Università degli Studi di Trento
Facoltà di Scienze Matematiche, Fisiche e Naturali

Doctoral school in Mathematics - Cycle XXIX

PhD Thesis

Mathematical modeling for epidemiological inference and public health support

Advisors:
Prof. Andrea Pugliese
University of Trento

Dr. Stefano Merler
Fondazione Bruno Kessler

Author:
Valentina Marziano

Contents

Introduction

Since the dawn of time infectious diseases have caused great damage to human societies. During the Middle Ages for example, the death toll of the black plague was as high as the 30-60% of human population [1]. After the introduction of the first vaccine, which dates back to 1796, when Edward Jenner successfully used cowpox material to induce immunity to smallpox, the effectiveness of prevention and control programs aimed at reducing the burden of diseases caused by infectious pathogens has significantly improved. Vaccination programs have played a crucial role in the efforts aimed at reducing the incidence of infectious diseases. As a result of vaccination policies, global eradication of smallpox has been achieved more than 30 years ago, polio is currently circulating only in few countries around the world and measles mortality has globally decreased by 79% between 2000 and 2015 [2]. However, large epidemics still represent a persisting major threat for public health in both developing and developed countries. According to the World Health Organization each year at least 15 millions people die because of infectious diseases [3]. Low-income countries are the most affected, with almost one third of the yearly deaths ascribable to acute lower respiratory infections, HIV, malaria and tuberculosis. The twenty-first century has only just begun, yet the world has already experienced at least four emergency situations caused by infectious diseases: the 2003 Severe Acute Respiratory Syndrome (SARS) epidemic, the 2009 H1N1 influenza pandemic, the 2014-2016 Ebola epidemic and, more recently, the emergence of Zika virus. In addition, in recent years, some of the most developed countries have failed to maintain a sufficiently high vaccine uptake to control vaccine preventable diseases [4], partially due to the growing skepticism about vaccines, suggesting that the battle against infectious diseases is still far from being won.

Mathematical modeling has been used since the beginning of the last century to identify the key general mechanisms driving the spread of infectious diseases [5]. Nowadays, it represents a powerful tool with a wide range of applications in epidemiology and public health [6]. One key role of mathematical modeling is the "understanding" of the past, which is an essential step to better inform the design of future public health control measures. Indeed, epidemiological data provide only fragmentary evidences about the diseases epidemiology and only partial information about the dynamics underlying what is observed. Mathematical models, informed with data, can help disentangling the complexity of transmission and provide explanations for the trends observed (e.g. recurrent epidemics, seasonality)

[7]. More importantly, mathematical modeling plays a crucial role in projecting the future trends of infectious disease dynamics and in estimating the impact of possible interventions thus, supporting health authorities in the choice of adequate strategies for the effective prevention and control of infectious cases. During the last decades, models have been often coupled with cost-effectiveness analyses which evaluate the economic sustainability of an intervention from a public health perspective. The economic evaluation of a number of vaccination programs have been used in the past to help inform public health decisions for a variety of infectious diseases; such as; meningo C and B [8, 9], pneumococcal disease [10], influenza [11, 12] and varicella [13, 14, 15, 16].

Over the past decades, the number of countries which uses mathematical modeling to support public health decisions has been increasing. According to a recent report published by the European Centre for Disease Prevention and Control (ECDC) about 70% of the European Union member states consider mathematical modelling as part of the process for planning routine vaccination strategies [17]. More recently, mathematical modeling has also played a key role to inform public health authorities during the emergence of new epidemics. During the 2014-2016 Ebola epidemic in West Africa, a Modeling Task Force was activated by the Centre for Disease Control and Prevention (CDC) to estimate in real-time key epidemiological quantities, useful to make public health decisions and allocate resources [18]. This experience has strengthened the awareness of health authorities about the importance of modeling tools, especially in the early stages of an outbreak and when data are lacking.

The increasing use of modeling by public health decision makers brings out the challenge of building models, whose results should be as much accurate and realistic as possible. To this aim, they need to be informed with data and to include the individual and population features which play a relevant role in the spread of the disease. A lot of studies have contributed in this direction, by shedding light on the role of heterogeneous mixing patterns by age [19], of human mobility [20, 21] or the sociodemographic structure of the population [22]. However, there are some issues that require further investigation.

One of them is represented by the inclusion in models of realistic demographic dynamics. Standard modeling approaches assume the population as either constant or characterized by a stable age-structure, and even when the projection of future epidemiological trends account for realistic demographic changes, models were calibrated under the assumption of constant demographic conditions. However, some recent modeling studies strongly have highlighted population demographic factors as relevant drivers in the dynamics of infectious disease and encourages the introduction of a new class of epidemiological models accounting for realistic demographic dynamics [23, 24].

Another topic that requires further investigation is the interplay between epidemics and human behavioral changes [25]. The capability of individual responses to alter transmission dynamics is well known [26, 27, 28], however a proper understanding and quantification of its impact is required in order to improve control efforts [29].

Finally, the difficulties recently experienced by most developed countries in the attain-

ment and maintenance of vaccination coverage levels that are sufficiently high to eliminate the disease have produced unexpected and novel epidemiological conditions, characterized by a low circulation of viruses and local interruption of disease transmission, making more difficult the interpretation of such complex dynamics. A better understanding of the epidemiological features of low-circulation settings is therefore necessary to fill important gaps in the basic knowledge on the epidemiology of infectious diseases and more adequate modeling tools should be provided to inform future intervention programs.

The works presented in this thesis attempt to shed light on some of the previously highlighted issues and to contribute to the development of the field of epidemiological modeling.

In Chapter 1 we explore the role of demographic changes in the transmission dynamics of varicella and Herpes Zoster. To this aim, a transmission model informed with country-specific demographic data is developed, and used to investigate the past dynamics of varicella and HZ in Spain and predict their future burden under illustrative demographic and vaccination scenarios.

In Chapter 2 the model developed in Chapter 1 is combined with a cost-effectiveness analysis to evaluate the joint impact of demographic changes and varicella and HZ vaccination strategies, which are currently under consideration, on the future epidemiology of varicella and HZ in Italy.

In Chapter 3 we use a computational model embedded in a Bayesian framework in order to detect and quantify the effect of spontaneous behavioral changes on the spatiotemporal dynamics of the 2009 H1N1 influenza pandemic in England.

In Chapter 4 we analyze the current epidemiology of measles in Italy to provide some insights into its epidemiological features in a low circulation setting. In particular, we use a detailed computational model, informed with regional heterogeneities in the age-specific seroprevalence profiles, to estimate where secondary infections occur and the attack rates that we can expect in different regions.

Chapter 1

The impact of demographic changes on the epidemiology of varicella and Herpes Zoster

1.1 Background

Varicella zoster virus (VZV) is responsible for two clinically different diseases. The first exposure to VZV usually occurs in childhood and causes varicella disease. After recovery, the virus remains dormant in sensory ganglia and in about 30% of cases it will reactivate later in life, causing herpes zoster (HZ) [30], an inflammatory skin disease associated with more serious morbidity [31]. In 1965, Hope-Simpson hypothesized that secondary exposures to VZV (also called "exogenous boosting events") boost the host's VZV-specific antibodies, thereby reducing the risk of developing HZ [32]. Later, it was found that the main component of immunity against HZ was cell-mediated [33, 34], but Hope-Simpson's exogenous boosting hypothesis has been corroborated by immunological studies showing an increase of VZV-specific T cells in people exposed to contacts with the virus [35, 36, 37, 38], as well as by epidemiological [39, 40] and modelling studies [40, 41, 42, 43]. Mathematical models of VZV transmission dynamics, all based on the exogenous boosting hypothesis, unanimously predict that mass immunization against varicella, which has been already introduced in some countries [44, 45], including some regions of Spain [46], may result in a temporary increase of HZ incidence [41, 42, 47, 48] as a consequence of the expected reduction of VZV re-exposures after vaccination. Nonetheless, empirical evidence coming from surveillance programmes of HZ provides ambiguous results, with some studies showing an increase of HZ following mass immunization and others not [49, 50]. However, an increase in HZ incidence has been detected in several areas long before the introduction of the varicella vaccine [51, 52, 53, 54] and, more recently, in areas without relevant VZV vaccination history [50, 55]: for example, in the Community of Madrid (where until 2005 the vaccine was introduced only for individuals at high risk of complications), the annual HZ incidence has risen from about 2.5 to about 3.6 cases per 1,000 individuals per year between 1997 and 2004, and the trend was statistically significant [55]. This suggests the presence of other

factors that might confound trend interpretations. One such factor may be represented by demographic changes. During the twentieth century, industrialized countries completed their fertility transitions [56], with a massive decline in birth rates and substantial population ageing. In particular, since 1980, several central and southern European countries have experienced sustained below-replacement fertility [57]. This work[1] aims at exploring the role of demographic changes in VZV transmission and reactivation dynamics by making use of a mathematical model. The developed model, informed with country-specific epidemiological and demographic data, is used to investigate the past dynamics of varicella and HZ in Spain, and to predict their future burden under different demographic and vaccination scenarios.

1.2 Methods

1.2.1 Data

Demographic changes for the period 1900-2009 were modelled using data on yearly birth rates [58], age-specific mortality rates [59] and age-specific migration flows [60, 61, 62] over time. Historical age structures of the Spanish population [58] were used to validate the demographic model. Available projections on mortality, birth and migration rates [63] relative to the period 2010-2050 were used for model predictions about future dynamics. For what concerns the transmission model, as no survey data on social mixing patterns [19, 64] are available for Spain, we parametrized age-specific contact patterns with estimates based on census-type data provided by a previously published study [65]. Calibration was done using age-specific VZV seroprevalence data in 1996 for Spain, made available by the European Seroepidemiological Network 2 [66], and age-specific data on HZ incidence from 2005 to 2006 [67]. Model estimates were also qualitatively compared against an independent dataset that included age-specific HZ incidence data in the years 1997-2004 [55].

1.2.2 The model

We use a stochastic individual-based age-structured epidemiological model for the natural history of varicella and HZ, explicitly accounting for demographic changes. Epidemiological transitions are modelled as in [43]: individuals are born susceptible to VZV and develop varicella upon contacts with VZV-infectious individuals. Contacts are assumed to be structured by age according to a mixing matrix previously estimated for Spain [65] and rescaled at each time step by the current age distribution of the population, following the approach proposed in [24]. After recovery, individuals gain lifelong immunity against varicella and become susceptible to HZ. The risk of HZ increases with age and time elapsed since the last exposure episode. We assume that the risk of HZ declines progressively upon repeated

[1] published in the Proceedings of the Royal Society B: Biological Sciences. Marziano et al. (2015)

contacts with VZV-infectious individuals, in order to take into account the accumulation of protection against reactivation that Hope-Simpson postulated to follow from successive re-exposures to VZV [31, 32]. Only a fraction of contacts that would result in varicella infection in susceptible individuals lead to an effective immune boosting in VZV-experienced individuals. After recovery from HZ, individuals are assumed to become lifelong immune to new episodes of VZV reactivation [40, 41, 42, 43]. Indeed, secondary HZ episodes, although possible, are relatively rare in immunocompetent individuals [31]. The vital dynamics of the population is explicitly modelled in order to reproduce the observed changes in the demographic structure of the Spanish population over the period 1900-2009, following the approach of a previously published model for Italy [23]. As in [23], the model is initialized in 1900 at both demographic and epidemiologic equilibrium, whereas for the following years, demographic dynamics are driven by variations of birth, age-specific mortality and migration rates. Postulating that the level of the birth rate is a main determinant of transmission in childhood infections, we believe the hypothesis of epidemiological equilibrium in 1900 is motivated by the fact that most of the fertility decline in Spain occurred after that date [68]. Full details about the developed model are reported in section A.1 of the appendix, together with a validation of the demographic model against observed longitudinal data on the age structure of the Spanish population.

1.2.3 Model calibration and validation

Model calibration was carried out by a Bayesian approach using uniform prior distributions on the six model parameters combined with the likelihood of varicella and zoster data. The latter is defined as the product between the binomial likelihood of the observed age-specific VZV serological profile in 1996 [66] and the Poisson likelihood of the observed age-specific HZ incidence in 2005-2006 [67]. Computation of posterior distributions was carried out by means of a Monte Carlo Markov chain approach with random-walk Metropolis-Hastings sampling with normal jump distributions. Convergence of MCMC (22,000 iterations) was assessed by considering several different starting points and by visual inspection, after a burn-in period of 2,000 iterations. Details on model calibration and the estimated posterior distribution of parameters are available in section A.1 of the appendix.

1.2.4 Demographic projections

The model is used to retrospectively analyse the impact of demographic changes on varicella and HZ (period 1900-2009) and to predict their future dynamics (period 2010-2050). The latter were simulated by using both published demographic projections [63] and theoretical scenarios, based for simplicity on temporal changes during the prediction period (2010-2050) of the birth rate only. Theoretical scenarios have been formulated by assuming that all other demographic indicators are kept equal to their 2009 values, based on the assumption that mortality and migration play a minor role on dynamics of varicella infection. In fact,

the circulation of infectious diseases depends on the fraction of susceptibles within the population, which, in the case of childhood infections, is shaped by the birth rate. As varicella occurs predominantly in children, either an increase or a decrease of the birth rate has the potential of leading respectively to higher or lower VZV circulation, which can affect the dynamics of HZ through the effect of boosting. Figure 1.1 shows the reported birth rate in Spain over the period 1900-2009, along with projected scenarios for the period 2010-2050. In particular, we considered a baseline projection scenario that assumes a constant birth rate; one that assumes a linear decrease of the birth rate up to a minimum of 0 in 2050 (denoted by “L" for “lowest"); and one assuming a linear increase with a doubling of the 2009 birth rate over the same period (denoted by “H" for “highest"). Although these scenarios are demographically naive, they represent useful illustrative limit cases for HZ. A more realistic scenario, based on published population projections made available from the United Nations [63] for Spain, has also been considered; this scenario includes changes in mortality rates and migration fluxes, and projects a decrease of the birth rate between 2010 and 2030, followed by a partial recovery between 2030 and 2050 (Figure 1.1).

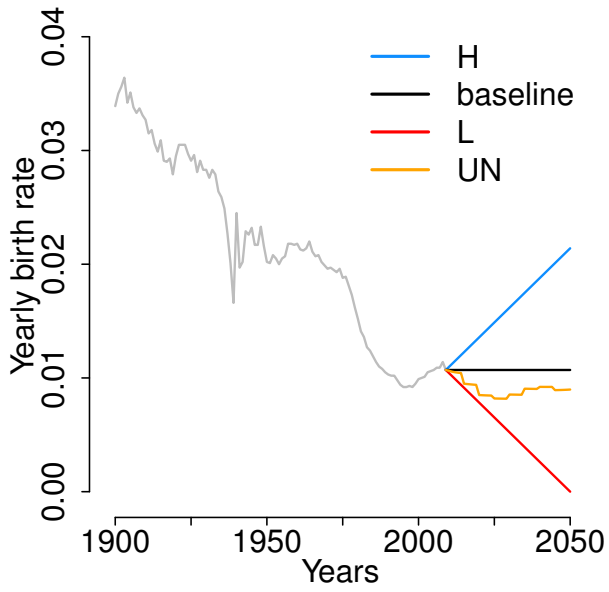

Figure 1.1: **Birth rate over time in Spain.** Yearly birth rate as observed in Spain during the period 1900-2009 (grey line) and as assumed by different prediction scenarios considered (2010-2050). Illustrative scenarios: H, baseline and L, Scenario based on published population projections of the UN [63].

1.2.5 Vaccination

The model is applied to evaluate the potential long-term effects of a varicella vaccination programme starting in 2010 on future HZ incidence under the baseline scenario for demographic projections. Vaccine is administered to children at the age of 15 months. We assume different combined levels of vaccine efficacy and coverage, namely 80%, 90% and 100%. Successfully immunized individuals are assumed to be fully protected against VZV infection (i.e. we do not explicitly consider breakthrough varicella infections) [69]. Moreover, vaccinated individuals can develop HZ from vaccine strain, although at a lower rate compared with natural VZV, as observed in [70].

1.3 Results

Figure 1.2 shows the ability of the model to fit the VZV serological profile observed in Spain in 1996. Observed and estimated 95% CI intersect for all younger age classes (age 20 years or below). The model also reproduces the observed age-specific HZ incidence in 2005-2006. The observed yearly incidence lies in the 95% CI of the estimated incidence for most age groups. Moreover, the model captures the decreasing trend of incidence in the elderly, which may be ascribed to exogenous boosting effects [42, 43].

1.3.1 Historical period (1900-2009)

Figure 1.3 shows model results for the period 1900-2009. According to the model, the overall decreasing trend of the birth rate observed since 1900 has led to a reduction of the fraction of VZV seropositive individuals at all ages (Figure 1.3**a**). For instance, the estimated fraction of children who have acquired varicella by 10 years of age dropped from 99.2% (95% CI 98.8-99.6) in 1900 to a minimum of 81.5% (95% CI 76.5-85.8) in 2004. Consequently, the mean age at varicella infection is estimated to increase from about 1.4 (95% CI 1.2-1.6) to a maximum of 5.6 (95% CI 4.9-6.2) years (Figure 1.3**b**). Most of this increase (from about 2.5 to 5.6 years) occurs during the prolonged epoch of low fertility started in the second half of the 1970s [58]. The modelled effect of changes in the birth rate on varicella is due to a depletion of new susceptible individuals that sustain the circulation of the infection, and occurs with a delay of a few years. Major fertility fluctuations observed in the past century are reflected in the estimated mean age at infection. For example, the increase in the mean age at varicella infection that is estimated around 1940 (Figure 1.3**b**) is a consequence of the temporary brisk drop of the birth rate during the Spanish civil war in 1936-1939 (Figure 1.1); similarly, the stabilization in the mean age estimated in the decade 2000-2009 is due to the recovery of the birth rate started in the late 1990s. As regards HZ, the model estimates a gradual increase of the incidence rates since 1900 (Figure 1.3**c**, inset). This trend, at least for recent years, has been documented for Spain by a surveillance study [55] reporting a remarkable growth in the total yearly HZ incidence over the period 1997-2004. In particular,

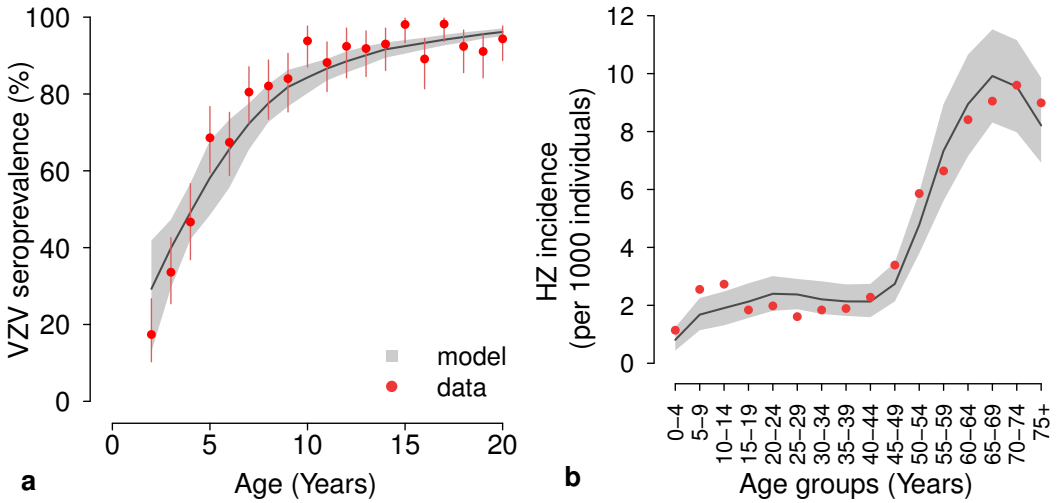

Figure 1.2: **VZV seroprevalence and HZ incidence. a** Age-specific VZV seroprevalence as observed in data from Spain in 1996 [66] and as estimated by the model. Vertical lines represent 95% CI of the data computed by exact binomial test. **b** HZ incidence by age group as observed in 2005-2006 [67] and as estimated by the model. In both panels, grey areas show the 95% CI of model estimates.

HZ incidence is approximately stable for younger individuals, with virtually all of the HZ growth concentrated at ages above 65 [55]. In comparison, our model estimates an increase in the total incidence of about 12% (up to 18% by assuming different contact patterns by age; see section A.5 of the appendix) over the same period, from 3.52 (95% CI 3.18-3.87) to 3.93 (95% CI 3.58-4.31) cases per 1000 individuals, also concentrated in older adults. These changes in HZ incidence are reflected in a progressive increase of the mean age at VZV reactivation from about 33.6 (95% CI 29.3-37.7) years in 1900 to 54.4 (95% CI 52.9-56.1) years in 2009 (Figure 1.3**d**). The estimated HZ growth is mainly ascribable to the declining birth rate via the complex dynamics of exogenous boosting. Indeed, the individual risk of HZ is higher when the number of experienced episodes of VZV re-exposures is low. The sustained decrease of the birth rate substantially reduced VZV circulation and frequency of re-exposures in the population, therefore contributing to increase the overall HZ risk at the population level. In other words, each cohort has been replaced over time by a new cohort with a lower level of protection, which has caused the increase of HZ incidence throughout the past century at virtually all ages. It is worth noting that, owing to these indirect and cumulative effects, the result of a changing birth rate becomes visible on HZ incidence after a significant delay of a few decades. For instance, the model estimates that despite the

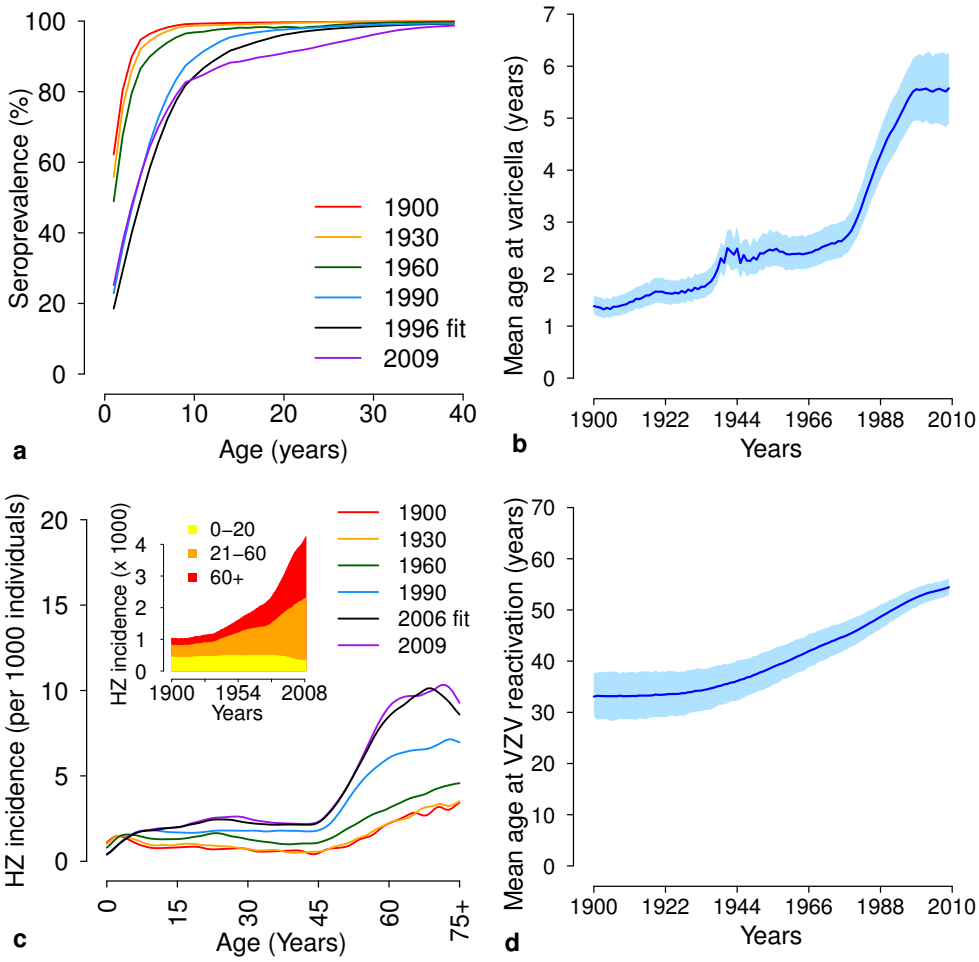

Figure 1.3: **Estimated impact of demographic changes on VZV epidemiology (1900-2009) a** Estimated age-specific VZV seroprevalence at different years. **b** Estimated mean age (and 95% CI, shaded areas) at varicella over time. **c** Estimated age-specific HZ incidence at different years. The inset shows the total HZ incidence over time and disaggregated by age group. **d** Estimated mean age at VZV reactivation (and 95% CI).

increasing birth rate in 2000-2009, HZ incidence keeps growing, although at a slower rate (Figure 1.3**c**, inset). The delay in the effect of birth rate changes on HZ incidence is intrinsic in the long time scales of VZV reactivation, which critically depends on the accumulation of individual immunological protection through VZV re-exposures over several decades.

1.3.2 Prediction period (2010-2050)

Model predictions on the future epidemiology of varicella are shown in Figure 1.4, under the three illustrative demographic scenarios presented in the Methods section. Results for scenarios based on realistic population projections [63] are qualitatively similar to the scenario where a constant birth rate is assumed and they are subjected to the same general considerations characterizing the analysis of the illustrative scenarios. They are therefore reported for brevity in section A.3 of the appendix. In the baseline scenario (where a constant birth rate is assumed), the predicted mean age at VZV infection (Figure 1.4**a**) and the serological profile (Figure 1.4**b**) remain substantially stable after 2010. Scenario L, which assumes a decreasing birth rate, results in a remarkable rise of the mean age at VZV infection (Figure 1.4**a**) and in a decrease in the seropositive fraction at all ages (Figure 1.4**c**). On the opposite, scenario H, characterized by a growing birth rate, results in a decrease of the mean age at VZV infection (1.4**a**) and a growing fraction of seropositive individuals at all ages (Figure 1.4**d**). The predicted impact of these scenarios on future HZ epidemiology is shown in Figure 1.5. In the baseline scenario, the model predicts an increase of the mean age at VZV reactivation, from 54.4 years (95% CI 52.9-56.1) in 2009 to 62.0 years (95% CI 60.7-63.2) in 2050 (Figure 1.5**a**). Correspondingly, the age-specific HZ incidence (Figure 1.5**b**) also increases, especially in adults and in the elderly, with peak values rising from about 10 to about 20 cases per 1000 individuals. This trend of HZ growth is consistent with the observed historical trend [55] and occurs despite the relative stability of the underlying varicella dynamics (Figure 1.4**a**) as a result of the delayed effect of the decreasing birth rate in the second half of the past century on the risk of HZ development. In scenario L, the increase in HZ incidence is inflated by the further decrease of VZV circulation predicted during the period 2010-2050 (Figure 1.5**c**). In scenario H, the increase of VZV circulation started in the first decade of the twenty-first century mitigates the growth of HZ incidence and eventually counterbalances it after a few decades (Figure 1.5**d**, inset).

1.3.3 Predictions under varicella vaccination

Figure 1.6 shows the age-specific and total HZ incidence over time in the case of a varicella vaccination programme starting in 2010 and assuming a combined level of vaccine effectiveness and coverage of 90%, under the baseline demographic scenario. The model predicts a further dramatic increase of HZ incidence in all non-vaccinated cohorts, as a consequence of the increased risk owing to the rapid decline of varicella circulation (see section A.4 of the appendix). In particular, the total HZ incidence is predicted to increase by 185% between 2010 and 2050, from 4.24 (95% CI 3.86-4.67) to 12.09 (95% CI 11.02-13.30) cases per 1000 individuals, compared with an 89% increase in the baseline scenario without vaccination (Figure 1.5**b**, inset). Similarly, the peak age-specific incidence soars to 37.5 cases per 1000 individuals in 2050 compared with a maximum of about 20 cases per 1000 individuals in the same year in the corresponding vaccination-free scenario. HZ incidence caused by the

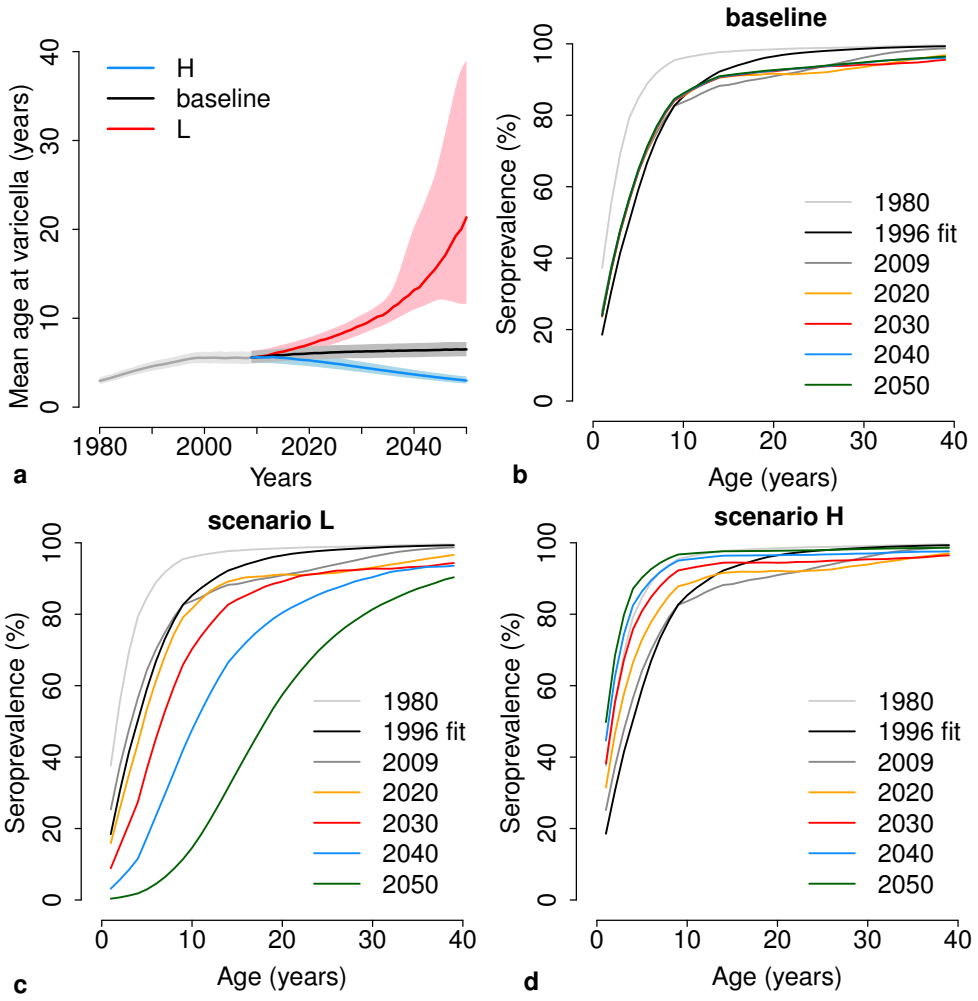

Figure 1.4: **Predicted impact of demographic changes on the future epidemiology of varicella (2010-2050) a** Predicted mean age (and 95% CI, shaded areas) at varicella over time. Different colors correspond to different projection scenarios. **b-d** Predicted age-specific VZV seroprevalence at different years for the baseline scenario (b) the "lowest birth rate" scenario L (c) and the "highest birth rate" scenario H.

reactivation of the vaccine strain is negligible, contributing less than 1.4% of all HZ cases that occurred throughout the prediction period. Results based on a different combined level of vaccine effectiveness and coverage are shown in section A.4 of the appendix.

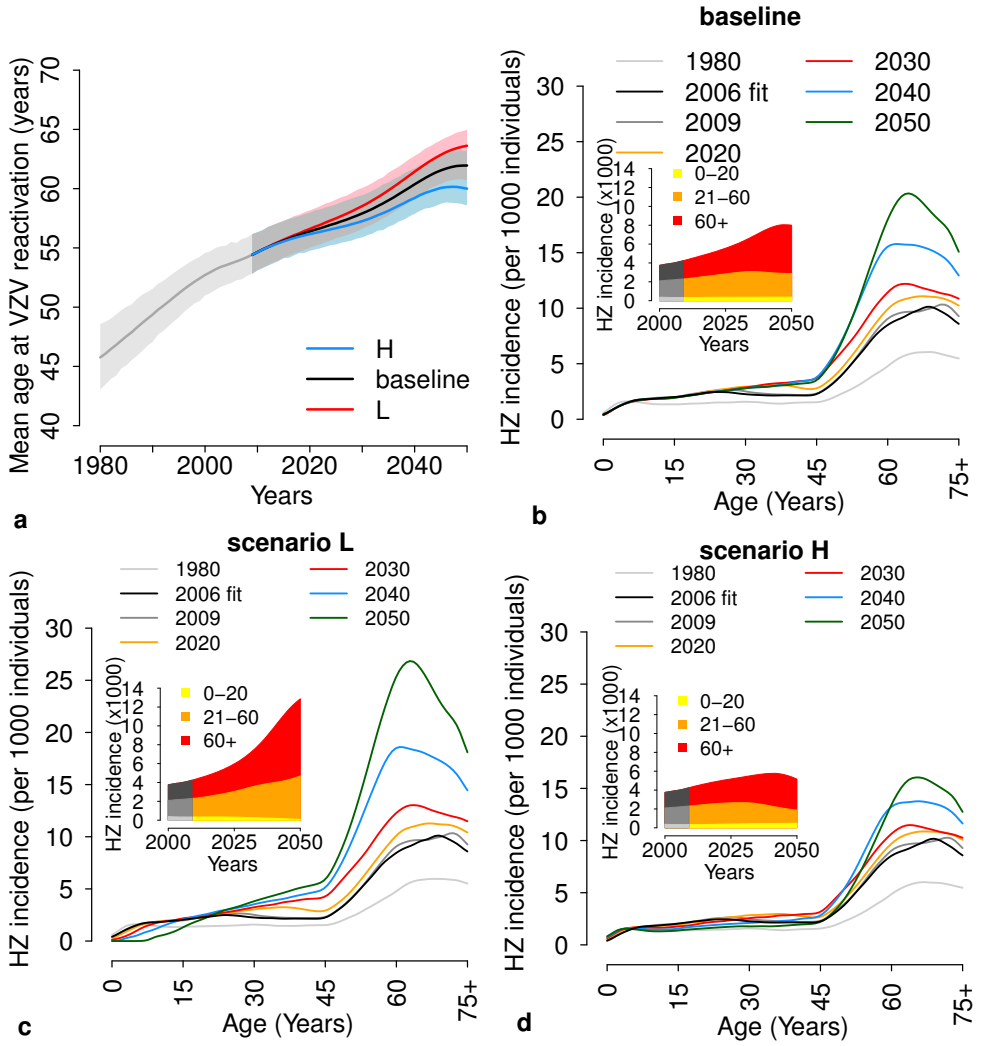

Figure 1.5: **Predicted impact of demographic changes on the future epidemiology of HZ (2010-2050) a** Predicted mean age (and 95% CI, shaded areas) at HZ over time. Different colors correspond to different projection scenarios. **b-d** Predicted age-specific HZ incidence at different years for the baseline scenario (b) the "lowest birth rate" scenario L (c) and the "highest birth rate" scenario H. The insets show the total HZ incidence over time disaggregated by age group.

