lower line) individuals as a function of time. We observe an oscillatory behaviour that is sustained in time. **(B)**: Detail of the oscillations. The red (top plot) and blue (middle plot) lines indicate the fraction of infected and protected individuals, respectively. In the bottom plot, the black dashed line plots the payoff of the *not protected* strategy (P_{np}), while the solid black line is the payoff of the *protected* strategy (P_p).

benefit of the prophylactic tool does not outweigh its cost ($P_{np} > P_p$). Accordingly, the protection level decreases over time, causing a surge in prevalence. Eventually, the increasing epidemic pressure makes adopting the prophylactic tool beneficial ($P_{np} < P_p$), adoption increases, and prevalence decreases. The repetition of this process sustains the limit cycle over time. In other words, it leads to a reproduction number that oscillates around 1, similar to the illustration in Fig. 3.1.

Given the presence of limit cycles, the natural question is where these cycles exist in the parameter phase space. Let us first focus on how the infection cost affects the amplitude of the limit cycle. As one would expect, on the left in Fig. 3.3A, an increasing infection cost decreases the prevalence level. However, as shown on the right in Fig. 3.3A, the limit cycle's relative

amplitude increases with the infection cost. This increase in the relative amplitude with the infection cost is further corroborated in Fig. 3.3B. Despite the increase in relative amplitude, the region in the $\lambda - \gamma$ phase space in which the limit cycle is present reduces with an increasing infection cost. In this phase space, the limit cycle only exists for low infectivity and high efficacy. Intuitively, we can understand this observation by considering the impact of the adoption of prophylactic tools on the epidemic evolution. The impact of

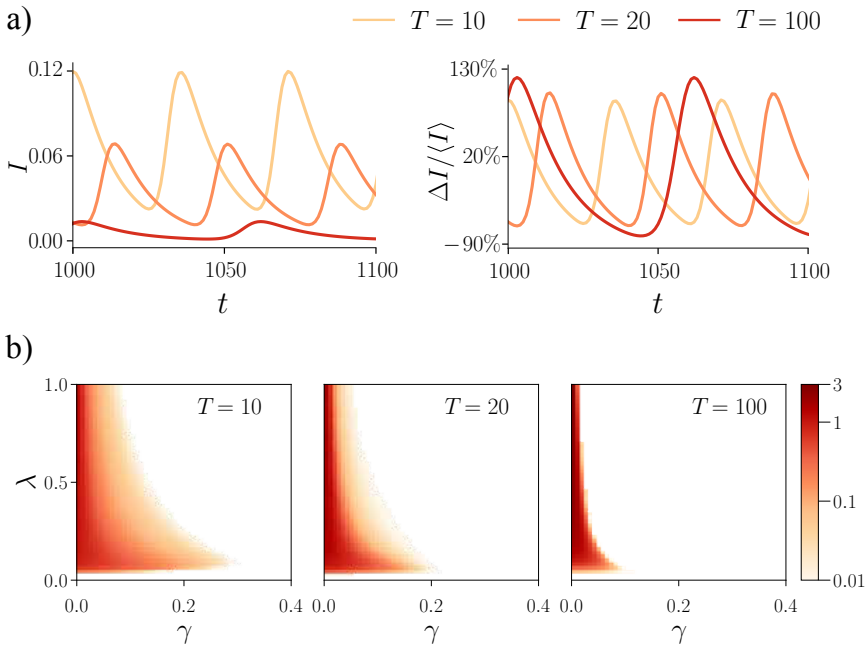


Figure 3.3: **(A)**: Amplitudes of the oscillations in the fraction of infected individuals as a function of time (darker colours represent higher values of T). The left plot depicts the absolute value of the amplitude, while the right one depicts the relative one, all of them for three values of T, the cost of infection. Higher values of T (higher cost) generate smaller oscillations (see left plot). On the right, we can see that all amplitudes of the oscillations are of the same order of magnitude when calculated relative to the fraction of infected individuals. **(B)**: Average relative amplitudes of the fraction of infected individuals in the steady state, for all ranges of γ (the probability of protection failure) and λ (the infectivity rate), for different values of T. We observe that as the infection cost increases, the area where the oscillations are present is reduced, but the oscillations themselves are larger. Numerical results of the risk-driven epidemic spreading model obtained on a power-law network of size $N = 2000$ and exponent 2.5. Parameters are $c = 1$, $\mu = 0.1$, $T = 10$, $\lambda = 0.05$, $\alpha = 0.1$ and $\gamma = 0.1$.

the prophylactic tool depends on its efficacy and the virus's transmissibility. To be more precise, the per-contact efficacy $1-\gamma$ relative to the transmissibility determines to which extent the adoption of a prophylactic tool reduces prevalence. Suppose the efficacy is low with respect to the transmissibility. In that case, the adoption of the prophylactic tool by a part of the population only has a limited impact on the population-level prevalence. Therefore, any temporal changes in the prophylaxis level will not trigger substantial variations in the prevalence. In other words, it essentially reduces the feedback loop between the behavioural and the epidemic dynamics, impeding a sustained limit cycle. Accordingly, the amplitudes are damped, and the system eventually converges to a stationary state.

In addition to the epidemiological and behavioural parameters, the time scale between the two dynamics defines whether the system either exhibits a limit cycle or reaches a stationary state. Fig. 3.4A & B show how increasing α leads to the emergence of a limit cycle. While the amplitudes smoothly in-

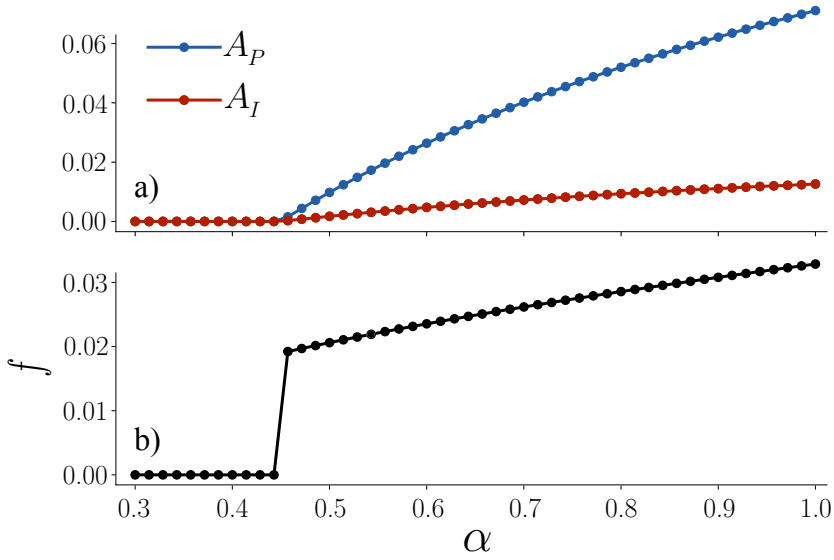


Figure 3.4: **(A)**: Amplitudes of the oscillations in the fraction of infected and protected individuals, respectively, as a function of the parameter α. **(B)**: Associated frequency f of the epidemic cycles. Numerical results of the risk-driven epidemic spreading model obtained on a power-law network of size $N = 2000$ and exponent 2.5. Parameters are $c = 1$, $\mu = 0.1$, $T = 10$, $\lambda = 0.15$, and $\gamma = 0.15$.

crease (A), the frequency of the limit cycle exhibits an abrupt transition (B). This result highlights that the behavioural dynamic needs to evolve rapidly to sustain the limit cycle. The phenomenology is similar to the one in the evolutionary vaccine game proposed in Ref. [270]. The rapid behavioural adaption leads to a constant overshooting in the uptake or the discontinuation of prophylactic tools with respect to the actual epidemic pressure. Due to the inertia of the epidemic dynamics, individuals decide according to an epidemic pressure which does not yet reflect the current uptake of prophylactic tools. In this sense, the delay until the behavioural adaption impacts the epidemic evolution sustains the limit cycle. Essentially, this delay acts in a similar way as the infection kernel in Ref. [275]. For this reason, considering explicitly the coupling between the behavioural and the epidemic dynamics, limit cycles are sustained even in a Markovian framework in which individuals have no memory of past states of the system.

3.1.3 *Time average of the macroscopic state*

After analysing the temporal evolution of the dynamics, let us now consider the time average of the macroscopic states, in particular, the prevalence I and the fraction of individuals adopting the prophylactic tool P. Fig. 3.5 shows that the system exhibits different dynamical regimes with respect to the dependence on λ. Independent of the efficacy, the prevalence increases with λ before any individual adopts the prophylactic tool ($P = 0$). However, after the initial increase in prevalence, the qualitative dependence on λ is shaped by γ. In the trivial case of 0 efficacy ($\gamma = 1$), prophylactic tools are not adopted, and we recover the standard second-order phase transition of the SIS model. If the prophylactic measure offers actual protection, the phenomenology is much richer. We find a non-monotonous dependence on λ for both I and P. For perfect protection ($\gamma = 0$), after the initial increase, I exhibits a maximum while P continuously increases. The case of an imperfect prophylactic tool is more delicate. As the virus becomes more transmissible, the actual benefit of adopting an imperfect prophylactic tool decreases. At some point, individuals get infected regardless of whether they protect themselves. Hence, we observe a non-monotonous dependence of P with respect to λ, and individuals eventually abandon the prophylactic behaviour.

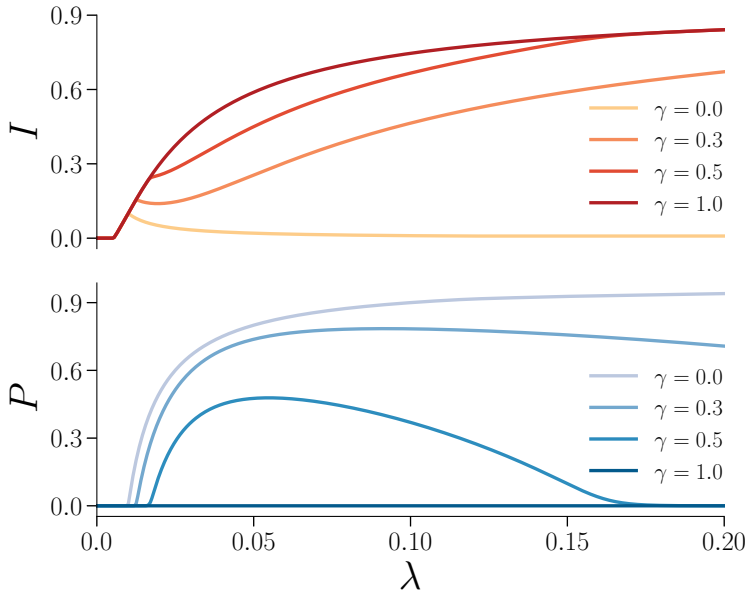


Figure 3.5: Fraction of infected individuals ($I = I_{np} + I_p$) and fraction of protected individuals ($P = S_p + I_p$) in the steady state as a function of the epidemic infection probability λ, for different values of the probability of preventive measures failing (γ). Numerical results of the proposed model obtained on a power-law network of size $N = 2000$ and exponent 2.5. Parameters are $c = 1$, $\mu = 0.1$, and $T = 10$.

3.1.4 *Calculating the epidemic threshold*

We pointed out before that individuals do not protect themselves if the prevalence is low. This behaviour suggests that, within our framework, the epidemic threshold remains unaffected by the voluntary uptake of the prophylactic tool. Intuitively, at the epidemic threshold, the epidemic pressure is effectively zero and thus also the benefit of the prophylactic tool. Accordingly, as highlighted in Sec. 2.5.1, the protection cost cannot be compensated, and individuals do not opt for prophylaxis. To prove rigorously that the epidemic threshold is the one of the standard SIS models, we consider the stationary state of the dynamical system. At equilibrium, the probability of an individual being protected, P^i, does not change between time steps, i.e. $P^i(t+1) = P^i(t)$. Imposing this condition, we find

$$P^i(t)\Gamma_{p\to np}(t) = \left(1 - P^i(t)\right)\Gamma_{np\to p}(t)\,. \tag{3.15}$$

Due to the Heaviside function in the transition probabilities $\Gamma_{p\to np}(t)$ and $\Gamma_{np\to p}(t)$, if both strategies do not yield the same payoff, only one of them is non zero. Therefore, assuming a non-trivial case where $P^i(t) \notin \{0,1\}$, the above condition can only be fulfilled if $P_p(t) = P_{np}(t)$. More specifically, this condition implies that in the stationary state, the behavioural and the epidemic dynamics are at equilibrium simultaneously. The condition in Eq. (3.15) translates as

$$\frac{c}{T} = \frac{I_{np}}{I_{np}+S_{np}} - \frac{I_p}{I_p+S_p}\,. \tag{3.16}$$

Let us first consider the case of perfect efficacy of the prophylactic measures. Since $I_p = 0$ for $\gamma = 0$, the condition in Eq. (3.16) can only be fulfilled if prevalence in the non-protected population equals the ratio c/T, which is certainly not true at the epidemic threshold. Due to the negative sign before the second term in the RHS, prevalence would need to be even higher in the non-protected population than c/T for $\gamma \neq 0$. More generally, at the epidemic threshold λ_c, the fraction of infected individuals is negligibly small, i.e. $I \approx \epsilon$. The condition is thus only verified for $c = 0$. Hence, a non-trivial equilibrium can only be reached at the epidemic threshold in the particular case without a protection cost. Discarding such artificiality, we conclude that the epidemic threshold is the one of the standard SIS models, i.e.

$$\lambda_c = \frac{\mu}{\lambda_{max}(A)}\,, \tag{3.17}$$

where $\lambda_{max}(A)$ refers to the spectral radius of the adjacency matrix A [173].

3.1.5 *Threshold for the behavioural adaption*

In analogy to the epidemic threshold, we define the protection threshold $\tilde{\lambda}$ for which individuals start to adapt their behaviour, and the fraction of protected individuals becomes non-zero. To calculate $\tilde{\lambda}$, we approximate the dynamics by considering a continuous-time well-mixed population of identical agents. The assumption of identical agents allows us to drop the index i and describe the dynamics through the four variables $S_p(t)$, $S_{np}(t)$, $I_p(t)$, and $I_{np}(t)$, where these quantities account for the fraction of agents in each com-

partment. Further, in continuous-time, the expressions of $q_{np}(t)$ and $q_p(t)$ in Eqs. (3.13) & (3.14) simplify to

$$q_{np}(t) = \lambda \left[\gamma I_p(t) + I_{np}(t)\right] \tag{3.18}$$
$$q_p(t) = \lambda\gamma \left[I_p(t) + I_{np}(t)\right] . \tag{3.19}$$

The differential equations governing the dynamics then read

$$\dot{S}_p = \mu I_p - \lambda\gamma S_p \left(I_p + I_{np}\right) + T_{np\to p} S_{np} - T_{p\to np} S_p \tag{3.20}$$
$$\dot{S}_{np} = \mu I_{np} - \lambda S_{np} \left(\gamma I_p + I_{np}\right) - T_{np\to p} S_{np} + T_{p\to np} S_p \tag{3.21}$$
$$\dot{I}_p = -\mu I_p + \lambda\gamma S_p \left(I_p + I_{np}\right) + T_{np\to p} I_{np} - T_{p\to np} I_p \tag{3.22}$$
$$\dot{I}_{np} = -\mu I_{np} + \lambda S_{np} \left(\gamma I_p + I_{np}\right) - T_{np\to p} I_{np} + T_{p\to np} I_p . \tag{3.23}$$

We see that the behavioural and epidemic dynamics are decoupled in the continuous-time description. Accordingly, as in the discrete-time model, we can consider the equilibrium condition for both dynamics separately. Further, since we focus on the protection threshold, we have $S_p \approx \epsilon$ and $I_p \approx \epsilon$. Accordingly, the equilibrium condition in Eq. (3.15) becomes

$$\begin{aligned}\frac{c}{T}\left(I_p + S_p\right) &= \frac{I_{np}}{I_{np} + S_{np}}\left(I_p + S_p\right) - I_p \\ &= I_{np}\left(I_p + S_p\right) - I_p + \mathcal{O}\left[\left(I_p + S_p\right)^2\right] .\end{aligned} \tag{3.24}$$

We will neglect second-order terms of S_p and I_p in the following. Accordingly, we can rewrite Eq. (3.15) as

$$I_p = \left(I_p + S_p\right)\left(I_{np} - \frac{c}{T}\right) . \tag{3.25}$$

Due to the conservation equation $I_p + S_p + I_{np} + S_{np} = 1$, we discard the redundant variable S_{np}. According to Eqs. (3.20) & (3.22), the equilibrium conditions for S_p and I_p become

$$S_p = \frac{I_p}{I_{np}} \frac{\mu}{\gamma\lambda} \tag{3.26}$$
$$I_p = I_{np} \frac{1 - I_{np} - \frac{\mu}{\lambda}}{I_{np}\left(1 + \gamma\right) - \gamma + \frac{\mu}{\gamma\lambda}} . \tag{3.27}$$

Inserting Eqs. (3.26) & (3.27) into Eq. (3.25), we find a quadratic equation for I_{np} as

$$I_{np}^2 - I_{np}\left(1 + \frac{c}{T} - \frac{\mu}{\gamma\lambda}\right) - \frac{c}{T}\frac{\mu}{\gamma\lambda} = 0\,. \tag{3.28}$$

The quadratic equation has only one non-negative solution, given by

$$I_{np} = \left[-1 - \frac{c}{T} + \frac{\mu}{\gamma\lambda} \sqrt{\underbrace{\left(1 + \frac{c}{T} - \frac{\mu}{\gamma\lambda}\right)^2 + 4\frac{c}{T}\frac{\mu}{\gamma\lambda}}_{\Delta}}\right]. \tag{3.29}$$

The protection threshold $\tilde{\lambda}$ is reached as I_p becomes positive. The denominator in Eq. (3.27) is always positive since

$$I_{np}\,(1+\gamma) - \gamma + \frac{\mu}{\gamma\lambda} = \frac{1}{2}\left[1 - \gamma + \frac{\mu}{\lambda}\left(\frac{1}{\gamma} - 1\right) + (1+\gamma)\left(\frac{c}{T} + \sqrt{\Delta}\right)\right] > 0\,. \tag{3.30}$$

Hence, the condition $I_p > 0$ is fulfilled whenever the numerator is positive, i.e. $I_{np} < 1 - \frac{\mu}{\lambda}$, where naturally the term $1 - \mu/\lambda$ is the prevalence of the SIS model in the absence of prophylactic behaviour. This condition further highlights why the epidemic threshold is unaffected by voluntary prophylaxis. In a well-mixed population, the epidemic threshold is given by $\lambda_c = \mu$. Therefore, at the epidemic threshold, the condition $I_p > 0$ reads as $I_{np} < 0$, which is never satisfied. Inserting the expression of I_{np} in Eq. (3.29) into the critical condition $I_{np} = 1 - \frac{\mu}{\tilde{\lambda}}$, we find after some tedious algebra the protection threshold $\tilde{\lambda}$ as

$$\tilde{\lambda}^{\pm} = \frac{2\mu}{1 - \frac{c}{T} \mp \sqrt{\left(1 - \frac{c}{T}\right)^2 - 4\frac{c}{T}\frac{\gamma}{1-\gamma}}}\,. \tag{3.31}$$

Interestingly, the protection threshold has two solutions: $\tilde{\lambda}^+$ and $\tilde{\lambda}^-$. The existence of two solutions is in line with our discussion of Fig. 3.5. The threshold $\tilde{\lambda}^-$ describes the critical value at which the epidemic pressure is sufficiently high such that the reduced infection risk outweighs the protection cost. In contrast, $\tilde{\lambda}^+$ represents the critical transmission probability at which individuals stop adopting the prophylactic tool since the infection risk is too high with respect to efficacy. Perfect efficacy is a particular case since prophylaxis fully protects individuals independently of the infection risk. Accordingly, the critical value $\tilde{\lambda}^+$ does not exist for $\gamma = 0$, i.e. $\lim_{\gamma\to 0} \tilde{\lambda}^+ = \infty$.

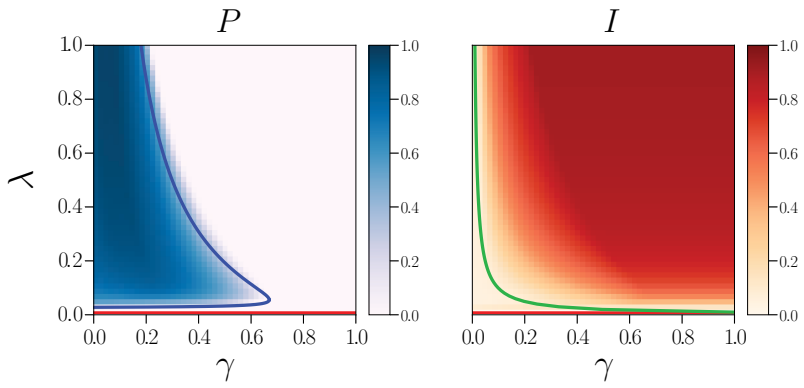


Figure 3.6: Phase-space diagrams of the prevalence on the number of Protected (left) and the number of Infected individuals (right). The red line in the left plot denotes the epidemic threshold of our model, as calculated in Eq. (3.17). The blue line on the left plot is the protection threshold obtained in Eq. (3.31). The green line indicates the epidemic threshold in the case of a fully protected population. Numerical results of the proposed model obtained on a power-law network of size $N = 2000$ and exponent 2.5. Parameters are $c = 1$, $\mu = 0.1$, and $T = 10$.

3.1.6 *Exploring the phase space*

In Fig. 3.6, we see that both the epidemic threshold and the protection threshold are in good agreement with the numerical solution of the recurrence equations. Please note that we rescaled $\tilde{\lambda}^{\pm}$ with respect to the average degree of the network. Additionally, we observe that a lower efficacy can actually promote the adoption of prophylaxis. The phenomenology behind this is detailed in Ref. [313]. Fig. 3.7 shows the time average of all four compartments. We see that as the transmission probability increases, the individuals transfer first from compartment S_{np} to S_p. Since the transmission probability is still relatively low compared to the efficacy of the prophylactic tool, the spread of the virus is still under control and prevalence (I_p and I_{np}) is low. As the transmission probability increases further, the fraction of protected individuals that are infected starts to increase substantially. Due to the increasing infection risk despite the prophylactic practice, individuals eventually abandon the prophylactic tool, and most of the population is in compartment I_{np}.

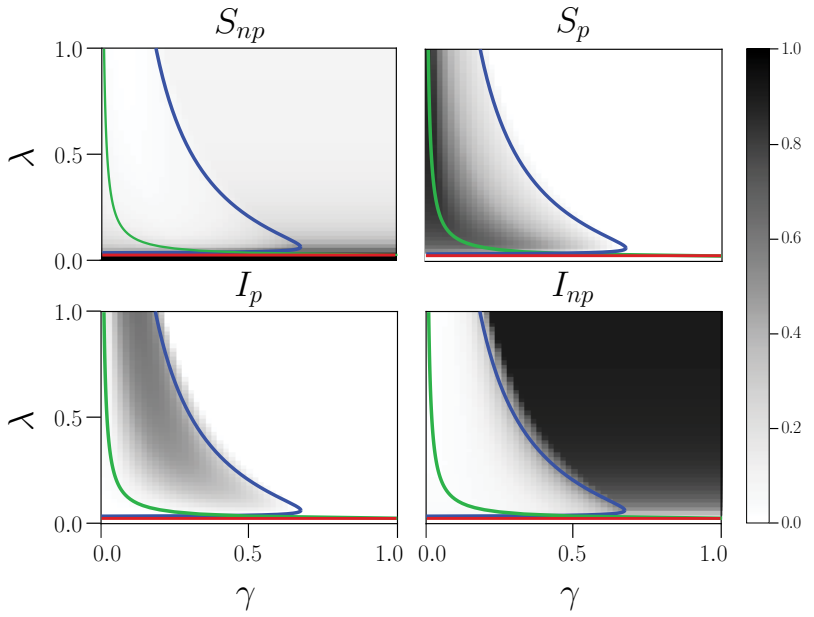


Figure 3.7: Phase-space diagrams for the fraction of the population in each of the four compartments of our model, at the steady state, for all range of λ and γ. Numerical results of the proposed model obtained on a power-law network of size $N = 2000$ and exponent 2.5. Parameters are $c = 1$, $\mu = 0.1$, and $T = 10$.

3.1.7 *Intervention policies*

We have now extensively analysed the phenomenology of the presented model. Naturally, the question arises whether we can leverage this model to gain insight into intervention policies. More specifically, we focus here on awareness campaigns that aim at promoting voluntary prophylaxis. We explicitly discard any policies that impose the use of a prophylactic tool through mandates. Furthermore, we discard technological advances that may increase adoption via improved effectiveness. Hence, considering the model's parameters, we realise that mainly the infection and protection costs can be influenced to promote adoption. While the protection cost can be reduced through increased availability or a price reduction, the infection cost can be increased through increased awareness about the severity of a disease. In contrast, the epidemic parameters μ and λ are inherent to the infectious disease and are not affected by the prophylactic tool.

As we pointed out before, the ratio c/T between protection and infection costs uniquely parametrises the behavioural dynamics, allowing us to subsume interventions that affect either the infection or the protection cost into the same category. Accordingly, even though we will only consider interventions that impact the infection cost, the same results hold with respect to the protection cost. Regarding the time-averaged quantities, changes in the infection cost affect the dynamics trivially; any increase in the infection cost will promote prophylaxis and thus reduce prevalence. However, the analysis is not straightforward if we consider the temporal evolution. To make an analogy with SARS-CoV-2, we all experienced how many governments focused on reducing the peak load on the healthcare system. In our framework, this equals reducing the peak prevalence of the limit cycle.

We will consider two types of intervention policies: (i) continuous campaigns that are sustained in time and (ii) pulsed campaigns that are only activated if prevalence is increasing. We assume that the campaigns make

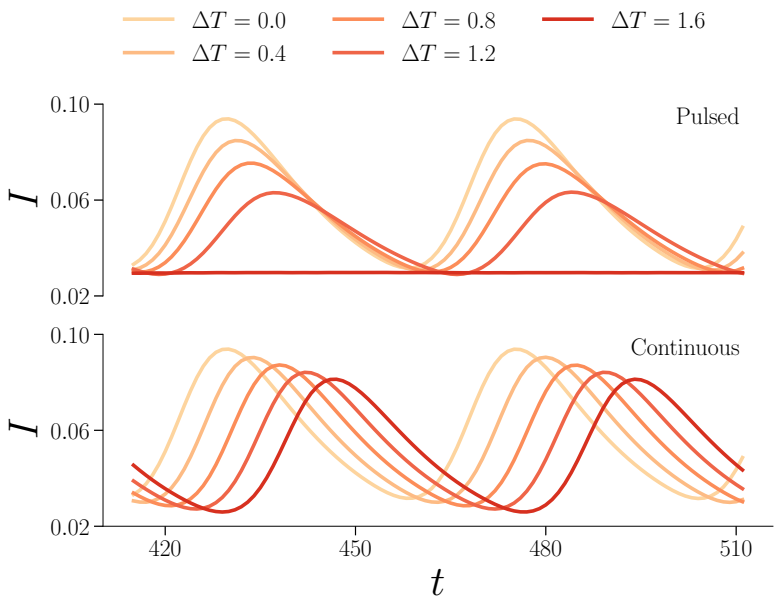


Figure 3.8: Time evolution of the infection prevalence in the presence of a pulsating awareness campaign (top) and a continuously sustained awareness campaign (bottom), for different values of the perceived risk increment ΔT. The contact network used here is a power-law network of size $N = 2000$ and exponent 2.5. Parameters are $c = 1$, $\mu = 0.1$, $\lambda = 0.05$, $\gamma = 0.1$, and $T = 10$. The pulsating campaign allows more efficient suppression of the peaks in the infection prevalence than the sustained campaign.

individuals more aware of the disease severity and the possible dangers of onward transmission in their social sphere. Specifically, within our model framework, the awareness campaigns increase the infection cost by T to $T + \Delta T$. While for the continuous campaigns, the infection cost is constantly $T + \Delta T$, the pulsed campaigns exhibit an infection cost of $T + \Delta T$ if prevalence increases and an infection cost of T if prevalence decreases. Further, to correctly bound the transition probabilities, we adapt the normalisation in the transition probabilities to $c + T + \Delta T$ for the continuous campaign and to $c + T + \Delta T\Theta(I(t) - I(t-1))$ for the pulsed one.

Surprisingly, we observe in Fig. 3.8 that the pulsed campaign performs better in reducing the peak prevalence than the continuous campaign. This result seems counterintuitive since a continuous awareness campaign requires more resources than a pulsed one. To elucidate the mechanism that makes pulsed campaigns more effective than continuous ones, we detail in Fig. 3.9 the evolution of I, P and the normalised payoff difference starting from $t = 0$. Initially, the epidemic is expanding. Hence, the pulsed campaign is active, and the dynamics evolve equally for both campaigns. As individuals increase prophylaxis, prevalence eventually starts decreasing (label 1). At this point, the pulsed campaign is deactivated (while the continuous campaign remains in place), decreasing the infection cost and, thus, the benefit of adopting prophylaxis. Therefore, since the payoff difference is directly proportional to the transition probabilities, prophylaxis increases faster in the continuous campaign than in the pulsed one. Due to the lower infection cost, abstaining from prophylactic behaviour becomes more beneficial than adoption (label 3), first, for the pulsed campaign and, later, for the continuous one. In addition to the delay, prophylaxis is less widespread at its peak for the pulsed campaign (label 4). In turn, the increased prophylaxis in the continuous campaign leads to a lower minimal prevalence than for the pulsed one (label 5). As prevalence increases again due to decreasing prophylaxis levels, the pulsed campaign is activated, increasing the payoff associated with strategy P (label 6). Eventually, due to the expanding epidemic, individuals start adopting prophylaxis again (label 7). We see that the minimal prophylaxis level in the pulsed campaign is substantially higher than in the continuous one. Hence, the pulsed campaign can control the expanding epidemic quicker, leading to a lower peak prevalence (label 8). From here, steps 1 to 8 are repeated periodically.

In a nutshell, both the continuous and pulsed campaigns sensitise the population about the disease severity, which translates into an increased infection cost. The pulsed campaign showed to damp the amplitude of the limit cycle

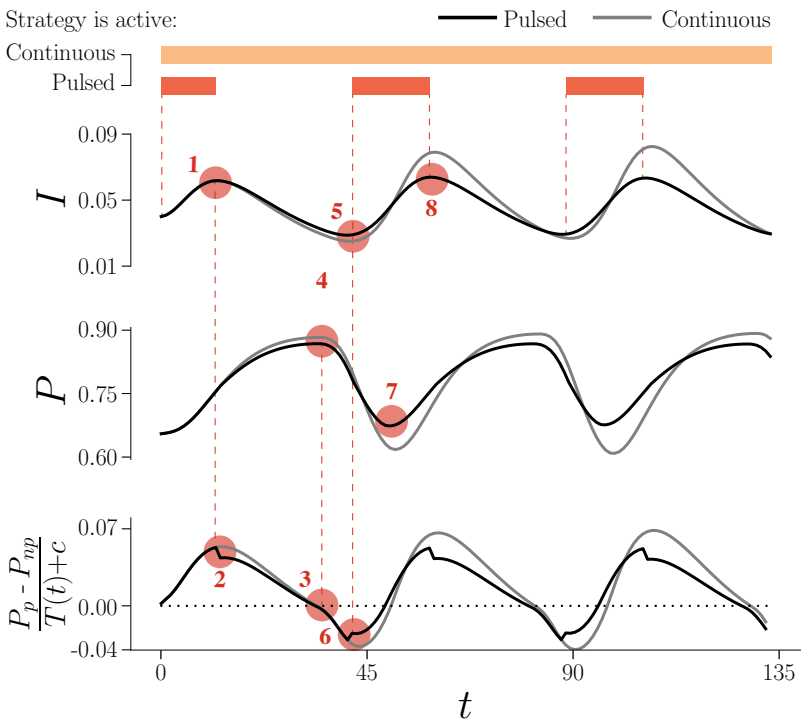


Figure 3.9: Infection prevalence (top), protection level (middle) and normalised payoff difference (bottom), as a function of time for $\Delta T = 1.2$. For the pulsed intervention, we define $T(t) = T + \Delta T \Theta[I(t) - I(t-1)]$. In the case of a continuous intervention, we have $T(t) = T + \Delta T$. The contact network is a power-law network of size $N = 2000$ and exponent 2.5. Parameters are $c = 1$, $\mu = 0.1$, $\lambda = 0.05$, $\gamma = 0.15$, and $T = 10$.

efficiently. Countercyclical on/off switching of the campaign reduces behavioural overshooting in response to the evolving epidemic. In other words, the pulsed campaign reduces the inertia in the coupling between the two dynamics, which drives and sustains the limit cycle. Accordingly, a pulsed campaign may be more beneficial to (i) save resources and (ii) reduce peak prevalence. It must be said, however, that the model assumes an immediate activation as prevalence starts to increase for the pulsed campaign. The experience in handling the SARS-CoV-2 epidemic has shown that authorities struggle to take action against an expanding epidemic if the prevalence is still relatively low. In many cases, health authorities hesitated during the early expansive phases and only took action after weeks of epidemic growth [298, 314, 315]. For this reason, it may be questionable how feasible the implementation of a pulsed campaign actually is.

3.2 DISEASES WITH A HETEROGENEOUS RISK OF SEVERE INFECTION

The model we introduced provides an explanation of how risk-based prophylaxis can lead to recurrent epidemic cycles. Further, we showed that individuals do not need any memory about past states of the system for limit cycles to be sustained. The dynamics is fully Markovian. We pointed out that such coupling between prophylaxis and the epidemic spreading may explain a part of the epidemic waves we experienced during the SARS-CoV-2 pandemic in many countries. Obviously, this model is a very crude simplification. It only considers voluntary adoption, while many countries introduced a series of mandates to control the spread of SARS-CoV-2 [314, 316, 317].

Additionally, the model assumes identical agents. Regarding the infection cost, assuming they are equal for the entire population is not reasonable for many diseases. For example, COVID-19 shows a strong age-dependent severity [326–328]. While the course of infection is mostly mild for younger individuals, elderly people are subject to an elevated fatality risk. In our framework, the age-dependent severity would translate into an age stratification of the model, in which age strata have different infection costs. We have previously shown, and as intuition suggests, that a higher infection cost results

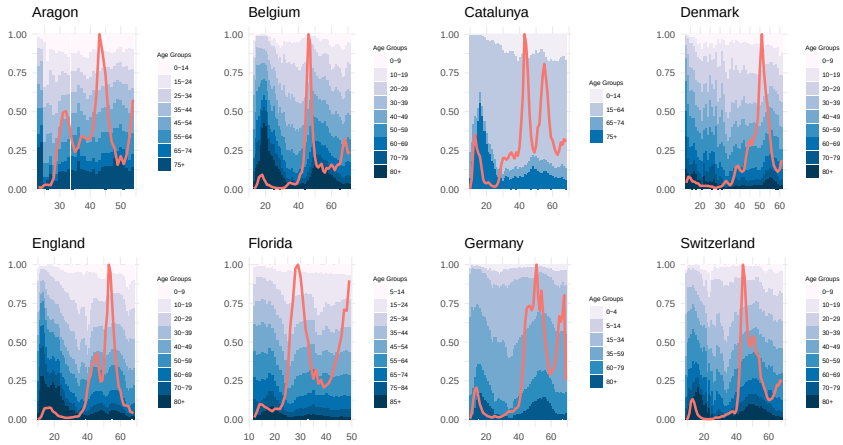


Figure 3.10: Weekly aggregated distribution of reported cases among the age groups for a variety of regions and countries [318–325]. The x-axis indicates time by enumerating the weeks of 2020 and continuing after 2021. The red line shows the evolution of the average weekly incidence, which serves as a proxy for prevalence, normalised to the peak incidence.

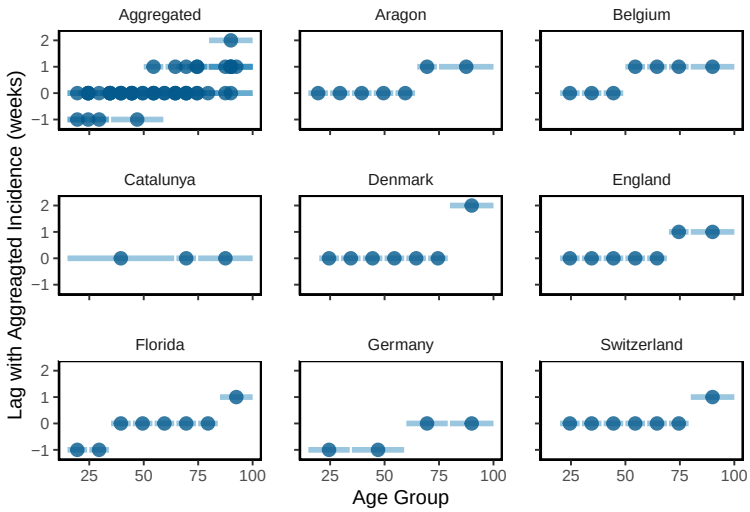


Figure 3.11: Temporal difference between the peak in the absolute number of infections of a particular age group and the entire population across all age groups. Lines show the range of the age groups, while dots indicate the mean between the limits. For all countries and regions, we show the second peak. An exception is Germany, whose second peak coincides with the Christmas holidays, which made reporting unreliable, wherefore we show the delay for the first wave. The top left plot reports the results from all countries and regions.

in more frequent prophylaxis. A series of studies have shown that elderly individuals actually adopted more intense and more frequent prophylaxis. The elderly showed intensified social distancing indicated by a stronger reduction of mobility [329–331], contacts [297, 306, 307, 332, 333], or credit card expenses [334] than the rest of the population.

The intensified social distancing, together with the generally lower number of contacts in the elderly, resulted in a lower seroprevalence than in the younger age groups [327]. Besides the overall infections, the case data exhibits an interesting temporal pattern with respect to age stratification. To be more precise, the data exhibits two phenomena: (i) the fraction of total cases attributed to the elderly increases towards the epidemic peak, and (ii) the absolute number of infections in the elderly have their maxima delayed with respect to the rest of the population. Fig. 3.10 illustrates the observation (i) for a series of countries and regions. We observe that the number of cases belonging to the older age groups steadily increases as the epidemic expands. Further, the maximal fraction of cases belonging to the older age

groups is reached substantially after the absolute number of infections has reached its peak (red curve). This difference directly connects with observation (ii) illustrated in Fig. 3.11. The data indicates that across countries and regions, the older age groups reach their maximal number of infections with a delay between one and two weeks with respect to the younger age groups. The same was observed for fatalities instead of case numbers [318, 335].

We have previously shown that a minimal model, which considers the coupling between the adoption of prophylaxis and the spread of an epidemic, can explain how recurrent epidemic waves may emerge. Naturally, we, therefore, ask ourselves whether we can adapt this previous model to exhibit not only recurrent waves but also the age-specific phenomenology discussed here. Our intuition is that introducing a heterogeneous infection cost is sufficient to explain the observed phenomenology. Therefore, in the following, we will adapt the model, introduce heterogeneity in the infection cost and analyse the implications for the dynamics.

3.2.1 *Differential equations for the dynamics*

In contrast to our previous modelling effort, the approach here attempts to explain a specific phenomenology that is exclusively observed in the SARS-CoV-2 epidemic. Since a SARS-CoV-2 infection generally leads to substantial immunity, we will consider an SIR model instead of an SIS. Additionally, for the time scale we consider here, one can assume that reinfections were not the main driver of the epidemic evolution [336], which allows us to simplify the model and disregard reinfections. We do not consider how contact tracing and isolation may affect the dynamics. Therefore, it is not necessary to distinguish between pre-symptomatic, asymptomatic and symptomatic infectiousness. The SIR model can capture the key ingredients of the epidemic dynamics. Additionally, the previous results have shown that the mean-field description of the model accurately captures the general behaviour. For the sake of simplicity, we will therefore consider a mean-field approach. Contacts are only structured according to the age-stratification, which is controlled by a contact matrix C.

Regarding the behavioural dynamics, we will only consider individual prophylaxis. As before, we disregard the role of central actors, such as health authorities and how their policies affect the population's behavioural response. Individual prophylaxis is manifold. It consists of physical distancing [337],

reduction of contacts [329], adoption of hygiene measures [338], regular ventilation [339], or the use of face masks [309], to name a few. However, from a mathematical point of view, all these different prophylactic tools simply reduce the transmission rate — either through a reduction of contacts or the transmission probability [18]. Hence, we subsume the different prophylactic practices into one category: the adoption of prophylactic measures. Adopting the prophylactic measures results in a reduced transmission rate β_i with efficacy ϵ_i for each age stratum i. We denote the two resulting strategies with P and NP for individuals who adopt those measures and individuals who do not, respectively.

We will slightly amend the definition of the payoffs. The dynamics of the SIR model do not reach an equilibrium as the SIS model. It is thus necessary to adapt the dynamics from Sec. 3.1. A natural choice for the initial dynamics is to introduce a small fraction of infected individuals in the absence of protection and analyse how individuals then adopt prophylactic measures in the face of an expanding epidemic. However, in the previous modelling effort, we defined the payoffs through the fraction of infected individuals for the respective strategy. The value of this fraction varies very rapidly if only a few individuals adopt strategy P. Therefore, we adapt the payoffs and multiply the infection costs by an effective epidemic pressure. For individuals adopting strategy NP, the epidemic pressure is simply the prevalence. In contrast, for individuals that adopt prophylaxis, we rescale the prevalence by the efficacy of the prophylactic measures. Precisely, the payoffs read

$$P_i^{NP} = -T_i \sum_{j=1}^{G} \sigma_j \left(I_j^{NP} + I_j^{P} \right) \tag{3.32}$$

$$P_i^{P} = -c_i - T_i(1-\epsilon_i) \sum_{j=1}^{G} \sigma_j \left(I_j^{NP} + I_j^{P} \right) . \tag{3.33}$$

The sum expresses the population's overall prevalence, where σ_i refers to the fraction of the population belonging to age stratum i. To formulate the model in all generality, we assume not only a different infection cost T_i for each age strata, but also separate protection costs c_i, transmission rates β_i,

and efficacies ϵ_i. As before, individuals adopt the most successful strategy according to the following transition probabilities:

$$\Gamma_i^{P\to NP} = \Theta\left(P_i^{NP} - P_i^P\right)\frac{P_i^{NP} - P_i^P}{c_i + T_i} \tag{3.34}$$

$$\Gamma_i^{NP\to P} = \Theta\left(P_i^{P} - P_i^{NP}\right)\frac{P_i^{P} - P_i^{NP}}{c_i + T_i}\,. \tag{3.35}$$

The transition probabilities are modulated by a selection rate α_i to control the time scale between the behavioural adaption and the epidemic dynamics. Coupling the behavioural model with the SIR spreading model leads to the following set of differential equations

$$\dot{S}_i^{NP} = -\beta_i S_i^{NP}\sum_{j=1}^{G} C_{ij}\left(I_j^{NP} + (1-\epsilon_j)I_j^P\right) + \alpha_i S_i^P \Gamma_i^{P\to NP} - \alpha_i S^{NP}\Gamma_i^{NP\to P} \tag{3.36}$$

$$\dot{S}_i^{P} = -\beta_i S_i^{P}(1-\epsilon_i)\sum_{j=1}^{G} C_{ij}\left(I_j^{NP} + (1-\epsilon_j)I_j^P\right) + \alpha_i S_i^{NP} \Gamma_i^{NP\to P} - \alpha_i S^{P}\Gamma_i^{P\to NP} \tag{3.37}$$

$$\dot{I}_i^{NP} = \beta_i S_i^{NP}\sum_{j=1}^{G} C_{ij}\left(I_j^{NP} + (1-\epsilon_j)I_j^P\right) - \mu_i I_i^{NP} \tag{3.38}$$

$$\dot{I}_i^{P} = \beta_i S_i^{P}(1-\epsilon_i)\sum_{j=1}^{G} C_{ij}\left(I_j^{NP} + \Gamma_i I_j^P\right) - \mu_i I_i^{P} \tag{3.39}$$

$$\dot{R}_i = \mu_i\left(I_i^{NP} + I_i^P\right)\,. \tag{3.40}$$

The parameter μ_i refers to the recovery rate in age-stratum i. G is the number of age strata. The variables S_i^P (I_i^P) and I_i^{NP} (I_i^{NP}) represent the fraction of susceptible (infected) individuals in age strata i that adopt strategy P and NP, respectively. As previously mentioned, the matrix entries C_{ij} control the interaction frequency between age group i and j. For the recovered department, we do not distinguish between strategies.

3.2.2 *Basic phenomenology of the model*

Before analysing the effect of the heterogeneous infection cost, let us first verify that, with the adapted payoffs, the model equally exhibits recurrent epidemic waves. We thus assume a homogeneous, well-mixed population with

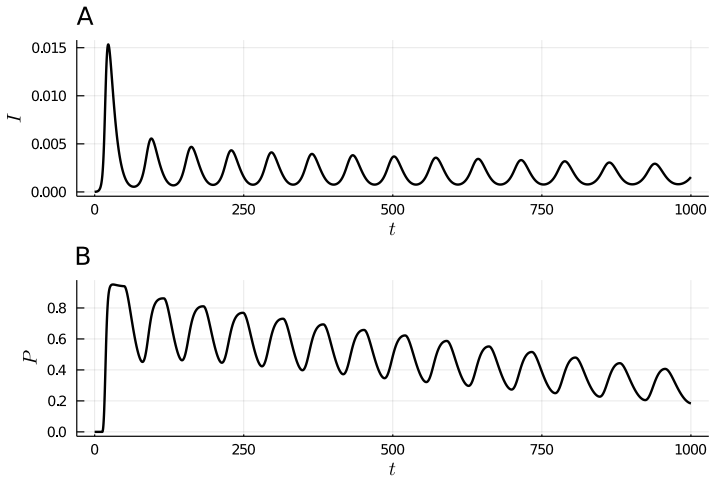


Figure 3.12: Prevalence or fraction of infected individuals, $I = I^{NP} + I^{P}$ **(A)** and fraction of protected individuals $P = S^{P} + I^{P}$ **(B)** evolving in time for a single realisation of the process, with initial conditions $I^{NP}(0) = 10^{-5}$ and the remaining fraction is attributed to the compartment $S^{NP}(0)$, with fixed $\alpha = 50$ and $T = 2000$. Remaining parameters are $\beta = 0.6$, $\mu = 0.2$, $\epsilon = 0.4$ and $c = 1$.

equal epidemiological and behavioural parameters. To study the system's dynamics, we numerically solve the differential equations. Fig. 3.12A & B show that the dynamics actually exhibits recurrent epidemic waves. Similarly to the original model, the epidemic evolution is driven by a time-varying prophylaxis level. Individuals constantly reduce and increase prophylaxis depending on the current epidemic pressure, which causes the reproduction number to be constantly "oscillating" around one.

Besides the temporal variation of the reproduction number, we have previously highlighted that the epidemic threshold is not affected by the voluntary adoption of prophylactic behaviour. The same holds in the current framework. Assuming that prophylaxis is absent at $t = 0$, the condition $P^{P}(t) > P^{NP}(t)$ must be fulfilled at some time t in order for individuals to adopt prophylaxis. This condition leads to a critical prevalence level I^{*} below which prophylaxis will not be adopted:

$$I^{*} = \frac{c}{\epsilon T}\,. \tag{3.41}$$

Accordingly, for prophylaxis to merge, the maximal prevalence I^{max} during the outbreak must exceed I^*. In the standard SIR model [18], the maximal prevalence is given by

$$I^{max} = 1 - \frac{1}{R_0}\left(1 + \ln(R_0)\right) . \tag{3.42}$$

The variable R_0 refers to the basic reproduction number of the disease. Combining the Eqs. (3.41) & (3.42), the ratio between the protection and infection cost must fulfil the following inequality for the behavioural response to materialise:

$$\frac{c}{T} < \epsilon \left[1 - \frac{1}{R_0}\left(1 + \ln(R_0)\right)\right] . \tag{3.43}$$

At the epidemic threshold, i.e. $\beta_c = \mu$, this condition is never fulfilled since the RHS is zero ($R_0 = 1$). Therefore, the epidemic threshold remains unaffected, as one would expect, given our previous results. Further, if one had several groups with different ratios c_i/T_i but equal transmission rates and random mixing across strata, the condition would equally hold. In such a case, the behavioural response would be triggered by the age strata with the lowest ratio c_i/T_i.

Fig. 3.12B shows that the protection level continuously decreases over time while the average prevalence stays relatively constant. The prevalence does vary less since the system evolves around the behavioural equilibrium $P^P = P^{NP}$. The structure of the payoffs fixes the prevalence at equilibrium by inducing the required prophylaxis level. As more individuals transfer into the recovered compartment, lower prophylaxis levels are necessary to maintain the prevalence at equilibrium, thus eventually lowering the average protection level.

Despite the relatively constant prevalence, we observe that the first wave exhibits a much higher peak prevalence than the subsequent ones. However, similarly to Ref. [340], whether the first wave exhibits the highest peak prevalence crucially depends on the parameter choice, particularly on the initial condition, as illustrated in Fig. 3.13. If the initial fraction of protected individuals (P_0) is low, the maximal prevalence is reached during the first wave. In contrast, if the initial protection is high, the maximal prevalence is reached during the second wave. For even higher protection, prevalence can rise over subsequent waves before it eventually decreases. The dependence on the initial condition may explain differences in how SARS-CoV-2 played out in different countries. In contrast to Italy, for example, countries like Switzerland

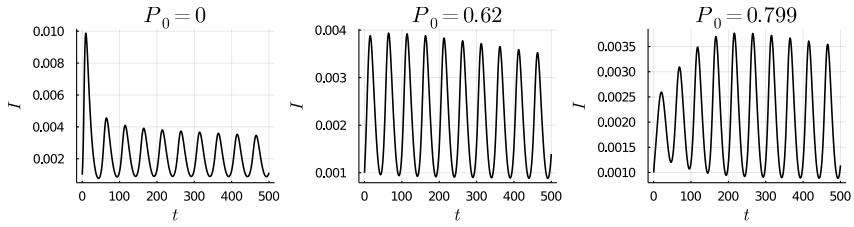


Figure 3.13: Early evolution of prevalence I for different values of $P_0 = S^{NP}(t = 0)$, the initial fraction of individuals that adopt strategy P. The fraction of initially infected individuals is $I^{NP}(0) = 10^{-5}$. The remaining fraction is attributed to the compartment $S^{NP}(0)$. Parameters are $\beta = 0.6$, $\mu = 0.2$, $\epsilon = 0.4$, $T = 1000$, $c = 1$ and $\alpha = 100$.

and Germany reached much higher prevalence in winter 2020/2021 than during the first wave in March 2020. Further, possibly due to the experiences in Italy, strong NPIs had been implemented in these countries relatively early compared to Italy, which would reflect the high protection level at the initial condition. In this sense, the dependence on the system's initial conditions would allow for an intuitive but partial explanation for the diverse early evolution of SARS-CoV-2 across countries.

3.2.3 *Added phenomenology in a heterogeneous population*

After verifying the basic phenomenology of the model in a homogeneous population, let us now consider the case of heterogeneous infection costs. As previously highlighted, a higher infection fatality rate (IFR) for the elderly, combined with stronger adherence to prophylactic measures, naturally suggests introducing different infection costs, T_i. To simplify things, we will separate the population into a high- and low-risk group that we refer to as young (Y) and old (O). Adding additional strata has not altered the phenomenology. To further reduce the complexity of the model, we assume the parameters c_i, β_i, μ_i, and ϵ_i are equal in the two groups. The two groups, Y and O, thus only differ in their infection cost. We parametrise the interaction between the two groups through a 2x2 contact matrix C. The parameter ν interpolates between random mixing, and complete modularity [73]. More precisely, given that the

fraction of the population σ_Y (σ_O) belongs to the group Y (O), the contact matrix is as follows

$$C_{YO} = \nu\sigma_O \tag{3.44}$$
$$C_{OY} = \nu\sigma_Y \tag{3.45}$$
$$C_{YY} = 1 - \nu\sigma_O \tag{3.46}$$
$$C_{OO} = 1 - \nu\sigma_Y\,, \tag{3.47}$$

where $\nu = 0$ corresponds to complete modularity, i.e. the two groups do not interact with each other. In contrast, $\nu = 1$ leads to random mixing. In the following, we assume that 20% of the population belongs to the group at risk, i.e. $\sigma_O = 0.2$ and $\sigma_Y = 0.8$.

In Figs. 3.14A & B, we observe that, as in the homogeneous population, recurrent epidemic waves emerge, which are driven by varying prophylaxis levels. The heterogeneous infection cost leads to higher protection levels in the group O and, thus, a lower prevalence level compared to the group Y. Furthermore, group O starts earlier to increase the prophylaxis level as the epidemic expands and group O decreases the prophylaxis level later as prevalence decreases. This difference is indicated by the dashed lines in Fig. 3.14D. The epidemic evolution is mainly driven by the group Y since they exhibit a much bigger variation of the prophylaxis level and the prevalence. Consequently, towards the epidemic peak, the rapidly increasing prevalence in group Y eventually spills over into group O. The reproduction number of the group O is below 1 at all time. Therefore, the prevalence in group O only increases due to the expanding epidemic in group Y, which leads to a delay in the prevalence peak of group O with respect to group Y as illustrated in Fig. 3.14D. In this sense, the model exhibits the same delay as in the data (Fig. 3.11).

Besides the delay, the model also exhibits the temporal variation in the case distribution (see Fig. 3.14E) that we observed in the data (Fig. 3.10). Considering Figs. 3.14C - E, we see how the fraction of cases belonging to the group O (Fig. 3.14E) is minimal as prevalence is low (Fig. 3.14D). At this point, the difference in the prophylaxis level between the two groups is very pronounced, wherefore the group Y is over-represented in the population level prevalence. In contrast, as prevalence increases (Fig. 3.14D), the difference in the prophylaxis level of the two groups Y and O is reduced (Fig. 3.14C). Accordingly, the epidemic's growth rates among both groups also become more

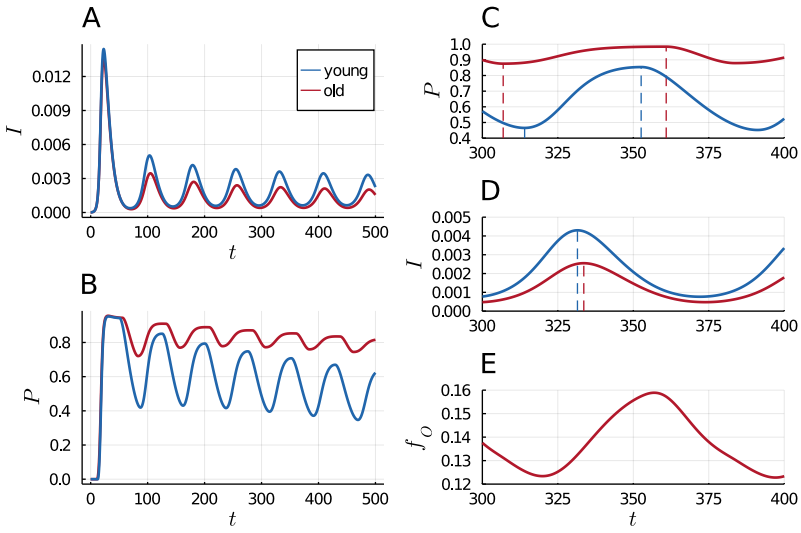


Figure 3.14: Left: prevalence or fraction of infected individuals, $I_i = I_i^{NP} + I_i^P$ **(A)** and fraction of protected individuals $P_i = S_i^P + I_i^P$ **(B)** in the old (red) and young (blue) groups, evolving in time, with initial conditions fixed as in Fig. 3.12 for both groups. The adoption rate is fixed at $\alpha = 50$, and the infection costs are $T_O = 2000$ and $T_Y = 500$, respectively. Right: fraction of protected individuals **(C)** and fraction of infected individuals **(D)** in both groups displayed in a shorter time window. In **(E)**, we show the fraction of cases belonging to the old group in time f_O. Parameters are $\beta = 0.6$, $\mu = 0.2$, $\epsilon = 0.4$ and $c = 1$.

similar (Fig. 3.14D). Thus, the fraction of cases belonging to the old group approaches the population distribution (Fig. 3.14E). The iteration of this process then leads to the repeating temporal pattern of the case distribution.

In a nutshell, the results presented in Fig. 3.14 demonstrate how the model exhibits an equivalent phenomenology to the data. In other words, the heterogeneous infection cost can explain both the delay in the epidemic peak and the temporal variation of the case distribution. We also verified that introducing different β_i does not alter the phenomenology. Further, different β_i but equal infection cost T_i does not lead to any of the two phenomena. Therefore, inside the framework of this model, the heterogeneous infection cost is the critical factor that qualitatively reproduces the observations in the data.

