Contents lists available at ScienceDirect

Physics Reports

Statistical physics of vaccination

Zhen Wang [a,b,*], Chris T. Bauch [c,**], Samit Bhattacharyya [d], Alberto d'Onofrio [e],
Piero Manfredi [f], Matjaž Perc [g,h], Nicola Perra [i], Marcel Salathé [j,k], Dawei Zhao [l]

[a] Qingdao University, Qingdao, Shandong, 266071, China

[b] Interdisciplinary Graduate School of Engineering Sciences, Kyushu University, Fukuoka, 816-8580, Japan

[c] Department of Applied Mathematics, University of Waterloo, Waterloo, ON N2L 3G1, Canada

[d] Department of Mathematics, School of Natural Sciences, Shiv Nadar University, India

[e] International Prevention Research Institute, 95 Cours Lafayette, 69006 Lyon, France

[f] Department of Economics and Management, University of Pisa, Italy

[g] Faculty of Natural Sciences and Mathematics, University of Maribor, Koroška cesta 160, SI-2000 Maribor, Slovenia

[h] Center for Applied Mathematics and Theoretical Physics, University of Maribor, Krekova 2, SI-2000 Maribor, Slovenia

[i] Centre for Business Network Analysis, Greenwich University, Park Raw, SE10 9LS, London, United Kingdom

[j] School of Life Sciences, EPFL, 1015, Lausanne, Switzerland

[k] School of Computer and Communication Sciences, EPFL, 1015, Lausanne, Switzerland

[l] Shandong Provincial Key Laboratory of Computer Networks, Shandong Computer Science Center (National Supercomputer Center in Jinan), Jinan 250014, China

ARTICLE INFO

ABSTRACT

Historically, infectious diseases caused considerable damage to human societies, and they continue to do so today. To help reduce their impact, mathematical models of disease transmission have been studied to help understand disease dynamics and inform prevention strategies. Vaccination – one of the most important preventive measures of modern times – is of great interest both theoretically and empirically. And in contrast to traditional approaches, recent research increasingly explores the pivotal implications of individual behavior and heterogeneous contact patterns in populations. Our report reviews the developmental arc of theoretical epidemiology with emphasis on vaccination, as it led from classical models assuming homogeneously mixing (mean-field) populations and ignoring human behavior, to recent models that account for behavioral feedback and/or population spatial/social structure. Many of the methods used originated in statistical physics, such as lattice and network models, and their associated analytical frameworks. Similarly, the feedback loop between vaccinating behavior and disease propagation forms a coupled nonlinear system with analogs in physics. We also review the new paradigm of digital epidemiology, wherein sources of digital data such as online social media are mined for high-resolution information on epidemiologically relevant individual behavior. Armed with the tools and concepts of statistical physics, and further assisted by new sources of digital data, models that capture nonlinear interactions between behavior and disease dynamics offer a novel way of modeling real-world phenomena, and can help improve health outcomes. We conclude the review by discussing open problems in the field and promising directions for future research.

* Corresponding author at: Qingdao University, Qingdao, Shandong, 266071, China.
** Corresponding author at: University of Waterloo, Waterloo, ON N2L 3G1, Canada.
E-mail addresses: zhenwang0@gmail.com (Z. Wang), cbauch@uwaterloo.ca (C.T. Bauch).

Contents

1. Introduction

1.1. Background

Infectious diseases have proven to be remarkably resilient foes of humanity. Despite major achievements in other areas, such as discovering gravitational waves and landing humans on the Moon, we still have not eradicated most vaccine-preventable infectious diseases [1]. Many of these diseases – measles, pertussis, influenza, and many others – have been burdening us for centuries, while new infectious diseases – such as Ebola, Zika virus, and SARS – continue to emerge from animal populations and make the jump to human populations, or spreading to new human populations due to climate change or other anthropogenic disturbances [2].

Sometimes, successful infection control is a matter of ensuring access to preventive technologies such as vaccines [3]. Barriers to access in lower-income countries remains an important reason why infectious disease burden remains high in many poor populations [3]. However, there are reasons to be cautiously optimistic that barriers to access will continue to fall, enabling more people to be vaccinated. For instance, measles was completely eliminated from South America in the 1990s [4], prompting many health researchers and policy experts to conceive of global measles eradication as a potentially realistic goal for the first time. Mathematical and statistical models, often based on methods and concepts borrowed from physics, have played a significant role in infection control by helping determine how to apply control measures in order to quickly and most effectively contain epidemics and helping predict and understand disease outbreak patterns, among their many other uses [5].

However, as our technologies advance and access to vaccines begin to become less important as a barrier, it is becoming clear that our own behavior may prevent local elimination or global eradication in many cases [6]. For instance, unfounded vaccine scares such as the ongoing measles–mumps–rubella vaccine autism scare which started in 1997, or the oral polio vaccine scare in northern Nigeria in 2003–2004 which arguably delayed global polio eradication by a decade, can enable a resurgence of disease in a population [7]. When the infection spreads to other populations in turn, local elimination or even global eradication can be threatened. Similarly, traditional infection control methods such as isolation, quarantine, and contact tracing require adequate levels of support from the public. When this is lacking, sometimes due to mismanagement of the situation by authorities or inadequate response of the international community, then infection control can fail, as was observed in the early stages of the 2014 West African Ebola outbreaks and in many other situations [8].

1.2. Motivation and objective

In summary, recent decades have seen the continual emergence of new infectious diseases together with growing examples where the technologies exist to eliminate long-established infections, but human behavior has prevented elimination. Moreover, more data on human behavior are available than ever before, at very high levels of spatial and temporal resolution through sources such as internet search engine data, Bluetooth technologies, and social online media.

These trends may partially explain the surging interest in incorporating human behavior directly into mathematical models of infectious disease transmission, in order to understand the two-way interplay between the dynamics of human behavior and the dynamics of infectious diseases. By accounting for human behavior, mathematical models may provide more insight, and better predictions, than models that neglect the enormous impact that behavior can have on infectious disease dynamics. For brevity, we will refer to such models as 'coupled behavior–disease models' through the remainder of this report, although models that only explore one direction of the coupling, from human dynamics to disease dynamics, or from disease dynamics to human dynamics, are also relevant to our report.

Much of this work relies upon methods and concepts from physics, either directly or indirectly. This is neither new nor surprising. For instance, exponential decay can describe radioactive decay of particles, or exponential growth can just as well describe the early growth of an invasive species. Hence, physics in general, and statistics physics in particular, is highly relevant to the development of disease transmission models that take human behavior into account.

The two objectives of our report are (1) to give a state-of-the art overview of the literature on mathematical models of the impact of human behavior on disease dynamics (and *vice versa*), with special emphasis on the connections to concepts and methods from physics; and (2) to provide an outlook on the future of the field, including open questions and promising directions for future research. We attempted to include as much of the relevant literature as possible, but we were also cognizant of the need to limit the list of references to a realistic number. Therefore we selected references that best captured the full breadth of research being undertaken in the field.

1.3. Outline of the report

In Section 2 of the report we provide a brief summary of basic ideas in epidemiology that are necessary to understand and develop mathematical models of infectious disease transmission. Classic mean-field models of infectious disease transmission are then described in Section 3. Subsequently we describe models of infectious disease transmission on networks and other spatially structured populations, including both simulation-based approaches (Section 4) and analytical approximations (Section 5) to network dynamics. Sections 3–5 are concerned with models of infectious disease transmission that do not explicitly include human behavior, since understanding such models is a necessary precursor to understanding models that explicitly do include human behavior. Section 6 provides a rationale for studying models of the interactions between human behavior and disease dynamics (some of which has already been described in the preceding subsections), while Section 7 introduces some basic concepts and methods in the modeling of human behavior. Section 8 provides an overview of mean-field models of coupled behavior–disease dynamics, while Section 9 provides an overview of such models in the context of networks and spatially-structured populations. The next two sections are concerned with 'traditional' sources of data on human behavior (Section 10) and digital sources of data on human behavior (Section 11) from which we can test and validate models of coupled behavior–disease dynamics. The report concludes with a discussion in Section 12.

2. Basic concepts in infectious disease epidemiology

Infectious disease epidemiology is the study of contagious parasites of humans. Parasitization is one of the oldest ecological relationships, and has been hypothesized to be a major driving factor in the evolution of a very broad range of species (the 'Red Queen hypothesis') [9]. Indeed, infectious diseases are likely to have been with humans for our entire existence. The Bible and ancient Greek historical writing contain apparent references to epidemics and case isolation [10,11]. This section describes the basic biology and ecology of infectious diseases, and how such diseases are controlled through interventions such as vaccines.

Parasites are traditionally divided into two main groups: microparasites and macroparasites. The distinction between them is their size and parasitic life cycle [12]. Macroparasites are larger forms of pathogens and include ticks, mites, helminthes and flukes [13]. Macroparasites grow in one host but reproduce by infective stages outside of the host organism [14]. They can be either external parasites (ectoparasitic) or internal parasites (endoparasitic) [15]. Comparatively, microparasites are small, generally single-cell organisms, and are either viruses, bacteria or protozoans [16,17]. They generally exhibit short generation times and a high rate of direct reproduction within the host organism [13]. For simplifying assumptions in modeling, we focus on the body of literature pertaining to microparasites, where extensive long-term data and a solid understanding of disease transmission dynamics have contributed to the development of well-parameterized models [18].

Infectious diseases (both micro- and macroparasitic) can be transmitted to susceptible hosts in a variety of ways. Routes of transmission are classified as direct and indirect transmission [19]. Direct transmission occurs when there is physical or close contact between an infected individual and a susceptible individual that permits transmission via droplets or aerosols [20].

Although some exceptions exist, the majority of microparasitic diseases are directly transmitted and exhibit short survival periods outside of the host organism [18]. Conversely, indirectly transmitted parasites are passed through vectors or are transmitted vehicularly via the environment [20]. Most macroparasitic diseases are indirectly transmitted, spending part of their life cycle outside of their host organism [18]. Indirect transmission occurs via another organism through a vector or via an intermediate host [21]. A vector is an organism that does not directly cause disease, but that transmits infection of pathogens from one host to another [22]. Vectors can be mechanical or biological. A mechanical vector transmits an infectious pathogen without it being altered while on the vector [23]. In contrast, biological vectors harbor pathogens within their bodies and deliver pathogens to new hosts in an active manger, usually a bite [24]. Biological vectors are usually continuously infected with the disease and in some cases are a required part of the organisms life cycle [23]. Increases in populations and urbanization, habitat loss, and international travel and trade all contribute to the spread of vectors and indirect transmission of pathogens [25]. Consequently, the mechanisms of transmission and transmissibility are key factors of pathogens that influence disease spread and what interventions, if any, are effective [26]. Timing of the onset of infectiousness relative to the onset of symptoms and duration of infectious periods are additional important factors in transmission dynamics [26].

Integral to understanding infectious disease dynamics is distinguishing three time periods involved in these processes: the latent period, defined as the time from infection to the time when a host is able to transmit the pathogen to another host [22]; the incubation period, defined as the time from infection to the onset of disease [27]; and the infectious period, defined as the period from the end of the pre-infectious period until the time when a host is no longer able to transmit the infection to other hosts [28]. The distinction between the incubation period and the latent period of a disease is integral when considering the effect of control methods based on symptomatic surveillance in a population. If the latent period is shorter than the incubation period (i.e., infectiousness proceeds with the onset of symptoms), then interventions targeted toward symptomatic individuals are generally ineffective [29]. Diseases whereby the onset of symptoms coincides with, or proceeds with infectiousness are more effective at being controlled using symptomatic surveillance methods [21]. Consequently, an important metric in the controllability of handling a disease outbreak requires the consideration of the proportion of transmission that takes place before the onset of symptoms occurs [29]. Both incubation period and latent period are generally modeled following an exponential, log-normal, Weibull, or gamma distribution [21].

Three important characteristics of pathogens that affect transmission of infectious diseases are infectivity, virulence and pathogenicity. In epidemiology, infectivity is the ability of a pathogen to cause infection in a susceptible host [21]. Infectivity can also vary between different strains of the same pathogen. The route of the infection, and the age and innate resistance of the host are key factors which contribute to the infectivity of a disease [30]. The gauge of infectivity is measured by the secondary attack rate (the proportion of exposed hosts who develop the disease [31]). Infectivity has been shown to positively correlate with virulence [32,33]. Virulence, although often used interchangeably with pathogenicity, refers to the severity of disease after infection occurs [21]. The capacity and vigor of the disease to cause severe and fatal cases of illness is virulence [34]. Pathogenicity refers to the ability of a pathogen to induce disease [21]. Pathogenicity and virulence are determined by the interaction between host and pathogen [20]. Characteristics of infectious pathogens that affect virulence include their ability to replicate, invade organisms and damage the host [20]. Host characteristics also affect the ability of a pathogen to cause disease, in that pathogenicity and virulence depend on the resistance and immunity system of the host, in addition to genetic factors, age, gender, and physiological condition [20].

Herd immunity is a form of indirect protection from infectious disease that is brought about by the presence of immune individuals in a population [35]. With certain diseases that spread from host to host, the level of population immunity is critical in determining the potential of an epidemic from occurring, and therefore the risk of infection for a susceptible host in a population [21]. In that transmission occurs between an infected and susceptible host, if the number of immune individuals is large, then the likelihood that a susceptible individual being in contact with an infected individual is reduced. Immunity can be conferred through natural immunity, previous exposure to the disease, or through vaccination [36]. The level of immunity required to attain herd immunity depends on the characteristics of the infectious disease. Generally a disease with high infectivity requires a higher proportion of immune individuals to achieve immunity than diseases with lower infectivity rates [36]. Additional considerations are: seasonality, to what effect it has on population mixing and the viability of the pathogen; the number of susceptible individuals, their age and geographic distribution; and social habits of the population with respect to mixing and type of contacts made [37]. Mass vaccination to induce herd immunity is common practice and has proven successful in preventing the spread of many infectious diseases globally [38]. Opposition to vaccination has posed a challenge to herd immunity, allowing preventable diseases to persist in or return to communities that have inadequate vaccination rates [39].

An integral application in infectious disease modeling is the need for evaluating intervention strategies and their effects on new and re-emerging pathogens. The development and widespread use of antibiotics was an important advance in the treatment of infectious diseases [26]. Antibiotics are a type of antimicrobial used in the treatment and prevention of bacterial infection, by killing or inhibiting the growth of bacteria [40]. Several antibiotics are effective against fungi and protozoans [41], but are not effective against viruses [42]. In many cases, their effectiveness and accessibility has led to their overuse, prompting bacteria to develop resistance [43–45]. Increasing travel and migration have also contributed to the growth of antibiotic resistance problems [46]. The consequences of antibiotic-resistant infections may result in more prolonged illnesses, longer hospitalization, increased risk of death and higher health care costs than do infections with antibiotic-sensitive strains of the same species [20].

Intervention	Examples	How They Work	When Used
Vaccines	Measles-mumps-rubella (MMR) vaccine, chickenpox vaccine, cervical cancer vaccine	Creates an immune memory of the pathogen so immune system can respond more vigorously when challenged with the pathogen; creates herd immunity at the population level whereby unvaccinated individuals derive indirect protection from vaccinated individuals	Used widely for prevention (prophylaxis), and sometimes in response to outbreaks
Antibiotic drugs, antiviral drugs	Penicillin (bacteria), M2 inhibitors and neuraminidase inhibitors (influenza)	Interferes with life cycle of virus or bacterium within the host	Used widely to treat individual infected patients, but may also be used for prophylaxis
Social distancing	Authorities close schools, individuals spontaneously avoid crowded places	Prevents transmission by limiting contact between susceptible and infectious individuals in the general population	Used when other measures are not available and/or in a severe epidemic (due to its disruptive nature)
Case isolation	Negative pressure rooms in hospitals, infected individuals staying home	Isolates infectious individuals so they do not transmit to others	Used widely, especially for severe infections
Quarantine	Asking housemates of an infected individual to remain in their house for a duration of time	Isolates individuals who may have been exposed to infection, and therefore potentially infectious themselves	Used for severe infections and/or when less disruptive interventions (e.g. drugs, vaccines) are not available
Barrier precautions	N95 respirator masks, surgical masks, gloves, condoms	Physically blocks infection sources, e.g. infected droplets	Used widely to greater or lesser extent
Contact tracing	Infectious individuals are queried regarding their contacts over the past few days	Contacts can be found and treated, quarantined, or isolated as needed	Used widely during outbreaks for infections spread through identifiable contacts (e.g. housemates, close contact infections)
Hygiene	Hand washing, boiling water, washing food	Physically removes pathogens	Used widely for a variety of infections

Fig. 1. Summary of commonly used interventions against infectious diseases. We note that this list is not exhaustive, and interventions are often used in combination, such as contract tracing and isolation.

A primary way to boost immunity in a population is through vaccination [47]. Vaccines are any biologically derived substance that produces a protective immune response when administered to a susceptible host [21]. Vaccine effectiveness is the result of the vaccine's direct effect, which refers to its ability to protect individuals from infection, and its indirect effect, which is its ability to reduce the spread of infection in a population [48,49]. Notably, no current vaccine is entirely effective. It is estimated that non-preventable primary vaccine failure rates range from 2 to 50 percent for licensed vaccines under ideal circumstances in clinical trials [50]. Vaccine effectiveness can be determined by comparing risk of the disease among vaccinated and nonvaccinated groups. Important variables to consider are that groups are comparable with respect to exposure, risk of infection, accessibility of the vaccine, and opportunity for diagnosis [21]. Vaccines, antibiotics, and other interventions are summarized in Fig. 1.

Social distancing is another strategy used to reduce the transmission of infectious diseases. Governments and public health agencies may promote targeted social distancing as a strategy, in the form of school or workplace closures. Social distancing also arises as individuals respond to news about disease spread, obtained from media reports, public announcements, rumors and individual experience [51]. Recognizing the importance of behavioral responses in populations has altered the way epidemiological modelers consider transmission [52–58]. Such models have shown that social distancing can be effective at reducing the attack rate of an epidemic [59,60] and that social distancing behavior is a plausible explanation for certain phenomena arising in real epidemics, such as multiple outbreaks or waves of infection [60,61].

Hygiene refers to environmental conditions and practices that contribute to maintaining health and prevent the spread of diseases [62]. Aiello and Larson [63] examined the epidemiological evidence for causal relation between hygiene and infections and found that personal and environmental hygiene significantly reduced the spread of infections. In particular, the importance of hand hygiene in preventing infections has been widely recognized. Hands are a key step in the transmission of many infectious agents, and the importance of hand hygiene and interrupting hand contact with body fluids

in preventing spread of infections is well documented [64–66]. Diseases related to hygiene in general disproportionately affect poorer members of society due primarily to accessibility of services and/or less polluted environments [67].

Despite the complexities of the biology of infectious diseases on the cellular and organismal levels, it is often possible to mathematically model the transmission of infectious diseases at the population level. In fact, some of the earliest research in mathematical biology concerned the analysis of population-level transmission of infectious diseases using differential equations [68]. The mathematical modeling of infectious disease transmission is the subject of the next section.

3. Basic concepts and methods in epidemiological modeling and vaccination: compartmental (mean-field) models

Epidemiological modeling has nowadays become a global discipline, playing a key role in policy decisions about prevention and control of communicable diseases. With hindsight, much of this success is due to a few, extraordinarily fecund, deterministic models, first of all the two classical susceptible–infective–removed (SIR, see Table A.1 in Appendix for this and other abbreviations used in our review) models for outbreaks and for endemic infections (and related extensions). These models represent the direct legacy of the intuitions of a few innovators who, at the beginning of the 1900*s*, imported the mass-action law of statistical physics to describe infection transmission. This intuition, which represents the building block of the transmission process of a communicable disease, namely the social mixing between the individuals of a community, as collisions between particles in a certain medium, has allowed mathematical epidemiology to take-off as a scientific discipline. This section reviews the key concepts and results of basic deterministic mean-field models of theoretical epidemiology, departing from an historical overview of the founding role of the mass-action principle (Section 3.1). Notably, the current *behavioral epidemiology* revolution stems, one hundred years after the birth of mathematical epidemiology, from a twofold awareness, first that humans, unlike particles, are thinking agents that might depart from the mass-action rule especially when the threats from infections are serious, and second that we now have easier access to the tools that are necessary to finally model such departures. The rest of the section is organized as follows: Section 3.2 introduces a basic compartmental modeling framework for infection transmission and the underlying key concepts; Section 3.3 reports a non-technical presentation of reproduction numbers, which are key summary parameters of transmission; the main results about models for infection with permanent immunity, namely the SIR and SEIR models for outbreaks and for endemic infections, follow in Sections 3.4 and 3.5 respectively; the effects of mass vaccination for SIR infections are reported in Section 3.6, while the consequences of immunity losses follow in Section 3.7; the section ends with a discussion of a number of extensions of basic models (Section 3.9), and of the epidemiological meanings and implications of the use of mean-field approximations (Section 3.8).

3.1. Mathematical epidemiology: a story shaped by statistical physics

The first application of mathematics to epidemiology dates back to Daniel Bernoulli [69] investigation of the gain in life expectancy that would be allowed by elimination of smallpox, then a major cause of humans' death. Though historically important and far-seeing – smallpox is so far the only infection ever eradicated in mankind history – Bernoulli's study cannot however be taken as the birthday of modern mathematical epidemiology, since it neglected the building block of the diffusion of a communicable disease, namely, the transmission process of an infective agent and its social spread in a population. The consideration of the transmission process took 150 further years i.e., following the nice historical reconstruction by Heesterbeek [70], the lapse of time necessary to build a solid scientific background to the idea of contagion. It is indeed only in the 1870s that the conclusive confirmation by Koch and Pasteur that infectious diseases are caused by living organisms, definitely paved the way to the explanation of observed *epidemic curves* (which today we call *incidence curves*), giving credit to a number of central earlier intuitions. The main one was the John Snow (Fig. 2) hypothesis that epidemics come to an end when the availability of susceptible individuals – the "epidemic fuel"– goes down. Some decades later, while analyzing UK measles incidence data, William Hamer [71] identified the biennial period of outbreaks in towns, and in particular explained the recurrence of measles epidemics by Snow's hypothesis, i.e. a new epidemics will start "when there is a sufficient accumulation of susceptible persons". In addition he computed from data a relationship of the form "New cases at time $t + 1$ = (New cases at time t) × (susceptibles at time t)". However, he failed to recognize this as a general mechanistic rule, nor did he draw any implications from it [70]. The key step finally turning mathematical epidemiology into a scientific discipline occurs when two fine innovators working independently but in strict touch, namely the 1908 Nobel Prize for medicine Sir Ronald Ross [72,73] and Anderson McKendrick [74] (Fig. 2), formulated and analyzed the first modern epidemic models [70]. It was in particular McKendrick who, by importing from a stronger domain of knowledge, namely statistical physics, the *mass-action* law of chemical reaction kinetics to describe infection transmission as the outcome of the social contact process between infective and susceptible individuals, sets-up the building block of modern epidemiology, thereby generating a true *Gestalt* shift in scientific knowledge. McKendrick's ground-breaking intuition is summarized in his metaphor of "collisions", that he applies to an animal infection: "The rate at which this epidemic will spread depends obviously on the number of infected animals, and also on the number of animals which remain to be infected, in other words the occurrence of a new infection depends on a collision between an infected and an uninfected animal." [74].

This observation makes McKendrick's seminal study a true – in contemporary language – transdisciplinary paper in statistical physics. He makes the counter-factual hypothesis that the population behaves as a collection of "particles" of

Fig. 2. Forerunners of mathematical epidemiology: John Snow (left), Sir Ronald Ross (center), Anderson McKendrick (right).

different types randomly moving in the environment, where each encounter ("collision") between a particle of *susceptible*-type and one of *infective*-type has a uniform probability that the S-type particle switches into I-type. Translating this into mathematical language for an infection without immunity in a stationary population of size N, letting Y and X respectively denote the numbers of infective and susceptible individuals ($X + Y = N$), and k' a proportionality constant, McKendrick derived the following ordinary differential equation (ODE) for the rate of epidemic spread

$$Y' = k'Y(N - Y), \tag{1}$$

as the first true model of modern epidemiology. Passing to the epidemiological fractions $S = X/N, I = Y/N, S + I = 1$ and defining $\hat{k} = k'N$ one gets

$$I' = \hat{k}I(1 - I). \tag{2}$$

This simple logistic equation, having $I = 1$ as a globally asymptotically stable (GAS) state reached by a sigmoid temporal pattern, implies that an infection without immunity spreading according to the mass action principle will eventually affect the entire population. McKendrick's *full contagion principle* represents the first principle of theoretical epidemiology.

By pursuing further his mass-action formulation, McKendrick definitely launched mathematical epidemiology in a famous series of papers with Kermack [75], where they introduced the celebrated *threshold theorem* for infections imparting permanent immunity, studied conditions for endemicity, and introduced and analyzed the first age-structure epidemiological models. The momentum impressed by Kermack and McKendrick work yielded, with a certain delay, an explosion of bio-mathematical studies, mostly devoted to the investigation of a number of variants of the key models introduced by the forerunners. For further historical details see [76].

There is a second major shift however, dating 1975–1980, when another group of innovators [77–79] started and systematically using mathematical models to interpret epidemiological data and trends aiming to deep the perspectives for infection control. In this phase, dramatically accelerated first by the HIV/AIDS epidemics in the 1980s, and in the 2000s by the emergence of influenza pandemic threats, mathematical models of infectious diseases definitely left the boundary of bio-mathematics and became a central supporting tools for public health decisions, such as determining the duration of travel restrictions or of school closure during a pandemic event, or the fraction of newborn to be immunized for a vaccine preventable infection, as measles. It is this fine tuning of epidemiological models that eventually raised the key question of behavioral epidemiology: given that humans are not "particles", which is then the role they play for infection transmission and prevention? [80].

3.2. Basic concepts: homogenous social mixing, infection incidence, force of infection

Regardless of their nature (e.g., directly vs sexually transmitted), for most human infectious diseases the individual life-course of infection is conveniently represented by a sequence of transitions between a few epidemiological classes, or *compartments*, triggered by specific processes and parameters [79,81]. For new pathogens, such as a pandemic flu virus, all individuals are initially *susceptible* to the infection (class S), which they acquire at a rate λ entering the *exposed* class (E), a latency or incubation phase where individuals are infected, i.e. the virus is replicating in their body, but not yet *infective*, which they become at some rate α. Individuals belonging to the infective class (I) will usually recover, at a rate γ, acquiring immunity and entering the *removed* class (R). This Susceptible–Exposed–Infective–Removed course, *SEIR* for brevity (Fig. 3), is shared for example by a number of vaccine-preventable infections of childhood yielding lifelong immunity, such as measles, mumps and rubella. For other, as pertussis and varicella, clinical reinfection through return to susceptibility (at some rate δ) is possible, so that their actual course would be *SEIRS*.
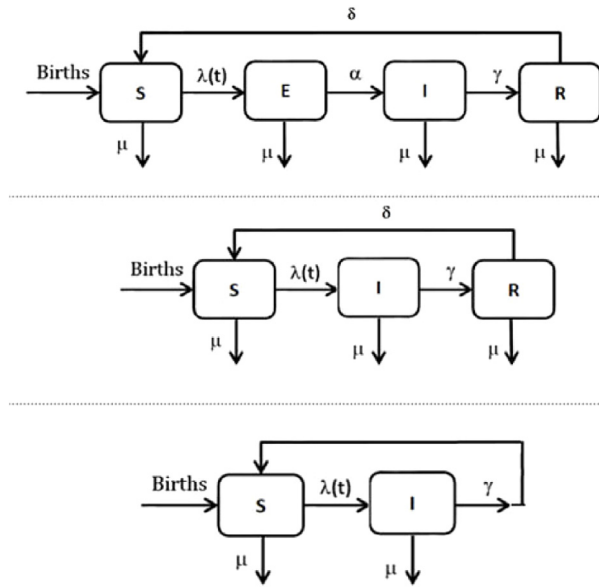
Fig. 3. Compartmental representation of the individual course for some important classes of infections: SEIRS model (top); SIRS model (middle); SIS model (bottom). Boxes identify epidemiological classes and arrows identify the processes governing transitions between classes, and in–out flows due to demographics.

Based on this representation of the individual course, epidemiological models seek to describe the dynamics of infections at the population level, by associating to each compartment a dynamic variable denoting the number (or the proportion) of individuals in that class at each instant of time. Denoting by N the total population size, we let X ($S = X/N$), H ($E = H/N$), Y ($I = Y/N$), Z ($R = Z/N$) to denote the numbers (proportions) of individuals who respectively occupy the various classes at time t, with $N = X + H + Y + Z$, and $S + E + I + R = 1$. The key process, namely infection spread, focuses on the number of transitions from compartment S to compartment E per unit of time, which is the absolute incidence (U) of new infections per time-units. In simple models the mass action formulation of incidence occurs by the concept of *homogeneous* (or random) *mixing* [79,81]. This refers to an infection spreading by *direct person-to-person contacts* in a closed homogenous social medium, the population, whose agents have an approximately identical social activity, summarized by the average number of contacts C per unit time, the same probability $I = Y/N$ that a given contact occurs with an infective individual, and the same constant probability β_I of infection transmission per single *adequate* social contact (i.e., per single social link in networks jargon) between a susceptible and an infective individual. Therefore, the expected instantaneous per-capita rate λ at which susceptible individuals acquire infection, called the *force of infection* (FOI), can be written as: $\lambda = \beta_I C I$ or, setting $\beta = \beta_I C$, as: $\lambda = \beta I$. The corresponding absolute incidence U is then given by

$$U = \lambda X = \beta (Y/N) X, \tag{3}$$

which is the *standard mass action* formulation of the incidence and the FOI [81,82]. It depends on the complex parameter β, representing the total *transmission rate* per unit of time, which embeds distinct sub-processes, namely social behavior (mean number and type of contacts per unit of time) as well as the biological process of contagion. Note that, assuming the frequencies S and I as the "molar concentrations" of the "particles" susceptible subjects and infectious, and the total population size N as the "Avogadro's number" of the model in study, then the change per time unit of such "molar concentrations" is equal to (U/N), which in turn is equal to βSI, i.e. the kinetic mass action law.

Also common in biomathematics is the *bilinear incidence* $U = \widetilde{\beta} XY$ [81,83], which, however, is less supported empirically [79], and moreover postulates that C is linearly increasing in population size which seems to be reasonable only in a few special cases, and also in the quite artificial case where the numerical value of $\widetilde{\beta}$ scales as $1/N$. This means that one can write $U = \widetilde{\beta} XY = \beta (Y/N) X$. In other words, in such cases the models, although formally built on the *bilinear incidence*, in reality are ruled by the *standard mass action* law for the contagion mechanism.

To complete the model formulation we make a few further hypotheses. First, we assume that all other epidemiological processes in Fig. 3, namely, transfers $E \rightarrow I, I \rightarrow R, R \rightarrow S$, occur at constant rates, meaning that the duration of sojourn in the involved states is exponentially distributed. This means for example that $\gamma = 1/D$, i.e. that the rate of recovery is the reciprocal of the average duration D of the infective phase. Second, we assume that the population is large, so that the law of large numbers rules out large chance fluctuations, and closed to migration, i.e. it evolves over time due the natural balance between exits due to mortality (from all compartments), at a per-capita rate $m(N)$, and entries due to new births, occurring at a per-capita rate $b(N)$ into the susceptible class only. Last, we rule out complications such as infection-related mortality, or vertical transmission from mother to foeta during pregnancy. We can then translate the SEIRS diagram in Fig. 3 into the

following *mean-field* system of ODEs

$$X' = b(N)N + \delta Z - m(N)X - \beta X \frac{Y}{N}, \tag{4}$$

$$H' = \beta X \frac{Y}{N} - \alpha H - m(N)H, \tag{5}$$

$$Y' = \alpha H - \gamma Y - m(N)Y, \tag{6}$$

$$Z' = \gamma Y - m(N)Z - \delta Z, \tag{7}$$

where the population size obeys

$$N' = (b(N) - m(N))N. \tag{8}$$

Given that the population evolution is known, one equation can be discarded, since e.g., $Z = N(t) - X - H - Y$. Assuming that the population N has a unique and globally stable equilibrium N_e such that $b(N_e) = m(N_e) = \mu$, and that the population is at its steady state N_e, yields

$$S' = \mu(1 - S) + \delta R - \beta SI, \tag{9}$$

$$E' = \beta SI - (\alpha + \mu)E, \tag{10}$$

$$I' = \alpha E - (\gamma + \mu)I, \tag{11}$$

complemented by $R = 1 - S - E - I$. Model of Eqs. (9)–(11) is the so-called deterministic SEIRS model [81].

By suitably restricting some parameters the SEIRS model can be specialized into various "classical" models appeared in the biomathematical literature. For example, by assuming a negligibly short duration of the exposed phase (i.e. $\alpha \gg 1$) and applying a quasi-steady state approximation

$$E \approx \frac{\beta}{\alpha + \mu} SI \tag{12}$$

yields

$$S' = \mu(1 - S) + \delta R - \beta SI, \tag{13}$$

$$I' \approx \frac{\alpha}{\alpha + \mu} \beta SI - (\gamma + \mu)I \approx \beta SI - (\gamma + \mu)I, \tag{14}$$

i.e. the SIRS model (see Fig. 3), to be complemented by the algebraic equation $R = 1 - S - I$. Further assuming that the immune phase is permanent (i.e. setting $\delta = 0$) yields the celebrated SIR model with vital dynamics

$$S' = \mu(1 - S) - \beta SI, \tag{15}$$

$$I' = \beta SI - (\gamma + \mu)I. \tag{16}$$

On the contrary, the assumption of absence of an immune phase (i.e. $\delta \gg 1$) brings to the SIS model (Fig. 3)

$$S' = \mu(1 - S) - \beta SI + \gamma I, \tag{17}$$

$$I' = \beta SI - (\gamma + \mu)I. \tag{18}$$

Finally, by considering epidemic models for which the epidemiological time scales are fast compared to demographic ones, i.e. by setting in the above model $\mu \approx 0$, one obtains models for outbreaks, which include the celebrated SIR epidemic model by Kermack and McKendrick [75].

3.3. Reproduction numbers of epidemic spread

What happens when a new infectious agent, such as a pandemic virus, enters (e.g. via an infected traveler) a population of susceptible individuals? Will the virus start spreading and causing an *outbreak*? And if it does, how fast and how many people will be eventually hit? Though detailed answers to such questions depend on the characteristics of the infection considered, a great deal of intuition can be obtained by a few summary parameters. The most well-known is the *basic reproduction number* (BRN) of infection, a dimensionless quantity universally denoted by $\mathcal{R}_0$, and intuitively defined as the expected number of secondary infection cases caused by a single *typical* infective case during his entire period of infectivity in a wholly susceptible population [79,82,84,85]. Under homogeneous social mixing $\mathcal{R}_0$ can be defined as

$$\mathcal{R}_0 = \beta_C CD = \beta D, \tag{19}$$

where D is the expected duration of the infective phase. In a large homogeneous population living a world without uncertainty, $\mathcal{R}_0$ represents, by definition, a threshold parameter stating whether the epidemics takes off, thereby invading the population ($\mathcal{R}_0 > 1$) or dies out (otherwise). In case of invasion $\mathcal{R}_0$ also represents the exponential growth factor

of infective cases counted per generations of infection. Let us consider for example an infection imparting immunity, as influenza, and take $\mathcal{R}_0 = 2$, a figure close to the estimates for 1918 Spanish pandemic flu [86]. The initial case (the "traveler", representing the *zero generation*) would cause $\mathcal{R}_0 = 2$ cases in the first generation, which would in turn cause $\mathcal{R}_0 = 2$ cases each one in the second generation, and so on, according to the geometric progression 2^n after n generations. This exponential growth according to

$$U_n = \mathcal{R}_0^n \tag{20}$$

holds for the initial *invasion* phase, that is as far as the depletion of susceptible subjects is negligible. The corresponding real time growth of infective cases will be $U(t) \approx e^{rt}$. Assuming a stationary distance between any two consecutive generations of infection, what today we call the infection *generation time* [87], and that this equals D, the duration of the infective phase, leads to the fundamental relation between generational growth and real time growth: $\mathcal{R}_0 = e^{rD}$, which enables to estimate $\mathcal{R}_0$ based on the observed rate of real time growth of cases and the infection generation time. Once epidemic spread has depleted the susceptible fraction S to levels well below 100%, the definition of $\mathcal{R}_0$ does not apply anymore, and the outbreak evolution is described by the quantity (still assuming homogenous mixing)

$$\mathcal{R}_E(t) = \mathcal{R}_0 S(t) \tag{21}$$

known as the *effective reproduction number* (ERN), often denoted by $\mathcal{R}_E$. The ERN captures the fact that contacts with immune people do not contribute to the further growth of the epidemics. Therefore, still under the Spanish flu assumption $\mathcal{R}_0 = 2$, when the epidemics has immunized 25% of the initial population, cases will temporary grow at the reduced generation speed $\mathcal{R}_E = 1.5$, meaning that on average each infective case will now generate only 1.5 secondary cases. With further progressing of the epidemics, $\mathcal{R}_E$ further declines, and when the immune proportion reaches 50%, implying $\mathcal{R}_E = 1$, the incidence curve reaches its maximum, after which the epidemics enters its declining phase, and eventually dies out.

Reproduction numbers allow to also understand some key effects of *mass vaccination*. For sake of simplicity we consider a *perfect* vaccine, i.e. one with 100% take and ensuring lifelong protection against infection. Under the threat of an outbreak, vaccination by a perfect vaccine of a fraction x of the population initially susceptible would decrease the transmission potential from $\mathcal{R}_0$ to $\mathcal{R}_V = \mathcal{R}_0(1 - x)$, the so-called *vaccine reproduction number* (VRN). Therefore, choosing x so as to bring the VRN below one, namely

$$x > 1 - 1/\mathcal{R}_0 = x_c, \tag{22}$$

would prevent the outbreak to occur. This argument extends to infections that prior to vaccination were persisting endemically. Assume that vaccination is administered to a constant proportion x of newborn individuals. If vaccination is continued throughout many consecutive birth cohorts, the susceptible fraction in the population will eventually follow $S \leq (1 - x)$, which in turn implies the inequality $\mathcal{R}_E = \mathcal{R}_0(1 - x)$. Therefore the condition $\mathcal{R}_0(1 - x) < 1$, yielding again $x > x_c$, represents in this case a correct condition of infection elimination (Anderson and May 1991), since it implies that the ERN is definitely forced below 1. The quantity x_c is termed the *critical vaccine coverage*. This result, indicating that a sufficiently large (but always smaller than 100%) vaccine uptake is capable to eliminate infection by breaking the transmission chain, is referred to as the *herd immunity* principle. The critical coverage x_c is increasing and concave in $\mathcal{R}_0$. Therefore, for measles, for which $\mathcal{R}_0$s was estimated to be about 15 in England and Wales prior to vaccination, the corresponding critical coverage is $x_c = 93.3\%$, requiring effective vaccination of nearly all newborns.

Both $\mathcal{R}_0$ and $\mathcal{R}_E$ actually are *model-dependent* parameters. That is, any well posed mean-field epidemiological model as those presented in Section 3.2 (i.e., SIS, SIR, SIRS, SEIR, SEIRS, etc.) will have its own $\mathcal{R}_0$ and $\mathcal{R}_E$, depending more or less complexly on the model parametric structure, as illustrated in forthcoming subsections. However, fully general routes to compute RNs are available for very broad families of epidemiological models [82,84,85,88]. The key underlying idea is that of *next generation operator* (NGO), which is the positive operator associating to any given distribution of infective cases, representing a certain generation of infection, the number of cases in the subsequent generation. Properties of positive operators ensure that iteration of the NGO will eventually bring the exponential growth (per generation) of infective cases at a factor - the *dominant eigenvalue* of the NGO - which is currently taken as the most general definition of RN in deterministic epidemiological models. Historical notes on the ubiquitous role of $\mathcal{R}_0$ in mathematical population dynamics are reported in [89].

3.4. SIR and SEIR models for epidemic outbreaks

What does the framework of Eqs. (9)–(11) further tells us about epidemic outbreaks? For example which is the outbreak's temporal profile? When is the epidemic peak reached? Which is the fraction of the population eventually infected? Let us depart for simplicity from an SIR communicable infection (i.e. we disregard the E class) that is introduced, by importation of a few infective cases, in a wholly susceptible population, that is with an initial susceptible fraction $S(0) \approx 1$. As we are interested here in short outbreaks – such as seasonal influenza, whose outbreaks last a few months, as opposed to tuberculosis, whose historical onset required many decades – we also neglect the population change, setting $\mu = 0$, which amounts to consider a *fixed* population. This is a strong assumption making such a model purely transient. We are thus led to the classical Kermack and McKendrick's SIR epidemic model [75]

$$S' = -\beta SI, \tag{23}$$

$$I' = \beta SI - \gamma I, \tag{24}$$

where $\gamma = 1/D$.

As expected from the presence of a communicable disease, the susceptible fraction S is decreasing, as reflected in Eq. (23). As for the dynamics of the infective fraction, or *prevalence*, $I(t)$, it is useful to rewrite Eq. (24) as

$$I' = \gamma I \left(\mathcal{R}_0 S - 1 \right), \tag{25}$$

where $\mathcal{R}_0 = \beta/\gamma$ indeed represents the BRN specific to the above model, given that $\beta = \beta_I C$, and $\gamma = 1/D$. $\mathcal{R}_0$ fully determines the behavior of the system. Indeed, since $S(t) < S(0) < 1$, the differential inequality $I' \leq \gamma I \left(\mathcal{R}_0 S(0) - 1 \right)$ implies that if

$$\mathcal{R}_0 \leq 1/S(0), \tag{26}$$

then *no outbreak occurs*. Note that the initial infective cases will cause a few other cases but the infection will not represent a threat since there is no increase in prevalence with respect to its initial level I_0. On the other hand, an outbreak is only possible if $I'(0) > 0$, i.e. if

$$\mathcal{R}_0 > 1/S(0). \tag{27}$$

Thus $\mathcal{R}_0$ is the key parameter which determines the *threshold character* of the system, as first noted by Kermack and McKendrick [75]. The result in Eq. (27), can be put in a number of informative ways. First, if the population is largely susceptible, i.e. it holds $S(0) \approx 1$, the threshold value for model of Eqs. (23)–(24) is 1, as claimed in the previous subsection. In the presence of a substantial immune component (i.e. $S(0) < 1$), as is typically the case for seasonal influenza, then the condition $\mathcal{R}_0 > 1/S(0)$ can be written as $\mathcal{R}_0 S(0) > 1$, i.e. $\mathcal{R}_E(0) > 1$, meaning that the occurrence of the outbreak requires an initial ERN greater than 1 (see previous subsection). A further noteworthy reading of condition in Eq. (27) is Kermack and McKendrick's original one [75]: $S(0) > 1/\mathcal{R}_0$, which, in line with John Snow's seminal intuition, states that an outbreak requires a large initial susceptible fraction.

The temporal dynamics of the outbreak under Eq. (27) shows, by linearization of Eq. (25), an initial *exponential* growth in prevalence according to $I(t) \approx I(0)\mathrm{Exp}(\gamma(\mathcal{R}_0 S(0) - 1)t)$, which will be faster the bigger the difference between the initial value of the ERN and its threshold. This initial exponential growth is followed by a non-linear deceleration due to the decline in the susceptible pool S, and consequently in the ERN $\mathcal{R}_{(t)} = \mathcal{R}_0 S(t)$. Moreover, at the time t_x when $S(t_x) = 1/\mathcal{R}_0$ it is $I'(t_x) = 0$, while for all $t > t_x$ it is $I'(t) < 0$, given that $S(t) < S(t_x)$ due to the decrease of $S(t)$. In words, the infective prevalence attains its maximum when the susceptible fraction S becomes equal to the inverse of the BRN, after which the epidemics enter its declining phase until it eventually dies out.

The maximum of the prevalence can be determined by the curve $I(S)$, which is obtained from Eqs. (23)–(24) by taking

$$\frac{dI}{dS} = \frac{1}{\mathcal{R}_0 S} - 1, \tag{28}$$

whose solution in $0 < S \leq S(0)$ (remind that $S(t)$ is decreasing) is

$$I(S) = I(0) + S(0) - S + \frac{1}{\mathcal{R}_0}\mathrm{Log}\left(\frac{S}{S(0)}\right). \tag{29}$$

This implies that the extremum of I is

$$I_{max} = I(0) + S(0) - \frac{\log(1 + \mathcal{R}_0 S(0))}{\mathcal{R}_0}. \tag{30}$$

Therefore, given a wholly susceptible initial population ($S(0) \approx 1$), the maximum prevalence is increasing in $\mathcal{R}_0$. For example, under the Spanish flu hypothesis $I_{max} \approx 45\%$, while for $\mathcal{R}_0 \approx 15$, as commonly observed for measles, then $I_{max} \approx 80\%$.

Moreover, the cumulative fraction hit by the infection at time t is given by

$$J(t) = S(0) - S(t) = I(t) - I(0) + R(t) - R(0). \tag{31}$$

Therefore the fraction eventually infected at the end of the outbreak, the so-called final *attack rate* (*AR*), is given by

$$AR = J(\infty) \approx 1 - S(\infty). \tag{32}$$

Setting in Eq. (29) $I = 0$, $S(0) + I(0) = 1$, $S_\infty \approx 1 - AR$ yields the following transcendental equation for the *AR*

$$AR = 1 - \exp(-\mathcal{R}_0 AR). \tag{33}$$

The AR is therefore an increasing and concave function of $\mathcal{R}_0$. For example, under the Spanish flu hypothesis ($\mathcal{R}_0 \approx 2$) the predicted AR is close to 80%. More important, Eq. (33) shows that for infections imparting immunity, the outbreak will never hit the entire population (though the AR may reach levels only slightly lower than 100% for measles-like values of $\mathcal{R}_0$). Also
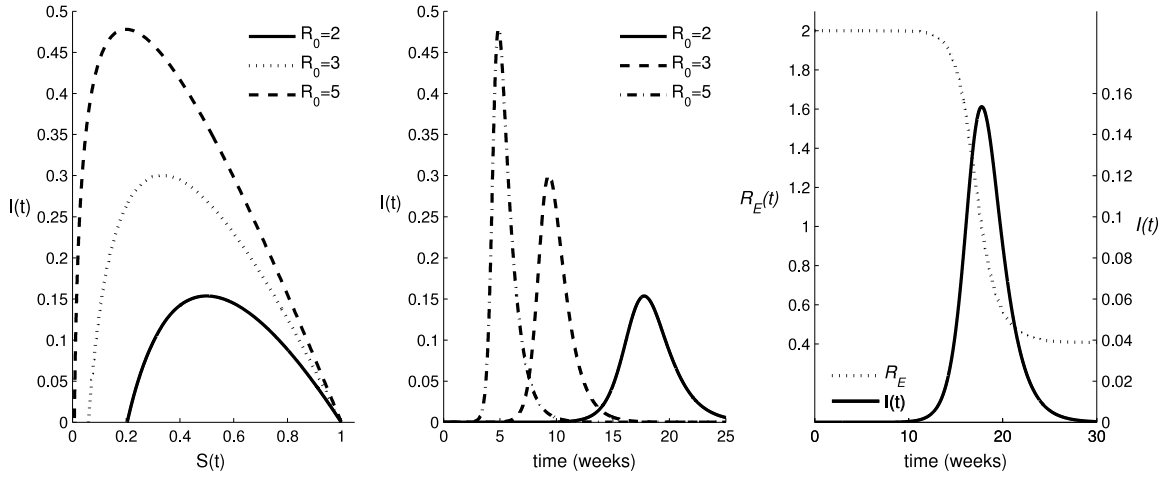
Fig. 4. Dynamics of the epidemic SIR model for different values of the BRN: $\mathcal{R}_0 = 2, 3, 5$. Left panel: dynamics in the S, I phase plane, based on Eq. (29). Central panel: time course of the infective prevalence $I(t)$. Right panel: time course of the infective fraction $I(t)$ and the ERN for $\mathcal{R}_0 = 2$. The duration of the infective phase $D = 1/\gamma$ is set to 3 days.

this *incomplete contagion* principle was first noted by Kermack and McKendrick [75]. Starting from Eq. (29), the case where an initial pool of immune subjects is present ($R_0 > 0$) – as in the case of seasonal influenza outbreaks – is straightforward.

The behavior of Eqs. (23)–(24) above the invasion threshold ($\mathcal{R}_0 > 1$) (Fig. 4) clarifies the nonlinear tuning of $\mathcal{R}_0$ on both the intensity (left panel, based on Eq. (29)) and the time-course of the outbreak (central panel). In particular the decline of the ERN following the decline in the susceptible pool during the outbreak modulates the changes in the form of the infective prevalence (right panel of Fig. 4).

Public health actions aimed at reducing the AR, or at slowing down the outbreak course, e.g. by vaccination and/or by reducing transmission, can be easily accommodated. For example, the effects of an awareness campaigns aimed at favoring behavioral changes affecting the contact rate, could be modeled by means of a time-varying decreasing transmission rate $\beta(t)$.

3.4.1. The SEIR model for outbreaks

The more realistic case where a time-lag between exposure and onset of infectivity exists due to the presence of an exposed, non-infective, phase, can be tackled by the epidemic SEIR model

$$S' = -\beta SI, \tag{34}$$

$$E' = \beta SI - \alpha E, \tag{35}$$

$$I' = \alpha E - \gamma I, \tag{36}$$

As for Eqs. (23)–(24), it holds $S(t) < S(0)$. The joint dynamics of (E, I) is, instead, more complex. However, observe that the subsystem (E, I) is cooperative [90] since

$$\frac{\partial E'}{\partial I} = \beta S > 0, \tag{37}$$

$$\frac{\partial I'}{\partial E} = \alpha > 0. \tag{38}$$

Thus the dynamics of $(E(t), I(t))$ is such that for all t it holds $(E(t), I(t)) < (U(t), W(t))$, where

$$U' = \beta S(0)W - \alpha U, \tag{39}$$

$$W' = \alpha U - \gamma W, \tag{40}$$

(with initial conditions $U(0) = E(0)$ and $W(0) = I(0)$), whose characteristic polynomial is $\lambda^2 + (\alpha + \gamma)\lambda + \alpha(\gamma - \beta S(0))$. Therefore, if $\mathcal{R}_0 < 1/S(0)$ the infection dies out also for SEIR model. Note also that the system of equations (39)–(40) is the one (approximately) governing the initial phase of the epidemics, when $S(t) \approx S(0)$. Thus, if $\mathcal{R}_0 > 1/S(0)$ then there is an initial phase of exponential growth. Finally, for those $t > 0$ such that $S(t)$ remains bigger than $1/\mathcal{R}_0$, it is $(E, I) > (U_1, W_1)$, where

$$U_1' = \beta \left(\epsilon + \frac{1}{\mathcal{R}_0} \right) W_1 - \alpha U_1, \tag{41}$$

$$W_1' = \alpha U_1 - \gamma W_1, \tag{42}$$
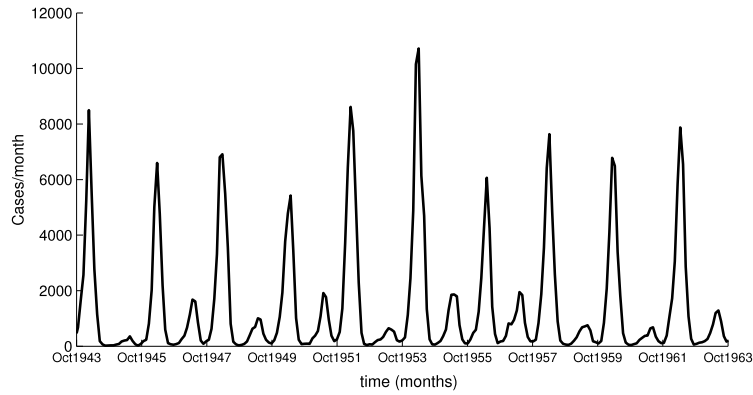
Fig. 5. Biennial pre-vaccination oscillations of measles in New York, 1943–1963.

(with initial conditions $U_1(0) = E(0)$ and $W_1(0) = I(0)$), which is unstable (characteristic polynomial: $\lambda^2 + (\alpha + \gamma)\lambda - \alpha\beta\epsilon$). In other words, as long as $S(t)$ remains bigger than $1/\mathcal{R}_0$, it follows that (E, I) is bigger than an increasing vector.

Summarizing, in vectorial sense, the BRN of the SIR transient model governs also the dynamics of the SEIR transient model. However, note that a vectorial increasing system, i.e. characterized by eigenvalues that have all positive real part, may have some of its components that can be transitorily decreasing. For example, if $E(0) = 0$ then initially $I(t)$ is decreasing: $I'(0) = -\gamma I(0) < 0$.

Finally, let us consider the case of a latent phase of average duration $1/\alpha$ much longer than the infective phase: $\alpha \ll \gamma$ and such that $E(0) = 0$, and that there is initially a single infectious subject $I(0) = 1/N$, so that $S(0) = 1 - 1/N \approx 1$. In such a case, in the time interval $[0, D]$ there is no new infectious subjects $I(t) = constant = 1/N$. It follows from

$$J = N \int_0^D \beta S(t) I(t) dt \approx \beta D = \mathcal{R}_0, \tag{43}$$

where the number J of secondary infections caused by one infectious subject during his entire infectious period in a wholly susceptible population is equal to the BRN.

3.5. Emergence and features of endemicity: the SIR and SEIR models with vital dynamics

The SIR and SEIR epidemic models unavoidably predict the extinction of the infection due to the depletion of the susceptible population. This means that further factors are necessary to sustain *endemicity*, that is the ability of a communicable disease to persist in a population in the long-term, with its peculiar patterns, such as the biennial pre-vaccination oscillations of measles in large UK and US cities (Fig. 5).

The first candidate among such factors is the birth process which allows to rebuild the susceptible pool after an outbreak. The simplest reference model is therefore the SIR model with *stationary vital dynamics* [79,81], which is obtained from Eqs. (9)–(10) by disregarding the exposed phase:

$$S' = \mu(1 - S) - \beta SI, \tag{44}$$

$$I' = \beta IS - (\gamma + \mu)I. \tag{45}$$

Note that system of equations (44)–(45) has a disease-free equilibrium $DFE = (1, 0)$. The BRN specific to this model is given by

$$\mathcal{R}_0 = \frac{\beta}{\gamma + \mu}. \tag{46}$$

Indeed, from $I' = (\gamma + \mu)I(\mathcal{R}_0 S - 1) < (\gamma + \mu)I(\mathcal{R}_0(1 - I) - 1)$ it follows that if $\mathcal{R}_0 \leq 1$ then $I(t) \to 0$ and $S(t) \to 1$, i.e. DFE is globally attractive. Moreover, linearizing Eq. (45) around the DFE gives, letting i to denote the linearized prevalence, $i' = (\gamma + \mu)(\mathcal{R}_0 - 1)i$, implying that if

$$\mathcal{R}_0 > 1$$

then the DFE is unstable and an introduction of infective subjects $i(0) \ll 1$ in a wholly susceptible population induces an outbreak, initially growing as $i(t) \approx i(0)\text{Exp}((\gamma + \mu)(\mathcal{R}_0 - 1)t)$. Moreover, for $\mathcal{R}_0 > 1$ an equilibrium with positive infective prevalence, termed an *endemic equilibrium*, appears

$$EE = (S_e, I_e) = \left(\frac{1}{\mathcal{R}_0}, \frac{\mu}{\mu + \gamma} \left(1 - \frac{1}{\mathcal{R}_0} \right) \right). \tag{47}$$
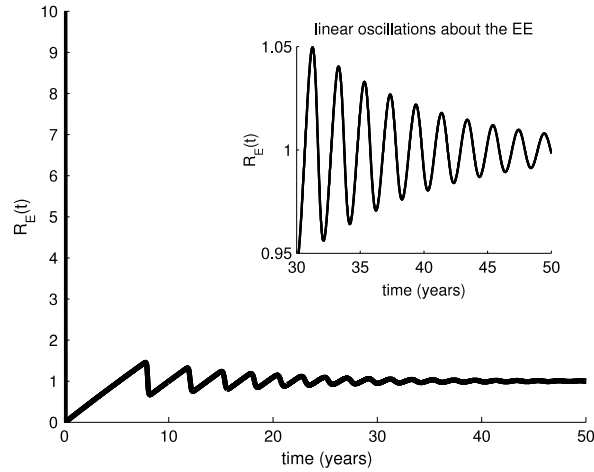
Fig. 6. The SIR model with vital dynamics: pre-vaccination dynamics of the effective reproduction number ($\mathcal{R}_E(t)$), with convergence to the endemic state. The pattern of linear oscillations about the *EE* is reported in the subgraph, showing the biennial inter-epidemic period. The demographic and epidemiological parameters are: $\mu = (1/75)$ year^{-1}, $\gamma = 52$ year^{-1}, $\mathcal{R}_0 = 15$.

Note that in the case where $\gamma \gg \mu$, common for childhood infections, it is $I_e < (\mu/\gamma) \ll 1$. The EE is globally attractive, as it follows by applying the Poincaré-Bendixon theorem with multiplying factor $1/I$

$$div\left(\frac{1}{I}(S', I')\right) = -\frac{\mu}{I} - \beta. \tag{48}$$

To sum up, technically speaking, $\mathcal{R}_0 > 1$ induces a *transcritical bifurcation* at $\mathcal{R}_0 = 1$, such that below threshold only the DFE exists and is GAS, while above threshold the endemic state appears, which inherits the GAS, and promotes an infective prevalence which is strictly increasing in $\mathcal{R}_0 > 1$. In words, if $\mathcal{R}_0 \leq 1$ then the infection is not able to self-sustain and dies out, whereas if $\mathcal{R}_0 > 1$ the infection will invade the population and remain endemic. The behavior of the endemic state is studied by linearizing around the EE and computing the associated characteristic polynomial

$$\lambda^2 + \mu\mathcal{R}_0\lambda + \mu(\mu + \gamma)(\mathcal{R}_0 - 1) = 0. \tag{49}$$

The associated discriminant is $\Delta = \mu^2\mathcal{R}_0^2 - 4\mu(\mu+\gamma)(\mathcal{R}_0-1) \approx -4\mu(\mu+\gamma)(\mathcal{R}_0-1) < 0$, implying dumping oscillations of prevalence about its endemic level with natural (pseudo-) period:

$$T = \frac{2\pi}{\sqrt{\mu(\mu + \gamma)(\mathcal{R}_0 - 1)}}. \tag{50}$$

In the case of measles, for which $1/\gamma \approx 1$ week, and estimates of $\mathcal{R}_0$ in the region of 15 [79] were common in western countries in the pre-vaccination epoch, by taking a life expectancy at birth of $1/\mu = 75$ years, one finds $T \approx 2.01$ years, therefore reproducing the measles inter-epidemic period observed in the UK and the US cities before the advent of vaccination.

The global dynamics of the system, particularly the interplay between infection and demographics, can be unfolded by a time-scale argument, departing from the introduction of the infection in a wholly susceptible population. Assume for simplicity that $\mathcal{R}_0$ is greater than 1 and large, and that epidemiological and demographic time scales are well-separated i.e. $\gamma \gg \mu$. In this case there is an initial phase where model dynamics is essentially described by model of Eqs. (23)–(24) for outbreaks, yielding a large initial epidemics, which exhausts most of the susceptible population, bringing the ERN (and prevalence) to very low levels, as clear by the thick line overlapped with the vertical axis in Fig. 6. This implies that thereafter the S-equation takes the approximate form $S' \approx \mu$, which has a twofold effect: it promotes a phase of rebuilding of the susceptible pool due to the new births on the one hand, while on the other hand it slows down the decline in prevalence, which avoids the infection to simply die out as in Eqs. (23)–(24). The growth in susceptible population will eventually end by a new, but less severe, epidemics, which will in turn be followed by a new phase of replenishment of the susceptible pool, and so on [79] until the endemic equilibrium is achieved (Fig. 6). The global dynamics then consists in continued alternation between epidemic phases, followed by epochs of demographic replenishment, which is how we know the *natural history* of many infectious diseases. The biennial period of quasi-linear oscillations about the EE is emphasized in the sub-graph of Fig. 6.

3.5.1. Periodic contact rates

The simple endemic SIR model with constant transmission rate β and stationary population dynamics plays a fundamental role as trigger of the studies on vaccine-preventable infections of childhood, since it well reproduces, by

minimal ingredients, their main qualitative and quantitative features such as the endemic oscillation and the observed inter-epidemic period [79,81]. However, the model fails to capture a number of further characteristics, ranging from the sustained biennial pre-vaccination oscillations of measles in US and UK cities (see again Fig. 5), to the more complicate dynamics observed in other contexts, both prior and after mass vaccination [91–93]. Motivated by the fact that both the weather and certain social processes, namely school-recruitment following the yearly school calendar, have one year periodicity, an important strain of the literature has focused on $1 - year$ periodic $\beta(t)$ in endemic SIR models. If the corresponding period T_β is around 1 year, then nonlinear resonance may appear, inducing phenomena such as biennial and multi-year periodicity, and chaos [91,93–99] that partially explain a number of observed epidemiological time-series.

The effect of periodic forcing in the contact rate on the stability of the disease-free equilibrium, is found by defining

$$\mathcal{R}_0(t) = \frac{\beta(t)}{\mu + \gamma}, \tag{51}$$

and operating as in the case of constant BRN, i.e. by differential inequality and linearization. Letting $\langle \mathcal{R}_0(t) \rangle$ to denote the average of $\mathcal{R}_0(t)$ over $[0, T_\beta]$, one easily gets that: 1) $\langle \mathcal{R}_0(t) \rangle \leq 1$ implies the global attractiveness of DFE; (2) $\langle \mathcal{R}_0(t) \rangle > 1$ implies the instability of the DFE. It is important to note that in the SIR model the stability and instability of DFE are independent of the specific shape of $\beta(t)$: they only depend on its average value. However, this independence of the behavior on the entire shape of $\beta(t)$ is a peculiarity of the SIR and of some other epidemic models, but it cannot be generalized.

The inclusion of a latency period in the SIR endemic model leads to the SEIR endemic model

$$S' = \mu(1 - S) - \beta SI, \tag{52}$$
$$E' = \beta IS - (\gamma + \alpha)E, \tag{53}$$
$$I' = \alpha E - (\gamma + \mu)I, \tag{54}$$

whose BRN is given by

$$\mathcal{R}_0^{SEIR} = \frac{\alpha}{\alpha + \mu} \frac{\beta}{\mu + \gamma}. \tag{55}$$

Note that if $\mu \ll \alpha$ then $\mathcal{R}_0^{SEIR} \approx \mathcal{R}_0^{SIR}$. It holds that (i) if $\mathcal{R}_0^{SEIR} \leq 1$ then the DFE is the unique equilibrium and it is GAS; (ii) if $\mathcal{R}_0^{SEIR} > 1$ then there is a unique EE, which is GAS. The latter property has been demonstrated by Li and Muldowney [100] as an application of their 3D extension of Poincaré-Bendixon theorem. As for the SIR case, nonlinear resonances are possible for a $1 - year$ periodic contact rate under parameter constellations characteristic of childhood infections.

Interestingly, in the SEIR model the deterministic periodic fluctuations of the contact rate have nontrivial effects also on the stability of the DFE. As regards the global stability of the DFE, also here the cooperativity of subsystem (E, I) implies that for all $t > 0(E(t), I(t)) < (U(t), W(t))$, where

$$U' = \beta(t)W - \alpha U, \tag{56}$$
$$W' = \alpha U - \gamma W, \tag{57}$$

with initial conditions $U(0) = E(0)$ and $W(0) = I(0)$. It follows that if both the eigenvalues of the Floquet's matrix F associated to Eqs. (56)–(57) are interior to the unit circle in the complex plane, then

$$(U(t), W(t)) \to (0, 0) \Rightarrow (E(t), I(t)) \to (0, 0) \tag{58}$$

in turn implying that $S(t) \to 1$. As regards the instability of DFE, by linearizing the SEIR model around the DFE one gets a linear time-varying system equivalent to Eqs. (56)–(57). Thus if the eigenvalues of the matrix F are outside the unit circle of the complex plane, then the DFE is unstable.

It is of interest to note that although $\mathcal{R}_0^{SEIR}(t) \approx \mathcal{R}_0^{SIR}(t)$, for the SEIR endemic model the stability of the DFE is deeply different from the SIR case. Indeed, in the SIR model the behavior of the DFE is uniquely determined by the average of the transmission rate $\beta(t)$, while in the SEIR case the stability of the DFE depends on the "whole shape" of $\beta(t)$.

The computation of Floquet's matrices is usually nontrivial. However, matrix F can be easily computed in the noteworthy case where $\beta(t)$ is switching between two constant values, which mimics well the realistic school-holiday scenario where students have a large number of contacts during the school period, which remarkably reduces during holidays [92,101]. Setting

$$\beta(t) = \begin{cases} \beta_1, & \text{if } t \in [0, T], \\ \beta_2, & \text{if } t \in (T, 1], \end{cases} \tag{59}$$

the linear time-varying system of period one year is equivalent to the following two systems

$$X'_1 = M_1 X_1; \, 0 \leq t \leq T, \tag{60}$$
$$X'_2 = M_2 X_2; \, T < t \leq 1, \tag{61}$$

where for $i = 1, 2$

$$M_i = \begin{bmatrix} -\alpha & \beta_i \\ \alpha & -\gamma \end{bmatrix}. \tag{62}$$

As a consequence F is the product of two matrix exponentials:

$$F = \text{Exp}(M_2(1 - T))\text{Exp}(M_1 T). \tag{63}$$

The generalization to more complex patterns of alternance between school and holidays can also easily handled.

In the general case, if one eliminates the first state variable in the above equivalent linear systems, one obtains the following *family* of Newton's equation [90]

$$w'' + (\gamma + \alpha)w' + (\gamma - \alpha\beta(t))w = 0. \tag{64}$$

Further, by setting $y(t) = w(t)\text{Exp}((\gamma + \alpha)t/2)$ [102], one yields the Hill family of equations

$$y'' + \left(-\frac{(\gamma + \alpha)^2}{4} + \gamma - \alpha\beta(t) \right) y = 0. \tag{65}$$

The families of Eqs. (64)–(65) are ubiquitous and widely studied in various fields of mathematical physics [102–108]. For example, in mechanics they model the periodic perturbations in the elastic force of damped or undamped oscillators respectively, which are at the base of several theories, such as that of parametric resonance [106–108]; while in applied electromagnetism Eq. (65) models the frequency modulation [104]. Analytic solutions for $t \in [0, 1]$ are known for a number of particular forms of $\beta(t)$. For example, if $\beta(t)$ is linearly increasing, the solutions of Eq. (65) in $t \in [nT, nT + 1)$ are Airy's functions, whereas if the contact rate is sinusoidally varying around an average value, the solutions of Eq. (65) are Mathieu's functions.

3.5.2. Nonlinear contact rates as a phenomenological route to behavior change

A limitation of mass-action-based SIR and SEIR models with and without vital dynamics, and of other epidemic models, is that they assume that the behavioral component is either constant (resulting in a constant β) or influenced by exogenous phenomena such as the school/work calendar (resulting in a periodic time-varying $\beta(t)$ with one-year period). Capasso and Serio were first in pointing out this shortcoming [109], and in stressing that agents might respond to the threat posed by the outbreak by changing their social behavior (i.e. reducing their average number of contacts C), and/or by adopting protections against transmission. They modeled these effects by a *prevalence-dependent* transmission rate of the form $\beta(I)$, $\beta'(I) < 0$, showing that behavior change may deeply modify the quantitative dynamics of an outbreak [109]. In a more general study on the SIR endemic model [110], it was stressed that prevalence-dependent behavior is influenced by a range of information variables, some of which can be made available to agents with time-lags. The inclusion of delayed prevalence-dependent behavior in the SIR endemic model can destabilize the endemic equilibrium by triggering sustained oscillations, and, in turn, these endogenous oscillations may interact with the one-year periodicity of the contact rate, resulting in chaotic fluctuations [110]. Finally, another seminal paper is Hadeler and Chavez [111], which models the spread of a vaccine/education preventable sexually transmitted disease in a population split in two groups: an inactive group with $\beta = 0$ and a *core* active group where $\beta > 0$. The model included a unidirectional prevalence-depending flux of subjects from the inactive to the *core* group, with transfer rate $r(I)$ decreasing with I.

3.6. Mass vaccination and herd immunity for vaccine preventable infections: the SIR model with vaccination at birth

In modern communities new vaccines, targeting different infections and different populations, appear at a continuously increasing rate, making the mosaic of characteristics of the different available vaccines quite articulated. Nonetheless, much of the main dynamic effects of mass vaccination in controlling communicable diseases can be understood by the following simple variation of the endemic SIR model of Eqs. (44)–(45)

$$S' = \mu(1 - x) - \mu S - \beta SI, \tag{66}$$

$$I' = \beta IS - (\gamma + \mu)I, \tag{67}$$

$$V' = \mu x - \mu V. \tag{68}$$

Model of Eqs. (66)–(68) (of course $R = 1 - S - I - V$) considers vaccination of a time-invariant, randomly chosen, proportion x of newborn susceptible individuals per unit of time by a *perfect vaccine*, i.e. one with 100% take and lifelong duration (see Section 3.3). Under a perfect vaccine the fraction successfully vaccinated coincides with the *vaccine coverage*, i.e. the fraction vaccinated. Vaccinated individuals are fully protected against infection (and ensuing disease) and therefore behave, in practice, as removed individuals. This does not need to be the case if, for example, the vaccine has waning immunity, or leaves some residual susceptibility – as is the case e.g. for the varicella vaccine – in which case further compartments are required.

After discarding the V equation, the analysis of Eqs. (66)–(67) goes along exactly the same lines as Eqs. (44)–(45). Since we are interested in infections that are endemic prior to vaccination, we assume $\mathcal{R}_0 > 1$, where $\mathcal{R}_0$ still represents the BRN. The key threshold quantity is now represented by the VRN

$$\mathcal{R}_V = \frac{\beta}{\gamma + \mu}(1 - x). \tag{69}$$

It can then be easily shown that Eqs. (66)–(67) have the $DFE = (1-x, 0)$, which is globally attractive if $\mathcal{R}_V \leq 1$, and unstable for $\mathcal{R}_V > 1$. For $\mathcal{R}_V > 1$ the model also has the globally attractive (by the same Poincaré-Bendixson argument) endemic state

$$EE_V = (S_e, I_e) = \left(\frac{1}{\mathcal{R}_0}, \frac{\mu}{\mu + \gamma} \left(1 - x - \frac{1}{\mathcal{R}_0} \right) \right). \tag{70}$$

This endemic state is also characterized by damped oscillations, with pseudo-period

$$T_V = \frac{2\pi}{\sqrt{\mu(\mu + \gamma)(\mathcal{R}_0(1 - x) - 1)}}. \tag{71}$$

Though the mathematical analysis is the same as for the basic endemic SIR model, there is a number of substantive implications of vaccination. Indeed, the threshold for the VRN implies the already met coverage threshold: $x_c = 1 - 1/\mathcal{R}_0$ (i.e. formula (22)), so that if $x \geq x_c$ then the infection is eliminated (technically, the introduction of mass vaccination above threshold causes the EE to disappear, leaving only the DFE, which becomes GAS). In the opposite scenario, i.e. if the coverage is insufficient to achieve elimination ($px \leq x_c$), the infection will continue to persist endemically, in the new long-term regime described by Eq. (70). This *post-vaccination* regime is characterized by the same susceptible proportion ($1/\mathcal{R}_0$) as in the pre-vaccination history, but by a different distribution of immune individuals between natural and vaccine-induced, and especially by a smaller prevalence and an increased inter-epidemic period with respect to the case where no vaccination campaign is enacted. The increase in T straightforwardly follows from the reduced replenishment of the susceptible class due to vaccination. For example, considering again a measles-type parametrization ($\mathcal{R}_0 = 15$, $1/\gamma = 1$ week, $\mu = (1/75)/\text{year}$), yielding a pre-vaccination inter-epidemic period of about 2 years, a mass vaccination program with $x = 0.9$ would increase the inter-epidemic period to a post-vaccination figure T_V in excess of 10 years.

The dynamic implications of the introduction of a mass vaccination for a measles-like infection are illustrated in Fig. 7 for different vaccination programs (respective coverages: $x = 0.80$, $x = 0.90$ and $x = 0.95$) initiated at time $t = 20$ years under the hypothesis that infection spread for $t < 20$ was evolving in its pre-vaccination regime, characterized by an almost biennial inter-epidemic period. As expected for $x = 0.95$ the infection is eliminated, since $x > x_c$. In the two remaining cases the infection remains endemic with a much lower prevalence but with an amazing increase in the inter-epidemic period. For example, for $x = 0.90$ the inter-epidemic period initially increases to a level in excess of 20 years before gradually contracting and converging in the long-term to the above predicted figure of $T_V \approx 10$ years. The inflation in the inter-epidemic period has far reaching consequences since it implies that unvaccinated individuals who did not acquire infection early in life might do so at somewhat later ages, where they might be exposed at a much larger risk of serious disease from the infection [79,81]. The increase in the average age at infection in the post-vaccination regime is the consequence in the decline in the individual risk of infection, as clear from the decline in the endemic FOI βI_e, which shifts from its pre-vaccination level of $\mu(\mathcal{R}_0 - 1)$ to $\mu(\mathcal{R}_0(1 - x) - 1)$. These phenomena are of importance, besides measles, for essentially all vaccine preventable infection, particularly for rubella [79,81].