1.4 Discussion

In the past century, a range of demographic changes has occurred in most industrialized countries [56], yielding a massive decrease of fertility and a progressive ageing of the popula-

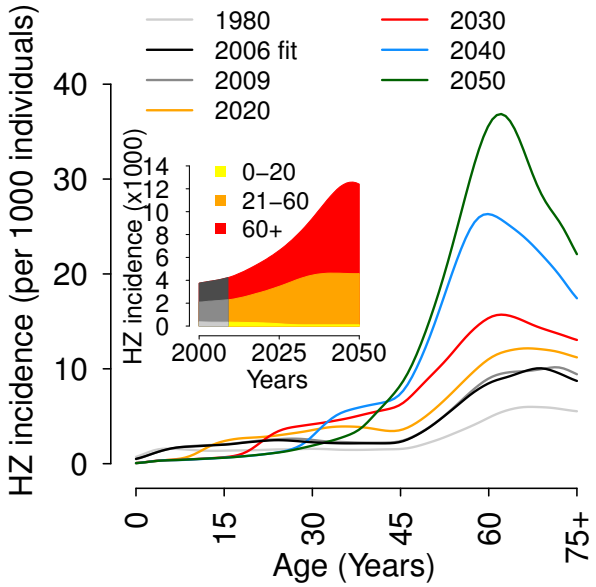

Figure 1.6: **Joint predicted impact of demographic changes and varicella vaccination on the future epidemiology of HZ (2010-2050)** . Predicted age-specific HZ incidence at different years under the baseline demographic scenarios and a vaccination programme starting in 2010 with 90% coverage and 100% life-long efficacy. The inset shows the total HZ incidence over time disaggregated by age group.

tion [58]. In this work, we investigate the impact of demographic dynamics on the epidemiology of varicella and HZ in Spain, by using an age-structured individual-based stochastic model of VZV transmission dynamics. The model is calibrated against the age-specific profiles of VZV seroprevalence and HZ incidence, and is able to qualitatively reproduce the age-specific increase in HZ observed between 1997 and 2004 [55]. In the model, a decrease in the birth rate results in a decrease of VZV circulation, owing to the depletion of susceptible individuals that fuel transmission. Under Hope-Simpson's exogenous boosting hypothesis, the reduced VZV circulation results in an increase of the reactivation risk. In particular, the model suggests that HZ incidence has been increasing in Spain since the beginning of the past century, in association with a strong reduction of varicella circulation, mainly caused by the progressive decline in fertility. While varicella-related epidemiological changes follow demographic changes with a delay of a few years, consequences on HZ incidence require several decades to reveal. As a consequence, under all demographic projections considered for the future, the model predicts a further increase of HZ incidence until at least 2040. This trend is expected to be eventually reverted only in the case of a massive, sustained

growth of the birth rate over the next decades.

Model results are a consequence of the assumption that varicella and HZ dynamics are linked through the exogenous boosting hypothesis. The existence of a relationship between VZV re-exposures and reduced HZ risk is largely supported by data [35, 36, 37, 39, 40], even though some opposing evidence [71] shows that the question is far from being completely elucidated and requires further study. Published mathematical models of VZV [40, 41, 42, 43] unanimously integrate the exogenous boosting hypothesis, as it represents a simple and plausible explanation of the observed age-specific HZ incidence patterns in several European countries [40, 41, 42, 43, 48, 55, 67, 72]. In this study, we chose to adopt a formulation that accounts for the cumulative increase of protection with repeated exposures, termed "progressive immunity" in a previously proposed study [43]. The progressive immunity model has been shown to accurately fit HZ data in several European countries, including Spain [43]. However, we found similar qualitative results when using a different formulation of the exogenous boosting hypothesis (following [42, 47, 48]), where immunity does not cumulate with successive re-exposure episodes (see secton A.6 in the appendix).

A limitation of this model concerns the representation of contact patterns for the transmission dynamics of VZV. Actual mixing patterns may change significantly over time, following changes in socio-demographic characteristics [65, 73] (e.g. household size and composition, characteristics of the educational system, school attendance levels). However, changes in contact patterns by age that occurred during the past century are difficult to reconstruct quantitatively, because neither historical data on mixing patterns themselves nor detailed historical epidemiological data useful to identify these patterns are available. In the absence of sufficient information on historical mixing patterns, and because the construction of time-dependent contact matrices is beyond the scope of this paper, we adopted recent estimates of age-specific mixing patterns [65], rescaled by the fraction of population by age group at each time step [24]. However, the main results of this work are robust with respect to this assumption. To support this claim, we performed a further analysis where mixing patterns are assumed to be homogeneous by age. Model estimates obtained under this alternative contact pattern scheme support the same qualitative conclusions with limited quantitative variations (see section A.5 in the appendix).

The choice of Spain for this study was dictated by the availability of longitudinal HZ incidence data [55], which show a strong trend of growth before the introduction of varicella vaccination. Similar trends have been reported in the pre-vaccine era in other geographical settings (e.g. USA [53], UK [54], Canada [51, 52, 54]), but factors underlying the growth of HZ have not been elucidated yet. Our study shows that demographic changes have contributed, at least partially, to this trend in Spain and may possibly be generalizable to other geographical settings that underwent comparable demographic changes.

Nonetheless, it is important to stress that, as shown in the literature [42, 43, 74], the epidemiology of Varicella Zoster Virus is characterized by a strong heterogeneity across different countries. In particular, epidemiological evidences coming from France show that this country has been characterized by a stable incidence of varicella and HZ during the

period 2005-2015 [75]. In order to test the flexibility of the model to reproduce stable trends in HZ incidence over time, we applied the proposed model to the French population and epidemiological data. Obtained results show that, since the fertility transition in France was milder than in other European countries (e.g. Spain and Italy), the epidemiology of varicella and HZ has been much more stable during the last decades (see Appendix A.7). However, the availability of age-specific varicella incidence data highlights a potential limitation of the proposed model in reproducing the observed epidemiological patterns. In particular, available data show a progressive increase of varicella incidence among young children since the 90s along with a decrease of the number of cases in the older age groups, suggesting a decrease in the average age at varicella. One possible explanation for this specific pattern is the potential effect on social mixing caused by the increase in school enrollment rates occurred during the last century [76]. The assumption on mixing patterns used in this work does not allow capturing these changes. A possible future refinement of the model could include time varying matrices for contacts at school, based on temporal changes in enrollment rates.

Results obtained in this work suggest that retrospective evaluations of the effects of varicella immunization on HZ epidemiology might have been confounded by pre-existing trends in HZ epidemiology triggered by demographic changes occurred in the past [49, 77]. Model simulations reproduce this situation by showing that, in the years immediately following the introduction of vaccination, the variation in HZ trends owing to vaccination may be subtle and hard to distinguish from that occurring in the pre-vaccine era. Therefore, we suggest that the approach proposed here may help to disentangle the effect of natural epidemiological processes from that of vaccination, eventually allowing a more accurate assessment of past and future interventions.

Appendix A

A.1 The model

In this study we use a stochastic individual-based model for Varicella-Zoster Virus (VZV) transmission and reactivation. Full details are provided hereafter.

Description

The population is stratified by 101 one-year age classes, $a \in \{0, 1, 2, ..., 99, 100+\}$. The model accounts for the vital dynamics of the Spanish population, by considering yearly variations in birth rate, mortality rates and migration fluxes. All individuals are born susceptible to VZV and acquire varicella with a time-dependent, age-specific force of infection (FOI) defined as:

$$\lambda_j(t) = \beta \sum_k C_{jk}(t) y_k(t) \tag{A.1}$$

where

- $\lambda_j(t)$ is the FOI for age class j at time t;
- β is the VZV transmission rate;
- $C_{jk}(t)$ is the contact matrix at time t, defined as the number of contacts of an individual in the age group j with individuals in the age group k. More specifically, it is computed at each time as $C_{jk}(t) = \tilde{C}_{jk}\pi_k(t)$, where estimates for contacts provided in [65], $\tilde{C}_{jk}$, are rescaled by the current fraction of individuals in the age group k over the total population, $\pi_k(t)$.
- $y_k(t)$ is the fraction of varicella-infectious individuals within the age group k at time t;

This form of the FOI was previously proposed in the classical modelling literature for measles [78, 79] and theoretically analysed [24] in the context of demographic changes. In formulation (A.1), we neglect the contribution of Herpes Zoster (HZ) infected individuals to the FOI, based on the observation that HZ is less infectious and much less frequent than

varicella [80]. VZV infected individuals recover at a constant rate γ. Once recovered from varicella, individuals gain lifelong immunity to VZV reinfection and become susceptible to HZ. The risk of developing HZ is reduced by subsequent re-exposures to VZV infectious individuals, called boosting events, that occur at a rate proportional to the FOI through a coefficient $z \in [0,1]$. This coefficient represents the assumption that only a fraction z of the contacts that would result in varicella infection in susceptibles triggers a boosting of the immune response against HZ.
HZ susceptible individuals reactivate VZV, thus developing HZ, according to a risk that depends on the number of VZV exposure episodes i, on the age of the host a and on the time elapsed since last VZV exposure τ. In particular, we assume the same functional form as in [43]:

$$\rho_i(a,\tau) = \rho_0 q^{(i-1)^2} e^{\theta_a (a-a_0)^+} e^{\theta_\tau \tau} \tag{A.2}$$

where

- $\rho_i(a,\tau)$ is the VZV reactivation risk;
- the term $q^{(i-1)^2}$, where $q \in [0,1]$, accounts for the reduction of HZ risk when the number of VZV exposures increases ($i = 1$ represents primary varicella infection);
- the exponential term $e^{\theta_a (a-a_0)^+}$, where $\theta_a > 0$ and $(a-a_0)^+ := \max(0, a-a_0)$, accounts for the increase of HZ risk due to immunosenescence of the host [34]. We assume that aging begins to have an effect on the risk of reactivation starting from age $a_0 = 45$ years [41, 43];
- the exponential term $e^{\theta_\tau \tau}$, where $\theta_\tau > 0$, accounts for the increase of HZ risk as the time elapsed since last VZV exposure episode increases;
- ρ_0 is the risk of developing HZ for individuals younger than 45 years of age ($a < a_0$), who have just recovered from varicella infection ($\tau = 0$, $i = 1$).

Individuals who have developed HZ are assumed to become permanently immune to VZV reactivation, based on the fact that recurrent HZ is relatively uncommon in immunocompetent persons [31], as detailed in the Methods section in the main text.
A schematic representation of the model is given in Figure A.1.

The model requires the tracking of HZ susceptibility states by age, time since last exposure to VZV and cumulative number of re-exposures. Considering 100 one-year age groups, a mean number of lifetime re-exposures of about 7 (see Section A.3) and a yearly discretization for the time since last exposure, the number of possible states for individuals susceptible to HZ amounts to at least $7 \times 100 \times 100/2 = 35,000$ (division by 2 accounts for the time since last exposure being always smaller than chronological age), which possibly underestimates the actual number of states, since some individuals may experience more than 7 re-exposure

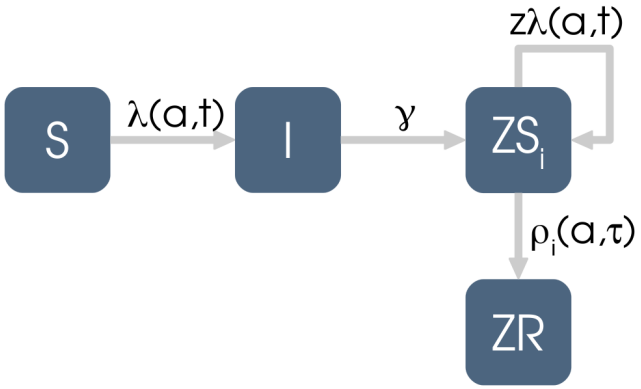

Figure A.1: **Flow diagram of the model. Varicella and HZ.** S: VZV susceptible individuals; I: varicella infectious individuals; ZS_i: HZ susceptible individuals with i episodes of previous exposure to VZV; ZR: individuals who have developed HZ. $\lambda(a,t)$: time and age-dependent force of infection; γ: constant recovery rate from varicella; z: boosting efficacy and $\rho_i(a,\tau)$: VZV reactivation rate.

episodes. In this study, the considered population size replicates the one on which the HZ calibration dataset was collected (the region of Navarra, about 600,000 individuals in 2009). Comparing population size with the number of HZ susceptibility states, it results that the mean population in each state is smaller than $600,000/35,000 \approx 17$. This number gets even smaller if we consider the possible underestimate of the number of HZ susceptibility states, that the model's population at the beginning of the century was about 1/3 of the current one, that not all individuals are HZ susceptibles and that the distribution of the population across model states is not homogeneous. Deterministic models assume a continuous variable for the population size over the various structural dimensions (represented here by chronological age, age since last exposure, and number of exposures): such assumption is only valid for a sufficiently large number of individuals in each compartment. For these reasons we preferred a stochastic individual-based model over a deterministic one.

Simulation algorithm

For each individual we take into account:

- age a, measured in years, $a \in \{0, ..., 100\}$;

- epidemiological status:$=\begin{cases} \text{Susceptible to VZV} \\ \text{Infected by varicella} \\ \text{HZ susceptible} \\ \text{Infected or recovered from HZ} \end{cases}$
- number of VZV exposures $\begin{cases} i = 0 & \text{Varicella susceptible individual} \\ i = 1 & \text{Primary varicella infection} \\ i > 1 & \text{Subsequent boosting episodes} \end{cases}$
- a^*, age of the individual at the time of last VZV exposure occurrence.

Time is discretized using a step of 1 week. At each time step the following events occur:

- Recoveries: the number of recoveries is obtained by sampling a Poisson distribution with parameter γI, where γ is the recovery rate and I is the current total number of infectious individuals.
- Primary varicella infections: the number of varicella infections is obtained, for each age group j, by sampling a Poisson distribution with parameter $\lambda_j S_j$ where λ_j is defined in Eq. (A.1) and S_j is the current total number of susceptibles in age group j. For each infected individual, i is set to 1 and a^* is updated with his current age.
- Boosting events: the number of boosting events is obtained, for each age group j, by sampling a Poisson distribution with parameter $z\lambda_j ZS_j$, where z is the parameter which accounts for boosting efficacy, λ_j is defined in Eq. (A.1) and ZS_j is the current total number of HZ susceptibles in age group j. For each boosted individual, i is incremented by 1 and a^* is updated with his current age.
- Births: the number of births is obtained by sampling a Poisson distribution with parameter $bN/52$, where b, N denote respectively the crude birth rate for the current year and the current total population.

The following events occur at a yearly frequency (at the end of each year):

- VZV reactivations: for each HZ susceptible individual with i VZV exposure episodes, age a, and time since last exposure $\tau = a - a^*$, VZV reactivation may occur according to a Bernoulli sample with probability $\rho_i(a, \tau)$, defined in Eq. (A.2).
- Deaths: for each individual of age a death may occur according to a Bernoulli sample with probability $\mu(a)$, representing age-specific mortality rate for the current year. $\mu(a)$ is obtained as the weighted mean of male- and female-specific mortality rates for age a, where the weight is the sex ratio. item Migrations: a number of individuals is added or removed from the population based on the relative age-specific migration

flux for the current year. In the case of a positive migration rate, the epidemiological status of added individuals is assigned according to the current epidemiological profile in the resident population.

- Aging: at the end of each year the age of all individuals is incremented by 1, with the exception of those belonging to the age class $a = 100$.

Model inputs

The model takes as input the following data:

- yearly birth rates [58] for the time period considered in the simulation (shown in Figure 1.1);
- annual sex- and age-specific mortality rates [59] for the considered period (Figure A.2**a-b**);
- yearly migration fluxes [61, 62] (Figure A.2**c**);
- age structure of the migrant population [60];
- age-specific contact matrix estimated for Spain in [65].

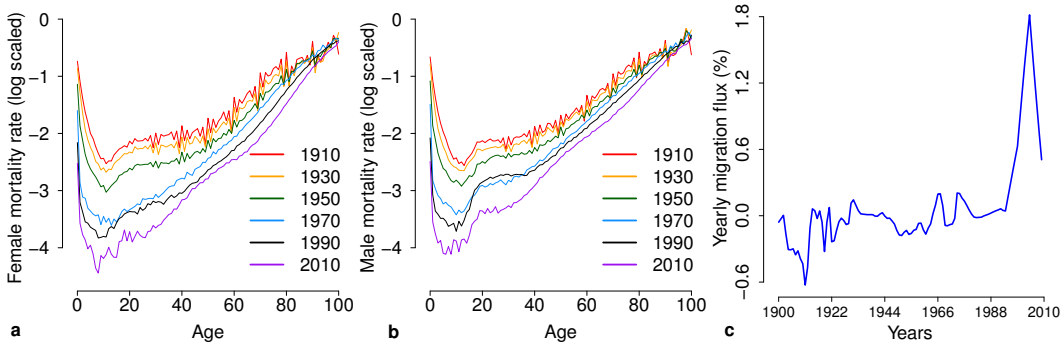

Figure A.2: **Historical demographic data.** **a** Age-specific mortality rate, females [59]. **b** Age-specific mortality rate, males [59]. **c** Yearly migration flux in % of the population [61, 62]

A.2 Model calibration

The model has six free parameters:

- the VZV transmission rate β;

- four parameters defining the VZV reactivation rate: $\rho_0, \theta_a, \theta_\tau, q$ (see Eq. (A.2));
- the boosting efficacy z.

The parameter vector is defined as $\Theta = (\beta, \rho_0, \theta_a, \theta_\tau, q, z)$. The posterior distribution of parameters was explored by using Markov Chain Monte Carlo (MCMC) applied to the likelihood of observed data, using uniform prior distributions. The likelihood is defined as the product between the binomial likelihood $\mathcal{L}_A$ of the observed VZV serological profile by age in 1996 [66] and the Poisson likelihood $\mathcal{L}_B$ of the observed HZ incidence profile by age in 2005-2006 [67]. Specifically,

$$\mathcal{L}_A(n_m, r_m|\Theta) = \prod_{m=1}^{M} \frac{n_m!}{r_m!(n_m - r_m)!} (p_m(\Theta))^{r_m} (1 - p_m(\Theta))^{n_m - r_m} \tag{A.3}$$

where

- M is the number of ages considered in the serological profile [66];
- n_m is the number of individuals of age m observed in [66];
- r_m is the number of seropositive individuals of age m observed in [66];
- $p_m(\Theta)$ is the VZV seroprevalence for age m, simulated by the model in 1996 with parameter set Θ.

and

$$\mathcal{L}_B(k_j|\Theta) = \prod_{j \in J} \frac{e^{-\eta_j(\Theta)} (\eta_j(\Theta))^{k_j}}{k_j!} \tag{A.4}$$

where

- J is the set of age groups in the HZ incidence profile [67];
- k_j is the total number of HZ cases in the j-th age group observed in [67];
- $\eta_j(\Theta)$ is the number of HZ cases in the age group j, simulated by the model in 2005-2006 in a population of the same size as that analyzed in [67], with parameter set Θ.

We determined the posterior distribution of Θ using random-walk Metropolis-Hastings sampling [81]. At each iteration, the algorithm evaluates the likelihood of a new candidate vector of parameters, that is accepted or not based on the standard Metropolis-Hastings algorithm. The values of a new candidate parameter vector are randomly sampled from normal distributions having mean equal to the current values and standard deviations given in Table 1.1. The algorithm is run for 22,000 iterations and a burn-in period of 2,000 iterations is considered. Convergence is checked by considering several different starting points

Table 1.1: **Standard deviations for parameters.**

Parameter	Standard Deviation	Range
β	$4.3 \cdot 10^{-2}$	$[0, +\infty)$
ρ_0	$7.0 \cdot 10^{-5}$	$[0, +\infty)$
θ_a	$8.8 \cdot 10^{-3}$	$[0, +\infty)$
θ_τ	$4.3 \cdot 10^{-3}$	$[0, +\infty)$
q	$6.3 \cdot 10^{-2}$	$[0, 1]$
z	$2.4 \cdot 10^{-2}$	$[0, 1]$

and by visual inspection. Estimated mean values and 95%CI of parameters are reported in Table 1.2. The reported mean values and credible intervals are computed from the selected realizations of the model, and thus account for both the stochasticity of model realizations and the uncertainty in model parameters estimates.

Table 1.2: **Estimates of parameters obtained for Spain.**

Parameter	Description	Mean	Unit	95% CI
β	VZV transmission rate	1.02×10^5	weeks^{-1}	$0.95 \times 10^5, 1.10 \times 10^5$
ρ_0	HZ risk at birth [1]	2.63×10^{-3}	years^{-1}	1.96×10^{-3} , 3.21×10^{-3}
θ_a	Rate of risk exponential growth with chronological age a [1]	11.19×10^{-2}	years^{-1}	7.53×10^{-2} , 14.53×10^{-2}
θ_τ	Rate of risk exponential growth with time since last exposure τ [1]	4.77×10^{-2}	years^{-1}	3.89×10^{-2} , 5.81×10^{-2}
q	Parameter regulating HZ risk reduction with increasing number of re-exposures to VZV [1]	0.64		0.46 , 0.81
z	Boosting efficacy	0.73		0.64 , 0.83

[1] $\rho_i(a, \tau) = \rho_0 q^{(i-1)^2} e^{\theta_a (a - a_0)^+} e^{\theta_\tau \tau}$

A.3 Additional results

Historical period (1900-2009)

The model is able to correctly reproduce the observed evolution of the population age structures [58], as shown in Figure A.3. The curves show downward peaks, reflecting the brisk drop of birth rate (Figure 1.1) during the Spanish Civil War (1936-1939).
The remarkable changes in the population age structure shown above have led, according to the model, to a reduction of VZV circulation. This has produced a reduction over time of the estimated mean number of VZV re-exposures collected by HZ susceptibles (Figure A.4), that has led to the increase in HZ incidence (Figure 1.3**c**).

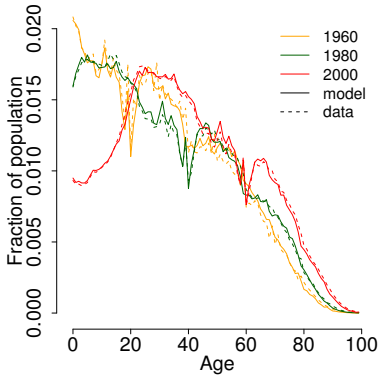

Figure A.3: Age distribution of the population at different years as observed [58] (dotted lines) and as predicted by the model (solid lines).

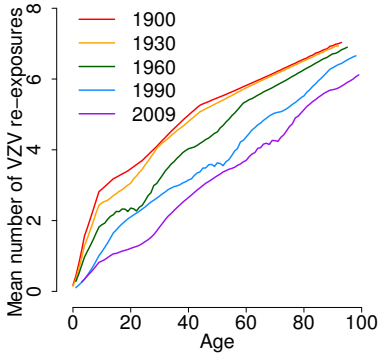

Figure A.4: Estimated mean number of VZV re-exposures by age at different years.

Prediction scenarios (2010-2050)

In this section we discuss further results obtained for the illustrative scenarios presented in the main text for Spain and results obtained for an official demographic projection scenario based on the 2012 UN population prospects [63].

Illustrative scenarios

The predicted evolution of the age structure of the Spanish population in the “Lowest-birth rate" (L, Figure 1.1) and“Highest-birth rate" (H, Figure 1.1) scenarios are shown in Figures A.5**a** and A.5**c**. Scenario L is characterized by a remarkable aging of the population, whereas in scenario H the fraction of young individuals in the population is predicted to

increase. Such changes lead respectively to a decrease (scenario L) and to an increase (scenario H) of VZV circulation, that influences the age-specific number of boosting events within the population. In particular, as shown in Figure A.5**b**, in scenario L the mean number of boosting events by age is predicted to decrease during the period 2010-2050 and this explains the predicted strong increase in HZ (see Figure 1.5**c** in the main text). On the other hand, as shown in Figure A.5**d**, the increasing VZV circulation in scenario H leads to an increase of the mean number of VZV re-exposures for young people and adults, which partially mitigates the HZ increase caused by demographic changes occurred during the last century (see Figure 1.5**d**).

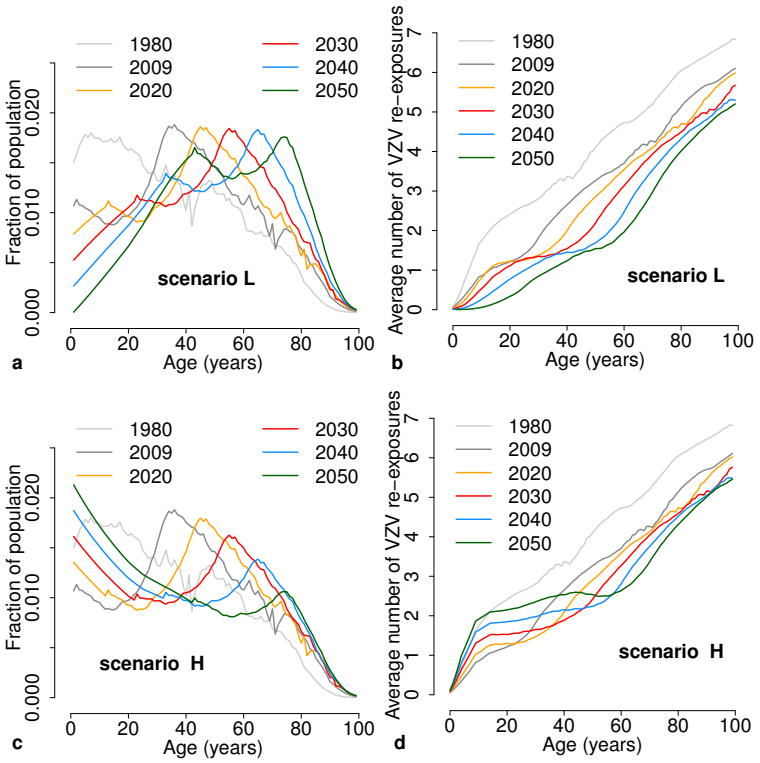

Figure A.5: **a** Age distribution of the population at different years in scenario L. **b** Mean number of VZV re-exposures by age at different years in scenario L. **c** As **a** but for scenario H. **d** As **b** but for scenario H.

Scenario based on UN demographic projections

In addition to the illustrative scenarios proposed in the main text we considered an official demographic projection scenario for Spain, based on the medium variant of the 2012 UN

population prospects [63]. This scenario predicts an initial decrease followed by a recovery of the birth rate between 2010 and 2050 (see Figure 1.1) and a progressive decrease of migration fluxes and mortality rates, especially in the elderly, during the same period.
As shown in Figure A.6, results obtained for this scenario are quite similar to the baseline, which assumes constant birth, mortality and migration rates for the whole prediction period. This is because the variation in the birth rate predicted during the period 2010-2050 by the realistic scenario considered is negligible when compared to that experienced during the last century and HZ dynamics are much less sensitive to changes in migration fluxes and mortality rates in adults and older individual, as they do not impact significantly on age groups involved in varicella transmission.
In summary, even under realistic demographic projections, an increase of HZ incidence should be expected in Spain as a consequence of the dramatic demographic changes that characterized the last century.

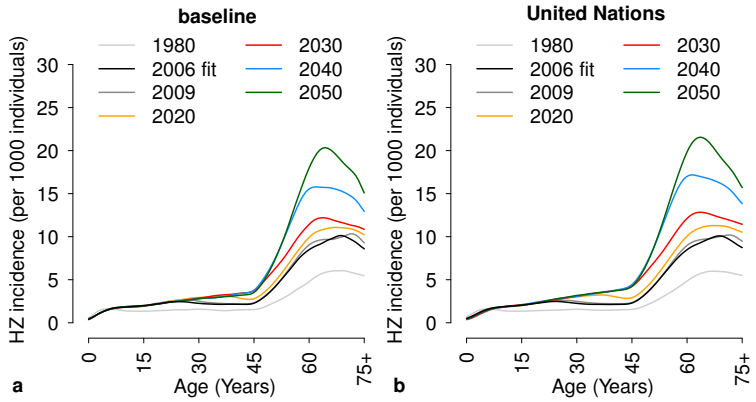

Figure A.6: **Predicted HZ incidence in Spain (2010-2050). Comparison between baseline and realistic scenario.** **a** Age-specific HZ incidence predicted by the model at different years for the baseline scenario. **b** As **a** but fot the demographic scenario based on UN population prospects [63].

A.4 Vaccination

In addition to results provided in the main text, we report in Figure A.7 the total incidence of varicella and HZ over time under a vaccination program starting in 2010 and targeting children of 15 months of age, considering different scenarios for the combined value of coverage and efficacy: 80%, 90% and 100%. Figure A.7 shows that coverage does not significantly affect the HZ incidence; however coverage at 80% is not sufficient to provide elimination of varicella in the long term.

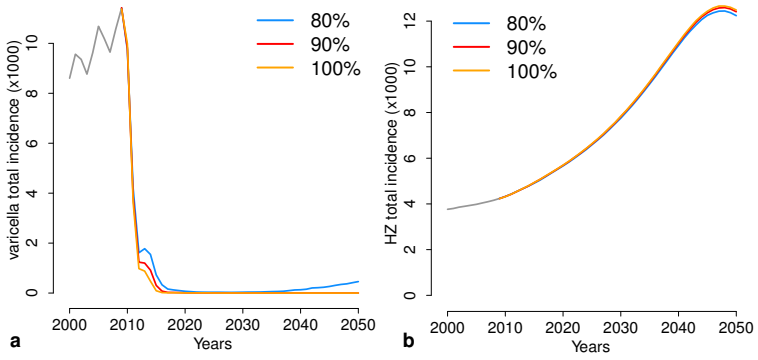

Figure A.7: **Model predictions under a vaccination program starting in 2010 and for different scenarios for the combined value of vaccine efficacy and coverage.** **a** Total varicella incidence over time. **b** Total HZ incidence over time.

A.5 Homogeneous mixing

To provide a comparison of model results under a different assumption on mixing patterns, we recalibrated the model by assuming homogeneous mixing (i. e. $C_{jk}(t) = 1$). We show that the model behavior is qualitatively similar to that reported in the main text, with a reduction over the past century of VZV circulation (Figure A.8**a**) and a growth in the mean age at varicella (A.8**b**), total and age-specific HZ incidence (A.8**c** and inset) and mean age at VZV reactivation (A.8**d**). In particular, under the homogeneous mixing assumption, the predicted growth of the total HZ incidence between 1997 and 2004 is higher (about 18%) than that predicted using age-specific contact matrices (about 12%). We conclude that the HZ growth predicted as a consequence of demographic changes is robust under different assumptions on mixing patterns.

A.6 A different formulation of exogenous boosting

The model was also recalibrated using a different formulation of the exogenous boosting hypothesis. In this alternative formulation, the mechanism of VZV reactivation is modeled as in [42, 47, 48]. In particular, once recovered from varicella individuals acquire lifelong immunity to varicella and temporary immunity to HZ. After an average time of $1/\delta$ the protection against HZ wanes and individuals become susceptible to VZV reactivation. HZ susceptibles may either be boosted or develop HZ. The risk of VZV reactivation has the following functional form [47]:

$$\rho(a) = \omega e^{-\phi a} + a^{\eta}\pi \qquad \omega, \phi, \eta, \pi > 0 \tag{A.5}$$

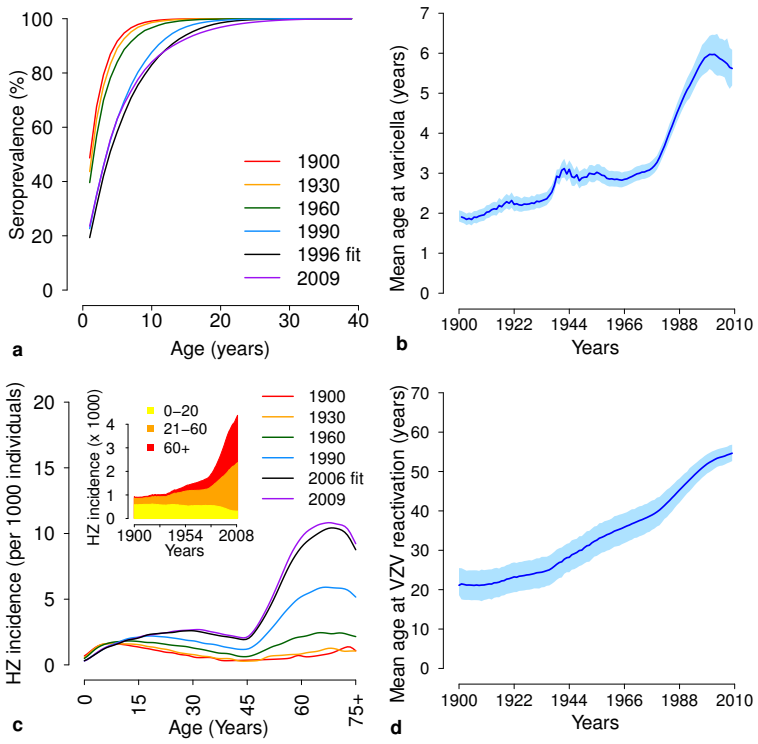

Figure A.8: **Estimated impact of demographic changes on VZV epidemiology by assuming homogeneous mixing. Historical period (1900-2009).** **a** Estimated age-specific VZV seroprevalence at different years. **b** Estimated mean age (and 95%CI, shaded areas) at varicella over time. **c** Estimated age-specific HZ incidence at different years. The inset shows the total HZ incidence over time and disaggregated by age group. **d** Estimated mean age (and 95%CI, shaded areas) at HZ over time.

The model is schematically described in Figure A.9.
Results obtained are, again, qualitatively analogous to those shown in the main text, with an increasing trend in the estimated total HZ incidence (Figure A.10**a**) over the last century. In this case, the estimated growth is smaller than that predicted by the progressive immunity model (reproposed here for convenience, Figure A.10**b**). We conclude, that the HZ growth predicted as a consequence of demographic changes is robust under different model formulations of exogenous boosting.

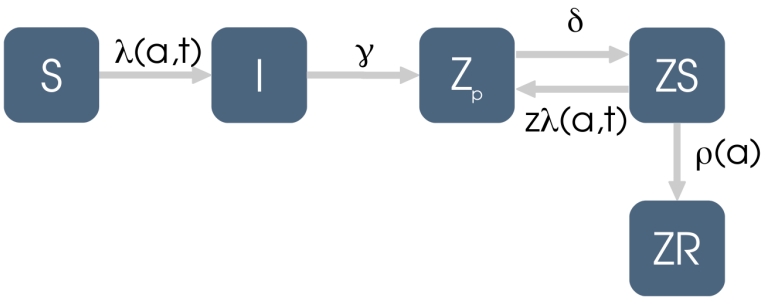

Figure A.9: **Flow diagram of the model based on [47].** S: VZV susceptible individuals; I: varicella infectious individuals; Z_p: individuals recovered from varicella and temporary protected against VZV reactivation; ZS: HZ susceptibles individuals; ZR: individuals who have developed HZ. $\lambda(a,t)$: time-dependent force of infection, γ: constant recovery rate from varicella, z: boosting efficacy, $1/\delta$: duration of immunity to HZ after exposure to varicella and $\rho(a)$: VZV reactivation rate.

