Annual Reviews in Control 52 (2021) 448–464

Contents lists available at ScienceDirect

Annual Reviews in Control

journal homepage: www.elsevier.com/locate/arcontrol

Review

Data-driven methods for present and future pandemics: Monitoring, modelling and managing ☆

Teodoro Alamo [a,*], Daniel G. Reina [b], Pablo Millán Gata [c], Victor M. Preciado [d], Giulia Giordano [e]

[a] Departamento de Ingeniería de Sistemas y Automática, Universidad de Sevilla, Escuela Superior de Ingenieros, Sevilla, Spain
[b] Departamento de Ingeniería Electrónica, Universidad de Sevilla, Escuela Superior de Ingenieros, Sevilla, Spain
[c] Departamento de Ingeniería, Universidad Loyola Andalucía, Seville, Spain
[d] Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, USA
[e] Department of Industrial Engineering, University of Trento, Trento, Italy

ARTICLE INFO

Keywords:
Pandemic control
Epidemiological models
Machine learning
Forecasting
Surveillance systems
Epidemic control
Optimal control
Model predictive control

ABSTRACT

This survey analyses the role of data-driven methodologies for pandemic modelling and control. We provide a roadmap from the access to epidemiological data sources to the control of epidemic phenomena. We review the available methodologies and discuss the challenges in the development of data-driven strategies to combat the spreading of infectious diseases. Our aim is to bring together several different disciplines required to provide a holistic approach to epidemic analysis, such as data science, epidemiology, and systems-and-control theory. A 3M-analysis is presented, whose three pillars are: Monitoring, Modelling and Managing. The focus is on the potential of data-driven schemes to address three different challenges raised by a pandemic: (i) monitoring the epidemic evolution and assessing the effectiveness of the adopted countermeasures; (ii) modelling and forecasting the spread of the epidemic; (iii) making timely decisions to manage, mitigate and suppress the contagion. For each step of this roadmap, we review consolidated theoretical approaches (including data-driven methodologies that have been shown to be successful in other contexts) and discuss their application to past or present epidemics, such as Covid-19, as well as their potential application to future epidemics.

1. Introduction

The 2019 coronavirus pandemic (Covid-19) is one of the most critical public health emergencies in recent human history. While facing this pandemic, governments, public institutions, healthcare professionals, and researches of different disciplines address the problem of effectively controlling the spread of the virus while minimizing the negative effects on both the economy and society. The challenges raised by this pandemic require a holistic approach. In this document, we analyse the interplay between data science, epidemiology and control theory, which is crucial to understand and manage the spread of diseases both in human and animal populations. In line with current epidemiological needs, this paper aims to review available methodologies, while anticipating the difficulties and challenges encountered in the development of data-driven strategies to combat pandemics. We consider the Covid-19 pandemic as a case study and summarize some lessons learned from this pandemic with the hope of improving our preparedness at handling future outbreaks.

In the context of epidemics outbreaks, data-driven tools are fundamental to: (i) monitor the spread of the epidemic and assess the potential impact of adopted countermeasures, not only from a healthcare perspective but also from a socioeconomic one; (ii) model and forecast the epidemic evolution; (iii) manage the epidemic by making timely decisions to mitigate and suppress the contagion. Optimal decision making in the context of a pandemic is a complex process involving a significant amount of uncertainty; at the same time, not reacting timely and with adequate intensity, even in the presence of overwhelming uncertainties, can lead to severe consequences. This survey provides a holistic roadmap that encompasses from the process of retrieving epidemiological data to the decision-making process aimed at controlling, mitigating and preventing the epidemic spread. A 3M-analysis

☆ The authors belong to the CONtrol COvid-19 Team, including more than 35 researches from universities of Spain, Italy, France, Germany, United Kingdom and Argentina. The main goal of the CONCO-Team is to develop data-driven methods to better understand and control the Covid-19 pandemic. This work was supported by the Agencia Estatal de Investigación (AEI)-Spain under Grant PID2019-106212RB-C41/AEI/10.13039/501100011033. Victor Preciado acknowledges the support of the US National Science Foundation under grants CAREER-ECCS-1651433 and NSF-III-200884556. Giulia Giordano acknowledges the support of the Strategic Grant MOSES at the University of Trento.

* Corresponding author.
E-mail address: talamo@us.es (T. Alamo).

https://doi.org/10.1016/j.arcontrol.2021.05.003
Received 30 December 2020; Received in revised form 24 May 2021; Accepted 27 May 2021
Available online 29 June 2021

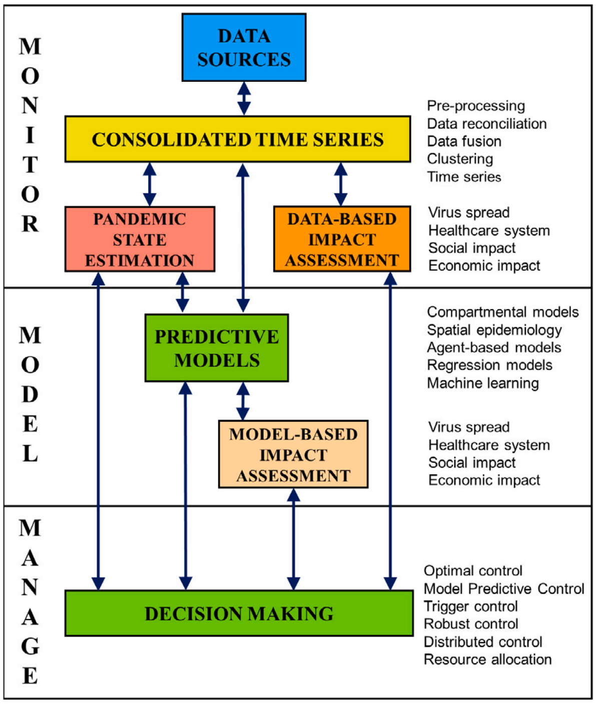

Fig. 1. 3M-Approach to data-driven control of an epidemic: Monitoring, Modelling and Managing.

is proposed, covering three main aspects: Monitoring, Modelling and Managing, as shown in Fig. 1. A more detailed document, focused on the Covid-19 pandemic, can be found in the preprint (Alamo, Reina and Millán, 2020). Each step of this roadmap is presented through a review of consolidated theoretical methods and a discussion of their potential to help us understand and control pandemics. When possible, examples of applications of these methodologies on past or current epidemics are provided. Data-driven methodologies that have proven successful in other biological contexts, or have been identified as promising solutions in the Covid-19 pandemic, are highlighted. This survey does not provide an exhaustive enumeration of methodologies, algorithms and applications. Instead, it is conceived to serve as a bridge between those disciplines required to develop a holistic approach to the epidemic, namely: data science, epidemiology, and control theory.

Data are a fundamental pillar to understand, model, forecast, and manage many of the aspects required to provide a comprehensive response against an epidemic, or pandemic, outbreak. There exist many different open data resources and institutions providing relevant information not only in terms of specific epidemiological variables but also of other auxiliary variables that facilitate the assessment of the effectiveness of the implemented interventions (see Alamo, Reina, Mammarella and Abella, 2020 for a review on open data resources and repositories for the Covid-19 case). Since the available epidemiological data suffer from severe limitations, methodologies to detect anomalies in the raw data and generate time-series with enhanced quality (like data reconciliation, data-fusion, data-clustering, signal processing, to name just a few) play a crucial role.

Another important aspect of the 3M-approach is the real-time surveillance of the epidemic, which can be implemented by monitoring mobility, using social media to assess the compliance to restrictions and recommendations, pro-active testing, contact-tracing, etc. The design and implementation of surveillance systems capable of early detecting secondary epidemic waves is also very important.

Modelling techniques are also fundamental in the fight against pandemics. Epidemiological models range from coarse compartmental models to complex networked and agent-based models. Fundamental parameters characterizing the dynamics of the virus can be identified using these models. Besides, data-driven parameter estimation provides mechanisms to forecast the epidemic evolution, as well as to anticipate the effectiveness of adopted interventions. However, fitting the models to the available data requires specific techniques because of critical issues like partial observation, non-linearities and non-identifiability. Sensitivity analysis, model selection and validation methodologies have to be implemented (Burnham & Anderson, 2010; Martcheva, 2015). Apart from the forecasting possibilities that epidemiological models offer, alternative forecasting techniques from the field of data science can be applied in this context. The choice ranges from simple linear parametric methods to complex deep-learning approaches. The methods can be parametric or non-parametric in nature. Some of these techniques provide probabilistic characterizations of the provided forecasts.

Several measures to mitigate the epidemic can be found in the literature, but one needs to be careful about their effectiveness (Xiao & Torok, 2020). Some measures, like an aggressive lockdown of an entire country, have a devastating effect on the economy and they might be adopted at very precise moments, preferably as early as possible and for short time periods. Other measures, like pro-active testing and contact-tracing, can be very effective while having a minor impact on the economy (Ferretti, Wymant, Kendall, Zhao, et al., 2020). In this direction, control theory provides a consolidated framework to formulate and solve many relevant decision-making problems (Nowzari, Preciado, & Pappas, 2016), such as the optimal allocation of resources (e.g. test reagents and vaccines) and the determination of the optimal time to implement certain interventions. The use of optimal control theory and (distributed) model predictive control has great potential in epidemic control. Mathematical tools from the fields of control theory and dynamic systems, such as bifurcation theory and Lyapunov theory, have been extensively used to characterize the different possible qualitative behaviours of epidemics.

This survey is organized as follows: Section 2 describes different methodologies to monitor the current state of a pandemic. An overview of different techniques to model an epidemic is provided in Section 3. The main forecasting techniques are described in Section 4. The question of how to assess the effectiveness of different non-pharmaceutical measures is analysed in Section 5. The decision making process and its link with control theory is addressed in Section 6. The review paper is finished with a section describing some conclusions and lessons learned.

2. Monitor– Estimation of the state of a pandemic

There is a plethora of indicators that can be monitored in order to contain a pandemic. This includes not only estimations of the current incidence of the disease in the population and the healthcare system, but also the (daily) surveillance of measures that directly or indirectly affect its spread, such as physical distancing and mobility, as well as testing and contact tracing. In order to design an effective response to an epidemic outbreak, it is of utmost importance to build up-to-date estimations of the epidemic state. This estimation process is hindered by the presence of an incubation period of the infectious disease, which introduces a time-delay between the beginning of a new infection and its potential detection. Another challenge in the estimation process is the presence of infectious but asymptomatic cases, which is an important transmission vector in the case of several pathogens, including HIV, Zika virus and SARS-CoV-2 (Ferretti et al., 2020). These (and other) challenges motivate the need for specific surveillance and estimation methodologies capable of using available information in order to design quick and effective control measures (Alamo, Reina, Gata, Preciado, & Giordano, 2021).

In this section, we cover the most relevant techniques to monitor the state of the pandemic, focusing on approaches oriented towards (i) real-time monitoring of different aspects of the pandemic (real-time epidemiology); (ii) early detection of infected cases and immune response estimation (pro-active testing); (iii) estimation of the current fraction of infected population, both symptomatic and asymptomatic (state estimation methods); (iv) early detection of new waves (epidemic wave surveillance).

2.1. Real-time epidemiology

The use of a large number of real-time data streams to infer the status and dynamics of a population's health presents enormous opportunities as well as significant scientific and technological challenges (Bettencourt, Ribeiro, Chowell, Lant, & Castillo-Chavez, 2007; Drew, Nguyen, Steves, Menni, Freydin, et al., 2020; Zeng, Chen, Castillo-Chavez, Lober, & Thurmond, 2010). Real-time epidemic data can vary widely in nature and origin (e.g., mobile phone data, social media data, IoT data and public health systems) (Alamo, Reina, Mammarella et al., 2020; Ting, Carin, Dzau, & Wong, 2020). During the Covid-19 pandemic, mobile phone data, when used properly and carefully, have provided invaluable information for supporting public health actions across early, middle, and late-stage pandemic phases (Oliver, Lepri, Sterly, Lambiotte, Deletaille et al., 2020). Voluntary installation of Covid-19 apps or web-based tools have allowed the active retrieval of data related to exposure and infections. The information stemming from these sources has provided real-time epidemiological data that have then been used to identify hot spots for outbreaks (Drew et al., 2020). Social media have also been relevant to assess the mobility of the population and its awareness with regard to physical distancing, as well as the state of the economy and many other key indicators (Zhou, Su, Pei, Zhang, Du et al., 2020).

Our ability to extract information regarding population mobility is essential to predict spatial transmission, identify risk areas, and decide control measures against the disease. Nowadays, the most effective tool to access real-time mobility data is through Big Data technologies and Geographic Information Systems (GIS). These systems have played a relevant role when addressing past epidemics like SARS and MERS (Peeri et al., 2020), providing efficient aggregation of multi-source big data, rapid visualization of epidemic information, spatial tracking of confirmed cases, surveillance of regional transmission and spatial segmentation of the epidemic risk (Wang, Ng, & Brook, 2020; Zhou, Su et al., 2020).

2.2. Proactive testing

Proactive testing is key in the control of infectious diseases, since it allows us to identify and isolate infected individuals. It also provides relevant information to identify risk areas, fraction of asymptomatic carriers, and attained levels of immunization in the population (Winter & Hegde, 2020; Yilmaz et al., 2020). There are different methodologies to approach proactive testing:

- **Risk-based approach**: In this approach, one must test first those individuals with the highest probability of being carriers of the disease (i.e. not only those with symptoms, but also those who have been heavily exposed to the disease). For example, healthcare workers are at high risk and can also be relevant transmission vectors. Second, test those individuals that have been exposed to a confirmed case according to contact tracing. Finally, test those individuals who have recently travelled to hot spots (Wang et al., 2020). The determination of hot spots can be done by means of government mobility surveillance or by personal software environments (Drew et al., 2020).
- **Voucher-based system**: In this system, people who test positive are given an anonymous voucher that they can share with a limited number of people whom they think might have infected. The recipients can use this voucher to book a test and receive their test results without ever revealing their identity. People receiving positive result are given vouchers to further backtrack the path of infection; see Nanni, Andrienko, Barabási, Boldrini, Bonchi, et al. (2021) and Roomp and Oliver (2020) for the Covid-19 case.
- **Serology studies**: Some tests (such as RT-PCR revealing viral load) are unable to detect past infection. Conversely, serological tests, carried out within the correct time frame after disease onset, can detect both active and past infections, since they detect antibodies produced in response to the disease. Serological analysis can be useful to identify clusters of cases, to retrospectively delineate transmission chains, to ascertain how long transmission has been ongoing, or to estimate the fraction of asymptomatic individuals in the population (Winter & Hegde, 2020).

2.3. State-space estimation methods

As we will see in the next section, dynamic state-space epidemiological models are fundamental to characterize how the virus spreads in a specific region and estimate time-varying epidemiological variables that are not directly measurable (Cazelles & Chau, 1997; Scharbarg, Moog, Mauduit, & Califano, 2020). Classical state-space estimation methods, like the Kalman filter (Riad, Scoglio, Cohnstaedt, & McVey, 2019), are employed to estimate the fraction of currently infected population. The objective of the Kalman filter is to update our knowledge about the state of the system whenever a new observation is available (Durbin & Koopman, 2012). Different modifications and generalizations of the Kalman filter have been developed and tailored to epidemic models. These methodologies are essential both to the estimation problem and to the inference of the parameters that describe the model (see Abreu & Dutra, 2020; Schön, Wills, & Ninness, 2011).