3.7. Return to susceptibility: SIS and SIRS models

If infection-acquired immunity is not lifelong, recovered subjects will eventually become again susceptible. A first possible scenario is the one where infective subjects directly return to susceptibility, yielding the so-called SIS model (Fig. 3)

$$S' = \mu(1 - S) - \beta(t)IS + \gamma I, \tag{72}$$
$$I' = \beta(t)IS - \mu I - \gamma I. \tag{73}$$

Since $S + I = 1$, one gets the scalar model

$$I' = (\mu + \gamma)I(\mathcal{R}_0(t)(1 - I) - 1), \tag{74}$$

where

$$\mathcal{R}_0(t) = \frac{\beta(t)}{\mu + \gamma} \tag{75}$$

and we still take $\beta(t)$ as a 1-year periodic function. Unsurprisingly, if $\langle \mathcal{R}_0(t) \rangle \leq 1$ then the DFE $I = 0$ is globally attractive, whereas if $\langle \mathcal{R}_0(t) \rangle > 1$ the DFE is unstable, and $I(t)$ tends to a one-year periodic solution, i.e. the model predicts that the
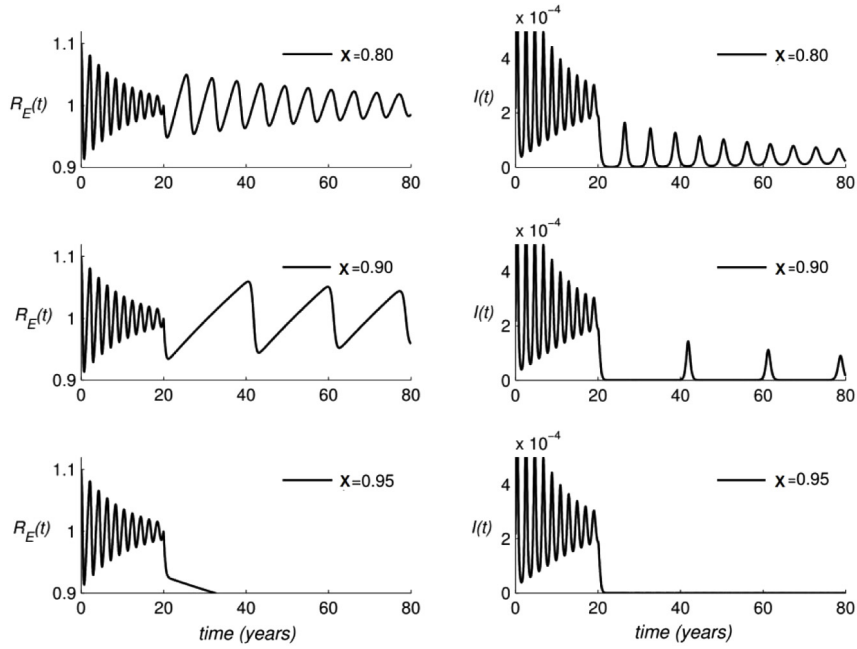
Fig. 7. The SIR model with vital dynamics: pre- and post- vaccination dynamics of the effective reproduction number $\mathcal{R}_E(t)$ (left panels) and the infective prevalence $I(t)$ (right panels) under different programs of vaccination at birth initiated at time $t = 20$ years with coverage of $x = 0.80$ (upper panels), $x = 0.90$ (central panels) and $x = 0.95$ (right panels) respectively. The demographic and epidemiological parameters are: $\mu = (1/75)$ year^{-1}, $\gamma = 52$ year^{-1}, $\mathcal{R}_0 = 15$.

disease remains endemic via a series of annual epidemic outbreaks. For a constant transmission rate β Eq. (74) is logistic, therefore implying a constant, globally attractive, endemic state:

$$I_{EE} = 1 - \frac{1}{\mathcal{R}_0}, \tag{76}$$

where the susceptible fraction is still given by the inverse of $\mathcal{R}_0$. Note that the establishment of a globally attractive EE also holds in absence of vital dynamics (i.e. for $\mu = 0$, suggesting that the loss of immunity *per se* is a further source of endemicity through rebuilding of the susceptible pool, as first noted by [75].

A second more general scenario occurs when return to susceptibility is buffered by a long sojourn time in the recovered class, yielding the SIRS endemic model (Fig. 3)

$$S' = \mu(1 - S) - \beta SI + \delta R, \tag{77}$$
$$I' = \beta IS - \gamma I - \mu I, \tag{78}$$
$$R' = \gamma I - \mu R - \delta R, \tag{79}$$

where $1/\delta$ is the average length of the immunity conferred by the disease. From $S = 1 - I - R$ one gets

$$I' = (\mu + \gamma)I(\mathcal{R}_0(t)(1 - I - R) - 1), \tag{80}$$
$$R' = \gamma I - (\mu + \delta)R, \tag{81}$$

where

$$\mathcal{R}_0(t) = \frac{\beta(t)}{\mu + \gamma}. \tag{82}$$

Again, the infection ultimate fate, namely extinction vs endemicity, depends on whether $\langle \mathcal{R}_0(t) \rangle$ is smaller or bigger than one. Under constant β, the threshold condition $\mathcal{R}_0 > 1$ allows the existence of an endemic state:

$$S_{EE} = \frac{1}{\mathcal{R}_0}, I_{EE} = \left(1 - \frac{1}{\mathcal{R}_0}\right)\frac{\mu + \delta}{\mu + \delta + \gamma}, \tag{83}$$

which can be shown to be globally attractive by applying the same Dulac–Bendixon test used for the SIR endemic model.

Mass vaccination by a perfect vaccine can be introduced in SIS and SIRS models along the same route followed for the SIR model, with analogous results. However, for infections with only temporary, or without, immunity, the hypothesis of a

perfect vaccine is mostly unattainable. If the vaccine is *imperfect*, several scenarios are possible i.e. vaccinated individuals might either (i) not produce any immune response, therefore remaining fully susceptible, or (ii) produce a weaker immune response. This weaker immune response might occur in the form of a temporary immunity which however typically wanes faster than natural one, eventually returning the individual to susceptibility, or in the form of reduced susceptibility, putting the individual at risk of *breakthrough* infection, at some reduced force of infection $\lambda = \sigma \beta I$ with $\sigma < 1$. For example, all these possibilities are known to occur for the current varicella vaccine [112]. This opens a more complicate scenario which makes infection elimination more difficult. Indeed, as it can be demonstrated even for simple SIS models with vaccination at birth, there is a range (x_c^0, x_c^*) of the vaccinated proportion p where a bistability phenomenon arises, i.e. where the disease-free equilibrium is locally stable but it coexists with two endemic equilibria: one unstable and one locally stable. Therefore, infection elimination requires that $x > x_c^*$, whereas if $x < x_c^0$ then there is only one endemic steady state, coexisting with the DFE which is unstable. This whole scenario is termed as backward bifurcation [113].

3.8. Remarks on the wise use of the mean-field principle in theoretical epidemiology

Thanks to the use of a few important deterministic models, theoretical epidemiology has become a key supporting tool to public health decision makers. These models, historically derived from classical physical chemistry and other deterministic applications of statistical mechanics, are, however, only approximations of the underlying real-world stochastic processes, obtained from the mean-field limits of stochastic birth and death processes analogous to those adopted in chemical kinetics [114].

In this subsection we therefore concisely remind some well-known tools of nonequilibrium statistical physics [114,115] with the aim to recall how deterministic epidemic models arise as limits on the underlying stochastic processes on the one hand, and their shortcomings, on the other hand.

For example, let us consider an epidemic (i.e. no vital dynamics) SIR infection process, with state variables $(X(t), Y(t), Z(t) = N - X(t) - Y(t))$. In this process there are two types of events: the *infection*, responsible for the $S \to I$ transition with infinitesimal probability

$$Prob\,((X(t+dt), Y(t+dt)) = (X(t) - 1, Y(t) + 1)) = \beta \frac{Y(t)}{N} X(t) dt, \tag{84}$$

and *recovery*, responsible for the $I \to R$ transition with infinitesimal probability

$$Prob\,((X(t+dt), Y(t+dt)) = (X(t), Y(t) - 1)) = \gamma Y(t) dt. \tag{85}$$

For the SIS model (still without vital dynamics), since $X(t) = N - Y(t)$, the $S \to I$ transition has probability $\beta(Y/N)(N - Y)$.

The behavior of the realizations of these epidemic stochastic processes can be simulated exactly by means of the Gillespie algorithm [116], whereas the probability density function of its state variables is described by a multivariate Master equation (ME) [114,115], which is bivariate for the epidemic SIR process, and univariate for the SIS process. For example, for the epidemic SIS model the ME reads

$$\frac{d}{dt} P(t, Y) = \beta \frac{Y-1}{N} \left(N - (Y - 1) \right) P(t, Y - 1) - \beta \frac{Y}{N} \left(N - Y \right) P(t, Y) + \gamma (Y + 1) P(t, Y + 1) - \gamma Y P(t, Y). \tag{86}$$

Taking the average $\langle Y(t) \rangle = \sum Y P(t, Y)$ from Eq. (86), we get

$$\frac{d}{dt} \langle Y(t) \rangle = (\beta - \gamma) \langle Y(t) \rangle - \beta \frac{1}{N} \langle Y^2(t) \rangle. \tag{87}$$

Essentially the mean field hypothesis applied in theoretical population models stands, for large populations, in neglecting the correlations between the state variables and assuming that the average of the product of two state variables is approximately equal to the product of the averages [117], for example in Eq. (87) above $\langle Y^2(t) \rangle \approx (\langle Y(t) \rangle)^2$.

However, due to the approximate nature of deterministic epidemic model and their impact of public health sciences, it is important to go beyond the mere application of the mean-field limit, and to understand under which conditions deterministic differential equations (whose state variables are continuous function of time) can safely be used in place of the originary "exact" stochastic formulation of the epidemic process (whose state variables are time-continuous stochastic processes taking integer non-negative values). Moreover, it is also important to have a feeling of the behavior of the system at least in the regime of small stochastic deviations from the deterministic behavior.

If the total population size is small, birth–death stochastic epidemic models must be used. On the other hand, when the size of the state variables becomes large (which is the case intuitively leading to the deterministic approximation), it is hard to simulate the process due to the crowding of events. In such a case, however, one can use a system-size approximations [114,115]. This methodology allows approximating, respectively, the solution of a multivariate ME with a multivariate Fokker–Planck equation (FPE) [114] (this approximate equation is sometime called Chemical FPE [114]). In turn, from the obtained FPE one can derive a corresponding system of Stochastic Differential equations (SDEs) that approximate the dynamics of the state variables vectors. In physical chemistry it has been proposed a reverse equivalent approach, the Chemical Langevin equation (CLE) [118], where first the approximating SDE system is obtained from the definition of the model, and

only then the FPE is written. Namely, a spatially homogeneous system of particles/subjects with state variable u character-ized by M "events" having probability $a_j(u)dt$ $j = 1, \ldots, M$ can be approximated by the following system of SDEs [118]

$$u'_i = \sum_j \nu_{i,j}(a_j(u) + \sqrt{a_j(u)}\xi_j(t)), \tag{88}$$

where the $\xi_j(t)$ are M uncorrelated white noises, and ν is the *stoichiometric* matrix where $\nu_{i,j} \in \{0, -1, +1\}$ depending on the changes induced by the jth event into the ith state variable. Note that here the original state vector, whose components were integer non-negative numbers, is approximated by a real vector.

For example, in the case of the SIR epidemic model, the CLE reads

$$X' = -\beta \frac{Y}{N} X - \xi_1(t)\sqrt{\frac{Y}{N}X}, \tag{89}$$

$$Y' = \beta \frac{Y}{N} X - \gamma Y + \xi_1(t)\sqrt{\frac{Y}{N}X} - \xi_2(t)\sqrt{\gamma Y}. \tag{90}$$

Since N is constant, we can pass to the fractions (S, I)

$$S' = -\beta SI - \xi_1(t)\frac{1}{\sqrt{N}}\sqrt{\beta SI}, \tag{91}$$

$$I' = \beta SI - \gamma I + \xi_1(t)\frac{1}{\sqrt{N}}\sqrt{\beta SI} - \xi_2(t)\frac{1}{\sqrt{N}}\sqrt{\gamma I}. \tag{92}$$

As a consequence, in the limit $N \to +\infty$ we obtain the deterministic model.

It is interesting to note that the use of the force of infection FOI $= \beta_1 Y$ leads, after passing to fractions, to a stochastic term $\xi_1(t)\sqrt{\beta_1 SI}$ which is not scaling as $1/\sqrt{N}$, and thus it does not disappear in the limit $N \to +\infty$. Moreover, note that an hypothetical model where the contact rate is decreasing to 0 with the population size N, would yield the following trivial uncoupled model of the type $S' = 0, I' = -\gamma I$.

As stressed by a few authors among which Gillespie [118], the accuracy of the approximation given by Eq. (88) or by those derived by the system-size expansion [114,115], depends on the fulfillment of specific conditions that are *time-varying* since they are state-variable dependent. We can roughly summarize these conditions as follows: for all times $a_j(u)dt$ must be suffi-ciently larger than $\sqrt{a_j(u)dt}$. This in turn implies that all state variables must be sufficiently large. Thus the above mentioned approximations require caution, since *a rigori* they are invalid even when only one of the state variables becomes small.

For the specific case of the SIR model, note that in the initial epidemic phase if one sets $X(t) \approx X(0) \approx N$ the ensuing stochastic process for $Y(t)$ is linear and its mean is given by $E[Y] = Y(0)Exp((\beta S(0) - \gamma)t)$ as in the approximation of the deterministic SIR epidemic model. The corresponding variance is $Var[Y] = Y(0)(e^{(\beta S(0)-\gamma)t} - 1)/(\beta S(0) - \gamma)$. Of course, this approximation is valid for a relatively short time-interval. However, as Bartlett noted in his pioneering work on measles [119], the most critical difference between deterministic and stochastic models arises in the intervals between epidemics (for example during the undamped oscillations caused by periodic changes in the contact rate) when the prevalence can become so small that extinction of the infective population almost certainly occurs in the stochastic formulation *also in presence of a large population* while deterministic models (e.g. the periodically forced SIR and SEIR models with vital dynamics) simply predict that system trajectories will be positive for all times. This is of relevance in presence of vaccination because stochasticity may cooperate with the further reduction in prevalence caused by vaccination in inducing infection elimination [120].

The study of stochasticity-induced infection extinction led Bartlett to introduce [119] the key concept of *critical community size* (CCS). This is the (heuristically defined) threshold over which, in absence of vaccination, the infection is likely to persist, and below which extinction is very likely. Over the CCS it is therefore safe to "read" the output of a deterministic model as an average behavior. For example pre-vaccination measles data indicate a CCS in the region of 250,000–400,000 [119,121]. The CCS is only weakly associated to the BRN, while it is strongly dependent on the duration of the infectious and latent periods. As for the latter, it was noted that simple stochastic models based on exponentially distributed durations, tend to severely overestimate the CCS [121], while instead adequate predictions can be obtained by resorting to more realistic, e.g. bell-shaped, distributions of the latent and infectious periods [121]. On the other hand, the effect of space on the CCS is extremely complex [120].

Stochastic epidemic models are characterized by a range of interesting phenomena not shown in their deterministic counterparts. The simplest is obviously stochastic extinction after importation, i.e. the fact that an infectious individual in-troduced in a fully susceptible population has a probability of recovering before being able to infect anyone (given by $1/\mathcal{R}_0$ $(1/\mathcal{R}_0^k)$ [120]), in which case no epidemics will occur, unlike the deterministic model. Other stochastic effects are more subtle. For example, in the SIR model with vital dynamics, the stochastic fluctuations of the total population $N(t)$ around its deterministic equilibrium (which can be approximated by the CLE: $N' = \xi_b(t)\sqrt{\mu N} - \xi_d(t)\sqrt{\mu N}$), may transform the damped oscillations around the endemic equilibrium of the corresponding deterministic model in sustained oscillations [122]. This phenomenon, termed *stochastic amplification*, is an instance of the more general phenomenon of *coherence reso-nance* [122]. The theoretical possibility of stochastically inducing the disease eradication with a limited amount of vaccines

via periodic delivery of vaccinations so that to induce resonances, and the related role of seasonal fluctuations of the contact rates, are investigated in [123] by means of methodologies of the theory of fluctuation-induced population extinction. Further details on the stochastic theory of epidemics can be found in the complementary books [124,125].

3.9. Space and beyond: mean-field metapopulations

Everytime a new epidemic focus appears, the first concern of public health is to assess whether this isolated focus will remain localized or will spatially spread. This leads to a key missing dimension in the previous simple models, namely *space*. The most straightforward way of modeling the geographical spread of infections is to add spatial diffusion to the basic SIR and SEIR model [83,126,127], leading to equations that are analogous to the reaction–diffusion equations of theoretical biochemistry [126,128]. This approach, which is proved useful to understand some basic aspects of spatial epidemic behavior by using tools of reaction–diffusion theory, has obvious shortcomings. The first one is that human beings have patterns of movement dramatically different from that of particles moving under Fickian diffusion [129,130]. This is a major complication leading to complex integro-differential equations where the diffusion part interacts with the reaction part, i.e. transmission [131]. Moreover, human movement often includes rapid commuting between areas with highly clustered population, such as cities [120,132].

To cope with the above problems, population ecology and mathematical epidemiology have increasingly used the *metapopulation*, or *multi-patch* approach, where a large populations is split in a finite number of *patches*, for example cities, where: (i) each city obeys its own deterministic "local" epidemic model, e.g. SIR-type; (ii) local dynamics are linked through population fluxes, e.g. due to commuting for work. Commuting plays a fundamental role in the geographical spread of infectious diseases in contemporary societies [120,132], by favoring the geographical spread and persistence of infections through the interlink between patches. The multi-patch approach was used in the pioneering work by Baroyan and Ravchev, who linked a large number of SIR epidemic models to simulate the spread of flu in URSS cities [127]. Multi-patch epidemiological models have since then developed fast both on the theoretical and the applied side, e.g. for predicting the world-wide spread of pandemic influenza [133–136].

As for the mathematical viewpoint, multipatch epidemic models show the same spatial phenomena we may observe in traditional reaction–diffusion models, namely: (i) emergent phenomena such as the onset of spatial heterogeneities by means of spatial instability of spatially homogeneous equilibrium points (Turing-like instabilities) or even of limit cycles [137]; (ii) the onset of traveling waves and solitons, representing the spatial displacement of epidemic outbreaks [120]; (iii) the emergence of various spatio-temporal asynchronous/synchronous complex patterns [120,138]. To these classes of phenomena, one has to add those induced by parametric heterogeneities between patches [120], which might mirror heterogeneities in the epidemiology (e.g. in social contacts), public health interventions (e.g. in vaccine uptake), demographics (e.g. population density and age-structure), as well as environmental (e.g. climate, temperature, etc.). As regards the onset of Turing-like patterns is concerned, let us consider the following system [137] for an infection with m epidemiological states ($m = 4$ for a SEIR system) spreading among n interacting patches. If patches are isolated, the local spread is modeled by n m-dimensional ODE systems of the form $u'_j = f(u_j)$. In presence of linear population flows between patches one has the following interlinked dynamical systems

$$u'_j = f(u_j) + \sum_{i=1}^{n} c_{i,j} M u_i, \tag{93}$$

where $u_i \in R^m$ are the state variables associated to patch i, $U = (u_1, \ldots, u_n)$ is the $m*n$ matrix of all state variables, M is a $m*m$ diagonal matrix reporting the different mobility coefficients of individuals in the different epidemiological compartments (e.g. the mobility of infectious subjects might be reduced with respect to that of susceptible and removed [127]), $\mathcal{C}$ is the $n*n$ *connection* matrix whose elements are the transmission rates between individuals of different patches.

A *spatially homogeneous* solution of Eq. (93) is one such that in all the patches the state vector is the same $u_j = s(t)$, with $s(t) \in R^m$. Linearizing around the spatially homogeneous solution and applying some similarity transformation one gets the following n independent systems in C^m [137]

$$\psi' = (Df(s(t)) + \lambda_h M) \psi, \quad h = 1, \ldots, n, \tag{94}$$

where λ_h are the eigenvalues of the connection matrix C. Eq. (94) is the discrete equivalent of the linearization equation arising in the study of the Turing bifurcation [126]. Notably, this methodology can also be applied in the case of nonlinear population fluxes [137].

Among multi-patch models, a fairly well-studied class is represented by the following *multi-group, cross-coupled* SEIR model ([137] and refs therein) in a constant population

$$S'_i = \mu(1 - S_i) - S_i \sum_{j=1}^{n} \beta_{ij} I_j, \tag{95}$$

$$E'_i = S_i \sum_{j=1}^{n} \beta_{ij} I_j - (\mu + \alpha) E_i, \tag{96}$$

$$I'_i = \alpha I_i - (\mu + \gamma)I_i, \tag{97}$$
$$R_i = 1 - S_i - E_i - I_i. \tag{98}$$

Unlike Eq. (93) where all contacts are local, through inter-patch movements, in Eq. (95) fluxes between patches are disregarded on the rationale that movements from own patch to other patches are short-lasting. The core of the description stands in the *spatial transmission matrix* β_{ij} assigning the per-capita rate at which a susceptible from patch i is infected by an infective individual belonging to patch j regardless of the patch where the at-risk contact occurred [120,137,139]. For Eq. (95) fairly general results are available, including a decoupling equivalent to those of Eq. (94) [137]. In particular, Eq. (95) was used to study the conditions under which the local (i.e., in each patch) oscillations of SIR and SEIR models become synchronized [137,139], as observed for measles in UK cities in the pre-vaccination era [92]. Notably, synchronization of unforced SIR and SEIR models arises at very low levels of patches connection, while seasonal forcing, besides inducing complicate behavior, allows maintenance of some phase differences in local dynamics [139], and yields more plausible prevalence levels at the inter-epidemic intervals compared to the simple SIR and SEIR models.

Introduction of vaccination can have a complex impact on the spatio-temporal dynamics of multi-patch systems [120]. For example metapopulation models can explain the observed de-correlation observed in the time-series of UK major cities after the introduction of measles vaccination [92,140].

3.9.1. Large scale meta-population networks

In very recent times, the increasing availability of data on both the spread of infectious diseases and human movements is allowing to finely model the dynamics of outbreaks on larger and larger scales, even world-wide [133]. This has led to a number of investigations on the role of the topology of the networks of inter-patches contacts as resulting from population mobility [141]. These analyses have shown that though the condition $\mathcal{R}_0 > 1$ still guarantees that the epidemics is not self-extinguishing and that there are patches here and there in the network where the infection is circulating, in order to have global macroscopic spread a second threshold has to be overcome [141,142]. This second threshold, larger than the first one, depends both on population mobility and on the topology of network. Not surprisingly, if the network is heterogenous, i.e. its degree distribution $P(k)$ (which summarizes the probability for a node of the underlying graph of having k links to other nodes of the networks) has a heavy tail, global epidemic spread is easier. However, for very large heterogeneous network structure the second threshold is close to the classical one, quite independent of population mobility. This can have profound implications as it implies that measures aimed at limiting international mobility to reduce infection spread might have little impact.

3.9.2. Metapopulation models beyond space: social and age-specific heterogeneities

Multi-group and metapopulation models are one of the most active and effective fields of investigation in the theoretical epidemiology of infectious diseases, as they *lato sensu* represent a convenient framework for modeling a much wider class of epidemiological phenomena than geographical spread only. This has been done by straightforwardly extending the interpretation of patches to different types of internally homogeneous groups, such as social groups, age groups, etc., still relying on appropriate contact or transmission matrix, having elements $\beta_{i,j}$ reflecting heterogeneous contact patterns between social or age groups. Though the first works in the field were following a more mathematical vein, i.e. the need to remove the homogeneous mixing hypothesis, the public health modeling revolution described in Section 3.1 is much the legacy of this intuition. Indeed, the multigroup SIS model analogous to the SEIR model of Eqs. (95)–(98) [137], first introduced by Lajmanovich and Yorke [143] to model the spread of an infection among societal groups with different social/sexual behavior, and later extended to model gonorrhea transmission in the US [78], has represented the departure point for the explosion of studies on the modeling of sexually transmitted infections that was subsequently triggered by the onset of the HIV/AIDS epidemics in the early eighties. Noteworthy extensions of Ref. [143] to heterogeneous networks of contacts are available in [144–146].

On the other hand, in the field of common childhood infections, such as measles and pertussis, the large availability of age-specific infection data triggered the development of the "standard" age-structured model [77,79,81]. The "standard" model is nowadays extensively used to investigate the nonlinear effects of mass vaccination programs, including the complicate patterns of increase in the average age at infection following the reduction in transmission, and the related perverse increase in the number of cases of serious disease, as it may be the case for rubella, measles and varicella [77,79,81]. The key ingredient of the "standard" model is represented by the age-specific transmission matrix of elements $\beta_{i,j}$, representing the per-capita transmission rates in contacts between individuals of different age groups. In a first phase these matrices were computed by resorting to simplifying hypotheses about contacts patterns between age groups, aimed to reduce the number of unknowns in the β_{ij} matrix from n^2 to n, thereby making them identifiable from standard age-specific epidemiological data. These approaches include the so-called Who Acquires Infection From Whom (WAIFW) matrices ([79] and refs therein), and the *proportionate* and *preferred* mixing hypotheses ([81] and refs therein). Recently, considerable advances were made thanks to studies aimed to directly observe population contact patterns [147–150], or to reconstruct them from official data [151,152].

Apart the anagraphical age of subjects, two other chrono-biological variables impact on the in-population dynamics of infectious diseases. The first is the so called *age since infection* [153], which is the time elapsed from the infection, a key

factor modulating both the contact rate of infected subjects, the disease duration and the disease-related risk of death. Quite remarkably, this variable was first introduced in mathematical epidemiology by Kermack and McKendrik in their 1927 paper [68], but later ignored for decades [153]. The second is the *age since vaccination* [154,155], which is the time elapsed since the vaccine inoculation, and which modulates both the vaccine-induced immunity [155] and the occurrence of vaccine-related side effects [154]. Note that if any of these temporal variables is modeled as a continuous variable, then the resulting model, due to the impact of these variables on the contact rate and other parameters, is not differential multi-patch but integro-differential, as in the seminal work [68].

Finally, a number of recent re-assessments of the main future challenges for multi-group and metapopulation studies can be found in [156–159].

4. Basic concepts and methods in (non-behavioral) epidemiological modeling and vaccination: network models

A real population is mix of heterogeneous individuals and the heterogeneity exists in several aspects, especially in contacts. The foundations of epidemiology and early epidemiological models were based on population wide random-mixing, which was well described in the last section. However, in practice, each individual has a finite set of contacts or individuals whom they can pass the infection. The ensemble of all such contacts forms a *network*. Structural pattern of this contact network among individuals allows models to compute the epidemic dynamics at the population scale from the individual-level process of spreading infections.

In the recent years, overwhelming data explosion in human sciences shows complex and heterogeneous connectivity patterns in a wide range of biological and social systems. Network theory provides a natural gateway to understand all these complex patterns. Ideas from network theory have also inspired research in many fields to study interactions between elements and their dynamics in areas like computer science, system biology, social sciences, economic and statistical physics. However, epidemiological study has also used potentiality of network theory like any other scientific disciplines. The complex properties of networks have a profound impact on studying equilibrium and non-equilibrium phenomena in epidemic processes. Dated back to mid-eighties, the researchers have shown that understanding the spreading process of an infection have considerable overlap with the studies of network properties [150,151]. The group of individuals and their connections in a population naturally defines a network in which disease spreads from one node (individual) to another distantly related node. So, insights from topological structure of the underlying network provides useful information that helps predicting spreading process including growth of infection at the first stage and subsequently its distribution over the entire contact network.

Although it has long been acknowledged that network theory is the key ingredient of epidemic modeling, recent abundance of data exemplifies the huge complexity in the spreading and persistence of infection in a population and thus calls for detailed theoretical understanding of the interplay between epidemic processes and networks. Previous works have shown that most real-world networks exhibit dynamic self-organization and are statistically heterogeneous—typically all marks of complex systems [160–168]. A regular lattice rarely represents the real-world networks of relevance for epidemic spreading. In a real population, there are few individuals that may act as hubs (or 'super-spreader'), whereas the majority of the population have very few interactions. Both social and infrastructure networks are organized in communities of tightly interconnected nodes of individuals [169]. Although randomness in the connection process of nodes is always present, organizing principles and correlations in the connectivity patterns define network structures that are deeply affecting the evolution and behavior of epidemic and contagion processes [144,170–172]. Furthermore, the complex features of networks often find their signature in statistical distributions which are generally heavy-tailed, skewed, and varying over several orders of magnitude [173]. The evidence of large-scale fluctuations, clustering and communities characterizes the connectivity patterns of real-world systems; and this has prompted the need for mathematical approaches capable to deal with the inherent complexity of networks.

However, the connection between infectious disease epidemiology and the network theory goes far beyond the general discussion. Understanding network topology may help identifying disease transmission routes, which may help designing strategies to control the disease. For instance, 'contact tracing' is a very powerful public health strategy during first phases of the spreading process. It relies on underlying transmission routes to control further spread without even any information on epidemiology of the infectious disease [166]. Thus studying network theory and how it can be applied to relate the infection spreading process may be a vital tool for understanding the disease and, designing control strategies. With a discussion of basic network metrics and topologies, here we review the development of current research at the interface of epidemiological studies and network theory including immunization to control the disease.

4.1. Networks: types and topologies

4.1.1. Definitions and notations

Graph theory [174] can be used to mathematically formalize networks [168]. A graph is a collection of points, called *vertices* (or *nodes*), and a set of connections, called *edges* (or *links*). Edges indicate link between vertices, representing the presence of an interaction or relations between those vertices. Interaction can be bidirectional which defines *undirected* networks, or it can be directional which defines *directed* networks or *digraphs*. We represent a graph of size N (say, with N

vertices, and L links) using $N \times N$ adjacency matrix A, with elements $a_{ij} = 1$ if an edge is connecting nodes i and j and zero otherwise. A is symmetric matrix in undirected graphs, and asymmetric in directed graphs.

- **Degree and Degree distribution:** The *degree* (or *connectivity*) k_i of vertex i in an undirected network is the number of edges emanating from i, i.e. $k_i = \sum_j a_{ij}$. In the case of directed networks, we have both *in-degree*, k^{in}, and *out-degree*, k^{out}, as the number of edges that end in i or start from i, respectively. The total degree, however, in directed network is defined by $k_i = k^{\text{in}} + k^{\text{out}}$. The *degree distribution* $P(k)$ of a network is the probability that a randomly chosen vertex has degree k. In a finite network, it denotes the fraction of vertices with degree exactly equal to k. In case of directed networks, there are instead two different distributions, the *out-degree* $P^{\text{out}}(k^{\text{out}})$ and the *in-degree* $P^{\text{in}}(k^{\text{in}})$ distributions, though in-degree and out-degree of a given vertex might not be independent. It is also useful to consider the *moments* of the degree distribution, $\langle k^n \rangle = \sum_k k^n P(k)$. The first moment $\langle k \rangle = 2L/N$ is called the *average degree*, provides information about the density of the network. The degree distribution defines classes of networks with similar statistical features. A network is called *sparse* if its number of links L grows at most linearly with the network size N; otherwise, it is called *dense*. In directed networks, since every edge contributes to one node in-degree and other node out-degree we have that $\langle k^{\text{in}} \rangle = \langle k^{\text{out}} \rangle$.

- **Connectedness, Shortest path length, and Diameter:** A path $P_{i,n}$ is a sequence of different edges $(i_j, i_{j+1}), j = 0, \ldots, n-1$, connecting vertices i_0 and i_n; the number of edges traversed, n, is also called the length of the path. A graph is *connected* if there exists a path connecting any two vertices in the graph. A component $\mathcal{C}$ of a graph is defined as a connected subgraph. The *shortest path* l_{ij} between two nodes i and j is defined as the length of the shortest path (though may not be unique) joining i and j, and the *average shortest path length* $\langle l \rangle$ is the average of the value of l_{ij} over all pairs of vertices in the network. The *diameter* of a network is the maximum value of all the pairwise shortest path lengths.

- **Degree correlations:** The *degree correlation* between two vertices is defined as the conditional probability $P(k'|k)$ that an edge departing from a vertex of degree k is connected to a vertex of degree k' [144]. A network is called *uncorrelated* if this conditional probability is independent of the originating vertex k. In this case, $P(k'|k)$ can be simply estimated as the ratio between the number of edges pointing to vertices of degree k', $k'P(k')N/2$, and the total number of edges, $L = \langle k \rangle N/2$, to yield $P^{\text{un}}(k'|k) = k'P(k')/\langle k \rangle$. The most popular measure of degree correlation is *Pearson correlation coefficient* r. Proposed by Mark Newman [175,176], the degree correlation coefficient is defined as

$$r = \sum_{jk} \frac{jk(e_{jk} - q_j q_k)}{\sigma^2}, \tag{99}$$

where

$$\sigma^2 = \sum_k k^2 q_k - \left[\sum_k k q_k \right]^2.$$

e_{ij} is the degree correlation matrix that represents the probability of finding node with degrees i and j at the two ends of a randomly selected link, and q_k is the probability that there is a degree-k node at the end of the randomly selected link. The degree correlation varies between $-1 \leq r \leq 1$. For instance, $r > 0$ the network is assortative, for $r = 0$ the network is neutral and for $r < 0$ the network is disassortative. In reality, social networks are assortative, whereas technological and biological networks are disassortative [176]. For example, for the scientific collaboration network we obtain $r = 0.13$, in line with its assortative nature; for the protein interaction network $r = -0.04$, supporting its disassortative nature and for the power grid we have $r = 0$. Another related measure of correlations is the *average degree* of the nearest neighbors of vertices of degree k, denoted by $\bar{k}_{nn}(k)$ which is formally defined as [144]

$$\bar{k}_{nn}(k) = \sum_{k'} k' P(k'|k).$$

- **Centrality and Betweenness:** The concept of *centrality* encodes the relative importance of a node inside a network. Together with the *degree* and the *closeness of a node* (defined as the inverse of the average distance from all other nodes), another standard measure of node centrality is *node betweenness*, which is obtained by counting the number of geodesics going through it. More precisely, the betweenness b_i of a node i, sometimes referred to as *load*, is defined as

$$b_i = \sum_{j,k \in N, j \neq k} \frac{n_{jk}(i)}{n_{jk}}, \tag{100}$$

where n_{jk} is the number of shortest paths connecting j and k, while $n_{jk}(i)$ is the number of shortest paths connecting j and k and passing through i. Betweenness distributions have been investigated in [177–184]. Betweenness–betweenness correlations and betweenness–degree correlations have been studied respectively in [185] and in [186]. The concept of betweenness can also be extended to edges. The *edge betweenness* is defined as the number of shortest paths between pairs of nodes that run through that edge [187].

- **Clustering:** *Clustering* refers to transitivity property of the network, i.e. the relative propensity of two nodes to be connected, provided that they share a common neighbor. The *clustering coefficient* C is defined as the ratio between the number of loops of length three in the network (i.e. triangles), and the number of connected triples (three nodes connected

by two edges). An alternative definition is also available by Watts and Strogatz [188]. A local measure c_i of clustering is defined as the ratio between the actual number of edges among the neighbors of a vertex i, e_i, and its maximum possible value, measuring thus directly the probability that two neighbors of vertex i are also neighbors of each other. The *mean clustering* of the network $\langle c \rangle$ is defined as the average of c_i over all vertices in the network. The clustering spectrum $\bar{c}(k)$ is defined as the average clustering coefficient of the vertices of degree k [185,186], satisfying $\langle c \rangle = \sum_k P(k)\bar{c}(k)$.

● **Community Structure:** A *community* or *cluster* (or *cohesive subgroup*) of a given network of vertices N and set of edges L is a subgraph $G'(N', L')$, whose nodes are tightly connected, i.e. cohesive. Since the structural cohesion of the nodes of G' can be quantified in several different ways, there are different formal definitions of community structures.

A related terminology of community structure of a network is *clique*. A *clique* is a maximal complete subgraph of three or more nodes all of which are adjacent to each other, and such that no other nodes exist adjacent to all of them. This definition can be extended by weakening the requirement of adjacency into a requirement of reachability: a *n-clique* is a maximal subgraph in which the largest geodesic distance between any two nodes is no greater than n. A detail description is given in [162,189].

4.1.2. Type of networks

Based on the qualitative and quantitative features of the degree distribution, the networks can be classified into a series of important types of networks, which are also useful for the epidemiological applications.

● **Random graph:** The simplest network model is the classical random graph model proposed by Erdös and Rényi (also named ER graph) [190,190–192]. A random graph model $G_p(N)$ is constructed from a set of N nodes in which each one of the $N(N-1)/2$ possible links is present with probability p. The degree distribution is given by a binomial form, which, in the limit of constant average degree (i.e. $p = \langle k \rangle / N$) and large N can be approximated by a Poisson distribution $P(k) = e^{-\langle k \rangle}\langle k \rangle^k / k!$. The clustering coefficient is given by $\langle c \rangle = p$, and the average shortest path length is $\langle l \rangle \sim \log N / \log \langle k \rangle$ [193]. This model is therefore adequate in the case of networks governed only by stochasticity, although $G_p(N)$ tends to a regular graph for large N and constant p. Extensions of the basic model to allow for other degree distributions lead to the class of models known as *generalized random graphs*, *random graphs with arbitrary degree distributions* and the *configuration model* [194].

● **Small-world (SW) network:** A more sophisticated and tractable model of a network with small proximity and high clustering coefficient is the *small-world network model* proposed by Watts and Strogatz (i.e. WS model) [173]. This network model is based on rewiring procedure of an ordered lattice. The initial configuration starts with a ring of N vertices, each one of which symmetrically connected to its $2m$ nearest neighbors. This represents the model with a structure which has large clustering coefficient and large average shortest path length. Starting from it, a fraction p of edges in the network are rewired, by visiting all m clock-wise edges of each vertex and reconnecting them, with probability p, to a randomly chosen node. There is another version of the model proposed by Monasson [195], where a fraction p of edges are added between randomly chosen pairs of vertices. The overall effect of the rewiring processes is to add long-range shortcuts, that, even for a small value of $p \sim N - 1$, greatly reduce the average shortest path length, while preserving a large clustering for not very large values of p. This model, although better suited for social networks with high clustering coefficient, has a degree distribution and centrality measures decaying exponentially fast away from the average value. The small-world model thus generates homogeneous networks where the average of each metric is a typical value shared, with little variations, by all nodes of the network.

● **Scale-free (SF) network—static and dynamic:** In contrast to random graph, and small-world network models, real-world networks are structured in a hierarchy of nodes with a few nodes having very large connectivity (the hub nodes), while the vast majority of nodes have much smaller degrees. In precise, these heterogeneous networks exhibit heavy-tailed degree distributions often approximated by a power-law behavior of the form $P(k) \sim k^{-\zeta}$, which implies a non-negligible probability of finding vertices with very large degree. The degree exponent ζ of many real-world networks takes a value between 2 and 3. In such cases, these network models are called *scale-free*. Several variations of the classical random graph models have been proposed in order to generate networks with a power-law degree distribution. For example, *configuration model* [177,196,197], *fitness model* by Caldarelli et al. [198], *threshold graph model* by Masuda et al. [199] and *gradient network models* by Toroczkai and Bassler [200], where directed graphs induced by local gradients of a scalar field distributed on the nodes.

So far, we have not discussed evolution process of these scale-free networks, i.e. nodes and links do not change over time. There are many examples of real networks in which the structural changes are ruled by the dynamical evolution of the system. Most well-known and useful evolving network model is *Barabási–Albert model* [201]. The Barabási–Albert (BA) network is a model of network growth inspired to the formation of the World Wide Web and is based on two basic ingredients: (i) *growth* (of nodes) and (ii) *preferential attachment* (connection of vertices). A description of the growth process and attachment function is discussed in [202]. Variations in preferential attachment function produced several other evolving network models like *Dorogovtsev–Mendes–Samukhin* (DMS) model (linear preferential attachment), *Krapivsky* model (nonlinear attachment probability), etc. [203]. The Barabási–Albert (BA) model has been generalized further by the same authors, and also by others including certain characteristics of underlying growth process of nodes [202].

● **Weighted networks:** In the real world settings, many networks exhibit varied intensity of connections. For example, social network, where existence of strong and weak ties between individuals are seen [204–207], metabolic reaction
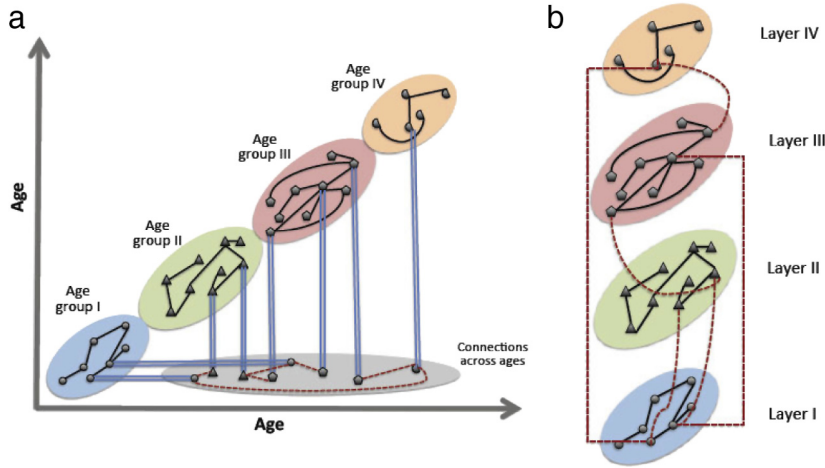
Fig. 8. Schematic illustration of age-dependent contact networks into a multilayer system. Network with different color shades in (a) indicates network of individuals in the same age group, where network in gray region denotes connections across different age groups. Blue lines point to the same individuals from colored regions to gray region. (b) The same network can be conceived as multilayer networks. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Definitions, measure and expressions of network properties of single-layer and multilayer networks.

Definition	Single-layer networks	Multilayer networks
Notation	$\mathcal{G} = (\mathcal{N}, \mathcal{L})$	$\mathcal{M} = (\mathcal{G}, \mathcal{C})$ layers: $\mathcal{G} = \{\mathcal{G}_\alpha : \alpha \in \{1, \ldots, M\}\}, \mathcal{G}_\alpha = (\mathcal{X}_\alpha, \mathcal{E}_\alpha)$ links between layers: $\mathcal{C} = \{\mathcal{E}_{\alpha\varphi} \subset \mathcal{X}_\alpha \times \mathcal{X}_\varphi ; \alpha \in \{1, \ldots, M\}, \alpha \neq \varphi\}$[a]
Degree & Centrality	degree: $k_i = \sum_{j\in\mathcal{N}} a_{ij}$ betweenness: $b_i = \sum_{j,k\in N, j\neq k} \frac{n_{jk}(i)}{n_{jk}}$ degree correlation coefficient $c_i = \frac{1}{\lambda_l} \sum_{i\in\mathcal{N}} a_{ij} c_j$	degree: $\mathbf{k_i} = (k_i^{[1]}, \ldots, k_i^{[M]}), k_i^{[\alpha]} = \sum_j a_{ij}^\alpha$ overlapping degree: $o_i = \sum_{\alpha=1}^M k_i^{[\alpha]}$ degree correlation coefficient $c_i = (c_i^{[1]}, \ldots, c_i^{[M]})$
Clustering coefficients	$c_i = \frac{2e_i}{k_i(k_i-1)}$	$C_\mathcal{M}(i) = \frac{2\sum_{\alpha=1}^M |\bar{E}_\alpha(i)|}{\sum_{\alpha=1}^M |\mathcal{N}_\alpha(i)|(|\mathcal{N}_\alpha(i)|-1)}$
Characteristic path length	$\frac{1}{N(N-1)} \sum_{i,j\in\mathcal{N}, i\neq j} d_{ij}$	$\frac{1}{N(N-1)} \sum_{u,v\in\mathcal{X}_\mathcal{M}, u\neq v} d_{uv}$

[a] α, φ denote layers in multilayer networks.

pathways, predator–prey interactions in food webs [208,209], different capabilities of transmitting electric signals in neural networks [210], unequal traffic on the Internet [211] or of the passengers in airline networks [212]. These types of complex systems can be explained with model networks in which each link carries a numerical value measuring the strength of the connection. Considering weights of the connections puts another dimension of complexity in the topology of weighted network. Similarly, we can also have networks with weighted nodes. Detail descriptions of the measuring and modeling weighted networks has been discussed in [162].

There are also other type of networks, which are important to study real world phenomenon. These networks include *spatial network* (whose nodes occupy a precise position in two or three-dimensional Euclidean space, and whose edges are real physical connections) [213–216], adaptive networks [217], temporal or time-varying networks [218–220] (for more details, please also see Section 11).

All the previous networks are single-layer. But there are many real world complex systems that can be conceived as multiple layers of such single networks. Most complex systems from physical, social, engineering, information and biological sciences include multiple subsystems and layers of connectivity, and they are often open, value-laden, directed, multilevel, multicomponent, reconfigurable systems of systems, and placed within unstable and changing environments. An example of multilayer networks conceived from age-dependent contact network is shown in Fig. 8. The recent development of *multilayer networks* in network theory provides radical views of understanding these multilayer systems. For more information about single-layer and multilayer networks, please see the comparison in Table 1. Descriptions and examples of more different type spatial and multilayer networks are given in [202].

4.2. Epidemic spread over networks

Most infectious diseases such as influenza, SARS, HIV spread through human populations by physical contact between infective and susceptible individuals. The pattern of these disease-causing contacts forms a network, which has substantial

influence on success of the spreading process. The traditional epidemic models such as classic **SIS** or **SIR** epidemic models do not explicitly incorporate the topological structure of the underlying contract network that facilitates the spread of a pathogen. Instead they assume that any individual can come into contact with any other individual (homogenous mixing hypothesis) and so all individuals have comparable number of contacts. It also assumes that probability of disease transmission is same across all individuals in the network. Both of these assumptions are limited in most of the diseases: in reality an individual can transmit a pathogen only to those it comes into contact with and also transmission occurs in varied probabilities. Hence pathogens spread can be assumed on complex contact networks.

Spreading epidemics on complex networks has been studied by many researchers including May and Lloyd [221]. Most detail and systematical study is done by Newman [222]. Considering an SIR model with uniform recovery rate γ, i.e. where infected nodes become removed at rate γ after infection, and infection rate β_I. Then the transmissibility T is defined as the probability that the infection will be transmitted from an infected node to a connected susceptible neighbor before recovery takes place. For continuous-time dynamics, the transmissibility computed as [222]:

$$T = 1 - \lim_{\delta t \to 0} (1 - \beta_I \delta t)^{\gamma/\delta t} = 1 - e^{-\gamma \beta_I}. \tag{101}$$

In general, β_I and γ will vary between individuals, and assume initially that these two quantities are independent and identically distributed (iid) random variables chosen from some appropriate distributions $P(\beta_I)$ and $P(\gamma)$. Hence, transmission of the disease between two individuals is simply the average T (i.e. homogeneous case) of T_{ij} over the distributions $P(\beta_I)$ and $P(\gamma)$, which is

$$T = \langle T_{ij} \rangle = 1 - \int_0^\infty P(\beta_I) P(\gamma) e^{-\gamma \beta_I} d\beta_I d\gamma. \tag{102}$$

Computing threshold transmissibility for spreading the disease in a special non-degenerate case of SIR model has been discussed in [223].

Now, starting from a single infective individual, disease spreads across the contact networks and a outbreak occurs. The ultimate size of the outbreak would be precisely the size of the cluster of vertices that can be reached from the initial vertex by traversing only occupied edges (i.e., edge in the graph across which the disease is transmitted, say with average probability T). This is precisely a similar model equivalent to a bond percolation model with bond occupation probability T on the graph representing the community. The connection between the spread of disease and percolation was in fact one of the original motivations for the percolation model itself. Below in the next subsection, we discuss how disease spread can be realized as the percolation problem on random graphs with arbitrary degree distributions discussing size of outbreaks, presence of an epidemic, transmissibility thresholds, etc.

4.2.1. Percolation theory and spreading process

Like mean-field method (mentioned in Section 3), which is a useful tool to analyze the transmission dynamics in homogeneous mixing population, percolation theory also plays a significant role in prediction of epidemic or vaccination dynamics, such as emerging threshold and phase transition, in network population. One convenient universal approach to deal percolation problem with bond occupation probability T on complex networks is considering *generating functions* [222]. Generating functions and the problem of percolation in networks have been discussed in many earlier studies [177,224–226]. We define generating functions $G_0(u)$ and $G_1(u)$ to describe a network as follows:

$$G_0(u) = \sum_{k=0}^{\infty} P(k) u^k, \tag{103}$$

$$G_1(u) = \frac{G_0'(u)}{G_0'(1)}, \tag{104}$$

where $P(k)$ is the degree distribution of random graph G with N vertices. We can calculate back the degree distribution $P(k)$ and also all the higher moments using $G_0(u)$ with

$$\langle k^n \rangle = \sum_k k^n P(k) = \left[\left(u \frac{d}{du} \right)^n G_0(u) \right]_{u=1}. \tag{105}$$

G_1 calculates the distribution of the degrees of vertices reached by following a randomly chosen edge. A modified set of generating functions is required to solve percolation problem with bond probability T: $G_0(u, T)$ and $G_1(u, T)$, where $G_i(u, T) = G_i(1 + (u - 1)T)$ [227,228]. $G_i(u, T)$, $i = 0, 1$ represents the distributions of the number of *occupied* edges attached to a vertex, as a function of the transmissibility T.

Outbreak size of a disease on a network precisely depends on the distribution of sizes of clusters of vertices together by occupied edges in the corresponding percolation model. Using the generation functions $G_0(u, T)$ and $G_1(u, T)$, we can find the mean outbreak size $\langle s \rangle$ in closed form as

$$\langle s \rangle = 1 + \frac{G_0'(1; T)}{1 - G_1'(1; T)}. \tag{106}$$

Table 2
Epidemic thresholds for **SI**, **SIS** and **SIR** models over Erdös and Rényi (ER) and Scale-Free (SF) networks.

Model type	Model expression	Characteristic time (τ_c)	Spreading rate (λ_c)
SI	$\frac{di_k}{dt} = \beta_I[1 - i_k(t)]k\theta_k(t)$	SF ($\zeta \geq 3$): $\frac{\langle k \rangle}{\beta_I(\langle k^2 \rangle - \langle k \rangle)}$ SF ($\zeta < 3$): 0 ER: $\frac{1}{\beta_I \langle k \rangle}$	SF ($\zeta > 3$): 0 SF ($\zeta \leq 3$): 0
SIS	$\frac{di_k}{dt} = \beta_I[1 - i_k(t)]k\theta_k(t) - \gamma i_k(t)$	SF: $\frac{\langle k \rangle}{(\beta_I \langle k^2 \rangle - \gamma \langle k \rangle)}$ ER: $\frac{1}{\beta_I(\langle k \rangle + 1) - \gamma}$	$\frac{\langle k \rangle}{\langle k^2 \rangle}$ $\frac{1}{\langle k \rangle + 1}$
SIR	$\frac{di_k}{dt} = \beta_I s_k(t)\theta_k(t) - \gamma i_k(t)$ $s_k(t) = 1 - i_k(t) - r_k(t)$	$\frac{\langle k \rangle}{(\beta_I \langle k^2 \rangle - (\gamma + \beta_I)\langle k \rangle)}$	$\frac{\langle k \rangle}{\langle k^2 \rangle - \langle k \rangle}$

i_k: infected nodes with degree k.
s_k: susceptible nodes with degree k.
r_k: recovered or immune nodes with degree k.
θ_k: fraction of the neighbors of node k that are infected.
β_I: transmission rate.
γ: recovery rate.
ζ: degree exponent SF network.

So, given the values of generating function G_i, we can completely evaluate this expression to get the mean outbreak size for any value of T and degree distribution. The outbreak size critically depends on the relation $TG'_1(1, T) = 1$. This point indicates the onset of an epidemic; it is the point at which the typical outbreak ceases to be confined to a finite number of individuals, and expands to fill significant fraction of the graph. This defines the critical transmissibility T_c, at which the critical transition occurs

$$T_c = \frac{1}{G'_1(1)} = \frac{\sum_k kP(k)}{\sum_k k(k-1)P(k)} = \frac{\langle k \rangle}{\langle k^2 \rangle - \langle k \rangle}. \tag{107}$$

For $T > T_c$, we have an epidemic, or "giant component" in the language of percolation. This equation is valid for degree of uncorrelated networks which have no loops in the clusters. However, in the case of heterogeneous networks with degree distribution $P(k) \sim k^{-\zeta}$, it was found that the percolation threshold tends to zero for $\zeta < 3$ in the limit of an infinite network size $N \to \infty$ [225]. That is, an epidemic can spread in this network regardless of how small the infection probability and how quick is the recovery process. In general, the expansion of the generating functions around the nonzero solution yields the scaling behavior of the order parameter in the vicinity of the critical point: $P_G(T) \sim (T - T_c)^{\beta_{I_c}}$, where the critical exponent β_{I_c} assumes following values

$$\beta_{I_c} = \begin{cases} 1/(3 - \zeta) & \text{for} \quad \zeta < 3, \\ 1/(\zeta - 3) & \text{for} \quad 3 < \zeta \leq 4, \\ 1 & \text{for} \quad \zeta \geq 4. \end{cases} \tag{108}$$

For the case $\zeta = 3$, a stretched exponential form $P_G(T) \sim e^{1/p}$ is expected. For more details about disease thresholds, we provide a brief summary in Table 2.

There are other approaches of modeling spreading process of infectious disease over networks such as *Degree-based mean-field theory* (DBMF) and *Individual-based mean-field theory* (IBMF). DBMF considers the dynamics of node of degree k to be infected or remain susceptible, and assume that all the nodes with degree k are statistically equivalent. The value of the epidemic threshold can be obtained by means of a linear stability analysis [144,223]. It was the first approach to the study of the SIS model in complex networks [170]. Global stability of the disease-free state in such model, the role of seasonal oscillations and the coexistence of an endemic equilibrium state via multistability have been investigated in [146]. In contrast, the state of the system in the disease model (such as SIS) in the IBMF framework is fully defined by a set of Bernoulli random variables $X_i(t) \in 0, 1 : X_i(t) = 0$ for a healthy, susceptible node and $X_i(t) = 1$ for an infected node; altogether it defines a 2^N Markov chain for a network with N nodes [227,228], which specifies exactly the time evolution of the disease model. However, the Markov chain approach complicates analytical calculations of the model. A discussion on both DBMF and IBMF can be found in [229].

4.2.2. Disease spreading in multilayer networks

In many situations, disease spread can be modeled over *interconnected networks* instead of a single-layer network (or called *monoplex* network). For example, spreading of an infection amongst groups of communities, or even groups of individuals of different age profiles Fig. 8. Studying the diffusion of pathogens over interconnected network or *multilayer networks* has got very recent attention in epidemiology. Multilayer networks are made of multiple layers, where each layer is a monoplex network, where the same nodes can appear in multiple layers and nodes on different layers can be connected to each other. One important aspect of spreading infection across multilayer networks is that it can also spread

from one layer to another. Generally, there are three possibilities for spreading process [230,231]: *same-node, inter-layer*, when infection switches layer but remains on the same node, e.g., when an infected individual travels from one city to another city; *other-node inter-layer*, when infection continues spreading to another node in another layer, e.g., spreading of infection between individuals in different age profiles through direct physical contact. In third type, *other-node intra-layer*, when infection spreads across the same community. In the context of interacting spreading processes in multilayer networks, two types of thresholds have recently been introduced, called *survival threshold* measuring if infection will survive and *absolute-dominance threshold* indicating whether it can completely remove another competing process [230].

Most of the works on modeling the dynamics of diffusion over multilayer networks have used epidemic models such as SIS, SIR [202,231–240]. In a similar approach with single-layer networks, the dynamics of epidemic spreading according to the SIS and SIR models over multilayer networks are described as a two- or three-state process, respectively. Thereafter, many extensions have been applied to SIR and SIS models. One of the most important extensions, Goldenberg et al. [241] proposed a discrete-time version of the SIR model called Independent Cascade Model (ICM), where time proceeds in discrete time steps. In this model, each infected node u at time t can infect each of its neighbors. If the infection succeeds, then neighbor v will become infected at step $t + 1$. ICM is often used in the literature on influence spreading. In [230], the authors extended this model to analyze the dynamics of multiple cascades over multiplex networks.

Unlike a monoplex network, the infection may diffuse over inter- and intra-layer connections at different speeds over multiple networks, meaning that we may have different infection rates (i.e., transmissibilities) across the links of each layer and also the links between the layers. Therefore, most of the works on spreading processes over multilayer networks [233,238,239,242–247] have extended epidemic-like models by considering different infection rates dependent on the types of the layers. A recent contribution in this context proposed a generalized epidemic mean-field (GEMF) framework, which was capable of modeling epidemic-like spreading processes with more complex states in multiplex network layers (compared to two or three states in the SIS and SIR models).

4.2.3. Simulation-based models

Simulation-based modeling, like analytical models described above, has similar appeal in epidemic theory [173,248–254]. There are many good examples of simulation based network models of infectious diseases such as sexually transmitted diseases (STD) that highlight the importance of network structure (e.g., role of core groups or interconnected individuals with a large number of contacts) in disease transmission and persistence [255–263].

There are several simulation-based works that attempt to model airborne disease outbreaks [252,264,265]. Such models consider large population, like of order $\sim 10^6$ with certain specified social connections among individuals. The networks used are generated by computer simulation to conform to several observed demographic and social characteristics such as given age distribution, household size, public crowding. Variable activities in the entire day are also assigned to the individuals such as working place, public gathering, or even spending time at home. Different activities reflect the change in contact pattern in different time of the day. Children are assigned to school, day-care center or play group where they interact (and therefore form connections) with other children. The simulation model puts emphasis particularly on transmission within households and family groups, which are probably the main routes of transmission for this disease. There are models that attempt similar tasks, using census data to determine interaction patterns estimating demographic and disease parameters [249]. Despite the number of approximations involved, the inherent stochasticity of such microsimulations allows a direct estimation of the variability between epidemics. For example, Meyers et al. [250] generates a network model to characterize the spread of SARS in Vancouver, British Columbia. They have used census data of Vancouver and using computer simulations generated a plausible contact network that reflects an urban setting. They had chosen $N = 1000$ households at random from the Vancouver household size distribution, which yields approximately 2600 people. Household members were given ages according to the census data of Vancouver. Kids were then assigned to schools according to school and class size distributions to occupations according to (un)employment data, to hospitals as patients and caregivers according to hospital employment and bed data, and to other public places. Within each location then random connections were created among individuals with variable probabilities at households, hospitals and schools, workplaces, and other public places. They have also chosen the parameters of the Poisson and power law networks so that all three networks share the same epidemic thresholds. Conventional epidemiological models shows clear threshold effects, that is, the disease is expected to die out when transmission probability falls below a critical threshold level. Olinky and Stone [266] have shown that epidemic propagation depends equally on the infection process and also on the network topology. Their recent analyses of infection processes on highly heterogeneous networks (e.g., scale-free networks) concludes that diseases spread and persist even for vanishingly small transmission probabilities.

However, network-based simulation models have certain pitfalls as well. As mentioned by Keeling and Eames [267], these models are limited to ascertain the sensitivity of the epidemiological results to the details of the network structure. For example, the study by Halloran et al. [264] may be limited to answer questions like whether the network is representative of an average American community, whether variation between communities will bias the results if large population sizes are considered and whether rare but epidemiologically important contact structures are missing from the network. It is difficult to answer such questions or gain an intuitive understanding of network structure if our experience is limited to simulations of sampled networks. However, a range of idealized networks and analytical tools have been developed that can reveal the elements of network structure, which are important determinants of epidemic dynamics.

4.3. Vaccinations over networks

Designing optimal vaccination strategies to prevent and control infection has been widely addressed in the epidemic modeling of infectious diseases. Vaccination over networks can be modeled as a site percolation problem [268]. Each vaccinated node can be regarded as a site which is removed from the network. The goal of the vaccination process thus is to reduce the transmissibility and pass the percolation threshold, which leads to minimization of the number of infected individuals in the network. Thus implementing the SIR model and vaccination can be considered as a site-bond percolation model, and the vaccination is considered successful if the network is below the percolation threshold [268,269]. However though, the epidemic threshold decreases with the standard deviation of the connectivity distribution in the heterogeneous networks [144,170,270]. This feature is paradoxically amplified in SF networks that have diverging connectivity fluctuations. In fact, epidemic processes in SF networks do not possess, in the limit of an infinite network, an epidemic threshold below which diseases cannot produce a major epidemic outbreak or the inset of an endemic state. SF networks are, therefore, prone to the spreading and the persistence of infections, whatever virulence the infective agent might possess. In view of this, it becomes a major task to public health policy makers designing optimal vaccination strategies oriented to minimize the risk of epidemic outbreaks on networks such as SF networks. However, population with heterogeneous connected networks may take different paths to extinction and thus control of infection. A recent paper by Hindes and Schwartz (2016) discussed such issues on heterogeneous random networks under various configuration [271].

The simplest hypothetical way to experiment the effect of vaccination over networks is the random introduction of immune nodes in the network population, in order to get a *uniform vaccination* density [270,272]. Immune nodes cannot become infected and, thus, do not transmit the infection to their neighbors. In this case, for a fixed spreading rate β_I, the relevant control parameter is the immunity x, defined as the fraction of vaccinated nodes present in the network. At the mean-field level, the presence of uniform vaccination will effectively reduce the spreading rate β_I by a factor $(1 - x)$; i.e. the probability of finding and infecting a susceptible and non-immune node. In homogeneous networks, such as the Watts–Strogatz (WS) model, it is easy to show that in the case of a constant β_I, the critical vaccination rate x_c is given by

$$x_c = 1 - \frac{\beta_{I_c}}{\beta_I}. \tag{109}$$

Thus, the critical vaccination that achieves eradication is related to the spreading rate and the epidemic threshold of the infection, which implies that the critical vaccination allows the complete eradication of the disease over the network.

However, in an heterogeneous population with a uniform vaccination scheme, it is necessary to vaccinate a fraction of the population larger than the estimate given by a simple (homogeneous) assumption [270]. In this case, it can be proved that optimal vaccination programs can eradicate the disease by vaccinating a smaller number of individuals. A straightforward way to reintroduce an intrinsic vaccination threshold in heterogeneous networks consists in using different fractions of vaccinated individuals according to their connectivity. Let us define x_k as the fraction of immune individuals with a given connectivity k and averaging x_k over the various connectivity classes, we have the critical vaccination threshold as

$$x_c = \sum_{k > \beta_I^{-1}} \left(1 - \frac{1}{\beta_I k} \right) P(k). \tag{110}$$

Scale-free networks can be considered as a limiting case of heterogeneous systems and it is natural to look for specifically devised vaccination strategies. This however, does not work on SF network. Because of the absence of any epidemic threshold in SF networks, it becomes impossible to find any critical vaccination rate that ensures the eradication. In scale-free (SF) networks, the previous equation takes the form

$$1 - x_c = \frac{1}{\beta_I} \frac{\langle k \rangle}{\langle k^2 \rangle}. \tag{111}$$

This shows that only a complete vaccination ($x_c = 1$) of the network works to eradicate the disease over scale free network with $\langle k^2 \rangle \to \infty$.

While uniform or proportional vaccination fails to protect contagion over scale-free networks due to its peculiar nature, it has been seen that scale-free networks are strongly affected by selective damage. For example, if a few of the most connected nodes are removed, the network suffers a dramatic reduction of its ability to carry information [224,225,268]. With this idea, one can design a *targeted vaccination* scheme to progressively make immune the most highly connected nodes, i.e., the ones more likely to spread the disease. In scale-free networks, it produces an significant increase of the network tolerance to infections at the price of a tiny fraction of immune individuals [269,273,274].

Suppose, a fraction x of all nodes with connectivity $k > k_c$ are vaccinated in a network. Then $x = \sum_{k > k_c} P(k)$. At the same time, this also implies that all links emanating from vaccinated individuals can be considered as if they were removed. So, if we assume that the fraction of links is effectively removed from the network, the new connectivity distribution after the vaccination of a fraction x of the most connected node is

$$P_x(k) = \sum_{q > k}^{k_c} P(q) \binom{q}{k} (1 - p)^k p^{q-k}. \tag{112}$$

This gives, calculating the first two moments, the critical fraction of vaccination needed to eradicate the infection given by the relation:

$$\frac{\langle k^2 \rangle_{x_c}}{\langle k \rangle_{x_c}} \equiv \frac{\langle k^2 \rangle_c}{\langle k \rangle_c}[1 - p(x_c)] + p(x_c) = \beta_I^{-1}, \tag{113}$$

where $\langle k \rangle_c = \sum_{k_{min}}^{k_c} kP(k)$, $\langle k^2 \rangle_c = \sum_{k_{min}}^{k_c} k^2 P(k)$. An explicit calculation for Barabási–Albert network model, the approximate solution for the vaccination threshold is given by

$$x_c \sim \exp(-2/k_{min}\beta_I). \tag{114}$$

This clearly indicates that the targeted vaccination program is extremely convenient in scale-free networks where the critical vaccination is exponentially small in a wide range of spreading rates β_I. Also in this case, the present result can be generalized for scale-free networks with arbitrary connectivity exponent ζ.

A comparative analysis of targeted vaccination over SW and BA models exhibit contrastive results [274]. In the case of the SW networks, the behavior of the prevalence as a function of x is equivalent in the uniform and targeted vaccination procedures. On the contrary, in the case of the BA networks, we observe a drastic variation in the prevalence behavior, though it is sensitive to random choice of immune nodes. This however, confirms that targeted strategies do not have a particular efficiency in systems with limited heterogeneity, but are highly sensitive to the network with higher degree of heterogeneity such as scale-free networks. Targeted vaccination of a small fraction of the most connected nodes in scale-free networks shows significant impact on spread process of infection.

Another target-based strategy is *acquaintance vaccination* that tries to target all of the most highly connected nodes for vaccination [269,274–277], but this strategy requires no knowledge of the node degrees or any other global knowledge. In this approach a random group of nodes are chosen and then a random set of their neighbors are selected for vaccination. The most highly connected nodes are far more likely to be in this group of neighbors. So immunizing this group results in targeting the most highly connected nodes, and it requires far less information about the network [277–279]. Another variant of this strategy again calls for the random selection of nodes but instead asks for one of their neighbors with a higher degree, or at least more than a given threshold degree and immunizes them [275]. These degree based strategies consistently require fewer nodes to be vaccinated and as such improve a network's chances against epidemic attacks [269,277].

This however, is just a brief summary about vaccination on networks. For more details about mathematical and physics framework, please see next section.

5. Non-behavioral epidemiological vaccination on networks

Due to the potential threat posed by the spreading of infectious disease, ample research has been devoted to the mitigation and prevention of epidemics. To date, one of the most popular and effective methods is network vaccination [229,274,280–284], where certain nodes in network are effectively immunized by a perfect vaccine, and are thus no longer able to transmit the disease to their neighbors [285]. More precisely, given a fixed effective spreading rate β_I/γ (β_I and γ are the infection rate and recovery rate respectively. Without loss of generality, $\gamma = 1$ is fixed since it only affects the definition of the time scale of epidemic spreading.), the relevant control parameter is the immunity rate x, which is simply defined as the fraction of vaccinated nodes in a network. The challenge is how to arrive at a minimal value of x, which is the so-called vaccination threshold x_c, to reduce the spreading rate of the disease β_I under the critical threshold β_{I_c}, such that the disease dies out completely [286–288].

This section is devoted to the systematic review of such non-behavioral vaccination techniques, where networks are effectively immunized by means of a targeted vaccination of certain nodes. We will review research done on single-layer and multilayer networks, on static and adaptive networks, and theoretical as well as empirical results.

5.1. Uniform (or random) vaccination

The simplest vaccination strategy that one can conceive, the implementation of which requires no preparation or information at all, is uniform or random vaccination [270,274]. Here a fraction x of nodes are randomly selected and then vaccinated. Accordingly, only the remaining $1 - x$ fraction of nodes contributes to the spreading of the disease. The effective number of neighbors that each susceptible node possesses thus decreases from its degree k to $k(1 - x)$. If one considers the standard SIS model as an example, then after vaccination the evolution equation of the probability $\rho_k^I(t)$ that a node with degree k is infected at time t is

$$\frac{d\rho_k^I(t)}{dt} = -\rho_k^I(t) + \beta_I k(1 - x)[1 - \rho_k^I(t)] \sum_{k'} P(k'|k)\rho_{k'}^I(t), \tag{115}$$

where the first term defines the rate at which infected nodes of degree k recover and become susceptible again, and in the second term $P(k'|k)$ refers to the conditional probability that an infected node with degree k' points to the node with degree k.

It can be observed that in the mean-field model, where individuals interact randomly with one another, the uniform vaccination is equivalent to the simple decrease of the spreading rate of the disease from β_I to $\beta_I(1 - x)$. Since eradicating an epidemic by uniform vaccination requires β_I to be below its critical epidemic threshold β_{I_c}, we obtain

$$\beta_I(1 - x_c) = \beta_{I_c}, \tag{116}$$

where x_c is the critical vaccination threshold, above which the density of infected individuals in the stationary state is zero.

However, in predominantly homogeneous networks, such as random or small-world networks, $\beta_{I_c} = 1/(\langle k \rangle + 1)$ for the SIS model and $\beta_{I_c} = 1/\langle k \rangle$ for the SIR model [144,280]. Based on Eq. (116), the vaccination threshold is

$$x_c = 1 - \frac{1}{\beta_I} \frac{1}{\langle k \rangle + 1} \tag{117}$$

for the SIS model, and

$$x_c = 1 - \frac{1}{\beta_I} \frac{1}{\langle k \rangle} \tag{118}$$

for the SIR model. Thus, if the uniform vaccination level is larger than x_c, homogeneous networks will be completely protected and no large epidemic outbreaks will be possible.

Conversely, the situation is much different for strongly heterogeneous networks, such as scale-free networks, where the epidemic threshold vanishes [225,289]. Here uniform vaccination will fail. In particular, the absence of an epidemic threshold ($\beta_{I_c} = 0$) in the thermodynamic limit implies that no matter how one rescales $\beta_I \to \beta_I(1 - x)$, the epidemic will not be stopped unless $x = 1$.

Based on the fact that in heterogenous networks $\beta_{I_c} = \langle k \rangle / \langle k^2 \rangle$ for the SIS model and $\beta_{I_c} = \langle k \rangle / (\langle k^2 \rangle - \langle k \rangle)$ for the SIR model [144,171,290], and by using Eq. (116), the threshold of uniform vaccination is given by

$$x_c = 1 - \frac{1}{\beta_I} \frac{\langle k \rangle}{\langle k^2 \rangle} \tag{119}$$

for the SIS model and

$$x_c = 1 - \frac{1}{\beta_I} \frac{\langle k \rangle}{\langle k^2 \rangle - \langle k \rangle} \tag{120}$$

for the SIR model.

In a scale-free network with degree exponent $\zeta < 3$, we have $\langle k^2 \rangle \to \infty$, which indicates $x_c \to 1$ for both the SIS and the SIR model [144,274]. In other words, a complete vaccination of the network is required to stop an epidemic. As emphasized before, uniform vaccination is largely inefficient for disease eradication in heterogenous networks [224,225]. This theoretical prediction is also consistent with the finding that many infectious diseases require that 80%–100% of the population be vaccinated to prevent an outbreak. For example, measles require 95% of the population be vaccinated [270], while for digital viruses the strategies relying on uniform vaccination call for practically 100% of the computers to install the appropriate antivirus software [144,225,291].

5.2. Targeted vaccination

5.2.1. Degree-based targeted vaccination: theoretical prediction of the vaccination threshold

The problems encountered with uniform vaccination are rooted in the vanishing epidemic threshold. Therefore, to successfully eradicate a disease in a heterogenous network, we must increase the epidemic threshold. This requires us to modify the underlying contact network to reduce the degree variance $\langle k^2 \rangle$. Evidently, large-degree nodes are known to be responsible for the large variance of heterogenous networks. Thus, if we vaccinate only the nodes whose degree exceeds a preselected threshold value k_c, it should be easy to decrease the variance and increase the epidemic threshold [274,281,292]. To this aim, Pastor-Satorras and Vespignani have proposed a degree-based targeted vaccination strategy in which the most highly connected nodes are progressively immunized, given that these are the ones that are most likely to spread the disease [274,275,277,283,292–299].

The vaccination of a fraction x of large-degree nodes can be regarded as the removal of nodes whose degree is larger than a certain value k_c. Accordingly, x thus can be implicitly defined as

$$x = \sum_{k=k_c+1}^{\infty} P(k). \tag{121}$$

Moreover, the average $\langle k \rangle_c$ and the second moment $\langle k^2 \rangle_c$ of the degree distribution with maximum degree k_c of the network are

$$\langle k \rangle_c = \sum_{k=k_{min}}^{k_c} kP(k), \tag{122}$$

and

$$\langle k^2 \rangle_c = \sum_{k=k_{min}}^{k_c} k^2 P(k), \tag{123}$$

where k_{min} denotes the minimal degree of the node in a network.

However, the removal of large-degree nodes will modify the degree distribution of the remaining nodes, owing to the deletion of the links between removed and remaining nodes. The probability that the link connected to the removed nodes is also removed equals the probability that the link points to a node whose degree is larger than k_c. We thus have

$$f = \sum_{k=k_c+1}^{\infty} \frac{kP(k)}{\langle k \rangle}, \tag{124}$$

and the degree distribution of the resulting network becomes [274,289]

$$P'(k') = \sum_{k=k'}^{\infty} \binom{k}{k'} f^{k-k'} (1-f)^{k'} P(k). \tag{125}$$

Finally, the average degree $\langle k' \rangle$ and the second moment $\langle k'^2 \rangle$ of the resulting network are [274,289]

$$\langle k' \rangle = \sum_{k'=k_{min}}^{k_c} k' P'(k') = (1-f)\langle k \rangle_c, \tag{126}$$

and

$$\langle k'^2 \rangle = \sum_{k'=k_{min}}^{k_c} k'^2 P'(k') = (1-f)^2 \langle k^2 \rangle_c + f(1-f)\langle k \rangle_c, \tag{127}$$

respectively.

For the SIS model, the epidemic threshold of the resulting network can be calculated as

$$\beta'_{I_c} = \frac{\langle k' \rangle}{\langle k'^2 \rangle} = \frac{(1-f)\langle k \rangle_c}{(1-f)^2 \langle k^2 \rangle_c + f(1-f)\langle k \rangle_c}. \tag{128}$$

Assuming $2 < \zeta < 3$ and combining Eqs. (122)–(124), we obtain [281]

$$\beta'_{I_c} = \left[\frac{3-\zeta}{\zeta-2} k_c^{3-\zeta} k_{min}^{\zeta-2} - \frac{3-\zeta}{\zeta-2} k_c^{5-2\zeta} k_{min}^{2\zeta-4} + k_c^{2-\zeta} k_{min}^{\zeta-2} \right]^{-1}. \tag{129}$$

A similar approach to the SIR model yields

$$\beta'_{I_c} = \left[\frac{3-\zeta}{\zeta-2} k_c^{3-\zeta} k_{min}^{\zeta-2} - \frac{3-\zeta}{\zeta-2} k_c^{5-2\zeta} k_{min}^{2\zeta-4} + k_c^{2-\zeta} k_{min}^{\zeta-2} - 1 \right]^{-1}. \tag{130}$$

In both the SIS and the SIR model, if $k_c \gg k_{min}$, we have [281]

$$\beta_{I_c} \approx \frac{\zeta-2}{3-\zeta} k_{min}^{2-\zeta} k_c^{\zeta-3}. \tag{131}$$

For here it follows that if more large-degree nodes are vaccinated, and thus k_c becomes smaller, the epidemic threshold will become larger. By vaccinating an appropriate fraction of large-degree nodes, we enable β_{I_c} to increase beyond the spreading rate β_I of the disease. Accordingly, the network is then protected completely.