3.2.4 *Dependence of the dynamics on the different parameters*

While the phenomenology is robust regarding the parameters, how pronounced it is depends on the latter's values. In the following, we will analyse the impact of the *infection cost*, *the mixing rate* and the *adoption rate*. Figs. 3.15A & C show the ratio between the maximum and minimum in time of the fraction of cases belonging to the age group O as a function of the mixing rate and the ratio of the infection costs between the two groups. We can see that the temporal variation in the case distribution has a non-monotonous dependence on the infection costs for all values of the mixing rate ν ($\nu = 1$

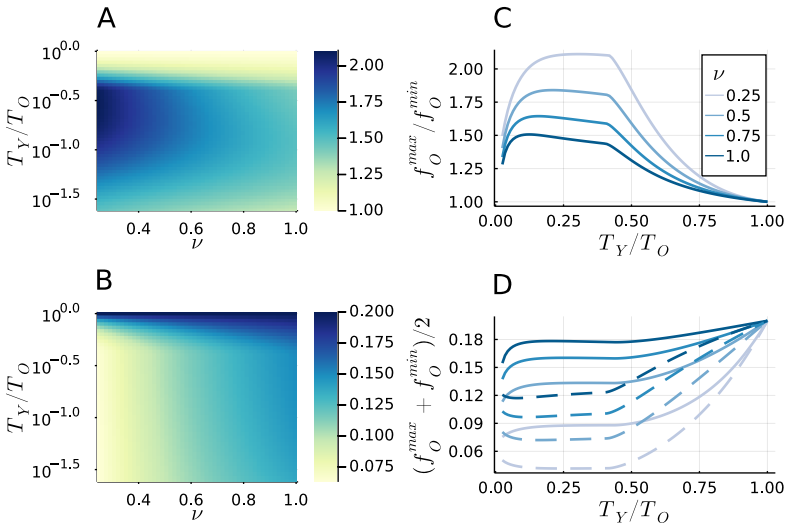


Figure 3.15: Top: Ratio between maximum, f_O^{max}, and minimum, f_O^{max}, in the variation of the case distribution, i.e. in the fraction of infected individuals belonging to the old group, f_O, in time depending on the mixing rate between the groups, ν, and the ratio of infection costs T_Y/T_O **(A)**. We also show the dependence on the infection cost in the young group, when the infection cost for the old group is fixed, $T_O = 2000$, for four choices of the mixing rate **(C)**. Bottom: Average between maximum, f_O^{max}, and minimum, f_O^{min}, in the variation of the case distribution, f_O depending on the mixing rate between the groups, ν, and the ratio of infection costs T_Y/T_O **(B)**. Further, we show that the maximal (solid) and minimal (dashed) fraction of cases belonging to the old group depends again on the infection cost of the young group T_Y, when the infection cost for the old group is fixed, $T_O = 2000$, for the same choices of the mixing rate **(D)**. The parameters are $\alpha = 0.50$, $\beta = 0.6$, $\mu = 0.2$, $\epsilon = 0.4$ and $c = 1$, and initial conditions $I^{NP}(0) = 0.0015$ in both groups.

corresponds to random mixing). If the infection costs are of similar value (high ratio T_Y/T_O), the dynamics approach the homogeneous case, and there is only slight variation in the case distribution. On the other hand, if the infection cost of the group Y is too low (low ratio T_Y/T_O), the increased prevalence causes the group O to keep prophylaxis for all time t. Thus, due to the limited behavioural evolution in the group O, there is only little temporal variation in the case distribution. Eventually, the balance between the two processes leads to the emergence of a maximum in the relative temporal variation of the case distribution. Further, the non-monotonous dependence of the ratio between infection costs becomes more pronounced as the mixing rate decreases. In other words, the better the elderly are shielded (lower ν), the more pronounced the temporal variation in the case distribution. This temporal variation may cause unexpected, additional stress on the healthcare system since the fraction of cases belonging to the elderly increases towards the epidemic peak. Accordingly, given the higher infection-hospitalisation risk of the elderly, the number of required hospital admissions at the epidemic peak may be substantially higher than those one would expect, considering the case distribution at low prevalence. Despite that the shielding increases the

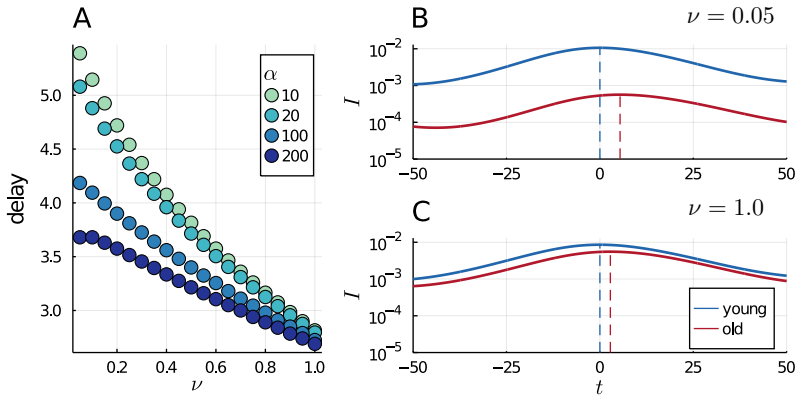


Figure 3.16: **(A)**: Dependence on the mixing rate for four choices of the adoption rate α. We use the same initial conditions as in Fig. 3.15. Other parameters are $T_O = 2000$, $T_Y = 500$, and in both groups $\alpha = 10$, $\beta = 0.6$, $\mu = 0.2$, $\epsilon = 0.4$ and $c = 1$. Right: zoom into the temporal evolution of the fraction of infected individuals for $\nu = 0.05$ **(B)** and random mixing, $\nu = 1.0$, **(C)**. Dashed lines indicate the position of the maximum. The adoption rate was fixed as $\alpha = 10$. Time, t, was shifted such that the peak of the age group young is located at zero.

relative variation in the case distribution, a lower ν always reduces the maximal and minimal fraction of cases that belong to the group O as shown in Fig. 3.15B & D.

Finally, Fig. 3.16A shows that the delay between the epidemic peaks monotonously decreases with the mixing as well as the adoption rate. We further illustrate in Fig. 3.16B & C how a more modular contact matrix (lower ν) delays the high prevalence in the group Y from propagating into the risk group O. In other words, better shielding of the risk population may cause an increased delay until the epidemic peak is reached in the age group O. Similarly, as in Sec. 3.1, a higher adoption rate α decreases the time interval at which recurrent epidemic waves emerge. Accordingly, it is not surprising that an increasing adoption rate reduces the delay between the epidemic peaks (Fig. 3.16B vs Fig. 3.16C) since less time is available for the prevalence to penetrate into the risk groups.

3.3 SUMMARY AND DISCUSSION

In this chapter, we have presented two relatively simple toy models that exhibit the basic phenomenology of recurrent epidemic waves, which was observed in a general matter (Sec. 3.1), and a temporal variation of the case distribution specific to SARS-CoV-2 (Sec. 3.2). The risk-based behavioural adaption in the presence of an epidemic can sustain recurrent epidemic waves over a wide range of parameters. Our models suggest that behavioural adaption does not suppress the epidemic but instead contains the epidemic via a reproduction number that varies around 1. These results are coherent with previous modelling efforts [273, 275–277], and are in line with the nature of the vaccination dilemma [266]. Overall, this naturally raises the question of whether a disease can actually be eradicated through the voluntary adoption of prophylactic tools or whether mandates imposed by a central actor such as the state are always necessary.

Compared to other approaches, the advantage of the proposed model here is its generality. It applies to a wide range of prophylactic tools. The framework can describe individual behaviour such as social distancing, reduction of contacts or the use of face masks. In contrast, many of the previous modelling efforts to capture the temporality of the evolution of a disease subject to a behavioural adaptation focused on specific prophylactic tools such as the use of vaccines [270] or the role of adaptive networks [282, 291]. It is worth

mentioning that, in the face of the SARS-CoV-2 epidemic, a series of studies were published, in which the behavioural adaption can lead to epidemic cycles [341–346]. These studies largely corroborate our results here.

Adding explicitly the behavioural dimension of the dynamics allowed us to explain the recurrent epidemic waves in a Markovian framework without memory. Approaches that bypassed the additional dimensions through an adaptation kernel require the system to have memory for cycles to emerge [275, 347]. In other words, the inertia of the interplay between disease and the behavioural dynamics in our model effectively acts as the system's memory. Only the dimensional reduction makes it necessary for the system to have memory to sustain epidemic cycles.

Our models may exhibit a phenomenology similar to that observed in the real world. However, we refer to them as toy models since they are not designed to match the data. Instead, their role is to deepen the basic understanding of the general mechanisms at play. Nevertheless, our extension of the basic model to incorporate heterogeneous infection costs has shown that even simple models can deepen the understanding in greater detail than one would think. Nevertheless, we are far from reproducing the data quantitatively and even further from making quantitative predictions about the dynamic interplay between human behaviour and the spread of an epidemic. As previously outlined, in the last chapter of this thesis, we will take a more quantitative approach and get our hands dirty with data. But first, let us focus on two examples in which human behaviour non-trivially impacts the spread of an epidemic, even if there is no dynamic interplay between behaviour and the epidemic.

4

HUMAN BEHAVIOUR AND EPIDEMIC CONTROL

In the previous chapter, we studied the interplay between behavioural adaption and the spread of an epidemic. The aspect of human behaviour we focused on explicitly resulted from the presence of the epidemic. However, even though the danger of epidemics is of great relevance for humans today and in their past [7], only a tiny part of human behaviour is a direct result of the presence of an epidemic [348, 349]. Most of the diverse aspects of human behaviour are not altered in the face of an epidemic. Nevertheless, many traits of human behaviour can still affect how an epidemic spreads, even though there might not be a direct dynamic interplay among them. In these cases, human behaviour does not co-evolve with the epidemic but crucially shapes how the epidemic spreads.

Understanding the impact of different behavioural traits on the epidemic spread is not only of interest to academics but also to health authorities in order to improve epidemic control, mainly through the use of prophylactic tools [350–353]. Moreover, deepening the knowledge on the impact of human behaviour on epidemic spread can facilitate health authorities to design efficient policies or, at least, more accurately estimate the impact of existing ones [354–356]. In line with such approaches, we will focus in this chapter on two fundamental aspects of human behaviour: homophily and heterogeneous contact patterns. To this aim, we will use simple mathematical models to study their impact on epidemic spreading for different prophylactic measures.

4.1 THE IMPACT OF HOMOPHILY ON EPIDEMIC CONTROL

The concept of homophily is often described by the popular expression "birds of a feather flock together" [357]. It refers to the fact that people mainly form

relationships and physically interact with others that share similar social characteristics. Decades of studies in sociology showed that categories like age, gender, or socio-economic class are robust predictors of interaction [357]. Within the framework of networks, it implies that similarity is associated with connection [358, 359]. In more modern terms, the metadata (categories) of the nodes (individuals) strongly correlates with the structure of the network [360, 361]. This correlation generally implies that the network is modular, i.e. that there is a community structure [73].

It is well established how modularity affects the spread of an epidemic on a network [144, 362, 363]. Less studied, though, is the case in which the metadata of the nodes also affects the epidemic dynamics. A typical example is the adoption of vaccines. In this case, the disease spreads on the contact network but is influenced by the nodes' metadata, indicating the individuals' vaccination status. Further, a series of studies have shown that the same predictors for connection are also predictors for the adoption of vaccines [364–366] and other prophylactic tools such as digital contact tracing apps [367–369]. In a broader scope, experimental results indicate that homophily significantly affects health behaviour [370, 371]. These findings suggest that prophylactic tools are not uniformly adopted in the population. Instead, adoption strongly correlates with the community structure induced by homophily.

In this section, we will first consider the case of digital proximity tracing and study how homophily affects its efficiency. Essentially, we will show how the presence of homophily can drive the system into three different dynamical regimes. In the second part, we then extend this formalism and show that the unveiled phenomenology also holds for prophylactic tools such as vaccines, the use of face masks or social distancing.

4.1.1 *Digital Proximity Tracing*

For many of us, the first time we heard the term *contact tracing* was in the context of the SARS-CoV-2 pandemic. Accordingly, one could think that contact tracing — the act of testing and isolating the contacts of infected individuals — is a recent invention. But, instead, contact tracing was used for a series of epidemics in the past, such as smallpox [372], tuberculosis [373], HIV [374], or Ebola [375], to name a few. Accordingly, many studies were done to investigate the efficiency of manual contact tracing [113, 372, 376–383].

What was new in the context of SARS-CoV-2 was the advent of digital proximity tracing (DPT) apps. Instead of manually collecting the contacts of infected individuals, an app records past contacts through the Bluetooth signal [384–387]. Hence, if an individual becomes infected, recent contacts can quickly be notified, tested and eventually isolated in an almost automatised way. However, for DPT to work, both the infected individual and the contacts must be adopters of this technology. Not surprisingly, around the implementation of DPT apps, many studies investigated the possible impact of DPT on the spread of SARS-CoV-2, and the necessary adoption levels to control the latter [367, 388–394]. Later, a series of theoretical studies that focused more on the physics behind DPT were published [395–398].

What has not been addressed is how the homophilic adoption of DPT apps affects the ability of DPT to control an epidemic. As we highlighted previously, the population does not adopt prophylactic tools uniformly. Studies on DPT have shown that adoption strongly correlates with age, income, and nationality [367–369]. Generally, adoption is particularly low among marginalised socio-economic classes [368]. Then, due to the similarity of social contacts (homophily), contacts of adopters are much more likely to be adopters than contacts of nonadopters. In Switzerland, for example, 70% of contacts among adopters were also found to be adopters, while the average national adoption was only around 20% [386]. Our goal is thus to build a minimal model that allows us to understand how the non-uniform adoption of DPT apps affects the spread of an epidemic.

4.1.1.1 Model equations

To describe the dynamics of DPT, we extend a model recently introduced by Bianconi et al. [396]. The model is relatively simple and analytically tractable, but it captures the main ingredients of the dynamics. For the epidemic spreading, the model considers a basic, discrete-time SIR model with transmission probability λ, infectious period τ and contact rate k. For convenience, we define $\beta = \lambda\tau$. To incorporate DPT, the model assumes that adopters infected by other adopters do not further transmit the disease. This assumption is equivalent to considering the case of perfect instantaneous contact tracing and self-isolation. For completeness, in Appendix A, we show that even when these assumptions are relaxed, the phenomenology is not altered.

Given the approach by Bianconi et al., it is necessary to incorporate the non-uniform distribution of DPT apps into the model. One could take the

approach of stratifying the population into all the different categories that are predictors for app adoption. However, this would yield a high model complexity and not allow for analytical tractability.

Since we are more interested in the phenomenology of the interplay between homophilic app adoption and the spread of an epidemic, we simplify the approach and subsume the diverse correlations in app adoption into a contact matrix K with parameter α. The parameter α interpolates from complete modularity ($\alpha = 0$) to random mixing ($\alpha = 1$) between adopters and nonadopters. We denote the entries of K as k_{ij} with $i, j \in \{\mathrm{A}, \mathrm{N}\}$, where A and N refer to adopters and nonadopters, respectively. The contact rate between adopters and nonadopters is given by $k_{\mathrm{AN}} = \alpha k(1-T)$, where T refers to the fraction of adopters in the population. The remaining entries of the contact matrix follow from the balance equation $Tk_{\mathrm{AN}} = (1-T)k_{\mathrm{NA}}$ and the average contacts $k = k_{\mathrm{AA}} + k_{\mathrm{AN}} = k_{\mathrm{NN}} + k_{\mathrm{NA}}$. Therefore, the entries of the matrix K are given by

$$k_{\mathrm{AN}} = \alpha(1-T)k \tag{4.1}$$

$$k_{\mathrm{NA}} = \alpha T k \tag{4.2}$$

$$k_{\mathrm{AA}} = [1 - \alpha(1-T)]\, k \tag{4.3}$$

$$k_{\mathrm{NN}} = [1 - \alpha T]\, k\,. \tag{4.4}$$

To describe the system's dynamics, we need to split the infectious compartment into three separate ones. First, we define $I_{\mathrm{N}}(t)$ as the number of newly infected nonadopters at generation t. Further, we define $I_{\mathrm{AA}}(t)$ and $I_{\mathrm{AN}}(t)$ as the number of newly infected adopters at generation t, which were infected by adopters and nonadopters, respectively. This distinction is necessary since adopters infected by other adopters do not further transmit the disease. For the susceptible compartment, such distinction is not necessary, and we only work with the number of adopters $S_{\mathrm{A}}(t)$ and nonadopters $S_{\mathrm{N}}(t)$ susceptible at generation t. With these definitions at hand, the equations describing the dynamics read

$$I_{\mathrm{N}}(t+1) = \beta\left[k_{\mathrm{NN}} I_{\mathrm{N}}(t) + k_{\mathrm{AN}} I_{\mathrm{AN}}(t)\right] \frac{S_{\mathrm{N}}(t)}{N_{\mathrm{N}}} \tag{4.5}$$

$$I_{\mathrm{AN}}(t+1) = \beta k_{\mathrm{NA}} I_{\mathrm{N}}(t) \frac{S_{\mathrm{A}}(t)}{N_{\mathrm{A}}} \tag{4.6}$$

$$I_{\mathrm{AA}}(t+1) = \beta k_{\mathrm{AA}} I_{\mathrm{AN}}(t) \frac{S_{\mathrm{A}}(t)}{N_{\mathrm{A}}} \tag{4.7}$$

$$S_N(t+1) = S_N(t) - I_N(t+1) \tag{4.8}$$
$$S_A(t+1) = S_A(t) - I_{AA}(t+1) - I_{AN}(t+1)\,, \tag{4.9}$$

where N_A and N_N refer to the total number of adopters and nonadopters, respectively, in the population. From there, we can calculate the final attack rate, i.e. the total number of recovered individuals over the course of the epidemic, by summing the number of newly infected individuals for all generations t.

4.1.1.2 Expression of the reproduction number

As the classical SIR model, the dynamical system does not allow for an explicit analytical expression for the final attack rate. Nevertheless, we can calculate the reproduction number of the dynamics through the next-generation matrix approach [211, 214]. The next-generation matrix NGM is then given by

$$\mathrm{NGM} = \beta \begin{pmatrix} k_{NN} & k_{AN} & 0 \\ k_{NA} & 0 & 0 \\ 0 & k_{AA} & 0 \end{pmatrix} . \tag{4.10}$$

To find the reproduction number, we insert the explicit expression of the entries of K and calculate the spectral radius of the matrix. One gets

$$R = \frac{R_0}{2}\left[1 - \alpha T + \sqrt{(1-\alpha T)^2 + 4\alpha^2 T(1-T)}\right] , \tag{4.11}$$

where R_0 refers to the basic reproduction number of the classical SIR model. As one would expect, R has a monotonous dependence on the adoption level T and the basic reproduction number R_0. More surprisingly, R exhibits a non-monotonous dependence on the mixing rate α. Straightforward calculations show that the criteria $\frac{dR}{d\alpha}|_{\alpha=0} < 0$ is always met. Accordingly, solving for $\frac{dR}{d\alpha} = 0$ we find the minimum of R with respect to the mixing rate α^* as

$$\alpha^* = 1 - \frac{\frac{2}{3} - T}{\frac{4}{3} - T}\,. \tag{4.12}$$

In the case that $\alpha^* \geq 1$, the reproduction number has its smallest value for random mixing between adopters and nonadopters, i.e. $\alpha = 1$. The condition $\alpha^* < 1$, implies that for all $T > T^* = 2/3$ no local minima exists in R with respect to α. Fig. 4.1A shows the dependence of the reproduction number on α and T. Besides the local minima, Eq. (4.11) provides a critical parameter

range in which eradication of the disease is possible, i.e. $R < 1$, by varying the mixing parameter α. More precisely, for $\alpha \in (\alpha_c^-, \alpha_c^+)$ the condition $R < 1$ holds, where

$$\alpha_c^{\pm} = \frac{1}{2R_0(1-T)} \left[1 \pm \sqrt{1 - 4\frac{1-T}{T}(R_0 - 1)} \right] . \tag{4.13}$$

The non-monotonicity of R with respect to α leads to the presence of two physical solutions $\alpha_c^{\pm}$. The existence of two solutions suggests that the disease can not only be eradicated as the mixing is increased from complete modularity towards random mixing but can also reemerge as the mixing rate crosses α_c^+.

4.1.1.3 Three dynamical regimes

Fig. 4.1B shows the final attack rate obtained numerically as a function of α and R_0. The solid line indicates the threshold α_c^-. While the attack rate monotonously increases as a function of R_0, we can distinguish three different

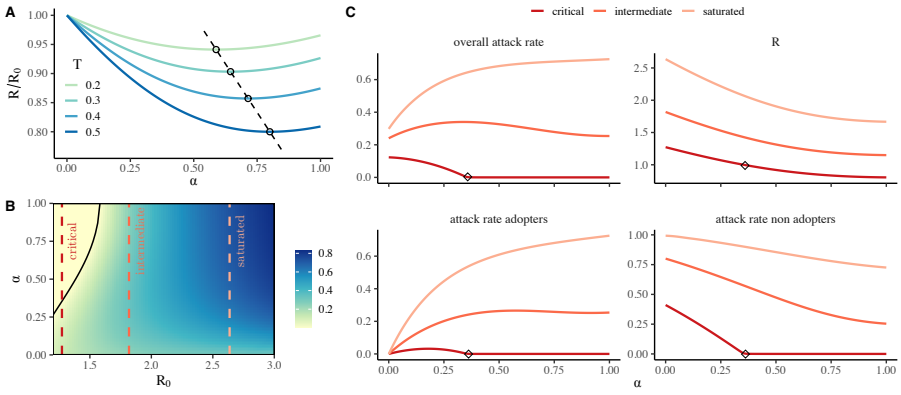


Figure 4.1: **(A)**: Reproduction number, R, normalised with respect to the basic reproduction number, R_0, for different values of adoption, T, below T^*. Dots indicate the minimum at α^*, while the dashed line shows its variation for $T \in [0.1, 0.55]$. **(B)**: Numerical solution of the dynamics for the attack rate as a function of α and R_0. The solid line indicates the threshold α_c^- for which $R = 1$. Coloured, dashed lines denote the dynamical regimes: critical, intermediate and saturated. Adoption was fixed as $T = 0.7$. For these parameter values, we have $\alpha_c^+ > 1$. **(C)**: Top panels show the attack rate and the reproduction number for the different regimes defined in B. The specific attack rates for adopters and nonadopters are reported in the bottom panels. Black diamonds indicate α_c^-, at which $R = 1$.

regimes for the dependence on α. The parameter R_0 then controls the regime in which the system is found. The coloured dashed lines indicate examples of these three regimes that we refer to as critical, intermediate and saturated. In the critical regime, close to the epidemic threshold, the attack rates monotonously decrease with respect to α. In contrast, in the intermediate regime, the attack rate exhibits a non-monotonous dependence on α. Finally, far from the epidemic threshold, in the saturated regime, the attack rates monotonously increase as a function of α.

The three distinct dynamical regimes emerge from the competition between the following two processes as the mixing (α) increases: a decreasing protection among adopters and an increasing protection among nonadopters. To illustrate these two processes, let us consider the case of complete separation between the cluster of adopters and nonadopters ($\alpha = 0$). In this situation, the epidemic spread is only slowed down in the cluster of adopters due to the presence of DPT apps since nonadopters do not interact with the latter. In contrast, as adopters and nonadopters start to interact with each other, the mixing rate increases. Nonadopters can then be considered free riders that take advantage of the protection provided by the adopters. The received protection is reflected by a decreasing attack rate among nonadopters as the mixing rate increases (Fig. 4.1C). At the same time, adopters become more vulnerable to infection as mixing increases. Individual adoption of DPT does not provide any protection since transmission can only be prevented if both the primary and the secondary infection concern an app adopter. Accordingly, as mixing increases, more transmission chains can arrive at adopters, increasing the attack rate among them as shown in Fig. 4.1C.

Given these two processes, the coverage T and the basic reproduction number of the disease (R_0) decide which one holds the upper hand. Far from the epidemic threshold, in the saturated regime, the protection that adopters can provide for nonadopters has little impact since the attack rate increases relatively slowly with the infectious pressure. Accordingly, the increased attack rate among adopters holds the upper hand and the overall attack rate increases with the mixing rate α. In contrast, close to the epidemic threshold, in the critical regime, the attack rate strongly varies with any change in the epidemic pressure. Hence, the protection provided by adopters to nonadopters compensates for the reduced protection among adopters and the overall attack rate decreases. Further, considering the attack rate among adopters (Fig. 4.1C), we see that, after the initial increase, the impact on the nonadopters eventually compensates for the lost protection among adopters leading

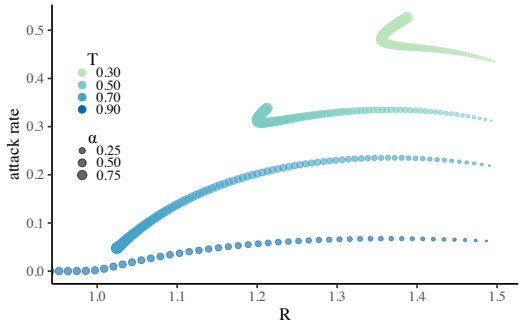


Figure 4.2: Attack rate as a function of the reproduction number, R, for different adoption values, T. The size of the points interpolates between $\alpha = 0$ and $\alpha = 1$. We fixed the basic reproduction number as $R_0 = 1.5$.

to a decrease in the attack rate as the mixing is increased further. Eventually, the system can be pushed below threshold, and the attack rate vanishes. Finally, in the intermediate regime, the system switches between the critical and saturated regime, causing the non-monotonous dependence on α.

In Figs. 4.1B & C the fraction of adopters T is fixed. To represent the effect of adoption, we fix the basic reproduction number and vary α for different values of T. Fig. 4.2 shows the attack rate as a function of the reproduction number. As highlighted previously, not only the basic reproduction number but also adoption can cause the system to change the dynamical regime. For $T = 0.9$, the system is in the critical regime. Then, for $T = 0.7$ and $T = 0.5$, the system is in the intermediate regime shown by the non-monotonous dependence of the attack rate on the mixing rate. Finally, for $T = 0.3$, the attack rate monotonously increases with the mixing rate, and thus, the system is in the saturated regime. Essentially, the adoption relative to the basic reproduction number defines whether the system is close or far from the epidemic threshold and thus pushes the system into one of the three dynamical regimes. Please note that for $T = 0.9$ and $T = 0.7$, the reproduction number exhibits a local minimum since $T > T^* = 2/3$.

4.1.1.4 Monte Carlo simulations

Up to this point, we assumed structured mean-field interactions. In other words, interactions occur randomly given the mixing rate α between adopters and nonadopters. However, real-world interactions exhibit more structure beyond the contact matrix K and can generally be represented as a net-

work [82]. Hence, we want to corroborate our theory through Monte-Carlo simulations on physical contact networks [144, 147, 399, 400]. The aim is to see whether the phenomenology, i.e. the three dynamical regimes, are also found when the dynamics evolves on a network. Since our model describes subsequent generations, we decided to aggregate real-world contact networks into static, binary ones. To this aim, we threshold the interaction weights, i.e. the duration of the interactions, while controlling that the network remains connected.

Given any empirical network, the question arises of how to distribute the apps on the network to reach a desired mixing rate α. To translate the concept of the mixing rate in the mean-field case to the contact network, let us first consider the probability that both individuals are either adopters or nonadopters during an interaction. We will refer to this probability as the homophily level h, which can be expressed as

$$h = \frac{1}{k}\left[Tk_{AA} + (1-T)\,k_{NN}\right] = 1 - 2\alpha T\,(1-T)\ . \tag{4.14}$$

Accordingly, we have established a relation between the homophily level h and the mixing rate α. Further, it is straightforward to calculate h on a network. Let us consider the adjacency matrix A and a vector v that defines whether node i is an adopter ($v_i = 1$) or not ($v_i = 0$). With these definitions, the homophily level can be expressed as

$$h = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} A_{ij}\left[v_i v_j + (1-v_i)(1-v_j)\right]}{\sum_{i=1}^{N}\sum_{j=1}^{N} A_{ij}}\ . \tag{4.15}$$

By inverting the relationship between h and α in Eq. 4.14, we can infer the value of α in the network via Eq. (4.15). We then employ an algorithm to distribute apps in the network until a fixed value of α is reached. More precisely, we start from a random distribution of apps ($\alpha = 1.0$) and randomly swap the roles of a pair of adopter and nonadopter. The swap is accepted if it lowers the value of α. We iterate this procedure until the desired value of α is reached within a tolerance of 0.02. Further, we check that the average degree of adopters and nonadopters is the same. Swaps are only accepted if it reduces the absolute value of the difference in average degree between the two groups. Without this constraint, the algorithm would induce pronounced

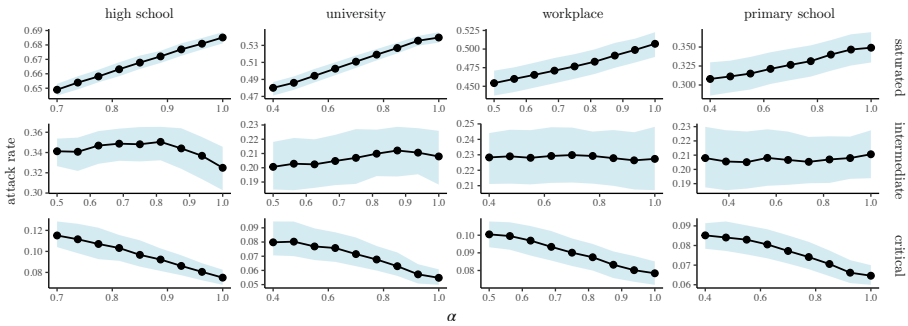


Figure 4.3: Results of Monte Carlo simulations for the dependence of the attack rate on α for the three dynamical regimes and four different real-world networks [144, 147, 399, 400]. Each column refers to a different network, while rows indicate the saturated, intermediate and critical regime, respectively, from top to bottom. Dots indicate the median value, whereas the ribbon indicates the first and third quartiles. Each point is obtained by averaging over 4×10^4 runs. From left to right, the networks have 780, 636, 212 and 226 nodes. For the saturated and critical regimes, we fixed $T = 0.30$ and $T = 0.90$, respectively. In the intermediate regime, T was fixed as 0.75, 0.72, 0.67 and 0.69 for the different networks (left to right). Specific to each network we set β as 0.030, 0.085, 0.110 and 0.090.

degree correlations in app adoption, impeding an isolated analysis of the impact of the mixing rate. Please note that one cannot reach any level of α in every network. The structure of the network and the adoption level impose a lower bound for the levels of α that can be reached [360, 361].

Fig. 4.3 reports the results of the Monte Carlo simulations. The curves, each one corresponding to a different value of T, show that the system switches between the three dynamical regimes, i.e. saturated, intermediate, and critical, as correctly identified by our mean-field model. For the smallest networks — workplace and high school — the intermediate regime exhibits little dependence on α. The small dependence is very likely an effect of stochastic fluctuations. Nevertheless, despite their small size, we can clearly distinguish three distinct regimes.

In a nutshell, the numerical results on real-world networks confirm the phenomenology observed in the minimal mean-field model. As we highlighted previously, assumptions regarding perfect contract tracing are relaxed in Appendix A. Namely, we show that delayed notification and incomplete notification of contacts do not alter the phenomenology. Overall, we explained the emergence of the three dynamical regimes by the requirement for clustered protection for DPT to work. Individual adoption does not reduce the infection

probability. Only notifications along contacts between adopters can prevent further transmission.

Naturally, the question is whether the phenomenology, i.e. the three dynamical regimes, could also arise in the context of other prophylactic tools. Based on our understanding, the prophylactic tool should also require clustered adoption to work effectively for the dynamical regimes to emerge. Accordingly, vaccines with 100% efficacy are an example where we would not expect the phenomenology. If a vaccine is 100% effective, individual adoption is sufficient to be protected independently of whether one's contacts are adopters too. However, for a vaccine that is not 100% effective, individual adoption may not be sufficient. If unvaccinated individuals surround an adopter, a vaccine may only marginally reduce the infection risk if its effectiveness is low. In such cases, clustered adoption could outperform random adoption of vaccines. This intuition leads us to consider the case of imperfect vaccines and other individual prophylactic tools that are not 100% effective, which will be presented in the next section.

4.1.2 *Vaccination and other prophylactic tools*

As we mentioned previously, similar to DPT apps, vaccines are not adopted uniformly in the population. While on a worldwide scale, lack of access impedes an equitable distribution of vaccines [401, 402], among western countries, vaccine hesitancy is the primary barrier [364, 403, 404]. Furthermore, studies have demonstrated that vaccine hesitancy correlates with age, socioeconomic status, education level, or ethnicity [365, 366, 405, 406]. Hence, as for DPT apps, due to homophily [357], the likelihood of physical interaction strongly correlates with the probability that both individuals are either adopters or not.

Evidence suggests that clusters of vaccine-hesitant individuals caused repeated measles outbreaks among high-income countries [364, 403, 407–410]. These recurrent outbreaks sparked many studies to understand the phenomena [411–415]. The high quality of vaccines against measles led these studies to assume an efficacy of almost 100%. In line with our intuition, these studies concluded that homophily, i.e. clustered vaccine adoption, is detrimental to epidemic control and may have sparked these subsequent outbreaks. In contrast, vaccines against influenza or the ones against various SARS-CoV-2 variants have little effectiveness: between 20% and 80% [416, 417]. The low

effectiveness of certain vaccines is the basis of our interest in revisiting previous studies and analysing the impact of homophily for a broader range of vaccine effectiveness.

4.1.2.1 Differential equations

We will take an approach very similar to the one we used to model the impact of homophily on DPT apps. We hope that the reader excuses the slight redundancy necessary for completeness. We consider the standard continuous-time SIR model with transmission probability β, recovery rate μ, and contact rate k. In the absence of vaccinated individuals, the basic reproduction number of the epidemic is thus given by $R_0 = (\beta k)/\mu$. Upon an encounter, an infected individual transmits the virus with a reduced probability $(1-\varepsilon)\beta$ to a vaccinated individual, where ε represents the vaccine efficacy. We assume that vaccines do not reduce the infectivity of infected individuals. Further, we consider the same contact structure as for DPT. The mixing parameter α interpolates between random mixing and the absence of interaction between vaccinated and unvaccinated individuals. Given the vaccine coverage V, the contact matrix has the following structure

$$k_{\mathrm{NV}} = \alpha V k \tag{4.16}$$
$$k_{\mathrm{NN}} = [1 - \alpha V]\, k \tag{4.17}$$
$$k_{\mathrm{VN}} = \alpha(1-V)k \tag{4.18}$$
$$k_{\mathrm{VV}} = [1 - \alpha(1-V)]\, k\,. \tag{4.19}$$

Further, we define $I_{\mathrm{V}}(t)$ and $I_{\mathrm{N}}(t)$ as the number of infectious individuals that are vaccinated and unvaccinated, respectively, at time t. The same holds for the number of susceptible individuals $S_{\mathrm{V}}(t)$ and $S_{\mathrm{N}}(t)$ at time t. With these definitions at hand, the differential equations of the dynamics read

$$\dot{I_{\mathrm{N}}}(t) = \beta\left[k_{\mathrm{NN}} I_{\mathrm{N}}(t) + k_{\mathrm{NV}} I_{\mathrm{V}}(t)\right] S_{\mathrm{N}}(t) - \mu I_{\mathrm{N}}(t) \tag{4.20}$$
$$\dot{I_{\mathrm{V}}}(t) = \beta\left[k_{\mathrm{VN}} I_{\mathrm{N}}(t) + k_{\mathrm{VV}} I_{\mathrm{V}}(t)\right] (1-\varepsilon) S_{\mathrm{V}}(t) - \mu I_{\mathrm{V}}(t) \tag{4.21}$$
$$\dot{S_{\mathrm{N}}}(t) = -\beta\left[k_{\mathrm{NN}} I_{\mathrm{N}}(t) + k_{\mathrm{NV}} I_{\mathrm{V}}(t)\right] S_{\mathrm{N}}(t) \tag{4.22}$$
$$\dot{S_{\mathrm{V}}}(t) = -\beta\left[k_{\mathrm{VN}} I_{\mathrm{N}}(t) + k_{\mathrm{VV}} I_{\mathrm{V}}(t)\right] (1-\varepsilon) S_{\mathrm{V}}(t)\,. \tag{4.23}$$

Through the next-generation matrix approach [211, 214] we calculate the reproduction number of the system and find

$$R = \frac{R_0}{2}\left[2 - \alpha + \alpha\varepsilon(1-V) - \varepsilon + \sqrt{[\alpha - \alpha\varepsilon(1-V) + \varepsilon]^2 - 4\alpha\varepsilon V}\right] . \quad (4.24)$$

It is straightforward to show that R depends monotonously on all figuring parameters. Hence, the dynamics differ qualitatively with respect to DPT, and R depends monotonously on the mixing parameter. The expression of R allows us to define the critical mixing rate α_c above which the disease is eradicated ($R < 1$) as

$$\alpha_c = \left(1 - \frac{1}{R_0}\right)\frac{1 - R_0(1-\varepsilon)}{1 - R_0(1-\varepsilon) - \varepsilon(1-V)} . \quad (4.25)$$

Since the dependence of R on α is monotonous, there is only one critical value and not two, as for DPT. Let us now consider the final attack rate to see whether the dynamics also differ qualitatively from DPT.

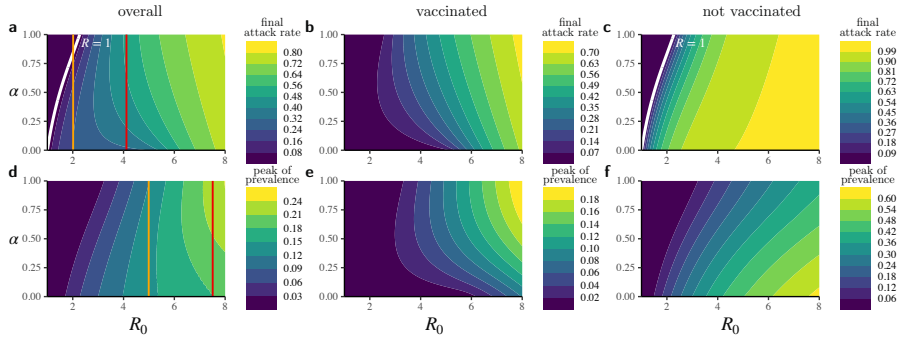


Figure 4.4: Final attack rate **((A)-(C))** and the peak of prevalence **((D)-(F))** as functions of the basic reproduction number, R_0, and the mixing parameter, α, as resulting from the numerical iteration of the differential equations governing the system dynamics (Eqs. (4.20)-(4.23)). Here, the vaccine coverage is $V = 0.7$ and the vaccine efficacy $\varepsilon = 0.8$. In order to highlight the competing processes at the base of the dynamics, we show, besides the results for the population overall **((A)** & **(D))**, those for the vaccinated individuals **((B)** & **(E))** and for the non-vaccinated ones **((C)** & **(F))**, separately. The white solid line indicates the critical curve, $\alpha_c \equiv \alpha_c(R_0)$, at which $R = 1$, as computed from Eq. (4.25). The orange and red vertical lines demarcate the boundaries between, respectively, the critical and the intermediate regime, and the intermediate and the saturated one, for both the final attack rate **(A)** and the peak of prevalence **(D)**.

4.1.2.2 The three dynamical regimes, again

In contrast to the reproduction number, the final attack rate exhibits the same three dynamical regimes as for DPT. Fig. 4.4A illustrates how the basic reproduction number lets the system switch between the three dynamical regimes, i.e. the critical, intermediate, and saturated. Furthermore, similarly as for DPT, the attack rate among nonadopters (Fig. 4.4C) monotonously decreases with α, while the dependence for adopters (Fig. 4.4B) is non-trivial. We observe the three regimes not only for the final attack rate but also for the peak prevalence (Fig. 4.4B). However, the system is not necessarily in the same dynamical regime for both the final attack rate and the peak prevalence. For example, at $R_0 = 5$, the final attack rate is in the intermediate regime while the peak prevalence is in the critical one (Fig. 4.4D).

Fig. 4.5 shows that the three regimes can also be accessed through the vaccine coverage V, or the vaccine efficacy ϵ. More precisely, increasing V or ϵ causes the system to transfer from the saturated to the critical regime passing through the intermediate one. Hence, for sufficiently high vaccine efficacy, the system is always in the critical regime for which mixing between vaccinated individuals and non-vaccinated ones is beneficial. Considering the case of a perfect vaccine, i.e. $\epsilon = 1$, the dynamical equations reduce to the standard SIR model with contact rate $k_{\mathrm{NN}} = (1 - \alpha V)k$ within the non-vaccinated population. $R_0 = (\beta k_{\mathrm{NN}})/\mu$ is thus linearly decreasing with respect to the

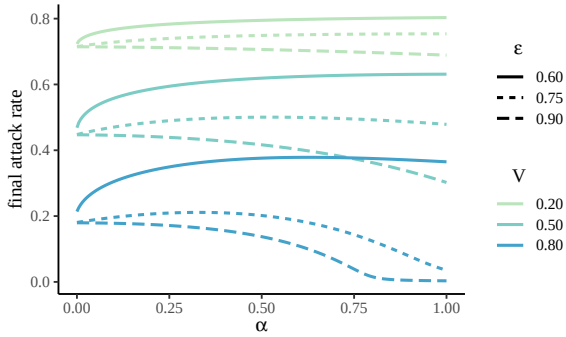


Figure 4.5: Attack rate as a function of the mixing parameter, α, for different combinations of vaccine coverage, V, and efficacy, ε, given $R_0 = 2.5$. As a complement to Fig. 4.4, this plot illustrates how all three dynamical regimes can be also explored by varying V or ε (or both).

mixing rate α, and the system is always in the critical regime where the attack rate monotonously decreases with α.

4.1.2.3 Monte Carlo simulations

As for DPT, we perform Monte Carlo simulations on a real-world physical contact network to corroborate the phenomenology we unveiled in the mean-field case. Specifically, we consider the Copenhagen Networks study where contacts were recorded through Bluetooth signal [147]. The data was collected during an entire month and was binned into 5-minute intervals with about 700 participants. We considered only those interactions where the associated signal strength indication (RSSI) was not lower than -74 dBm corresponding to an interaction range of 2 meters [418]. We iterated the temporal network for the Monte Carlo simulations of the epidemic dynamics until all individuals were either susceptible or recovered.

To infer the mixing parameter α in the network, we need to slightly adapt how we calculate the homophily level h since we consider a temporal network this time. With the temporal adjacency matrix $A(t)$ and the vaccination status vector v, the homophily level is now given by

$$h = \frac{\sum_t \sum_{i=1}^{N} \sum_{j=1}^{N} A_{ij}(t) \left[v_i v_j + (1 - v_i)(1 - v_j) \right]}{\sum_t \sum_{i=1}^{N} \sum_{j=1}^{N} A_{ij}(t)} . \tag{4.26}$$

In other words, we calculate the time-averaged homophily level. As before, by inverting the relation in Eq. (4.14), we can infer the mixing parameter α for a given h. Further, we distribute vaccines with the same algorithm we used to distribute DPT apps. The only difference is that we now control for the average strength instead of the average degree. Therefore, swaps are only allowed if the absolute difference in average strength between vaccinated nodes and not-vaccinated ones is reduced. The results of the Monte Carlo simulations for the final attack rate and peak prevalence are reported in Fig. 4.6. The final attack rate exhibits the three dynamical regimes for three values of ε. However, the particular constraints of this network, which do not allow for arbitrary small values of α, at least with the algorithm we employ to distribute the vaccines, make the saturated regime less evident. Further, the case $\epsilon = 0.6$ shows an example where the attack rate and the peak prevalence are not in the same dynamical regime as already found in the mean-field case.

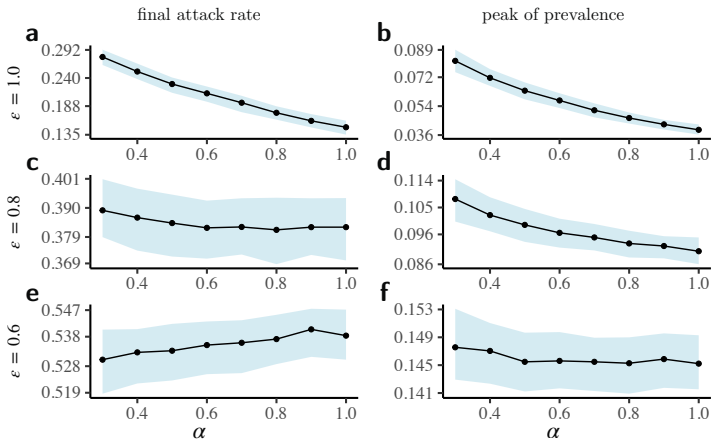


Figure 4.6: Attack rate (left) and the peak prevalence (right) as functions of the mixing parameter, α, for decreasing (from top to bottom) efficacy, ε, as resulting from the numerical simulations performed on top of a real-world temporal contact network (see the main text for details). Dots indicate the median value, whereas the ribbon indicates the first and third quartiles. Each point is obtained by averaging over 2×10^4 runs. We fixed $V = 0.5$, $\mu = 4.6 \times 10^{-4}$ (corresponding to a mean infection time of 7.5 days) and $\beta = 1.142 \times 10^{-1}$, yielding $R_0 = 6$ from the estimation $R_0 = \beta\kappa/\mu$, where $\kappa = \overline{s^2}/\bar{s}$, being $\bar{s}$ ($\overline{s^2}$) the network average of the (squared) number of contacts per timestamp.

4.1.2.4 Practical implications and applicability to other prophylactic tools

Depending on data availability, this framework can be applied to different populations and at various levels of resolution. Many epidemic models only implicitly consider the mixing between vaccinated and non-vaccinated individuals through age-stratification [419, 420]. Our findings indicate that further sub-group-specific mixing levels may impact the epidemic dynamics both quantitatively and qualitatively. Detailed empirical studies on the interaction rates between vaccinated and non-vaccinated individuals would provide a further tool to inform models and interpret epidemiological data. Such more realistic models could improve the guidance of policymakers regarding the impact of specific non-pharmaceutical interventions.

An example with respect to SARS-CoV-2 is the introduction of the *green pass* (proof of vaccination) as a requirement to enter the workplace or restaurants. One intention behind the green pass is nudging non-vaccinated individuals to adopt the vaccine or at least to reduce their number of contacts [421]. At the same time, the green pass reshapes social contacts [422]. Besides

reducing the number of contacts of non-vaccinated individuals, it effectively reduces the mixing between vaccinated and non-vaccinated individuals. Our results indicate that models that do not account for the change in the mixing rate may overestimate the green passes impact in the case of a low reproduction number.

The model presented here focuses on the distribution of vaccines. Nevertheless, the differential equations of the dynamics are the same for any prophylactic tool that causes a reduction of the transmission probability [18]. Accordingly, the three dynamical regimes and the results drawn from them naturally extend to prophylactic practices such as the use of face masks or social distancing. All these examples, from DPT to social distancing, have in common that their individual adoption is insufficient to provide protection in all environments. Collective, clustered adoption is necessary to reduce the infection risk effectively. The presented phenomenology is robust not only across a variety of prophylactic tools but also with respect to model assumptions. Without being aware then, different research groups simultaneously studied the impact of homophily on epidemic control through prophylaxis [423–427]. For example, Fefferman et al. demonstrated that the presence of homophily also adds interesting phenomenology regarding the temporal evolution of the system beyond the final attack rate [425]. With respect to the final attack rate, these studies unveiled the same phenomenology as was presented here for both DPT and vaccine adoption, even though different assumptions are made, as they consider, for example, random networks [423, 427], while we focused on the mean-field case and real-world networks.

Further, Hiraoka et al. assumed an all-or-nothing vaccine [423] — complete protection to a subset ϵ of the vaccinated individuals — in contrast to the leaky vaccine [312, 428] — partial protection ϵ upon contact to every individual — considered here. Despite this difference, they have found a very similar phenomenology. Hence, imperfect protection is at the root of the phenomenology regardless of the type of vaccine. In the next section, we will present another example where perfect vaccines and non-perfect ones differ in their qualitative behaviour. However, this time, whether we consider an all-or-nothing vaccine or a leaky one will change things drastically.

4.2 DEGREE-BASED IMMUNISATION STRATEGIES REVISITED

We highlighted before that humans tend to interact with individuals with similar social characteristics. At the same time, individual success in making social connections and maintaining them is heterogeneously distributed in the population [429]. The heterogeneity is even more pronounced if one considers sexual contacts. The distribution of the number of sexual contacts in the population was shown to be fat-tailed and well captured by a power law [126, 430]. These features imply that while most individuals have only a few sexual contacts, a small fraction participates in most sexual interactions. The resulting heterogeneity in the contact distribution drives the spread of an epidemic beyond the average degree since it enables super-spreading events to occur. As previously mentioned, for example, Gonorrhea can only spread due to the heterogeneity in the sexual contact distribution [18]. If every individual had the average number of sexual contacts in the population, the disease would die out.

Hence, it is not surprising that the idea to focus on individuals with many contacts to control an epidemic emerged in the context of sexually transmitted diseases during the early 80s. More specifically, the most effective control strategies were found to be the ones that aim to promote condom use and regular screening among the most sexually active individuals [59, 79]. Beyond the scope of sexually transmitted diseases, this basic idea to focus on individuals with a high contact frequency was extended to other prophylactic tools such as vaccines. During the 80s, a set of studies focused on the possible eradication of measles through the employment of vaccines [431, 432]. Due to different levels of immunity and activity across age groups, understanding the contact patterns between age groups was crucial in this context.

The problem of optimal vaccine allocation also quickly raised interest in the field of complex networks. Pastor-Satorras and Vespignani provided the first complete analytical analysis on how one can leverage the heterogeneity in the degree distribution to ameliorate the impact of vaccines [433]. Eventually, researchers realised that if the main goal is to reduce infections, one can leverage different techniques developed in statistical physics to tackle the issue of optimal vaccine allocation on a network [434, 435]. These techniques could be employed due to the existing mapping of the SIR model to a percolation problem [188]. Further, the percolation problem, and by transitivity, optimal vaccine allocation, is closely related to the design of resilient infrastructure networks [436–438]. The generality and relevance of this problem

sparked a series of studies that did not only focus on the degree of a node but made use of other network metrics such as betweenness or other centrality measures [439–441]. These approaches leverage heuristics to solve the NP-hard optimisation problem. From an analytical point of view, message-passing and the cavity method have proven useful to immunise networks [434, 435] efficiently. Furthermore, several extensions exist, where researchers considered, for example, temporal networks [442] or immunised links instead of nodes [180, 443].

From a practical point of view, implementing most of these advanced strategies to allocate vaccines is simply not feasible. While one may have access to all the nodes and edges present in infrastructure networks [444], privacy concerns prevent such data collection in the context of human contact networks [445]. Hence, in practical settings, strategies that focus on the more frequently interacting individuals are often still the best one can do.

Most studies that focused on degree-based immunisation strategies considered the vaccine to be perfect, corresponding to an efficacy of 100% [433, 446, 447]. However, as we pointed out earlier, many vaccines and prophylactic tools have an efficacy nowhere near a 100% [416, 417, 448]. Therefore, we asked ourselves whether targeting the more frequently interacting individuals is still the best option to control an epidemic if efficacy is not 100%. Intuitively, a higher risk of exposure to the pathogen, i.e. more contacts, enhances the probability of prevention to fail (breakthrough infection) [312, 428]. Accordingly, there could be a trade-off between the probability of breakthrough infections and the ability to prevent super-spreading events when focusing on the individuals with the most contacts. We will blend an analytical approach with a numerical analysis to investigate whether this trade-off exists.

4.2.1 *Model setup*

To model the spread of a pathogen in a population with different interaction rates, represented by a degree distribution p_k, we consider a standard SIS compartmental model in the heterogeneous mean-field approach, i.e. the annealed network formalism (Sec. 2.2.3.2). The annealed network formalism factors out all dynamical correlations but is analytically tractable [131]. The absence of dynamical correlation is not an issue since we solely want to focus on the impact of the degree heterogeneity. Further, we consider a leaky prevention, with efficacy ϵ in decreasing the transmission probability λ. Let

us define y_k (x_k) as the probability that an individual in degree class k who does (not) receive protection is infected. The variable g_k is the probability that an individual in degree class k receives protection. Finally, given a recovery rate μ, the equations of the heterogeneous mean-field in the presence of prophylaxis read

$$\dot{x}_k = -\mu x_k + \frac{\lambda}{\langle k \rangle} k(1 - x_k)\xi \tag{4.27}$$

$$\dot{y}_k = -\mu y_k + \frac{\lambda}{\langle k \rangle}(1 - \epsilon)k(1 - y_k)\xi \tag{4.28}$$

$$\xi = \sum_k k p_k \left[(1 - g_k)x_k + g_k y_k\right] . \tag{4.29}$$

ξ is an auxiliary variable that simplifies the calculations later on and represents the probability of entering in contact with an infected individual. To simplify further, we rescale time and define the reduced transmission rate $\hat{\lambda} = \lambda/(\mu\langle k \rangle)$, where $\langle k \rangle$ is the average degree.

4.2.2 *Optimal Distribution Strategy*

The overall prevalence in the population is expressed as

$$I = \sum_k p_k \left[(1 - g_k)x_k + g_k y_k\right] . \tag{4.30}$$

Now, given a fixed amount of prevention, optimising the distribution is equivalent to minimising overall prevalence with respect to the functional $I[g]$. From an analytical point of view, this minimisation problem is not tractable. Hence, we take another approach and consider the limit case where only a small amount of prevention is available. More precisely, we measure the impact of distributing prevention to a few individuals in degree class k on the overall prevalence. This approach essentially results in a linear response function $f(x)$ that we define as the change in the number of infected individuals NI if a small amount of prevention $Np_k g_k$ is provided in degree class k. Hence, $f(k)$ can be written as

$$f(k) \sim -\left.\frac{\mathrm{d}(NI)}{\mathrm{d}(Np_k g_k)}\right|_{g=0} . \tag{4.31}$$

We put the minus sign just for convenience to work with a positive valued function. We see that the population size N cancels out, and the response function reads

$$f(k) = -\frac{1}{p_k}\left.\frac{\mathrm{d}I}{\mathrm{d}g_k}\right|_{g=0}. \tag{4.32}$$

If a risk-based strategy is optimal, $f(k)$ would monotonously increase with k. In contrast, the existence of a trade-off would result in a maximum in $f(k)$ for a finite value of k. Let us go ahead and insert the expression of I into Eq. (4.32) leading to

$$f(k) = \overbrace{(x_k - y_k)|_{g=0}}^{F_{dir}(k)} + \overbrace{\frac{1}{p_k}\sum_m p_m \left.\frac{\mathrm{d}x_m}{\mathrm{d}g_k}\right|_{g=0}}^{F_{indir}(k)}. \tag{4.33}$$

As denoted above, $f(k)$ has two terms $F_{dir}(k)$ and $F_{indir}(k)$. $F_{dir}(k)$ quantifies the reduction in the infection probability among those that receive prevention. In contrast, $F_{indir}(k)$ quantifies the indirect effect of the prevention campaign, i.e. the reduction in the infection risk for those that do not receive prevention. Now, to make explicit the expression of $F_{dir}(k)$ and $F_{indir}(k)$ requires us to calculate x_k, y_k, and the derivative $\frac{\mathrm{d}x_m}{\mathrm{d}g_k}$.

4.2.2.1 Individual reduction of the infection risk

First, at the equilibrium ($\dot{x}_k = \dot{y}_k = 0$), one can use Eq. (4.28) to write y_k as a function of x_k:

$$y_k = \frac{1-\epsilon}{1-\epsilon x_k} x_k. \tag{4.34}$$

Further, given that x_k has to be evaluated at $g = 0$ (no prevention), from Eq. (4.27) (see also sec. 2.2.2) we can express x_k recursively as

$$x_k = \frac{z\hat{\lambda}k}{1 + z\hat{\lambda}k}, \tag{4.35}$$

where we defined z as

$$z = \langle kx \rangle = \sum_k p_k k x_k \,. \tag{4.36}$$

The term z represents an individual's expected number of at-risk contacts. More generally, angle brackets will refer to averages with respect to the de-

gree distribution. The two relations in Eq. (4.34) & (4.35) allows us already to express $F_{dir}(k)$ as

$$F_{dir}(k) = \frac{\epsilon\hat{\lambda}zk}{\left[1+\hat{\lambda}zk\right]\left[1+(1-\epsilon)\hat{\lambda}zk\right]}\,. \tag{4.37}$$

Considering the shape of $F_{dir}(k)$, we note the possibility of a maximum at $k = k^*_{dir} \sim 1/\sqrt{1-\epsilon}$ (Fig. 4.7A). In other words, for individuals that receive protection, their infection risk is maximally reduced if they are in degree class k^*_{dir}. Hence, we have found the postulated trade-off regarding the direct impact on the individuals given prevention. Further, given the scaling of k^*_{dir}, the maximum only exists for $\epsilon < 1$. If prevention is perfect, one should always prioritise immunising highly connected nodes to maximise the direct reduction in the individual infection risk since breakthrough infections are impossible.

4.2.2.2 Impact on the population level

To see whether the same trade-off holds at the population level, i.e. including also individuals that are not given protection, we need to express the derivative $\frac{\mathrm{d}x_m}{\mathrm{d}g_k}$ first. Let us thus consider the equilibrium condition $\dot{x}_k = 0$ and take the derivative with respect to g_m on both sides, and we find

$$-\frac{dx_k}{dg_m} + \hat{\lambda}k\left[-\frac{dx_k}{dg_m}z + (1-x_k)\frac{d\xi}{dg_m}\right] = 0\,. \tag{4.38}$$

We implicitly assume that all the terms should be evaluated at $g = 0$. We then can compute the derivative of ξ given its definition in Eq. (4.29):

$$\frac{d\xi}{dg_m} = mp_m(y_m - x_m) + \sum_{k'} k' p_{k'} \frac{dx_{k'}}{dg_m}\,. \tag{4.39}$$

Inserting this into Eq. (4.38), we find

$$\sum_{k'}\left[\hat{\lambda}k(1-x_k)k'p_{k'} - \delta_{kk'}(1+z\hat{\lambda}k)\right]\frac{dx_{k'}}{dg_m} = -\hat{\lambda}k(1-x_k)mp_m(y_m - x_m)\,. \tag{4.40}$$

This equation constitutes a linear system for the matrix $J_{km} = \frac{\mathrm{d}x_k}{\mathrm{d}g_m}$. By defining the auxiliary variables $u_k = \hat{\lambda}k(1-x_k)$, $v_k = kp_k$, $w_k = kp_k(y_k - x_k)$, and $D_{kk'} = (1 + z\hat{\lambda}k)\delta_{kk'}$, we can rewrite Eq. (4.40) as

$$(uv^T - D)J = -uw^T \,. \tag{4.41}$$

We explicit J by making use of the Sherman-Morrison formula [449] as

$$J = \frac{1}{1 - v^T D^{-1} u} D^{-1} u w^T \,. \tag{4.42}$$

Inserting the expression for u, v, w and D gives us the explicit expression of J and hence $F_{indir}(k)$ as

$$F_{indir}(k) = \frac{\psi\hat{\lambda}}{1-\phi} k F_{dir}(k) \,, \tag{4.43}$$

where we defined

$$\phi = \hat{\lambda}\left\langle \left(\frac{k}{1+z\hat{\lambda}k}\right)^2 \right\rangle \quad \text{and} \quad \psi = \left\langle \frac{k}{(1+z\hat{\lambda}k)^2} \right\rangle \,. \tag{4.44}$$

With respect to $F_{dir}(k)$, $F_{indir}(k)$ has an additional factor k. The additional factor is intuitive since the individual reduction in the infection risk linearly reduces the risk of onward transmission, but with a proportionality factor k, the number of contacts of the protected individual. Due to this scaling factor, $F_{indir}(k)$ does not have a maximum and monotonously increases with the degree k (Fig. 4.7A). Accordingly, with respect to the impact on the rest of the population, targeting hubs is always the best strategy, as it always provides the highest indirect benefit.