2.4. Epidemic wave surveillance

Infectious diseases often lead to recurring epidemic waves interspersed with periods of low-level transmission, as observed, for example, in the "Spanish" flu (Reid, Taubenberger, & Fanning, 2001), Influenza (Vega, Lozano, Meerhoff, Snacken, Beauté, et al., 2015) and Covid-19 (Glass, 2020). In this context, it is crucial to implement a surveillance system able to detect, or even predict, recurring epidemic waves, so as to enable an immediate response aiming to reduce the potential burden of the outbreak. Detecting outbreaks requires methodologies able to process huge amount of data stemming from various surveillance systems (Althouse, Scarpino, Meyers, Ayers, et al., 2015; Dubrawski, 2011; Elliot, Harcourt, Hughes, Loveridge, Morbey, et al., 2020) and determine whether the spread of the virus has surpassed a threshold requiring mitigation measures; see, e.g. Lazarus (2010). A large body of literature focuses on epidemiological detection problems, since many infectious diseases undergo considerable seasonal fluctuations with peaks seriously impacting the healthcare systems (Sparks, 2013; Unkel, Farrington, Garthwaite, Robertson, & Andrews, 2012). National surveillance systems are implemented world-wide to rapidly detect outbreaks of influenza-like illnesses, and assess the effectiveness of influenza vaccines (Thompson, Comanor, & Shay, 2006; Vega et al., 2015). Specific methodologies to determine the baseline influenza activity and epidemic thresholds have been proposed and implemented (Vega, Lozano, Meerhoff, Snacken, Mott, et al., 2013). These methods aim at reducing false alarms and detection lags. Outbreak detection can be implemented in different ways that range from simple predictors based on moving average filters (Farrington, Andrews, Beale, & Catchpole, 1996) and fusion methods (Dubrawski, 2011) to complex spatial and temporal analyses (Batlle, Bruna, Fernandez-Granda, & Preciado, 2020; Chan & King, 2011).

In the early phases of a new pandemic, such as the recent Covid-19, the detection of recurring epidemic waves is particularly challenging because: (i) historical seasonal data are lacking, (ii) determining the current fraction of infected population can be difficult when many asymptomatic infected are present, and (iii) determining baselines and thresholds requires a precise characterization of the regional (time-varying) reproduction number.

3. Model– Epidemiological models

Mathematical epidemiology is a well-established field aiming to model the spread of diseases both in human and animal populations (Martcheva, 2015; Rothman, Greenland, & Lash, 2008; Thrusfield, 2018). Given the high complexity of these phenomena, models are key to understand epidemiological patterns and support decision making processes (Heesterbeek, Anderson, Andreasen, Bansal, De Angelis, et al., 2015). There are in-host models that take into account the complexity of virus–host dynamics at the microscopic scale, describing how the pathogen interacts with cellular biomolecular processes and with the immune system, and between-host models that describe how the epidemic spreads within a population at the macroscopic scale, by considering the contagion either at an aggregate level (compartmental models) or through agent-based networked models of the population. Approaches for epidemic multi-scale modelling, which include the interplay between immunological and epidemiological phenomena, are very recent and mostly rely on partial differential equations, sometimes reduced to small-size ordinary differential-equation systems, see e.g. Almocera and Hernandez-Vargas (2019), Almocera, Nguyen, and Hernandez-Vargas (2018), Barbarossa and Röst (2015), Cai, Tuncer, and Martcheva (2017), Feng, Cen, Zhao, and Velasco-Hernandez (2015), Gandolfi, Pugliese, and Sinisgalli (2015), Gulbudak and Browne (2020) and Hart, Maini, Yates, and Thompson (2020). Multi-scale epidemic modelling with an interdisciplinary approach integrating epidemiology, immunology, economy and mathematics is advocated in Bellomo, Bingham, Chaplain, Dosi, Forni, et al. (2020).

3.1. Time-response and viral shedding

In-host infection dynamics capture the interplay between virus and host. Models describing the dynamics of the immune response (Castiglione & Celada, 2015) in the presence of an infectious disease have been proposed for influenza (Handel, Longini, & Antia, 2010; Li, McCaw, & Cao, 2021; Yan, Cao, Heffernan, McVernon, Quinn, et al., 2017; Zarnitsyna et al., 2016) and generic viral infections (Moore et al., 2020). Very recently an immunological description for Covid-19 has been provided (Matricardi, Dal Negro, & Nisini, 2020) and has enabled the characterization of virus–host dynamics for SARS-CoV-2 (Abuin, Anderson, Ferramosca, Hernandez-Vargas, & Gonzalez, 2020; Hernandez-Vargas & Velasco-Hernandez, 2020).

The evolution of a disease and its infectiousness over time can be characterized through some key epidemiological parameters (see e.g. Heffernan, Smith, & Wahl, 2005; Hellewell, Abbott, Gimma, Bosse, Jarvis, et al., 2020; Vink, Bootsma, & Wallinga, 2014; Wallinga & Lipsitch, 2007):

- **Latency time**: Time during which an individual is infected but not yet infectious. For Covid-19, initial estimates are of 3–4 days (Li, Pei, Chen, Song Zhang et al., 2020).
- **Incubation time**: Time between infection and onset of symptoms. For Covid-19, the median incubation period is estimated to be 5.1 days, and 97.5% of those who develop symptoms will do so within 11.5 days of infection (Lauer, Grantz, Bi, Jones, Zheng, et al., 2020); the median time from the onset of symptoms to death is close to 3 weeks (Zhou et al., 2020).
- **Serial interval**: Time between symptom onsets of successive cases in a transmission chain (Vink et al., 2014). For Covid-19, initial estimates of the median serial interval yield a value of around 4 days, which is shorter than its median incubation period (Nishiura, Linton, & Akhmetzhanov, 2020); this implies that a substantial proportion of secondary transmission may occur prior to illness onset (He, Lau, Wu, Deng, Wang, et al., 2020).
- **Infectiousness profile**: It characterizes the infectiousness of an infected individual over time. For Covid-19, the median duration of viral shedding estimation was 20 days in survivors, while the most prolonged observed duration of viral shedding in survivors was 37 days (Zhou, Yu et al., 2020).
- **Basic reproduction number** $\mathcal{R}_0$: It represents the average number of new infections generated by an infectious person at the early stages of the outbreak, when everyone is susceptible, and no countermeasures have been taken (Heffernan et al., 2005; Liu, Gayle, Wilder-Smith, & Rocklöv, 2020; Wallinga & Lipsitch, 2007). For the original strain of SARS-CoV-2, first estimations range from 2.24 to 3.58 (Zhao, Lin, Ran, Musa, Yang, et al., 2020); the effect of temperature and humidity in this parameter is addressed in different studies, see e.g. Mecenas, Travassos da Rosa Moreira Bastos, Rosário Vallinoto, and Normando (0000).

The basic reproduction number, along with the serial interval, can be used to estimate the number of infections that are caused by a single case in a given time period. Without any control measure, at the early stages of the outbreak, more than 400 people can be infected by a single Covid-19 case in one month (Nicola, O'Neill, Sohrabi, Khan, Agha, et al., 2020). Estimates of the basic reproductive number are of interest during an outbreak because they provide information about the level of intervention required to interrupt transmission and about the potential final size of the outbreak (Heffernan et al., 2005).

The aforementioned parameters are often inferred from epidemiological models, once they have been fitted to available data on the number of confirmed infection cases and deaths (Rothman et al., 2008; Wallinga & Lipsitch, 2007).

3.2. Simple compartmental models

Compartmental models partition a population into different groups, called *compartments*, associated with mutually exclusive stages of the disease. Each compartment is associated with a variable that counts the individuals who are in that stage of the infection (Brauer, 2008).

The simplest compartmental models are the SI, SIS, and SIR models, introduced by Kermack and McKendrick at the beginning of the 20th century (Kermack & McKendrick, 1927). The SIR model includes three compartments: *Susceptible* (S), representing healthy individuals susceptible of getting infected, *Infected* (I), and *Recovered/Removed* (R). For possibly fatal diseases, this last compartment can take into account both recovered (with permanent immunity) and deceased individuals; however, for low mortality rate diseases, including only recovered individuals can be a good approximation.

The SIR model describes the dynamics of an epidemic according to the following set of nonlinear differential equations:

$$\frac{dS(t)}{dt} = -\beta S(t)I(t), \tag{1}$$
$$\frac{dI(t)}{dt} = \beta S(t)I(t) - \mu I(t), \tag{2}$$
$$\frac{dR(t)}{dt} = \mu I(t), \tag{3}$$

where β is the infection rate, while μ is the recovery rate; the variables S, I and R represent the fraction of susceptible, infected and recovered (or removed) individuals within the population, and $S(t)+I(t)+R(t) = 1$ at all times t. At the onset of a new epidemic, S equals approximately the entire population, and thus from (2) it holds that $I(t) = I_0 e^{(\beta-\mu)t} = I_0 e^{\mu(\mathcal{R}_0-1)t}$, where I_0 represents the initial number of infected $I_0 = I(0)$ and $\mathcal{R}_0 = \beta/\mu$ is the *basic reproduction number*, i.e. the average number of secondary cases produced by an infectious individual when $S \approx 1$. Clearly, when $\mathcal{R}_0$ is greater than 1, there is an exponential increase in the number of infected individuals during the early days of the epidemic. The same equation can also be used to estimate the point at which the number of newly infected individuals begins to decrease, $S(t) = 1/\mathcal{R}_0$. At this point, the given population has reached what is known as *herd immunity* (Fine, Eames, & Heymann, 2011).

To account for the latency time, an extended version of the SIR model, called the SEIR model, includes an extra compartment for *Exposed* (E) individuals, who have been infected but are not yet infectious, and are transitioning into the Infectious compartment at a fixed rate.

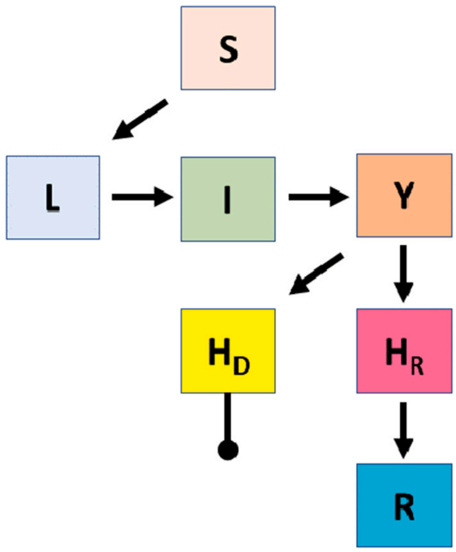

Fig. 2. Illustration of an extended compartmental epidemic model with seven compartments used in Riley, Fraser, Donnelly, Ghani, Abu-Raddad, et al. (2003) to model SARS : Susceptible (S), Latent (L), Asymptomatic and potentially infectious (I), Symptomatic Diagnosed (Y), Hospitalized that die (H_D), Hospitalized that recover (H_R) and Recovered R .

3.3. Extended compartmental models

To model the specific dynamics of a given infectious disease, extended compartmental models including additional compartments and transitions are often proposed. In particular, it is possible to consider symptomatic and asymptomatic compartments, vaccinated and unvaccinated, the possibility of reinfection after recovery, quarantined individuals, hospitalized, etc. Comprehensive books, surveys and works on compartmental models and their extensions are (Anderson & May, 1991; Brauer & Castillo-Chavez, 2012; Breda, Diekmann, de Graaf, Pugliese, & Vermiglio, 2012; Capasso & Serio, 1978; Diekmann & Heesterbeek, 2000; Gumel et al., 2004; Hethcote, 2000).

The number of compartments required to model a disease depends on a variety of factors. For example, when modelling the dynamics of a new disease, for which no vaccine is available, it makes no sense to consider the vaccinated group. However, in other cases, as when modelling seasonal influenza, it is relevant to distinguish between vaccinated and unvaccinated populations (Brauer, 2008). Many diseases, like malaria, West Nile virus, etc., are transmitted not directly from human to human but by infected animals (usually insects) (Taylor, Latham, & Woolhouse, 2001). For these cases, the corresponding animal compartments are included in the model. Another relevant factor influencing what compartments to include in a model is the quantity and quality of available data. Complex models require more data to fit the parameters, so in the early stages of a new disease outbreak simple compartmental models are often employed.

Many applications of extended compartmental models can be found in the literature. For example, in Riley et al. (2003), the authors use a dynamical compartmental model to analyse the effective transmission rate of the SARS epidemic in Hong Kong. The model consists of 7 compartments: Susceptible individuals (S) become infected and enter a latent state (L). They then progress to a short asymptomatic and potentially infectious state (I) followed by a symptomatic state that leads to diagnose (Y) and hospitalization. It is assumed that every symptomatic case is eventually hospitalized and either dies (H_D) or, after treatment in the hospital (H_R), recovers (R) (see Fig. 2). In Chowell, Blumberg, Simonsen, Miller, and Viboud (2014), a stochastic SEIR model is used to estimate the basic reproduction number of MERS-CoV in the Arabian Peninsula, distinguishing between cases transmitted by animals and secondary cases.

In the spread of the Covid-19 pandemic, asymptomatic infected individuals play a crucial role (see Ferretti et al., 2020; Giordano et al., 2020); the large prevalence of asymptomatic infections makes it harder to detect all cases and, thus, timely break the contagion chain. In Giordano et al. (2020), a SIDARTHE model with eight compartments is proposed. This model distinguishes between asymptomatic and symptomatic infected, as well as detected and undetected infection cases, and partitions the population into Susceptible, Infected (asymptomatic infected, undetected), Diagnosed (asymptomatic infected, detected), Ailing (symptomatic infected, undetected), Recognized (symptomatic infected, detected), Threatened (infected with life-threatening symptoms, detected), Healed (recovered) and Extinct (dead) individuals. In Ferretti et al. (2020), the epidemic model includes a transmission rate β that takes into account the contributions of asymptomatic, presymptomatic and symptomatic transmissions, as well as environmental transmission. In both works, the results indicate that the contribution of asymptomatic infected to $\mathcal{R}_0$ is higher than that of symptomatic infected and other transmission modalities. In fact, symptomatic infected are often rapidly detected and isolated.

3.3.1. Age-structured models

Age-structured epidemic models incorporate heterogeneous, age-dependent contact rates between individuals (Del Valle, Hyman, & Chitnis, 2013). In Safi, Gumel, and Elbasha (2013) and Xue-Zhi, Gupur, and Guang-Tian (2001), stability results for different age-structured SEIR models are given. For Covid-19, an age-structured model, aiming at estimating the effect of physical distancing measures in Wuhan, is presented in Prem, Liu, Russell, Kucharski, Eggo, et al. (2020). In Salje, Kiem, Lefrancq, Courtejoie, et al. (2020), a stratified approach is used to model the epidemic in France.

3.4. Seasonal behaviour

Some works have studied the influence of temperature and humidity on the spread of viruses (Grassly & Fraser, 2006; Waikhom, Jain, & Tegar, 2016). In the case of Covid-19, it has been reported that both variables have an effect on the basic reproduction number $\mathcal{R}_0$ (Iqbal et al., 2020; Mecenas et al., 0000). This influence might be included in the epidemic models to capture the seasonal behaviour of Covid-19; for instance, by considering the parameters β and μ as functions of both temperature and relative humidity. Yet, it remains unclear under which circumstances seasonal and geographic variations in climate can substantially alter the dynamics of a given pandemic, specially in the case of high susceptibility (Baker, Yang, Vecchi, Metcalf, & Grenfell, 2020).

3.5. Spatial epidemiology

Compartmental models are well-suited to describe the evolution of epidemics in a single, well-mixed population where each individual is assumed to interact with every other at a common rate (homogeneous contacts). While this can be a reasonable approximation in some contexts, it is not appropriate to study the global spread of a pandemic over a large, geographically dispersed population. In the last decades, compartmental models have been successfully extended to spatial epidemiological models in order to analyse spreading phenomena where spatial patterns need to be more accurately described. Graphs and networks have often been used to achieve this, see for instance (House, 2012; Keeling & Eames, 2005; Kiss, Miller, & Simon, 2017; Lewien & Chapman, 2019; Mei, Mohagheghi, Zampieri, & Bullo, 2017; Nowzari et al., 2016; Ogura & Preciado, 2016a, 2016b; Paré, Beck, & Başar, 2020; Paré, Beck, & Nedić, 2018; Pastor-Satorras, Castellano, Van Mieghem, & Vespignani, 2015; Zino, Rizzo, & Porfiri, 2020). Three widely used classes of models are described in the following sub-subsections.