In order to assess the efficiency of degree-based targeted vaccination, it is useful to derive the explicit expression for the vaccination threshold in an uncorrelated network with a power-law degree distribution $P(k) = ck^{-\zeta}$ and $c \approx (\zeta-1)/k_{min}^{-\zeta+1}$, where $\zeta = 3$ [274]. In this case, we obtain $x = k_{min}^2 k_c^{-2}$, $f = x^{1/2}$, $\langle k \rangle_c = 2k_{min}$, and $\langle k^2 \rangle_c = 2k_{min}^2 \ln(x^{-1/2})$ based on Eqs. (121)–(124). The critical vaccination threshold x_c needed to eradicate the disease is then

$$\frac{\langle k' \rangle}{\langle k'^2 \rangle} = \beta_I. \tag{132}$$

By inserting above values into Eq. (132) and combining Eqs. (126)–(127), we obtain the approximate solution for the vaccination threshold in the case of targeted vaccination [274]
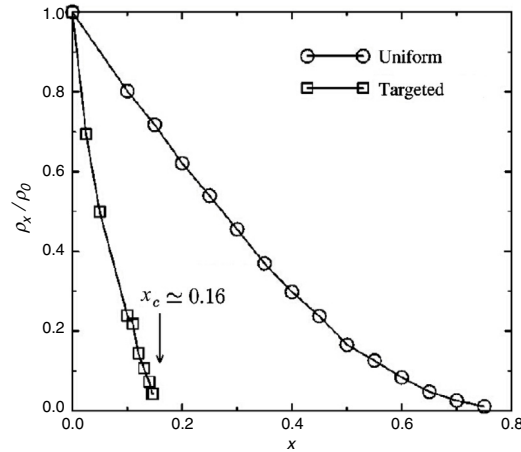
$$x_c \sim \exp(-2/k_{min}\beta_I). \tag{133}$$

Fig. 9. Reduced prevalence of disease ρ_x/ρ_0 for the SIS model on a scale-free network, as obtained with random (circles) and targeted (squares) vaccination at a fixed spreading rate $\beta_l = 0.25$, respectively. A linear extrapolation from the largest values of x yields an estimation of the threshold $x_c \simeq 0.16$ for targeted vaccination.
Source: Adapted with permission from Ref. [274].
© 2002, by the American Physical Society.

This equation indicates that the degree-based targeted vaccination strategy can be very effective, giving a critical vaccination threshold that is exponentially small for a wide range of spreading rates.

This theoretical prediction can be tested numerically with simulations of the SIS model on a scale-free network. Results presented in Fig. 9 compare the efficiency of targeted vaccination and random vaccination, respectively. It can be observed that for random vaccination the disease prevalence decays very slowly and vanishes only in the $x \rightarrow 1$ limit. Conversely, for targeted vaccination there is a very sharp drop toward the vaccination threshold (indicated by the arrow), above which the system is infection-free. A linear regression from the largest value of x yields an approximate vaccination threshold $x_c \simeq 0.16$ in this case, which proves that scale-free networks are very sensitive to degree-based targeted vaccination even if a very small fraction of large-degree nodes is selected for vaccination.

5.2.2. Centrality-based targeted vaccination: definitions and simulation-based models

The procedure of degree-based targeted vaccination is equivalent to altering the structure of the network. Namely, by vaccinating the large-degree nodes, we fragment the contact network, thus making it more difficult for the disease to reach the nodes in the other components. Based on this principle, it is possible to devise alternative vaccination strategies which take advantage of the heterogenous connectivity patterns in order to achieve a high level of tolerance toward infections [274]. Accordingly, other centrality indexes, such as betweenness, closeness, eigenvalue, or PageRank, can be used to design targeted vaccination strategies [288,293,294,296,297,297,300–310]. In what follows, we will briefly survey these centrality indexes and their applications in targeted vaccination.

Betweenness centrality. Betweenness centrality is defined as the number of shortest paths between pairs of nodes that pass through a given node [311–313]. More precisely, if σ_{vw} is the total number of shortest paths from node v to node w and $\sigma_{vw}(i)$ is the number of those shortest paths that pass through the node i, the betweenness of i is given by $b_i = \sum_{v \neq w \neq i} \frac{\sigma_{vw}(i)}{\sigma_{vw}}$.

Since nodes with large betweenness centrality are often bridges between different communities in the network [304,314], the vaccination of these nodes has a good potential to break up a large network into smaller parts. Hence, if an epidemic starts in a particular component of the network, it cannot infect nodes in other components. Additionally, betweenness centrality is an effective measure to identify high risk individuals, as it determines the volume of flow passing through each node [294]. That is also one of the main reasons why largest betweenness-based vaccination is considered to be the most effective targeted vaccination algorithm [315].

Random-walk centrality. Betweenness is, in some sense, a measure of the influence a node has over the spread of the epidemic through the network. By counting shortest paths, betweenness-based vaccination simply assumes that epidemics spread only along those shortest paths, which, however, is inconsistent with most empirical observations. To relax this hypothesis, Newman proposed the random-walk centrality that involves the contributions from essentially all the paths between nodes, though it still gives more weight to short paths [316]. The random-walk centrality of a node i is a measure based on random walks, counting how often the node i is traversed by a random walk between any pair of nodes s and v

$$r_i = \sum_{s<v} I_{sv}(i), \tag{134}$$

where $I_{sv}(i) = \frac{1}{2} \sum_j A_{ij} |T_{is} - T_{iv} - T_{js} + T_{jv}|$ for $i \neq s, v$, and T_{is} is the element in the voltage matrix. This measure is particularly useful for finding high-centrality vertices that do not happen to lie on geodesic paths or on the paths formed

by maximum-flow cut-sets, which are still important for the epidemic spreading [316]. Salathé and Jones have proven that vaccination strategies based on random-walk centrality can result in the lowest number of infected cases if the vaccination coverage is low [304].

Closeness centrality. Closeness centrality is based on the assumption that nodes with a short distance to other nodes can spread the disease very effectively across the network [294,311]. It is defined as

$$c_i = \frac{1}{\sum_{j \neq i} l_{ij}},$$

(135)

where l_{ij} is the distance between nodes i and j. This measure gives a large centrality to nodes which have small shortest path distance to the other nodes. Therefore, immunizing largest-closeness nodes not only vaccinates high risk individuals, but also postpones epidemic spreading through the network [311,312].

Eigenvector centrality. Degree centrality awards one "centrality point" for every network neighbor a node has. In reality, however, all the neighbors are not equivalent. Under many circumstances, a node's importance in one network increases via building connections to other nodes that are themselves important, which is the concept behind eigenvector centrality [296,315,317,318]. Instead of awarding just one point for each neighbor, eigenvector centrality gives each node a score proportional to the sum of the score of its neighbors. That is, the eigenvector centrality e_i of a node i is proportional to the sum of the eigenvector centrality of the nodes it is connected to. It is defined as

$$e_i = \Lambda^{-1} \sum_j a_{ij} e_j,$$

(136)

where a_{ij} denotes the adjacency matrix of the network, Λ is the largest eigenvalue and e_i is the ith component of the eigenvector associated with Λ of the network.

The impact of eigenvector centrality on epidemic control has been evaluated by [317,319]. Tomovski and Kocarev used nonlinear system stability analysis of the SIS model and proved that the epidemic threshold equals to the reciprocal value of the largest eigenvalue of the network adjacency matrix [319], that is

$$\beta_{l_c} = \frac{1}{\Lambda}.$$

(137)

Therefore, the risk of infection through networks can be reduced by decreasing Λ. Along the same lines, Masuda defined I_i as the decrement of Λ owing to the removal of node i [320]

$$I_i = -\frac{\Delta \Lambda}{\Lambda} \approx \frac{e_i^2}{\sum_j e_j^2}.$$

(138)

From here it is clear that the vaccination of a node with a large e_i (I_i) will significantly increase the epidemic threshold, and thus the epidemic will be effectively suppressed.

PageRank centrality. The PageRank centrality [307], introduced by Google for webpage ranking, is given by

$$p_i = c \sum_{j \in B_i} \frac{p_j}{N_j} + (1 - c)1/N,$$

(139)

where i is a web page, B_i represents the set of pages that point to i, and N_j denotes the number of pages that j points to. Moreover, c is a random jumping factor which is introduced by assuming that the web surfer will browse the web pages along the links with probability c, and leave the current page and open a random page with probability $1-c$. Similar to eigenvector centrality, PageRank supposes that the importance of a node is determined by both the quantity and the quality of the nodes that point to it. In terms of epidemic spreading, nodes with large PageRank centrality are more likely to be infected or infect others [296]. Their vaccination can thus eliminate a large number of potentially very effective disease transmission routes. Since nodes with large PageRank centrality have many low-degree neighbors, largest PageRank vaccination will immunize influential nodes whose vaccination choice can strongly protect their neighbors [296].

To obtain a better understanding of all these proposed vaccination schemes, Shams et al. compared the efficiency of different centrality indexes in terms of the vaccination threshold x_c, and plotted it as a function of the average degree $\langle k \rangle$ for different network models [315]. As shown in Fig. 10, it can be observed that HP (for the meaning of the abbreviation please refer to the caption of Fig. 10) delivers the best performance on all the networks, HB is second-best, while HE and HC deliver the worst performance. Despite of the weak performance of HD in small-world networks, the approach performs well on scale-free and ER random networks.

Moreover, we point out that targeted vaccination techniques can be improved by adaptive strategies, in which the centrality is recalculated for the network of unvaccinated nodes at each step of the vaccination process [313]. For example, degree-based vaccination with dynamic re-ranking entails an immediate update of node degrees, and then the nodes with the largest number of unvaccinated neighbors are vaccinated at each time step [296]. When a node has been vaccinated, the adaptive betweenness centrality is recalculated for the remaining network composed of unvaccinated nodes
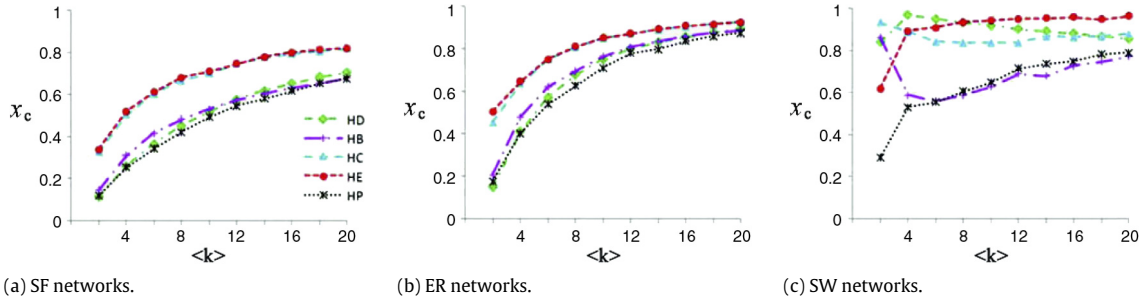
Fig. 10. The critical threshold of vaccination x_c, as a function of the average degree $\langle k \rangle$ on different networks. The legend refers to the following vaccination strategies: targeted vaccination based on degree (HD), betweenness (HB), closeness (HC), eigenvalue (HE), and PageRank (HP).
Source: Reprinted figure from Ref. [315].

[296,297,305,313]. Although such adaptive approaches are more efficient from the theoretical viewpoint, they usually require more information, and the complexity of the procedure might make a practical implementation difficult for all but very small populations [305].

5.3. Vaccination without global knowledge

Although targeted strategies are in principle very effective, an important drawback in terms of practical applications in the real world is that they require complete knowledge of the network structure. Only so it is possible to identify and then vaccinate the most influential nodes as determined by the employed metric. However, the complete structure of the network is seldom, if ever, known, and often it is also not well-defined. In social networks, for example, the number of links an individual has depends strongly on the criteria based on which the network is constructed. To overcome the lack of global knowledge of the network structure, several different strategies have been proposed that require information only about the local structure of the network. The efficiency of these vaccination strategies can nevertheless be very similar to that of targeted vaccination, as reviewed in what follows.

5.3.1. Acquaintance vaccination: theoretical prediction of the vaccination threshold

The so-called acquaintance vaccination was first proposed by Cohen et al. [269,277], in which a fraction p of nodes is selected at random, with the constraint that each node is required to have at least one connection to another individual in the network. Subsequently, this neighbor, rather than the originally selected node, is vaccinated. The strategy thus requires no knowledge of the degree of nodes or other global knowledge about the network.

More precisely, the probability that a node with degree k is selected for vaccination is $kP(k)/(N\langle k \rangle)$. This quantifies the known fact that randomly selected acquaintances have, on average, a higher degree than randomly selected nodes [176,321]. Based on percolation theory, Cohen et al. developed a theoretical framework to determine the critical fraction p_c and vaccination threshold x_c. They followed a possible branch in the course of the epidemic, starting from a random link of the spanning cluster. That is, they studied the possible spreading of the epidemic by considering nodes that are not vaccinated, and are therefore susceptible to the epidemic and may become infected [277]. If $n_l(k)$ denotes the number of nodes with degree k in some layer l (hop-distance from the starting point), the number $n_{l+1}(k')$ of nodes with degree k' that are susceptible in layer $l + 1$ is given by

$$n_{l+1}(k') = \sum_k n_l(k)(k-1)p(k'|k, s_k)p(s_{k'}|k', k, s_k), \tag{140}$$

where $k - 1$ means that the node with degree k in layer l has $k - 1$ new neighbors in layer $l + 1$ (excluding the one through which we arrived). Moreover, $p(k'|k, s_k)$ is the probability of reaching a node with degree k' by following a link from a susceptible node of degree k, while s_k is the probability that a node with degree k is susceptible. Lastly, $p(s_{k'}|k', k, s_k)$ is the probability that the reached node is also susceptible.

Based on Bayes' rule, it holds that

$$p(k'|k, s_k) = \frac{p(s_k|k, k')p(k'|k)}{p(s_k|k)}. \tag{141}$$

Furthermore, assuming that the network is uncorrelated, we have

$$\phi(k') = p(k'|k) = k'P(k')/\langle k \rangle, \tag{142}$$

which is independent of k. The probability that the acquaintance is not selected in one particular vaccination attempt by a random node of degree k, is $1 - 1/(Nk)$, and in all Np attempts, it reads

$$v_p(k) = \left(1 - \frac{1}{Nk}\right)^{Np} \approx e^{-p/k}. \tag{143}$$
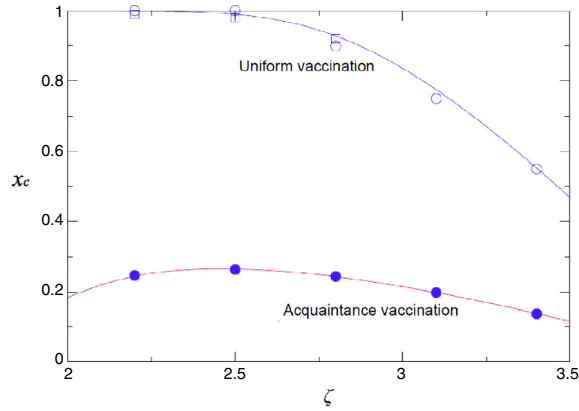
Fig. 11. The critical vaccination threshold x_c as a function of the degree exponent ζ of the contact network on which the disease spreads via the SIS model. The curves correspond to two distinct vaccination strategies: uniform vaccination (top curve) and acquaintance vaccination (lower curve). The continuous lines represent the analytical results while the symbols represent simulation data for $N = 10^6$ and $k_{min} = 1$.
Source: Adapted with permission from Ref. [277].
© 2003, by the American Physical Society.

However, if the neighbor's degree is not known, the average probability becomes $v_p = \langle v_p(k) \rangle$. Then the probability that a node with degree k' is susceptible (not vaccinated) is

$$p(s_{k'}|k') = \langle v_p(k) \rangle^{k'}, \tag{144}$$

if no other information on its neighbors exists. But when the degree of one neighbor is known to be k', we have

$$p(s_k|k, k') = e^{-p/k'} \times \langle e^{-p/k} \rangle^{k-1}. \tag{145}$$

Since the fact that a neighbor with known degree is vaccinated does not provide any further information about a node's probability of vaccination, it follows $p(s_k|k, k') = p(s_k|k, k', s_{k'})$.

Substituting above results in Eq. (140), we obtain

$$n_{l+1}(k') = n_l(k') \sum_k \phi(k)(k-1)v_p^{k-2}e^{-2p/k}. \tag{146}$$

If the sum in the above equation is larger than 1, the branching process will continue forever, which is referred to as the percolating phase. On the contrary, if it is smaller than 1, the vaccination is subcritical and the epidemic will be contained. Thus, we obtain a relation for p_c, which is

$$\sum_k \frac{P(k)k(k-1)}{\langle k \rangle} v_{p_c}^{k-2}e^{-2p_c/k} = 1, \tag{147}$$

and the vaccination threshold is easily obtained from the fraction of nodes which are not susceptible, namely

$$x_c = 1 - \sum_k P(k)p(s_k|k) = 1 - \sum_k P(k)v_{p_c}^k. \tag{148}$$

Results presented in Fig. 11 show the critical threshold that is required to eradicate a disease in a scale-free network. It can be observed that acquaintance vaccination is more efficient than random vaccination for both high and low ζ values, corresponding to homogeneous and heterogenous networks, respectively.

5.3.2. Other strategies: definitions and simulation-based models

In addition to acquaintance vaccination reviewed above, similarly efficient strategies can be devised based on local information about the structure of the network. An example is the so-called random walk vaccination strategy [322], in which a random walker diffuses through the network, and every node that it visits is vaccinated until a given fraction of the population is immune [323]. Given that a random walker visits a node with degree k_i with the probability proportional to k_i, this strategy could lead to the same effectiveness as achieved by acquaintance vaccination [229]. Another alternative to acquaintance vaccination is common acquaintance vaccination, where only the common neighbors of the randomly chosen nodes are selected for vaccination [276]. This may alleviate inefficiencies related to the fact that, of course, only very few of the randomly selected nodes and their neighbors will be the hubs of the network. Indeed, common acquaintance vaccination proves to be somewhat more effective than the simpler acquaintance vaccination.

The acquaintance vaccination strategy can also be improved by taking into account additional information about the local structure of the network. For example, if each node has the local knowledge about the degree of its nearest neighbors,
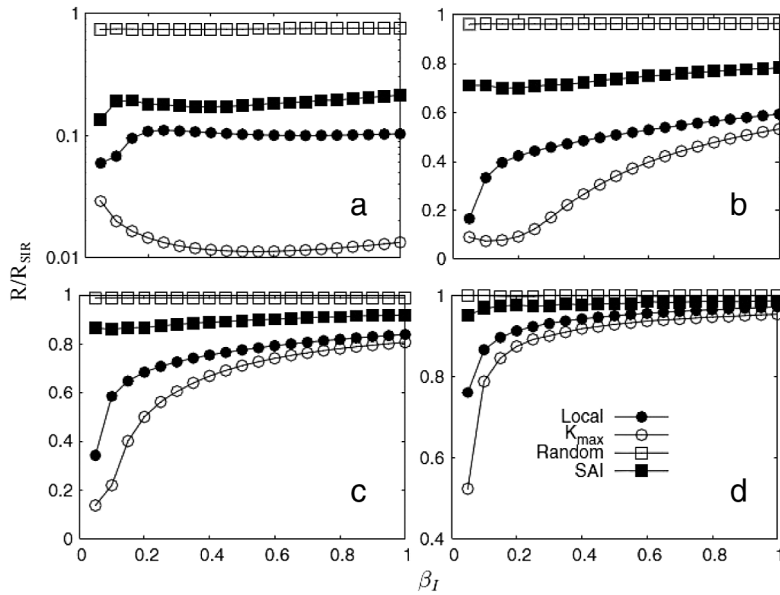
Fig. 12. Comparison of different vaccination strategies for the Internet router map. R/R_{SIR} represents the ratio between the epidemic incidence (R) of the four proposed vaccination strategies and the epidemic incidence (R_{SIR}) without vaccination. The legend refers to the following vaccination strategies: D-steps vaccination (Local), targeted vaccination (K_{max}), random vaccination (Random) and single acquaintance vaccination (SAI). In each case, 1% of the non-vaccinated nodes were initially infected at random. The distance considered in the local algorithm is: (a) $D = 1$; (b) $D = 2$; (c) $D = 3$; (d) $D = 5$.
Source: Adapted with permission from Ref. [326]. With permission of Springer.

and selects the neighbor with the degree to be vaccinated, then the efficiency in comparison to acquaintance vaccination is significantly improved [324]. Similarly, the random walk vaccination strategy can also be improved via a bias favoring the exploration of large-degree nodes during the random walk process [323,325].

Along the same line, if even more local information is available, for example if nodes have the knowledge about the degree of their neighbors within short paths of length D, then the so-called D-steps vaccination strategy can be employed [326,327]. In particular, for every node i, simply look for the largest-degree node within the distance D and vaccinate it. In case several nodes within the distance have the same largest degree, then choose one of them uniformly at random.

Based on these considerations, in [326] detailed numerical simulations of four different vaccination schemes, including D-steps vaccination, degree-based targeted vaccination, random vaccination, and acquaintance vaccination, were performed and tested in terms of their efficiency. As shown in Fig. 12, targeted vaccination expectedly produces the best results, and this irrespective of the interaction topology. Also understandably, the performance of random vaccination is not affected by the structure of the network. Turning to local algorithms, it is found that the vaccination scheme based on the covering algorithm performs better than acquaintance vaccination, even for small values of D. In fact, the D-steps vaccination is only outperformed by the targeted procedure and lies between the most efficient and the acquaintance vaccination scheme for all the values of D and β_I. Moreover, from a practical point of view, the covering strategy could be a good policy since it balances the degree of local knowledge and the efficiency of the vaccination. As most empirical network topologies are neither completely known nor completely unknown, this method thus allows us to fine-tune the value of D on a case-by-case basis, i.e., according to the degree of local knowledge of the network.

In addition, Gao et al. evaluated the relationship between the threshold x_c of the aforementioned vaccination strategies and the power-law exponent of the degree distribution of the network, which is shown in Fig. 13 [283]. A low level of x_c with a small variation indicates that the immunization strategy is robust. Results presented in Fig. 13 also show that the robustness of the D-steps vaccination strategy is close to that of targeted vaccination.

Vaccination plays an obviously important role in the study of epidemic transmission and prevention. Based on the above-reviewed research, we can conclude that the efficiency of vaccination depends significantly on the amount of information we have about the structure of the network, and whether this information concerns the global or the local structure. In the rare case information about the global structure is available, targeted vaccination works best. If only information about the local structure is available, then the D-steps vaccination seems to provide the optimal compromise between efficiency and flexibility. Table 3 gives a systematic summary of the reviewed vaccination strategies, for a quick overview.

5.4. Vaccination on other types of networks

Up to now, we have focused mainly on research concerning vaccination on traditional network models. This research provides vital insight in terms of the efficiency of vaccination under different circumstances, yet empirical networks usually have particular topology properties that require special care. This includes taking into account community structure
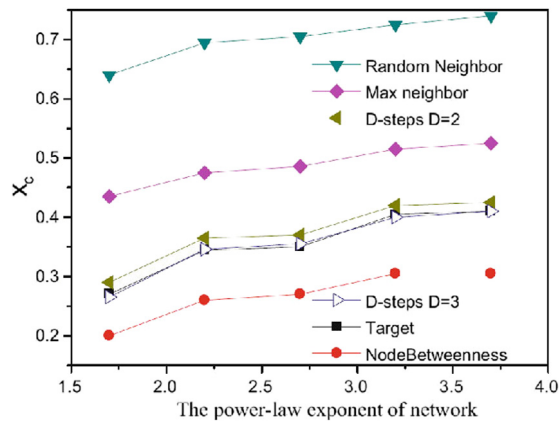
Fig. 13. The critical vaccination threshold x_c as a function of the power-law exponent of the degree distribution of the network. The legend refers to the following vaccination strategies: acquaintance vaccination (Random Neighbor and Max neighbor), D-steps vaccination (D-steps $D = 2$ and D-steps $D = 3$), node degree-based targeted vaccination (Target) and node betweenness-based targeted vaccination (NodeBetweenness).
Source: Adapted with permission from Ref. [283]. With permission of Springer.

Table 3
Summary of non-behavioral epidemiological vaccination strategies, with details concerning vaccinated nodes, network models, and the required knowledge about the structure of the network.

Vaccination strategy	Vaccinated nodes	Applied network models	Required knowledge of topology
Uniform (or random) vaccination	Random chosen nodes	Homogeneous networks less heterogeneous networks	No any need of network information
Targeted vaccination	Nodes with high centrality[a]	Heterogeneous networks networks with special centrality	Global knowledge of network topology
Acquaintance vaccination[b]	Random neighbor of random chosen node[c]	Homogeneous networks heterogeneous networks	Local knowledge of network topology

  [a] Centrality includes degree centrality, betweenness centrality, random-walk centrality, closeness centrality, eigenvector centrality and PageRank centrality.
  [b] Except for simple acquaintance vaccination, it also includes common acquaintance vaccination, random walk vaccination, D-steps vaccination strategy.
  [c] Correspondingly, it also involves common neighbor of random chosen nodes, nodes visited by random walker, largest-degree node within the distance D of random chosen node.

[169,187,328–331], changes in the network structure over time [52,217,332–336], as well as multilayer properties that are typical for a large plethora of real-world networks [202,230,239,240,337–346]. In what follows, we will review vaccination programs that take into account the particular aspects of these properties on various types of networks.

5.4.1. Vaccination on community networks

Community structure is ubiquitous in a variety of real complex systems, such as human contact networks and social networks. The hallmark of this fact is that the connections among members of the same community are much more common than connections among members of different communities [169,187,328–331]. For networks with community structure, the weak ties that connect a pair of nodes belonging to different communities, usually named the bridge nodes, therefore typically provide pathways for information and disease to propagate from one community to the other. These nodes thus play a more important role in the spreading of disease than the nodes with no or fewer inter-community links [347,348]. However, their importance is not necessarily reflected by their degree centrality. In this regard, identifying the bridge nodes in community networks is crucial in preventing epidemic outbreaks [320,349]. Identification of community bridges can be partially achieved by betweenness centrality or random-walk centrality, since both locate nodes on the basis of important paths of epidemic spreading. On a similar note, Chen et al. proposed a vaccination strategy based on an equal graph partitioning algorithm to identify the minimum group that separates a network into several clusters of approximately equal size [293]. The role of the minimum separator group is actually similar to that of the bridge nodes connecting different communities.

Salathé et al. [304] subsequently proposed a dedicated community bridge finder (CBF) algorithm that searches for the bridge nodes, thereby requiring only local structural information about the structure of the network. Their research showed that such a method is more efficient than vaccination strategies targeting different kinds of hubs. The algorithm is based on the self-avoiding walk, which starts from a randomly chosen node v_0, and then follows the procedure outlined in Fig. 14. After t steps ($t \geq 2$), the set of all the visited nodes is denoted by $\{v_{t'}\}$ for $t' = 0, 1, \ldots, t$. Thus, v_t is the node where the walker is located after t steps, and v_{t-1} is the node visited at step $t - 1$. The first process is to examine whether the node v_t
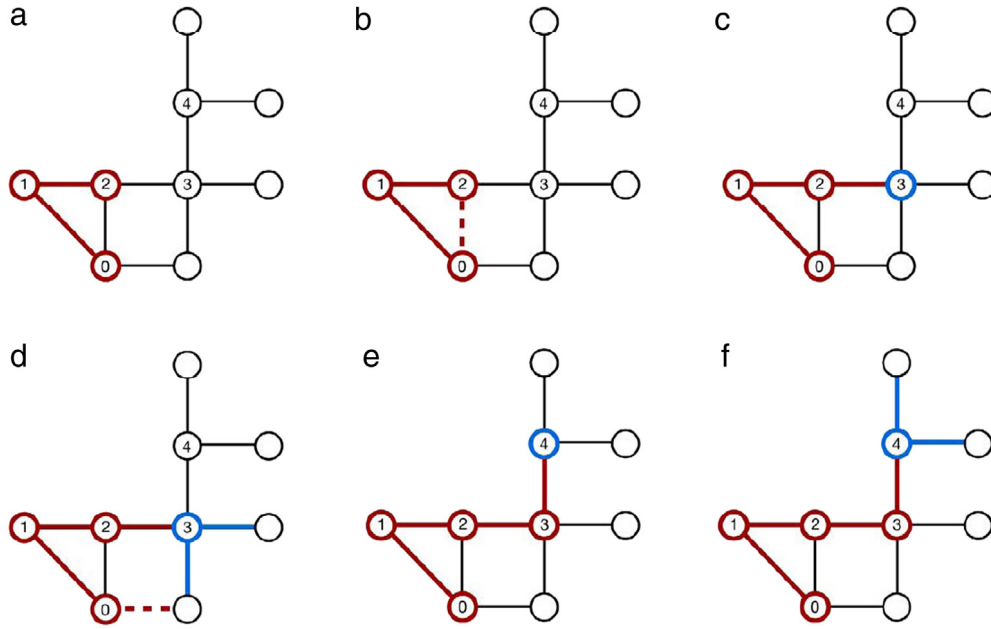
Fig. 14. A schematic presentation of the community bridge finder (CBF) algorithm. (a) A random walker follows the path starting from v_0 to v_1 and v_2, at which point it starts checking the connections from v_2 to v_0 and v_1. (b) Since there is more than one connection ($v_2 - v_1$ and $v_2 - v_0$), the walker turns to v_3. (c) Except for the obvious $v_3 - v_2$, there are no connections from v_3 to any of the previously visited nodes, so v_2 is a potential target. (d) The algorithm then picks two random neighbors of v_3 to check connections with previously visited nodes and finds one (to v_0). (e) Hence, v_2 is dismissed as a potential target, and the random walker comes to v_4. Again, v_4 does not connect to any previously visited node (except to v_3), and thus v_3 is identified as a potential target. (f) Again, two random neighboring nodes are picked to check connections with previously visited nodes. Since no back connections can be found, v_3 is identified as a target and vaccinated.
Source: Reprinted figure from Ref. [304].

has a link or several links to the nodes in the set $\{v_{t'}\}$, other than the link between v_t and v_{t-1}. If yes, the self-avoiding walk proceeds toward step $t + 1$; otherwise v_{t-1} will be considered as a possible target of the bridge node. To determine whether the node v_{t-1} is one bridge node, two nodes are randomly chosen among all the possible nodes that the walker could visit in step $t+1$, i.e. two neighbors of the node v_t are randomly chosen (except for the node v_{t-1} due to the self-avoiding restriction of the walk). If there exists a path from any of the two chosen nodes back to any node in the set $\{v_{t'}\}$, then the node v_{t-1} will not be a bridge node and the walker moves to node v_{t+1}. If there exists no path back to the set $\{v_{t'}\}$ from both chosen nodes, v_{t-1} is regarded as a bridge node that connects two communities and is vaccinated. Then a new self-avoiding walk starts and the above procedure is repeated until the desired vaccination ratio is attained. An important idea behind the CBF algorithm is actually that a community is formed by a circle of close friends. Thus, when two randomly chosen neighbors of v_t cannot trace back to the community that v_{t-1} belongs to, the link between v_{t-1} and v_t is likely to be a bridge between two communities, and the node v_{t-1} is hence a bridge node.

It has been found that community networks typically exhibit a heterogeneous distribution in the number of weak ties originating from the bridge nodes [184,350]. Inspired by this finding, Gong et al. turned their attention to a particular kind of bridge nodes, named bridge hubs, which connect a community to many other communities, i.e., bridge hubs have a large number of weak ties [349]. They proposed a vaccination strategy named bridge-hub detector (BHD) to identify the bridge hubs. This algorithm extends the self-avoiding searching scheme to examining the overlap and the existence of links from all the neighbors of the last node back to the union of the friendship circles of all the nodes in the trail of the walk. A pair of nodes, a bridge node and a bridge hub, are searched for vaccination via a self-avoiding walk. Compared with other local vaccination strategies such as acquaintance vaccination and CBF, this strategy is proven to be more effective in preventing the epidemic [349].

Also of note, Masuda recently proposed a vaccination strategy based on the eigenvector centrality for community-structured networks [320]. In his work, a module based strategy technique is employed for measuring the contribution of each node to the weighted community network in order to preferentially vaccinate nodes that bridge important communities. The number of links that these nodes share gives the weight of a link between any two communities. The influence of a node is thus related to the importance of the community together with its connectivity to other communities.

5.4.2. Vaccination on adaptive networks

From the epidemiological viewpoint, when an infectious disease occurs in a population, individuals are likely to adopt some self-protection measures to protect themselves from the disease. For example, susceptible people may break their contacts with infected partners, which can significantly alter the structure of the contact network, thus influencing the
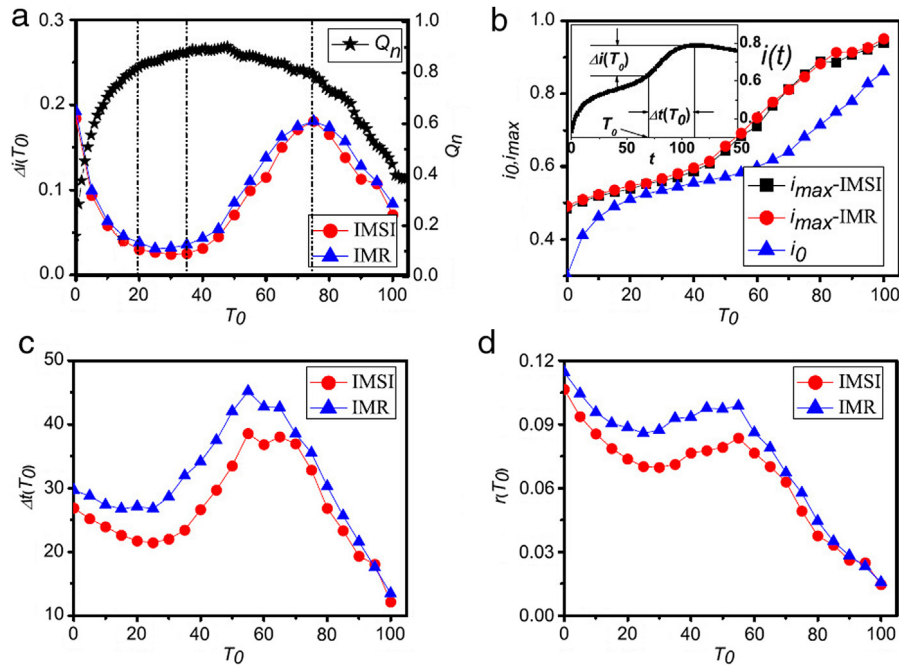
Fig. 15. The effect of IMR strategy and IMSI strategy versus the starting time T_0 of quarantine for (a) $\Delta i_0(T_0)$, (b) i_0, i_{max}, (c) $\Delta t(T_0)$ and (d) $r(T_0)$.
Source: Reprinted figure from Ref. [351].

pathway of the epidemic spreading [52,333–335,351–354]. Both SIS and SIR models have been studied on adaptive networks in which susceptible nodes rewire their links adaptively from infected neighbors toward other non-infected nodes [52,333]. Such adaptation typically increases the epidemic threshold and reduces the number of infectious cases, thus contributing favorably to the containment of epidemic spreading.

Moreover on the subject, Shaw et al. have studied vaccine control of disease spreading on an adaptive network to model the specifics of disease avoidance behavior [335]. Control was implemented by adding a Poisson-distributed vaccination to the susceptible individuals. Their research showed that vaccine control is much more effective in adaptive networks than static networks due to the feedback loop between the adaptive rewiring and the administration of the vaccine. On the other hand, Yang et al. have focused on the transient process rather than the steady state, and found that strong community structure (S nodes community and I nodes community, connected by SI links) is induced by the rewiring mechanism in the early stages of the epidemic spreading [351]. Subsequently, they proposed and examined various vaccination strategies that build on this fact, targeting in particular such SI links. In particular, the vaccination of randomly selected S bridge nodes from SI links (IMSI) and the vaccination of randomly selected I bridge nodes from SI links (ISSI) were considered. To evaluate the efficiency of IMSI and ISSI strategies, the vaccination of randomly selected S nodes from the network (IMR) and the vaccination of randomly selected I nodes from the network (ISR) were also employed.

In terms of their results, in Figs. 15 and 16, T_0 denotes the starting time of vaccination, $\Delta i_0(T_0)$ is the difference between the maximal density of infected individuals $i_{max}(T_0)$ that can be reached after vaccination and the density of infected individuals $i_0(T_0)$ at time T_0, i.e. $\Delta i_0(T_0) = i_{max}(T_0) - i_0(T_0)$. $\Delta t(T_0)$ represents the time interval of this process, and $r(T_0)$ denotes the total percentage of vaccinated nodes in $\Delta t(T_0)$. In Fig. 15, the impact of the IMR strategy is almost equal to that of the IMSI strategy. Because most S nodes are located on SI links when an epidemic prevails in an adaptive network, randomly vaccinating S nodes can give rise to similar effects as vaccinating S nodes on SI links. Nevertheless, for ISR and ISSI, quarantining I nodes on SI links is significantly more effective than quarantining I nodes randomly, since the former can efficiently cut the pathways along which the disease could be invading the S cluster (Fig. 16). Because of this fact, the ISSI strategy is more efficient in controlling the epidemic and has a larger optimal region. More importantly, the consideration of these different vaccination strategies reveals a counterintuitive conclusion: "the earlier the better" does not necessarily hold for the implementation of vaccination measures. Optimal results can be expected more realistically if smart vaccination strategies are implemented when the community structure emerges after the initial rewiring of links that sets in due to an epidemic.

5.4.3. Vaccination on multiplex networks

Multiplex networks, as a typical kind of multilayer network structure, can be regarded as the combination of several network layers, which contain the same nodes (or share at least some fraction of nodes) yet different intra-layer connections [202,230,239,240,337–343,355–357]. For each node, we term its counterpart in every layer as a 'replica'.
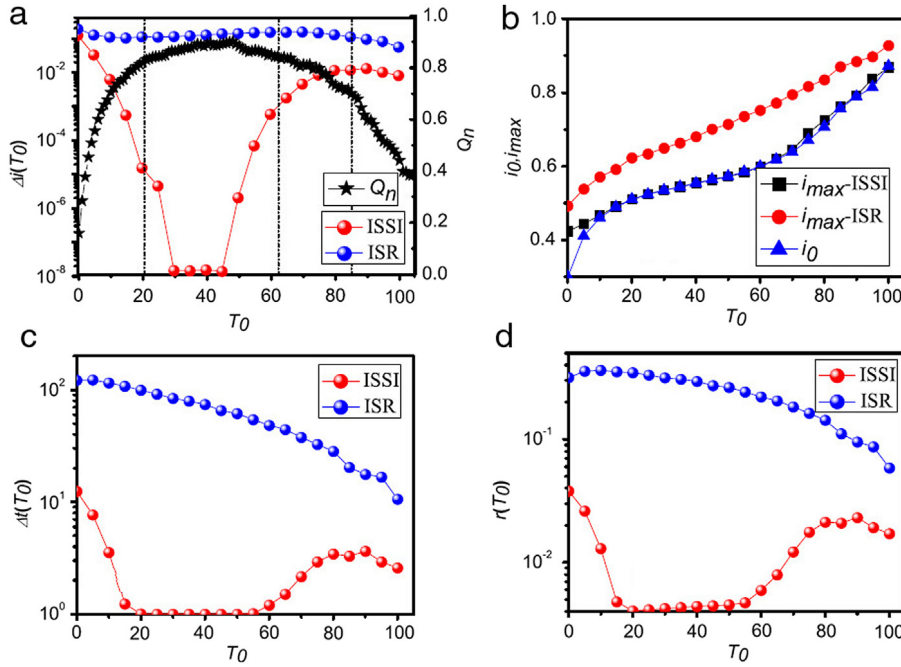
Fig. 16. The effect of ISR strategy and ISSI strategy versus the starting time T_0 of quarantine for (a) $\Delta i_0(T_0)$, (b) i_0, i_{max}, (c) $\Delta t(T_0)$ and (d) $r(T_0)$.
Source: Reprinted figure from Ref. [351].

Along the above definition, may real-world social and engineering systems, such as online social networks [231,358–364], airport traffic networks [357,365,366], biological metabolic networks [367,368], and scientific collaboration and citation networks [369], can be put into the framework of multiplex networks. Multiplex networks naturally shed new light into the research of epidemiology (see Refs. [202,230] for a comprehensive understanding). For example, the well known acquired immune deficiency syndrome (AIDS) usually propagates via three types of ways, namely sexual activity, blood, and breast milk. Different transmission routes could be mapped into different network layers which contain the same individuals yet different network topologies and dynamic properties [239]. Moreover, coupling methods from different layers will greatly affect the onset of disease and its outbreak threshold [239]. In what follows, we will focus on recent research concerning vaccination in multiplex networks.

To distinguish the node of a multiplex network and its replica in each network layer, Zhao et al. proposed the terminology of the multiplex node and the layer node [343]. The former refers to the node which has neighbors in each layer via different intra-layer connections, like node 3 of Fig. 17(a), whose neighbors are nodes 1, 5 and 4 via blue and black links. While the latter is the partial case of a multiplex node and just involves the local connections of a node in one given layer. For example, if we only consider blue (black) links in Fig. 17(b), node 3 is the layer node of Layer-1 (Layer-2). Correspondingly, vaccination on multiplex networks could be classified into multiplex node-based vaccination and layer node-based vaccination. Multiplex node-based vaccination means that all the replicas of one node simultaneously possess immunity in all the layers, while layer node-based vaccination provides protection solely to nodes in a given network layer.

Similar to the vaccination strategies on single-layer networks, Refs. [343,370–372] proposed the corresponding vaccination strategies (including random vaccination, targeted vaccination, and acquaintance vaccination) on multiplex networks. Here, however, the aim of vaccination is specified to layer nodes and multiplex nodes. These vaccination strategies are determined by the vaccinated probability of layer nodes and multiplex nodes, which are summarized in Table 4.

Multiplex node-based vaccination refers to the case where a fraction of multiplex nodes is vaccinated. If $x(\overrightarrow{k_j})$ is defined as the vaccinated probability of multiplex node j, the generating function of the joint degree distribution under multiplex node-based vaccination is given by

$$G_0(\overrightarrow{u}) = \sum_{\overrightarrow{k_j}} p(\overrightarrow{k_j})(1 - x(\overrightarrow{k_j})) \prod_{i=1}^{m} u_i^{k_{j_i}}. \tag{149}$$

The generating function of the remaining joint degree distribution by following a randomly chosen link of layer i is defined as

$$G_1^{(i)}(\overrightarrow{u}) = \frac{1}{z_i} \frac{\partial}{\partial u_i} G_0(\overrightarrow{u}). \tag{150}$$

Table 4
Vaccinated probability of layer nodes and multiplex nodes for different vaccination strategies.

Vaccination type & probability	Layer node-based vaccination	Multiplex node-based vaccination
Random vaccination	$x(k_{j_i}) = x_i$, for $\forall j$	$x(\overrightarrow{k_j}) = x$, for $\forall j$
Targeted vaccination	$x(k_{j_i}) = \begin{cases} 1, & \text{if } k_{j_i} > k_{ci} \\ f_i, & \text{if } k_{j_i} = k_{ci} \\ 0, & \text{if } k_{j_i} < k_{ci} \end{cases}$	$x(\overrightarrow{k_j}) = \begin{cases} 1, & \text{if } K_j > K_c \\ f, & \text{if } K_j = K_c \\ 0, & \text{if } K_j < K_c \end{cases}$ where $K_j = \sum_{i=1}^{m} \beta_{Ii} k_{j_i}$
Acquaintance vaccination	$x(k_{j_i}) = 1 - (1 - q_i \sum_{\overrightarrow{k_j'}} \frac{k_{j_i}' p(\overrightarrow{k_j'})}{z_i} \frac{1}{k_{j_i}'})^{k_{j_i}}$	$x(\overrightarrow{k_j}) = 1 - \prod_{i=1,\dots,m} (1 - \frac{q}{m} \sum_{\overrightarrow{k_j'}} \frac{k_{j_i}' p(\overrightarrow{k_j'})}{z_i} \frac{1}{k_{j_i}'})^{k_{j_i}}$

$\overrightarrow{k_j} = (k_{j_1}, k_{j_2}, \dots, k_{j_m})$: the degree expression of multiplex node j; k_{j_i} refers to its degree in layer i.
$x(k_{j_i})$: the vaccinated probability of layer node j in layer i with degree k_{j_i}.
$x(\overrightarrow{k_j})$: the vaccinated probability of multiplex node j.
β_{Ii}: the infectious rate of the disease in layer i.
k_{ci}: the cutoff degree for vaccination of layer node in layer i.
K_j: the spreading degree of multiplex node j.
K_c: the cutoff spreading degree for vaccination of multiplex node.
z_i: the average degree of layer i.
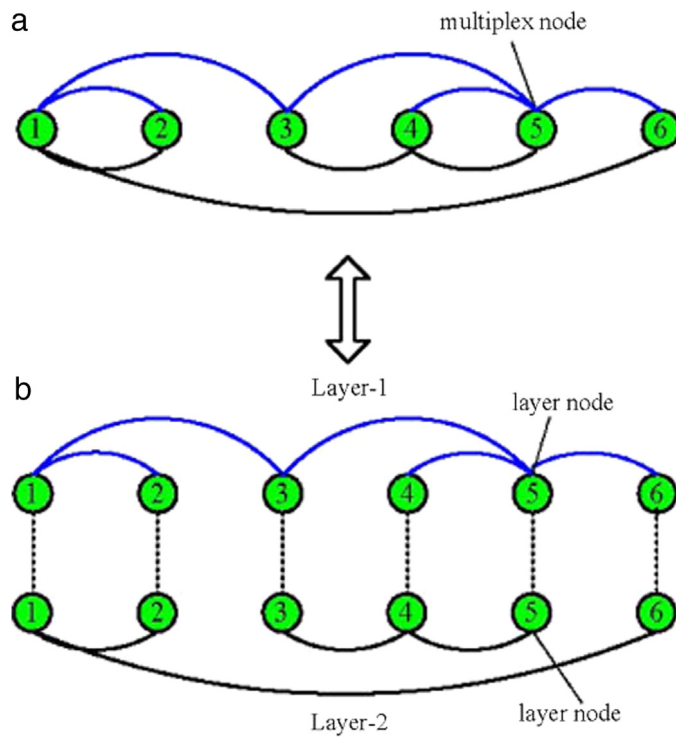
Fig. 17. Schematic illustration of a multiplex configuration. (a) In a multiplex network, six nodes are connected through two different types of links, i.e., blue and black links. In this configuration, each node simultaneously belongs to two layers, and it is thus termed as a multiplex node. (b) Multiplex networks are composed of several, for example, two layers, each of which is composed by the same type of links. The term layer node indicates that the node is connected only through the one type of link in the corresponding layer. See main text for details with regards to the importance of this classification for vaccination strategies.
Source: Reprinted figure from Ref. [355].

Then, the probability υ_i that a multiplex node connects to the infected cluster by following a random link of layer i is given by the coupled self-consistency equation

$$\upsilon_i = G_1^{(i)}(\overrightarrow{1}) - G_1^{(i)}(\overrightarrow{1 - \beta_I \upsilon}), \tag{151}$$

where $\overrightarrow{1 - \beta_I \upsilon} = (1 - \beta_{I1} \upsilon_1, \dots, 1 - \beta_{Im} \upsilon_m)$. It is worth mentioning that the cluster of multiplex networks is defined as a set of connected multiplex nodes. A pair of multiplex nodes is regarded as having a connection if there exists at least one type of link between them.

Thus, the existence of an epidemic regime under multiplex node-based vaccination just requires the largest eigenvalue Λ of the Jacobian matrix of Eq. (151) at $(0, 0, \ldots, 0)$ to be larger than unity [373]. Consequently, the critical vaccination threshold x_c^M will be the value of immunity $x^M = \sum p(\vec{k_j}) x(\vec{k_j})$ satisfying $\Lambda = 1$ in Eq. (151).

Layer node-based vaccination, in contrast, aims to vaccinate layer nodes. To calculate the critical vaccination threshold, we just need to replace $G_0(\vec{u})$ by

$$G'_0(\vec{u}) = \sum_{\vec{k_j}} p(\vec{k_j}) \prod_{i=1}^m (x_i(k_{j_i}) + (1 - x_i(k_{j_i})) u_i^{k_{j_i}}), \tag{152}$$

in Eq. (150), where $x(k_{j_i})$ is the vaccinated probability of layer node j in layer i with degree k_{j_i}. Since the vaccination of layer nodes is performed independently in each layer, the critical vaccination threshold becomes a vector $(x_1^L, \ldots, x_m^L)_c$ which can be obtained when the immunity $x_i^L = \sum_{k_{j_i}} p_i(k_{j_i}) x_i(k_{j_i})$ is satisfying $\Lambda = 1$ in Eq. (151).

According to the above theoretical analysis, computer simulations have revealed that both multiplex node-based random vaccination and layer node-based random vaccination strategies are more effective in controlling disease spreading on multiplex ER networks [343]. On the contrary, multiplex node-based targeted vaccination and layer node-based targeted vaccination strategies provide better protection on multiplex SF networks. As on single-layer networks, random vaccination strategies work well only in largely homogeneous networks, while targeted vaccination strategies are required for useful results on heterogeneous networks.

With regards to the efficiency of acquaintance vaccination in multiplex networks, we note that it is closely related to the correlation between network layers [370]. Interestingly, as the correlation coefficient increases, the vaccination threshold decreases for multiplex node-based acquaintance vaccination, but slowly increases under the layer node-based framework. Moreover, Buono et al. considered vaccination of only one layer and studied its effect on all the other layers while disregarding degree–degree correlations among them [371]. Though, relative to the case of no vaccination, the size of the epidemic is drastically reduced in the layer where the vaccination strategy is applied, the targeted strategy is still not as efficient as in single-layer networks. Thus, the selection of the vaccination strategy has a major effect on the layer where it is employed, but does not efficiently protect the individuals on other layers. Similar conclusions have also been reported in Refs. [372,374], thus corroborating the need to reconsider single-layer vaccination strategies when applied to multilayer networks.

6. Rationale for studying behavior–disease dynamics

As demonstrated by the foregoing sections of this review, epidemiological models that do not capture behavioral dynamics have been very successful over the past century, in terms of yielding rich theoretical opportunities for research as well as in terms of providing insights into epidemiological mechanisms and how to improve infection control. However, the impact of human behavior on disease dynamics is ubiquitous, and vaccinating behavior is no exception.

This section is concerned with explaining the rationale for studying and modeling coupled behavior–disease dynamics for vaccines, especially on networks. We start with a brief summary of vaccine use in public health, discuss why vaccinating behavior may become an increasingly important part of disease control in coming decades, and then discuss how models can help us to better understand and predict behavior–disease dynamics.

The twentieth century saw enormous progress in public health, and especially in terms of preventing and treating infectious diseases. For instance, medical use of antibiotics began during World War II, and antibiotics have since changed how bacterial infections are treated and saved many lives. Improvements in sanitation and hygiene are likewise credited with preventing many cases of cholera, typhoid fever, yellow fever, dysentery, tuberculosis, and malaria by providing cleaner drinking water and living conditions and improved hygienic practice [375]. The use of vaccines, although already available for smallpox since the eighteenth century, expanded considerably in the twentieth century as vaccines were invented for more infectious diseases, and were distributed to more people. Safe and effective vaccines now exist for many of the most common infectious diseases that impose significant morbidity and mortality upon populations, such as smallpox, measles, pertussis, influenza, hepatitis A and B, chickenpox, diphtheria, and others [3]. Global smallpox eradication was achieved in large part due to the success of ring vaccination [376].

The use of vaccines has been estimated to save millions of children's lives per year in the 1990s alone [3]. In the 2000s, special immunization activities (SIAs) – ambitious large-scale immunization campaigns in the world's poorest regions and especially in sub-Saharan Africa – reduced the number of deaths due to measles from an estimated 766,000 per year in 2000 to an estimated 164,000 per year in 2008 [377]. Global eradication of measles in the coming decades is becoming seen as a real possibility for the first time in history [378]. Thus, although challenges still remain, it is becoming significantly easier to get vaccines administered to the people who need them. There are, however, important exceptions. For instance, the challenges of administering vaccines in war zones have contributed to a large outbreak of polio in Syria as recently as 2015 [379]. At the same time in Pakistan, violence against WHO vaccinators has stymied efforts to combat polio [380]. Despite this, global polio eradication is coming closer [381].

While some countries with resource limitations struggle to contain vaccine-preventable infectious diseases, other comparatively wealthy countries are also struggling with infectious disease control, albeit for different reasons. In countries

Fig. 18. Satirical cartoon by James Gillray capturing public fears of the smallpox vaccine in the 18th century, the first "vaccine scare". Edward Jenner stands in the middle, administering the smallpox vaccine. Vaccinated persons grow cow parts out of various parts of their bodies (the smallpox vaccine of the time was based on the virus that causes cowpox, hence the fear among some that the vaccine would turn humans into cow-like hybrids). A painting of the Worshippers of the Golden Calf hangs in the background.
Source: Wikimedia Commons.

such as the United Kingdom, Switzerland, and Germany, measles circulation in the 2000s and 2010s has been relatively widespread compared to previous decades, and has often re-established endemicity in some populations due to parental vaccine refusal or vaccine hesitancy on account of unfounded concerns about the link between the measles–mumps–rubella (MMR) vaccine and autism [382–384]. Similarly, polio was on the cusp of global eradication and had been limited to northern Nigeria when an oral polio vaccine scare in 2003 caused resurgence of the disease and spread of polio to countries like Syria and Pakistan that had previously eliminated the infection [385]. These are only two examples where population resistance to vaccines due to unfounded concerns has led to reduced vaccine coverage and resurgent infections, even in resource-rich countries where it is easy for children to get vaccinated, if their parents want it.

Logistic, financial, technological, and administrative barriers to vaccination are likely to continue receding in the coming decades, enabling it to become easier to improve vaccine coverage everywhere in the world. As this happens, we speculate that vaccine refusal and vaccine hesitancy such as recently observed in the case of MMR vaccine will become an important barrier – and perhaps even the most important barrier – to global eradication of vaccine-preventable infections. Many factors contribute to the decision to become vaccinated [6,383,386]. Also, vaccine refusal is not a new phenomenon, and even the first vaccine to be invented – the smallpox vaccine – was subject to considerable resistance (see Fig. 18). However, growing herd immunity generated by vaccine programs themselves can bring about the preconditions for vaccine refusal.

Vaccine refusal is often construed in the context of "free-riding": individuals who do not vaccinate are consciously free-riding on the herd immunity provided by those who do vaccinate. Debate on the usefulness of coupled behavior–disease models often centers on whether individuals consciously free-ride when they make vaccinating decisions or not. However, it is not necessary for individuals to consciously free-ride in order for vaccine refusal to be linked to growing herd immunity, or for behavior–disease models to be useful (although it is certainly possible that some individuals will consciously free-ride when making vaccinating decisions [387]). It is well-known that perceived infection risk and past exposure to infections are determinants of whether or not individuals will choose to be vaccinated [386]. In particular, a high perceived infection risk and/or a past history of infection are positively associated with decisions to accept vaccination. Therefore, lack of recent infection, such as caused by vaccine-generated herd immunity, can decrease the motivation to become vaccinated. Parents who have never seen a case of measles infection leading to hospitalization will tend to under-estimate the risk and/or probability of measles infection. Similarly, lack of a recent outbreak can cause carelessness or de-prioritization on the part of physicians, and therefore lower vaccine uptake [388]. Whether or not non-vaccinators consciously free-ride is secondary in some respects. What matters is that vaccine-generated herd immunity brings about lack of recent infections in a population, which can in turn cause vaccinating to become less of a priority, as well as creating fertile conditions for false rumors about vaccine risk to spread. Vaccine programs are victims of their own success.

This is also echoed in observed behavior for pertussis and MMR vaccines. When declining vaccine coverage caused in partly by vaccine refusal leads to an outbreak, vaccine coverage responds by increasing during and after the outbreak

[389,390]. This vividly demonstrates the coupled nature of disease dynamics and vaccinating choices, completely aside from whether we need to characterize this behavior as free-riding, or whether non-vaccinating individuals are aware that is what they are doing. Individuals may not answer truthfully when surveyed about their motives for becoming vaccinated or not becoming vaccinated, or more importantly they may answer very differently in surveys conducted during an outbreak, than during a time of elimination. However, when we observe vaccinating behavior responding to changes in disease prevalence, as in the Disneyland, California measles outbreak of 2014 [391] and other examples, we see the proof in the behavioral pudding that vaccinating behavior and disease dynamics are strongly linked to one another, particularly near the elimination threshold.

The Disneyland, California measles outbreak and similar outbreaks illustrate the important role played by social factors in determining individual vaccinating choices. The influences of peer groups and medical professional opinions on vaccine decision-making are well-documented [6,392]. Social contact networks can be mapped out using Bluetooth and other technologies (see Section 11 for more details) [393]. Moreover, social influence has often been described as a contagious process in itself, as ideas and behaviors spread through networks [394,395]. Therefore, social processes can be very naturally modeled by defining processes on networks representing social contacts, and the methods of statistical physics concerned with the properties of networks could therefore be very useful [396].

Because vaccine-generated herd immunity tends to create complacency toward vaccine programs in the way we have described above, we expect vaccine hesitancy and outright vaccine refusal or vaccine "scares" to become more common as eradication and elimination thresholds are approached for more vaccine-preventable infectious diseases. Moreover, we have seen how the significant role of social context in influencing individual decision-making is very naturally represented using the network paradigm. The study of coupled behavior–disease dynamics on networks could therefore help us address this potentially growing problem [397]. In particular, mathematical models, often based on methods borrowed from physics, can be a useful way of better understanding and predicting these coupled dynamics.

Human behavior is doubtless more difficult to mathematically model than the motion of a single billiard ball, for instance, or many of the slightly more complex systems studied by physicists. However, at some levels of organization (in particular, the aggregate population level) it is often possible to develop simple models that have some predictive power. For instance, mathematical models incorporating behavior–disease interactions have been fitted to time series data of infectious disease prevalence, and model selection approaches such as Aikaike Information Criterion (AIC) have shown that models that include behavior can explain the data more effectively compared to models that neglect behavior, with little or no parsimony penalty [390,398,399]. Such models have also shown retrospective predictive power [390]. Other model fitting and data analysis exercises without mechanistic representations of behavior have likewise shown the importance of accounting for behavior when explaining observed epidemic patterns [400,401]. Similarly, the results of experimental games and surveys have been useful in the parametrization of coupled behavior–disease models [402,403]. The advent of online social media promises another source of data on both individual-level and population-level human behavior and this will be explored in greater depth in a later section.

Hence, there appears to be a strong public health and scientific rationale for studying behavior–disease interactions. The existing research shows that behavior and disease are strongly coupled to another. This coupling is particularly important near the elimination threshold, hence vaccine refusal may become an increasingly important barrier to achieving elimination and eradication, as logistic barriers recede into the distance. Moreover, models show promise to help us better understand coupled behavior–disease dynamics, and comparing models with data have often shown that behavior is important for understanding observed epidemic patterns and that behavioral models may have predictive power. Finally, these models can show interesting dynamics that are similar to those studied in physics – such as phase transitions, oscillatory solutions, and spatio-temporal patterns – and therefore similar methodologies to those used in physics may be helpful in their study. In the next section, we will introduce some of the basic concepts of behavioral modeling that have been applied to study coupled behavior–disease interactions.

7. Basic concepts in behavioral modeling

Much of the recent research on modeling the coupled dynamics of vaccinating behavior and disease dynamics has been conducted by investigators with training in the natural sciences, including physics, applied mathematics and epidemiology. These investigators have borrowed concepts from the social sciences in order to find behavioral models to couple with the existing infection transmission models they are more familiar with. These behavioral models can be broken down into three main categories: phenomenological models, game theoretical models, and psychological models. However, we note that the distinctions between these categories are not always well defined, and the delineations depends to some extent on subjective opinions. We also note that our goal here is to characterize different ways of doing behavioral modeling, rather than different ways of modeling coupled behavior–disease interactions *per se* (see Refs. [404,405] for classifications of coupled behavior–disease models).

7.1. Phenomenological models

Phenomenological approaches to modeling behavior mathematically describe the observed effects without positing mechanisms behind the observed behavior. For instance, arguably the first model to capture the impact of adaptive human

behavior on disease transmission used the mean-field ordinary differential equations [406]

$$\frac{dS}{dt} = -g(I)S,$$
$$\frac{dI}{dt} = g(I)S, -\gamma I \qquad (153)$$
$$\frac{dR}{dt} = \gamma I,$$

where S is the number of susceptible individuals, I is the number of infectious individuals, R is the number of recovered individuals, γ is the per capita recovery rate, and $g(I)$ is the force of infection. The form of $g(I)$ was chosen to reflect the phenomenological impact of unspecified psychological effects. One of the functional forms explored in Ref. [406] included

$$g(I) = \frac{\beta I}{1 + \beta \delta I}, \qquad (154)$$

which saturates at high prevalence of infection, I. Such saturation of the transmission rate at high levels of prevalence could occur if individuals reacted to high prevalence by reducing their contact rate through hand-washing or other means, which is a plausible assumption for a sufficiently dangerous infectious disease. We describe this model as phenomenological because, although it is motivated by psychological factors, the equations only describe the impact of such psychological factors on the transmission rate without actually modeling specific psychological processes such as individual cognition or social learning.

Similarly, as noted in the previous section, a rapidly growing body of research uses the contagion metaphor to describe the transmission of ideas and behaviors through social networks [394,395]. We classify these approaches as phenomenological because they describe the effect of behaviors and ideas being transmitted between individuals, without positing the mechanisms that might bring about this transfer. Contagion of ideas is a metaphor, and caution must be exercised when applying metaphors to complex social phenomena. Vaccine scare and others panics undoubtedly share common features with contagious processes, and using this metaphor can simplify and facilitate analysis of social processes. Moreover, the contagion metaphor comes naturally to investigators experienced in modeling contagious diseases. However, the downside of such metaphor is that it runs the risk of being too facile for some applications, by ignoring important subtleties of real-world social behavior [407]. For instance, some social science literature distinguishes between descriptive and injunctive social norms [408]. Descriptive social norms are those which describe what other individuals do, whereas injunctive social norms describe what other individuals approve or disapprove [408]. It is moreover possible to distinguish between social learning and social norms. Social learning theory emphasizes the social aspects of cognitive processes: individuals do not learn many behaviors by cogitating in isolation from others, but rather they learn through a more efficient process of imitating others [409]. An individual may use social learning to choose among descriptive or injunctive social norms that seem to be most acceptable, or most successful. It is not immediately clear how the contagion metaphor applies to these nuanced distinctions of how we learn ideas and behaviors from others. Indeed, empirical research seems to indicate that the contagion metaphor (unsurprisingly) is not universally applicable [407]. Moreover, many of the usual criticisms of memetics also apply to using the contagion metaphor for the spread of ideas [410].

However, in many contexts, these limitations of the contagion metaphor may not matter. Science works by adopting the simplest explanation that captures the observed behavior (Occam's Razor) and there are some contexts in which the contagion metaphor may be perfectly adequate to explain the observations. This is often the value of phenomenological models over mechanistic models. For instance, to prove that behavior influences the spread of infectious diseases, it is sufficient to construct a phenomenological model that captures the impact of behavior and to show that it is superior to models that do not capture behavior [390,398]. And, precisely because phenomenological models are often simpler in structure, they may make it easier to understand how simple mechanisms such as information spread through networks can give rise to complex emergent phenomena, such as control of infection spread in a spatially structured population [411]. In other contexts, mechanistic models may be desirable, and in the next two subsections we discuss two broad categories of mechanistic models: game theoretical models and psychological models.

7.2. Game theoretical models

Game theory is the formalization of strategic interactions between individuals in a group of two or more individuals [412]. Game theory is well exemplified by the Prisoner's Dilemma game (so called because it was originally explained as a decision process made by two prisoners being interrogated in separate jail cells who must decide whether to confess their crime, or stick to their story of being innocent). Each player can choose to either cooperate ('C') or defect ('D'), and receives payoffs depending on what s/he chooses, and what her/his opponent chooses (hence, the strategic interaction of the game). If you, (the focal player) cooperates and your opponent also cooperates, then you receive \$3 for instance (Fig. 19). But if you cooperate and your opponent defects, you get no money. On the other hand, if you defect and your opponent cooperates, you get a huge \$10 payoff, but if your opponent also chooses to defect, you only get \$1. Your opponent has the same payoff outcomes, and therefore the payoff matrix is called symmetric.

Opponent's Strategy

		C	D
Focal Player's Strategy	C	$3	$0
	D	$10	$1

Fig. 19. Payoff matrix describing the Prisoner's Dilemma, a classic example from game theory. Players must choose whether to Cooperate (C) or Defect (D) against their opponent, with the payoffs for taking these two different courses of action also dependent on what their opponent chooses. The money values are the payoff to the focal player, and the payoffs for the opponent are symmetrical hence payoffs for the opponent are the same, if they are treated as the focal player instead.

What should you – the focal player – do in this situation? If your opponent cooperates, then it is optimal for you to defect since you get $10 instead of $3. If your opponent defects, then it is also optimal for you to defect, since you get $1 instead of $0. Therefore, regardless of what your opponent does, you should choose 'D', defect. Your opponent is thinking the same way, so s/he also chooses to defect. Therefore you both defect and you get a payoff of $1 each. This is a shame, since if you had both cooperated, you would each get $3, but the selfishly optimizing logic of maximizing personal gain in anticipation of what your opponent will do prevents this socially (Pareto) optimal outcome. The Defect–Defect outcome is known as a Nash equilibrium, since neither player has an incentive to unilaterally change his or her strategy away from Defect and therefore we expect the Nash equilibrium strategy choices to be stable.

The Prisoner's Dilemma has been used to illustrate the clash between what is optimal for the individual and what is optimal for the group in a variety of contexts, including vaccination [413–415]. In the case of high levels of vaccine coverage, non-vaccinators are the Defectors, who benefit from herd immunity generated by vaccinators without paying any real or perceived cost of vaccination, while vaccinators are the Cooperators. However, this analogy only applies at high levels of vaccine coverage, since at low levels of vaccine coverage it is optimal for individuals to accept at least some vaccination.

Game theoretical models of interactions between vaccinating behavior and disease dynamics typically define strategies such as Vaccinator or Non-vaccinator, define payoffs for those strategies, and attempt to solve for a Nash equilibrium [415–417] or the analogue from evolutionary biology, the evolutionarily stable state (ESS) [418]. The approaches may also attempt to determine whether the Nash equilibrium is convergently stable, which indicates whether players will evolve, over time, toward the Nash equilibrium. However, classical game theory may not be well suited to describing real-world vaccinating behavior. It assumes that individuals are selfishly rational optimizers, but in fact, effects like social learning, bounded rationality, and imperfect information are very important in vaccinating decisions [6,392], and this can impact predicted behavior [390].

Research has explored the statistical physics of games such as the Prisoner's Dilemma or the Snowdrift Game played out on lattices [419–422], subsequently expanded to networks [423–430]. Likewise, game theoretical concepts have been applied to vaccination decisions on networks [431–434]. However, the resemblance of this work with 'game theory' gradually fades, as investigators use increasingly computational approaches in their work (and less rigorous proof of Nash equilibria) as a result of including more realistic cognitive processes. Such cognitive processes better reflect the importance of social influences and social norms, and the reality of imperfect information which makes individuals use 'rules of thumb' to make decisions, rather than assuming individuals can predict the future. In the end, the only resemblance to 'game theory' may be the use of utility functions where individuals make decisions (either non-strategic, explicitly strategic, or implicitly strategic) based on how much utility (a measure of preference) they will receive for different actions such as vaccinating or not vaccinating [431–434]. Hence, these more complex game theoretically-inspired models blur the distinction between 'game theoretical' models and 'psychological' models, the latter of which is the focus of the next subsection.

7.3. Psychological models

Some models of the interaction between vaccinating choices and disease dynamics invoke psychological theory in order to formulate mechanisms underlying vaccinating decisions. An example of a model that uses social and psychological assumptions is the following SIR model modified to include mechanistic modeling of vaccinating behavior according to an evolutionary game theoretical approach [435,436]:

$$\frac{dS}{dt} = \mu(1-x) - \beta SI - \mu,$$

$$\frac{dI}{dt} = \beta SI - \gamma I - \mu S,$$

$$\frac{dx}{dt} = \hbar x(1-x)(-1 + \omega I),$$

(155)

where S is the proportion of susceptible individuals, I is the proportion of infectious individuals, x is the proportion of individuals who adopt the vaccinator strategy, μ is the per capita birth/death rate, β is the transmission rate, γ is the recovery rate, $\hbar$ is a rescaled parameter that essentially captures the rate of social learning, and ω is a rescaled parameter that essentially captures the relative perceived risk of the disease compared to the vaccine. In this model, individuals are not perfectly rational optimizers because they only switch strategies through social learning ($\hbar$), after they have interacted with individuals playing a different strategy and moreover, they make the 'rule of thumb' assumption that their probability of eventually being infected depends on the current proportion of infectious persons, I, rather than assuming that they could solve the SIR equations explicitly and figure out their chances of becoming eventually infected. This approach can be straightforwardly extended to spatial settings, where individuals imitate successful strategies of their network neighbors [432]. A similar approach has also been adopted in models of coupled human–environment systems [437,438].

Other approaches incorporate specific theories from psychology and social psychology in order to formulate mechanisms in coupled behavior–disease models. For example, random walk subjective expected utility (SEU) theory posits that individual decision-making can be modeled as a biased random walk in a one-dimensional space of possible states with two decision outcomes on either end of the space. Individuals move toward one decision or the other depending on the probability of certain outcomes of their decisions, and the subjective utility associated with those outcomes, as formulated through a decision tree [439]. Once a certain threshold has been crossed close to one end of the one dimensional space or the other, the individual has finally decided on a course of action, and takes it. This has been applied to vaccinating decisions for influenza on a contact network, to explore how decisions about influenza vaccination and decisions about contact precautions interact with one another through their mutual coupled influence on disease dynamics [440].

Another relevant theory from psychology is prospect theory, which mathematically formalizes the phenomenon whereby humans tend to over-estimate the probabilities of very rare events. Since vaccine adverse events and disease complications are rare events, the relevant of prospect theory to vaccinating decisions is obvious. Prospect theory has been applied in mean-field models of coupled vaccinating behavior and disease dynamics for pediatric infectious diseases, and has been found to significantly impact model predictions [441].

A third example is the application of Dempster–Shafer theory, which is a general statistical framework (closely related to Bayesian statistics) for dealing with decision-making under uncertainty [442]. This theory has been incorporated into models of coupled dynamics of vaccinating decisions and disease dynamics, where investigators have used the theory to explore how awareness of disease-related events and vaccine-related events spreads through a group of socially connected individuals [443].

8. Behavior-vaccination dynamics in well-mixed (mean-field) populations

The dynamics and control of vaccine-preventable infectious diseases under voluntary vaccination can be quite complex, because of the role of human decisions. These are in turn based on information and *rumors* gathered on: (i) the *perceived* risks of getting the disease; (ii) its spread; (iii) the *perceived* risks of being affected by vaccine-related side effects. This complexity becomes apparent in the modeling phase, due to the intrinsic difficulty of quantitatively representing human behavior, and is mirrored by the features of the resulting mean-field models, whose solutions are radically different from those of the basic model with mandatory vaccination illustrated in Section 3. Here we review the recent literature on the dynamic implications of the introduction of human behavior in mean-field SIR-like models for vaccine preventable infections under voluntary vaccination. The presentation broadly follows the categorization proposed in [444] (also refer to Section 7 for classification), contrasting *phenomenological*, or *behavior-implicit* models of vaccinating behavior, with *behavior-explicit* ones, such as *psychological* or *game-theoretical* models. However, this categorization is not always effective from the modeling viewpoint since, under certain conditions, the various approaches can become mathematically indistinguishable. More mathematically focused reviews can be found in [444–446].

8.1. Phenomenological SIR models with information-dependent vaccine uptake

A general phenomenological framework is represented by the SIR model with *information-dependent* vaccination introduced in [447–450]. The underlying rationale is that the possibly complex behavioral rules parents follow when deciding whether to vaccinate or not their children will often translate into simple relationships between the individual probability to vaccinate, i.e. the *vaccine demand*, and the gathered information on infection spread. These models therefore postulate relatively simple feedback rules whereby trends of infection (and of vaccine adverse effects, VAE) can affect the current vaccine uptake, eventually feeding-back on infection dynamics.

These ideas led to the following modified SIR model with vaccination by a perfect vaccine [447] (note that for basic demographic and epidemiological quantities we use the notations adopted in Section 3)

$$S' = \mu \left(1 - x(M)\right) - \mu S - \beta(t)SI, \tag{156}$$

$$I' = I(\beta(t) S - (\mu + \gamma)), \tag{157}$$

where the vaccine demand $x(.)$ is an increasing function of a new variable, M, which summarizes the information available on current and/or past disease trend, used e.g. to formulate expectations about future risks. Here we considered uniquely

vaccination of newborns, however models considering 'delayed' vaccination strategies targeted also to later ages, i.e. of the type

$$S' = \mu(1 - S) - \psi(M)S - \beta(t)SI, \tag{158}$$

where $\phi(M)$ is an increasing function of M, have a qualitative behavior that is similar to the one of the model (156)–(157) [447].

Focusing on the perceived risk of disease as the driving force of vaccination decision, M might be any continuous and increasing function g of the *current* infection prevalence or incidence. For linear g it would hold $M = h(t)\beta(t)SI$ (with $h(t) > 0$) or $M = k(t)I$ (with $k(t) > 0$). Namely, the latter expression defines the perceived risk of serious disease as the product of the perceived risk of infection, proportional to infection prevalence I, times a constant perceived risk of serious disease given infection.

More realistically M also depends on past values of state variables, due e.g. to time-delays in the reporting/acquisition of information, and mainly to the fact that parents might also include their experience of past disease trends, etc. In this case, M can be modeled as follows

$$M(t) = \int_{-\infty}^{t} g(I(\tau))K(t - \tau)d\tau, \tag{159}$$

where the *delaying kernel* K [451] is a probability density function. Besides the trivial case $K(t) = \delta(t)$ (yielding back the unlagged case), the literature has focused on kernels allowing reduction to ordinary differential equations [447–450,452], in particular on the *exponentially fading memory* kernel $K(t) = a\exp(-at)$, which describe an exponential decay of the "memory" of agents, with characteristic memory length $T = 1/a$ [451]. Another kernel investigated in [450] is the *acquisition-fading* kernel $K(t) = a_1 a_2(a_2 - a_1)^{-1}(e^{-a_1 t} - e^{-a_2 t})$. This unimodal kernel is null at $t = 0$, and it mimics not only the exponential decay of memory, but also an initial phase of information acquisition.