4.2.2.3 Analytical expression of the linear response function

If we observe a trade-off when considering both effects depends on whether $F_{dir}(k)$ primes over $F_{indir}(k)$. Actually, making the expression of $f(k)$ in Eq. (4.33) explicit, we find

$$f(k) = \frac{\epsilon\hat{\lambda}zk}{(1+\hat{\lambda}zk)\left[1+(1-\epsilon)\hat{\lambda}zk\right]} \left(1 + \frac{\hat{\lambda}\psi}{1-\phi}k\right) \,. \tag{4.45}$$

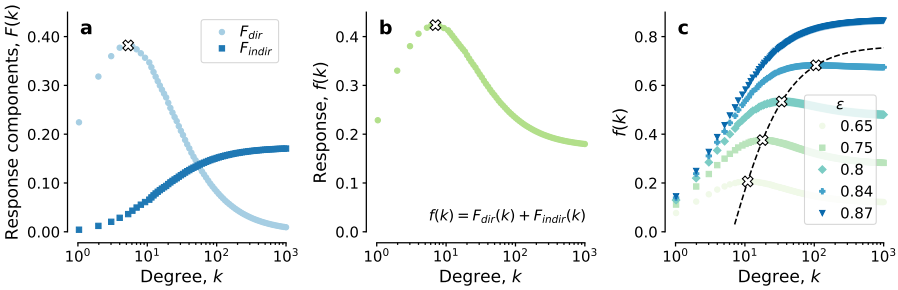


Figure 4.7: **(A)**: Terms $F_{dir}(k)$ (Eq. (4.37)) and $F_{indir}(k)$ (Eq. (4.43)) of the response function. The cross indicates k^*_{dir}, maximum of F_{dir}. Reduced transmissibility is $\hat{\lambda} = 0.25$; degree distribution is a negative binomial with mean 2.0, coefficient of variation (standard deviation over mean) 4.7; efficacy is $\epsilon = 0.5$. **(B)**: Response function $f(k)$, which is the sum of the terms in (A). The cross indicates k^*, maximum of f. **(C)**: Response function $f(k)$ for different values of ϵ. The dashed lines and the crosses indicate k^*. Reduced transmissibility is $\hat{\lambda} = 2$; degree distribution is a negative binomial with mean 2.0 and coefficient of variation 4.7.

To analyse whether $f(k)$ has a maximum, let us consider the derivative with respect to k:

$$f'(k) \sim k^2\hat{\lambda}^2 z\left[\psi(1-\phi)(2-\epsilon) + z(1-\epsilon)\right] - 2\hat{\lambda}\psi(1-\phi)k + 1\,, \tag{4.46}$$

where we dropped a strictly positive term on the RHS. At $k = 0$, we have $f'(k) > 0$. Hence, $\lim_{k\to\infty} f(k) < 0$ is a sufficient condition for $f(k)$ to have a maximum in $\mathbb{R}^+$. This condition translates to

$$z(1-\phi) - 2\psi > \epsilon\left[z(1-\phi) - \psi\right]\,, \tag{4.47}$$

leading to a critical efficacy ϵ_c below which the maximum exists:

$$\epsilon < \epsilon_c = \frac{z(1-\phi) - 2\psi}{z(1-\phi) - \psi}\,. \tag{4.48}$$

We find the maximum by setting $f'(k) = 0$ as

$$k^* = \frac{1 + \sqrt{\left(\frac{z(1-\phi)}{\psi} - 1\right)\left(\frac{z(1-\phi)}{\psi}(1-\epsilon) - 1\right)}}{\hat{\lambda} z\left[\frac{z(1-\phi)}{\psi}(1-\epsilon) - (2-\epsilon)\right]}\,. \tag{4.49}$$

Accordingly, the trade-off exists also when taking into account both $F_{dir}(k)$ and $F_{indir}(k)$. Fig. 4.7B illustrates how $f(k)$ reaches its maxima at k^*. However, the maximum k^* only exists whenever $\epsilon < \epsilon_c$. Fig. 4.7C illustrates that as efficacy increases, the maxima disappears and $f(k)$ monotonously increases with k. In this sense, ϵ_c distinguishes a *high-efficacy* ($\epsilon > \epsilon_c$) region from a *low-efficacy* region ($\epsilon < \epsilon_c$). While in the high-efficacy region, one should adopt a risk-based strategy (i.e. immunising the hubs), targeting individuals in degree class k^* has the strongest impact on community prevalence in the low-efficacy region. Analysing the dependence of k^* and ϵ_c on the different parameters requires, at least partially, a numerical approach. The expressions z, ϕ and ψ, cannot be calculated analytically. Instead, we numerically solve the differential equations without prevention ($g = 0$), enabling us to evaluate the expressions [253, 450, 451].

As a first step, we focus on the impact of prevalence on k^* and ϵ_c. To this aim, we numerically optimise the transmission probability λ until the desired prevalence is reached. We observe in Fig. 4.8A that higher prevalence leads to higher ϵ_c and, in the low-efficacy region, to lower k^*. Further, lower efficacy leads to a decrease in k^*. These results are intuitive since higher prevalence or lower efficacy make breakthrough infections more likely. Besides prevalence, the heterogeneity in the degree distribution will also affect the dynamics. Fig. 4.8B indicates how ϵ_c increases with the network heterogeneity. Similarly, increasing heterogeneity in the network elevates connectivity between hubs and favours breakthrough infections. Accordingly, as the heterogeneity increases, ϵ_c increases too.

4.2.2.4 The impact of prevalence and degree heterogeneity

To this point, the findings on how prevalence or heterogeneity affects the dynamics are mainly numerical. However, we can also show this analytically if we assume a degree distribution that follows a power law. Let us first consider how z is affected by the heterogeneity of the degree distribution. As previously mentioned, z is expressed as $z = \sum_k p_k k x_k$. Accordingly, since the degree and the infection probability are positively correlated (Eq. (4.36)), a lower bound for z is $I\langle k\rangle$, which would correspond to a homogeneous population. Instead, broad degree distributions lead to $z > I\langle k\rangle$. More explicitly, from Eqs. (4.35) & (4.36) it follows that

$$\hat{\lambda} \sum_k p_k \frac{k^2}{1 + z\hat{\lambda}k} = 1. \tag{4.50}$$

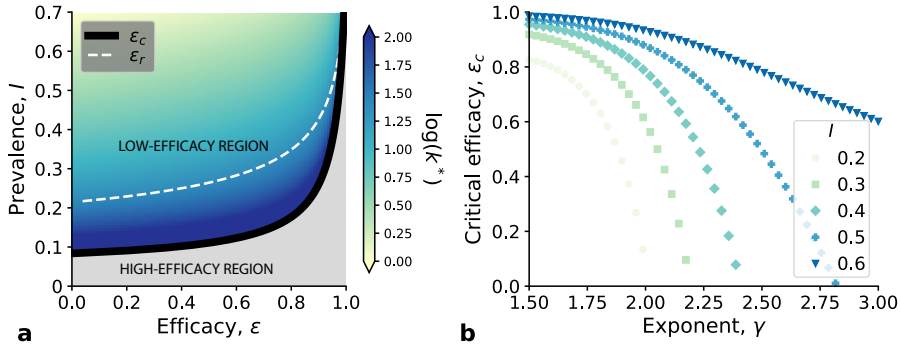


Figure 4.8: **(A)**: k^* as a function of efficacy ϵ, and baseline disease prevalence (prevalence without prevention). Finite values of k^* mark the low-efficacy region. The grey area indicates $k^* \to \infty$: the high-efficacy region. The solid black lines show critical efficacy ϵ_c; the dashed black line shows ϵ_r. The degree distribution is a power law with the exponent 2.0. The reduced transmissibility $\hat{\lambda}$ is numerically set to match the corresponding baseline prevalence. **(B)**: Critical efficacy ϵ_c as a function of the heterogeneity of the contact network, measured as the exponents of the power-law degree distribution. Different curves mark different values of baseline prevalence.

To simplify things, we will assume the degree distribution follows a power-law with exponent γ, i.e. $p_k \sim k^{-\gamma}$. If we approximate the sum on k with an integral, Eq. (4.50) becomes

$$ {}_2F_1(1, \gamma-2, \gamma-1, -\frac{1}{z\hat{\lambda}}) = z\frac{\gamma-2}{\gamma-1}, \tag{4.51} $$

where ${}_2F_1$ is the hypergeometric function. At $\gamma = 2$, ${}_2F_1$ has a simple pole, i.e. ${}_2F_1(1, \gamma-2, \gamma-1, -\frac{1}{z\hat{\lambda}}) \approx \frac{1}{z(\gamma-2)}$. Accordingly, close to $\gamma = 2$, Eq. (4.51) becomes

$$ \frac{1}{[z(\gamma-2)]^2} \approx \frac{1}{\gamma-1}\,. \tag{4.52} $$

Since the RHS is finite at $\gamma = 2$, z needs to scale as $z \sim 1/(\gamma-2)$ to cancel the divergence at $\gamma = 2$, which proves that z increases as the network becomes more heterogeneous (decreasing γ). Further, z monotonously increases with the prevalence in the population (higher λ). Having dealt with the depend-

ence of z on prevalence and the heterogeneity γ, we can now approximate ϕ and ψ for large z:

$$\phi \approx \frac{1}{z^2\hat{\lambda}} \tag{4.53}$$

$$\psi \approx \frac{k^{-1}}{z^2\hat{\lambda}^2}\,. \tag{4.54}$$

Hence, in the limit for large z, k^* and ϵ_c become

$$k^* \approx \frac{1}{\hat{\lambda}z}\frac{1+\sqrt{1-\epsilon}}{1-\epsilon} \to 0 \tag{4.55}$$

$$\epsilon_c \approx 1 - \frac{k^{-1}}{z^3\hat{\lambda}^2} \to 1\,. \tag{4.56}$$

We thus observe that ϵ_c and k^* increase and decrease, respectively, with z. Therefore, based on our analysis of z, we conclude that higher prevalence and network heterogeneity increase ϵ_c and cause k^* to decrease in the low-efficacy region.

4.2.2.5 Immunisation close to the epidemic threshold

Since high prevalence is necessary for the trade-off to occur, at the epidemic threshold, targeting hubs is always the best strategy. We can see this by linearising Eqs. (4.27) & (4.28) around the disease-free equilibrium:

$$\dot{x}_k = -\mu x_k + \frac{\lambda}{\langle k\rangle}k\xi \tag{4.57}$$

$$\dot{y}_k = -\mu y_k + \frac{\lambda}{\langle k\rangle}(1-\epsilon)k\xi. \tag{4.58}$$

We can find an equation for the evolution of ξ by multiplying the first line by $kp_k(1-g_k)$, the second line by kp_kg_k, add both equations together and sum over the degrees k:

$$\dot{\xi} = \left[-\mu + \frac{\lambda}{\langle k\rangle}\left(\langle k^2\rangle - \langle gk^2\rangle\right)\right]\xi. \tag{4.59}$$

Hence, we notice that the disease-free equilibrium is not stable, i.e. $\dot{\xi} > 0$, if

$$-\mu + \frac{\lambda}{\langle k\rangle}\left(\langle k^2\rangle - \langle gk^2\rangle\right) > 0\,, \tag{4.60}$$

leading to the epidemic threshold

$$\lambda_c = \lambda_0 \left(1 - \frac{\epsilon \langle g k^2 \rangle}{\langle k^2 \rangle}\right)^{-1}, \tag{4.61}$$

where λ_0 refers to the standard epidemic threshold in the absence of protection, i.e. $\lambda_0 = \mu \langle k \rangle / \langle k^2 \rangle$. In analogy to our approach for the prevalence, we define the response function $f_\lambda(k)$ as

$$f_\lambda(k) = \frac{1}{p_k} \frac{\mathrm{d}\lambda_c}{\mathrm{d}g_k}\bigg|_{g=0}. \tag{4.62}$$

We did not introduce a minus sign here since prevention increases the epidemic threshold. After some algebra, one finds

$$f_\lambda(k) = \frac{\epsilon \lambda_0}{\langle k \rangle} k^2. \tag{4.63}$$

In contrast to $f(k)$, $f_\lambda(k)$ is monotonously increasing with k, proving that, independent of the efficacy ϵ, targeting hubs is always the best strategy at the epidemic threshold.

4.2.2.6 Random distribution can outperform a risk-based strategy

The results on the epidemic threshold further highlight that prevalence crucially separates the high-efficacy from the low-efficacy region. In this sense, if one wanted to leverage our results in a distribution campaign, accurate knowledge of the prevalence is necessary. Further, it could be challenging to precisely determine the number of contacts to distribute prevention to individuals in degree class k^*. For example, in the context of sexually transmitted diseases, the stigma involved often hinders information collection [452]. For this reason, we propose an easier strategy to implement, which is the *uniform distribution of prevention*. Uniform distribution randomly gives prevention to individuals and hence is mathematically described by $\langle f \rangle$, which we evaluate numerically. We then compare the uniform distribution with a risk-based strategy that is quantified as $f(\infty) = \lim_{k\to\infty} f(k)$ and can be calculated as

$$f(\infty) = \frac{\psi}{z(1-\phi)} \frac{\epsilon}{1-\epsilon}. \tag{4.64}$$

In analogy to ϵ_c, we define ϵ_r as the efficacy value for which $\langle f \rangle > f(\infty)$. The white dashed line in Fig. 4.8A indicates ϵ_r. We see that $\epsilon_r > \epsilon_c$. Nevertheless,

there is a parameter region where uniform distribution outperforms a risk-based strategy. The existence of this parameter region opens the pathway for a distribution strategy that is easy to implement and equitable while still performing well in a high-prevalence setting.

4.2.3 *PrEP distribution as a possible application*

Naturally, the question that arises is which virus has such a high prevalence for our theory to apply. Further, a prevention tool should exist with relatively low efficacy, and we have shown that a heterogeneous degree distribution promotes the trade-off to occur. Taking into account all these properties, the distribution of *pre-exposure prophylaxis* (PrEP) to prevent HIV infections in communities of men who have sex with men (MSM) is a possible application. In particular, in African countries, MSM communities exhibit elevated HIV prevalence, with estimates ranging between 5% and 40% [453]. Further, sexual contact networks in MSM communities have shown to be very heterogeneous. Different studies concluded that a power law with an exponent between 1.5 and 2.0 is a good approximation of the actual degree distribution [430, 454, 455]. Due to the elevated prevalence, huge efforts were made in the recent past to control the spread of HIV among MSM through PrEP distribution. If taken regularly, PrEP has almost an efficacy of 100% [456]. However, in many countries, stigma still surrounds MSM and HIV [452]. The stigma can hinder both regular access to PrEP as well as access to accurate information on how to use PrEP. In many African countries, both of these issues have been shown to cause a lack of adherence [457]. For PrEP to be effective, the medicamentation must be taken daily. Otherwise, the efficacy of PrEP is substantially lower. Studies have shown that in different African countries, due to the lack of adherence, the efficacy of PrEP is only between 40% and 80% [456]. In this sense, HIV prevention in African MSM communities through PrEP distribution fulfils all three criteria for our theory to apply: high prevalence, a heterogeneous degree distribution and low efficacy. Random distribution is particularly interesting in these settings since the stigma involved often makes it difficult to closely engage with the affected individuals [452]. In such circumstances, implementing a risk-based strategy can be very challenging since one cannot identify the most at-risk individuals. Therefore, the uniform distribution we showed here could be a good alternative that is easier to implement while performing equally well or even better.

However, a detailed analysis should forego any implementation of a uniform distribution in a specific region. Firstly, a robust estimation of the prevalence is necessary. Second, one must infer or approximate the degree distribution in the MSM population. Further, besides looking at the degree distribution, one should also assess whether interactions are recurrent and with which frequency new partners are acquired [83, 458]. With such information at hand, agent-based simulations would then probably provide the most accurate estimation of the impact of the different distribution strategies [459]. Additionally, here we focused only on the optimal distribution if only a small amount of PrEP is available. This approach allowed us to consider the linear response function and made the problem analytically tractable. However, if it is possible to distribute PrEP to a substantial part of the population, optimisation of the functional as presented in Eq. (4.30) would be necessary. In this sense, the approach presented here is a first step towards developing optimal distribution strategies in high-prevalence settings. Hopefully, our findings can spark new studies that focus on specific environments in order to implement these distribution strategies in the future.

4.2.4 *Emerging Phenomenology due to imperfect vaccines*

If we set aside the real-world applications and focus on the theoretical implications of our findings, we found, after the case of homophilic vaccine adoption, another example where an imperfect vaccine behaves qualitatively differently from a perfect one. Similarly to the three dynamical regimes, the trade-off between breakthrough infections and the population level impact can only take place if efficacy is not 100%. However, as we pointed out, for the homophilic adoption, the three dynamical regimes arise independently on whether one considers leaky or all-or-nothing protection. In contrast, the trade-off we unveiled here can only arise for leaky protection. For an all-or-nothing protection, the relative individual effectiveness is not affected by the number of contacts. The all-or-nothing case is equivalent to perfect protection with lower adoption. Therefore, we discovered a case where leaky and all-or-nothing protection differ not only quantitatively but also phenomenologically [312, 419, 428]. Hence, one should also be aware of the biological details of a prevention mechanism to assess whether it is leaky or all-or-nothing.

4.3 SUMMARY AND DISCUSSION

In this chapter, we have considered two facets of human behaviour and analysed their impact on epidemic control. First, we have shown how the preference of humans to interact with others that share similar characteristics — homophily — non-trivially affects the impact of prophylactic tools. Contrary to previous findings, our results indicated that if efficacy is not 100%, homophily is not always beneficial but can also be detrimental or non-monotonously affect epidemic control. The phenomenology holds for a wide range of prophylactic tools such as vaccines, the use of masks, social distancing, or DPT. Second, we considered optimal strategies for prophylaxis distribution for populations whose contacts follow heterogeneous degree distributions. We have shown that other than what standard theory tells us, a risk-based strategy (targeting hubs) is not always the best option. Instead, if efficacy is low with respect to the prevalence, an optimal finite degree exists to target. Further, if prevalence is even higher or efficacy is even lower, uniform distribution may outperform a risk-based approach. Both findings share that the phenomenology only emerges if the prevention tool's efficacy is not 100%. Furthermore, the results highlighted that, when considering the correlation between metadata and network structure, one can still explore new phenomenology [460–464].

From a more technical point of view, this and the previous chapter share a similar approach, but differ in one aspect. Here, we considered an empirically grounded aspect of human behaviour and analysed through minimal mathematical modelling what this may imply for epidemic control. In contrast, in Chapter 3, we started from the phenomenology and tried to understand through modelling how human behaviour may be at its cause. However, both approaches share that they focus mainly on the qualitative behaviour, the phenomenology. Next, we will take a more quantitative approach and do actual real-world modelling in the context of SARS-CoV-2. We will show how modelling can be used to make predictions, but more importantly, how we can infer aspects of human behaviour or characteristics of the virus.

UNIVERSITAT ROVIRA I VIRGILI
MODELLING THE INTERPLAY BETWEEN HUMAN BEHAVIOUR AND THE SPREAD OF INFECTIOUS DISEASES: FROM TOY MODELS TO QUANTITATIVE APPROACHES
Benjamin Franz Josef Steinegger

5

THE SPREAD OF SARS-COV-2

With the worldwide spread of SARS-CoV-2 in early 2020, epidemic models were suddenly discussed in the public sphere. As health authorities consulted epidemiologists to provide them with projections of the epidemics' evolution, model predictions became part of the daily argy-bargy in politics [354]. Unfortunately, the public discussion mainly focused on the predictions of epidemic models. While some ridiculed modellers as a kind of modern-day Nostradamus, others almost thirsted for pessimistic predictions. In this heated debate, it was often forgotten that epidemiological models serve for much more than to predict the epidemics' trajectory. One could even argue that epidemiological models are much more helpful for everything other than making predictions. For example, epidemiological models allowed us to infer many of the characteristics of the virus, such as its generation time [232], infectious period [62], or reproduction number [57, 465]. Further, with the emergence of new variants, models were used to infer the increase in transmissibility and virulence of the mutated strains [466, 467]. Also, epidemiological models were used to retrospectively assess what actually occurred [468, 469]. Such assessments allowed to estimate the impact of past NPIs [314, 316], eventually enabling the improvement of the containment strategy in the future. Especially in early 2020, the reporting of cases, and hence data availability, was so poor that the evolution of the first wave could only be accurately assessed in hindsight [470, 471]. The rapid increase in available information required a constant adaptation of the modelling efforts and the conclusions drawn from them. Updating previously drawn conclusions, and maybe even contradicting them, appeared to be very challenging to communicate to the general public [472]. Probably, this was one of the few occasions where the inner workings of scientific production, in particular, its historicity [473], were at full display in real-time for the general public.

All the above also holds for the modelling efforts regarding SARS-CoV-2 we did in our research group. Starting in February 2020, our models and assumptions were adapted multiple times. Hence, it should not be surprising to the reader if this chapter presents the contributions with a critical eye. Furthermore, temporal constraints made it necessary to find a trade-off between pragmatism and scientific perfection. In this chapter, we will first present the model we used to understand the spatial spread of SARS-CoV-2 in Spain during the spring of 2020. The second study leverages case data that became available in the summer of 2020, which allowed us to assess the evolution of the first wave retrospectively. The retrospective analysis will shed light on our difficulties in modelling the epidemic in real-time. We will conclude the chapter by building a model for the spread of the Delta variant during the summer of 2021 in Catalonia. The spread of the Delta variant is of particular interest since, besides the emergence of the new variant, vaccines were newly available and restrictions had just been lifted. Eventually, this chapter should illustrate how epidemic models can be used to predict the epidemic evolution, infer characteristics of the virus, and highlight how human behaviour shaped the disease propagation.

5.1 A SPATIOTEMPORAL MODEL FOR THE PROPAGATION OF SARS-COV-2

On December 31 2019, China officially communicated to the WHO about the possible emergence of a new Coronavirus. Soon after, on January 23 2020, China implemented a travel ban that affected millions of people [474]. In hindsight, we know that, at this point, the virus had already spread to a substantial part of the globe [195]. However, in early 2020, European countries detected their first cases between the end of January and the beginning of February. Furthermore, the constant reports of newly detected cases in Spain from February onward substantiated the possibility of a large-scale propagation of SARS-CoV-2. These reports prompted us to start our modelling efforts, which were, at least initially, mainly aimed at understanding the spatiotemporal patterns of the epidemic that one would likely observe. Hence, human mobility was at the core as we built the epidemiological model that we will present in this section.

5.1.1 *Basic model equations*

To model the spread of SARS-CoV-2 we use the MMCA [175, 460], which we presented in Sec. 2.2.3.4. The MMCA has the particular advantage that it is in discrete-time, reducing the computational cost. We extended the standard approaches and introduced additional compartments that are tailored for SARS-CoV-2. To incorporate mobility, we use the approach introduced in Ref. [208], which was developed for recurrent mobility, i.e. commuting. In this framework (Sec. 2.2.4), a set of variables $\{\rho_i^{m,g}(t)\}$ represent the probability that individuals, with residence in patch i, belonging to age strata g, are in state m. The different states and the transitions among them are illustrated in Fig. 5.1. We consider the following compartments: susceptible (S), exposed (E), pre-symptomatic infectious (A), symptomatic infectious (I), to be admitted in ICU (pre-hospitalised in ICU, P_H), fatal prognosis (predeceased, P_D), admitted in ICU that will recover (H_R) or decease (H_D), recovered (R), and deceased (D). Compartments P_H and P_D are introduced to account for the delays until individuals are admitted to the ICU or until death without admission to an ICU, respectively. The compartments H_R and H_D serve to distinguish individuals who will recover and decease, respectively, after admission to an ICU.

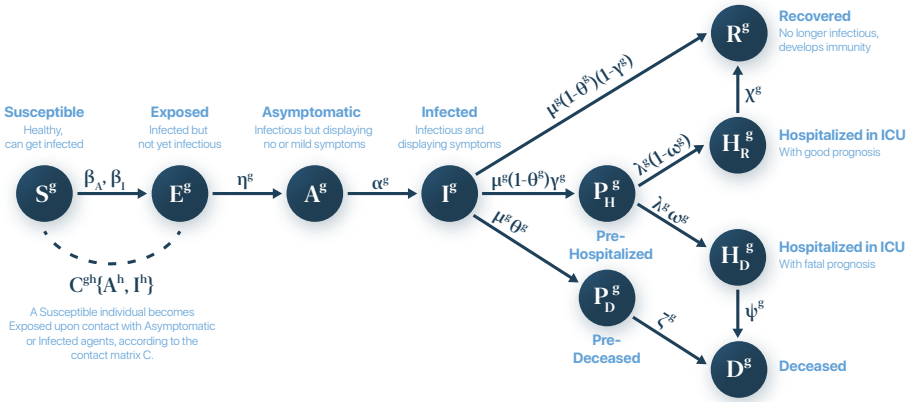


Figure 5.1: The acronyms correspond to: susceptible (S^g), exposed (E^g), asymptomatic infectious (A^g), symptomatic infectious (I^g), pre-hospitalised in ICU (P_H^g), pre-deceased (P_D^g), in ICU before recovery (H_R^g), in ICU before death (H_D^g), deceased (D^g), and recovered (R^g), where g denotes the age stratum of all compartments. The arrows indicate the transition probabilities.

The rationale for the transition between the compartments is the following. Susceptible individuals get infected through contact with asymptomatic or infected agents with a probability $\Pi_i^g(t)$ (details in Sec. 5.1.2) entering the exposed compartment. Please note that asymptomatic refers here to pre-symptomatic infectiousness. Exposed individuals become infectious but are still asymptomatic at a rate η^g. Individuals then develop symptoms at a rate α^g. After the infectious compartments, three paths emerge for the clinical part at a rate μ^g. Individuals either recover, are admitted to an ICU or decease without admission. Individuals die without ICU admission with a probability θ^g and rate ζ^g. Otherwise, individuals require ICU admission with probability γ^g and rate λ^g. Once admitted to an ICU, individuals decease with probability ω^g and rate ψ^g. If individuals do not decease, they exit the ICU with a rate χ^g. As already pointed out, the compartment P_D, and similarly P_H, serves only as an intermediate to guarantee the desired times between infection and the clinical events. Appendix B details the choices for these parameters.

As mentioned above, we separate the population into N_G age strata accounting for the age-dependent risk for a severe course of infection [326–328]. More precisely, we use the following three age states ($N_G = 3$): young Y, with age up to 25; adults M, with age between 26 and 65; and older people O, with age larger than 65. Accordingly, we split the population of N individuals into N_P patches, with n_i individuals residing in patch i. Further, n_i^g individuals of age stratum g live in patch i. Hence,

$$N = \sum_{g=1}^{N_G} \sum_{i=1}^{N_P} n_i^g = \sum_{i=1}^{N_P} n_i = \sum_{g=1}^{N_G} n^g \,, \tag{5.1}$$

where n^g is the total population of age stratum g. As previously mentioned, $\{\rho_i^m(t)\}$ accounts for the probability that individuals of age stratum g assigned to patch i are in state m at time t, where $m \in \{S, E, A, I, P_H, P_D, H_R, H_D, R, D\}$ and $g \in \{Y, M, O\}$. With these definitions, the temporal evolution of the system is given by

$$\rho_i^{S,g}(t+1) = \rho_i^{S,g}(t)\,(1 - \Pi_i^g(t)) \tag{5.2}$$

$$\rho_i^{E,g}(t+1) = \rho_i^{S,g}(t)\,\Pi_i^g(t) + (1-\eta^g)\,\rho_i^{E,g}(t) \tag{5.3}$$

$$\rho_i^{A,g}(t+1) = \eta^g\,\rho_i^{E,g}(t) + (1-\alpha^g)\,\rho_i^{A,g}(t) \tag{5.4}$$

$$\rho_i^{I,g}(t+1) = \alpha^g\,\rho_i^{A,g}(t) + (1-\mu^g)\,\rho_i^{I,g}(t) \tag{5.5}$$

$$\rho_i^{P_D,g}(t+1) = \mu^g\,\theta^g\,\rho_i^{I,g}(t) + (1-\zeta^g)\,\rho_i^{P_D,g}(t) \tag{5.6}$$

$$\rho_i^{P_H,g}(t+1) = \mu^g\,(1-\theta^g)\,\gamma^g\,\rho_i^{I,g}(t) + (1-\lambda^g)\,\rho_i^{P_H,g}(t) \tag{5.7}$$

$$\rho_i^{R,g}(t+1) = \mu^g\,(1-\theta^g)\,(1-\gamma^g)\,\rho_i^{I,g}(t) + \chi^g\,\rho_i^{H_R,g}(t) + \rho_i^{R,g}(t) \tag{5.8}$$

$$\rho_i^{H_D,g}(t+1) = \lambda^g\,\omega^g\,\rho_i^{P_H,g}(t) + (1-\psi^g)\,\rho_i^{H_D,g}(t) \tag{5.9}$$

$$\rho_i^{H_R,g}(t+1) = \lambda^g\,(1-\omega^g)\,\rho_i^{P_H,g}(t) + (1-\chi^g)\,\rho_i^{H_R,g}(t) \tag{5.10}$$

$$\rho_i^{D,g}(t+1) = \zeta^g\,\rho_i^{P_D,g}(t) + \psi^g\,\rho_i^{H_D,g}(t) + \rho_i^{D,g}(t). \tag{5.11}$$

Above, $\Pi_i^g(t)$ encodes the probability that a susceptible agent belonging to age stratum g and patch i contracts the disease during time step t. Time steps correspond to a day.

5.1.2 *Modelling the mobility between patches*

Due to the age-stratification, the metapopulation approach here requires a slight adaptation from Ref. [208]. The patches correspond to the municipalities in Spain. The mobility data refers to the daily commuting patterns between municipalities. The mobility pattern is encoded in a matrix R, whose elements R_{ij}^g refer to the fraction of agents in patch i and age stratum g that commute daily to patch j. Given the mobility matrix, the probability $\Pi_i^g(t)$ that a susceptible agent belonging to age stratum g and patch i gets infected, can be written as

$$\Pi_i^g(t) = (1-p^g)\,P_i^g(t) + p^g \sum_{j=1}^{N_P} R_{ij}^g\,P_j^g(t)\,, \tag{5.12}$$

where p^g denotes the mobility level of individuals in age stratum g. If individuals follow the usual commuting patterns, we set $p^g = 1$. In contrast, $p^g = 0$ implies that all individuals stay in their patch of origin. Later we will show how we used the mobility parameter to model the impact of confinement measures. P_i^g encodes the probability that an individual contracts the disease when in patch i. To express this probability, we need to make some further considerations. We will consider an age-specific contact matrix C that controls the mixing between age strata. Additionally, we assume that the number of contacts increases monotonously according to a function $f(x)$

with the density of individuals in the patch. With these definitions, P_i^g is expressed as

$$P_i^g(t) = 1 - \prod_{h=1}^{N_G} \prod_{j=1}^{N_P} \prod_{m\in\{A,I\}} (1-\beta_m)^{z^g \langle k^g \rangle f\left(\frac{\tilde{n}_i}{s_i}\right) C^{gh} \frac{n_{j\to i}^{m,h}(t)}{\tilde{n}_i^h}} . \tag{5.13}$$

The parameters β_A and β_I correspond to the transmission probability upon contact with a pre-symptomatic and symptomatic individual, respectively. The exponent expresses the number of contacts made by an agent of age stratum g in patch i with infectious individuals — compartments A and I — of age stratum h belonging to patch j. The three products then iterate through all patches, age strata, and symptomatic and asymptomatic individuals.

The term $z^g \langle k^g \rangle f(\tilde{n}_i/s_i)$ represents the total number of contacts that an agent from age stratum g makes inside patch i. The function $f(x)$ controls the scaling for the density, as s_i is the area of patch i and $\tilde{n}_i$ is its effective population. z^g serves as a normalisation factor, such that the average number of contacts of age stratum g in the entire country is $\langle k^g \rangle$, which is fixed according to empirical studies. For this reason, z^g is given by

$$z^g = \frac{n^g}{\sum_{i=1}^{N_P} f\left(\frac{\tilde{n}_i}{s_i}\right) \tilde{n}_i^g} . \tag{5.14}$$

The effective population in patch $\tilde{n}_i$, i.e the number of people present during the day, is given by

$$\tilde{n}_i = \sum_{g=1}^{N_G} \tilde{n}_i^g , \tag{5.15}$$

with

$$\tilde{n}_i^g = \sum_{j=1}^{N_P} M_{ji}^g \, n_j^g . \tag{5.16}$$

For convenience, we define the effective mobility matrices

$$M_{ji}^g = (1-p^g)\,\delta_{ij} + p^g R_{ji}^g . \tag{5.17}$$

The function $f(x)$, which governs the dependence of the number of contacts on the density in each patch, is selected according to Ref. [475] as

$$f(x) = 1 + (1 - e^{-\xi x}) . \tag{5.18}$$

Finally, the last term in the exponent in Eq. (5.13) encodes the probability that an interaction involves a contagious individual. This probability is expressed as the ratio between infectious individuals $n_{j\to i}^{m,h}(t)$ ($m \in \{A, I\}$) and the total number of individuals present in $\tilde{n}_i^h$. Further, we distinguish here according to age strata h, patch of origin j and whether individuals are pre-symptomatic or symptomatic. In this sense, $n_{j\to i}^{m,h}(t)$ can be expressed as

$$n_{j\to i}^{m,h}(t) = n_j^h \, \rho_j^{m,h}(t) \, M_{ji}^h \,. \tag{5.19}$$

Since the dynamics is in discrete time, the models' computational cost is relatively low. Nevertheless, the model reproduces well the key characteristics with regards to the temporal evolution of the relevant epidemiological quantities. We used this model to anticipate the spread of SARS-CoV-2 in Spain early on in February 2020. Unfortunately, data availability was really poor at this point, and the regional distribution of the confirmed cases was not yet directly accessible. Calibrating our model with the confirmed cases in every region of Spain required screening a dozen local newspapers. We made these early efforts public on a personal webpage. However, we will not focus on the details of this early work here but instead on our efforts to introduce the impact of confinement on the model. As the virus started to affect an increasing part of the population during February/March of 2020, implementing confinement measures became an option for many countries to prevent an overload of the health system. In Spain, discussions around a lockdown started around the beginning of March 2020. The lockdown was then introduced on March 15 [476]. These events forced us to reconsider our model. In the following, we will outline how the model was adapted in order to account for the presence of confinement measures.

5.1.3 *Incorporating confinement into the model*

The most immediate effects of the confinement is a reduction in mobility and contacts [329, 477], which are represented by p^g and $\langle k^g \rangle$ in our framework. So let us assume that confinement is introduced at time t and affects a percentage κ_0^g in age group g. Accordingly, the mobility level p^g becomes

$$p^g(t) = (1 - \kappa_0^g(t)\,\Theta(t - t_c))\, p^g \,, \tag{5.20}$$

where $\Theta(t)$ is the Heaviside function. In other words, individuals affected by the confinement do no longer commute and stay in their patch of origin. Please note that, as a byproduct, z^g and $\tilde{n}_i^g$ also become time-dependent (Eqs. (5.14) & (5.16)).

Modelling the impact on the number of contacts is slightly more complex since it varies across age groups and strongly depends on the type of confinement that is introduced. In line with the restrictions in place during the first wave in Spain, we assume that the lockdown entails the closure of restaurants, bars and non-essential stores. Accordingly, physical contacts only occur at home and the workplace. The study that informs the mixing between age strata in our model separates contacts into three categories: home, workplace and leisure [72]. Hence, we can directly leverage this separation to formulate the impact of the lockdown on the average number of contacts in each age stratum. For the age stratum O, we assume that nobody is commuting to work, wherefore contacts take only place at home. Thus, after the introduction of confinement, the average number of contacts in the age group O $\langle k^O \rangle$ is given by $\langle k_{\mathrm{h}}^O \rangle$, where $\langle k_{\mathrm{h}}^O \rangle$ refers to the contacts taking place inside the household. In contrast to the elderly, a fraction of the population in age stratum Y and M continued to commute to work every day. This implies that a fraction $\kappa_0^g(t)$ has also contacts at the workplace in addition to the ones at home. Accordingly, after the introduction of the lockdown, the average degree $\langle k^g(t) \rangle$ for individuals belonging to age strata $g \in \{Y, M\}$ becomes

$$\langle k_c^g \rangle(t) = \kappa_0^g(t)\, k_{\mathrm{h}}^g + (1 - \kappa_0^g(t))(1-\delta)\, k_{\mathrm{h+w}}^g \,, \tag{5.21}$$

where k_{h}^g and $k_{\mathrm{h+w}}^g$ refer to the contacts made by confined (home) and non confined (home and work) individuals. The parameter δ accounts for the social distancing that is adopted in work settings. Eventually, the time-dependent average number of contacts for agents in age strata g can be expressed as

$$\langle k^g \rangle(t) = (1 - \Theta(t - t_c)) \langle k^g \rangle + \Theta(t - t_c) \langle k_c^g \rangle. \tag{5.22}$$

The time-dependence of $\langle k^g \rangle$ impacts the probability that individuals from age group g and patch i get infected at time i as

$$P_i^g(t) = 1 - \prod_{h=1}^{N_G} \prod_{j=1}^{N_P} \prod_{m \in \{A,I\}} (1 - \beta_m)^{z^g(t)\, \langle k^g \rangle(t)\, f\left(\frac{\tilde{n}_i(t)}{s_i}\right) C^{gh}\, \frac{n_{j\to i}^{m,h}(t)}{\tilde{n}_i^h(t)}} \,. \tag{5.23}$$

This equation completes how we describe the impact of the reduction in mobility and contacts inside the model framework. What we have not included so far is the impact of household isolation. To be more explicit, the confinement in place reshapes the physical contact network; it fragments it. For households where nobody leaves for work, it is improbable that someone gets infected since the virus has no entry point. From a mathematical point of view, one can see this presence of totally isolated households as a reduction of the pool of susceptible individuals. As a matter of fact, Maier et al. have shown that the isolation of individuals — treated as a reduction of the pool of susceptible individuals — is able to explain the linear growth in confirmed cases observed in Chinese provinces during early 2020 [478]. They proposed an elegant effective theory that incorporates a network effect — the isolation of households — into a mean-field model. The simplicity of their approach allowed us to take advantage of their theory and incorporate it with small changes into our model. A relevant quantity in this context is the fraction of individuals in patch i that live in a household in which none of its members are infected. We denote this by the quantity CH_i and can be expressed as

$$CH_i(t) = \left(\frac{1}{n_i} \sum_{g=1}^{N_G} \left(1 - \rho_i^{E,g}(t) - \rho_i^{A,g}(t) - \rho_i^{I,g}(t) \right) n_i^g \right)^{\sigma}, \tag{5.24}$$

where σ is the average household size. We made here the crude approximation that ages are randomly mixed inside households. With the quantity $CH_i(t)$, we can express how the fraction of susceptible individuals is reduced, while the isolated individuals are withheld in the new compartment $\rho_i^{CH,g}(t)$. These equations then read

$$\begin{aligned}
\rho_i^{S,g}(t+1) &= \left(\rho_i^{S,g}(t) + \rho_i^{CH,g}(t) \right) \left(1 - \Pi_i^g(t) \right) \times \\
&\quad \left(1 - \Theta(t - t_c) \left(1 - \phi \right) \kappa_0^g(t) \, CH_i(t) \right) && (5.25) \\
\rho_i^{E,g}(t+1) &= \left(\rho_i^{S,g}(t) + \rho_i^{CH,g}(t) \right) \Pi_i^g(t) \times \\
&\quad \left(1 - \Theta(t - t_c) \left(1 - \phi \right) \kappa_0^g(t) \, CH_i(t) \right) + \\
&\quad \left(1 - \eta^g \right) \rho_i^{E,g}(t) && (5.26) \\
\rho_i^{CH,g}(t+1) &= \left(\rho_i^{S,g}(t) + \rho_i^{CH,g}(t) \right) \times \\
&\quad \left(1 - \phi \right) \kappa_0^g(t) \, CH_i(t) \, \Theta(t - t_c) \, . && (5.27)
\end{aligned}$$

Put simply, we removed confined individuals that do not live with any infected individuals from the susceptible population. However, we introduced a

household permeability parameter ϕ that accounts for the non-perfect isolation that originates from essential activities such as buying groceries, drugs, etc. We will detail later how the different parameters are estimated when adjusting the model to the data. However, first, we focus on how confinement affects the reproduction number within our modelling framework.

5.1.4 *Expression of the reproduction number*

The ultimate goal of introducing the lockdown was to contain the epidemic, i.e. to push the reproduction number $\mathcal{R}$ below 1, and hence to reduce the pressure on the hospitals. In this sense, having an idea of the necessary reduction in mobility κ_0^c to achieve $\mathcal{R} < 1$ facilitates the design of the intervention. Additionally, the temporal evolution of the reproduction number gives detailed insight into when and how strongly different NPIs impacted the evolution of the epidemics. This correspondence prompted us to find an analytical description of the reproduction number inside our model framework.

The most straightforward and also most established way to express the reproduction number would be through the next-generation matrix [211, 214]. However, as pointed out in Sec. 2.2.5, the NGM assumes that the distribution of newly infected individuals is the one at equilibrium. More precisely, the effective reproduction number is the average of the entries of the NGM weighted by the normalised eigenvector, which describes the relative case distribution in each patch or age stratum. Nevertheless, in the presence of NPIs, data indicated that mobility levels changed on a daily basis. Therefore, given a generation time of about 5 days [232], it does not seem reasonable that the transmission dynamics reach such an equilibrium. It is thus desirable to find an alternative expression of the reproduction number which does not make the assumption of equilibrium. To do so, we can replace the leading eigenvector with the case distribution at time t inside the model. Further, this approach will allow us to define a reproduction number specific to each age stratum and patch.

The next-generation matrix expresses the instantaneous reproduction number. The instantaneous reproduction number at time t is defined as the number of secondary infections caused by an average individual infected at time t assuming that the environment is constant during his infectious period [216] (Sec. 2.3). To illustrate this, let us assume that the lockdown is introduced at time t_c. Accordingly, for all $t < t_c$ the instantaneous reproduction number

does not take into account the implementation of the lockdown. It assumes that the environment is constant. Thus, contacts do not vary over infection. However, given that infectiousness does not start directly after exposure to the virus, someone that gets infected a day before the introduction of the lockdown will only start to infect others once the lockdown is in place. In this sense, the generations infected at $t < t_c$ already have a reduced number of secondary cases and, hence, a lower reproduction number. The quantity that does take into account the changing environment during the course of infection is the case reproduction number [74]. Its definition is equivalent to the one of the instantaneous reproduction number, except that the assumption of a constant environment is not made. In the following, we will focus on the case reproduction number. Nevertheless, recovering the instantaneous reproduction number from our approach will be straightforward.

Now, to find the expression of $\mathcal{R}$, let us first simplify things and solely consider the different components it entails. As mentioned in Chapter 1, the reproduction number can be expressed as

$$\mathcal{R} = \tau_I \, \beta \, \langle k \rangle \, \rho_S \, . \tag{5.28}$$

Where τ_I is the infectious period, β the transmission probability, $\langle k \rangle$ the average contact rate, and ρ_S the fraction of susceptible individuals in the population. The translation of this to our example is not that straightforward, since we have a time dependence of the contact rate and the fraction of susceptible individuals. Further, we have two distinct probabilities for pre-symptomatic and symptomatic transmission. Eventually, we also have different patches and age strata. As a first step, we will forget about all the former and focus on how to deal with the presence of age strata and different patches. In line with the global reproduction number, we define $\mathcal{R}_i^g$, the reproduction number of patch i and age stratum g, as the number of secondary cases produced by an infected individual belonging to patch i and age stratum g. Therefore, $\mathcal{R}_i^g$ is given by

$$\mathcal{R}_i^g = \tau^g \, \langle \beta^g \rangle \sum_{j=1}^{N_P} \sum_{h=1}^{N_G} k_{ij}^{gh} \, \tilde{\rho}_j^{S,h} \, . \tag{5.29}$$

Since we do not consider the distinct infectivity of pre-symptomatics yet, we used the average infectivity $\langle \beta^g \rangle$ during the infectious period. The quantity k_{ij}^{gh} expresses the average number of daily contacts an individual belonging to patch i and age stratum g makes with individuals in patch j belonging to

age stratum h. The expression in Eq. (5.29) is equivalent to the mean-field one but takes explicitly into account the variation of the contact structure across patches and age strata. According to the definitions we made, the effective fraction of susceptible individuals in patch j and belonging to age stratum h, $\tilde{\rho}_j^{S,h}(t)$, can be expressed as

$$\tilde{\rho}_j^{S,h}(t) = \frac{1}{\tilde{n}_j^h(t)} \sum_{k=1}^{N_P} n_{k\to j}^{S,h}(t)\,, \tag{5.30}$$

with

$$n_{k\to j}^{S,h}(t) = n_k^h\, \rho_k^{S,h}(t)\, M_{kj}^h(t)\,. \tag{5.31}$$

Up to now, we have neglected the temporal variation of the daily contacts and the fraction of susceptible individuals, i.e. the environment. In this sense, since the environment has stayed constant, the expression we have so far corresponds to the instantaneous reproduction number. In contrast, to express the case reproduction number at time t, we need to analyse how many secondary cases a primary case that gets infected at t produces during the course of infection. To this aim, we will first mathematically describe how individuals transition through the compartments E, A and I (course of infection). Let us define the quantities $\zeta^{E,g}(t+1)$, $\zeta^{A,g}(t)$, $\zeta^{I,g}(t)$ as the probability for an individual to be in the respective compartments t time steps after being exposed to the virus, i.e. entering the compartment E. Given the model equations, the equations governing the evolution of these probabilities read

$$\zeta^{E,g}(t+1) = (1-\eta^g)\,\zeta^{E,g}(t) \tag{5.32}$$

$$\zeta^{A,g}(t+1) = \eta^g\,\zeta^{E,g}(t) + (1-\alpha^g)\,\zeta^{A,g}(t) \tag{5.33}$$

$$\zeta^{I,g}(t+1) = \alpha^g\,\zeta^{A,g}(t) + (1-\mu^g)\,\zeta^{I,g}(t)\,. \tag{5.34}$$

Further, since the individual enters the compartment E at $t=0$, the initial conditions are $\zeta^{E,g}(0) = 1.0$, $\zeta^{A,g}(0) = 0.0$ and $\zeta^{I,g}(0) = 0.0$. Then, solving the recursive Eqs. (5.32)-(5.34) explicitly, we find

$$\zeta^{E,g}(t) = (1-\eta^g)^t \tag{5.35}$$

$$\zeta^{A,g}(t) = \frac{\eta^g}{\alpha^g-\eta^g}\left[(1-\eta^g)^t - (1-\alpha^g)^t\right] \tag{5.36}$$

$$\zeta^{I,g}(t) = \frac{\eta^g\,\alpha^g}{(\alpha^g-\eta^g)\,(\mu^g-\eta^g)\,(\mu^g-\alpha^g)} \times\left[(\mu^g-\alpha^g)\,(1-\eta^g)^t + (\eta^g-\mu^g)\,(1-\alpha^g)^t\right.$$

$$+ \left(\alpha^g - \eta^g\right)\left(1-\mu^g\right)^t\Big] \,. \tag{5.37}$$

These probabilities allow us to define how infectious an individual is on average t time steps after having been infected, namely $\zeta^{A,g}(s-t)\,\beta_A + \zeta^{I,g}(s-t)\,\beta_I$. With this at hand, we can express the case reproduction number for patch i and age stratum g as

$$\begin{aligned}\mathcal{R}_i^g(t) \;&=\; \sum_{s=t}^{\infty}\Big(\zeta^{A,g}(s-t)\,\beta_A + \zeta^{I,g}(s-t)\,\beta_I\Big) \\ &\quad\times \sum_{j=1}^{N_P}\sum_{h=1}^{N_G} k_{ij}^{gh}(s)\,\tilde{\rho}_j^{S,h}(s)\,,\end{aligned} \tag{5.38}$$

where

$$k_{ij}^{gh}(t) = z^g(t)\,\langle k^g\rangle(t)\,f\left(\frac{\tilde{n}_j(t)}{s_j}\right) C^{gh}\,M_{ij}^g(t)\,. \tag{5.39}$$

To calculate the global reproduction number, the only thing left to do is to average $\mathcal{R}_i^g(t)$ with respect to the incidence distribution at time t. More precisely, we can express $\mathcal{R}(t)$ as

$$\mathcal{R}(t) = \frac{\displaystyle\sum_{m\in\{S,CH\}}\sum_{i=1}^{N_P}\sum_{g=1}^{N_G}\Big(\rho_i^{m,g}(t-1)-\rho_i^{m,g}(t)\Big)\,n_i^g\,\mathcal{R}_i^g(t)}{\displaystyle\sum_{m\in\{S,CH\}}\sum_{i=1}^{N_P}\sum_{g=1}^{N_G}\Big(\rho_i^{m,g}(t-1)-\rho_i^{m,g}(t)\Big)\,n_i^g}\,. \tag{5.40}$$

As previously pointed out, the weighted average here is equivalent to calculating the reproduction number through the NGM. The only difference is that the weighting is done according to the case distribution in the model and not according to the equilibrium distribution described by the leading eigenvector. In a similar way, one can now define a reproduction number for age stratum g as

$$\mathcal{R}^g(t) = \frac{\displaystyle\sum_{m\in\{S,CH\}}\sum_{i=1}^{N_P}\Big(\rho_i^{m,g}(t-1)-\rho_i^{m,g}(t)\Big)\,n_i^g\,\mathcal{R}_i^g(t)}{\displaystyle\sum_{m\in\{S,CH\}}\sum_{i=1}^{N_P}\Big(\rho_i^{m,g}(t-1)-\rho_i^{m,g}(t)\Big)\,n_i^g}\,, \tag{5.41}$$

or for a patch i as

$$\mathcal{R}^i(t) = \frac{\sum_{m\in\{S,CH\}}\sum_{g=1}^{N_G}\left(\rho_i^{m,g}(t-1)-\rho_i^{m,g}(t)\right)n_i^g\,\mathcal{R}_i^g(t)}{\sum_{m\in\{S,CH\}}\sum_{g=1}^{N_G}\left(\rho_i^{m,g}(t-1)-\rho_i^{m,g}(t)\right)n_i^g}\,. \tag{5.42}$$

5.1.5 *Critical confinement level*

The expression we provided above mainly serves to follow the temporal evolution of the reproduction number in a computationally not very costly way. However, the relatively complicated structure does not allow for an intuitive understanding of how the different parameters shape the reproduction number. As highlighted before, with respect to the lockdown, a crucial quantity is the critical confinement κ_0^c that contains the epidemic, i.e. $\mathcal{R} < 1$. To provide an analytical expression of κ_0^c and gain intuition on how confinement impacts $\mathcal{R}$, we forget about age strata and different patches and consider a well-mixed population of equal agents. In this case, the reproduction number can be expressed as

$$\begin{aligned}\mathcal{R} &= \rho^S\left(\frac{\beta^A}{\alpha}+\frac{\beta^I}{\mu}\right)\left(1-\kappa_0\left(1-\phi\right)\langle CH\rangle\right)\\ &\quad\times\left(\kappa_0\langle k_{\mathrm{h}}\rangle+\left(1-\delta\right)\left(1-\kappa_0\right)\langle k_{\mathrm{h+w}}\rangle\right)\,.\end{aligned} \tag{5.43}$$

One can see that $\mathcal{R}$ has a quadratic dependence on κ_0^c. Setting $\mathcal{R}=1$, we find the following equation for the critical confinement

$$\begin{aligned}&(\kappa_0^c)^2\left(1-\phi\right)\langle CH\rangle\left(\langle k_{\mathrm{h+w}}\rangle\left(1-\delta\right)-\langle k_{\mathrm{h}}\rangle\right)\\ &\qquad-\kappa_0^c\left(\langle k_{\mathrm{h+w}}\rangle\left(1-\delta\right)\left(1+\langle CH\rangle\left(1-\phi\right)\right)-\langle k_{\mathrm{h}}\rangle\right)\\ &\qquad\qquad+\langle k_{\mathrm{h+w}}\rangle\left(1-\delta\right)-\frac{1}{\rho^S\left(\frac{\beta^A}{\alpha}+\frac{\beta^I}{\mu}\right)}=0\,.\end{aligned} \tag{5.44}$$

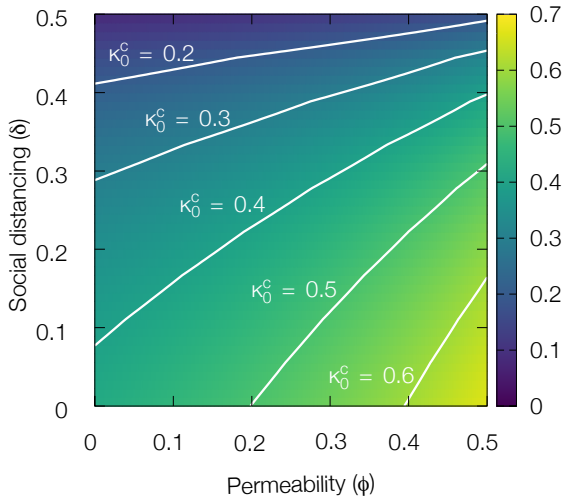


Figure 5.2: Critical confinement κ_0^c computed from Eq. ((5.49)) needed to ensure that the system reaches the bending regime, i.e. $\mathcal{R} = 1$, as a function of the household permeability ϕ and the social distance δ.

Defining the following coefficients

$$A = (1-\phi)\langle CH\rangle\left(\langle k_{\mathrm{h+w}}\rangle(1-\delta) - \langle k_{\mathrm{h}}\rangle\right), \tag{5.45}$$

$$B = \langle k_{\mathrm{h+w}}\rangle(1-\delta)\left(1+\langle CH\rangle(1-\phi)\right) - \langle k_{\mathrm{h}}\rangle, \tag{5.46}$$

$$C = \langle k_{\mathrm{h+w}}\rangle(1-\delta) - \frac{1}{\rho^S\left(\frac{\beta^A}{\alpha}+\frac{\beta^I}{\mu}\right)}, \tag{5.47}$$

we can rewrite Eq. (5.44) as

$$A\,(\kappa_0^c)^2 - B\,\kappa_0^c + C = 0\,. \tag{5.48}$$

The solution of the above equation is then given by

$$\kappa_0^c = \frac{B \pm \sqrt{B^2 - 4\,A\,C}}{2\,A}\,. \tag{5.49}$$

where only the physical solution $\kappa_0^c \in [0,1]$ should be retained. To better illustrate the impact of the household permeability and social distancing, we show in Fig. 5.2 their effect on κ_0^c. As one would expect, the dependence on both parameters is monotonous. While the calculation here is mostly qualitative, the quantitative results give a good intuition about the necessary confinement for low levels of social distancing (low δ) and intense confinement (low

ϕ). In the following, we will adjust the model to the epidemiological data and highlight that the dynamics was precisely in this regime of low θ and ϕ.

5.1.6 *Data adjustment*

We have presented all the details of the model employed to describe the spread of SARS-CoV-2's first wave in Spain during the spring of 2020. Accordingly, the remaining step is to adjust the model to the epidemiological data. However, the epidemiological data was one of the main issues. The reader may have noticed that our model only includes ICU admissions; hospital admissions are not explicitly modelled. Unfortunately, in early 2020, reliable hospital admission data was unavailable. Further, after working for a while with ICU data, we realised that some regions reported admissions while others occupations and some even switched between the two multiple times. Therefore, the only robust data stream we were left with was the daily deaths [479]. Nevertheless, we will see in the next section that even the daily deaths were subjected to substantial underreporting.

Now, to adjust the model to the daily fatalities, we fix a series of parameters from the literature. Especially, parameters specific to SARS-CoV-2 such as the transition rates and probabilities between the clinical compartments are fixed according to the literature. An overview of all these parameters is given in Appendix B. In contrast, we fitted ϕ and δ that define the impact of confinement. Further, we fitted the initial reproduction number of the disease, i.e. the transmission probability and β_I as well as the transition rates η^g, α^g and μ^g. Furthermore, we also estimated the rate at which individuals exit the pre-deceased, η^g, and the pre-hospitalised in ICU, λ^g, compartment. As we previously highlighted, the contact matrix C was fixed according to Ref. [72]. Mobility data, R^g_{ij}, comes from an empirical survey regarding the commuting patterns in Spain [480]. The population structure for each municipality, n^g_i and s_i, is given by the statistical agency of Spain [481]. We estimate the confinement level directly from the reduction in mobility reported by the authorities [480]. Since only work-related mobility was allowed during the lockdown, the reduction in mobility reflects well the population under confinement. The mobility is relative to a reference day on November 19, 2019. We assume that the time series are equal for all age groups, i.e. $\kappa_0^Y = \kappa_0^M = \kappa_0^O = \kappa_0$. Further, confinement measures are introduced in the model on March 14, t_c, as the lockdown was announced.

Before fitting the parameters to the daily deaths, we must distribute the initial seeds spatially. To this aim, we consider all the reported cases until March 3 and use them as initial seeds at the start of our simulation on February 9. However, due to underreporting, the seeds are not sufficient to explain the observed fatalities in some regions. Accordingly, we add infectious individuals in the capital of each Comunidad Autónoma (CA) and fit the initial seed through the Nelder-Mead method [482]. Instead of fitting all the seeds in every capital simultaneously, we iterated each CA and adjusted the number of initially infected individuals. The presence of mobility implies that introducing a seed in one CA also affects the spread in another CA. To account for this issue, we repeated the iterative fit of the seeds in each CA three times.

Given the distribution of the seeds, we then proceeded to adjust the epidemiological parameters — β_I, η^g, α^g, and μ^g. We used approximate Bayesian computation (ABC) [238]. More specifically, we took advantage of the implementation of ABC in Julia package ApproxBayes [483]. We fitted the model to the time series of nationwide daily deaths until April 8 [479]. The objective function for the ABC was the logarithmic least squares error. The posterior of the inference is shown in Fig. B.1. We then validated the model's predictive power with the daily deaths from April 8 forward. Fig. 5.3A shows the adjustment of the fit and the validation. The model seems to describe the evolution of daily deaths reasonably well. Please note that the horizontal bars indicate the different intensities of the lockdown that were in place. In LM, all work-related mobility was allowed, while only essential workers were allowed to commute in a stricter scenario (EM). In ECS, in addition to commuting for work, individuals were allowed to go for a walk or sports. Finally, in R1 some commercial activity was reopened.