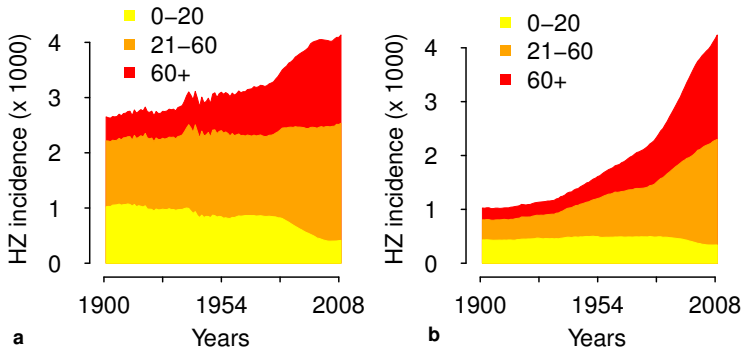

Figure A.10: **Total HZ incidence and disaggregated by age group. a** As predicted by the alternative model formulation based on a mechanism of temporary immunity. **b** As predicted by progressive immunity model.

A.7 The case of France

Methods

We applied the model presented in Chapter 1 to investigate the epidemiology of varicella and HZ in France during the period 1850-2015.

Model inputs

The model take as input the following data:

- yearly birth rates [59] for the time period considered in the simulation (Figure A.11**a**);
- annual sex- and age-specific mortality rates [59] for the considered period (e.g. Figure A.11**b**);
- yearly migration fluxes [82] (Figure A.11**c**);
- age structure of the migrant population [60];
- age-specific contact matrix estimated for France in [83].

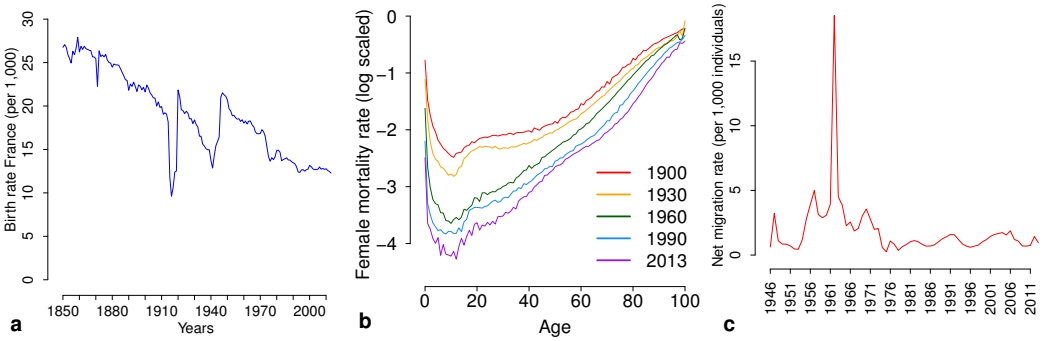

Figure A.11: **Historical demographic data.** **a** Yearly birth rate [59] **b** Age-specific mortality rate, females [59]. **c** Yearly migration flux in % of the population [82]

Data used for calibration

Model calibration was performed analogously to the case of Spain, by using age-specific VZV seroprevalence data in 2003, based on [84], and age-specific data on HZ incidence in 2015 [75] for France. Model estimates were qualitatively compared against age-specific HZ incidence data between 2005-2015 and age-specific varicella incidence data between 1991-2015, both provided by [75].

Results

The model accurately reproduces the observed changes in the age distribution of the France between 1850-2015 [59], see Figure A.12. The model is able to reproduce the age-specific VZV seroprevalence observed in France in 2003 (Figure A.13**a**); the 95% CI of model estimates (grey area) overlap for almost all age groups with the data (red lines). Nonetheless,

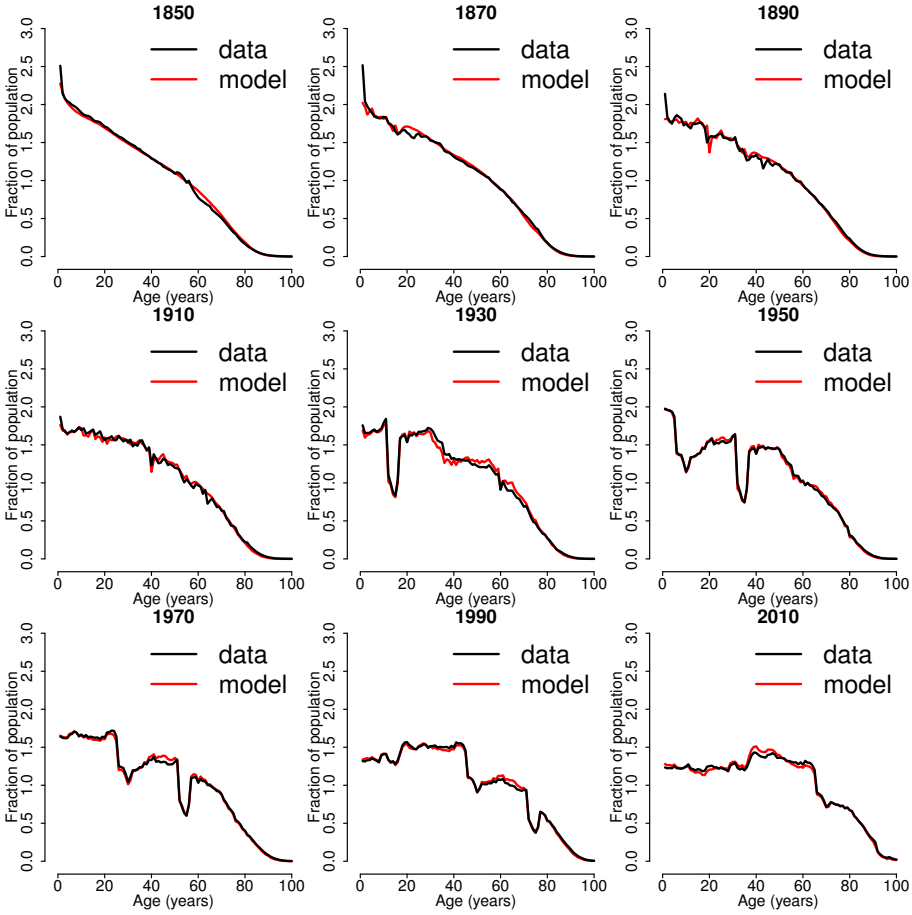

Figure A.12: Age distribution of the French population at different years as observed [59] (black) and as predicted by the model (red)

it partially fails in reproducing the sudden increase of the seroprevalence observed between 3 and 4 years of age, which is likely ascribable to the pre-primary school enrollment. Figure A.13**b** show the capability of the model to reproduce the age-specific HZ incidence observed in 2015, including the decreasing incidence in the elderly, which is a feature of both French and Spanish HZ incidence data.

Historical period (1850-2015)

Historical demographic data reported in Figure A.11 show that during the last century France experienced a decrease in the birth rate less remarkable than the one observed in Spain (-39% vs -73% decrease between 1900 and 2000). According to model estimates, demographic changes occurred in France resulted in an relatively stable VZV circulation,

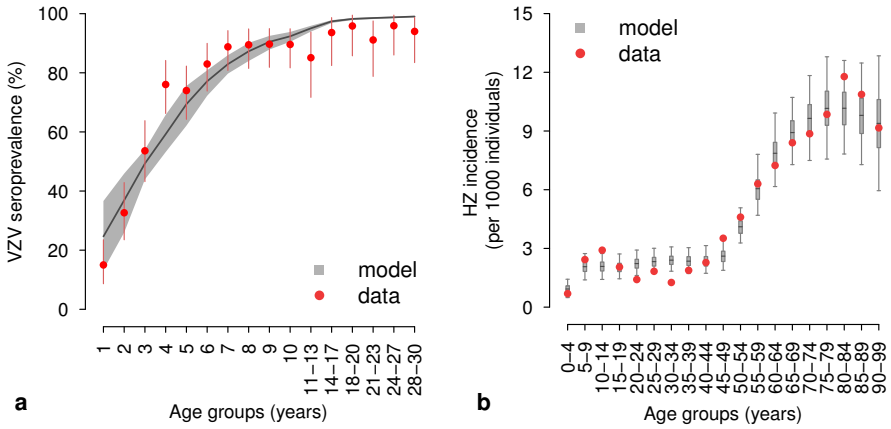

Figure A.13: **VZV seroprevalence and HZ incidence. a** Age-specific VZV seroprevalence as observed in data from France in 2003 [84] and as estimated by the model. Vertical lines represent 95% CI of the data computed by exact binomial test. **b** HZ incidence by age group as observed in 2015 [75] and as estimated by the model. In both panels, grey areas show the 95% CI of model estimates.

as shown by the estimated evolution of the serological profile over time (Figure A.14**a**) and by the more contained increase in the mean age at varicella compared to the case of Spain (Figure A.14**b**). Moreover, since the frequency of boosting events does not substantially change over time, the estimated increase in HZ incidence is mild (see Figure A.14 **c**/**d**).

Epidemiological validation

The French Sentinelles Network provides estimates for the incidence of varicella between 1991-2015 and the incidence of HZ between 2005-2015. We use this data to validate modeling results obtained during the corresponding periods.
The model estimates a stable overall and age-specific HZ incidence during 2005-2015, which is qualitatively and quantitatively compliant with what observed in the data (see Figure A.15**a** and Figure A.16). The total incidence of varicella observed in France between 1991 and 2015 show seasonal waves, however, its mean is qualitatively stable (red, Figure A.15**b**). The predicted total incidence of varicella during 1991-2015, slightly overestimate the observed one, though the overall temporal trend is rather consistent with data (blue, Figure A.15**b**). However, as shown in Figure A.17 the model fails to capture the age-specific varicella incidence observed between 1991 and 2015, especially the increasing number of cases in early childhood (1-3 years of age), where most of varicella occurs.

The analysis performed shows that in the case of France demographic changes are

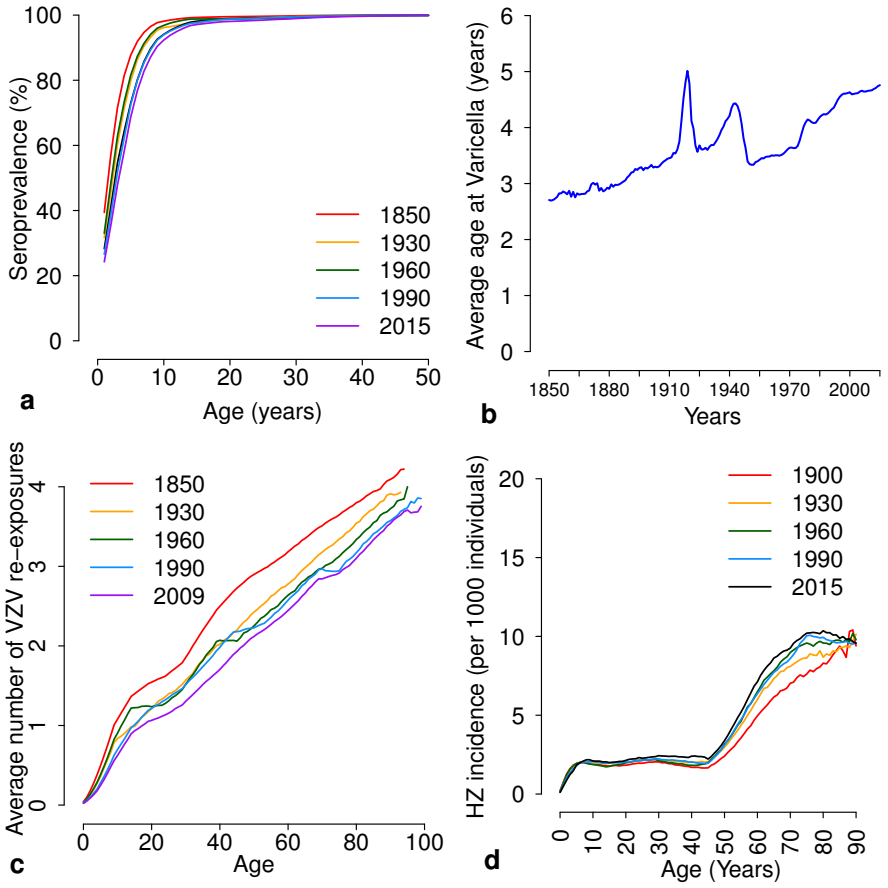

Figure A.14: **Estimated impact of demographic changes on VZV epidemiology in France (1850-2015) a** Estimated age-specific VZV seroprevalence at different years. **b** Estimated mean age at varicella over time. **c** Estimated mean number of VZV re-exposures by age at different years in scenario L. **d** Estimated age-specific HZ incidence at different years.

sufficient to explain the temporal trends of overall varicella and HZ incidence. In particular, the country was characterized by a stable incidence of varicella and HZ during the last decades. However, data highlight that a stable incidence at the population level can mask changes in the distribution of cases among different age groups, as is the case of varicella in France. This suggest the presence of other factors, beyond demographic changes, that can influence the dynamics of VZV. One possible crucial factor shaping varicella incidence over time may be represented by changes in enrollment rates of pre-primary school and nursery, which likely occurred during the last decades. Indeed, school attendance play a relevant

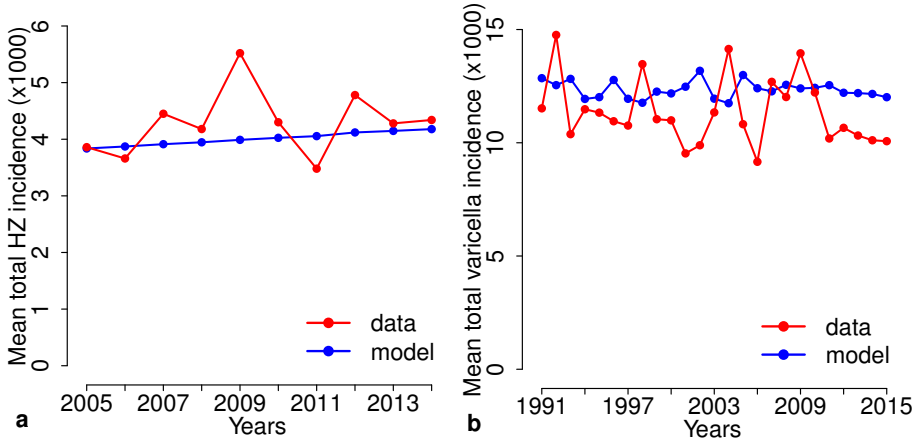

Figure A.15: **Total varicella and HZ incidence in France.** **a** Total HZ incidence (per 1,000 individuals) as observed in the data [75] (red) and as estimated by the model (blue) over the period 2005-2015. **b** Total varicella incidence (per 1,000 individuals) as observed in the data [75] (red) and as estimated by the model (blue) over the period 1991-2015.

role in the spread of childhood diseases, affecting contacts among children, which represent the greatest fraction of susceptibles in the population.
A possible refinement of the model may include information on changing enrollment rates by rescaling the contacts occurring at school.

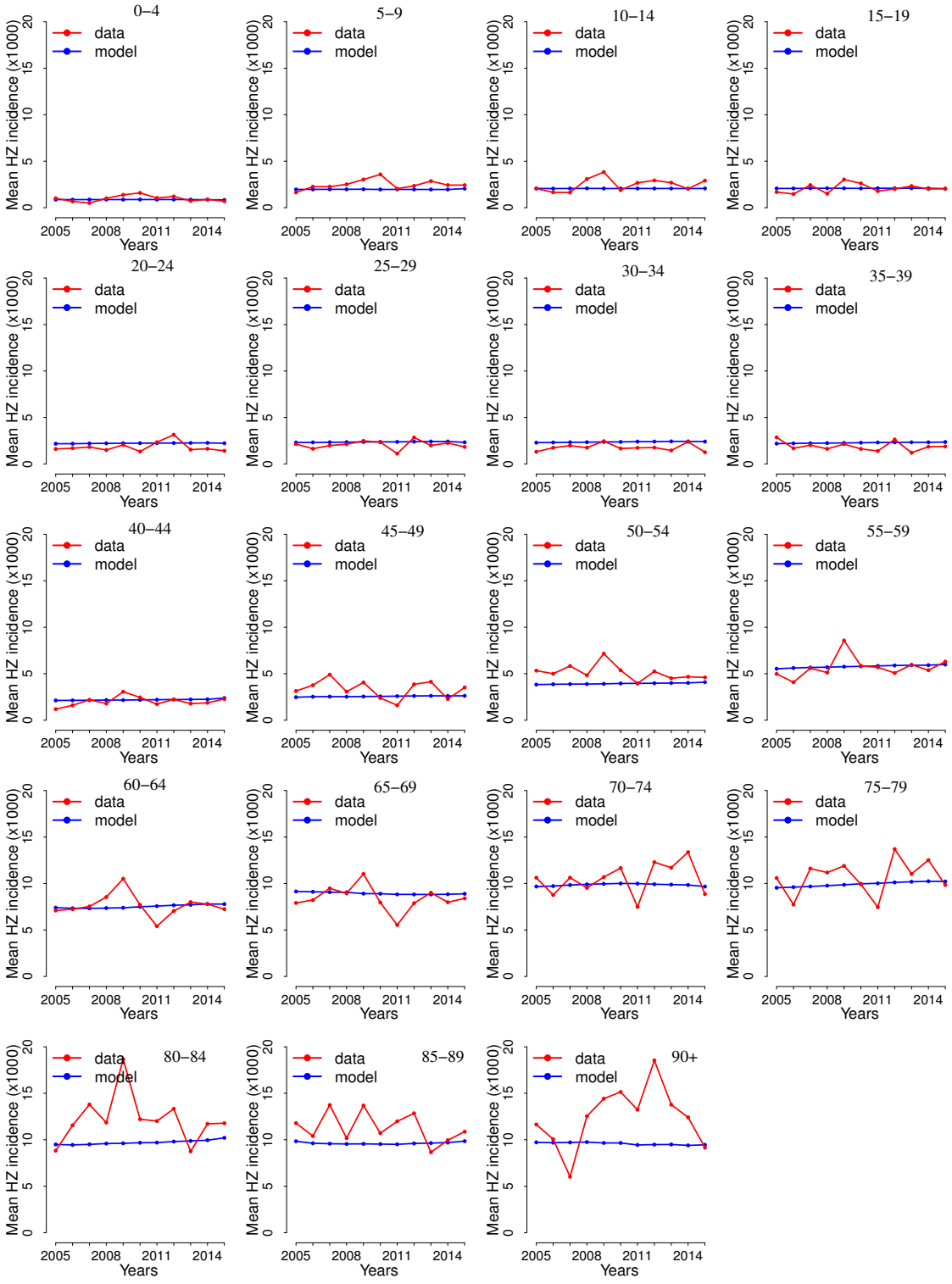

Figure A.16: Age-specific HZ incidence (per 1,000 individuals) as observed in the data [75] (red) and as estimated by the model (blue) over the period 2005-2015.

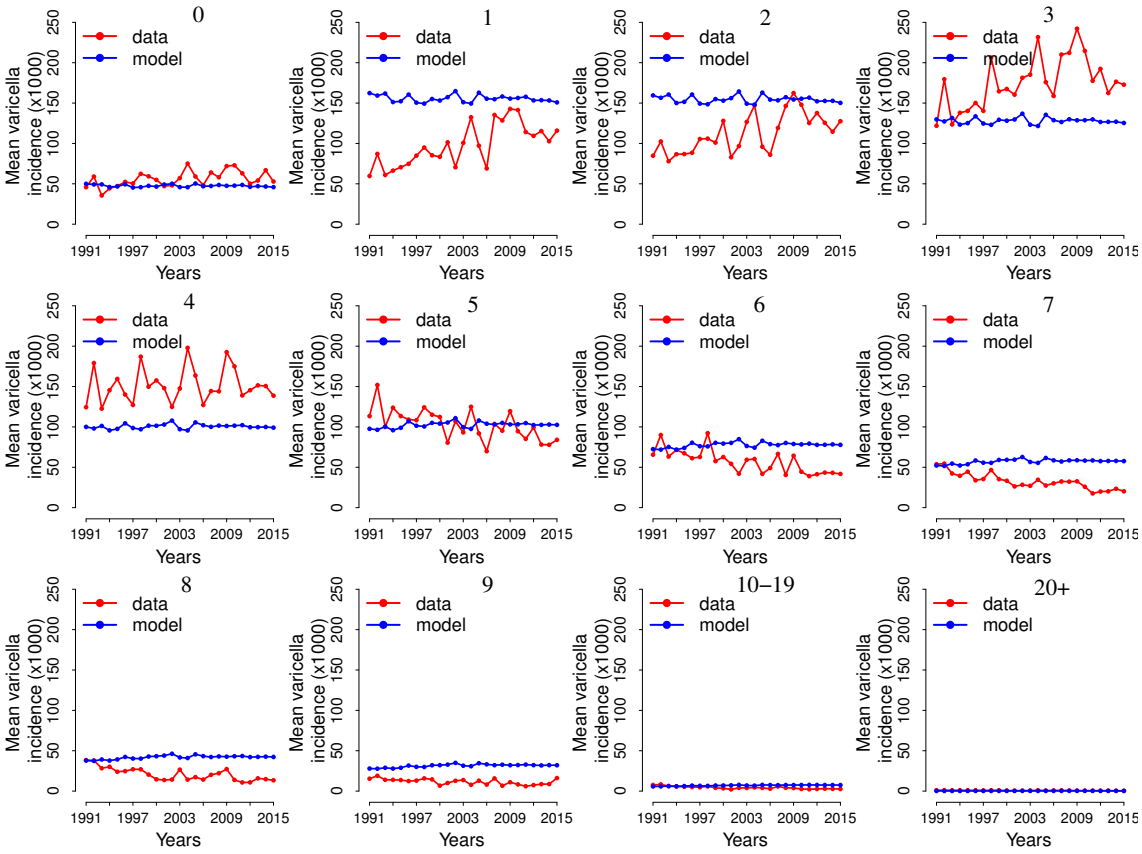

Figure A.17: Age-specific varicella incidence (per 1,000 individuals) as observed in the data [75] (red) and as estimated by the model (blue) over the period 1991-2015.

Chapter 2

A dynamic cost-effectiveness analysis of varicella and Herpes Zoster vaccination programs

2.1 Background

Varicella zoster virus (VZV) is a DNA virus belonging to the Herpesviridae family that affects only humans. Infection by VZV could result in two distinct diseases: varicella, which is a highly communicable and widespread childhood disease, and Herpes Zoster (HZ), caused by reactivation of VZV that remains latent in the dorsal root ganglia after primary varicella infection. Exposures to varicella-infected individuals (also called "boosting events") are believed to boost the VZV-specific immune response thereby inhibiting VZV reactivation [32, 35, 36, 37, 38]. At present, chickenpox remains the most frequent notified disease in Italy affecting almost the entire birth cohort and continuing to show an endemic-epidemic pattern. Although it is usually a mild disease with a relative low percentage of complications, especially in immunocompetent children, varicella is highly contagious and may develop more severe consequences and disabling symptoms in adults. Reactivation of the virus is more likely to occur in immunocompromised patients or the elderly. HZ is characterized by a vesicular eruption along the course of the nerve and is often associated with pain. The most common complication of HZ is the postherpetic neuralgia (PHN), a persistent and intractable chronic pain that occurs in 5 to 30% of HZ cases [74, 85]. Opposite to varicella, HZ is not a notifiable disease in Italy.
A live attenuated vaccine against varicella was developed in 1974 and it has been introduced in some countries starting from 1995 [86]. However, in many developed countries its introduction into the national schedule still represents a matter of discussion. First, VZV mass immunization may potentially increase the incidence of more severe varicella cases among adults [87]. Another main concern is that the lack of boosting events caused by the reduction of varicella circulation may lead to an increase of HZ incidence in the post-vaccination period. Results coming from surveillance programs of HZ in countries that have introduced VZV mass immunization do not provide univocal evidences. Some countries detected an

increase of HZ following mass varicella immunization and others did not observe any effect on HZ incidence [49, 50]. Mathematical models for VZV unanimously predict an increase of HZ incidence as a consequence of the reduction of boosting associated to varicella vaccination, however the magnitude of such increase may depend on modeling assumptions on the mechanism of VZV reactivation [41, 42, 47, 48], whose biology has not been elucidated yet [35, 36, 37, 39, 40, 71]. Moreover, as shown in Chapter 1, an increase in HZ incidence should be expected in the next decades as a consequence of past demographic changes and their effect may be hard to disentangle from that of varicella immunization.
Finally, in more recent years, a vaccine against HZ has been licensed [88] and it has been recommended in some countries both in combination with varicella vaccination (e.g. US) and alone (e.g. France, UK), yielding the evaluation of post-vaccination trends even more complex.
Cost-effectiveness analyses is nowadays considered an essential tool to be used to allocate scarce public health resources. Through these techniques we are able to estimate the net health benefit and costs associated with a given intervention and taking into consideration all positive and negative, direct and indirect, effects that result from the introduction of a program. In particular, the economic evaluation of a number of vaccination strategies has been used in the past to help inform public health decisions for a variety of infectious diseases such as meningo C and B [8, 9], pneumococcal disease [10], influenza [11, 12] and varicella [13, 14, 15, 16]. Modelling and economic analyses on varicella vaccination, in particular, has been conducted for a number of countries and results appear to be hindered by the assumptions made on the existence of the boosting mechanism [89]. More specifically, varicella vaccination results economically acceptable when the model does not consider the indirect effect of the vaccine on the epidemiology of zoster. When, conversely, when the model takes into account the potential increase of zoster incidence as a consequence of varicella vaccination, the cost-effectiveness of the program becomes questionable [89].
In this study we use the individual-based stochastic model introduced in Chapter 1 to evaluate the joint impact of demographic changes and different varicella and HZ vaccination strategies on the future epidemiology of varicella and HZ in Italy. In particular, in order to take into account the uncertainty regarding the mechanism of VZV reactivation, we consider two different modeling assumptions on exogenous boosting. The first one has been more traditionally used in the literature [47], whereas the second one more close to the biology has been shown to fit better the age profile for HZ incidence in several countires, including Italy [43]. The final aim of this work is to provide a picture of the effectiveness and cost-effectiveness of different varicella and HZ vaccination strategies, including both the uncertainty regarding the biological mechanism of boosting, the one regarding the economic parameters and the underlying effect of demographic changes.

2.2 Methods

In this study we use a stochastic individual-based model for varicella and HZ epidemiology, adapted from the previous one developed for Spain [90]. We investigate the effect of demographic changes on the past and future epidemiology of varicella and HZ in Italy and perform a cost-effectiveness analysis of varicella and HZ vaccination, taken individually or combined.

2.2.1 Epidemiological and demographic data

Demographic changes for the period 1900-2014 are modelled using data on yearly variations of birth rate, age-specific mortality rates and net migration fluxes over time specific for Italy [59, 91]. Demographic changes during the prediction period (2015-2100) are modeled using projections on births, mortality and migration available from the 2015 Revision of the World Population prospects [92]. As in [90], we assume heterogeneous mixing patterns by age according to the synthetic contact matrices estimated for Italy in [65]. Epidemiological parameters are estimated using data on the Italian VZV seroprevalence profile observed in 1996 [66] and age-specific data on HZ incidence in 2004 [72].

2.2.2 The model

Natural history of varicella and HZ

Two different modeling formulations on the effect of boosting are considered. The first one is based on [32, 43] and will be referred in the rest of the chapter as *"Progressive Immunity"* (PI). The second one, which was introduced in [47], will be denoted as *"Temporary Immunity"* (TI).
Both models have been already introduced in Chapter 1, to which we refer for all details. Briefly, in both models maternal antibodies confer protection against varicella infection to new-borns for 6 months on average, after which children become susceptible to VZV natural (i.e. wild-type) infection (S). Susceptible individuals are exposed to a time- and age-dependent force of infection. Varicella infected individuals (I) recover after 3 weeks on average and acquire life-long immunity against varicella.
In Model PI, after recovery from varicella individuals become susceptible to HZ (ZS). The rate of VZV reactivation depends on: i) the individual chronological age, ii) the number of re-exposures to VZV already experienced and iii) the time elapsed since last exposure to VZV (see Equation A.2 and Figure A.1). According to this formulation, multiple exposures to VZV-infectious individuals lead to a progressive decrease of the HZ risk [32, 43].
In Model TI, individuals recovered from varicella acquire temporary protection against HZ (Z_p). After an average time of $\frac{1}{\delta}$ the acquired immunity wanes and individuals become susceptible to HZ (ZS). The VZV reactivation rate depends on the individual chronological age (see Equation A.5 and Figure A.9).

In both models, HZ susceptibles may either be boosted or develop HZ and, in the latter case, they become lifelong immune to HZ (ZR).

Varicella and HZ vaccination

The varicella vaccine is administered with a coverage c, starting from the year 2015. Vaccine failure is assumed to occur at vaccine administration with a probability $(1-p)$, where p is the probability to be immunized for targeted susceptible individuals. As a consequence of vaccine failure, individuals are still susceptible to VZV (S*) and they can experience a milder varicella infection (the so-called breakthrough varicella). Individuals infected with breakthrough varicella may transmit the disease, nevertheless they are assumed to be half contagious as natural cases [93].
Analogously to the case of natural varicella, under Model PI individuals recovered from breakthrough varicella become susceptible to HZ (ZS*), whereas under Model TI they become temporarily immune against HZ (Z_p^*). In both models, successfully immunized individuals are assumed to gain lifelong protection against VZV infection (VP), though they can develop HZ from the vaccine strain. Vaccinated individuals (VP), including those that have experienced vaccine failure (ZS*), experience VZV reactivation at a lower rate compared to unvaccinated. In particular, the VZV reactivation rate for vaccinated individuals is rescaled by a factor χ, whose value is based on the literature [70].
In both models, all individuals that are susceptible to HZ are exposed to an age- and time-dependent force of boosting.
The HZ vaccine is administered to a proportion c_{HZ} of the HZ susceptible individuals targeted by the program (ZS, ZS* and VP). Successfully immunized individuals acquire life-long immunuty to HZ, whereas those experiencing vaccine failure, which occurs with a probability $1-p_{HZ}$, remain susceptible to HZ.
A schematic representation of the models is given Figure 2.1 (Model PI) and Figure 2.2 (Model TI).

2.2.3 Demographic and vaccination scenarios

No vaccination. As a baseline, we consider a model without varicella nor HZ vaccination. The evolution of varicella and HZ epidemiology under this model is only driven by the changing demography.

In order to investigate the cost-effectiveness of varicella and HZ vaccination we consider eight different scenarios and compare them to the case of no vaccination. In particular, following the guidelines in PNPV 2016-2018 [94], we implement varicella vaccination with a 2-dose schedule (first dose at 15 months of age, second dose at 5-6 years of age) and HZ vaccination programs targeted either at 50 or 65 years old individuals.

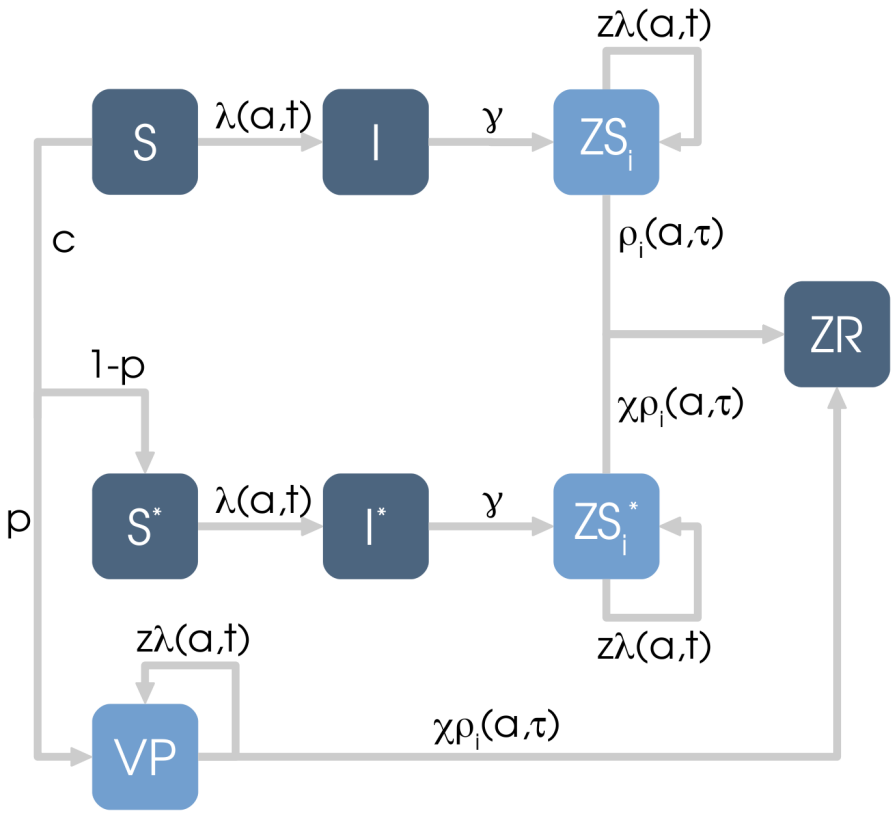

Figure 2.1: **Flow diagram of the model Progressive Immunity (PI). Varicella and HZ natural history and vaccination.** S: individuals susceptible to natural varicella; I: individuals infected with natural varicella; ZS_i: individuals susceptible to natural HZ with i episodes of previous exposure to VZV; S*: individuals susceptible to breakthrough varicella; I*: individuals infected with breakthrough varicella; ZS_i^*: individuals susceptible to HZ after infection with breakthrough varicella with i episodes of previous exposure to VZV; VP: individuals successfully immunized to varicella and susceptible to HZ from varicella vaccine strain; ZR: individuals who have developed HZ. $\lambda(a,t)$: time and age-dependent force of infection; γ: constant recovery rate from varicella; z: boosting efficacy, $\rho_i(a,\tau)$: VZV reactivation rate, c: varicella vaccination coverage, p: varicella vaccine efficacy; χ scale factor for the VZV reactivation rate of vaccinated individuals. HZ susceptible individuals (light blue compartments) move to ZR with a probability $c_{HZ}p_{HZ}$ where c_{HZ} is the HZ vaccination coverage and p_{HZ} is the HZ vaccine efficacy.