Finally, function x can be written as [447,449,450]:

$$x(M) = x_0 + x_1(M), \quad 0 < x_0 < 1, \tag{160}$$

where x_0 is a non-null baseline constant coverage, mirroring awareness of disease severity also in absence of spread, and $x_1(M)$ is a sufficiently regular increasing function, with $x(M) \leq 1$.

For the sake of simplicity, from now on we shall assume that the information index M only depends on (current or past values of) I, the so called *prevalence-dependent* case [444,453]. In the current case this assumption implies that epochs of increasing prevalence would rapidly promote an increase in vaccine uptake in children (and vice-versa).

Two main substantive questions arise about model of Eqs. (156)–(160), namely how prevalence-dependent vaccination affects infection control on the one hand, and, on the other hand, how it might affect the dynamical pattern of infection and vaccination, e.g. by triggering oscillations. From now on we always assume that in absence of vaccination the infection is endemic, i.e. $\mathcal{R}_0 > 1$.

8.1.1. Implications of information-dependent vaccination: elimination "mission impossible"

The basic SIR model with constant mandatory vaccination x at birth shows a simple threshold behavior governed by the *vaccine reproduction number* (VRN) $\mathcal{R}_V = (1 - x)\mathcal{R}_0$ such that if $\mathcal{R}_V \leq 1$ (i.e. $x \geq x_c$) then the infection can be eliminated, otherwise the disease remains endemic. This simple result no longer holds under prevalence-dependent vaccination which makes elimination impossible. This is a general fact holding regardless of the form of the information index M, of the memory kernel $K(.)$ and of the nature of the transmission rate $\beta(t)$. This may be inferred by investigating the local as well as the global stability of the disease-free equilibrium (DFE): $DFE = (1-x_0, 0, 0)$ of model in Eqs. (156)–(160), yielding the following global eradication condition

$$B = \frac{1 - x_0}{\mu + \gamma} \frac{1}{\theta} \int_0^{\theta} \beta(u)du \leq 1, \tag{161}$$

(where B represents the average VRN when the vaccine uptake is set to the baseline coverage x_0) while if $B > 1$ the DFE is unstable. For constant β the condition $B \leq 1$ is equivalent to the condition $x_0 > x_c$, in other words elimination can only be achieved if the baseline vaccine uptake lies above the elimination threshold. Actually, it is a documented fact that western countries have been able to maintain very high coverages – i.e. above the estimated elimination threshold – against e.g. polio and diphtheria for decades. This however has been possible under situations where these vaccinations were essentially mandatory. Whether similar results could be maintained under fully voluntary vaccination regimes seems to be quite remote, given that scenarios where prevalence is very small would also make the perceived reward of vaccination to vanish, so that people would start escaping from vaccination.

8.1.2. Memory-triggered oscillations

Given that infection elimination is ruled out, the question turns into the types of dynamics that might be triggered by prevalence-dependent vaccination. For $\mathcal{R}_0(1 - x_0) > 1$ and constant β, the system of Eq. (156)–(160) has a unique endemic equilibrium $EE = (S_e, I_e, M_e)$.

In the unlagged case the system becomes 2-dimensional and it is easy to prove that the endemic state is globally asymptotically stable (GAS), so that only damped oscillations are possible [447]. Instead, in the delayed case, stable oscillations – via Hopf bifurcation of the endemic state – already appear under the simplest pattern of delay, namely the exponentially fading memory. In particular, assuming as key parameter the inverse of the average memory/delay length $a = 1/T$, and taking for simplicity $g(I) = kI$, it holds that [447] if and only if

$$(\beta I_e + \mu)^2 - \beta I_e \mu k x_1'(M_e) + 2(\beta I_e + \mu)\sqrt{\beta I_e(\gamma + \mu)} < 0, \tag{162}$$

there is a range $[a_1, a_2]$ for a such that: (i) for $a \in (a_1, a_2)$ Yakubovich oscillations [454] occur via Hopf bifurcations at a_1 and a_2; (ii) for $a \notin [a_1, a_2]$ the endemic equilibrium EE is locally stable. The global stability of EE, suggested by numerical simulations, is nontrivial and it has been investigated in [455] by means of the Li–Muldowney geometric theorem, which extends the Dulac–Bendixon theorem.

Note that Eq. (162) shows that sustained oscillations can appear only if $x_1'(M_e)$ is sufficiently steep, that is only if the behavioral response of vaccine uptake to changes in the perceived risks is sufficiently intense. In this case oscillations occur in an appropriate intermediate window of the average delay $1/a$. These results continue to hold for more realistic patterns of delay, such as acquisition-fading and Erlangian kernels [154,447–450], for generic increasing $g(I)$ functions [449], in presence of disease-related mortality and a non-constant population [449], and when the incidence of VAE is taken as the key determinant of the vaccine demand [154]. Finally, as shown in [456], oscillations in behavior-depending vaccination models are also observed in absence of memory mechanisms but in presence of disease-intrinsic delays, such as in the SEIR model (the latent period being, of course, equivalent to an exponentially distributed delay in the onset of infectivity).

8.1.3. Dynamic pattern for a measles-like infection

We illustrate some of the above described dynamic oscillatory patterns, under the parameter values adopted in Section 3.6. As for behavioral rules we choose $g(S, I) = I$, and a Michaelis–Menten-type [128] coverage function $x_1(M) = (1 - x_0 - \varepsilon)DM/(DM + 1)$ where $D > 0$ is a shape parameter tuning the reactivity of x_1, and $1 - x_0 - \varepsilon$ represents the *maximal coverage* arising for large DM values. Behavioral parameters are chosen in the oscillatory region, with the average delay $T = 1/a$ set to 4 months, $x_0 = 0.75$, $\varepsilon = 0.01$ and $D = 15 \times 10^4$. The selected values of ε and D allow, in principle, the vaccine uptake to reach values close to 100% during epochs of high perceived risk.

The model is initialized at time $t = 0$ at the endemic state determined by a constant coverage set at the baseline level x_0. Fig. 20 reports the transient (left-hand side) and long-term (right-hand side) time paths of the ERN $R_E(t) = \mathcal{R}_0 S(t)$ (top), of the infective prevalence (medium), and of the overall vaccine uptake $x(M) = x_0 + x_1(M)$ (bottom), jointly with its time average. State variables converge to a stable limit cycle in about 150 years, with a long-term inter-epidemic period around 12 years, i.e. about 2.5 times the pseudo-period of the SIR model with constant coverage $x_0 = 0.75$. The most striking fact is that though the vaccine uptake $x(M)$ reaches levels as high as 96% during epochs of high perceived risk, this is totally insufficient for elimination, as witnessed by the average long-term coverage, which remains below 80%, i.e. well below the critical coverage $x_c \approx 0.93$. This oscillatory nature appears to be intrinsic to SIR systems with delayed prevalence-dependent vaccine demand: epochs of increasing prevalence will yield, after a certain period of time, an increase in the *demand* for vaccines, which will in turn reduce infection prevalence, therefore eventually feeding-back on the vaccine uptake and so on, as first noted in [453].

8.1.4. Further extensions and remarks on phenomenological models

The information-dependent framework is a flexible one from several standpoints. As mentioned above, a model analogous to the one presented here has been applied to investigate the effects of VAE [154] which, in the current western situation where incidence of most infections is persistently low, appear to represent the most important component of the vaccine demand. In particular, in [154] the key determinant of the vaccine demand was the information on the incidence of VAE. Role of VAE as vaccination determinants will be considered more in detail in the *behavior-explicit* models of next subsection. However, many other applications are possible. For example the framework in [447] was also used to consider the possibility that parents adopt the strategy of delaying the age at vaccination of their children, later extended in a more general behavior-explicit setting in [457].

More important, the phenomenological model of Eqs. (156)–(159) represents a general form in that, as we have remarked, also *behavior-explicit* (e.g., game theoretic) formulations for vaccine uptake eventually collapse into the same structure, once the vaccine uptake function resulting from the adopted formulation is set within the framework of a dynamic epidemiological model [452,453]. For example, game-theoretic approaches (discussed in Section 8.3) often yield step-wise *best response* behavioral functions. The simplest example for the vaccination problem would be *not to vaccinate* with probability one when the risk of infection is below a certain threshold, *to vaccinate* above that threshold, and to vaccinate with a certain probability when the risk is exactly at threshold. However, a number of reasons (e.g., bounded rationality)
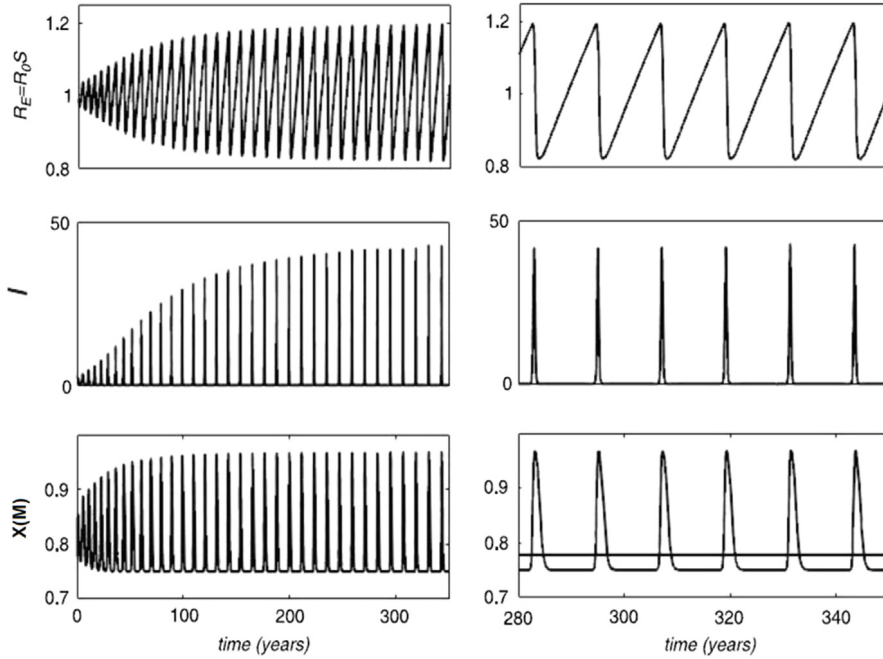
Fig. 20. Dynamics of model in Eqs. (156)–(160) for $g(S, I) = I$ under an exponentially fading kernel ($T = 4$ months) and a Michaelis–Menten-type prevalence-dependent component of vaccine uptake ($D = 500$). Transient (left-hand side) and long-term (right-hand side) paths of: (top) effective reproduction number $R_E = R_0 S$; (middle) infective prevalence I (normalized to its equilibrium value); and (bottom) overall vaccine uptake $x(M)$ (compared to its time average given by the flat line in the right-bottom graph). R_0 is equal to 15.
Source: Reprinted figure from Ref. [450].

suggest to approximate the best response function by a corresponding *smoothed responses* [458]. The shape of smooth response function is typically analogous to that of function $x_1(.)$ used here.

8.2. Behavior-explicit psychological models of vaccinating behavior

Most available psychological models of vaccinating behavior in previous sections relied on what we will term *Imitation Game Dynamics* (IGD), an important class of models developed within *Evolutionary Game Dynamics Theory* (EGDT) [459–461], which is a powerful approach to modeling behavior by setting ideas from game theory into a population dynamics framework [459–461]. *Imitation* represent a main avenue through which the different strategies enacted by players can spread within a population, through mutual learning between individuals [462]. The first application of IGD to investigate vaccinating behavior and its feedback on infection dynamics and control was developed in a seminal paper by Bauch [462], later extended [463] by a more general formulation which will prove useful to link the class of imitation-based models with the phenomenological models of previous section.

8.2.1. Imitation dynamics

IGD is usually presented using the framework of game theory, both in its economic and in its bioevolutionary standpoints. However, we note that at its very core there is again a mass-action diffusion process based on statistical physics [461]. Let us consider a binary decision problem, i.e. one for which only two, mutually exclusive, strategies are played: strategy V and N (these might for example be *to Vaccinate*, and *Not to Vaccinate*). Let us also assume that individuals can switch from the strategy they are currently playing to the alternative one by *learning and imitating* after social encounters with subjects that are playing the other strategy. Considering a population of constant size, and indicating with y_V and y_N the fractions of the N and V sub-populations ($y_V + y_N = 1$), leads to the following mean-field *double infection* model:

$$y'_V = -K_{VN} y_N y_V + K_{NV} y_N y_V, \tag{163}$$

$$y'_N = -K_{NV} y_N y_V + K_{VN} y_N y_V, \tag{164}$$

where the coefficients K_{NV}, K_{VN} have a clear "epidemiological" interpretations as the *strategy-specific* transmission rates following social contacts with individuals playing the other strategy. Following EGDT [461], these coefficients are proportional to the perceived benefit, or *payoff*, arising from strategy switching. Since $y_N = 1 - y_V$, it follows

$$y'_V = (K_{NV} - K_{VN}) y_V (1 - y_V). \tag{165}$$

Note that if $K_{NV} - K_{VN}$ is strictly positive (negative) in all the state space, strategy V (N) will eventually prevail, by spreading to the entire population with a logistic-type dynamics, as is intuitive given that the quantity $K_{NV} - K_{VN} > 0$ must be proportional to the perceived *net payoff gain* when switching from strategy N to V. More interesting, dynamic situations arise when the net payoff $K_{NV} - K_{VN}$ has non-constant sign, which happens when it is made dependent on other state variables of the system, as is the case for vaccination games [399,457,462–465].

8.2.2. Basic SIR models with imitation-based vaccinating behavior

Unlike previous subsection where the vaccine uptake at birth x was taken as a "static" function of I (and/or of S), we assume that x is a state variable determined by the switches of parents between strategy N (not to vaccinate) and V (vaccinate) according to the above described IGD. Therefore, $x(t)$ actually represents the proportion of parents favorable to vaccination at time t, which is taken as a proxy for the actual emerging vaccine uptake. Setting $y_V = x$, yields the following family of mean-field SIR models with vaccination [110,463]

$$S' = \mu(1-x) - \mu S - \beta SI, \tag{166}$$

$$I' = \beta SI - (\mu + \gamma)I, \tag{167}$$

$$x' = \kappa_1 \Delta E x (1-x), \tag{168}$$

where the term $K_{NV} - K_{VN}$ has been represented by the product between the scale coefficient κ_1, tuning the speed of the imitation process, and the net expected *payoff gain* of vaccination, ΔE. The latter can in turn be defined as the difference $e_V - e_N$ between the *vaccinator payoff* and the *non vaccinator payoff*, or alternatively as $C_V - C_N$ where $C_i = -e_i$ represents the corresponding *cost*.

Note preliminarily that irrespective of the specific form of $\Delta E(t)$, the family of models in Eqs. (166)–(168), always has the following three equilibria: (i) a disease-free state with no vaccinators $A = (1, 0, 0)$, which is unstable due to the above assumption that $\mathcal{R}_0 > 1$; (ii) a *pure-vaccinator* disease-free equilibrium $B = (0, 0, 1)$ where everyone is vaccinated; (iii) the pre-vaccination endemic equilibrium $C = \left(S_{SIR}, I_{SIR}, x_{SIR}\right) = (\mathcal{R}_0^{-1}, \mu(1 - \mathcal{R}_0^{-1})/(\mu + \gamma), 0)$. On the other hand, the stability of B and C as well as the existence of further equilibria (induced by the behavioral component) depend on the chosen form of the perceived payoff gain $\Delta E(t)$.

8.2.3. The baseline SIR model with imitation-based vaccinating behavior

The seminal study [462], motivated by need to understand the causes and dynamics of *vaccine scares*, such as the pertussis or the MMR vaccine scares, rests on two main hypotheses, namely: (a) a prevalence-dependent cost of not vaccinating (see Fig. 21), of the form $C_N = r_I mI(t)$, where $mI(t)$ is the current perceived risk of infection (taken as an estimate of the actual FOI $\lambda = \beta I$), and $r_I > 0$ is the perceived (conditional) risk of serious disease given infection; and (b) a constant cost of vaccinating $C_V = r_V > 0$, representing the perceived risk of suffering a VAE per single vaccination. Hence

$$\Delta E(t) = r_I mI(t) - r_V = r_V \left(\vartheta I(t) - 1\right), \tag{169}$$

where $\vartheta = mr_I/r_V$ is a measure of the *relative cost* of disease with respect to that of vaccination. Based on these hypotheses, a number of substantive results were shown [462]. First, the pure-vaccinator equilibrium B is unstable, reflecting the fact that in a wholly vaccinated population, prevalence is equal to zero and therefore the payoff of vaccination is strictly negative ($\Delta E = -r_V$), removing incentive to vaccinate, so that vaccination will start declining. Second, there is a *partial vaccinator* endemic state D, that is an endemic state with positive vaccine uptake $D = \left(\mathcal{R}_0^{-1}, \vartheta^{-1}, \widehat{x}\right)$, where $\widehat{x} = (1 + \gamma/\mu)\left(I_{SIR} - \vartheta^{-1}\right)$. The equilibrium coverage $\widehat{x}$ is meaningful only for values of the behavioral parameters ensuring that $I_D < I_{SIR}$, i.e. such that the corresponding endemic prevalence is lower than the prevalence observed in absence of any vaccination. This in turn requires $\vartheta > I_{SIR}^{-1}$, meaning that D exists only for sufficiently high levels of the relative cost of disease. Finally, at $\vartheta = \vartheta_0 = I_{SIR}^{-1}$ there is a transcritical bifurcation between C and D. In words, given that endemic infection for which at some stage a vaccine becomes available, then: (i) if the relative cost of the infection is low (i.e. the perceived cost of vaccination is high with respect to infection) then the vaccine would not be adopted due to lack of incentive; (ii) if the relative cost of infection is high then the vaccine would spread in the population, therefore destabilizing the pre-vaccination endemic state C, and sustaining the appearance of the new, locally stable, endemic state D with positive vaccination. In turn, the stability of D depends on $\kappa_1(\vartheta - \vartheta_0)$, showing in particular that appropriately large values of the imitation coefficient can yield sustained oscillations by Hopf bifurcation of D, indicating that fast social learning is a source of oscillations.

Previous results add robust evidence, based on a behavior explicit model, to those reported in Section 8.1, indicating that prevalence-dependent vaccinating behavior represents a serious threat for programs aimed to infection elimination.

8.2.4. An alternative model with myopic perception of the cost of vaccine adverse events

The hypothesis made in previous subsection that the perceived risk of vaccinating is constant, implies that the public correctly evaluates this risk, as the ratio between total reported VAEs for a given infection, and the total number of vaccinations for that infection per unit of time. However, due to the current high levels of herd immunity in modern industrialized countries, the perceived risk of serious disease from many infections is currently very low, while the large
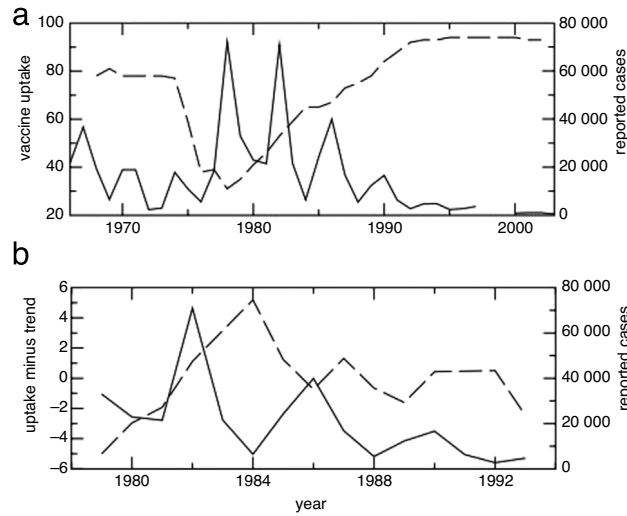
Fig. 21. Trends in pertussis vaccine uptake (dashed line) and case notifications (solid) in England and Wales 1967–2003 motivating the idea of prevalence-dependence in vaccine uptake: (a) absolute levels, (b) with vaccine uptake represented as deviations from a linear least-squares fit. Note in particular that when coverage went abruptly down due to vaccine scare (mid seventies), large epidemics restarted, possibly uprising the perceived risk of infection, and eventually restoring vaccine uptake at high levels. Nonetheless symptoms of prevalence-dependence persist, though at a lower scale, even after the overall relapse, as evidenced in (b). Moreover, the initial (pre-1975) persistently high coverage, bringing incidence to low and declining levels, might have sharply increased the number of VAE, eventually triggering the vaccine scare.
Source: Reprinted figure from Ref. [462].

number of vaccinations administered unavoidably yields a steady flow of VAEs [466]. The VAE burden might then make the perceived risk of suffering vaccine side effects (VSE) higher than risks perceived from infection.

Therefore, in [463] it was assumed that the public *myopically* evaluates the risk of VSEs by available information on the total number of VAE, which is proportional to actual vaccine uptake x. In particular, it was investigated the case $C_V = \alpha x$, defining the perceived cost of vaccination as the product of the (perceived) probability of being vaccinated, times a perceived probability of suffering VSE given vaccination. Similarly, C_N is taken as an increasing function of infection prevalence: $C_N = h_1(I(t))$. Note that the case $h_1(0) > 0$ is a realistic one as it might represent a scenario of local elimination where however infection re-emergence (e.g. by importation of cases) is feared.

The resulting dynamic equation for $x(t)$ is

$$x' = \kappa \left(h(I) - x \right) x \left(1 - x \right),$$ (170)

where $\kappa = \kappa_1 \alpha$, and $h_1(\cdot) = h(\cdot)/\alpha$.

Interestingly, if imitation dynamics is fast compared to infection dynamics, i.e. $\kappa \gg (\mu + \gamma)$, then a quasi-stead state approximation yields

$$x(t) \approx \min \left(h \left(I(t) \right), 1 \right),$$ (171)

i.e. the information-dependent model with current information (i.e. Eqs. (156)–(160)) presented in the previous subsection is recovered. In other words, the phenomenological model [447] is a special case of the imitation-based case when the imitation process is fast.

Comparing with the baseline model of previous Section 8.2.3, we note that the alternative model of Eqs. (166)–(170) also has the equilibria A (unstable), B (unstable) and C. Unlike Section 8.2.3 [462], the pre-vaccination endemic equilibrium C is always unstable, reflecting the fact that introduction of a vaccine will always be successful, due to the low perceived cost of vaccination at low coverage levels. Moreover, two further equilibria are induced by the inclusion of behavior i.e., a disease-free equilibrium E with positive vaccine uptake (*disease free partial vaccinator* equilibrium) $E = (1 - h(0), 0, h(0))$, and the partial vaccinator endemic state D

$$D = \left(\mathcal{R}_0^{-1}, I_e, h\left(I_e \right) \right),$$ (172)

where $\mathcal{R}_0 = \beta/(\mu + \gamma)$ is the basic reproduction number of the infection, and I_e is the unique solution of the equation

$$h(I) = 1 - \mathcal{R}_0^{-1} - \frac{\mu + \gamma}{\mu} I.$$ (173)

Eq. (171) alone can explain the behavior of the disease-free equilibrium E. Indeed, from $x' \geq \kappa (h(0) - x)x(1 - x)$ it follows that for large times $x(t) \geq \min(h(0), 1)$ and in turn: if $h(0) \geq x_c$ then $x(t) \to 1$, i.e. E is globally attractive. Moreover, linearizing at E it follows that if $h(0) < x_c$ then E is unstable. Therefore, in case of myopic evaluation of VAE
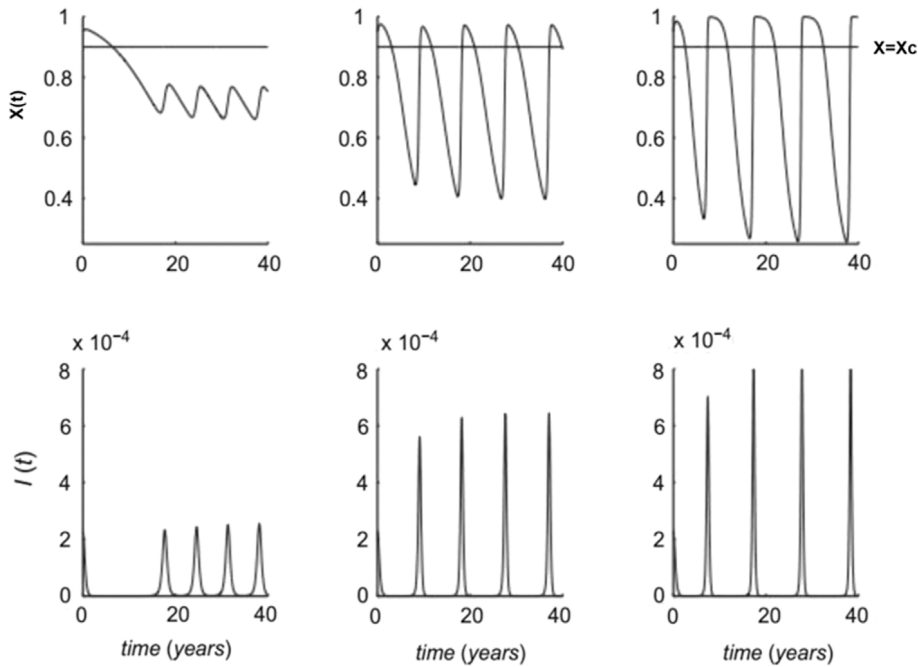
Fig. 22. Oscillatory dynamics of the SIR model with vaccination decisions governed by Eq. (170) for a linear $h(I)$ of the form: $h(I) = \theta I, \theta > 0$. The graphs report the dynamics of the vaccinating proportion $x(t)$ (top) and of infection prevalence $I(t)$ (bottom), for different values of the rescaled imitation coefficient κ: $\kappa = 0.0005$ day^{-1} (left), $\kappa = 0.002$ day^{-1} (center), $\kappa = 0.0035$ day^{-1} (right). The system is initialized from the pre-vaccination endemic levels of variables S, I and from $x(0) = 0.95$. Demographic and epidemiological parameters are: $\mu = (1/75)$ year^{-1}, $\gamma = 1/7$ day^{-1}, $\mathcal{R}_0 = 10$, $=15{,}000$, $\kappa = 0.002$ day^{-1}.
Source: Reprinted figure from Ref. [463]. With permission from Elsevier.

elimination turns out to be possible, but only if the perceived cost of disease associated with infection re-emergence is so large to bring the vaccine uptake in excess of the elimination threshold.

As all equilibria are independent of κ, it is natural to choose κ as a bifurcation parameter. Not surprisingly, the dynamics around the partial vaccinator endemic state D depends on the steepness of function $h(I)$ at D. Namely, it exists a threshold:

$$th = \frac{(\mu + \beta I_e)\left((\mu + \beta I_e) + 2\sqrt{\beta I_e(\mu + \gamma)}\right)}{\beta I_e \mu}$$

such that (i) if $h'(I_e) < th$ then equilibrium D is locally stable irrespective of the imitation speed κ; (ii) if $h'(I_e) > th$ then there is a range of values (κ_1, κ_2) for the imitation speed κ such that for $\kappa \in [\kappa_1, \kappa_2]$ and the system's orbits $x(t) = (S(t), I(t), x(t))$ are oscillatory (i.e. recurrent epidemics are predicted) in the sense of Yakubovich [454], which onset via Hopf bifurcations at κ_1 and at κ_2.

Note that unlike the model reported in Section 8.2.3 where sustained oscillations around the endemic state were triggered by large values of the imitation coefficient κ [462], in the present model oscillations occur in an intermediate window of values of κ. This suggests that both slow and fast imitation might be stabilizing forces, quite similarly to what is reported in Section 8.1.2 for the phenomenological model. This is not surprising since $x(t)$ depends on $I(t)$ via a differential equation, i.e. $x(t)$ "follows" $h(I(t))$ with an appropriate delay.

Possible dynamics of the system for some parametric constellations promoting sustained oscillations are illustrated in Fig. 22 for a measles-like infection. A noticeable fact is that large levels of the imitation coefficient κ, besides triggering very long inter-epidemic periods, can promote long-lasting epochs where the vaccine uptake x lies above the elimination threshold x_c.

Modeling irrational behaviors. A positive aspect of the imitation game-based seminal models [110,462,463] is that they parsimoniously describe, by means of a classical tool of game theory, the complex social learning underlying the parental decision to vaccinate their children. Moreover, broadly speaking, the imitation process represent a first deviation from full rational behavior. However, models [110,462,463] are based on an important but implicit assumption, i.e. that the "force" driving the vaccine non-acceptance is para-rational. Indeed either it is constant (and fully dictated by a rational thought, as in [462]) or it depends on the information on vaccine side effects (as in [110,463]). In reality in doing this the models [110,462,463] miss the fully irrational behavior underlying the periods of vaccine scares.

In [464] Bauch and Bhattacharyya go an important step further by assuming that the "anti-vaccination force" (which depends on the vaccine risk perception) is an empirical function mimicking the rumors-driven sudden rise, stay and decay of vaccine scare. As far as the "force" favoring the vaccine uptake (i.e. the feedback from disease prevalence) they assume

that it is proportional to the time-series of notified cases. The resulting dynamics for the fraction of vaccinated newborns $x(t)$ is thus of the type:

$$x' = \kappa x (1-x)(uL(t) - \omega(t)), \tag{174}$$

where $L(t)$ is an interpolation of notified cases, u is a proportionality constant and $\omega(t)$ is an appropriate unimodal function. Interestingly, the authors consider also other simplified models,

(a) absence of social learning:

$$x(t) = uL(t) - \omega(t); \tag{175}$$

(b) absence of disease driven feedback:

$$x(t) = 1 - \omega(t); \tag{176}$$

(c) absence of pro-vaccination "force":

$$x' = \kappa x (1-x)(-\omega(t)); \tag{177}$$

(d) explanatory model with force pro-vaccine depending on the state variable I (as in [462]):

$$x' = \kappa x (1-x)(I(t) - \omega(t)), \tag{178}$$

complemented by an SIR model with vital dynamics, of course.

They applied the various models, and various models of the function $\omega(t)$, to available data concerning the whole cells pertussis vaccine scare and to data on the MMR vaccine scare. These multiple comparison not only were done in order to find the best fitting and predicting model, but also in order to disentangle the effect of each component: the social learning, the kind of curve $\omega(t)$ modeling the vaccine scare; the type of disease-related feedback. They obtained the following interesting results: (i) the best fit to available data were obtained by the model of Eq. (178) including both the disease-related feedback and the social learning mechanism; (ii) this model also produced good predictions, up to 10 years in the case of pertussis data; (iii) the model of Eq. (174) performed better than the simplified models.

8.2.5. Delaying vaccination as a further strategy

A wide empirical evidence suggests that parents often adopt the strategy of delaying their children's *age at vaccination* against common vaccine preventable infections ([457] and references therein), mostly because they fear that the immune system of a newborn is still too weak to adequately tolerate immunization. Bhattacharyya and Bauch [457] proposed a minimal 3-strategy model, with two possible vaccination rounds, the first at birth (age 0), and the second one at a fixed subsequent age a_1 set to four years [457], always by a *perfect* vaccine. This is remindful of the two-rounds MMR (measles–mumps–rubella) vaccination, with an early first dose and a second one offered during the childhood to reach unvaccinated or non-successfully vaccinated children. Parents can opt between the following strategies: (i) *timely vaccinator* (V) i.e. vaccinate at the first round; (ii) *delayer* (*Del*) i.e. vaccinate at the second round; (ii) *Non vaccinators* (N). The three corresponding fractions of the population are denoted as x_V, x_{Del} and $x_N = 1 - x_V - x_{Del}$. The adopted model is a two-age classes meta-population version of the SIR endemic model, with exponential transitions between age groups and inhomogeneous age-specific mixing (see Section 3.9).

Individual learn and switch between strategies according to an IGD driven by the constant perceived risks of VAE from vaccination at ages 0 and a_1 respectively, denoted by $r_{V,1} < r_{V,2}$, and by the direct prevalence-dependent perceived risk of disease in the two age groups [462]. The latter depends on the corresponding risks η_i ($i = 1, 2$) of acquiring infection in age groups $i = 1, 2$, that are modeled as follows $\eta_i = \chi_i I_i / (\rho_1 + I_1)$, namely as increasing and saturating functions of the (perceived) group-specific infective prevalence I_i ($i = 1, 2$) to mirror two stylized facts, namely that (i) social contacts relevant for transmission are largely assortative i.e., tend to mostly occur with individuals of the same age group [147], and (ii) agents form their perceptions of risk according to simple heuristics. As for payoffs: (i) the timely vaccinator payoff only includes the vaccination cost at age zero, i.e. $C_V = -r_{V,1}$; (ii) the delayer payoff includes the sum of the cost of acquiring disease in period 1, plus the cost of vaccination at age 2, which arises only if he/she escaped infection during period 1, i.e. $C_{Del} = -r_I \eta_1 - r_V (1 - \eta_1)$, (iii) the non-vaccinator payoff is $C_N = -r_I \eta_1 - r_I \eta_2 (1 - \eta_1)$.

The resulting dynamics of state variables x_V and x_{Del} (the dynamics of x_N is given by $x_N = 1 - x_V - x_{Del}$) obey

$$x'_V = \kappa_v (\Delta E_{V,Del} x_V x_{Del} + x_V (1 - x_V - x_{Del}) \Delta E_{V,N}), \tag{179}$$

$$x'_{Del} = \kappa_{Del} (x_{Del} (1 - x_V - x_{Del}) \Delta E_{Del,N} - x_{Del} x_V \Delta E_{Del,V}), \tag{180}$$

where (similarly to the previously illustrated IGD-based models) the quantities $\Delta E_{i,j} = e_i - e_j = C_j - C_i$ ($i, j = V, Del, N$) represent the pairwise perceived payoff gain resulting from social encounters between individuals belonging to group i and group j, resulting from the difference of the related costs. Quite reasonably Eqs. (179)–(180) postulate a simplified mixing pattern between players of various strategies (compared to the mixing relevant for infection transmission), as indeed parents of children eligible for vaccination represent a restricted subgroup of the population in age group 2. The full model is obtained by adding to Eqs. (179)–(180) the equations for susceptible and infective individuals in age groups 1,2.

The qualitative behavior of the model is more complicated than the baseline model, although basic intuitions from Section 8.2.3 explain well the main results. For example, in addition to the unstable *infection free-all vaccinators* equilibrium, there is an unstable endemic *pure delayer* equilibrium with $I_2 = 0$. This further confirms the impossibility of disease elimination under prevalence-dependent cost of infection.

Several mutually exclusive other locally stable endemic states exist, including the pre-vaccination equilibrium and various *mixed* states with sub-optimal vaccine uptake combining different proportions of individuals following the different strategies. Interesting oscillatory regimens were observed in the simulations: individuals switch between strategies depending on the pattern of prevalence, with frequently observed anti-phase oscillations between strategies *V* and *Del*, e.g. epochs of high prevalence incentivate timely vaccination, which in turn reduce prevalence, thereby promoting the delayer attitude, which in turn favors a restart of prevalence in age group 1, and so on. However, inclusion of the third strategy has a stabilizing effect with respect to the basic (*V*–*N*) strategy of Section 8.2.3 [462]. The model also offers some insight on the role of the delayer strategy during epochs of vaccine scare.

8.2.6. Dynamical effects of social norms

Previous models assuming *prevalence-dependent* vaccinating behavior predict that elimination is impossible since vaccine uptake eventually sets around some equilibrium levels below the critical threshold x_c. However, in western countries vaccination coverage for certain diseases such as poliomyelitis and diphtheria attained and maintained very high levels, ensuring persistent local elimination, even in countries with voluntary vaccination. This suggests that further behavioral mechanisms are at work. A number of psychological theories relevant for health decisions such as the theory of planned behavior [467] and its extensions, have stressed the importance of social pressure as a determinant of individual behavior. In particular, our own perceptions of what persons relevant to us might think we ought to do are relevant—the so-called *injunctive social norms* (ISNs) [468] might affect vaccinating behavior in a number of ways, i.e. by building self-legitimation for non-vaccinators, by enforcing their cohesion, or by increasing the *conformism* of vaccinators as a majority.

Oraby et al. [399] included ISNs into the IGD model of Section 8.2.3 by assuming, in line with the game-theoretical work by [469], that ISNs add a new component to payoffs which is proportional to *group pressure*, i.e. to how many other people in the population are also playing that strategy. The perceived payoffs for nonvaccinators and vaccinators become: $C_N = -r_i m I(t) + \delta_0 (1-x)$, and $C_V = -r_V + \delta_0 x$, where δ_0 tunes the effects of social pressure on payoffs, taken as symmetric. Note that social pressure has the potential to mitigate, or even to sign-reverse, the costs of the two strategies.

After some rescaling, this yields the following equation for the vaccinators proportion x

$$x' = \kappa_1 x (1-x)(I - \omega - \delta(1-2x)), \qquad (181)$$

which, differently from Section 8.2.3, emphasizes the relative cost of vaccination and the relative strength of social pressure with respect to the cost of disease, given by $\omega = r_V/(r_i m I)$ and $\delta = \delta_0/(r_i m I)$ respectively.

With respect to the model in Section 8.2.3 considering only the direct costs of vaccination or not, the inclusion of ISNs yields a far richer dynamics. First, besides the four basic equilibria A, B, C, D observed in Section 8.2.3, there is a fifth equilibrium E induced by social pressure, where $S_E = 1 - x_E$, $I_E = 0$, $x_E = 2^{-1}(1 + \omega/\delta)$, with $x_E > 1/2$. However, this *disease free partial vaccinator* equilibrium, which exists only if $\delta < \omega$ (i.e. if the strength of group pressure is smaller than the relative cost of vaccination) is always unstable. Second, the pre-vaccination endemic state $C = (S_{SIR}, I_{SIR}, x_{SIR} = 0)$ is locally stable if $I_{SIR} < \omega + \delta$, i.e. if the perceived cost of disease is small. Third, if $\delta > \omega$ ($\delta < \omega$), i.e. if the strength of group pressure is larger (smaller) than the cost of vaccination, then the pure vaccinator equilibrium B is locally stable (unstable). As a consequence of the above two stability conditions, both B and C can coexist in their stability ranges, i.e. the system exhibits bistability and its mathematical behavior will depend on the initial conditions (see Fig. 23). Bistability phenomena also involve the *partial vaccinator* endemic state D (see Fig. 23). Finally, oscillations due to destabilization of D are possible, as in Section 8.2.3.

Overall, including ISNs widely expands the spectrum of qualitative mathematical behaviors w.r.t. the baseline case where payoffs only include direct costs of the alternative strategies. In particular, ISNs can mitigate the main perverse consequence of prevalence-dependence, namely the "elimination impossible" result, by allowing the pure vaccinator equilibrium to become stable at large levels of social pressure. Other effects are more controversial, as social pressures can induce bistability phenomena in a large portion of the parameter space, with the well-known amplification of uncertainty due to both the imprecise measurement of initial state and the equilibrium switch that can be induced by extrinsic stochasticity. Thus, depending on the context, social norms can either support or hinder vaccination goals. Finally, inclusion of ISNs allows, under a few further assumptions, an excellent fit of prevalence and coverage data of pertussis in the UK since the seventies, including the epoch of the pertussis vaccine scare and the subsequent epoch of coverage relapse (see Fig. 24).

8.2.7. Implications of bounded rationality in vaccination decisions

Vaccinating (or not) is a typical example of *decision under uncertainty* conditions, for which bounded rationality approaches, such as Kahneman and Tversky *prospect theory* [470] (see Section 7.3) typically work better than classical rational calculus. Oraby and Bauch [465] have investigated the dynamic implications of bounded rationality in vaccinating behavior by an SIR model in which vaccination decisions are formulated according to the ideas of prospect theory. The chosen coupled behavior–disease model is formulated as an IGD, i.e. parents still learn about other's strategies through imitation,
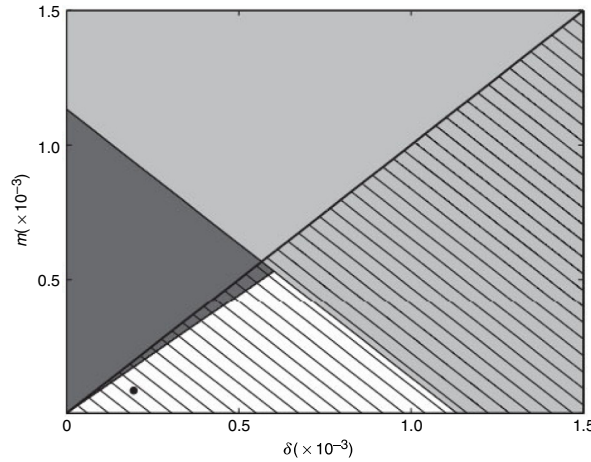
Fig. 23. Bistability of equilibria in the model with imitation dynamics and ISNs: regions of the (δ, m) parameters plane for the stability of the equilibria B (disease-free full vaccinator equilibrium), C (endemic pre-vaccination equilibrium), and D (partial vaccinator endemic state). Equilibrium B is stable in the hatched region below the thick black diagonal line, C is stable in the light gray region, D is stable in the dark gray region on the left. Details on other behavioral parameters and basic demographic and epidemiologic parameters are reported in [399].
Source: Reprinted figure from Ref. [399].

also including the effects of ISNs (see the previous subsection) and a non-perfect vaccine. The x-equation for the proportion of parents choosing strategy V is formulated as

$$x' = \kappa_1 x(1-x)F(\pi_{V|N}(I, x)) - F(\pi_{N|V}(I, x)), \qquad (182)$$

where $F(.)$ is a suitable probability distribution function, and $\pi_{V|N}(I, x)$ (alt. $\pi_{N|V}(I, x)$) is the net reward that a parent playing strategy N (alt. V) perceives to gain from switching to strategy V (alt. N), given current conditions on prevalence and vaccine uptake I, x. The map F transforms net rewards to a probability scale, so that the term $\Delta_F = F(\pi_{V|N}(I, x)) - F(\pi_{N|V}(I, x))$, replacing the *sure* perceived payoff gain of previous models, represents the *probability gain* arising from the switch between strategies. The perceived net rewards $\pi_{V|N}(I, x), \pi_{N|V}(I, x)$ are defined as

$$\pi_{V|N}(I, x) = \pi_N(V) - \pi_N(N) + (\delta_V x - \delta_N(1-x)), \qquad (183)$$

$$\pi_{N|V}(I, x) = \pi_V(N) - \pi_V(V) + (\delta_N(1-x) - \delta_V x). \qquad (184)$$

In Eqs. (183)–(184), besides the social norm components (analogous to those described in the previous section but taken asymmetric for sake of generality), the true core is represented by the differences $\Delta\pi_{V|N} = \pi_N(V) - \pi_N(N)$, and $\Delta\pi_{N|V} = \pi_V(N) - \pi_V(V)$, where the quantities $\pi_N(V), \pi_N(N), \pi_V(N), \pi_V(V)$ represent the expected utilities arising from each of the four possible *prospects* faced by agents. In [465] it is hypothesized that agents classify the possible realizations of infection in a four item scale $r_{i,I}$ ($i = 1, \ldots, 4$) i.e. (i) *mild* infection, (ii) *moderate*, (iii) *morbid*, up to (iv) *death* due to disease/vaccine adverse events (VAE), scaled to the base-case of absence of infection, with corresponding probabilities $p(r_{i,I})$. Analogous classification $r_{i,VAE}$ ($i = 1, \ldots, 4$) is done for VAE. At each realization it is attached a utility evaluation $U_I(r_{i,I})$ as well as a weight $W(x(r_{i,I}))$ to the corresponding probability. The various prospects are defined consequently. For example, agents currently playing N who continue to play N can only experience the consequences of infection, so their prospect includes the utility outcomes $U_I(r_{i,I})$ with expected utility $\pi_N(N) = \sum U_I(r_{i,I})W(x(r_{i,I}))$. On the other hand agents currently playing N who consider to switch to V will include in their prospect both the utility outcomes from VAE, $U_{VAE}(r_{i,VAE})$, as well as those due to infection, as a consequence of vaccine imperfection, and so on. The key step then lies in assigning the utility evaluations and corresponding weights, according to the principles of bounded rationality, i.e. accounting for the two different possible parents' attitudes, namely *risk averting* vs *risk seeking*, that classically cause departures from full rationality, depending on a number of parameters reflecting *boundedly rational* behavior. This is done by appropriate parametric specifications of U and W [465] including *perfect rationality* as a sub-case.

The resulting SIR vaccination model based on Eqs. (182) and (183)–(184) is quite richer compared to Section 8.2.6 as it includes, besides bounded rationality of vaccination decisions, also vaccine imperfection and asymmetric effects of social pressure. In particular, the underlying perfect rationality model is the extension of the model of previous section accounting for vaccine imperfection and asymmetric social pressure.

The ensuing equilibrium structure differs from that in 8.2.6 due to the appearance of a sixth endemic state F, which is a *pure vaccinator* endemic state forced by vaccine imperfection. In particular (i) the *disease free partial vaccinator* equilibrium E remains always unstable, as in 8.2.6, (ii) the pure vaccinator, disease-free equilibrium B is stable for infections that are not highly transmissible, if the vaccinators social pressure is relatively strong compared to the vaccine cost, while the *pure vaccinator endemic state F* is stable if and only if vaccinator pressure is large and the vaccine efficacy is low.
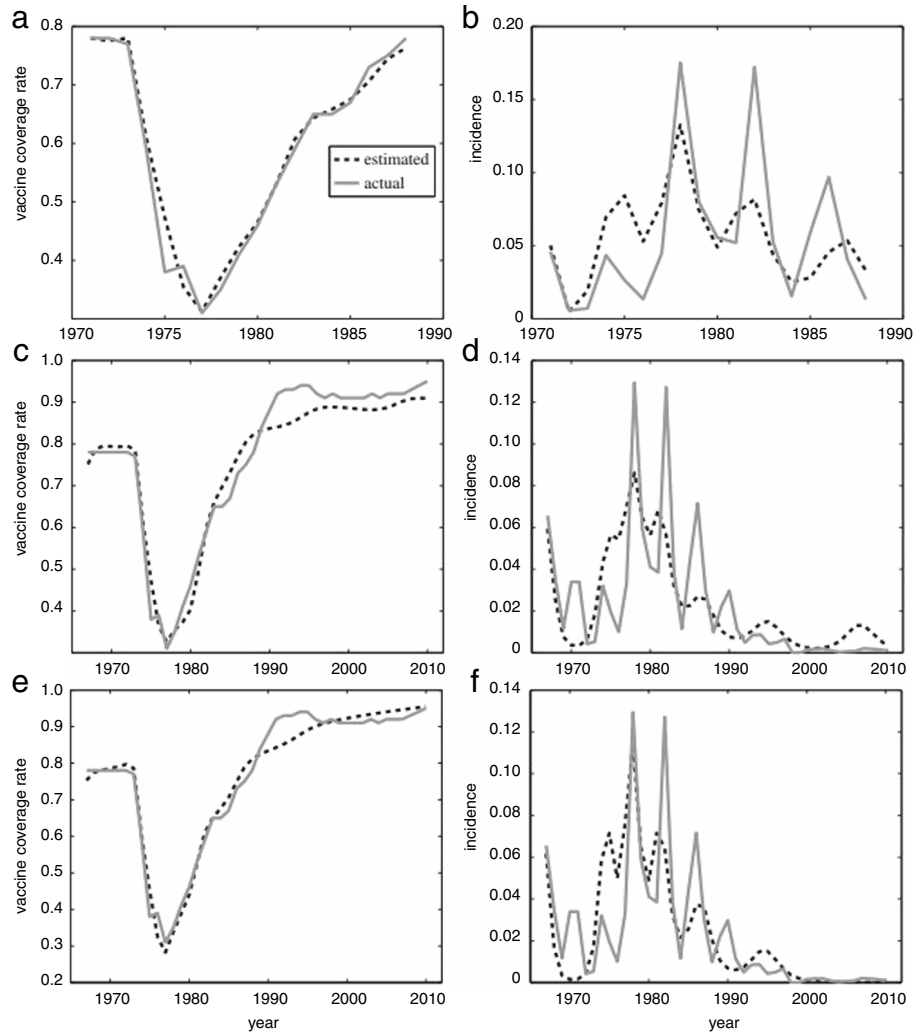
Fig. 24. Empirical and modeled (a, c, e) pertussis vaccine coverage and (b, d, f) pertussis incidence in the UK, (a, b) from 1971 to 1988 without injunctive social norms, (c, d) from 1967 to 2010 without injunctive social norms and (e, f) from 1967 to 2010 with injunctive social norms. The solid line is the empirical data and the dashed line is the best-fitting model.
Source: Reprinted figure from Ref. [399].

To sum up, the effects of bounded rationality are mostly dynamical. Indeed, while the dynamics observed in the underlying rational model are much alike those in Section 8.2.6, with bistability regions having simple forms, the inclusion of bounded rationality yields a zoo of dramatic dynamical changes, which include complicate nonlinear shapes of the stability regions, and steady oscillations about the pure vaccinator disease-free equilibrium *B*. In addition, boundedly rational behavior makes it harder to eliminate infection, because it can offset the beneficial effects of the social pressure of the vaccinators group.

8.2.8. Modeling public communication on disease and vaccine risks

A parsimonious explanation of some observed features of vaccination systems can be obtained by removing an unrealistic feature of models based on IGD, namely that switching between strategies is entirely determined by learning from other agents (i.e., parents of children to be vaccinated) during social encounters. Indeed, the pure-imitation model poorly represents the process of information provision in modern countries because, even when a given vaccination is fully voluntary. Public health systems will nonetheless keep the role of main suppliers of the relevant information on risks associated to diseases and vaccines, as suggested by a body of evidence [110]. In this subsection we report an extension of the basic imitation-based framework of Sections 8.2.3–8.2.4, which accounts also for perceptions based on information (about risks) supplied by the public health systems [110]. In this new framework, remindful of Bass classic model for information diffusion [471], switches between *N* and *V* strategies is determined by the balance between private information, exchanged through inter-personal communication between parents of children eligible for vaccination during their social contacts,

and public information, communicated by the public health authorities through media and related channels. This allows to amend the equations for the dynamics of switch between the two strategies as follows

$$y'_V = G(t)y_N - K_{VN}y_Ny_V + k_{NV}y_NY_V, \tag{185}$$

$$y'_N = -G(t)y_N - k_{NV}y_NY_V + K_{VN}y_Ny_V, \tag{186}$$

where $G(t)$ represents the rate of transition from strategy N to strategy V induced by mediatic information (which is independent of the level of individuals' social activity), yielding

$$y'_V = G(t)(1 - y_V) + (K_{NV} - K_{VN})y_V(1 - y_V), \tag{187}$$

where $K_{NV} = \kappa\theta(I)$ and $K_{VN} = \kappa\alpha(x)$. In the simplest case one can assume that $G(t)$ is a strictly positive constant, i.e. $G(t) = k_G > 0$. The underlying idea is that the information provided by public health systems aims to convey a different perception of risks related to disease and vaccination compared to what might happen in private social encounters. In particular, a "wise" public communication aims to convince that (i) vaccines are highly safe, with a very low, constant, risk of VAE, and that (ii) risks associated with the disease are prevalence-independent.

Setting $x = y_V$ and simplifying Eq. (187) yields the following dynamic equation for the proportion favorable to strategy V

$$x' = \kappa(1 - x)\left((\theta(I) - \alpha(x))x + \psi\right), \tag{188}$$

where parameter $\psi = G/\kappa$ summarizes the effectiveness of the public *effort* (information, education, availability of vaccination infrastructures, subsides to vaccination staff, etc.) in affecting perceptions on vaccines and disease.

The properties of the ensuing behavior–disease model arising by coupling Eq. (188) with Eqs. (166)–(167) depend in a critical manner on parameter ψ. Briefly, there exist appropriate thresholds $\psi_c < \psi_1$ such that

- For high levels of public effort ($\psi \geq \psi_1$) only the pure vaccinator equilibrium B exists and it is globally asymptotically stable (GAS).
- For intermediate levels of public effort (say $\psi_c < \psi < \psi_1$) the pure vaccinator equilibrium is unstable, but there is a *disease free partial vaccinator* equilibrium E, with vaccine uptake above the elimination threshold, which is GAS.
- For low levels of public effort ($\psi \leq \psi_c$) the disease free partial vaccinator equilibrium E exists but is unstable, and the *partial vaccinator* endemic state D appears ensuring endemicity, though it is not necessarily stable i.e. oscillations may occur.

In a control perspective, if the infection is endemic at D in a scenario where vaccination is voluntary with public intervention absent, it is possible to increase the equilibrium coverage by increasing the public effort in providing information about benefits of vaccination. Suitable further increases in public effort can allow the equilibrium vaccine uptake to expand until the endemic state D disappears by exchanging its stability with the disease free partial vaccinator E, thus achieving elimination. Further increases in ψ can push E to collapse on the pure vaccinator equilibrium B.

These results indicate that public intervention can completely mitigate the negative implications of prevalence-dependent, pure-imitation models (Sections 8.2.3–8.2.4), particularly removing the impossibility to eliminate the infection. A simple calibration to Italian data on measles coverage shows that this model explains the recent Italian story, characterized by a fast increase in measles vaccine uptake following substantial efforts by the public health system, largely better than models including imitation only [110].

8.3. Behavior-explicit game-theoretic models of vaccinating behavior

The first application of classical game theory and epidemiological modeling to investigate the implications of vaccination decisions dealt with the conflict between group interest and self-interest in relation to smallpox vaccination upon reintroduction following a terroristic attack [472]. A companion paper [473] applied the same approach to mass vaccination against common vaccine-preventable childhood infections under voluntary vaccination. The latter work gave much momentum to subsequent research in the field. In line with the previous subsections, this subsection reviews this seminal work and a few subsequent fundamental contributions focusing on childhood vaccination [452,473–475]. The flavor of the presentation is slightly different however, as the focus of these approaches, though framed within SIR models for vaccine-preventable infections, is more on classic game theoretic questions, namely existence of Nash equilibria for individuals' vaccine uptake and comparisons with the level which is optimum for the community as a whole, rather than on dynamic properties.

8.3.1. The baseline equilibrium model

The work in [473] integrates ideas from the endemic SIR model with vaccination at birth (see Section 3.6) into a game theoretical framework to investigate the interplay between individual vaccination decisions and the population-level dynamics of infection in a voluntary vaccination scenario. Game theory is the natural framework given that the vaccination strategy of each specific agent (i.e. parents), whose sum defines the vaccine coverage at the population level, is influenced by other parents' decisions in the sense that the corresponding payoff will depend on the strategies played by all other players.
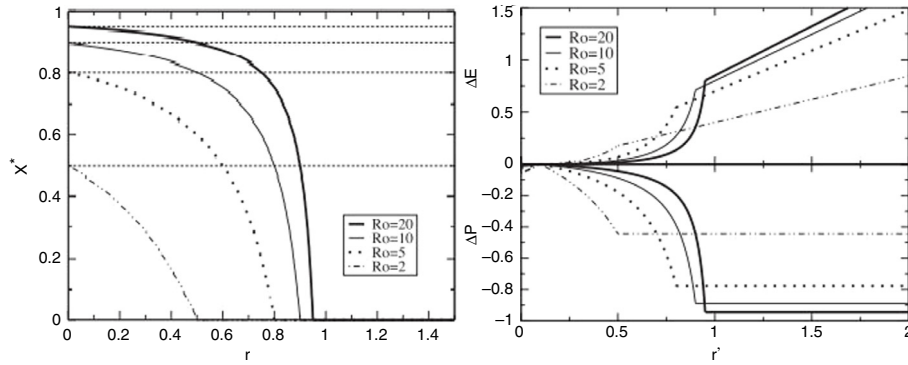
Fig. 25. The baseline equilibrium model of the vaccination game. Left panel: patterns of the vaccination coverage x^* predicted by the CSNE as a function of the relative risk $r = r_V/r_I$ of vaccination for different values of the BRN $\mathcal{R}_0$. The dashed horizontal lines drawn at the level of the critical coverage $x_c = x^*(r = 0)$ (for the different levels of $\mathcal{R}_0$) show the impossibility to eliminate the infection in a regime where individuals vaccinate according to self-interest driven by direct perceived risks. Right panel: impact of a vaccine scare as represented by a large increase of the relative risk of vaccination from a very low level $r^0 \ll x_c$ to a new level r', for different levels of the BRN $\mathcal{R}_0$. The graph reports the trend of the payoff gain ΔE and of the change ΔP in vaccine uptake for different levels of r'. For moderate levels of r', namely $r' < x_c$, the impact is moderate but becomes dramatic when r' exceeds x_c causing the CSNE coverage dropping to zero.
Source: Reprinted figure from Ref. [473].

Letting the individual strategy be represented by the parents' probability to vaccinate x, authors then seek a *convergently stable Nash equilibrium* (CSNE) solution x^*, which is expected to be observed in a real population of players.

The key determinants of individual strategies are the perceived *relative cost of vaccination* r, defined as the ratio between the perceived risk of VAE (r_V) and the perceived risk of disease given infection (r_I), and the probability π_x of experiencing infection under a given vaccination coverage x. For the latter it is assumed that parents correctly know the underlying epidemic model and estimate π_x by the corresponding lifetime probability of infection at endemic equilibrium for an unvaccinated individual, given by

$$\pi_x = 1 - \frac{1}{(1-x)\mathcal{R}_0}. \tag{189}$$

Note that the scenario of full absence of vaccinators ($x = 0$) yields $\pi_0 = x_c$.

The authors show that there is always a unique CSNE solution x^*. For high levels of the relative perceived risk of vaccination, namely for $r > \pi_0$, i.e. for $r > x_c$, the CNSE corresponds to the situation where no one vaccinates ($x^* = 0$), i.e. the pure non-vaccinator strategy. On the other hand, if $r < \pi_0$, i.e. $r < x_c$, the CNSE solution predicts a strictly positive probability to vaccinate x^*. The condition $r < \pi_0$ implies

$$(1 - r)\mathcal{R}_0 < 1. \tag{190}$$

Under the above constraint, the CSNE is

$$x^* = 1 - \frac{1}{(1-r)\mathcal{R}_0}, \tag{191}$$

which is such that

$$x^*(r) < x_c \tag{192}$$

(unless $r = 0$). In other words, under voluntary vaccination elimination turns out to be impossible if individual only take their self-interest into account (Fig. 25, left panel). Moreover, the case where r exceeds the elimination threshold ($r > x_c$), a fact possibly observable during periods of vaccine scare, the CNSE predicts absence of any vaccinator at all. A key point stressed in [473] is that when the perceived relative risk switches from an initial level $r^0 < x_c$ to another one r' greater than x_c, then the resulting payoff gain is increasing with r' and it has no plateau. In other words, the vaccine scare can go unlimited. In [473] it is also considered the effect of measures induced by the public health systems to counteract the vaccine scare, i.e. a reduction of the relative risk from a value in excess of x_c down to another value below x_c. The authors obtained that the incentive to vaccinate diminishes as the vaccine coverage approaches to its Nash equilibrium. Summarizing, a vaccine scare easily induces a substantial drop in vaccination, whereas restoring the pre-scare coverage levels is difficult. These findings are illustrated in the right panel of Fig. 25.

8.3.2. Best individual vs community vaccination response in a heterogeneous population

In [452] a general game-theoretical approach was adopted, aimed to relate the population vaccine demand to individual decisions in a behaviorally heterogeneous population sharing a common information 'signal' about the dynamic spread of an SIR-type infection. In their model the individuals choose between strategy *V* (vaccinate their children) and *N* (not

to vaccinate) based on perceived benefits and costs of vaccinating, in turn based on the information signal σ carrying information about risks of acquiring the infection. This information signal has the same meaning (and form) of the information index M in the phenomenological models of Section 8.1. The population is assumed to be behaviorally heterogeneous, and its internal structure is described by a single real variable y. Individuals having type y choose strategy V (to vaccinate) with probability $\chi(y, \sigma)$ and strategy N with probability $1 - \chi(y, \sigma)$. Benefits of vaccination are summarized by the utility function of strategy V which depends on σ and y: $U_V(y, \sigma)$, and is increasing in σ. On the other hand, the utility of vaccine refusal (strategy N) is assumed to also depend on the population average vaccine uptake

$$\bar{\chi} = \int_y \chi(y, \sigma) dy, \tag{193}$$

i.e. $U_N(y, \sigma, \bar{\chi})$, where U_N is decreasing in $\bar{\chi}$, mirroring a free-riding effect (compare this assumption with Sections 8.2.4–8.2.8 where the perceived cost of vaccination was increasing in the average vaccine uptake).

By taking the individual's expected utility

$$U(y, \sigma, \bar{\chi}) = \chi(y, \sigma) U_V(y, \sigma) + (1 - \chi(y, \sigma)) U_N(y, \sigma, \bar{\chi}), \tag{194}$$

one notes that the resulting individual's optimal strategy will necessarily depend on the strategies adopted by other individuals in the population i.e., the above problem defines a game.

The search for the *best response strategy* χ^* for the previous problem has to be framed into the theory of set-valued maps [476], yielding

$$\chi^*(y, \sigma, \bar{\chi}) = \begin{cases} 0 & \text{if } U_V(y, \sigma) < U_N(y, \sigma, \bar{\chi}), \\ 1 & \text{if } U_V(y, \sigma) > U_N(y, \sigma, \bar{\chi}), \\ [0, 1] & \text{if } U_V(y, \sigma) = U_N(y, \sigma, \bar{\chi}). \end{cases} \tag{195}$$

In order to completely characterize the Nash equilibrium, it is needed to fulfill a self-consistency condition concerning the average vaccine uptake at equilibrium:

$$\bar{\chi}^* \in \int \chi^*(y, \sigma, \bar{\chi}^*) dy, \tag{196}$$

i.e.

$$\int \inf \chi^*(y, \sigma, \bar{\chi}^*) dy < \bar{\chi}^* < \int \sup \chi^*(y, \sigma, \bar{\chi}^*) dy. \tag{197}$$

This condition ensures that there exists a Nash equilibrium and that the population's average response is the same for all Nash equilibria.

On the other hand, at the population level, a benevolent social planner will seek the optimal community strategy by maximizing the *community utility*

$$\widehat{U}(\chi(.)) = \int (U(y, \sigma, \bar{\chi}) \chi(y, \sigma) U_N(y, \sigma) + (1 - \chi(y; \sigma)) U_V(y, \sigma, \bar{\chi})) \, dy. \tag{198}$$

By elementary functional analysis it is easy to verify that the above individual-wise optimum χ^* and the community-wise optimum χ^C are such that for all y and given σ

$$\chi^*(y, \sigma) < \chi^C(y, \sigma). \tag{199}$$

In other words, the Nash equilibrium for individuals will always be characterized by "less vaccination" than the corresponding community-based equilibrium, unless the Nash equilibrium is universal vaccination, confirming in a very general manner a long-standing result of the literature, namely that the pursuit of optimal self-interest leads to vaccine uptakes that are systematically lower than those that are optimal at community level.

Note that both the optimal χ^* and its average value $\bar{\chi}^*$ depend implicitly on the information signal σ, which is a function of the state variables. For example in [452] it is assumed that σ is an increasing and saturating function of infection prevalence I: $\sigma(t) = \beta I / (\mu + \beta I)$. Note also that $\bar{\chi}^*$ results to be an increasing function of prevalence if the signal σ has this property.

The implications of the above described approach for infection transmission and control are then investigated by setting the resulting vaccination strategies within the framework of an SIR model with vaccination. Three main cases are considered: (a) the vaccine coverage x is equal to the average Nash equilibrium response $x(t) = \bar{\chi}^*(I)$; (b) the vaccine coverage adapts dynamically to the differences between its departures from the average Nash response, according to the equation $x'(t) = \alpha(\widehat{\chi}^*(I) - x)$; (c) the vaccine coverage adapts dynamically according to an IGD where the switch rate from N to V strategy is determined by the payoff gain of vaccination (which, as it can be easily seen, is constant) and the inverse switch is determined by the perceived vaccine cost (which depends on $\sigma(I)$). Unsurprisingly, model (a) (compare with the phenomenological model in Section 8.1) leads to a unique locally stable endemic equilibrium with partial vaccination. Similarly in model (b), where x 'follows' $\bar{\chi}^*$ with an exponential delay, the endemic equilibrium can be unstabilized by varying α, yielding sustained oscillations. This confirms the already stated fact that phenomenological and game-theoretic coupled behavior–infection models are often indistinguishable when the game-theoretic solution is simply set, heuristically, within the framework of a dynamic epidemiological model.

8.3.3. Markov decision processes show that the interaction between vaccine-believers and vaccine-skeptics yields sub-optimal measles vaccine uptake

Within game-theoretic approaches several efforts [474,475,477] have been devoted to the investigation of the population level implications, of individual choices determined within the framework of Markov decision processes theory (MDPT). An attempt of systematization of the subject in reference to vaccination games was developed in [474]. The framework rests on three interacting building blocks, namely a model for individual decisions, a model for changes at the individual scale ($\mathcal{I}$-scale), and a mean-field epidemiological model for describing the dynamics at the population scale ($\mathcal{P}$-scale). Individuals handle uncertainty by maximizing their (inter-temporal) expected utility, on the assumption that they are able to predict – in probabilistic terms – their future evolution in terms of both their future states $\mathcal{Y}(t)$ and the ensuing utility flow, that follow from taking a given strategy (i.e., vaccinating a child at a certain age, or not vaccinating him), or action, $A(t)$. Path $\mathcal{Y}(t)$ is a stochastic process which depends on the adopted strategy as well as on the strategies that have been adopted by other individuals, that is the average strategy at the population level. For example, for those not vaccinating the future utility path depends on the risk of catching infection which under homogeneous mixing depends on the overall coverage prevailing in the population. The $\mathcal{P}$-scale dynamics of infection is specified by an appropriate mean-field model, e.g. an SIR model, where the actual coverage is determined by the average of individuals' strategies. Finally, the $\mathcal{I}$-scale dynamics is specified by taking a suitable continuous-time Markov chain embedded in the $\mathcal{P}$-scale model. In [475] this framework was applied to investigate the consequences of a typical phenomenon in current vaccination systems, namely the coexistence of well-identified and stable groups valuing differently the costs arising from infection and vaccination.

The simplest scenario, introduced in the static game-theory framework in [478], is represented by the conflict between two groups, namely vaccine *skeptics* (group 1) and *believers* (group 2), with different asymmetric perceived costs of infection and VAE: $C_{V,1} > C_{V,2}, C_{I,1} < C_{I,2}$. Let q_n ($n = 1, 2, q_1 + q_2 = 1$) denote the relative size of each group (assumed to be constant), ϕ_n the vaccination *strategy* adopted by individuals in group n and $\overline{\phi}_n$ the average coverage prevailing in group n. The $\mathcal{P}$-scale model is represented by the following simple two groups' SIRV model with a class vaccinated at birth by a perfect vaccine, in a constant population, where the two groups differ only in the average coverage:

$$S_n' = \mu(1 - \overline{\phi}_n)q_n - \beta S_n(I_1 +_I 2), \tag{200}$$

$$I_n' = \beta S_n(I_1 + I_2) - (\gamma + \mu)I_n, \tag{201}$$

$$V_n' = \mu\overline{\phi}_n q_n - \mu V_n, \tag{202}$$

where S_n, I_n, V_n respectively represent the susceptible, infective, removed and vaccinated proportions in each group, with the removed compartment obeying $R_n = q_n - (S_n + I_n + V_n)$. The rates determining the various state transitions in Eqs. (200)–(202) determine the associated Markov chain for state transition at the $\mathcal{I}$-scale, under the further assumption that agents estimate the FOI faced by their children by its endemic equilibrium value $\lambda = \lambda_e = \beta(I_{1,e} + I_{2,e})$, which agents know from the knowledge of the $\mathcal{P}$-scale model. It holds $\lambda = \mu(\mathcal{R}_V - 1)$ where $\mathcal{R}_V = \mathcal{R}_0(1 - \overline{\phi}_1 q_1 - \overline{\phi}_2 q_2)$ is the vaccine reproduction number (Section 3.6), meaning that individuals correctly perceive the long-term risk as a function of the overall coverage prevailing in the community. The corresponding system for the individual's state probabilities has the form $\mathcal{Y}_n' = \mathcal{Q}_n\mathcal{Y}_n$, driven by the 4×4 transition matrix $\mathcal{Q}_n$:

$$\mathcal{Q}_n = \begin{pmatrix} -(\lambda + \mu) & 0 & 0 & 0 \\ 0 & \lambda & -(\gamma + \mu) & 0 \\ 0 & 0 & -\mu & 0 \\ 0 & 0 & 0 & -\mu \end{pmatrix}$$

with initial probability distribution, actually representing the probabilities that the generic individual occupies at the moment of birth any of the four possible epidemiological states, given by the vector $\mathcal{Y}_n(0) = [1 - \overline{\phi}_n, 0, \overline{\phi}_n, 0]$. Note that the transition matrix is the same for both groups, since perceptions about epidemiological and demographic hazards are assumed to be the same, so that difference in $\mathcal{I}$-scale dynamics is only due to the decision to vaccinate or not made at birth. Given the adopted individual strategy ϕ, and the ensuing $\mathcal{I}$-scale path $\mathcal{Y}_n(t)$, individuals belonging to group n have a (discounted) utility flow

$$U_n = F_n\mathcal{X}_n(0) + \int_0^\infty e^{-rt}f_n^T\mathcal{X}_n(t)dt, \tag{203}$$

where vector $f_n = [0, 0, -C_{V,n}, 0]^T$, ($n = 1, 2$) includes the costs associated with occupancy of various states at any time, while $F_n = [0, -C_{I,n}, 0, 0]^T$ includes the costs associated with transitions occurring at age 0, which occur only in the event of vaccination. Integrating Eq. (203) yields $U_n = (F_n^T + f_n^T(rI - Q_n))\mathcal{X}_n(0)$, which expands as

$$U_n = -\phi_n C_{V,n} - (1 - \phi_n)C_{I,n}(r + \mu + \gamma)^{-1}\frac{\lambda}{r + \mu + \lambda}. \tag{204}$$

The mathematical analysis shows that the Nash strategy, i.e. the *best response curve* $\widehat{\phi}_n$ arising from the maximization of expected utility of Eq. (204), has the following form: (i) not to vaccinate ($\widehat{\phi}_n = 0$) for levels of the perceived (equilibrium) FOI λ strictly below a group-specific threshold $\lambda_{\mathcal{T},n}$, (ii) to always vaccinate ($\widehat{\phi}_n = 1$) for values of λ strictly above $\lambda_{\mathcal{T},n}$,
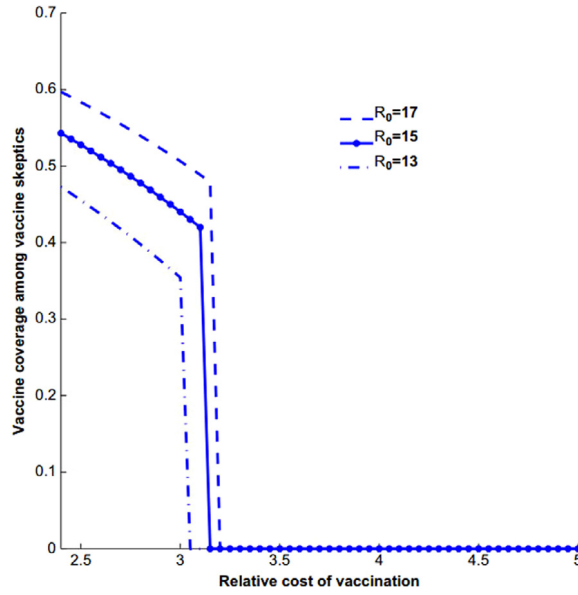
Fig. 26. Pattern of the best Nash response in terms of vaccine uptake among vaccine skeptics, drawn as a function of the cost of vaccination relative to the cost of infection ($C_{V,1}/C_{I,1}$) for different levels of the BRN $\mathcal{R}_0$ consistent with measles infection ($\mathcal{R}_0 = 13, 15, 17$), and for a proportion of skeptics ($q_1 = 0.30$), showing the dramatic fall of the optimal vaccine uptake when the relative perceived risk exceeds a threshold level.
Source: Reprinted figure from Ref. [475]. With permission from Elsevier.

and (iii) to choose some (unspecified) intermediate level $\widetilde{\phi}_n$, $0 < \widetilde{\phi}_n < 1$ for $\lambda = \lambda_{\mathcal{T},n}$. The threshold $\lambda_{\mathcal{T},n}$ depends on the group's perceived costs

$$\lambda_{\mathcal{T},n} = \frac{C_{V,n}(r + \mu + \gamma)(r + \mu)}{C_{I,n} - C_{V,n}(r + \mu + \gamma)},$$

and it is obviously larger for *skeptics*.

Previous results allow a range of possibilities (ranging from both groups not vaccinating to partial coverage to both groups vaccinating) yielding a number of substantive implications. In particular, the community vaccine uptake determined by the Nash strategy, which is always lower than the community uptake based on the maximization of the (expected) average population utility (see also previous subsection), exhibits a sharp negative relationship with the proportion of vaccine skeptics in the population. Under a measles-like parametrization Nash strategies therefore imply that even moderate proportions of vaccine skeptics can bring the overall coverage sharply below the critical level. Moreover, the Nash vaccine uptake among vaccine skeptics shows a dramatic responsiveness to the relative cost of vaccination $C_{V,1}/C_{I,1}$ i.e. while for low relative vaccination cost the coverage among vaccine skeptics can be high, there is a level of vaccination cost in correspondence of which coverage abruptly falls to zero, as illustrated in Fig. 26.

8.4. Other contributions to mean-field coupled disease–behavior models

In the previous subsections an attempt was made to present some key modeling approaches, and related dynamic implications, from the recent literature on mean-field coupled behavior–disease models. No-doubt, deterministic mean-field models represented the vast majority of contributions in the first epoch of the behavioral epidemiology of infectious diseases. Nonetheless, in order to keep the presentation adequately self-contained we restricted the choice of the materials, focusing on SIR-like models for vaccine preventable infections, such as measles and pertussis. This was partly motivated by the fact that many seminal ideas arose from the conflict between public and self-interest in relation to common childhood vaccinations. Nonetheless it is to be acknowledged that this choice has left-out several excellent contributions to the field. We therefore devote this sub-section to depict some of the historical profile of behavioral epidemiology in order to recall the other many areas and contributions to coupled mean-field models for disease and vaccinating behavior. After the seminal work by Capasso and Serio [406] and a few other phenomenological models mainly focusing on the effects of including nonlinear forces of infection in basic deterministic epidemiological model, behavioral epidemiology developed in the 1980–1990s along two main veins. The first one focused on the already mentioned conflict between public and self-interest in relation to common childhood vaccinations, and the free-rider problem [416,453,479,480]. These works prepared the humus for the outbreak of contributions presented in this subsection. The second vein yielded to the literature on behavioral change triggered by the HIV/AIDS threat since the 1980s. As noted in [80] "the combination of a long incubation period, with difficult and costly treatment, and the lack of a vaccine, have made instilling preventive behavior through the dissemination of information on risky behavior with respect to sexual or intra-venous drug use the main control strategy, especially

in poor resource settings". In this situation, where reliable data on individuals' responses to the spread of epidemics were mostly missing, mathematical modeling played a pioneering role in the understanding of the effects of behavior change on HIV dynamics, including the effect of prevalence-dependent switching to lower risk groups, reducing contact rates after screening or treatment, prevalence-dependent sexual mixing patterns, including the warning that availability of effective therapies and protective vaccine might increases disease severity by raising at-risk behavior [111,481–486]. Though most among the cited papers, which were pioneers in an endless list, considered phenomenological models, there have been also instances of more structured approaches to behavior, such as [487,488]. This literature on behavior change in response to the HIV epidemics threat posed, among other things, the bricks for the current vast literature on behavioral responses to pandemics threat.