Besides the daily fatalities, we also validate the model with the daily cases [479]. This comparison requires correcting for underreporting. We do so by rescaling the reported cases with the total underreporting estimates until May 4 [484]. The seroprevalence study concluded that around 5% of the Spanish population got infected, while the reported cases correspond to only 0.5%. Accordingly, only about 1 out of 10 infections were diagnosed and reported. Nevertheless, rescaling the reported cases, we see in Fig. 5.3B that the time series agrees rather well with the model predictions. However, the predicted cases peak slightly earlier than the reported ones. Next to the adjustment to the national time series in Figs. 5.3C & D, we can see that the model predicts the spatial variations fairly well. For the daily fatalities and the daily cases

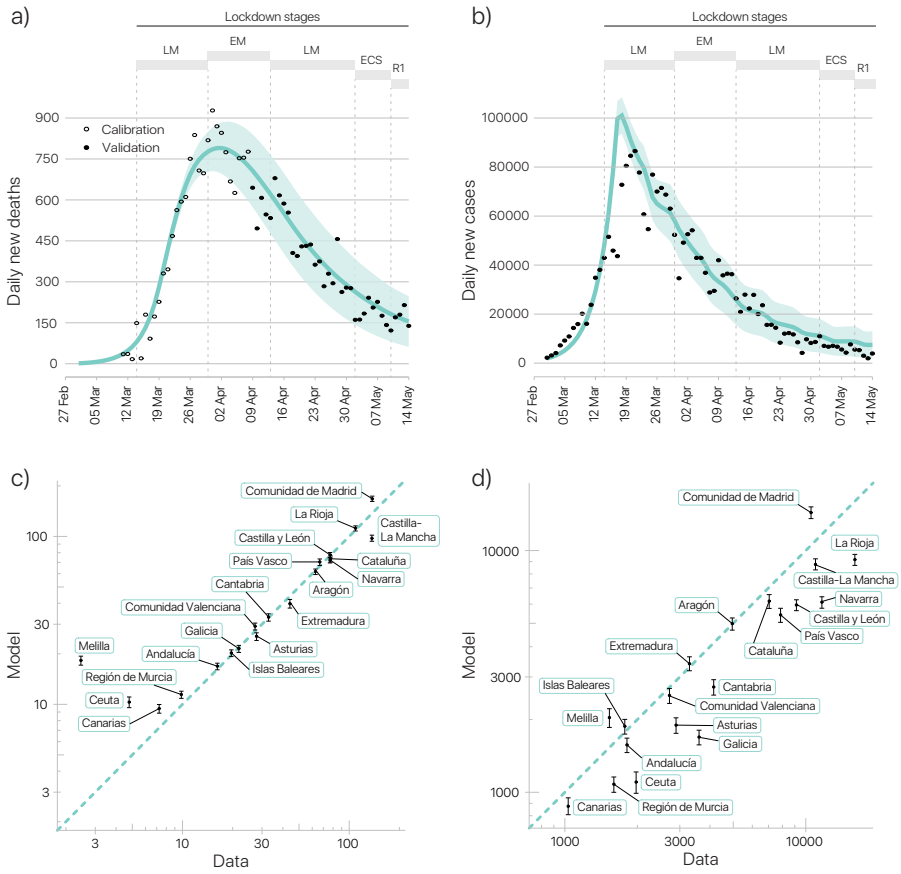


Figure 5.3: Top: Solid lines show model predictions for the daily fatalities **(A)** and the daily number of new symptomatic individuals **(B)** whereas dots correspond to real data. The shadowed areas represent the 95% prediction interval. Bottom: Correlation between the accumulated number of deaths ($r = 0.98$) **(C)** and the accumulated number of reported cases ($r = 0.90$) **(D)** per 100000 inhabitants observed by May 15 and those predicted by the model. The plot shows this correlation at the level of Comunidades Autónomas of Spain. Error bars cover the 95% prediction interval.

in the different CCAA (plural of CA), the model output shows a correlation coefficient with the death and case data of 0.98 and 0.9, respectively.

Eventually, we can also look at the reproduction number's evolution. In Fig. 5.4, we present the evolution of $\mathcal{R}(t)$ along with the reported values of $\kappa_0(t)$. We see that the confinement level between 0.5 and 0.8 was sufficient to drive the reproduction number below one. The reproduction number is already below 1 early on since we consider the case reproduction number

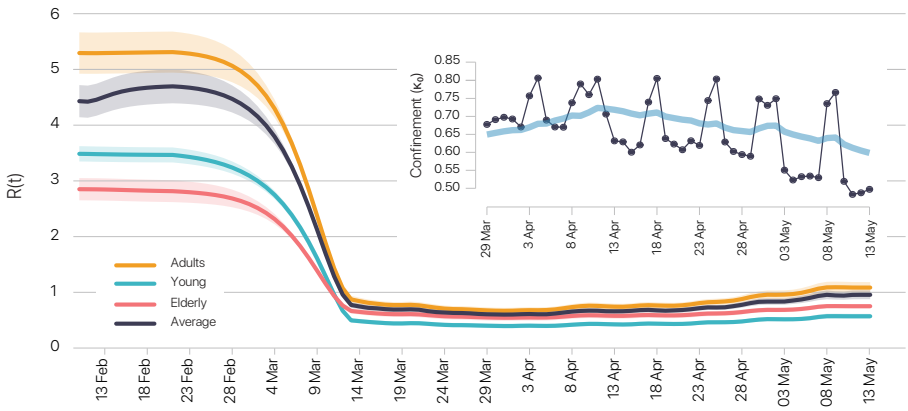


Figure 5.4: Temporal evolution of the effective reproduction number $\mathcal{R}(t)$ for each age stratum and its average computed according to Eq. ((5.40)). Inset: The black line represents the temporal evolution of the degree of confinement $\kappa_0(t)$ extracted from data daily updated by INE [480]. The blue line shows the average over a 7 days time window to smooth out fluctuations observed in the data. To improve the visibility of the inset, we do not represent the degrees of confinement from March 14 until March 28.

and not the instantaneous one. Accordingly, changes in the contact structure induced by the implementation of the lockdown are already reflected in the reproduction number various days before their manifestation. Further, as mobility increases towards the end of the lockdown, the reproduction number of adults surpasses 1. However, since $\mathcal{R}^Y$ and $\mathcal{R}^O$ are still below 1, the overall reproduction number does not exceed 1. The same could occur with respect to the different patches. Even though some patches may show $\mathcal{R}_i > 1$, the overall reproduction number can still be below 1.

5.1.7 *Discussion*

The low value of the fitted household permeability ($\phi = 0.174$ (CrI : $0.079 - 0.269$)) indicates that confinement was relatively strong. We find a value of $\delta = 0.207$ (CrI : $0.053 - 0.359$) for the social distancing parameter. Given a confinement level between 0.5 and 0.8, we see in Fig. 5.2 that the dynamics was precisely in this regime of low ϕ and δ. In this sense, the simple qualitative consideration of one patch accurately predicts that $\kappa_0 \in [0.5, 0.8]$ is sufficient to contain the epidemics and push $\mathcal{R}$ below 1. With a mobility reduction in this range, we observe that the first lockdown was sufficient to

contain the epidemic. The strengthening of the lockdown on March 29 (EM), i.e. the closure of all non-essential business, seems only to have resulted in a slight reduction in $\mathcal{R}$. However, bear in mind that we came to this conclusion after using data until April 8, which is almost two weeks after the start of the intensified confinement.

In contrast, at the time that the strengthening of the lockdown was discussed, fitting the model to the daily fatalities, one could not conclusively say that $\mathcal{R} < 1$. In contrast, our results indicated that strengthening the lockdown — higher κ_0 — would be necessary to contain the epidemic. Partially, these findings motivated a series of experts to write an open letter that requested the strengthening of the lockdown [485]. Two factors may have contributed to this prudent judgement. First, predicting the peak incidence is subject to substantial uncertainty. Castro et al. perfectly highlighted in a Bayesian approach that, through pure data fitting, the prediction on whether an intervention is sufficient can only be made many days after its introduction [486]. If one utilises a mechanistic model for the relation between mobility and the reduction in contacts, as we do here, these conclusions may only apply to a lesser extent. Nevertheless, our approach also requires fitting the parameters ϕ and δ. In this regard, it is to mention that we also could have summarised the contributions of ϕ and δ in one effective parameter, which reduces the transmission probability [297, 314]. Even though both parameters make sense from a conceptual/mechanistic point of view, working with an effective theory would have facilitated the fitting of the model.

Further, another improvement could have been to describe the dynamics at the same resolution level as the one of the available data. Unfortunately, case data was only available at the level of a CA. Accordingly, it was impossible to validate the results at the level of municipalities. Working at the level of a CA would have also substantially reduced the computational complexity. However, as pointed out in the beginning, the model was initially designed to understand the spatial spread of SARS-CoV-2 when almost no data was available. At this point, and for this aim, our approach was reasonable. However, later, as the objective shifted towards prediction, the temporal constraints and the exceptional circumstances made the adequate continuous adaption of the methodology difficult.

After quite a detour, the second point that made the decision-making process challenging was the substantial reporting delay. As a matter of fact, looking at Fig. 5.3B, we see that cases were increasing almost until the end of March, even though the reproduction was already below 1 around two weeks

before. The discrepancy between the two stems from the time between infection and detection and, even more importantly, until the detected case was reported in the national statistics. In the model adjustment to the reported cases, we somehow contoured this problem by assuming that individuals are only detected once they exit the infectious compartment. This approach allowed us to "artificially" introduce the reporting delay in the model and compare it with the reported cases. Nevertheless, as highlighted before, model predictions peak slightly before the reported data. Although we were aware, back in March 2020, that a reporting delay existed, we were not aware of how pronounced it was. Accordingly, despite confinement being in place for almost two weeks, we still observed increasing daily fatalities and reported cases. In this reality, it was not straightforward to convince oneself that the worst of the epidemic was over, i.e. that $\mathcal{R}$ was already below 1 for some time. The substantial reporting delay became apparent once the Spanish national centre of epidemiology, the CNE, made more detailed data about the first wave in the summer of 2020 available. Only the availability of this data made an in-depth analysis of the first wave possible, which is what we will do in the next section.

5.2 RETROSPECTIVE ANALYSIS OF THE FIRST WAVE

After the end of the lockdown, on May 2, case numbers stayed low for almost the entire summer of 2020. The Centro Nacional de Epidemiología (CNE) took advantage of this relatively calm period to revisit the data that was collected during the first wave. Thanks to this effort, data for hospital admissions became available, case data at the resolution level of provinces was made public, and the information regarding the day of symptom onset of detected cases became accessible [487]. Additionally, the collaboration between the ministry of transport and the Barcelona Supercomputing Centre made large-scale mobile phone data available [488]. While during our modelling efforts presented in the section before, we had only daily, nationally aggregated mobility data at our disposal, the newly available data showed the evolution of mobility for every hour and municipality.

The availability of this more detailed data motivated us to reconsider, with the required distance, the factors that shaped the spread of the virus. Similar retrospective analyses were performed in many other countries [468, 489–492]. As previously pointed out, they mainly serve to evaluate the impact and effi-

ciency of the different NPIs, which may later contribute to the design of more efficient containment strategies [314, 316]. These studies are paradigmatic examples where modelling is not used to make predictions about the future but rather to analyse past events. Here, we contrast the evolution of mobility with epidemiological data such as case numbers, hospital admissions and fatalities. From a methodological point of view, we blend a direct analysis of the epidemiological data with a model-based inference. The results indicate that the reproduction number had already substantially decreased before the implementation of the lockdown. Eventually, we leverage the modelling efforts to build counterfactual scenarios and quantify the impact of the epidemic response that preceded the national lockdown.

5.2.1 *The evolution of mobility*

Mobility was shown to be a good indicator of the epidemic evolution [477, 489, 493]. In particular, in the absence of efficient contact tracing, social distancing or the use of face masks, the reduction of contacts is the only option to contain the spread of SARS-CoV-2 [477]. As leveraged in the model before, a reduction in the contact rate generally leads to a reduction of the mobility level [231, 466, 477]. In this sense, during the first wave, when contact tracing and mask mandates were not yet implemented, mobility was the primary tool to analyse the impact of NPIs. The mobility data we use here stems from different major network providers and consists of anonymised individual trajectories of about 13 million individuals. Using sociodemographic indicators, the ministry of transport inferred an Origin-Destination matrix from this raw data in collaboration with the Barcelona supercomputing centre [488]. The mobility reduction is calculated with respect to the corresponding day in a reference period (February 14-20).

Fig. 5.5A depicts the evolution of mobility on a national level. The data indicates that mobility substantially decreased from March 10 onward, previous to the lockdown. The decrease in mobility coincides with the introduction of the first regional NPIs in Spain. Madrid and the province of Alava (Basque Country) closed the entire educational system on March 10 [476]. The decrease in mobility further coincides with an increase in awareness of SARS-CoV-2 as indicated by the increase of COVID-19-related Google searches (Fig. 5.6) and Tweets [494] from March 9 onward. This finding let us hypothesise that the interplay between local measures and the display of

the epidemic in other places, such as Italy, increased individual caution and led to a reduction in mobility. Eventually, mobility was at about 50% of the pre-pandemic level after the implementation of the lockdown, which rose to 60% towards the end of the lockdown.

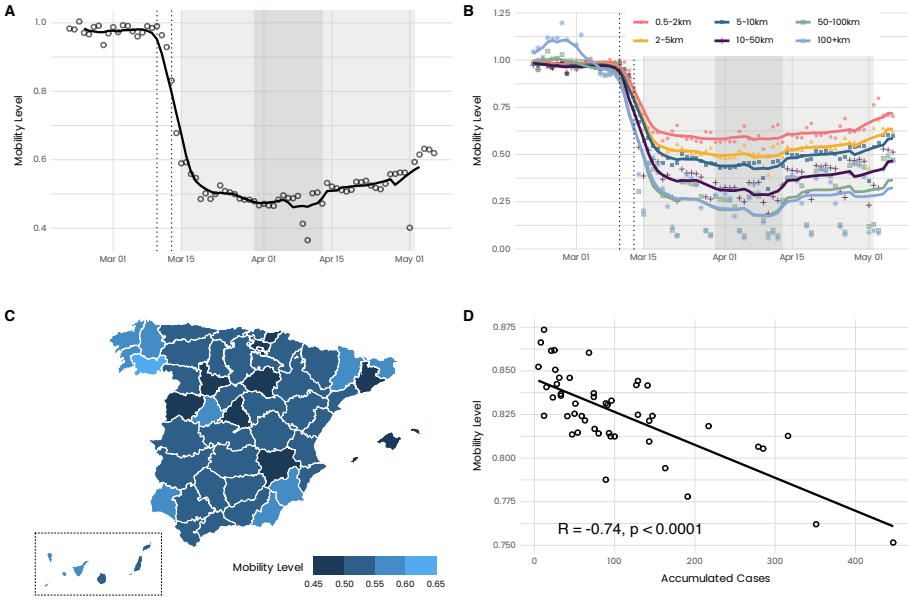


Figure 5.5: **(A)**: Average, nationwide aggregated mobility before and during the lockdown. Dots indicate the data points while the solid line shows a rolling centred 7 day average. Vertical dashed lines indicate the first NPIs introduced in Spain on March 10 (school closures in Madrid and Basque Country) and the announcement of the state of alarm on March 13. The shaded area in light grey indicates lockdowns 1 and 3. In dark grey, we indicate lockdown 2 where, in addition, all non-essential economic activity was shut down. **(B)**: Nationwide aggregated mobility but separating the distances of trips. Dots indicate data points, and the solid lines represent a rolling centred 7 day average. In general, long-distance trips show a much stronger reduction than shorter trips. **(C)**: Mobility level during lockdown (March 15 – May 2) for the total number of trips in the provinces (administrative subdivisions of CCAA) of Spain. Ceuta and Melilla are not shown. Please note that the Canary Islands (islands at the bottom) were moved to be visible. **(D)**: Correlation between the accumulated number of cases until lockdown and the mobility level. Points represent the provinces of Spain. We excluded the provinces of Madrid and Barcelona since they represent statistical outliers due to their high number of cases. R and p denote the Pearson correlation coefficient and the associated p-value. Similarly, the Spearman correlation coefficient is found as -0.7 with $p < 0.0001$.

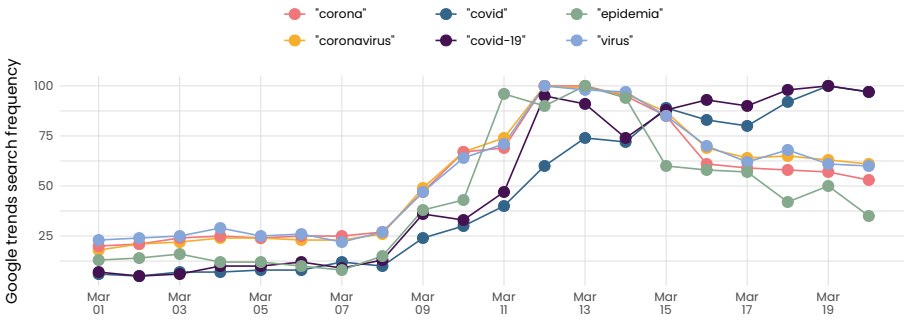


Figure 5.6: Frequency of different COVID-19 related search queries according to Google Trends. Note that Google Trends normalises the frequency of queries between 0 and 100 during the considered period. All keywords see a sharp increase from March 9 onward. On the same day, the Community of Madrid and the Basque Country announced the closure of educational centres. These closures correspond to the first NPIs introduced in Spain [476].

Beyond the national, aggregated evolution, mobility reduction is very heterogeneous. For example, long-distance trips were only about 15-20% of the pre-pandemic level and exhibited thus a much more significant reduction than short trips (Fig. 5.5B), which is consistent with findings from other countries [329, 495]. The mobility level also exhibits spatial heterogeneity (Fig. 5.5C). Madrid, the strongest hit province in terms of infections, also exhibits the strongest mobility reduction. Further, highly affected provinces such as Álava and Barcelona are among the five provinces with the strongest reduction in mobility. Quantitatively speaking, we find a correlation of $R = -0.74$ between the mobility level and the number of reported cases previous to the lockdown (Fig. 5.5D). During the lockdown, the correlation between reported cases and mobility is substantially reduced (Fig. C.2 in Appendix C). The lower correlation seems reasonable, as, during the lockdown, mobility reduction was mainly mandated and not voluntary. Interestingly, we find a higher correlation for the absolute number of reported cases than relative to population size. The same holds for hospital admissions, ICU admissions and fatalities (Fig. C.3 in Appendix C).

5.2.2 *Evolution of the reproduction number*

A natural way to test whether the mobility reduction affected the course of the epidemic is to contrast the evolution of mobility with the reproduction

number. In this case, it is most convenient to consider the instantaneous reproduction number and not the case reproduction number [216]. Since the instantaneous reproduction number assumes a constant environment during the course of infection, any changes in the epidemic dynamics are immediately and sharply reflected in its evolution. In contrast, as previously mentioned, the case reproduction at time t also considers future changes in the transmission dynamics [74]. Accounting for the future changes essentially smooths out any abrupt variations in the epidemic dynamics. For this reason, we will focus on the instantaneous reproduction number.

5.2.2.1 Estimation of the exposure times

As highlighted in sec. 2.3.3, the estimation of the reproduction number first requires inferring the time of infection of the reported cases, here referred to as exposure times [225, 233]. More precisely, the inference of the exposure times consists of a deconvolution of the detection date with the time between exposure and detection [233]. Typically, this time interval does not correspond to a fixed amount of days but instead follows a distribution [18]. However, the distribution that describes the time between exposure and detection is generally challenging to determine since the time of infection is only rarely precisely known. Nevertheless, symptomatic individuals can report the day that symptoms started (symptom onset). Accordingly, given that the incubation period is known, we can deconvolute the incubation period from the time series of symptom onset to infer the exposure times. The incubation period is inferred from transmission chains, where one has detailed knowledge about the time of infection. For SARS-CoV-2, the incubation period was shown to be between 4 and 6 days and is generally well described by a gamma distribution [62, 496].

For the above reasons, the time series of symptom onset made available in the summer of 2020 by the CNE [487] is particularly helpful in making a retrospective analysis. It allows us to estimate the exposure times and, later, the reproduction number. Multiple methods exist to perform the deconvolution [234–236]. Here, we will use the methods proposed by Becker et al., which is based on a maximum likelihood approach [234, 497]. The algorithm, referred to as backprojNP [498], is already implemented in the surveillance package in R [499]. Credible intervals are calculated through bootstrap with 1000 samples ($B = 1000$). Further, the estimation of the exposure times is smoothed through a centred rolling average of 7 days ($k = 6$). Finally, we

fix the incubation period as a gamma distribution with mean 5.2 days and a standard deviation of 2.8 days according to Ref. [232].

Fig. 5.7A illustrates the exposure times compared to the day of symptom onset and the reporting date. We see that the exposure times substantially precede the reporting date. More precisely, the two curves peak with a delay of 16 days, consistent with the estimation in Ref. [500]. Considering the incubation period of about 5 days, this highlights the substantial delay from symptom onset until individuals got tested, results were received, and eventual positive cases were reported. The delay of 16 days is about 3 days longer than what was estimated for the United States [501]. We repeated the same

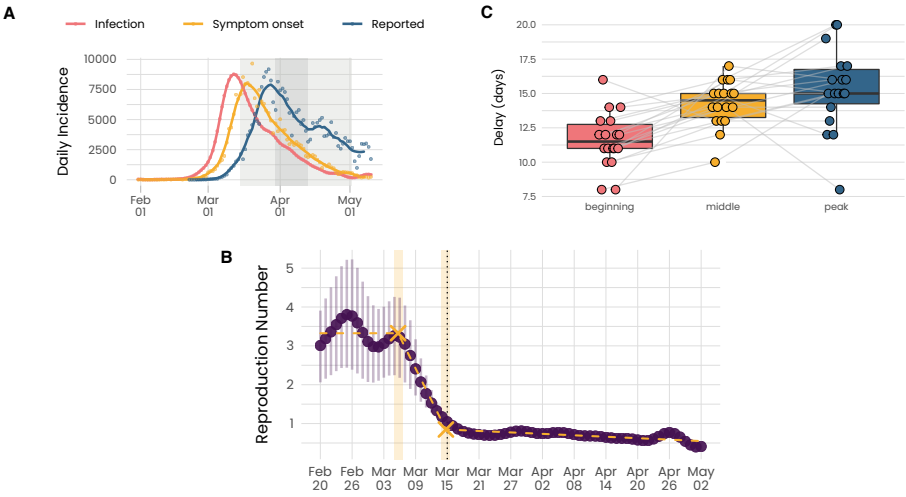


Figure 5.7: **(A)**: Cases when they were reported (blue), the onset of symptoms (yellow) and the reconstructed exposure times. Dots indicate data points, and the solid line is a rolling centred 7 day average. The shaded area for the exposure times represents 95% credible intervals. **(B)**: Estimation of Rt. Vertical bars indicate 95% credible intervals. The dashed, yellow line represents the three linear segments we identified. The shaded yellow area shows the 95% credible interval for the two breakpoints. The two breakpoints are between March 5 (CI: 4 – 6) and March 15 (CI: 14 – 16), where the latter coincides with the implementation of the lockdown. **(C)**: Delay between the exposure time and the reporting date. The position is defined with respect to the peak. To be more precise, we show the time difference when both curves reached x% of their peak value. The three points –beginning, middle and peak– correspond to 5, 50 and 100 % of daily infections compared to the peak incidence. Each point corresponds to a CA. Grey lines indicate how the delay evolves for each CA. We note that the delay steadily increases towards the peak in almost all CCAA.

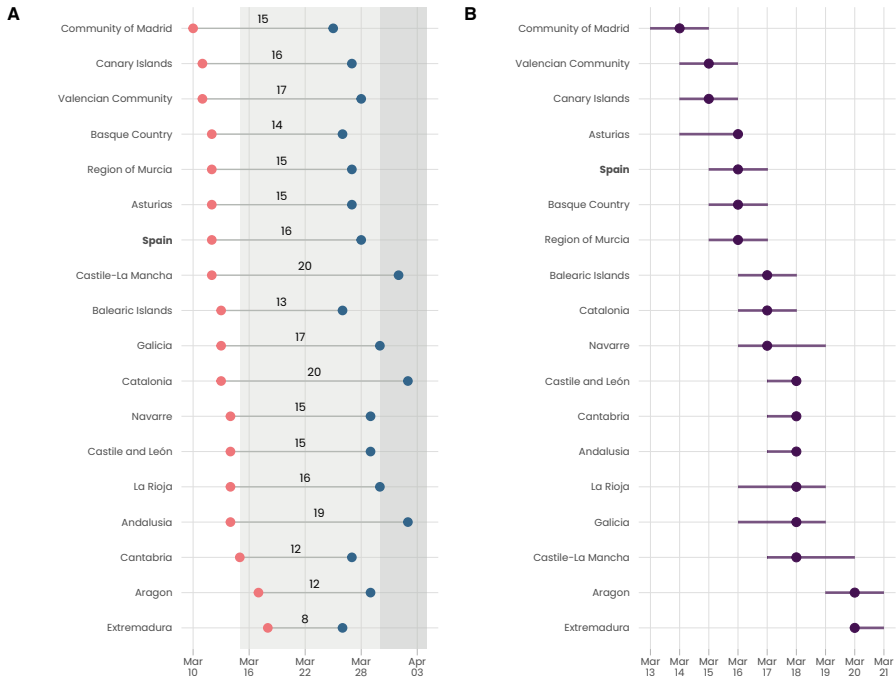


Figure 5.8: **(A)**: Blue and red dots indicate the day the curve of the reported cases and exposure times reached their respective peak. Number in between denote the difference in days between these dates. **(B)**: The day Rt was first below one in the different CCAA. Horizontal bars indicate 95% credible intervals.

analysis for each CCAA, which is presented in Fig. 5.8A and Fig. C.4 in Appendix C. Among the CCAA, the delay between exposure and reporting varies from 8 days in Extremadura to 20 days in Catalonia.

5.2.2.2 Estimation of the reproduction number

We can now estimate the reproduction number with the exposure times at hand. We use the package EpiEstim to infer the reproduction number from the median of the exposure times [218–220]. For the infectivity profile, we fix the generation time according to Ganyani et al. as a gamma distribution with mean 5.2 (CI: 3.78 – 6.78) days and 1.72 (CI: 0.91 – 3.93) days standard deviation [232]. The choice of the generation time is consistent with the incubation period we assumed to infer the exposure times [496]. Ganyani et al. used the same incubation period to estimate the generation time. We assume a standard deviation for the mean and the standard deviation of the genera-

tion time of 1.0 days and 1.2 days, respectively. However, we bound the values for the mean and standard deviation by the estimations of Ganyani et al. We considered a centred rolling average of 7 days to estimate the reproduction number. Moreover, we obtained the confidence intervals through bootstrap taking 100 samples of the generation time distribution and considering 100 posteriors for each of these samples (n1 = 100, n2 = 100).

Fig. 5.7B shows the inferred evolution of the reproduction number. The results indicate that the epidemic spread initially with a reproduction number of around 3. More surprisingly, the reproduction number started to decrease around March 5/6 and was below one shortly after the implementation of the lockdown between March 15 and 17. The same can be observed in most CCAA (Fig. C.1 in Appendix C). In Fig. 5.8B, we can see that all CCAA had a reproduction number below one between March 13 and 17. The temporal evolution of the reproduction number in the different CCAA is shown in Fig. C.1 in Appendix C. Due to the high case numbers, Madrid strongly dominates the temporal evolution of the reproduction number at the national level.

The decrease in the reproduction number at the national level around March 5/6 is surprising since it precedes the decrease in mobility that starts on March 9. As we previously pointed out, without the use of face masks or efficient contact tracing in place, the primary way to slow down the spread of the epidemic is a reduction of contacts. Further, such reduction in contacts should be reflected in a reduction of mobility, as individuals go out less to meet friends or work from home. Another possible explication for a decrease in the reproduction number preceding the one in mobility could be a saturation in the test capacity. If the detection rate stays constant over time, the estimation of the reproduction number is not affected by the underreporting [233]. However, suppose the test capacity is finite, and the epidemic continuously expands. In that case, the detection rate will start to decrease at some point, which, if the estimation is based on the reported cases (as we do here), would cause a decrease in the reproduction number.

5.2.2.3 Increasing reporting delay towards the epidemic peak

We can find an indication of the saturation of the test capacity by looking at the evolution of the delay between infection and reporting. Fig. 5.7C compares the delay between the two time series at different points with respect to the epidemic peaks. More precisely, we find an initial delay of 12 days that

increased to 16 days towards the epidemic peak. Further, the delay between exposure and notification increased in almost all CCAA as shown in Fig. C.4 in Appendix C. The only exceptions are Aragon and Extremadura. Interestingly, the CA with the lowest delay is also the last, with a reproduction number below 1. This further supports the hypothesis that the testing facilities got under increased stress, leading to an anticipated decrease in the reproduction number.

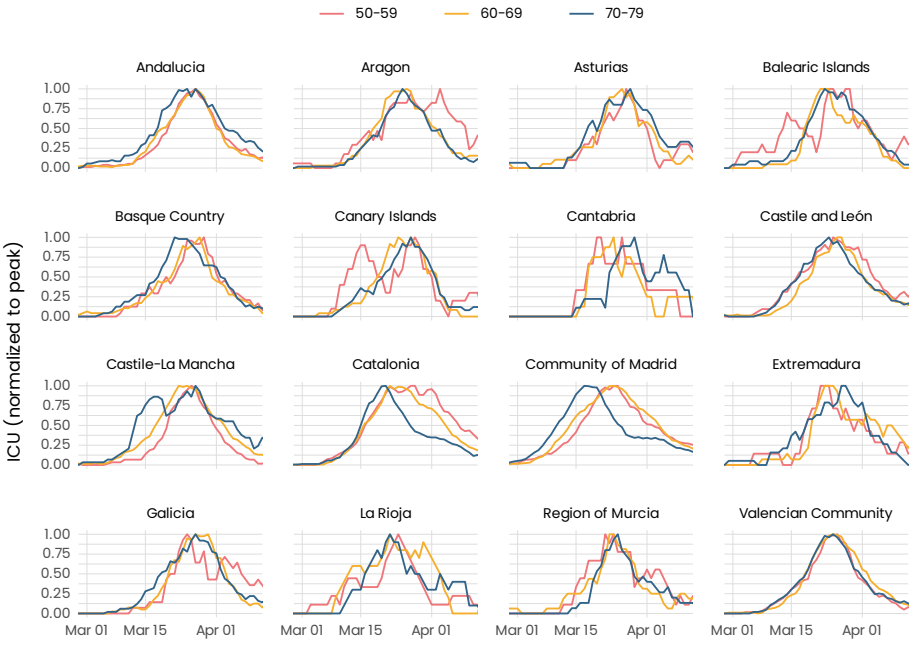


Figure 5.9: Seven day centered rolling average for the evolution of ICU admission for different age groups separated with respect to the CCAA [487]. We excluded CCAA where numbers were too low to exhibit a robust evolution. Admissions were normalized with respect to their peak value. We observe that in the strongest hit CCAA, Catalonia and Madrid, the admissions of individuals in the age group 70–79 peak earlier and decrease much faster. In contrast, in most other CCAA, the evolution of admission among different age groups is very similar. It seems reasonable that changing admission criteria lead to this fast decrease. We could not find any other explanation for this discrepancy. Among others, this inconsistency caused us to not include ICU admission as a data stream for our model-based inference.

5.2.3 *Inference with a minimal epidemic model*

A way to corroborate our intuition regarding the origin of the early decrease is to consider more robust data streams, which are not subjected to underreporting to the same extent. Typical examples are hospital admissions, ICU admissions and deaths. If there is no overload in the hospitals such that admission criteria change in time, these indicators should not show any underreporting. Nevertheless, even those data streams must be taken with a grain of salt for the first wave in Spain. As we highlighted earlier, the reporting of ICU admissions was very inconsistent across the CCAA. Some reported occupation, others admissions and a few even switched between the two from time to time. Further, let us look at the temporal evolution of ICU admissions (Fig. 5.9). We observe that the two strongest hit CCAA, Madrid and Catalonia, show an anticipated peak of the age group 70—79 with respect to the younger age groups and a rapid decrease after that. To our understanding, the only explanation for such a pronounced variation in the qualitative evolution is a change in admission criteria. For the above reasons, we refrained from considering ICU admissions.

An analysis of the excess deaths showed that even the number of deaths was subject to substantial underreporting. Many residents in care homes died without a diagnosis and thus did not appear in the statistics. Official reports suggested that around 28'000 individuals died during the first wave [487],

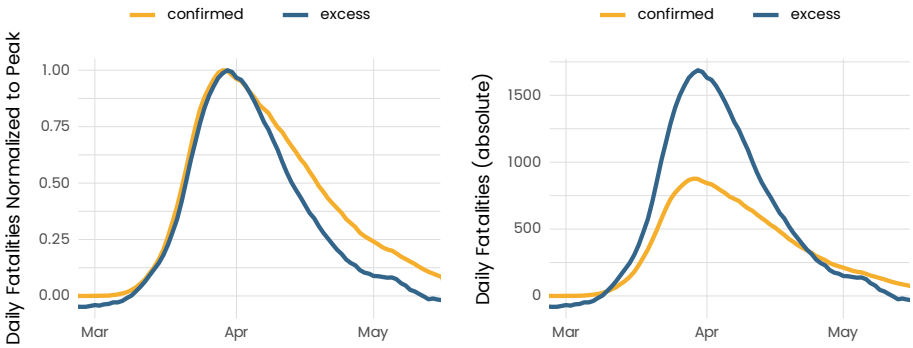


Figure 5.10: Evolution of the daily fatalities. The left panel shows the evolution normalised with respect to the peak value, while the right panel shows the fatalities in absolute terms. We applied a 7-day centred rolling average to the data. We show reported deaths [487], and excess deaths [502]. While the rise in fatalities is almost equivalent between the two curves, the daily excess deaths decrease much faster than the reported fatalities.

while in the same period, there were excess deaths of about 50'000 [502]. Assuming that all excess deaths can be attributed to COVID-19, which may not be the case, this would amount to an underreporting of 40%. However, if we look at the temporal evolution of reported and excess deaths (Fig. 5.10), we observe that the evolution of the two time series is pretty congruent. The main difference is a faster decrease in excess deaths after the epidemic peak. This faster decrease seems intuitive since, after the epidemic peak, health authorities were under less pressure, which should improve medical attendance. Despite this faster decrease, we consider the daily deaths as a sufficiently robust data stream for the analysis. Finally, we will also consider hospital admissions whose evolution does not show unusual features.

5.2.3.1 Identifying linear segments of the reproduction number

Given the two data streams (hospital admissions and deaths) we consider in the analysis, we may now decide how to analyse them. The most straightforward approach would be to infer the reproduction number from both time series directly but separately. This can be done by, first, convoluting the incubation period with the distribution for the time between symptom onset until either hospital admission or deaths, second, deconvolute the time series with the resulting distributions, and, third, inferring the reproduction number with EpiEstim [233]. The downside of this approach is that we cannot do a joint estimation, where we leverage both time series simultaneously. More specifically, the joint estimation improves robustness and reduces the influence of possible biases in one data stream on the overall results.

Such a joint estimation is possible if we consider a model that simultaneously describes the dynamics of hospital admissions and daily deaths. In this case, the reproduction number can be inferred by comparing the model output to the reported data, jointly leveraging both data streams. Therefore, one needs to design a model that describes the epidemic dynamics and decide how to parametrise the reproduction number. We will focus on the latter first. Our principal interest is to determine when the reproduction number started to decrease, i.e., whether there was a decrease that preceded the one in mobility. In this sense, we can tailor the parametrisation of the reproduction number to this aim. Intuitively, the reproduction number should follow three segments: An initial constant value as the virus mainly spread undetected in Spain, a rapid decrease after the initiation of the epidemic response, and finally, a relatively constant value during the lockdown or a slight de-

crease due to the intensification of the confinement. Let us look at the shape of the reproduction number that we inferred from the reported cases to test whether this is a reasonable hypothesis.

To analyse whether these three segments can describe the evolution of the reproduction number, we adjust precisely this hypothesised functional form with three segments to the reproduction number inferred from the reported cases. To this aim, we use the segmented package that implements a method proposed by Muggeo, which relies on a maximum likelihood estimation [503, 504]. We assume the reproduction to be piece-wise linear, with three segments and two breakpoints, where the first segment is constant. The yellow line in Fig. 5.7B indicates the results from the maximum likelihood estimation. Interestingly, the second breakpoint coincides with the implementation of the lockdown on March 15 (CI: 14 —16). Further, the first breakpoint corresponds to the start of the decrease in the reproduction number that we had already identified on March 5 (CI: 4 – 6). We can also see that during the lockdown, in the third segment, the reproduction number decreases very little. The limited reduction with our finding in sec. 5.1 that the reinforced confinement had little impact on the evolution of the epidemic. Overall, we can say that the three segments reasonably well describe the evolution of the reproduction number.

5.2.3.2 Model definition

With respect to the three segments we adjusted before to the estimated reproduction number, we will make two additional assumptions for the model parametrisation, which both reduce the number of parameters to estimate. First, since there was only a slight variation, we will assume that the third segment, representing the period during the lockdown, is constant. Second, we fixed the second breakpoint on March 15, the day the lockdown was introduced. We make this assumption because the second breakpoint coincides with the introduction of the lockdown in estimated segments. Furthermore, it also seems reasonable to assume that the reproduction number reached a relatively constant value as the lockdown was introduced since the mobility restrictions did not substantially change anymore from there onward. As part of the sensitivity analysis, we will relax this second assumption showing that conclusions drawn from the analysis are not altered.

Accordingly, we parametrise the reproduction number for the epidemic model with an initial constant value R_1, a breakpoint BP that initiates the

epidemic response, and a constant value R_2 during the lockdown. In this framework, the inferred value of BP will indicate whether the early decrease in the reproduction number preceded the one in mobility.

We will denote the parametrised form of the reproduction number for the model as Rt^{M}. We stay as simple as possible for the epidemic model while respecting the correct distributions for the waiting times between infection, symptom onset, hospital admission and death. More precisely, we chose a discrete-time, Non-Markovian model that assumes a well-mixed population without age stratification [229]. The model is similar to the approach used to infer the reproduction number by EpiEstim [225]. Given a population size N and the generation time distribution $w(t)$, the daily incidence at time t, I_t, evolves as

$$I_t = \mathrm{Rt}^{\mathrm{M}} \left(1 - \frac{1}{N} \sum_{\tau=1}^{t-1} I_\tau\right) \sum_{\tau=1}^{t} I_{t-\tau}\, w(\tau)\,. \tag{5.50}$$

The first term here corresponds to the increasing immunity in the population due to previous infections as the virus spreads. From there, one can convolute the daily incidence with the corresponding distributions to receive the time series for symptom onset S_t, hospital admissions H_t, and fatalities F_t. Given the incubation period, p_t, and time between symptom onset to hospital admission, h_t, and death, d_t, the remaining equations of the model read

$$S_t = \sum_{\tau=0}^{t-1} I_\tau p_{t-\tau} \tag{5.51}$$

$$H_t = \mathrm{IHR} \times \sum_{\tau=0}^{t-1} S_\tau h_{t-\tau} \tag{5.52}$$

$$F_t = \mathrm{IFR} \times \sum_{\tau=0}^{t-1} S_\tau d_{t-\tau}\,. \tag{5.53}$$

IHR and IFR refer to the infection-hospitalisation ratio and infection-fatality ratio, respectively. As previously mentioned, we adjusted the model to the hospital admissions and fatalities. Besides R1, R2 and BP, there is an additional parameter for the initial number of infected individuals, I0, as the model dynamics start on February 10. The details of the model fitting are presented in the following section.

5.2.3.3 Details of the model inference

The comparison between the model and the data is done in a Bayesian approach, where the likelihood is described by a negative binomial distribution. Regarding the parameters, we fix the same incubation period as for the inference of the exposure times [496]. Similarly, we choose the generation time as for the estimation of the reproduction number [232]. We assume that the time between symptom onset and hospital admission/death follows a gamma distribution. Since case line data is not available, we fix the shape factor of these distributions as 2.5 and 2.2, respectively, according to Hawryluk et al. [505]. The scale parameter is then fixed according to the median time the health authorities reported from symptom onset to hospital admission (6 days) and death (11 days) [479], which resulted in scale parameters of 2.68 and 5.85, respectively. However, the first adjustments showed that the time series of daily deaths was not accurately adjusted by using the distribution mentioned above. Accordingly, we decided to add the shape and scale parameter of the distribution for the time between symptom onset and death to the free parameters. The IHR and IFR are calculated from the nationwide seroprevalence study [484]. We divide the total number of infected individuals, estimated in this study, by the accumulated number of hospital admissions and deaths, giving an IHR and IFR of 4.54% and 1.18%, respectively.

Motivated by the estimation of the reproduction number in sec. 5.1, we fix the prior of R1 as a normal distribution with mean 4.5. We set the standard deviation as 1.0. The prior of R2 is a uniform distribution between 0.4 and 1.0. Similarly, the prior of BP is flat between 15 and 0 days before lockdown. The prior for the initial number of infected individuals I0 is uniform between 1 and 5000. We set the prior for the dispersion parameter for both time series as a normal distribution with mean 10.0 and standard deviation 5.0. For the time distribution between symptom onset and death, we choose a Gaussian prior with the mean corresponding to our initial estimation and a standard deviation of 0.1. The form of the inferred distribution from symptom onset to death is shown in Fig. C.5 in Appendix C.

We run 6 chains with 4000 iterations, where 2000 iterations are used for warm-up. Gelman-Rubin convergence statistics, i.e. potential scale reduction factors were all smaller than 1.001 [255], and trace plots indicated all convergence. The inference is performed in Stan [250, 251]. Posteriors are shown in Fig. C.6 in Appendix C.

5.2.3.4 Estimated form of the reproduction number

Figs. 5.11A & B show the model's fit to the daily hospital admission and deaths, respectively. Looking at the inferred parameters, we find R1 = 3.27 (CrI: 3.01 – 3.61) and R2 = 0.66 (CrI: 0.64 – 0.67). Both R1 and R2 are consistent with the estimation from the reported cases, Rt. With the assumed generation time, this results in a doubling time of 2.55 days (CrI: 2.32 – 2.78) during the free phase before lockdown and a half-life time of 9.11 (CrI: 8.72 – 9.56) during lockdown [506]. The relatively short doubling time indicates how difficult it was to realistically evaluate the epidemiological situation during the first wave, given the substantial reporting delay we uncovered previously. Nevertheless, the low value of R2 and the corresponding half-life time suggest how efficient the lockdown was in reducing the incidence. Further, we find an initial number of infected individuals of 1140 (CrI: 773 – 1599) at the start of the model dynamics on February 10. The high number of initially infected individuals further highlights how the virus was able to silently spread through the population in early 2020 in Spain since the first case of local transmission was only reported on February 28 [476]. The early underdetection is also consistent with an excess in influenza-attributed deaths in February 2020 in Catalonia [507].

Let us now turn our attention to the initiation of the epidemic response, i.e. the day the reproduction number started to decrease. The results suggest that the first breakpoint is March 10 (CrI: 8 — 12). A decrease initiated on March 10 contrasts with the one found from the reported cases on March 5/6. Further, a decrease that started on March 10 is consistent with the initial decrease in mobility presented in Fig. 5.5A. Additionally, in Fig. 5.11D, the mobility reduction is calculated as a weighted average over all provinces with respect to the reported case numbers instead of the population size. The weighted average is directly inspired by the NGM approach [214] and the model presented in sec. 5.1. We performed the same weighted average with other epidemiological indicators such as hospital admissions, ICU admissions and deaths, but the temporal evolution did not substantially change. Finally, we note that there is an additional, even though less rapid, decrease in mobility after the introduction of the lockdown. It is difficult to pinpoint how this may have contributed to an additional decrease in the reproduction number. Nevertheless, our results from the sensitivity analysis (Fig. C.7 in Appendix C) indicate that the reproduction number reached a stable value as the lockdown was introduced, or, at least, shortly after that.

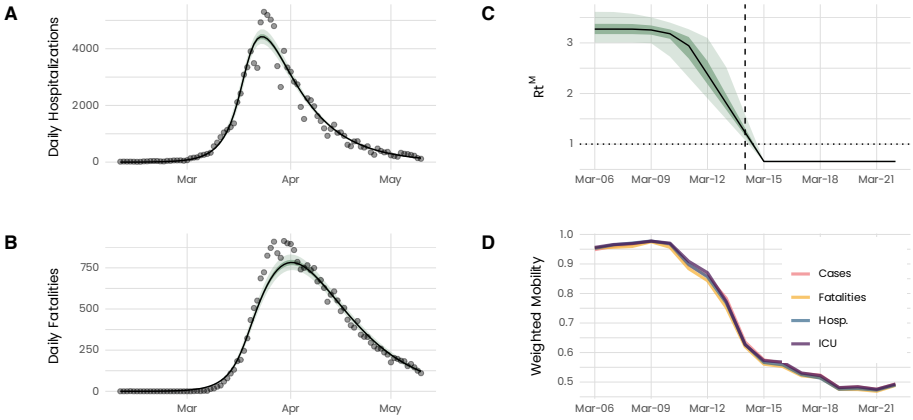


Figure 5.11: **(A)** & **(B)**: Model adjustment to the daily hospitalisations and fatalities, respectively. The green and light-green shaded areas represent 50% and 95% credible intervals. The solid line represents the median. **(C)**: Inferred evolution of Rt^M. The dashed vertical line indicates the implementation of the lockdown. **(D)**: Aggregated mobility level of the Spanish provinces before the lockdown and shortly afterwards. Instead of averaging by population size, we average by different daily epidemiological indicators to visualize the impact of the mobility reduction on the epidemic evolution. This is in analogy with the definition of Rt [218]. We observe that mobility and Rt^M started to decrease around the same time.

Overall, these results, in particular the estimated value of BP, support the hypothesis that the early decrease on March 5/6 in the reproduction number, which was inferred from the reported case data, is a result of saturation in the test capacity. Further, the possible saturation is reflected in a decrease in the detection rate towards the introduction of the lockdown, shown in Fig. 5.12A. We define here the detection rate as the ratio between the reconstructed infections from the reported cases (exposure times) and the infections estimated from our model. We estimate that the detection rate was initially only around 5%, increased then to almost 10% before experiencing a sudden drop towards March 15.

5.2.4 *Counterfactual scenarios*

A broad spectrum of opinions exists regarding what would have been the optimal epidemic response. In this context, one of the relevant questions is what would have occurred had we acted earlier or later. A series of studies addressed this by evaluating the impact of earlier, or later introduction of the

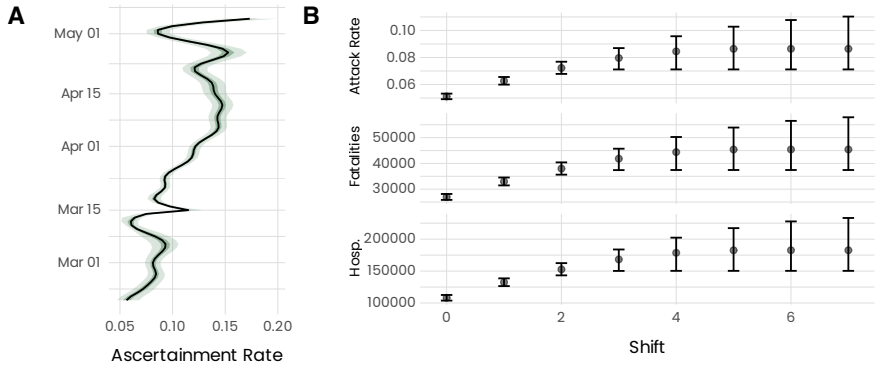


Figure 5.12: **(A)**: Evolution of ascertainment in time. Ascertainment is defined as the ratio between the reported cases and the incidence from the adjusted model. We compare the date of infection (exposure time) instead of the reporting date. We note a sudden decrease from March 5/6 towards lockdown. Later, testing capacity was massively increased, which increased ascertainment (Fig. S19). **(B)**: Counterfactual scenarios for the attack rate, the death toll and the total number of hospitalisations. The counterfactual scenarios consist of shifting BP by x days. No further shift is applied if the BP exceeds March 15, i.e. lockdown. We note how the response previous to the implementation of the lockdown substantially contributed to limiting the impact of the epidemic.

lockdown [468, 501, 508]. In these studies, the counterfactual scenario mainly consisted of a shift of the entire epidemic response by a given amount of days. However, suppose the reproduction number is initially constant as in our approach. In that case, the impact of a shift in the epidemic response is entirely determined by R1, i.e. the initial doubling time. More precisely, neglecting the depletion of susceptible individuals, if one shifts the entire response by the duration of the doubling time, then all the epidemiological indicators — attack rate, hospital admission, and fatalities — will double.

Since our analysis focuses on the epidemic response that preceded the lockdown, we will consider a counterfactual scenario that aims to quantify the latter's impact on the epidemic. As a matter of fact, given the parametrisation of Rt^{M}, this counterfactual scenario is straightforward to implement. Instead of shifting the entire epidemic response, we will only shift the initiation of the epidemic response, i.e. BP. This scenario allows us to estimate the possible outcome of the first wave if the reduction in the reproduction number started at a later time. In the case that the shift of BP exceeds the introduction of the lockdown, BP is kept fixed on March 15. In other words, our scenario

assumes that the lockdown is introduced with identical intensity on March 15, i.e. identical R2. Similarly, the virus spreads with the same velocity, i.e. same R1, through the population previous to the epidemic response. The only difference in the counterfactual scenario is that the epidemic response is initiated later (shift of BP).

The results, shown in Fig. 5.12B, estimate that, in the absence of an epidemic response previous to the lockdown (maximal shift), 8.6% (CrI: 7.1 – 11.0) of the population would have been infected instead of 5% according to the seroprevalence study [484]. The increased attack rate results in 45'400 (CrI: 37'400 – 58'000) fatalities and 182'600 (CrI: 150'400 – 233'800) hospital admissions compared to 27'800 and 107'600 reported [487]. Overall, our results indicate that the pandemic response before the lockdown substantially contributed to limiting the impact of SARS-CoV-2 during the first wave in Spain.

5.2.5 *Limitations of the analysis*

The analysis presented has several limitations. First, we considered the most simple epidemic model possible. We mainly focused on respecting the correct timings, i.e. to consider the accurate distributions for the generation time, the incubation period or the time between symptom onset and hospital admission/death. For this reason, we chose a Non-Markovian model, where we can directly introduce the empirical waiting times and are not restricted to exponential ones. It was crucial to accurately reflect the waiting times in the modelling approach since, with the epidemic model, we primarily wanted to determine the start of the epidemic response. One could extend the work by stratifying the population according to age or geography instead of assuming a well-mixed population. However, this approach would require inferring the accurate distribution for the stratified waiting times. They vary substantially with respect to age and region [505]. In our case, we only have access to the national median regarding the time between symptom onset and hospital admission/death.

Another limitation is that our analysis is based on correlation. Therefore, we cannot make a causal connection between the decrease in mobility and the decrease in the reproduction number, even though it seems intuitive. Similarly, we cannot pinpoint the exact reasons that caused the decrease in mobility before the lockdown. We theorise that this early decrease was a con-

sequence of individual awareness which was fuelled by news from Italy, as well as other countries, and the regional NPIs that were introduced from March 10 onward. The correlation we found between the mobility level previous to the lockdown and the number of reported cases in the different provinces further indicates that individual awareness drove the mobility reduction at least partially. However, we cannot exclude that other factors may be responsible for this correlation.

5.2.6 *Implications of the findings*

Of our findings, the substantial delay between infection and reporting date has the most direct consequences. As pointed out in sec. 5.1, with such delay, one can see increasing case numbers for almost two weeks after the reproduction number has crossed 1. While the reproduction number was below 1 from March 15 onward, when the lockdown was introduced, case numbers were growing almost until the end of April. Accordingly, this can partially explain why the authorities decided on March 28 to halt all non-essential economic activity from March 30 onward. However, our results here and in section sec. 5.1 indicate that the reinforcement of the lockdown did not lead to a substantial further reduction in the reproduction number.

More generally, our results question a mechanistic understanding of NPIs. We found evidence that the reduction in the reproduction number before the lockdown substantially contributed to reducing COVID-19-related hospital admissions and deaths. While regional NPIs certainly played a part, our results indicate that individual awareness contributed to the initial decrease. Please note that our results do not imply that the lockdown was ineffective at containing the epidemic. Our estimation of the haft-life time during lockdown has clearly shown how quickly case numbers were reduced. Instead, the results indicate an epidemic response previous to the national lockdown and beyond the containment policies. In our model in sec. 5.1, we only considered the mobility reduction starting from the announcement of the lockdown on March 14. In other words, we attributed the entire mobility reduction to the national lockdown. However, retrospectively, we have evidence that this was not the case. Mobility and the reproduction number had already decreased before the lockdown.

Nevertheless, despite this discrepancy, we could accurately fit the model to the data in sec. 5.1, at least according to a visual inspection. Model selection

is a general issue for fitting epidemiological data. Often, many model choices can reasonably well reproduce the data. There are many free parameters to fit relatively simple time series. Hence, especially without explicit model selection, an accurate adjustment does not necessarily mean that one also considers the true underlying mechanisms. For example, many approaches that model the impact of NPIs assume abrupt changes in the transmission probability, i.e. the reproduction number, every time an NPI is introduced [314, 316, 490, 509]. Accordingly, the model implicitly assumes that only the introduction of NPIs can change the course of the epidemic. Also, these approaches generally reproduce the data perfectly well. However, we have shown an example here where not the entire reduction in mobility and the reproduction number can be attributed to the mandated reduction during the lockdown. Even though such an approach could perfectly match the data, as shown in Ref. [314]. Accordingly, these approaches tend to overestimate the impact of NPIs without leaving room for the complex interplay of individual awareness, fear of infection and measures introduced by health authorities. These factors were not only observed in the context of SARS-CoV-2 [300, 489] but also for the 2009 H1N1 swine flu pandemic [510] or the 2003 SARS pandemic [511], to name a few. In the next section, we will address this issue by providing a modelling and inference framework that allows for a retrospective analysis without such strong assumptions regarding the impact of NPIs. The proposed approach will prove helpful in understanding the course of the epidemic in summer 2021 in Catalonia as a vaccination campaign, a new variant, and the easing of restrictions all coincided together.

5.3 THE DELTA VARIANT IN CATALONIA DURING SUMMER 2021

To this point, our modelling efforts focused on the first wave of SARS-CoV-2 during early 2020. In this period, almost the entire globe experienced important outbreaks. However, also during the rest of 2020, the virus continued to spread all around the globe. In particular, the colder weather in autumn sparked a substantial acceleration of the epidemic in the northern hemisphere [295, 296]. Until December 31 2020, there were over 80 million confirmed cases worldwide [294]. This massive amount of infections accelerated the evolution of the virus and hence increased the probability of the emergence of new variants [8]. The first more transmissible variant emerged was the Alpha variant (B.1.1.7) [466]. The first cases of the Alpha variant emerged probably in

September in England. The first detection occurred in November 2020, and the variant was declared a variant of concern (VOC) a month later [512]. Due to the increased transmissibility, the Alpha variant rapidly replaced the original strain, first in the United Kingdom and in early 2021 in continental Europe.

As illustrated in Fig. 5.13, the Alpha variant led to new outbreaks in a series of countries. The timing of the emergence of the Alpha variant was delicate since the vaccination campaign was within reach before this more transmissible and more virulent variant destroyed the probably too optimistic hopes of a summer 2021 without COVID-19. Additionally, the Alpha variant also added uncertainty on whether vaccine efficacy would still be as elevated as against the original strain. Luckily, the first studies quickly showed that the vaccines still gave substantial protection, particularly against a severe course of infection [513]. However, the vaccination campaign that, in most countries, started around February 2021 arrived slightly too late to actually contain the spread of the Alpha variant. The early doses administered to the elderly may

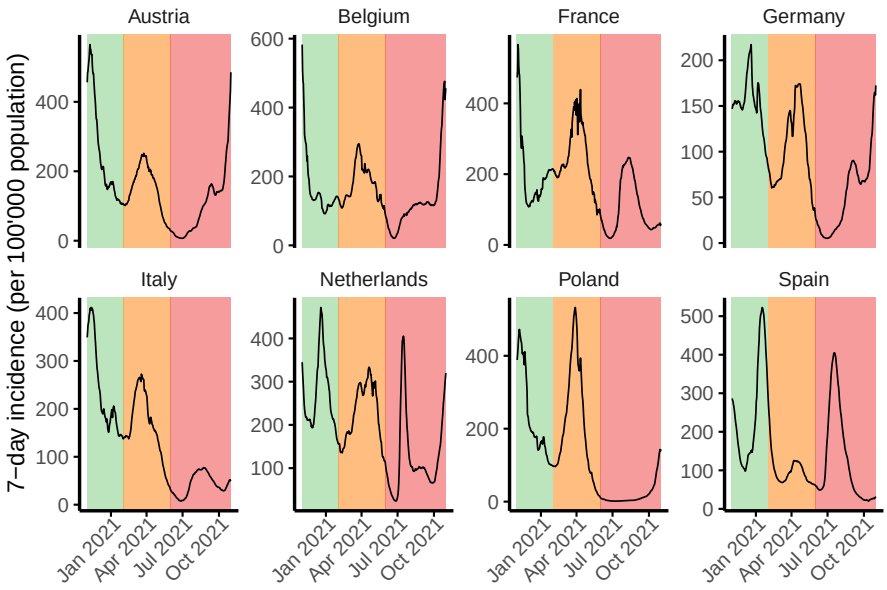


Figure 5.13: Evolution of the reported cases for a series of European countries [294]. Colours indicate periods where different virus variants were dominant. Green indicates the original strain, orange is the Alpha variant, and red is the Delta variant.

have helped reduce the number of deaths, but the vaccine coverage was too low to substantially slow down the spread of the virus. Nevertheless, most European countries managed to contain the spread of the Alpha variant relatively well through continued mask use, social distancing and capacity limitations in restaurants [294]. Unfortunately, as the case numbers were decreasing in spring 2021, the Delta variant, a new, more transmissible and more virulent variant, was already looming to spread around Europe.

The first cases of the Delta variant were detected in India in late 2020 [514]. The increased transmissibility and virulence of the Delta variant sparked a massive outbreak of SARS-CoV-2 in India during the spring of 2021. The number of daily reported deaths amounted to over 3000, and the actual one was probably a multiple of that [294]. In light of the extraordinary disease burden in India and the information gathered by scientists, the WHO declared the Delta variant as a VOC in May 2021 [515]. Due to the strongly reduced air traffic and general mobility between Europe and the rest of the world, it took the Delta variant much longer to arrive in continental Europe than the Alpha variant [194]. However, as the first cases started to appear in April 2021, data indicated that the Delta variant would quickly outcompete the Alpha variant [516]. The Delta variant then led to significant outbreaks across many countries during the summer of 2021, as we illustrate in Fig. 5.13. The spread of the Delta variant coincided with the vaccine rollout. The distribution of the vaccine coverage in the population was very heterogeneous due to the prioritisation of the older age groups. Furthermore, the spread of the Delta variant also coincided with the easing of restrictions. Many countries removed NPIs after the decrease in infections in the spring of 2021. This mixture what it particularly difficult to predict the variant's impact.