3.5.1. Meta-population models

Meta-population models integrate two types of dynamics: one related to the disease, typically driven by a compartmental model, and the other to the mobility of individuals (agent-based model) across the sub-populations that build the meta-population under analysis (Ball, Britton, House, Isham, Mollison, et al., 2015; Grenfell & Harwood, 1997). As a representative example, in Brockmann and Helbing (2013) the authors introduce the notion of effective distance to capture the spatio-temporal dynamics of epidemics, combining the SIR model of $n = 1, 2, \ldots, p$ populations with mobility among them. The resulting model for each population is

$$\begin{aligned}
\frac{dS_n(t)}{dt} &= -\beta S_n(t) I_n(t) + \sum_{m \neq n} (w_{nm} S_m - w_{mn} S_n) \\
\frac{dI_n(t)}{dt} &= \beta S_n(t) I_n(t) - \mu I_n(t) \\
&\quad + \sum_{m \neq n} (w_{nm} I_m - w_{mn} I_n), \\
\frac{dR_n(t)}{dt} &= \mu I_n(t) + \sum_{m \neq n} (w_{nm} R_m - w_{mn} R_n),
\end{aligned}$$

where w_{nm} is the per capita traffic flux from population m to population n. In Aleta and Moreno (2020), the authors use a SEIR compartmental model together with stochastic data-driven simulations to capture the mobility in all Spanish provinces. The work focuses on evaluating the effectiveness of containment measures in Spain on February 28th, when a few dozen cases of Covid-19 had been detected. Meta-population models to capture the spatio-temporal dynamics of the Covid-19 epidemics in Italy have been proposed in Bertuzzo, Mari, Pasetto, Miccoli, Casagrandi, et al. (2020), Della Rossa, Salzano, Di Meglio, De Lellis, Coraggio, et al. (2020) and Gatto et al. (2020). By capturing both temporal and spatial evolution of epidemics, meta-population models are also capable of forecasting the effectiveness of mobility restrictions.

3.5.2. Social networks models

Social network models consider that transmission can only occur along linked or connected individuals (El-Sayed, Scarborough, Seemann, & Galea, 2012), which allows to explicitly model heterogeneity in contact patterns. Small-world networks have been used in combination with compartmental models to model disease transmission of SARS (Small & Tse, 2005) and Covid-19 (Thurner, Klimek, & Hanel, 2020), and also to assess the efficacy of contact tracing (Kiss, Green, & Kao, 2006). In general, network models produce a more accurate prediction of the disease spread (Paré et al., 2020). In particular, the use of homogeneous compartmental models in populations with heterogeneous contacts tends to underestimate disease burden early in the outbreak and overestimate it towards the end, although for certain kinds of networks compartmental models can be modified to prevent this problem (Bansal, Grenfell, & Meyers, 2007). Another interesting aspect of studying epidemic spreads with network models is the observation of a percolation phase transition (House, 2012; Pastor-Satorras et al., 2015), i.e., an abrupt change in the global dynamics of the epidemics. Percolation theory has been widely studied in random networks (Albert & Barabási, 2002). In the context of epidemic modelling, the transition phase occurs where isolated clusters of infected people join to form a giant component that is able to infect many people (Harding, Spinney, & Prokopenko, 2020).

3.5.3. Self-exciting spatio-temporal point processes

In epidemiology, it is natural to register each new infection event with a pair (t, x) in which t refers to time and x to location. The underlying stochastic model for this kind of data is called spatio-temporal point process (Diggle, 2006). Since each infection event potentially causes new ones, an epidemic can be modelled as a self-exciting spatio-temporal point process in which the rate of infections depends on the past history of the process (Reinhart et al., 2018; Zino et al., 2020). In this setting, the objective is to estimate an intensity function which predicts the rate of infections at any spatial location x and time t (Diggle, 2006; Waller, 2010). This modelling framework, which constitutes a generalization of Hawkes processes (Hawkes, 1971), permits the incorporation of the distributions of the duration of incubation, pre-symptomatic and asymptomatic phases, along with the modulating effect of time-varying counter-measures and detection efforts (Garetto, Leonardi, & Torrisi, 2021).

3.6. Computer-based models

Computer-based simulation methods to predict the spread of epidemics can take into account numerous factors, such as heterogeneous behavioural patterns, mobility patterns, both at long and short scales, demographics, epidemiological data, or disease-specific mechanisms (Helbing, Brockmann, Chadefaux, Donnay, Blanke, et al., 2015; Marathe & Vullikanti, 2013). The real-world accuracy of mathematical and computational models used in epidemiology has been considerably improved by the integration of large-scale data sets and explicit simulations of entire populations down to the scale of single individuals. These computational tools have recently gained importance in the field of infectious disease epidemiology, by providing rationales and quantitative analysis to support decision-making and policy-making processes (Tizzoni, Bajardi, Poletto, Ramasco, Balcan, et al., 2012). As a representative example, the Global Epidemic and Mobility simulation framework (GLEAM) allows performing stochastic simulations of a global epidemic with different global–local mobility patterns, as well as data regarding demographics or hospitalization (Van den Broeck, Gioannini, Gonçalves, Quaggiotto, Colizza, et al., 2011).

However, detailed simulation-based methods depend on a significant number of parameters, which need to be chosen and fixed for a specific simulation. This is especially difficult in the early days of an epidemic outbreak. Furthermore, these approaches might not reveal which factors are actually relevant in the spread of epidemics. Simpler data-driven tools have also been developed to overcome these difficulties (Marathe & Vullikanti, 2013).

3.7. Modelling the effect of containment measures

Controlling an emerging infectious disease requires both the prompt implementation of countermeasures and the rapid assessment of their efficacy (Brauner, Mindermann, Sharma, Johnston, Salvatier, et al., 2020; Cauchemez, Boëlle, Thomas, & Valleron, 2006; Chowell, Fenimore, Castillo-Garsow, & Castillo-Chavez, 2003; Guan et al., 2020; Gumel et al., 2004; Haug, Geyrhofer, Londei, Dervic, Desvars-Larrive, et al., 2020). In what follows, we enumerate the most relevant non-pharmaceutical interventions, focusing on different research works that assess their efficacy.

- **Quarantine**: Quarantine of diagnosed cases, or probably infected, is crucial in every epidemic outbreak. In order to model the effect of quarantine, specific compartments are included in the epidemic models for SARs (Chowell et al., 2003; Gumel et al., 2004). If a significant fraction of the infected population is not diagnosed (or diagnosed with a significant delay), then the modelling is harder and non-diagnosed groups are included in the models (Ansumali, Kaushal, Kumar, Prakash, & Vidyasagar, 2020; Giordano et al., 2020; May & Anderson, 1987).
  Quarantine of a whole population (i.e., lockdown) is the most extreme measure in the scope of physical distancing/mobility restrictions. The extreme impact of Covid-19 yield to the quarantine of the epicentre of the pandemic (Wuhan) on January 24th, 2020, and the same measures were subsequently adopted in different countries of Europe and America (Gatto et al., 2020). In this case, the effect of a lockdown can be modelled by means of time-varying epidemic models, see e.g. Calafiore, Novara, and Possieri (2020).

- **Physical distancing**: Physical (or social) distancing is another measure promoted by governments, public and private institutions in an attempt to reduce disease transmission (Morato, Bastos, Cajueiro, & Normey-Rico, 2020; Prather, Wang, & Schooley, 2020; Prem et al., 2020). Population-wide wearing of masks, capacity reduction on public transport, reducing or stopping the activity in educational institutions or factories are examples of this. In Maharaj and Kleczkowski (2012), the authors conduct a simulation-based analysis to determine the effects of physical distancing both in public health and in the economy. Two social network models (regular and small-world networks) are combined with a compartmental SIR model, and the economic impact takes into account the costs of individuals falling ill and the cost of a reduction in social contacts.
- **Mobility restrictions**: Governments often introduce long-range or local mobility restrictions aimed at reducing disease transmission. Spatial epidemiology is particularly useful to model the effects of such measures. For instance, in Thurner et al. (2020), the authors show, by means of a small-world network model, that the onset of mobility restrictions influences the final size of the outbreak, which is well below the levels of herd immunity.
- **Proactive testing**: Proactive testing of asymptomatic individuals is very relevant for the monitoring and control of the Covid-19 pandemic (World Health Organization (WHO), 2020), since it allows to isolate infectious individuals and implement contact tracing strategies, which have been shown to be crucial for an effective control of the pandemic (Giordano et al., 2020).
- **Contact tracing**: Contact tracing is a widely used epidemic control measure that aims to identify and isolate infected individuals by following the social contacts of individuals that are known to be infectious. A review of contact-tracing based epidemic models for SARS and MERS can be found in Kwok et al. (2019). In Kiss et al. (2006), a small-world, free-scale network model is combined with a compartmental model to assess the efficacy of contact tracing.

3.8. Fitting epidemic models to data

Dynamic epidemiological models rely on a set of parameters that have to be tuned in order to provide realistic predictions and/or infer essential features, such as the (time-varying) effective reproduction number (Cori, Ferguson, Fraser, & Cauchemez, 2013), or the latent period. Fitting epidemic models to data is a fundamental problem in epidemiology that can be approached in different ways. We can distinguish between classical methods, in which the parameters of the model are unknown but fixed, and Bayesian methods, in which they are assumed to be random variables (Kypraios, Neal, & Prangle, 2017). Another classification follows from the accessibility to the populations considered in the compartments of the model:

- **Full access** to the evolution of the number of cases in each compartment: In most models, the parameters that determine the dynamics multiply linear or bi-linear terms, depending on the current number of cases in each compartment. This means that a (vector) equality constraint, that depends (bi-) linearly on the parameters to fit, can be obtained at each sample time. In the case of linear constraints, standard linear identification techniques, such as least-square methods, can be applied to estimate the parameters that best fit the model to the data. See, for example, Allman and Rhodes (2004) and Martcheva (2015, Chapter 6).
- **Partial access** to the number of cases in each compartment: In many situations, there are no available time series for one or more of the groups considered in the model. This complicates the data-fitting process considerably because it is no longer possible to obtain, in a simple way, the equality constraints described in the full access case. The standard approach in this case is to resort to non-linear identification techniques (see Abreu & Dutra, 2020; Schön et al., 2011). In this context, Monte Carlo based methods (e.g. Markov Chain Monte Carlo and Sequential Monte Carlo algorithms) play a crucial role in addressing the challenges that lie in reconciling predictions and observations (McKinley, Cook, & Deardon, 2009).

3.8.1. Sensitivity analysis

Sensitivity analysis (SA) is the study of how the uncertainty in the output of a model (numerical or otherwise) can be apportioned to different sources of uncertainty in the model input (Saltelli, 2002). See the review paper (Qian & Mahdi, 2020) on the use of this technique in the context of biological sciences. A monovariate and multivariate sensitivity analysis for a data-fitted SARS model is given in Álvarez, Donado-Campos, and Morilla (2015). The use of SA is common in many research papers on modelling Covid-19 (see e.g. Fang, Nie, & Penny, 2020; Salje et al., 2020).

3.8.2. Validation and model selection

The ultimate test of the validity of any model is that its behaviour is in accord with real data. Because of the simplifications introduced in any mathematical model of a biological system, we must expect some divergence between the results of a model and reality, even for the most carefully collected data and most detailed model. Different questions arise in this context: (i) How can we determine if a model describes data well? (ii) How can we determine the parameter values in a model that are appropriate for describing real data? These questions are too broad to have a single answer (Allman & Rhodes, 2004; Vittinghoff, Glidden, Shiboski, & McCulloch, 2012).

Epidemic models depend on their data calibration. However, many possible models are potentially suited to analyse the spread of a pandemic in a given moment. The models are inherently linked to the goal for which they were envisaged. For a given goal (for example second outbreak detection), different models can be considered. Model selection techniques are used on a regular basis in epidemiology (Portet, 2020). They address the problem of choosing, among a set of candidate models, the most suitable for a given purpose (Burnham & Anderson, 2010). The selection is based on different aspects: (i) How the calibrated model is able to reconcile and match observations and (ii) the complexity of the model. Under similar adjustment to observations, simpler models are preferred since they are more robust from an information-theoretic point of view (Huyvaert, Burnham, & Anderson, 2011).

There are often different sets of parameters yielding a similar fit to data, but providing significantly different estimations of the main characteristics of the spread of the epidemic (like peak size, reproduction number, etc.). This issue is known as non-identifiability (Gustafsson, Soltesz, Jaldén, & Bernhardsson, 2020; Roda, Varughese, Han, & Li, 2020). Identifiability issues may lead to inferences that are driven more by prior assumptions than by the data themselves (Lintusaari, Gutmann, Kaski, & Corander, 2016). There are some approaches to address this difficulty. The first one is to resort to simplified models (SIR and SEIR models, for example) in which the number of parameters to adjust is small, see Postnikov (2020) and Roda et al. (2020). The second one is to use data from different regions in a not aggregated way, which reduces the probability of parametric over-fitting (Fiacchini & Alamir, 2021). In this context, model selection theory provides systematic methodologies to determine which model structure best suits the purposes of the model (Burnham & Anderson, 2010; Portet, 2020).

4. Model– Forecasting

The task of forecasting a time series can be stated as a supervised learning problem in which a number of temporal variables (also called predictors or *features* in the machine learning literature) are used to learn a model able to predict the future value of an output variable of interest (Bishop, 2006). In our context, we focus on forecasting methods aiming to predict the future evolution of epidemiological variables (Chowell, Tariq, & Hyman, 2019; Suárez, Pérez, Rivera, & Martínez, 2017). We find in the literature numerous approaches to forecast temporal variables describing the evolution of Covid-19 (Calafiore et al., 2020; Petropoulos & Makridakis, 2020; Tayarani-N, 2020), from black-box approaches to estimates based on learning the internal parameters of compartmental epidemic models. Forecasting in the context of global pandemics faces many difficulties (Ioannidis, Cripps, & Tanner, 2020) and requires the implementation of validation and sensitivity analysis (Burnham & Anderson, 2010). We now introduce some considerations that should be taken into account in order to select and train a suitable forecasting model.

First, we start with some statistical considerations:

- Frequentist versus Bayesian statistical methods: In the former, probabilities are assigned according to experiment repetition and occurrence. In the latter, the parameters of a model are learned using Bayes' theorem and prior knowledge about the probability distributions of unknown variables (Bonamente, 2013).
- Parametric versus non-parametric approaches: In the former, we assume a parametric function mapping past variables input into future predictions. This function contains several unknown parameters that are learned using historic time series. In the non-parametric approach, we do not assume such a parametric function (Malley, Kruppa, Dasgupta, Malley, & Ziegler, 2012); for example, one can make future predictions for a given time series by analysing the behaviour of historic past behaviours resembling the behaviour of the time series under consideration.

Other considerations to keep in mind are:

- The model should be trained with reliable data. If the available data is poor, the forecasts produced will be unreliable. In this direction, data-cleaning techniques such as data reconciliation, standardization, filtering, and outlier detection should be utilized to improve the quality of the input data collected (Albuquerque & Biegler, 1996).
- The amount of data collected should be appropriate for the forecasting technique under consideration. For instance, black-box models, such as deep learning, require vast amounts of data compared with compartmental models (Torrealba-Rodriguez, Conde-Gutiérrez, & Hernández-Javier, 2020; Yang, Zeng, Wang, Wong, Liang, et al., 2020); therefore, while dealing with relatively short time series, making predictions using compartmental models is more appropriate than using deep learning (and other black-box techniques).
- Learning procedures should include training, validation, and test phases executed separately. In other words, available data set should be divided into three parts, each one used for a different purpose. In the training stage, model parameters are learned using training data. In the validation step, one adjusts model hyper-parameters and performs comparisons with other competing approaches. Finally, the final test of a model should be carried out with data that has not been used during the training or validation phases (Burnham & Anderson, 2010).
- Interpretability of the model. While deep learning (and other black-box techniques) may produce high-quality predictions, the obtained model is hard to interpret; in other words, we typically do not have an intuitive understanding of why the model is making a prediction (Arrieta, Díaz-Rodríguez, Del Ser, Bennetot, Tabik, et al., 2020). However, when policy-makers make critical decisions based on the forecast of a model, it is important for them to understand why the model is behaving in a certain way. Therefore, it is sometimes reasonable to use more interpretable models, with parameters having a clear physical/biological interpretation, even at the expense of having a lower performance than with black-box approaches.