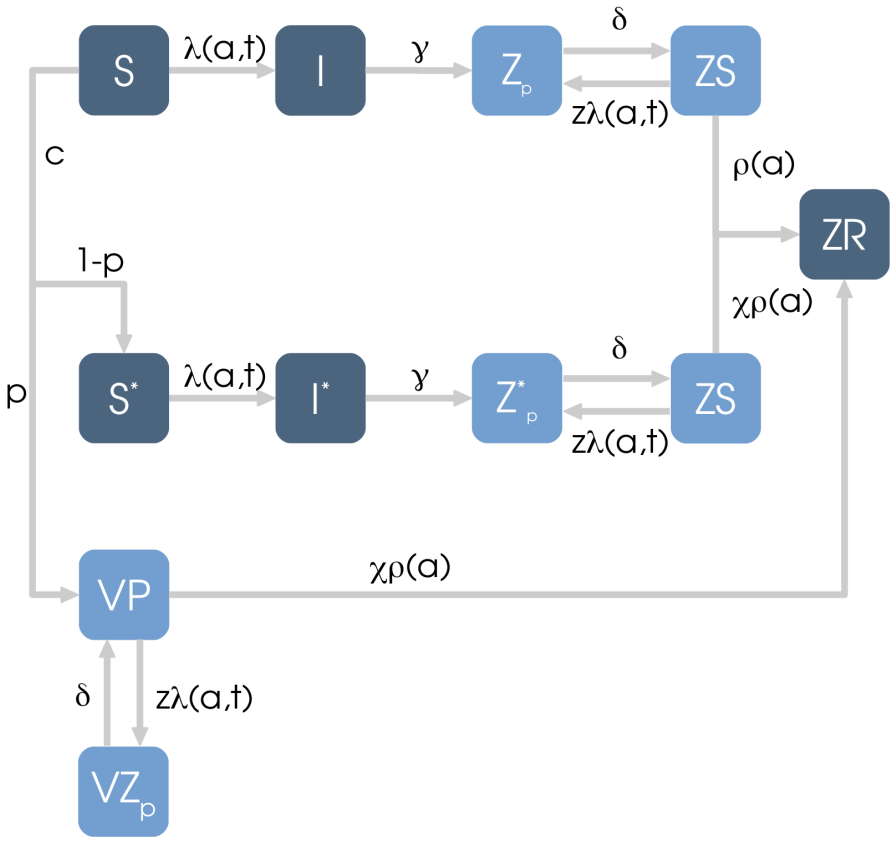

Figure 2.2: **Flow diagram of the model Temporary Immunity (TI). Varicella and HZ natural history and vaccination.** S: individuals susceptible to natural varicella; I: individuals infected with natural varicella; Z_p: individuals recovered from natural varicella and temporary protected against VZV reactivation; ZS: individuals susceptible to natural HZ; S*: individuals susceptible to breakthrough varicella; I*: individuals infected with breakthrough varicella; Z_p^*: individuals recovered from breakthrough varicella and temporary protected against VZV reactivation; ZS*: individuals susceptible to HZ after infection with breakthrough varicella; VP: individuals successfully immunized to varicella and temporary immune to HZ; VP*: individuals susceptible to HZ from vaccine strain; ZR: individuals who have developed HZ. $\lambda(a,t)$: time and age-dependent force of infection; γ: constant recovery rate from varicella; z: boosting efficacy, $1/\delta$: duration of immunity to HZ after exposure to varicella; $\rho(a)$: VZV reactivation rate, c: varicella vaccination coverage, p: varicella vaccine efficacy; χ scale factor for the VZV reactivation rate of vaccinated individuals. HZ susceptible individuals (light blue compartments) move to ZR with a probability $c_{HZ}p_{HZ}$ where c_{HZ} is the HZ vaccination coverage and p_{HZ} is the HZ vaccine efficacy.

Varicella vaccination only. We consider two extreme scenarios, chosen with respect to their effects on the epidemiology of HZ. The first scenario, called "worst-case scenario", is the one having the greatest effect on HZ as it tends to eliminate VZV circulation, thus removing the exogenous boosting effect. This scenario presents a varicella vaccination coverage and efficacy of 95% (combined for the first and second dose), and it assumes the highest estimate obtained in [70] for the relative risk of VZV reactivation among vaccinated individuals, $\chi = 0.23$. The second scenario, called "best-case scenario", is the one having the least effect on HZ as it maintains a moderate level of VZV circulation and, in turn, a higher protective effect of exogenous boosting. This scenario presents a varicella vaccination coverage of 70%, a varicella vaccine efficacy of 65%, and it assumes the lowest estimate obtained in [70] for the relative risk of VZV reactivation among vaccinated individuals, $\chi = 0.08$.

HZ vaccination only. We consider two possible scenarios that vary in terms of age at vaccination. For both scenarios we assume an HZ vaccination coverage of 60% and a vaccine efficacy of 70%. For what concerns the age at vaccination, in the first scenario the vaccine is administered at 50 years of age (we will refer to it as "Zoster50"), whereas in the second at 65 years (we will refer to it as "Zoster65").

Combined varicella and HZ vaccination. The last four scenarios combine the two varicella vaccination scenarios ("worst-case scenario" and "best-case scenario") with the two HZ vaccination scenarios considered ("Zoster50" and "Zoster65").
A summary of the scenario considered is reported in Table 2.1.

Table 2.1: **Description of the vaccination scenarios considered.** The coverage and efficacy of the varicella vaccine are denoted respectively by c and p. χ is the scale factor for the VZV reactivation rate among vaccinated. c_{HZ} and p_{HZ} are respectively the coverage and the efficacy of the HZ vaccine.

Scenario	c	p	χ	c_{HZ}	p_{HZ}	**Age at HZ vaccination**
Worst-case	95%	95%	0.23	-	-	-
Best-case	70%	65%	0.08	-	-	-
Zoster50	-	-	-	60%	70%	50y
Zoster65	-	-	-	60%	70%	65y
Worst-case+Zoster50	95%	95%	0.23	60%	70%	50y
Worst-case+Zoster65	95%	95%	0.23	60%	70%	65y
Best-case+Zoster50	70%	65%	0.08	60%	70%	50y
Best-case+Zoster65	70%	65%	0.08	60%	70%	65y

2.2.4 Model calibration

Analogously to the model presented in Chapter 1, the model is calibrated by using Monte Carlo Markov chain applied to the product of the binomial likelihood of the VZV seroprevalence profile in 1996 [66] and the Poisson likelihood of the observed age-specific HZ incidence in 2004 [72].

2.2.5 Cost-effectiveness analysis

A cost-utility economic framework is applied to the outcome of the epidemiological models. In particular, the epidemiological and economic impact of the eight intervention scenarios, with respect to the no vaccination strategy, are reported in terms of averted deaths, QALYs gained, and net costs. More specifically, the number of averted deaths due to a specific vaccination scenario $\bar{S}$ ($AD_{\bar{S}}$) vs no vaccination is defined as the sum (over all ages) of the averted deaths due to infection with natural varicella ($AD_{\bar{S}}^{NV}$), breakthrough varicella ($AD_{\bar{S}}^{BV}$) and Herpes Zoster ($AD_{\bar{S}}^{HZ}$), which are computed as follows:

$$AD_{\bar{S}}^{NV} = D_{novacc}^{NV} - D_{\bar{S}}^{NV} = [NPV_{novacc}^{NV} - NPV_{\bar{S}}^{NV}] \times HR^{NV} \times CFR^{NV}$$

$$AD_{\bar{S}}^{BV} = D_{novacc}^{BV} - D_{\bar{S}}^{BV} = [NPV_{novacc}^{BV} - NPV_{\bar{S}}^{BV}] \times HR^{BV} \times CFR^{BV}$$

$$AD_{\bar{S}}^{HZ} = D_{novacc}^{HZ} - D_{\bar{S}}^{HZ} = [NPV_{novacc}^{HZ} - NPV_{\bar{S}}^{HV}] \times HR^{HZ} \times CFR^{HZ}$$

where D is the number of deaths under a specific vaccination scenario. $NPV_{\bar{S}}$ is the net present value of infection cases under a specific vaccination scenario and is computed as:

$$NPV_{\bar{S}} = \sum_{t=1}^{T} \frac{C_t}{(1+r)^t}$$

where C_t is the number of infection cases at time t, T is the time horizon considered and r is the discount rate. CFR is the case fatality rate per hospitalized individual and HR is the hospitalization rate per case of infection. The averted deaths are calculated under the assumption that only people who are hospitalized for varicella or HZ carry the risk of dying because of the infection.
The QALY gain due to a specific vaccination scenario $\bar{S}$ ($QG_{\bar{S}}$) vs no vaccination, is defined as the sum (over all ages) of the QALY gain due to infection with natural varicella ($QG_{\bar{S}}^{NV}$), breakthrough varicella ($QG_{\bar{S}}^{BV}$) and Herpes Zoster ($QG_{\bar{S}}^{HZ}$), which are computed as follows:

$$QG_{\bar{S}}^{NV} = QL_{novacc}^{NV} - QL_{\bar{S}}^{NV} = [NPV_{novacc}^{NV} - NPV_{\bar{S}}^{NV}] \times [QL_m + QL_d \times HR^{NV} \times CFR^{NV}]$$

$$QG_{\bar{S}}^{BV} = QL_{novacc}^{BV} - QL_{\bar{S}}^{BV} = [NPV_{novacc}^{BV} - NPV_{\bar{S}}^{BV}] \times [QL_m + QL_d \times HR^{BV} \times CFR^{BV}]$$

$$QG_{\bar{S}}^{HZ} = QL_{novacc}^{HZ} - QL_{\bar{S}}^{HZ} = [NPV_{novacc}^{HZ} - NPV_{\bar{S}}^{HZ}] \times [QL_m + QL_d \times HR^{HZ} \times CFR^{HZ}]$$

where QL is the total QALY loss associated with a specific vaccination scenario,QL_m is the QALY loss per case of infection due to morbidity, and QL_d is the QALY loss per case of infection due to death calculated on the basis of age-specific life expectancies (experienced only by people who have been hospitalized).
The net total costs associated to a specific vaccination scenario $\bar{S}$ ($NC_{\bar{S}}$) vs no vaccination, are defined as the sum of the net costs associated to infection with natural varicella ($NC_{\bar{S}}^{NV}$), breakthrough varicella ($NC_{\bar{S}}^{BV}$) and Herpes Zoster ($NC_{\bar{S}}^{HZ}$), which are computed as follows:

$$NC_{\bar{S}}^{NV} = TC_{\bar{S}}^{NV} - TC_{novacc}^{NV} = [NPV_{novacc}^{NV} - NPV_{\bar{S}}^{NV}] \times [C_{GP} + C_H \times HR] + NPV_{vd} \times C_V]$$

$$NC_{\bar{S}}^{BV} = TC_{\bar{S}}^{BV} - TC_{novacc}^{BV} = [NPV_{novacc}^{BV} - NPV_{\bar{S}}^{BV}] \times [C_{GP} + C_H \times HR] + NPV_{vd} \times C_V]$$

$$NC_{\bar{S}}^{HZ} = TC_{\bar{S}}^{HZ} - TC_{novacc}^{HZ} = [NPV_{novacc}^{HZ} - NPV_{\bar{S}}^{HZ}] \times [C_{GP} + C_H \times HR] + NPV_{vd} \times C_V]$$

where TC are the total costs associated with a specific vaccination scenario and add up to the total cost of disease and the total cost of vaccination. Direct costs (GP, hospital and treatment) and costs of the vaccination program are considered. In particular, the cost of disease per case of infection amounts to C_{GP}, the total outpatient cost (cost of GP visit and cost of treatment given by GP), and C_H, the total inpatient cost, experienced only by those who are hospitalized. The cost of vaccination per administered vaccine dose (NPV_{vd}) amounts to C_V, which combines the cost of vaccination and the respective administration costs.
A stochastic simulation of a certain vaccination scenario $\bar{S}$ is defined cost-effective if

$$K \times QG_{\bar{S}} - NC_{\bar{S}} > 0 \tag{2.1}$$

where K, taken equal to 40,000 €, is the maximum value that the payer (in our case, the National Health System) is willing to pay for an additional QALY gained. According to the formula, a vaccination scenario is deemed cost-effective with respect to the no vaccination scenario whenever the benefits, namely, the economic value of the QALY gained, dominate over the net costs. Cost-effectiveness results are produced for three different time horizons (25, 50 and 86 years) and discount rates are fixed at 1.5% and 3% per year. The percentage of runs that are cost-effective is computed as the percentage of simulations that satisfies Equation 2.1. In this way, we take into account the uncertainty intrinsic to stochasticity and parameter estimation.
In order to assess the robustness of our results with respect to the economic parameters, a probabilistic sensitivity analysis was also performed, whereby we included uncertainties in the economic and quality of life (QoL) parameters, generated under a Bayesian framework. Epidemiological and economic parameters used to perform the cost-effectiveness analysis are summarized in Table B.1-B.3 reported in the appendix.

2.3 Results

Both models reproduce well the VZV serological profile observed in Italy in 1996, including the drop in the percentage of seropositives after the first year of age, which is ascribable to the waning of maternal antibodies (see Figure 2.3 **a**/**c**). The mean estimate obtained by the model (black line) lies within the 95% CI of the data for all considered age groups. For what concerns HZ, both models are capable to reproduce the HZ incidence by age observed in 2004 (Figures 2.3 **b**/**d**), even though with some differences. Model PI captures well the incidence of HZ in adults and elderly, whereas it slightly overestimates the HZ incidence in individuals younger than 30 years. In particular, Model PI accurately reproduce the slope of the HZ increase occurring between 40 and 80 years and it is able to capture the decreasing trend in HZ incidence observed after 80 years of age.
Model TI is more precise in reproducing the incidence of HZ in individuals younger than 30 years, whereas it performs slightly worse both in capturing the increasing trend in the age-groups 40-80 years and the decrease of HZ incidence after 80 years of age.

2.3.1 VZV epidemiology in Italy (1900-2015)

The model was firstly used to investigate the past epidemiology of VZV infection and reactivation under the effect of demographic changes.
As shown in Figure B.1 in the appendix during the last century the Italian population underwent a fertility transition comparable to that observed in Spain [90]. According to both models (PI and TI) the overall decreasing trend of the birth rate observed in Italy since 1900 has led to a reduction of the fraction of VZV seropositive individuals at all ages and to an increase in the estimated mean age at varicella infection (see Figure B.2 in the appendix). The decrease in varicella circulation over time has resulted in an overall reduction in the frequency of boosting events, whose impact on HZ incidence is different depending on the formulation used to model exogenous boosting. In particular, as shown in Figure 2.4, according to Model PI the reduction of boosting results in a 138% increase in the total HZ incidence between 1900 to 2015, whereas the one estimated by Model TI is less remarkable (53%). On the other hand, the increase in the estimated mean age at VZV reactivation, which is mainly ascribable to population aging, has the same magnitude in both models (see Figure B.3 in the appendix). These results suggest that the dynamics of HZ incidence are much more sensitive to changes in varicella circulation, which in the past were mainly driven by demographic changes, when we model boosting according to PI.

2.3.2 VZV epidemiology in Italy (2015-2100)

In all scenarios considered for the future we model changes in the age distribution of the population using projections on the yearly birth rate and mortality rates by age provided by the UN for Italy [92]. According to the 2015 UN World Population Prospects the Italian

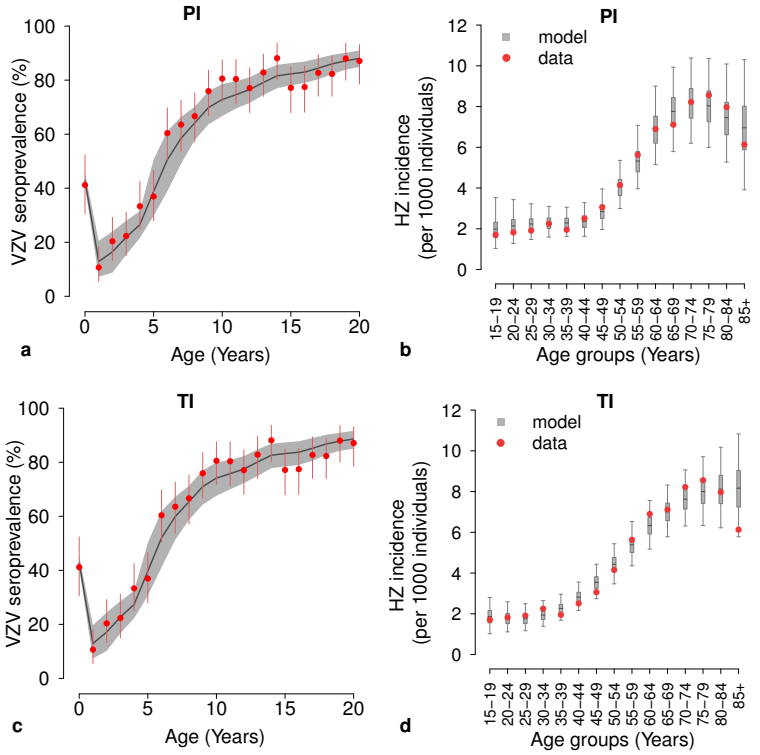

Figure 2.3: **VZV seroprevalence and HZ incidence.** **a** Age-specific VZV seroprevalence as observed in data from Italy in 1996 [66] and as estimated by the model PI. Grey areas show the 95% CI of model estimates, vertical lines represent 95% CI of the data computed by exact binomial test. **b** HZ incidence by age group as observed in 2004 [72] and as estimated by the model PI. **c** As **a** but as estimated by the model TI. **d** As **b** but as estimated by the model TI.

birth rate will remain approximatively stable during the period 2015-2100 (see Figure B.1 in the appendix), whereas the age-specific mortality rate will experience a progressive decrease during the same period.

No vaccination

Results obtained by both models show a stable epidemiology of varicella during the prediction period (see Figure 2.5 **a**/**c**) which is explained by the relative stability of the birth rate. Analogously to what observed in the case of Spain, under both models the incidence of HZ will continue its growth and it will stabilize only after some decades (see Figure 2.5

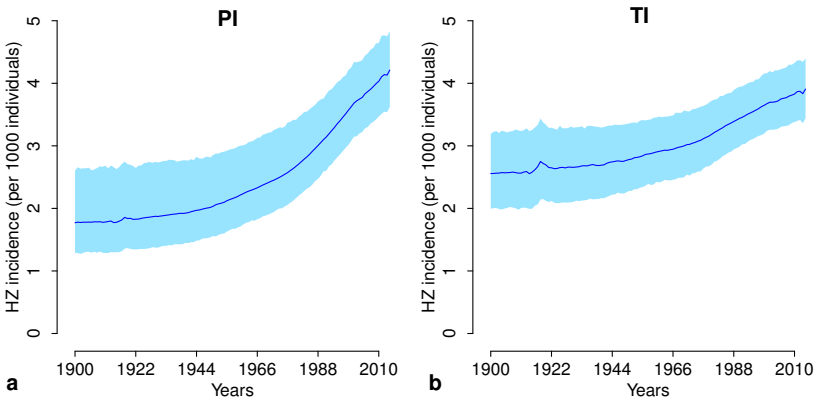

Figure 2.4: **Total HZ incidence over time. a** Total incidence of HZ over the period 1900-2015 as estimated by the model PI. **b** As **a** but as estimated by the model TI.

b/**d**). Such growth is ascribable to two factors, the first one is population aging that acts equally in both models, the second one is the delayed effect of past changes in the boosting effect. As previously noticed, the latter play a more relevant role in Model PI, as opposed to Model TI, and it is responsible for the quantitative difference in the predicted HZ incidence increase. In particular, during the period 2015-2100, Model PI estimates a maximum increase in the total HZ incidence of about 56% and it will stabilize in the long term at a 48% higher value than in 2015. On the other hand, the maximum increase estimated by Model TI amounts to 11% and stabilizes in the long term at a value 3% higher than in 2015.

Varicella vaccination only

Predictions from both models suggest that in the worst-case scenario, which assumes a high varicella vaccination coverage (95%) and high vaccine efficacy (95%), we can expect a sudden and enduring reduction in varicella incidence (see Figure B.4 **a**/**b** in the appendix). As a consequence, starting from 2015 the frequency of VZV re-exposures among HZ susceptibles experiences a remarkable drop. As shown in Figure 2.6 **a**/**b**, according to both models, the lack of boosting will result in an increase of natural HZ incidence in the first decades after the introduction of vaccination. After this period, since varicella vaccination is effective in reducing the fraction of individuals susceptible to natural HZ, the incidence of HZ is expected to decrease. However, in the long term, the incidence of natural HZ will progressively be replaced by the HZ cases among vaccinated individuals, i.e. from breakthrough or vaccine strain.
In the best-case scenario we assume a lower varicella vaccine coverage (70%) and efficacy (65%). This vaccination strategy results less effective in reducing the incidence of varicella

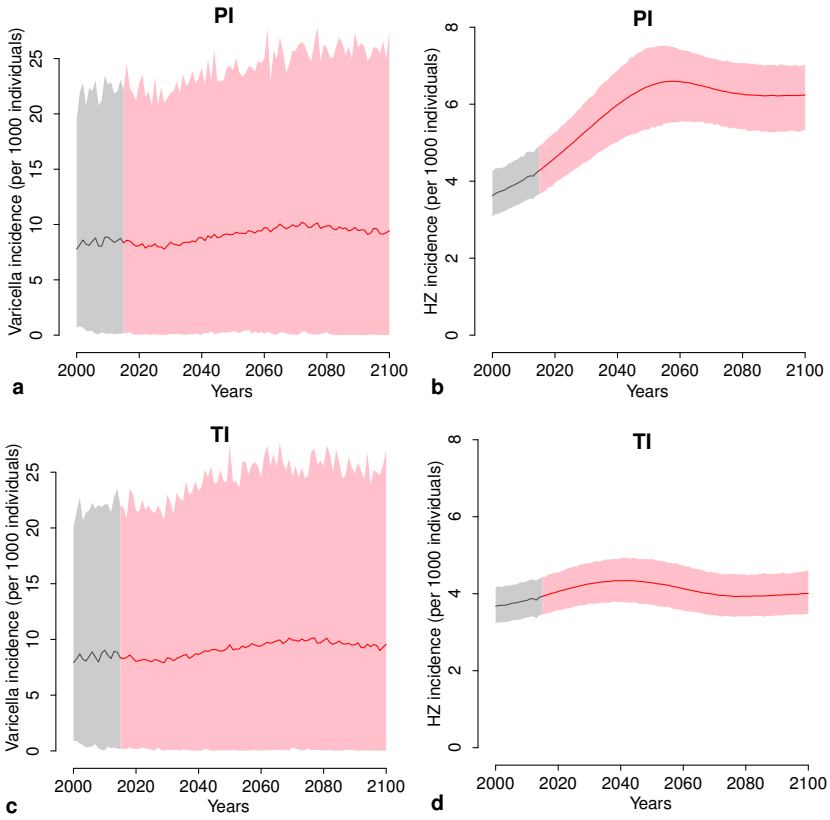

Figure 2.5: **Varicella and HZ incidence in the scenario with no vaccination (2015-2100). a** Total incidence of varicella as estimated by the model PI during the period 2015-2100 in the scenario without nor varicella or HZ vaccination. **b** Total incidence of HZ as estimated by the model PI during the period 2015-2100 in the scenario without nor varicella or HZ vaccination. **c** As **a** but as estimated by the model TI. **d** As **b** but as estimated by the model TI.

that will keep circulating both as natural and breakthrough (see Figure B.4 **c**/**d**). As a consequence, the reduction of boosting after 2015 is milder than in the worst-case scenario and this results in a more contained increase of HZ incidence in the post-vaccination period. Moreover, since in this scenario natural varicella continues circulating , natural HZ will play an important role even in the long term (See Figure 2.6 **c**/**d**). The magnitude of the predicted changes in HZ incidence depends on the boosting formulation considered. Moreover, it is worth noting that the predicted changes in HZ incidence are not purely ascribable to varicella immunization, but encompass also the increase of HZ due to demographic changes. It is thus crucial to interpret results obtained in light of the dynamics predicted in the absence of vaccination, when the dynamics are influenced by demographic processes only.

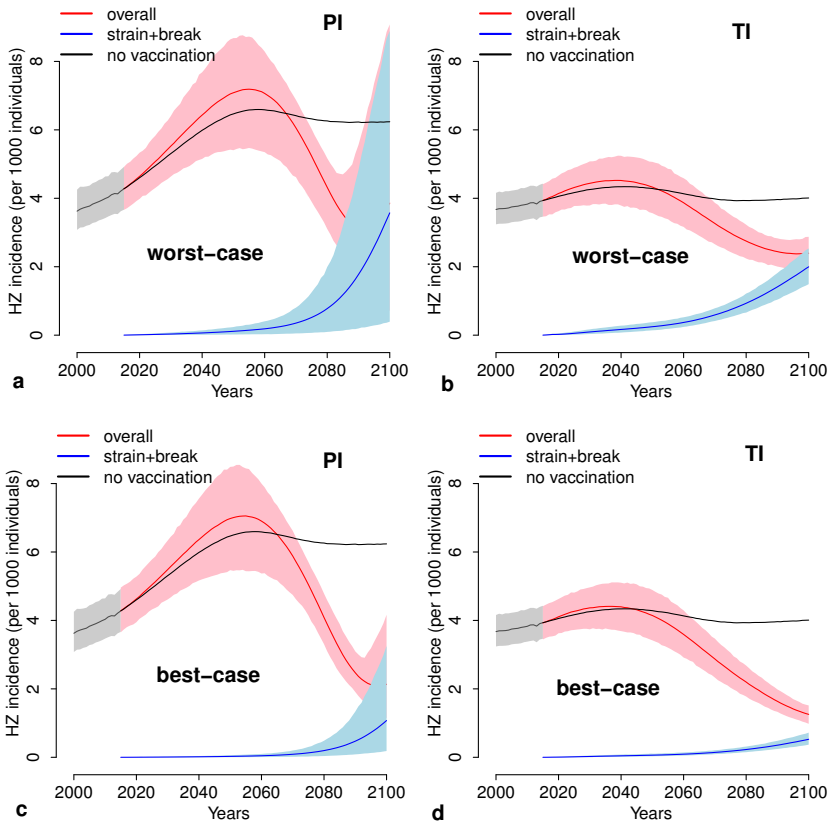

Figure 2.6: **Total HZ incidence in the scenarios with varicella vaccination only (2015-2100). a** Total incidence of HZ overall (red line) and total incidence of HZ developed after infection with breakthrough varicella or from varicella strain (blue line) as estimated by the model PI during the period 2015-2100 in the worst-case scenario for varicella vaccination only (95% coverage and 95% efficacy). **b** As **a** but as estimated by the model TI. **c** As **a** but as estimated by the model PI in the best-case scenario for varicella vaccination only (70% coverage and 65% efficacy). **d** As **c** but as estimated by the model TI.

For instance, in the worst case scenario Model PI estimates a maximum increase in the total HZ incidence of about 71% with respect to 2015, however only the 20% of this increase may be directly ascribed to varicella immunization.

HZ vaccination only and combined with varicella vaccination

The introduction of an HZ vaccination program can be used to mitigate the increase of HZ incidence both in the absence or presence of varicella mass immunization. For instance, according to Model PI, the HZ vaccination scenario "Zoster65" would reduce by 26% the

long term incidence of HZ expected in the scenario without interventions. If we combine "Zoster65" with the "worst case scenario" for varicella immunization, the peak incidence expected would be reduced by 30%.

2.3.3 Cost-effectiveness analysis

Firstly, we consider the whole epidemiological and demographic uncertainty in the net present values of the predicted cases of varicella and HZ, as generated by the mathematical model, and we combine it with the "observed" values for the economic and quality of life parameters (Base case analysis, Table B.2-B.3 in the appendix). In this way, the uncertainty in the results is due exclusively to the demographic and epidemiological data. Figures 2.7-2.10 report for each epidemiological model (PI and TI) the results of the cost-effectiveness analysis, under the base case, and considering three time horizons (25, 50, and 86 years). Results are presented in terms of averted deaths, QALYs gained, net costs and percentage of cost-effective runs. The implementation of a vaccination program including

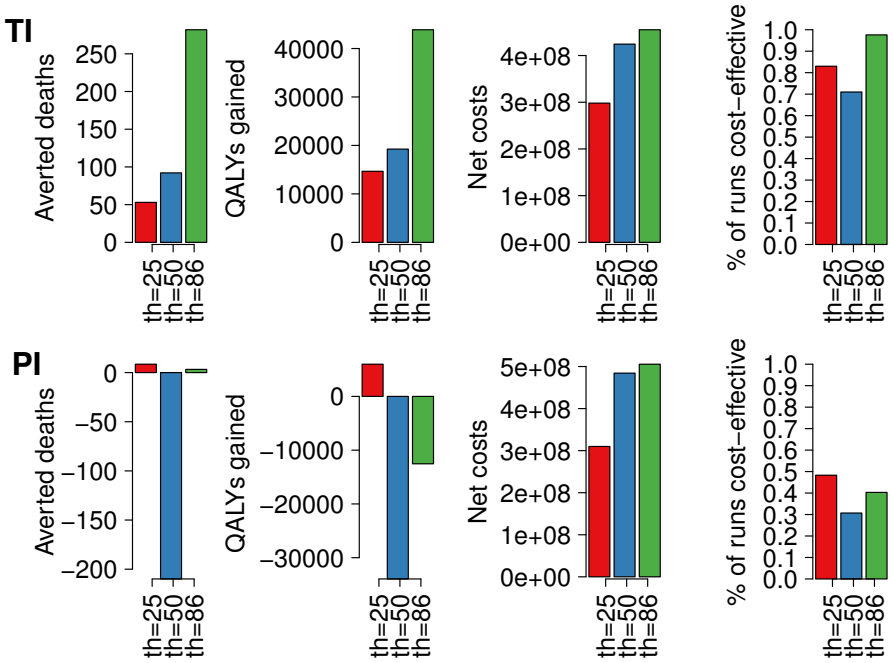

Figure 2.7: **Cost-effectiveness analysis results for the worst-case scenario of varicella vaccination alone vs no vaccination, under the base case.** Median number (over 2,000 stochastic simulations) of the averted deaths, QALYs gained, net costs and percentage of runs that are cost-effective, for the three considered time horizons as obtained by using model TI (top row) and PI (bottom row).

only varicella vaccination is not deemed cost-effective, especially under the epidemiological

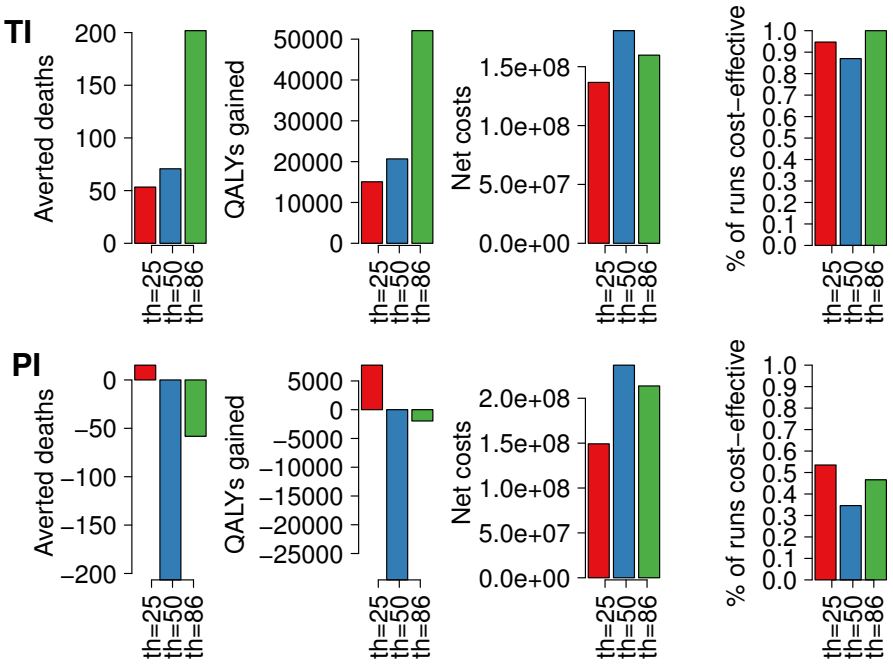

Figure 2.8: **Cost-effectiveness analysis results for the best-case scenario of varicella vaccination alone vs no vaccination, under the base case.** Median number (over 2,000 stochastic simulations) of the averted deaths, QALYs gained, net costs and percentage of runs that are cost-effective, for the three considered time horizons as obtained by using model TI (top row) and PI (bottom row).

predictions based on model PI. Figures 2.7 and 2.8 (bottom row) show that, compared to no vaccination, both worst- and best-case scenarios would result in a negative balance in terms of averted deaths and QALYs gained, especially after 50 years since the implementation of the program. When looking at the probability of the program of being cost-effective, we notice that less than 50% of runs are cost-effective for all time horizons, in particular for the time horizons of 50 years. The fact that the worst situation is reported for the middle time horizon can be explained by the fact that 50 years after the implementation of the vaccination program, the predicted HZ incidence is at the highest level, starting to decrease after having reached the peak (see for instance Figure 2.6 **a**/**c**). Hence, a vaccination program that aims only at reducing VZV circulation has also the drawback of reducing the boosting effect and thus increasing HZ incidence in the medium term. If we focus on the difference between the worst- and the best-case scenarios (Figures 2.7 and 2.8, bottom row), we see that the former, characterized by higher coverage and efficacy and higher percentage of vaccinated that will get HZ, performs worse in terms of averted deaths and QALYs gained (more deaths under the vaccination scenario and less QALYs gained), and also in terms of

costs (higher costs mainly due to higher varicella vaccination coverage and higher cost of disease for increased HZ cases).
The results of HZ vaccination alone, at either 50 or 65 years, are shown in Figures B.5**a** in the appendix and in Figure 2.9. In both models (PI and TI) and for all the time horizons, HZ vaccination is always cost-effective. In particular, both programs perform very well in terms of averted deaths and QALYs gained, with benefits increasing over the time horizons, but also in terms of net costs with respect to no vaccination. By protecting older people from getting HZ, the vaccination programs prevent the potential high costs of disease at older ages (more hospitalizations and higher risk of death) and the reduction in the related QoL and deaths.
Moreover, vaccinating at 65 years of age results better than at 50 years, since the benefits

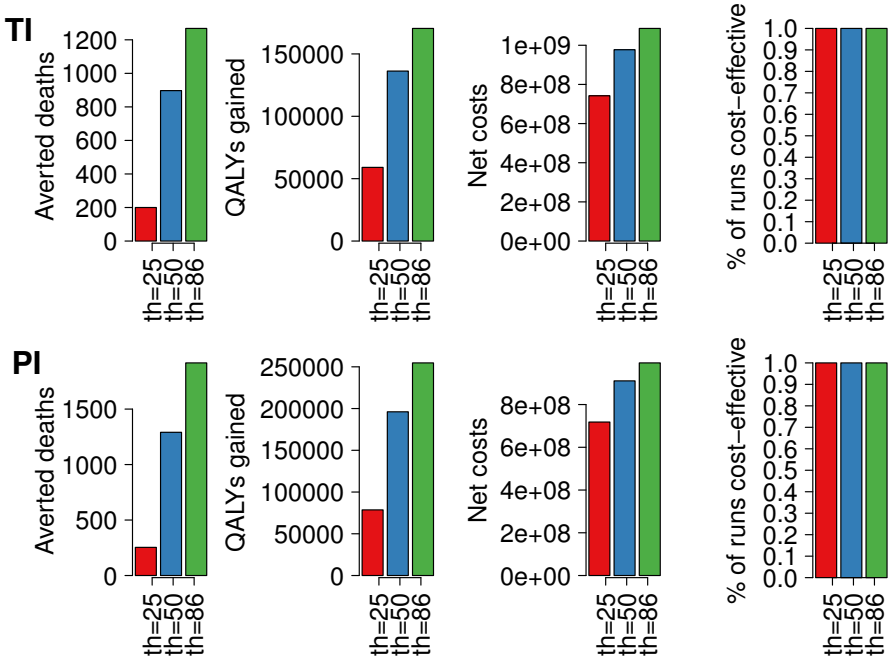

Figure 2.9: **Cost-effectiveness analysis results for the scenario with HZ vaccination alone targeted at 65 years old vs no vaccination, under the base case.** Median number (over 2,000 stochastic simulations) of the averted deaths, QALYs gained, net costs and percentage of runs that are cost-effective, for the three considered time horizons as obtained by using model TI (top row) and PI (bottom row).

in terms of averted deaths and QALYs gained increase with the age at vaccination, being the incidence higher, while the net costs remain more or less the same.
Finally, when combining varicella vaccination with HZ vaccination, we manage to partially compensate the drawbacks of varicella vaccination with the positive effects of HZ vaccina-

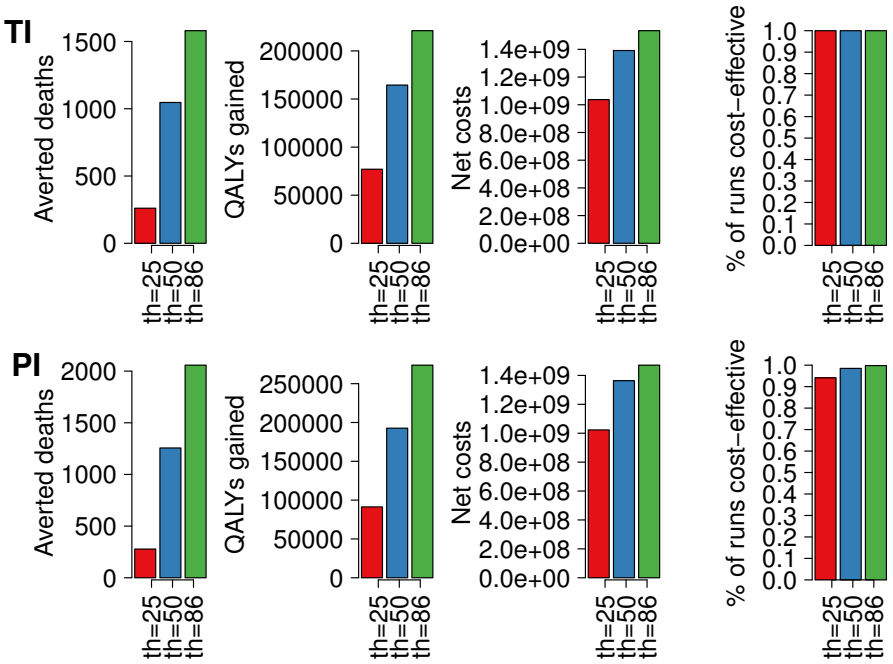

Figure 2.10: **Cost-effectiveness analysis results for the worst-case scenario of varicella vaccination and HZ vaccination at 65 years vs no vaccination, under the base case.** Median number (over 2,000 stochastic simulations) of the averted deaths, QALYs gained, net costs and percentage of runs that are cost-effective, for the three considered time horizons as obtained by using model TI (top row) and PI (bottom row).

tion, thus obtaining a higher percentage of cost-effective runs (see Figure 2.10 and Figures B.5-B.6). As a matter of fact, the increase in the medium term of HZ incidence, predicted by model PI following varicella vaccination, is counterbalanced by the protective effect of HZ vaccination among the elderly. In this way, vaccinating the people more at risk of developing HZ, we manage to maintain a positive balance, increasing over the time horizons, in terms of averted deaths and QALY gained, but also in terms of net costs. The result is that the percentage of cost-effective runs is very high, also after 25 years, mostly if we vaccinate at 65 years (e.g. see difference between Figure 2.10 and B.5**b** in the appendix, bottom row).