What we described above also holds for the Spanish CA Catalonia. Case numbers and hospital admissions rapidly increased in Catalonia between July and August 2021 (Figs. 5.14A & B) when the vaccination level was still relatively low (Fig. 5.14C). This rapid increase in reported cases took place after the end of the national state of alarm on May 9, and the opening of nightclubs on June 21 [517]. Further, on June 23, the regional festivities of "Sant Joan" took place. The bonfires of Sant Joan are a tradition in Catalan-speaking regions like Catalonia, Valencia, or Mallorca [518]. These festivities include musical performances, fireworks, and large public gatherings. The main festivities occur during the evening of June 23; the following day, June 24, is a public holiday. Since June 24 was a Thursday in 2021, many individuals took

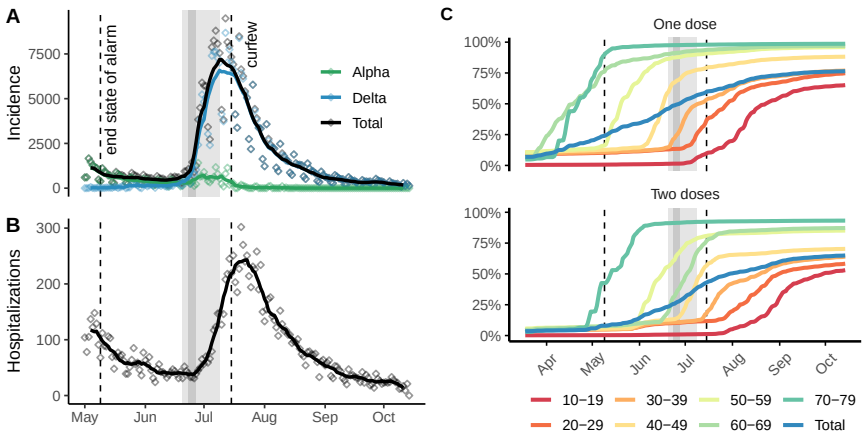


Figure 5.14: **(A)**: Daily incidence from May 1st onward. The solid lines indicate a centred 7-day rolling average. Additionally, the daily incidence is separated into cases of the Alpha and Delta variants. The dashed lines indicate the end of the state of alarm and the implementation of a curfew from 1am to 6am. The grey shaded area indicates the period from June 21 until July 9 as nightclubs were opened. The dark grey area indicates the long weekend of Sant Joan. **(B)**: Daily hospitalisation from May 1st onward. **(C)**: Evolution of the fraction of vaccinated individuals in the various age groups. As nightclubs opened (grey-shaded area) only the population older than 50 had substantial protection from vaccination. Administration of the first vaccine dose in the age group 30-39 was initiated almost simultaneously as nightclubs were opened. The age group 0-9 was omitted since no vaccines were administered in this age group.

advantage of a long weekend [1]. Eventually, the rapid rise in reported infections and hospital admission caused the Catalan health authorities to close the night clubs on July 9 and further intensify the measures by introducing a curfew from 1 am until 6 am on July 15. The curfew was lifted on August 19 by the court since the judges considered that the Catalan government does not have the legal authority to introduce a curfew. The authorities opened night clubs on October 8, requiring the vaccination certificate to enter [517].

The events described above motivated us to employ an epidemiological model to analyse the driving factors that led to the infection wave in Catalonia during the summer of 2021. The model takes into account the population's time-dependent vaccination status, the presence of two variants — Alpha and Delta variant — and the varying interaction rates of the popula-

1 Historically, Saint John's Eve, starting at sunset on June 23, is the eve of celebration before the Feast Day of Saint John the Baptist (Sant Joan in Catalan). It closely coincides with the June solstice, also referred to as Midsummer in the Northern Hemisphere [519].

tion. As we highlighted in the conclusion of the previous section, behavioural change, i.e. a change in the interaction rates, occurs not only in light of NPIs but also voluntarily. For this reason, we will outline a model framework here that does not make any strong assumptions on how NPIs impact the spread of the virus and allows the contact rates to vary even if no NPIs are introduced or lifted. Our results will show that the introduced NPIs had a limited direct effect, and contact rates had substantially decreased before their introduction.

5.3.1 *Model description*

We use a continuous-time, age-stratified compartmental model whose evolution can be described through a set of ordinary differential equations. The model includes the following compartments: susceptible, latent infection, infectious, admitted to the hospital, admitted to the ICU, deceased and recovered. The compartments are stratified into different age groups from 0 to 79 years by steps of 10 years. We excluded the age group 80+. Including this age group would have required a tailored approach for the dynamics in care homes for which we do not dispose of the necessary data [72, 468]. Further, the age group 80+ acts rather as a sink than a driver of the epidemic. Hence, their exclusion should not substantially impact the dynamics. The relative differences between age groups regarding the risk of hospital admission, ICU admission or decease in ICU are fixed according to Ref. [468], which are estimates for the original strain. We further fix the absolute value for the probability of hospital admission [468]. In contrast, the absolute value, which acts as a scaling factor while maintaining the relative differences with respect to age, was left as a free parameter in the inference process for the ICU admission and probability of death.

Besides stratification according to age, the model separates between vaccination status (unvaccinated, 1 dose, 2 doses) and infection by virus variant (Alpha, Delta). Individuals do not transition only between epidemic compartments but also between vaccine compartments. The data with respect to the daily number of doses administered directly informs these transitions. The efficacy of the vaccines against the Alpha and Delta variants regarding infectiousness, susceptibility, and probability of hospital/ICU admission are fixed according to Refs. [420, 513, 520–522]. We set the increase in severity (risk of hospital admission) of the Alpha variant with respect to the normal strain as

1.42 [523]. The further increase in severity of the Delta variant in comparison to the Alpha variant is fixed as 1.85 [524]. The transmission advantage of the Delta variant with respect to the Delta variant is left as a free parameter.

5.3.1.1 Modelling time-varying contact rates

We highlighted before that it would be desirable to have a model framework which does not require strong assumptions on how NPIs impact the spread of the virus. Accordingly, instead of assuming change points in the contact rates at the introduction of NPIs [314, 316], we would favour inferring a more continuous evolution of the contact rates that does not discard the possibility of voluntary behavioural changes beyond NPIs [300]. In this sense, such an approach would be closer to the methodology we employed to infer the reproduction number with EpiEstim [220, 225]. While EpiEstim estimates a value of the reproduction number for each day, we will infer the contact rate $\beta^i(t)$ for each day t and every age stratum g. In other words, the daily contact rates of the different age strata are free parameters in the inference framework. Effectively, $\beta^i(t)$ corresponds to a transmission rate but assuming equal infectivity and per-contact susceptibility across age strata, $\beta^i(t)$ is directly proportional to the contact/interaction rate. For this reason, we refer to $\beta^i(t)$ as the contact/interaction rate. Now, while the contact rates vary daily, to control the complexity of the model, we will assume that the relative mixing between age strata, controlled through the contact matrix C [72], stays constant. Neglecting that we distinguish between variants and vaccination status, the number of susceptible individuals in age group $S_i(t)$ at time t evolves as

$$\dot{S}_i(t) = -\beta^i(t) S_i(t) \sum_{i=1}^{M} C_{ij} \frac{I_j(t)}{N_j} ,$$

where $I_j(t)$ referes to the number of infected individuals in age strata j at time t. M represents the number of age strata and N_i the number of individuals in age strata I. The complete model equations can be found in Appendix D.

From the contact rates, the reproduction number can be calculated in analogy to what we outlined in sec. 5.1. To do so, one needs to consider the presence of different variants and the vaccination status of individuals. The detailed expression is presented in Appendix D. Accordingly, the approach we propose here acts very similar to EpiEstim, but with a more complex compartmental structure tailored to the epidemiological situation. To prevent overfitting and guarantee smoothness, EpiEstim employs a rolling average.

Here, in contrast, we consider a Gaussian smoothing prior. In the next section, which outlines the inference framework, we will detail how we included this Gaussian smoothing prior to the model.

5.3.2 *Inference framework*

We adjust the model to five data streams: the daily incidence, $\mathcal{I}_d^i(t)$, the daily number of hospital admissions, $\mathcal{H}_d^i(t)$, the daily ICU admissions, $\mathcal{ICU}_d^i(t)$, and the daily deaths, $\mathcal{D}_d^i(t)$, in each age group, and the sequencing data. The subscript d denotes that the variable refers to the data. We will use the subscript m to refer to quantities generated from the model. The sequencing data consists of the daily number of Delta sequences detected, $s_d^\delta(t)$ and the total number of sequences analyzed, $s_d^{\text{tot}}(t)$. We have sequencing data available from May 1 until October 15. The sequencing data was received from the Sistema de Notificació Microbiològica de Catalunya (SNMC). On average, 50 sequences were analyzed every day. We discarded the detected variants that were neither the Alpha nor the Delta variant, as they only represented 6% of the cases. Individuals get detected as they exit the infectious compartment. We fix the detection rate constant over the period considered. We do not stratify ICU admission and fatalities by age due to the low number of cases.

We assume that the first four data streams originate from the same negative binomial process. The dispersion parameter ϕ is left as an auxiliary variable of the inference. The contribution of the sequencing data is modelled through a binomial observation process. Accordingly, the contribution to the

log-likelihood $\mathcal{L}$ of the comparison between the model and the data $\mathcal{L}_{d;m}$ is given by:

$$\begin{aligned}\mathcal{L}_{d;m} = &\sum_{t=1}^{t_{\text{end}}}\sum_{i=1}^{M}\log \text{NB}\left(\mathcal{I}_d^i(t);\ \mathcal{I}_m^i(t),\phi\right)\\ &+\sum_{t=1}^{t_{\text{end}}}\sum_{i=1}^{M}\log \text{NB}\left(\mathcal{H}_d^i(t);\ \mathcal{H}_m^i(t),\phi\right)\\ &+\sum_{t=1}^{t_{\text{end}}}\log \text{NB}\left(\sum_{i=1}^{M}\mathcal{ICU}_d^i(t);\ \sum_{i=1}^{M}\mathcal{ICU}_m^i(t),\phi\right)\\ &+\sum_{t=1}^{t_{\text{end}}}\log \text{NB}\left(\sum_{i=1}^{M}\mathcal{D}_d^i(t);\ \sum_{i=1}^{M}\mathcal{D}_m^i(t),\phi\right)\\ &+\sum_{t=1}^{t_{\text{end}}}\log \text{B}\left(s_d^{\delta}(t);\ s_d^{\text{tot}}(t);\ \frac{\sum_{i=1}^{M}\mathcal{I}_m^{\delta,i}(t)}{\sum_{i=1}^{M}\mathcal{I}_m^{i}(t)}\right).\end{aligned} \tag{5.54}$$

As previously mentioned, the interaction rates $\beta^i(t)$ are also inferred. To prevent overfitting, we add a Gaussian smoothing prior. The data indicates that the reproduction number rapidly rose as nightclubs opened, coinciding with the "Sant Joan" festivities. Outside this period, the reproduction number seems to have evolved very smoothly. We, therefore, consider two separate standard deviations for the Gaussian smoothing prior as the night clubs were open, σ_o, and closed, σ_c, respectively. Both standard deviations are fixed as the average change in the reproduction number measured from the reported cases through EpiEstim in the respective periods. Values are fixed to $\sigma_o = 0.12$ and $\sigma_c = 0.037$, respectively. Defining as t_o and t_c the days at which night clubs opened and closed, respectively, the contribution of the smoothing prior to the log-likelihood $\mathcal{L}_{\text{smooth}}$ is expressed as

$$\begin{aligned}\mathcal{L}_{\text{smooth}} = &\sum_{t=2}^{t_o-1}\sum_{i=1}^{M}\log\mathcal{N}\left(\beta^i(t)-\beta^i(t-1);\ 0,\sigma_c\right)\\ &+\sum_{t_o}^{t_c-1}\sum_{i=1}^{M}\log\mathcal{N}\left(\beta^i(t)-\beta^i(t-1);\ 0,\sigma_o\right)\\ &+\sum_{t_c}^{t_{\text{end}}}\sum_{i=1}^{M}\log\mathcal{N}\left(\beta^i(t)-\beta^i(t-1);\ 0,\sigma_c\right).\end{aligned} \tag{5.55}$$

We fix the priors of the auxiliary variable ϕ as $\phi^{-1} \sim \text{Exp}(5.0)$. The prior of the detection rate for age group i is fixed with a Cauchy distribution, i.e. $d_r^i \sim \mathcal{C}(0.5, 1.0)$. The rest of the priors are presented in Table D.5 in Appendix D. In

the subsequent section (sec. 5.3.2.1), we will discuss two additional priors for the initial number of infected individuals and the interaction rates at $t = 1$. Subsuming the contributions of the priors to the log-likelihood into $\mathcal{L}_{\text{priors}}$, the overall log-likelihood is then given by

$$\mathcal{L} = \mathcal{L}_{d;m} + \mathcal{L}_{\text{smooth}} + \mathcal{L}_{\text{priors}} \,. \tag{5.56}$$

5.3.2.1 Implementation

We implement the dynamics, as well as the inference framework, in Stan [249, 251]. We solve the differential equations with the Euler method. Each day consists of five timesteps. To initialize the model, we run the dynamics from April 1 onward, while the model-to-data comparison starts on May 1. In this initial period, the interaction rates do not vary and are kept at $\beta^i(t = 0)$. The number of initially infected individuals in each age group, $I^i_{0,m}$, follows the distribution of reported cases. To this aim, we calculate the average number of reported cases in the seven days after April 1 in each age group, $I^i_{0,d}$. Correcting by the detection rate, $I^i_{0,m}$ is expressed as

$$I^i_{0,m} = I^{\text{tot}}_{0,m} \frac{\frac{I^i_{0,d}}{d^i_r}}{\sum_{j=1}^{M} \frac{I^i_{0,d}}{d^j_r}} \,. \tag{5.57}$$

The variable $I^{\text{tot}}_{0,m}$ is a free parameter of the model and represents the total number of initially infected individuals with the Alpha variant in the model. We fix the prior of $I^{\text{tot}}_{0,m}$ as:

$$I^{\text{tot}}_{0,m} \sim \mathcal{N}\left((\tau_E + \tau_I) \sum_{j=1}^{M} \frac{I^j_{0,d}}{d^j_r}, f \times (\tau_E + \tau_I) \sum_{j=1}^{M} \frac{I^j_{0,d}}{d^j_r} \right) . \tag{5.58}$$

Initially, infected individuals are distributed proportionally to the waiting times in the latent ($\tau_E/(\tau_E + \tau_I)$) and infectious compartment ($\tau_I/(\tau_E + \tau_I)$) of the Alpha variant. In this way, at the mean of the prior distribution, the number of detected cases in the model matches the reported cases. Furthermore, the number of individuals leaving equals the one entering the infectious compartment. We fixed $f = 0.2$. Since case numbers were decreasing in the month previous to May 1, we fixed the prior of $\beta^1(t = 0)$ as $\beta^1(t = 0) \sim \mathcal{N}(0.3, 0.2)$. For the Delta variant, we assume the same relative

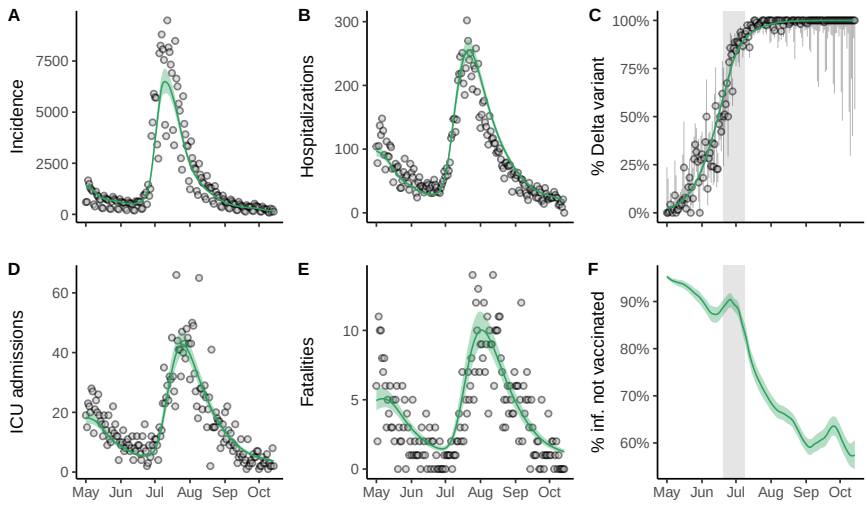


Figure 5.15: Adjustment to the data for daily incidence (A), hospitalisations (B), ICU admissions (D) and fatalities (E). Panel **C** compares the sequencing data with the model. It shows the fraction of infections that stem from the Delta variant. Vertical bars indicate the 95% credible interval assuming a uniform prior. **F**. Fraction of infected individuals that were not vaccinated. The still very low vaccination coverage as the nightlife was opened (grey shaded area) led to a high proportion of infected individuals that were not vaccinated in that period.

distribution over age groups and adjust the total number of infected individuals through the parameter $I_{0,m}^{\delta,\text{tot}} \sim \mathcal{N}(100, 80)$.

We ran four chains for a total of 2000 steps, of which 1000 were used for warm-up. Trace plots for time-independent variables were inspected and showed convergence. The Gelman-Rubin convergence statistics, i.e., potential scale reduction factors, were all smaller than 1.01 [255].

5.3.3 *Adjustment to the data*

Fig. 5.15 shows the good agreement of the model with respect to the reported cases (A), hospitalisations (B), ICU admissions (D) and fatalities (E). Age-stratified comparisons for the reported cases and hospitalisations are shown in Appendix D. The Delta variant's transmission advantage is 1.52 (CrI: 1.50-1.54). The estimation is in line with other estimations [420, 525, 526]. How the transmission advantage depends on the choice of the generation time is shown in Appendix D. As the model dynamics start, on April 1, we find 53

(CrI: 37-75) cases that were infected with the Delta variant. In contrast, no case of the Delta variant was detected until April 2022. The discrepancy between the reported infections from the Delta variant and our estimation highlights how limited sequencing of, on average, 50 and maximally 200 samples per day makes the early detection and prevention of introducing a new VOC unfeasible. The comparison of the model to the sequencing data is shown in Fig. 5.15C. The data and the model highlight that the Delta variant was already on the verge of becoming the dominant variant as the nightclubs were opened in Catalonia (grey-shaded area). Further, the emergence of the variant coincided with a relatively small protection provided by the vaccines at that point. As nightclubs were opened, we estimate that around 90% of the infected individuals were non-vaccinated (Fig. 5.15D).

5.3.4 *Inferred contact rates*

Having shown that our model fits the data reasonably well, let us now consider the inferred evolution of the contact rates. Overall, we find rapidly increasing interaction rates among the younger age strata 10-19, 20-29 and 30-39 (Fig. 5.16A). More precisely, there is a constant increase in contacts from the beginning of June onward in each of these three age groups. Interestingly, the contact rates peak on the long weekend of Sant Joan (dark-grey shaded area). In particular, the age group 20-29 presents a pronounced increase in social activity towards this weekend. However, after Sant Joan, the interaction rates rapidly decrease, even before the closure of nightclubs and the introduction of the curfew, i.e. before any NPIs. Similarly to the age strata 40-49 and 50-59, the remaining age strata – 0-9, 60-69 and 70-79 (Appendix D) – do not exhibit any substantial increase in the number of contacts around Sant Joan. In contrast, from the middle of July onward, we observe an increase in social activity across almost all age strata.

The later increase in the contact rate did not lead to a substantial rise in the age-specific reproduction numbers (Fig. 5.16B). It appears that the administration of vaccine doses, in particular in the younger age groups, effectively reduced the impact of increasing contact rates. Accordingly, the population could normalise social interactions due to the protection provided by the vaccines. However, the impact of waning immunity, which the model does not include [527, 528], introduces some uncertainty. At least for the older age groups that received their second dose between April and May, the estim-

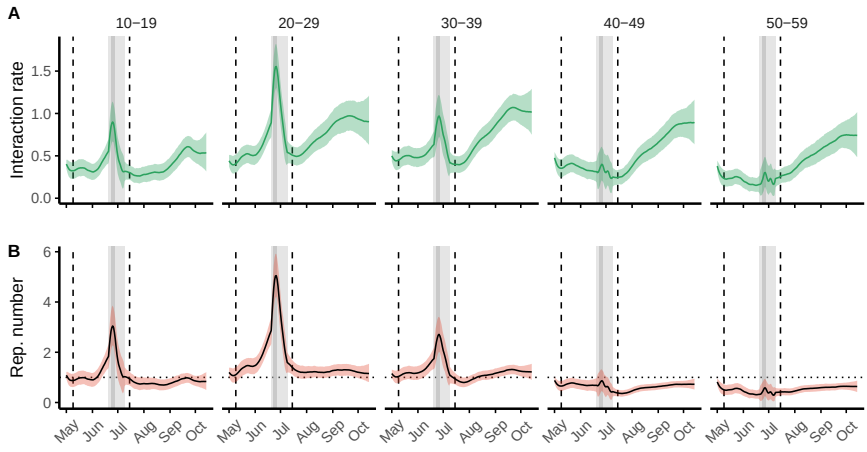


Figure 5.16: Inferred interaction rate (A) and reproduction number (B) for the age groups 10-19, 20-29, 30-39, 40-49 and 50-59. We omitted the remaining age groups to improve visibility. For completeness, they are shown in Figs. S4-S5 in the SI. The peak in the interaction rate and the reproduction number coincide with Sant Joan's festivities (dark-grey shaded area). The dotted line indicates a reproduction number of 1.

ated increase in the contact rates in the model could be an artefact waning vaccine-induced immunity.

As the earlier peak in the contact rates was attained around Sant Joan, we also observe a pronounced peak in the reproduction numbers. For example, the age group 20-29 reaches a peak reproduction number of 5.0 (CrI: 4.2-5.9). The concurrent increase of the contact rates and the reproduction number during this period is a direct consequence of the small number of vaccine doses that had been administered in the younger age strata until that point in time (Fig. 5.14B).

We have highlighted in sec. 5.2 that mobility was generally a good indicator for the reduction in contacts along which the disease can be transmitted during the SARS-CoV-2 pandemic [477, 489, 493]. However, as time progressed, the correlation between the epidemic evolution became less evident [477]. The use of face masks and the increased understanding of applying physical distancing made the relationship less evident. Beyond the mere reduction in contacts, other factors now influenced the evolution of the epidemic. The same holds for the evolution of mobility in Catalonia, shown in Fig. 5.17. Even though we can observe an increase in outdoor mobility (parks) towards Sant Joan, the present decrease in the other mobility categories impedes any

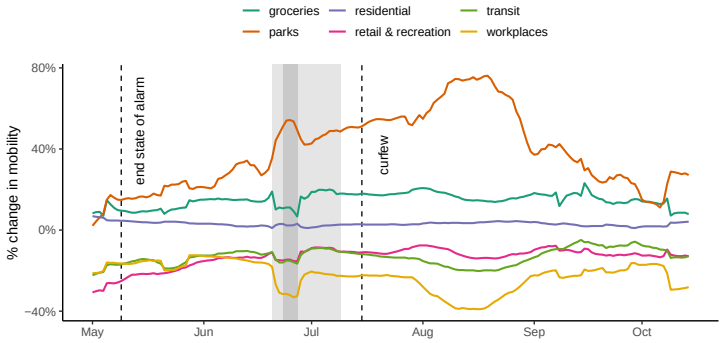


Figure 5.17: Evolution of mobility in Catalonia during the considered period according to the Google mobility report [529]. Google directly provides separated typrd of mobility.

conclusion on whether Sant Joan accelerated the spread of the virus or not. Similarly, the increase in outdoor mobility during the summer holidays is concurrent with a decrease in work-related mobility. In this sense, aggregated mobility data can only give little information on the epidemic evolution during this period. More detailed information would be necessary on the types of contact with which the different mobility categories are associated [148, 149].

5.3.5 *Considering two scenarios*

Let us look at the evolution of the overall reproduction number (Fig. 5.18A). We can see how strongly the younger age groups dominated the epidemic dynamics due to their higher contact rates. In line with the contact rates, the overall reproduction number peaks during the long weekend of Sant Joan, reaching a maximal value of 3.4 (CrI: 3.0-3.9). As highlighted for the age-specific reproduction numbers, we only observe a slight decrease in the overall reproduction number as the nightclubs are closed. Similarly, we do not find any substantial impact of the curfews' implementation. These findings further highlight that the NPIs put in place seem not to have substantially impacted the epidemic's evolution.

To further corroborate the reproduction number inferred from the model, we compare it with the direct estimation from EpiEstim, showing good qualitative agreement (Fig. 5.18A). The earlier increase and later decrease of the

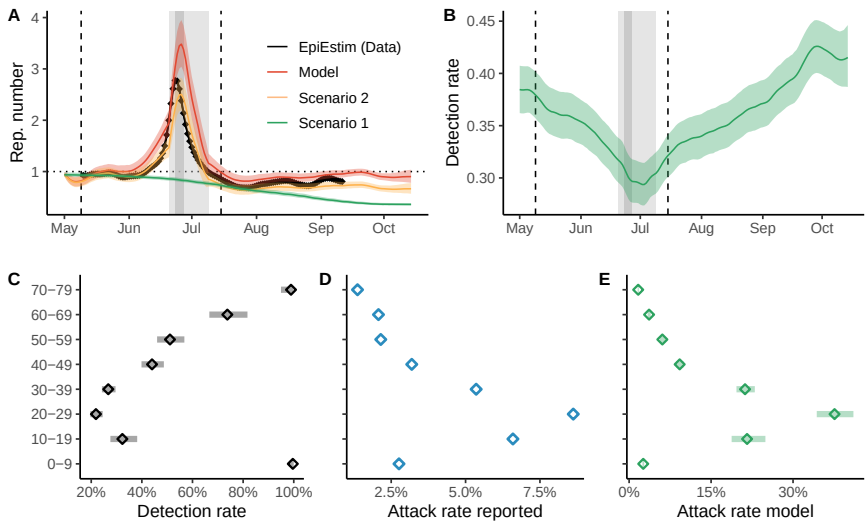


Figure 5.18: **(A)**: Overall reproduction number estimated from the model (red) and directly from the data via EpiEstim (black). Additionally, we show the overall reproduction number for Scenario 1 (green) and Scenario 2 (orange). The dark- and light-grey shaded area indicate the weekend of Sant Joan and the interval in which nightclubs were allowed to open, respectively. The dotted line indicates a reproduction number of 1. **(B)**: Evolution of the overall detection rate in time. The evolution is driven by changing infection patterns with respect to the age groups. **(C)**: Age-specific attack rates in the model. Markers show the median value, while horizontal bars indicate 95% CrI. **(D)**: Same to **(C)** but using the reported cases. **(E)**: Age-specific detection rates inferred through the model adjustment.

reproduction number derived from the model could be a consequence of the Gaussian smoothing prior we assumed for the interaction rates. For the details on the estimation of the reproduction number via EpiEstim, please see Appendix D.

Now, let us consider two scenarios to disentangle the impact of the vaccines, the varying contact rates and the presence of the Delta variant on the evolution of the reproduction number. In scenario 1, we rerun the dynamics assuming that interaction rates stay constant and that only the Alpha variant is present. In other words, we fix $\beta^i(t) = \beta^i(t = 0)\ \forall t$ and set the number of individuals infected with the Delta variant at $t = 0$ to 0.0. The remaining parameters are all left as estimated from the inference. In this way, solely the vaccine rollout determines the evolution of the epidemic. In scenario 2, we further include the changing interaction rates but still discard the pres-

ence of the Delta variant. Then, calculating the reproduction number for both scenarios allows us to quantify the different contributions.

Looking at the evolution of the reproduction number for scenarios 1 and 2 (Fig. 5.18A), we observe that the increasing contact rates were the main factor for the expansion of the epidemic. Even without the Delta variant (Scenario 2), our results indicate that the reproduction number would have been substantially above 1. In contrast, if the interaction rates had stayed constant, the vaccine rollout would have caused a continuous reduction of the reproduction number (Scenario 1). Therefore, the presence of the Delta variant has accelerated the epidemic expansion, but the main driver was the variation in the interaction rates.

5.3.6 *Estimation of the detection and attack rate*

We find a substantial variation in the detection rate across the age strata (Fig. 5.18C). The detection rate of the age group 0-9 stands out at almost 100%. The elevated detection rate is probably a consequence of the age-specific IHR we use, which stems from the first and second waves in England in 2020. The increased experience in dealing with SARS-CoV-2 infections among children possibly reduced the number of precautionary hospital admissions. We find the lowest detection rate for the age group 20-29 with 21% (CrI: 19-24%). The variation with age of the detection rate, together with the temporal evolution of the contact rates, causes the overall detection rate to vary substantially in time (Fig. 5.18B). On May 1, we find a detection rate of 38% (CrI: 36-41%) that subsequently decreases to 29% (CrI: 27-31%) at the beginning of July and then rises to 41% (CrI: 39-44%). In this sense, the rapid rise in reported cases at the beginning of July coincides with the lowest detection rate.

The evolution of the overall detection rate highlights how the fast-evolving epidemics among younger individuals made it increasingly difficult for health authorities to follow the epidemics. Nevertheless, the estimated detection rates are substantially higher – approximately double – than during the first wave of spring 2020 in Spain [484]. Except for the age groups 0-9 and 10-19, the detection rates decreased in line with the heterogeneous probability of developing symptoms [328]. Infected individuals in the 70-79 group were three times more likely to be detected than the ones in the 20-29. The variation in the detection rate across age groups highlights that monitoring

an epidemic becomes increasingly tricky as primarily younger individuals with milder symptoms are affected.

Similarly to the detection rate, the attack rate substantially varies across age groups. The reported cases correspond to an attack rate that varies between 1.4% for the age group 70-79 to 8.6% for the age group 20-29 (Fig. 5.18D). These differences are amplified if we consider the heterogeneity in the detection rates (Fig. 5.18E). In the model, we find an attack rate of 1.6% (CrI: 1.5-1.7%) and 37% (CrI: 34-41%) for the age groups 70-79 and 20-29, respectively. These findings suggest that from May 1 until October 15, a substantial fraction of the population between 10 and 40 years old got infected. The exact percentages strongly depend on the choice of the age-specific IHR. In the sensitivity analysis (Appendix D), we consider a more substantial increase in virulence of the Delta variant, which reduces our estimate of the attack rates in the age group 20-29 to 27.4% (CrI: 24.9-30.3%). Nevertheless, these values are a multiple of the attack rates during the first wave in 2020 among these age groups [484].

5.3.7 *Limitations of the study*

One of the limitations of the study is the uncertain parametrisation. In particular, the absence of age-specific estimations of the IHR for the Alpha variant and Delta variant required us to rescale estimations for the wild-type with the inferred increase in virulence of the Alpha variant and Delta variant. However, the estimations for the increase in virulence are subject to substantial uncertainty, wherefore we included it in the sensitivity analysis. In this sense, while we can safely conclude that a large part of the younger population got infected, we cannot estimate the precise percentage. Additionally, due to the absence of age-specific estimations of the IHR for Catalunya, we used estimations from the United Kingdom [468]. The variation of the IHR across regions [505] could add further uncertainty to our estimations. Considering natural immunity from infection did not substantially alter our findings, as shown in Appendix D.

Further limitations of the study are the contact patterns. Due to an absence of additional data, we are required to parametrise the model with a pre-pandemic contact matrix [72]. However, the substantial variation in the contact rates in the period under consideration suggests that the contact patterns had shifted. Furthermore, the variation in the mixing pattern could be

modelled if we had data available for Catalonia that is, for example, collected by the CoMix study [530]. The CoMix study leverages surveys to track the changes in the mixing patterns over time and was shown to be a good predictor of the epidemic's evolution.

From a more methodological point of view, one could also consider different geographical regions separately [297] or employ a meta-population model encoding individual mobility [531]. Unfortunately, such data are unavailable, as the Spanish government's project to monitor mobility flows through large-scale mobile phone data was abandoned on May 10. Nevertheless, the epidemic evolution was very similar across all regions, indicating that geographical aggregation is reasonable.

5.3.8 *Final thoughts*

Overall, our results indicate that, after the state of alarm, an increase in the contact rate among the younger population towards Sant Joan led to a very pronounced peak in the number of infections. The infections among the younger population that were not yet vaccinated eventually spilt over to the older ones, which caused a substantial number of hospital admissions. In contrast, as contacts increased again after the younger population had access to the vaccine, the reproduction number only minimally changed, and no material increase in infections was observed.

Interestingly, we found that the contact rate decreased before the closure of nightclubs and the introduction of the curfew. Furthermore, our results do not indicate any substantial impact of the two NPIs on the evolution of the contact rate. After more than a year of experience with the pandemic, the population may not have followed guidelines and NPIs as firmly as in early 2020. Often this phenomenon is referred to as pandemic fatigue [532, 533], and first studies tried to incorporate it to improve the design of NPIs [534]. These findings further challenge the mechanistic view of NPIs discussed in sec. 5.2.

Precisely due to the various factors beyond the actual policy that determine the impact of NPIs, we developed the model and inference framework presented here. Compared to other approaches that assume sharp change points in the reproduction number for the introduction of every NPI [314, 316], our framework relies on only a few assumptions. The primary assumption is the Gaussian smoothing prior, which prevents overfitting and guarantees smooth-

ness. Alternatively, there are a series of other approaches that could impose smoothness [225, 236].

Even though the inference makes relatively few assumptions, the approach is undoubtedly not model free. The inference framework is coupled to a relatively extensive compartmental model tailored to SARS-CoV-2. The possibility to flexibly adapt the epidemic model is the generality and appeal of our approach. More specifically, the compartmental model can be easily adapted to the characteristics of other viruses or other additional available data streams and can incorporate, as the recently proposed approach by Green et al. [535], any heterogeneities in the population. The framework then serves, similarly to EpiEstim [220, 225], to infer the evolution of the contact rates, i.e. the reproduction number, but at the desired resolution level with respect to age and geography.

The inferred evolution of the reproduction number can then serve to understand and quantify the impact of epidemiological factors, such as the presence of a new variant or the influence of behavioural changes on the contact patterns. As in sec. 5.2, the challenge remains to relate any changes in the contact patterns to different events. External sources to the analysis, such as data on mobility or empirical surveys like CoMix, can help in this regard [530]. In particular, if we had more detailed mobility data available than the one provided by the Google mobility report [529], we could try to associate changes in mobility in specific social settings, such as the workplace, restaurants, or nightclubs, with the spread of the epidemic. Nevertheless, the approach is also based on correlation rather than causality. However, instead of explicitly modelling the behavioural response, for example, with mobility as a proxy in sec. 5.1, this more data-driven approach allows to incorporate a broader behavioural spectrum. Although we leverage the available data, we purposely call this a "more" data-driven approach instead of a purely data-driven approach since it still largely relies on empirical knowledge of the characteristics of the virus, which is used to create the compartmental mathematical.

5.4 SUMMARY AND DISCUSSION

Contrary to the studies presented in chapters 3 & 4, which took a qualitative approach, we employed a series of models in this chapter to describe the spread of SARS-CoV-2 quantitatively. We have highlighted in sec. 5.1 the

difficulties that come along with predicting the future evolution of an epidemic, particularly if faced with substantial reporting delays and data uncertainty. However, the retrospective assessments of the first wave (5.2) and the Delta wave (5.3) have shown how models can be leveraged to infer aspects of human behaviour or characteristics of a virus variant. The findings in both secs. 5.2 & 5.3 added evidence [300, 510, 511] that the behavioural reaction beyond the introduction of NPIs crucially shapes the evolution of an epidemic. In particular, the results in sec. 5.2 regarding the correlation between the reduction in mobility and the number of confirmed cases support the risk-based approach we took to describe behavioural feedback in chapter 3.

However, the advances from our modelling efforts, which we presented in secs. 5.1 & 5.3, are not purely empirical but as well methodological. To our knowledge, it is the first time an MMCA approach describes the evolution of a real-world epidemic. Until this point, the MMCA was mainly employed in theoretical settings [175, 177, 208]. Our description of the impact of confinement provides a mechanistic and intuitive understanding with respect to the impact of lockdown on the epidemic dynamics. Furthermore, the approach is relatively easy to use and understand due to its discrete-time nature. In contrast, the framework used to model the Delta variant is closer to a black box. Given the reported epidemiological data, contact rates are inferred with only a few assumptions. On the downside, results can only be interpreted through correlation with external indicators such as mobility. In this sense, comparing the approaches in secs. 5.1 & 5.3 highlights the existing trade-off between flexibility and interpretability. In the conclusions of the thesis, which will be the next and last chapter, we will discuss the opportunities of more data-driven approaches and their possible drawbacks in more detail.

6 CONCLUSIONS

This thesis explored diverse aspects of human behaviour that either interact with or unidirectionally shape the spread of an epidemic. While Chapters 3 & 4 employed simple toy models, Chapter 5 focused on more quantitative approaches. Both approaches are equally valid, but they serve different purposes. Toy models mostly act as a mathematical formalisation of a thought experiment. They allow us to lay out a hypothesis and elaborate on its consequences in great detail. In contrast, quantitative approaches, beyond prediction, serve to add empirical evidence regarding the characteristics of the virus, the drivers behind the epidemic spread, the impact of previously introduced NPIs, or the changes in human behaviour. Furthermore, quantitative models can also be used to estimate the possible impact of containment measures, for example, studies that aimed to analyse whether SARS-CoV-2 could be controlled through digital proximity tracing [367, 388–394]. Nowadays, mathematical epidemiology typically focuses on more quantitative approaches, while the complex systems community is mainly interested in understanding the phenomenology provided by toy models. One of the contributions of this thesis is to reduce this gap, working on both ends of the spectrum simultaneously.

During this process, in Chapter 3, we outlined how risk-based adoption of prophylactic tools can lead to sustained epidemic cycles. It was hypothesised that a risk-driven behavioural adaption might be the cause behind recurrent epidemic waves for measles [270], syphilis [291], or more recently, SARS-CoV-2 [341, 342, 345]. The framework is relatively flexible and describes any prophylactic tool that reduces the transmission probability or the contact rate. Explicitly introducing the behavioural dimension allowed us to explain the recurrent epidemic waves in a Markovian framework without memory. Similarly to the memory kernel in Ref. [275], the mechanism of rapid behavioural adaption of our model can be essentially understood as having a system with

dynamic memory, as individuals adapt their behaviour based on information which does not reflect the current but past infection risk. From a more theoretical point of view, the dynamics we studied is a typical example of an evolutionary game with environmental feedback [308]. In our context, the evolutionary game corresponds to whether individuals adopt prophylaxis, and the environmental feedback comes from the evolving epidemic. In addition, we have shown that the risk-based voluntary adoption of prophylaxis can only contain the spread of a virus but cannot eradicate it. In mathematical terms, this means that the epidemic threshold remains unaffected by voluntary prophylaxis. The unaffected epidemic threshold is a general feature of models which assume a prevalence-based behavioural adaption. It calls into question whether infectious diseases can be eradicated through voluntary measures or central actors that impose mandates are always necessary. However, social pressure (complex contagion), beyond the mere prevalence-based adoption, is a mechanism that could sustain prophylaxis even in times of low prevalence [281, 536, 537] and lead to the eradication of the virus. An extension of our framework, including the effects of social pressure in the payoff structure, could be interesting for future work.

We further developed this framework to include a heterogeneous population structure. Our focus was primarily on the impact of the heterogeneous risk of severe infection exhibited by SARS-CoV-2, which vastly correlates with age [326–328]. Incorporating these differences across age groups into the model allowed us to explain two phenomena observed in different countries and regions in the reported case numbers: the temporal variation of the case distribution across age groups and the delay in the epidemic peak of the older age groups with respect to the rest of the population. The rationale behind these findings is that older individuals adhere to prophylaxis measures more strictly due to their higher infection cost, i.e. a higher risk of a severe infection. In contrast, younger individuals intensify prophylaxis only when the infection risk (prevalence) increases substantially. Hence, towards the epidemic peak, the behaviour of the younger and older individuals is similar, and the case distribution of the infections thus approaches the one of the population structure. Furthermore, since the younger individuals intensify prophylaxis later than the older ones, they drive the epidemic waves, and the infection risk eventually "spills over" to the older population, causing the observed delay in epidemic peaks. The relatively simple extension we proposed is an illustrative example of how toy models can be used to explain complex phenomenology in the data. Nevertheless, it remains a toy model, where in-

terpretability is more critical than empirical accuracy. In both approaches in Chapter 3, we analysed a hypothesis — the risk-based adoption of prophylaxis — and verified that the resulting dynamics could explain the observed phenomenology rather than reproduce the actual data.

However, toy models are not only valuable to do hypothesis testing with respect to the observed phenomenology, they also serve to analyse what type of phenomenology an empirical finding may imply in different scenarios. In Chapter 4, we considered two such empirical findings — homophily [357] and heterogeneity in the contact pattern [126] — and analysed their impact on epidemic control through prophylaxis. The phenomenology we encountered in both cases is inherently linked with the imperfection of the prophylactic tool. In the first part (Sec. 4.1), we have shown that the correlation between adoption and the contact structure, as suggested by the presence of homophily, can be either beneficial, detrimental, or affect the final attack rate non-trivially, depending on the quality of the prophylactic tool and the infectivity of the disease. The corresponding three dynamical regimes were first found for digital proximity tracing (DPT). In the case of DPT, the three dynamical regimes emerge since contact tracing apps do only slow down the spread of the disease if both the infected individual and the contacts are adopters. Accordingly, as adopters and non-adopters mix together, two phenomena coexist. The first one, which can be deemed beneficial, is that non-adopters profit from the protection of adopters. On the other hand, the number of contacts among adopters – the contacts along which onward transmission can be prevented – is reduced. Then, the competition between the two phenomena leads to the three previous dynamical regimes. Suppose the infectivity (basic reproduction number) is significant with respect to the coverage of DPT. In that case, clustered adoption is necessary to have any effect on the course of the epidemic, and homophily is beneficial to the efficacy of DPT. In contrast, if the basic reproduction number is small relative to the adoption, mixing between adopters and non-adopters is always beneficial since the high app coverage can provide protection for the entire population and not only the adopters. In between, the efficacy of DPT is non-monotonously affected by the mixing, and a local maximum exists.

In contrast to DPT, using individual prophylactic tools such as vaccines does not require the adoption of both the infected individuals and the contacts for it to work. However, if efficacy is low with respect to the infectivity of the disease, only a clustered adoption can provide actual protection. For this reason, we recovered the same three dynamical regimes in the case of

vaccines or any other prophylactic tool that reduces the transmission probability. If the vaccine's effectiveness is 100%, though, mixing is always beneficial since individual adoption already provides complete protection. The distinct phenomenology for perfect vaccines also explains why, previously, the common belief was that mixing between vaccine adopters and non-adopters is always beneficial. Studies which supported this finding mainly focused on childhood diseases for which vaccines have very high efficacy [411–415]. In this sense, the three dynamical regimes can only emerge if the vaccine, or the prophylactic tool, provides insufficient protection.

Similarly, we have shown in Sec. 4.2 that, in contrast to previous belief [433, 446, 447], protecting the individuals with the most contacts is not always the most efficient strategy if the efficacy is not 100%. Our results indicate a trade-off between the population-level impact to prevent super-spreading events and the probability of a breakthrough infection at the individual level. To be more precise, if the infectivity is high compared to the efficacy, there exists a contact rate for which protection maximally reduces the individual infection risk. In contrast, the reduction in the infection risk of the remaining population monotonously increases with the contact rate. Nevertheless, combining both the individual and the population level impact, a specific value of the contact rate still exists, which maximally reduces the prevalence. Accordingly, a risk-based strategy focusing on the most frequently interacting individuals is not always optimal. Our analysis has shown that the combination of high prevalence, heterogeneity in the contact structure, and low efficacy in the prevention measure promote the trade-off to occur. Since, for perfect protection, breakthrough infections do not occur, the risk-based strategy is always the best one in this case. Accordingly, as in the case of homophilic adoption, the phenomenology unveiled here is directly induced by the imperfection of the prevention tool.

A possible application of our findings would be to develop strategies for PrEP distribution in the MSM community. In many African countries, MSM communities show high HIV prevalence [453], substantial heterogeneity in the number of sexual interactions [430], and a relatively low PrEP efficacy due to irregular use [456]. Since much stigma surrounds HIV and the MSM community [452], it can be challenging to reach the concerned individuals, effectively hindering any sophisticated distribution strategy. For this reason, we have also shown that, in some conditions, random distribution can even outperform a risk-based approach. However, quantitative analyses tailored to the characteristics of a specific location are necessary before implementing

such strategies. Such an analysis is feasible if data on the population structure and partner acquisition patterns in the considered location is available. In this case, also agent-based models may be an option to incorporate all the empirical findings. Similarly, for the homophilic adoption of vaccines or DPT apps, if sufficient data on adoption patterns were available, it would be straightforward to incorporate homophily into quantitative epidemiological models. The possibility of building more realistic models separates the approaches presented in Chapter 4 from the ones outlined in Chapter 3. While we are still far from incorporating the dynamic behavioural response into quantitative models, any static aspects of human behaviour that are empirically measured can be easily utilised in epidemiological models. In this sense, the toy models employed in Chapter 4 also served to explore new directions along which quantitative efforts could be undertaken in the future.

Finally, in Chapter 5, we employed a series of quantitative models to describe the spread of SARS-CoV-2 in Spain. While Sec. 5.1 mainly focused on predictive modelling, Secs. 5.2 & 5.3 assessed the course of the epidemic retrospectively. The predictive modelling was centred around the first wave in Spain during early 2020. We first built a metapopulation model that incorporated commuting patterns to describe the spatial spread of SARS-CoV-2. It is probably the first time an MMCA approach describes an actual real-world epidemic [175, 460]. The MMCA approach is useful as it is in discrete-time and is thus computationally not very costly. We then adapted the model to incorporate the impact of confinement measures. We used the mobility level as a proxy for the reduction in contacts and the isolation of households inspired by Ref. [478]. Overall, the description of confinement allowed for a mechanistic understanding of how it impacts the epidemic dynamics. The model framework enabled us to algebraically express the reproduction number of the whole population, specific age strata, and each geographical patch. The reproduction number is defined in analogy with the next-generation matrix (NGM) approach [211, 214], but allows us to consider any case distribution across age strata and patches. It does not require the dynamics to reach the equilibrium distribution predicted by the NGM approach. Finally, the reproduction number's expression allows tracking the epidemic's evolution as the model is adjusted to the data.

We have also highlighted the difficulties encountered when adjusting the model in real-time. Many data streams were unreliable due to inconsistent reporting, and others suffered substantial reporting delays. The reporting delays increased the difficulty of outlining the future evolution of the epi-

demic, besides the already inherently challenging task of predicting the epidemic peak [486]. Accordingly, we needed data until the beginning of April – almost a week after the intensification of the lockdown – to predict the epidemic's decline. In hindsight, we know that the reproduction number was already below one shortly after the introduction of the first lockdown.

We observed the same delay during the retrospective analysis of the first wave in Sec. 5.1, for which we had additional data available, in particular, the date of symptom onset for the reported cases [487]. We estimated that the time lapse between the peak in infection was reached and the moment when it was reflected in the reported case numbers was around 16 days. The estimation of the reproduction number suggested that the epidemic slowed down several days before the implementation of the national lockdown on March 15. Similarly, large-scale mobility data [488] suggests a decrease in social activity from March 9/10 onward. Furthermore, we found a substantial correlation between the decrease in mobility and the accumulated cases across the Spanish provinces. This correlation is further evidence for the prevalence-based prophylaxis we considered in Chapter 3.

Since the estimated reproduction number started to decrease before mobility did, we used a minimal model to make a joint estimation of the reproduction number from the hospital admissions and deaths. The results from the inference are in line with the evolution of mobility. The alignment hence suggests that the reporting rate was decreasing towards lockdown. Furthermore, the minimal model allowed us to build counterfactual scenarios that quantify the epidemic response previous to the lockdown. The analysis suggests that the behavioural response previous to the lockdown reduced the attack rate, hospital admissions and deaths by around 30%. It is not easy to pinpoint what exactly caused the initial behavioural response. However, we assume it is an interplay between the news from Italy, the rising case numbers in Spain, and local containment measures. The different dynamics at play generally question the mechanistic view of NPIs, which often attribute the entire reduction in contacts to their implementation [314, 316, 490, 509].

Similarly, in the analysis of the spread of the Delta variant in Catalonia during the summer of 2021 (Sec. 5.3), our results indicated changes in the contact rates before implementing any NPIs. More precisely, among the younger population, contact rates constantly increased and reached their peak during the regional festivities of Sant Joan. Afterwards, they decreased before the Catalan authorities closed nightclubs and introduced a curfew. The decoupled evolution of NPIs and contact rates further indicates that at least

one year after the pandemic's start, behavioural changes beyond NPIs can crucially shape the epidemic evolution [300, 532, 533]. While the increase in contact rates was the main driver behind the epidemic expansion, the introduction of the more transmissible Delta variant and the relatively low vaccination coverage among the younger generation further accelerated the pathogen's spread. As contact rates increased again across all age groups, the reproduction number stayed below one since the younger age groups also had access to the vaccine. The differentiated outcome of increasing contact rates before and after the vaccination of the younger age groups highlights how delicate it can be to design opening strategies.

Beyond the empirical findings, the study is also interesting from a methodological point of view. We only used a minimal assumption, a Gaussian smoothing prior, to infer the evolution of the contact rates and the reproduction number. In contrast to EpiEstim [220, 225], the approach can deal with any heterogeneity in the population, as well as with the presence of different variants and vaccination statuses. A similar extension has recently been proposed to account for heterogeneities in the population when estimating the reproduction number in Ref. [535]. The only downside of our approach is that we assume a constant mixing matrix, which is not the case, as contact rates for each age group change daily. In this sense, future work could try to incorporate clever parametrisations for the contact matrix [18], allowing the inclusion of varying mixing patterns in the inference process. A brute force approach seems not viable since the number of parameters scales quadratically with the number of age groups. Overall, the framework is flexible and does not require making concrete assumptions on how and whether NPIs impact the evolution of the epidemic. It thus offers the possibility to retrospectively assess how changes in human behaviour, voluntary or not, impacted the course of an epidemic.

If we compare the frameworks used to model the first wave (Sec. 5.1) and the one used to describe the spread of the Delta variant (Sec. 5.3), we note that the latter is much closer to a so-called black box. Contact rates are directly inferred from the likelihood expression and the data through MCMC sampling. Accordingly, even though the framework offers excellent accuracy, the results may not always be easily interpretable. The evolution of the reproduction number can only be interpreted by correlating their evolution with other external events (e.g. the Sant Joan holiday), the introduction of NPIs or one can try to leverage mobility data. In contrast, the approach proposed to model the first wave mechanistically relates the reduction in mobility to the

changes in the contact rate. Such an approach may be less flexible but more interpretable since contact rates are causally linked to the level of mobility. Furthermore, only a few parameters need to be estimated instead of thousands. Relatively simple mechanistic models may seem outdated nowadays, mainly since computationally intensive approaches have often been at the forefront during the SARS-CoV-2 epidemic [148, 191, 367, 420, 538]. Furthermore, their accuracy might not be outstanding, but their explanatory power as an effective theory will probably make them powerful assets in the future.

On the other end of the spectrum, computational advances in the past decade have led to the explosion of the use of artificial intelligence, in particular deep neural networks. However, deep neural networks are an even bigger black box than the inference framework presented in Sec. 5.3. To this day, deep learning is only of anecdotal use in mathematical epidemiology [539]. During the SARS-CoV-2 pandemic, deep learning was mainly utilised to improve diagnosis by developing new diagnostic tools [540, 541]. In contrast, few examples exist where deep learning was used in actual modelling [539, 542]. Therefore, it is difficult to predict how it will impact the field of mathematical epidemiology. In the immediate future, deep learning will probably mainly serve to improve predictive modelling [539, 542] as accuracy is more relevant in this context than interpretability. Also, ensemble modelling has proven beneficial to ameliorate prediction [543]. In contrast, retrospective assessments always require at least some interpretability, as one is interested in unveiling the driving factors behind the observed dynamics. The SARS-CoV-2 pandemic generated a vast amount of data on how epidemics evolve and their associated behavioural reaction. Analysing this data, which was partly done in this thesis, offers a never seen opportunity to deepen the understanding of the interplay between human behaviour and the spread of infectious diseases. The fast-evolving field of symbolic regression could open new opportunities to detect or approximate the hidden laws in the newly available data [544].

In a more traditional approach, one can leverage the data to develop new simple models explaining the observed phenomenology, which we did in Chapter 3. Naturally, the question is whether one can translate these approaches into a more quantitative realm. If human behaviour does not evolve with the epidemic as in Chapter 4, quantitative models, or at least agent-based simulations, can always be built upon data availability. In these cases, the problem is not complex but only complicated. In contrast, the dynamic interplay between human behaviour and the spread of infectious diseases has

so many aspects and driving factors that, until now, all the approaches are crude simplifications. The example studied in Chapter 3 neglected the impact of NPIs or seasonality of SARS-CoV-2 and solely considered the prevalence-based prophylaxis. However, without this simplification, the model would hardly be interpretable and analytically tractable. For this reason, behavioural models employed in infectious disease modelling resort primarily to a qualitative approach.

Decoding the complexity of human behaviour, or at least some of its traits, would open the pathway to building quantitative models that could ameliorate the retrospective assessment but, more importantly, improve the predictions of epidemics. Put more pointedly; we could advance from building conditional scenarios to making actual predictions. In the context of SARS-CoV-2, scenarios typically make different assumptions about how individuals will reduce their contacts. Accordingly, the inclusions of the behavioural reaction – i.e. the decrease in contacts – in the epidemic models would advance us from scenario analysis to prediction. However, even assuming that human behaviour could be perfectly modelled, the public communication of a prediction may still undermine its accuracy. As we mentioned in the introduction, predictive modelling can act as a self-defeating prophecy since behavioural changes can alter the course of the epidemic. Accordingly, long-term predictive modelling of epidemics may be intrinsically doomed to fail.

Despite the possibility of prediction being an unsolvable task, a deeper understanding of what traits and aspects of human behaviour shape the propagation of an epidemic is still helpful in developing better containment policies. Although in many cases, it is not a lack of scientific understanding but instead a mix of economic and political barriers that hinder the eradication of diseases [545]. The necessary studies to unveil how human behaviour shapes the spread of infectious diseases are inherently interdisciplinary. This thesis tried to bridge the gap between simple toy models often used in complex systems science and quantitative approaches more widely used in mathematical epidemiology. Working at the boundaries of the two fields yielded advances in understanding phenomenology, added empirical evidence, and proposed novel methodology. However, establishing a fundamental theory of coevolution, at both the cultural and biological time scale, between human behaviour and the spread of infectious diseases will require the collaboration between sociologists, behavioural economists, historians, epidemiologists and virologists, to name a few. Calls for such endeavour were made in the past [6]. However, like the author's project mentioned in the introduction,

initiatives in this direction have not always been supported. Nevertheless, the urgency of the SARS-CoV-2 pandemic may contribute to the future success of grant proposals for these interdisciplinary projects. Hopefully, the advances of these efforts will support humanity in its quest to become resilient against infectious diseases and not let death triumph so often.

Appendices

UNIVERSITAT ROVIRA I VIRGILI
MODELLING THE INTERPLAY BETWEEN HUMAN BEHAVIOUR AND THE SPREAD OF INFECTIOUS DISEASES: FROM TOY MODELS TO QUANTITATIVE APPROACHES
Benjamin Franz Josef Steinegger

A

IMPERFECT DPT

The primary analysis is based on the assumption that adopters infected by adopters do not further transmit the disease. In other words, contact tracing reaches all individuals and is immediate. However, in any practical application, this assumption does not hold. Individuals may not isolate after notification, or secondary cases may have already transmitted the disease due to the delay between the detection of the primary case and the notification of its contacts.

To test the robustness of our results, we thus incorporate both these factors into the model, which allows us to analyse whether they alter the phenomenology. More specifically, we denote with $\epsilon \in [0,1]$ the probability of a notified adopter to not self-isolate. Furthermore, $\gamma \in [0,1]$ represents the impact of contact tracing on the infectious period τ. The larger γ, the larger the delay between infection and notification, hence longer the period, $\gamma\tau$, during which the infectious adopter can infect. We recover instantaneous tracing for $\gamma = 0$. In contrast, $\gamma = 1$ implies that adopters are not notified before the end of their infectious period. Accordingly, the parameter ϵ enters the equations only for $\gamma < 1$.