5. Model– Impact assessment tools

In order to design effective control strategies, it is important to define the control goals first. In the context of the current pandemic, the ultimate goal is to maintain the spread of the virus within an adequate threshold (e.g., a low level of infection cases Priesemann et al., 2021), while minimizing the economic and social impacts of the interventions. Once this goal is quantified in terms of a cost function, we should then consider the types of interventions that can be taken to achieve our goals, as well as their associated costs. For example, there are several non-pharmaceutical interventions that can be used before a vaccine is widely available, such as physical distancing, border closures, school closures, isolation of symptomatic individuals, among others (see Section 3.7). Each of these interventions has an associated economic and social cost that should be considered while making a decision.

In order to use disciplined decision-making techniques, like the ones described below, one needs to clearly state the control objectives in a precise, quantitative form. Furthermore, it is necessary to quantify the impact and costs of all possible interventions, as well as their actuation limits (Brauner et al., 2020; Cauchemez et al., 2006). In this direction, we can quantify the impact of our actions by using suitable indexes such as the mean reproductive number, the mortality index, or the unemployment rate or public debt, to name just a few. Once the decision-maker has decided how to use these indexes to measure the impact and cost of potential actions, the decision-making process can be stated as a formal optimization problem with constraints. For example, the goal could be the minimization of a weighted index measuring the economic and social impact of our actions while keeping the reproductive number smaller than one.

We would like to remark that the numerical estimation of certain indexes is not an easy task because they require the design of data-driven strategies to assess the effect of each potential decision on different indexes. This could be done by means of predictive models and forecasting schemes analysed in the previous sections. In some cases, quantifying the effect of one intervention over the spread of an epidemic is a non-trivial task, since multiple interventions are typically present at the same time (Haushofer, Jessica, & Metcalf, 2020). In these scenarios, correlation analyses, like Pearson Correlation Coefficient (PCC), can be a naive way to assess causalities. Whenever possible, a reliable approach to establish causalities is to perform Randomized Control Trials (RTC) (Donner & Klar, 1994; Haushofer et al., 2020). In an RTC, a subset of randomly chosen individuals receives an intervention, while the rest of individuals receives no intervention. A standard statistical analysis of the observed results can be used to reliably evaluate the impact of this intervention. In the following subsections, we discuss a collection of indexes that could be included in the decision-making process of managing a pandemic.

5.1. Spread of the virus and reproductive number

It is natural to express the effectiveness of control strategies in terms of the effective reproductive number $\mathcal{R}(t)$. As introduced in Section 3, the basic reproduction number $\mathcal{R}_0$ determines the potential of an epidemic to spread exponentially at its early stage by measuring the number of secondary infections induced by a typical infectious individual in a population when everyone is susceptible. In contrast, when an epidemic is ongoing, the effective reproduction number, denoted by $\mathcal{R}(t)$, is used to quantify the average number of secondary infections per infectious case in a population with both susceptible and non-susceptible hosts. The effective reproduction number can be used to assess the ability of available control measures to contain the spread of an epidemic. By implementing interventions able to maintain $\mathcal{R}(t)$ below 1, the incidence of new infections decreases and the spread of epidemics fades with time. In Cori et al. (2013), the authors presented a software tool that was validated with 5 different epidemics, including SARS and influenza. This tool can be used to estimate the daily reproductive number $\mathcal{R}(t)$ and its variation in the presence of vaccination and super-spreading events.

For Covid-19, a numerical analysis of the effective reproductive number can be found in Fang et al. (2020), where, using real data and a SEIR model, the authors estimate $\mathcal{R}(t)$ in Wuhan and quantify the effectiveness of government measures. Based on the number of deaths, in Flaxman et al. (2020), the Imperial College Covid-19 Response Team used a semi-mechanistic Bayesian model to estimate the evolution of $\mathcal{R}(t)$ when non-pharmaceutical measures, such as physical distancing, self-isolation, school closure, public events banned, and complete lock-down, were recommended/enforced.

Limitations in the use of $\mathcal{R}(t)$ as an assessment tool stem from the unreliability of available data sources. As a result, determining the real value of $\mathcal{R}(t)$ is difficult. Other indirect measures, like the number of deaths, ICU cases, saturation of healthcare systems can also be employed to assess the current epidemic burden, as described in the next subsection.

5.2. Healthcare systems capacity

The capacity of a country to prevent, detect, and respond to epidemic outbreaks varies widely across countries. The preparedness and resilience of a healthcare system is a particularly relevant factor to analyse the future impact of an infectious outbreak in the population (Kandel, Chungong, Omaar, & Xing, 2020). The capacity of a healthcare system to continue delivering the same level (quantity, quality and equity) of basic healthcare services and protection to the population can severely degrade during an epidemic outbreak (Blanchet, Nam, Ramalingam, & Pozo-Martin, 2017; Emanuel, Persad, Upshur, Thome, Parker, et al., 2020). At the early stages of the Covid-19 outbreak, its virulence and high contagiousness quickly saturated the healthcare system of many cities around the world, resulting in higher mortality rates (Lai et al., 2020; Miller, Becker, Grenfell, & Metcalf, 2020). Furthermore, in countries with low capacity, like African and South American countries, saturation levels are reached even with a significantly smaller number of cases (Morato et al., 2020; Velavan & Meyer, 2020).

To limit the saturation of healthcare systems and plan resource distribution effectively, tools that assess the effect of different interventions on the magnitude and timing of the epidemic peak during first and secondary outbreaks (see Sections 3 and 4) are fundamental. However, precise tools to forecast these peaks are challenging to obtain, due to the limitations of the available data and the time-varying nature of the mitigation efforts and potential seasonal behaviour of a pandemic. Another issue is the uncertain adherence of the population to the interventions (see next subsection). In order to partially circumvent these issues, forecasts of cumulative disease burden are often looked for. While missing the intensity and timing of the peaks, these projections can at least allow to identify areas with heavy present and/or future pandemic incidence.

5.3. Adherence to interventions and social impact

Analyses of the relationship between risk perception and preventive behaviours can be found in the social epidemiology literature (Berkman, Kawachi, & Glymour, 2014; Lep, Babnik, & Hacin Beyazoglu, 2020). Moreover, the level of belief in the effectiveness of recommended behaviours and trust in authorities are important predictors of adherence to preventive behaviour (see the survey paper Bish & Michie, 2010), which is fundamental to deploy effective containment strategies (Moran et al., 2016). Here, we review some of the methodologies that could be helpful to design indexes aiming to monitor the adherence of the population to interventions and the social burden of the pandemic.

- **Social network analysis**: Online social networks, such as Facebook and Twitter, can be used to assess the impact of an infectious disease on society. People post in these social networks their feelings and worries. In Shanthakumar, Anand, and Ramesh (2020), 530,206 tweets in the USA were analysed to measure the social impact of Covid-19. The hashtags were classified into six categories, including general covid, quarantine, school closures, panic buying, lockdowns, frustration and hope. Thus, the number of tweets in each category can be used as a metric of social impact and overall sentiment. Similarly, Weibo microblogging social network was used in Li et al. (2020) to study the propagation of situational information related to Covid-19 in China. In Jiang, Chen, Yan, Lerman, and Ferrara (2020), the political polarization with regards to Covid-19 in the United States was analysed using a large Twitter dataset.
- **Search engines**: Online searches made by citizens in search engines, such as Google, Bing, or Baidu, can be used to measure the social impact of the epidemic in different locations. Normally, people try to find information about unknown diseases, drugs, vaccines, and treatments on the Internet. Along this line, the authors of Ginsberg et al. (2009) found a correlation between the relative frequency of certain queries in Google and the percentage of physician visits in which a patient presents influenza-like symptoms. Furthermore, other works have performed similar studies for other epidemics like Influenza Virus A (H1N1) (Cook, Conrad, Fowlkes, & Mohebbi, 2011). Regarding Covid-19, in Qin, Sun, Wang, Wu, Chen, et al. (2020), the Baidu engine is used to estimate the number of new cases of Covid-19 in China by the number of searches of five keywords, such as dry cough, fever, chest distress, coronavirus, and pneumonia. These five keywords showed a high correlation with the number of new cases.
- **News**: The number and the content of posts in online newspapers can also be used to assess the spread of the virus. Along this line, in Zheng, Du, Wang, Zhang, Cui, et al. (2020), Natural Language Processing (NLP) is used to extract the relevant features of news media in China.
- **Online questionnaires**: Another tool for measuring the social impact of a sanitary emergence is through online questionnaires such as Oliver, Barber, Roomp and Roomp (2020) (Spain, 146,728 participants), Qiu et al. (2020) (China, 52,730 participants) and Kleinberg, van der Vegt, and Mozes (2020) (UK, 2500 participants), which were implemented for the Covid-19 pandemic. These questionnaires allow to rapidly ask citizens multiple questions related to adherence to interventions, as well as psychological, social and economic impact, among other aspects. The main difficulty is to spread the questionnaires throughout the population, although social networks and web-based tools help to reach a large amount of population.
- **Mobility**: One of the most relevant indexes to understand the spread of a pandemic is mobility (Tizzoni et al., 2014). See Section 8.4 in Alamo, Reina, Mammarella et al. (2020) for a relation of mobility data sets in the context of Covid-19. The reduction of mobility is not only due to the imposed quarantines and lockdowns by governments but also due to the increasing population's fear of getting infected. In Engle, Stromme, and Zhou (2020), a perceived risk index of contracting Covid-19 is defined. This metric measures the individuals' perception of risk, and it is

determined by several variables, such as prevalence in both local and neighbouring locations, as well as population demographics. The results in Engle et al. (2020) indicate that a rise of local infection rate from 0% to 0.003% reduces mobility by 2.31%.

6. Manage– Managing and decision making

Deciding which of the far-reaching social and economic restrictions are the most effective to contain the spread of a disease, as well as the conditions under which they can be safely lifted, is one of the main goals of data-driven decision approaches to combat pandemics. Unlike an unmitigated pandemic, which spreads through the susceptible population out of control and eventually fades out, a mitigated pandemic presents waves. For example, a first wave grows when a very transmissible disease appears and decreases due to, for example, social distancing measures. However, as soon as social distancing measures are relaxed, a new wave can appear as long as we have a large number of individuals susceptible to the infection. To avoid recurrent waves, it is important to put in place surveillance systems and reactive mechanisms to reduce the potential burden of secondary epidemic waves. The decision-making process in this context is complex for many reasons:

- The presence of uncertainty in some crucial parameters characterizing the spread, such as seasonality, extent and duration of immunity of a new pandemic outbreak (Cobey, 2020; Kissler, Tedijanto, Goldstein, Grad, & Lipsitch, 2020).
- The difficulties in assessing the quantitative effect of a specific set of mitigation interventions on the effective reproduction number (Haushofer et al., 2020).
- The possibility of significant non-symptomatic transmission (as in the case of Covid-19), which renders some interventions less effective (Nishiura et al., 2020; Prather et al., 2020).
- The different regional incidence and adherence to interventions, which motivates spatially distributed decisions (Della Rossa et al., 2020; Sélley, Besenyei, Kiss, & Simon, 2015).
- The limited capacity of healthcare systems and the logistic challenges to address mass testing and mass vaccination.
- The necessity to mitigate the spread of the epidemic and, at the same time, reduce the socioeconomic impact.
- The time-delay induced by the incubation period of the disease, as well as the testing system, which does not allow for a prompt evaluation of the effect of the implemented actions.
- The difficulties of assessing in a quantitative way the disruptive effects of the undertaken measures on relevant macroeconomic variables.

In what follows, we analyse under which circumstances the epidemic can be mitigated (controllability of the pandemic). After that, we also discuss some methodologies that have been applied to combat infectious diseases, including the Covid-19 pandemic, and that could potentially be applied in the context of future pandemics. See also the review papers (Bussell, Dangerfield, Gilligan, & Cunniffe, 2019; Nowzari et al., 2016) for the use of control theory in the context of disease control, or (Ansumali et al., 2020; Paré et al., 2020; Preciado, Zargham, Enyioha, Jadbabaie, & Pappas, 2014) for the stability analysis of an epidemic.

6.1. Controllability of the pandemic

In this subsection, we review the most important factors determining the controllability of a pandemic: the aspects that have a relevant impact on the effective reproduction number. We link them with standard epidemic threshold theorems (e.g. Becker, 1977; Kermack & McKendrick, 1927; Whittle, 1955).

The epidemic threshold theorem of Kermack and McKendrick (1927), stated in 1927, and in particular its stochastic form as given by Whittle (1955) are fundamental to predict the size and nature of an infectious disease outbreak. The theorem indicates that, in homogeneously mixed communities, major epidemics can be prevented by keeping the product of the size of the susceptible population, the infection rate, and the mean duration of the infectious period, sufficiently small (Becker, 1977). We now discuss how to have an impact on each of these factors by means of control actions.

- **Size of the susceptible population:** The most effective way to reduce the susceptible population is by means of vaccines: vaccination campaigns increase herd immunity to a level that prevents further spread of the disease (Giordano et al., 2021; Scherer & McLean, 2002). Protection against an infectious disease can either be achieved by widespread vaccination or by repeated waves of infection over the years, until a large enough fraction of the population is immunized (Graham, 2020). However, an issue is the duration of the acquired immunity (Kissler et al., 2020), which in some infectious diseases, like the seasonal influenza, is not long enough to prevent recurring seasonal peaks (Cobey, 2020).
- **Infection rate:** This factor can be reduced by means of different control actions like physical distancing, mobility constraints or prohibition of certain activities (Kraemer, Yang, Gutierrez, Wu, Klein, et al., 2020; Ngonghala, Iboi, Eikenberry, Scotch, MacIntyre, et al., 2020). Depending on the seasonality and the specific demographic characteristics of a given population, the implemented measures can exhibit a time-varying effect on the infection rate (Cori et al., 2013; Fang et al., 2020). This might cause flows from tropical to temperate regions and back in each hemisphere's respective winter, limiting opportunities for global disease declines (Cobey, 2020) and implying that surveillance methods to detect a seasonal peak should be put in place.
- **Mean duration of the infectious period:** An effective way to reduce the infectious period consists in detecting infected cases and setting them into quarantine (Chowell et al., 2003). Challenges are posed by relatively short latent periods and by the presence of many asymptomatic cases, as in the Covid-19 pandemic; then, the impact of quarantine measures depends very much on how fast the detection is taking place. It has been shown that the probability of effectively controlling the outbreak decreases with long delays from symptom onset to isolation (Ferretti et al., 2020; Hellewell et al., 2020). A large prevalence of asymptomatic cases is indeed an issue due to the significant probability that transmission occurs before the onset of symptoms (when the median latent delay is smaller than the median incubation time), hence before the infection can be detected (Ferretti et al., 2020; Giordano et al., 2020).

6.2. Optimal allocation of limited resources

During a major health crisis, policy makers face the problem of optimally allocating limited resources, such as intensive care beds, ventilators, tests, high-filtration masks and Individual Protection Equipment (IPE), medicines, vaccines, etc. (Brandeau, Zaric, & Richter, 2003; Zaric & Brandeau, 2002). This fact has led to the problem of how to ethically and consistently allocate resources (Emanuel et al., 2020). In this context, the term "resource allocation problem" extends to issues such as where and when to allocate available resources.

A rigorous and precise allocation method should lead to the formulation of an optimization problem, composed of a mathematical formulation and efficient algorithms to obtain its numerical solution (Hansen & Day, 2011). In the mathematical model, resource allocations are the decision variables while the objectives are encoded in cost functions and equality or inequality constraints. For example, in Brandeau et al. (2003) and Zaric and Brandeau (2002), budget allocation models for multiple populations are provided. in Preciado, Zargham, Enyioha, Jadbabaie, and Pappas (2013), a network model is used to optimally allocate vaccines to eradicate an initial epidemic outbreak using linear matrix inequalities. An extension of this work to the case of directed and weighted networks can be found in Nowzari, Preciado, and Pappas (2015) and Preciado et al. (2014), where geometric programming was proposed to find an optimal solution. The same authors extend this last result to more general compartmental models in Nowzari, Preciado, and Pappas (2017). See also Hayhoe, Barreras, and Preciado (2020) for an application of geometric programming and multi-task learning in the context of Covid-19.