2.3.4 Probabilistic sensitivity analysis

We also performed a probabilistic sensitivity analysis (PSA) to investigate the effect of the uncertainty in the economic and quality of life (QoL) parameters, generated under a Bayesian framework. In general, we assumed a Gamma distribution for the continuous positive parameters (costs and number of GP visits), and a Beta distribution for the prob-

ability parameters, that are bounded between 0 and 1. More details on the distribution assumed for each parameter are given in Table B.1-B.3 in the appendix.
To assess the effect of the economic and QoL uncertainty, we employed the following procedure. Each single realisaton of the economic and QoL model parameters is combined to all the stochastic runs of the epidemiological (and demographic) model, and the median values of the number of averted deaths, QALYs gained, net costs, and percentage of cost-effective runs are derived. Replicating this procedure for all the runs of the economic and QoL model, we managed to take into account the economic and QoL uncertainty, which is then reported as credible intervals around the median values of the outputs. In the appendix we report the results of the cost-effectiveness analysis under the PSA, for the two different mathematical models (PI and TI), and for three time horizons (25, 50, and 86 years), in terms of the four outputs presented above (Figures B.7-B.10). Comparing the resulting outputs of the PSA with those reported for the base case analysis, we notice that there is consistency among the median values of the outputs. The effect of the uncertainty in the economic and QoL model is evident mostly for the scenarios with varicella vaccination only, for which we found large credible intervals for all the outputs. In particular, we highlight the large economic and QoL uncertainty for the averted deaths and QALYs gained under the longest time horizon (86 years).

2.4 Conclusions

About twenty years after the first licensing of the varicella vaccine, public health authorities of several countries are still discussing its introduction into the national vaccination schedules. The first concern is that varicella immunization programs could lead to an increase in the incidence of varicella among adults, which is usually more severe than childhood infection [87]. Second, the efficacy of the vaccine strongly depends on the number of doses [95] and after vaccine failure people may still develop a milder form of the disease, i.e. breakthrough varicella. However, the third and greatest cause of concern is that the reduction of varicella circulation may lead to an increase of HZ incidence, which is generally characterized by more severe and persisting symptoms, debilitating sequelae and a higher overall costs for treatment compared to varicella [72, 96]. More recently, a vaccine against HZ has been licensed and introduced in some countries, e.g. France and the UK, with the aim of containing its burden even in the absence of varicella vaccination [97, 98]. However, if on one side the French public health authorities are not considering the introduction of varicella vaccination, in the UK they are evaluating the possibility of starting a varicella vaccination program targeted at 1 or 3 years old children.
One of the tools that is nowadays used to evaluate vaccination policies is the cost- effectiveness analysis. This technique allows to take into account several factors, like for instance the direct and indirect effects of the vaccine, the avertable burden of disease as well as the costs of implementing the intervention. Past works on the cost-effectiveness of varicella

vaccination have already consistently shown that the economic acceptability of the vaccine introduction in several countries primarily depends on the in- or exclusion of exogenous boosting in the model [89]. However, not only the analysis has not been performed in the Italian setting, but also the combined effect of the boosting mechanism and the ongoing demographic changes in the population structure (e.g. aging) has never been considered. In this work we perform a cost-effectiveness analysis of varicella and HZ vaccination in Italy, including the following innovative aspects. First, we test the robustness of our results by using two different modeling assumptions about the effect of exogenous boosting on the risk of developing HZ [43, 47]. Second, in both models we include realistic changes of the age distribution of the population, in order to take into account trends in VZV epidemiology that are not directly ascribable to the immunization programs.

Our findings confirm that the magnitude of the increase of HZ incidence after the introduction of varicella vaccination depends on the formulation used to model exogenous boosting. In particular, when assuming "progressive immunity" (PI) the dynamics of HZ are much more sensitive to the changes in varicella circulation than when assuming "temporary immunity" (TI), i.e. we obtain a more remarkable increase of HZ. According to the cost-effectiveness analysis performed, these quantitative difference results critical from the public health decision point of view especially when considering the introduction of varicella vaccination alone. In particular, the extremely highly efficacious varicella vaccination scenario considered here (95% efficacy and 95% coverage) results cost-effective in the short, medium and long-term only when using model TI, whereas the percentage of cost-effective runs for model PI is less than 50% for all considered time horizons. On the other hand, HZ vaccination alone results highly cost-effective, independently of the model used, both if the vaccine is administered at 50y or 65y and for all time horizons considered. Finally, the combination of varicella and HZ vaccination are highly cost-effective for all models' formulation but, especially, when the HZ vaccine is given at 65 years. These results are robust to changes in economic an QoL parameters and consistent with the outcome from the base case analysis. To conclude, the analysis performed highlights that quantitative predictions about the impact of different immunization strategies on the incidence of HZ strongly depend on the modeling assumption on exogenous boosting. However, the cost-effectiveness analysis results robustly show that, disregarding the way in which we model boosting, an HZ vaccination program combined with a varicella vaccination program has a high probability of being cost-effective even considering all the different sources of uncertainty in demographic, epidemiological and economic model parameters.

Appendix B

B.1 Additional figures and tables

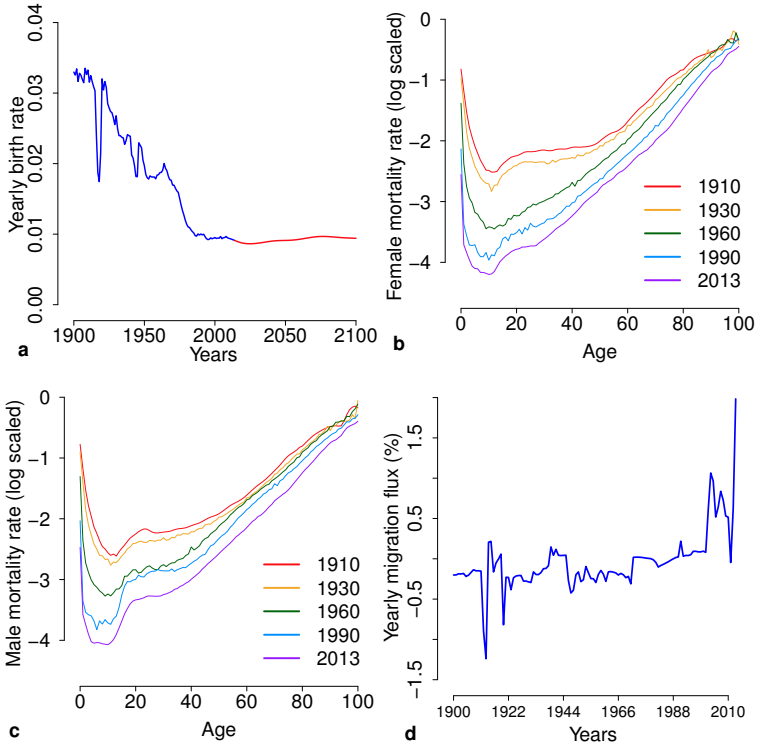

Figure B.1: **Historical demographic data in Italy and projections for the future.** **a** Birth rate over time in Italy. Yearly birth rate as observed in Italy during the period 1900-2014 [59, 91] (grey line) and as predicted by the UN during 2015-2100 [92] **b** Age-specific mortality rate, females [59]. **c** Age-specific mortality rate, males [59]. **d** Yearly migration flux in % of the population [91]

Table B.1: **Epidemiological parameters of the model.** Base case values and variability due to uncertainty.

Parameter	Base case	Distribution [95% CI]	References
NV incidence (per 1,000 PY)	8.43 (variable by age)	Fixed	Model prediction
HZ incidence (per 1,000 PY)	3.92 (variable by age)	Fixed	Model prediction
Proportion of HZ cases developing PHN	6.2% (variable by age)	Beta	[72]
Hospitalizations rate for NV (per 1,000 cases)	2.36 (variable by age)	Binomial	[a]
BV vs NV hospitalization rate	25%	Beta [16.2%-35.2%]	[99]
Hospitalizations rate for HZ (per 1,000 cases)	13.12 (variable by age)	Binomial	[a]
Hospitalizations rate for PHN (per 1,000 cases)	40.77 (variable by age)	Binomial	[a]
Case fatality rate for NV (per 1,000 cases)	4.01 (variable by age)	Beta	[b]
BV vs NV case fatality rate	0.5%	Beta [0.2%-1%]	[100, 101]
HZ case fatality rate (per 1,000 hospitalized)	12.70 (variable by age)	Beta	[b]
Number of GP consultations per NV varicella case (≤14y)	2	Gamma [1.6-2.4]	[100]
Number of GP consultations per NV varicella case (>14y)	1	Gamma [0.6-1.4]	[100]
Number of GP consultations per BV varicella case	0.5	Gamma [0.4-0.6]	[100]

[a] Average number of hospitalizations by age due to varicella, HZ and PHN (Hospital Discharge Register, HDR, 2001-2012) divided by the pre-vaccination incidence generated by the epidemiological model. The incidence of PHN is calculated by multiplying the incidence of HZ estimated by the model by the probability of developing PHN.

[b] Average number of deaths by age due to varicella (Italian National Health Institute, ISS 2001-2012) and HZ (EU detailed mortality database, DMDB 2001-2012) divided by the respective estimates of the hospitalization rates.

Table B.2: **Quality of life measures.** Base case values and variability due to uncertainty.

Parameter	Base case	Distribution [95% CI]	References
QALY loss NV (≤14y)	0.0036	Beta [0.003-0.005]	[48]
QALY loss NV (>14y)	0.0052	Beta [0.004-0.006]	[48]
BV vs NV QALY loss	60%	Beta [35.8%-81.9%]	[48]
QALY loss HZ (>14y)	0.075 (variable by age)	Beta	[102]
Background mortality rates	Variable by age	Fixed	[c]
Population	Variable by age	Fixed	[c]

[c] Population by age and national mortality rates were obtained from National Italian Institute of Statistics [91]

Table B.3: **Economic and vaccination parameters.** Base case values and variability due to uncertainty.

Parameter	Base case	Distribution [95% CI]	References
GP consultation costs for NV (≤14y)	20.90 €	Gamma [18.1-23.9]	[103]
GP consultation costs for NV (>14y)	14.24 €	Gamma [13.4-17.2]	[104]
GP treatment costs for NV (≤14y)	9.98 €	Gamma [7.8-12.6]	[105]
GP treatment costs for NV (>14y)	20.93 €	Gamma [18.1-24.1]	[104]
Total hospitalization costs for NV (≤14y)	338.83 €(variable by age)	Gamma	HDR[d]
Total hospitalization costs for NV (>14y)	399.72 €(variable by age)	Gamma	HDR[d]
Outpatient costs for HZ	122.68 €	Gamma [9.3-388.3]	[72]
Outpatient costs for PHN	446.10 €	Gamma [12.6-1552.8]	[72]
Total hospitalization costs for HZ (≤50y)	226.71 €(variable by age)	Gamma	HDR[d]
Total hospitalization costs for HZ (>50y)	219.87 €(variable by age)	Gamma	HDR[d]
Total hospitalization costs for PHN (≤50y)	61.77 €(variable by age)	Gamma	HDR[d]
Total hospitalization costs for PHN (>50y)	44.27 €(variable by age)	Gamma	HDR[d]
Cost per dose of varicella vaccination	35.00 €	Fixed	Invitation for bid
Cost per dose of HZ vaccination	87.00 €	Fixed	Invitation for bid
Administration cost per dose of vaccination	6.21 €	Fixed	[100]

[d] Hospital Discharge Register, 2001-2012

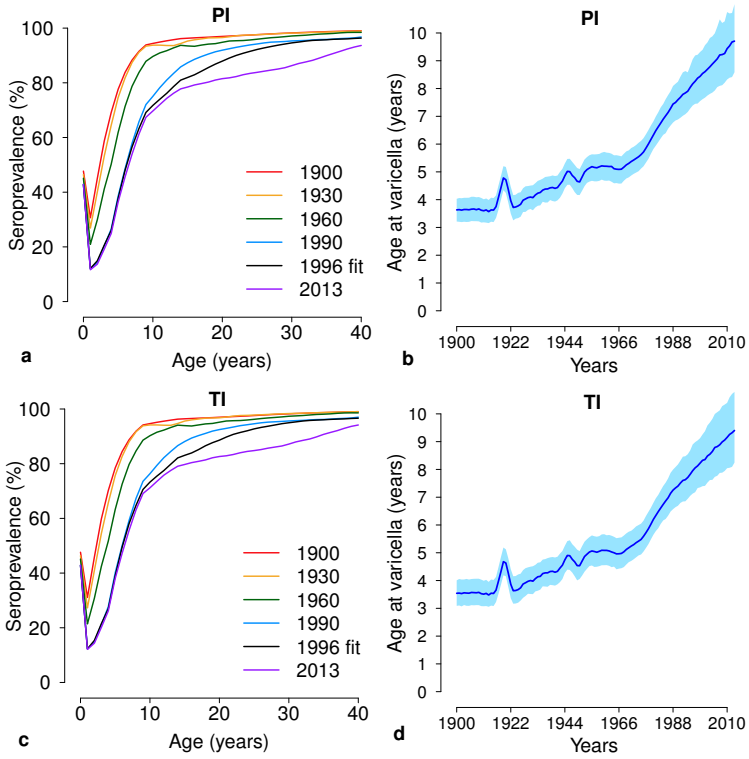

Figure B.2: **Estimated impact of demographic changes on varicella epidemiology in Italy (1900-2009) a** Age-specific VZV seroprevalence at different years as estimated by the model PI. **b** Mean age (and 95% CI, shaded areas) at varicella over time as estimates by the model PI. **c** As **a** but as estimated by the model TI. **d** As **b** but as estimated by the model TI.

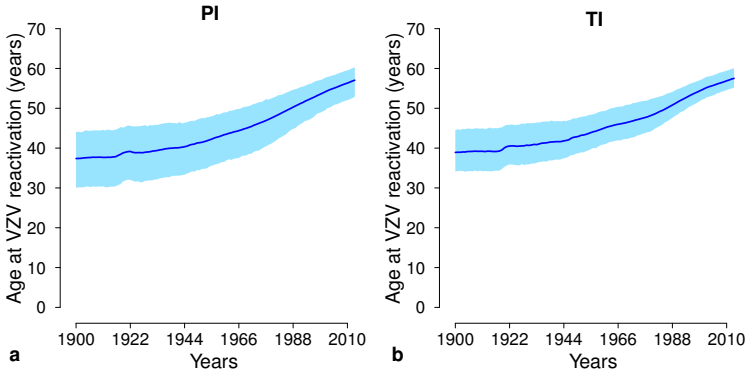

Figure B.3: **Age at VZV reactivation in Italy. a** Mean age at VZV reactivation (and 95% CI) as estimated by the model PI. **b** As **a** but as estimated by the model TI.

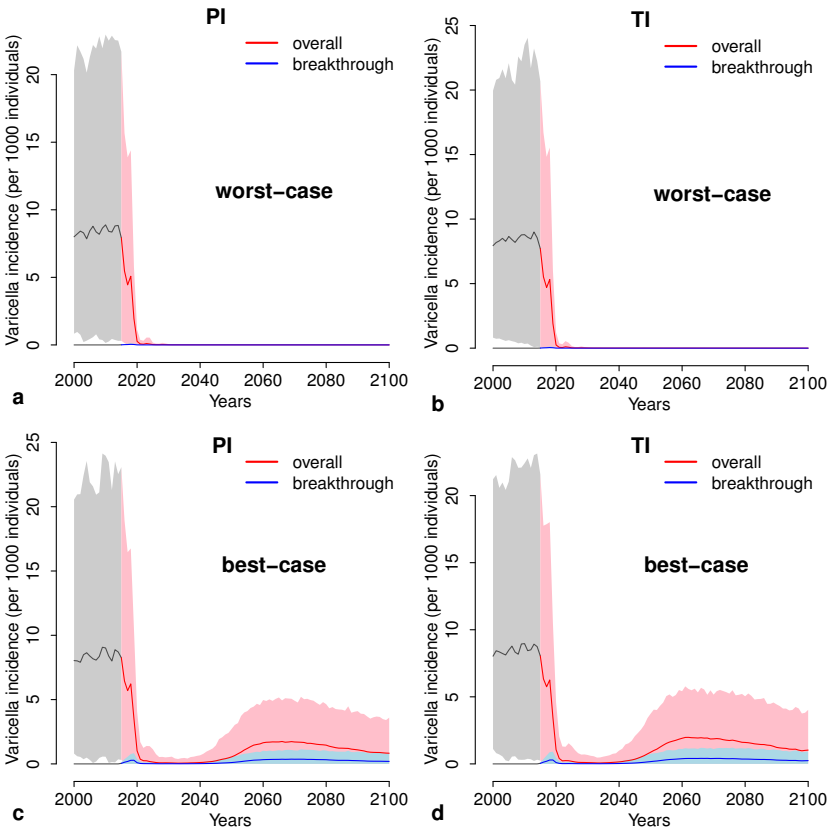

Figure B.4: **Total varicella incidence in the scenarios with varicella vaccination only (2015-2100). a** Total incidence of varicella overall, i.e. natural and breakthrough, (red line) and breakthrough varicella alone (blue line) as estimated by the model PI during the period 2015-2100 in the worst-case scenario for varicella vaccination only (95% coverage and 95% efficacy). **b** As **a** but as estimated by the model TI. **c** As **a** but as estimated by the model PI in the best-case scenario for varicella vaccination only (70% coverage and 65% efficacy). **d** As **c** but as estimated by the model TI.

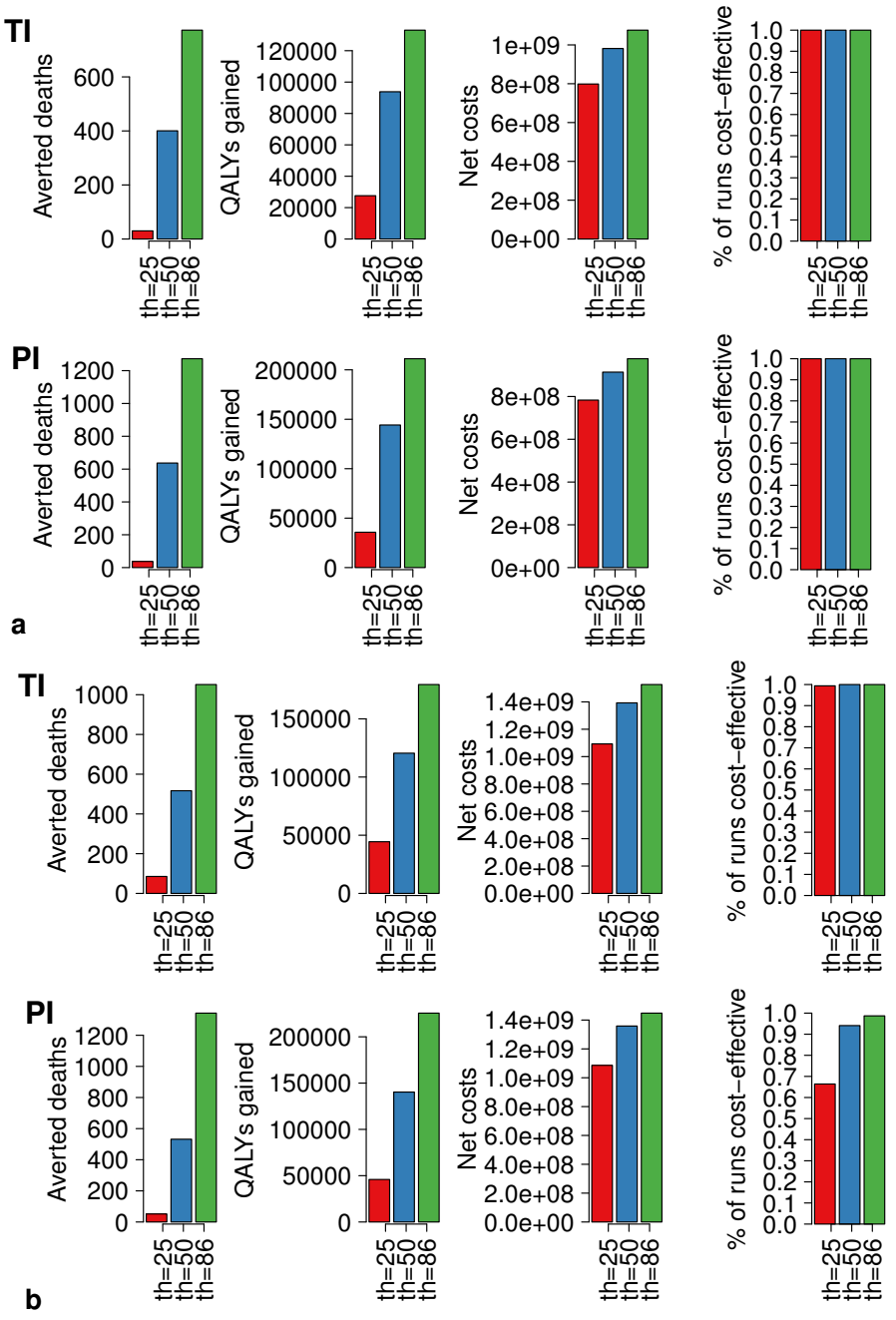

Figure B.5: **Cost-effectiveness analysis results, under the base case.** Median number (over 2,000 stochastic simulations) of the averted deaths, QALYs gained, net costs and percentage of runs that are cost-effective, for the three considered time horizons as obtained by using model TI (top row) and PI (bottom row). **a** "Zoster50" scenario vs no vaccination **b** "Worst-case+Zoster50" vs no vaccination.

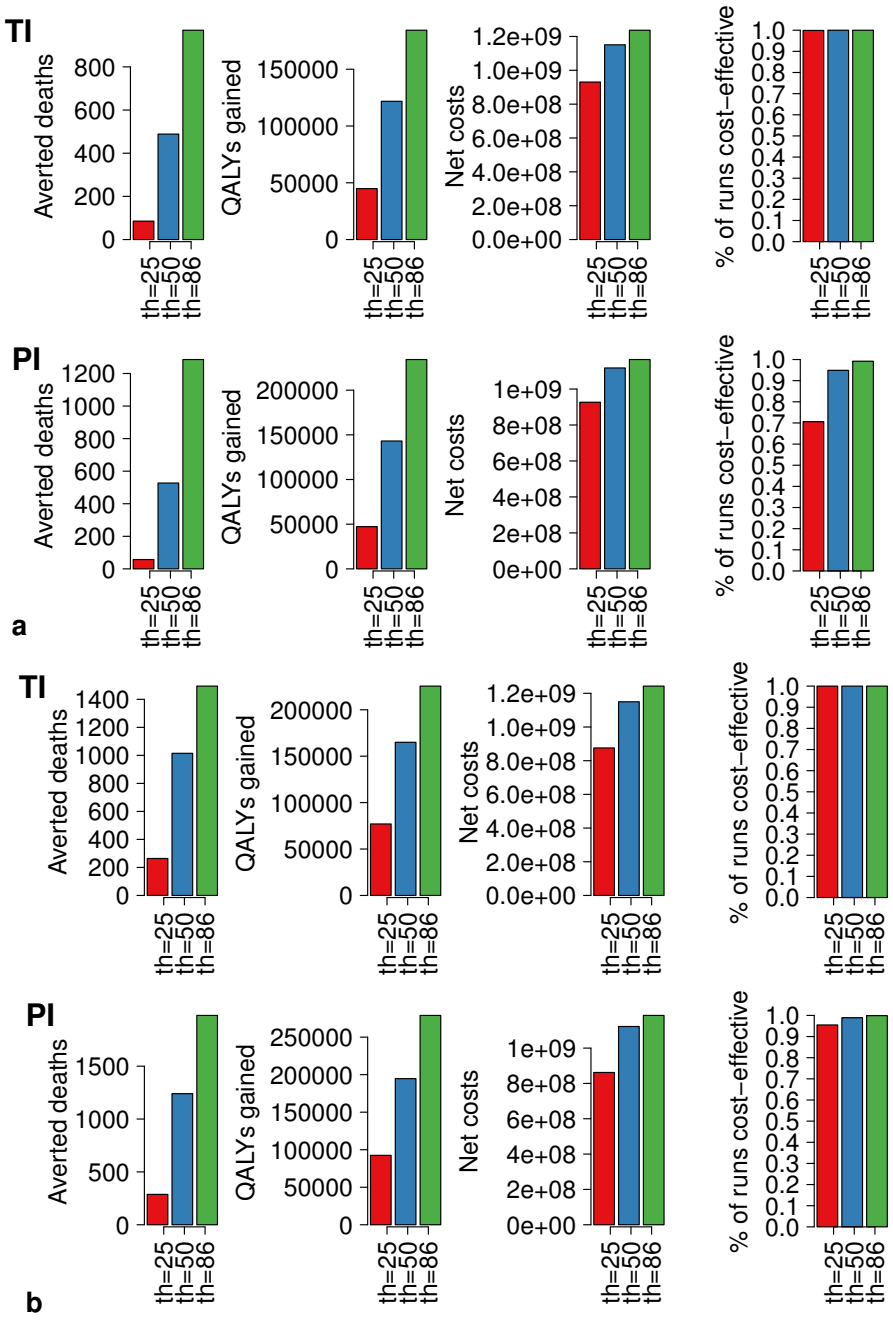

Figure B.6: **Cost-effectiveness analysis results, under the base case.** Median number (over 2,000 stochastic simulations) of the averted deaths, QALYs gained, net costs and percentage of runs that are cost-effective, for the three considered time horizons as obtained by using model TI (top row) and PI (bottom row). **a** "Best-case+Zoster50" scenario vs no vaccination **b** "Best-case+Zoster65" vs no vaccination.