With these assumptions, individuals in the compartment I_{AA} can also transmit the disease. We can interpret γI_{AA} as the individuals who never receive the notification, whereas a fraction ϵ of those that did receive it instantaneously ($(1-\gamma)I_{\mathrm{AA}}$), do not self-isolate. Therefore, Eqs. (4.5)-(4.7) become

$$I_{\mathrm{N}}(t+1) = \beta[k_{\mathrm{NN}} I_{\mathrm{N}}(t) + k_{\mathrm{AN}} I_{\mathrm{AN}}(t) + (\gamma + \epsilon(1-\gamma))\, k_{\mathrm{AN}} I_{\mathrm{AA}}(t)] \frac{S_{\mathrm{N}}(t)}{N_{\mathrm{N}}} \tag{A.1}$$

$$I_{\mathrm{AN}}(t+1) = \beta k_{\mathrm{NA}} I_{\mathrm{N}}(t) \frac{S_{\mathrm{A}}(t)}{N_{\mathrm{A}}} \tag{A.2}$$

$$I_{\mathrm{AA}}(t+1) = \beta \left[k_{\mathrm{AA}} I_{\mathrm{AN}}(t) + (\gamma + \epsilon(1-\gamma))\, k_{\mathrm{AA}} I_{\mathrm{AA}}(t)\right] \frac{S_{\mathrm{A}}(t)}{N_{\mathrm{A}}}\,. \tag{A.3}$$

The equations of the number of susceptible individuals is not affected. If we define $\xi \equiv \gamma + \epsilon(1-\gamma)$, one clearly sees from Eqs. (A.1)-(A.3) that an individual in compartment I_{AA} has a reduced infectivity $\beta\xi \leq \beta$. Accordingly, any change in ϵ and/or γ leads simply to rescaling the infectiousness of the individuals in compartment I_{AA}. We find the next-generation matrix associated to the Eqs. (A.1)-(A.3) as

$$\mathrm{NGM} = \beta \begin{pmatrix} k_{NN} & k_{AN} & \xi k_{AN} \\ k_{NA} & 0 & 0 \\ 0 & k_{AA} & \xi k_{AA} \end{pmatrix}. \tag{A.4}$$

Consistently, the original next-generation matrix, given in Eq. (11), is recovered for $\xi = 0$, i.e., $\epsilon = \gamma = 0$. A straightforward calculation of the eigenvalues of the NGM shows its spectral radius, i.e., the effective reproduction number R, to be given by

$$\begin{aligned} R &= \frac{\beta}{2}\left[k_{NN} + \xi k_{AA} + \sqrt{(k_{NN} + \xi k_{AA})^2 + 4\left(k_{AN}k_{NA} - \xi k_{NN}k_{AA}\right)}\right] \\ &= \frac{R_0}{2}\Big[1 - \alpha T + \xi(1 - \alpha(1-T)) \\ &\quad + \sqrt{\left[1 - \alpha T - \xi\left(1 - \alpha\left(1-T\right)\right)\right]^2 + 4\alpha^2 T(1-T)}\Big]. \end{aligned} \tag{A.5}$$

With this expression at hand, let us proceed and do the same calculations as in the main text. We first impose $\frac{dR}{d\alpha} = 0$ to prove the existence of $\alpha = \alpha^*$, for which R reaches its minimum. As in the main text, the analysis yields a second-order equation with a single positive solution $\alpha = \alpha^*$, reading

$$\alpha^* = \frac{T - \xi\left[T + 1 - \xi\left(2 - T\right) + \xi^2\left(1 - T\right)\right] + \sqrt{\left(T + \xi - \xi T\right)^2\left(1 - \xi\right)^3}}{\left(1 - \xi\right)\left[\left(4 - 3T\right)T - 2\xi T\left(1 - T\right) + \xi^2\left(1 - T\right)^2\right]}. \tag{A.6}$$

We see that Eq. (A.6) correctly reduces to Eq. (4.12) when $\xi = 0$. Imposing $\alpha^* \leq 1$, we find the upper bound T^* for the app coverage T, below which $\alpha^* < 1$. It reads,

$$T^* = \frac{1}{3 + \xi}\left(1 + \frac{1 + \xi}{\sqrt{1 - \xi}}\right), \tag{A.7}$$

reducing to $T^* = 2/3$ for $\xi = 0$. Curiously, $T^* \leq 1$ if and only if $\xi \leq \left(\sqrt{5} - 1\right)/2 = 1/\phi \approx 0.618$ where ϕ is the golden ratio. Consequently, for greater values of ξ, $\alpha^* < 1$ for any $T \in [0, 1]$, meaning the minimum

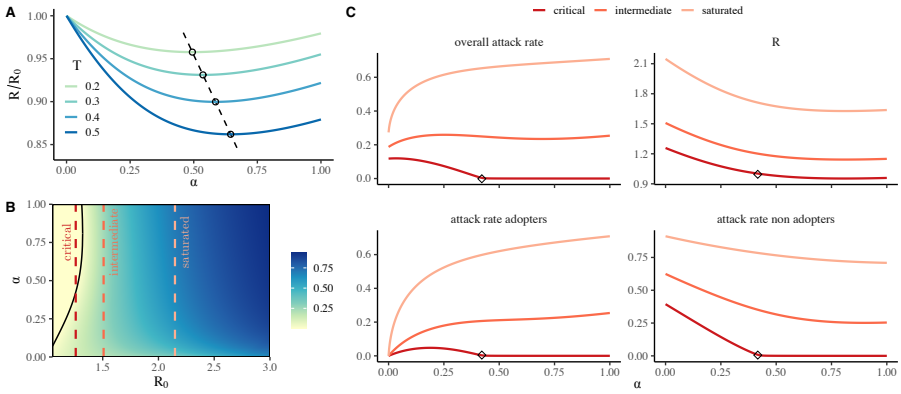


Figure A.1: Analogous illustration to Fig. 4.1 in the main text, but for $\gamma = 0.30$ and $\epsilon = 0.20$, giving $\xi = 0.44$. **(A)**: Reproduction number, R, normalized with respect to the basic reproduction number, R_0, for different values of adoption, T, below T^*. Dots indicate the minimum at α^*, while the dashed line shows its variation for $T \in [0.1, 0.55]$. **(B)**: Attack rate as a function of α and R_0. The solid line indicates the threshold $\alpha_c^{\pm}$ for which $R = 1$. Coloured, dashed lines denote the dynamical regimes: critical, intermediate and saturated. Adoption was fixed as $T = 0.7$. **(C)**: Top panels show the attack rate and the reproduction number for the different regimes defined in B. The specific attack rates for adopters and non-adopters are reported in the bottom panels. Black diamonds indicate α_c^-, at which $R = 1$.

for R is always met. We find that the disease is eradicated (i.e., $R < 1$) if $\alpha \in (\alpha_c^-, \alpha_c^+)$, with

$$\alpha_c^{\pm} = \frac{T + \xi(1-T) - \xi R_0}{2R_0 T(1-T)(1-\xi)} \pm \frac{\sqrt{[T + \xi(1-T) - \xi R_0]^2 - 4T(1-T)(1-\xi)\left[R_0(1+\xi) - 1 - \xi R_0^2\right]}}{2R_0 T(1-T)(1-\xi)}, \tag{A.8}$$

being the two roots of the equation $R = 1$. Like Fig. 4.1 in the main text, Fig. A.1 resumes the findings of the above analysis. The imperfections affect the efficacy of DPT, but the overall phenomenology is not altered.

UNIVERSITAT ROVIRA I VIRGILI
MODELLING THE INTERPLAY BETWEEN HUMAN BEHAVIOUR AND THE SPREAD OF INFECTIOUS DISEASES: FROM TOY MODELS TO QUANTITATIVE APPROACHES
Benjamin Franz Josef Steinegger

B

PARAMETERS USED IN SEC. 5.1

This Appendix describes the model parameters we fixed according to the literature in sec. 5.1 and reports the estimated values from the inference.

Table B.1: Epidemic parameters of the model and their estimations with 95% credible intervals for COVID-19 in Spain, with strata separation $g \in \{Y, M, O\}$ when appropriate.

Symbol	Description	States	Estimates	Assignment
β_I	Infectivity symptomatic	$S \to E$	0.075 CrI: 0.068 – 0.082	Calibrated
β_A	Infectivity asymptomatic	$S \to E$	$0.5\,\beta_I$	Assumed
η^g	Exposed rate	E	2.444 days^{-1} CrI: 1.823 – 3.064	Calibrated
α^g	Asymptomatic rate	A	$(5.671, 2.756, 2.756)$ days^{-1} CrI^Y : 5.046 – 6.356 CrI^M: 2.135 – 3.377 CrI^O : 2.135 – 3.377	Calibrated
μ^g	Infectious rate	I	$(1, 3.915, 3.915)$ days^{-1} CrI: 3.470 – 4.360	Y: Assumed M, O: Calibrated

Table B.2: Parameters of the model related to geographic and population data, including mobility, and their values for Spain.

Symbol	Description	Estimates for $g \in \{Y, M, O\}$ in Spain
n_i^g	Region population	Data provided by INE [481]
s_i	Region surface	Data provided by INE [481]
R_{ij}^g	Mobility matrix	Data provided by INE [481]
$\langle k^g \rangle$	Average total number of contacts	$(11.8, 13.3, 6.8)$ [72]
$\langle k_h^g \rangle$	Average number of contacts at home	$(3.1, 3.2, 3.3)$ [72]
$\langle k_w^g \rangle$	Average number of contacts at work	$(1.8, 5.2, 0.0)$ [72]
C^{gh}	Contacts-by-age matrix	$\begin{pmatrix} 0.5980 & 0.3849 & 0.0171 \\ 0.2440 & 0.7210 & 0.0350 \\ 0.1919 & 0.5705 & 0.2376 \end{pmatrix}$ [72]
ξ	Density factor	0.01 km^2
p^g	Mobility factor	$(1.0, 1.0, 1.0)$
σ	Average household size	2.5 [481]

Table B.3: Clinical parameters of the model and their estimations with 95% credible intervals for COVID-19 in Spain, with strata separation $g \in \{Y, M, O\}$ when appropriate.

Symbol	Description	States	Estimates	References
θ^g	Direct death probability	$I \to P_D$	$(0.0, 0.008, 0.047)$	[479, 484]
γ^g	ICU probability	$I \to P_H$	$(0.0003, 0.003, 0.026)$	[479, 484]
ζ^g	Pre-deceased rate	P_D	7.084 days^{-1} CrI: 6.640 – 7.537	[479]
λ^g	Pre-hospitalized in ICU rate	P_H	4.084 days^{-1} CrI: 3.640 – 4.537	[479]
ω^g	Fatality probability in ICU	$P_H \to H_D$	0.3	[479]
ψ^g	Death rate in ICU	H_D	7 days^{-1}	[546]
χ^g	ICU discharge rate	H_R	21 days^{-1}	[547]

Table B.4: Parameters describing the confinement and their estimations with 95% credible intervals for COVID-19 in Spain, with strata separation $g \in \{Y, M, O\}$ when appropriate.

Symbol	Description	Estimates	Assignment
κ_0^g	Mobility reduction	Time-varying	Estimated from INE [480]
ϕ	Household permeability	0.174 CrI: 0.079 – 0.269	Calibrated
δ	Social distancing	0.207 Calibrated CrI: 0.053 – 0.359	

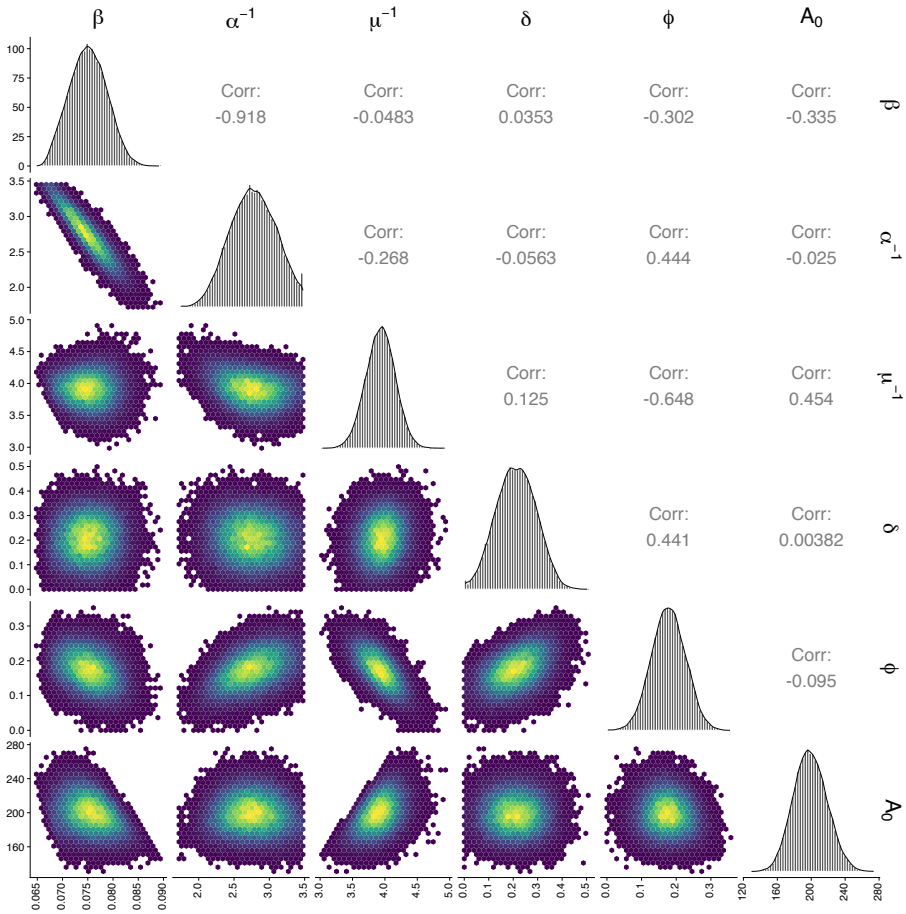


Figure B.1: The table's diagonal contains the posterior distributions for the parameters obtained after model calibration. The lower triangular part shows the scatter plot for every pairwise combination of parameters in the posterior parameter space. In contrast, the upper triangular one represents the entries of the covariance matrix for the parameters involved in the calibration.

C

ADDITIONAL FIGURES FOR SEC. 5.2

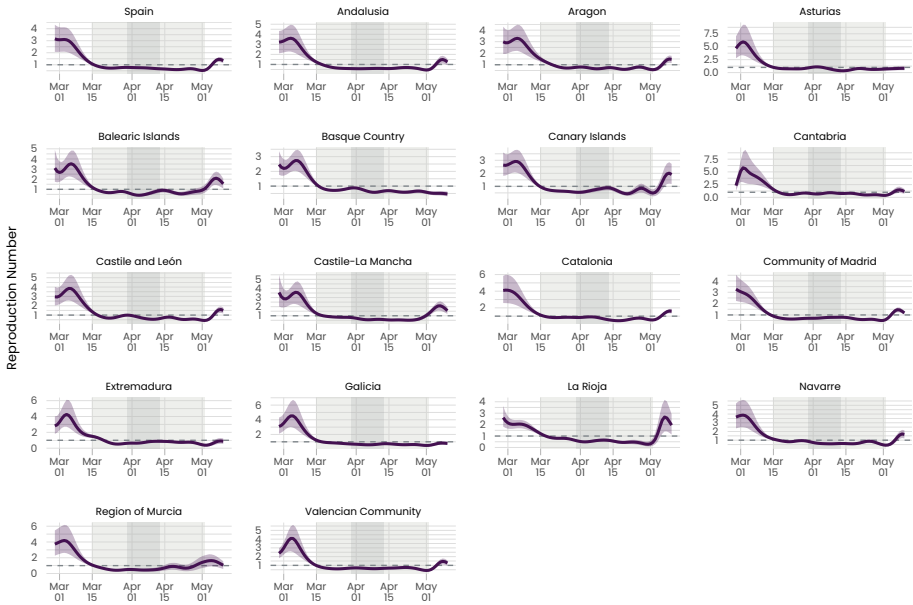


Figure C.1: Time series for the inferred Rts for Spain and the CCAA. Solid lines indicate a centred, seven-day rolling average. In red, we show the reconstructed exposure times. The shaded area in light grey indicates lockdowns 1 and 3. In dark grey, we indicate lockdown 2, where, in addition, all non-essential economic activity was shut down.

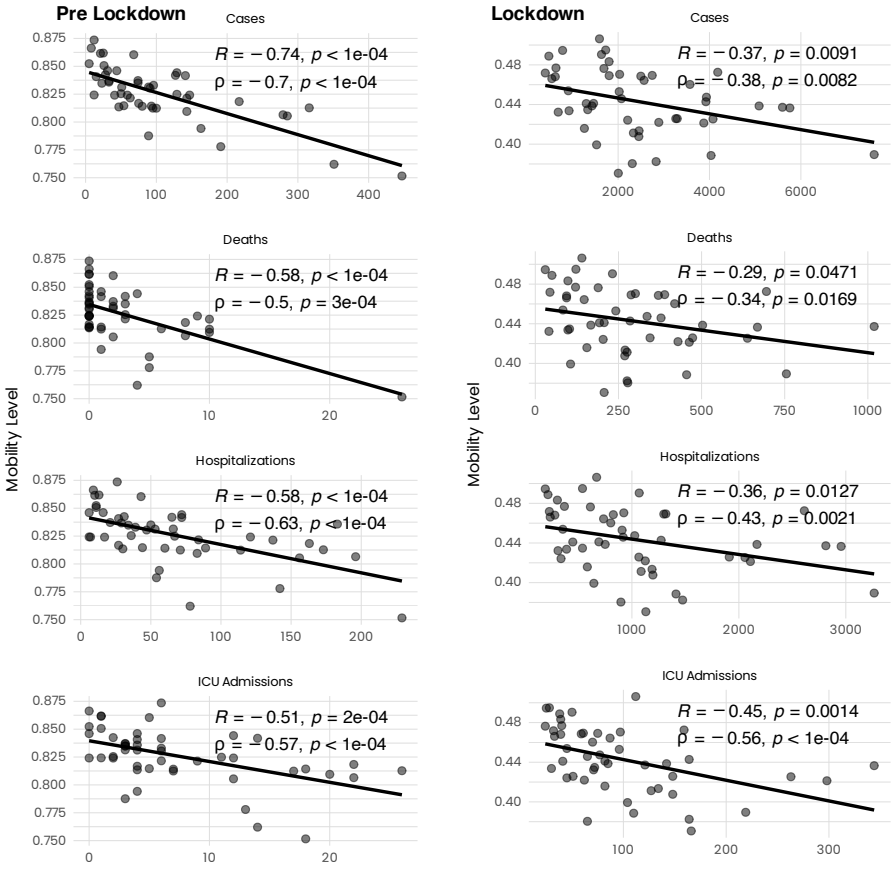


Figure C.2: Correlation between the mobility level [488] and accumulated number of reported cases, deaths, hospitalizations and ICU admissions [487]. Each data point corresponds to a province. The correlation was calculated through the Pearson R and Spearman ρ correlation coefficient. We excluded Ceuta and Melilla due to their size. Barcelona and Madrid were excluded since they are statistical outliers. The left side focuses on the week before the lockdown from March 9 to 15. The right side aggregates the lockdown period from March 15 to May 5. We observe that correlation between cases and mobility is reduced during lockdown compared to pre-lockdown. This reduction may be explained due to the stay-at-home order that reduced the effect of individual voluntary behaviour.

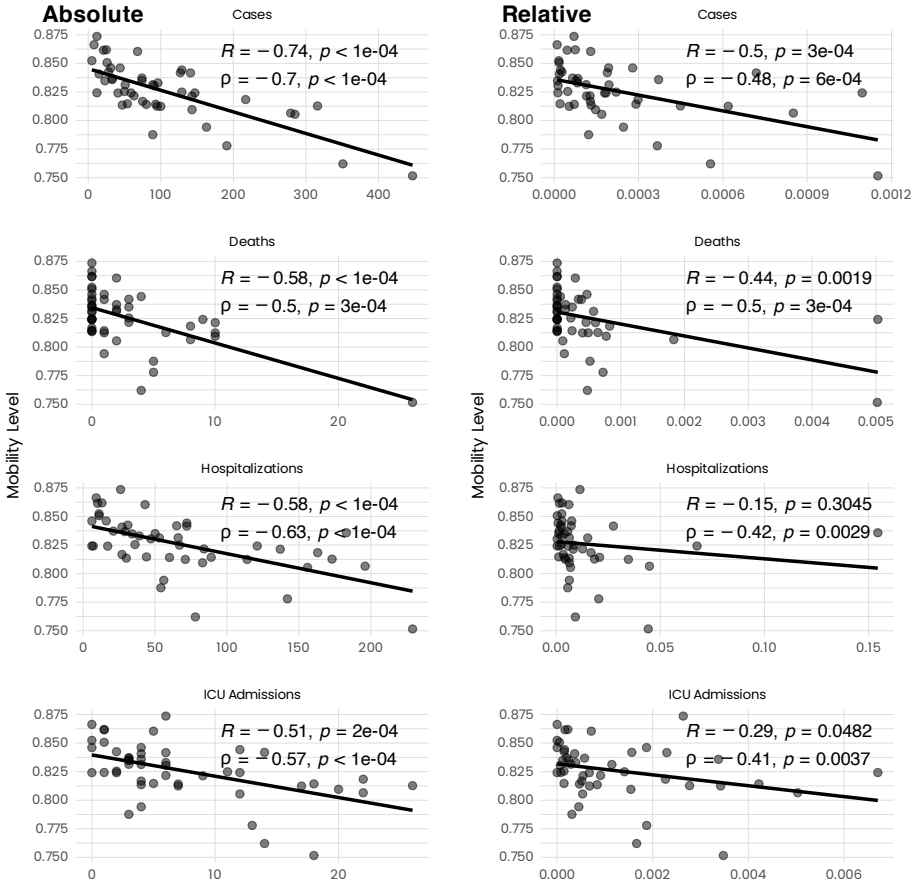


Figure C.3: Correlation between the mobility level [488] and accumulated number of reported cases, deaths, hospitalizations and ICU admissions [487] during the week previous to lockdown from March 9–15. Each data point corresponds to a province. The correlation was calculated through the Pearson R and Spearman ρ correlation coefficient. We excluded Ceuta and Melilla due to their size. Barcelona and Madrid were excluded since they are statistical outliers. The left side takes absolute values of the epidemiological indicators, while the right normalizes them with respect to the population. Interestingly, we find a stronger correlation regarding all indicators for absolute rather than relative case numbers.

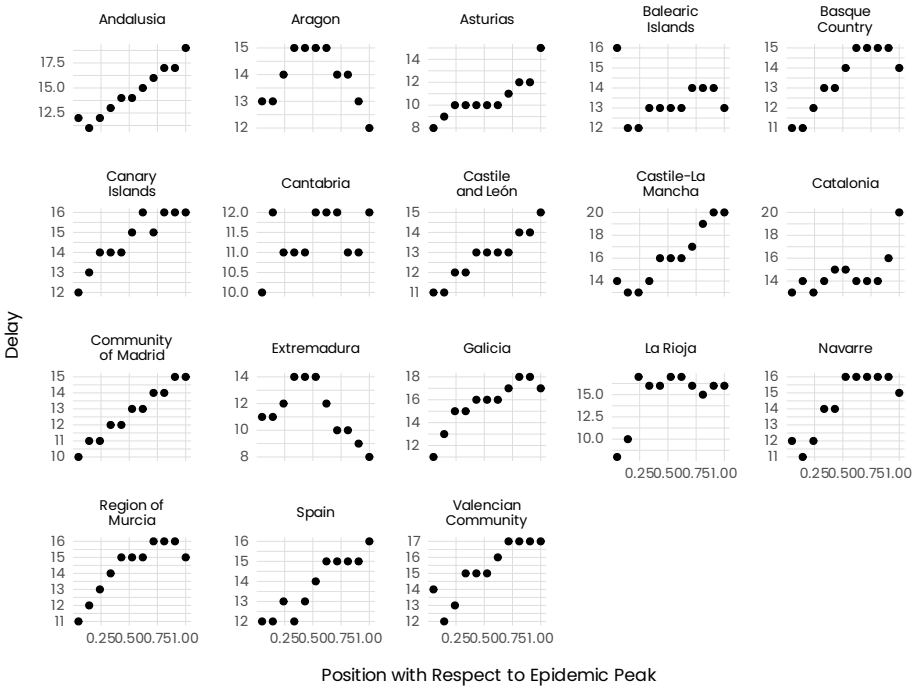


Figure C.4: Delay between the exposure time and the reporting date. The position is defined with respect to the peak. To be more precise, we show the time difference when both curves reached x% of their peak value. Each point corresponds to a CCAA. We note that the delay steadily increases towards the peak in almost all CCAA. Exceptions are Aragon and Extremadura.

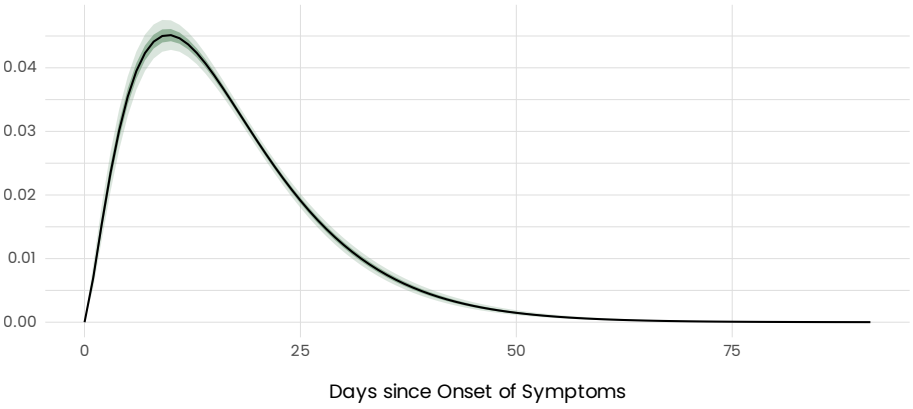


Figure C.5: Posterior for the distribution between the onset of symptoms and death. The median of the distribution is found as 14.5 days (CrI: 13.7 – 15.4). The shape and scale factors are given by 2.36 (CrI: 2.18 – 2.42) and 7.1 (CrI: 6.5 – 7.9).

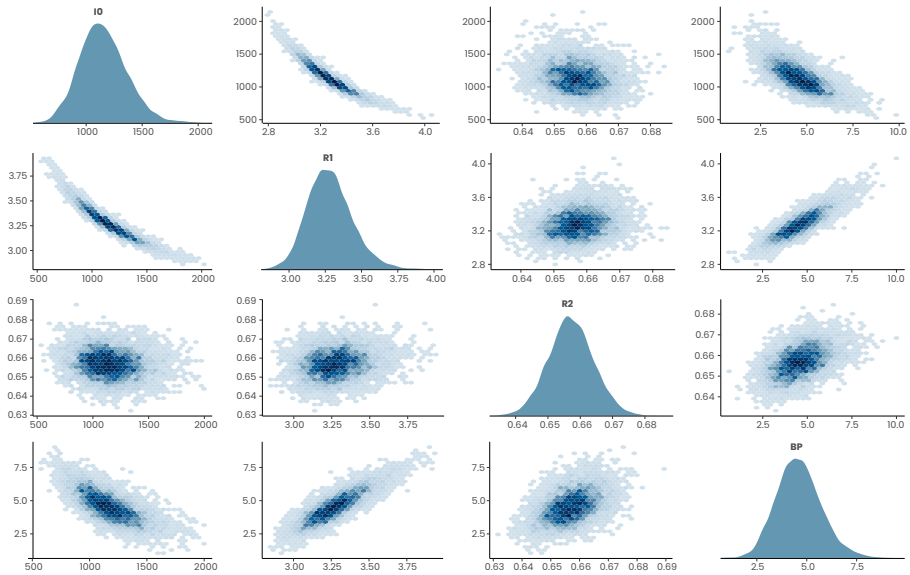


Figure C.6: Posterior distributions and correlations for the parameters I0, R1, R2 and BP that define how many days before lockdown the epidemic response was initiated.

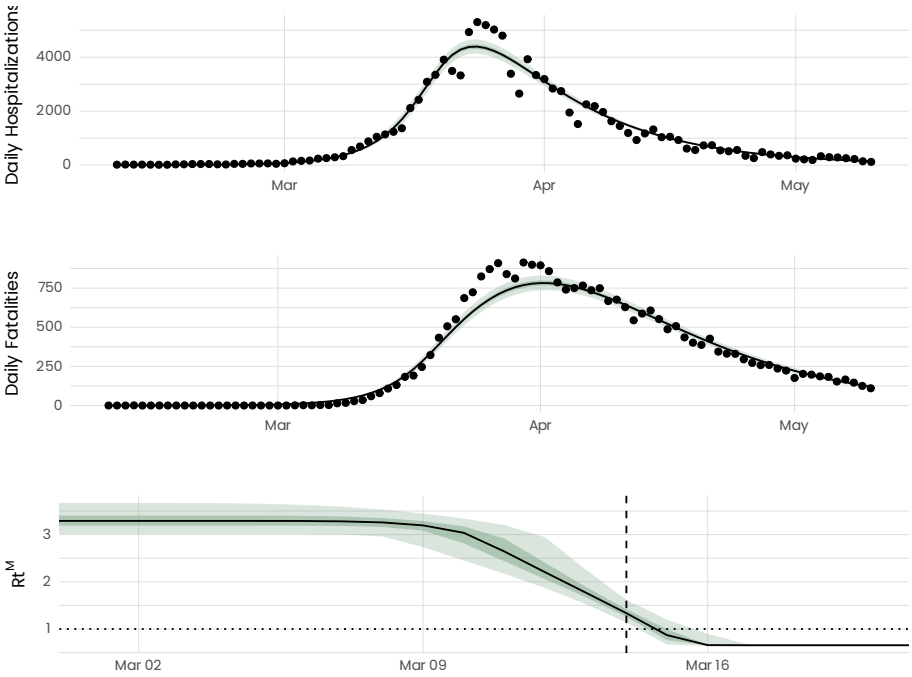


Figure C.7: Top and middle panels show the adjustment for the daily hospitalizations and fatalities as Rt^M if the second breakpoint is left as a free parameter. The bottom panel shows the inferred form of Rt^M. We find the inferred parameters I0, R1 and R2 as 1120 (CrI: 210 – 1590), 3.29 (CrI: 3.01 – 3.67) and 0.66 (CrI: 0.64 – 0.67), respectively. The decrease was initiated on March 10 (CrI: 6 – 12). Rt^M reaches a stable value between 0 and 2 days after the lockdown was implemented. The median is found half a day after the implementation of the lockdown. The relatively small credible interval supports our assumption that Rt^M reached a stable value on the day of lockdown. Furthermore, the rest of the inferred parameters are very similar to the ones presented in the main text.

D

SUPPLEMENTARY INFORMATION FOR SEC. 5.3

EPIDEMIOLOGICAL MODEL

We use a standard age-stratified compartmental model to describe the epidemic dynamics. The different compartments are shown in Fig. D.1. Individuals that are susceptible (S) and get infected transfer first to the latent state ($E_{1/2}$). The latent state is modelled with an Erlang distribution with shape factor $k = 2$. Subsequently, individuals become infectious (I). Infectious individuals either recover (R) or transfer to the compartment (P_H) that precedes hospitalization. The compartment is separated into individuals that will be then admitted to an ICU (H_{ICU}) and the ones that will recover (R) without ICU admission (H_R). Since we did not include the age group 80+ and thus do not explicitly consider the dynamics in care homes, we rule out the possibility of decease without ICU admission. The ICU compartment is separated to describe individuals that will recover (ICU_R) and the ones that will decease (ICU_D) and thus transfer to compartment D.

Furthermore, the compartments S, $E_{1/2}$ and I are separated into individuals that received one, two or no vaccine dose. Individuals pass to the vaccinated states according to the official number of daily first ($\nu_1^i(t)$) and second doses ($\nu_2^i(t)$) that are administered in every age group. We assume that susceptible, as well as individuals that were previously infected, are vaccinated. Depending on the vaccination status, individuals have a different transmission probability, susceptibility and risk of hospital admission (Table D.1). The vaccine is assumed to take effect 15 days after vaccination. Once hospitalized, the probability of being admitted to an ICU or death is independent of the individual's vaccination status.

Besides the stratification by age and vaccination status, we also distinguish whether the infection occurred through the Alpha or Delta variants. Individu-

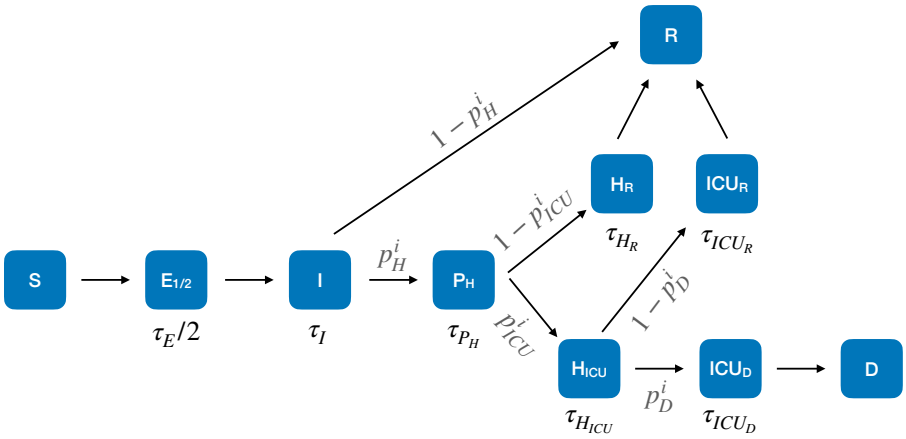


Figure D.1: Compartmental model for the dynamics. Black letters indicate the average time spent in each compartment, while grey letters represent transition probabilities.

als that were infected with the Delta variant have an increased transmission probability f_δ and hospitalization risk ζ_δ with respect to the Alpha variant.

The infection force depends on the contact matrix C and an age- and time-dependent interaction rate $\beta^i(t)$. We assume that the contact matrix (relative distribution of contacts) stays constant during the period considered here and is taken from Ref. [72]. The term $\beta^i(t)$ accounts for the age-dependent interaction rates. We assume that infectiousness and susceptibility are independent of age and thus lead to a rescaling of $\beta^i(t)$. The age-dependent infection hospitalization rate IHR is fixed according to the wild-type (Table D.2). We then correct the increase of IHR in the case of the Alpha/Delta variant with the factors ζ_α and ζ_δ respectively [523, 524]. For the risk of ICU admission and decease, we fix the relative differences between age groups according to Ref. [468] (Table D.3). The respective scaling factors, p_{ICU} and p_D, are left as free parameters and adjusted in the fit. We average the age-dependent quantities taken from Ref. [468] according to the census data in Catalonia [548] to match the age stratification considered here. The model's parameters are detailed in Tables D.1-D.6. The model's differential equations are reported in Eqs. (D.1)-(D.17).

PARAMETERS OF THE MODEL

Table D.1: Fixed parameters of the model that are non-age-dependent.

Variable	Description	Value
τ_E	Latent period	2.6 days [232]
τ_I	Infectious period (equals a generation time of $\tau_E + \tau_I$ = 5.2 days as in Ref. [232])	2.6 days
τ_{H_R}	Time in Hospital before recovery	8.4 days [549]
τ_{ICU_R}	Time in ICU before discharge	12.4 days [549]
ζ_α	Increased probability of hospitalizations of the Alpha variant with respect to the wild-type.	1.42 [523]
ζ_δ	Increased probability of hospitalizations of the Delta variant with respect to the Alpha variant.	1.85 [524]
$\gamma_{V_1}^{\{\alpha,\delta\}}$	Efficacy of one dose to prevent transmission if infected with the Alpha/Delta variant.	0.45 / 0.4 [420, 520]
$\gamma_{V_2}^{\{\alpha,\delta\}}$	Efficacy of two doses to prevent transmission if infected with the Alpha/Delta variant.	0.45 / 0.4 [420, 520]
$\sigma_{V_1}^{\{\alpha,\delta\}}$	Efficacy of one dose to prevent infection against the Alpha/Delta variant.	0.487 / 0.307 [521]
$\sigma_{V_2}^{\{\alpha,\delta\}}$	Efficacy of two doses to prevent infection against the Alpha/Delta variant.	0.937 / 0.88 [521]
$\varepsilon_{V_1}^{\{\alpha,\delta\}}$	Efficacy of one dose to prevent hospitalization for the Alpha/Delta variant.	0.8 / 0.95 [513, 522]
$\varepsilon_{V_2}^{\{\alpha,\delta\}}$	Efficacy of two doses to prevent hospitalization for the Alpha/Delta variant.	0.8 / 0.95 [513, 522]
M	Number of age groups.	8

Table D.2: Age-dependent IHR (p_H^i) [420, 548].

Age Group i	p_H^i
0-9	0.01500
10-19	0.00064
20-29	0.00180
30-39	0.00520
40-49	0.01500
50-59	0.03100
60-69	0.09800
70-79	0.27

Table D.3: Age-dependent relative risk of ICU admission after hospitalization (r_{ICU}^i) and decease after ICU admission (r_D^i). The absolute probabilities are then given by $p_{ICU}^i = p_{ICU} r_{ICU}^i$ and $p_D^i = p_D r_D^i$ [420, 548].

Age Group i	r_{ICU}^i	r_D^i
0-9	0.27	0.28
10-19	0.36	0.29
20-29	0.47	0.32
30-39	0.61	0.37
40-49	0.81	0.48
50-59	0.98	0.65
60-69	0.92	0.85
70-79	0.54	0.98

Table D.4: Adjusted parameters of the model which are age-dependent.

Variable	Description	Prior
ρ^i	Detection rate for the i-th age compartment.	Cauchy(0.5,1)
$\beta^i(t)$	Interaction rate for the i-th age compartment at time t.	Cauchy(0.5,2)

Table D.5: Adjusted parameters of the model which are age-independent.

Variable	Description	Prior	Value
f_δ	Transmission advantage of the Delta variant.	$\mathcal{N}(1.5, 0.5)$ [420, 525, 526, 550]	1.52 (CrI: 1.50-1.54)
p_{ICU}	Scaling factor for the age-dependent probability to ICU admission after hospitalization.	$\mathcal{N}(0.24, 0.06)$ [468]	0.24 (CrI: 0.22-0.26)
p_D	Scaling factor for the age-dependent probability to decease after ICU admission.	$\mathcal{N}(0.67, 0.05)$ [468]	0.4 (CrI: 0.36-0.45)
τ_{P_H}	Days between the infectious compartment and hospital admission.	$\mathcal{N}(5.0, 5.0)$ [505]	10.0 (CrI: 8.3-11.1)
$\tau_{H_{ICU}}$	Days between hospital admission and transfer to ICU.	$\mathcal{N}(2.0, 1.0)$ [505]	3.4 (CrI: 1.8-5.0)
τ_{ICU_D}	Days between ICU admission and decease.	$\mathcal{N}(10.0, 4.0)$ [505]	6.7 (CrI: 3.8-9.1)
I_0^α	Number of initially infected individuals with Alpha variant.	$\mathcal{N}([\tau_E + \tau_I] \times I_0^\delta, 0.2 \times [\tau_E + \tau_I] \times I_0^\delta)$	29059 (CrI: 24001 - 34770)
I_0^δ	Number of initially infected individuals with the Delta variant.	$\mathcal{N}(100.0, 80.0)$	53 (CrI: 37-75)
ϕ	Dispersion parameter for the negative binomial in the log-likelihood (See Eq.(5.54)).	$\phi^{-1} \sim \text{Exp}(5.0)$	6.5 (CrI: 6.1-7.0)

VARIABLES OF THE MODEL

Table D.6: Variables of the model.

Variable	Description
$S^i_{\{NV,V_1,V_2\}}(t)$	Number of susceptible individuals in age group i at time t, either non vaccinated (NV) or with one dose (V_1) or fully vaccinated (V_2).
$E^{\{\alpha,\delta\},i}_{\{1,2\},\{NV,V1,V2\}}(t)$	Number of individuals in the $NV/V_1/V_2$ vaccination state, belonging to age group i that were infected with the Alpha/Delta variant and are in the two respective latent states (1 or 2) at time t.
$I^{\{\alpha,\delta\},i}_{\{NV,V_1,V_2\}}(t)$	Number of individuals in the $NV/V_1/V_2$ vaccination state, belonging to age group i, that were infected with the Alpha/Delta variant and are infectious at time t.
$P^i_H(t)$	Number of individuals in age group i that will subsequently be admitted to the hospital.
$H^i_R(t)$	Number of individuals in age group i that are in the hospital at time t but will recover without being admitted to the ICU.
$H^i_{ICU}(t)$	Number of individuals in age group i that are in the hospital at time t and will subsequently be admitted to the ICU.
$ICU^i_{\{R,D\}}(t)$	Number of individuals in age group i that are in the ICU at time t and will recover/decease.
$D^i(t)$	Number of individuals in age group i that deceased until time t.
$R^i(t)$	Number of individuals that recovered from infection until time t.
$R^i_{NV}(t)$	Number of recovered not vaccinated individuals that were infected until time t.
$R^i_{V_1}(t)$	Number of recovered individuals that were infected while being vaccinated with one dose until time t.
$\lambda^{\{\alpha,\delta\},i}_{\{NV,V_1,V_2\}}(t)$	Infection force at time t of the Alpha/Delta variant for individuals belonging to age group i, in the $NV/V_1/V_2$ vaccination state.

MODEL EQUATIONS

$$\dot{S}^i_{NV}(t) = -\Big[\sum_\theta \lambda^{\theta,i}_{NV}\Big] S^i_{NV}(t) - \nu^i_1(t)\frac{S^i_{NV}(t)}{S^i_{NV}(t)+R^i_{NV}(t)} \tag{D.1}$$

$$\dot{S}^i_{V_1}(t) = -\Big[\sum_\theta \lambda^{\theta,i}_{V_1}\Big] S^i_{V_1}(t) + \nu^i_1(t)\frac{S^i_{NV}(t)}{S^i_{NV}(t)+R^i_{NV}(t)} - \nu^i_2(t)\frac{S^i_{V_1}(t)}{S^i_{V_1}(t)+R^i_{V_1}(t)} \tag{D.2}$$

$$\dot{S}^i_{V_2}(t) = -\Big[\sum_\theta \lambda^{\theta,i}_{V_2}\Big] S^i_{V_2}(t) + \nu^i_2(t)\frac{S^i_{V_1}(t)}{S^i_{V_1}(t)+R^i_{V_1}(t)} \tag{D.3}$$

$$\dot{E}^{\theta,i}_{1,\omega}(t) = \lambda^{\theta,i}_\omega(t) S^i_\omega(t) - \frac{2}{\tau_E} E^{\theta,i}_{1,\omega}(t) \tag{D.4}$$

$$\dot{E}^{\theta,i}_{2,\omega}(t) = \frac{2}{\tau_E} E^{\theta,i}_{1,\omega}(t) - \frac{2}{\tau_E} E^{\theta,i}_{2,\omega}(t) \tag{D.5}$$

$$\dot{I}^{\theta,i}_\omega(t) = \frac{2}{\tau_E} E^{\theta,i}_{2,\omega}(t) - \frac{1}{\tau_I} I^{\theta,i}_\omega(t) \tag{D.6}$$

$$\dot{P}^i_H(t) = \frac{1}{\tau_I} p^i_H \zeta_\alpha \Big[\sum_\omega (1-\epsilon^\alpha_\omega) I^{\alpha,i}_\omega(t) + \zeta_\delta \sum_\omega (1-\epsilon^\delta_\omega) I^{\delta,i}_\omega(t)\Big] \tag{D.7}$$

$$\dot{H}^i_R(t) = \frac{1}{\tau_{P_H}}(1-p^i_{ICU}) P^i_H(t) - \frac{1}{\tau_{H_R}} H^i_R(t) \tag{D.8}$$

$$\dot{H}^i_{ICU}(t) = \frac{1}{\tau_{P_H}} p^i_{ICU} P^i_H(t) - \frac{1}{\tau_{H_{ICU}}} H^i_{ICU}(t) \tag{D.9}$$

$$\dot{ICU}^i_R(t) = \frac{1}{\tau_{H_{ICU}}}(1-p^i_D) H^i_{ICU}(t) - \frac{1}{\tau_{ICU_R}} ICU^i_R(t) \tag{D.10}$$

$$\dot{ICU}^i_D(t) = \frac{1}{\tau_{H_{ICU}}} p^i_D H^i_{ICU}(t) - \frac{1}{\tau_{ICU_D}} ICU^i_D(t) \tag{D.11}$$

$$\dot{D}(t) = \frac{1}{\tau_{ICU_D}} ICU^i_D(t) \tag{D.12}$$

$$\begin{aligned}\dot{R}(t) = {} & \frac{1}{\tau_I}\Big[\sum_\omega \big(1-p^i_h\zeta_\alpha(1-\epsilon^\alpha_\omega)\big) I^{\alpha,i}_\omega(t) + \big(1-p^i_h\zeta_\alpha\zeta_\delta(1-\epsilon^\delta_\omega)\big) I^{\delta,i}_\omega(t)\Big] \\ & + \frac{1}{\tau_{H_R}} H^i_R(t) + \frac{1}{\tau_{ICU_R}} ICU^i_R(t)\end{aligned} \tag{D.13}$$

$$\dot{R}^i_{NV}(t) = \frac{1}{\tau_I}\sum_\theta I^{\theta,i}_{NV}(t) \tag{D.14}$$

$$\dot{R}^i_{V_1}(t) = \frac{1}{\tau_I}\sum_\theta I^{\theta,i}_{V_1}(t) \tag{D.15}$$

$$\lambda^{\theta,i}_{NV}(t) = \beta^i(t) f_\theta \sum_{j,\omega} \frac{C_{ij}}{N_j}(1-\gamma^\theta_\omega) I^{\theta,j}_\omega(t) \tag{D.16}$$

$$\lambda^{\theta,i}_{\omega\neq NV}(t) = (1-\sigma^\theta_{\omega\neq NV})\lambda^{\theta,i}_{NV}(t), \tag{D.17}$$

where $\theta \in \{\alpha, \delta\}$ indicates the variant, $\omega \in \{NV, V_1, V_2\}$ indicates the vaccination state and $i \in \{1, \dots, M\}$ indicates the age group. We introduced the rescaled efficacy to prevent hospitalization $\epsilon_\omega^\theta = (\varepsilon_\omega^\theta - \sigma_\omega^\theta)/(1 - \sigma_\omega^\theta)$. We also set for convenience $\varepsilon_{NV}^\theta = \gamma_{NV}^\theta = \sigma_{NV}^\theta = 0$ for each value of θ, and $f_\alpha = 1$.

ESTIMATING THE REPRODUCTION NUMBER

From the model

The dynamical states of the model allow us to calculate the instantaneous reproduction number $Rn^i(t)$ at time t respective to the age group i. We express $Rn^i(t)$ in the same way as in sec. 5.1 by considering the secondary infectious created by an average individual in age group i that becomes infectious at time t. The expression of $Rn^i(t)$ is given by:

$$Rn^i(t) = \beta^i(t) \sum_{\theta,\omega_1,\omega_2,j} f_\theta (1 - \gamma_\omega^{\theta_1})(1 - \sigma_{\omega_2}^\theta) \frac{\mathcal{I}_{\omega_1}^{\theta,i}(t)}{\sum_{\Omega,\Theta} \mathcal{I}_\Omega^{\Theta,i}(t)} S_{\omega_2}^j(t) \frac{C_{ij}}{N_j} . \tag{D.18}$$

The term $Ic_{\omega'}^{\theta,i}(t)$ corresponds to the incidence of variant θ in age-stratum i, with vaccination status ω, and is given by

$$\mathcal{I}_m^{\delta,i}(t) = \frac{\rho^i}{\tau_I} \sum_\omega \int_{t-1}^t I_\omega^{\theta=\delta,i}(y) dy . \tag{D.19}$$

The overall reproduction number $Rn(t)$ is then calculated from the average of the $Rn^i(t)$, weighted by the respective $\mathcal{I}^i(t) = \sum_{\theta,\omega} \mathcal{I}_{\omega'}^{\theta,i}(t)$. That is, $\mathcal{I}(t) = \sum_{i=1}^M \mathcal{I}^i(t)$ being the total number of infectious individuals, it reads

$$Rn(t) = \frac{1}{\mathcal{I}(t)} \sum_i \mathcal{I}^i(t) Rn^i(t) . \tag{D.20}$$

From the reported cases

We estimate the reproduction number from the reported cases through EpiEstim [220, 225]. EpiEstim estimates the reproduction number based on the generation time distribution. To accurately compare the results from our model with the reproduction number inferred from the reported cases, we use the generation time distribution of the epidemic model. The generation

time can be computed based on the approach of Ref. [227]. Given the random latent time, X, and a random infectious time, Y, the probability density function of the generation time distribution, $p_G(t)$, is given by the convolution $g * h_s$. Here, g is the probability density function of X, while h_p is expressed as

$$h_s(t) = \frac{1 - F(t)}{\bar{Y}} \tag{D.21}$$

with $F(t)$ being the cumulative distribution function of Y and $\bar{Y}$ the expected value. Based on this approach, we find

$$p_G(t) = \frac{4}{\tau_E^2 \tau_I} \left[\frac{1}{\eta^2} \left(e^{-\frac{t}{\tau_I}} - e^{-2\frac{t}{\tau_E}} \right) - \frac{t}{\eta} e^{-2\frac{t}{\tau_E}} \right], \tag{D.22}$$

where we defined $\eta = \frac{2}{\tau_E} - \frac{1}{\tau_I}$.

To estimate the reproduction number with EpiEstim, it is necessary first to infer the infection times (sec. 2.3.3). In our model, individuals are detected once they exit the infectious compartment. Accordingly, the distribution of the time between infection and detection, p_{ID}, is given by the convolution between the probability density function of the latent time, X, and the infectious time, Y, respectively. Computing this convolution we find $p_{ID} = p_G(t)$. Given p_{ID}, we infer the infection times from the reported cases through a maximum likelihood deconvolution approach [497, 498], as in sec. 5.2. We fixed the smoothing factor $k = 6$, corresponding to a centred rolling average of 7 days. The deconvoluted time series was then used as an input for EpiEstim to estimate the reproduction number with generation time $p_G(t)$.

ADDITIONAL FIGURES OF THE MAIN ANALYSIS

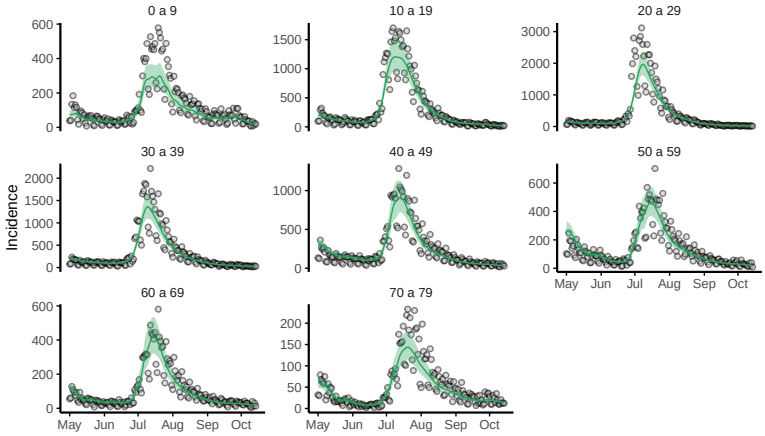


Figure D.2: Adjustment of the model to the reported cases (incidence) for all age groups.

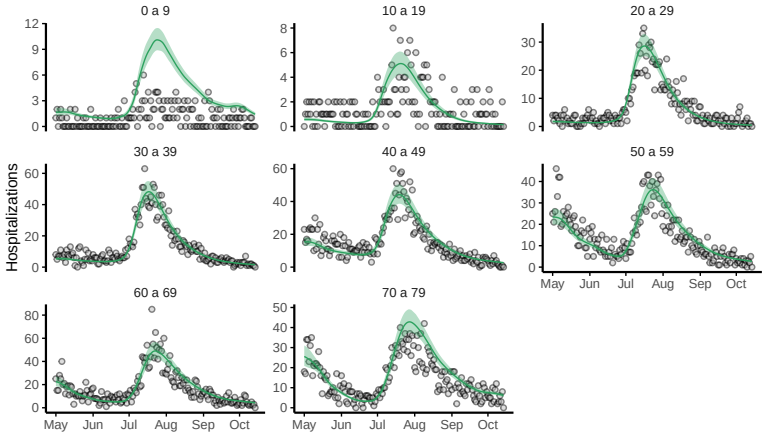


Figure D.3: Adjustment of the model to the daily hospitalizations for all age groups.

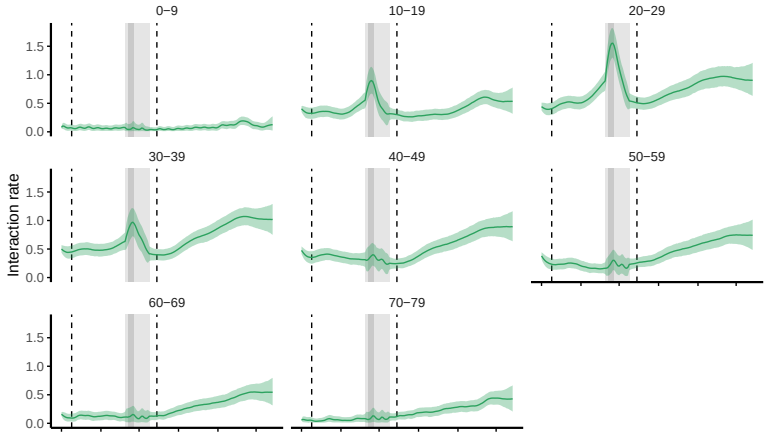


Figure D.4: Interaction rate inferred from the model for all age groups.

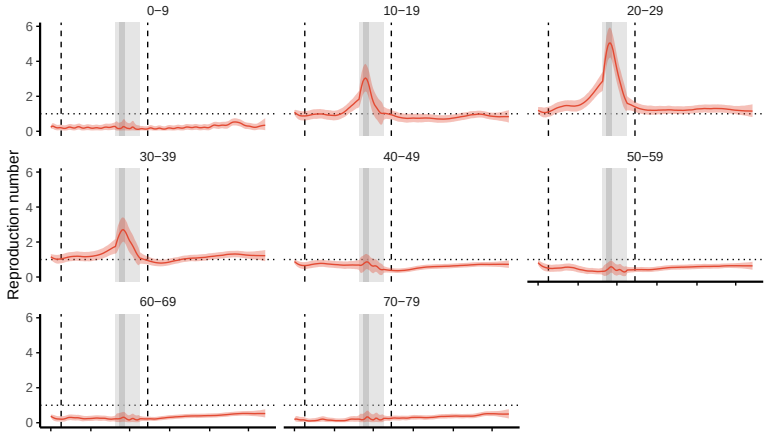


Figure D.5: Reproduction number inferred from the model for all age groups.

SENSITIVITY ANALYSIS

Generation Time

The transmission advantage of the Delta variant strongly depends on the generation time [551]. For this reason, we explore a set of different values of the generation time. Motivated by the range of existing estimates of the generation time [552–554], we repeat the numerical analysis with generation times of ± 1 compared to the value of 5.2 chosen for the main analysis [232]. There is also evidence that the Delta variant exhibits a shorter generation time than the alpha variant [555] (4.6 vs. 5.5). We also explore this scenario, and the results are shown in Fig. D.6. The estimated transmissibility advantage ranges between 1.41 (CrI: 1.39-1.43) and 1.63 (CrI: 1.60-1.66).

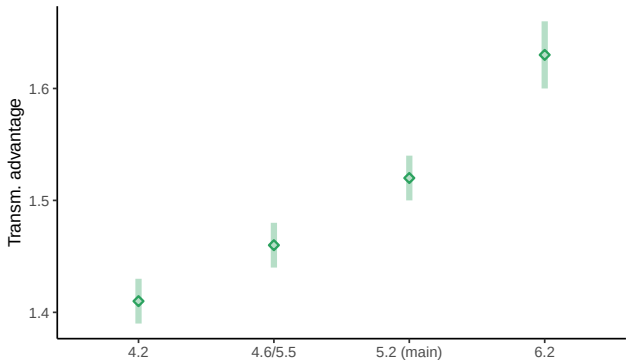


Figure D.6: Estimated transmission of the Delta variant for different choices of the generation time. The entry 4.6/5.5 corresponds to a generation time of 4.6 and 5.5 for the Delta and Alpha variant [555], respectively .

Increased severity Delta variant

Given the hospital admissions, the IHR fixes the attack rate. However, the increase in severity of the Delta variant with respect to the Alpha variant is subject to substantial uncertainty. In the main analysis, we fixed the increase in severity as 1.85 according to Ref. [524]. However, other estimates exist that report values up to 2.83 [556, 557]. We report the attack rates with an increase in severity of 2.83 in Fig. D.7. We observe that the attack rate in the age strata 20-29 reduces to 27.4% (CrI: 24.9-30.3%).

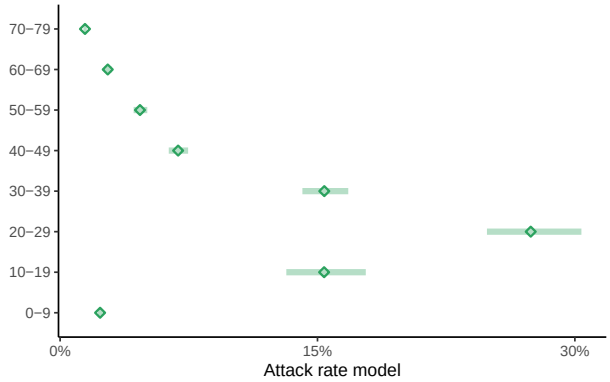


Figure D.7: Model-inferred attack rates for the different age-strata as we fix the increase of severity of the Delta variant at 2.83, according to Ref. [557].

Previous immunity through infection

We did not include existing immunity from previous infections in the analysis presented in the main text. We excluded it because there is still considerable uncertainty on how much such immunity last [558], and even more so if we consider reinfections across different variants and the wild-type. Additionally, including all these factors would require a complete reconstruction of the infections since the beginning of the pandemic. To account for this uncertainty, we consider here a case where all previously infected individuals have complete immunity. We fix the previous immunity according to a study that estimated seroprevalence in Catalonia between June 2020 and November 2020 [559]. The overall seroprevalence was estimated to be around 15%. Taking estimations from the end of 2020 neglects infections that occurred between November 2020 and May 2021. However, assuming complete immunity should compensate substantially for the lower seroprevalence. Figs. D.8-D.10 repeat the analysis presented in the main text incorporating immunity-induced infection. We do not observe any substantial difference compared to the analysis in the main text. Therefore, the conclusions drawn in the main text do also hold when previous immunity is included.

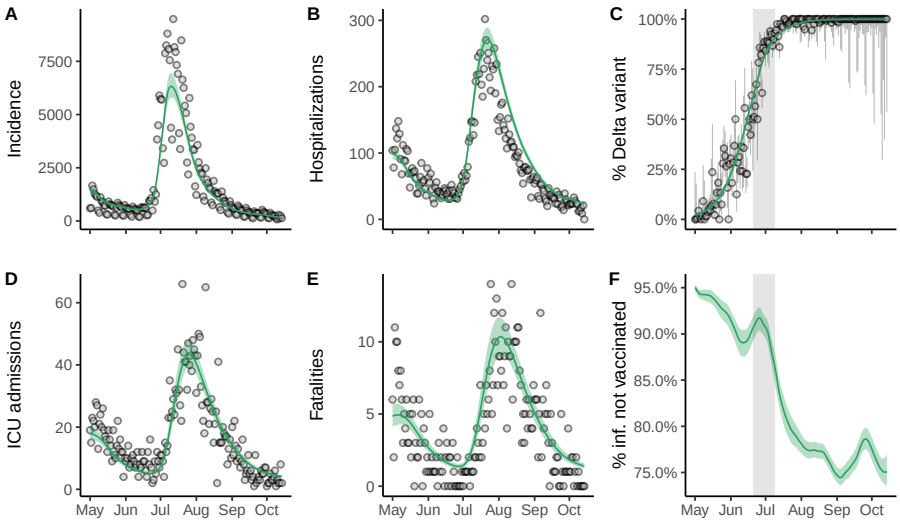


Figure D.8: Adjustment to the data for daily incidence **(A)**, hospitalizations **(B)**, ICU admissions **(D)** and fatalities **(E)**. Panel **C** compares the sequencing data with the model. It shows the fraction of infections that stems from the Delta variant. Vertical bars indicate the 95% credible interval assuming a uniform prior. **(F)**: Fraction of infected individuals that were not vaccinated. The still very low vaccination coverage as the nightlife was opened (grey shaded area) led to a high proportion of infected individuals that were not vaccinated in that period.