In Lampariello and Sagratella (2021) an optimization problem is formulated to find the number of tests to be performed in the different Italian regions in order to maximize the overall detection capabilities. The problem is a quadratic, convex optimization program. In Gollier and Grossner (2020), a group testing (Walter, Hildreth, & Beaty, 1980) approach is considered, and it is shown how the optimization of the group size can save between 85% and 95% of tests with respect to individual testing. See also Yilmaz et al. (2020) for a strategy that optimizes testing resources in the context of the Covid-19 pandemic.

Estimation, forecasting, and impact assessment techniques are often used to allocate resources, as they enable decision-makers to predict imbalances between supply and demand and to evaluate the overall efficiency of different alternatives of allocation. In Emanuel et al. (2020), the authors propose fair resource allocation guidelines in the time of Covid-19, which can be a reference for future pandemics. These guidelines come from four fundamental values: (i) maximizing the benefits, (ii) treating people equally, (iii) promoting instrumental value, and (iv) giving priority to the worst off. As a result, these guidelines are condensed in some recommendations:

1. To maximize the number of saved lives and life-years, with the latter metric subordinated to the former.
2. To prioritize critical interventions for healthcare workers and others who take care of sick patients because of their instrumental value.
3. For patients with similar prognoses, equality should be invoked and operationalized through random allocation.
4. To distinguish priorities depending on the interventions and the scientific evidence (e.g. vaccines could be prioritized for older persons while allocation ICU resources depending on prognosis might mean giving priority to younger patients).
5. People who participate in research to prove the safety and effectiveness of vaccines and therapeutics should receive some priority for interventions.

6.3. Trigger control

A strategy to modulate the intensity of non-pharmaceutical interventions consists in implementing a trigger mechanism to maintain the effective reproduction number close to one, avoiding the saturation of the healthcare system while reducing, when possible, the economic and social burden of the pandemic (Bin, Cheung, Crisostomi, Ferraro, Lhachemi, et al., 2021; Cauchemez et al., 2006; Della Rossa et al., 2020; Preciado et al., 2014). The on-line surveillance of the pandemic permits to estimate the time-varying value of the effective reproduction number. Three cases are possible:

The effective reproduction number is largely under 1: in this case, one could consider lifting one, or more non-pharmaceutical measures. However, other criteria should be met in order to implement a reduction on the confinements measures in a safe way (Priesemann et al., 2021). The three criteria highlighted by the European Commission to decide on the lifting of confinement measures for Covid-19 (European Commission, 2020) are:

1. Epidemiological criteria showing that the spread of the disease has significantly decreased and stabilized for a sustained period of time. This can, for example, be indicated by a sustained reduction in the number of new infections, hospitalizations and patients in intensive care.
2. Sufficient health system capacity, in terms of, for instance, occupancy rate for Intensive Care Units; adequate number of hospital beds; access to pharmaceutical products required in intensive care units; reconstitution of stocks of equipment; access to care, in particular for vulnerable groups; availability of primary care structures, as well as sufficient staff with appropriate skills to care for patients discharged from hospitals or maintained at home and to engage in measures to lift confinement (testing for example). This criterion is essential as it indicates that the different national healthcare systems can cope with future increases in cases after lifting the measures. At the same time, hospitals are likely to face a backlog of elective interventions that had been temporarily postponed during the pandemic peak. Therefore, healthcare systems should have recovered sufficient capacity in general, and not only related to the management of Covid-19.
3. Appropriate monitoring capacity, including large-scale testing capacity to detect and monitor the spread of the virus combined with contact tracing and possibilities to isolate people in case of resurgence and further spread of infections. Antibody detection capacities, e.g. in the case of Covid-19, provide complementary data on the share of the population that has successfully overcome the disease and eventually measure the acquired immunity.

The effective reproduction number has increased to a level clearly above 1: this would demand, in most cases, extremely prompt strengthening of the mitigation interventions. The stringency of the new measures should guarantee that the healthcare system is not overwhelmed by a new epidemic wave. This requires the implementation of forecasting tools that help decision-makers to determine the most suitable set of mitigating measures.

The effective reproductive factor is close to 1: in this case, a deeper analysis is required. The decision on whether to keep the same set of current mitigation measures or not will depend on the current fraction of infected population, the healthcare system capacity, and the potentiality of implementing in a short period of time a mitigating intervention, which is capable of bringing the effective reproductive number to admissible values. That is, the decision could be determined by the worst-case cost of delaying in one week the implementation of new measures. It is worth stressing that preemptive actions are always preferable: the earlier a countermeasure is adopted, the better in terms of its efficacy and potential to save lives (Giordano et al., 2021).

In order to develop a timely and appropriate response, different methodologies from the field of control theory are available (see the review paper Nowzari et al., 2016). Relying on Pontryagin's maximum principle, optimal control approaches have been proposed to design optimal treatment plans, or vaccination plans, that minimize the cost of the epidemics, including both the cost of infection and the cost of treatment or vaccination (Bloem, Alpcan, & Başar, 2009; Forster & Gilligan, 2007; Hansen & Day, 2011; Morton & Wickwire, 1974). Robust control approaches have also been proposed to control the spreading of infectious diseases, seen as uncertain dynamical systems (Lee & Leitmann, 1994; Leitmann, 1998). We provide more details on optimal control approaches in the following subsections.

6.4. Optimal control theory

Optimal control theory (Liberzon, 2012) can be applied to reduce in an effective way the burden of an epidemic (Lenhart & Workman, 2007; Martcheva, 2015, Chapter 9). The dynamic optimization techniques of the calculus of variations and of optimal control theory provide methods for solving planning problems in continuous time. The solution is a continuous function (or a set of functions) indicating the optimal path to be followed by the variables through time or space (Kamien &

Schwartz, 2012). We present here a common formulation of a continuous dynamical optimization problem (Hartl, Sethi, & Vickson, 1995, Section 2):

$$\begin{aligned} \min_{x(\cdot),u(\cdot)} \quad & S(x(T),T) + \int_0^T F(x(T),u(t),t)dt \\ s.t. \quad & x(0) = x_0, \\ & \dot{x} = f(x(t),u(t),t), \\ & g(x(t),u(t),t) \geq 0, && (4) \\ & h(x(t),t) \geq 0, && (5) \\ & a(x(T),T) \geq 0, && (6) \\ & b(x(T),T) = 0. && (7) \end{aligned}$$

In an epidemic control problem $x(t)$ represents the state of the pandemic at time t (for example, in terms of the populations of the different compartments), $u(t)$ is the control action which can be stated in a direct way (intensity of the interventions, number of vaccines, treatments), or in an indirect way (infection rate, immunologic protection, recovery rate). The differential equation $\dot{x}(\cdot) = f(\cdot,\cdot,\cdot)$ represents the epidemic model, inequality (4) allows us to incorporate (mixed) constraints on $x(\cdot)$ and $u(\cdot)$ whereas the (pure) constraint (5) can be used to impose limits on the size of the components of $x(\cdot)$. Finally, (6) and (7) are terminal constraints. The question of existence of optimal pairs $(x^*(\cdot), u^*(\cdot))$ was studied in Cesari (1965) and Filippov (1962). See also Hartl et al. (1995, Section 3) and the references therein.

Pontryagin's maximum principle provides necessary conditions that characterize the optimal solutions in the presence of inequality constraints (Kirk, 2004; Liberzon, 2012). These necessary conditions become sufficient under certain convexity conditions on the objective and constraint functions (Kamien & Schwartz, 1971; Mangasarian, 1966). In general, the solution of the optimal problem in the presence of nonlinear dynamics and constraints requires iterative numerical methods to solve the so-called Hamiltonian system, which is a two-point boundary value problem, plus a maximum (minimum) condition of the Hamiltonian (see e.g. Kirk, 2004, Chapter 6).

We now describe some examples of the use of optimal control theory in epidemic control. In Zaman, Kang, and Jung (2008), the dynamic optimal vaccination strategy for a SIR epidemic model is described. The optimal solution is obtained using a forward–backward iterative method with a Runge–Kutta fourth-order solver. An example of how to deploy scarce resources for disease control when epidemics occur in different but interconnected regions is presented in Rowthorn, Laxminarayan, and Gilligan (2009). The authors solve the optimal control problem of minimizing the total level of infection when the control actions are bounded.

In Youssef and Scoglio (2013) the authors apply Pontryagin's Theorem to obtain an optimal Bang-Bang strategy to minimize the total number of infection cases during the spread of SIR epidemics in contact networks. Optimal control theory is employed to design the best policies to control the spread of seasonal and novel A-H1N1 strains in Prosper, Saucedo, Thompson, Torres-Garcia, Wang, et al. (2011). An example of the use of optimal control theory to control the present Covid-19 pandemic is presented in Hayhoe et al. (2020) and Mandal et al. (2020), where the authors design an optimal strategy, for a five compartmental model, in order to minimize the number of infected cases while minimizing the cost of non-pharmaceutical interventions.

6.5. Model predictive control

Model predictive control (MPC) provides optimal solutions to a control decision problem subject to constraints (Camacho & Bordons, 2013; Rawlings, Mayne, & Diehl, 2017). MPC is a receding horizon methodology that involves repeatedly solving a constrained optimization problem, using predictions of future costs, disturbances, and constraints over a moving time horizon. In epidemic control, the aforementioned optimization problem is solved daily, or weekly, in order to decide the optimal control action (for example, the intensity of mitigation interventions, or the optimal allocation of resources). The output of the model predictive controller is adaptive in the sense that it takes into consideration the latest available information on the state of the pandemic (Bussell et al., 2019; Sélley et al., 2015). See, for example, Alleman, Torfs, and Nopens (2020), Köhler et al. (2020) and Morato et al. (2020) for MPC formulations that address the control of the Covid-19 pandemic. See Carli, Cavone, Epicoco, Scarabaggio, and Dotoli (2020) for a review paper on the application of MPC in the context of Covid-19 pandemic.

Because of the spatial clustered distribution of an epidemic, it is possible to apply specific control techniques from the field of distributed model predictive control (Christofides, Scattolini, Muñoz de la Peña, & Liu, 2013; Maestre & Negenborn, 2014). For example, non-linear model predictive control can be used to control the epidemics by solely acting upon the individuals' contact pattern or network (Sélley et al., 2015). Another example of distributed MPC in the control of epidemics is given in Köhler, Enyioha, and Allgöwer (2018), where the problem of dynamically allocating limited resources (vaccines and antidotes) to control an epidemic spreading process over a network is addressed.

6.6. Multi-objective control

Pareto optimality is used in multi-objective control problems with counter-balanced objectives. For instance, in a counter-balanced bi-objective problem, improving one objective implies to worsen the other one. Pareto optimality is based on the Pareto dominance, which defines that one solution dominates another one if it is strictly superior in all the objectives. Thus, the goal of the optimization algorithm is to find the Pareto front, which includes all dominant solutions of the control problem. Therefore, there is a set of optimal solutions instead of one optimal solution. The Pareto front is a useful tool for decision-makers that allows to visualize all the possible optimal solutions (for two objectives is a curve, for three objectives a plane, and so forth) and to evaluate the trade-off between different strategies. In the context of epidemic control (Sharomi & Malik, 2017), Pareto optimality has been used in Yousefpour, Jahanshahi, and Bekiros (2020) in a bi-objective control problem, the goals are related to epidemic measures like the number of cases and economic costs.

7. Conclusions

This review has presented a roadmap for controlling present and future pandemics from a data-driven perspective, based on three pillars: Monitoring, Modelling, and Managing. We have highlighted the interplay between data science, epidemiology, and control theory to address the different challenges raised by a pandemic.

Methodologies and approaches proposed for previous epidemics and the present Covid-19 pandemic have been reviewed, without claiming exhaustiveness, given the huge and continuously growing literature on this subject. Although the relevant body of literature is extremely large and many approaches have been studied in the past, further research is still needed. Implementing effective control strategies to mitigate a pandemic is difficult because of various reasons: (i) the unavoidable uncertainty affecting some crucial parameters that characterize the spread, including compliance issues due to the unpredictable human behaviour, (ii) the difficulties in assessing the quantitative effect of mitigating interventions, (iii) the impossibility of obtaining a prompt evaluation of the effect of the implemented interventions, due to the intrinsic time-delay, and at the same time the critical importance of acting quickly, due to the exponential nature of the spreading phenomenon: even a small delay in interventions can lead to a much heavier healthcare burden and a much larger death toll (see e.g. Giordano et al., 2021).

The first step for modelling different aspects of the pandemic is the processing of the available raw data to obtain consolidated time-series. In order to obtain predictive models, which are crucial for the decision-making process, we have discussed several techniques from epidemiology and machine learning. We have described the most relevant modelling and forecasting approaches, focusing on the adjustment of the prediction models to the available data, model selection and validation processes.

Different surveillance systems able to detect, or anticipate, possible recurring epidemic waves have been surveyed. These systems enable an immediate response that reduces the potential burden of the outbreak. Different methods from control theory can be applied to provide an optimal, robust and adaptive response to the time-varying incidence of an epidemic. These methods can be applied to the optimal allocation of resources, useful for testing campaigns and vaccination plans, and to determine trigger control schemes that modulate the stringency of the adopted interventions. We have reviewed the control-theory literature focused on the analysis and the design of feedback structures for the efficient control of an epidemic. Besides, we have also mentioned some techniques from distributed model predictive control that can be applied to control the temporal and spatial evolution of an infectious disease.

Preventing and controlling pandemics will be an increasingly important challenge in the future, due to the likelihood of new virus spillovers resulting from the increasing ecological footprint of humans. The systems and control community has powerful tools available to contribute and take on this fundamental challenge.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Abreu, D., & Dutra, A. (2020). Uncertainty estimation in equality-constrained MAP and maximum likelihood estimation with applications to system identification and state estimation. *Automatica*, *116*, Article 108935.

Abuin, P., Anderson, A., Ferramosca, A., Hernandez-Vargas, E. A., & Gonzalez, A. H. (2020). Characterization of SARS-CoV-2 dynamics in the host. *Annual Reviews in Control*, *50*, 457–468.

Alamo, T., Reina, D. G., Gata, P. M., Preciado, V. M., & Giordano, G. (2021). Challenges and future directions in pandemic control. *IEEE Control Systems Letters (L-CSS)*.

Alamo, T., Reina, D. G., Mammarella, M., & Abella, A. (2020). Covid-19: Open-data resources for monitoring, modeling, and forecasting the epidemic. *Electronics*, *9*(5), 827.

Alamo, T., Reina, D., & Millán, P. (2020). Data-driven methods to monitor, model, forecast and control Covid-19 pandemic: Leveraging data science, epidemiology and control theory. arxiv preprint arXiv:2006.01731.

Albert, R., & Barabási, A. L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics*, *74*(1), 47–97.

Albuquerque, J. S., & Biegler, L. T. (1996). Data reconciliation and gross-error detection for dynamic systems. *AIChE Journal*, *42*(10), 2841–2856.

Aleta, A., & Moreno, Y. (2020). Evaluation of the potential incidence of COVID-19 and effectiveness of containment measures in Spain: a data-driven approach. *BMC Medicine*, *18*, 1–12.

Alleman, T., Torfs, E., & Nopens, I. (2020). *Covid-19: from model prediction to model predictive control*: *Technical Report*, Ghent University.

Allman, E. S., & Rhodes, J. A. (2004). *Mathematical models in biology. an introduction*. Cambridge University Press.

Almocera, A. E. S., & Hernandez-Vargas, E. A. (2019). Coupling multiscale within-host dynamics and between-host transmission with recovery (SIR) dynamics. *Mathematical Biosciences*, *309*, 34–41.