In particular, sticking on mean-field models of vaccinating behavior, the current phase has seen an explosion of contributions. A number of papers have applied the Markov decision process theory (DMPT) machinery presented in Section 8.3.3 to formulate game-theoretic models of vaccinating behavior in relation to both seasonal [402,403] and pandemic influenza vaccination [489]. Mean-field analyses were carried out for models investigating adaptive influenza vaccination decisions throughout several influenza seasons based on the inductive vaccination game [490,491]. Still, DMPT models for childhood immunization were considered for rubella [492] and varicella [492]. Mean field epidemiological models with vaccination choices derived from classical utility-based economic approaches were used in [493] to demonstrate the conflict between individual and social optimum, and the nature of externalities arising from vaccination. Deeper utility maximization frameworks were used to investigate a number of discrete-time coupled disease–behavior models considering both risky behavior and vaccination decisions in [494,495]. In particular [494,495] revisited Francis' work [480], showing that (a) the free market solution always yields inefficiently low levels of vaccine uptake, and that (b) higher coverage can only result from frictions such as vaccine imperfection, agents heterogeneity, and in general from strategic interactions. In [496] the implications of vaccination choices were investigated combining a game-theoretic approach with the framework of projected dynamical systems. A coupled behavior–disease model within a public choice economic framework was used in [497] to investigate the effects of the different types of externalities introduced by interventions against infections. An SIR model with bounded-rational vaccination decisions alternative to the model by Oraby and Bauch [465] presented in Section 8.2.7 has been developed and analyzed in [498]. This massive work has prepared the ground for the next step, i.e. moving toward more complicated and possibly more realistic, network models, as reported in the next sections.

9. Behavior-vaccination dynamics in structured, networked populations

In Sections 6–8, we have reviewed the progress made in the realm of behavioral vaccination, which is directly linked to public health measures aimed at eradicating vaccine-preventable infectious diseases. However, the majority of such research relies on the assumption of a homogeneous mixing population, where individuals interact randomly with one another, and thus a susceptible agent is equally likely to acquire the infection from any infected agent in the population. Despite this simplification, as reviewed above, research along these lines has provided important insights into the spreading properties of highly transmissible diseases, and has outlined effective ways for their containment. Nevertheless, the effectiveness and the predictive value of the theory relying on the assumption of a homogeneous mixing population falls sharply for close contact epidemics, which, for example, are brought about by sexually transmitted diseases [6,499–501] or the smallpox [414,502,503]. To resolve this weakness, the assumption of a homogeneous mixing population needs to be replaced by an actual social network that describes the contacts among all involved. In this framework, the nodes of a network represent individuals, while the edges connecting them represent contacts that enable the disease to be transmitted. Actually, in Sections 4–5, we have summarized the recent development of infectious diseases and vaccination on networks, where the dimension of individual behavior is neglected. Here, from a new viewpoint, we will systematically survey how social contact patterns influence the interplay between individual vaccination behavior and disease dynamics, i.e., we will review behavior-vaccination dynamics in structured, networked populations.

9.1. Behavior-vaccination dynamics on static networks

Here we review research concerning behavior-vaccination dynamics on classical static networks, including lattices, small-world networks (SW) [173,504], Erdös–Rényi random graphs (ER) [505], and Barabási–Albert scale-free networks (BA or SF) [506]. Under different microscopic rules, like payoff-driven imitation or opinion-associated dynamics, the vaccination decisions of individuals become dependent on the strategies (i.e., to vaccinate or not) of their neighbors, the perceived risk of infection in their neighborhood, the perceived safety and efficiency of the vaccine, as well as on the financial costs that are associated with vaccination and disease infection. This gives rise to collective phenomena that are akin to the universal behavior of particles near phase transitions points in classical models of statistical physics, and this in turn has far reaching implications for the successful mitigation of epidemic spreading.

9.1.1. Impact of imitation dynamics: The universal "double-edged sword" effect

Like the public goods dilemma [507–509], voluntary vaccination is often regarded as a problem of the commons [415]. In particular, if sufficiently few people vaccinate, the "tragedy of the commons" is likely [510]. As the vaccination level
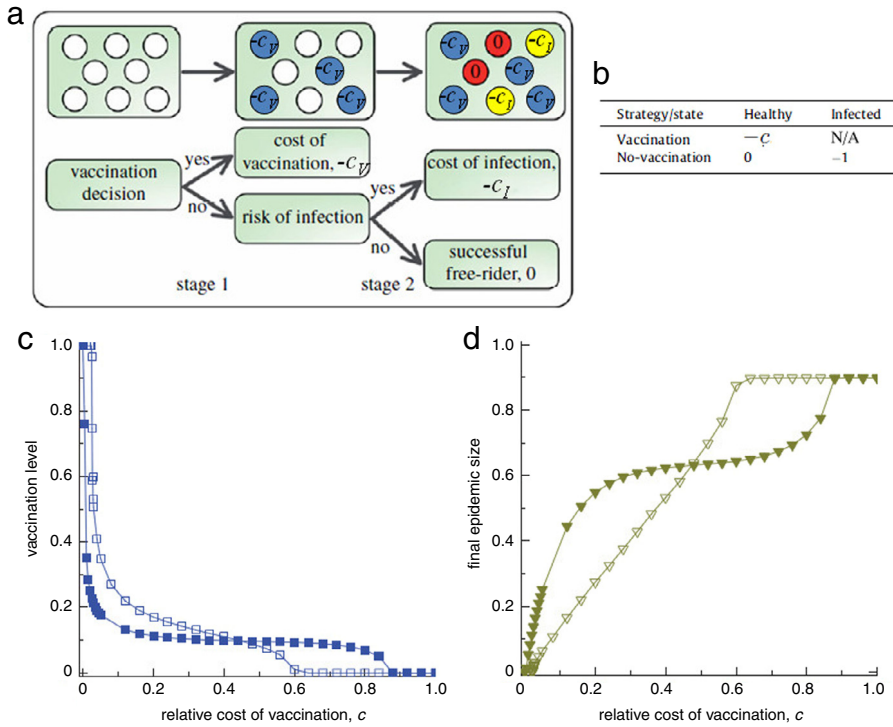
Fig. 27. (a) Schematic illustration of the voluntary vaccination dilemma game, which is further divided into two stages. The first stage is the public vaccination campaign, where vaccinated agents (blue) bear the vaccination cost C_V in return for perfect immunity. During the second stage, the epidemic strain breaks out, and each unvaccinated individual faces the risk of infection. Infected agents (yellow) bear the cost of infection C_I. Unvaccinated individuals who remain healthy are free-riding on the vaccination efforts of others, and they are indirectly protected through "herd immunity". (b) The payoffs for different strategies (to vaccinate or not) and outcomes (healthy or infected). The fractions of vaccinated (c) and infected (d) agents are shown as a function of the relative cost of vaccination c for the selection strength $\kappa^s = 1$ (open symbols) and 10 (filled symbols), as obtained on a square lattice. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
Source: Reprinted figure from Ref. [432].

increases and herd immunity is achieved, individual often forget the risk of some infectious disease and they lost the memory of their outbreaks. Increasingly often there is no memory at all of outbreaks of some infectious diseases. As a consequence, many decide to no longer vaccinate, and thus to effectively free-ride on the efforts of others who still vaccinate their offspring. With such behavior on the uprise, the herd immunity gets lost and the outbreak probability increases significantly. To mathematically describe such a dilemma, Fu et al. have recently integrated the vaccination dynamics and epidemiological modeling into an agent-based evolutionary vaccination game, where vaccination and non-vaccination are mapped to strategies [432]. More precisely, the vaccination dilemma is modeled as a two-stage game (see Fig. 27(a) for a schematic illustration), with the first stage corresponding to a vaccination campaign, and the second stage corresponding to an outbreak of the virus. During the first stage, each vaccinated person acquires perfect immunity at a cost C_V, which covers the actual cost of the vaccine, the perceived risk, adverse long-term health impact (if any), and other intangibles. During the second stage, each infected individual bears the cost C_I, which covers all expenses related to the infection, including the cost of medication, missed work days, and so on. There are also those lucky individuals who remain uninfected during stage two despite being unvaccinated (marked by red circles in Fig. 27(a)). These free-riders avoid all the costs and are thus better off as everybody else, benefiting of course from herd immunity. For simplicity but without loss of generality, to reduce the dimensionality of the parameter space, the aforementioned costs can be re-scaled by defining the relative cost of vaccination as $c = C_V/C_I$ ($0 < c < 1$ since $C_V < C_I$). The payoff function of an individual i thus becomes

$$\Pi_i = \begin{cases} -c, & \text{vaccinated;} \\ -1, & \text{infected;} \\ 0, & \text{free-riding.} \end{cases} \tag{205}$$

It is worth pointing out that the payoff Π_i is negative, simply because agents need to bear the cost rather than to enjoy a profit. Actually the most profitable outcome is a net zero payoff.

Once the epidemic ends, individuals update their strategies for the next season based on their payoffs. To accommodate imperfect information, errors in judgment, and irrational decision making in humans [511], the Fermi function

$$W_{i \to j} = f(\Pi_j - \Pi_i) = \frac{1}{1 + \exp[-\kappa^s(\Pi_j - \Pi_i)]}, \tag{206}$$

can be used to determine whether an individual i adopts the strategy of one randomly selected neighbor j or not. Here Π_i and Π_j are their payoffs, while κ^s ($0 < \kappa^s < \infty$) denotes the strength of selection [423,512,513] (note that κ^s is different from the imitation coefficient κ for well-mixed population in Section 8). Small κ^s values correspond to weak selection, which means that individuals are less sensitive to the payoff difference. Conversely, large κ^s values correspond to strong selection, which means that the level of uncertainty during the strategy adoption process is small.

An important result is that, in comparison to the homogeneous mixing case [435,463], networks act as a "double-edged sword". On the one hand, networks promote vaccination if the cost of vaccinating is low, but on the other hand, beyond a certain threshold value of the cost, they act as strong suppressors of vaccination. Also, the more heterogeneous the interaction network, the more pronounced this effect. Interestingly, a similar paradox exists also for different selection strengths, as shown in Fig. 27(b) and (c). In particular, weak selection guarantees a larger low-cost threshold c_L, below which the epidemic is completely eradicated. As the selection strength increases, the number of free-riders starts going up, which in turn breaks apart the clusters of those who vaccinate. However, when the vaccination cost is sufficient large, the high-cost threshold c_H, above which nobody vaccinates, rises with an increasing selection strength. This means that the coexistence phase with vaccinators and non-vaccinators simultaneously present in the population widens with larger κ^s values. Due to the special role of hubs, these phenomena become the more pronounced the more heterogeneous a network is. In the light of this research, this behavior corresponds to a scenario where voluntary vaccination provided herd immunity for some time but then it fails [491,514].

Similarly to other binary-option games [423,490], the above-reviewed vaccination dilemma game can also be studied with the mean-field approximation approach [397,515,516], as follows. On an uncorrelated and fully connected graph, we define x as the fraction of vaccinated individuals and $r(x)$ as the probability that a susceptible individual gets infected in the population (in general, the whole population is initially divided into two classes: vaccinated agents with fraction x, and susceptible subjectives with fraction $1 - x$. $r(x)$ depending on x is equivalent to being a fraction of susceptible people). Each individual can take the states and payoffs: vaccinated (V, $\Pi_V = -c$), infected (I, $\Pi_I = -1$), non-vaccinated and healthy (or free-riding) (H, $\Pi_H = 0$) (i.e., Eq. (205)). It is then possible to derive how the fraction of vaccinated individuals x changes over time. Namely, when an individual with state V turns to compartment H or I, x drops as

$$x^- = x(1-x)\{[(1-r(x))]W_{V \to H} + r(x)W_{V \to I}\}, \tag{207}$$

where $1 - x$ is the fraction of neighbors holding the opposite strategy, $W_{V \to H}$ ($W_{V \to I}$) is the probability that individuals from the compartment V change to the compartment H (I). Accordingly, x rises as

$$x^+ = x(1-x)\{[(1-r(x))]W_{H \to V} + r(x)W_{I \to V}\}. \tag{208}$$

Taken together, the time evolution of x is

$$\begin{aligned}
\frac{dx}{dt} &= x^+ - x^- \\
&= x(1-x)\{[(1-r(x))]W_{H \to V} + r(x)W_{I \to V}\} - x(1-x)\{[(1-r(x))]W_{V \to H} + r(x)W_{V \to I}\}.
\end{aligned} \tag{209}$$

If the strategy transition probability is determined by Eq. (206), the above equation becomes

$$\begin{aligned}
\frac{dx}{dt} &= x(1-x)\left\{[(1-r(x))]\tanh\left[\frac{\kappa^s}{2}(\Pi_V - \Pi_H)\right] + r(x)\tanh\left[\frac{\kappa^s}{2}(\Pi_V - \Pi_I)\right]\right\} \\
&= x(1-x)\left\{[(1-r(x))]\tanh\left[\frac{\kappa^s}{2}(-c)\right] + r(x)\tanh\left[\frac{\kappa^s}{2}(1-c)\right]\right\}.
\end{aligned} \tag{210}$$

Assuming the right-hand side of this equation equal to zero, we can get the equilibrium vaccination level. Then, the equilibrium fraction of infected individuals, i.e., the size of the epidemic, is expected to be $f_I = r(x)$, satisfying the self-consistent equation $f_I = (1-x)(1 - e^{\mathcal{R}_0 f_I})$, where $\mathcal{R}_0$ is the basic reproductive number of the epidemic (it is usually related with degree variation or degree distribution of networks [290], but here equal to the case of homogeneous mixing population owing to full-connections in network population).

In a social dilemma game, whether or not to switch the vaccination strategy depends not only on the personal success of each individual, but also on the success of others [517,518]. Using this as motivation, one recent work extended the decision-making process on the basis of personal imitation and social influence (or peer pressure) [515]. The pairwise comparison rule for the strategy-updating becomes

$$W_{i \to j} = f(\Pi_j - \Pi_i) = (k_o/k_i)^\alpha \frac{1}{1 + \exp[-\kappa^s(\Pi_j - \Pi_i)]}, \tag{211}$$

where k_o is the number of neighbors possessing the opposite strategy of individual i, k_i is the number of all neighbors of i (i.e., the degree of i), and α determines the so-called peer pressure. Here $\alpha = 0$ recovers the setup used in [432], while positive α values make individuals more prone to sticking with the strategy that is representative for their neighbors.
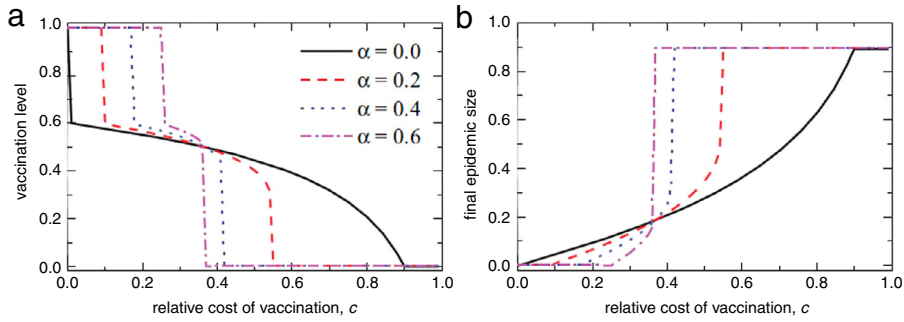
Fig. 28. The fraction of vaccinated (a) and infected (b) agents is shown as a function of the vaccination cost c, as obtained with the mean-field approximation of the vaccination dilemma game. The value of α defines peer pressure, where $\alpha = 0$ recovers the setup used in [432], while positive α values introduce social influence in the imitation processes.
Source: Reprinted figure with permission from Ref. [515].

Accordingly, the mean-field equations for the vaccination level (i.e., Eq. (209)) become

$$\frac{dx}{dt} = x^+ - x^-$$
$$= (1-x)x^{1+\alpha}\{[(1-r(x))]W_{H\to V} + r(x)W_{I\to V}\} - x(1-x)^{1+\alpha}\{[(1-r(x))]W_{V\to H} + r(x)W_{V\to I}\}. \tag{212}$$

Interestingly, compared with the classical version of the model [432], the "double-edged sword" effect occurs again. Peer pressure (i.e., $\alpha > 0$) promotes vaccination coverage as its cost is small, but plummets when costs exceed a critical value, irrespective of the structure of the interaction network. The larger the value of α, the narrower the range of the coexistence phase for vaccination and non-vaccination. These findings can be quantitatively validated also by the mean-field approximation, as shown in Fig. 28. If the vaccination is costly, peer pressure promotes the advantage of vaccinators, and thus forces their neighbors into following. This gives rise to larger clusters of vaccinators, which become unable to effectively prevent the spreading of the disease across the network. Only as the cost of vaccination decreases sufficiently is this cluster formation impaired, and then randomly distributed vaccinators are able to cover wider parts of the network to provide effective protection for their peers.

Along similar lines, the generality of the "double-edged sword" effect was further confirmed with a revised imitation dynamics. Recently, Zhang et al. considered a more rational decision-making model from two aspects [519]. In the first place, the vaccination behavior of individuals was endowed with memory for their past experiences. The payoff function in Eq. (206) was thus reformulated as

$$\Pi_i(t) = \Pi_i(t) + w\Pi_i(t-1), \quad t \geq 1, \tag{213}$$

where w ($0 \leq w \leq 1$) is the memory weight. Secondly, a perceived risk threshold ρ was introduced to avoid the blind imitation of individuals. That is, only when the infection risk surpasses this threshold, the willingness to vaccinate increases. Despite of these differences in the model, the results have shown that, as the cost of the vaccine is small, the coverage of the vaccination increases with increasing w and ρ, leading to lower onsets of an epidemic. If the vaccination is costly, the results also remain qualitatively unchanged, and the main observations are attributed to the formation and decomposition of vaccination clusters, as described in [515].

When modeling voluntary vaccination programs, we can choose different decision-making rules that govern the microscopic dynamics, and these different rules also frequently lead to different outcomes on the global scale. Social learning or imitation, as reviewed in this subsection, is a popular and realistic rule, which however is conducive to the emergence of clusters in structured populations. In vaccination dilemmas, this ultimately leads to a persistent "double-edged sword" effect, where below a certain cost vaccination is effective, while above this threshold it becomes ineffective. In what follows, we will review results obtained also with other decision-making rules at the microscopic level, and outline differences among them at larger scales.

9.1.2. Impact of interaction network topology

In network epidemiology research, it is thoroughly established that the structure of the interaction network plays an important role in the spread of infectious disease [171,202,214,222,229,230,337,520–523]. For example, epidemic spreading on scale-free networks is difficult to control, and with the lack of an outbreak threshold, the disease can propagate even with a low infection rate [521]. However, this holds if vaccination is absent. If an effective vaccination program is incorporated into the disease dynamics the outcome can be different. To account for this, a new modeling framework for evolutionary vaccination decision-making was proposed [524], where individuals resolved the vaccination dilemma in terms of the trade-offs between the infection risk, the side effects, and the cost of vaccination. More precisely, if a susceptible individual has
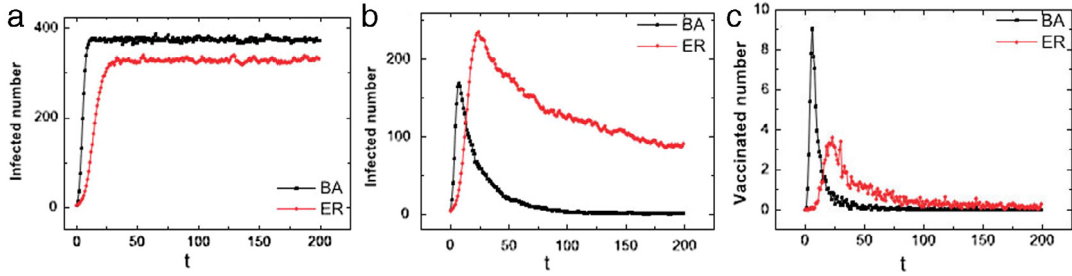
Fig. 29. Time evolution of the number of infected individuals without (a) and with (b) voluntary vaccination, and the number of vaccinated agents (c) on BA (or SF) and ER networks. The parameter values are $c_1 = 1$ and $c_2 = 0.7$.
Source: Reprinted figure with permission from Ref. [524]. © IOP Publishing & Deutsche Physikalische Gesellschaft. CC BY-NC-SA. Reproduced by permission of IOP Publishing.

k_{inf} infected neighbors, the total infection probability will be

$$\lambda_I = 1 - (1 - \beta_I)^{k_{inf}}, \tag{214}$$

where β_I is the infection rate for a single contact between a susceptible and an infected individual.

Similarly, the total perceived infection probability is

$$\lambda_{perc} = 1 - (1 - \beta_{perc})^{k_{inf}}, \tag{215}$$

where if there is full knowledge for the infection risk $\beta_{perc} = \beta_I$ and $\lambda_{perc} = \lambda_I$.

If the susceptible agent does not choose to vaccinate, it bears the perceived cost

$$P_N = c_1 \lambda_{perc}, \quad or \quad P_N = c_1 \lambda_I. \tag{216}$$

Conversely, missing out on vaccination yields

$$P_V = c_2, \tag{217}$$

where c_1 and c_2 are related to the risk of infection, to potential side effects, and to the cost of vaccination.

Unlike with imitation considered in the previous subsection (see Eq. (206)), here individuals adjust their strategies in agreement with another kind of decision-making rule that governs the microscopic dynamics, namely according to their self-interest or cost minimization. Thus, individuals try either to minimize their cost or to maximize their benefits (i.e. self-interest rule). The strategy s_i of individual i will thus be

$$s_i = f(P_V - P_N) = \begin{cases} \text{vaccination}, & \text{if } P_V - P_N < 0; \\ \text{non-vaccination}, & \text{otherwise}. \end{cases} \tag{218}$$

Contrary to the result that the outbreak of disease on SF networks is more likely than on ER random graphs [229,520,521], in the realm of the model above it turns out that, if individuals are inclined to vaccinate at the initial stage of an outbreak, the epidemic is better constrained on the SF network (see Fig. 29). Namely, on SF networks, the majority of middle- and large-degree nodes have a higher tendency to vaccinate. To verify this fact, the authors in [524] provide a mathematical proof for the threshold degree of vaccination. Taking into account the vaccination condition given by Eq. (218), one obtains the threshold

$$P_N > P_V \Rightarrow \quad c_1 \lambda_I > c_2 \Rightarrow \quad 1 - (1 - \beta_I)^{k_{inf}} > \frac{c_2}{c_1} \Rightarrow \quad k_{inf} \geq \left[\log_{1-\beta_I} \frac{c_1 - c_2}{c_1} \right] + 1. \tag{219}$$

Using $\beta_I = 0.2$, $c_1 = 1$ and $c_2 = 0.7$ as an example, we have $k_{inf} \geq 6$. For network of $\langle k \rangle = 6$, the vaccinated candidates are very limited on ER graphs, but more nodes prefer to vaccinate to reduce the loss because of the existence of middle- and large-degree nodes in SF networks. The vaccination program on SF networks, in some sense, is like the targeted vaccination in the mandatory model [274,277], yet shows an adaptive character as well. In addition to permanently effective vaccination, temporarily effective vaccination (taking into account that the vaccine for some disease like influenza and hepatitis B will lose its efficiency after a given period of time) has also been considered. Though in this case there exist vaccination oscillations, the spread of disease ultimately can be controlled on complex networks. This further confirms that the larger inclination for hub nodes to get vaccinated plays an important role and can critically affect the success of vaccination campaigns, which ought to be considered in the design of incentives for vaccination.

Despite of the inspiring progress, there does not exist a direct comparison between a perfect and an imperfect vaccine in [524]. Aiming to clarify this, Cardillo et al. introduced a tunable parameter τ to modulate the quality of the vaccine in influenza-type disease, the vaccine being perfect at $\tau = 0$ and useless at $\tau = 1$ [525]. When the vaccine is perfect, the research revealed that SF networks outperform ER graphs, since the overall vaccinated (infected) number is larger (smaller) on SF networks. This result is in agreement with the general conclusions reported in [524]. However, if the vaccine is

imperfect, a crossover effect occurs such that the ER graphs are more effective than SF networks in enhancing vaccination. In [525], these results are mainly attributed to the competition of two processes taking place. One is the propagation of the diseases, and second is the uptake of vaccination behavior. If the vaccine works perfectly, or even if the vaccine is imperfect but the infection rate is small, the vaccination behavior spreads faster on SF networks than on ER graphs (i.e., the vaccination onset starts earlier on SF networks), as the natural epidemic threshold is smaller on SF networks [229,520,521]. However, if the vaccine is imperfect and the infection rate is large, the outcome of the competition between disease and vaccination spreading reverses. Namely, under such conditions, the smaller benefits provided by the imperfect vaccine cause agents to forgo vaccination, which naturally promotes the propagation of the disease.

Taken together, these findings significantly enrich our understanding of behavior-vaccination dynamics on static, complex networks, and can relevantly inform on optimal vaccination strategies.

9.1.3. External incentive programs

The role of subsidy policy: a paradox. Except for internal mechanisms, external incentive programs also play a vital role in the design of behavioral vaccination [526,527]. For example, historically, subsidy policy showed its particular effectiveness in terms of controlling epidemic spreading, which is of significance from the socioeconomic perspective [528–532]. However, how different subsidy policies might work in structured populations is still an open issue. Motivated by this, Zhang et al. recently considered two categories of subsidy policies and explored their influence on different networks, and how they affect vaccination decisions and disease dynamics [533]. The first class is the so-called free subsidy policy, where the total amount of subsidy Λ is distributed to a certain fraction of individuals, who take vaccination without any personal cost. The (net) payoff function (i.e. Eq. (205)) thus becomes

$$\Pi_i = \begin{cases} 0, & \text{subsidized-vaccinating;} \\ -c, & \text{vaccinated;} \\ -1, & \text{infected;} \\ 0, & \text{free-riding} \end{cases} \tag{220}$$

where the number (fraction) of subsidized vaccinators or freely vaccinated individuals is $N_{SV} = \Lambda/c$ ($p = N_{SV}/N$, where N is the size of the network). In this case, it is worth noting that, since the number of freely vaccinated individuals keeps constant in each vaccination stage, the selection process of them does not need to be repeated. If one individual is initially chosen as the recipient (to receive the vaccine for free), it will remain a recipient forever. The second class is partial-offset subsidy policy, where each vaccinated individual is offset by a certain amount ξ, resulting in the actual vaccination cost $c(1 - \xi)$. Correspondingly, the (net) payoff becomes

$$\Pi_i = \begin{cases} (\xi - 1)c, & \text{subsidized-vaccinating;} \\ -1, & \text{infected;} \\ 0, & \text{free-riding.} \end{cases} \tag{221}$$

Similar to Eq. (209), the mean-field approximation for the fraction of vaccinated individuals x for both the free subsidy policy and the partial-offset subsidy policy is given as

$$\frac{dx}{dt} = x^+ - x^-$$
$$= x(1 - x)\{[(1 - r(x))]W_{H \to V} + r(x)W_{I \to V}\} - (x - p)(1 - x)\{[(1 - r(x))]W_{V \to H} + r(x)W_{V \to I}\}, \tag{222}$$

and

$$\frac{dx}{dt} = x^+ - x^-$$
$$= x(1 - x)\{[(1 - r(x))]W_{H \to SV} + r(x)W_{I \to SV}\} - x(1 - x)\{[(1 - r(x))]W_{SV \to H} + r(x)W_{SV \to I}\}, \tag{223}$$

where SV is the subsidized-vaccinating state and $x - p$ is the fraction of individuals who can change their vaccination strategies.

Based on the imitation dynamics (i.e. Eq. (206)), it is found that the free subsidy policy in general leads to higher vaccination coverage than the partial-offset subsidy policy, indicating that the former case is more effective in controlling epidemic outbreaks. This phenomenon can also be verified by means of the mean-field approximation (i.e., Eqs. (222) and (223)). In the free subsidy case, the recipients act as role models, and are thus more appealing than the non-subsidized vaccinators in attracting their neighbors to take vaccination. This is expressed qualitatively in terms of the higher fraction of imitated vaccinators in the neighborhoods of recipients. Although freely vaccinated recipients (red) are randomly distributed on networks, it is quite remarkable how they are able to excite their neighbors and motivated them to imitate their behavior (see Fig. 30). This gives rise to vaccination clusters (green), which naturally results in a more effective control in comparison to the partial-offset subsidy policy. Snapshots in Fig. 30 provide a compelling visual in favor of this argument.

As can be inferred from 9.1.1, the updating dynamics based on individual behavior choices can directly decide the final vaccination coverage and the average epidemic size. When studying the impact of subsidies [533], the authors considered
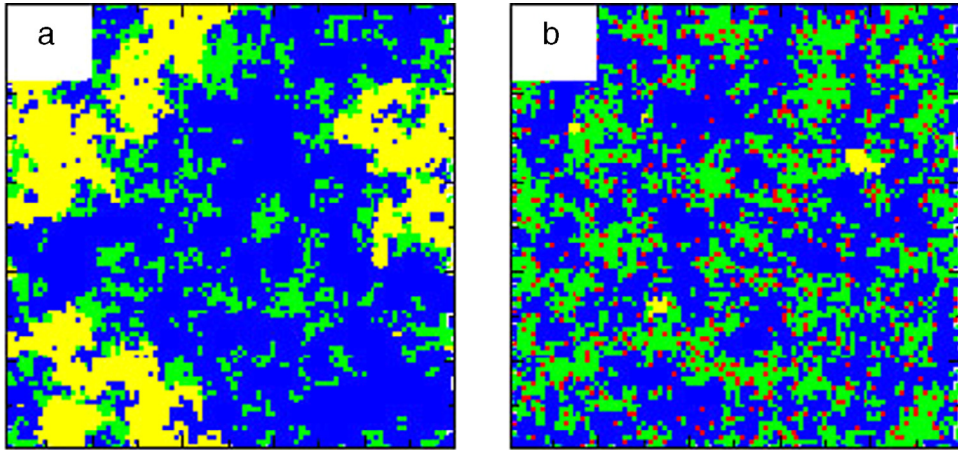
Fig. 30. Characteristic snapshots of the stationary state configuration for partial-offset (a) and free (b) subsidy policies on the square lattice. Blue, green, yellow, and red points represent free-riders (not vaccinated and healthy), vaccinators, infected individuals, and freely vaccinated recipients, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
Source: Reprinted figure from Ref. [533].

also another Fermi-like updating rule, but one that does not involve imitation [528]. Such updating is known as myopic [534], because individuals update their vaccination preference by balancing their own advantages and disadvantages of vaccination (like the self-interest rule), rather than considering input from their neighbors. In particular, an individual first gets the perceived costs P_N and P_V, and then opts for vaccination according to the probability

$$w_C = f(P_V - P_N) = \frac{1}{1 + \exp[-\kappa^s(P_V - P_N)]}.$$ (224)

In other words, the strategy s_i of individual i is governed as

$$s_i = f(P_V - P_N) = \begin{cases} \text{vaccination}, & \text{if } q < 1/\{1 + \exp[-\kappa^s(P_V - P_N)]\}; \\ \text{non-vaccination}, & \text{otherwise}; \end{cases}$$ (225)

where q is a randomly distributed real number in [0, 1].

This myopic updating dynamics leads to completely different results in terms of the effect of different subsidy policies. Namely, in this case the partial-offset subsidy policy works better than the free subsidy policy. Under the latter, there is polarization between subsidized vaccinators and general agents, which is very non-beneficial for the emergence of herd immunity. Owing to the unequal distribution of subsidy, the perceived infection risk of those recipients greatly decreases, but non-recipients are more prone to taking a risk (i.e., the willingness of vaccination among non-recipients declines) after balancing the advantages and disadvantages of vaccination. In the partial-offset subsidy case, the cost of vaccination is evenly reduced for each vaccinated individual due to the endowment of the policy, which provides natural conditions for the formation of herd immunity. In addition to group-optimal vaccination, the study revealed that a moderate subsidy rate can generate the maximal socioeconomic benefit under the partial-offset subsidy policy, which is a particular meaningful message for the public health authorities.

Considering these results, it becomes clear that the optimization [535–538] of the subsidy policy is indeed a difficult yet very important issue, which can mean the difference between a successful and an unsuccessful vaccination campaign. For the optimal outcome, it is important to carefully consider the mechanisms of individual decision-making, as well as to actually test on a representative sample how people respond to different external incentives in realistic situations.

The impact of committed individuals, or zealots. In real life, many people hold rigid viewpoints on certain aspects of reality, regardless of peer pressure, incentives, and media. Such people are known as committed individuals or, shorter, zealots. The impact of zealots on various aspects of social dynamics has been studied often before, for example in opinion dynamics [539–542] or in evolutionary social dilemmas [543]. Of course, zealots are also found with regards to vaccination. Since some agents are much more sensitive to the infection risk than average, they consistently adopt the vaccinating strategy, i.e., the committed vaccinators do not engage in the decision-making process [544]. In essence, the setup of committed vaccinators is the same as the free subsidy policy (the mean-field analysis is the same as put forward with Eq. (222)), needing only one additional parameter f_C to determine their fraction. Owing to the effect of "steadfast role models", committed vaccinators stimulate vaccination uptake and can thus contribute significantly to the prevention of infectious diseases. The larger the value of f_C, the more obvious this effect. As such, committed vaccinators persuade their neighbors to follow their behavior and form vaccinating clusters, which is, as highlighted often above, the typical effect of imitation [432,515,533].

Of course, zealots can also be committed non-vaccinators. It is thus necessary to look at the down side too, that is, considering committed vaccinators as good role models, but also considering committed non-vaccinators as bad role models.

Table 5
The payoffs for different strategies and states. Here Π_V, Π_S and Π_L are the (net) payoffs for vaccination, self-protection and laissez-faire, respectively. The superscripts H and I denote healthy and infected states, while w is the probability of a susceptible to become infected. For more details, please refer to [546].

Strategy & State	Fraction	Payoffs
Vaccinated & Healthy	ρ_V	$\Pi_V = -c$
Self-protected & Healthy	$\rho_S^H = \rho_S[\delta + (1-\delta)(1-w)]$	$\Pi_S^H = -b$
Self-protected & Infected	$\rho_S^I = \rho_S(1-\delta)w$	$\Pi_S^I = -b-1$
Laissez-faire & Healthy	$\rho_L^H = (1-\rho_V-\rho_S)(1-w)$	$\Pi_L^H = 0$
Laissez-faire & Infected	$\rho_L^I = (1-\rho_V-\rho_S)w$	$\Pi_L^I = -1$

This has been done in [545], and the results were as follows. On lattices, the presence of committed vaccinators (non-vaccinators) exerts a positive (negative) effect on the epidemic spreading. Moreover, on SF networks the presence of committed vaccinators cannot prevent an epidemic for small and intermediate vaccine cost, predominantly because the hub nodes consistently prefer to reject vaccination. In this respect, the effect of committed individuals depends not only on their fraction, but also on the network structure and the cost of vaccination.

9.1.4. Multiple-strategy dilemma framework: A counter-intuitive observation

When an infectious diseases threatens, vaccination is often the first and the most common line of defense to protect the public. However, due to high economic cost, religious beliefs, as well as possible side effect [547–549], people sometimes prefer alternative, and frequently also cheaper self-protection measures [110,406], such as avoiding contact with risk groups or wearing face masks to at least temporarily mitigate the risks. To take this into account, three different strategies, namely vaccination, self-protection, and laissez-faire (i.e., neither vaccination nor self-protection), were considered in the realm of an evolutionary epidemic game [546]. Assuming that the vaccine provides perfect protection at a cost c, self-protection is effective with probability δ at a lower cost b ($b < c$). Moreover, with probability $1 - \delta$ a self-protected individual turns to laissez-faire. That is, self-protection is an alternative in between traditional vaccination and laissez-faire. Analogous to the binary-choice vaccination game [432], the payoff of this new framework is portrayed in Table 5, where ρ_V, ρ_S and ρ_L are the factions of vaccinated, self-protected and laissez-faire individuals, respectively. According to the mean-field approximation framework, the time evolution of the fraction of strategies is governed by the following equations

$$\frac{d\rho_V}{dt} = (\rho_V \leftrightarrows \rho_S^H) + (\rho_V \leftrightarrows \rho_S^I) + (\rho_V \leftrightarrows \rho_L^H) + (\rho_V \leftrightarrows \rho_L^I), \tag{226}$$

and

$$\frac{d\rho_S}{dt} = (\rho_S^H \leftrightarrows \rho_V) + (\rho_S^H \leftrightarrows \rho_L^H) + (\rho_S^H \leftrightarrows \rho_L^I) + (\rho_S^I \leftrightarrows \rho_V) + (\rho_S^I \leftrightarrows \rho_L^H) + (\rho_S^I \leftrightarrows \rho_L^I), \tag{227}$$

where

$$
\begin{aligned}
\rho_A \leftrightarrows \rho_B &= (\rho_B \to \rho_A) - (\rho_A \to \rho_B) \\
&= \rho_A\rho_B(W_{B\to A} - W_{A\to B}) \\
&= \rho_A\rho_B\left\{\frac{1}{1+\exp[-\kappa^s(\Pi_A - \Pi_B)]} - \frac{1}{1+\exp[-\kappa^s(\Pi_B - \Pi_A)]}\right\} \\
&= \rho_A\rho_B\tanh\left[\frac{\kappa^s}{2}(\Pi_A - \Pi_B)\right].
\end{aligned}
\tag{228}
$$

Here A and B represent the final states, and the expressions for the fractions and payoffs of each state are from Table 5. Since the preservation condition $\rho_V + \rho_S + \rho_L = 1$ holds, it is straightforward to get the equilibrium fraction of all three strategies from the above equations.

With this setup, the authors in [546] discovered a counterintuitive phenomenon. Namely, intermediate values of δ lead to a larger epidemic size and a higher collective cost, which is similar to the well-known Braess's paradox [550,551]. More precisely, the whole range of δ is divided into three regions, as shown in Fig. 31(a). In the left region, owning to the low efficiency, no one self-protects, which returns the traditional binary selection model [432]. Infected laissez-faire individuals (dark red) and uninfected laissez-faire individuals (i.e., free-riders, light red) are strictly separated by vaccinators and form percolating clusters (see Fig. 31(d)). With middle region, for intermediate values of δ (i.e., the region between both dash lines), the payoff advantage of self-protection starts to play a noticeable role, so that vaccination and partial laissez-faire are subverted. However, this leaves the heard with insufficient protection, ultimately resulting in a larger epidemic and the fragmentation of percolating free-rider clusters. Another counterintuitive fact is that more self-protected individuals (whose individual payoff is higher) generate a lower (rather than higher) total payoff of the population. Thus, the collective payoff and the epidemic size are negatively correlated with the spread of self-protected individuals, as shown in Fig. 31(b) and (c). In
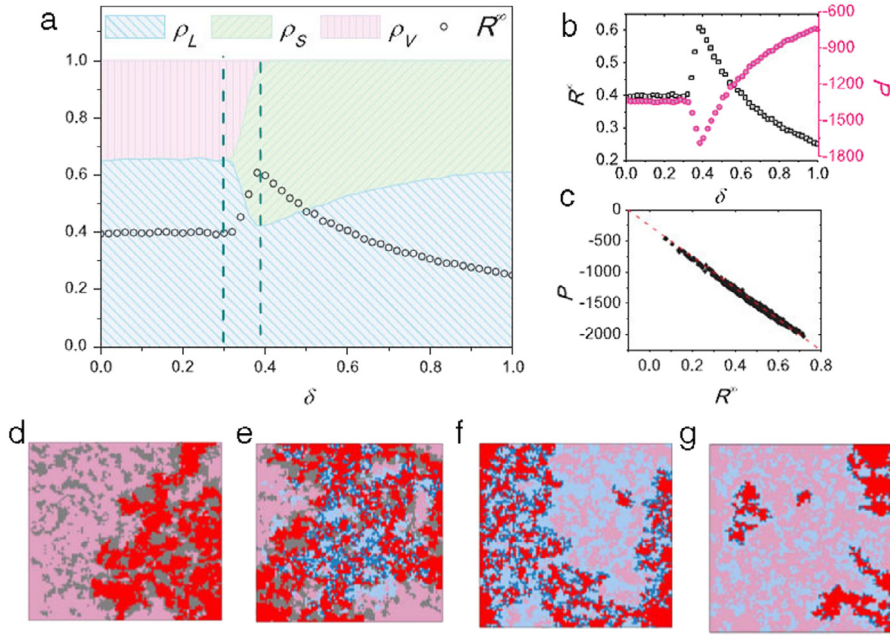
Fig. 31. (a) The fraction of strategies and the epidemic size R^∞ are shown as a function of the successful rate of self-protection δ. (b) The epidemic size R^∞ and the total payoff P are shown as a function of δ. (c) Correlation between the population payoff P and the epidemic size R^∞. (d)–(g) Typical snapshots of the steady state for $\delta = 0.2, 0.35, 0.5$ and 0.95, respectively. Gray, light red, dark red, light blue and dark blue denote vaccinated, uninfected laissez-faire, infected laissez-faire, uninfected self-protective, and infected self-protective individuals, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
Source: Reprinted figure from Ref. [546].

the third region, when δ is sufficiently large, self-protection is sufficiently effective. Epidemics decline and vaccination is not a viable option. Under such circumstances, the percolating cluster of free-riders (light red) recovers due to herd immunity.

The universality of this interesting phenomenon can be further extended to other networks, and it can be qualitatively validated by means of the mean-field approximation. In heterogeneous networks, due primarily to the hubs, those that become infected and laissez-faire individuals fare worse still than they do on more homogeneous networks where the differences in degree are smaller. Similarly counterintuitive phenomena can also be observed with myopic updating (self-interest rule) [552], although the differences with the more commonly considered imitation rule [432,515,524,533] are significant. This should be taken into account, specifically when considering competing strategies in vaccination programs.

9.2. Behavior-vaccination dynamics on empirical networks

Different from the network models reviewed above, realistic contact networks usually exhibit assortative and disassortative mixing patterns [168,175,553] and community structure [554,555], both of which significantly impacts the coupled dynamics of behavior and vaccination taking place on them. In this subsection we review recent research devoted to behavior-vaccination dynamics on empirical networks.

9.2.1. How does feedback influence well-known passive and pro-active policies?

The voluntary ring vaccination strategy. In a homogeneous mixing population, it is a well-recognized fact that voluntary vaccination (without any economic incentives) is difficult or impossible to control, and thus cannot be used effectively to eradicate vaccine-preventable disease [402,415,447,490,491,529,557–559]. As vaccination coverage levels increase, herd immunity can provide greater protection. However, soon the low infection risk reduces the enthusiasm of individual vaccination and leads to the reoccurrence of high disease prevalence. That is to say, voluntary policy generates oscillations in the uptake of vaccination [435], which is particular true for highly transmissible diseases. However, the history of smallpox teaches us that complete eradication by means of voluntary vaccination is possible [376,381,560]. Aiming to elucidate this observation, Perisic et al. have integrated disease dynamics with a voluntary ring vaccination policy in a social network [431].

Based on the perceived risk of infection and benefit of vaccination, they obtained the perceived cost functions of non-vaccination P_N and vaccination P_V (see Fig. 32(a) for the transition probability of individual states):

$$P_N = (1 - \lambda_{perc})a + \lambda_{perc}[(1 - d_{inf})L], \tag{229}$$

$$P_V = \varsigma[(1 - d_{vac})L] + (1 - \varsigma)\{(1 - d_{vac})[(1 - \lambda_{perc})a + \lambda_{perc}(1 - d_{inf})L]\}, \tag{230}$$
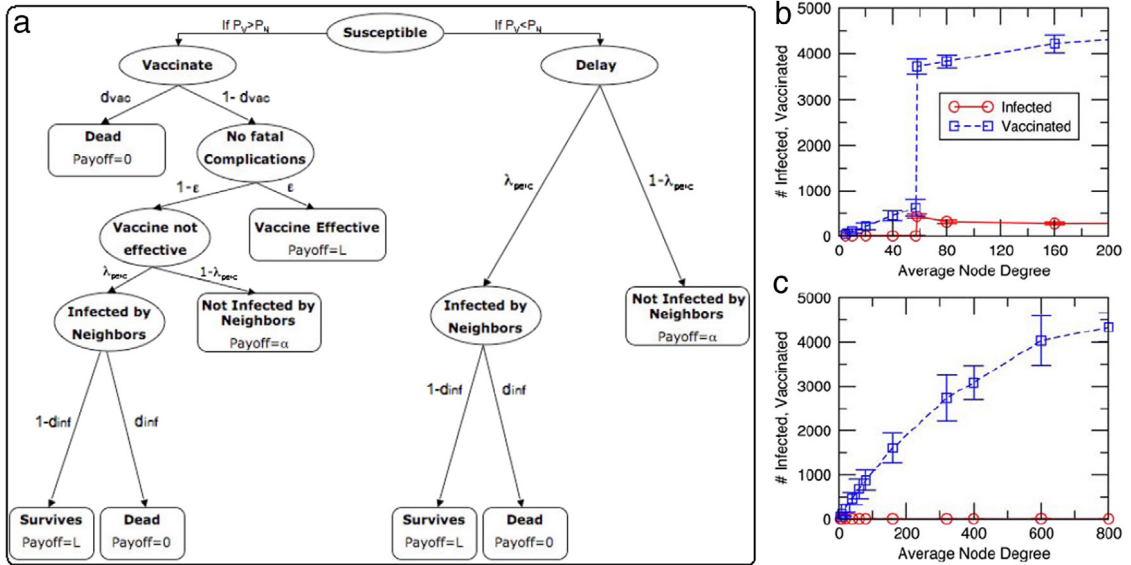
Fig. 32. (a) Schematic illustrations of the perceived costs of individuals. Ellipses represent transition states while boxes represent formal states. Parameters on arrows denote transition probabilities and expressions in boxes denote costs for entering that state. For the meaning of used parameters we refer to [431,556]. The numbers of infected and vaccinated agents are shown as a function of the average neighborhood size in the theoretical model (b) and the smallpox model (c). The error bars represent two standard deviations from the mean across 20 simulations per datum.
Source: Reprinted figure from Refs. [431,556].

where a (L) is the payoff for individuals with continued susceptibility (lifelong immunity), ς is vaccine efficacy and d_{inf} (d_{vac}) represents the probability of death due to infection (vaccine-related complications). In particular, P_V consists of two terms: the first term denotes the total payoff if vaccine is efficacious; while the second is the total payoff once vaccine loses efficacy. The decision-making process was in accordance with rational self-interest rule (i.e., Eq. (218)). If the cost to vaccinate is smaller than the cost not to vaccinate on a given day (i.e., $P_V < P_N$), individuals would turn to vaccination. Interestingly, there is a threshold in the neighborhood size, which guaranteed "first-order" phase transition of vaccination behavior, like the explosive phenomena of collective dynamics in statistical physics [561,562]. From Fig. 32(b), it is clear that voluntary ring vaccination quickly contains the outbreak until the average neighborhood size is larger than 57. Due to effective control, the total number of infected and vaccinated agents remains very small below this threshold. Interestingly, as the neighborhood size increases (while the average force of infection is held constant), the model recovers the homogeneous mixing scenario [402,415,447,557,558], where each infected agent could influence more neighbors. As a result, infection breaks through the imperfect rings of vaccinated individuals, percolates through the whole network and leads to a considerable final epidemic size and numerous vaccinated. To further validate the accuracy of these theoretical predictions, the authors in [431] also carried out the investigation for smallpox-specific disease history and vaccine properties, with similar results (see Fig. 32(c)). Their efforts thus illustrate the significant difference between continuous population and spatial structure population in coupled behavior-vaccination dynamics. On the other hand, it suggests that the design of voluntary vaccination, as a primary control measure, to eradicate outbreak of some close contact diseases such as smallpox should rely on the number of social contacts and on the local structure of the population.

Though voluntary vaccination policy enables eradication of the disease in contact-structured networks, the foregoing model made a couple of simplifying assumptions regarding the natural history of the disease, pre-existing immunity, and human behavior. To relax some of these simplifications, the authors further investigated the dynamics via weighing infection risk (i.e., the probability of death due to disease) versus vaccine risk (i.e., the probability of death due to vaccine) [556]. Of significance, they have shown a counterintuitive phenomenon for the coupled dynamics, which was then divided into three phases (see Fig. 33). Phase I exhibits a negligible final size of the epidemic and a small number of vaccinated individuals. In this phase, any node with an infected neighbor would vaccinate immediately, preventing disease transmission. Thus, the infection risk will be fast, and effectively controlled under a voluntary ring vaccination. Then, in phase II, nobody is likely to vaccinate due to low disease risk or high vaccine risk (this phase is not shown in Fig. 33(b)). This finally results in a large final epidemic size. For intermediate epidemic risk and vaccine risk (i.e., phase III, the area between two dash lines), though a great number of agents prefer to vaccinate, the disease still percolates through the whole population due to imperfect protection. Except for a deterministic decision process, the authors also tested the effect of a probabilistic setup. Accordingly, even if $P_V < P_N$, individuals may still opt to vaccinate with a given probability. If this probability is low, then very few people are allowed to vaccinate and infection takes over the population (see Fig. 33(c)). As the probability increases, the outbreak is effectively contained by means of the voluntary ring vaccination. Interestingly, the continuous enhancement of vaccination probability quickly reduces the essential number of those that need to be vaccinated for better controlling the spread of the
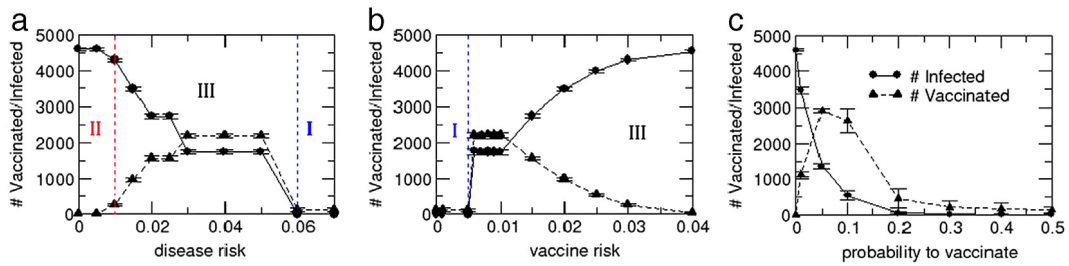
Fig. 33. The numbers of infected and vaccinated individuals are shown as a function of disease risk (a), i.e., the probability of death due to disease), vaccine risk (b), i.e., the probability of death due to vaccine) and probability to vaccinate (c). Note that (a)–(b) are further divided into three phases. The error bars represent two standard deviations from the mean across 20 simulations per datum.
Source: Reprinted figure from Ref. [556].

disease (namely, there is a vaccination peak). Altogether, the study [556] revealed that the final outcome of rational self-interest dynamics under a voluntary vaccination policy is closely related with contact patterns, with the perceived vaccine risk, as well as with the risk of contracting the disease.

The superspreader vaccination strategy. In the context of network vaccination (i.e. Section 5), it has been thoroughly established that targeted vaccination toward superspreaders can significantly reduce the spread of disease [274,290,343,370,371]. However, the assumption that superspreaders (who typically have many connections and are responsible for the majority of secondary infections [563–567]) always accept vaccination does not capture the realistic transmission dynamics with behavioral response, especially for seasonal influenza-like vaccination. In this regard, Wells et al. analyzed how behavioral feedback affects the effectiveness of vaccination strategies [434]. In addition to the passive vaccination as the baseline strategy (whose update is governed by a function of payoff difference $f(P_V - P_N)$), individuals can also implement one of four pro-active strategies: (1) random vaccination (RV) targeting a randomly chosen individual; (2) nearest neighbor vaccination (NN) aiming to a randomly chosen individual and one of their neighbors; (3) chain vaccination (CV) targeting either a randomly chosen individual or a neighbor of an individual targeted the previous day; (4) improved nearest neighbor vaccination (INN) targeting a randomly chosen individual and one of their most connected neighbors.

Surprisingly, it was revealed that the proposed superspreader vaccination strategies provide little or no improvements, compared to traditional ad-hoc delivery of vaccination [324,568]. The behavioral feedback facilitates policy resistance, which nearly completely undermines the effectiveness of superspreader strategies. Actually, superspreader strategies enhance the vaccination level among superspreaders. However, this effort is offset by the low coverage among non-superspreaders (representing the majority of nodes in a heterogeneous network), which naturally leads to a low overall level of success. These observations indicate that selecting connection heterogeneity as an important criterion of implementing superspreader strategies is not sufficiently accurate. For example, individuals with higher infectiousness or a longer period of infectiousness may not coincide with highly connected individuals [296,566]. Moreover, to improve the readiness to vaccinate in the population, an economic incentive to the targeted individual has also been proposed. As expected, the undermined benefit of pro-vaccination strategies can be somewhat recovered under a substantial economic stimulus. Thus, the design of vaccination strategies targeting superspreaders must consider the possible policy resistance caused by strong effects of behavioral feedback, and mitigate it whenever possible.

The mixed vaccination strategy. Up to now, we have acquired a thorough understanding for the interrelationship between vaccination decision and disease dynamics: via discrete, memoryless strategies, human decision-making can profoundly influence epidemic incidence as well as vaccination coverage [397,404,415,569]. However, as the cooperation strategy in evolutionary games [570,571], vaccination sometimes can be regarded as one mixed strategy, i.e., choosing vaccination with a given probability and taking the non-vaccination stance with the remaining probability. It is interesting to consider how does this setup affect the vaccination program and the spread of disease.

To address this, Cornforth et al. defined the vaccination strategy v_k of an individual with degree k as the probability that it will opt for vaccination [285]. The total perceived payoff of this individual is thus given by

$$\pi_k(v_k) = u - v_k(C_V + o_k C_R) - (1 - v_k)\alpha_k C_I, \tag{231}$$

where u is the baseline payoff corresponding to the value of a healthy life over the course of an epidemic, while C_V (C_I) represents the monetary cost of vaccination (infection). Because the influenza vaccine is only partially effective, some vaccinated agents remain susceptible to the infection. Here, C_R denotes the cost of infection to a partially resistant (immune) agent, while o_k (α_k) is the perceived probability of becoming infected if vaccinated (non-vaccinated). Thus, different from respective expressions of discrete strategies [432,524,546], the second term in the right hand side of Eq. (231) is the perceived payoff of vaccination selection, and the third accounts for the perceived payoff of non-vaccinated behavior. It is also worth mentioning that the estimation of perceived risk depends on the information of past epidemic seasons.

Regardless of the properties of the interaction networks, it was shown in [285] that the likelihood of vaccination increases with degree (because the infection risk also increases with degree), which is the same with prediction of voluntary ring vaccination [431]. Nevertheless, the degree threshold for vaccination is closely related with the overall structure of the
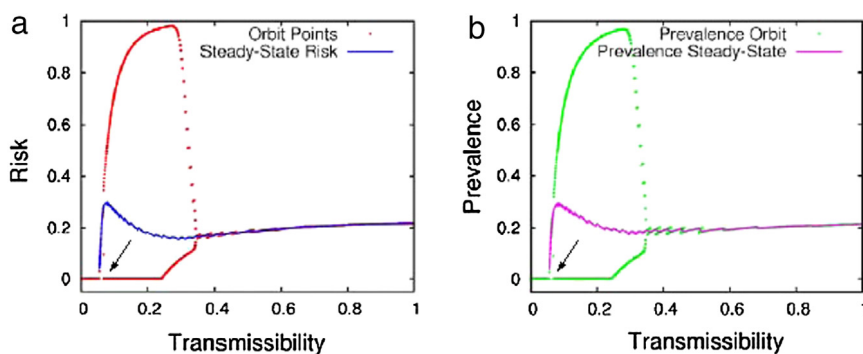
Fig. 34. The impact of transmissibility on risk (a) and disease prevalence (b) (steady-states and unstable limit cycles) in urban networks. Before the threshold (indicated by black arrows), there is no epidemic risk and vaccination. With transmissibility increasing, these two first diverge and then end at the same state again.
Source: Reprinted figure from Ref. [285].

network, and increases with the heterogeneity of the degree distribution (though the perceived fraction of vaccination does not change monotonically with network heterogeneity). Interestingly, it was also found that vaccination relies critically on the transmissibility of disease [285]. Results presented in Fig. 34 show how the steady-state risk/prevalence (corresponding to Nash Equilibria) and unstable limit cycles vary as a function of transmissibility. For small transmissibility, there is no epidemic risk and thus no vaccination under the threshold (indicated by black arrows). As transmissibility increases, epidemic risk/incidence is predicted to converge to a stable Nash Equilibrium. However, in reality there are considerable oscillations from one season to the next. This is caused by the polarization of vaccination behavior: non-vaccination is always followed by wide-spread disease and consequently high vaccination coverage in the next season. At a high level of transmissibility, the oscillatory pattern in vaccination behavior goes back to the steady state. These oscillations seem empirically inconsistent with the relatively stable vaccination pattern of influenza. To elucidate this discrepancy, it is further unveiled that increasing the number of prior seasons that individuals recall during the decision-making process enables these oscillations collapsing onto a stable vaccination strategy.

9.2.2. Social clustering of behavior and opinions—assortativity of (non-)vaccination

Empirical evidences—clusters driven by strategy and opinion updating. When reviewing results presented in 9.1, we have emphasized that the updating dynamics plays a crucial role in the final size of epidemic and vaccination coverage. Notably, during the decision-making process, there exist different criteria and factors that affect each particular choice, such as payoff maximization [524], stochastic imitation [432], and social impact theory [572]. While it is clear that taking into account one or the other criterion leads to different evolutionary outcomes, few works have focused on the impact of joint action. In this regard, Mbah et al. incorporated imitation dynamics (i.e., Eq. (206)) and self-interest dynamics (i.e., Eq. (218)) into a coupled epidemiological model on different contact networks [433]. In particular, individuals could make their vaccination decision either based on imitation or based on payoff maximization, with probability Θ and $1 - \Theta$, respectively.

Surprisingly, in this framework the "double-edge sword" effect appears again. Namely, imitation behavior (i.e., $\Theta = 1$) increases vaccination coverage for relatively low costs of vaccination, yet impedes the incentive to vaccinate at high costs (see Fig. 35(a)). While in terms of completely eradicating the epidemic, imitation behavior fails and even enlarges the ultimate epidemic size, irrespective of vaccination costs (Fig. 35(b)). This unexpected outcome is related to behavioral clusters resulting from imitation dynamics, where vaccinators tend to contact vaccinators, and non-vaccinators tend to contact non-vaccinators (Fig. 35(c)). Although vaccination levels increase in some regions, clusters of non-vaccinators (susceptible individuals) exacerbate transmission and elevate the probability of an outbreak. In this regard, imitation behavior may shed new light into the high incidence of measles in many countries with high vaccine uptake [242,573–576]. Realistic policies should thus consider the impact of non-vaccinated clusters, and empirical data on vaccination behavior should be explored also in terms of geolocating potential hotspots of non-vaccinated individuals by public health authorities.

Besides payoff driven updating, opinion dynamics is another possible explanation for the large outbreaks of some vaccine-preventable diseases such as measles in many high-vaccination countries [577,578]. Recently, Salathé et al. assigned either a positive or a negative opinion about vaccination to each node [573], where positive opinion means that the agent will take up vaccination immediately, while individual with negative opinion reject vaccination due to personal beliefs, such as the safety and efficiency of vaccine, religious beliefs, and related considerations. Subsequently, an opinion formation process occurs based on the dissimilarity index and the strength of opinion formation. Interestingly, such a process leads to clusters of non-vaccinated individuals, since agents with negative opinion are more likely to contact same-minded counterparts. The invasion of a single infected agent into such a susceptible clusters may have dire consequences and dramatically decrease herd immunity. Moreover, the research also found that, when the vaccination coverage is close to or at the level required to provide herd immunity, this effect of opinion formation (or clustering) facilitating outbreaks is the strongest. This in turn
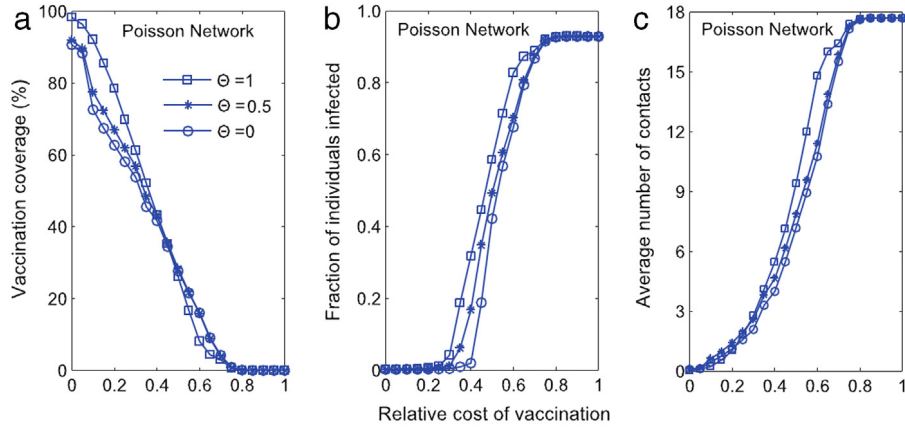
Fig. 35. Vaccination coverage (a), epidemic size (b), and average number of contacts between non-vaccinators are shown as a function of the relative cost of vaccination in passion (or homogeneous) networks. The fraction of individuals with imitation behavior is Φ, while the remaining $1 - \Phi$ follow a payoff maximization strategy.
Source: Reprinted figure from Ref. [433].

indicates that the estimated coverage level to avert an outbreak of measles may be much higher than realistic demands [548,575].

Quantitative characterization. Relative to the above primarily qualitative endeavors, Salathé et al. also defined a quantitative framework for the mixing patterns of vaccination behavior and opinions related to vaccinating [579]. Based on publicly available data from Twitter, they firstly constructed a directed social network of information flow, where sentiment toward new influenza A (H1N1)-like vaccination was classified into four types: positive, negative, neutral, and irrelevant. Then, building on the Pearson coefficient [176,520,580], the assortativity coefficient r is given as (relative to Eq. (99), the following formula is actually another expression of correlation coefficient, and their detailed comparison has been proved in [581])

$$r = \frac{\sum_i e_{ii} - \sum_i a_i b_i}{1 - \sum_i a_i b_i},$$
(232)

where

$$a_i = \sum_j e_{ij}, \quad b_j = \sum_i e_{ij}$$
(233)

and e_{ij} is the fraction of edges connecting nodes of types i and j. Moreover, here $r = 0$ means random mixing of opinions, while positive (negative) r values introduce assortative (disassortative) patterns, where nodes are more likely to connect with nodes of the same (different) type. Interestingly, the research found that vaccination is positively assortative with the sentiment score, which means that agents expressing positive (negative) sentiment are more often connected and take (reject) vaccination. In this sense, opinion distributions in different communities (each of which usually shares common opinion) profoundly affect the ultimate size of an epidemic. Results presented in Fig. 36 show that the likelihood of an outbreak is enlarged, if unprotected agents are positively assorted and form susceptible clusters (due to the clusters of negative sentiment). For further details on the role of vaccination sentiment in realistic networks, we also refer to Sections 10 and 11 for more details.

The aforementioned studies have proved that positively assortative non-vaccinated individuals in contact networks can lead to an enlarged likelihood of large disease outbreaks [433,573,579]. To extend the universality of this finding, Barclay et al. investigated whether such a pattern existed in influenza vaccination of contact networks among high schoolers [582]. Contact network were found to be positively assortative (measured by Eq. (232)) with respect to influenza vaccination in that vaccinated (non-vaccinated) individuals are more inclined to link with their like. Although the overall coverage level may be temporarily high, clusters of non-vaccinated individuals will destroy the herd immunity and may take over the largest component of the network in the long term. Relative to the states of randomly assigned vaccination, positive assortativity thus increases the likelihood of large disease outbreaks. Of particular importance, non-vaccinated males contribute more than non-vaccinated females toward positive mixing patterns. Another interesting point is that the effect of non-vaccination clusters (or positive assortativity) becomes more visible if contact durations are longer or combined.

9.2.3. Belief-decision models and dual-perspective systems

As surveyed in 9.2.2, except for payoff-driven vaccination decision-making (i.e., the game-theoretical approach) [432,524,528], vaccination choices can also be driven by the awareness concerning the disease, and the perception one
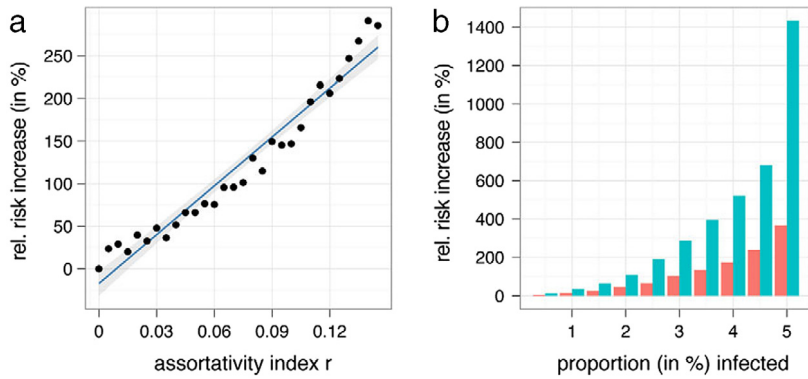
Fig. 36. (a) The impact of the assortativity index r on the relative risk increase (compared to risk at $r \sim 0$) of disease outbreaks. Blue line shows the best fit of linear regression (confidence interval based on standard error). (b) Relative risk increase (compared to risk at $r \sim 0$) of disease outbreaks of a given fraction of the population for two values of the assortativity index $r = 0.075$ (red) and $r = 0.145$ (green). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
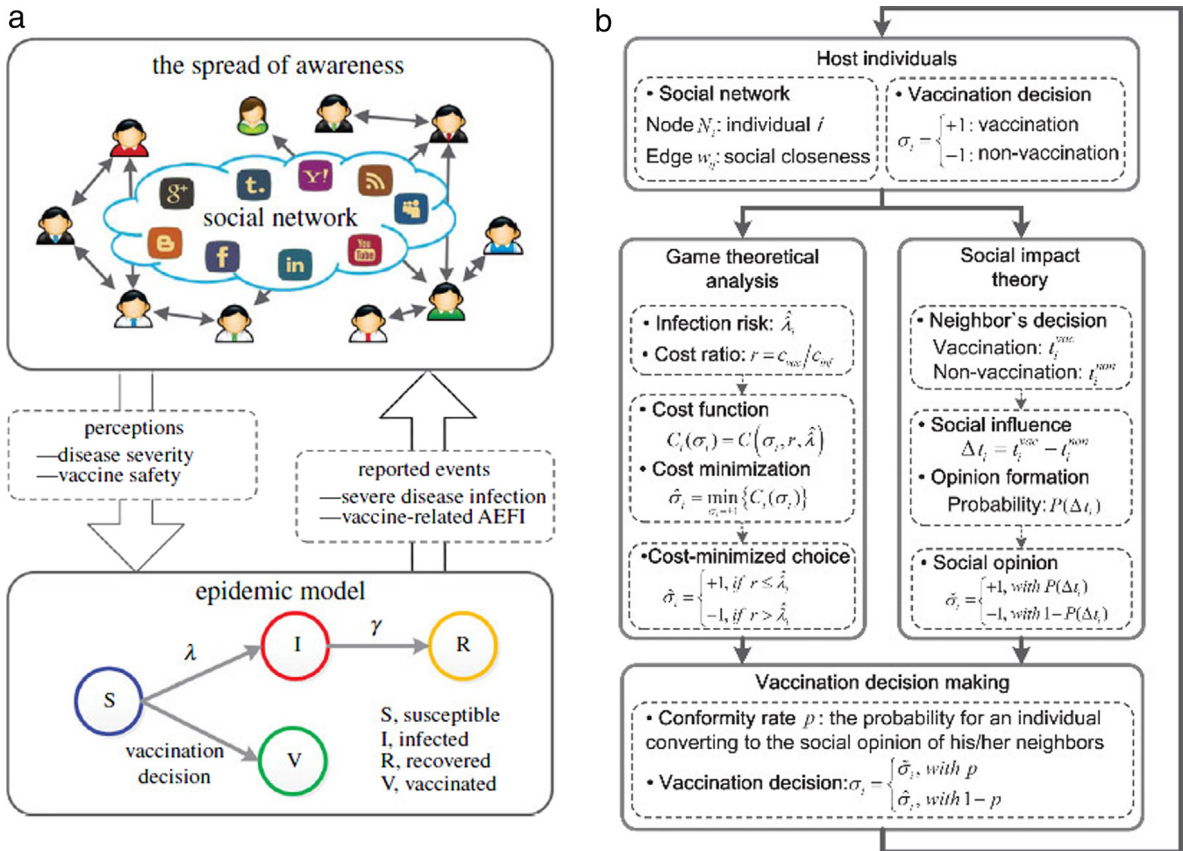*Source:* Reprinted figure from Ref. [579].

Fig. 37. (a) Schematic illustration of a belief-decision model for a vaccination campaign. A social network is used to describe the structure of interactions among individuals. There exists a feedback between the awareness and the vaccination choices of individuals. Namely, vaccination decisions depend on the perceptions about disease severity and vaccine safety. (b) Schematic illustration of a dual-perspective view on modeling vaccination decision making. Here, vaccination decisions (i.e., to vaccinate $+1$, or not -1) are in accordance with either cost minimization or the impact of social influence. The conformity rate p is the probability that an individual finally converts to the formalized social opinion, or otherwise follows a cost-minimization choice.
Source: Reprinted figure from Refs. [443,572].

has about the vaccine and its safety. Actually, awareness has been widely identified as a significant factor in other disease prevention measures such as social distancing [352,411,583–585]. Motivated by this fact, Xia et al. designed a belief-decision framework (see Fig. 37(a) for a schematic illustration), where three possible vaccination decisions, i.e., to accept, to reject, and yet to decide (or no decision), are assorted to a set of belief values [443]. The belief function for vaccination is given as

follows

$$
\begin{cases}
m(\text{Yes}) \in [0, 1], \\
m(\text{No}) \in [0, 1], \\
m(\Theta) = 1 - m(\text{Yes}) - m(\text{No}),
\end{cases}
\tag{234}
$$

where $m(\text{Yes})$ ($m(\text{No})$) means the probability of taking (rejecting) vaccination immediately, while $m(\Theta)$ represents no firm decision and postponing the decision to the next time step. Owning to the spread of awareness of disease severity and vaccine safety, individuals update their belief values based on the extended Dempster–Shafer theory (a theory of belief functions or a mathematical theory of evidence, it is a general framework for reasoning with uncertainty, with understood connections to other frameworks such as probability, possibility and imprecise probability theories. In details, such a theory allows one to combine evidence from different sources and arrive at a degree of belief that takes into account all the available evidence) [442], which incorporates information of connected neighbors. Interestingly, the authors found that the reporting rates of disease- and vaccine-related events mainly affect the vaccination coverage. While another key factor, the awareness fading coefficient, decides the time of individuals' vaccine administration.

After the above finding, another question naturally poses itself, namely, if both the game-theoretical approach and opinion dynamics are put together in a unified framework, how does this dual-perspective system influence the vaccination uptake and potential epidemic outbreaks? A computational model to evaluate the impact of vaccination associated cost and social opinions on vaccination decision-making (i.e., the interplay between payoff optimization and social influence) has therefore been proposed and studied by Xia et al. [572]. In more detail, the cost minimization scenario assumed that individuals (with mixed strategy) take up vaccination if the relative cost c of the vaccine is larger than the perceived infection risk λ_{perc} (i.e. $c > \lambda_{perc}$), where

$$
\lambda_{perc} = \beta_I \cdot \left(\frac{n_{non}}{n_{vac} + n_{non}} \right)
\tag{235}
$$

and n_{vac} (n_{non}) denotes the number of neighbors being vaccinated (not vaccinated). In the opposite case, individuals give up vaccination. Moreover, according to the social impact theory [586,587], to either vaccinate or not can be regarded as two types of opinions, which are acceptance and rejection, respectively. Individuals then update their opinions in terms of the Fermi-like function, where a normalized discrepancy between viewpoints replaces the previously used payoff difference [432]. Finally, the conformity rate p is introduced to control the tendency of individuals toward social opinions (see Fig. 37(b) for an illustration). When the conformity rate is relatively small (i.e., decision are mainly cost-based), the effectiveness of vaccination declines with larger costs, which is as reported before in the realm of a purely game-theoretical model [432]. However, when individuals follow social opinion (i.e., $p = 1$), the resulting vaccination converges to a certain level, depending on individuals' initial level of willingness rather than the associated cost. For intermediate conformity rates (i.e., the joint impact of cost-minimization and opinion formation), the ultimate vaccination coverage and epidemic incidence are completely determined by the competition between vaccine cost and initial vaccination willingness. Thus, in addition to the determinants associated with the costs and benefits of vaccination, a voluntary vaccination program should carefully consider also the social aspects and opinion formation in contact networks.

9.2.4. Competing controlling measures (going beyond traditional binary selection)

As mentioned in Section 9.1.4, apart from vaccination, there might also exist non-pharmaceutical interventions that may help mitigate the risks and seem of particular importance before effective vaccines are widely available [52,54,59, 411,502,588,589]. However, the majority of existing research considers either vaccinating behavior or non-pharmaceutical interventions separately. The effectiveness of the combination of both types of interventions on social or empirical networks may thus, at least to some extent, be beyond our knowledge (although both are used in some disease such as influenza). Aiming for a better understanding of this issue, Andrews et al. recently proposed a model of competing interventions, which assesses the decision on the basis of subjective expected utility theory [440]. It is worth mentioning that, unlike by vaccination, susceptible/infectious individuals will implement non-pharmaceutical interventions on different neighbors.