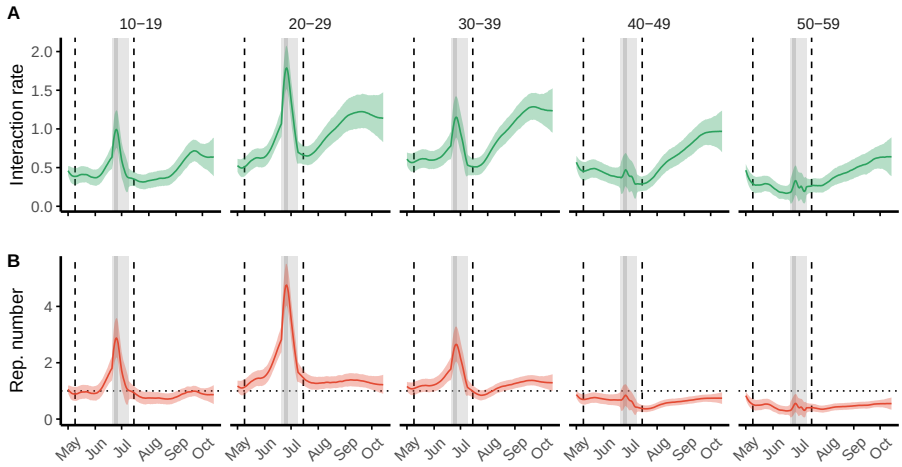


Figure D.9: Inferred interaction rate **(A)** and reproduction number **(B)** for the age groups 10-19, 20-29, 30-39, 40-49 and 50-59. We omitted the remaining age groups to improve visibility. For completeness, they are shown in Figs. S4-S5 in the SI. The peak in the interaction rate and the reproduction number coincide with the festivities of Sant Joan (dark-grey shaded area). The dotted line indicates a reproduction number of 1.

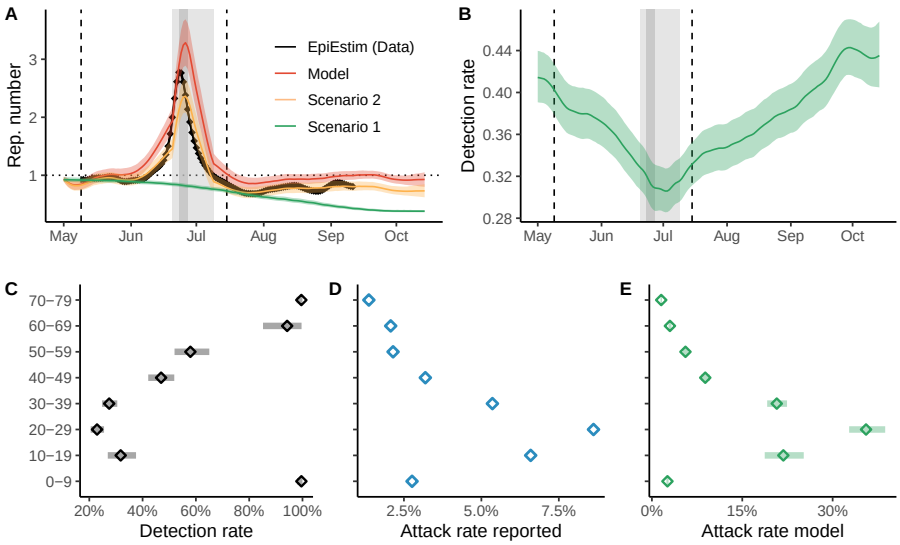


Figure D.10: **(A)**: Overall reproduction number estimated from the model (red) and directly from the data via EpiEstim (black). Additionally, we show the overall reproduction number for Scenario 1 (green) and Scenario 2 (orange). The dark- and light-grey shaded area indicate the weekend of Sant Joan and the interval in which nightclubs were allowed to open, respectively. The dotted line indicates a reproduction number of 1. **(B)**: Evolution of the overall detection rate in time. The evolution is driven by changing infection patterns with respect to the age groups. **(C)**: Age-specific attack rates in the model. Markers show the median value, while horizontal bars indicate 95% CrI. **(D)**: Same to **C** but using the reported cases. **(E)**: Age-specific detection rates inferred through the model adjustment.

BIBLIOGRAPHY

1. Vos, T., Lim, S. S., Abbafati, C., Abbas, K. M., Abbasi, M., Abbasifard, M., Abbasi-Kangevari, M., Abbastabar, H., Abd-Allah, F., Abdelalim, A. *et al.* Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *The Lancet* **396,** 1204–1222 (2020).
2. Bruegel, P. *The Triumph of Death* https://www.museodelprado.es/en/the-collection/art-work/the-triumph-of-death/. [Online; accessed 24-July-2022].
3. Schwartz, M. & Perrot, A. *Robert Koch und Louis Pasteur: Duell zweier Giganten* 1st ed. (Theiss, Konrad, 2015).
4. Woolhouse, M., Scott, F., Hudson, Z., Howey, R. & Chase-Topping, M. Human viruses: discovery and emergence. *Philosophical Transactions of the Royal Society B: Biological Sciences* **367,** 2864–2871 (2012).
5. Plotkin, S. History of vaccination. *Proceedings of the National Academy of Sciences* **111,** 12283–12287 (2014).
6. Ferguson, N. Capturing human behaviour. *Nature* **446,** 733–733 (2007).
7. McNeill, W. H. *Plagues and Peoples* 1st Edition (Anchor, 1976).
8. Solé, R. & Elena, S. F. *Viruses as Complex Adaptive Systems* (Princeton University Press, 2019).
9. Willett, B. J. *et al.* SARS-CoV-2 Omicron is an immune escape variant with an altered cell entry pathway. *Nature Microbiology* **7,** 1161–1179 (2022).
10. Ripperger, S. P., Stockmaier, S. & Carter, G. G. Tracking sickness effects on social encounters via continuous proximity sensing in wild vampire bats. *Behavioral Ecology* **31,** 1296–1302 (2020).
11. Poirotte, C., Massol, F., Herbert, A., Willaume, E., Bomo, P. M., Kappeler, P. M. & Charpentier, M. J. E. Mandrills use olfaction to socially avoid parasitized conspecifics. *Science Advances* **3,** e1601721 (2017).
12. Castillo-Chavez, C., Bichara, D. & Morin, B. R. Perspectives on the role of mobility, behavior, and time scales in the spread of diseases. *Proceedings of the National Academy of Sciences* **113,** 14582–14588 (2016).
13. Spyrou, M. A. *et al.* The source of the Black Death in fourteenth-century central Eurasia. *Nature* **606,** 718–724 (2022).
14. Solé, R., Sardanyés, J. & Elena, S. F. Phase transitions in virology. *Reports on Progress in Physics. Physical Society (Great Britain)* **84** (2021).
15. Wittgenstein, L. *Philosophical investigations* (John Wiley & Sons, 2010).
16. Artime, O. & De Domenico, M. From the origin of life to pandemics: emergent phenomena in complex systems. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **380,** 20200410 (2022).
17. Solé, R. *Phase Transitions* (Princeton University Press, 2011).
18. Keeling, M. J. & Rohani, P. *Modeling Infectious Diseases in Humans and Animals* (Princeton University Press, 2008).
19. Funk, S., Salathé, M. & Jansen, V. A. Modelling the influence of human behaviour on the spread of infectious diseases: a review. *Journal of the Royal Society Interface* **7,** 1247–1256 (2010).
20. Bauch, C. T. & Galvani, A. P. Social factors in epidemiology. *Science* **342,** 47–49 (2013).

21. Wang, Z., Andrews, M. A., Wu, Z.-X., Wang, L. & Bauch, C. T. Coupled disease–behavior dynamics on complex networks: A review. *Physics of life reviews* **15,** 1–29 (2015).

22. Wang, Z., Bauch, C. T., Bhattacharyya, S., d'Onofrio, A., Manfredi, P., Perc, M., Perra, N., Salathé, M. & Zhao, D. Statistical physics of vaccination. *Physics Reports* **664,** 1–113 (2016).

23. Blower, S. & Bernoulli, D. An attempt at a new analysis of the mortality caused by smallpox and of the advantages of inoculation to prevent it. 1766. *Reviews in Medical Virology* **14,** 275–288 (2004).

24. Boylston, A. The origins of inoculation. *Journal of the Royal Society of Medicine* **105,** 309–313 (2012).

25. Dietz, K. & Heesterbeek, J. A. P. Daniel Bernoulli's epidemiological model revisited. *Mathematical Biosciences* **180,** 1–21 (2002).

26. Euler, L. *Briefwechsel mit Daniel Bernoulli - Teil II: Briefwechsel 1744–1778, Anhänge, Register* (eds Fellmann, E. A. & Mikhajlov, G. K.) (Birkhäuser, 2016).

27. Heesterbeek, H. in *Ecological Paradigms Lost* (eds Cuddington, K. & Beisner, B. E.) 81–105 (Academic Press, 2005).

28. Farr, W. On the cattle plague. *Journal of Social Sciences* (1866).

29. Brownlee, J. Historical note on Farr's theory of the epidemic. *British Medical Journal* **2,** 250–252 (1915).

30. Snow, J. On Continuous Molecular Changes, More Particularly in Their Relation to Epidemic Diseases. *Reviews of Infectious Diseases* **7,** 441–447 (1985).

31. Greenwood, M. Obituary of Sir William Hamer. *British Medical Journal* **2,** 154–155 (1936).

32. Hamer, W. H. *The Milroy Lectures on epidemic disease in England — The evidence of variability and of persistency of type* (Bedford Press, 1906).

33. Nye, E. R. & E, G. M. *Ronald Ross: Malariologist and Polymath* (Palgrave Macmillan, 1997).

34. Gani, J. in *Statisticians of the Centuries* (eds Heyde, C. C., Seneta, E., Crépel, P., Fienberg, S. E. & Gani, J.) 323–327 (Springer, 2001).

35. Ross, R. An application of the theory of probabilities to the study of a priori pathometry.—Part I. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **92,** 204–230 (1916).

36. Ross, R. & Hudson, H. P. An Application of the Theory of Probabilities to the Study of a priori Pathometry. Part II. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **93,** 212–225 (1917).

37. Ross, R. & Hudson, H. P. An application of the theory of probabilities to the study of a priori pathometry.—Part III. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **93,** 225–240 (1917).

38. McKendrick, A. G. The rise and fall of epidemics. *Paludism* **1,** 54–66 (1912).

39. Kermack, W. O. & McKendrick, A. G. A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **115,** 700–721 (1927).

40. Kermack, W. O. & McKendrick, A. G. Contributions to the mathematical theory of epidemics. II. —The problem of endemicity. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **138,** 55–83 (1932).

41. Kermack, W. O. & McKendrick, A. G. Contributions to the mathematical theory of epidemics. III.—Further studies of the problem of endemicity. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **141,** 94–122 (1933).

42. Érdi, P. & Tóth, J. *Mathematical Models of Chemical Reactions: Theory and Applications of Deterministic and Stochastic Models* (Manchester University Press, 1989).

43. Diekmann, O., Metz, J. & Heesterbeek, J. in *Epidemic Models: Their Structure and Relation to Data* (ed Mollison, D.) (Cambridge University Press, 1995).

44. McCallum, H., Barlow, N. & Hone, J. How should pathogen transmission be modelled? *Trends in Ecology & Evolution* **16,** 295–300 (2001).

45. Bjørnstad, O. N., Finkenstädt, B. F. & Grenfell, B. T. Dynamics of Measles Epidemics: Estimating Scaling of Transmission Rates Using a Time Series SIR Model. *Ecological Monographs* **72,** 169–184 (2002).

46. Hazarie, S., Soriano-Paños, D., Arenas, A., Gómez-Gardeñes, J. & Ghoshal, G. Interplay between population density and mobility in determining the spread of epidemics in cities. *Communications Physics* **4,** 1–10 (2021).

47. Nightingale, E. S. *et al.* The importance of saturating density dependence for population-level predictions of SARS-CoV-2 resurgence compared with density-independent or linearly density-dependent models, England, 23 March to 31 July 2020. *Eurosurveillance* **26,** 2001809 (2021).

48. Rader, B. *et al.* Crowding and the shape of COVID-19 epidemics. *Nature Medicine* **26,** 1829–1834 (2020).

49. Yeomans, J. M. *Statistical Mechanics of Phase Transitions* (Oxford University Press, 1993).

50. Strogatz, S. H. *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering, Second Edition* (CRC Press, 2014).

51. Heesterbeek, J. A Brief History of R0 and a Recipe for its Calculation. *Acta Biotheoretica* **50,** 189–204 (2002).

52. Lotka, A. J. The measure of net fertility. *Journal of the Washington Academy of Sciences* **15,** 469–472 (1925).

53. Dietz, K. The estimation of the basic reproduction number for infectious diseases. *Statistical Methods in Medical Research* **2,** 23–41 (1993).

54. Lotka, A. J. Analytical Note on Certain Rhythmic Relations in Organic Systems. *Proceedings of the National Academy of Sciences of the United States of America* **6,** 410–415 (1920).

55. Lotka, A. J. Quantitative Studies in Epidemiology. *Nature* **88,** 497–498 (1912).

56. Smith, D. L., Battle, K. E., Hay, S. I., Barker, C. M., Scott, T. W. & McKenzie, F. E. Ross, Macdonald, and a Theory for the Dynamics and Control of Mosquito-Transmitted Pathogens. *PLoS Pathogens* **8,** e1002588 (2012).

57. Kucharski, A. J., Russell, T. W., Diamond, C., Liu, Y., Edmunds, J., Funk, S., Eggo, R. M., Sun, F., Jit, M., Munday, J. D. *et al.* Early dynamics of transmission and control of COVID-19: a mathematical modelling study. *The Lancet Infectious Diseases* **20,** 553–558 (2020).

58. Arita, I., Wickett, J. & Fenner, F. Impact of population density on immunization programmes. *The Journal of Hygiene* **96,** 459–466 (1986).

59. Anderson, R. M. & May, R. M. *Infectious Diseases of Humans: Dynamics and Control* Reprint Edition (Oxford University Press, 1992).

60. Kenyon, C., Lynen, L., Florence, E., Caluwaerts, S., Vandenbruaene, M., Apers, L., Soentjens, P., Van Esbroeck, M. & Bottieau, E. Syphilis reinfections pose problems for syphilis diagnosis in Antwerp, Belgium - 1992 to 2012. *Euro Surveillance* **19,** 20958 (2014).

61. Vynnycky, E. & White, R. G. *An Introduction to Infectious Disease Modelling* 1st ed. (Oxford University Press, 2010).

62. He, X. *et al.* Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nature Medicine* **26,** 672–675 (2020).

63. Milne, G., Hames, T., Scotton, C., Gent, N., Johnsen, A., Anderson, R. M. & Ward, T. Does infection with or vaccination against SARS-CoV-2 lead to lasting immunity? *The Lancet Respiratory Medicine* **9,** 1450–1466 (2021).

64. Petrova, V. N. & Russell, C. A. The evolution of seasonal influenza viruses. *Nature Reviews Microbiology* **16,** 47–60 (2018).

65. Keeling, M. J. & Eames, K. T. Networks and epidemic models. *Journal of the Royal Society Interface* **2,** 295–307 (2005).

66. Bartlett, M. S. Deterministic and Stochastic Models for Recurrent Epidemics. *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Volume 4: Contributions to Biology and Problems of Health* **3.4,** 81–110 (1956).

67. London, W. P. & Yorke, J. A. Recurrent outbreaks of measles, chickenpox and mumps. I. Seasonal variation in contact rates. *American Journal of Epidemiology* **98,** 453–468 (1973).

68. Sutherland, I. & Fayers, P. M. Effect of measles vaccination on incidence of measles in the community. *British Medical Journal* **1,** 698–702 (1971).

69. Yorke, J. A., Nathanson, N., Pianigiani, G. & Martin, J. Seasonality and the requirements for perpetuation and eradication of viruses in populations. *American Journal of Epidemiology* **109,** 103–123 (1979).

70. Schenzle, D. An age-structured model of pre- and post-vaccination measles transmission. *IMA journal of mathematics applied in medicine and biology* **1,** 169–191 (1984).

71. Dietz, K. Transmission and control of arbovirus diseases. *In Proceedings of the Society for Industrial and Applied Mathematics, Epidemiology: Philadelphia* (1974).

72. Prem, K., Cook, A. R. & Jit, M. Projecting social contact matrices in 152 countries using contact surveys and demographic data. *PLOS Computational Biology* **13,** e1005697 (2017).

73. Newman, M. E. J. Modularity and community structure in networks. *Proceedings of the National Academy of Sciences* **103,** 8577–8582 (2006).

74. Wallinga, J., Teunis, P. & Kretzschmar, M. Using Data on Social Contacts to Estimate Age-specific Transmission Parameters for Respiratory-spread Infectious Agents. *American Journal of Epidemiology* **164,** 936–944 (2006).

75. Mossong, J. *et al.* Social Contacts and Mixing Patterns Relevant to the Spread of Infectious Diseases. *PLOS Medicine* **5,** e74 (2008).

76. Prem, K., Zandvoort, K. v., Klepac, P., Eggo, R. M., Davies, N. G., Cook, A. R. & Jit, M. Projecting contact matrices in 177 geographical regions: An update and comparison with empirical data for the COVID-19 era. *PLOS Computational Biology* **17,** e1009098 (2021).

77. Mistry, D. *et al.* Inferring high-resolution human mixing patterns for disease modeling. *Nature Communications* **12,** 323 (2021).

78. Cooke, K. L. & Yorke, J. A. Some equations modelling growth processes and gonorrhea epidemics. *Mathematical Biosciences* **16,** 75–101 (1973).

79. Hethcote, H. W. & Yorke, J. A. *Gonorrhea Transmission Dynamics and Control* (Springer, 1984).

80. Anderson, R. M., Medley, G. F., May, R. M. & Johnson, A. M. A preliminary study of the transmission dynamics of the human immunodeficiency virus (HIV), the causative agent of AIDS. *IMA journal of mathematics applied in medicine and biology* **3,** 229–263 (1986).

81. Garnett, G. P. & Anderson, R. M. Factors controlling the spread of HIV in heterosexual communities in developing countries: patterns of mixing between different age and sexual activity classes. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **342,** 137–159 (1993).

82. Newman, M. *Networks: An Introduction* (Oxford University Press, 2010).

83. Hansson, D., Strömdahl, S., Leung, K. Y. & Britton, T. Introducing pre-exposure prophylaxis to prevent HIV acquisition among men who have sex with men in Sweden: insights from a mathematical pair formation model. *BMJ Open* **10,** e033852 (2020).

84. Lajmanovich, A. & Yorke, J. A. A deterministic model for gonorrhea in a nonhomogeneous population. *Mathematical Biosciences* **28,** 221–236 (1976).

85. Jacquez, J. A., Simon, C. P., Koopman, J., Sattenspiel, L. & Perry, T. Modeling and analyzing HIV transmission: the effect of contact patterns. *Mathematical Biosciences* **92,** 119–199 (1988).

86. Busenberg, S. & Castillo-Chavez, C. A General Solution of the Problem of Mixing of Subpopulations and its Application to Risk- and Age-Structured Epidemic Models for the Spread of AIDS. *Mathematical Medicine and Biology: A Journal of the IMA* **8,** 1–29 (1991).

87. Garnett, G. P. & Anderson, R. M. Sexually transmitted diseases and sexual behavior: insights from mathematical models. *The Journal of Infectious Diseases* **174,** S150–161 (1996).

88. Ghani, A. C., Swinton, J. & Garnett, G. P. The role of sexual partnership networks in the epidemiology of gonorrhea. *Sexually Transmitted Diseases* **24,** 45–56 (1997).

89. Suen, S.-c., Goldhaber-Fiebert, J. D. & Brandeau, M. L. Risk Stratification in Compartmental Epidemic Models: Where to Draw the Line? *Journal of Theoretical Biology* **428,** 1–17 (2017).

90. Klovdahl, A. S. Social networks and the spread of infectious diseases: The AIDS example. *Social Science & Medicine* **21,** 1203–1216 (1985).

91. May, R. M. & Anderson, R. M. Transmission dynamics of HIV infection. *Nature* **326,** 137–142 (1987).

92. Latkin, C., Mandell, W., Oziemkowska, M., Celentano, D., Vlahov, D., Ensminger, M. & Knowlton, A. Using social network analysis to study patterns of drug use among urban drug users at high risk for HIV/AIDS. *Drug and Alcohol Dependence* **38,** 1–9 (1995).

93. Morris, M., Zavisca, J. & Dean, L. Social and sexual networks: their role in the spread of HIV/AIDS among young gay men. *AIDS education and prevention: official publication of the International Society for AIDS Education* **7,** 24–35 (1995).

94. Latkin, C. A. Outreach in natural settings: the use of peer leaders for HIV prevention among injecting drug users' networks. *Public Health Reports (Washington, D.C.: 1974)* **113,** 151–159 (1998).

95. Laumann, E. O. & Youm, Y. Racial/ethnic group differences in the prevalence of sexually transmitted diseases in the United States: a network explanation. *Sexually Transmitted Diseases* **26,** 250–261 (1999).

96. Bell, D. C., Atkinson, J. S. & Carlson, J. W. Centrality measures for disease transmission networks. *Social Networks* **21,** 1–21 (1999).

97. Klovdahl, A. S. Networks and pathogens. *Sexually Transmitted Diseases* **28,** 25–28 (2001).

98. Moreno, J. L. *Who Shall Survive?* (Beacon House, 1953).

99. Leinhardt, S. *Social Networks: A Developing Paradigm* (Academic Press, 1977).

100. Scott, J. Social Network Analysis. *Sociology* **22,** 109–127 (1988).

101. Wassermann, S. & Faust, K. *Social network analysis: Methods and applications* (Cambridge University Press, 1994).

102. Zachary, W. W. An Information Flow Model for Conflict and Fission in Small Groups. *Journal of Anthropological Research* **33,** 452–473 (1977).

103. Granovetter, M. S. The Strength of Weak Ties. *American Journal of Sociology* **78,** 1360–1380 (1973).

104. Cook, K. S., Emerson, R. M., Gillmore, M. R. & Yamagishi, T. The Distribution of Power in Exchange Networks: Theory and Experimental Results. *American Journal of Sociology* **89,** 275–305 (1983).

105. Coleman, J. S. *Introduction To Mathematical Sociology* (Collier - MacMillan, 1964).

106. Harary, F., Norman, R. Z. & Cartwright, D. *Structural Models: Introduction to the Theory of Directed Graphs* (John Wiley & Sons Inc, 1965).

107. Bollobas, B. *Graph Theory: An Introductory Course* (Springer, 2012).

108. Appel, K. & Haken, W. Every planar map is four colorable. Part I: Discharging. *Illinois Journal of Mathematics* **21,** 429–490 (1977).

109. Appel, K., Haken, W. & Koch, J. Every planar map is four colorable. Part II: Reducibility. *Illinois Journal of Mathematics* **21,** 491–567 (1977).

110. Friedkin, N. E. Theoretical Foundations for Centrality Measures. *American Journal of Sociology* **96,** 1478–1504 (1991).

111. Holland, P. W. & Leinhardt, S. Transitivity in Structural Models of Small Groups. *Comparative Group Studies* **2,** 107–124 (1971).

112. Holland, P. W., Laskey, K. B. & Leinhardt, S. Stochastic blockmodels: First steps. *Social Networks* **5,** 109–137 (1983).

113. Kretzschmar, M., van Duynhoven, Y. T. & Severijnen, A. J. Modeling prevention strategies for gonorrhea and chlamydia using stochastic network simulations. *American Journal of Epidemiology* **144,** 306–317 (1996).

114. Korenromp, E. L., van Vliet, C., Bakker, R., de Vlas, S. J. & Habbema, J. D. F. HIV spread and partnership reduction for different patterns of sexual behaviour - a study with the microsimulation model STDSIM. *Mathematical Population Studies* **8,** 135–173 (2000).

115. Chick, S. E., Adams, A. L. & Koopman, J. S. Analysis and simulation of a stochastic, discrete-individual model of STD transmission with partnership concurrency. *Mathematical Biosciences* **166,** 45–68 (2000).

116. Andersson, H. Epidemics in a population with social structures. *Mathematical Biosciences* **140,** 79–84 (1997).

117. Estrada, E. *The Structure of Complex Networks: Theory and Applications* (Oxford University Press, 2011).

118. Latora, V., Nicosia, V. & Russo, G. *Complex Networks: Principles, Methods and Applications* (2017).

119. Ashcroft, N. W. & Mermin, N. D. *Solid State Physics* (Holt-Saunders, 1976).

120. Bollobás, B. *Random Graphs* 2nd ed. (Cambridge University Press, 2001).

121. Watts, D. J. & Strogatz, S. H. Collective dynamics of 'small-world' networks. *Nature* **393,** 440–442 (1998).

122. Travers, J. & Milgram, S. An Experimental Study of the Small World Problem. *Sociometry* **32,** 425–443 (1969).

123. Mermin, N. & Ashcroft, N. *Solid State Physics* (Brooks/Cole, 1976).

124. Mézard, M. & Montanari, A. *Information, Physics, and Computation* Illustrated Edition (Oxford University Press, 2009).

125. Albert, R., Jeong, H. & Barabási, A.-L. Diameter of the World-Wide Web. *Nature* **401,** 130–131 (1999).

126. Liljeros, F., Edling, C. R., Amaral, L. A. N., Stanley, H. E. & Åberg, Y. The web of human sexual contacts. *Nature* **411,** 907–908 (2001).

127. Barabási, A.-L. & Albert, R. Emergence of Scaling in Random Networks. *Science* **286,** 509–512 (1999).

128. Caldarelli, G., Capocci, A., De Los Rios, P. & Muñoz, M. A. Scale-Free Networks from Varying Vertex Intrinsic Fitness. *Physical Review Letters* **89,** 258702 (2002).

129. Jackson, M. O. & Zenou, Y. in *Handbook of Game Theory with Economic Applications* (eds Young, H. P. & Zamir, S.) 95–163 (Elsevier, 2015).

130. Arenas, A., Díaz-Guilera, A., Kurths, J., Moreno, Y. & Zhou, C. Synchronization in complex networks. *Physics Reports* **469,** 93–153 (2008).

131. Pastor-Satorras, R., Castellano, C., Van Mieghem, P. & Vespignani, A. Epidemic processes in complex networks. *Reviews of Modern Physics* **87,** 925–979 (2015).

132. Danon, L., Ford, A. P., House, T., Jewell, C. P., Keeling, M. J., Roberts, G. O., Ross, J. V. & Vernon, M. C. Networks and the epidemiology of infectious disease. *Interdisciplinary Perspectives on Infectious Diseases* **2011,** 284909 (2011).

133. Brin, S. & Page, L. The anatomy of a large-scale hypertextual Web search engine. *Computer Networks and ISDN Systems. Proceedings of the Seventh International World Wide Web Conference* **30,** 107–117 (1998).

134. Guimerà, R. & Nunes Amaral, L. A. Functional cartography of complex metabolic networks. *Nature* **433,** 895–900 (2005).

135. Pagani, G. A. & Aiello, M. The Power Grid as a complex network: A survey. *Physica A: Statistical Mechanics and its Applications* **392,** 2688–2700 (2013).

136. Eames, K., Bansal, S., Frost, S. & Riley, S. Six challenges in measuring contact networks for use in modelling. *Epidemics. Challenges in Modelling Infectious Disease Dynamics* **10,** 72–77 (2015).

137. Meyers, L., Newman, M & Pourbohloul, B. Predicting epidemics on directed contact networks. *Journal of Theoretical Biology* **240,** 400–18 (2006).

138. Li, C., Wang, H. & Van Mieghem, P. Epidemic threshold in directed networks. *Physical Review E* **88,** 062802 (2013).

139. Eames, K. T. D., Read, J. M. & Edmunds, W. J. Epidemic prediction and control in weighted networks. *Epidemics* **1,** 70–76 (2009).

140. Kivelä, M., Arenas, A., Barthelemy, M., Gleeson, J. P., Moreno, Y. & Porter, M. A. Multilayer networks. *Journal of Complex Networks* **2,** 203–271 (2014).

141. Battiston, F., Cencetti, G., Iacopini, I., Latora, V., Lucas, M., Patania, A., Young, J.-G. & Petri, G. Networks beyond pairwise interactions: Structure and dynamics. *Physics Reports* **874,** 1–92 (2020).

142. Read, J. M., Eames, K. T. & Edmunds, W. J. Dynamic social networks and the implications for the spread of infectious disease. *Journal of the Royal Society Interface* **5,** 1001–1007 (2008).

143. Cattuto, C., Van den Broeck, W., Barrat, A., Colizza, V., Pinton, J. & Vespignani, A. Dynamics of Person-to-Person Interactions from Distributed RFID Sensor Networks. *PLOS ONE* **5,** e11596 (2010).

144. Salathé, M., Kazandjieva, M., Lee, J. W., Levis, P., Feldman, M. W. & Jones, J. H. A high-resolution human contact network for infectious disease transmission. *Proceedings of the National Academy of Sciences of the United States of America* **107,** 22020–22025 (2010).

145. Stehlé, J. *et al.* High-Resolution Measurements of Face-to-Face Contact Patterns in a Primary School. *PLOS ONE* **6,** e23176 (2011).

146. Isella, L. *et al.* Close Encounters in a Pediatric Ward: Measuring Face-to-Face Proximity and Mixing Patterns with Wearable Sensors. *PLOS ONE* **6,** e17144 (2011).

147. Sapiezynski, P., Stopczynski, A., Lassen, D. D. & Lehmann, S. Interaction data from the Copenhagen Networks Study. *Scientific Data* **6,** 315 (2019).

148. Crawford, F. W. *et al.* Impact of close interpersonal contact on COVID-19 incidence: Evidence from 1 year of mobile device data. *Science Advances* **8,** eabi5499 (2022).

149. Pung, R., Firth, J. A., Spurgin, L. G., Lee, V. J. & Kucharski, A. J. Using high-resolution contact networks to evaluate SARS-CoV-2 transmission and control in large-scale multi-day events. *Nature Communications* **13,** 1956 (2022).

150. Erdős, P. & Rényi, A. On Random Graphs I. *Publicationes Mathematicae Debrecen* **6,** 290–297.

151. Clauset, A., Shalizi, C. R. & Newman, M. E. J. Power-Law Distributions in Empirical Data. *SIAM Review* **51,** 661–703 (2009).

152. Faloutsos, M., Faloutsos, P. & Faloutsos, C. On power-law relationships of the internet topology. *ACM SIGCOMM computer communication review* **29,** 251–262 (1999).

153. Price, D. D. S. A general theory of bibliometric and other cumulative advantage processes. *Journal of the American Society for Information Science* **27,** 292–306 (1976).

154. Lotka, A. J. The frequency distribution of scientific productivity. *Journal of the Washington Academy of Sciences* **16,** 317–323 (1926).

155. Dorogovtsev, S. N., Mendes, J. F. & Samukhin, A. N. Structure of growing networks with preferential linking. *Physical Review Letters* **85,** 4633–4636 (2000).

156. Murray, W. H. The application of epidemiology to computer viruses. *Computers & Security* **7,** 139–145 (1988).

157. Pastor-Satorras, R. & Vespignani, A. Epidemic Spreading in Scale-Free Networks. *Physical Review Letters* **86,** 3200–3203 (2001).

158. Pastor-Satorras, R. & Vespignani, A. Epidemic dynamics and endemic states in complex networks. *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics* **63,** 066117 (2001).

159. Marro, J. & Dickman, R. *Nonequilibrium Phase Transitions in Lattice Models* (Cambridge University Press, 1999).

160. Stumpf, M. P. H. & Porter, M. A. Critical Truths About Power Laws. *Science* **335,** 665–666 (2012).

161. Broido, A. D. & Clauset, A. Scale-free networks are rare. *Nature Communications* **10,** 1017 (2019).

162. Holme, P. Rare and everywhere: Perspectives on scale-free networks. *Nature Communications* **10,** 1016 (2019).

163. Voitalov, I., van der Hoorn, P., van der Hofstad, R. & Krioukov, D. Scale-free networks well done. *Physical Review Research* **1,** 033034 (2019).

164. Serafino, M., Cimini, G., Maritan, A., Rinaldo, A., Suweis, S., Banavar, J. R. & Caldarelli, G. True scale-free networks hidden by finite size effects. *Proceedings of the National Academy of Sciences of the United States of America* **118,** e2013825118 (2021).

165. Sharkey, K. J. Deterministic epidemiological models at the individual level. *Journal of Mathematical Biology* **57,** 311–331 (2008).

166. Schwartz, N. & Stone, L. Exact epidemic analysis for the star topology. *Physical Review E* **87,** 042815 (2013).

167. Chung, F., Lu, L. & Vu, V. Spectra of random graphs with given expected degrees. *Proceedings of the National Academy of Sciences* **100,** 6313–6318 (2003).

168. Castellano, C. & Pastor-Satorras, R. Thresholds for Epidemic Spreading in Networks. *Physical Review Letters* **105,** 218701 (2010).

169. Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. & Teller, E. Equation of State Calculations by Fast Computing Machines. *The Journal of Chemical Physics* **21,** 1087–1092 (1953).

170. Kroese, D. P., Brereton, T., Taimre, T. & Botev, Z. I. Why the Monte Carlo method is so important today. *WIREs Computational Statistics* **6,** 386–392 (2014).

171. Gillespie, D. T. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of Computational Physics* **22,** 403–434 (1976).

172. Gillespie, D. T. Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry* **81,** 2340–2361 (1977).

173. Wang, Y., Chakrabarti, D., Wang, C. & Faloutsos, C. *Epidemic spreading in real networks: an eigenvalue viewpoint* in *22nd International Symposium on Reliable Distributed Systems, 2003. Proceedings.* (2003), 25–34.

174. Chakrabarti, D., Wang, Y., Wang, C., Leskovec, J. & Faloutsos, C. Epidemic thresholds in real networks. *ACM Transactions on Information and System Security* **10,** 1:1–1:26 (2008).

175. Gómez, S., Arenas, A., Borge-Holthoefer, J., Meloni, S. & Moreno, Y. Discrete-time Markov chain approach to contact-based disease spreading in complex networks. *EPL (Europhysics Letters)* **89,** 38009 (2010).

176. En'Ko, P. On the course of epidemics of some infectious diseases. *International Journal of Epidemiology* **18,** 749–755 (1989).

177. Gómez, S., Gómez-Gardenes, J., Moreno, Y. & Arenas, A. Nonperturbative heterogeneous mean-field approach to epidemic spreading in complex networks. *Physical Review E* **84,** 036105 (2011).

178. Simon, P. L., Taylor, M. & Kiss, I. Z. Exact epidemic models on graphs using graph-automorphism driven lumping. *Journal of Mathematical Biology* **62,** 479–508 (2011).

179. Mieghem, P., Omic, J. & Kooij, R. Virus Spread in Networks. *Networking, IEEE/ACM Transactions on* **17,** 1 –14 (2009).

180. Matamalas, J. T., Arenas, A. & Gómez, S. Effective approach to epidemic containment using link equations in complex networks. *Science Advances* **4,** eaau4212 (2018).

181. Matamalas, J. T., Gómez, S. & Arenas, A. Abrupt phase transition of epidemic spreading in simplicial complexes. *Physical Review Research* **2,** 012049 (2020).

182. Burgio, G., Arenas, A., Gómez, S. & Matamalas, J. T. Network clique cover approximation to analyze complex contagions through group interactions. *Communications Physics* **4,** 1–10 (2021).

183. Iacopini, I., Petri, G., Barrat, A. & Latora, V. Simplicial models of social contagion. *Nature Communications* **10,** 2485 (2019).

184. Valdano, E., Ferreri, L., Poletto, C. & Colizza, V. Analytical Computation of the Epidemic Threshold on Temporal Networks. *Physical Review X* **5,** 021005 (2015).

185. Valdano, E., Fiorentin, M. R., Poletto, C. & Colizza, V. Epidemic Threshold in Continuous-Time Evolving Networks. *Physical Review Letters* **120,** 068302.

186. Keeling, M. J. The effects of local spatial structure on epidemiological invasions. *Proceedings of the Royal Society B: Biological Sciences* **266,** 859–867 (1999).

187. Gillespie, C. S. Moment-closure approximations for mass-action models. *IET Systems Biology* **3,** 52–58 (2009).

188. Moore, C. & Newman, M. E. Epidemics and percolation in small-world networks. *Physical Review. E, Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics* **61,** 5678–5682 (2000).

189. Newman, M. E. J. Spread of epidemic disease on networks. *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics* **66,** 016128 (2002).

190. Kenah, E. & Robins, J. M. Second look at the spread of epidemics on networks. *Physical review. E, Statistical, nonlinear, and soft matter physics* **76,** 036113 (2007).

191. Aleta, A. *et al.* Modelling the impact of testing, contact tracing and household quarantine on second waves of COVID-19. *Nature Human Behaviour* **4,** 964–971 (2020).

192. Pullano, G., Pinotti, F., Valdano, E., Boëlle, P.-Y., Poletto, C. & Colizza, V. Novel coronavirus (2019-nCoV) early-stage importation risk to Europe, January 2020. *Euro Surveillance* **25** (2020).

193. Gilbert, M. *et al.* Preparedness and vulnerability of African countries against importations of COVID-19: a modelling study. *The Lancet* **395,** 871–877 (2020).

194. Chinazzi, M. *et al.* The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science* **368,** 395–400 (2020).

195. Davis, J. T. *et al.* Cryptic transmission of SARS-CoV-2 and the first COVID-19 wave. *Nature* **600,** 127–132 (2021).

196. Balcan, D., Colizza, V., Gonçalves, B., Hu, H., Ramasco, J. J. & Vespignani, A. Multiscale mobility networks and the spatial spreading of infectious diseases. *Proceedings of the National Academy of Sciences* **106,** 21484–21489 (2009).

197. Brockmann, D. & Helbing, D. The Hidden Geometry of Complex, Network-Driven Contagion Phenomena. *Science* **342,** 1337–1342 (2013).

198. Levins, R. Some Demographic and Genetic Consequences of Environmental Heterogeneity for Biological Control1. *Bulletin of the Entomological Society of America* **15,** 237–240 (1969).

199. Hanski, I. *Metapopulation Ecology* (Oxford University Press, 1999).

200. Hethcote, H. W. An immunization model for a heterogeneous population. *Theoretical Population Biology* **14,** 338–349 (1978).

201. Keeling, M. J., Bjørnstad, O. N. & Grenfell, B. T. in *Ecology, Genetics and Evolution of Metapopulations* (eds Hanski, I. & Gaggiotti, O. E.) 415–445 (Academic Press, 2004).

202. Belik, V., Geisel, T. & Brockmann, D. Natural human mobility patterns and spatial spread of infectious diseases. *Physical Review X* **1,** 011001 (2011).

203. Wesolowski, A., Eagle, N., Tatem, A. J., Smith, D. L., Noor, A. M., Snow, R. W. & Buckee, C. O. Quantifying the impact of human mobility on malaria. *Science* **338,** 267–270 (2012).

204. Tizzoni, M., Bajardi, P., Decuyper, A., Kon Kam King, G., Schneider, C. M., Blondel, V., Smoreda, Z., González, M. C. & Colizza, V. On the use of human mobility proxies for modeling epidemics. *PLOS Computational Biology* **10,** e1003716 (2014).

205. Finger, F., Genolet, T., Mari, L., de Magny, G. C., Manga, N. M., Rinaldo, A. & Bertuzzo, E. Mobile phone data highlights the role of mass gatherings in the spreading of cholera outbreaks. *Proceedings of the National Academy of Sciences* **113,** 6421–6426 (2016).

206. Oliver, N. *et al.* Mobile phone data for informing public health actions across the COVID-19 pandemic life cycle. *Science Advances* **6,** eabc0764 (2020).

207. Song, C., Qu, Z., Blumm, N. & Barabási, A.-L. Limits of predictability in human mobility. *Science* **327,** 1018–1021 (2010).

208. Gómez-Gardeñes, J., Soriano-Panos, D. & Arenas, A. Critical regimes driven by recurrent mobility patterns of reaction–diffusion processes in networks. *Nature Physics* **14,** 391–395 (2018).

209. Granell, C. & Mucha, P. J. Epidemic spreading in localized environments with recurrent mobility patterns. *Physical Review E* **97,** 052302 (2018).

210. Valgañón, P., Soriano-Paños, D., Arenas, A. & Gómez-Gardeñes, J. Contagion–diffusion processes with recurrent mobility patterns of distinguishable agents. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **32,** 043102 (2022).

211. Diekmann, O., Heesterbeek, J. A. P. & Metz, J. A. On the definition and the computation of the basic reproduction ratio R 0 in models for infectious diseases in heterogeneous populations. *Journal of Mathematical Biology* **28,** 365–382 (1990).

212. Van den Driessche, P. & Watmough, J. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Mathematical Biosciences* **180,** 29–48 (2002).

213. Van den Driessche, P. & Watmough, J. in *Mathematical Epidemiology* (eds Brauer, F., van den Driessche, P. & Wu, J.) 159–178 (Springer, 2008).

214. Diekmann, O., Heesterbeek, J. & Roberts, M. G. The construction of next-generation matrices for compartmental epidemic models. *Journal of the Royal Society Interface* **7,** 873–885 (2010).

215. Maier, B. F., Wiedermann, M., Burdinski, A., Klamser, P., Jenny, M. A., Betsch, C. & Brockmann, D. Germany's current COVID-19 crisis is mainly driven by the unvaccinated. *medRxiv 2021.11.24.21266831* (2021).

216. Fraser, C. Estimating individual and household reproduction numbers in an emerging epidemic. *PLOS ONE* **2,** e758 (2007).

217. Wallinga, J. & Teunis, P. Different epidemic curves for severe acute respiratory syndrome reveal similar impacts of control measures. *American Journal of epidemiology* **160,** 509–516 (2004).

218. Cori, A., Ferguson, N. M., Fraser, C. & Cauchemez, S. A new framework and software to estimate time-varying reproduction numbers during epidemics. *American journal of epidemiology* **178,** 1505–1512 (2013).

219. Thompson, R., Stockwin, J., van Gaalen, R. D., Polonsky, J., Kamvar, Z., Demarsh, P., Dahlqwist, E., Li, S., Miguel, E., Jombart, T. *et al.* Improved inference of time-varying reproduction numbers during infectious disease outbreaks. *Epidemics* **29,** 100356 (2019).

220. Cori, A. *EpiEstim: Estimate time varying reproduction numbers from epidemic curves* `https://CRAN.R-project.org/package=EpiEstim`. R package version 2.2-3. (2020).

221. Lloyd-Smith, J. O., Schreiber, S. J., Kopp, P. E. & Getz, W. M. Superspreading and the effect of individual variation on disease emergence. *Nature* **438,** 355–359 (2005).

222. Nielsen, B. F., Simonsen, L. & Sneppen, K. COVID-19 superspreading suggests mitigation by social network modulation. *Physical Review Letters* **126,** 118301 (2021).

223. Fine, P. E. The interval between successive cases of an infectious disease. *American journal of epidemiology* **158,** 1039–1047 (2003).

224. Lipsitch, M., Cohen, T., Cooper, B., Robins, J. M., Ma, S., James, L., Gopalakrishna, G., Chew, S. K., Tan, C. C., Samore, M. H. *et al.* Transmission dynamics and control of severe acute respiratory syndrome. *science* **300,** 1966–1970 (2003).

225. Cori, A., Valleron, A.-J., Carrat, F., Tomba, G. S., Thomas, G. & Boëlle, P.-Y. Estimating influenza latency and infectious period durations using viral excretion data. *Epidemics* **4,** 132–138 (2012).

226. Jones, T. C., Biele, G., Mühlemann, B., Veith, T., Schneider, J., Beheim-Schwarzbach, J., Bleicker, T., Tesch, J., Schmidt, M. L., Sander, L. E. *et al.* Estimating infectiousness throughout SARS-CoV-2 infection course. *Science* **373,** eabi5273 (2021).

227. Svensson, Å. A note on generation times in epidemic models. *Mathematical Biosciences* **208,** 300–311 (2007).

228. Park, S. W., Sun, K., Champredon, D., Li, M., Bolker, B. M., Earn, D. J., Weitz, J. S., Grenfell, B. T. & Dushoff, J. Forward-looking serial intervals correctly link epidemic growth to reproduction numbers. *Proceedings of the National Academy of Sciences* **118,** e2011548118 (2021).

229. Diekmann, O., Othmer, H. G., Planqué, R. & Bootsma, M. C. The discrete-time Kermack–McKendrick model: A versatile and computationally attractive framework for modeling epidemics. *Proceedings of the National Academy of Sciences* **118,** e2106332118 (2021).

230. Champredon, D., Dushoff, J. & Earn, D. J. Equivalence of the Erlang-distributed SEIR epidemic model and the renewal equation. *SIAM Journal on Applied Mathematics* **78,** 3258–3278 (2018).

231. Wong, G. N., Weiner, Z. J., Tkachenko, A. V., Elbanna, A., Maslov, S. & Goldenfeld, N. Modeling COVID-19 dynamics in Illinois under nonpharmaceutical interventions. *Physical Review X* **10,** 041033 (2020).

232. Ganyani, T., Kremer, C., Chen, D., Torneri, A., Faes, C., Wallinga, J. & Hens, N. Estimating the generation interval for coronavirus disease (COVID-19) based on symptom onset data, March 2020. *Eurosurveillance* **25,** 2000257 (2020).

233. Gostic, K. M., McGough, L., Baskerville, E. B., Abbott, S., Joshi, K., Tedijanto, C., Kahn, R., Niehus, R., Hay, J. A., De Salazar, P. M. *et al.* Practical considerations for measuring the effective reproductive number, Rt. *PLOS Computational Biology* **16,** e1008409 (2020).

234. Becker, N. G. & Marschner, I. C. A method for estimating the age-specific relative risk of HIV infection from aids incidence data. *Biometrika* **80,** 165–178 (1993).

235. Goldstein, E., Dushoff, J., Ma, J., Plotkin, J. B., Earn, D. J. & Lipsitch, M. Reconstructing influenza incidence by deconvolution of daily mortality time series. *Proceedings of the National Academy of Sciences* **106,** 21825–21829 (2009).

236. Scire, J., Huisman, J. S., Grosu, A., Angst, D. C., Li, J., Maathuis, M. H., Bonhoeffer, S. & Stadler, T. estimateR: An R package to estimate and monitor the effective reproductive number. *medRxiv 2022.06.30.22277095* (2022).

237. Mollison, D. & Denis, M. *Epidemic models: their structure and relation to data* (Cambridge University Press, 1995).

238. Toni, T., Welch, D., Strelkowa, N., Ipsen, A. & Stumpf, M. P. Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. *Journal of the Royal Society Interface* **6,** 187–202 (2009).

239. Chowell, G. Fitting dynamic models to epidemic outbreaks with quantified uncertainty: A primer for parameter uncertainty, identifiability, and forecasts. *Infectious Disease Modelling* **2,** 379–398 (2017).

240. Cole, S. R., Chu, H. & Greenland, S. Maximum likelihood, profile likelihood, and penalized likelihood: a primer. *American journal of epidemiology* **179,** 252–260 (2014).

241. O'Neill, P. D. A tutorial introduction to Bayesian inference for stochastic epidemic models using Markov chain Monte Carlo methods. *Mathematical Biosciences* **180,** 103–114 (2002).

242. Van de Schoot, R., Depaoli, S., King, R., Kramer, B., Märtens, K., Tadesse, M. G., Vannucci, M., Gelman, A., Veen, D., Willemsen, J. *et al.* Bayesian statistics and modelling. *Nature Reviews Methods Primers* **1,** 1–26 (2021).

243. Andrieu, C., De Freitas, N., Doucet, A. & Jordan, M. I. An introduction to MCMC for machine learning. *Machine learning* **50,** 5–43 (2003).

244. Hastings, W. K. Monte Carlo Sampling Methods Using Markov Chains and Their Applications. *Biometrika* **57,** 97–109 (1970).

245. Betancourt, M. *A Conceptual Introduction to Hamiltonian Monte Carlo* (2017).

246. Brooks, S., Gelman, A., Jones, G. & Meng, X.-L. *Handbook of Markov Chain Monte Carlo* (CRC press, 2011).

247. Lloyd-Smith, J. O. Maximum likelihood estimation of the negative binomial dispersion parameter for highly overdispersed data, with applications to infectious diseases. *PLOS ONE* **2,** e180 (2007).

248. King, A. A., Domenech de Cellès, M., Magpantay, F. M. & Rohani, P. Avoidable errors in the modelling of outbreaks of emerging pathogens, with special reference to Ebola. *Proceedings of the Royal Society B: Biological Sciences* **282,** 20150347 (2015).

249. Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P. & Riddell, A. Stan: A probabilistic programming language. *Journal of statistical software* **76** (2017).

250. Stan Development Team. *RStan: the R interface to Stan* R package version 2.17.3. (2018).

251. Stan Development Team. *The Stan Core Library* Version 2.18.0. (2018).

252. Ge, H., Xu, K. & Ghahramani, Z. *Turing: a language for flexible probabilistic inference* in *International Conference on Artificial Intelligence and Statistics, AISTATS 2018, 9-11 April 2018, Playa Blanca, Lanzarote, Canary Islands, Spain* (2018), 1682–1690.

253. Bezanson, J., Edelman, A., Karpinski, S. & Shah, V. B. Julia: A fresh approach to numerical computing. *SIAM review* **59,** 65–98 (2017).

254. Foreman-Mackey, D., Hogg, D. W., Lang, D. & Goodman, J. emcee: The MCMC Hammer. *PASP* **125,** 306 (2013).

255. Gelman, A. & Rubin, D. B. Inference from iterative simulation using multiple sequences. *Statistical science,* 457–472 (1992).

256. Grinsztajn, L., Semenova, E., Margossian, C. C. & Riou, J. Bayesian workflow for disease transmission modeling in Stan. *Statistics in Medicine* **40,** 6209–6234 (2021).

257. Funk, S., Bansal, S., Bauch, C. T., Eames, K. T., Edmunds, W. J., Galvani, A. P. & Klepac, P. Nine challenges in incorporating the dynamics of behaviour in infectious diseases models. *Epidemics* **10,** 21–25 (2015).

258. Tognotti, E. Lessons from the history of quarantine, from plague to influenza A. *Emerging Infectious Diseases* **19,** 254 (2013).

259. Ahituv, A., Hotz, V. J. & Philipson, T. The responsiveness of the demand for condoms to the local prevalence of AIDS. *Journal of Human Resources,* 869–897 (1996).

260. SteelFisher, G. K., Blendon, R. J., Bekheit, M. M. & Lubell, K. The public's response to the 2009 H1N1 influenza pandemic. *New England Journal of Medicine* **362,** e65 (2010).

261. Neumann, J. v. Zur Theorie der Gesellschaftsspiele. *Mathematische annalen* **100,** 295–320 (1928).

262. Von Neumann, J. & Morgenstern, O. *Theory of Games and Economic Behavior* (Princeton University Press, 1947).

263. Hargreaves-Heap, S. & Varoufakis, Y. *Game Theory: A Critical Introduction* (Routledge, 1995).

264. Schelling, T. C. *Arms and Influence* (Yale University Press, 1966).

265. Philipson, T. Private vaccination and public health: an empirical examination for US measles. *Journal of Human Resources,* 611–630 (1996).

266. Bauch, C. T. & Earn, D. J. Vaccination and the theory of games. *Proceedings of the National Academy of Sciences* **101,** 13391–13394 (2004).

267. Smith, J. M. *Evolution and the Theory of Games* (Cambridge University Press, 1982).

268. Taylor, P. D. & Jonker, L. B. Evolutionary stable strategies and game dynamics. *Mathematical Biosciences* **40,** 145–156 (1978).

269. Nowak, M. A. *Evolutionary Dynamics: Exploring the Equations of Life* (Harvard University Press, 2006).

270. Bauch, C. T. Imitation dynamics predict vaccinating behaviour. *Proceedings of the Royal Society B: Biological Sciences* **272,** 1669–1675 (2005).

271. Fu, F., Rosenbloom, D. I., Wang, L. & Nowak, M. A. Imitation dynamics of vaccination behaviour on social networks. *Proceedings of the Royal Society B: Biological Sciences* **278,** 42–49 (2011).

272. Cardillo, A., Reyes-Suárez, C., Naranjo, F. & Gómez-Gardees, J. Evolutionary vaccination dilemma in complex networks. *Physical Review E* **88,** 032803 (2013).

273. Funk, S., Gilad, E., Watkins, C. & Jansen, V. A. The spread of awareness and its impact on epidemic outbreaks. *Proceedings of the National Academy of Sciences* **106,** 6872–6877 (2009).

274. Perra, N., Balcan, D., Gonçalves, B. & Vespignani, A. Towards a characterization of behavior-disease models. *PLOS ONE* **6,** e23084 (2011).

275. d'Onofrio, A. & Manfredi, P. Information-related changes in contact patterns may trigger oscillations in the endemic prevalence of infectious diseases. *Journal of Theoretical Biology* **256,** 473–478 (2009).

276. Cao, L. Infection dynamics in structured populations with disease awareness based on neighborhood contact history. *The European Physical Journal B* **87,** 1–10 (2014).

277. Moinet, A., Pastor-Satorras, R. & Barrat, A. Effect of risk perception on epidemic spreading in temporal networks. *Physical Review E* **97,** 012313 (2018).

278. Funk, S., Gilad, E & Jansen, V. A. Endemic disease, awareness, and local behavioural response. *Journal of Theoretical Biology* **264,** 501–509 (2010).

279. Granell, C., Gómez, S. & Arenas, A. Dynamical interplay between awareness and epidemic spreading in multiplex networks. *Physical Review Letters* **111,** 128701 (2013).

280. Granell, C., Gómez, S. & Arenas, A. Competing spreading processes on multiplex networks: awareness and epidemics. *Physical Review E* **90,** 012808 (2014).

281. Centola, D. *How Behavior Spreads* (Princeton University Press, 2018).

282. Gross, T., D'Lima, C. J. D. & Blasius, B. Epidemic dynamics on an adaptive network. *Physical Review Letters* **96,** 208701 (2006).

283. Shaw, L. B. & Schwartz, I. B. Fluctuating epidemics on adaptive networks. *Physical Review E* **77,** 066101 (2008).

284. Marceau, V., Noël, P.-A., Hébert-Dufresne, L., Allard, A. & Dubé, L. J. Adaptive networks: Coevolution of disease and topology. *Physical Review E* **82,** 036116 (2010).

285. Sherborne, N, Blyuss, K. & Kiss, I. Bursting endemic bubbles in an adaptive network. *Physical Review E* **97,** 042306 (2018).

286. Finkenstädt, B. F. & Grenfell, B. T. Time series modelling of childhood diseases: a dynamical systems approach. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* **49,** 187–205 (2000).

287. Pitzer, V. E., Viboud, C., Simonsen, L., Steiner, C., Panozzo, C. A., Alonso, W. J., Miller, M. A., Glass, R. I., Glasser, J. W., Parashar, U. D. *et al.* Demographic variability, vaccination, and the spatiotemporal dynamics of rotavirus epidemics. *Science* **325,** 290–294 (2009).

288. Shaman, J. & Karspeck, A. Forecasting seasonal outbreaks of influenza. *Proceedings of the National Academy of Sciences* **109,** 20425–20430 (2012).

289. Łuksza, M. & Lässig, M. A predictive fitness model for influenza. *Nature* **507,** 57–61 (2014).

290. Fanget, N. Pertussis: a tale of two vaccines. *Nature - Milestones in Vaccines* (2020).

291. Althouse, B. M. & Hébert-Dufresne, L. Epidemic cycles driven by host behaviour. *Journal of the Royal Society Interface* **11,** 20140575 (2014).

292. Grassly, N. C., Fraser, C. & Garnett, G. P. Host immunity and synchronized epidemics of syphilis across the United States. *Nature* **433,** 417–421 (2005).

293. Breban, R., Supervie, V., Okano, J. T., Vardavas, R. & Blower, S. Is there any evidence that syphilis epidemics cycle? *The Lancet Infectious Diseases* **8,** 577–581 (2008).

294. Dong, E., Du, H. & Gardner, L. An interactive web-based dashboard to track COVID-19 in real time. *The Lancet Infectious Diseases* **20,** 533–534 (2020).

295. Merow, C. & Urban, M. C. Seasonality and uncertainty in global COVID-19 growth rates. *Proceedings of the National Academy of Sciences* **117,** 27456–27464 (2020).

296. Fontal, A., Bouma, M. J., San-José, A., López, L., Pascual, M. & Rodó, X. Climatic signatures in the different COVID-19 pandemic waves across both hemispheres. *Nature Computational Science* **1,** 655–665 (2021).

297. Pullano, G., Di Domenico, L., Sabbatini, C. E., Valdano, E., Turbelin, C., Debin, M., Guerrisi, C., Kengne-Kuetche, C., Souty, C., Hanslik, T. *et al.* Underdetection of cases of COVID-19 in France threatens epidemic control. *Nature* **590,** 134–139 (2021).

298. Hodcroft, E. B., Zuber, M., Nadeau, S., Vaughan, T. G., Crawford, K. H., Althaus, C. L., Reichmuth, M. L., Bowen, J. E., Walls, A. C., Corti, D. *et al.* Spread of a SARS-CoV-2 variant through Europe in the summer of 2020. *Nature* **595,** 707–712 (2021).

299. Mazzoli, M., Valdano, E. & Colizza, V. Projecting the COVID-19 epidemic risk in France for the summer 2021. *Journal of Travel Medicine* **28,** taab129 (2021).

300. Yan, Y., Malik, A. A., Bayham, J., Fenichel, E. P., Couzens, C. & Omer, S. B. Measuring voluntary and policy-induced social distancing behavior during the COVID-19 pandemic. *Proceedings of the National Academy of Sciences* **118,** e2008814118 (2021).

301. Chen, F., Jiang, M., Rabidoux, S. & Robinson, S. Public avoidance and epidemics: insights from an economic model. *Journal of Theoretical Biology* **278,** 107–119 (2011).