Almocera, A. E. S., Nguyen, V. K., & Hernandez-Vargas, E. A. (2018). Multiscale model within-host and between-host for viral infectious diseases. *Journal of Mathematical Biology*, *77*, 1035–1057.

Althouse, B. M., Scarpino, S. V., Meyers, L. A., Ayers, et al. (2015). Enhancing disease surveillance with novel data streams: challenges and opportunities. *EPJ Data Science*, *4*(1), 1–8.

Álvarez, E., Donado-Campos, J., & Morilla, F. (2015). New coronavirus outbreak. Lessons learned from the severe acute respiratory syndrome epidemic. *Epidemiology and Infection*, *143*(13), 2882–2893.

Anderson, R. M., & May, R. M. (1991). *Infectious diseases of humans: Dynamics and control*. Oxford University Press.

Ansumali, S., Kaushal, S., Kumar, A., Prakash, M. K., & Vidyasagar, M. (2020). Modelling a pandemic with asymptomatic patients, impact of lockdown and herd immunity, with applications to SARS-CoV-2. *Annual Reviews in Control*, *50*, 432–447.

Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., et al. (2020). Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, *58*, 82–115.

Baker, R. E., Yang, W., Vecchi, G. A., Metcalf, C. J. E., & Grenfell, B. T. (2020). Susceptible supply limits the role of climate in the early SARS-CoV-2 pandemic. *Science*, *369*(6501), 315–319.

Ball, F., Britton, T., House, T., Isham, V., Mollison, D., et al. (2015). Seven challenges for metapopulation models of epidemics, including households models. *Epidemics*, *10*, 63–67.

Bansal, S., Grenfell, B. T., & Meyers, L. A. (2007). When individual behaviour matters: Homogeneous and network models in epidemiology. *Journal of the Royal Society Interface*, *4*(16), 879–891.

Barbarossa, M. V., & Röst, G. (2015). Immuno-epidemiology of a population structured by immune status: a mathematical study of waning immunity and immune system boosting. *Journal of Mathematical Biology*, *71*, 1737–1770.

Batlle, P., Bruna, J., Fernandez-Granda, C., & Preciado, V. M. (2020). Adaptive test allocation for outbreak detection and tracking in social contact networks. arXiv preprint arXiv:2011.01998.

Becker, N. G. (1977). On a general stochastic epidemic model. *Theoretical Population Biology*, *11*(1), 23–36.

Bellomo, N., Bingham, R., Chaplain, M. A. J., Dosi, G., Forni, G., et al. (2020). A multiscale model of virus pandemic: Heterogeneous interactive entities in a globally connected world. *Mathematical Models & Methods in Applied Sciences*, *30*(08), 1591–1651.

Berkman, L. F., Kawachi, I., & Glymour, M. M. (2014). *Social epidemiology*. Oxford University Press.

Bertuzzo, E., Mari, L., Pasetto, D., Miccoli, S., Casagrandi, R., et al. (2020). The geography of COVID-19 spread in Italy and implications for the relaxation of confinement measures. *Nature Communications*, *11*(4264).

Bettencourt, L. M., Ribeiro, R. M., Chowell, G., Lant, T., & Castillo-Chavez, C. (2007). Towards real time epidemiology: data assimilation, modeling and anomaly detection of health surveillance data streams. In *NSF workshop on intelligence and security informatics* (pp. 79–90). Springer.

Bin, M., Cheung, P., Crisostomi, E., Ferraro, P., Lhachemi, H., et al. (2021). Post-lockdown abatement of COVID-19 by fast periodic switching. *PLoS Computational Biology*, *17*(1), Article e1008604.

Bish, A., & Michie, S. (2010). Demographic and attitudinal determinants of protective behaviours during a pandemic: A review. *British Journal of Health Psychology*, *15*(4), 797–824.

Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.

Blanchet, K., Nam, S. L., Ramalingam, B., & Pozo-Martin, F. (2017). Governance and capacity to manage resilience of health systems: towards a new conceptual framework. *International Journal of Health Policy and Management*, *6*(8), 431–435.

Bloem, M., Alpcan, T., & Başar, T. (2009). Optimal and robust epidemic response for multiple networks. *Control Engineering Practice*, *17*(5), 525–533.

Bonamente, M. (2013). *Statistics and analysis of scientific data*. Springer.

Brandeau, M. L., Zaric, G. S., & Richter, A. (2003). Resource allocation for control of infectious diseases in multiple independent populations: Beyond cost-effectiveness analysis. *Journal of Health Economics*, *22*(4), 575–598.

Brauer, F. (2008). Compartmental models in epidemiology. In F. Brauer, P. van den Driessche, & J. Wu (Eds.), *Mathematical epidemiology* (pp. 19–79). Springer Berlin Heidelberg.

Brauer, F., & Castillo-Chavez, C. (2012). *Mathematical models in population biology and epidemiology* (2nd ed.). Springer.

Brauner, J. M., Mindermann, S., Sharma, M., Johnston, D., Salvatier, et al. (2020). Inferring the effectiveness of government interventions against COVID-19. *Science*, eabd9338. http://dx.doi.org/10.1126/science.abd9338.

Breda, D., Diekmann, O., de Graaf, W. F., Pugliese, A., & Vermiglio, R. (2012). On the formulation of epidemic models (an appraisal of Kermack and McKendrick). *Journal of Biological Dynamics*, *6*(sup2), 103–117.

Brockmann, D., & Helbing, D. (2013). The hidden geometry of complex, network-driven contagion phenomena. *Science*, *342*(6164), 1337–1342.

Burnham, K. P., & Anderson, D. R. (2010). *Model selection and multimodel inference: A practical information theoretic approach* (2nd ed.). Springer.

Bussell, E. H., Dangerfield, C. E., Gilligan, C. A., & Cunniffe, N. J. (2019). Applying optimal control theory to complex epidemiological models to inform real-world disease management. *Philosophical Transactions of the Royal Society B*, *374*(1776), Article 20180284.

Cai, L., Tuncer, N., & Martcheva, M. (2017). How does within-host dynamics affect population-level dynamics? insights from an immuno-epidemiological model of malaria. *Mathematical Methods in the Applied Sciences*, *40*(18), 6424–6450.

Calafiore, G. C., Novara, C., & Possieri, C. (2020). A time-varying SIRD model for the COVID-19 contagion in Italy. *Annual Reviews in Control*, *50*, 361–372.

Camacho, E. F., & Bordons, C. (2013). *Model predictive control*. Springer Science & Business Media.

Capasso, V., & Serio, G. (1978). A generalization of the Kermack-McKendrick deterministic epidemic model. *Mathematical Biosciences*, *42*(1), 43–61.

Carli, R., Cavone, G., Epicoco, N., Scarabaggio, P., & Dotoli, M. (2020). Model predictive control to mitigate the COVID-19 outbreak in a multi-region scenario. *Annual Reviews in Control*, 373–393.

Castiglione, F., & Celada, F. (2015). *Immune system modeling and simulation*. CRC Press, Taylor & Francis Group.

Cauchemez, S., Boëlle, P.-Y., Thomas, G., & Valleron, A.-J. (2006). Estimating in real time the efficacy of measures to control emerging communicable diseases. *American Journal of Epidemiology*, *164*(6), 591–597.

Cazelles, B., & Chau, N. P. (1997). Using the Kalman filter and dynamic models to assess the changing HIV/AIDS epidemic. *Mathematical Biosciences*, *140*(2), 131–154.

Cesari, L. (1965). Existence theorems for optimal solutions in pontryagin and Lagrange problems. *Journal of the Society for Industrial and Applied Mathematics Series A Control*, *3*(3), 475–498.

Chan, T., & King, C. (2011). Surveillance and epidemiology of infectious diseases using spatial and temporal clustering methods. In D. Zeng, H. Chen, C. Castillo-Chavez, W. Lober, & M. Thurmond (Eds.), *Infectious disease informatics and biosurveillance* (pp. 207–234). Springer.

Chowell, G., Blumberg, S., Simonsen, L., Miller, M. A., & Viboud, C. (2014). Synthesizing data and models for the spread of MERS-CoV, 2013: Key role of index cases and hospital transmission. *Epidemics*, *9*, 40–51.

Chowell, G., Fenimore, P. W., Castillo-Garsow, M. A., & Castillo-Chavez, C. (2003). SARS outbreaks in Ontario, Hong Kong and Singapore: the role of diagnosis and isolation as a control mechanism. *Journal of Theoretical Biology*, *224(1)*, 1–8.

Chowell, G., Tariq, A., & Hyman, J. M. (2019). A novel sub-epidemic modeling framework for short-term forecasting epidemic waves. *BMC Medicine*, *17*, 164.

Christofides, P. D., Scattolini, R., Muñoz de la Peña, D., & Liu, J. (2013). Distributed model predictive control: A tutorial review and future research directions. *Computers & Chemical Engineering*, *51*, 21–41.

Cobey, S. (2020). Modeling infectious disease dynamics. *Science*, *368*(6492), 713–714.

Cook, S., Conrad, C., Fowlkes, A. L., & Mohebbi, M. H. (2011). Assessing Google Flu trends performance in the United States during the 2009 influenza virus A (H1N1) pandemic. *PLoS ONE*, *6*(8), Article e2361.

Cori, A., Ferguson, N. M., Fraser, C., & Cauchemez, S. (2013). A new framework and software to estimate time-varying reproduction numbers during epidemics. *American Journal of Epidemiology*, *178*(9), 1505–1512.

Del Valle, S., Hyman, J. M., & Chitnis, N. (2013). Mathematical models of contact patterns between age groups for predicting the spread of infectious diseases. *Mathematical Biosciences and Engineering*, *10*(5–6), 1475–1497.

Della Rossa, F., Salzano, D., Di Meglio, A., De Lellis, F., Coraggio, M., et al. (2020). A network model of Italy shows that intermittent regional strategies can alleviate the COVID-19 epidemic. *Nature Communications*, *11*(5106), 1–9.

Diekmann, O., & Heesterbeek, J. A. P. (2000). *Mathematical epidemiology of infectious diseases: Model building, analysis and interpretation*. Wiley.

Diggle, P. J. (2006). *Spatio-Temporal Point Processes: Methods and Applications, Vol. 107* (p. 1). Chapman & Hall.

Donner, A., & Klar, N. (1994). Cluster randomization trials in epidemiology: theory and application. *Journal of Statistical Planning and Inference*, *42*(1), 37–56.

Drew, D. A., Nguyen, L. H., Steves, C. J., Menni, C., Freydin, M., et al. (2020). Rapid implementation of mobile technology for real-time epidemiology of COVID-19. *Science*, *368*(6497), 1362–1367.

Dubrawski, A. (2011). Detection of events in multiple streams of surveillance data. In D. Zeng, H. Chen, C. Castillo-Chavez, W. Lober, & M. Thurmond (Eds.), *Infectious disease informatics and biosurveillance* (pp. 145–171). Springer.

Durbin, J., & Koopman, S. J. (2012). *Time series analysis by state space methods* (2nd ed.). Oxford University Press.

El-Sayed, A. M., Scarborough, P., Seemann, L., & Galea, S. (2012). Social network analysis and agent-based modeling in social epidemiology. *Epidemiologic Perspectives and Innovations*, *9*(1), 1–9.

Elliot, A. J., Harcourt, S. E., Hughes, H. E., Loveridge, P., Morbey, R. A., et al. (2020). The COVID-19 pandemic: A new challenge for syndromic surveillance. *Epidemiology and Infection*, *148, e122*, 1–5.

Emanuel, E. J., Persad, G., Upshur, R., Thome, B., Parker, M., et al. (2020). Fair allocation of scarce medical resources in the time of Covid-19. *New England Journal of Medicine*, 1–7.

Engle, S., Stromme, J., & Zhou, A. (2020). Staying at home: Mobility effects of COVID-19. *SSRN Electronic Journal*, http://dx.doi.org/10.2139/ssrn.3565703.

European Commission (2020). *Joint European Roadmap towards lifting COVID-19 containment measures*: *Technical Report*, (pp. 1–15). European Commission.

Fang, Y., Nie, Y., & Penny, M. (2020). Transmission dynamics of the COVID-19 outbreak and effectiveness of government interventions: A data-driven analysis. *Journal of Medical Virology*, 1–15.

Farrington, C., Andrews, N. J., Beale, A., & Catchpole, M. (1996). A statistical algorithm for the early detection of outbreaks of infectious disease. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, *159*(3), 547–563.

Feng, Z., Cen, X., Zhao, Y., & Velasco-Hernandez, J. X. (2015). Coupled within-host and between-host dynamics and evolution of virulence. *Mathematical Biosciences*, *270*, 204–212.

Ferretti, L., Wymant, C., Kendall, M., Zhao, L., et al. (2020). Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science*, *368*(6491), eabb6936.

Fiacchini, M., & Alamir, M. (2021). The Ockham's razor applied to COVID-19 model fitting french data. *Annual Reviews in Control*.

Filippov, A. (1962). On certain questions in the theory of optimal control. *Journal of the Society for Industrial and Applied Mathematics Series A Control*, *1*(1), 76–84.

Fine, P., Eames, K., & Heymann, D. L. (2011). "Herd immunity": a rough guide. *Clinical Infectious Diseases*, *52*(7), 911–916.

Flaxman, S., Mishra, S., Gandy, A., Unwin, H. J. T., Mellan, T. A., Coupland, H., et al. (2020). Estimating the effects of non-pharmaceutical interventions on COVID-19 in europe. *Nature*, *584*(7820), 257–261.

Forster, G. A., & Gilligan, C. A. (2007). Optimizing the control of disease infestations at the landscape scale. *Proceedings of the National Academy of Sciences*, *104*(12), 4984–4989.

Gandolfi, A., Pugliese, A., & Sinisgalli, C. (2015). Epidemic dynamics and host immune response: a nested approach. *Journal of Mathematical Biology*, *70*(3), 399–435.

Garetto, M., Leonardi, E., & Torrisi, G. L. (2021). A time-modulated hawkes process to model the spread of COVID-19 and the impact of countermeasures. *Annual Reviews in Control*, http://dx.doi.org/10.1016/j.arcontrol.2021.02.002.

Gatto, M., Bertuzzo, E., Mari, L., Miccoli, S., Carraro, L., Casagrandi, R., et al. (2020). Spread and dynamics of the COVID-19 epidemic in Italy: Effects of emergency containment measures. *Proceedings of the National Academy of Sciences*, *117*(19), 10484–10491.

Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S., & Brilliant, L. (2009). Detecting influenza epidemics using search engine query data. *Nature*, *457*(7232), 1012–1014.

Giordano, G., Blanchini, F., Bruno, R., Colaneri, P., Di Filippo, A., Di Matteo, A., et al. (2020). Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy. *Nature Medicine*, *26*, 855–860.

Giordano, G., Colaneri, M., Di Filippo, A., Blanchini, F., Bolzern, P., De Nicolao, G., et al. (2021). Modeling vaccination rollouts, SARS-CoV-2 variants and the requirement for non-pharmaceutical interventions in Italy. *Nature Medicine*, http://dx.doi.org/10.1038/s41591-021-01334-5.

Glass, D. H. (2020). European and US lockdowns and second waves during the COVID-19 pandemic. *Mathematical Biosciences*, *330*, Article 108472.

Gollier, C., & Grossner, O. (2020). Group testing against Covid-19. *Covid Economics*, *2*.

Graham, B. S. (2020). Rapid COVID-19 vaccine development. *Science*, *368*(6494), 945–946.

Grassly, N. C., & Fraser, C. (2006). Seasonal infectious disease epidemiology. *Proceedings of the Royal Society B: Biological Sciences*, *273*(1600), 2541–2550.

Grenfell, B., & Harwood, J. (1997). (Meta) population dynamics of infectious diseases. *Trends in Ecology & Evolution*, *12*(10), 395–399.

Guan, L., Prieur, C., Zhang, L., Prieur, C., Georges, D., & Bellemain, P. (2020). Transport effect of COVID-19 pandemic in France. *Annual Reviews in Control*, *50*, 394–408.