Of note, there exists a tradeoff between two measures. If the vaccination coverage increases, the decay of infections and the resulting changes in transmission patterns enable a decrease in the practices of non-pharmaceutical interventions. Similarly, if non-pharmaceutical interventions become more common, vaccination coverage will decrease by roughly a similar extent. Even if changing the utility or the perceived risk, it seems impossible to simultaneously enhance the effectiveness of both measures. Moreover, it was also revealed that efforts to reduce spreading by expanding non-pharmaceutical interventions are almost completely mitigated by the resulting drop in vaccine coverage, whereas efforts to expand vaccine uptake are only weakly affected by the response of non-pharmaceutical interventions. Thus, (lack of) vaccination interferes more strongly than non-pharmaceutical interventions as the primary determinant of influenza incidence. Along this line, when considering multiple intervention programs, officials should pay particular attention to the potential interference among them, in particular looking out for (dis)assortativity between vaccination and non-pharmaceutical interventions in real-life situations. It is important to communicate to the public that these are complementary rather than exclusive measures to mitigate an outbreak.

9.3. Behavior-vaccination dynamics on temporal networks

In the above subsections, we have reviewed the remarkable advances in understanding vaccination and the spread of epidemics in structured populations, which were described by different complex, yet static networks. While many empirical observations could be explained in this theoretical framework [431,432,524,579], it is nevertheless a fact that most social interactions among people change over time. Accordingly, a more apt description involves using time-varying or temporal networks [52,217,590–593]. Such adjustments may of course be unrelated to vaccination and epidemics, or they can happen precisely because of this. Either way, changes in the interaction structure of social networks affect how we vaccinate and how effective our measures are aimed at mitigating epidemics. In this subsection, we therefore cover an important aspect of behavior-vaccination dynamics in structured, networked populations, namely that on temporal networks. To that effect, we consider two representative examples.

9.3.1. Outcome inelasticity and outcome variability on adaptive networks

Early on in epidemiology, it has been pointed out that adaptations of various kinds can lead to a series of interesting phenomena, such as degree mixing patterns, oscillations, hysteresis, and phase transition between disease-free and endemic states [52]. Along this line one might wonder, if a vaccination program is placed onto an adaptive network, how do the epidemic incidence and vaccination uptake co-evolve with time? To address this question, Morsky et al. have recently proposed and studied an edge turnover effect on networks [594]. Naturally, if an individual in a network is vaccinated (not vaccinated) and dies from adverse side effects (an infection), any links from and to it are removed from the network (technically, since the links are undirected, simply an edge is removed). To maintain the balance in the system, i.e, the average number of edges in the network, the same number of edges are re-introduced randomly among pairs of individuals, whereby not allowing multiple connections. Interestingly, such a relatively simple and straightforward setup gives rise to three vaccination regimes. These are successful ring vaccination, widespread vaccination, and no vaccination, all owing to a different frequency of edge turnovers that are related to the death of individuals in the network. Moreover, it was shown that stochastic effects can lead to a large variability in outcomes. For the same parameter values, some realizations guarantee fast control, while others cannot prevent a fast-spreading epidemics. This indicates that the outcome of voluntary vaccination on adaptive networks may be highly unpredictable, precisely because the interaction structure changes over time. Thus, many arguments put forward above for static networks simply do not hold. There exist inherently nonlinear feedback mechanisms between behavior and disease dynamics, which can result in policy resistance or outcome inelasticity, neither of which would be observed on static networks in the traditional models that did not take strategic interactions or behavioral responses to disease dynamics into account [229,520]. This gives rise to extreme events that we are only beginning to uncover and study, and which certainly shed new light into just how complex successful vaccination policies might turn out to be.

9.3.2. The impact of closeness driven time-varying topology

Along similar lines, in [595] authors proposed a closeness-driven time-varying framework. In an instantaneous network, every individual connects to an acquaintance with probability p, or with probability $1 - p$ to a random node [596]. After an instantaneous network is formed in this way, the vaccination game proceeds with an imitation rule at the heart of the microscopic dynamics [432]. These two steps, i.e., the creation of the instantaneous network and the evolutionary vaccination game proceed repeatedly until a steady state is reached. In the absence of vaccination there is monotonous increase in the size of the epidemics. Conversely, if vaccination is introduced, the epidemic incidence reaches a peak and then drops due to the enhanced crisis awareness. More importantly, the increase of p greatly weakens the vaccination enthusiasm of individuals, since the links among acquaintances provide convenient conditions for herd immunity. These results indicate that closeness likely plays an important role in decreasing both the epidemic propagation and vaccination.

Although existing research already conveys clearly that temporal networks bring about novel phenomena in the mitigation of epidemics and the promotion of vaccination, there is still ample room for further research, for example related to the importance of different time scales that describe changes in the network, in the epidemic transmission, and in the vaccination uptake, where differences can markedly determine the fate of the population [597,598]. Further research on behavior-vaccination dynamics on temporal networks also requires better integration with empirical data, which we cover in more detail in Section 11 [599,600].

9.4. Behavior-vaccination dynamics on multilayer networks

In addition to static and temporal networks, the most recent frontier in network science are the so-called interdependent or multilayer networks [202,337,601,602]. The main premise behind multilayer networks is that a single layer of interactions is likely insufficiently accurate to describe the complexity of our social interactions. Arguably, each individual is likely a member in many social networks, including online social networks, and all these interactions shape markedly our perception of vaccination in the broadest possible sense. Although one may argue that all these interactions could be placed under the umbrella of a single network, the multilayer network perspective nevertheless creates better conditions for properly studying such social systems.
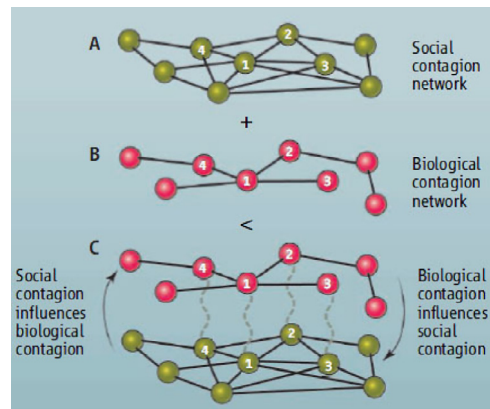
Fig. 38. Schematic illustration of behavior-vaccination dynamics on multilayer networks, which consist of social and biological contagion networks. The former layer supports the propagation of information related to the disease, such as opinions and sentiments, while the latter layer allows the actual spreading of the virus. More importantly, the resulting interplay on the new architecture leads to complex dynamics that is not present in either network layer in isolation, such as the rise and fall of vaccine coverage during a scare scenario.
Source: From Ref. [608]. Reprinted with permission from AAAS.

Not surprisingly, multilayer networks have recently been the subject of many novel frameworks to study various dynamical processes, such as epidemic spreading [215,230,371,603], diffusion [341], voting [604], synchronization [605], and the evolution of cooperation in game-theoretical models [606,607] (see [202,337] for systematic reviews). In this subsection, we therefore cover the still very young field of behavior-vaccination dynamics on multilayer networks.

9.4.1. The general framework and an example

With respect to coupled disease–behavior dynamics on multilayer networks, one recent work firstly provided a brief, conceptual framework, where the system is composed of two networks: the social contagion network [609,610] and the biological contagion network [608]. The former is assumed to support the spreading of opinions and sentiments related with vaccination, while the latter allows the propagation of the virus through local intra-layer connections. When both networks are coupled together, a feedback loop between layers spontaneously emerges, which generates more complex dynamics beyond either isolated network layer (see Fig. 38 for a schematic illustration). For instance, the vaccination opinions on the social contagion network may play a positive or a negative role in the epidemic trajectory on the biological contagion network and vice versa.

Along this main idea, Fukuda et al. recently proposed a multilayer infrastructure, which is composed of a disease transmission network and an information propagation network through which individuals update their vaccination strategies and health information [611]. Compared with the evolutionary outcomes in the traditional single-layer network [432], it has been discovered that multilayer infrastructures typically suppress the vaccination coverage, resulting in relatively larger and more common epidemics. Although middle-degree and large-degree nodes show higher vaccination enthusiasm, numerous small-degree nodes make the propagation of vaccination very difficult, which is further exacerbated by the propagation of negative opinions in the other layer. Overall, successful vaccination policies in multilayer networks require even more care and consideration to account for multiple channels along which potentially discouraging information about vaccinating may spread, as we review also in the next subsection.

9.4.2. Parallel, interacting dynamics on different networks—The impact of opinion clusters

As in the case of [433,573,579,582], when opinions of individuals toward vaccination are influenced by their neighbors or social groups on networks, the susceptible/unvaccinated clusters can readily arise, which directly affects the likelihood of disease outbreak. Even with very high vaccination coverage, outbreaks can still occur via these clusters [433,573]. However, empirical evidence suggests that the transmission of infections is sometimes separated from decision-making. One typical example for this reasoning are childhood diseases, for which the primary risk group are young children, while the decisions about vaccination are of course taken by their parents [549,613–615]. Motivated by this fact, already in 2009 Eames used a multilayer network setup, where one network supports the opinion formation among parents (i.e. parent network), while the network of children allows infection transmission among offspring (i.e. offspring network) [612]. When parents and children are distributed on each network layer, offspring share their parents' locations (denoted by the inter-layer links) and intra-layer connections, whose fraction is controlled by the parameter F. $F = 1$ means that each intra-layer link in the offspring network also appears in the parent network, i.e. both network layers are identical; while $F = 0$ indicates that parent and offspring networks have no shared intra-layer connection. Using an identical opinion updating dynamics as in [573], it was reported that even if the vaccination level is extremely high, the opinion formation process greatly enhances the outbreak likelihood, relative to the case without opinion dynamics (see Fig. 39(a)). Because the opinion of not to vaccinate is inclined to clustering among parents, this naturally gives rise to susceptible offspring clusters and thus creates significant
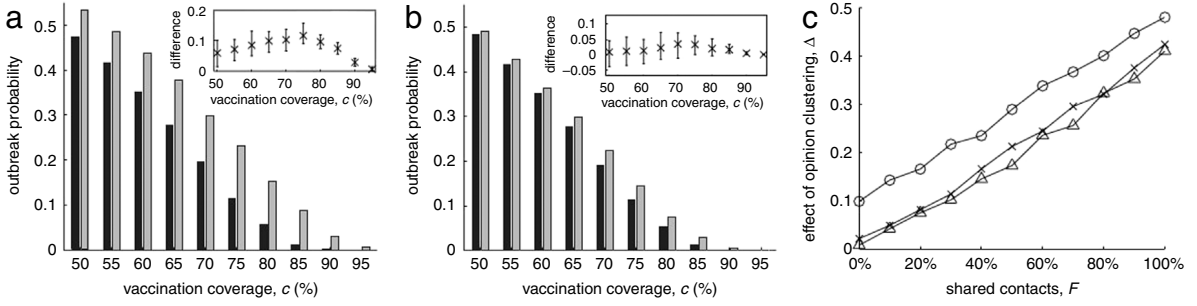
Fig. 39. Outbreak probability with (gray) and without (black) opinion formation in dependence on the vaccination coverage. The insets show the difference of outbreak probability between gray and black vertical lines. The parent and offspring networks share $F = 1$ (a) and $F = 0.2$ (b) of the intra-layer connections. (c) The impact of shared connections F on the effect of opinion clustering Δ. Circles, crosses and triangles represent the probability of forming a intra-layer connection equal to 0.1, 0.2 and unlimited, respectively.
Source: Reprinted figure from Ref. [612].

infection risks. Interestingly, the above observation still remains intact, though less dramatic, when the overlap fraction is low, as shown in Fig. 39(b).

To quantify the difference between outbreak probability with and without opinion formation, in [612] the author further defined the effect of opinion clustering

$$\Delta = \frac{\sum\limits_{\text{vaccination levels,} c} \{\text{Prob (Outbreak with opinion)} - \text{Prob. (Outbreak without opinion)}\}}{\sum\limits_{\text{vaccination levels,} c} \text{Prob. (Outbreak without opinion)}}. \tag{236}$$

As shown in Fig. 39(c), the effect of social influence monotonously increases with the amount of overlap between the two networks. When the two networks share many contacts, clusters of parents who choose not to vaccine are more likely to have interacting children, generating clusters of unvaccinated offspring and thus ultimately enhancing the risk of an infection, which is then likely followed by an outbreak. In this sense, non-overlapping intra-layer connections between the networks of parents and their offspring seem to be an important factor for determining the desired vaccination coverage for a successful mitigation of disease.

While the above-reviewed research already delivered inspiring results, just as with temporal networks, the study of behavior-vaccination dynamics on multilayer networks is still very much a developing field with vast possibilities for innovative future research. The evolution of links between different network layers, as well as their weight in determining the course of vaccination in comparison to within-layer links are still very much under-explored subjects, just as potential extension of the basic theoretical framework to more than three layers.

So far, we have surveyed most achievements about behavior-vaccination dynamics in networked populations, which, as a whole, provides novel insight into real-world disease prevention under vaccination policy. To obtain a clearer understanding, Table 6 provides a summary of these new characteristics. We also notice because temporal networks and multilayer networks are two novel frameworks in network science, more efforts from statistical physics and mathematics are still needed to enrich the content of behavior-vaccination dynamics in future. Moreover, another fantastic point is that, except for infectious diseases, complex contagion [364] and corresponding behavioral vaccination strategy [395] deserve great attention as well, especially combining with lots of emerging data in this realm.

10. Measuring and assessing vaccination levels: traditional approaches

Due to the overwhelming public health benefits of vaccination, governments and public health institutions around the globe have invested substantial efforts to immunize their populations against vaccine-preventable diseases. The population dynamics of vaccine-preventable diseases depend strongly on the magnitude and distribution as well as the efficacy of vaccines in the population, but measuring these indicators is not always easy. First, measuring the vaccination coverage – the percentage of the population that has been vaccinated – is subject to numerous difficulties and biases. Second, even when vaccination coverages can be measured accurately, measuring the distribution of vaccination in the population – i.e. which parts of the population have been vaccinated – remains challenging. Third, the effect of vaccination on population immunity will ultimately depend crucially on the effectiveness of the vaccine, a hard-to-measure but crucial indicator for many public health policies. In this section, we review the challenges associated with assessing the vaccination status of a population. These challenges are not only relevant from a data perspective, but they are also crucial to inform modeling efforts.

10.1. Measuring vaccination coverage

In an ideal world, vaccination coverage would be easy to assess. Every individual in a population who receives a vaccine would be centrally registered, allowing for an always accurate calculation of the percentage of the population vaccinated.

Table 6
Summary of behavior-vaccination in network populations.

Used networks	Employed methods	Novel characteristics & effects of behavior-vaccination	Observed physics phenomena
Lattice ER graphs SW networks SF networks	Monte Carlo simulation mean-field theory stochastic processes	① Updating dynamics and network topology profoundly affect vaccination level and epidemic incidence; ② external incentives have the dual role in vaccination uptake; ③ there may exist tradeoff in the competition of multiple prevention strategies.	① Second-order phase transition of vaccination/infection; ② self-organization and percolating of non-/vaccination clusters; ③ positive feedback between vaccination and epidemic
Social contact networks Urban networks Technological networks	Monte Carlo simulation data analysis network creation social influence theory	① Efficiency of behavior-vaccination is closely related to contact patterns, vaccine/disease risk and external incentives; ② the emergence of non-vaccination clusters could lead to outbreaks in spite of high vaccination level; ③ opinion/awareness plays a significant role in vaccination campaign and disease prevention.	① 'First-order' phase transition of vaccination/infection; ② assortative pattern of strategies/opinions; ③ positive feedback between opinion/awareness and vaccination
Temporal networks Adaptive networks	Monte Carlo simulation Network creation	Outcome of voluntary vaccination becomes highly unpredictable	Nonlinear feedback between vaccination and disease
Multilayer networks	Monte Carlo simulation Social influence theory	Opinion/awareness on one layer may impede/accelerate disease spreading on other layers	Positive/negative feedback among dynamics of different layers

Fig. 40. A vaccination paper card, showing the various improvisations made over time. (Photo courtesy of Carolina Danovaro, Pan American Health Organization.)
Source: Reprinted figure from Ref. [616].

In practice, numerous problems stand in the way. On the numerator side, the number of people vaccinated is rendered inaccurate by the non-continuity of registration systems, by migration flows, and record duplications, as well as other challenges. On the denominator side, estimates can be equally inaccurate for a number of reasons. The various challenges are particularly difficult in low- and middle-income countries, where registry systems are often manual, patchy, or entirely absent.

The potential inaccuracy in estimates at the administrative levels means that measures from alternative data sources need to complement the numbers. Surveys in particular are a well-established source for alternative vaccination coverage estimates. Globally, there are four types of surveys that are commonly employed: Demographic and Health Surveys (DHS), Multiple Indicator Cluster Surveys (MICS), the Expanded Programme on Immunization cluster survey (EPI), and surveys using the Lot Quality Assurance Sampling (LQAS) method. These common surveys have different characteristics with respect to the primary objectives, the sampling scheme, household selection, total sample size, questionnaire length, duration, and so on [616]. They all suffer from a number of biases, of which two are typically highlighted: (i) a selection bias due to sampling methods that are non-representative or reflective of poor field practices, and (ii) an information bias generally – but not exclusively – due to inaccurate data recording manual records involving a paper card retained in the household (Fig. 40). Many of these problems are expected to worsen in the future because of the increased complexity of immunization schedules.

In recent times, biomarker surveys have been added to the methods to measure vaccination coverage. While the presence of antibodies is a good marker of immunity, it is typically impossible to distinguish whether their presence is due to

vaccination or to natural infection. Furthermore, while natural infection often causes life-long immunity, most current vaccines require booster shots because of the waning levels of antibodies following vaccination [617]. On the other hand, the case is reversed for some pathogens such as tetanus, where natural infection causes a short-lived immunity while vaccination can provide life-long immunity [618]. Some vaccines, such as the Hepatitis B vaccine, cause the presence of antibodies that are not created following natural infection [619], providing a clear measurement of vaccination, but this is the exception rather than the rule. And because of waning antibody levels, the number of doses can still not be reliably estimated [620]. To complicate matters further, re-exposure to pathogens can trigger an increase of antibody levels ("immune boosting"), often without developing symptoms, as appears to be the case in whooping cough [621–624]. Indeed, it has been suggested that immune boosting must be more easily triggered than primary infection to account for age-incidence data found in the recent increase in whooping cough cases in teenagers and adults [625].

10.2. Vaccine efficacy and effectiveness

Many studies looking at the effects of vaccination on disease dynamics at the population level assume that vaccines have 100% efficacy. In other words, they assume that in the vaccinated group, the disease incidence following vaccination is zero. This assumption is generally wrong. At its most basic level, vaccine efficacy can be defined as the opposite of the relative risk (RR) among vaccinated individuals [626]. It corresponds to the relative reduction of disease attack rate in the unvaccinated group (ARu), compared to that of a vaccinated group (ARv), all else being equal:

$$\frac{ARu - ARv}{ARu} = 1 - RR. \tag{237}$$

In order to comply with the "all else being equal" requirement, vaccine efficacy is usually assessed through double-blind, randomized, clinical controlled trials, representing best case scenarios. The vaccine efficacy so established is usually a requirement for a new vaccine to be licensed by the relevant governmental authorities.

Vaccine efficacies vary widely. In influenza, where a new vaccine is developed each year to prevent seasonal influenza, vaccine efficacy can vary dramatically depending on the match of the vaccine with the dominant circulating flu strains. A recent meta-analysis [627] has shown that influenza vaccine efficacy has been measured to be between 8% and 93%, depending on the vaccine type, the year, and the study population. The pooled efficacy of the trivalent inactivated vaccine in adults of age 18–65 was 59%. For the vaccine preventing Human Papillomavirus (HPV) infections, meta-analyses have shown a vaccine efficacy of 78%–90% [628]. The efficacy of an acellular pertussis vaccine among adolescents and adults was shown to be 92% [629].

Vaccine effectiveness needs to be distinguished from vaccine efficacy. While vaccine effectiveness measures the same outcome of interest – relative reduction of disease attack rate – it does so in an observational way under field conditions, rather than in clinically controlled experiments. Vaccine effectiveness is therefore influenced by many factors, including, but not limited to, vaccine efficacy. These factors include vaccine characteristics as well as host factors such as age, previous vaccinations and infections, etc. It is easy to see that all the challenges relate to vaccine coverage surveillance outlined above apply also to studies of vaccine effectiveness. Because vaccine effectiveness studies do not require clinically controlled trials, there are more data available on vaccine effectiveness than on vaccine efficacy. Vaccine efficacy is often assessed during or after outbreaks. For example, recent data obtained from a large-scale outbreak of measles in a German public school was analyzed resulting in an estimate of the measles vaccine effectiveness of over 99% (following two doses of vaccination) [630].

Vaccine effectiveness and vaccination coverage combine to create the dynamics of infectious diseases observed in populations. In all populations, there will be non-vaccinated individuals, and there will be individuals for which the vaccine does not prevent disease. During a disease outbreak, we can therefore expect both unvaccinated and vaccinated individuals to fall ill. Under the assumption of random mixing, the fraction of unvaccinated individuals among individuals with disease can be easily calculated. The susceptible population is the sum of the number of unvaccinated individuals U, and the vaccinated individuals who remain susceptible despite vaccination, V, i.e. $S = U + V$. If these two types of susceptible individuals get infected with the same probability, then the fraction of unvaccinated individuals among diseased individuals will be $U/(U + V)$. Expressed in terms of vaccination coverage x and vaccine effectiveness τ, individuals can be susceptible because they did not get vaccinated $(1 - x)$ or because they did get vaccinated but the vaccination did not render the individual immune ($x(1 - \tau)$). Thus, the total susceptible fraction of the population is $1 - x + x(1 - \tau) = 1 - \tau x$, and the fraction of unvaccinated individuals among diseased individuals is

$$\frac{x(1 - \tau)}{1 - \tau x}. \tag{238}$$

This relationship can be leveraged for multiple purposes. First, it can be used to estimate vaccine effectiveness. If the proportion of cases occurring in vaccinated and unvaccinated as well as the vaccination coverage x is known, vaccine effectiveness τ can be easily calculated [631]. Secondly, it can be used to estimate vaccine coverage, if the proportion of cases occurring in vaccinated and unvaccinated as well as the vaccination effectiveness τ is known. This can be particularly helpful in situations where the vaccine effectiveness is well known, such as in the case of measles [632]. An increase in the fraction of cases among unvaccinated individuals may seem counter-intuitive and worrisome, but indeed is most likely

an indicator of increasing vaccination coverages. The opposite case is just as true, and particularly relevant in situations of increasing vaccination hesitancy.

Overall, the ability to measure, or at least accurately estimate vaccination and immunization levels in a population are of crucial importance to public health, and the modeling efforts. As outlined above, the practical challenges have historically been quite substantial, and continue to pose problems. At the same time, the advance and widespread adoption of technology, and in particular also consumer technology such as the internet and mobile computing devices, can help supplement these efforts. In the next chapter, we will discuss the approaches based on a new set of methods broadly encapsulated under the term digital epidemiology.

11. Digital epidemiology

As discussed in many of the previous sections, the ability to timely monitor the unfolding of outbreaks, as well as the capacity to capture our contacts, movements and communications are keys to characterize the spreading of infectious diseases. This observation is quite intuitive and clear; perhaps since the very first epidemiological studies. However, major challenges in the data collection process have impeded developments for many years. Let us think first about surveillance systems. In many countries, health departments have the mandate to collect epidemiological data. In the USA for example, the surveillance for influenza like illnesses (ILI) is conducted by a network of about 3000 healthcare providers [633]. As we noted in Section 10 traditional surveillance approaches are affected by many limitations [634]. Data is typically affected by large delays and backlogs. The geographical granularity of the sample is heterogeneous across the country. Furthermore, selection biases might play a major role. Indeed, only a fraction of infected individuals visit a doctor. These cases are typically associated with complications or legal requirements (leave from work). Data collections challenges have been even more pronounced in the observation of human dynamics relevant for diseases spreading. Indeed, the characterization of the way we interact, move and communicate has been relegated to ad-hoc small-scale studies for many years. The inherent complexity of such processes strongly limited the attempts to study their properties in these settings [635,636].

The digital revolution helped, and is helping, to lift such limitations [637]. Nowadays more than 3 billion of people had access to the Internet. About 7 billion phone subscriptions have been activated, and more than 20% of them are associated to smartphones. As result a large fraction of our activities is digital, often online. Communication using social networks, emails, blog posts, and other forms of digital media are now commonplace. Countless online resources are now used to access information that researchers can efficiently mine using search engines. Furthermore, the miniaturization of devices created a wide range of wearable able to measure our interactions, movements, as well as vital signs [638].

It did not take long before researchers saw the potential of these new technologies for epidemiology [639–641]. Indeed, the large amount of data we generate, or that we can collect, contains crucial epidemiological indicators. Just to mention few examples, think people speaking about their health on social networks, or searching for diseases' symptoms on search engines. Furthermore, think about the uses of GPS data to characterize human mobility at different geographical scales, of RFID (radio frequency identification devices) tags to identify the features of human face-to-face contacts, or of mobile apps to gather health information from users that sign-up. This unprecedented wealth of data is at the basis for the development, test, and refinement of many modeling approaches discussed in the previous sections.

In the following, we will review recent developments in this arena (also see Table 7 for a systematic summary in advance). In doing so, we will consider three main approaches according to the data collection process. We will first consider participatory methodologies that rely on individuals voluntary contribution to the data collection. We will then present passive approaches that identify and extract epidemiological relevant indicators from the digital traces of our daily activities. We will then conclude presenting mixed approaches of the previous two.

11.1. Participatory digital epidemiology

The widespread diffusion of Internet connections and the miniaturization of devises is providing unprecedented opportunities to collect epidemiological relevant information with the direct participation of large number of individuals.

11.1.1. Participatory platforms for digital surveillance

Mobile and web-based applications can be effectively used to monitor the emergence and unfolding of outbreaks in near real-time [642,643]. FluNearYou [644,645], and InfluenzaNet [646,647] are two examples. The first platform engages individuals living in USA or Canada. The second instead is based in Europe and includes 11 different countries. The main idea behind both initiatives is to complement the classic ILI surveillance systems crowdsourcing the data collection. Indeed, users that sign-up become sentinels. Every week, during the influenza season, they receive a request to answers few questions covering geographical, behavioral and medical information. In other words, the two platforms are systems for digital surveillance. Remarkably, several studies have shown that the ILI incidence estimated using these platforms is highly correlated with the classic ILI surveillance data [645,647,648]. With minimal participation from the subscribed users they allow to overcome several limitations of the classic systems. The data collection is inexpensive, and scalable. The data can be visualized in near real-time, providing interactive maps describing the current unfolding of ILIs that are readily available for modeling efforts at the sub-population level as those described in Sections 3 and 8. Another crucial advantage of these

Table 7
Summary of approaches in digital epidemiology.

Method	Data collection process	Relevance for epidemiology	Main advantages	Main disadvantages
Participatory digital epidemiology	Surveillance platforms and wearable devices	Estimation of disease indicators and characterization of face-to-face interactions	Near real-time estimations, access to different geographical granularities, detailed characterization of social interactions responsible for the spreading of human to human diseases	Sample size is typically small, reliance on users' participation
Non participatory digital epidemiology	Social media and mobile phones	Estimation of disease indicators, characterization of social interactions and human mobility	Large sample size, cheap data collection process, near real-time analysis, access to different geographical granularities, characterization of human mobility and its impact in disease spreading	Correlations might be spurious, the interactions observed might not be relevant for disease spreading, keyword selection not trivial, changes in platform design might influence users' behaviors
Mixed methods	Social media and users participation	Estimation of disease indicators and discovery of new outbreaks	Near real-time estimations, data classification improved by users' participation	Performance tied to users participation

approaches is the reduction of back-logs and delays in cases reporting. This issues is particular relevant during the initial phases of the flu season when its severity and general characteristics are mostly unknown. The delays and back-logs of the traditional surveillance introduce a posteriori change in the projections making forecasts a difficult task. In other words, the "ground truth" used to inform models is far to be stable and reliable. Instead, data from digital surveillance can be used to inform realistic and geographically resolved epidemiological models with timely and high resolution intelligence that can be adopted to define precise initial conditions necessary for the simulations [649].

Another important aspect of the data collection is the ability to access behavioral information as for example the vaccination status of each user. On one side this allows geographically resolved real-time evaluation of vaccination rates complementing those gathered via traditional approaches (i.e. Section 10) and might be used to refine the input for the modeling efforts described in Sections 3 and 8. On the other side, it allows to estimate the vaccination efficacy overcoming the issues of classic observational studies [650]. These are often cohort studies performed considering, retrospectively, health records of vaccinated and not vaccinated patients. Clearly, the same selection biases, black-log and delays we saw afflict clinical surveillance are a problem also in this case. Digital surveillance approaches, although not perfect, clearly mitigate these issues. For example, while selection biases can be still a problem, well crafted questionnaires can alleviate them. Again, the main advantage is the scalability of such methods, and the possibility to perform near real-time analysis.

When considering participatory platform for digital surveillance, it is important to stress that these application have been developed to complement, not substitute, classic surveillance systems. Indeed, the detailed characterization of circulating strains is crucial and can be achieved, at the moment, just in clinical settings. However, the penetration of mobile and connected devices in developing countries is offering unprecedented opportunities to gather richer and timely epidemiological data in cases where classic surveillance methods are typically limited.

11.1.2. Measuring and understanding close proximity interactions

The data and collection methods we just introduced, provide critical information to estimate diseases prevalence and vaccination rates. However, they do not provide any explicit information about the spreading pathways. From a modeling standpoint, such data can inform just mean-field modeling efforts (see Sections 3 and 8). Our close proximity contacts are responsible for the spreading of infectious diseases. As discussed in Section 4 social networks are far from homogeneous and their nontrivial properties have critical effects on the spreading of infectious diseases. For many years, empirical studies devoted to characterize their features have been centered on surveys and diaries [651]. Potentially, these methods allow to gather information regarding the number, location (work, home, or other places), duration, and frequency of contacts as well as the health status of the respondent [652,653]. However, they are affected by several limitations as the scalability in terms of costs and individuals engagement [652], the introduction of biases by the self-reporting of events/contacts [652,654–656], and the introduction of biases by the survey design [652].

The incredible advances in technology as the reduction of costs and sizes of devices allowed the development of a wide new spectrum of methods to measure close proximity contacts. In the following, we will focus on those that are participatory in nature: individuals are asked to voluntarily participate to a data collection process where their close interactions are

measured via some device as smartphones or RFID tags [393,657–670]. According to the technology used, the data collection can span few hours, days, or months. Smartphones allow longitudinal data collections. However, they can record proximity interactions (via bluetooth or WiFi) that are not necessarily relevant for the spreading of infectious diseases [651,659–661, 664,666,667,670]. Furthermore, as we will see later, smartphones allow to access/collect a rich set of data that can be used to better understand different aspects of human dynamics ranging from digital interactions to mobility patterns [670]. RFID tags instead can be tuned to detect actual face-to-face interactions. However, they require (i) participants to wear specifically designed devices (ii) the presence of stationary nodes deployed to collect information about the location of each interaction [651,671]. For these reasons, the data collection with this technology is limited to short time periods and controlled settings.

In general, all these tools have major advantages respect to traditional methods. First, technology makes their deployment cheap, simple and scalable. Second, they provide an objective measurement of close proximity contacts. Few recent comparative studies have shed some light in the difference between contact networks estimated by traditional approaches (surveys and diaries) and those gathered via wearables [655,672]. The findings show that traditional methods miss the large majority of short contacts, and overestimate the duration of all interactions. Furthermore, the engagement of individuals with traditional methods is reduced respect to those based on wearables. Notwithstanding these limitations, traditional methods reproduce with good accuracy strong ties that are likely to be responsible to sustain the local unfolding of infectious diseases. However, they provide an incomplete picture. Indeed, they miss weak ties that have been identified as crucial connections to bridge separated clusters of individuals, thus fundamental for large scale outbreaks [347,672,673].

The analysis of the data collected with wearable devices allows to characterize with precision the features of close proximity contacts. The first important observation is that contact networks are time-varying [218,651,674,675]. The distribution of individuals activity, the duration and time intervals of contacts are typically heterogeneous implying the absence of characteristic scales, i.e. averages are not good descriptors. Consequently, the role played by each individual in the spreading is heterogeneous. Furthermore, for diseases in which the probability of transmission is function of the time spent in contact, different interactions might be associated with very different transmission potentials. In particular, the large majority of contacts are short and thus associated with small transmission probabilities. However, some contacts have duration that is much larger than the average and thus can be extremely important for the spreading. The heterogeneity in time intervals between contacts, burstiness [676], has been associated with a rich, case-dependent, phenomenology that can either slow down [677–682] or speed up [683,684] the spreading. Interestingly, the dynamics ruling the evolution of contacts change depending on the time-scale considered [651,674]. In other words, the mechanisms driving the formation of contacts at the time-scale of minutes are different than those at the time scale of months [674]. This is of particular importance when studying the spreading of infectious diseases [229,592,595,683,685–703]. Indeed, contacts dynamics that affect the spreading of transmissible illnesses are those unfolding at comparable time-scale respect to the disease [218,651,675,685]. For example, the characteristics of contacts networks at the time-scale of days are relevant for the spreading of ILIs. In order to show more clearly this crucial point we will now consider a simple, yet not trivial, model of temporal human interactions. This modeling framework is based on the empirical observation that the propensity of individuals of being engaged in social acts, _activity_, is heterogeneous [595]. Interestingly, such observations have been performed in a wide range of real datasets describing different types of human interactions or human dynamics ranging from conversations on Twitter to R&D alliances between firms [595,704]. For a given time interval Δt the activity of each node i can be measured as the fraction between the number of interactions made by i, n_i, divided by the total number of interactions made by all the nodes:

$$a_i = \frac{n_i}{\sum\limits_l n_l}.$$ (239)

While the value of a_i might change in time [705], observation in real data shown that the distribution of activity is virtually independent of the choice of Δt [595,703,704]. This candidates the activity as good variable to describe some important aspects of time-varying networks. The modeling framework we are describing starts from such evidence and generates so called activity-driven networks [595]. Each node is assigned to an activity a extracted from a distribution $F(a)$. At any time step t the network G_t is build starting from N disconnected vertices. In their simplest form the creation of activity networks is as follows:

- With probability $a_i \Delta t$ the node i becomes active and generates m links connected to m randomly selected nodes
- With probability $1 - a_i \Delta t$ the node will not be active. However, it will still be able to receive connections from other active vertices.

To reproduce empirical observations, activities are extracted from heavy-tailed distribution as for example $F(a) \sim a^{-\zeta}$ with $a \in [\epsilon; 1]$. The average number of active nodes at each time step is $N \langle a \rangle$ that is in general much smaller than N. Consequently, the large number of nodes will be not active and the topology of each G_t will be based on set of mostly disconnected stars centered around active nodes. On the other end, it is possible to prove that integrating links over sufficient T time-steps such that $k/N \ll 1$ and $T/N \ll 1$ the resulting network will have a degree distribution that follows the activity distribution [595,695]. In other words, the heterogeneity in the number of contacts integrated over time is driven by the heterogeneity in the propensity of nodes to be engaged in social acts. In this model, _hubs_ emerge in time due to their constant engagement

rather than due to some first mover advantages as in classic preferential attachment models. In general, the full dynamics of the network and its ensuing structure is completely encoded in the activity distribution. Interestingly, it is possible to derive the epidemic evolution equation in which a spreading process and the network dynamics are coupled together. We will consider an SIS model, but the same results holds for an SIR. As done throughout the review, let us define β_I the infection probability per contact, and γ the recovery rate. At a mean-field level, we can assume that nodes with the same activity a are statistically equivalent. In these settings, we can describe the epidemic process considering the number of infected individuals in the class of activity rate a, at time t, as I_a^t. The number of infected individuals of class a at time $t + \Delta t$ is given by

$$I_a^{t+\Delta t} = I_a^t - \gamma \Delta t I_a^t + \beta_I m (N_a^t - I_a^t) a \Delta t \int da' \frac{I_{a'}^t}{N} + \beta_I m (N_a^t - I_a^t) \int da' \frac{I_{a'}^t a' \Delta t}{N}, \tag{240}$$

where N_a is the total number of individuals with activity a [595]. In Eq. (240), the second term on the right hand side refers to infected nodes that recover, the third term takes into account the probability that a susceptible node of class a is active and acquires the infection getting a connection from any other infected individual, while the last term takes into account the probability that a susceptible node, independently of his activity, gets a connection from any infected active individual [595]. In order to solve Eq. (240) we can consider the total number of infectious nodes in the system by summing over all the activity classes:

$$\int da I_a^{t+\Delta t} = I^{t+\Delta t} = I^t - \gamma \Delta t I^t + \beta_I m \langle a \rangle \Delta t I^t + \beta_I m \theta^t \Delta t, \tag{241}$$

where $\theta^t = \int da' I_{a'}^t a'$ and we have dropped all second order terms in the activity rate I_a^t. Indeed, we are interested in the early stages of the spreading where the number of infected individuals is much smaller than the number of susceptible nodes. In this regime the equation can be linearized. In order to obtain an closed expression for θ we multiply both sides of Eq. (240) by a and integrate over all activities, obtaining the auxiliary equation

$$\theta^{t+\Delta t} = \theta^t - \gamma \Delta t \theta^t + \beta_I m \langle a^2 \rangle I^t \Delta t + \beta_I m \langle a \rangle \theta^t \Delta t. \tag{242}$$

In the continuous time limit we obtain the following closed system of differential equations

$$\partial_t I = -\gamma I + \beta_I m \langle a \rangle I + \beta_I m \theta, \tag{243}$$

$$\partial_t \theta = -\gamma \theta + \beta_I m \langle a^2 \rangle I + \beta_I m \langle a \rangle \theta, \tag{244}$$

whose Jacobian matrix has eigenvalues

$$\Lambda_{(1,2)} = \langle a \rangle \beta_I m - \gamma \pm \beta_I m \sqrt{\langle a^2 \rangle}. \tag{245}$$

The epidemic threshold for the system is obtained requiring the largest eigenvalue to be larger than 0, which leads to the condition for the presence of an endemic state [595]:

$$\frac{\beta_I}{\gamma} > \frac{1}{m} \frac{1}{\langle a \rangle + \sqrt{\langle a^2 \rangle}}. \tag{246}$$

The epidemic threshold can be derived considering that the per capita infection rate is given by the infection probability per contact, β_I, times the contact rate, $\langle k \rangle$: $\lambda_I = \beta_I \langle k \rangle$. It is easy to prove [595] that $\langle k \rangle = 2m \langle a \rangle$, and consequently

$$\frac{\lambda_I}{\gamma} > \frac{2 \langle a \rangle}{\langle a \rangle + \sqrt{\langle a^2 \rangle}}. \tag{247}$$

The threshold is function of the first and second moments of the activity distribution thus it takes into account the dynamics of interactions. Furthermore, the epidemic threshold is not function of the time-aggregated network presentation. It depends just on the interaction rate of nodes. This results show the importance of time-scales. Indeed the spreading condition is dependent on the interplay between the time-scales of the network and spreading process. It is important to remember how in the case of uncorrelated annealed networks the threshold is function of the first and second moment of the degree distribution (see Section 4 for the detailed derivation). In Fig. 41, we show the density of infected individuals in the endemic state as a function of λ_I/γ for three different networks [595]. In green diamonds and blue squares we show the results obtained running an SIS model in static networks generated aggregating links of an activity-driven network over $T = 40$ and $T = 20$ time-steps respectively. In red circles instead, we show the results obtained running an SIS model on activity-driven networks in which both links and disease evolve in time. Clearly, the results are quite different and show the effects of dynamical connectivity patterns on disease spreading. In particular, the threshold in static networks is significantly smaller than in their time-varying networks counterparts. Indeed, in static networks the disease can spread using all links activated in a certain time period independently of their activation time or order. In time-varying graphs instead, the disease spreads over a different, more homogeneous, topology in which at each time-step just few links are available for the contagion.
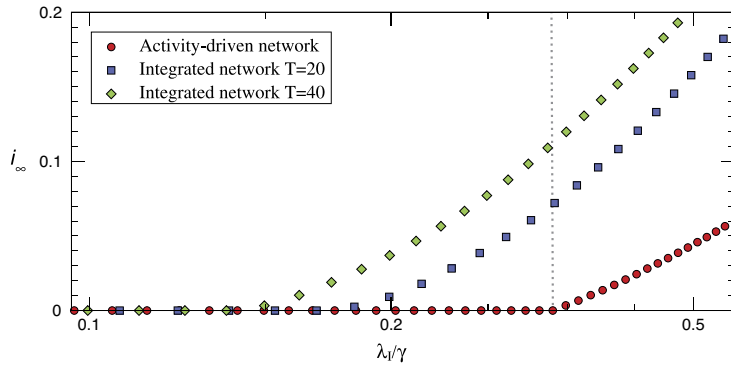
Fig. 41. Density of infected nodes, i_∞, in the stationary state, obtained from numerical simulations of the SIS model on activity driven networks and two others resulting from an integration of the model over 20 and 40 time steps, respectively. $N = 10^6$, $m = 5$, $F(a) \propto a^{-\zeta}$ with $\zeta = 2.1$ and $\epsilon \le a \le 1$ with $\epsilon = 10^{-2}$. Each point represents an average over 10^2 independent simulations. The dashed line marks the analytical epidemic threshold in Eq. (247).

Consequently, the disease will need a larger per capita infection rate or smaller recovery rate in order to percolate in the system.

In their simplest form[1] activity-driven networks consider a very simple and unrealistic link formation mechanism. Indeed, empirical contact networks are characterized by several types of correlations as moderate modularity, transitivity and different types of assortativity [393,582,670,678,707]. In order words, individuals do not form random connections but tend to connect in groups/communities [189]. For example, a study conducted in a high-school showed significant levels of clustering and different types of assortativity. In particular, the project collected data for three days in a high-school with a population of about 1000 individuals (students, teacher, and staff) [582]. Each day about one third of them volunteered to wear a tag to collect their close proximity contacts and shared their flu vaccination status via an online survey. The assortativity measured considering the vaccination was significantly more positive than what expected by chance [582]. Furthermore, the levels of assortativity increased for longer contact duration and combining all data collected across the three days. The mechanisms driving such features could be associated to differences in vaccination coverage and assortativity within sub-networks such as age, gender, role, and ethnicity [582]. Some hints supporting these hypothesis are found in the coverage respect to gender and role. Around 48% of female reported to be vaccinated respect to 33% of male. Around 52% of staff reported to be vaccinated respect to 39% of students. Remarkably, epidemics spreading simulations suggest that the observed features of contacts networks might induce larger average outbreak size and increase the probability of such outbreaks [582]. More in general, the presence of correlations in links' creation might have contrasting effects on disease propagation. Locally the presence of dense connections might promote the propagation (especially in SIS models), the low connectivity between communities might hamper and slows down the global spreading [194,590,696,708,709].

Summarizing, empirical contact networks are characterized by heterogeneous activity, duration and time intervals between interactions as well as different types of correlations shaping the creations of links. Interestingly, a recent and simple model able to reproduce many of these properties has been developed [710,711]. The basic ingredient of the model is the idea that each node i is characterized by an *attractiveness* b_i. This is a proxy for social appeal/status. In the model, attractiveness are extracted from a uniform distribution with $b \in (0, 1]$. In order to mimic close proximity interactions, nodes move in a box of fixed size performing a biased random walk. More precisely, at each time step t, each node i moves, with probability $p_i(t)$, a step of length r along a direction given by a randomly chosen angle. Instead, with probability $1 - p_i(t)$, the node does not move. The probability $p_i(t)$ is defined as

$$p_i(t) = 1 - \max_{j \in E_j(t)} b_j, \tag{248}$$

where E_j are the nodes within a distance d from i at time t. In these settings, the node i will create a connection with all the nodes in its proximity, and it will be driven to stop and interact with a probability that is proportional to the attractiveness of the most attractive node in the circle of radius d. The model considers also that nodes might not all be present, and willing to make connections, at the same time. To include this aspect, nodes might be active or inactive. At each time step, an inactive node i become active with probability t_i, and an active node becomes inactive with probability $1 - t_i$. These probabilities are extracted from an uniform distribution with $t_i \in [0, 1]$. Although, nodes do not keep memory of past interactions and despite its extreme simplicity, the model is able to reproduce contact duration, time interval between contacts, and weight distributions of real contact networks. Also, it captures the evolution of the average degree as function of time and the distribution of groups sizes. It is of particular interest consider that a simple biased random walk in 2D driven by a random distribution of attractiveness is able to capture highly not trivial properties of empirical close proximity networks.

---

[1] The models have been extended to include more realistic link's formation mechanisms [590,706].

The study of close proximity contacts shows robust statistical patterns from one day to the next [707]. Interestingly, the same holds considering different years (for different students) [707]. However, the interactions of each individuals are not always the same. In fact, the study of the renewal of contacts in different days shows significant variations in the ego-net of each individuals [707]. At this time-scale the distribution of number of contacts does not show large heterogeneities [167,393,636,670,678,707]. As we saw above in the analysis of the epidemic threshold in activity-driven networks, this observation is of particular importance. Indeed, as we discussed in details in Section 4 heterogeneities in degree distributions make systems extremely fragile to the spreading of infectious diseases (lowering the epidemic threshold) and at the same time make targeted vaccination strategies extremely efficient [167,280]. Consequently, contacts networks at the time-scale of the day are more resilient to the spreading of infectious diseases, but also less susceptible to targeted vaccination strategies than networks whose contacts are integrated over longer time-scales, i.e. months or years. These observations clearly point out the importance of better understanding the features of close proximity contacts in order to devise efficient vaccination strategies [393,592,593,712,713]. To this end, let us extend what discussed in Section 5 for annealed and static networks by considering the efficiency of different vaccination protocols on time-varying networks. We will focus on the case of activity-driven networks [592]. First, let us consider a random strategy (RS) in which a fraction of nodes x is vaccinated. In this case, it is easy to modify the Eq. (241) and obtain [592]

$$\frac{\lambda_I}{\gamma} \geq \xi^{RS} \equiv \frac{1}{1-x} \frac{2\langle a \rangle}{\langle a \rangle + \sqrt{\langle a^2 \rangle}} = \frac{\xi}{1-x}. \tag{249}$$

In words, when a fraction x of nodes is randomly vaccinated, the epidemic threshold can be written as the threshold with no intervention, ξ, rescaled by the number of nodes still available to the spreading process. Indeed, vaccinating random nodes is equivalent to rescale the per capita spreading rate by the fraction of available nodes $\lambda_I \to \lambda_I(1-x)$. As observed in Section 4 this is the same in static and annealed networks. Fig. 42 shows the ratio between the number of infected nodes in the endemic state with random and no interventions, I_∞/I_0, as function of x. As clear from the plot, the random strategy (green triangle) requires more than 50% of vaccinated to reduce the endemic state to less than 20% respect to the scenario without interventions. Let us now consider a targeted strategy (TS) in which xN nodes are vaccinated in decreasing order of activity. This method is equivalent to fix a value a_c so that any node with activity $a \geq a_c$ is immune to the contagion process.[2] Also, for this scheme, it is possible to derive the analytic expression for the epidemic threshold [592]

$$\frac{\lambda_I}{\gamma} \geq \xi^{TS} \equiv \frac{2\langle a \rangle}{\langle a \rangle^c + \sqrt{(1-x)\langle a^2 \rangle^c}}, \tag{250}$$

where $\langle a^n \rangle^c = \int_\epsilon^{a_c} a^n F(a) da$ describes the moments of the activity distribution discounting the vaccinated nodes. Eq. (250) is not a simple rescaling of the original threshold expression and implies a drastic change in the behavior of the contagion process. Interestingly, vaccinating a very small fraction of the most active nodes is enough to stop the contagion process. Indeed, as clear from Fig. 42 vaccinating just the top 0.04% of nodes is enough to halt the disease [592]. As discussed in Sections 4–5, the network-wide knowledge required to implement targeted control strategies is generally not available [714]. In the case of time-varying networks, this issue is even more pronounced as the characterization of nodes depends on how long it is possible to observe the network dynamics [593]. Understanding the efficiency of strategies that use partial information about the network is crucial. A possible implementation of this is based on egocentric samples (ES) on the networks [592]. In particular, a fraction w of randomly selected nodes might act as "probes". During an observation time T, their egonet generated by their interactions in the network can be monitored. After the observation window, each probe is asked to randomly select a node in the egonet that is then vaccinated. In this strategy, the probability of vaccination for one node with activity a after a single time step is:

$$P_a = aw \int da' \frac{mN_{a'}}{N} + w \int da' a' \frac{mN_{a'}}{N} \frac{1}{m}. \tag{251}$$

The first term on the right considers the probability that a node of class a is active and reaches one of the probes; the second term, instead, takes into account the probability that one node of class a is connected by an active probe. Solving the integrals in Eq. (251), we can write $P_a = w(am + \langle a \rangle)$. Starting from this consideration, the probability of vaccination after t time steps is $P_a^t = 1 - (1 - P_a)^t$. The total number of nodes protected from the disease can be obtained summing over all the activity classes: $R^T = \sum_a N_a P_a^T = \sum_a N_a \left[1 - (1 - P_a)^T\right]$.[3] Using this expression, it is possible to obtain the epidemic threshold [592]:

$$\frac{\lambda_I}{\gamma} \geq \xi^{ES} \equiv \frac{2\langle a \rangle}{\Psi_1^T + \sqrt{\Psi_0^T \Psi_2^T}}, \tag{252}$$

---

[2] The values of x and a_c are linked by the relation $x = \int_{a_c}^1 F(a) da$.

[3] The equation for P_a does not consider the depletion of nodes in each class due to the vaccination process. The formulation is then a good approximation for small w and T, when the probability that a probe is selected more than once is very small.
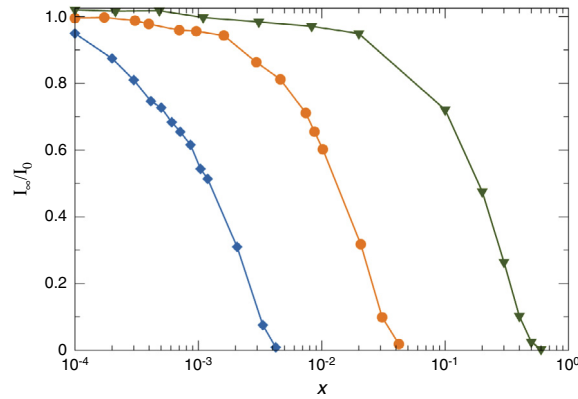
Fig. 42. Comparison of the stationary state of an SIS model with and without control strategy, I_∞/I_0, as a function of x when $\lambda_I/\gamma = 0.81$. In green triangles, we consider the random strategy, in blue diamonds the targeted strategy, and in orange circles the egocentric strategy. Each plot is made averaging 10^2 independent simulations started with 1% of random seeds. $N = 10^4$, $m = 3$, $\epsilon = 10^{-3}$, activity distributed as $F(a) \sim a^{-2.2}$.
Source: Adapted with permission from Ref. [592].
© 2014, by the American Physical Society.

where $\Psi_n^T = \int da \; a^n (1 - P_a)^T F(a)$. This last integral is a function of the observing time window T. As clear from Fig. 42 this strategy is much more efficient than the random one, although not as performant as the targeted one. It is important to stress how the efficiency of this strategy is due to the ability to reach active nodes by a local exploration done observing the systems for few time steps.

Another particularly important application of understanding close proximity contacts is the characterization of pathogen transmissions in hospitals settings [668,715–718]. Indeed, nosocomial infections are a major problem for the safety and health of patients and induce high economical costs [719]. However, due to challenges in the data collection, very little was understood until very recently. Wearables provide an unique opportunity in this context. To the present day, RFID tags have been used to identify or estimate pathways of transmission [715–717] and to help the test and develop prevention as well as containment measures [668,718].

11.2. Non participatory digital epidemiology

Our daily interactions with technology generate an unprecedented amount of data that can be harvested for digital epidemiology applications. Contrary to what we discussed above, here we will focus on methods that use data not generated, nor stored, explicitly for epidemiological studies. In other words, we will review recent non participatory approaches in which data is mined in search for proxies of epidemiological relevant behaviors/indicators.

11.2.1. Social media based methods

Social media provides unique opportunities to communicate, gather and spread information. The incredible progresses of technology and the consequent diffusion of mobile smart devices boosts the penetration and relevance of such rich medium. It did not take much time before our queries on search engines [720], communication on online social networks [721–725], visits to Wikipedia pages [726–728], cancellations of online reservations [729], satellite images of parking lots of hospitals [730], or online sources [731,732] started to be adopted for epidemiological applications. Indeed, all these behaviors and interactions between us or with technology can be extremely relevant to model, characterize, anticipate, and forecast the unfolding of infectious diseases [639].

While the details behind each methodology are case specific and clearly beyond the scope of this review, the basic underlying ideas, as well as the associated pitfalls, are a common feature. In order to better understand them, we will describe the prototypical example of Google Flu Trends (GFT) [733]. The service premiered in late 2008 and used our queries on Google to predict the unfolding of the seasonal influenza (and in some countries also dengue) in 29 countries. As we will describe later in more details, despite the initial glory, the service has been now discontinued and it has been reshaped in the form of a data sharing collaboration with few institutes [734]. The basic intuition behind GFT is extremely simple and powerful: traditional surveillance data (as for example the percentage of ILI visits) can be estimated and predicted by mining search engine queries. In particular, the epidemiological real data describing the progress of ILIs in a specific country, P, is estimated using proxy data, Q, describing the fraction of ILI related queries. More formally, the value of P at a given week is obtained via a linear model:

$$\tilde{P} = \vartheta_0 + \vartheta_1 \tilde{Q} + \varrho, \tag{253}$$

where the tilde refers to the logit of each quantity ($\tilde{y} = \log\left(\frac{y}{1-y}\right)$), the ϑs are fitting parameters, and ϱ is an error term. In other words, GFT obtains an estimation of P considering the fraction of ILI relevant queries, Q. Interestingly,

this quantity is evaluated mining about hundred billions of queries on Google's search engine. The methodology focuses just aggregated signals, no information/features about users are considered. Furthermore, the classification of queries as ILI relevant is performed measuring, independently, the correlation between 50 millions queries and historic surveillance data. Remarkably, this unsupervised method, returns as most relevant queries, searches associated to influenza, influenza complications, and remedies [720]. It is important to mention that the expression reported above could be generalized to include a set of Q variables:

$$\tilde{P} = \vartheta_0 + \sum_{i=1,n} \vartheta_i \tilde{Q}_i + \varrho, \tag{254}$$

each one associated to a different query [723]. However, as noted in the original paper, the performance of the method was higher adding up all the signal in a signal explanatory variable [720]. The initial results reported were extremely impressive: GFT had a mean correlation with surveillance data of about 0.97 and was able to predict it one/two weeks ahead. It is worth thinking about this a bit more. A simple linear regression model fed with an unparalleled dataset of queries to a search engine is able to estimate with great accuracy information about the unfolding of ILI.

Clearly the methodology described is quite simple. The strength of the approach is the unique and unprecedented dataset used. For many different reasons this data cannot be easily shared. Consequently, researchers applied, extended, and improved the GFT archetype to other data streams that are instead accessible. Examples are Twitter and Wikipedia. Indeed, in both cases large samples and/or relevant complete information can been gathered via their standardized APIs. Methods using Twitter data rely on the assumption that people post tweet, information about their health status. In particular, tweets are mined in search of ILI related terms that are fed to regression models [722–725,735]. The relevant terms are typically provided a priory or learned via machine learning techniques. Indeed, researchers do not have, typically, the computational resources of Google to test independently each n-gram used on Twitter. An increasing fraction of tweets contains geographical information that can be adopted to create models at different granularity going from cities [724] to countries [722,723,725,735]. Methods based on Wikipedia consider pages view of particular entries as data proxies [726–728]. The underlying assumption is the same of GFT: users visit specific pages to gather information about diseases, their symptoms, and possible medications. The approaches presented so far have been not focused just on ILI, but they considered a wide range of transmissible diseases going from dengue to ebola [726]. As we observed in the case of participatory digital surveillance data, the results of GFT or similar tools based on different data proxies can serve as input for modeling efforts based on unstructured (sub)populations (see Section 3). A nice summary the results of such approaches in providing actual predictions for the 2013–2014 flu season in the USA is detailed in Ref. [736].

Social media offers unprecedented opportunities to access information and communicate. Unfortunately, they can also be used effectively to spread misinformation, myths, or conspiracy theories [737]. In the realm of public health this phenomenon has been extremely clear. Indeed, as discussed in Sections 7–9 vaccination coverage has been affected by the spreading of myths (as the link between vaccines and autism or sterilization) or miscommunication and poor policy making from public officials (as in the case of Nigeria boycott of polio vaccination) [738,739]. Social media has changed significantly the way people acquire health information. We transitioned from classic doctor–patient to more patient–patient or patient–social media interactions [7]. On one side, this allows digital epidemiology based on social media to work. On the other, it makes us susceptible to the spreading of unfounded and unscientific information.

Many experts are now referring to the decrease of vaccination rates as a *confidence crisis* linked to the spreading of unscientific myths on social media, but also to socio-economical, socio-cultural, psychological, and political factors [7,387,403,740–742]. The features of online discussions about these and other topics can be mapped and as discussed in Section 9 used to model behavior-vaccination dynamics [737,738,743–746]. Interestingly, different opinions (pro and anti vaccination) induce effects of polarization in which individuals of different ideas do not interact much. As consequence, communication networks cluster in separated communities, *eco-chambers*, that exacerbates the challenges in spreading factual and expert-based knowledge. Furthermore, a recent study on the diffusion of sentiments regarding vaccines on Twitter suggests other peculiar dynamics [746]. First, individuals that are connected with large number of opinionated users are inhibited from expressing their own opinions. Second, exposure to negative sentiments is contagious while exposure to positive sentiments is not. All together these results indicate that peer influence and social contagion are content dependent and far from being trivial. It is important to mention how social media and more in general technology are also used to improve vaccine communication and coverage via texts, emails, calls, apps, and web services [747–749].

Despite good performances, all digital epidemiological methods based on social media are affected by several limitations (see Table 8). Also in this case, we will consider GFT as example to better understand the shortcomings and potential pitfalls of these approaches. GFT has been released in November 2008 showing great results. However, the tool was not able, despite being updated, to predict the unfolding of the pandemic in 2009 [750,751]. Furthermore, GFT overestimated (i) 92% of the weeks in the seasons 2011 and 2012, (ii) the peak in 2013 (the prediction was more than double the real value). As we wrote above the service has been now discontinued [734] and it is limited to data sharing with research partners. The reasons behind these bad performances are several. Firstly, GFT adopted simple correlations between online queries and ILI surveillance data. In doing so, it captured other phenomena correlated with ILI as for example winter patterns. The pandemic in 2009 peaked much earlier (late October early November) than the classic flu. Probably, this is what impeded GFT to capture its unfolding. Secondly, Google search engine is highly dynamical. Changes are constantly implemented. Some of them might modify users' searching behaviors, as for example the auto-complete suggestions. It is important to consider

Table 8
Summary of the main advantages and disadvantages of social media based methods in digital epidemiology.

Main advantages	Main disadvantages
Sample size	Possible biases in sample
Access to different geographical granularities	Correlations might be spurious
Near real-time analysis	Simple word matching might not sufficient
Cheaper data collection	n-gram selection computationally expensive
Rich set of information available	Changes in platform design and functionalities might influence users behaviors

all of these points as they are likely to affect all digital tools for epidemiology. Indeed, they are typical shortcomings and issues associated to all non participatory data collections. In particular, simple correlations might not be enough to describe complex phenomena unfolding at different scales in time and spaces respect to those used for training purposes. More in general, correlation does not guarantee causation. A recent study on Twitter shows that the time series of the word zombie is a very good predictor of the seasonal flu [723]. Furthermore, matching searches or keywords to context (in this case disease) without taking in account that the same word can be used in discussions completely unrelated to illness might introduce spurious correlations or decrease the signal to noise ratio [752]. Notwithstanding, it is fundamental to acknowledge the potential of these tools in the fight against infectious diseases. They offer unprecedented opportunities to estimate epidemic proxies at high geographical resolution or in locations where surveillance data is not available. Furthermore, these data can be used to inform realistic epidemic models [649,753] as well as a large number of more theoretical models as those described in Sections 4–9.

11.2.2. Mobile phone based methods

Mobile phones are one of the most diffuse technological device. About 96% of the worldwide population has one [754]. For billing purposes phone carriers collect call data records (CDRs) that can be mined and used to study our communications, social networks, and mobility (see Ref. [755] for a recent review). The last two phenomena are of epidemiological relevance. In particular, contact networks estimated via CDRs can be used to characterize human interactions at large and long-term scales well beyond those accessible via participatory experiments mentioned above. Although, not all connections mediated by mobile phones are relevant for disease transmission, recent studies clearly show the possibility of inferring close proximity networks from CDRs [756,757]. The temporal features describing mobile phone networks are similar to those of actual closed proximity interactions [590,679,680,758]. However, CDRs can provide just contact networks integrated in time. Indeed, the empirical timestamp of each interaction is not per se relevant for disease spreading. Interestingly, also these, time-aggregated and static, networks share several statistical features with close proximity contacts. In particular, high clustering, transitivity, and assortativity have been reported [755]. Furthermore, integrated mobile phone datasets are typically characterized by heavy-tailed distributions of contacts and weights. However, their heterogeneity are, also in this case, function of the time window used to aggregate data [759]. From a modeling stand point, CDRs perfectly lend themselves as input and test bed for the different modeling efforts beyond the homogeneous mixing approximation discussed in Section 4. Human mobility is another crucial ingredient for spreading of infectious disease. CDRs provide unprecedented opportunity to study this human phenomena. The large body of literature on the subject identifies three main characteristics of our movements [755,760]. Firstly, the exploration of new places is very slow. Secondly, our movement is relatively predictable. Thirdly, the way each one of us moves is rather unique. The understanding of such features has been crucial for modeling and understanding of the spreading of infectious diseases. For example approaches informed with data from CDRs have been used to map and control the spreading of Malaria [761,762], Ebola [763], and ILIs [249,764]. In order to better understand the potential and importance of this data we will briefly summarize the methods and results recently reported in Ref. [761]. The diffusion of Malaria is the results of a complex process that involves the interaction between its parasite and two hosts: mosquitoes and humans [761]. In particular, the movement of people increases the dispersal of the parasite beyond the reach of mosquitoes. CDRs in combination with spatially resolved information about Malaria prevalence have been used to define how human movements drive parasite importation between different regions of Kenya. In particular, CDRs were used to quantify the travel patterns of about 15 millions of people (around 34% of the total population) over one year. The calls of such individuals have been mapped to one of the 12,000 cells towers spread across 692 settlements [761]. In order to define the primary settlement for each person, the researchers considered a majority rules according to the time spent in each location. Furthermore, they collected information, such as destination and time, about the movements of individuals made out of their primary settlement. This data has been combined with an simple epidemic transmission model based on high resolution maps of Malaria prevalence. The fusion of these two data sources allowed to define the networks of parasite movements across the country as well as source (net emitters of people and parasites) and sink (net receivers of people and parasites) regions [761]. These results provided unprecedented lens to characterize Malaria dynamics and spreading. More importantly, they suggested a novel set of strategies for containment that targets

large flows of human movements associated with high risk of importations. These encode many aspects of the complex dynamics driving the spreading and go well beyond the typical based on local transmission hot spots or routine surveillance in high-risk areas [761].

11.3. Mixed approaches

We distinguished among the various methods and tools of digital epidemiology considering the strategy used for data collection. So far we presented active (participatory) or passive (non participatory) approaches. This division is quite clear and includes most of the notable tools. However, there are some that are a mix between the two. In the case of *Health Map* [765] the mix is quite asymmetric and it could have been classified as non participatory tool. This project is a data grinder. It considers a wide range of data sources such as different online data streams, eyewitness reports, expert-curated discussion, and official reports [765–767]. It has been lunched in 2006 with the goal of monitor, nowcast, and predict the spreading of infectious disease. The data is fetched, classified, and visualized 24/7/365. The system is automatized and allows to process and digest incredible amount of relevant information providing comprehensive reports that are used to facilitate the detection or monitoring of public health threats across the world. The Health Map infrastructure is behind an impressive number of projects and specific applications [768,769]. Examples include Health Map Flu Trends in which all the different data streams are combined to provide forecasts of flu activity in the USA [770,771], Dengue Map that provides real-time alerts of dengue outbreaks [772], or Health Map Vaccine which connects the population with vaccines [749,773]. As for the participatory surveillance systems described above, the data provided by Health Map does not provided any information about the contacts of individuals. It provided instead, data at the (sub)population level that can be used to develop unstructured compartmental models (see Sections 3 and 8).

Another application that mixes data from social media to active participation of users is Crowdbreaks [774]. This tool mines Twitter in search of tweets of epidemiological relevance. Crowdbreaks is a system for digital surveillance able to provide geographically resolved information about diseases in the USA. The selection of relevant tweets process is what makes it unique. Indeed, the task is crowdsourced to users visiting the webpage. In particular, it adopts the wisdom of the crowd to classify tweets as relevant, not relevant, or potentially relevant to a particular disease. With the collaboration and direct engagement of people the performance of machine learning methods used to automatically detect relevant tweets increases in time. The system, is currently in a beta version during which the contribution of individuals is crucial to help make it operational.

12. Conclusion, discussion and future research

This review has summarized the rich opportunities and challenges for studying theoretical and empirical aspects of the complex interactions between vaccinating behavior and disease dynamics, as well as the many opportunities to apply methods and concepts from statistical physics to these interactions. This was first seen in the contrast between mean-field models that do not incorporate human behavior versus mean-field models that do incorporate them (Sections 3 and 8). Mean-field models continue to play an invaluable role in the development and application of the behavioral epidemiology of infectious diseases in at least two respects. Firstly, they have shown the intrinsic complexity of coupled behavior–disease dynamics: coupled behavior–disease models generate emergent dynamics that are not present in models where vaccinating behavior is ignored, and coupled models also exhibit more dynamical outcomes in general. For instance, these models can exhibit oscillations in disease outbreaks and vaccine coverage even when the corresponding compartmental model does not exhibit any such dynamics. By incorporating human behavior (either explicitly or implicitly) we can ask questions that are not possible with models lacking behavior, such as how do different modes of communication (e.g. word-of-mouth versus mass media) influence the probability of eradication of a vaccine-preventable infection? Secondly, coupled behavior–disease models have demonstrated the ability to explain observed data better compared to simple behavior-free epidemiological models, in cases where data availability permits such a comparison. As a result, mean-field behavior–disease models show potential for informing public health policy makers of modern industrialized countries planning the introduction of a new vaccine; planning the response to a decline in vaccination coverage following a rumor; or more generally to strategically act to increase the degree of resilience of public vaccination systems in face of the threat posed by human behavior in a world where risks from most common infections are typically perceived as "minor".

Future research on mean-field behavior–disease models should aim to further exploit benefits and potentialities of simple models, such as representing simple and parsimonious environments for coupling behavior–disease dynamics and allowing predictions at a high level of abstraction (e.g. population-level prevalence and coverage). As such, a first line of inquiry should use mean-field models for investigating the implications of further social, economic, or psychological theories of behavior relevant for vaccination in order to achieve a more detailed understanding of the "typically observed" behavior (see e.g. [399]). A second line should aim to more extensively develop the – just initiated – empirical use of mean-field models by applying them over different datasets (e.g. different countries) and different infections (e.g. measles vs pertussis). This would allow to test the role of different behavioral factors and improve our understanding of why such factors promoted different behavioral responses to the same rumor, such as observed in the MMR vaccine-autism rumor. A third line of inquiry should aim to systematically include stochasticity in simple behavior–disease models, with the aim to preserve

parsimony but at the same time to better account for the fact that most challenges due to vaccine refusal will appear under circumstances of disease rarity, i.e. when infections become sub-critical due to efficient control, thereby lowering the perceived payoff of vaccination. These situations are of special relevance for large scale elimination plans, as the World Health Organization plans for measles elimination, where local infection rarity is expected to become the rule of infection dynamics and chance effects become fundamental. Last, straightforwardly incorporating behavioral responses in the realistic age-structured models typically used to evaluate the impact of intervention programs would help such models account for possibly the most important source of uncertainty about the future success of vaccination programs—human behavior. We emphasize that these suggestions do not exhaust the possibilities for further development of mean-field behavior–disease models.

Mean-field models assume that the population can be approximated as a continuum, and they are often deterministic in nature. However, for many applications, it may not be possible to ignore the discrete, individual-based nature of real populations. Section 5 – on models for disease dynamics and vaccinating on networks that do not explicitly represent behavior dynamics – showed that even without taking individual behavior into account, understanding networks of individual contacts in a population has a critical role to play in shaping our understanding of epidemiological processes. It is vitally important that network structure is taken into account if we attempt to predict population-level dynamics from individual-level observations. In addition, many intervention strategies such as contact tracing or ring vaccination can only be accurately captured and modeled using network-based approaches. The emergence of network modeling tools enables different strategies to be tested in an artificial environment. Finally, vaccination on time-varying networks is another interesting research topic in this area, requiring high-resolution temporal data to track and understand the control of disease and the dynamic evolution of networks.

Given that network models lend themselves so easily to representation of individuals, it is perhaps no surprise that the recent growth in network-based coupled behavior–disease models has been so rapid, as described in Section 9. As with the mean-field models, many behaviors were noted in network behavior–disease models that were not possible in network models lacking behavior, such as the difficulty of eradicating infection. Moreover, the importance of local structure and stochasticity take on new meaning, in networks where nodes actively adapt to their local epidemiological environment. In Section 9, we provided an overview of current results about behavioral vaccination in networked populations. Starting from the simple single-layer networks, we first reported the influence of benefit-inspired imitation dynamics, network topology, external incentives as well as additional protection measures on individual vaccination willingness and total vaccination level. We then demonstrated how social clusters expand and induce a large outbreak probability even if there exists extremely high vaccination coverage on empirical networks. It is worth mentioning that opinions about vaccination have been well identified as another type of method for protection measure updating. Finally, we briefly discussed the recent progress of behavioral vaccination on the novel framework of temporal and multilayer networks, where the links among nodes are time-varying or multiple. Along this line, behavioral vaccination on metapopulation networks deserves our considerable attention in future, because this topology supports the human mobility and information exchange among different patches, which will directly affect how human update their measures and seems closer to realistic situations. This framework, on the one hand, will further enrich content of behavioral vaccination (especially the feedback loop between human behavior and disease) with mobile fashion, similar to reaction–diffusion processes in physical chemistry. On the other hand, such a framework can be well-integrated with empirical data, such as the control of an infectious disease pandemic driven by air travel [775]. As mentioned above, (benefit-inspired or opinion-based) updating dynamics plays a crucial role in the final vaccination level and epidemic incidence. Most existing literature assumes that everyone has the same updating criterion, namely, identical cognition process for infection risk, vaccine risk and benefit brought by vaccination, which in fact greatly differs among individuals due to social norms, education background, age structure, etc. This point may attract great interest in future research. Risk perception on networks is another unexplored field of voluntary vaccination. The vast majority of models assume that information transmission of infection risk is completely local (i.e. coming from the contacting neighbors) and neglects information dissemination to the whole population or network, which is ubiquitous in reality, due to public media. Thus, how to incorporate local and global information into one modeling framework will be of particular significance in future.

Other unexplored problems concerning behavior–disease interactions on vaccination networks also merit great attention in the future. For instance, these approaches could be extended to other types of vaccination on multilayer networks, like local vaccination, or targeted vaccination based on various centrality indices. However, the development of a theoretical framework for multilayer networks still lies at the beginning stages: there are no uniform definitions of these centrality indices, which indicates that such a task needs in-depth interdisciplinary collaborations, especially with network scientists. Another scenario that must attract more attention is when multilayer networks support different dynamical processes on each layer. An infectious disease can spread simultaneously on different layers, or disease and awareness can diffuse in their respective layers. In such cases, how to implement vaccination becomes a novel challenge. Additionally, one of the most sophisticated and commonly used models for non-behavioral infectious disease modeling is the meta-population network, where individuals can move between different patches due to migration or commuting, as in many real-world populations. Vaccination on these and other social structures using endogenized behavioral dynamics remains little explored. Finally, one of main needs in theoretical epidemiology of infectious diseases is the development of new analytical or approximated theoretical physics methodologies in order to make simulation-independent inferences concerning individual-based models. Indeed, inference uniquely based on simulations can sometime hide the underlying big picture of a complex system.

We suggest that Ising models, percolation theory, and related approaches from statistical physics may be useful for studying coupled behavior–disease dynamics on networks, just as these approaches has proven useful in the study of (non-behavioral) epidemiological processes on lattices [222,250,776].

The ultimate standard in any scientific discipline should be testing hypotheses against data, and the standard for models of coupled behavior–disease dynamics should be no different. The last few sections of our review explored existing data sources like surveys and vaccine coverage records, which could be used to parameterize or validate such models. Measuring and assessing immunization levels remains a difficult problem. As more health systems globally are moving into the digital area, we can expect an improvement in the quality of the data. Furthermore, novel biomedical approaches may make it easier to assess whether individuals could mount an immune response, the expected waning time, and how immunity came about (vaccination vs. infection). Such individual-level characteristics are highly relevant to population-level infectious disease models. Nevertheless, as problems of vaccinations are particularly pronounced in low-income countries, the general challenges will remain in the foreseeable future. Surprisingly, the problems described here – uncertainty in vaccination coverage, and variability in vaccine efficacy and effectiveness – are only rarely taken into account in theoretical studies of infectious disease dynamics. We hope that in the future, these real world issues will get more attention in the modeling literature.

The last few sections of our review also explored new digital data sources that could facilitate a new era in the theoretical and empirical development of coupled behavior–disease models. Digital epidemiology is a very new and promising area of research: a natural evolution of epidemiology in the digital era. Such data sources lend themselves very naturally to network coupled behavior–disease models. Recent advances in mobile phone technology, information from twitter, Facebook, and GPS location enables us to determine individual connections, or even track the movement of people in real time. This would help us to build full and comprehensive networks for many airborne diseases such as influenza. It will also track how these network topologies change in the face of a severe epidemic, or even how individual behaves when an intervention measure is implemented. The vibrant body of literature on digital epidemiology shows incredible potential as well as future directions and challenges. The size and heterogeneity of the datasets we will face is destined to increase dramatically. This will require the development of new algorithms and techniques to effectively collect, store and analyze large-scale epidemiological data. In particular, new approaches, specifically designed for epidemiology, will be needed to filter, classify, and model massive data streams. Cheaper and more scalable devices will be required to improve our understanding of face-to-face interactions and thus of spreading pathways. From a theoretical standpoint, new frameworks able to further characterize the complexity of human contacts and their effects on epidemic spreading will be needed. In particular, it will be crucial to deepen our understanding of the temporal, and high-order structural features of social interactions. Furthermore, a better understanding of the dynamics of social contagion will be necessary to characterize behavioral changes induced by spreading of information/misinformation that affect vaccination rates and more in general disease spreading. The future of digital epidemiology is also tied to the development of new approaches for data access, sharing and privacy. Indeed, the ethical and legal challenges that arise from mining human digital traces are in general far from be understood or well regulated. Digital epidemiology is a truly interdisciplinary endeavor that currently connects a wide range of disciplines ranging from data science and mathematics to network science and medicine. A new breed of practitioners able to bridge all of these disciplines are necessary to face the challenges ahead. Furthermore, the prospect of being able to relate individual-level behavioral dynamics to population-level behavioral dynamics with unprecedented spatial and temporal resolution raises the possibility of attempting to develop a statistical physics of behavior. The extensive literature on non-equilibrium thermodynamic theory may prove useful in this endeavor [777].