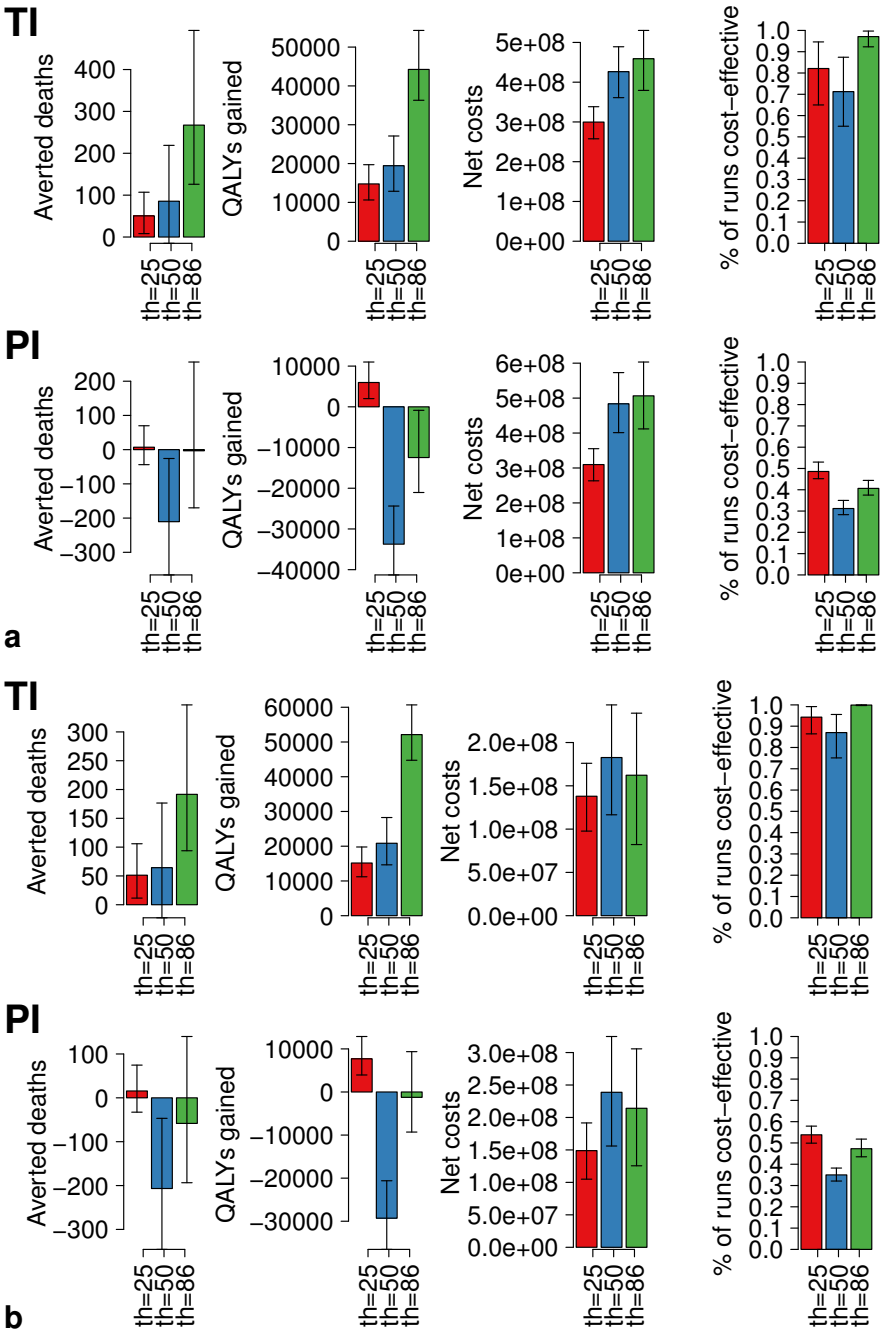

Figure B.7: **Cost-effectiveness analysis results for varicella vaccination alone vs no vaccination, under the PSA.** Median number (over 2,000 stochastic simulations) of the averted deaths, QALYs gained, net costs and percentage of runs that are cost-effective, for the three considered time horizons as obtained by using model TI (top row) and PI (bottom row). **a** Worst-case scenario **b** Best-case scenario

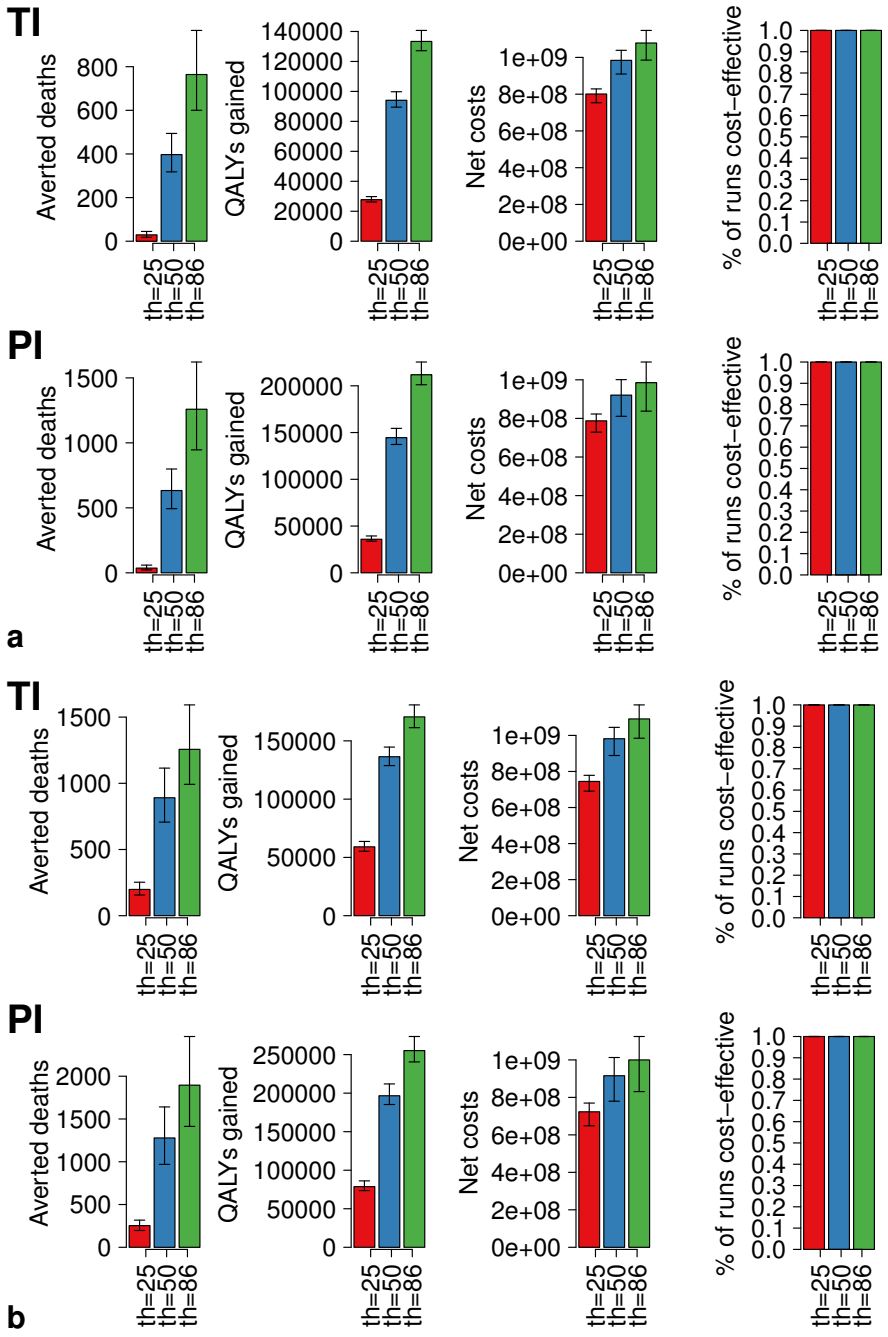

Figure B.8: **Cost-effectiveness analysis results for HZ vaccination alone vs no vaccination, under the PSA.** Median number (over 2,000 stochastic simulations) of the averted deaths, QALYs gained, net costs and percentage of runs that are cost-effective, for the three considered time horizons as obtained by using model TI (top row) and PI (bottom row). **a** Zoster50 scenario **b** Zoster65 scenario

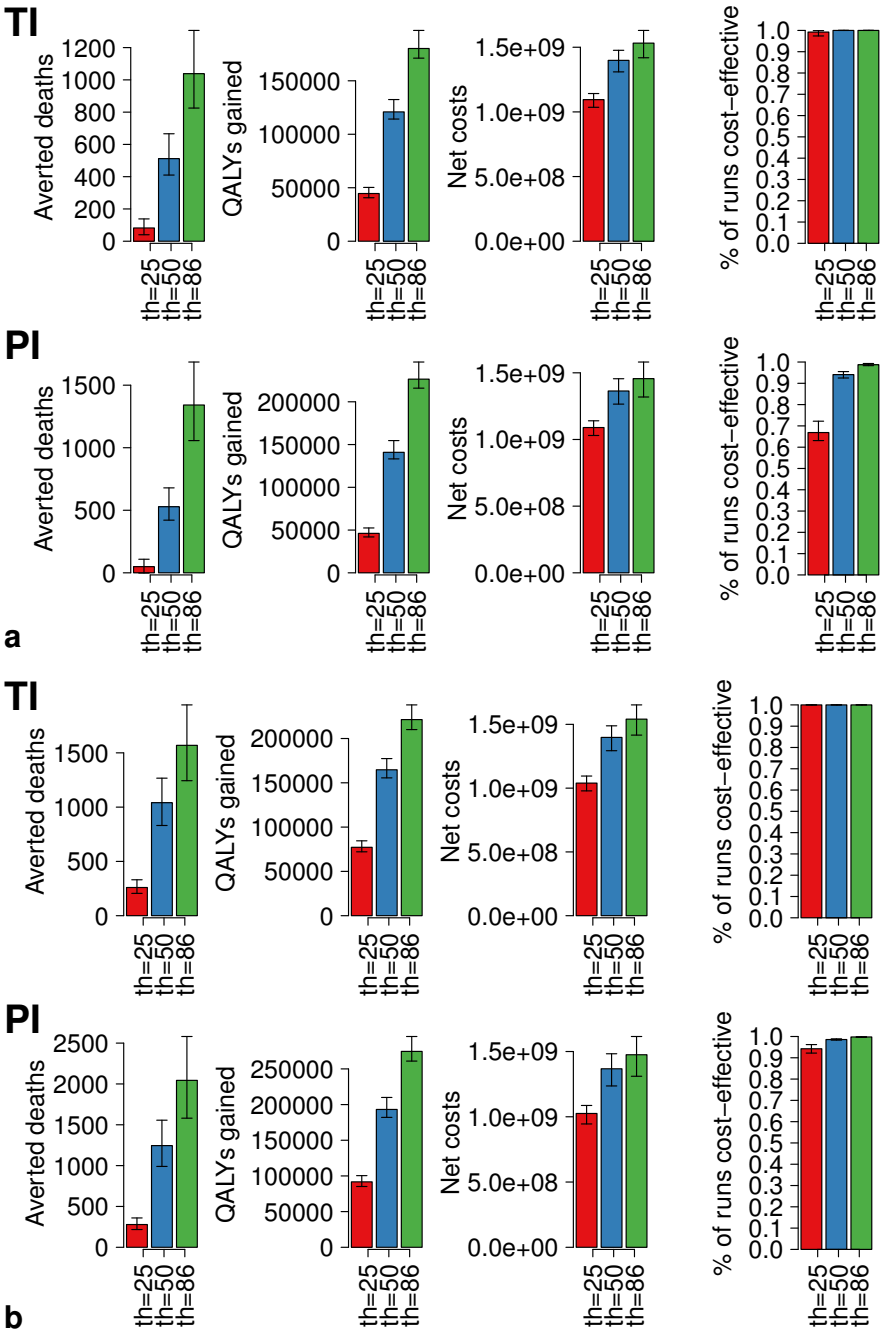

Figure B.9: **Cost-effectiveness analysis results for varicella and HZ vaccination vs no vaccination, under the PSA.** Median number (over 2,000 stochastic simulations) of the averted deaths, QALYs gained, net costs and percentage of runs that are cost-effective, for the three considered time horizons as obtained by using model TI (top row) and PI (bottom row). **a** Zoster50+Worst-case scenario **b** Zoster65+Worst-case scenario

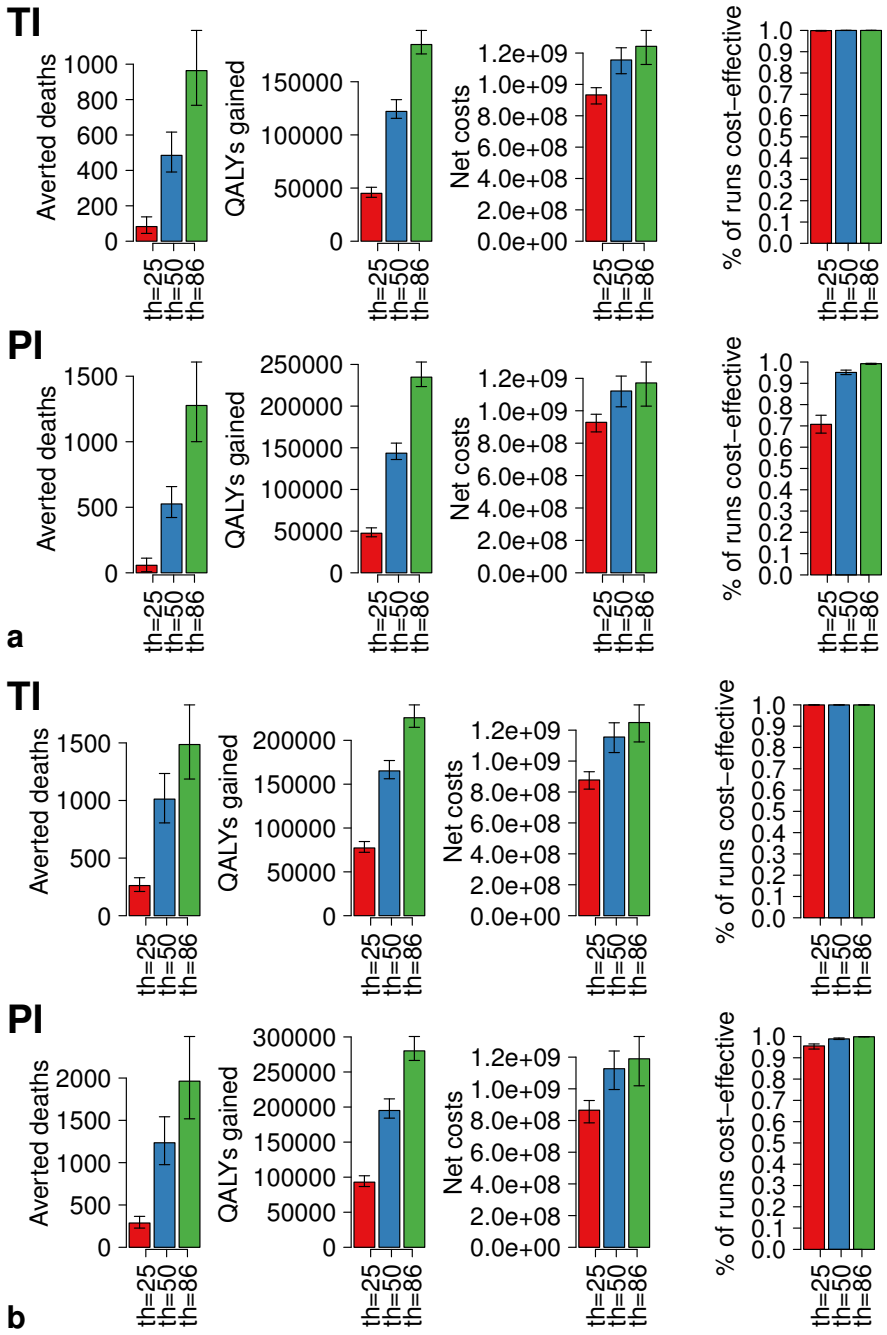

Figure B.10: **Cost-effectiveness analysis results for varicella and HZ vaccination vs no vaccination, under the PSA.** Median number (over 2,000 stochastic simulations) of the averted deaths, QALYs gained, net costs and percentage of runs that are cost-effective, for the three considered time horizons as obtained by using model TI (top row) and PI (bottom row). **a** Zoster50+Best-case scenario **b** Zoster65+Best-case scenario

Chapter 3

Detecting the signature of spontaneous human responses to a pandemic threat

3.1 Background

In March 2009 a novel H1N1pdm influenza virus emerged in Mexico and started spreading globally [106]. The first cases in Europe, mainly travellers coming back from infected areas (Mexico and US), were recorded at the end of April 2009 [107]. At the beginning of June 2009, more than 70 countries had been reached by the infection and the World Health Organization declared a pandemic. The spatial diffusion of 2009 H1N1 influenza showed significant geographic heterogeneities [108, 109]; in Europe, the pandemic progressed from West to East [110, 111] and about 66% of cases affecting European countries up to the end of June 2009 were reported in the United Kingdom [112]. Unlike all other European countries, which were characterized by moderate transmission in spring and summer and by a single fall/winter wave, in 2009 the UK experienced two waves, the first one in late spring/summer and the second one in the fall. The intensive air connections between the UK and the US and the late closure of schools for summer holidays have been identified as the main determinants of the first wave [111, 113]. Data collected since 2009, together with modeling studies, have provided a clear picture of H1N1 epidemiology at the national and international scale. However, some notable patterns of spread at the sub-country level warrant further investigation. In particular, during the first wave of the pandemic, H1N1 cases were heterogeneously dispersed within England. The regions of London and West Midlands experienced early large school-based outbreaks [114, 115] and a more rapid increase of GP consultation rates for ILI with respect to the other regions [107] - this evidence was later confirmed by a cross-sectional serological study [116]. On the contrary, the second wave spread in a much more homogeneous way, affecting most of the regions of England soon after the reopening of schools after the summer break [117]. Here we aim to investigate whether the observed spatiotemporal dynamics of the epidemic was shaped by spontaneous behavioral changes.

The effect of human response to an epidemic/pandemic threat has been widely theorized in the last decade [29]. In the 2014-2015 Ebola epidemic in West Africa, its impact has been clearly visible, especially in terms of population compliance to the intervention strategies proposed by health authorities (e.g., contact tracing and safe burial practices), which have been the main drivers of the elimination process [118, 119, 120]. For instance, in some districts of Guinea, the local population did not allow contact tracers to follow identified contacts of index cases [121]. Moreover, human behavior has the capability to drastically alter the mobility pattern of the population (see for instance the marked drop of air flow to the Hong Kong region recorded during the 2003 SARS epidemic [122]). However, the role of human response in shaping the epidemic spread is not always as immediate to detect. In particular, providing quantitative estimates of the effects of human behavior on the spread of influenza represents a hard challenge. This is the reason why most of the works aimed at estimating these effects are mainly theoretical [123, 124, 125, 126].
In this work a model-based Bayesian analysis of the spread of 2009 H1N1pdm influenza virus provides evidence on human response to the pandemic threat and quantifies its effect in shaping the epidemic spread.

3.2 Methods

3.2.1 The model

The model used in this work is a stochastic spatially-explicit individual-based model for England adapted from previous models developed for Europe [22, 111, 127]. The model can be seen as the combination of two different layers: i) a model of the socio-demographic structure, ii) an infection transmission model.

Sociodemographic model

The socio-demographic structure of the model is the same as in [65] and we refer to this work for all technical details.
Briefly, socio-demographic data on the age structure of the English population are used to generate about 47 millions individuals, which are distributed in a grid of 3894 cells proportionally to the population density, according to the Gridded Population of the World, version 3 (GPW v3). Each individual is characterized by an age, belongs to an household and - if of school age - to a school.
Households are allocated using an heuristic model previously proposed in [65], which reproduces data on household size, composition and age of households member by size specific for the United Kingdom, as provided by the Statistical Office of the European Commission (`http://ec.europa.eu/eurostat`), and preserves realistic age difference between members of the same household. As shown in [65], the age distribution of the population, the household size, the household composition by size and the age distribution of household members

by size simulated by using this model match the observed ones.
Schools are allocated proportionally to population density and the size is determined using data on school size by level specific for the England [128]. School age children (5-18 years) are assigned to a school of the corresponding education level using the resource competition algorithm introduced in [129], which accounts for the population density of the considered geographic area. The mean home-to-school distance obtained by using this algorithm results 4 km, in perfect agreement with the estimate provided by the Department of Transport [130].

Seeding of infection

The number of cases imported at each time-step is based on the actual time-series of travel-related cases reported in [131]. Some studies on 2009 H1N1 pandemic influenza have highlighted that H1N1 importations in the early phase of the 2009 pandemic were strongly correlated with air passenger flows. We thus determine the geographic location of imported cases by randomly distribute cases among all English regions (i.e. NUTS1) proportionally to air passenger arrivals. In particular, we use data on incoming air passenger flows during the period April-June 2009 provided by the Civil Authority [132] to determine the fraction of incoming passengers specific for each region of England. As shown in Figure 3.1, the fraction of incoming passengers is not proportional to population density, with Greater London accounting for about 70% of the total volume. For each time step, the number of imported cases is distributed among regions proportionally to the corresponding fraction of incoming passengers and, within each region, cases are assigned to a cell proportionally to population density.

Infection transmission model

The transmission of influenza infection follows a discrete-time individual-based SEIR model with time step $\Delta t = 1$ day. The model explicitly considers transmission in schools, households and "other settings", which accounts for contacts occurring in the general community (e.g. workplaces, means of transports, free-time activities).
At any time step of the simulation each susceptible individual i has a probability $p_i = 1 - e^{-\Delta t \lambda_i(t)}$ of becoming infected. The probability of infection, which is re-computed at each time step, depends on the individual risk of infection $\lambda_i(t)$ that takes into account the contribution of infection sources in each of the considered settings:

- contacts with infectious members of his/her household;
- contacts with infectious school-mates (if individual i is of school age);
- contacts with infectious people in "other settings".

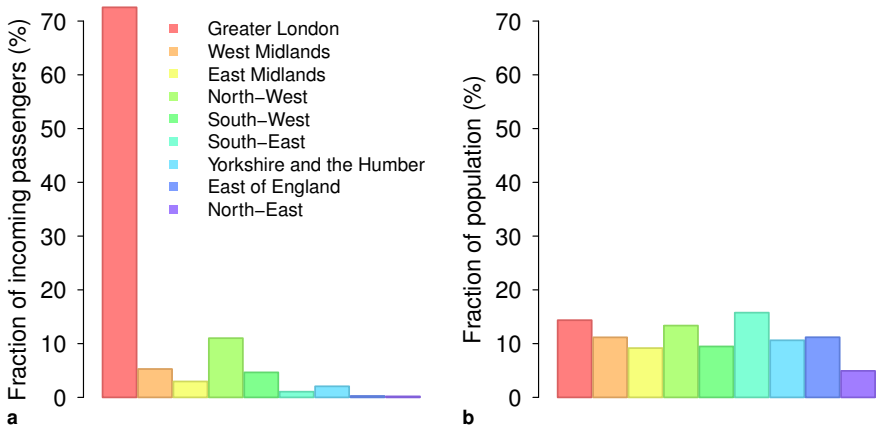

Figure 3.1: **Incoming air passengers and population density.** **a** Fraction of incoming air passengers by region (%) based on data provided by the Civil Aviation Authority [132] on incoming air flows by airport. **b** Fraction of population by region (i.e. NUTS1) of England as obtained from Eurostat.

Transmission in “other settings" explicitly depends on the geographical distance, which is modeled by a kernel function K.
More precisely, the risk of infection is defined by the following equation:

$$\lambda_i = \rho(a_i)\left[\sum_{\{k=1,\dots,N|H_k=H_i\}} \frac{I_k\beta_h}{(n_i-1)^\delta} + \sum_{\{k=1,\dots,N|S_k=S_i\}} \frac{I_k\beta_s}{m_i} + \sum_{\{k=1,\dots,N\}} \frac{I_k\beta_r K(d_{ik})}{\sum_{\{j=1,\dots,N\}} K(d_{jk})}\right]$$

where

- N is the population of England.
- H_i is the index of the household of individual i, n_i is the number of members of the household and $\delta = 1.63$, as estimated in [133].
- S_i is the index of the school where individual i eventually studies and m_i is the school size.
- $I_k = 1$ if individual k is infected, 0 otherwise.
- a_i is the age of individual i.
- $\rho(a_i)$ is the age-dependent susceptibility to infection. For the sake of simplicity, we divide the population into two groups: children ($a_i < 15$ years), for which $\rho(a_i) = 1$, and adults ($a_i \geq 15$ years), for which $\rho(a_i) = \sigma \in [0, 1]$. We call σ relative susceptibility of adults with respect to children.

- β_h is the transmission rate within households (day^{-1}).

- β_s is the transmission rate within schools (day^{-1}).

- β_r is the transmission rate in “other settings" (day^{-1}).

- $K(d_{ij})$ is a decreasing function of the geographical distance d_{ij} between the households of individuals i and j, modelling human mobility patterns. In particular, we assume that function K has the following form:

$$K(d_{ij}) = \begin{cases} \frac{1}{1+d_{ij}^{\alpha_1}} & \text{up to week 33, 2009} \\ \frac{1}{1+d_{ij}^{\alpha_2}} & \text{from week 34, 2009} \end{cases} \tag{3.1}$$

  i.e. function K is shaped by one parameter during the first wave, α_1, and one parameter during the second way, α_2, where parameters α_1 and α_2 are estimated.

The model explicitly includes school closure for holidays, during which transmission in schools is interrupted and transmission in “other settings" is multiplied by a factor $\tau \in [0, +\infty)$. At any time step Δt of the simulation exposed individuals become infectious at a rate $\Delta t/T_L$, where T_L is the latent period - assumed equal to the incubation period - that lasts on average 1.5 days. Infectious individuals recover at a rate $\Delta t/T_I$, where T_I is the infectious period, which lasts on average 1.6 days. Recovered individuals are assumed to be fully protected.
Simulations are initialized by including age-specific pre-pandemic immunity, as measured in [116]. In particular, for each age class the corresponding proportion of initially immune individuals is randomly sampled over the population.

3.2.2 Model calibration

Model calibration is performed using Markov Chain Monte Carlo (MCMC) sampling applied to the product of the binomial likelihoods ($\mathcal{L}_1$ and $\mathcal{L}_2$) of the age-specific prevalence of H1N1 antibodies observed in England respectively in August 2009 [116] (after the first wave) and in the period January-April 2010 [117] (after the second wave). Data are disaggregated into the groupings of regions specified in Table 3.1.

Table 3.1: Grouping of regions in the datasets used for model calibration.

Calibration dataset	**Group A**	**Group B**
Post-first wave [116]	London	East Midlands
	West Midlands	North-West
		South-East
		South-West
Post-second wave [117]	London	East Midlands
		North-West
		South-East
		South-West
		West Midlands

The model has the following free parameters: three transmission rates (in households, schools and "other settings"), relative susceptibility of adults with respect to children, a multiplying factor for the transmission in "other settings", one parameter regulating kernel $K(d)$ during the first wave (up to week 33, 2009) and one for the second wave (from week 34, 2009 until the end of the epidemic). More specifically, the vector of free parameters is defined as

$$\Theta = (\beta_h, \beta_s, \beta_r, \sigma, \tau, \alpha_1, \alpha_2)$$

and the two likelihood functions are defined as follows:

$$\begin{aligned}\mathcal{L}_1(n_m^{\mathrm{A}}, r_m^{\mathrm{A}}, n_m^{\mathrm{B}}, r_m^{\mathrm{B}}|\Theta) = \prod_{m\in M} & \frac{n_m^{\mathrm{A}}!}{r_m^{\mathrm{A}}!(n_m^{\mathrm{A}} - r_m^{\mathrm{A}})!}(p_m^{\mathrm{A}}(\Theta))^{r_m^{\mathrm{A}}}(1 - p_m^{\mathrm{A}}(\Theta))^{n_m^{\mathrm{A}} - r_m^{\mathrm{A}}} \\ & \frac{n_m^{\mathrm{B}}!}{r_m^{\mathrm{B}}!(n_m^{\mathrm{B}} - r_m^{\mathrm{B}})!}(p_m^{\mathrm{B}}(\Theta))^{r_m^{\mathrm{B}}}(1 - p_m^{\mathrm{B}}(\Theta))^{n_m^{\mathrm{B}} - r_m^{\mathrm{B}}}\end{aligned} \tag{3.2}$$

where

- M is the set of age groups considered in [116];
- n_m^{A} is the number of individuals in the m-th age group observed in Group A in the post-first wave dataset [116];
- r_m^{A} is the number of seropositive individuals (haemagglutination inhibition titre 1:32 or more) in the m-th age group observed in Group A in the post-first wave dataset [116];
- $p_m^{\mathrm{A}}(\Theta)$ is the seroprevalence in the m-th age group simulated by the model with parameter set Θ in Group A after the first wave.

- n_m^{B} is the number of individuals in the m-th age group observed in Group B in the post-first wave dataset [116];
- r_m^{B} is the number of seropositive individuals (haemagglutination inhibition titre 1:32 or more) in the m-th age group observed in Group B in the post-first wave dataset [116];
- $p_m^{\mathrm{B}}(\Theta)$ is the seroprevalence in the m-th age group simulated by the model with parameter set Θ in Group B after the first wave.

The likelihood $\mathcal{L}_2$ is defined analogously to $\mathcal{L}_1$ using the post-second wave dataset [117] and the respective grouping of regions (see Table 3.1).
The posterior distribution of Θ is determined using random-walk Metropolis-Hastings sampling [81]. In particular, we performed 21,000 simulations and considered a burn-in period of 2,000 iterations. We assume no *a priori* knowledge on model parameters (i.e. flat prior distributions), except for β_h, for which we assume a prior uniform distribution in $[0, 2]$. Indeed, when assuming $\beta_h = 2$ we obtain a secondary attack rate in households that is more than twice the one obtained for the same pandemic in [133]. Convergence was checked by considering different starting points and by visual inspection (see for instance Figure 3.2).

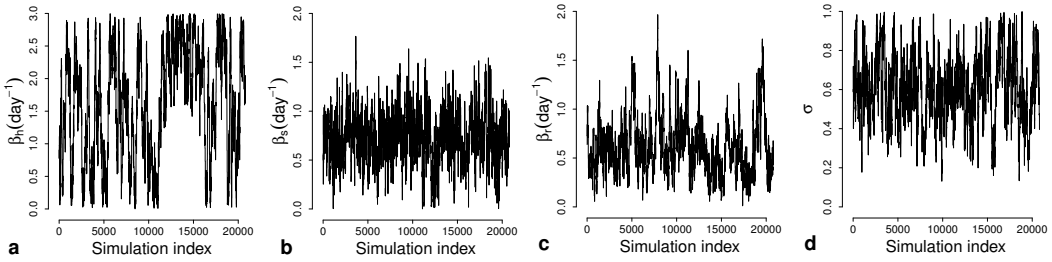

Figure 3.2: **MCMC output at each iteration. a** Transmission rate within households (day^{-1}). **b** Transmission rate within schools (day^{-1}). **c** Transmission rate in "other settings" (day^{-1}). **d** Relative susceptibility of adults (≥ 15 years) with respect to children.

3.3 Results

We analyze seroprevalence data for England at the regional spatial resolution by using the spatially explicit individual-based model of influenza transmission introduced in previous section. In particular, the model explicitly accounts for transmission in household and school, and the transmission in the general community is assumed to depend on the geographic distance between individuals. In this work, we consider two parameters regulating the distance-dependent component of the force of infection (one for each epidemic

wave) and estimate their posterior distributions. This allows us to evaluate whether the distance-dependent component of the force of infection varied over time.

3.3.1 Epidemiological characterization of the 2009 H1N1 pandemic in England.

According to seroprevalence data [116, 117] and our modeling analysis, a substantial proportion of adults and elderly were already protected against H1N1 infection and most of cases in 2009 occurred among school-age children (Figure 3.3**a**/**b**). We found that the low proportion of cases among adults and elderly cannot be the result of contact patterns structure (adults and elderly have relatively less contacts than children and young adults [19, 65] - this is entailed in model structure) and pre-pandemic immunity only. Model simulations suggest that individuals older than 15 years were less susceptible to infection than younger ones: relative susceptibility to infection of adults 0.61 (95%CI: 0.28-0.94), see Figure 3.3**c**. Such estimate is in line with previous findings where adults were found to be approximately half as susceptible as younger individuals [106, 133, 134, 135].
A second major determinant of epidemic spread is the reproduction number (i.e., the average number of secondary cases generated by an index case). We estimated a relatively low transmissibility of the H1N1pdm virus compared to previous pandemics: effective reproduction number 1.45 (95% CI: 1.36–1.54) for the first wave and 1.30 (95% CI: 1.08-1.52) early on in the second wave (Figure 3.3**d**). Both estimates are in line with results reported in previous independent studies and reviewed in [136].

The limited transmissibility of H1N1pdm virus is not sufficient to justify the low overall attack rate measured during the first wave; rather, it is a consequence of the drop of transmission associated to the closure of schools for summer holidays, even though we found an increase in the transmission in 'other settings' (representing all contacts except those occurring in household and school - see Sec. Material and Methods), possibly ascribable to increased activity of students outside schools. In particular, we estimated an increase of 1.2 (95% CI: 0.6-1.9) (Figure 3.3**e**), which is well aligned with findings obtained from the analysis of seasonal influenza in France [137].
Although schools remained closed during summer (as well in fall and winter breaks), they had a major role in the spread of infection. In particular, considering the whole pandemic, we estimated that 17.8% (95% CI: 2.7-35.6) of the infections are linked to contacts at school - this is remarkable considering that the fraction of school-age individuals in England corresponds to only 20% of the population. Moreover, we estimated that 34.7% (95% CI: 4.7-54.1) of infections occurred in households, and 47.5% (95% CI: 19.7-79.2) in 'other settings' (Figure 3.3**f**). This figure remained quite constant over the course of the pandemic (Figure 3.3**f**). Comparable values have been obtained in a previous work on the 2009 H1N1 pandemic in Italy [138], and for past influenza pandemics as well [139].
Results reported so far demonstrate that the model well reproduces the main epidemiological features of the 2009 influenza pandemic in England.

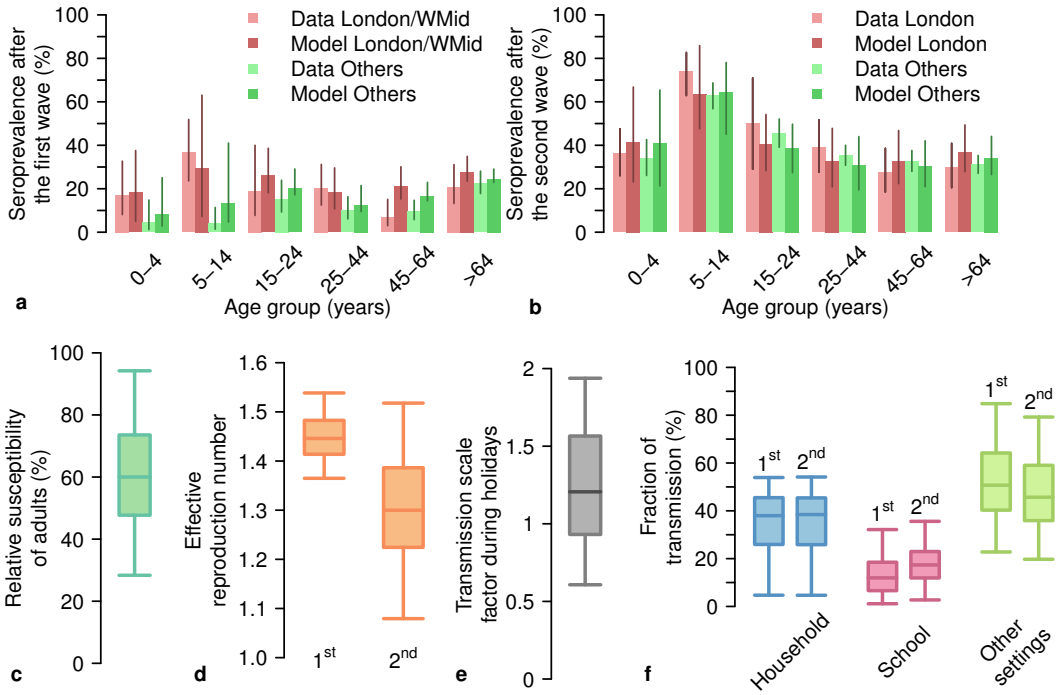

Figure 3.3: **Epidemiological characteristics of the 2009 H1N1 influenza pandemic in England.** **a** Age-specific seroprevalence (mean, 95% CI) by age group and region as reported in [116, 117] (proportion of serum samples with haemagglutination inhibition titre 1:32 or more) and as estimated by the model, as of August 2009 (i.e., at the end of the first epidemic wave). **b** as **a**, but as of January 2010 (i.e., at the end of the second wave). **c** Posterior distribution (median, 50% CI, 95% CI) of the relative susceptibility to infection of adults (individuals aged 15 years or more) with respect to children. **d** Boxplots (median, 50% CI, 95% CI) of the reproduction number as estimated from the exponential growth of model simulations over the early phase of the first and of the second wave (see section C.1 in the appendix for details). **e** Posterior distribution (median, 50% CI, 95% CI) of the increase of the transmission rate in 'other setting' during holidays. **f** Boxplots (median, 50% CI, 95% CI) of the fraction of cases by setting after the first and after the second wave.

3.3.2 Distance dependent component of the force of infection and distance of secondary infections.

The distance dependent component of the force of infection is driven by a kernel function distance $K(d) = 1/(1 + d^{\alpha})$, where d is the geographic distance and α is the parameter regulating the kernel (details in section 3.2). Basically, the kernel defines the distance at which an infectious individual generates secondary infections through contacts in 'other

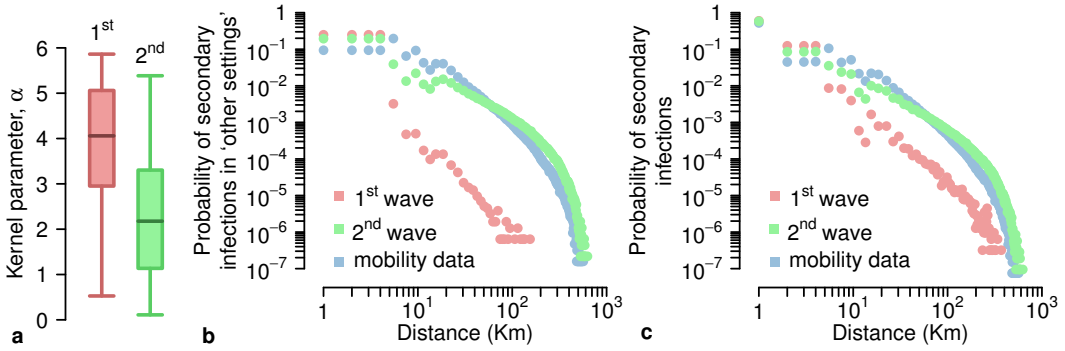

Figure 3.4: **Distance-dependent force of infection.** **a** Posterior distribution (median, 50% CI, 95% CI) of the parameter regulating the distance kernel describing transmission in 'other settings' during the two waves. **b** Probability distribution of the distance at which secondary infections are generated through contacts in 'other settings' as estimated for the two influenza waves and as derived from Ferguson et al. [139], in which the kernel is slightly different and has been calibrated on UK commuting data: $\tilde{K}(d) = 1/(1 + (d/a)^b)$, where d represents the distance, $a = 4$ km, and $b = 3$. Note that the distance of secondary infections depends on both kernel function and population density (which is responsible for the bumps in the curves). The curves were obtained by averaging over 100 simulations run by using median values of the posterior distributions of model parameters; ticks on the x-axis correspond to distances between two consecutive ticks so that for instance 1 corresponds to secondary infections occurring between 0 km (e.g., infection between members of the same household) and 1 km. **c** Probability distribution of the distance at which secondary infections are generated (i.e., the distributions account for infection occurring in all settings).

settings' (i.e., all contacts except those occurring in household and school).