Gulbudak, H., & Browne, C. J. (2020). Infection severity across scales in multi-strain immuno-epidemiological dengue model structured by host antibody level. *Journal of Mathematical Biology*, *80*(6), 1803–1843.

Gumel, A. B., Ruan, S., Day, T., Watmough, J., Brauer, F., van den Driessche, P., et al. (2004). Modelling strategies for controlling SARS outbreaks. *Proceedings of the Royal Society of London, Series B*, *271*(1554), 2223–2232.

Gustafsson, F., Soltesz, K., Jaldén, J., & Bernhardsson, B. (2020). Identifiability of non-pharmaceutical intervention effects on Covid-19 Spread in Europe. In *59th IEEE conference on decision and control*.

Handel, A., Longini, I. M., & Antia, R. (2010). Towards a quantitative understanding of the within-host dynamics of influenza a infections. *Journal of the Royal Society Interface*, *7*(42), 35–47.

Hansen, E., & Day, T. (2011). Optimal control of epidemics with limited resources. *Journal of Mathematical Biology*, *62*, 423–451.

Harding, N., Spinney, R. E., & Prokopenko, M. (2020). Phase transitions in spatial connectivity during influenza pandemics. *Entropy*, *22*(2), 133.

Hart, W. S., Maini, P. K., Yates, C. A., & Thompson, R. N. (2020). A theoretical framework for transitioning from patient-level to population-scale epidemiological dynamics: influenza a as a case study. *Journal of the Royal Society Interface*, *17(166)*, Article 20200230.

Hartl, R. F., Sethi, S. P., & Vickson, R. G. (1995). Survey of the maximum principles for optimal control problems with state constraints. *SIAM Review*, *37*(2), 181–218.

Haug, N., Geyrhofer, L., Londei, A., Dervic, E., Desvars-Larrive, A., et al. (2020). Ranking the effectiveness of worldwide COVID-19 government interventions. *Nature Human Behaviour*, *4*, 1303–3318.

Haushofer, J., Jessica, C., & Metcalf, E. (2020). Which interventions work best in a pandemic?. *Science*, *368*(6495), 1063–1065.

Hawkes, A. G. (1971). Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, *58*(1), 83–90.

Hayhoe, M., Barreras, F., & Preciado, V. (2020). Data-driven control of the COVID-19 outbreak via non-pharmaceutical interventions: A geometric programming approach. arXiv preprint arXiv:2011.01392.

He, X., Lau, E. H., Wu, P., Deng, X., Wang, J., et al. (2020). Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nature Medicine*, 1–4.

Heesterbeek, H., Anderson, R. M., Andreasen, V., Bansal, S., De Angelis, D., et al. (2015). Modeling infectious disease dynamics in the complex landscape of global health. *Science*, *347*(6227).

Heffernan, J. M., Smith, R. J., & Wahl, L. M. (2005). Perspectives on the basic reproductive ratio. *Journal of the Royal Society Interface*, *2*(4), 281–293.

Helbing, D., Brockmann, D., Chadefaux, T., Donnay, K., Blanke, U., et al. (2015). Saving human lives: What complexity science and information systems can contribute. *Journal of Statistical Physics*, *158*, 735–781.

Hellewell, J., Abbott, S., Gimma, A., Bosse, N. I., Jarvis, C. I., et al. (2020). Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. *The Lancet Global Health*, *8*(4), e488–e496.

Hernandez-Vargas, E. A., & Velasco-Hernandez, J. X. (2020). In-host mathematical modelling of COVID-19 in humans. *Annual Reviews in Control*, *50*, 448–456.

Hethcote, H. W. (2000). Mathematics of infectious diseases. *SIAM Review*, *42*(4), 599–653.

House, T. (2012). Modelling epidemics on networks. *Contemporary Physics*, *53*(3), 213–225.

Huyvaert, K. P., Burnham, K. P., & Anderson, D. R. (2011). AIC model selection and multimodel inference in behavioral ecology: some background, observations, and comparisons. *Behavioral Ecology and Sociobiology*, *65*, 23–35.

Ioannidis, J. P., Cripps, S., & Tanner, M. A. (2020). Forecasting for COVID-19 has failed. *International Journal of Forecasting*, http://dx.doi.org/10.1016/j.ijforecast.2020.08.004.

Iqbal, M. M., Abid, I., Hussain, S., Shahzad, N., Waqas, M. S., & Iqbal, M. J. (2020). The effects of regional climatic condition on the spread of COVID-19 at global scale. *Science of the Total Environment*, *739*, Article 140101.

Jiang, J., Chen, E., Yan, S., Lerman, K., & Ferrara, E. (2020). Political polarization drives online conversations about COVID-19 in the United States. *Human Behavior and Emerging Technologies*, *2*(3), 200–211.

Kamien, M. I., & Schwartz, N. L. (1971). Sufficient conditions in optimal control theory. *Journal of Economic Theory*, *3*(2), 207–214.

Kamien, M. I., & Schwartz, N. L. (2012). *Dynamic optimization: The calculus of variations and optimal control in economics and management* (4th ed.). Dover.

Kandel, N., Chungong, S., Omaar, A., & Xing, J. (2020). Health security capacities in the context of COVID-19 outbreak: an analysis of international health regulations annual report data from 182 countries. *The Lancet*, *395*(10229), 1047–1053.

Keeling, M. J., & Eames, K. T. (2005). Networks and epidemic models. *Journal of the Royal Society Interface*, *2*(4), 295–307.

Kermack, W. O., & McKendrick, A. G. (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London. Series A*, *115*(772), 700–721.

Kirk, D. E. (2004). *Optimal control theory. An introduction*. Dover.

Kiss, I. Z., Green, D. M., & Kao, R. R. (2006). Infectious disease control using contact tracing in random and scale-free networks. *Journal of the Royal Society Interface*, *3*(6), 55–62.

Kiss, I. Z., Miller, J., & Simon, P. L. (2017). *Mathematics of epidemics on networks: From exact to approximate models*. Springer.

Kissler, S. M., Tedijanto, C., Goldstein, E., Grad, Y. H., & Lipsitch, M. (2020). Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. *Science*, *368*(6493), 860–868.

Kleinberg, B., van der Vegt, I., & Mozes, M. (2020). Measuring emotions in the COVID-19 real world worry dataset. In *Proceedings of the 1st workshop on NLP for COVID-19 at ACL 2020*. Association for Computational Linguistics.

Köhler, J., Enyioha, C., & Allgöwer, F. (2018). Dynamic resource allocation to control epidemic outbreaks. a model predictive control approach. *Proceedings of the American Control Conference*, *2018-June*, 1546–1551.

Köhler, J., Schwenkel, L., Koch, A., Berberich, J., Pauli, P., & Allgöwer, F. (2020). Robust and optimal predictive control of the COVID-19 outbreak. *Annual Reviews in Control*.

Kraemer, M. U., Yang, C.-H., Gutierrez, B., Wu, C.-H., Klein, B., et al. (2020). The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science*, eabb4218.

Kwok, K. O., Tang, A., Wei, V. W., Park, W. H., Yeoh, E. K., & Riley, S. (2019). Epidemic models of contact tracing: Systematic review of transmission studies of severe acute respiratory syndrome and middle east respiratory syndrome. *Computational and Structural Biotechnology Journal*, *17*, 186–194.

Kypraios, T., Neal, P., & Prangle, D. (2017). A tutorial introduction to Bayesian inference for stochastic epidemic models using approximate Bayesian computation. *Mathematical Biosciences*, *287*, 42–53.

Lai, C.-C., Wang, C.-Y., Wang, Y.-H., Hsueh, S.-C., Ko, W.-C., & Hsueh, P.-R. (2020). Global epidemiology of coronavirus disease 2019 (COVID-19): disease incidence, daily cumulative index, mortality, and their association with country healthcare resources and economic status. *International Journal of Antimicrobial Agents*, *55*(4), Article 105946.

Lampariello, L., & Sagratella, S. (2021). Effectively managing diagnostic tests to monitor the COVID-19 outbreak in Italy. *Operations Research for Health Care*, *28*, Article 100287.

Lauer, S. A., Grantz, K. H., Bi, Q., Jones, F. K., Zheng, Q., et al. (2020). The incubation period of coronavirus disease 2019 (COVID-19). From publicly reported confirmed cases: Estimation and application. *Annals of Internal Medicine*, *172*(9), 577–582.

Lazarus, R. (2010). Automated, high-throughput surveillance systems for public health. In V. Sintchenko (Ed.), *Infectious disease informatics* (pp. 323–344). Springer.

Lee, C. S., & Leitmann, G. (1994). Control strategies for an endemic disease in the presence of uncertainty. In R. P. Agarwal (Ed.), *Recent trends in optimization theory and applications* (pp. 221–238).

Leitmann, G. (1998). The use of screening for the control of an endemic disease. In B. Birkhäuser (Ed.), *International series of numerical mathematics*: *vol. 124*, *Variational calculus, optimal control and applications* (pp. 291–300).

Lenhart, S., & Workman, J. T. (2007). *Optimal control applied to biological models* (p. 261). Chapman & Hall/CRC.

Lep, Z., Babnik, K., & Hacin Beyazoglu, K. (2020). Emotional responses and self-protective behavior within days of the COVID-19 outbreak: The promoting role of information credibility. *Frontiers in Psychology*, *11*, 1846.

Lewien, P., & Chapman, A. (2019). Time-scale separation on networks for multi-city epidemics. In *2019 IEEE 58th conference on decision and control (CDC)* (pp. 746–751).

Li, K., McCaw, J. M., & Cao, P. (2021). Modelling within-host macrophage dynamics in influenza virus infection. *Journal of Theoretical Biology*, *508*, Article 110492.

Li, R., Pei, S., Chen, B., Song, Y., Zhang, T., et al. (2020). Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2). *Science*, *368*(6490), 489–493.

Li, L., Zhang, Q., Wang, X., Zhang, J., Wang, T., Gao, T.-L., et al. (2020). Characterizing the propagation of situational information in social media during COVID-19 epidemic: A case study on weibo. *IEEE Transactions on Computational Social Systems*, *7*(2), 556–562.

Liberzon, D. (2012). *Calculus of variations and optimal control theory: A concise introduction*. Princeton University Press.

Lintusaari, J., Gutmann, M. U., Kaski, S., & Corander, J. (2016). On the identifiability of transmission dynamic models for infectious diseases. *Genetics*, *202*(3), 911–918.

Liu, Y., Gayle, A., Wilder-Smith, A., & Rocklöv, J. (2020). The reproductive number of COVID-19 is higher compared to SARS coronavirus. *Journal of Travel Medicine*, *27*(2), 1–4.

Maestre, J. M., & Negenborn, R. R. (2014). *Distributed model predictive control made easy, Vol. 69* (p. 601). Springer.

Maharaj, S., & Kleczkowski, A. (2012). Controlling epidemic spread by social distancing: do it well or not at all. *BMC Public Health*, *12*, 679.

Malley, J. D., Kruppa, J., Dasgupta, A., Malley, K. G., & Ziegler, A. (2012). Probability machines: consistent probability estimation using nonparametric learning machines. *Methods of Information in Medicine*, *51*(1), 74.

Mandal, M., Jana, S., Nandi, S. K., Khatua, A., Adak, S., & Kar, T. (2020). A model based study on the dynamics of COVID-19: Prediction and control. *Chaos, Solitons & Fractals*, *136*, Article 109889.

Mangasarian, O. L. (1966). Sufficient conditions for the optimal control of nonlinear systems. *SIAM Journal on Control*, *4*(1), 139–152.

Marathe, M., & Vullikanti, A. K. S. (2013). Computational epidemiology. *Communications of the ACM*, *56*(7), 88–96.

Martcheva, M. (2015). *An introduction to mathematical epidemiology*. Springer Science and Business Media LLC.

Matricardi, P. M., Dal Negro, R. W., & Nisini, R. (2020). The first, holistic immunological model of COVID-19: Implications for prevention, diagnosis, and public health measures. *Pediatric Allergy and Immunology*, *31*(5), 454–470.

May, R. M., & Anderson, R. M. (1987). Transmission dynamics of HIV infection. *Nature*, *326*(6109), 137–142.

McKinley, T., Cook, A. R., & Deardon, R. (2009). Inference in epidemic models without likelihoods. *International Journal of Biostatistics*, *5*(1), 1–37.

Mecenas, P., Travassos da Rosa Moreira Bastos, R., Rosário Vallinoto, A., & Normando, D. Effects of temperature and humidity on the spread of COVID-19: A systematic review. *PLoS ONE, 15*, (9), e0238339.

Mei, W., Mohagheghi, S., Zampieri, S., & Bullo, F. (2017). On the dynamics of deterministic epidemic propagation over networks. *Annual Reviews in Control*, *44*, 116–128.

Miller, I. F., Becker, A. D., Grenfell, B. T., & Metcalf, C. J. E. (2020). Disease and healthcare burden of COVID-19 in the United States. *Nature Medicine*, *26*, 1212–1217.

Moore, J. R., Ahmed, H., Manicassamy, B., Garcia-Sastre, A., Handel, A., & Antia, R. (2020). Varying inoculum dose to assess the roles of the immune response and target cell depletion by the pathogen in control of acute viral infections. *Bulletin of Mathematical Biology*, *62*(35).

Moran, K. R., Fairchild, G., Generous, N., Hickmann, K., Osthus, D., Priedhorsky, R., et al. (2016). Epidemic forecasting is messier than weather forecasting: The role of human behavior and internet data streams in epidemic forecast. *The Journal of Infectious Diseases*, *214*(suppl_4), S404–S408.

Morato, M. M., Bastos, S. B., Cajueiro, D. O., & Normey-Rico, J. E. (2020). An optimal predictive control strategy for COVID-19 (SARS-CoV-2) social distancing policies in Brazil. *Annual Reviews in Control*, *50*, 417–431.

Morton, R., & Wickwire, K. H. (1974). On the optimal control of a deterministic epidemic. *Advances in Applied Probability*, *6*(4), 622–635.

Nanni, M., Andrienko, G., Barabási, A.-L., Boldrini, C., Bonchi, F., et al. (2021). Give more data, awareness and control to individual citizens, and they will help COVID-19 containment. *Ethics and Information Technology*, 1–6.

Ngonghala, C. N., Iboi, E., Eikenberry, S., Scotch, M., MacIntyre, C. R., et al. (2020). Mathematical assessment of the impact of non-pharmaceutical interventions on curtailing the 2019 novel coronavirus. *Mathematical Biosciences*, *325*, Article 108364.

Nicola, M., O'Neill, N., Sohrabi, C., Khan, M., Agha, M., et al. (2020). Evidence based management guideline for the COVID-19 pandemic - Review article. *International Journal of Surgery*, *77*, 206–216.

Nishiura, H., Linton, N. M., & Akhmetzhanov, A. R. (2020). Serial interval of novel coronavirus (COVID-19) infections. *International Journal of Infectious Diseases*, *93*, 284–286.

Nowzari, C., Preciado, V. M., & Pappas, G. J. (2015). Optimal resource allocation for control of networked epidemic models. *IEEE Transactions on Control of Network Systems*, *4*(2), 159–169.

Nowzari, C., Preciado, V., & Pappas, G. (2016). Analysis and control of epidemics: A survey of spreading processes on complex networks. *IEEE Control Systems*, *36*(1), 26–46.

Nowzari, C., Preciado, V. M., & Pappas, G. J. (2017). Optimal resource allocation for control of networked epidemic models. *IEEE Transactions on Control of Network Systems*, *4*(2), 159–169.

Ogura, M., & Preciado, V. M. (2016a). Epidemic processes over adaptive state-dependent networks. *Physical Review E*, *93*(6), Article 062316.

Ogura, M., & Preciado, V. M. (2016b). Stability of spreading processes over time-varying large-scale networks. *IEEE Transactions on Network Science and Engineering*, *3*(1), 44–57.

Oliver, N., Barber, J. X., Roomp, K., & Roomp, K. (2020). Assessing the impact of the COVID-19 pandemic in Spain: Large-scale, online, self-reported population survey. *Journal of Medical Internet Research*, *22*(9), Article e21319.