302. Poletti, P., Ajelli, M. & Merler, S. The effect of risk perception on the 2009 H1N1 pandemic influenza dynamics. *PLOS ONE* **6,** e16460 (2011).

303. Poletti, P., Ajelli, M. & Merler, S. Risk perception and effectiveness of uncoordinated behavioral responses in an emerging epidemic. *Mathematical Biosciences* **238,** 80–89 (2012).

304. Anderson, J. E., Wilson, R., Doll, L., Jones, T. S. & Barker, P. Condom use and HIV risk behaviors among US adults: Data from a national survey. *Family Planning Perspectives,* 24–28 (1999).

305. Prata, N., Morris, L., Mazive, E., Vahidnia, F. & Stehr, M. Relationship between HIV risk perception and condom use: evidence from a population-based survey in Mozambique. *International Family Planning Perspectives,* 192–200 (2006).

306. Franzen, A. & Wöhner, F. Coronavirus risk perception and compliance with social distancing measures in a sample of young adults: Evidence from Switzerland. *PLOS ONE* **16,** e0247447 (2021).

307. Wambua, J., Hermans, L., Coletti, P., Verelst, F., Willem, L., Jarvis, C. I., Gimma, A., Wong, K. L., Lajot, A., Demarest, S. *et al.* The influence of risk perceptions on close contact frequency during the SARS-CoV-2 pandemic. *Scientific Reports* **12,** 1–12 (2022).

308. Tilman, A. R., Plotkin, J. B. & Akçay, E. Evolutionary games with environmental feedbacks. *Nature Communications* **11,** 1–11 (2020).

309. Howard, J., Huang, A., Li, Z., Tufekci, Z., Zdimal, V., van der Westhuizen, H.-M., von Delft, A., Price, A., Fridman, L., Tang, L.-H. *et al.* An evidence review of face masks against COVID-19. *Proceedings of the National Academy of Sciences* **118,** e2014564118 (2021).

310. Fazio, R. H., Ruisch, B. C., Moore, C. A., Granados Samayoa, J. A., Boggs, S. T. & Ladanyi, J. T. Social distancing decreases an individual's likelihood of contracting COVID-19. *Proceedings of the National Academy of Sciences* **118,** e2023131118 (2021).

311. Johnson, W. D., O'Leary, A. & Flores, S. A. Per-partner condom effectiveness against HIV for men who have sex with men. *AIDS* **32,** 1499–1505 (2018).

312. Edlefsen, P. T. Leaky vaccines protect highly exposed recipients at a lower rate: Implications for vaccine efficacy estimation and sieve analysis. *Computational and Mathematical Methods in Medicine* **2014** (2014).

313. Steinegger, B., Cardillo, A., De Los Rios, P., Gómez-Gardeñes, J. & Arenas, A. Interplay between cost and benefits triggers nontrivial vaccination uptake. *Physical Review E* **97,** 032308 (2018).

314. Flaxman, S., Mishra, S., Gandy, A., Unwin, H. J. T., Mellan, T. A., Coupland, H., Whittaker, C., Zhu, H., Berah, T., Eaton, J. W. *et al.* Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature* **584,** 257–261 (2020).

315. Davis, J. T., Chinazzi, M., Perra, N., Mu, K., Pastore y Piontti, A., Ajelli, M., Dean, N. E., Gioannini, C., Litvinova, M., Merler, S. *et al.* Cryptic transmission of SARS-CoV-2 and the first COVID-19 wave. *Nature* **600,** 127–132 (2021).

316. Haug, N., Geyrhofer, L., Londei, A., Dervic, E., Desvars-Larrive, A., Loreto, V., Pinior, B., Thurner, S. & Klimek, P. Ranking the effectiveness of worldwide COVID-19 government interventions. *Nature Human Behaviour* **4,** 1303–1312 (2020).

317. Brauner, J. M. *et al.* Inferring the effectiveness of government interventions against COVID-19. *Science* **371,** eabd9338 (2021).

318. *Dades COVID* https://dadescovid.cat/descarregues. [Online; accessed 19-April-2021].

319. *Covid-19 Schweiz* https://www.covid19.admin.ch/de/overview. [Online; accessed 19-April-2021].

320. *Florida COVID-19 Case Line Data* https://open-fdoh.hub.arcgis.com/datasets/florida-covid19-case-line-data. [Online; accessed 19-April-2021].

321. *RKI COVID-19* https://npgeo-corona-npgeo-de.hub.arcgis.com/datasets/dd4580c810204019a7b8eb3e0b329dd6_0. [Online; accessed 19-April-2021].

322. *Statistics Denmark* https://www.statistikbanken.dk/statbank5a/SelectVarVal/Define.asp?Maintable=SMIT2&PLanguage=1. [Online; accessed 19-April-2021].

323. *Aragón Open Data* https://opendata.aragon.es/datos/catalogo/dataset/publicaciones-y-anuncios-relacionados-con-el-coronavirus-en-aragon. [Online; accessed 19-April-2021].

324. *COVID-19 in Belgium* https://data.world/liz-friedman/covid-19-in-belgium. [Online; accessed 19-April-2021].

325. *GOV.UK Coronavirus (COVID-19) in the UK* https://coronavirus.data.gov.uk/details/download. [Online; accessed 19-April-2021].

326. O'Driscoll, M., Ribeiro Dos Santos, G., Wang, L., Cummings, D. A. T., Azman, A. S., Paireau, J., Fontanet, A., Cauchemez, S. & Salje, H. Age-specific mortality and immunity patterns of SARS-CoV-2. *Nature* **590,** 140–145 (2021).

327. Pastor-Barriuso, R. *et al.* Infection fatality risk for SARS-CoV-2 in community dwelling population of Spain: nationwide seroepidemiological study. *BMJ (Clinical research ed.)* **371,** m4509 (2020).

328. Davies, N. G. *et al.* Age-dependent effects in the transmission and control of COVID-19 epidemics. *Nature Medicine* **26,** 1205–1211 (2020).

329. Pullano, G., Valdano, E., Scarpa, N., Rubrichi, S. & Colizza, V. Evaluating the effect of demographic factors, socioeconomic factors, and risk aversion on mobility during the COVID-19 epidemic in France under lockdown: a population-based study. *The Lancet Digital Health* **2,** e638–e649 (2020).

330. Chan, H. F., Skali, A., Savage, D. A., Stadelmann, D. & Torgler, B. Risk attitudes and human mobility during the COVID-19 pandemic. *Scientific Reports* **10,** 1–13 (2020).

331. Gauvin, L., Bajardi, P., Pepe, E., Lake, B., Privitera, F. & Tizzoni, M. Socio-economic determinants of mobility responses during the first wave of COVID-19 in Italy: from provinces to neighbourhoods. *Journal of the Royal Society Interface* **18,** 20210092 (2021).

332. *Santé publique France. Covid-19: une enquête pour suivre l'évolution des comportements et de la santé mentale pendant l'épidémie* https://www.santepubliquefrance.fr/etudes-et-enquetes/covid-19-une-enquete-pour-suivre-l-evolution-des-comportements-et-de-la-sante-mentale-pendant-l-epidemie. [Online; accessed 24-July-2022].

333. Bíró, A., Branyiczki, R. & Elek, P. Time patterns of precautionary health behaviours during an easing phase of the COVID-19 pandemic in Europe. *European Journal of Ageing*, 1–12 (2021).

334. Eichenbaum, M., Godinho de Matos, M., Lima, F., T. Rebelo, S. & Trabandt, M. How Do People Respond to Small Probability Events with Large, Negative Consequences? *CEPR Discussion Paper* **No. DP15373** (2020).

335. Mena, G. E., Martinez, P. P., Mahmud, A. S., Marquet, P. A., Buckee, C. O. & Santillana, M. Socioeconomic status determines COVID-19 incidence and related mortality in Santiago, Chile. *Science* **372,** eabg5298 (2021).

336. Nordström, P., Ballin, M. & Nordström, A. Risk of SARS-CoV-2 reinfection and COVID-19 hospitalisation in individuals with natural and hybrid immunity: a retrospective, total population cohort study in Sweden. *The Lancet Infectious Diseases* **22,** 781–790 (2022).

337. Kutter, J. S., de Meulder, D., Bestebroer, T. M., Lexmond, P., Mulders, A., Richard, M., Fouchier, R. A. & Herfst, S. SARS-CoV and SARS-CoV-2 are transmitted through the air between ferrets over more than one meter distance. *Nature Communications* **12,** 1–8 (2021).

338. Machida, M., Nakamura, I., Saito, R., Nakaya, T., Hanibuchi, T., Takamiya, T., Odagiri, Y., Fukushima, N., Kikuchi, H., Kojima, T. *et al.* Adoption of personal protective measures by ordinary citizens during the COVID-19 outbreak in Japan. *International Journal of Infectious Diseases* **94,** 139–144 (2020).

339. Lasser, J., Sorger, J., Richter, L., Thurner, S., Schmid, D. & Klimek, P. Assessing the impact of SARS-CoV-2 prevention measures in Austrian schools using agent-based simulations and cluster tracing data. *Nature Communications* **13,** 1–17 (2022).

340. Poletti, P., Caprile, B., Ajelli, M., Pugliese, A. & Merler, S. Spontaneous behavioural changes in response to epidemics. *Journal of Theoretical Biology* **260,** 31–40 (2009).

341. Weitz, J. S., Park, S. W., Eksin, C. & Dushoff, J. Awareness-driven behavior changes can shift the shape of epidemics away from peaks and toward plateaus, shoulders, and oscillations. *Proceedings of the National Academy of Sciences* **117,** 32764–32771 (2020).

342. Tkachenko, A. V., Maslov, S., Wang, T., Elbana, A., Wong, G. N. & Goldenfeld, N. Stochastic social behavior coupled to COVID-19 dynamics leads to waves, plateaus, and an endemic state. *Elife* **10,** e68341 (2021).

343. Ye, M., Zino, L., Rizzo, A. & Cao, M. Game-theoretic modeling of collective decision making during epidemics. *Physical Review E* **104,** 024314 (2021).

344. Khazaei, H., Paarporn, K., Garcia, A. & Eksin, C. *Disease spread coupled with evolutionary social distancing dynamics can lead to growing oscillations* in *2021 60th IEEE Conference on Decision and Control (CDC)* (2021), 4280–4286.

345. Manrubia, S. & Zanette, D. H. Individual risk-aversion responses tune epidemics to critical transmissibility (R= 1). *Royal Society Open Science* **9,** 211667 (2022).

346. Khanjanianpak, M., Azimi-Tafreshi, N., Arenas, A. & Gómez-Gardeñes, J. Emergence of protective behaviour under different risk perceptions to disease spreading. *Philosophical Transactions of the Royal Society A* **380,** 20200412 (2022).

347. d'Onofrio, A., Manfredi, P. & Salinelli, E. Vaccinating behaviour, information, and the dynamics of SIR vaccine preventable diseases. *Theoretical Population Biology* **71,** 301–317 (2007).

348. Tomasello, M. *The Cultural Origins of Human Cognition* Reprint Edition (Harvard University Press, 2001).

349. Tomasello, M. *A Natural History of Human Morality* (Harvard University Press, 2016).

350. Abbott, A. & Pearson, H. Fear of human pandemic grows as bird flu sweeps through Asia. *Nature* **427,** 472 (2004).

351. Galvani, A. P., Pandey, A., Fitzpatrick, M. C., Medlock, J. & Gray, G. E. Defining control of HIV epidemics. *The Lancet HIV* **5,** e667–e670 (2018).

352. Jacob, S. T., Crozier, I., Fischer, W. A., Hewlett, A., Kraft, C. S., Vega, M.-A. d. L., Soka, M. J., Wahl, V., Griffiths, A., Bollinger, L. *et al.* Ebola virus disease. *Nature Reviews Disease Primers* **6,** 1–31 (2020).

353. Elmokashfi, A., Sundnes, J., Kvalbein, A., Naumova, V., Reinemo, S.-A., Florvaag, P. M., Stensland, H. K. & Lysne, O. Nationwide rollout reveals efficacy of epidemic control through digital contact tracing. *Nature Communications* **12,** 1–8 (2021).

354. Hadley, L., Challenor, P., Dent, C., Isham, V., Mollison, D., Robertson, D. A., Swallow, B. & Webb, C. R. Challenges on the interaction of models and policy for pandemic control. *Epidemics* **37,** 100499 (2021).

355. Kretzschmar, M. E., Ashby, B., Fearon, E., Overton, C. E., Panovska-Griffiths, J., Pellis, L., Quaife, M., Rozhnova, G., Scarabel, F., Stage, H. B. *et al.* Challenges for modelling interventions for future pandemics. *Epidemics* **38,** 100546 (2022).

356. Marion, G., Hadley, L., Isham, V., Mollison, D., Panovska-Griffiths, J., Pellis, L., Tomba, G. S., Scarabel, F., Swallow, B., Trapman, P. *et al.* Modelling: understanding pandemics and how to control them. *Epidemics,* 100588 (2022).

357. McPherson, M., Smith-Lovin, L. & Cook, J. M. Birds of a feather: Homophily in social networks. *Annual Review of Sociology* **27,** 415–444 (2001).

358. Kossinets, G. & Watts, D. J. Empirical analysis of an evolving social network. *Science* **311,** 88–90 (2006).

359. Kossinets, G. & Watts, D. J. Origins of homophily in an evolving social network. *American Journal of Sociology* **115,** 405–450 (2009).

360. Newman, M. E. Mixing patterns in networks. *Physical Review E* **67,** 026126 (2003).

361. Cinelli, M., Peel, L., Iovanella, A. & Delvenne, J.-C. Network constraints on the mixing patterns of binary node metadata. *Physical Review E* **102,** 062310 (2020).

362. Karsai, M., Kivelä, M., Pan, R. K., Kaski, K., Kertész, J., Barabási, A.-L. & Saramäki, J. Small but slow world: How network topology and burstiness slow down spreading. *Physical Review E* **83,** 025102 (2011).

363. Nadini, M., Sun, K., Ubaldi, E., Starnini, M., Rizzo, A. & Perra, N. Epidemic spreading in modular time-varying networks. *Scientific Reports* **8,** 1–11 (2018).

364. Parker, A. A., Staggs, W., Dayan, G. H., Ortega-Sánchez, I. R., Rota, P. A., Lowe, L., Boardman, P., Teclaw, R., Graves, C. & LeBaron, C. W. Implications of a 2005 measles outbreak in Indiana for sustained elimination of measles in the United States. *New England Journal of Medicine* **355,** 447–455 (2006).

365. MacDonald, N. E. & on Vaccine Hesitancy, S. W. G. Vaccine hesitancy: Definition, scope and determinants. *Vaccine* **33,** 4161–4164 (2015).

366. Cascini, F., Pantovic, A., Al-Ajlouni, Y., Failla, G. & Ricciardi, W. Attitudes, acceptance and hesitancy among the general population worldwide to receive the COVID-19 vaccines and their contributing factors: A systematic review. *EClinicalMedicine* **40,** 101113 (2021).

367. Moreno López, J. A., García, B. A., Bentkowski, P., Bioglio, L., Pinotti, F., Boëlle, P.-Y., Barrat, A., Colizza, V. & Poletto, C. Anatomy of digital contact tracing: Role of age, transmission setting, adoption and case detection. *Science Advances* (2021).

368. Von Wyl, V., Höglinger, M., Sieber, C., Kaufmann, M., Moser, A., Serra-Burriel, M., Ballouz, T., Menges, D., Frei, A. & Puhan, M. A. Drivers of Acceptance of COVID-19 Proximity Tracing Apps in Switzerland: Panel Survey Analysis. *JMIR Public Health Surveill* **7,** e25701 (2021).

369. Munzert, S., Selb, P., Gohdes, A., Stoetzer, L. F. & Lowe, W. Tracking and promoting the usage of a COVID-19 contact tracing app. *Nature Human Behaviour* (2021).

370. Centola, D. The spread of behavior in an online social network experiment. *Science* **329,** 1194–1197 (2010).

371. Centola, D. An experimental study of homophily in the adoption of health behavior. *Science* **334,** 1269–1272 (2011).

372. Eichner, M. Case isolation and contact tracing can prevent the spread of smallpox. *American Journal of Epidemiology* **158,** 118–128 (2003).

373. Fox, G. J., Barry, S. E., Britton, W. J. & Marks, G. B. Contact investigation for tuberculosis: a systematic review and meta-analysis. *The European Respiratory Journal* **41,** 140–156 (2013).

374. Ramstedt, K., Hallhagen, G., Lundin, B. I., Håkansson, C., Johannisson, G., Löwhagen, G. B., Norkrans, G. & Giesecke, J. Contact tracing for human immunodeficiency virus (HIV) infection. *Sexually Transmitted Diseases* **17,** 37–41 (1990).

375. Swanson, K. C., Altare, C., Wesseh, C. S., Nyenswah, T., Ahmed, T., Eyal, N., Hamblion, E. L., Lessler, J., Peters, D. H. & Altmann, M. Contact tracing performance during the Ebola epidemic in Liberia, 2014-2015. *PLOS Neglected Tropical Diseases* **12** (2018).

376. Howell, M. R., Kassler, W. J. & Haddix, A. Partner Notification to Prevent Pelvic Inflammatory Disease in Women. *Sexually Transmitted Diseases* **24,** 287–292 (1997).

377. Fraser, C., Riley, S., Anderson, R. M. & Ferguson, N. M. Factors that make an infectious disease outbreak controllable. *Proceedings of the National Academy of Sciences of the United States of America* **101,** 6146–6151 (2004).

378. Klinkenberg, D., Fraser, C. & Heesterbeek, H. The effectiveness of contact tracing in emerging epidemics. *PLOS ONE* **1** (2006).

379. Porco, T. C., Holbrook, K., Fernyak, S. E., Portnoy, D., Reiter, R. & Aragón, T. J. Logistics of community smallpox control through contact tracing and ring vaccination: A stochastic network model. *BMC Public Health* **4** (2004).

380. Müller, J., Kretzschmar, M. & Dietz, K. Contact tracing in stochastic and deterministic epidemic models. *Mathematical Biosciences* **164,** 39–64 (2000).

381. Kiss, I. Z., Green, D. M. & Kao, R. R. Infectious disease control using contact tracing in random and scale-free networks. *Journal of the Royal Society Interface* **3,** 55–62 (2006).

382. Eames, K. T. & Keeling, M. J. Contact tracing and disease control. *Proceedings of the Royal Society B: Biological Sciences* **270,** 2565–2571 (2003).

383. Becker, N. G., Glass, K., Li, Z. & Aldis, G. K. Controlling emerging infectious diseases like SARS. *Mathematical Biosciences* **193,** 205–221 (2005).

384. Rodríguez, P. *et al.* A population-based controlled experiment assessing the epidemiological impact of digital contact tracing. *Nature Communications* **12,** 1–6 (2021).

385. Braithwaite, I., Callender, T., Bullock, M. & Aldridge, R. W. Automated and partly automated contact tracing: a systematic review to inform the control of COVID-19. *The Lancet Digital Health* **2,** e607–e621 (2020).

386. Salathé, M. *et al.* Early evidence of effectiveness of digital contact tracing for SARS-CoV-2 in Switzerland. *Swiss Medical Weekly* **150,** w20457 (2020).

387. Wymant, C. *et al.* The epidemiological impact of the NHS COVID-19 app. *Nature* **594,** 408–412 (2021).

388. Kucharski, A. J., Klepac, P., Conlan, A. J., Kissler, S. M., Tang, M. L., Fry, H., Gog, J. R., Edmunds, W. J., Emery, J. C., Medley, G. *et al.* Effectiveness of isolation, testing, contact tracing, and physical distancing on reducing transmission of SARS-CoV-2 in different settings: a mathematical modelling study. *The Lancet Infectious Diseases* **20,** 1151–1160 (2020).

389. Bradshaw, W. J., Alley, E. C., Huggins, J. H., Lloyd, A. L. & Esvelt, K. M. Bidirectional contact tracing could dramatically improve COVID-19 control. *Nature Communications* **12** (2021).

390. Aleta, A., Martin-Corral, D., y Piontti, A. P., Ajelli, M., Litvinova, M., Chinazzi, M., Dean, N. E., Halloran, M. E., Longini Jr, I. M., Merler, S. *et al.* Modelling the impact of testing, contact tracing and household quarantine on second waves of COVID-19. *Nature Human Behaviour* **4,** 964–971 (2020).

391. Ferretti, L., Wymant, C., Kendall, M., Zhao, L., Nurtay, A., Abeler-Dörner, L., Parker, M., Bonsall, D. & Fraser, C. Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science* **368,** eabb6936 (2020).

392. Barrat, A., Cattuto, C., Kivelä, M., Lehmann, S. & Saramäki, J. Effect of manual and digital contact tracing on COVID-19 outbreaks: a study on empirical contact data. *Journal of the Royal Society Interface* **18,** 20201000 (2021).

393. Hellewell, J. *et al.* Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. *The Lancet Global Health* **8,** e488–e496 (2020).

394. Cencetti, G., Santin, G., Longa, A., Pigani, E., Barrat, A., Cattuto, C., Lehmann, S., Salathé, M. & Lepri, B. Digital proximity tracing on empirical contact networks for pandemic control. *Nature Communications* **12,** 1655 (2021).

395. Reyna-Lara, A., Soriano-Paños, D., Gómez, S., Granell, C., Matamalas, J. T., Steinegger, B., Arenas, A. & Gómez-Gardeñes, J. Virus spread versus contact tracing: Two competing contagion processes. *Physical Review Research* **3,** 013163 (1 2021).

396. Bianconi, G., Sun, H., Rapisardi, G. & Arenas, A. Message-passing approach to epidemic tracing and mitigation with apps. *Physical Review Research* **3,** L012014 (1 2021).

397. Kryven, I. & Stegehuis, C. Contact tracing in configuration models. *Journal of Physics: Complexity* **2,** 025004 (2021).

398. Kojaku, S., Hébert-Dufresne, L., Mones, E., Lehmann, S. & Ahn, Y.-Y. The effectiveness of backward contact tracing in networks. *Nature Physics,* 1–7 (2021).

399. Génois, M. & Barrat, A. Can co-location be used as a proxy for face-to-face contacts? *EPJ Data Science* **7,** 11 (2018).

400. Stehlé, J. *et al.* High-resolution measurements of face-to-face contact patterns in a primary school. *PLOS ONE* **6,** e23176 (2011).

401. Smith, J., Lipsitch, M. & Almond, J. W. Vaccine production, distribution, access, and uptake. *The Lancet* **378,** 428–438 (2011).

402. Wouters, O. J., Shadlen, K. C., Salcher-Konrad, M., Pollard, A. J., Larson, H. J., Teerawattananon, Y. & Jit, M. Challenges in ensuring global access to COVID-19 vaccines: production, affordability, allocation, and deployment. *The Lancet* **397,** 1023–1034 (2021).

403. May, T. & Silverman, R. D. "Clustering of exemptions" as a collective action threat to herd immunity. *Vaccine* **21,** 1048–1051 (2003).

404. Editorial. Vaccine hesitancy: a generation at risk. *The Lancet Child & Adolescent Health* **3,** 281 (2019).

405. Robertson, E., Reeve, K. S., Niedzwiedz, C. L., Moore, J., Blake, M., Green, M., Katikireddi, S. V. & Benzeval, M. J. Predictors of COVID-19 vaccine hesitancy in the UK household longitudinal study. *Brain, Behavior, and Immunity* **94,** 41–50 (2021).

406. Larson, H. J., Jarrett, C., Eckersberger, E., Smith, D. M. D. & Paterson, P. Understanding vaccine hesitancy around vaccines and vaccination from a global perspective: a systematic review of published literature, 2007-2012. *Vaccine* **32,** 2150–2159 (2014).

407. Richard, J. l, Masserey-Spicher, V., Santibanez, S. & Mankertz, A. Measles outbreak in Switzerland–an update relevant for the European football championship (EURO 2008). *Eurosurveillance* **13,** 8043 (2008).

408. Omer, S. B., Enger, K. S., Moulton, L. H., Halsey, N. A., Stokley, S. & Salmon, D. A. Geographic clustering of nonmedical exemptions to school immunization requirements and associations with geographic clustering of pertussis. *American Journal of Epidemiology* **168,** 1389–1396 (2008).

409. Atwell, J. E., Van Otterloo, J., Zipprich, J., Winter, K., Harriman, K., Salmon, D. A., Halsey, N. A. & Omer, S. B. Nonmedical vaccine exemptions and pertussis in California, 2010. *Pediatrics* **132,** 624–630 (2013).

410. Lieu, T. A., Ray, G. T., Klein, N. P., Chung, C. & Kulldorff, M. Geographic clusters in underimmunization and vaccine refusal. *Pediatrics* **135,** 280–289 (2015).

411. Mbah, M. L. N., Liu, J., Bauch, C. T., Tekel, Y. I., Medlock, J., Meyers, L. A. & Galvani, A. P. The impact of imitation on vaccination behavior in social contact networks. *PLOS Computational Biology* **8,** e1002469 (2012).

412. Salathé, M. & Bonhoeffer, S. The effect of opinion clustering on disease outbreaks. *Journal of the Royal Society, Interface* **5,** 1505–1508 (2008).

413. Liu, F., Enanoria, W. T. A., Zipprich, J., Blumberg, S., Harriman, K., Ackley, S. F., Wheaton, W. D., Allpress, J. L. & Porco, T. C. The role of vaccination coverage, individual behaviors, and the public health response in the control of measles epidemics: an agent-based simulation for California. *BMC Public Health* **15,** 447 (2015).

414. Glasser, J. W., Feng, Z., Omer, S. B., Smith, P. J. & Rodewald, L. E. The effect of heterogeneity in uptake of the measles, mumps, and rubella vaccine on the potential for outbreaks of measles: a modelling study. *The Lancet Infectious Diseases* **16,** 599–605 (2016).

415. Kuylen, E., Willem, L., Broeckhove, J., Beutels, P. & Hens, N. Clustering of susceptible individuals within households can drive measles outbreaks: an individual-based model exploration. *Scientific Reports* **10,** 19645 (2020).

416. Centers for Disease Control and Prevention. *CDC Seasonal Flu Vaccine Effectiveness Studies* https://www.cdc.gov/flu/vaccines-work/effectiveness-studies.htm. [Online; accessed 01-December-2021]. (2021).

417. Pouwels, K. B. *et al.* Effect of Delta variant on viral burden and vaccine effectiveness against new SARS-CoV-2 infections in the UK. *Nature Medicine* **27,** 2127–2135 (2021).

418. Sekara, V. & Lehmann, S. The strength of friendship ties in proximity sensor data. *PLOS ONE* **9,** e100915 (2014).

419. Bubar, K. M., Reinholt, K., Kissler, S. M., Lipsitch, M., Cobey, S., Grad, Y. H. & Larremore, D. B. Model-informed COVID-19 vaccine prioritization strategies by age and serostatus. *Science* **371,** 916–921 (2021).

420. Sonabend, R. *et al.* Non-pharmaceutical interventions, vaccination, and the SARS-CoV-2 delta variant in England: a mathematical modelling study. *The Lancet* **398,** 1825–1835 (2021).

421. Wilf-Miron, R., Myers, V. & Saban, M. Incentivizing Vaccination Uptake: The "Green Pass" Proposal in Israel. *JAMA* **325,** 1503–1504 (2021).

422. Dada, S., Battles, H., Pilbeam, C., Singh, B., Solomon, T. & Gobat, N. Learning from the past and present: social science implications for COVID-19 immunity-based documentation. *Humanities and Social Sciences Communications* **8,** 219 (2021).

423. Hiraoka, T., Rizi, A. K., Kivelä, M. & Saramäki, J. Herd immunity and epidemic size in networks with vaccination homophily. *Physical Review E* **105,** L052301 (2022).

424. Watanabe, H. & Hasegawa, T. Impact of assortative mixing by mask-wearing on the propagation of epidemics in networks. *Physica A: Statistical Mechanics and its Applications* **603,** 127760 (2022).

425. Fefferman, N. H., Silk, M. J., Pasquale, D. K. & Moody, J. Homophily in risk and behavior complicate understanding the COVID-19 epidemic curve. *medRxiv 2021.03.16.21253708* (2021).

426. Young, M. J., Silk, M. J., Pritchard, A. J. & Fefferman, N. H. Diversity in valuing social contact and risk tolerance leading to the emergence of homophily in populations facing infectious threats. *Physical Review E* **105** (2022).

427. Rizi, A. K., Faqeeh, A., Badie-Modiri, A. & Kivelä, M. Epidemic spreading and digital contact tracing: Effects of heterogeneous mixing and quarantine failures. *Physical Review E* **105,** 044313 (2022).

428. Gomes, M. G. M., Gordon, S. B. & Lalloo, D. G. Clinical trials: The mathematics of falling vaccine efficacy with rising disease incidence. *Vaccine* **34,** 3007–3009 (2016).

429. Feld, S. L. Why your friends have more friends than you do. *American Journal of Sociology* **96,** 1464–1477 (1991).

430. Schneeberger, A., Mercer, C. H., Gregson, S. A., Ferguson, N. M., Nyamukapa, C. A., Anderson, R. M., Johnson, A. M. & Garnett, G. P. Scale-free networks and sexually transmitted diseases: a description of observed patterns of sexual contacts in Britain and Zimbabwe. *Sexually transmitted diseases,* 380–387 (2004).

431. Anderson, R. & May, R. Control of communicable diseases by age-specific immunisation schedules. *The Lancet* **319,** 160 (1982).

432. Anderson, R. & May, R. Age-related changes in the rate of disease transmission: implications for the design of vaccination programmes. *Epidemiology & Infection* **94,** 365–436 (1985).

433. Pastor-Satorras, R. & Vespignani, A. Immunization of complex networks. *Physical Review E* **65,** 036104 (2002).

434. Altarelli, F., Braunstein, A., Dall'Asta, L., Wakeling, J. R. & Zecchina, R. Containing epidemic outbreaks by message-passing techniques. *Physical Review X* **4,** 021024 (2014).

435. Morone, F. & Makse, H. A. Influence maximization in complex networks through optimal percolation. *Nature* **524,** 65–68 (2015).

436. Albert, R., Jeong, H. & Barabási, A.-L. Error and attack tolerance of complex networks. *Nature* **406,** 378–382 (2000).

437. Callaway, D. S., Newman, M. E., Strogatz, S. H. & Watts, D. J. Network robustness and fragility: Percolation on random graphs. *Physical Review Letters* **85,** 5468 (2000).

438. Cohen, R., Erez, K., Ben-Avraham, D. & Havlin, S. Breakdown of the internet under intentional attack. *Physical Review Letters* **86,** 3682 (2001).

439. Restrepo, J. G., Ott, E. & Hunt, B. R. Characterizing the dynamical importance of network nodes and links. *Physical Review Letters* **97,** 094102 (2006).

440. Chen, Y., Paul, G., Havlin, S., Liljeros, F. & Stanley, H. E. Finding a better immunization strategy. *Physical Review Letters* **101,** 058701 (2008).

441. Kitsak, M., Gallos, L. K., Havlin, S., Liljeros, F., Muchnik, L., Stanley, H. E. & Makse, H. A. Identification of influential spreaders in complex networks. *Nature Physics* **6,** 888–893 (2010).

442. Liu, S., Perra, N., Karsai, M. & Vespignani, A. Controlling contagion processes in activity driven networks. *Physical Review Letters* **112,** 118702 (2014).

443. Van Mieghem, P., Stevanović, D., Kuipers, F., Li, C., Van De Bovenkamp, R., Liu, D. & Wang, H. Decreasing the spectral radius of a graph by link removals. *Physical Review E* **84,** 016101 (2011).

444. Buldyrev, S. V., Parshani, R., Paul, G., Stanley, H. E. & Havlin, S. Catastrophic cascade of failures in interdependent networks. *Nature* **464,** 1025–1028 (2010).

445. Buckee, C. O., Balsari, S., Chan, J., Crosas, M., Dominici, F., Gasser, U., Grad, Y. H., Grenfell, B., Halloran, M. E., Kraemer, M. U. *et al.* Aggregated mobility data could help fight COVID-19. *Science* **368,** 145–146 (2020).

446. Cohen, R., Havlin, S. & Ben-Avraham, D. Efficient immunization strategies for computer networks and populations. *Physical Review Letters* **91,** 247901 (2003).

447. Holme, P. Efficient local strategies for vaccination and network attack. *EPL (Europhysics Letters)* **68,** 908 (2004).

448. Grant, R. M. *et al.* Preexposure Chemoprophylaxis for HIV Prevention in Men Who Have Sex with Men. *New England Journal of Medicine* **363,** 2587–2599 (2010).

449. Sherman, J. & Morrison, W. J. Adjustment of an Inverse Matrix Corresponding to a Change in One Element of a Given Matrix. *The Annals of Mathematical Statistics* **21,** 124–127 (1950).

450. Rackauckas, C. & Nie, Q. Differentialequations.jl–a performant and feature-rich ecosystem for solving differential equations in julia. *Journal of Open Research Software* **5** (2017).

451. Mogensen, P. K. & Riseth, A. N. Optim: A mathematical optimization package for Julia. *Journal of Open Source Software* **3,** 615 (2018).

452. Wei, C. & Raymond, H. F. Pre-exposure prophylaxis for men who have sex with men in China: challenges for routine implementation. *Journal of the International AIDS Society* **21,** e25166 (2018).

453. UNAIDS. *Key Population Atlas* [Online; accessed 20-September-2021]. https://aidsinfo.unaids.org/ (2021).

454. Aghaizu, A., Wayal, S., Nardone, A., Parsons, V., Copas, A., Mercey, D., Hart, G., Gilson, R. & Johnson, A. M. Sexual behaviours, HIV testing, and the proportion of men at risk of transmitting and acquiring HIV in London, UK, 2000–13: a serial cross-sectional study. *The Lancet HIV* **3,** e431–e440 (2016).

455. Whittles, L. K., White, P. J. & Didelot, X. A dynamic power-law sexual network model of gonorrhoea outbreaks. *PLOS Computational Biology* **15,** e1006748 (2019).

456. Buchbinder, S. P. Maximizing the Benefits of HIV Preexposure Prophylaxis. *TTopics in Antiviral Medicine* **25,** 138–142 (2018).

457. Mugo, N. R., Ngure, K., Kiragu, M., Irungu, E. & Kilonzo, N. PrEP for Africa: What we have learnt and what is needed to move to program implementation. *Current Opinion in HIV and AIDS* **11,** 80 (2016).

458. Pines, H. A., Karris, M. Y. & Little, S. J. Sexual partner concurrency among partners reported by MSM with recent HIV infection. *AIDS and Behavior* **21,** 3026–3034 (2017).

459. Gantenberg, J. R., King, M., Montgomery, M. C., Galárraga, O., Prosperi, M., Chan, P. A. & Marshall, B. D. Improving the impact of HIV pre-exposure prophylaxis implementation in small urban centers among men who have sex with men: An agent-based modelling study. *PLOS ONE* **13,** e0199915 (2018).

460. Gómez-Gardeñes, J., Gómez, S., Arenas, A. & Moreno, Y. Explosive synchronization transitions in scale-free networks. *Physical Review Letters* **106,** 128701 (2011).

461. Newman, M. E. J. & Clauset, A. Structure and inference in annotated networks. *Nature Communications* **7,** 11863 (2016).

462. Peel, L., Larremore, D. B. & Clauset, A. The ground truth about metadata and community detection in networks. *Science Advances* **3,** e1602548 (2017).

463. Artime, O. & De Domenico, M. Percolation on feature-enriched interconnected systems. *Nature Communications* **12,** 2478 (2021).

464. Fajardo-Fontiveros, O., Guimerà, R. & Sales-Pardo, M. Node Metadata Can Produce Predictability Crossovers in Network Inference Problems. *Physical Review X* **12,** 011010 (2022).

465. Riou, J. & Althaus, C. L. Pattern of early human-to-human transmission of Wuhan 2019 novel coronavirus (2019-nCoV), December 2019 to January 2020. *Eurosurveillance* **25,** 2000058 (2020).

466. Davies, N. G., Abbott, S., Barnard, R. C., Jarvis, C. I., Kucharski, A. J., Munday, J. D., Pearson, C. A., Russell, T. W., Tully, D. C., Washburne, A. D. *et al.* Estimated transmissibility and impact of SARS-CoV-2 lineage B. 1.1. 7 in England. *Science* **372,** eabg3055 (2021).

467. Faria, N. R., Mellan, T. A., Whittaker, C., Claro, I. M., Candido, D. d. S., Mishra, S., Crispim, M. A., Sales, F. C., Hawryluk, I., McCrone, J. T. *et al.* Genomics and epidemiology of the P. 1 SARS-CoV-2 lineage in Manaus, Brazil. *Science* **372,** 815–821 (2021).

468. Knock, E. S. *et al.* Key epidemiological drivers and impact of interventions in the 2020 SARS-CoV-2 epidemic in England. *Science Translational Medicine* **13** (2021).

469. Marziano, V. *et al.* Retrospective analysis of the Italian exit strategy from COVID-19 lockdown. *Proceedings of the National Academy of Sciences USA* **118** (2021).

470. García-Basteiro, A. L. *et al.* The need for an independent evaluation of the COVID-19 response in Spain. *The Lancet* **396,** 529–530 (2020).

471. García-Basteiro, A. L., Legido-Quigley, H. & signatories, Evaluation of the COVID-19 response in Spain: principles and requirements. *The Lancet Public Health* **5,** e575 (2020).

472. McCabe, R., Kont, M. D., Schmit, N., Whittaker, C., Løchen, A., Walker, P. G., Ghani, A. C., Ferguson, N. M., White, P. J., Donnelly, C. A. *et al.* Communicating uncertainty in epidemic models. *Epidemics* **37,** 100520 (2021).

473. Kuhn, T. S. *The Structure of Scientific Revolutions* (Chicago University Press, 1970).

474. Wikipedia. *COVID-19 — Wikipedia, The Free Encyclopedia* http://en.wikipedia.org/w/index.php?title=COVID-19&oldid=1100209224. [Online; accessed 27-July-2022].

475. Hu, H., Nigmatulina, K. & Eckhoff, P. The scaling of contact rates with population density for the infectious disease models. *Mathematical Biosciences* **244,** 125 –134 (2013).

476. Wikipedia. *Anexo:Cronología de la pandemia de COVID-19 en España — Wikipedia, The Free Encyclopedia* http://es.wikipedia.org/w/index.php?title=Anexo%3ACronolog%C3%ADa%20de%20la%20pandemia%20de%20COVID-19%20en%20Espa%C3%B1a&oldid=144249837. [Online; accessed 27-July-2022].

477. Nouvellet, P., Bhatia, S., Cori, A., Ainslie, K. E., Baguelin, M., Bhatt, S., Boonyasiri, A., Brazeau, N. F., Cattarino, L., Cooper, L. V. *et al.* Reduction in mobility and COVID-19 transmission. *Nature Communications* **12,** 1–9 (2021).

478. Maier, B. F. & Brockmann, D. Effective containment explains subexponential growth in recent confirmed COVID-19 cases in China. *Science* **368,** 742–746 (2020).

479. Red Nacional de Vigilancia Epidemiológica. *Informe nº 31. Situación de COVID-19 en España a 14 de mayo de 2020* tech. rep. (Instituto de Salud Carlos III, 2020).

480. *Instituto Nacional de Estadistica. Mobility reduction.* https://www.ine.es/en/covid/covid_movilidad_en.htm. [Online; accessed 17-June-2020].

481. *Instituto Nacional de Estadistica. Spain* https://www.ine.es. [Online; accessed 26-February-2022].

482. Nelder, J. A. & Mead, R. A Simplex Method for Function Minimization. *The Computer Journal* **7,** 308–313 (1965).

483. Williams, M. J. *ApproxBayes.jl (v0.3.2)* https://github.com/marcjwilliams1/ApproxBayes.jl. (2020).

484. Pollán, M., Pérez-Gómez, B., Pastor-Barriuso, R., Oteo, J., Hernán, M. A., Pérez-Olmeda, M., Sanmartín, J. L., Fernández-García, A., Cruz, I., de Larrea, N. F. *et al.* Prevalence of SARS-CoV-2 in Spain (ENE-COVID): a nationwide, population-based seroepidemiological study. *The Lancet* **396,** 535–544 (2020).

485. Mitjà, O., Arenas, À., Rodó, X., Tobias, A., Brew, J. & Benlloch, J. M. Experts' request to the Spanish Government: move Spain towards complete lockdown. *The Lancet* **395,** 1193–1194 (2020).

486. Castro, M., Ares, S., Cuesta, J. A. & Manrubia, S. The turning point and end of an expanding epidemic cannot be precisely forecast. *Proceedings of the National Academy of Sciences* **117,** 26190–26196 (2020).

487. Centro Nacional de Epidemilogía. *COVID-19 en España* https://cnecovid.isciii.es/covid19/#documentaci%C3%B3n-y-datos. [Online; accessed 20-November-2020].

488. Ministerio de transportes, movilidad, y agenda urbana. *Estudio de movilidad con Big Data* https://www.mitma.gob.es/ministerio/covid-19/evolucion-movilidad-big-data.

489. Badr, H. S., Du, H., Marshall, M., Dong, E., Squire, M. M. & Gardner, L. M. Association between mobility patterns and COVID-19 transmission in the USA: a mathematical modelling study. *The Lancet Infectious Diseases* **20,** 1247–1254 (2020).

490. Dehning, J., Zierenberg, J., Spitzner, F. P., Wibral, M., Neto, J. P., Wilczek, M. & Priesemann, V. Inferring change points in the spread of COVID-19 reveals the effectiveness of interventions. *Science* **369,** eabb9789 (2020).

491. Lemaitre, J. C., Perez-Saez, J., Azman, A. S., Rinaldo, A. & Fellay, J. Assessing the impact of non-pharmaceutical interventions on SARS-CoV-2 transmission in Switzerland. *Swiss Medical Weekly* **150** (2020).

492. Marziano, V., Guzzetta, G., Rondinone, B. M., Boccuni, F., Riccardo, F., Bella, A., Poletti, P., Trentini, F., Pezzotti, P., Brusaferro, S. *et al.* Retrospective analysis of the Italian exit strategy from COVID-19 lockdown. *Proceedings of the National Academy of Sciences* **118,** e2019617118 (2021).

493. Vanni, F., Lambert, D., Palatella, L. & Grigolini, P. On the use of aggregated human mobility data to estimate the reproduction number. *Scientific Reports* **11,** 1–10 (2021).

494. Esquirol, B., Prignano, L., Díaz-Guilera, A. & Cozzo, E. *Characterizing Twitter users behaviour during the Spanish COVID-19 first wave, arXiv:2012.06550* (2020).

495. Schlosser, F., Maier, B. F., Jack, O., Hinrichs, D., Zachariae, A. & Brockmann, D. COVID-19 lockdown induces disease-mitigating structural changes in mobility networks. *Proceedings of the National Academy of Sciences* **117,** 32883–32890 (2020).

496. Zhang, J. *et al.* Evolving epidemiology and transmission dynamics of coronavirus disease 2019 outside Hubei province, China: a descriptive and modelling study. *The Lancet Infectious Diseases* **20,** 793–802 (2020).

497. Becker, N. G., Watson, L. F. & Carlin, J. B. A method of non-parametric back-projection and its application to aids data. *Statistics in Medicine* **10,** 1527–1542 (1991).

498. Michael Höhle, D. S. B. *Non-parametric back-projection of incidence cases to exposure time* https://cran.r-project.org/web/packages/BayesianTools/index.html. (2019).

499. Meyer, S., Held, L. & Höhle, M. Spatio-Temporal Analysis of Epidemic Phenomena Using the R Package surveillance. *Journal of Statistical Software* **77,** 1–55 (2017).

500. García-García, D., Vigo, M. I., Fonfría, E. S., Herrador, Z., Navarro, M. & Bordehore, C. Retrospective methodology to estimate daily infections from deaths (REMEDID) in COVID-19: the Spain case study. *Scientific Reports* **11,** 11274 (2021).

501. Pei, S., Kandula, S. & Shaman, J. Differential effects of intervention timing on COVID-19 spread in the United States. *Science Advances* **6,** eabd6370 (2020).

502. Instituto Nacional de Estadística. *Estimación del número de defunciones semanales (EDeS) durante el brote de covid-19* https://www.ine.es/experimental/defunciones/experimental_defunciones.htm. [Online; accessed 25-July-2021]. (2020).

503. Muggeo, V. M. Estimating regression models with unknown break-points. *Statistics in Medicine* **22,** 3055–3071 (2003).

504. Muggeo, V. M. *segmented: an R Package to Fit Regression Models with Broken-Line Relationships* https://cran.r-project.org/doc/Rnews/. (2008).

505. Hawryluk, I. *et al.* Inference of COVID-19 epidemiological distributions from Brazilian hospital data: Inference of COVID-19 epidemiological distributions from Brazilian hospital data. *Journal of the Royal Society Interface* **17** (2020).

506. Wallinga, J. & Lipsitch, M. How generation intervals shape the relationship between growth rates and reproductive numbers. *Proceedings of the Royal Society B: Biological Sciences* **274,** 599–604 (2007).

507. Coma Redon, E., Mora, N., Prats-Uribe, A., Fina Avilés, F., Prieto-Alhambra, D. & Medina, M. Excess cases of influenza and the coronavirus epidemic in Catalonia: a time-series analysis of primary-care electronic medical records covering over 6 million people. *BMJ Open* **10,** e039369 (2020).

508. Althaus, C. L., Probst, D., Hauser, A. & Riou, J. Time is of the essence: containment of the SARS-CoV-2 epidemic in Switzerland from February to May 2020. *medRxiv 2020.07.21.20158014* (2020).

509. Hsiang, S., Allen, D., Annan-Phan, S., Bell, K., Bolliger, I., Chong, T., Druckenmiller, H., Huang, L. Y., Hultgren, A., Krasovich, E. *et al.* The effect of large-scale anti-contagion policies on the COVID-19 pandemic. *Nature* **584,** 262–267 (2020).

510. Bayham, J., Kuminoff, N. V., Gunn, Q. & Fenichel, E. P. Measured voluntary avoidance behaviour during the 2009 A/H1N1 epidemic. *Proceedings of the Royal Society B: Biological Sciences* **282,** 20150814 (2015).

511. Lau, J. T., Yang, X., Pang, E., Tsui, H., Wong, E. & Wing, Y. K. SARS-related perceptions in Hong Kong. *Emerging Infectious Diseases* **11,** 417 (2005).

512. Public Health England. *Investigation of novel SARS-CoV-2 variant: variant of concern 2020 12/01* https://www.gov.uk/government/publications/investigation-of-novel-sars-cov-2-variant-variant-of-concern-20201201. [Online; accessed 02-February-2021]. (2020).

513. Abu-Raddad, L. J., Chemaitelly, H., Butt, A. A. & for COVID-19 Vaccination, N. S. G. Effectiveness of the BNT162b2 Covid-19 Vaccine against the B.1.1.7 and B.1.351 Variants. *New England Journal of Medicine* **385,** 187–189 (2021).

514. Cherian, S., Potdar, V., Jadhav, S., Yadav, P., Gupta, N., Das, M., Rakshit, P., Singh, S., Abraham, P., Panda, S. *et al.* SARS-CoV-2 spike mutations, L452R, T478K, E484Q and P681R, in the second wave of COVID-19 in Maharashtra, India. *Microorganisms* **9,** 1542 (2021).

515. Nebehay, S. & Farge, E. *WHO classifies India variant as being of global concern* https://www.reuters.com/business/healthcare-pharmaceuticals/who-designates-india-variant-being-global-concern-2021-05-10/. [Online; accessed 27-July-2022]. (2021).

516. Campbell, F., Archer, B., Laurenson-Schafer, H., Jinnai, Y., Konings, F., Batra, N., Pavlin, B., Vandemaele, K., Van Kerkhove, M. D., Jombart, T. *et al.* Increased transmissibility and global spread of SARS-CoV-2 variants of concern as at June 2021. *Eurosurveillance* **26,** 2100509 (2021).

517. Wikipedia. *Pandèmia de COVID-19 a Catalunya — Wikipedia, The Free Encyclopedia* http://ca.wikipedia.org/w/index.php?title=Pand%C3%A8mia%20de%20COVID-19%20a%20Catalunya&oldid=28489476. [Online; accessed 27-October-2021].

518. Wikipedia. *Bonfires of Saint John* https://en.wikipedia.org/wiki/Bonfires_of_Saint_John. [Online; accessed 17-January-2022].

519. Wikipedia. *Saint John's Eve* https://en.wikipedia.org/wiki/Saint_John%27s_Eve. [Online; accessed 17-January-2022].

520. Lopez Bernal, J. *et al.* Effectiveness of the Pfizer-BioNTech and Oxford-AstraZeneca vaccines on covid-19 related symptoms, hospital admissions, and mortality in older adults in England: test negative case-control study. *BMJ* **373,** n1088 (2021).

521. Lopez Bernal, J. *et al.* Effectiveness of Covid-19 Vaccines against the B.1.617.2 (Delta) Variant. *New England Journal of Medicine* **385,** 585–594 (2021).

522. Tang, P. *et al.* BNT162b2 and mRNA-1273 COVID-19 vaccine effectiveness against the SARS-CoV-2 Delta variant in Qatar. *Nature Medicine* **27,** 2136–2143 (2021).

523. Bager, P. *et al.* Risk of hospitalisation associated with infection with SARS-CoV-2 lineage B.1.1.7 in Denmark: an observational cohort study. *The Lancet Infectious Diseases* **21,** 1507–1517 (2021).

524. Sheikh, A., McMenamin, J., Taylor, B., Robertson, C., Scotland, P. H. & the EAVE II Collaborators. SARS-CoV-2 Delta VOC in Scotland: demographics, risk of hospital admission, and vaccine effectiveness. *The Lancet* **397,** 2461–2462 (2021).

525. Li, B., Deng, A., Li, K., Hu, Y., Li, Z., Shi, Y., Xiong, Q., Liu, Z., Guo, Q., Zou, L. *et al.* Viral infection and transmission in a large, well-traced outbreak caused by the SARS-CoV-2 Delta variant. *Nature Communications* **13,** 1–9 (2022).

526. Alizon, S., Haim-Boukobza, S., Foulongne, V., Verdurme, L., Trombert-Paolantoni, S., Lecorche, E., Roquebert, B. & Sofonea, M. T. Rapid spread of the SARS-CoV-2 Delta variant in some French regions, June 2021. *Eurosurveillance* **26** (2021).

527. Goldberg, Y., Mandel, M., Bar-On, Y. M., Bodenheimer, O., Freedman, L., Haas, E. J., Milo, R., Alroy-Preis, S., Ash, N. & Huppert, A. Waning Immunity after the BNT162b2 Vaccine in Israel. *The New England Journal of Medicine* **385,** e85 (2021).

528. Chemaitelly, H. *et al.* Waning of BNT162b2 Vaccine Protection against SARS-CoV-2 Infection in Qatar. *The New England Journal of Medicine* **385,** e83 (2021).

529. Google LLC. *Google COVID-19 Community Mobility Reports* https://www.google.com/covid19/mobility/. [Online; accessed 20-July-2022].

530. Verelst, F., Hermans, L., Vercruysse, S., Gimma, A., Coletti, P., Backer, J. A., Wong, K. L., Wambua, J., van Zandvoort, K., Willem, L. *et al.* SOCRATES-CoMix: a platform for timely and open-source contact mixing data during and in between COVID-19 surges and interventions in over 20 European countries. *BMC Medicine* **19,** 1–7 (2021).

531. Citron, D. T., Guerra, C. A., Dolgert, A. J., Wu, S. L., Henry, J. M., Sánchez C, H. M. & Smith, D. L. Comparing metapopulation dynamics of infectious diseases under different models of human movement. *Proceedings of the National Academy of Sciences of the United States of America* **118,** e2007488118 (2021).

532. Petherick, A., Goldszmidt, R., Andrade, E. B., Furst, R., Hale, T., Pott, A. & Wood, A. A worldwide assessment of changes in adherence to COVID-19 protective behaviours and hypothesized pandemic fatigue. *Nature Human Behaviour* **5,** 1145–1160 (2021).

533. Delussu, F., Tizzoni, M. & Gauvin, L. Evidence of pandemic fatigue associated with stricter tiered COVID-19 restrictions. *PLOS Digital Health* **1,** e0000035 (2022).

534. Di Domenico, L., Sabbatini, C. E., Boëlle, P.-Y., Poletto, C., Crépey, P., Paireau, J., Cauchemez, S., Beck, F., Noel, H., Lévy-Bruhl, D. *et al.* Adherence and sustainability of interventions informing optimal control against the COVID-19 pandemic. *Communications medicine* **1,** 1–13 (2021).

535. Green, W. D., Ferguson, N. M. & Cori, A. Inferring the reproduction number using the renewal equation in heterogeneous epidemics. *Journal of the Royal Society Interface* **19,** 20210429 (2022).

536. Betsch, C., Korn, L., Sprengholz, P., Felgendreff, L., Eitze, S., Schmid, P. & Böhm, R. Social and behavioral consequences of mask policies during the COVID-19 pandemic. *Proceedings of the National Academy of Sciences* **117,** 21851–21853 (2020).

537. Qiu, Z., Espinoza, B., Vasconcelos, V. V., Chen, C., Constantino, S. M., Crabtree, S. A., Yang, L., Vullikanti, A., Chen, J., Weibull, J. *et al.* Understanding the coevolution of mask wearing and epidemics: A network perspective. *Proceedings of the National Academy of Sciences* **119,** e2123355119 (2022).

538. Ferguson, N., Laydon, D., Nedjati Gilani, G., Imai, N., Ainslie, K., Baguelin, M., Bhatia, S., Boonyasiri, A., Cucunuba Perez, Z., Cuomo-Dannenburg, G. *et al. Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand* https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/covid-19/report-9-impact-of-npis-on-covid-19/. [Online; accessed 02-February-2021]. (2020).

539. Murphy, C., Laurence, E. & Allard, A. Deep learning of contagion dynamics on complex networks. *Nature Communications* **12,** 1–11 (2021).

540. Zhao, W., Jiang, W. & Qiu, X. Deep learning for COVID-19 detection based on CT images. *Scientific Reports* **11,** 1–12 (2021).

541. Chieregato, M., Frangiamore, F., Morassi, M., Baresi, C., Nici, S., Bassetti, C., Bnà, C. & Galelli, M. A hybrid machine learning/deep learning COVID-19 severity predictive model from CT images and clinical data. *Scientific Reports* **12,** 1–15 (2022).

542. Kapoor, A., Ben, X., Liu, L., Perozzi, B., Barnes, M., Blais, M. & O'Banion, S. *Examining COVID-19 Forecasting using Spatio-Temporal Graph Neural Networks, arXiv:2007.03113* (2020).

543. Paireau, J., Andronico, A., Hozé, N., Layan, M., Crepey, P., Roumagnac, A., Lavielle, M., Boëlle, P.-Y. & Cauchemez, S. An ensemble model based on early predictors to forecast COVID-19 health care demand in France. *Proceedings of the National Academy of Sciences* **119,** e2103302119 (2022).

544. Guimerà, R., Reichardt, I., Aguilar-Mogas, A., Massucci, F. A., Miranda, M., Pallarès, J. & Sales-Pardo, M. A Bayesian machine scientist to aid in the solution of challenging scientific problems. *Science Advances* **6,** eaav6971 (2020).

545. Hopkins, D. R. Disease eradication. *New England Journal of Medicine* **368,** 54–63 (2013).

546. Wilson, N., Kvalsvig, A., Barnard, L. T. & Baker, M. G. Case-Fatality Risk Estimates for COVID-19 Calculated by Using a Lag Time for Fatality. *Emerging Infectious Diseases* **26,** 1339–1441 (2020).

547. Bhatraju, P. K. *et al.* Covid-19 in Critically Ill Patients in the Seattle Region - Case Series. *New England Journal of Medicine* **382,** 2012–2022 (2020).

548. Instituto de Estadística de Cataluña. *Población a 1 de enero. Por sexo y grupos de edad* [Online; accessed 30-September-2021]. https://www.idescat.cat/pub/?id=aec&n=253&lang=es (2021).

549. Vekaria, B. *et al.* Hospital length of stay for COVID-19 patients: Data-driven methods for forward planning. *BMC Infectious Diseases* **21,** 700 (2021).

550. Tegally, H. *et al.* Rapid replacement of the Beta variant by the Delta variant in South Africa. *medRxiv 2021.09.23.21264018* (2021).

551. Marée, A. F., Keulen, W., Boucher, C. A. & De Boer, R. J. Estimating relative fitness in viral competition experiments. *Journal of Virology* **74,** 11067–11072 (2000).

552. Cereda, D. *et al.* The early phase of the COVID-19 epidemic in Lombardy, Italy. *Epidemics* **37,** 100528 (2021).

553. Griffin, J., Casey, M., Collins, a., Hunt, K., McEvoy, D., Byrne, A., McAloon, C., Barber, A., Lane, E. A. & More, S. Rapid review of available evidence on the serial interval and generation time of COVID-19. *BMJ Open* **10,** e040263 (2020).

554. Lau, Y. C., Tsang, T. K., Kennedy-Shaffer, L., Kahn, R., Lau, E. H. Y., Chen, D., Wong, J. Y., Ali, S. T., Wu, P. & Cowling, B. J. Joint Estimation Of Generation Time And Incubation Period For Coronavirus Disease (Covid-19). *The Journal of Infectious Diseases,* jiab424 (2021).

555. Hart, W. S., Miller, E., Andrews, N. J., Waight, P., Maini, P. K., Funk, S. & Thompson, R. N. Generation time of the alpha and delta SARS-CoV-2 variants: an epidemiological analysis. *The Lancet Infectious Diseases* **22,** 603–610 (2022).

556. Twohig, K. A. *et al.* Hospital admission and emergency care attendance risk for SARS-CoV-2 delta (B.1.617.2) compared with alpha (B.1.1.7) variants of concern: a cohort study. *The Lancet Infectious Diseases,* S1473–3099(21)00475–8 (2021).

557. Bager, P., Wohlfahrt, J., Rasmussen, M., Albertsen, M. & Krause, T. G. Hospitalisation associated with SARS-CoV-2 delta variant in Denmark. *The Lancet Infectious Diseases* **21,** 1351 (2021).

558. Centers for Disease Control and Prevention. *Science Brief: SARS-CoV-2 Infection-induced and Vaccine-induced Immunity* https://www.cdc.gov/coronavirus/2019-ncov/science/science-briefs/vaccine-induced-immunity.html. [Online; accessed 15-December-2021]. (2021).

559. Karachaliou, M. *et al.* Infection induced SARS-CoV-2 seroprevalence and heterogeneity of antibody responses in a general population cohort study in Catalonia Spain. *Scientific Reports* **11,** 21571 (2021).