We estimated a significantly lower value (p-value< 0.0001) of the parameter α in the first wave than in the second one: median $\alpha_1 = 4.1$ (95%CI: 0.5-5.9) and $\alpha_2 = 2.2$ (95%CI: 0.1-5.4), respectively (Figure 3.4**a**). This can be interpreted as a high force of infection due to close distance contacts (roughly inside a radius of 5 km) and remarkably lower at higher distances during the first wave, and to a less sharp decline in the fall wave (Figure 3.4**b**). In particular, the mean distance of secondary infections linked to contacts in 'other settings' increases from 2.1 km (standard deviation, SD=0.67) during the early epidemic phase to 10.0 km (SD=29.1) later on. The latter well complies with the value obtained by using an alternative kernel calibrated on commuting data [139], namely 10.9 km (SD=20.1), see also Figure 3.4**b**. If we consider the mean distance of secondary infections irrespective of the setting where the infection occurs, the figure becomes: 1.4 km (SD=3.8) during the first wave, 5.3 km (SD=20.7) during the second wave, and 5.8 km (SD=15.6) if considering the kernel parameterized on commuting data (Figure 3.4**c**). Again, our results highlight that i) secondary infections occurred at a markedly lower distance during the initial phase of the

pandemic, and ii) during the second wave the estimated distances well compare with those obtained by considering a model calibrated on commuting flows.

A possible interpretation for the obtained findings can be a spontaneous behavioral change occurred in the population as a response to the pandemic threat. In particular, the lower distance of secondary infections in the early pandemic phase could have been determined either by a reduction of human mobility or by precautionary behavior aimed at decreasing the number of potentially infectious contacts at high distance from the place of residence/study; both could have been caused by an initial overestimation of the risk of infection. This initial overestimation has been detected in Italy [26] and Mexico [140] as well, suggesting that this behavior could have been common to several countries.

One might argue that a change in the mobility pattern is not necessary to reproduce the observed pattern of spread, or that the observed pattern of spread could be explained by assuming a number of imported cases much larger than observed. Therefore, in order to test the robustness of our findings, we introduced two alternative models (details in section C.2 in the appendix). Briefly, the first alternative model assumes a unique value for the kernel parameter (Model 1); the second one assumes a unique value for the kernel parameter and the daily number of imported cases is assumed to be a free parameter (Model 2). According to the deviance information criterion (DIC), the baseline model should be preferred to both alternative models - DIC= 390.9 for the baseline model, DIC=434.8 for Model 1, and DIC=438.5 for Model 2. Moreover, both Model 1 and 2 show a spatiotemporal spread in some regions of England (e.g., North-East) that appears to be less compliant with observed evidence. Finally, the estimated number of imported cases from Model 2 largely overestimates the observed data. Overall, the performed sensitivity analysis reveals that the spatiotemporal pattern of the 2009 pandemic in England can be explained by a behavioral change remarkably decreasing the distance of secondary infections. According to our modeling analysis, it is more likely that a behavioral response occurred in the early phase of the pandemic only (baseline model); however, it is also possible that the human response lasted for a longer period of time (Model 1).

3.3.3 Geographic spread of the pandemic.

Both the analysis performed by the Strategic Health Authorities [141] and serological data [116, 117] on the first wave reveal very different incidence rates among regions. Such a heterogeneous pattern is confirmed also by model predictions (Figure 3.5**a**). At the end of the pandemic, however, both model predictions reported in Figure 3.5**a** and seroprevalence rates show a rather homogeneous pattern across all England [117] (see also Figure 3.3**b**). The same pattern can be found also by looking at age-specific prevalence rates (Figure 3.5**b**).

At the national level, the pandemic clearly showed two waves of infection (Figure 3.5**c**). According to our modeling analysis and in agreement with previous investigations [142], the second wave was characterized by a markedly higher peak weekly incidence (mean 28.2,

95%CI: 14.2-43.9 cases per 1,000 individuals, to be compared with 10.8, 95%CI: 1.2-26.4 cases per 1,000 individuals for the first wave - Figure 3.5**c**). The crude number of ILI notified cases shows the opposite pattern that, however, is the result of a markedly higher (three to ten times higher) GP consultation rate during the first wave [135, 142]. By

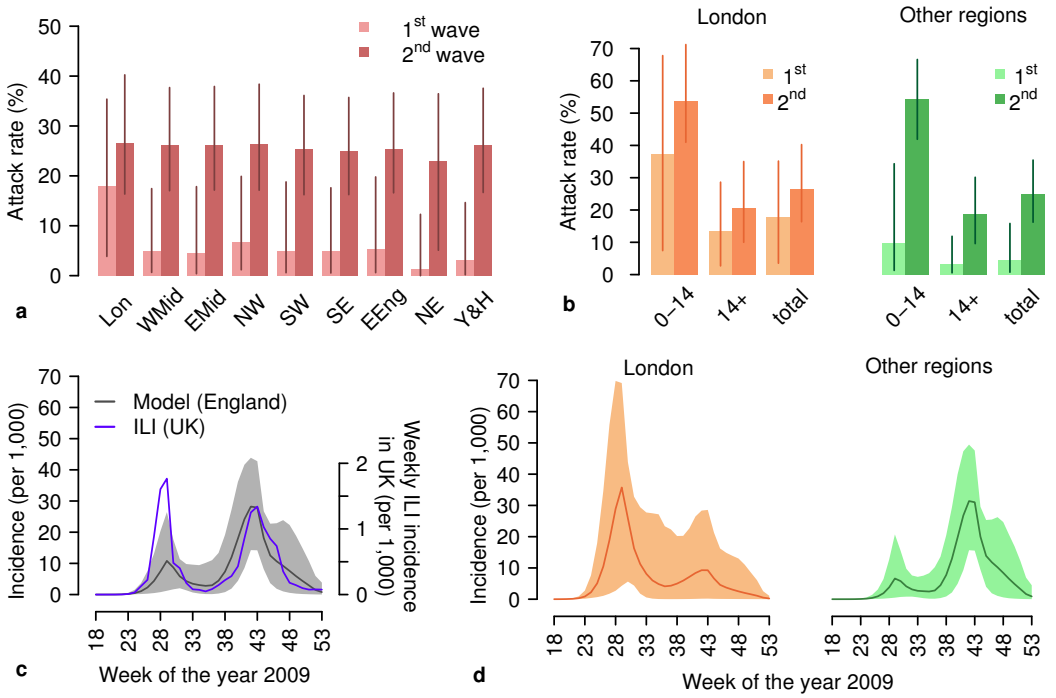

Figure 3.5: **Spatio-temporal dynamics of 2009 H1N1 influenza pandemic in England.** **a** Estimated attack rate (mean and 95%CI) by region of England at the end of the two epidemic waves. **b** Estimated attack rate by age (mean and 95%CI) in London and in all other England regions at the end of the first wave and at the end of the second wave. **c** Weekly incidence of new reported ILI cases in the UK [111] and weekly incidence of new infections estimated by model simulations (mean and 95%CI) for England. Note that the comparison can be considered representative for the timing of the epidemic only, and not for the absolute magnitude of incidence, as: i) ILI cases refer to the entire UK, and ii) they are affected by underreporting (which was estimated to be remarkably lower during the first epidemic wave [135, 142] than in the second one). **d** Weekly incidence of new infections in London and in all other regions of England (grouped together) as estimated by the model (mean and 95%CI).

looking more closely at the sub-national scale, the observed dynamics at the national scale is determined by the sum of two very different dynamics at regional level: London suffered a major epidemic wave during the summer and a more moderate one during the fall; the opposite pattern can be observed in the other regions of England (Figure 3.5**d**). Such a

pattern is clearly visible also by looking at the dynamics of weekly and cumulative incidence at the resolution of single geographic cell (Figure 3.6**a**/**b**). In particular, it is apparent that the fall wave (school reopening, week 37, 2009) spread on the territory in a remarkably more homogeneous way and at a higher pace than the summer one. These patterns are driven by the detected change in the distance of secondary infections: a lower distance during the early phase corresponds to a lower rate of spatial diffusion of the infection, while the estimated larger value in subsequent phases corresponds to a quicker and much more homogeneous geographic spread.

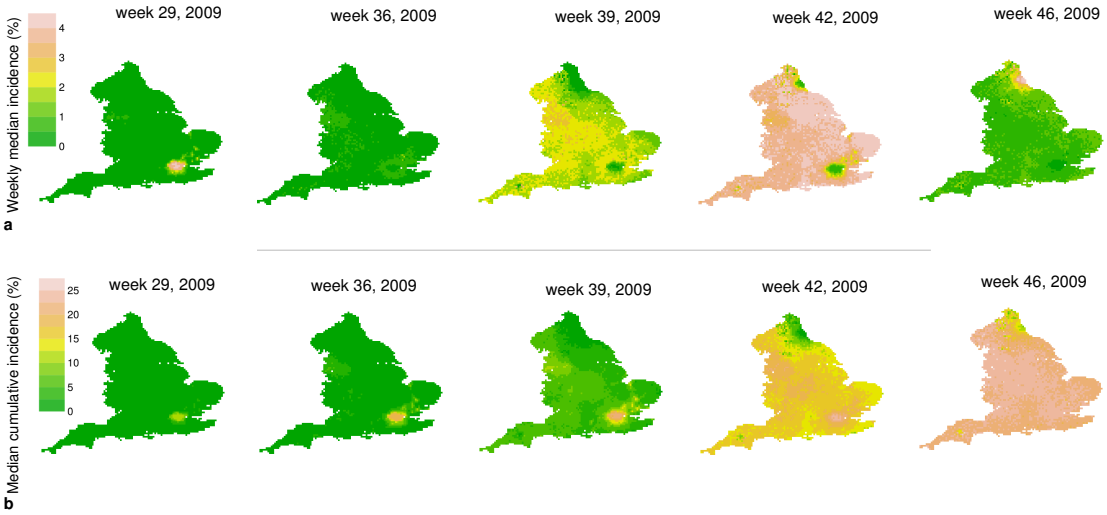

Figure 3.6: **Simulated geographic spread of 2009 H1N1 influenza pandemic in England.** **a** Simulated weekly incidence of new infections in each single cell (median over 2,000 simulations). **b** Simulated cumulative weekly incidence of new infections in each single cell (median over 2,000 simulations).

3.4 Discussion

In this work we use a computational model embedded in a Bayesian framework in order to detect the effect of spontaneous behavioral changes in the dynamics of the 2009 H1N1 influenza pandemic in England. Our findings provide a clear picture of the epidemiology of the 2009 pandemic, well compliant with the existing knowledge. Children younger than 15 years were the most affected age group; our analysis confirms that this was due not only to their lack of pre-existing immunity and their higher contact rates [19, 65], but also to their higher (about twice as much) susceptibility to infection than adults. We estimated a relatively low value of the reproduction number (around 1.4) in both waves. We also provided

quantitative estimates of the proportion of cases by setting that is crucial for determining the effectiveness of intervention options, such as school closure [143], highlighting a major contribution of schools to the overall transmission. Consistently with serum sample data [117], we estimated a highly heterogeneous geographic spread across England: i) London shows a marked first wave, while other regions show little evidence of a summer wave, ii) the fall wave is remarkably lower in London than elsewhere, and iii) London shows a slightly higher attack rate at the end of the epidemic.

What is notably novel in this study is the detection of a clear signal that the distance at which secondary infections were generated changed over time. In particular, during the initial phase of the epidemic, when attention of the public on the pandemic threat was higher, we found that infections were generated at a much closer distance than later on (i.e., 1.4 km vs 5.3 km). In the following phases, we estimate that the radius of diffusion of the epidemic became comparable with the distance traveled by commuters. A possible explanation for such a pattern lies in behavioral changes, spontaneously emerged as a response to the pandemic threat. During an emerging epidemic people may engage precautionary behaviors that alter the transmission dynamics of the disease [126, 144]. According to survey results [145, 146], in the case of an influenza pandemic a large proportion of people is willing to avoid crowded places, especially public transportation. For instance, a worker who usually commutes to reach his workplace, would likely continue to travel the same distance to reach the workplace even during a pandemic, but he might decide avoiding crowded environments near the workplace (such as pubs, and restaurants) where he usually goes outside business hours. Our findings suggest that in the early phase of the pandemic behavioral changes may have led to either a reduction of mobility in absolute terms or, more likely, to a decrease in the number of potentially infectious contacts at high distance from the place of residence/study. Therefore, our results support the idea already presented in [147] that the geographic spread of influenza might be inaccurately described by raw commuting fluxes.

Our results, although pointing in the direction of a human behavioral adaptation to a pandemic threat, are not conclusive. In fact, other factors could have been responsible for the changes in the force of infection that we measured. For instance, the infection was mainly spread by children, who are used to travel less frequently and at shorter distances than adults [148, 149, 150]; this would entail a more marked decline of the force of infection with distance. However, this argument does not provide any explanation for the difference in the estimated kernel parameters for the two waves, especially if considering that seroprevalence data [116, 117] clearly show that children were the most affected age group in both waves. Climatic factors such as absolute humidity have been shown to influence the transmissibility of the virus [151]. However, given the limited size of the considered geographic area and the relative homogeneity of the territory in this respect, we do not expect climatic factors to be responsible for the observed highly heterogeneous geographic pattern of the first wave and for the subsequent homogeneous spread. Clearly, data quality and (un)availability (far from being flawless) impose further limitations to the model. First, concerning the initial seeding of the infection, we compute the region-specific probability

of importing cases by using data on the total volume of incoming passengers by airport, disregarding their origin and final destination. Second, our model neglects cross-border infections that, given the geographic features of England, may have been imported from Scotland or through the Channel Tunnel, i.e. from France and continental Europe. However, we do not expect them to have played a major role in the spread of the 2009 pandemic as England was the first European region with sustained transmission. A further limitation of the model is that transmission in workplaces is not explicitly modeled; in fact, its contribution is entailed in the 'other settings' component of the force of infection. However, the contribution of workplaces during the 2009 pandemic has been shown to be marginal [138]. Finally, it should be remarked that describing infection transmission through a distance kernel neglects most details of human mobility. However, it has been shown that it gives an adequate description, especially in a setting like that of the UK [152].

In conclusion, our results help shedding light on the epidemiology of the 2009 pandemic and provide a possible explanation for the initially heterogeneous spatial spread of the epidemic within England, followed by a highly homogeneous one. This analysis calls for a deeper understanding of human interactions and movements under the pressure posed by an epidemic threat. As preparedness plans for pandemic influenza are currently based on mobility patterns not accounting for human response, we believe that our findings would be instrumental for the design of more effective control strategies.

Appendix C

C.1 Computation of $R_{\mathbf{e}}$

The effective reproduction number R_{e} represents the average number of infections generated by an infectious individual in a partly immune population. We estimated its value in the first and in the second wave from the corresponding epidemic growth, following a technique already employed in the literature [111, 139, 153]. Briefly, the exponential growth rates r_1 and r_2 of the two epidemic waves are estimated by fitting a linear model to the logarithm of the incidence in the corresponding time window (week 18 to week 33 for the first wave, from week 34 on for the second). Then, the effective reproduction number is computed as:

$$R_e^i = (1 + r_i T_L)(1 + r_i T_I) \qquad i \in \{1, 2\}$$

where T_L = 1.5 days and T_I = 1.6 days are respectively the average duration of the latent and infectious period.

C.2 Alternative models

We considered two alternative models in order to evaluate whether the difference in the kernel parameter between the two epidemic waves that we obtained with the baseline model is necessary to interpret the spatiotemporal pattern observed in the 2009 H1N1 pandemic in England.

The first model (Model 1) is analogous to the baseline except for the assumption of a unique kernel parameter for both epidemic waves (α=α_1=α_2). The model was recalibrated and we obtained a median estimate for the kernel parameter α = 3.56 (95%CI: 0.1-5.9), which falls between the values α_1 and α_2 obtained respectively for the first and for the second wave in the baseline model (see Figure C.1**a**). According to Model 1, infections linked to contacts in "other settings" occur at a mean distance of 2.2 Km (SD =1.58), which is greater, but close to the value estimated by the baseline model in the first wave (see Figure C.1**b**). Analogously, as shown in Figure C.1**c**, if we consider infections due to contacts occurring in all considered settings, the mean distance of secondary infections becomes 1.8 Km (SD=5.6), which is similar to the one obtained by the baseline model. In both cases, the mean distances obtained by Model 1 are remarkably lower than the one estimated by using the model

calibrated on commuting flows (values reported in the main text). Results obtained suggest that the spatiotemporal pattern of the 2009 pandemic may be alternatively explained by a slighter behavioral response to the pandemic threat, when compared to the one obtained by the baseline model during the first wave, but lasting the entire course of the pandemic. However, when looking at the geographic spread of infection, results highlight that with such a high value for the kernel parameter α the spread of the epidemic is too slow to reach a peak in all regions, e.g. North-East (see Figure C.1**d**).
In the second model (Model 2) we assume, as we did for Model 1, a unique kernel parameter (α=α_1=α_2), moreover, we add the number of daily imported cases to the set of free parameters in order to complete the time-series of travel-related cases reported in [131] till the end of the pandemic. More specifically, at each-time step (1 day) the number of imported cases is sampled from a Poisson distribution with parameter $\lambda = -1 + 10^{\delta}$, where $\delta \in [0, +\infty]$ and the geographic location of imported cases is determined analogously to the baseline model (see Section 3.2.1). The median estimate for the kernel parameter obtained in this case is α = 4.0 (95%CI: 0.6-5.9), which is very close to the value obtained by the baseline model in the first wave (see Figure C.2**a**). The resulting daily number of imported cases is much higher than the number of travel-related cases reported during the first months of the pandemic (see Figure C.2**b**). According to Model 2, the epidemic spreads very slowly both in the first and second wave, with a mean distance of secondary infections of 2.1 Km (SD=0.65) in “other settings" and 1.8 Km (SD=5.5) overall (see Figure C.2**c**/**d**). Thus, a very high number of imported cases is needed in order to reach an homogeneous spread over the territory before the end of the epidemic. It is worth nothing that, since the distance of secondary infections estimated by Model 2 is very low, even such a massive import of cases does not allow the epidemic to reach a peak in all regions, e.g. North-West (see Figure C.2**e**).
The models presented in this section provide alternative interpretations of the spatiotemporal pattern observed during the 2009 H1N1 pandemic in England. However, according to the deviance information criterion (DIC) the baseline model has a DIC score smaller than both the alternative models and should therefore be preferred (see Table C.1).

Table C.1: Deviance information criterion (DIC) scores as obtained for the baseline and the alternatives models.

Model	**DIC score**
Baseline	390.9
Model 1	434.8
Model 2	438.5

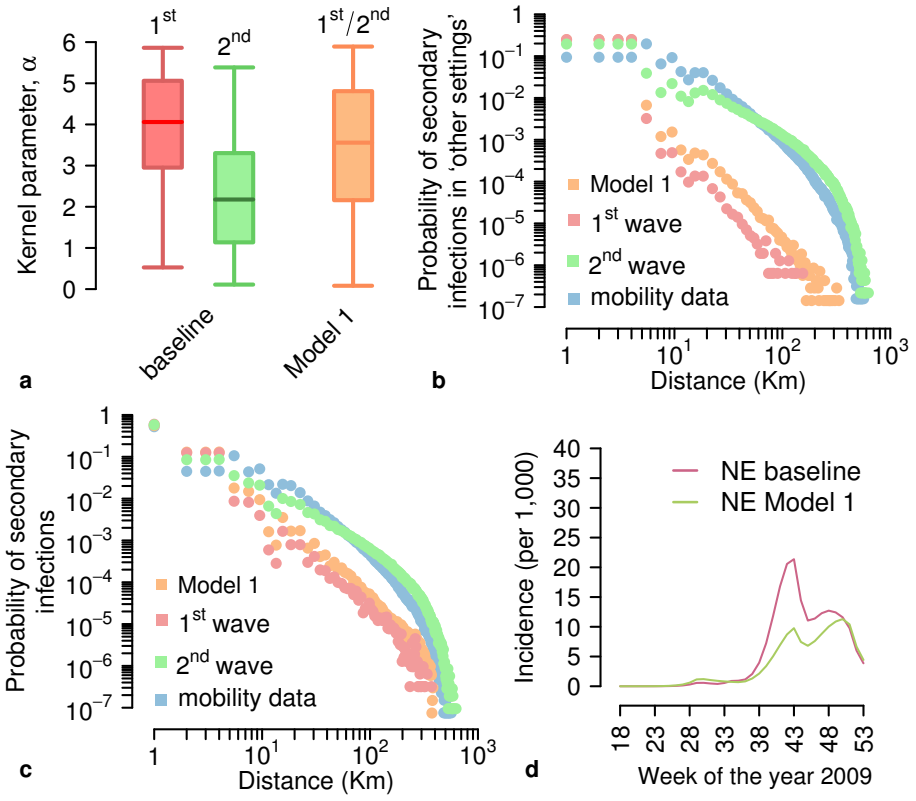

Figure C.1: **Model 1.** **a** Posterior distribution (median, 50% CI, 95% CI) of the parameters regulating the distance kernel describing transmission in "other settings" as obtained by the baseline model and by Model 1. **b** Probability distribution of the distance at which secondary infections are generated through contacts in "other settings" as estimated for the two influenza waves by Model 1, by the baseline model and as derived by using the kernel calibrated on UK commuting data [139]. The curves were obtained by averaging over 100 simulations run by using median values of the posterior distribution of model parameters; ticks on the x-axis correspond to distances between two consecutive ticks so that for instance 1 correspond to secondary infections occurring between 0 km and 1 km. **c** Probability distribution of the distance at which secondary infections are generated (i.e. the distribution accounts for infections occurring in all settings). **d** Mean weekly incidence of new infections in North-East as estimated by Model 1 and by the baseline model.

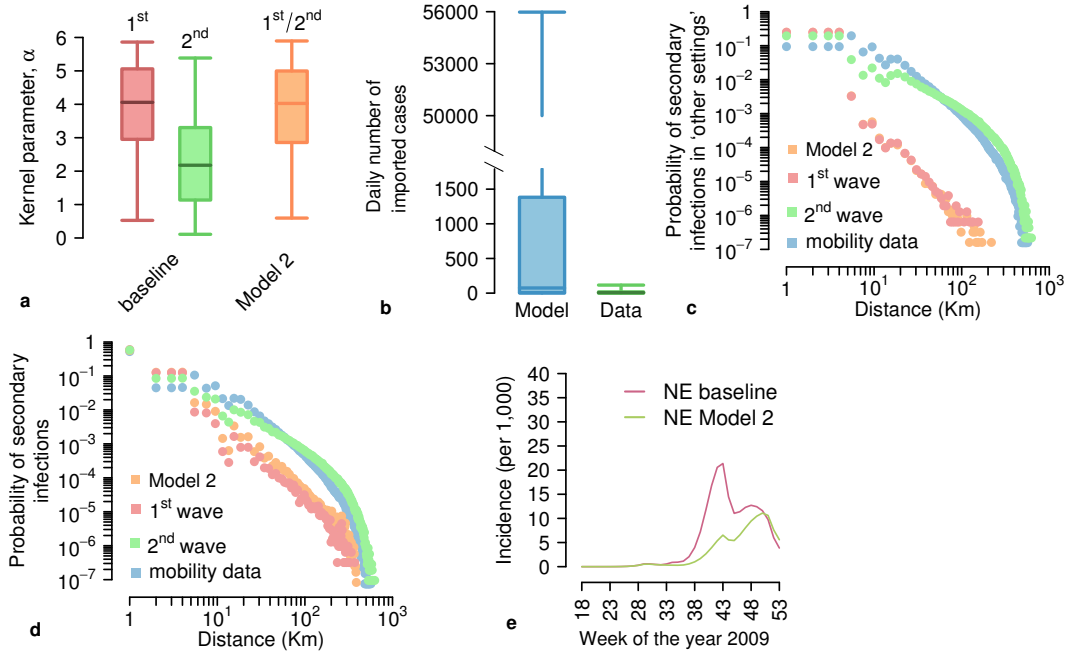

Figure C.2: **Model 2.** **a** Posterior distribution (median, 50% CI, 95% CI) of the parameters regulating the distance kernel describing transmission in "other settings" as obtained by the baseline model and by Model 2. **b** Posterior distribution (median, 50% CI, 95% CI) of the daily number of imported cases as obtained by Model 2 and as reported in the time-series of travel-related cases [131]. **c** Probability distribution of the distance at which secondary infections are generated through contacts in "other settings" as estimated for the two influenza waves by Model 2, by the baseline model and as derived by using the kernel calibrated on UK commuting data [139]. The curves were obtained by averaging over 100 simulations run by using median values of the posterior distribution of model parameters; ticks on the x-axis correspond to distances between two consecutive ticks so that for instance 1 correspond to secondary infections occurring between 0 km and 1 km. **d** Probability distribution of the distance at which secondary infections are generated (i.e. the distribution accounts for infections occurring in all settings). **e** Mean weekly incidence of new infections in North-East as estimated by Model 2 and by the baseline model.

Chapter 4

Measles epidemiology in low circulation settings

4.1 Background

Measles is a highly contagious disease caused by a virus in the paramyxovirus family, which is transmitted through the air or direct contact. Despite the availability of a vaccine since 1976, measles still represents one of the main causes of death among young children. The elimination of measles is possible [154], even though it requires to reach and maintain very high vaccination coverages ($\geq 95\%$) with a two-dose vaccine schedule. The World Health Organization (WHO) established the year 2015 as a target for measles elimination in the European region, including Italy [155]. However, data collected in between July 2015 and June 2016 show that measles elimination is far to be reached and the WHO had to establish a new elimination goal by the year 2020 [156]. Indeed, during the last year 1818 cases of measles were reported by 30 EU/EEA countries and Italy accounted for 31% of them. The introduction of single-antigen measles vaccine in Italy dates back to 1979 and a measles-mumps-rubella (MMR) vaccine has been recommended since the early 1990s. Between 1979 and 1999 routine vaccination with a single dose of measles vaccine at 15 months of age was recommended. In 1999, the age of administration of the first dose of MMR vaccine was lowered to 12 months and a two-dose schedule was recommended in regions that had reached a coverage level of 80% or higher for the first dose. The Italian National Health System was fully decentralized in 2001 and each of the 21 Italian regions is responsible for the organization and delivery of its own health services, including vaccinations, as long as they are in line with recommendations issued by the Ministry of Health. In 2003, a National Plan for the Elimination of Measles and Congenital Rubella (NPEMCR) was implemented and approved by all regions and a two-dose schedule was introduced in all of them; the first at age 12-15 months and the second dose at 5-6 years or 11-12 years [157]. Measles coverage data in children, which are routinely collected by the Italian Ministry of Health, show regional vaccine coverages well below the level necessary to interrupt endemic

transmission. Moreover, the proportion of parents that decide not to vaccinate their children is worryingly increasing. The mean national coverage for the first dose of MMR vaccine has decreased from about 90.0% in 2012 (ranging from 71.5% to 96.7% in the country's 21 regions) to 85% in 2015, with some regional coverages below 80% and the highest one around 90% [158]. As for the second dose, the mean national coverage in children aged 5-6y (2008 birth cohort) is about 83% (regional range, 59.8% to 91.4%) [158]. In order to implement effective vaccination strategies aimed at disease elimination it is important to understand the features of measles epidemiology in settings where the circulation of the virus is low. Indeed, most of the countries that have failed in achieving measles elimination are characterized by transmission dynamics altered by the introduction of vaccination. The long-lasting suboptimal coverages have produced groups of susceptibles, heterogeneously distributed among the population that sporadically cause even large outbreaks. The aim of this study is to provide some insights in this direction by focusing on the Italian setting. In particular:

1. we provide updated estimates of the age-specific seroprevalence profiles in 2014-2015 in the 21 Italian regions, starting from the last available measles seroprevalence data, collected in 2003-2004 [159] and taking into account the regional heterogeneities in measles vaccination coverages observed during the last 10 years.

2. we analyze data on the number of measles cases reported during the period 2013-2015 (source, Italian Ministry of Health) together with the reported epidemiological links to obtain information about the serial interval of measles, i.e. the time between symptom onset of an index case and a secondary case.

3. we use the serial interval estimate to reconstruct the missing epidemiological links.

4. we inform a detailed computational model with the information obtained and we use it to estimate where secondary infections occur and the attack rates that we can expect in different regions.

4.2 Methods

4.2.1 Data

Serological data collected in 2003-2004 [159] and regional MMR vaccine coverages observed from 2004 to 2014 [158] are used to estimate regional serological profiles by age in 2014.
Data reported to the Italian National Surveillance System for measles over the period 2013-2015 are analyzed to estimate the serial interval.
Demographic data on the age distribution of the population in 2014, household and schools structure specific for Italy [91] are used to allocate the population of the computational model. The regional age-specific serological profiles in 2014, based on [158, 159], are used

to initialize the immunological status of the population. The distribution of outbreaks by size obtained after the reconstruction of epidemiological links in data provided by the Italian National Surveillance System is used to calibrate the epidemiological parameters of the model.

4.2.2 Age-specific serological profile of the Italian population in 2014

The last serological profiles by age available for Italy are based on a serosurvey conducted in 2004 [159]. Age-specific seroprevalences are provided in [159] for three geographical groups: Northern regions (we will refer to this macro-region as "North"), Central regions ("Centre"), Southern regions and Islands ("South"). The serological profiles by age for the three macro-regions are reported in Figure 4.1**a**, whereas the region members of each group are reported in Table 4.1.

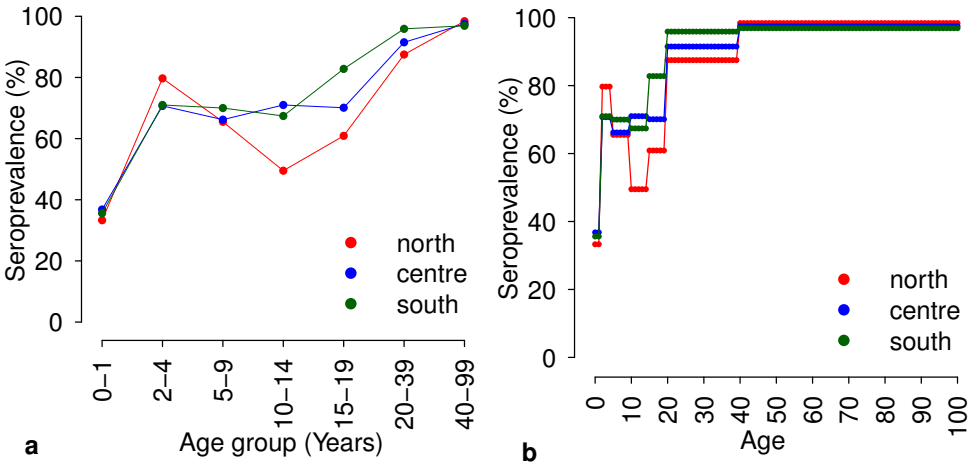

Figure 4.1: **Measles age-specific serological profiles by geographical area in 2004. a** Measles seroprevalence by age-group reported in [159] **b** Measles seroprevalence extended to one-year age groups, based on [159].

Table 4.1: Groups of regions considered in the serosurvey [159].

Macro-region	**Regions**
North	Piedmont, Lombardy, AP Trento, AP Bolzano, Veneto, Friuli-Venezia-Giulia, Liguria, Emilia-Romagna, Aosta Valley
Center	Tuscany, Umbria, Marche, Lazio
South	Abruzzo, Molise, Apulia, Calabria, Sicily, Sardinia, Campania, Basilicata

In order to estimate the age-specific serological profile by region in 2014, we developed a simple demographic model, informed with regional coverage. Specifically, we initialize the age-specific serological profile in 2004 for each region according to [159], by assuming that the aggregated values are representative for each of the one-year age subgroups (see Figure 4.1**b**). Then, for each year (from 2005 on) we:

- shift the seroprevalence by one year in order to account for aging;
- set the seroprevalence of newborns (0 year-old individuals) to 50% in order to account for the immunity conferred by maternal antibodies;
- set the seroprevalence of 1-year old individuals to 0%, in order to account for the waning of maternal antibodies;
- set the seroprevalence of 2-year old individuals equal to the vaccine coverage specific for that year and that region (see Figure 4.2).

4.2.3 Serial interval estimation

During the period from January 2013 to December 2015, 4336 cases of measles were reported to the Italian National Surveillance System. The dataset provides for each case the date, the region of reporting. Moreover, for a subset of cases an epidemiological link to other cases is specified. Data on cluster of size two, i.e. couple of cases for which a certain epidemiological link has been identified in the dataset, can be used to provide an estimate of the serial interval. In particular, by assuming that the date of reporting usually corresponds to the date of clinical onset, the difference between the date of reporting of the primary and the secondary case represents a good approximation for the serial interval. We approximate the distribution of the data with a Gamma distribution with shape α and rate β, by using maximum likelihood [160].

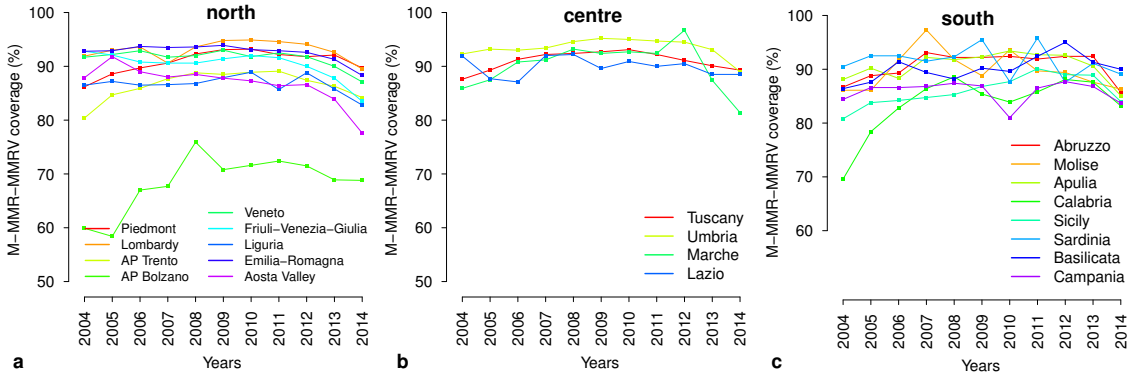

Figure 4.2: **Regional measles vaccine coverages during 2004-2014.** Measles vaccine coverage by region as reported in [158]. The coverage includes the one of measles vaccine administered alone (M) or combined with other vaccines, i.e. measles-mumps-rubella (MMR) or measles-mumps-rubella-varicella (MMRV).