Oliver, N., Lepri, B., Sterly, H., Lambiotte, R., Deletaille, S., et al. (2020). Mobile phone data for informing public health actions across the COVID-19 pandemic life cycle. *Science Advances*, *0764*, eabc0764.

Paré, P. E., Beck, C. L., & Başar, T. (2020). Modeling, estimation, and analysis of epidemics over networks: An overview. *Annual Reviews in Control*, *50*, 345–360.

Paré, P. E., Beck, C. L., & Nedić, A. (2018). Epidemic processes over time-varying networks. *IEEE Transactions on Control of Network Systems*, *5*(3), 1322–1334.

Pastor-Satorras, R., Castellano, C., Van Mieghem, P., & Vespignani, A. (2015). Epidemic processes in complex networks. *Reviews of Modern Physics*, *87*, 925–979.

Peeri, N. C., Shrestha, N., Rahman, M. S., Zaki, R., Tan, Z., Bibi, S., et al. (2020). The SARS, MERS and novel coronavirus (COVID-19) epidemics, the newest and biggest global health threats: what lessons have we learned?. *International Journal of Epidemiology*, *49*(3), 717–726.

Petropoulos, F., & Makridakis, S. (2020). Forecasting the novel coronavirus COVID-19. *PloS One*, *15*(3), Article e0231236.

Portet, S. (2020). A primer on model selection using the Akaike Information Criterion. *Infectious Disease Modelling*, *5*, 111–128.

Postnikov, E. B. (2020). Estimation of COVID-19 dynamics "on a back-of-envelope": Does the simplest SIR model provide quantitative parameters and predictions?. *Chaos, Solitons & Fractals*, *135*, Article 109841.

Prather, K. A., Wang, C. C., & Schooley, R. T. (2020). Reducing transmission of SARS-CoV-2. *Science*, *368*(6498), 1422–1424.

Preciado, V. M., Zargham, M., Enyioha, C., Jadbabaie, A., & Pappas, G. (2013). Optimal vaccine allocation to control epidemic outbreaks in arbitrary networks. In *52nd IEEE conference on decision and control* (pp. 7486–7491). IEEE.

Preciado, V. M., Zargham, M., Enyioha, C., Jadbabaie, A., & Pappas, G. J. (2014). Optimal resource allocation for network protection against spreading processes. *IEEE Transactions on Control of Network Systems*, *1*(1), 99–108.

Prem, K., Liu, Y., Russell, T. W., Kucharski, A. J., Eggo, R. M., et al. (2020). The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: a modelling study. *The Lancet Public Health*, *5*, e261–e270.

Priesemann, V., Brinkmann, M. M., Ciesek, S., Cuschieri, S., Czypionka, T., Giordano, G., et al. (2021). Calling for pan-European commitment for rapid and sustained reduction in SARS-CoV-2 infections. *The Lancet*, *397*(10269), 92–93.

Prosper, O., Saucedo, O., Thompson, D., Torres-Garcia, G., Wang, X., et al. (2011). Modeling control strategies for concurrent epidemics of seasonal and pandemic H1N1 influenza. *Mathematical Biosciences and Engineering*, *8*(1), 141–170.

Qian, G., & Mahdi, A. (2020). Sensitivity analysis methods in the biomedical sciences. *Mathematical Biosciences*, *323*, Article 108306.

Qin, L., Sun, Q., Wang, Y., Wu, K.-F., Chen, M., et al. (2020). Prediction of number of cases of 2019 novel coronavirus (COVID-19) using social media search index. *International Journal of Environmental Research and Public Health*, *17*(7), 2365.

Qiu, J., Shen, B., Zhao, M., Wang, Z., Xie, B., & Xu, Y. (2020). A nationwide survey of psychological distress among Chinese people in the COVID-19 epidemic: implications and policy recommendations. *General Psychiatry*, *33(2)*.

Rawlings, J. B., Mayne, D. Q., & Diehl, M. (2017). *Model predictive control: Theory, computation, and design*. WI: Nob Hill Publishing Madison.

Reid, A. H., Taubenberger, J. K., & Fanning, T. G. (2001). The 1918 Spanish influenza: integrating history and biology. *Microbes and Infection*, *3*(1), 81–87.

Reinhart, A., et al. (2018). A review of self-exciting spatio-temporal point processes and their applications. *Statistical Science*, *33*(3), 299–318.

Riad, M. H., Scoglio, C. M., Cohnstaedt, L. W., & McVey, D. S. (2019). Short-term forecast and dual state-parameter estimation for Japanese encephalitis transmission using ensemble Kalman filter. In *Proceedings of the American control conference* (pp. 3444–3449). IEEE.

Riley, S., Fraser, C., Donnelly, C. A., Ghani, A. C., Abu-Raddad, L. J., et al. (2003). Transmission dynamics of the etiological agent of SARS in Hong Kong: Impact of public health interventions. *Science*, *300*(5627), 1961–1966.

Roda, W. C., Varughese, M. B., Han, D., & Li, M. Y. (2020). Why is it difficult to accurately predict the COVID-19 epidemic?. *Infectious Disease Modelling*, *5*, 271–281.

Roomp, K., & Oliver, N. (2020). ACDC-tracing: Towards anonymous citizen-driven contact tracing. arxiv preprint arXiv:2004.07463.

Rothman, K. J., Greenland, S., & Lash, T. L. (2008). *Modern epidemiology*. Lippincott Williams & Wilkins.

Rowthorn, R. E., Laxminarayan, R., & Gilligan, C. A. (2009). Optimal control of epidemics in metapopulations. *Journal of the Royal Society Interface*, *6*(41), 1135–1144.

Safi, M. A., Gumel, A. B., & Elbasha, E. H. (2013). Qualitative analysis of an age-structured SEIR epidemic model with treatment. *Applied Mathematics and Computation*, *219*(22), 10627–10642.

Salje, H., Kiem, C. T., Lefrancq, N., Courtejoie, N., et al. (2020). Estimating the burden of SARS-CoV-2 in France. *Science*, *369*(6500), 208–211.

Saltelli, A. (2002). Sensitivity analysis for importance assessment. *Risk Analysis*, *22*(3), 579–590.

Scharbarg, E., Moog, C. H., Mauduit, N., & Califano, C. (2020). From the hospital scale to nationwide: observability and identification of models for the COVID-19 epidemic waves. *Annual Reviews in Control*, *50*, 409–416.

Scherer, A., & McLean, A. (2002). Mathematical models of vaccination. *British Medical Bulletin*, *62*(1), 187–199.

Schön, T., Wills, A., & Ninness, B. (2011). System identification of nonlinear state-space models. *Automatica*, *47*, 39–49.

Sélley, F., Besenyei, Á., Kiss, I. Z., & Simon, P. L. (2015). Dynamic control of modern, network-based epidemic models. *SIAM Journal on Applied Dynamical Systems*, *14*(1), 168–187.

Shanthakumar, S. G., Anand, S., & Ramesh, A. (2020). Understanding the socio-economic disruption in the United States during COVID-19's early days. arXiv preprint arXiv:2004.05451.

Sharomi, O., & Malik, T. (2017). Optimal control in epidemiology. *Annals of Operations Research*, *251*(1–2), 55–71.

Small, M., & Tse, C. K. (2005). Clustering model for transmission of the SARS virus: Application to epidemic control and risk assessment. *Physica A: Statistical Mechanics and its Applications*, *351*(2–4), 499–511.

Sparks, R. (2013). Challenges in designing a disease surveillance plan: What we have and what we need? *IIE Transactions on Healthcare Systems Engineering*, *3*(3), 181–192.

Suárez, E., Pérez, C. M., Rivera, R., & Martínez, M. N. (2017). *Applications of regression models in epidemiology*. John Wiley & Sons.

Tayarani-N, M.-H. (2020). Applications of artificial intelligence in battling against Covid-19: A literature review. *Chaos, Solitons & Fractals*, Article 110338.

Taylor, L. H., Latham, S. M., & Woolhouse, M. E. (2001). Risk factors for human disease emergence. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *356*(1411), 983–989.

Thompson, W., Comanor, L., & Shay, D. K. (2006). Epidemiology of seasonal influenza: Use of surveillance data and statistical models to estimate the burden of disease. *The Journal of Infectious Diseases*, *194*, 82–91.

Thrusfield, M. (2018). *Veterinary epidemiology*. John Wiley & Sons.

Thurner, S., Klimek, P., & Hanel, R. (2020). A network-based explanation of why most COVID-19 infection curves are linear. *Proceedings of the National Academy of Sciences*, *117*(37), 22684–22689.

Ting, D. S. W., Carin, L., Dzau, V., & Wong, T. Y. (2020). Digital technology and COVID-19. *Nature Medicine*, *26*(4), 459–461.

Tizzoni, M., Bajardi, P., Decuyper, A., King, G. K. K., Schneider, C. M., Blondel, V., et al. (2014). On the use of human mobility proxies for modeling epidemics. *PLOS Computational Biology*, *10*(7), Article e1003716.

Tizzoni, M., Bajardi, P., Poletto, C., Ramasco, J. J., Balcan, D., et al. (2012). Real-time numerical forecast of global epidemic spreading: case study of 2009 a/H1N1pdm. *BMC Medicine*, *10*(165), 1–31.

Torrealba-Rodriguez, O., Conde-Gutiérrez, R., & Hernández-Javier, A. (2020). Modeling and prediction of COVID-19 in Mexico applying mathematical and computational models. *Chaos, Solitons & Fractals*, Article 109946.

Unkel, S., Farrington, C. P., Garthwaite, P. H., Robertson, C., & Andrews, N. (2012). Statistical methods for the prospective detection of infectious disease outbreaks: A review. *Journal of the Royal Statistical Society. Series A: Statistics in Society*, *175*(1), 49–82.

Van den Broeck, W., Gioannini, C., Gonçalves, B., Quaggiotto, M., Colizza, V., et al. (2011). The GLEaMviz computational tool, a publicly available software to explore realistic epidemic spreading scenarios at the global scale. *BMC Infectious Diseases*, *11*(1), 1–14.

Vega, T., Lozano, J. E., Meerhoff, T., Snacken, R., Beauté, J., et al. (2015). Influenza surveillance in europe: comparing intensity levels calculated using the moving epidemic method. *Influenza and Other Respiratory Viruses*, *9*(5), 234–246.

Vega, T., Lozano, J. E., Meerhoff, T., Snacken, R., Mott, J., et al. (2013). Influenza surveillance in Europe: Establishing epidemic thresholds by the moving epidemic method. *Influenza and Other Respiratory Viruses*, *7*(4), 546–558.

Velavan, T. P., & Meyer, C. G. (2020). The COVID-19 epidemic. *Tropical Medicine & International Health*, *25*(3), 278–280.

Vink, M. A., Bootsma, M. C. J., & Wallinga, J. (2014). Serial intervals of respiratory infectious diseases: a systematic review and analysis. *American Journal of Epidemiology*, *180*(9), 865–875.

Vittinghoff, E., Glidden, D. V., Shiboski, S. C., & McCulloch, C. E. (2012). *Regression methods in biostatistics. Linear, logistic, survival, and repeated measures models*. Springer.

Waikhom, P., Jain, R., & Tegar, S. (2016). Sensitivity and stability analysis of a delayed stochastic epidemic model with temperature gradients. *Modeling Earth Systems and Environment*, *2*(49), 1–18.

Waller, L. (2010). Point process models and methods in spatial epidemiology. In A. E. Gelfand, P. Diggle, P. Guttorp, & M. Fuentes (Eds.), *Handbook of spatial statistics* (pp. 403–423). CRC press.

Wallinga, J., & Lipsitch, M. (2007). How generation intervals shape the relationship between growth rates and reproductive numbers. *Proceedings of the Royal Society B: Biological Sciences*, *274*(1609), 599–604.

Walter, S. D., Hildreth, S. W., & Beaty, B. J. (1980). Estimation of infection rates in populations of organisms using pools of variable size. *American Journal of Epidemiology*, *112*(1), 124–128.

Wang, C. J., Ng, C. Y., & Brook, R. H. (2020). Response to COVID-19 in Taiwan: big data analytics, new technology, and proactive testing. *Jama*, *323*(14), 1341–1342.

Whittle, P. (1955). The outcome of a stochastic epidemic–A note on Bailey's paper. *Biometrika*, *42*(1–2), 116–122.

Winter, A. K., & Hegde, S. T. (2020). The important role of serology for COVID-19 control. *The Lancet Infectious Diseases*, *20*(7), 758–759.

World Health Organization (WHO) (2020). *Considerations in the investigation of cases and clusters of COVID-19*: *Technical Report 2nd April*, (pp. 1–4). World Health Organization.

Xiao, Y., & Torok, M. E. (2020). Taking the right measures to control COVID-19. *The Lancet Infectious Diseases*, *20*(5), 523–524.

Xue-Zhi, L., Gupur, G., & Guang-Tian, Z. (2001). Threshold and stability results for an age-structured SEIR epidemic model. *Computers and Mathematics with Applications*, *42*, 883–907.

Yan, A. W., Cao, P., Heffernan, J. M., McVernon, J., Quinn, K. M., et al. (2017). Modelling cross-reactivity and memory in the cellular adaptive immune response to influenza infection in the host. *Journal of Theoretical Biology*, *413*, 34–49.

Yang, Z., Zeng, Z., Wang, K., Wong, S.-S., Liang, W., et al. (2020). Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions. *Journal of Thoracic Disease*, *12*(3), 165–174.

Yilmaz, S., Dudkina, E., Bin, M., Crisostomi, E., Ferraro, P., Murray-Smith, R., et al. (2020). Kemeny-based testing for COVID-19. *PLoS ONE*, *15*(11), Article e0242401.

Yousefpour, A., Jahanshahi, H., & Bekiros, S. (2020). Optimal policies for control of the novel coronavirus (COVID-19). *Chaos, Solitons & Fractals*, *136*, Article 109883.

Youssef, M., & Scoglio, C. (2013). Mitigation of epidemics in contact networks through optimal contact adaptation. *Mathematical Biosciences and Engineering*, *10*(4), 1227–1251.

Zaman, G., Kang, Y. H., & Jung, H. (2008). Stability analysis and optimal vaccination of an SIR epidemic model. *BioSystems*, *93*, 240–249.

Zaric, G. S., & Brandeau, M. L. (2002). Dynamic resource allocation for epidemic control in multiple populations. *IMA Journal of Mathematics Applied in Medicine and Biology*, *19*(4), 235–255.

Zarnitsyna, V. I., Handel, A., McMaster, S. R., Hayward, S. L., Kohlmeier, J. E., & Antia, R. (2016). Mathematical model reveals the role of memory CD8 T cell populations in recall responses to influenza. *Frontiers in Immunology*, *7*, 165.

Zeng, D., Chen, H., Castillo-Chavez, C., Lober, W., & Thurmond, M. (2010). *Infectious disease informatics and biosurveillance*. Springer.

Zhao, S., Lin, Q., Ran, J., Musa, S. S., Yang, G., et al. (2020). Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: A data-driven analysis in the early phase of the outbreak. *International Journal of Infectious Diseases*, *92*, 214–217.

Zheng, N., Du, S., Wang, J., Zhang, H., Cui, W., et al. (2020). Predicting COVID-19 in China using hybrid AI model. *IEEE Transactions on Cybernetics*, *50*(7), 2891–2904.

Zhou, C., Su, F., Pei, T., Zhang, A., Du, Y., et al. (2020). COVID-19: Challenges to GIS with big data. *Geography and Sustainability*, *1*(1), 77–87.

Zhou, F., Yu, T., Du, R., Fan, G., Liu, Y., Liu, Z., et al. (2020). Clinical course and risk factors for mortality of adult patients with COVID-19 in Wuhan, China: a retrospective cohort study. *The Lancet*, *395*(10229), 1054–1062.

Zino, L., Rizzo, A., & Porfiri, M. (2020). Analysis and control of epidemics in temporal networks with self-excitement and behavioral changes. *European Journal of Control*, *54*, 1–11.