Clearly, the opportunities for new theory and new applications, as well as the potential awards of success, are enormous in the field of coupled behavior–disease modeling. But how can the field move forward on these opportunities? Review papers often include a discussion paragraph on how more interdisciplinary interactions are needed. We fall to the temptation of including such a paragraph here, but we will explain exactly why interdisciplinary interactions are valuable in the context of this particular field. The field itself emerged from interdisciplinary interactions when researchers – usually trained in other areas like physics or mathematics – took an interest in disease systems. Despite this, there remains a need for more interdisciplinary interactions than currently take place. For instance, physicists, mathematicians, and economists often develop models of behavior–disease interactions in parallel with one another and apparently without interfacing with the knowledge of the other fields. This not only risks duplication, but also means that research fails to benefit from cross-fertilization of ideas. Interactions with epidemiologists and public health researchers could also be more pervasive than they currently are, in order to ensure that these models are ultimately applied to improving population health outcomes in the real world. In a similar vein, the scope of interdisciplinary interactions needs to expand to include new fields. For instance, digital data and machine learning hold significant promise for helping us better understand behavior–disease interactions and parameterize coupled behavior–disease models. However, properly harnessing these new sources of data will require help from statisticians and computer science experts.

Interdisciplinary interactions are stymied by different terminology used by differently trained researchers. For instance, the theoretical study of interactions between disease transmission and behavior has been variously referred to using terms such as "behavioral epidemiology", "economic epidemiology", "coupled disease–behavior systems" and other terms although in some cases, it should be noted that these terms do not completely overlap because they are used differently to marry different sets of concepts. Similarly, economists use the word "rational" to describe selfish, socially blind behavior, although other fields would not describe such behavior as rational in its usual sense. The problem of inconsistency in

language could be addressed through workshops devoted to bringing together different fields, which would also help the more general lack of interaction between different disciplines. A degree of coordination as exhibited by the nomenclature committee of the International Astronomical Union for studying interactions between behavior and disease is probably impossible due to the nature of the subject, but efforts to bring together different types of experts could at least accomplish some of the benefits of greater coordination.

Shakespeare famously ended his play "A Midsummer Night's Dream" with an apology from the spirit Puck to the audience: "If we shadows have offended, Think but this, and all is mended—That you have but slumbered here While these visions did appear". We also wish to apologize to the readers in advance, for having inadvertently missed some excellent papers in the literature, having unintentionally misrepresented other papers that we did include, or being completely wrong about where the field should go next. However, there is no need for the reader to pretend our review was something they dreamed while slumbering. Instead, please write us if you feel our review was lacking in some respect, or better yet, publish some research that shows us the value of other perspectives on the field. If nothing else, we hope that our review has convinced the readers that coupled behavior–disease modeling is a rapidly growing area that promises exciting theoretical developments and empirical applications in the coming decade, especially if we draw on the accumulated knowledge of many decades of statistical physics.

Table A.1
Abbreviations of the used phrases in our report.

Abbreviations	Full names	First appeared
SIR	Susceptible–Infectious–removed	
SIS	Susceptible–Infectious–Susceptible	
SEIR	Susceptible–Exposed–Infective–Removed	
SEIRS	Susceptible–Exposed–Infective–Removed–Susceptible	
ODE	ordinary differential equation	
GAS	globally asymptotically stable	
AIDS	acquired immune deficiency syndrome	
FOI	force of infection	
BRN	basic reproduction number	
ERN	effective reproduction number	
VRN	vaccine reproduction number	Section 3
NGO	next generation operator	
AR	attack rate	
DFE	disease-free equilibrium	
EE	endemic equilibrium	
WAIFW matrices	Who Acquires Infection From Whom matrices	
ME	Master equation	
CLE	Chemical Langevin equation	
FPE	Fokker–Planck equation	
SDE	Stochastic Differential equation	
CCS	critical community size	
ER graph	Erdös–Rényi graph	
SW network	small-world network	
WS model	Watts and Strogatz model	
SF network	scale-free network	
BA network	Barabási–Albert network	
DMS model	Dorogovtsev–Mendes–Samukhin model	Section 4
DBMF	Degree-based mean-field theory	
IBMF	Individual-based mean-field theory	
ICM	Independent cascade model	
GEMF framework	generalized epidemic mean-field framework	
STD	sexually transmitted diseases	
CBF	community bridge finder	
BHD	bridge-hub detector	Section 5
SIAs	special immunization activities	
MMR	measles–mumps–rubella	Section 6
AIC	Akaike Information Criterion	
SEU	subjective expected utility	
ESS	evolutionarily stable state	Section 7
VAE	vaccine adverse effects	
IGD	Imitation Game Dynamics	
EGDT	Evolutionary Game Dynamics Theory	
VSE	vaccine side effects	Section 8
ISNs	injunctive social norms	
CSNE	convergently stable Nash equilibrium	
MDPT	Markov decision processes theory	

Table A.1 (*continued*)

Abbreviations	Full names	First appeared
RV	random vaccination	
NN	nearest neighbor vaccination	Section 9
CV	chain vaccination	
INN	improved nearest neighbor vaccination	
DHS	Demographic and Health Surveys	
MICS	Multiple Indicator Cluster Surveys	
EPI	Expanded Programme on Immunization	Section 10
LQAS	Lot Quality Assurance Sampling	
HPV	Human Papillomavirus	
ILI	influenza like illnesses	
RFID	radio frequency identification devices	
ES	egocentric samples	
GFT	Google Flu Trends	Section 11
CDRs	call data records	
RS	random strategy	
TS	targeted strategy	

Acknowledgments

Appendix

Here, we provide a summary of abbreviations of the used terminologies in the main text.

References

[1] D. Henderson, P. Klepac, Lessons from the eradication of smallpox: an interview with DA Henderson, Phil. Trans. R. Soc. B 368 (1623) (2013) 20130113.
[2] D.M. Morens, G.K. Folkers, A.S. Fauci, The challenge of emerging and re-emerging infectious diseases, Nature 430 (6996) (2004) 242–249.
[3] P. Bonanni, Demographic impact of vaccination: a review, Vaccine 17 (1999) S120–S125.
[4] C.A. De Quadros, J.M. Olivé, B.S. Hersh, M.A. Strassburg, D.A. Henderson, D. Brandling-Bennett, G.A. Alleyne, Measles elimination in the Americas: evolving strategies, JAMA 275 (3) (1996) 224–229.
[5] H. Heesterbeek, R.M. Anderson, V. Andreasen, S. Bansal, C. De Angelis, C. Dye, K.T. Eames, W.J. Edmunds, S.D. Frost, S. Funk, et al., Modeling infectious disease dynamics in the complex landscape of global health, Science 347 (6227) (2015) aaa4339.
[6] L.A. Sturm, R.M. Mays, G.D. Zimet, Parental beliefs and decision making about child and adolescent immunization: from polio to sexually transmitted infections, J. Dev. Behav. Pediatr. 26 (6) (2005) 441–452.
[7] H.J. Larson, L.Z. Cooper, J. Eskola, S.L. Katz, S. Ratzan, Addressing the vaccine confidence gap, Lancet 378 (9790) (2011) 526–535.
[8] R.S. Dhillon, J.D. Kelly, Community trust and the ebola endgame, New Engl. J. Med. 373 (9) (2015) 787–789.
[9] L. Van Valen, A new evolutionary law, Evol. Theory 1 (1973) 1–30.
[10] D.L. Page, Thucydides' description of the great plague at Athens, Class. Q. (N.S.) 3 (3–4) (1953) 97–119.
[11] J.R. Bennett, R. Bennett, The Diseases of the Bible, Vol. 9, Religious Tract Society, 1891.
[12] P.J. Brown, Microparasites and macroparasites, Cult. Anthr. 2 (1) (1987) 155–171.
[13] R.M. Anderson, R.M. May, Population biology of infectious diseases: Part I, Nature (280) (1979) 361–367.
[14] S. Morand, B.R. Krasnov, R. Poulin, et al., Micromammals and Macroparasites, Springer, 2006.
[15] S.K. Collinge, C. Ray, Disease Ecology: Community Structure and Pathogen Dynamics, Oxford University Press, 2006.
[16] R.M. Anderson, R.M. May, The population dynamics of microparasites and their invertebrate hosts, Phil. Trans. R. Soc. B 291 (1054) (1981) 451–524.
[17] J. Swinton, M. Woolhouse, M. Begon, A. Dobson, E. Ferroglio, B. Grenfell, V. Guberti, R. Hails, J. Heesterbeek, A. Lavazza, et al., Microparasite transmission and persistence, Ecol. Wildl. Dis. (2002) 83–101.
[18] M.J. Keeling, P. Rohani, Modeling Infectious Diseases in Humans and Animals, Princeton University Press, 2008.
[19] J. Giesecke, et al., Modern Infectious Disease Epidemiology, Edward Arnold (Publisher) Ltd., 1994.
[20] A. Krämer, M. Akmatov, M. Kretzschmar, Principles of Infectious Disease Epidemiology, Springer, 2010.
[21] K.E. Nelson, C.M. Williams, Infectious Disease Epidemiology: Theory and Practice, Jones & Bartlett Publishers, 2014.
[22] M.S. Porta, S. Greenland, M. Hernán, I. dos Santos Silva, J.M. Last, A Dictionary of Epidemiology, Oxford University Press, 2014.
[23] A.R. Spickler, Emerging and Exotic Diseases of Animals, CFSPH Iowa State University, 2010.
[24] J. Goddard, Infectious Diseases and Arthropods, Springer Science & Business Media, 2009.

[25] D. Rogers, M. Packer, Vector-borne diseases, models, and global change, Lancet 342 (8882) (1993) 1282–1284.
[26] K.H. Mayer, H. Pizer, The Social Ecology of Infectious Diseases, Academic Press, 2011.
[27] R.M. Merrill, Introduction to Epidemiology, Jones & Bartlett Publishers, 2013.
[28] E. Vynnycky, R. White, An Introduction to Infectious Disease Modelling, Oxford University Press, 2010.
[29] C. Fraser, S. Riley, R.M. Anderson, N.M. Ferguson, Factors that make an infectious disease outbreak controllable, Proc. Natl. Acad. Sci. USA 101 (16) (2004) 6146–6151.
[30] M. Thrusfield, Veterinary Epidemiology, Elsevier, 2013.
[31] J.C. Thomas, D.J. Weber, Epidemiologic Methods for the Study of Infectious Diseases, Oxford University Press New York, NY, USA, 2001.
[32] C. Cheng-Mayer, D. Seto, M. Tateno, J.A. Levy, Biologic features of HIV-1 that correlate with virulence in the host, Science 240 (4848) (1988) 80–82.
[33] A.D. Stewart, J.M. Logsdon, S.E. Kelley, An empirical study of the evolution of virulence under both horizontal and vertical transmission, Evolution 59 (4) (2005) 730–739.
[34] T.C. Timmreck, An Introduction to Epidemiology, Jones & Bartlett Learning, 2002.
[35] P.E. Fine, Herd immunity: history, theory, practice, Epidemiol. Rev. 15 (2) (1993) 265–302.
[36] S. Boslaugh, Encyclopedia of Epidemiology, Sage Publications, 2007.
[37] J.P. Fox, L. Elveback, W. Scott, L. Gatewood, E. Ackerman, Herd immunity: basic concept and relevance to public health immunization practices, Amer. J. Epidemiol. 94 (3) (1971) 179–189.
[38] S. Plotkin, Mass Vaccination: Global Aspects-progress and Obstacles, Vol. 304, Springer Science & Business Media, 2006.
[39] E.J. Gangarosa, A. Galazka, C. Wolfe, L. Phillips, E. Miller, R. Chen, R. Gangarosa, Impact of anti-vaccine movements on pertussis control: the untold story, Lancet 351 (9099) (1998) 356–361.
[40] M.A. Kohanski, D.J. Dwyer, J.J. Collins, How antibiotics kill bacteria: from targets to networks, Nat. Rev. Microbiol. 8 (6) (2010) 423–435.
[41] V. Lorian, Antibiotics in Laboratory Medicine, Lippincott Williams & Wilkins, 2005.
[42] R. Gonzales, D.C. Malone, J.H. Maselli, M.A. Sande, Excessive antibiotic use for acute respiratory infections in the United States, Clin. Infect. Dis. 33 (6) (2001) 757–762.
[43] H. Seppälä, T. Klaukka, J. Vuopio-Varkila, A. Muotiala, H. Helenius, K. Lager, P. Huovinen, The effect of changes in the consumption of macrolide antibiotics on erythromycin resistance in group A streptococci in Finland, New Engl. J. Med. 337 (7) (1997) 441–446.
[44] M. Blaser, Antibiotic overuse: stop the killing of beneficial bacteria, Nature 476 (7361) (2011) 393–394.
[45] J. Davies, D. Davies, Origins and evolution of antibiotic resistance, Microbiol. Mol. Biol. Rev. 74 (3) (2010) 417–433.
[46] R. Zhang, K. Eggleston, V. Rotimi, R.J. Zeckhauser, Antibiotic resistance as a global threat: evidence from China, Kuwait and the United States, Glob. Health 2 (6) (2006) 1–14.
[47] I. Stephenson, K.G. Nicholson, A. Colegate, A. Podda, J. Wood, E. Ypma, M. Zambon, Boosting immunity to influenza H5N1 with MF59-adjuvanted H5N3 A/Duck/Singapore/97 vaccine in a primed human population, Vaccine 21 (15) (2003) 1687–1693.
[48] G.W. Comstock, Vaccine evaluation by case-control or prospective studies, Amer. J. Epidemiol. 131 (2) (1990) 205–207.
[49] W.A. Orenstein, R.H. Bernier, A.R. Hinman, Assessing vaccine efficacy in the field further observations, Epidemiol. Rev. 10 (1) (1988) 212–241.
[50] R.T. Chen, W.A. Orenstein, Epidemiologic methods in immunization programs, Epidemiol. Rev. 18 (2) (1996) 99–117.
[51] R.J. Blendon, J.M. Benson, C.M. DesRoches, E. Raleigh, K. Taylor-Clark, The public's response to severe acute respiratory syndrome in Toronto and the United States, Clin. Infect. Dis. 38 (7) (2004) 925–931.
[52] T. Gross, C.J.D.D. Lima, B. Blasius, Epidemic dynamics on an adaptive network, Phys. Rev. Lett. 96 (20) (2006) 208701.
[53] S. Funk, E. Gilad, C. Watkins, V.A. Jansen, The spread of awareness and its impact on epidemic outbreaks, Proc. Natl. Acad. Sci. 106 (16) (2009) 6872–6877.
[54] D.H. Zanette, S. Risau-Gusmán, Infection spreading in a population with evolving contacts, J. Biol. Phys. 34 (1–2) (2008) 135–148.
[55] Y. Yu, G. Xiao, On early detection of strong infections in complex networks, J. Phys. A: Math. Gen. 47 (6) (2014) 065101.
[56] C. Xia, L. Wang, S. Sun, J. Wang, An SIR model with infection delay and propagation vector in complex networks, Nonlinear Dynam. 69 (3) (2012) 927–934.
[57] G. Sun, Z. Wu, Z. Wang, Z. Jin, Influence of isolation degree of spatial patterns on persistence of populations, Nonlinear Dynam. 83 (1–2) (2016) 811–819.
[58] G. Sun, S. Wang, Q. Ren, Z. Jin, Y. Wu, Effects of time delay and space on herbivore dynamics: linking inducible defenses of plants to herbivore outbreak, Sci. Rep. 5 (2015) 11246.
[59] R.J. Glass, L.M. Glass, W.E. Beyeler, H.J. Min, et al., Targeted social distancing design for pandemic influenza, Emerg. Infect. Dis. 12 (11) (2006) 1671–1681.
[60] P. Poletti, B. Caprile, M. Ajelli, A. Pugliese, S. Merler, Spontaneous behavioural changes in response to epidemics, J. Theoret. Biol. 260 (1) (2009) 31–40.
[61] P. Caley, D.J. Philp, K. McCracken, Quantifying social distancing arising from pandemic influenza, J. R. Soc. Interface 5 (23) (2008) 631–639.
[62] M.L. Cohen, Changing patterns of infectious disease, Nature 406 (6797) (2000) 762–767.
[63] A.E. Aiello, E.L. Larson, What is the evidence for a causal link between hygiene and infections? Lancet Infect. Dis. 2 (2) (2002) 103–110.
[64] D. Pittet, S. Hugonnet, S. Harbarth, P. Mourouga, V. Sauvan, S. Touveneau, T.V. Perneger, et al., Effectiveness of a hospital-wide programme to improve compliance with hand hygiene, Lancet 356 (9238) (2000) 1307–1312.
[65] B. Allegranzi, D. Pittet, Role of hand hygiene in healthcare-associated infection prevention, J. Hosp. Infect. 73 (4) (2009) 305–315.
[66] L. Fewtrell, R.B. Kaufmann, D. Kay, W. Enanoria, L. Haller, J.M. Colford, Water, sanitation, and hygiene interventions to reduce diarrhoea in less developed countries: a systematic review and meta-analysis, Lancet Infect. Dis. 5 (1) (2005) 42–52.
[67] A. Prüss, D. Kay, L. Fewtrell, J. Bartram, et al., Estimating the burden of disease from water, sanitation, and hygiene at a global level, Environ. Health Perspect. 110 (5) (2002) 537–542.
[68] W.O. Kermack, A.G. McKendrick, A contribution to the mathematical theory of epidemics, in: Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, vol. 115, The Royal Society, 1927, pp. 700–721.
[69] D. Bernoulli, Essai d'une nouvelle analyse de la mortalité cause par la petite vérole et des avantages de l'inoculation pour la prévenir. Histoire de l'académie royale des sciences avec les mémoires de mathématique et de physique tirés des registres de cette académie. Paris 1766 (année 1760), in: S. Haberman, T. Sibbett (Eds.), History of Actuarial Science, vol. VIII, Multiple Decrement and Multiple State Models, William Pickering, 1995.
[70] H. Heesterbeek, The law of mass-action in epidemiology: a historical perspective, in: K. Cuddington, B. Beisner (Eds.), Ecological Paradigms Lost: Routes of Theory Change, Academic Press, 2005, pp. 81–106.
[71] W.H. Hamer, The Milroy Lectures on Epidemic Disease in England: the Evidence of Variability and of Persistency of Type, Bedford Press, 1906.
[72] R. Ross, An application of the theory of probabilities to the study of a priori pathometry. Part I, Proc. R. Soc. A, Contain. Pap. Math. Phys. Charact. 92 (638) (1916) 204–230.
[73] R. Ross, H.P. Hudson, An application of the theory of probabilities to the study of a priori pathometry. Part II, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 93 (650) (1917) 212–225.
[74] A. McKendrick, The rise and fall of epidemics, Paludism 1 (1912) 54–66.
[75] M. Kermark, A. Mckendrick, Contributions to the mathematical theory of epidemics. Part I, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 115 (5) (1927) 700–721.
[76] K. Dietz, D. Schenzle, Mathematical models for infectious disease statistics, in: A. Atkinson (Ed.), A Celebration of Statistics, Springer, 1985, pp. 167–204.
[77] K. Dietz, The evaluation of rubella vaccination strategies, in: R. Hiorns, D. Cooke (Eds.), The Mathematical Theory of the Dynamics of Biological Populations, Vol. 2, Academic, New York, 1981, pp. 81–98.
[78] H. Hethcote, J. Yorke, Gonorrhea. Transmission Dynamics and Control, Springer-Verlag, 1984.
[79] R.M. Anderson, R.M. May, Infectious Diseases of Humans: Dynamics and Control, Oxford University Press, 1992.

[80] C. Bauch, A. d'Onofrio, P. Manfredi, Behavioral epidemiology of infectious diseases: an overview, in: P. Manfredi, A. d'Onofrio (Eds.), Modeling the Interplay Between Human Behavior and the Spread of Infectious Diseases, Springer, 2013, pp. 1–19.
[81] H.W. Hethcote, The mathematics of infectious diseases, SIAM Rev. 42 (4) (2000) 599–653.
[82] O. Diekmann, J. Heesterbeek, Mathematical epidemiology of infectious diseases: model building, analysis and interpretation (2000).
[83] V. Capasso, The Mathematical Structure of Epidemic Systems, Springer, 2008.
[84] O. Diekmann, J. Heesterbeek, J.A. Metz, On the definition and the computation of the basic reproduction ratio r0 in models for infectious diseases in heterogeneous populations, J. Math. Biol. 28 (4) (1990) 365–382.
[85] O. Diekmann, H. Heesterbeek, T. Britton, Mathematical Tools for Understanding Infectious Disease Dynamics, Princeton University Press, 2012.
[86] N.M. Ferguson, D.A. Cummings, S. Cauchemez, C. Fraser, S. Riley, A. Meeyai, S. Iamsirithaworn, D.S. Burke, Strategies for containing an emerging influenza pandemic in Southeast Asia, Nature 437 (7056) (2005) 209–214.
[87] J. Wallinga, M. Lipsitch, How generation intervals shape the relationship between growth rates and reproductive numbers, Phil. Trans. R. Soc. B 274 (1609) (2007) 599–604.
[88] P. Van den Driessche, J. Watmough, Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission, Math. Biosci. 180 (1) (2002) 29–48.
[89] J. Heesterbeek, A brief history of R 0 and a recipe for its calculation, Acta Biotheor. 50 (3) (2002) 189–204.
[90] A. d'Onofrio, Stability properties of pulse vaccination strategy in SEIR epidemic model, Math. Biosci. 179 (1) (2002) 57–72.
[91] L.F. Olsen, G.L. Truty, W.M. Schaffer, Oscillations and chaos in epidemics: A nonlinear dynamic study of six childhood diseases in Copenhagen, Denmark, Theor. Popul. Biol. 33 (3) (1988) 344–370.
[92] D.J. Earn, P. Rohani, B.M. Bolker, B.T. Grenfell, A simple model for complex dynamical transitions in epidemics, Science 287 (5453) (2000) 667–670.
[93] B.D. Dalziel, O.N. Bjørnstad, W.G. van Panhuis, D.S. Burke, C.J.E. Metcalf, B.T. Grenfell, Persistent chaos of measles epidemics in the prevaccination united states caused by a small change in seasonal transmission patterns, PLoS Comput. Biol. 12 (2) (2016) e1004655.
[94] Z. Grossman, I. Gumowski, K. Dietz, The incidence of infectious diseases under the influence of seasonal fluctuations, Anal. Approach (1977) 525–546.
[95] J.L. Aron, I.B. Schwartz, Seasonality and period-doubling bifurcations in an epidemic model, J. Theoret. Biol. 110 (4) (1984) 665–679.
[96] I.B. Schwartz, H.L. Smith, Infinite subharmonic bifurcation in an SEIR epidemic model, J. Math. Biol. 18 (3) (1983) 233–253.
[97] H.L. Smith, Multiple stable subharmonics for a periodic epidemic model, J. Math. Biol. 17 (2) (1983) 179–190.
[98] N. Bacaer, X. Abdurahman, Resonance of the epidemic threshold in a periodic environment, J. Math. Biol. 57 (5) (2008) 649–673.
[99] C. Rebelo, A. Margheri, N. Bacaër, Persistence in seasonally forced epidemiological models, J. Math. Biol. 64 (6) (2012) 933–949.
[100] M.Y. Li, J.S. Muldowney, Global stability for the SEIR model in epidemiology, Math. Biosci. 125 (2) (1995) 155–164.
[101] M.J. Keeling, P. Rohani, B.T. Grenfell, Seasonally forced disease dynamics explored as switching between attractors, Physica D 148 (3) (2001) 317–335.
[102] N. Minorsky, Introduction to Non-linear Mechanics: Topological Methods, Analytical Methods, Non-linear Resonance, Relaxation Oscillations, JW Edwards, 1947.
[103] F.M. Arscott, An Introduction to Periodic Differential Equations, Pergamon, 1964.
[104] N. McLachlan, Theory and Application of Mathieu Functions, Clarendon Press, 1947.
[105] M. Farkas, Periodic Motions, Springer, 1994.
[106] L. Cesari, Asymptotic Behavior and Stability Problems in Ordinary Differential Equations, Springer, 1971.
[107] V.I. Arnold, Ordinary Differential Equations, Springer, 1992.
[108] L.D. Landau, E.M. Lifshitz, Course of Theoretical Physics, in: Mechanics, vol. 1, Pergamon Press, 1965.
[109] V. Capasso, G. Serio, A generalization of the Kermack-McKendrick deterministic epidemic model, Math. Biosci. 42 (1) (1978) 43–61.
[110] A. d'Onofrio, P. Manfredi, Information-related changes in contact patterns may trigger oscillations in the endemic prevalence of infectious diseases, J. Theoret. Biol. 256 (3) (2009) 473–478.
[111] K.P. Hadeler, C. Castillo-Chávez, A core group model for disease transmission, Math. Biosci. 128 (1) (1995) 41–55.
[112] A.S. Lopez, D. Guris, L. Zimmerman, L. Gladden, T. Moore, D.T. Haselow, V.N. Loparev, D.S. Schmid, A.O. Jumaan, S.L. Snow, One dose of varicella vaccine does not prevent school outbreaks: is it time for a second dose? Pediatrics 117 (6) (2006) e1070–e1077.
[113] C.M. Kribs-Zaleta, J.X. Velasco-Hernandez, A simple vaccination model with multiple endemic states, Math. Biosci. 164 (2) (2000) 183–201.
[114] C. Gardiner, Stochastic Methods: A Handbook for the Natural and Social Sciences, fourth ed., in: Springer Series in Synergetics, Springer, 2009.
[115] N.G. Van Kampen, Stochastic Processes in Physics and Chemistry, third ed., Elsevier, 2007.
[116] D.T. Gillespie, Exact stochastic simulation of coupled chemical reactions, J. Phys. Chem. 81 (25) (1977) 2340–2361.
[117] A.J. McKane, T.J. Newman, Stochastic models in population biology and their deterministic analogs, Phys. Rev. E 70 (2004) 041902. http://dx.doi.org/10.1103/PhysRevE.70.041902, URL http://link.aps.org/doi/10.1103/PhysRevE.70.041902.
[118] D.T. Gillespie, The chemical langevin equation, J. Chem. Phys. 113 (1) (2000) 297–306.
[119] M. Bartlett, Deterministic and stochastic models for recurrent epidemics, in: Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, vol. 4, 1956, pp. 81–109.
[120] M.J. Keeling, P. Rohani, Modeling Infectious Diseases in Humans and Animals, Princeton University Press, 2008.
[121] M.J. Keeling, B. Grenfell, Disease extinction and community size: modeling the persistence of measles, Science 275 (5296) (1997) 65–67.
[122] R. Kuske, L.F. Gordillo, P. Greenwood, Sustained oscillations via coherence resonance in SIR epidemic model, J. Theoret. Biol. 245 (3) (2007) 459–469.
[123] M. Khasin, M. Dykman, B. Meerson, Speeding up disease extinction with a limited amount of vaccine, Phys. Rev. E 81 (5) (2010) 051925.
[124] H. Andersson, T. Britton, Stochastic Epidemic Models and their Statistical Analysis, Vol. 151, Springer Science & Business Media, 2012.
[125] L.J. Allen, Stochastic Population and Epidemic Models -Persistence and Extinction, Springer, 2015.
[126] J.D. Murray, Mathematical Biology, third ed., in: Spatial Models and Biomedical Applications, vol. II, Springer-Verlag, 2002.
[127] N.T. Bailey, The Mathematical Theory of Infectious Diseases and its Applications, Charles Griffin & Company Ltd, 1975.
[128] J.D. Murray, Mathematical Biology, third ed., in: An Introduction, vol. I, Springer-Verlag, 2002.
[129] M.C. Gonzalez, C.A. Hidalgo, A.-L. Barabasi, Understanding individual human mobility patterns, Nature 453 (7196) (2008) 779–782.
[130] D. Balcan, A. Vespignani, Phase transitions in contagion processes mediated by recurrent mobility patterns, Nat. Phys. 7 (7) (2011) 581–586.
[131] V. Mendez, S. Fedotov, W. Horsthemke, Reaction-transport Systems: Mesoscopic Foundations, Fronts, and Spatial Instabilities, Springer Science & Business Media, 2010.
[132] M.J. Keeling, P. Rohani, Estimating spatial coupling in epidemiological systems: a mechanistic approach, Ecol. Lett. 5 (1) (2002) 20–29.
[133] P. Bajardi, C. Poletto, J.J. Ramasco, M. Tizzoni, V. Colizza, A. Vespignani, Human mobility networks, travel restrictions, and the global spread of 2009 H1N1 pandemic, PLoS One 6 (1) (2011) e16591.
[134] L. Wang, Y. Zhang, T. Huang, X. Li, Estimating the value of containment strategies in delaying the arrival time of an influenza pandemic: A case study of travel restriction and patient isolation, Phys. Rev. E 86 (3) (2012) 032901.
[135] K. Nah, S. Otsuki, G. Chowell, H. Nishiura, Predicting the international spread of Middle East respiratory syndrome (MERS), BMC Infect. Dis. 16 (1) (2016) 1.
[136] K. Nah, K. Mizumoto, Y. Miyamatsu, Y. Yasuda, R. Kinoshita, H. Nishiura, Estimating risks of importation and local transmission of Zika virus infection, PeerJ 4 (2016) e1904.
[137] V.A. Jansen, A.L. Lloyd, Local stability analysis of spatially homogeneous solutions of multi-patch systems, J. Math. Biol. 41 (3) (2000) 232–252.
[138] A.L. Lloyd, R.M. May, Synchronicity, chaos and population cycles: spatial coherence in an uncertain world, Trends Ecol. Evol. 14 (11) (1999) 417–418.
[139] A.L. Lloyd, R.M. May, Spatial heterogeneity in epidemic models, J. Theoret. Biol. 179 (1996) 1–11.
[140] B. Bolker, B. Grenfell, Impact of vaccination on the spatial correlation and persistence of measles dynamics, Proc. Natl. Acad. Sci. 93 (22) (1996) 12648–12653.
[141] V. Colizza, A. Vespignani, Invasion threshold in heterogeneous metapopulation networks, Phys. Rev. Lett. 99 (14) (2007) 148701.

[142] V. Colizza, A. Vespignani, Epidemic modeling in metapopulation systems with heterogeneous coupling pattern: Theory and simulations, J. Theor. Biol. 251 (3) (2008) 450–467.
[143] A. Lajmanovich, J.A. Yorke, A deterministic model for gonorrhea in a nonhomogeneous population, Math. Biosci. 28 (3) (1976) 221–236.
[144] R. Pastor-Satorras, A. Vespignani, Epidemic spreading in scale-free networks, Phys. Rev. Lett. 86 (14) (2001) 3200.
[145] R. Olinky, L. Stone, Unexpected epidemic thresholds in heterogeneous networks: The role of disease transmission, Phys. Rev. E 70 (2004) 030902.
[146] A. d'Onofrio, A note on the global behaviour of the network-based SIS epidemic model, Nonlinear Anal. RWA 9 (4) (2008) 1567–1572.
[147] J. Mossong, N. Hens, M. Jit, P. Beutels, K. Auranen, R. Mikolajczyk, M. Massari, S. Salmaso, G.S. Tomba, J. Wallinga, et al., Social contacts and mixing patterns relevant to the spread of infectious diseases, PLoS Med. 5 (3) (2008) e74.
[148] J. Wallinga, P. Teunis, M. Kretzschmar, Using data on social contacts to estimate age-specific transmission parameters for respiratory-spread infectious agents, Amer. J. Epidemiol. 164 (10) (2006) 936–944.
[149] L. Wang, Z. Wang, Y. Zhang, X. Li, How human location-specific contact patterns impact spatial transmission between populations? Sci. Rep. 3 (2013) 1468.
[150] L. Wang, Y. Zhang, Z. Wang, X. Li, The impact of human location-specific contact pattern on the sir epidemic transmission between populations, Internat. J. Bifur. Chaos 23 (05) (2013) 1350095.
[151] E. Zagheni, F.C. Billari, P. Manfredi, A. Melegaro, J. Mossong, W.J. Edmunds, Using time-use data to parameterize models for the spread of close-contact infectious diseases, Amer. J. Epidemiol. 168 (9) (2008) 1082–1090.
[152] L. Fumanelli, M. Ajelli, P. Manfredi, A. Vespignani, S. Merler, Inferring the structure of social contacts from demographic data in the analysis of infectious diseases spread, PLoS Comput. Biol. 8 (9) (2012) e1002673.
[153] H.R. Thieme, Mathematics in Population Biology, Princeton University Press, 2003.
[154] A. d'Onofrio, P. Manfredi, Vaccine demand driven by vaccine side effects: dynamic implications for SIR diseases, J. Theoret. Biol. 264 (2) (2010) 237–252.
[155] M. Iannelli, M. Martcheva, X.-Z. Li, Strain replacement in an epidemic model with super-infection and perfect vaccination, Math. Biosci. 195 (1) (2005) 23–46.
[156] F. Ball, T. Britton, T. House, V. Isham, D. Mollison, L. Pellis, G.S. Tomba, Seven challenges for metapopulation models of epidemics, including households models, Epidemics 10 (2015) 63–67.
[157] M. Roberts, V. Andreasen, A. Lloyd, L. Pellis, Nine challenges for deterministic epidemic models, Epidemics 10 (2015) 49–53.
[158] C.J.E. Metcalf, V. Andreasen, O.N. Bjørnstad, K. Eames, W.J. Edmunds, S. Funk, T.D. Hollingsworth, J. Lessler, C. Viboud, B.T. Grenfell, Seven challenges in modeling vaccine preventable diseases, Epidemics 10 (2015) 11–15.
[159] J.O. Lloyd-Smith, D. Mollison, C. Metcalf, P. Klepac, J. Heesterbeek, Challenges in modelling infectious disease dynamics: preface, Epidemics 10 (2015) iii.
[160] R. Albert, A.-L. Barabási, Statistical mechanics of complex networks, Rev. Modern Phys. 74 (1) (2002) 47.
[161] A. Baronchelli, R. Ferrer-i Cancho, R. Pastor-Satorras, N. Chater, M.H. Christiansen, Networks in cognitive science, Trends Cogn. Sci. 17 (7) (2013) 348–360.
[162] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, D.-U. Hwang, Complex networks: Structure and dynamics, Phys. Rep. 424 (4) (2006) 175–308.
[163] G. Caldarelli, et al., Scale-free networks: complex webs in nature and technology, OUP Catalogue.
[164] R. Cohen, S. Havlin, Complex Networks: Structure, Robustness and Function, Cambridge University Press, 2010.
[165] L.d.F. Costa, F.A. Rodrigues, G. Travieso, P.R. Villas Boas, Characterization of complex networks: A survey of measurements, Adv. Phys. 56 (1) (2007) 167–242.
[166] S.N. Dorogovtsev, J.F. Mendes, Evolution of networks, Adv. Phys. 51 (4) (2002) 1079–1187.
[167] M. Newman, Networks: An Introduction, OUP Oxford, 2010.
[168] M.E. Newman, The structure and function of complex networks, SIAM Rev. 45 (2) (2003) 167–256.
[169] M. Girvan, M.E. Newman, Community structure in social and biological networks, Proc. Natl. Acad. Sci. 99 (12) (2002) 7821–7826.
[170] R. Pastor-Satorras, A. Vespignani, Epidemic dynamics and endemic states in complex networks, Phys. Rev. E 63 (6) (2001) 066117.
[171] Y. Moreno, R. Pastor-Satorras, A. Vespignani, Epidemic outbreaks in complex heterogeneous networks, Eur. Phys. J. B 26 (4) (2002) 521–529.
[172] M. Kretzschmar, M. Morris, Measures of concurrency in networks and the spread of infectious disease, Math. Biosci. 133 (2) (1996) 165–195.
[173] D.J. Watts, S.H. Strogatz, Collective dynamics of small-world networks, Nature 393 (6684) (1998) 440–442.
[174] O. Ore, Y. Ore, Theory of Graphs, Vol. 38, American Mathematical Society Providence, 1962.
[175] M.E.J. Newman, Mixing patterns in networks, Phys. Rev. E 67 (2003) 026126.
[176] M.E. Newman, Assortative mixing in networks, Phys. Rev. Lett. 89 (20) (2002) 208701.
[177] M. Molloy, B. Reed, A critical point for random graphs with a given degree sequence, Random Struct. Algorithms 6 (2–3) (1995) 161–180.
[178] M. Molloy, B. Reed, The size of the giant component of a random graph with a given degree sequence, Combin. Probab. Comput. 7 (03) (1998) 295–305.
[179] M.E. Newman, S.H. Strogatz, D.J. Watts, Random graphs with arbitrary degree distributions and their applications, Phys. Rev. E 64 (2) (2001) 026118.
[180] M.E. Newman, D.J. Watts, S.H. Strogatz, Random graph models of social networks, Proc. Natl. Acad. Sci. 99 (Suppl. 1) (2002) 2566–2572.
[181] L.A. Adamic, R.M. Lukose, A.R. Puniyani, B.A. Huberman, Search in power-law networks, Phys. Rev. E 64 (4) (2001) 046135.
[182] K.-I. Goh, B. Kahng, D. Kim, Universal behavior of load distribution in scale-free networks, Phys. Rev. Lett. 87 (27) (2001) 278701.
[183] M. Barthelemy, Betweenness centrality in large complex networks, Eur. Phys. J. B 38 (2) (2004) 163–168.
[184] R. Guimera, S. Mossa, A. Turtschi, L.A. Amaral, The worldwide air transportation network: Anomalous centrality, community structure, and cities' global roles, Proc. Natl. Acad. Sci. 102 (22) (2005) 7794–7799.
[185] E. Ravasz, A.-L. Barabási, Hierarchical organization in complex networks, Phys. Rev. E 67 (2) (2003) 026112.
[186] A. Vázquez, R. Pastor-Satorras, A. Vespignani, Large-scale topological and dynamical properties of the Internet, Phys. Rev. E 65 (6) (2002) 066130.
[187] M.E. Newman, M. Girvan, Finding and evaluating community structure in networks, Phys. Rev. E 69 (2) (2004) 026113.
[188] S.H. Strogatz, Exploring complex networks, Nature 410 (6825) (2001) 268–276. http://dx.doi.org/10.1038/35065725.
[189] S. Fortunato, Community detection in graphs, Phys. Rep. 486 (3) (2010) 75–174.
[190] P. Erdős, A. Rényi, On random graphs, Publ. Math. Debrecen 6 (1959) 290–297.
[191] E.N. Gilbert, Random graphs, Ann. Math. Stat. 30 (4) (1959) 1141–1144.
[192] R. Solomonoff, A. Rapoport, Connectivity of random nets, Bull. Math. Biophys. 13 (2) (1951) 107–117.
[193] S.N. Dorogovtsev, J.F. Mendes, Evolution of Networks: From Biological Nets to the Internet and WWW, OUP Oxford, 2013.
[194] M.E. Newman, Properties of highly clustered networks, Phys. Rev. E 68 (2) (2003) 026121.
[195] R. Monasson, Diffusion, localization and dispersion relations on "small-world" lattices, Eur. Phys. J. B 12 (4) (1999) 555–567.
[196] E.A. Bender, E.R. Canfield, The asymptotic number of labeled graphs with given degree sequences, J. Combin. Theory Ser. A 24 (3) (1978) 296–307.
[197] W. Aiello, F. Chung, L. Lu, A random graph model for massive graphs, in: Proceedings of the Thirty-second Annual ACM Symposium on Theory of Computing, ACM, 2000, pp. 171–180.
[198] G. Caldarelli, A. Capocci, P. De Los Rios, M.A. Munoz, Scale-free networks from varying vertex intrinsic fitness, Phys. Rev. Lett. 89 (25) (2002) 258702.
[199] N. Masuda, K. Aihara, Global and local synchrony of coupled neurons in small-world networks, Biol. Cybernet. 90 (4) (2004) 302–309.
[200] Z. Toroczkai, K.E. Bassler, Network dynamics: Jamming is limited in scale-free systems, Nature 428 (6984) (2004) 716–716.
[201] A.-L. Barabási, H. Jeong, Z. Néda, E. Ravasz, A. Schubert, T. Vicsek, Evolution of the social network of scientific collaborations, Physica A 311 (3) (2002) 590–614.
[202] S. Boccaletti, G. Bianconi, R. Criado, C.I. Del Genio, J. Gómez-Gardeñes, M. Romance, I. Sendiña-Nadal, Z. Wang, M. Zanin, The structure and dynamics of multilayer networks, Phys. Rep. 544 (1) (2014) 1–122.
[203] S.N. Dorogovtsev, J.F.F. Mendes, A.N. Samukhin, Structure of growing networks with preferential linking, Phys. Rev. Lett. 85 (21) (2000) 4633.

[204] M. Marchiori, V. Latora, Harmony in the small-world, Physica A 285 (3) (2000) 539–546.
[205] V. Latora, M. Marchiori, Efficient behavior of small-world networks, Phys. Rev. Lett. 87 (19) (2001) 198701.
[206] A.E. Krause, K.A. Frank, D.M. Mason, R.E. Ulanowicz, W.W. Taylor, Compartments revealed in food-web structure, Nature 426 (6964) (2003) 282–285.
[207] V. Latora, M. Marchiori, Economic small-world behavior in weighted networks, Eur. Phys. J. B 32 (2) (2003) 249–263.
[208] G.A. Polis, Ecology: Stability is woven by complex webs, Nature 395 (6704) (1998) 744–745.
[209] K. McCann, A. Hastings, G.R. Huxel, Weak trophic interactions and the balance of nature, Nature 395 (6704) (1998) 794–798.
[210] O. Sporns, G. Tononi, G.M. Edelman, Connectivity and complexity: the relationship between neuroanatomy and brain dynamics, Neural Netw. 13 (8) (2000) 909–922.
[211] A. Barrat, M. Barthelemy, R. Pastor-Satorras, A. Vespignani, The architecture of complex weighted networks, Proc. Natl. Acad. Sci. USA 101 (11) (2004) 3747–3752.
[212] W. Li, X. Cai, Statistical analysis of airport network of China, Phys. Rev. E 69 (4) (2004) 046106.
[213] D. Li, Y. Jiang, R. Kang, H. Shlomo, Spatial correlation analysis of cascading failures: congestions and blackouts, Sci. Rep. 4 (2014) 5381.
[214] D. Li, K. Kosmas, B. Armin, H. Shlomo, Dimension of spatially embedded networks, Nat. Phys. 7 (6) (2011) 481–484.
[215] G.-Q. Sun, M. Jusup, Z. Jin, Y. Wang, Z. Wang, Pattern transitions in spatial epidemics: Mechanisms and emergent properties, Phys. Life Rev. (2016) http://dx.doi.org/10.1016/j.plrev.2016.08.002.
[216] L. Wang, X. Li, Y. Zhang, Y. Zhang, K. Zhang, Volution of scaling emergence in large-scale spatial epidemic spreading, PLoS One 6 (7) (2011) e21197.
[217] T. Gross, B. Blasius, Adaptive coevolutionary networks: a review, J. R. Soc. Interface 5 (20) (2008) 259–271.
[218] P. Holme, J. Saramäki, Temporal networks, Phys. Rep. 519 (3) (2011) 97–125.
[219] V. Nicosia, V. Latora, Measuring and modelling correlations in multiplex networks, Phys. Rev. E 92 (2015) 032805.
[220] Y. Zhang, L. Wang, Y. Zhang, X. Li, Towards a temporal network analysis of interactive WiFi users, Europhys. Lett. 98 (6) (2012) 68002.
[221] R.M. May, A.L. Lloyd, Infection dynamics on scale-free networks, Phys. Rev. E 64 (6) (2001) 066112.
[222] M.E. Newman, Spread of epidemic disease on networks, Phys. Rev. E 66 (1) (2002) 016128.
[223] M. Boguná, R. Pastor-Satorras, Epidemic spreading in correlated complex networks, Phys. Rev. E 66 (4) (2002) 047104.
[224] D.S. Callaway, M.E. Newman, S.H. Strogatz, D.J. Watts, Network robustness and fragility: Percolation on random graphs, Phys. Rev. Lett. 85 (25) (2000) 5468.
[225] R. Cohen, K. Erez, D. Ben-Avraham, S. Havlin, Resilience of the Internet to random breakdowns, Phys. Rev. Lett. 85 (21) (2000) 4626.
[226] R. Parshani, S. Carmi, S. Havlin, Epidemic threshold for the susceptible-infectious-susceptible model on random networks, Phys. Rev. Lett. 104 (25) (2010) 258701.
[227] P. Van Mieghem, E. Cator, Epidemics in networks with nodal self-infection and the epidemic threshold, Phys. Rev. E 86 (1) (2012) 016116.
[228] P. Van Mieghem, J. Omic, R. Kooij, Virus spread in networks, IEEE/ACM Trans. Netw. 17 (1) (2009) 1–14.
[229] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, A. Vespignani, Epidemic processes in complex networks, Rev. Modern Phys. 87 (3) (2015) 925.
[230] M. Salehi, R. Sharma, M. Marzolla, M. Magnani, P. Siyari, D. Montesi, Spreading processes in multilayer networks, IEEE Trans. Netw. Sci. Eng. 2 (2) (2015) 65–83.
[231] B. Min, K.-I. Goh, Layer-crossing overhead and information spreading in multiplex social networks, Bull. Amer. Phys. Soc. 59 (1) (2014).
[232] M. Magnani, L. Rossi, The ml-model for multi-layer social networks, in: Advances in Social Networks Analysis and Mining, ASONAM, 2011 International Conference on, IEEE, 2011, pp. 5–12.
[233] H. Wang, Q. Li, G. D'Agostino, S. Havlin, H.E. Stanley, P. Van Mieghem, Effect of the interconnected network structure on the epidemic threshold, Phys. Rev. E 88 (2) (2013) 022801.
[234] K.-M. Lee, J.Y. Kim, S. Lee, K.-I. Goh, Multiplex networks, in: Networks of Networks: The Last Frontier of Complexity, Springer, 2014, pp. 53–72.
[235] M. Berlingerio, M. Coscia, F. Giannotti, A. Monreale, D. Pedreschi, Multidimensional networks: foundations of structural analysis, World Wide Web 16 (5–6) (2013) 567–593.
[236] S. Wasserman, K. Faust, Social Network Analysis: Methods and Applications, Vol. 8, Cambridge University Press, 1994.
[237] J. Gao, D. Li, S. Havlin, From a single network to a network of networks, Natl. Sci. Rev. 1 (3) (2014) 346–356.
[238] M. Dickison, S. Havlin, H.E. Stanley, Epidemics on interconnected networks, Phys. Rev. E 85 (6) (2012) 066109.
[239] D. Zhao, L. Li, H. Peng, Q. Luo, Y. Yang, Multiple routes transmitted epidemics on multiplex networks, Phys. Lett. A 378 (10) (2014) 770–776.
[240] W. Wang, M. Tang, H.-F. Zhang, H. Gao, Y. Do, Z.-H. Liu, Epidemic spreading on weighted networks with general weight and degree distributions, Phys. Rev. E 90 (4) (2014) 042803.
[241] J. Goldenberg, B. Libai, E. Muller, Talk of the network: A complex systems look at the underlying process of word-of-mouth, Mark. Lett. 12 (3) (2001) 211–223.
[242] Y. Wang, G. Xiao, L. Wong, X. Fu, S. Ma, T. Cheng, Effects of fear factors in disease propagation, J. Phys. A: Math. Gen. 44 (35) (2011) 355101.
[243] D. Qian, O. Yagan, L. Yang, J. Zhang, Diffusion of real-time information in social-physical networks, in: Global Communications Conference, GLOBECOM, 2012 IEEE, IEEE, 2012, pp. 2072–2077.
[244] V. Marceau, P.-A. Noël, L. Hébert-Dufresne, A. Allard, L.J. Dubé, Modeling the dynamical interaction between epidemics on overlay networks, Phys. Rev. E 84 (2) (2011) 026105.
[245] C. Buono, L.G. Alvarez-Zuzek, P.A. Macri, L.A. Braunstein, Epidemics in partially overlapped multiplex networks, PLoS One 9 (3) (2014) e92200.
[246] S. Funk, V.A. Jansen, Interacting epidemics on overlay networks, Phys. Rev. E 81 (3) (2010) 036118.
[247] J. Sanz, C.-Y. Xia, S. Meloni, Y. Moreno, Dynamics of interacting diseases, Phys. Rev. X 4 (4) (2014) 041005.
[248] K.T. Eames, M.J. Keeling, Modeling dynamic and network heterogeneities in the spread of sexually transmitted diseases, Proc. Natl. Acad. Sci. 99 (20) (2002) 13330–13335.
[249] S. Eubank, H. Guclu, V.A. Kumar, M.V. Marathe, A. Srinivasan, Z. Toroczkai, N. Wang, Modelling disease outbreaks in realistic urban social networks, Nature 429 (6988) (2004) 180–184.
[250] L.A. Meyers, B. Pourbohloul, M.E. Newman, D.M. Skowronski, R.C. Brunham, Network theory and SARS: predicting outbreak diversity, J. Theoret. Biol. 232 (1) (2005) 71–81.
[251] J.M. Read, M.J. Keeling, Disease evolution on networks: the role of contact structure, Proc. R. Soc. Lond. Biol. 270 (1516) (2003) 699–708.
[252] J. Wallinga, W.J. Edmunds, M. Kretzschmar, Perspective: human contact patterns and the spread of airborne infectious diseases, TIM 7 (9) (1999) 372–377.
[253] D.L. Chao, M.E. Halloran, V.J. Obenchain, I.M. Longini Jr., FluTE, a publicly available stochastic influenza epidemic simulation model, PLoS Comput. Biol. 6 (1) (2010) e1000656.
[254] S. Merler, M. Ajelli, The role of population heterogeneity and human mobility in the spread of pandemic influenza, Proc. R. Soc. Lond. Biol. 277 (1681) (2010) 557–565.
[255] G.P. Garnett, R.M. Anderson, Sexually transmitted diseases and sexual behavior: insights from mathematical models, J. Infect. Dis. 174 (Suppl. 2) (1996) S150–S161.
[256] A.C. Ghani, J. Swinton, G.P. Garnett, The role of sexual partnership networks in the epidemiology of gonorrhea, Sex. Transm. Dis. 24 (1) (1997) 45–56.
[257] M. Morris, M. Kretzschmar, Concurrent partnerships and the spread of HIV, AIDS 11 (5) (1997) 641–648.
[258] J.J. Potterat, R.B. Rothenberg, S.Q. Muth, Network structural dynamics and infectious disease propagation, Int. J. STD AIDS 10 (3) (1999) 182–185.
[259] A.S. Klovdahl, Networks and pathogens, Sex. Transm. Dis. 28 (1) (2001) 25–28.
[260] R. Rothenberg, How a net works: implications of network structure for the persistence and control of sexually transmitted diseases and hiv, Sex. Transm. Dis. 28 (2) (2001) 63–68.
[261] P. McElroy, R. Rothenberg, R. Varghese, R. Woodruff, G. Minns, S. Muth, L. Lambert, R. Ridzon, A network-informed approach to investigating a tuberculosis outbreak: implications for enhancing contact investigations, Int. J. Tuberc. Lung Dis. 7 (Suppl. 3) (2003) S486–S493.
[262] B. Szendroi, G. Csányi, Polynomial epidemics and clustering in contact networks, Proc. R. Soc. Lond. Biol. 271 (Suppl. 5) (2004) S364–S366.

[263] I.A. Doherty, N.S. Padian, C. Marlow, S.O. Aral, Determinants and consequences of sexual networks as they affect the spread of sexually transmitted infections, J. Infec. Dis. 191 (Suppl. 1) (2005) S42–S54.
[264] M.E. Halloran, I.M. Longini, A. Nizam, Y. Yang, Containing bioterrorist smallpox, Science 298 (5597) (2002) 1428–1432.
[265] S. Cohen, W.J. Doyle, D.P. Skoner, B.S. Rabin, J.M. Gwaltney, Social ties and susceptibility to the common cold, JAMA 277 (24) (1997) 1940–1944.
[266] R. Olinky, L. Stone, Unexpected epidemic thresholds in heterogeneous networks: The role of disease transmission, Phys. Rev. E 70 (3) (2004) 030902.
[267] M.J. Keeling, K.T. Eames, Networks and epidemic models, J. R. Soc. Interface 2 (4) (2005) 295–307.
[268] R. Albert, H. Jeong, A.-L. Barabási, Error and attack tolerance of complex networks, Nature 406 (6794) (2000) 378–382.
[269] N. Madar, T. Kalisky, R. Cohen, D. ben Avraham, S. Havlin, Immunization and epidemic dynamics in complex networks, Eur. Phys. J. B 38 (2) (2004) 269–276.
[270] R.M. Anderson, R.M. May, B. Anderson, Infectious Diseases of Humans: Dynamics and Control, Vol. 28, Wiley Online Library, 1992.
[271] J. Hindes, I. Schwartz, Epidemic extinction and control in heterogeneous networks, Phys. Rev. Lett. 117 (2016) 028302.
[272] J. Heesterbeek, Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation, Vol. 5, John Wiley & Sons, 2000.
[273] A.-L. Barabási, et al., Scale-free networks: a decade and beyond, Science 325 (5939) (2009) 412.
[274] R. Pastor-Satorras, A. Vespignani, Immunization of complex networks, Phys. Rev. E 65 (3) (2002) 036104.
[275] L.K. Gallos, F. Liljeros, P. Argyrakis, A. Bunde, S. Havlin, Improving immunization strategies, Phys. Rev. E 75 (4) (2007) 045104.
[276] P. Liu, H. Miao, Q. Li, A common acquaintance immunization strategy for complex network, in: Computer and Information Science, 2009, ICIS 2009. Eighth IEEE/ACIS International Conference on, IEEE, 2009, pp. 713–717.
[277] R. Cohen, S. Havlin, D. Ben-Avraham, Efficient immunization strategies for computer networks and populations, Phys. Rev. Lett. 91 (24) (2003) 247901.
[278] N.A. Christakis, J.H. Fowler, Social network sensors for early detection of contagious outbreaks, PLoS One 5 (9) (2010) e12948.
[279] K. Krieger, Focus: Vaccinate thy neighbor, Physics 12 (2003) 23.
[280] A. Barrat, M. Barthelemy, A. Vespignani, Dynamical Processes on Complex Networks, Cambridge University Press, 2008.
[281] A.-L. Barabási, Network Science, Cambridge University Press, 2016.
[282] S. Bornholdt, H.G. Schuster, Handbook of Graphs and Networks: From the Genome to the Internet, John Wiley & Sons, Inc., 2003.
[283] C. Gao, J. Liu, N. Zhong, Network immunization with distributed autonomy-oriented entities, IEEE Trans. Parallel Distrib. Syst. 22 (7) (2011) 1222–1229.
[284] Y. Liu, Y. Deng, M. Jusup, Z. Wang, A biologically inspired immunization strategy for network epidemiology, J. Theoret. Biol. 400 (2016) 92–102.
[285] D.M. Cornforth, T.C. Reluga, E. Shim, C.T. Bauch, A.P. Galvani, L.A. Meyers, Erratic flu vaccination emerges from short-sighted behavior in contact networks, PLoS Comput. Biol. 7 (1) (2011) e1001062.
[286] Y. Wang, D. Chakrabarti, C. Wang, C. Faloutsos, Epidemic spreading in real networks: An eigenvalue viewpoint, in: Reliable Distributed Systems, 2003, Proceedings, 22nd International Symposium on, IEEE, 2003, pp. 25–34.
[287] D. Chakrabarti, Y. Wang, C. Wang, J. Leskovec, C. Faloutsos, Epidemic thresholds in real networks, ACM Trans. Inf. Syst. Secur. 10 (4) (2008) 1.
[288] J.G. Restrepo, E. Ott, B.R. Hunt, Characterizing the dynamical importance of network nodes and links, Phys. Rev. Lett. 97 (9) (2006) 094102.
[289] R. Cohen, K. Erez, D. Ben-Avraham, S. Havlin, Breakdown of the Internet under intentional attack, Phys. Rev. Lett. 86 (16) (2001) 3682.
[290] A.L. Lloyd, R.M. May, How viruses spread among computers and people, Science 292 (5520) (2001) 1316.
[291] H. Jeong, B. Tombor, R. Albert, Z.N. Oltvai, A.-L. Barabási, The large-scale organization of metabolic networks, Nature 407 (6804) (2000) 651–654.
[292] Z. Dezső, A.-L. Barabási, Halting viruses in scale-free networks, Phys. Rev. E 65 (5) (2002) 055103.
[293] Y. Chen, G. Paul, S. Havlin, F. Liljeros, H.E. Stanley, Finding a better immunization strategy, Phys. Rev. Lett. 101 (5) (2008) 058701.
[294] S.P. Borgatti, Centrality and network flow, Social Networks 27 (1) (2005) 55–71.
[295] K.T. Eames, J.M. Read, W.J. Edmunds, Epidemic prediction and control in weighted networks, Epidemics 1 (1) (2009) 70–76.
[296] J.C. Miller, J.M. Hyman, Effective vaccination strategies for realistic social networks, Physica A 386 (2) (2007) 780–785.
[297] C.M. Schneider, T. Mihaljev, S. Havlin, H.J. Herrmann, Suppressing epidemics with a limited amount of immunization units, Phys. Rev. E 84 (6) (2011) 061911.
[298] B. Vidondo, M. Schwehm, A. Bühlmann, M. Eichner, Finding and removing highly connected individuals using suboptimal vaccines, BMC Infect. Dis. 12 (1) (2012) 51.
[299] Y. Wang, G. Xiao, J. Hu, T. Cheng, L. Wang, Imperfect targeted immunization in scale-free networks, Physica A 388 (12) (2009) 2535–2546.
[300] D.-M. B. Friedl, J. Heidemann, et al., A critical review of centrality measures in social networks, Bus. Inf. Syst. Eng. 2 (6) (2010) 371–385.
[301] L. Hébert-Dufresne, A. Allard, J.-G. Young, L.J. Dubé, Global efficiency of local immunization on complex networks, Sci. Rep. 3 (2013) 2171.
[302] V. Latora, M. Marchiori, A measure of centrality based on network efficiency, New J. Phys. 9 (6) (2007) 188.
[303] M. Ventresca, D. Aleman, Evaluation of strategies to mitigate contagion spread using social network characteristics, Social Networks 35 (1) (2013) 75–88.
[304] M. Salathé, J.H. Jones, Dynamics and control of diseases in networks with community structure, PLoS Comput. Biol. 6 (4) (2010) e1000736.
[305] C.M. Schneider, T. Mihaljev, H.J. Herrmann, Inverse targeting–an effective immunization strategy, Europhys. Lett. 98 (4) (2012) 46002.
[306] P. Bonacich, Power and centrality: A family of measures, Amer. J. Sociol. (1987) 1170–1182.
[307] L. Page, S. Brin, R. Motwani, T. Winograd, The PageRank citation ranking: bringing order to the web (1999–66).
[308] H. Yang, B. Wang, Immunization of traffic-driven epidemic spreading, Physica A 443 (2016) 86–90.
[309] H. Yang, W. Wang, Y. Lai, Y. Xie, B. Wang, Control of epidemic spreading on complex networks by local traffic dynamics, Phys. Rev. E 84 (4) (2011) 045101.
[310] W. Du, B. Liang, G. Yan, O. Lordan, X. Cao, Identifying vital edges in Chinese air route network via memetic algorithm, 2016, arXiv preprint, arXiv:1608.00142.
[311] L.C. Freeman, Centrality in social networks conceptual clarification, Social Networks 1 (3) (1978) 215–239.
[312] R.M. Christley, G. Pinchbeck, R. Bowers, D. Clancy, N. French, R. Bennett, J. Turner, Infection in social networks: using network analysis to identify high-risk individuals, Amer. J. Epidemiol. 162 (10) (2005) 1024–1031.
[313] P. Holme, B.J. Kim, C.N. Yoon, S.K. Han, Attack vulnerability of complex networks, Phys. Rev. E 65 (5) (2002) 056109.
[314] Y. Yu, T.Y. Berger-Wolf, J. Saia, et al., Finding spread blockers in dynamic networks, in: Advances in Social Network Mining and Analysis, Springer, 2010, pp. 55–76.
[315] B. Shams, Using network properties to evaluate targeted immunization algorithms, Netw. Biol. 4 (3) (2014) 74.
[316] M.E. Newman, A measure of betweenness centrality based on random walks, Social Networks 27 (1) (2005) 39–54.
[317] P. Bonacich, Factoring and weighting approaches to status scores and clique identification, J. Math. Sociol. 2 (1) (1972) 113–120.
[318] P. Bonacich, Some unique properties of eigenvector centrality, Social Networks 29 (4) (2007) 555–564.
[319] I. Tomovski, L. Kocarev, Simple algorithm for virus spreading control on complex networks, IEEE Trans. Circuits Syst. I 59 (4) (2012) 763–771.
[320] N. Masuda, Immunization of networks with community structure, New J. Phys. 11 (12) (2009) 123018.
[321] S.L. Feld, Why your friends have more friends than you do, Amer. J. Sociol. (1991) 1464–1477.
[322] J.D. Noh, H. Rieger, Random walks on complex networks, Phys. Rev. Lett. 92 (11) (2004) 118701.
[323] H. Ke, T. Yi, Immunization for scale-free networks by random walker, Chin. Phys. 15 (12) (2006) 2782.
[324] P. Holme, Efficient local strategies for vaccination and network attack, Europhys. Lett. 68 (6) (2004) 908.
[325] S. Lee, S.-H. Yook, Y. Kim, Centrality measure of complex networks using biased random walks, Eur. Phys. J. B 68 (2) (2009) 277–281.
[326] J. Gómez-Gardenes, P. Echenique, Y. Moreno, Immunization of real complex communication networks, Eur. Phys. J. B 49 (2) (2006) 259–264.
[327] P. Echenique, J. Gómez-Gardeñes, Y. Moreno, A. Vázquez, Distance-d covering problems in scale-free networks with degree correlations, Phys. Rev. E 71 (3) (2005) 035102.
[328] G. Palla, I. Derényi, I. Farkas, T. Vicsek, Uncovering the overlapping community structure of complex networks in nature and society, Nature 435 (7043) (2005) 814–818.

[329] M.E. Newman, Modularity and community structure in networks, Proc. Natl. Acad. Sci. 103 (23) (2006) 8577–8582.
[330] Z. Liu, B. Hu, Epidemic spreading in community networks, Europhys. Lett. 72 (2) (2005) 315.
[331] X. Wu, Z. Liu, How community structure influences epidemic spread in social networks, Physica A 387 (2) (2008) 623–630.
[332] T. Gross, H. Sayama, Adaptive Networks, Springer, 2009.
[333] L.B. Shaw, I.B. Schwartz, Fluctuating epidemics on adaptive networks, Phys. Rev. E 77 (6) (2008) 066101.
[334] V. Marceau, P.-A. Noël, L. Hébert-Dufresne, A. Allard, L.J. Dubé, Adaptive networks: Coevolution of disease and topology, Phys. Rev. E 82 (3) (2010) 036116.
[335] L.B. Shaw, I.B. Schwartz, Enhanced vaccine control of epidemics in adaptive networks, Phys. Rev. E 81 (4) (2010) 046120.
[336] C. Gao, J. Liu, Modeling and restraining mobile virus propagation, IEEE Trans. Mobile Comput. 12 (3) (2013) 529–541.
[337] M. Kivelä, A. Arenas, M. Barthelemy, J.P. Gleeson, Y. Moreno, M.A. Porter, Multilayer networks, J. Complex Netw. 2 (3) (2014) 203–271.
[338] W. Wang, Q.-H. Liu, S.-M. Cai, M. Tang, L.A. Braunstein, H.E. Stanley, Suppressing disease spreading by using information diffusion on multiplex networks, Sci. Rep. 6 (2016) 29259.
[339] W. Wang, M. Tang, H. Yang, Y. Do, Y.-C. Lai, G. Lee, Asymmetrically interacting spreading dynamics on complex layered networks, Sci. Rep. 4 (2014) 5097.
[340] Q.-H. Liu, W. Wang, M. Tang, H.-F. Zhang, Impacts of complex behavioral responses on asymmetric interacting spreading dynamics in multiplex networks, Sci. Rep. 6 (2016) 25617.
[341] S. Gomez, A. Diaz-Guilera, J. Gomez-Gardeñes, C.J. Perez-Vicente, Y. Moreno, A. Arenas, Diffusion dynamics on multiplex networks, Phys. Rev. Lett. 110 (2) (2013) 028701.
[342] E. Cozzo, R.A. Banos, S. Meloni, Y. Moreno, Contact-based social contagion in multiplex networks, Phys. Rev. E 88 (5) (2013) 050801.
[343] D. Zhao, L. Wang, S. Li, Z. Wang, L. Wang, B. Gao, Immunization of epidemics in multiplex networks, PLoS One 9 (11) (2014) e112018.
[344] D. Zhao, Z. Wang, G. Xiao, B. Gao, L. Wang, The robustness of interdependent networks under the interplay between cascading failures and virus propagation, Europhys. Lett. 115 (2016) 58004.
[345] B. Gao, D. Zhao, Competing spreading processes and immunization in multiplex networks, 2016, arXiv preprint, arXiv:1608.01038.
[346] D. Zhao, L. Wang, L. Xu, Z. Wang, Finding another yourself in multiplex networks, Appl. Math. Comput. 266 (2015) 599–604.
[347] J.-P. Onnela, J. Saramäki, J. Hyvönen, G. Szabó, D. Lazer, K. Kaski, J. Kertész, A.-L. Barabási, Structure and tie strengths in mobile communication networks, Proc. Natl. Acad. Sci. 104 (18) (2007) 7332–7336.
[348] J. Zhao, J. Wu, K. Xu, Weak ties: subtle role of information diffusion in online social networks, Phys. Rev. E 82 (1) (2010) 016105.
[349] K. Gong, M. Tang, P.M. Hui, H.F. Zhang, D. Younghae, Y.-C. Lai, An efficient immunization strategy for community networks, PLoS One 8 (12) (2013) e83489.
[350] R. Guimera, L.A.N. Amaral, Functional cartography of complex metabolic networks, Nature 433 (7028) (2005) 895–900.
[351] H. Yang, M. Tang, H.-F. Zhang, Efficient community-based control strategies in adaptive networks, New J. Phys. 14 (12) (2012) 123017.
[352] Z. Ruan, M. Tang, Z. Liu, Epidemic spreading with information-driven vaccination, Phys. Rev. E 86 (3) (2012) 036117.
[353] D. Zhao, H. Peng, L. Li, Y. Yang, S. Li, An efficient patch dissemination strategy for mobile networks, Math. Probl. Eng. 2013 (2013) 896187.
[354] J. Zhou, G. Xiao, S. Cheong, X. Fu, L. Wong, S. Ma, T. Cheng, Epidemic reemergence in adaptive complex networks, Phys. Rev. E 85 (3) (2012) 036107.
[355] D.-w. Zhao, L.-h. Wang, Y.-f. Zhi, J. Zhang, Z. Wang, The robustness of multiplex networks under layer node-based attack, Sci. Rep. 6 (2016) 24304.
[356] H. Peng, N. Wei, L. Li, W. Xie, Y. Yang, Models and synchronization of time-delayed complex dynamical networks with multi-links based on adaptive control, Phys. Lett. A 374 (23) (2010) 2335–2339.
[357] W.-B. Du, X.-L. Zhou, O. Lordan, Z. Wang, C. Zhao, Y.-B. Zhu, Analysis of the chinese airline network as multi-layer networks, Transp. Res. E Logist. Transp. Rev. 89 (2016) 108–116.
[358] Y. Hu, S. Havlin, H.A. Makse, Conditions for viral influence spreading through multiplex correlated social networks, Phys. Rev. X 4 (2) (2014) 021031.
[359] T.J. Zagenczyk, R.L. Purvis, Multiplex social network ties and psychological climate, in: Academy of Management Proceedings, vol. 2015, Academy of Management, 2015, p. 17574.
[360] T. Ma, J. Zhou, M. Tang, Y. Tian, A. Al-Dhelaan, M. Al-Rodhaan, S. Lee, Social network and tag sources based augmenting collaborative recommender system, IEICE Trans. Inf. Syst. E98-D (4) (2015) 902–910.
[361] S. Xie, Y. Wang, Construction of tree network with limited delivery latency in homogeneous wireless sensor networks, Wirel. Pers. Commun. 78 (1) (2014) 231–246.
[362] J. Shen, H. Tan, J. Wang, J. Wang, S. Lee, A novel routing protocol providing good transmission reliability in underwater sensor networks, J. Internet Technol. 16 (1) (2015) 171–178.
[363] P. Guo, J. Wang, X.H. Geng, C.S. Kim, J.-U. Kim, A variable threshold-value authentication architecture for wireless mesh networks, J. Internet Technol. 15 (6) (2014) 929–936.
[364] D. Centola, The spread of behavior in an online social network experiment, Science 329 (5996) (2010) 1194–1197.
[365] S. Hosseini, K. Barker, J.E. Ramirez-Marquez, A review of definitions and measures of system resilience, Reliab. Eng. Syst. Saf. 145 (2016) 47–61.
[366] W.-B. Du, X.-L. Zhou, M. Jusup, Z. Wang, Physics of transportation: towards optimal capacity using the multilayer network framework, Sci. Rep. 6 (2016) 19059.
[367] M. De Domenico, S. Sasai, A. Arenas, Mapping multiplex hubs in human functional brain network, arXiv preprint arXiv:1603.05897.
[368] O. Sporns, R.F. Betzel, Modular brain networks, Ann. Rev. Psych. 67 (2016) 613–640.
[369] F. Battiston, J. Iacovacci, V. Nicosia, G. Bianconi, V. Latora, Emergence of multiplex communities in collaboration networks, arXiv preprint arXiv:1506.01280.
[370] Z. Wang, D.-W. Zhao, L. Wang, G.-Q. Sun, Z. Jin, Immunity of multiplex networks via acquaintance vaccination, Europhys. Lett. 112 (4) (2015) 48002.
[371] C. Buono, L.A. Braunstein, Immunization strategy for epidemic spreading on multilayer networks, Europhys. Lett. 109 (2) (2015) 26001.
[372] L.G.A. Zuzek, C. Buono, L.A. Braunstein, Epidemic spreading and immunization strategy in multiplex networks, in: Journal of Physics: Conference Series, vol. 640, IOP Publishing, 2015, p. 012007.
[373] B. Min, S. Do Yi, K.-M. Lee, K.-I. Goh, Network robustness of multiplex networks with interlayer degree correlations, Phys. Rev. E 89 (4) (2014) 042811.
[374] A. Vaidya, Modeling the spread and control of a contagion on multiplex social networks, in: 2016 AAAS Annual Meeting (February 11–15, 2016), aaas, 2016.
[375] G. Fink, I. Günther, K. Hill, The effect of water and sanitation on child health: evidence from the demographic and health surveys 1986–2007, Int. J. Epidemiol. 40 (5) (2011) 1196–1204.
[376] F. Fenner, D.A. Henderson, I. Arita, Z. Jezek, I.D. Ladnyi, et al., Smallpox and its Eradication, World Health Organization, Geneva, 1988.
[377] A. Dabbagh, M. Gacic-Dobo, E. Simons, D. Featherstone, P. Strebel, J. Okwo-Bele, E. Hoekstra, M. Chopra, A. Uzicanin, S. Cochi, et al., Global measles mortality, 2000–2008, Morbidity and Mortality Weekly Report 58 (47) (2009) 1321–1326.
[378] A. Levin, C. Burgess, L.P. Garrison, C. Bauch, J. Babigumira, E. Simons, A. Dabbagh, Global eradication of measles: an epidemiologic and economic evaluation, J. Infect. Dis. 204 (Suppl. 1) (2011) S98–S106.
[379] S. Cousins, Syrian crisis: health experts say more can be done, Lancet 385 (9972) (2015) 931–934.
[380] P.V. Ganapathiraju, C.B. Morssink, J. Plumb, Endgame for polio eradication? options for overcoming social and political factors in the progress to eradicating polio, Global Public Health 10 (4) (2015) 463–473.
[381] S.L. Cochi, A. Freeman, S. Guirguis, H. Jafari, B. Aylward, Global polio eradication initiative: lessons learned and legacy, J. Infect. Dis. 210 (Suppl. 1) (2014) S540–S546.
[382] H. De Melker, R. Pebody, W. Edmunds, D. Lévy-Bruhl, M. Valle, M. Rota, S. Salmaso, S. Van Den Hof, G. Berbers, P. Saliou, et al., The seroepidemiology of measles in Western Europe, Epidemiol. Infect. 126 (02) (2001) 249–259.
[383] K.F. Brown, J.S. Kroll, M.J. Hudson, M. Ramsay, J. Green, S.J. Long, C.A. Vincent, G. Fraser, N. Sevdalis, Factors underlying parental decisions about combination childhood vaccinations including MMR: a systematic review, Vaccine 28 (26) (2010) 4235–4248.

[384] H.J. Larson, C. Jarrett, E. Eckersberger, D.M. Smith, P. Paterson, Understanding vaccine hesitancy around vaccines and vaccination from a global perspective: A systematic review of published literature, 2007–2012, Vaccine 32 (19) (2014) 2150–2159.
[385] L. Serpell, J. Green, Parental decision-making in childhood vaccination, Vaccine 24 (19) (2006) 4041–4046.
[386] G.B. Chapman, E.J. Coups, Predictors of influenza vaccine acceptance among healthy adults, Prev. Med. 29 (4) (1999) 249–262.
[387] Y. Ibuka, M. Li, J. Vietri, G.B. Chapman, A.P. Galvani, Free-riding behavior in vaccination decisions: An experimental study, PLoS One 9 (1) (2014) e87164.
[388] B. Swennen, P. Van Damme, A. Vellinga, Y. Coppieters, A. Depoorter, Analysis of factors influencing vaccine uptake: perspectives from belgium, Vaccine 20 (2001) S5–S7.
[389] K.P. Goldstein, T.J. Philipson, H. Joo, R.S. Daum, The effect of epidemic measles on immunization rates, JAMA 276 (1) (1996) 56–58.
[390] C.T. Bauch, S. Bhattacharyya, Evolutionary game theory and social learning can determine how vaccine scares unfold, PLoS Comput. Biol. 8 (4) (2012) e1002452.
[391] CDPHA, Immunization rates in child care and schools, california department of public health, https://www.cdph.ca.gov/programs/immunize/Pages/ImmunizationLevels.aspx, 2016.
[392] J.D. Allen, M.K. Othus, R.C. Shelton, Y. Li, N. Norman, L. Tom, M.G. del Carmen, Parental decision making about the HPV vaccine, Cancer Epidemiol. Biomarkers Prevent. 19 (9) (2010) 2187–2198.
[393] M. Salathé, M. Kazandjieva, J.W. Lee, P. Levis, M.W. Feldman, J.H. Jones, A high-resolution human contact network for infectious disease transmission, Proc. Natl. Acad. Sci. 107 (51) (2010) 22020–22025.
[394] N.A. Christakis, J.H. Fowler, The spread of obesity in a large social network over 32 years, New Engl. J. Med. 357 (4) (2007) 370–379.
[395] E. Campbell, M. Salathé, Complex social contagion makes networks more vulnerable to disease outbreaks, Sci. Rep. 3 (2013) 1905.
[396] S. Bansal, B.T. Grenfell, L.A. Meyers, When individual behaviour matters: homogeneous and network models in epidemiology, J. R. Soc. Interface 4 (16) (2007) 879–891.
[397] Z. Wang, M.A. Andrews, Z.-X. Wu, L. Wang, C.T. Bauch, Coupled disease–behavior dynamics on complex networks: A review, Phys. Life Rev. 15 (2015) 1–29.
[398] D. He, J. Dushoff, T. Day, J. Ma, D.J. Earn, Inferring the causes of the three waves of the 1918 influenza pandemic in England and Wales, Proc. R. Soc. Lond. Biol. 280 (1766) (2013) 20131345.
[399] T. Oraby, V. Thampi, C.T. Bauch, The influence of social norms on the dynamics of vaccinating behaviour for paediatric infectious diseases, Proc. R. Soc. Lond. Biol. 281 (1780) (2014) 20133172.
[400] G. Chowell, C. Ammon, N. Hengartner, J. Hyman, Transmission dynamics of the great influenza pandemic of 1918 in Geneva, Switzerland: assessing the effects of hypothetical interventions, J. Theoret. Biol. 241 (2) (2006) 193–204.
[401] M.C. Bootsma, N.M. Ferguson, The effect of public health measures on the 1918 influenza pandemic in US cities, Proc. Natl. Acad. Sci. 104 (18) (2007) 7588–7593.
[402] A.P. Galvani, T.C. Reluga, G.B. Chapman, Long-standing influenza vaccination policy is in accord with individual self-interest but not with the utilitarian optimum, Proc. Natl. Acad. Sci. 104 (13) (2007) 5692–5697.
[403] E. Shim, G.B. Chapman, J.P. Townsend, A.P. Galvani, The influence of altruism on influenza vaccination decisions, J. R. Soc. Interface (2012) rsif20120115.
[404] S. Funk, M. Salathé, V.A. Jansen, Modelling the influence of human behaviour on the spread of infectious diseases: a review, J. R. Soc. Interface 7 (50) (2010) 1247–1256.
[405] N. Perra, D. Balcan, B. Gonçalves, A. Vespignani, Towards a characterization of behavior-disease models, PLoS One 6 (8) (2011) e23084.
[406] V. Capasso, G. Serio, A generalization of the Kermack-McKendrick deterministic epidemic model, Math. Biosci. 42 (1) (1978) 43–61.
[407] A. Alshamsi, F. Pianesi, B. Lepri, A. Pentland, I. Rahwan, Beyond contagion: reality mining reveals complex patterns of social influence, PLoS One 10 (8) (2015) e0135740.
[408] R.B. Cialdini, R.R. Reno, C.A. Kallgren, A focus theory of normative conduct: recycling the concept of norms to reduce littering in public places, J. Pers. Soc. Psychol. 58 (6) (1990) 1015.
[409] A. Bandura, R.H. Walters, Social Learning and Personality Development, vol. 14, JSTOR, 1963.
[410] S. Atran, The trouble with memes, Hum. Nature 12 (4) (2001) 351–381.
[411] S. Funk, E. Gilad, C. Watkins, V.A. Jansen, The spread of awareness and its impact on epidemic outbreaks, Proc. Natl. Acad. Sci. USA 106 (16) (2009) 6872–6877.
[412] J. Von Neumann, O. Morgenstern, Theory of Games and Economic Behavior, Princeton University Press, 2007.
[413] S. May, Enhanced: simple rules with complex dynamics, Science 287 (5453) (2000) 601–602.
[414] C.T. Bauch, A.P. Galvani, D.J. Earn, Group interest versus self-interest in smallpox vaccination policy, Proc. Natl. Acad. Sci. USA 100 (18) (2003) 10564–10567.
[415] C.T. Bauch, D.J. Earn, Vaccination and the theory of games, Proc. Natl. Acad. Sci. U.S.A. 101 (36) (2004) 13391–13394.
[416] D.L. Brito, E. Sheshinski, M.D. Intriligator, Externalities and compulsory vaccinations, J. Publ. Econ. 45 (1) (1991) 69–90.
[417] G. Heal, H. Kunreuther, The vaccination game, Risk Management and Decision Processes Center Working Paper (05-10).
[418] J.M. Smith, Evolution and the Theory of Games, Cambridge University Press, 1982.
[419] M.A. Nowak, R.M. May, Evolutionary games and spatial chaos, Nature 359 (6398) (1992) 826–829.
[420] G. Szabó, C. Tőke, Evolutionary prisoner's dilemma game on a square lattice, Phys. Rev. E 58 (1) (1998) 69.
[421] C. Hauert, M. Doebeli, Spatial structure often inhibits the evolution of cooperation in the snowdrift game, Nature 428 (6983) (2004) 643–646.
[422] C.-Y. Xia, S. Meloni, M. Perc, M. Yamir, Dynamic instability of cooperation due to diverse activity patterns in evolutionary social dilemmas, Europhys. Lett. 109 (5) (2015) 58002.
[423] G. Szabó, G. Fath, Evolutionary games on graphs, Phys. Rep. 446 (4) (2007) 97–216.
[424] J. Tanimoto, Dilemma solving by the coevolution of networks and strategy in a 2×2 game, Phys. Rev. E 76 (2) (2007) 021126.
[425] Y. Deng, Deng entropy, Chaos Solitons Fractals 91 (2016) 549–553.
[426] Y. Deng, Generalized evidence theory, Appl. Intell. 43 (3) (2015) 530–543.
[427] X. Deng, Z. Wang, Q. Liu, Y. Deng, S. Mahadevan, A belief-based evolutionarily stable strategy, J. Theoret. Biol. 361 (2014) 81–86.
[428] X. Deng, Q. Zhang, Y. Deng, Z. Wang, A novel framework of classical and quantum prisoner's dilemma games on coupled networks, Sci. Rep. 6 (2016) 23024.
[429] K. Huang, X. Zheng, Z. Li, Y. Yang, Understanding cooperative behavior based on the coevolution of game strategy and link weight, Sci. Rep. 5 (2015) 14783.
[430] K. Huang, Y. Cheng, X. Zheng, Y. Yang, Cooperative behavior evolution of small groups on interconnected networks, Chaos Solitons Fractals 80 (2015) 90–95.
[431] A. Perisic, C.T. Bauch, Social contact networks and disease eradicability under voluntary vaccination, PLoS Comput. Biol. 5 (2) (2009) e1000280.
[432] F. Fu, D.I. Rosenbloom, L. Wang, M.A. Nowak, Imitation dynamics of vaccination behaviour on social networks, Proc. R. Soc. B 278 (1702) (2011) 42–49.
[433] M.L.N. Mbah, J. Liu, C.T. f, Y.I. Tekel, J. Medlock, L.A. Meyers, A.P. Galvani, The impact of imitation on vaccination behavior in social contact networks, PLoS Comput. Biol. 8 (4) (2012) e1002469.
[434] C.R. Wells, E.Y. Klein, C.T. Bauch, Policy resistance undermines superspreader vaccination strategies for influenza, PLoS Comput. Biol. 9 (3) (2013) e1002945.
[435] C.T. Bauch, Imitation dynamics predict vaccinating behaviour, Proc. R. Soc. B 272 (1573) (2005) 1669–1675.
[436] T.C. Reluga, C.T. Bauch, A.P. Galvani, Evolving public perceptions and stability in vaccine uptake, Math. Biosci. 204 (2) (2006) 185–198.
[437] C. Innes, M. Anand, C.T. Bauch, The impact of human–environment interactions on the stability of forest-grassland mosaic ecosystems, Sci. Rep. 3 (2013) 2689.

[438] L.-A. Barlow, J. Cecile, C.T. Bauch, M. Anand, Modelling interactions between forest pest invasions and human decisions regarding firewood transport restrictions, PLoS One 9 (4) (2014) e90511.

[439] J.R. Busemeyer, J.T. Townsend, Decision field theory: a dynamic-cognitive approach to decision making in an uncertain environment, Psychol. Rev. 100 (3) (1993) 432.

[440] M.A. Andrews, C.T. Bauch, Disease interventions can interfere with one another through disease-behaviour interactions, PLoS Comput. Biol. 11 (6) (2015) e1004291.

[441] T. Oraby, C.T. Bauch, Bounded rationality alters the dynamics of paediatric immunization acceptance, Sci. Rep. 5 (2015) 10724.

[442] G. Shafer, et al., A Mathematical Theory of Evidence, vol. 1, Princeton University Press, 1976.

[443] S. Xia, J. Liu, A belief-based model for characterizing the spread of awareness and its impacts on individuals vaccination decisions, J. R. Soc. Interface 11 (94) (2014) 20140013.

[444] S. Funk, M. Salathé, V.A. Jansen, Modelling the influence of human behaviour on the spread of infectious diseases: a review, J. R. Soc. Interface 7 (50) (2010) 1247–1256.

[445] P. Manfredi, A. d'Onofrio (Eds.), Modeling the Interplay between Human Behavior and the Spread of Infectious Diseases, Springer Science & Business Media, 2013.

[446] C.T. Bauch, A.P. Galvani, Epidemiology. social factors in epidemiology, Science (New York, NY) 342 (6154) (2013) 47–49.

[447] A. d'Onofrio, P. Manfredi, E. Salinelli, Vaccinating behaviour, information, and the dynamics of sir vaccine preventable diseases, Theoret. Popul. Biol. 71 (3) (2007) 301–317.

[448] A. d'Onofrio, P. Manfredi, Bifurcation thresholds in an sir model with information-dependent vaccination, Math. Model. Nat. Phenom. 2 (01) (2007) 26–43.

[449] A. d'Onofrio, P. Manfredi, E. Salinelli, Fatal SIR diseases and rational exemption to vaccination, Math. Med. Biol. 25 (4) (2008) 337–357.

[450] A. d'Onofrio, P. Manfredi, E. Salinelli, Vaccinating behaviour and the dynamics of vaccine preventable infections, in: A. d'Onofrio, P. Manfredi (Eds.), Modeling the Interplay Between Human Behavior and the Spread of Infectious Diseases, Springer, 2013, pp. 267–287.

[451] N. MacDonald, Biological Delay Systems: Linear Stability Theory, Cambridge University Press, 1989.

[452] T.C. Reluga, C.T. Bauch, A.P. Galvani, Evolving public perceptions and stability in vaccine uptake, Math. Biosci. 204 (2) (2006) 185–198.

[453] P.-Y. Geoffard, T. Philipson, Disease eradication: private versus public vaccination, Amer. Econ. Rev. 87 (1) (1997) 222–230.

[454] D.V. Efimov, A.L. Fradkov, Yakubovich's oscillatority of circadian oscillations models, Math. Biosci. 216 (2) (2008) 187–191.

[455] B. Buonomo, A. d'Onofrio, D. Lacitignola, Global stability of an SIR epidemic model with information dependent vaccination, Math. Biosci. 216 (1) (2008) 9–16.

[456] B. Buonomo, A. d'Onofrio, D. Lacitignola, Modeling of pseudo-rational exemption to vaccination for SEIR diseases, J. Math. Anal. Appl. 404 (2) (2013) 385–398.

[457] S. Bhattacharyya, C. Bauch, A game dynamic model for delayer strategies in vaccinating behaviour for pediatric infectious diseases, J. Theoret. Biol. 267 (3) (2010) 276–282.

[458] F. Xu, R. Cressman, Disease control through voluntary vaccination decisions based on the smoothed best response, Comput. Math. Methods Med. 2014 (2014) 825734.

[459] M. Nowak, Evolutionary Dynamics: Exploring the Equations of Life, Belknap Press of Hardvard University Press, Cambridge, Massachusets, 2006.

[460] M. Nowak, Five rules for the evolution of cooperation, Science 8 (5805) (2006) 1560–1563.

[461] J. Hofbauer, K. Sigmund, Evolutionary Games and Population Dynamics, Cambridge university press, 1998.

[462] C.T. Bauch, Imitation dynamics predict vaccinating behaviour, Proc. R. Soc. Lond. Biol. 272 (1573) (2005) 1669–1675.

[463] A. d'Onofrio, P. Manfredi, P. Poletti, The impact of vaccine side effects on the natural history of immunization programmes: an imitation-game approach, J. Theoret. Biol. 273 (1) (2011) 63–71.

[464] C.T. Bauch, S. Bhattacharyya, Evolutionary game theory and social learning can determine how vaccine scares unfold, PLoS Comput. Biol. 8 (4) (2012) e1002452.

[465] T. Oraby, C.T. Bauch, Bounded rationality alters the dynamics of paediatric immunization acceptance, Sci. Rep. 5 (2015) 10724.

[466] Center for disease control, vaccine safety, 2016. http://www.cdc.gov/vaccinesafety/ensuringsafety/monitoring/vaers/. (Accessed 3 June 2016).

[467] I. Ajzen, The theory of planned behavior, Organ. Behav Hum. Decis. Process. 50 (2) (1991) 179–211.

[468] R.B. Cialdini, M.R. Trost, Social Influence: Social Norms, Conformity and Compliance, McGraw-Hill, 1998.

[469] D. Helbing, A. Johansson, Cooperation, norms, and revolutions: a unified game-theoretical approach, PLoS One 5 (10) (2010) e12530.

[470] D. Kahneman, A. Tversky, Prospect theory: An analysis of decision under risk, Econometrica (1979) 263–291.

[471] F. Bass, A new product growth for model consumer durables, Manag. Sci. 15 (1969) 215–227.

[472] C.T. Bauch, A.P. Galvani, D.J. Earn, Group interest versus self-interest in smallpox vaccination policy, Proc. Natl. Acad. Sci. 100 (18) (2003) 10564–10567.

[473] C.T. Bauch, D.J. Earn, Vaccination and the theory of games, Proc. Natl. Acad. Sci. USA 101 (36) (2004) 13391–13394.

[474] T.C. Reluga, A.P. Galvani, A general approach for population games with application to vaccination, Math. Biosci. 230 (2) (2011) 67–78.

[475] E. Shim, J.J. Grefenstette, S.M. Albert, B.E. Cakouros, D.S. Burke, A game dynamic model for vaccine skeptics and vaccine believers: measles as an example, J. Theoret. Biol. 295 (2012) 194–203.

[476] J.-P. Aubin, H. Frankowska, Set-valued Analysis, Springer Science & Business Media, 2009.

[477] T.C. Reluga, An sis epidemiology game with two subpopulations, J. Biol. Dyn. 3 (5) (2009) 515–531.

[478] P. Manfredi, P. Della Posta, A. d'Onofrio, E. Salinelli, F. Centrone, C. Meo, P. Poletti, Optimal vaccination choice, vaccination games, and rational exemption: an appraisal, Vaccine 28 (1) (2009) 98–109.

[479] P.E. Fine, J.A. Clarkson, Individual versus public priorities in the determination of optimal vaccination policies, Amer. J. Epidemiol. 124 (6) (1986) 1012–1020.

[480] P.J. Francis, Dynamic epidemiology and the market for vaccinations, J. Publ. Econ. 63 (3) (1997) 383–406.

[481] G. Scalia-Tomba, The effect of structural behavior change on the spread of hiv in a one-sex population, Math. Biosci. 107 (2) (1991) 547–555.

[482] J. Li, Effects of behavior change on the spread of AIDS epidemic, Math. Comput. Modelling 16 (6–7) (1992) 103–111.

[483] H. Stigum, P. Magnus, L.S. Bakketeig, Effect of changing partnership formation rates on the spread of sexually transmitted diseases and human immunodeficiency virus, Amer. J. Epidemiol. 145 (7) (1997) 644–652.

[484] J.X. Velasco-Hernandez, Y.-H. Hsieh, Modelling the effect of treatment and behavioral change in hiv transmission dynamics, J. Math. Biol. 32 (3) (1994) 233–249.

[485] Y.-H. Hsieh, A two-sex model for treatment of HIV/AIDS and behaviour change in a population of varying size, Math. Med. Biol. 13 (3) (1996) 151–173.

[486] J.X. Velasco-Hernández, F. Brauer, C. Castillo-Chavez, Effects of treatment and prevalence-dependent recruitment on the dynamics of a fatal disease, Math. Med. Biol. 13 (3) (1996) 175–192.

[487] M. Kremer, Integrating behavioral choice into epidemiological models of AIDS, Q. J. Econ. 111 (2) (1996) 549–573.

[488] M. Kremer, C. Morcom, The effect of changing sexual activity on HIV prevalence, Math. Biosci. 151 (1) (1998) 99–122.

[489] E. Shim, L. Meyers, A.P. Galvani, Optimal H1N1 vaccination strategies based on self-interest versus group interest, BMC Public Health 11 (1) (2011) 1.

[490] R. Breban, R. Vardavas, S. Blower, Mean-field analysis of an inductive reasoning game: application to influenza vaccination, Phys. Rev. E 76 (3) (2007) 031127.

[491] R. Vardavas, R. Breban, S. Blower, Can influenza epidemics be prevented by voluntary vaccination? PLoS Comput. Biol. 3 (5) (2007) e85.

[492] E. Shim, B. Kochin, A. Galvani, Insights from epidemiological game theory into gender-specific vaccination against rubella, Math. Biosci. Eng. 6 (4) (2009) 839–854.

[493] M. Gersovitz, J.S. Hammer, The economical control of infectious diseases, Econ. J. 114 (492) (2004) 1–27.

[494] F.H. Chen, A susceptible-infected epidemic model with voluntary vaccinations, J. Math. Biol. 53 (2) (2006) 253–272.
[495] F. Chen, F. Toxvaerd, The economics of vaccination, J. Theoret. Biol. 363 (2014) 105–117.
[496] M.-G. Cojocaru, C.T. Bauch, M.D. Johnston, Dynamics of vaccination strategies via projected dynamical systems, Bull. Math. Biol. 69 (5) (2007) 1453–1476.
[497] B.M. Althouse, T.C. Bergstrom, C.T. Bergstrom, A public choice framework for controlling transmissible and evolving diseases, Proc. Natl. Acad. Sci. 107 (Suppl. 1) (2010) 1696–1701.
[498] M. Voinson, S. Billiard, A. Alvergne, Beyond rational decision-making: Modelling the influence of cognitive biases on the dynamics of vaccination coverage, PLoS One 10 (11) (2015) e0142990.
[499] N. Crepaz, T.A. Hart, G. Marks, Highly active antiretroviral therapy and sexual risk behavior: a meta-analytic review, JAMA 292 (2) (2004) 224–236.
[500] J. Gómez-Gardeñes, V. Latora, Y. Moreno, E. Profumo, Spreading of sexually transmitted diseases in heterosexual populations, Proc. Natl. Acad. Sci. 105 (5) (2008) 1399–1404.
[501] C. Bauch, D. Rand, A moment closure model for sexually transmitted disease transmission through a concurrent partnership network, Proc. R. Soc. Lond. Biol. 267 (1456) (2000) 2019–2027.
[502] S. Del Valle, H. Hethcote, J. Hyman, C. Castillo-Chavez, Effects of behavioral changes in a smallpox attack model, Math. Biosci. 195 (2) (2005) 228–251.
[503] M. Kretzschmar, S. Van den Hof, J. Wallinga, J. Van Wijngaarden, Ring vaccination and smallpox control, Emerg. Infect. Dis. 10 (5) (2004) 832–841.
[504] M.E. Newman, D.J. Watts, Scaling and percolation in the small-world network model, Phys. Rev. E 60 (6) (1999) 7332.
[505] P. Erdös, A. Rényi, On the evolution of random graphs, Publ. Math. Inst. Hungar. Acad. Sci. 5 (1960) 17–61.
[506] A.-L. Barabási, R. Albert, Emergence of scaling in random networks, Science 286 (5439) (1999) 509–512.
[507] D.G. Rand, A. Dreber, T. Ellingsen, D. Fudenberg, M.A. Nowak, Positive interactions promote public cooperation, Science 325 (5945) (2009) 1272–1275.
[508] Z. Wang, S. Kokubo, M. Jusup, J. Tanimoto, Universal scaling for the dilemma strength in evolutionary games, Phys. Life Rev. 14 (2015) 1–30.
[509] M. Perc, J. Gómez-Gardeñes, A. Szolnoki, L.M. Floría, Y. Moreno, Evolutionary dynamics of group interactions on structured populations: a review, J. R. Soc. Interface 10 (80) (2013) 20120997.
[510] G. Hardin, The tragedy of the commons, Science 162 (3859) (1968) 1243–1248.
[511] M. Pingle, Imitation versus rationality: An experimental perspective on decision making, J. Socio-Econ. 24 (2) (1995) 281–315.
[512] L.E. Blume, The statistical mechanics of strategic interaction, Games Econom. Behav. 5 (3) (1993) 387–424.
[513] M. Perc, J. Gómez-Gardeñes, A. Szolnoki, L.M. Floría, Y. Moreno, Evolutionary dynamics of group interactions on structured populations: A review, J. R. Soc. Interface 6 (2013) 20120997.
[514] J. Colgrove, State of Immunity: the Politics of Vaccination in Twentieth-Century America, vol. 16, Univ of California Press, 2006.
[515] Z.-X. Wu, H.-F. Zhang, Peer pressure is a double-edged sword in vaccination dynamics, Europhys. Lett. 104 (1) (2013) 10002.
[516] X.-T. Liu, Z.-X. Wu, L. Zhang, Impact of committed individuals on vaccination behavior, Phys. Rev. E 86 (5) (2012) 051132.
[517] D.J. Watts, P.S. Dodds, Influentials, networks, and public opinion formation, J. Consumer Res. 34 (4) (2007) 441–458.
[518] A. Szolnoki, Z. Wang, M. Perc, Wisdom of groups promotes cooperation in evolutionary social dilemmas, Sci. Rep. 2 (2012) 576.
[519] H. Zhang, F. Fu, W. Zhang, B. Wang, Rational behavior is a double-edged sword when considering voluntary vaccination, Physica A 391 (20) (2012) 4807–4815.
[520] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, D.-U. Hwang, Complex networks: structure and dynamics, Phys. Rep. 424 (4–5) (2006) 175–308. http://dx.doi.org/10.1016/j.physrep.2005.10.009.
[521] R. Pastor-Satorras, A. Vespignani, Epidemic spreading in scale-free networks, Phys. Rev. Lett. 86 (14) (2001) 3200–3203.
[522] D. Li, P. Qin, H. Wang, C.L. Liu, Y. Jiang, Epidemics on interconnected lattices, Europhys. Lett. 105 (6) (2014) 68004.
[523] C.-y. Xia, Z. Wang, J. Sanz, S. Meloni, Y. Moreno, Effects of delayed recovery and nonuniform transmission on the spreading of diseases in complex networks, Physica A 392 (7) (2013) 1577–1585.
[524] H. Zhang, J. Zhang, C. Zhou, M. Small, B. Wang, Hub nodes inhibit the outbreak of epidemic under voluntary vaccination, New J. Phys. 12 (2) (2010) 023015.
[525] A. Cardillo, C. Reyes-Suárez, F. Naranjo, J. Gómez-Gardeñes, Evolutionary vaccination dilemma in complex networks, Phys. Rev. E 88 (3) (2013) 032803.
[526] D.T. Jamison, External finance of immunization programs: Time for a change in paradigm? Sci. Tech. Publ. Pan Am. Health Organ. (2004) 325–332.
[527] A. d'Onofrio, P. Manfredi, P. Poletti, The interplay of public intervention and private choices in determining the outcome of vaccination programmes, PLoS One 7 (10) (2012) e45653.
[528] H.-F. Zhang, Z.-X. Wu, X.-K. Xu, M. Small, L. Wang, B.-H. Wang, Impacts of subsidy policies on vaccination decisions in contact networks, Phys. Rev. E 88 (1) (2013) 012813.
[529] P.-Y. Geoffard, T. Philipson, Disease eradication: private versus public vaccination, Am. Econ. Rev. 87 (1) (1997) 222–230.
[530] A.J. Culyer, J.P. Newhouse, Handbook of Health Economics, Elsevier, 2000.
[531] M. Gersovitz, J.S. Hammer, Tax/subsidy policies toward vector-borne infectious diseases, J. Public Econ. 89 (4) (2005) 647–674.
[532] M. Gersovitz, J.S. Hammer, Infectious diseases, public policy, and the marriage of economics and epidemiology, World Bank Res. Observer 18 (2) (2003) 129–157.
[533] H.-F. Zhang, Z.-X. Wu, M. Tang, Y.-C. Lai, Effects of behavioral response and vaccination policy on epidemic spreading-an approach based on evolutionary-game dynamics, Sci. Rep. 4 (2014) 5666.
[534] Z. Wang, A. Szolnoki, M. Perc, If players are sparse social dilemmas are too: Importance of percolation for evolution of cooperation, Sci. Rep. 2 (2012) 369.
[535] Y. Gao, W. Du, G. Yan, Selectively-informed particle swarm optimization, Sci. Rep. 5 (2015) 9295.
[536] L. Li, H. Peng, J. Kurths, Y. Yang, H.J. Schellnhuber, Chaos-order transition in foraging behavior of ants, Proc. Natl. Acad. Sci. 111 (23) (2014) 8392–8397.
[537] W. Du, W. Ying, G. Yan, Y. Zhu, X. Cao, Heterogeneous strategy particle swarm optimization, IEEE Trans. Circuits Syst. II (2016). http://dx.doi.org/10.1109/TCSII.2016.2595597.
[538] W. Du, X. Cao, M. Hu, W. Wang, Asymmetric cost in snowdrift game on scale-free networks, Europhys. Lett. 87 (6) (2009) 60004.
[539] M. Mobilia, Does a single zealot affect an infinite group of voters? Phys. Rev. Lett. 91 (2) (2003) 028701.
[540] S. Galam, Heterogeneous beliefs, segregation, and extremism in the making of public opinions, Phys. Rev. E 71 (4) (2005) 046123.
[541] J. Xie, S. Sreenivasan, G. Korniss, W. Zhang, C. Lim, B.K. Szymanski, Social consensus through the influence of committed minorities, Phys. Rev. E 84 (1) (2011) 011130.
[542] P. Singh, S. Sreenivasan, B.K. Szymanski, G. Korniss, Accelerating consensus on coevolving networks: the effect of committed individuals, Phys. Rev. E 85 (4) (2012) 046104.
[543] N. Masuda, Evolution of cooperation driven by zealots, Sci. Rep. 2 (2012) 646.
[544] X.-T. Liu, Z.-X. Wu, L. Zhang, Impact of committed individuals on vaccination behavior, Phys. Rev. E 86 (5) (2012) 051132.
[545] E. Fukuda, J. Tanimoto, Effects of stubborn decision-makers on vaccination and disease propagation in social networks, Int. J. Autom. Logist. 2 (1–2) (2016) 78–92.
[546] H.-F. Zhang, Z. Yang, Z.-X. Wu, B.-H. Wang, T. Zhou, Braess's paradox in epidemic game: Better condition results in less payoff, Sci. Rep. 3 (2013) 3292.
[547] P. Schimit, L. Monteiro, A vaccination game based on public health actions and personal decisions, Ecol. Model. 222 (9) (2011) 1651–1655.
[548] A.A. Parker, W. Staggs, G.H. Dayan, I.R. Ortega-Sánchez, P.A. Rota, L. Lowe, P. Boardman, R. Teclaw, C. Graves, C.W. LeBaron, Implications of a 2005 measles outbreak in Indiana for sustained elimination of measles in the United States, N. Engl. J. Med. 355 (5) (2006) 447–455.
[549] T. May, R.D. Silverman, Clustering of exemptions as a collective action threat to herd immunity, Vaccine 21 (11) (2003) 1048–1051.
[550] P.-D.D.D. Braess, Über ein Paradoxon aus der Verkehrsplanung, Unternehmensforschung 12 (1) (1968) 258–268.
[551] M. Pala, S. Baltazar, P. Liu, H. Sellier, B. Hackens, F. Martins, V. Bayot, X. Wallart, L. Desplanque, S. Huant, Transport inefficiency in branched-out mesoscopic networks: An analog of the Braess paradox, Phys. Rev. Lett. 108 (7) (2012) 076802.

[552] Z. Wang, H. Zhang, Z. Wang, Multiple effects of self-protection on the spreading of epidemics, Chaos Solitons Fractals 61 (2014) 1–7.
[553] M.E.J. Newman, Networks: An Introduction, Oxford University Press, New York, 2010.
[554] M.E.J. Newman, Detecting community structure in networks, Eur. Phys. J. B 38 (2) (2004) 321–330.
[555] S. Fortunato, Community detection in graphs, Phys. Rep. 486 (2010) 75–174. http://dx.doi.org/10.1016/j.physrep.2009.11.002.
[556] A. Perisic, C.T. Bauch, A simulation analysis to characterize the dynamics of vaccinating behaviour on contact networks, BMC Infect. Dis. 9 (1) (2009) 1.
[557] P.E. Fine, J.A. Clarkson, Individual versus public priorities in the determination of optimal vaccination policies, Amer. J. Epidemiol. 124 (6) (1986) 1012–1020.
[558] C.T. Bauch, A.P. Galvani, D.J. Earn, Group interest versus self-interest in smallpox vaccination policy, Proc. Natl. Acad. Sci. 100 (18) (2003) 10564–10567.
[559] S. Barrett, The smallpox eradication game, Public Choice 130 (1–2) (2007) 179–207.
[560] L.B. Brilliant, The Management of Smallpox Eradication in India, University of Michigan Press Ann Arbor, MI, 1985.
[561] J. Gómez-Gardeñes, S. Gómez, A. Arenas, Y. Moreno, Explosive synchronization transitions in scale-free networks, Phys. Rev. Lett. 106 (12) (2011) 128701.
[562] D. Achlioptas, R.M. D'Souza, J. Spencer, Explosive percolation in random networks, Science 323 (5920) (2009) 1453–1455.
[563] J.T. Kemper, On the identification of superspreaders for infectious disease, Math. Biosci. 48 (1) (1980) 111–127.
[564] M.E. Woolhouse, C. Dye, J.-F. Etard, T. Smith, J. Charlwood, G. Garnett, P. Hagan, J. Hii, P. Ndhlovu, R. Quinnell, et al., Heterogeneities in the transmission of infectious agents: implications for the design of control programs, Proc. Natl. Acad. Sci. 94 (1) (1997) 338–342.
[565] A.P. Galvani, R.M. May, Epidemiology: dimensions of superspreading, Nature 438 (7066) (2005) 293–295.
[566] J.O. Lloyd-Smith, S.J. Schreiber, P.E. Kopp, W.M. Getz, Superspreading and the effect of individual variation on disease emergence, Nature 438 (7066) (2005) 355–359.
[567] R.A. Stein, Lessons from outbreaks of H1N1 influenza, Ann. Intern. Med. 151 (1) (2009) 59–62.
[568] R. Cohen, S. Havlin, D. Ben-Avraham, Efficient immunization strategies for computer networks and populations, Phys. Rev. Lett. 91 (2003) 247901.
[569] L. Wang, X. Li, Spatial epidemiology of networked metapopulation: An overview, Chin. Sci. Bull. 59 (28) (2014) 3511–3522.
[570] A. Szolnoki, G. Szabó, Cooperation enhanced by inhomogeneous activity of teaching for evolutionary Prisoner's Dilemma games, Europhys. Lett. 77 (3) (2007) 30004.
[571] S. Kokubo, Z. Wang, J. Tanimoto, Spatial reciprocity for discrete, continuous and mixed strategy setups, Appl. Math. Comput. 259 (2015) 552–568.
[572] S. Xia, J. Liu, A computational approach to characterizing the impact of social influence on individuals vaccination decision making, PLoS One 8 (4) (2013) e60373.
[573] M. Salathé, S. Bonhoeffer, The effect of opinion clustering on disease outbreaks, J. R. Soc. Interface 5 (29) (2008) 1505–1508.
[574] D. Schmid, H. Holzmann, S. Abele, S. Kasper, S. König, S. Meusburger, H. Hrabcik, A. Luckner-Hornischer, E. Bechter, A. DeMartin, et al., An ongoing multi-state outbreak of measles linked to non-immune anthroposophic communities in Austria, Germany, and Norway, March-April 2008, Eurosurveillance 13 (16).
[575] J. Richard, V. Masserey-Spicher, S. Santibanez, A. Mankertz, Measles outbreak in Switzerland-an update relevant for the European football championship (EURO 2008), 13 (1–3) (2008) 1–4.
[576] T. Zhang, X. Fu, S. Ma, G. Xiao, L. Wong, C. Kwoh, M. Lees, G. Lee, T. Hung, Evaluating temporal factors in combined interventions of workforce shift and school closure for mitigating the spread of influenza, PLoS One 7 (3) (2012) e32203.
[577] R. Anderson, R. May, B. Anderson, Infectious Diseases of Humans: Dynamics and Control, Oxford University Press, Oxford, 1992.
[578] W.A. Orenstein, K.L. Samuel, A.R. Hinman, Summary and conclusions: measles elimination meeting, 16–17 march 2000, J. Infect. Dis. 189 (Suppl. 1) (2004) S43–S47.
[579] M. Salathé, S. Khandelwal, Assessing vaccination sentiments with online social media: implications for infectious disease dynamics and control, PLoS Comput. Biol. 7 (10) (2011) e1002199.
[580] S.-J. Wang, Z. Wang, T. Jin, S. Boccaletti, Emergence of disassortative mixing from pruning nodes in growing scale-free networks, Sci. Rep. 4 (2014) 7536.
[581] J. Qu, S.-J. Wang, M. Jusup, Z. Wang, Effects of random rewiring on the degree correlation of scale-free networks, Sci. Rep. 5 (2015) 15450.
[582] V.C. Barclay, T. Smieszek, J. He, G. Cao, J.J. Rainey, H. Gao, A. Uzicanin, M. Salathé, Positive network assortativity of influenza vaccination at a high school: implications for outbreak risk and herd immunity, PLoS One 9 (2) (2014) e87042.
[583] Q. Wu, X. Fu, M. Small, X.-J. Xu, The impact of awareness on epidemic spreading in networks, Chaos 22 (1) (2012) 013101.
[584] L. Valdez, P.A. Macri, L.A. Braunstein, Intermittent social distancing strategy for epidemic control, Phys. Rev. E 85 (3) (2012) 036108.
[585] H.-F. Zhang, J.-R. Xie, M. Tang, Y.-C. Lai, Suppression of epidemic spreading in complex networks by local information based behavioral responses, Chaos 24 (4) (2014) 043106.
[586] A. Nowak, J. Szamrej, B. Latané, From private attitude to public opinion: A dynamic theory of social impact, Psychol. Rev. 97 (3) (1990) 362.
[587] J. Lyst, K. Kacperski, F. Schweitzer, Social impact models of opinion dynamics, Annu. Rev. Comput. Phys. 9 (2002) 253–273.
[588] J.K. Kelso, G.J. Milne, H. Kelly, Simulation suggests that rapid activation of social distancing can arrest epidemic development due to a novel strain of influenza, BMC Public Health 9 (1) (2009) 117.
[589] T.C. Reluga, Game theory of social distancing in response to an epidemic, PLoS Comput. Biol. 6 (5) (2010) e1000793.
[590] M. Karsai, N. Perra, A. Vespignani, Time varying networks and the weakness of strong ties, Sci. Rep. 4 (2014) 4001.
[591] P. Holme, J. Saramäki, Temporal networks, Phys. Rep. 519 (3) (2012) 97–125.
[592] S. Liu, N. Perra, M. Karsai, A. Vespignani, Controlling contagion processes in activity driven networks, Phys. Rev. Lett. 112 (11) (2014) 118702.
[593] M. Starnini, A. Machens, C. Cattuto, A. Barrat, R. Pastor-Satorras, Immunization strategies for epidemic processes in time-varying contact networks, J. Theoret. Biol. 337 (2013) 89–100.
[594] B. Morsky, C.T. Bauch, Outcome inelasticity and outcome variability in behaviour-incidence models: an example from an seir infection on a dynamic network, Comput. Math. Methods Med. 2012 (2012) 652562.
[595] N. Perra, B. Gonçalves, R. Pastor-Satorras, A. Vespignani, Activity driven modeling of time varying networks, Sci. Rep. 2 (2012) 469.
[596] D. Han, M. Sun, An evolutionary vaccination game in the modified activity driven network by considering the closeness, Physica A 443 (2016) 49–57.
[597] J. Paulsen, D. Langbehn, J. Stout, E. Aylward, C. Ross, M. Nance, M. Guttman, S. Johnson, M. MacDonald, L. Beglinger, et al., Detection of huntingtons disease decades before diagnosis: the predict-hd study, J. Neurol. Neurosurg. Psychiatry 79 (8) (2008) 874–880.
[598] B. Guerra, J. Gomez-Gardenes, V. Latora, Adaptive growing networks coevolving with the spread of diseases, Int. J. Bifurcation Chaos 22 (07) (2012) 1250168.
[599] D. Brockmann, D. Helbing, The hidden geometry of complex, network-driven contagion phenomena, Science 342 (6164) (2013) 1337–1342.
[600] D. Balcan, V. Colizza, B. Gonçalves, H. Hu, J.J. Ramasco, A. Vespignani, Multiscale mobility networks and the spatial spreading of infectious diseases, Proc. Natl. Acad. Sci. 106 (51) (2009) 21484–21489.
[601] T. Poggio, F. Girosi, Regularization algorithms for learning that are equivalent to multilayer networks, Science 247 (4945) (1990) 978–982.
[602] K. Hornik, M. Stinchcombe, H. White, Multilayer feedforward networks are universal approximators, Neural Net. 2 (5) (1989) 359–366.
[603] Y. Wang, G. Xiao, Epidemics spreading in interconnected complex networks, Phys. Lett. A 376 (42) (2012) 2689–2696.
[604] A. Halu, K. Zhao, A. Baronchelli, G. Bianconi, Connect and win: The role of social networks in political elections, Europhys. Lett. 102 (1) (2013) 16002.
[605] X. Zhang, S. Boccaletti, S. Guan, Z. Liu, Explosive synchronization in adaptive and multilayer networks, Phys. Rev. Lett. 114 (3) (2015) 038701.
[606] Z. Wang, L. Wang, A. Szolnoki, M. Perc, Evolutionary games on multilayer networks: a colloquium, Eur. Phys. J. B 88 (5) (2015) 1–15.
[607] Z. Wang, L. Wang, M. Perc, Degree mixing in multilayer networks impedes the evolution of cooperation, Phys. Rev. E 89 (5) (2014) 052813.
[608] C.T. Bauch, A.P. Galvani, Social factors in epidemiology, Science 342 (6154) (2013) 47–49.

[609] W. Wang, M. Tang, P. Shu, Z. Wang, Dynamics of social contagions with heterogeneous adoption thresholds:crossover phenomena in phase transition, New J. Phys. 18 (2016) 013029.

[610] W. Wang, M. Tang, H.-F. Zhang, Y.-C. Lai, Dynamics of social contagions with memory of nonredundant information, Phys. Rev. E 92 (2015) 012820.

[611] E. Fukuda, J. Tanimoto, M. Akimoto, Influence of breaking the symmetry between disease transmission and information propagation networks on stepwise decisions concerning vaccination, Chaos Solitons Fractals 80 (2015) 47–55.

[612] K.T. Eames, Networks of influence and infection: parental choices and childhood disease, J. R. Soc. Interface 6 (38) (2009) 811–814.

[613] J. Leask, S. Chapman, P. Hawe, M. Burgess, What maintains parental support for vaccination when challenged by anti-vaccination messages? a qualitative study, Vaccine 24 (49) (2006) 7238–7245.

[614] K. Abbasi, MMR and the value of word of mouth in social networks, J. R. Soc. Med. 101 (5) (2008) 215–216.

[615] R. Heathcock, C. Watts, Measles outbreaks in London, United Kingdom-a preliminary report, Euro Surveill 13 (15) (2008) 1.

[616] F.T. Cutts, H.S. Izurieta, D.A. Rhoda, Measuring coverage in MNCH: design, implementation, and interpretation challenges associated with tracking vaccination coverage using household surveys, PLoS Med. 10 (5) (2013) e1001404.

[617] M.K. Slifka, I. Amanna, How advances in immunology provide insight into improving vaccine efficacy, Vaccine 32 (25) (2014) 2948–2957.

[618] K.S. Wagner, J.M. White, N.J. Andrews, R. Borrow, E. Stanford, E. Newton, R.G. Pebody, Immunity to tetanus and diphtheria in the UK in 2009, Vaccine 30 (49) (2012) 7111–7117.

[619] F. Cutts, A. Hall, Vaccines for neonatal viral infections: hepatitis B vaccine, Expert Rev. Vaccines 3 (4) (2004) 349–352.

[620] M.D. Tapia, M.F. Pasetti, L. Cuberos, S.O. Sow, M.N. Doumbia, M. Bagayogo, K.L. Kotloff, M.M. Levine, Measurement of tetanus antitoxin in oral fluid: a tool to conduct serosurveys, Pediatr. Infect. Dis. J. 25 (9) (2006) 819–825.

[621] R. Pebody, N. Gay, A. Giammanco, S. Baron, J. Schellekens, A. Tischer, R.-M. Ölander, N. Andrews, W. Edmunds, H. Lecoeur, et al., The seroepidemiology of bordetella pertussis infection in Western Europe, Epidemiol. Infect. 133 (01) (2005) 159–171.

[622] G. Saemann-Ischenko, B. Tillmanns, K. Kösters, M. Riffelmann, W.C. Von Koenig, Stability of antibodies to Bordetella antigens in German adults, Eur. J. Clin. Microbiol. Infect. Dis. 20 (12) (2001) 850–853.

[623] L.A. Cattaneo, G.W. Reed, D.H. Haase, M.J. Wills, K.M. Edwards, The seroepidemiology of bordetella pertussis infections: a study of persons ages 1–65 years, J. Infect. Dis. 173 (5) (1996) 1256–1259.

[624] P. Teunis, O. Van Der Heijden, H. De Melker, J. Schellekens, F. Versteegh, M. Kretzschmar, Kinetics of the igg antibody response to pertussis toxin after infection with B. pertussis, Epidemiol. Infect. 129 (03) (2002) 479–489.

[625] J.S. Lavine, A.A. King, O.N. Bjørnstad, Natural immune boosting in pertussis dynamics and the potential for long-term vaccine failure, Proc. Natl. Acad. Sci. 108 (17) (2011) 7259–7264.

[626] G.A. Weinberg, P.G. Szilagyi, Vaccine epidemiology: efficacy, effectiveness, and the translational research roadmap, J. Infect. Dis. 201 (11) (2010) 1607–1610.

[627] M.T. Osterholm, N.S. Kelley, A. Sommer, E.A. Belongia, Efficacy and effectiveness of influenza vaccines: a systematic review and meta-analysis, Lancet Infect. Dis. 12 (1) (2012) 36–44.

[628] G. La Torre, C. de Waure, G. Chiaradia, A. Mannocci, W. Ricciardi, HPV vaccine efficacy in preventing persistent cervical HPV infection: a systematic review and meta-analysis, Vaccine 25 (50) (2007) 8352–8358.

[629] J.I. Ward, J.D. Cherry, S.-J. Chang, S. Partridge, H. Lee, J. Treanor, D.P. Greenberg, W. Keitel, S. Barenkamp, D.I. Bernstein, et al., Efficacy of an acellular pertussis vaccine among adolescents and adults, N. Engl. J. Med. 353 (15) (2005) 1555–1563.

[630] M. van Boven, M. Kretzschmar, J. Wallinga, P.D. O'Neill, O. Wichmann, S. Hahné, Estimation of measles vaccine efficacy and critical vaccination coverage in a highly vaccinated population, J. R. Soc. Interface 7 (52) (2010) 1537–1544.

[631] W.A. Orenstein, R.H. Bernier, T.J. Dondero, A.R. Hinman, J. Marks, K. Bart, B. Sirotkin, Field evaluation of vaccine efficacy, Bull. World. Health. Organ. 63 (6) (1985) 1055.

[632] C.L. Althaus, M. Salathé, Measles vaccination coverage and cases among vaccinated persons, Emerg. Infect. Diseases 21 (8) (2015) 1480.

[633] Overview of influenza surveillance in the united states, http://www.cdc.gov/flu/weekly/overview.htm.

[634] R. Chunara, M.S. Smolinski, J.S. Brownstein, Why we need crowdsourced data in infectious disease surveillance, Current Infect. Dis. Rep. 15 (4) (2013) 316–319.

[635] A. Vespignani, et al., Predicting the behavior of techno-social systems, Science 325 (5939) (2009) 425.

[636] B. Gonçalves, N. Perra, Social Phenomena: From Data Analysis to Models, Springer, 2015.

[637] D. Lazer, A.S. Pentland, L. Adamic, S. Aral, A.L. Barabasi, D. Brewer, N. Christakis, N. Contractor, J. Fowler, M. Gutmann, et al., Computational social science, Science (New York, NY) 323 (5915) (2009) 721.

[638] L. Piwek, D.A. Ellis, S. Andrews, A. Joinson, The rise of consumer health wearables: Promises and barriers, PLoS Med. 13 (2) (2016) http://dx.doi.org/10.1371/journal.pmed.1001953, e1001953.

[639] M. Salathé, L. Bengtsson, J. Bodnar, D. Brewer, J.S. Brownstein, C. Buckee, E.M. Campbell, C. Cattuto, S. Khandelwal, P.L. Mabry, A. Vespignani, Digital epidemiology, PLoS Comput. Biol. 8 (2012) 7.

[640] R. Chunara, C.C. Freifeld, J.S. Brownstein, New technologies for reporting real-time emergent infections, Parasitol. 139 (14) (2012) 1843–1851.

[641] D.M. Hartley, N.P. Nelson, R. Arthur, P. Barboza, N. Collier, N. Lightfoot, J. Linge, E. Goot, A. Mawudeku, L. Madoff, et al., An overview of Internet biosurveillance, Clin. Microbiol. Infect. 19 (11) (2013) 1006–1013.

[642] O.P. Wójcik, J.S. Brownstein, R. Chunara, M.A. Johansson, Public health for the people: participatory infectious disease surveillance in the digital age, Emerg. Themes Epidem. 11 (1) (2014) 7.

[643] T.A. Kass-Hout, H. Alhinnawi, Social media in public health, Br. Med. Bull. 108 (1) (2013) 5–24.

[644] Flu near you, http://fluneaYou.org.

[645] M.S. Smolinski, A.W. Crawley, K. Baltrusaitis, R. Chunara, J.M. Olsen, O. Wójcik, M. Santillana, A. Nguyen, J.S. Brownstein, Flu near you: crowdsourced symptom reporting spanning 2 influenza seasons, Am. J. Public Health. 105 (10) (2015) 2124–2130.

[646] Influenzanet, https://www.influenzanet.eu/.

[647] D. Paolotti, A. Carnahan, V. Colizza, K. Eames, J. Edmunds, G. Gomes, C. Koppeschaar, M. Rehn, R. Smallenburg, C. Turbelin, S. Van Noort, A. Vespignani, Web-based participatory surveillance of infectious diseases: the influenzanet participatory surveillance experience, Clin. Microbiol. Infect. 20 (1) (2014) 17–21.

[648] R. Chunara, E. Goldstein, O. Patterson-Lomba, J.S. Brownstein, Estimating influenza attack rates in the United States using a participatory cohort, Sci. Rep. 5 (2015) 9540.

[649] Q. Zhang, C. Gioannini, D. Paolotti, N. Perra, D. Perrotta, M. Quaggiotto, M. Tizzoni, A. Vespignani, Social data mining and seasonal influenza forecasts: the fluoutlook platform, in: Machine Learning and Knowledge Discovery in Databases, Springer International Publishing, 2015, pp. 237–240.

[650] W.J. Edmunds, S. Funk, Using the internet to estimate influenza vaccine effectiveness, Exp. Rev. Vaccines 11 (9) (2012) 1027.

[651] A. Barrat, C. Cattuto, Face-to-face interactions, in: Social Phenomena, Springer International Publishing, 2015, pp. 37–57.

[652] L. Danon, J.M. Read, T.A. House, M.C. Vernon, M.J. Keeling, Social encounter networks: characterizing Great Britain, Proc. R. Soc. Lond. Biol. 280 (1765) (2013) 20131037.

[653] K. Van Kerckhove, N. Hens, W.J. Edmunds, K.T. Eames, The impact of illness on social networks: Implications for transmission and control of influenza, Am. J. Epidemiol. 178 (11) (2013) 1655–1662.

[654] T. Smieszek, E. Burri, R. Scherzinger, R.W. Scholz, Collecting close-contact social mixing data with contact diaries: reporting errors and biases, Epidemiol. Infect. 140 (04) (2012) 744–752.

[655] T. Smieszek, V.C. Barclay, I. Seeni, J.J. Rainey, H. Gao, A. Uzicanin, M. Salathé, How should social mixing be measured: comparing web-based survey and sensor-based methods, BMC Infect. Dis. 14 (1) (2014) 136.

[656] J.M. Read, K.T. Eames, W.J. Edmunds, Dynamic social networks and the implications for the spread of infectious disease, J. R. Soc. Interface 5 (26) (2008) 1001–1007.

[657] P. Hui, A. Chaintreau, J. Scott, R. Gass, J. Crowcroft, C. Diot, Pocket switched networks and human mobility in conference environments, in: Proceedings of the 2005 ACM SIGCOMM Workshop on Delay-tolerant Networking, ACM, 2005, pp. 244–251.

[658] E. O'Neill, V. Kostakos, T. Kindberg, A. Penn, D.S. Fraser, T. Jones, et al., Instrumenting the city: Developing methods for observing and understanding the digital cityscape, in: UbiComp 2006: Ubiquitous Computing, Springer, 2006, pp. 315–332.

[659] N. Eagle, A. Pentland, Reality mining: sensing complex social systems, Pers. Ubiquitous Comput. 10 (4) (2006) 255–268.

[660] A. Pentland, T. Heibeck, Honest Signals, MIT Oress Cambridge, MA, 2008.

[661] M. Raento, A. Oulasvirta, N. Eagle, Smartphones an emerging tool for social scientists, Sociol. Methods Res. 37 (3) (2009) 426–454.

[662] M.S. Hashemian, K.G. Stanley, N. Osgood, Flunet: Automated tracking of contacts during flu season, in: Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt), 2010 Proceedings of the 8th International Symposium on, IEEE, 2010, pp. 348–353.

[663] C. Cattuto, W. Van den Broeck, A. Barrat, V. Colizza, J.-F. Pinton, A. Vespignani, Dynamics of person-to-person interactions from distributed rfid sensor networks, PLoS One 5 (7) (2010) e11596.

[664] N. Kiukkonen, J. Blom, O. Dousse, D. Gatica-Perez, J. Laurila, Towards rich mobile phone datasets: Lausanne data collection campaign, Proc. ICPS, Berlin.

[665] S. Liu, A. Striegel, Accurate extraction of face-to-face proximity using smartphones and bluetooth, in: Computer Communications and Networks, ICCCN, 2011 Proceedings of 20th International Conference on, IEEE, 2011, pp. 1–5.

[666] D.O. Olguín, A. Madan, M. Cebrian, A.S. Pentland, Mobile sensing technologies and computational methods for collective intelligence, in: Next Generation Data Technologies for Collective Computational Intelligence, Springer, 2011, pp. 575–597.

[667] N. Aharony, W. Pan, C. Ip, I. Khayal, A. Pentland, Social fMRI: Investigating and shaping social mechanisms in the real world, Pervasive Mobile Comput. 7 (6) (2011) 643–659.

[668] T. Hornbeck, D. Naylor, A.M. Segre, G. Thomas, T. Herman, P.M. Polgreen, Using sensor networks to study the effect of peripatetic healthcare workers on the spread of hospital-associated infections, J. Infect. Dis. (2012) jis542.

[669] A. Striegel, S. Liu, L. Meng, C. Poellabauer, D. Hachen, O. Lizardo, Lessons learned from the netsense smartphone study, in: ACM SIGCOMM Computer Communication Review, vol. 43, ACM, 2013, pp. 51–56.

[670] A. Stopczynski, V. Sekara, P. Sapiezynski, A. Cuttone, M.M. Madsen, J.E. Larsen, S. Lehmann, Measuring large-scale social networks with high resolution, PLoS One 9 (4) (2014) e95978.

[671] M.A. Kazandjieva, J.W. Lee, M. Salathé, M.W. Feldman, J.H. Jones, P. Levis, Experiences in measuring a human contact network for epidemiology research, in: Proceedings of the 6th Workshop on Hot Topics in Embedded Networked Sensors, ACM, 2010, p. 7.

[672] R. Mastrandrea, J. Fournet, A. Barrat, Contact patterns in a high school: a comparison between data collected using wearable sensors, contact diaries and friendship surveys, PLoS One 10 (9) (2015) e0136497.

[673] P. Jensen, M. Morini, M. Karsai, T. Venturini, A. Vespignani, M. Jacomy, J.-P. Cointet, P. Mercklé, E. Fleury, Detecting global bridges in networks, J. Complex Netw. (2015) cnv022.

[674] V. Sekara, A. Stopczynski, S. Lehmann, The fundamental structures of dynamic social networks, arXiv preprint, arXiv:1506.04704.

[675] P. Holme, Modern temporal network theory: a colloquium, Eur. Phys. J. B 88 (9) (2015) 1–30.

[676] A.-L. Barabasi, The origin of bursts and heavy tails in human dynamics, Nature 435 (7039) (2005) 207–211.

[677] B. Min, K.-I. Goh, A. Vazquez, Spreading dynamics following bursty human activity patterns, Phys. Rev. E 83 (3) (2011) 036102.

[678] J. Stehlé, N. Voirin, A. Barrat, C. Cattuto, L. Isella, J.-F. Pinton, M. Quaggiotto, W. Van den Broeck, C. Régis, B. Lina, et al., High-resolution measurements of face-to-face contact patterns in a primary school, PLoS One 6 (8) (2011) e23176.

[679] M. Karsai, M. Kivelä, R. Pan, K. Kaski, J. Kertész, A.-L. Barabási, J. Saramäki, Small but slow world: how network topology and burstiness slow down spreading, Phys. Rev. E 83 (2) (2011) 025102.

[680] G. Miritello, E. Moro, R. Lara, Dynamical strength of social ties in information spreading, Phys. Rev. E 83 (4) (2011) 045102.

[681] A. Vazquez, B. Racz, A. Lukacs, A.-L. Barabasi, Impact of non-poissonian activity patterns on spreading processes, Phys. Rev. Lett. 98 (15) (2007) 158702.

[682] B. Karrer, M.E. Newman, Message passing approach for general epidemic models, Phys. Rev. E 82 (1) (2010) 016101.

[683] L.E. Rocha, F. Liljeros, P. Holme, Simulated epidemics in an empirical spatiotemporal network of 50,185 sexual contacts, PLoS Comput. Biol. 7 (3) (2011) e1001109.

[684] T. Takaguchi, N. Masuda, P. Holme, Bursty communication patterns facilitate spreading in a threshold-based epidemic dynamics, PLoS One 8 (7) (2013) e68629.

[685] M. Morris, M. Kretzschmar, Concurrent partnerships and transmission dynamics in networks, Social Networks 17 (3) (1995) 299–318.

[686] J. Stehlé, N. Voirin, A. Barrat, C. Cattuto, V. Colizza, L. Isella, C. Régis, J.-F. Pinton, N. Khanafer, W. Van den Broeck, et al., Simulation of an SEIR infectious disease model on the dynamic contact network of conference attendees, BMC Med. 9 (1) (2011) 87.

[687] N. Masuda, P. Holme, Predicting and controlling infectious disease epidemics using temporal networks, F1000 prime reports 5 (2013) 6.

[688] J. Moody, The importance of relationship timing for diffusion, Social Forces 81 (1) (2002) 25–56.

[689] S.-Y. Liu, A. Baronchelli, N. Perra, Contagion dynamics in time-varying metapopulation networks, Phys. Rev. E 87 (3) (2013) 032805.

[690] C.S. Riolo, J.S. Koopman, S.E. Chick, Methods and measures for the description of epidemiologic contact networks, J. Urban Health 78 (3) (2001) 446–457.

[691] N. Fefferman, K. Ng, How disease models in static networks can fail to approximate disease in dynamic networks, Phys. Rev. E 76 (3) (2007) 031919.

[692] Y. Zhu, D. Li, W. Guo, F. Zhang, Effect of heterogeneity of vertex activation on epidemic spreading in temporal networks, Math. Probl. Eng. 2014 (2014) 409510.

[693] D.X. Horváth, J. Kertész, Spreading dynamics on networks: the role of burstiness, topology and non-stationarity, New J. Phys. 16 (7) (2014) 073037.

[694] A. Machens, F. Gesualdo, C. Rizzo, A.E. Tozzi, A. Barrat, C. Cattuto, An infectious disease model on empirical networks of human contact: bridging the gap between dynamic network data and contact matrices, BMC Infect. Dis. 13 (1) (2013) 1.

[695] M. Starnini, R. Pastor-Satorras, Temporal percolation in activity-driven networks, Phys. Rev. E 89 (3) (2014) 032807.

[696] K. Sun, A. Baronchelli, N. Perra, Contrasting effects of strong ties on SIR and SIS processes in temporal networks, Eur. Phys. J. B 88 (12) (2015) 1–8.

[697] D. Han, M. Sun, D. Li, Epidemic process on activity-driven modular networks, Physica A 432 (2015) 354–362.

[698] A. Sunny, B. Kotnis, J. Kuri, Dynamics of history-dependent epidemics in temporal networks, Phys. Rev. E 92 (2) (2015) 022811.

[699] P. Holme, N. Masuda, The basic reproduction number as a predictor for epidemic outbreaks in temporal networks, PLoS One 10 (3) (2015) e0120567.

[700] F. Liljeros, J. Giesecke, P. Holme, The contact network of inpatients in a regional healthcare system. a longitudinal case study, Math. Popul. Stud. 14 (4) (2007) 269–284.

[701] P. Holme, F. Liljeros, Birth and death of links control disease spreading in empirical contact networks, Sci. Rep. 4 (2014) 4999.

[702] D.J. Toth, M. Leecaster, W.B. Pettey, A.V. Gundlapalli, H. Gao, J.J. Rainey, A. Uzicanin, M.H. Samore, The role of heterogeneity in contact timing and duration in network models of influenza spread in schools, J. R. Soc. Interface 12 (108) (2015) 20150279.

[703] B. Ribeiro, N. Perra, A. Baronchelli, Quantifying the effect of temporal resolution on time-varying networks, Sci. Rep. 3 (2013) 3006.

[704] M.V. Tomasello, N. Perra, C.J. Tessone, M. Karsai, F. Schweitzer, The role of endogenous and exogenous mechanisms in the formation of r&d networks, Sci. Rep. 4 (2014) 5679.

[705] A. Moinet, M. Starnini, R. Pastor-Satorras, Burstiness and aging in social temporal networks, Phys. Rev. Lett. 114 (10) (2015) 108701.

[706] E. Ubaldi, N. Perra, M. Karsai, A. Vezzani, R. Burioni, A. Vespignani, Asymptotic theory for the dynamic of networks with heterogenous social capital allocation, arXiv preprint arXiv:1509.04563.

[707] J. Fournet, A. Barrat, Contact patterns among high school students, PLoS One 9 (9) (2014) e107878.

[708] M. Keeling, The implications of network structure for epidemic dynamics, Theoret. Popul. Biol. 67 (1) (2005) 1–8.

[709] J.C. Miller, Spread of infectious disease through clustered populations, J. R. Soc. Interface (2009) rsif–2008.

[710] M. Starnini, A. Baronchelli, R. Pastor-Satorras, Modeling human dynamics of face-to-face interaction networks, Phys. Rev. Lett. 110 (2013) 168701.

[711] M. Starnini, A. Baronchelli, R. Pastor-Satorras, Model reproduces individual, group and collective dynamics of human contact networks, arXiv preprint, arXiv:1409.0507.

[712] X. Sun, Z. Lu, X. Zhang, M. Salathé, G. Cao, Targeted vaccination based on a wireless sensor system, in: Pervasive Computing and Communications, PerCom, 2015 IEEE International Conference on, IEEE, 2015, pp. 215–220.

[713] M. Génois, C.L. Vestergaard, J. Fournet, A. Panisson, I. Bonmarin, A. Barrat, Data on face-to-face contacts in an office building suggests a low-cost vaccination strategy based on community linkers, arXiv preprint, arXiv:1409.7017.

[714] R. Cohen, S. Havlin, D. ben Avraham, Efficient immunization strategies for computer networks and populations, Phys. Rev. Lett. 91.

[715] N. Voirin, C. Payet, A. Barrat, C. Cattuto, N. Khanafer, C. Régis, B. Kim, B. Comte, J.-S. Casalegno, B. Lina, et al., Combining high-resolution contact data with virological data to investigate influenza transmission in a tertiary care hospital, Infec. Control Hosp. Epidemiol. 36 (03) (2015) 254–260.

[716] P. Vanhems, A. Barrat, C. Cattuto, J.-F. Pinton, N. Khanafer, C. Régis, B.-a. Kim, B. Comte, N. Voirin, Estimating potential infection transmission routes in hospital wards using wearable proximity sensors, PLoS One 8 (9) (2013) e73970.

[717] T. Obadia, R. Silhol, L. Opatowski, L. Temime, J. Legrand, A.C. Thiébaut, J.-L. Herrmann, É. Fleury, D. Guillemot, P.-Y. Boëlle, et al., Detailed contact data and the dissemination of staphylococcus aureus in hospitals, PLoS Comput. Biol. 11 (3) (2015) e1004170.

[718] R. Mastrandrea, A. Soto-Aladro, P. Brouqui, A. Barrat, Enhancing the evaluation of pathogen transmission risk in a hospital by merging hand-hygiene compliance and contact data: a proof-of-concept study, BMC Res. Notes 8 (1) (2015) 426.

[719] S.E. Cosgrove, The relationship between antimicrobial resistance and patient outcomes: mortality, length of hospital stay, and health care costs, Clin. Infect. Dis. 42 (Suppl. 2) (2006) S82–S89.

[720] J. Ginsberg, M.H. Mohebbi, R.S. Patel, L. Brammer, M.S. Smolinski, L. Brilliant, Detecting influenza epidemics using search engine query data, Nature 457 (7232) (2008) 1012–1014.

[721] A. Signorini, A.M. Segre, P.M. Polgreen, The use of Twitter to track levels of disease activity and public concern in the US during the influenza A H1N1 pandemic, PLoS One 6 (5) (2011) e19467.

[722] M.J. Paul, M. Dredze, You are what you Tweet: Analyzing Twitter for public health, in: ICWSM, 2011, pp. 265–272.

[723] T. Bodnar, M. Salathé, Validating models for disease detection using Twitter, in: Proceedings of the 22nd International Conference on World Wide Web Companion, International World Wide Web Conferences Steering Committee, 2013, pp. 699–702.

[724] D.A. Broniatowski, M.J. Paul, M. Dredze, National and local influenza surveillance through Twitter: An analysis of the 2012–2013 influenza epidemic, PLoS One 8 (12) (2013) e83672.

[725] A. Culotta, Towards detecting influenza epidemics by analyzing Twitter messages, in: Proceedings of the First Workshop on Social Media Analytics, ACM, 2010, pp. 115–122.

[726] N. Generous, G. Fairchild, A. Deshpande, S.Y. Del Valle, R. Priedhorsky, Detecting epidemics using Wikipedia article views: A demonstration of feasibility with language as location proxy, arXiv preprint, arXiv:1405.3612.

[727] D.J. McIver, J.S. Brownstein, Wikipedia usage estimates prevalence of influenza-like illness in the United States in near real-time, PLoS Comput. Biol. 10 (4) (2014) e1003581.

[728] K.S. Hickmann, G. Fairchild, R. Priedhorsky, N. Generous, J.M. Hyman, A. Deshpande, S.Y. Del Valle, Forecasting the 2013–2014 influenza season using Wikipedia, PLoS Comput. Biol. 11 (5) (2015) e1004239.

[729] E.O. Nsoesie, D.L. Buckeridge, J.S. Brownstein, Guess who is not coming to dinner? Evaluating online restaurant reservations for disease surveillance, J. Med. Internet Res. 16 (1) (2014).

[730] P. Butler, N. Ramakrishnan, E.O. Nsoesie, J.S. Brownstein, Satellite imagery analysis: What can hospital parking lots tell us about a disease outbreak? Computer 47 (4) (2014) 94–97.

[731] N. Collier, S. Doan, A. Kawazoe, R.M. Goodwin, M. Conway, Y. Tateno, Q.-H. Ngo, D. Dien, A. Kawtrakul, K. Takeuchi, et al., BioCaster: detecting public health rumors with a Web-based text mining system, Bioinformatics 24 (24) (2008) 2940–2941.

[732] M. Torii, L. Yin, T. Nguyen, C.T. Mazumdar, H. Liu, D.M. Hartley, N.P. Nelson, An exploratory study of a text classification framework for Internet-based surveillance of emerging epidemics, Int. j. Med. Inform. 80 (1) (2011) 56–66.

[733] Google flu trends, http://www.google.org/flutrends/.

[734] The next chapter of google flu trends, http://googleresearch.blogspot.co.uk/2015/08/the-next-chapter-for-flu-trends.html.

[735] A. Signorini, P.M. Polgreen, A.M. Segre, Using Twitter to estimate H1N1 influenza activity, in: 9th Annual Conference of the International Society for Disease Surveillance.

[736] M. Biggerstaff, D. Alper, M. Dredze, S. Fox, I.C.-H. Fung, K.S. Hickmann, B. Lewis, R. Rosenfeld, J. Shaman, M.-H. Tsou, et al., Results from the centers for disease control and prevention's predict the 2013–2014 Influenza Season Challenge, BMC Infec. Dis. 16 (1) (2016) 1.

[737] M. Del Vicario, A. Bessi, F. Zollo, F. Petroni, A. Scala, G. Caldarelli, H.E. Stanley, W. Quattrociocchi, The spreading of misinformation online, Proc. Natl. Acad. Sci. 113 (3) (2016) 554–559.

[738] H.J. Larson, D.M. Smith, P. Paterson, M. Cumming, E. Eckersberger, C.C. Freifeld, I. Ghinai, C. Jarrett, L. Paushter, J.S. Brownstein, et al., Measuring vaccine confidence: analysis of data obtained by a media surveillance system used to analyse public concerns about vaccines, Lancet Infect. Dis. 13 (7) (2013) 606–613.

[739] A. Kata, Anti-vaccine activists, Web 2.0, and the postmodern paradigm–an overview of tactics and tropes used online by the anti-vaccination movement, Vaccine 30 (25) (2012) 3778–3789.

[740] J.C. Hershey, D.A. Asch, T. Thumasathit, J. Meszaros, V.V. Waters, The roles of altruism, free riding, and bandwagoning in vaccination decisions, Org. Behav. Hum. Decis. Process. 59 (2) (1994) 177–187.

[741] R. Menzies, P. McIntyre, Vaccine preventable diseases and vaccination policy for indigenous populations, Epidemiol. Rev. 28 (1) (2006) 71–80.

[742] G.B. Chapman, M. Li, J. Vietri, Y. Ibuka, D. Thomas, H. Yoon, A.P. Galvani, Using game theory to examine incentives in influenza vaccination behavior, Psychol. Sci. (2012) 0956797612437606.

[743] M. Salathé, S. Khandelwal, Assessing vaccination sentiments with online social media: Implications for infectious disease dynamics and control, PLoS Comput. Biol. 7 (10) (2011) http://dx.doi.org/10.1371/journal.pcbi.1002199.

[744] M. Conover, J. Ratkiewicz, M.R. Francisco, B. Gonçalves, F. Menczer, A. Flammini, Political polarization on Twitter, in: Proceedings of the Fifth International AAAI Conference on Weblogs and Social Medi, vol. 133, 2011, pp. 89–96.

[745] W. Quattrociocchi, G. Caldarelli, A. Scala, Opinion dynamics on interacting networks: media competition and social influence, Sci. Rep. 4 (2014) 4938.

[746] M. Salathé, D.Q. Vu, S. Khandelwal, D.R. Hunter, The dynamics of health behavior sentiments on a large online social network, EPJ Data Sci. 2 (1) (2013) 1–12.

[747] M.S. Stockwell, A.G. Fiks, Utilizing health information technology to improve vaccine communication and coverage, Hum. Vaccines Immunother. 9 (8) (2013) 1802–1811.

[748] H.O. Witteman, B.J. Zikmund-Fisher, The defining characteristics of Web 2.0 and their potential influence in the online vaccination debate, Vaccine 30 (25) (2012) 3734–3740.

[749] J.E. Huston, S.R. Mekaru, S. Kluberg, J.S. Brownstein, Searching the web for influenza vaccines: Healthmap vaccine finder, Amer. J. Public Health 105 (8) (2015) e134–e139.

[750] D. Lazer, R. Kennedy, G. King, A. Vespignani, The parable of google flu: Traps in big data analysis, Science 343 (6176) (2014) 1203–1205. http://dx.doi.org/10.1126/science.1248506, URL http://www.sciencemag.org/content/343/6176/1203.short. arXiv:http://www.sciencemag.org/content/343/6176/1203.full.pdf.

[751] D.R. Olson, K.J. Konty, M. Paladini, C. Viboud, L. Simonsen, Reassessing google flu trends data for detection of seasonal and pandemic influenza: a comparative epidemiological study at three geographic scales, PLoS Comput. Biol. 9 (10) (2013) e1003256.

[752] C. Weeg, H.A. Schwartz, S. Hill, R.M. Merchant, C. Arango, L. Ungar, Using Twitter to measure public discussion of diseases: a case study, JMIR Public Health Surveill. 1 (1) (2015) e6.

[753] J. Shaman, A. Karspeck, Forecasting seasonal outbreaks of influenza, Proc. Natl. Acad. Sci. 109 (50) (2012) 20425–20430. http://dx.doi.org/10.1073/pnas.1208772109, URL http://www.pnas.org/content/109/50/20425.abstract. arXiv:http://www.pnas.org/content/109/50/20425.full.pdf+html.
[754] The world in 2014: ICT facts and figures. International telecommunication union, http://www.itu.int/.
[755] V.D. Blondel, A. Decuyper, G. Krings, A survey of results on mobile phone datasets analysis, EPJ Data Sci. 4 (1) (2015) 1–55.
[756] N. Eagle, A.S. Pentland, D. Lazer, Inferring friendship network structure by using mobile phone data, Proc. Natl. Acad. Sci. 106 (36) (2009) 15274–15278.
[757] F. Calabrese, Z. Smoreda, V.D. Blondel, C. Ratti, Interplay between telecommunications and face-to-face interactions: A study using mobile phone data, PLoS One 6 (7) (2011) e20814.
[758] M. Karsai, K. Kaski, A.-L. Barabási, J. Kertész, Universal features of correlated bursty behaviour, Sci. Rep. 2 (2012) 397.
[759] G. Krings, M. Karsai, S. Bernhardsson, V.D. Blondel, J. Saramäki, Effects of time window size and placement on the structure of an aggregated communication network, EPJ Data Sci. 1 (4) (2012) 1–16.
[760] J.L. Toole, Y.-A. de Montjoye, M.C. González, A.S. Pentland, Modeling and understanding intrinsic characteristics of human mobility, in: Social Phenomena, Springer International Publishing, 2015, pp. 15–35.
[761] A. Wesolowski, N. Eagle, A.J. Tatem, D.L. Smith, A.M. Noor, R.W. Snow, C.O. Buckee, Quantifying the impact of human mobility on malaria, Science 338 (6104) (2012) 267–270.
[762] D.K. Pindolia, A.J. Garcia, A. Wesolowski, D.L. Smith, C.O. Buckee, A.M. Noor, R.W. Snow, A.J. Tatem, Human movement data for malaria control and elimination strategic planning, Malar. J. 11 (1) (2012) 205.
[763] A. Wesolowski, C.O. Buckee, L. Bengtsson, E. Wetter, X. Lu, A.J. Tatem, Commentary: containing the Ebola outbreak-the potential and challenge of mobile network data, PLoS Curr. 6.
[764] M. Tizzoni, P. Bajardi, A. Decuyper, G.K.K. King, C.M. Schneider, V. Blondel, Z. Smoreda, M.C. González, V. Colizza, On the use of human mobility proxies for modeling epidemics, PLoS Comput. Biol. 10 (7) (2014) e1003716.
[765] Health map, http://healthmap.org/.
[766] C.C. Freifeld, K.D. Mandl, B.Y. Reis, J.S. Brownstein, Healthmap: global infectious disease monitoring through automated classification and visualization of internet media reports, J. Amer. Med. Inform. Assoc. 15 (2) (2008) 150–157.
[767] J. Brownstein, C. Freifeld, et al., HealthMap: the development of automated real-time internet surveillance for epidemic intelligence, Euro Surveill. 12 (11) (2007) E071129.
[768] Health map projects, http://healthmap.org/site/projects.
[769] Health map publications, http://www.healthmap.org/site/about/publications.
[770] Health map flu trends, http://www.healthmap.org/flutrends/.
[771] M. Santillana, A.T. Nguyen, M. Dredze, M.J. Paul, E.O. Nsoesie, J.S. Brownstein, Combining search, social media, and traditional data sources to improve influenza surveillance, PLoS Comput. Biol. 11 (10) (2015) e1004513.
[772] Dengue map, http://www.healthmap.org/dengue/en/.
[773] Health map vaccine, http://flushot.healthmap.org/.
[774] Crowdbreaks, http://www.crowdbreaks.com/.
[775] V. Colizza, A. Barrat, M. Barthélemy, A. Vespignani, The role of the airline transportation network in the prediction and predictability of global epidemics, Proc. Natl. Acad. Sci. USA 103 (7) (2006) 2015–2020.
[776] R. Durrett, Epidemic Models: their Structure and Relation to Data, vol. 5, Cambridge University Press, 1995.
[777] H.J. Kreuzer, Nonequilibrium Thermodynamics and its Statistical Foundations, vol. 1, Clarendon Press, Oxford, New York, 1981, p. 455.