4.2.4 Reconstructing epidemiological links

About 60% of measles cases reported to the Italian National Surveillance System between 2013 and 2015 do not have any information about the source of infection. The reasons for this gap can be various. First, since the data are collected at regional level and then reported to the national authority, the amount of information may depend from the availability of resources to employ in tracing the contacts, which can be very different among regions. Second, the main purpose of case reporting is to have information about the incidence of the disease and not to identify links between cases. Thus, local authorities are free to decide whether performing further epidemiological investigations. Such considerations suggest that the absence of an indication about epidemiological links does not guarantee that the case considered is an index case, but possibly that the link has not been detected (e.g. contacts in the general community between persons who do not know each other). This is confirmed by the fact that a fraction of cases reported in the dataset without no indications on epidemiological links are spatio-temporally compatible with other cases (e.g. black line, Figure 4.3). Moreover, there are clusters of cases identified as disjoint in the dataset, that could represent a unique outbreak (colored lines in Figure 4.3).

The distribution of the serial interval obtained by using the information on clusters of size two is used to infer "probabilistically" the missing epidemiological links. In particular, for each region and each reported case (starting from the second) we consider all the cases which were reported previously and determine whether they are linked or not by applying the following procedure:

- we compute the interval between the two dates of reporting;
- if the time interval lies inside the 95% CI of the estimated distribution of the serial

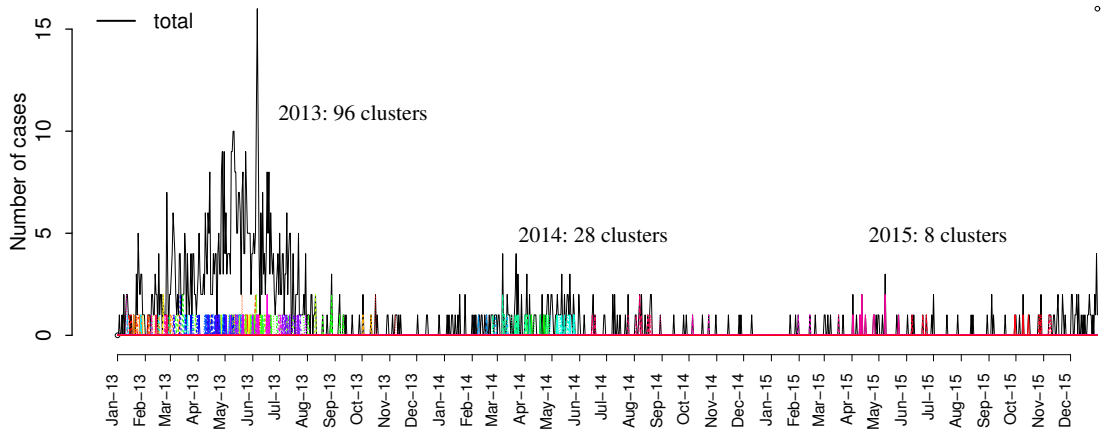

Figure 4.3: **Measles cases and epidemiological links reported in Lombardy (2013-2015).** Number of cases reported in Lombardy by date, from January 2013 to December 2015. Different colors corresponds to different clusters, i.e. groups of cases for which an epidemiological link is reported in the original dataset

interval the two cases are assumed to be epidemiologically linked;

- whenever the case results linked to multiple cases, the epidemiological link is assigned probabilistically (i.e. by random sampling among all compatible links).

This procedure allows us to obtain a new distribution of outbreaks by size that will be used to calibrate the computational model (see next section).

4.2.5 Modeling measles transmission in Italy

The model

The model used is an individual-based, stochastic model of measles transmission in Italy, adapted from previous models developed for other infectious diseases. e.g. influenza, in Europe [22, 111]. Individuals are distributed in households, according to data on household structure specific for Italy, and they are assigned to a school if they are of school-age. In particular we consider four education levels (from kindergartens to high school) and assign each student to a school of the level corresponding to his age. The transmission of infection is modeled through a discrete-time SEIR model (1 day time-step) with an exponentially distributed latent period T_L - assumed equal to the incubation period - and an exponentially distributed infectious period T_I, in such a way to obtain a gamma-distributed generation time of mean $T_L + T_I$. In particular, since under the assumption that all people need the same time to develop symptoms the serial interval and the generation time coincide, we will

set $T_L + T_I$ equal to the mean serial interval estimated from the data (see Section 4.2.3). At any time step of the simulation each susceptible individual i has a probability $p_i = 1 - e^{\Delta t \lambda_i(t)}$ of becoming infected, where $\lambda_i(t)$ is the individual risk of infection and is defined as:

$$\lambda_i(t) = \sum_{\{k=1,...,N|H_k=H_i\}} I_k(t)\beta_h + \sum_{\{k=1,...,N|S_k=S_i\}} \frac{I_k(t)\beta_s}{m_i} + \sum_{\{k=1,...,N|R_k=R_i\}} \frac{I_k(t)\beta_r}{n_k}$$

where

- N is the population of Italy.
- H_i is the index of the household of individual i.
- S_i is the index of the school where individual i eventually studies and m_i is the school size.
- n_i is the number of people living in region R_i.
- $I_k(t) = 1$ if individual k is infected at time t, 0 otherwise.
- β_h is the transmission rate within households (day^{-1}).
- β_s is the transmission rate within schools (day^{-1}).
- β_r is the transmission rate in the "general community" (day^{-1}).

The immunological status of individuals is initialized according to the estimates of the age-specific seroprevalence profile in 2014 obtained for each of the 21 Italian regions (see Section 4.2.2). Model simulation are initialized with one infected individual randomly chosen among the Italian population.

Model calibration

The free parameters of the model are the transmission rates in households, schools and the "general community". Model calibration was performed by using Markov chain Monte Carlo (MCMC) sampling applied to the multinomial likelihood of the distribution of outbreaks by size, as obtained from the cases reported during 2013-2015 with the procedure explained in Section 4.2.4. In particular, for each parameter set $\Theta = (\beta_h, \beta_s, \beta_r)$, we run 10,000 stochastic simulations by seeding infection randomly within the population and we count the number of secondary infections generated in each simulation (if any). After the 10,000 runs, we compute the distribution of outbreaks by size generated by that specific parameter set and the multinomial likelihood as:

$$\mathcal{L}(n_1, ..., n_K|\Theta) = N! \prod_{k=1}^{K} \frac{(p_k(\Theta))^{n_k}}{n_k!} \tag{4.1}$$

where

- K is the number of groups of outbreak sizes considered in the data;
- n_k is the number of outbreaks of size K observed in the data;
- N is the total number of outbreaks observed in the data;
- $p_k(\Theta)$ is the probability of observing an outbreak of size K according to the model with parameter set Θ.

We assume no a priori knowledge (i.e. flat prior distributions) on the transmission rates in schools and in the "general community" (β_s and β_r), whereas for the β_h we assume a prior uniform distribution $[0, 5]$. Indeed, when assuming $\beta_h = 5$, the probability of a susceptible individual living in the same household of a case to escape from infection is about 10^{-4}. There is no need to consider value of β_h larger than 5, indeed, they would indiscriminately result in a probability for a susceptible member of the household of the case of getting infected almost equal to 1.
To determine the posterior distribution of Θ we use random-walk Metropolis-Hastings sampling. In particular, we performed 400,000 MCMC simulation and checked convergence by assuming different starting points and visual inspection. An example of the chains obtained is shown in Figure 4.4.

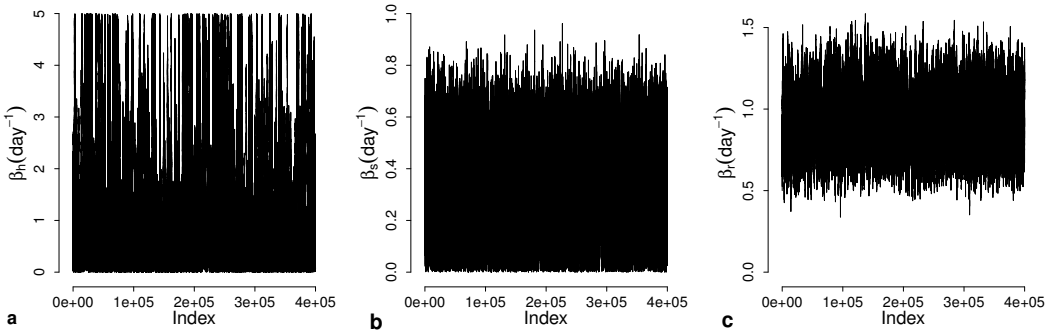

Figure 4.4: **MCMC output at each iteration.** **a** Transmission rate within households (day^{-1}). **b** Transmission rate within schools (day^{-1}). **c** Transmission rate in the general community (day^{-}1).

4.3 Results

4.3.1 Regional age-specific serological profile 2014

As shown in Figure 4.1, serological data collected in 2004 show a lower fraction of measles seropositives in Northern regions than in Central or Southern ones, especially in the age group 10-19 years. By means of the procedure described in Section 4.2.2 we obtain updated

serological profiles by age for each of the 21 Italian regions, which are reported in Figure 4.5 and in the appendix. Results obtained highlight that the measles susceptibility profile by age is strongly heterogeneous among Italian regions. These differences depend first on the starting point, i.e. whether the considered region belonged to the "North", "Centre" or "South" group in the 2004 serological survey. Indeed, regions originally belonging to the "North" group are characterized by a higher proportion of susceptibles among young adults and children with respect to Central and Southern regions, which is a consequence of the shift of the seroprevalence gaps observed in 2004. For instance, in 2014 we estimate for Emilia-Romagna ("North") a fraction of susceptibles between 0 and 35 years of about 27%, compared to 22% in Umbria ("Center") and 20% in Sardinia ("South"), see Figure 4.5, top row. On the other hand, susceptibility of children and infants mainly depends on vaccination

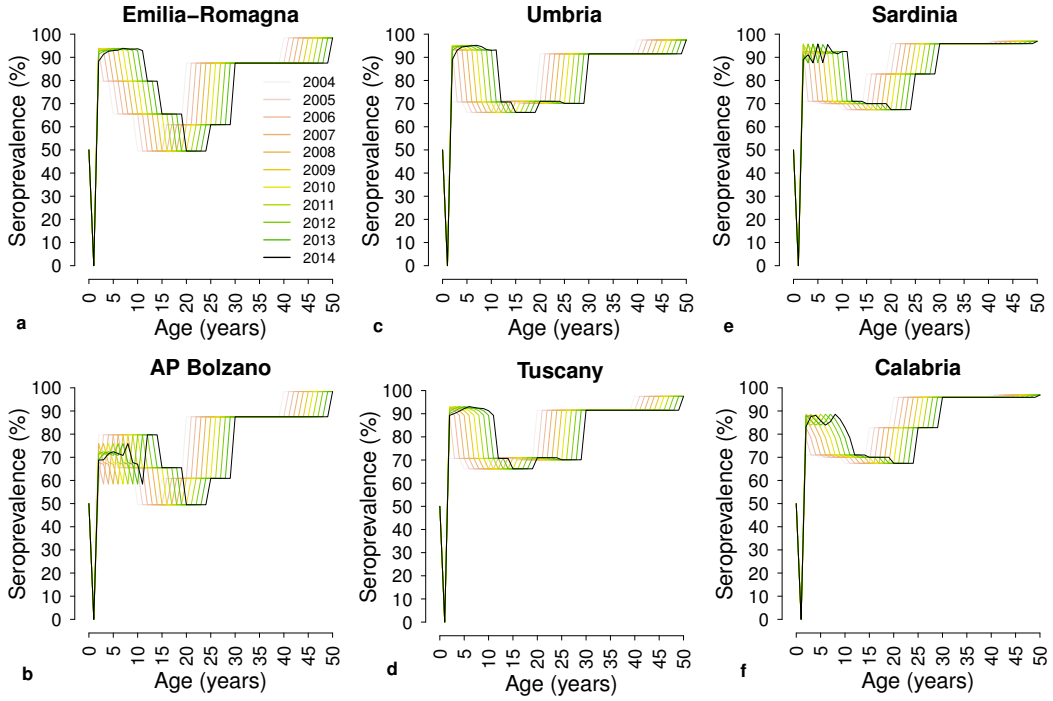

Figure 4.5: **Age-specific measles seroprevalence in 2014.** Emilia-Romagna and AP Bolzano belong to the group "North" in the serosurvey [159], Umbria and Tuscany to the "Centre", Sardinia and Calabria to the "South".

coverages over the period 2004-2014, which in some cases are remarkably different also between regions belonging to the same group. For instance, the proportion of susceptibles individuals of school age (6-13 years) is much higher in the AP Bolzano, 28.6%, than in Emilia-Romagna, 9.9% (see Figure 4.2**a** and 4.5**a**/**b**). On the other hand, the proportion of susceptible school-aged children is more uniform, especially among Central regions (13.3%

in Umbria vs 11.6% in Tuscany), due to less remarkable differences in vaccination coverages during the last decade (e.g. see Figure 4.2**b** and 4.5**c**/**d**).

4.3.2 Serial interval

Figure 4.6**a** shows the distribution of the serial interval obtained by analyzing the clusters of size two in the data (mean: 11.4 days, 95%CI: 3.3-19 days). The best fitting Gamma distribution (red line in Figure 4.6**a**) as obtained by maximizing the likelihood has a shape α = 8.65 and a rate β = 0.76. The mean serial interval according to the Gamma distribution results $\frac{\alpha}{\beta}$ = 11.4 days (95%CI 5.1-20.1 days). This value well complies with previous estimates reported in the literature [161].

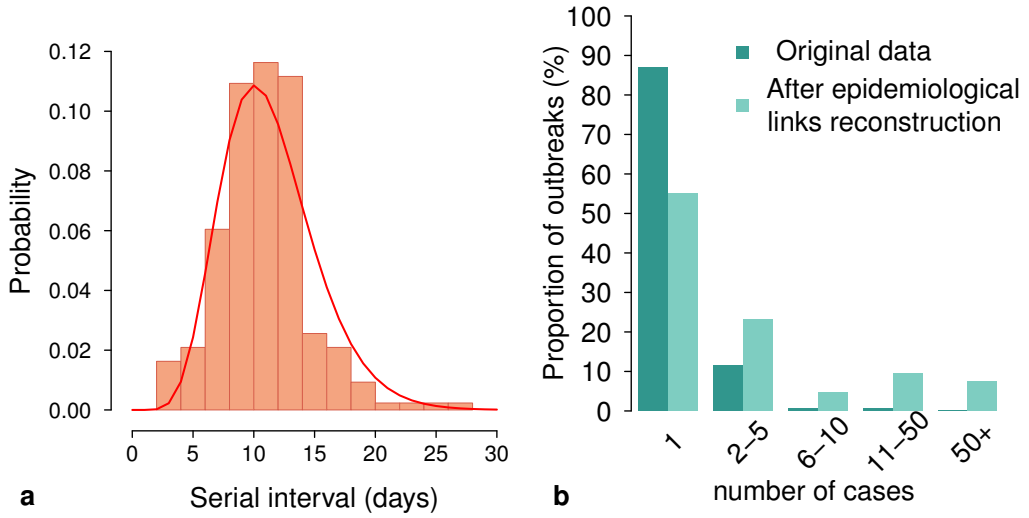

Figure 4.6: **Serial interval and distribution of outbreaks by size.** **a** Distribution of the serial interval of measles as observed in cluster of size two in the dataset (histogram) and as estimated by fitting a Gamma distribution to the data (red line). **b** Proportion of outbreaks by size as observed in the dataset and as estimated by reconstructing the missing espidemiological links.

4.3.3 Reconstructing epidemiological links

The information inferred on the distribution of the serial interval is used to reconstruct missing epidemiological links with the procedure explained in Section 4.2.4. In this way, we obtain for each region new clusterings of reported cases, e.g. compare Figure 4.7 with Figure 4.3, that result in a new distribution of outbreaks by size. The distribution of outbreaks by size in the original dataset is compared to the one obtained after the process in Figure 4.6**b**.

In particular, the number of different outbreaks detected in the original data is 426, of

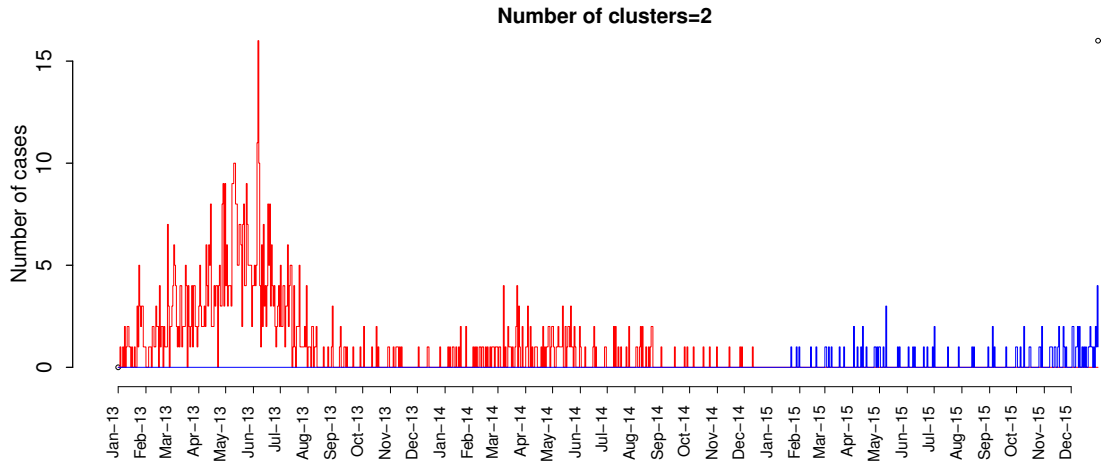

Figure 4.7: **Measles cases reported in Lombardy and inferred epidemiological links.** Number of cases reported in Lombardy by date, from January 2013 to December 2015. Different colors corresponds to different clusters, as obtained by reconstructing the epidemiological links.

which about 85% were outbreaks without secondary infections (i.e. single cases), about 10% consisted of 2 to 5 cases and only a small fraction of outbreaks had more than 5 cases. After the reconstruction of epidemiological links, the 4336 reported cases result grouped in 189 different outbreaks, of which about the 55% consists only of the index case, about the 25% of 2-5 cases and the remaining 20% is distributed among outbreaks with more than 5 cases. It is worth noting that according to our analysis a non negligible proportion ($\approx 8\%$) of the outbreaks consists of more than 50 cases.

4.3.4 Modeling measles epidemiology in Italy

Figure 4.8**a** compared the distribution of outbreaks by size obtained by the computational model and the one estimated in the previous section, which was used for model calibration. In particular, after the introduction of an infected individual within a population characterized by the serological profile by age of Italy in 2014, the model predicts a 49.8% (95%CI 46.1-54.3%) of observing no secondary infections. The probability of generating relatively small outbreaks (1-4 secondary cases) is about 28.5% (26.3-30.3%), decreases to 6.6% (5.5-7-5%) for outbreaks up to 10 cases and 7.4% (5.0-9.0%) for outbreaks up to 50 cases. However, according to the model, 8 in 100 imported cases may trigger outbreaks with more than 50 cases. Our analysis additionally suggests, that - whenever secondary cases are generated - most of them are ascribable to contacts in the "general community" (67.3%, 95%CI 53.3-81.6%), whereas contacts in households and schools account respectively for 17.6% (9.7-22.6) and 15.1% (4.2-27.3%) of transmission (see Figure 4.8**b**). One possible ex-

planation for the marginal role played by transmission in households and at schools is that, in a highly protected population, the fraction of susceptible individuals in these settings may rapidly go to zero.

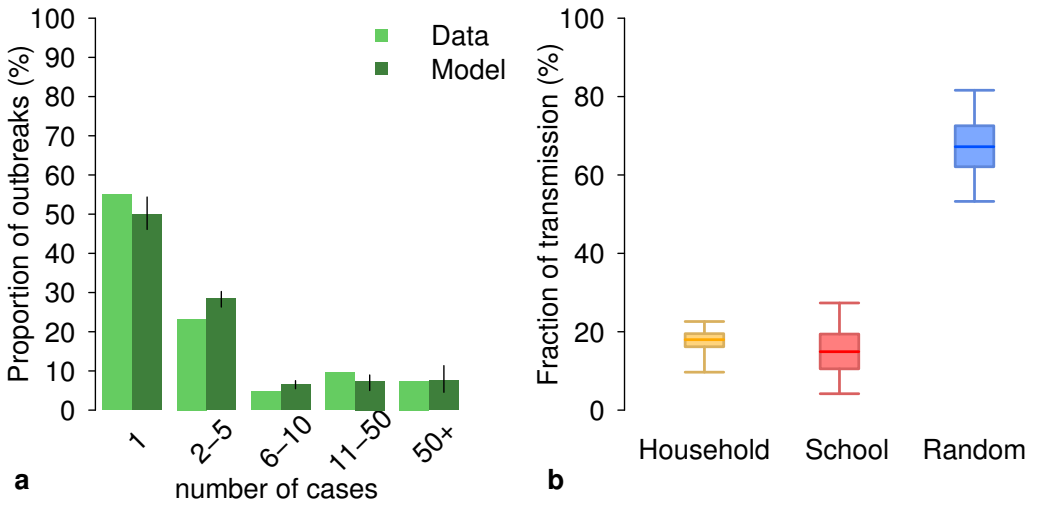

Figure 4.8: **Distribution of outbreaks by size and fraction of cases by setting.** **a** Proportion of outbreaks by size as obtained after the reconstruction of epidemiological links (light green) and as estimated by the epidemiological model (mean, dark green, 95% CI, black). **b** Boxplots (median, 50% CI, 95% CI) of the fraction of cases by setting.

Nonetheless, the probability of an index case to produce large outbreaks depends on the regional susceptibility profiles. For instance the probability of outbreaks with more than 50 cases ranges from 2.8% in Basilicata to 37.5% in the AP Bolzano. Another quantity that is strongly region-dependent is the average attack rate for outbreaks with more than 50 cases, that can range from about 0.3% in Sardinia to 3.4% in AP Bolzano (see Figure 4.9). It is worth nothing that the estimated attack rates could be an overestimation, given that we are assuming homogeneous mixing in the "general community" within each region.

4.4 Conclusions

Measles is a highly infectious disease that is preventable through a vaccine made available in 1976. Mass immunization programs have been implemented since the early 90's in several countries and the WHO had set the year 2015 as elimination goal for measles in Europe. However, data collected in the recent years show that the objective is far to be reached. The countries that have failed in achieving measles elimination, e.g. Italy, are characterized by a low circulation of the virus that is sufficient to cause sporadically large

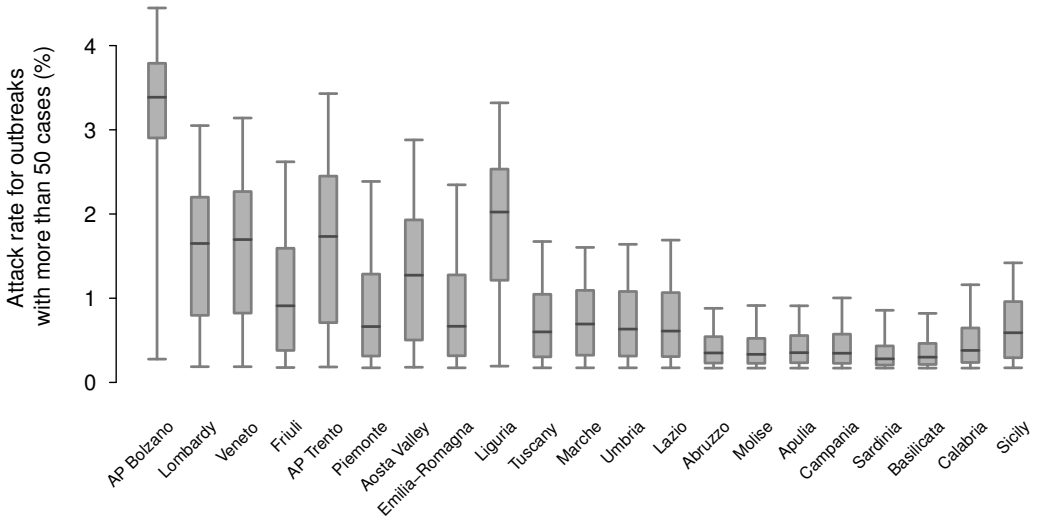

Figure 4.9: Attack rate (%) in outbreaks with more than 50 secondary cases by region.

outbreaks. Several studies on mathematical modeling of measles, either deterministic or stochastic, have been published starting from late 80s, providing pictures of the transmission dynamics of the disease and a better understanding of the size and timing of possible outbreaks [162, 163, 164, 165]. In Italy some measles modeling studies were conducted [166, 167, 168], since then, the epidemiological situation of measles has changed and the transmission dynamics of measles are affected by several years of suboptimal vaccination coverages. The implementation of effective vaccination strategies aimed at disease elimination needs a better understanding of the epidemiological features of measles in settings where the circulation of the virus is low. The aim of this study is to contribute to fill these gaps by focusing on the case of Italy. Firstly, we provide an updated estimate of measles seroprevalence by age in all Italian regions by using serological data collected in 2004 and regional measles vaccination coverages over the period 2004-2014. Results obtained show that there is remarkable heterogeneity in measles susceptibility profiles by age between regions. Second, we estimate a serial interval of 11.4 days, in good agreement with previous findings [161], by analyzing data on measles cases reported during the period 2013-2015 and the reported epidemiological links. Third, we use this information to reconstruct the epidemiological links potentially missing in the original dataset. Finally, we use a computational model to estimate where secondary infections occur and the attack rates that we can expect in different regions. Modeling results suggest that when importing an index case, most of new infections are linked to contacts in the "general community". In addition, given the heterogeneities in the regional susceptibility profiles by age, there are regions which are more likely to experience large outbreaks (more than 50 cases). For instance, the highest probability of an imported case to generate more than 50 secondary cases and the highest

attack rate were estimated in the AP of Bolzano, respectively 37.5% and 3.4%.
The analysis performed has several limitations. First, when we update the age-specific serological profile of 2004 to the year 2014 we start from estimates that, as observed in [159], may be biased by the non-detection of seropositivity among vaccinated individuals, which seem to be characterized by a lower level of antibodies with respect to people immune after natural infection. Second, when estimating the age-specific seroprevalence we do not consider the immunity acquired by people that got infected during the period 2004-2014 that, however, represent a negligible fraction of the population. Third, the estimation of the serial interval is based on clusters of size two identified in the cases of measles reported between 2013 and 2015. A refined version of the model could include information on multiple epidemiological links in the serial interval estimation [169]. Finally, we assume that all individuals have the same level of infectivity. A future sensitivity analysis could include the individual infectivity into the set of free parameters of the computational model.
Results obtained are not conclusive, even though they provide insights into some key epidemiological features of measles in the low-circulation Italian setting. In particular, our findings highlight that the transmission dynamics of measles in the country are strongly region-dependent and suggest that, in order to achieve measles elimination, intervention strategies tailored to the different regions should be considered, rather than nationwide.

Appendix D

D.1 Additional figures

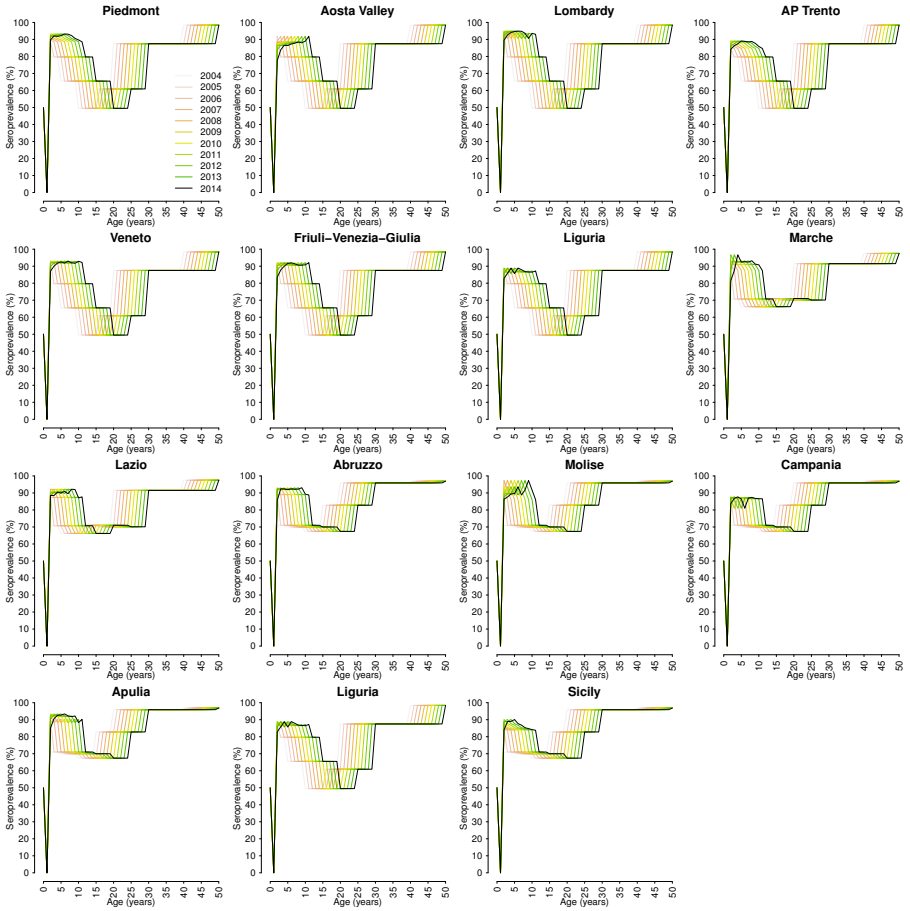

Figure D.1: Measles age-specific serological profiles by region as estimated in 2014.

Conclusions

As recently pointed out by the European Center for Disease Prevention and Control (ECDC) public health policy makers are increasingly turning to mathematical modeling to support their decisions [17]. This trend has been calling for the introduction of a new class of models that not only are able to explain qualitatively the dynamics of infectious diseases, but also have the capability to provide quantitatively accurate and reliable results. To this aim models are becoming more and more detailed and informed with data. However, there is still much to be done in order to capture the individual and population features that shape the spread of infectious diseases. This thesis addresses some issues in epidemiological modeling that warrant further investigation.

In Chapter 1 we introduce an age-structured individual-based stochastic model of Varicella Zoster Virus (VZV) transmission, whose main novelty is the inclusion of realistic population dynamics over the last century. This chapter represents an attempt to answer the need pointed out by recent studies [23, 24, 170] for a better understanding of the role of demographic processes in shaping the circulation of infectious diseases.

In Chapter 2 we use the model for VZV transmission developed in Chapter 1 to evaluate the effectiveness of varicella and HZ vaccination programs in Italy. With a view to the support of public health decisions, the epidemiological model is coupled with a cost-effectiveness analysis. To the best of our knowledge, this work represents the first attempt to evaluate the post-vaccination trends in varicella and HZ, both from an epidemiological and economic perspective, in light of the underlying effect of demographic processes. Another novelty of this study is that we take into account the uncertainty regarding the mechanism of VZV reactivation, by comparing results obtained using two different modeling assumptions on exogenous boosting.

In Chapter 3 we retrospectively analyze the spatiotemporal dynamics of the 2009 H1N1 influenza pandemic in England, by using a spatially-explicit model of influenza transmission, accounting for socio-demographic and disease natural history data. The aim of this work is to investigate whether the observed spatiotemporal dynamics of the epidemic was shaped by a spontaneous behavioral response to the pandemic threat. This chapter, repre-

sents an attempt to contribute to the challenge of understanding and quantifying the effect of human behavioral changes on the spread of epidemics [25, 29].

In Chapter 4 we investigate the current epidemiology of measles in Italy, by using a detailed computational model for measles transmission, informed with regional heterogeneities in the age-specific seroprevalence profiles. The analysis performed in this chapter tries to fill some of the existing gaps in the knowledge of the epidemiological features of vaccine preventable diseases in frameworks characterized by a low circulation of the virus.

In the past century, most industrialized countries underwent a massive decline of fertility and a progressive aging of the population [56]. We investigated the impact of demographic dynamics on the epidemiology of varicella and HZ in Spain, by using an age-structured individual-based stochastic model of VZV transmission dynamics, calibrated against the age-specific profiles of VZV seroprevalence and HZ incidence (Chapter 1). The model qualitatively reproduces the remarkable growth of HZ incidence observed in Spain between 1997 and 2004, before the introduction of vaccination [55]. We demonstrate that this growth may be partially ascribed to the reduction of varicella circulation that followed the overall decline of the birth rate in the twentieth century. Model predictions further suggest that, even under the most optimistic projections, HZ incidence will continue its rise for some decades, thereby confounding the interpretation of the effects of varicella immunization on HZ. Our findings confirm the important role of demographic processes in the epidemiology of infectious diseases. Moreover, the model proposed may help to disentangle the effect of natural epidemiological processes from that of vaccination, thereby allowing a more accurate assessment of past and future interventions. To this aim we employed it to evaluate the joint impact of demographic changes and different varicella and HZ vaccination strategies on the future epidemiology of varicella and HZ in Italy (Chapter 2). In particular, we performed a cost-effectiveness analysis of varicella and HZ vaccination by using two different modeling assumptions about the effect of exogenous boosting on the risk of developing HZ. Results obtained show that the increase of HZ incidence predicted after the introduction of varicella vaccination strongly depends on the formulation used to model exogenous boosting. In particular, when assuming "progressive immunity" the dynamics of HZ is more affected by changes in varicella circulation, i.e. we obtain a more remarkable increase of HZ than when assuming "temporary immunity". According to the cost-effectiveness analysis performed, these quantitative difference become critical from the public health economical perspective only when considering the introduction of varicella vaccination alone. On the other hand, the introduction of HZ vaccination alone, either targeted at 50 or 60 years old, results highly cost-effective, independently of the model used. Finally, results obtained suggest that when considering the combination of varicella and HZ vaccination, the latter partially compensates the negative drawbacks of varicella vaccination on the incidence of HZ. In particular, all scenarios considered result highly cost-effective for both modeling formulations but, especially, when the HZ vaccine is given at 65 years. The analysis per-

formed highlights that the choice of the exogenous boosting assumption is a key factor in the quantitative prediction of the impact of different immunization strategies on the incidence of HZ. However, the cost-effectiveness analysis robustly show that, disregarding the way in which we model boosting, an HZ vaccination program eventually combined with a varicella vaccination program has a high probability of being cost-effective.

We introduced a computational model embedded in a Bayesian framework in order to detect the effect of spontaneous behavioral changes in the spatiotemporal dynamics of the 2009 H1N1 influenza pandemic in England (Chapter 3). Our findings provide a clear picture of the epidemiology of the 2009 pandemic, well compliant with the existing knowledge. Consistently with serum sample data, we estimate a highly heterogeneous pattern of geographic spread in the first few months of the epidemic, with London experiencing a marked first wave during spring/summer, and an accelerated and much more homogeneous spread in the subsequent phases. The main novelty of this study is the detection of a clear signal that the distance at which secondary infections were generated changed over time. We estimate that the mean distance of secondary infections was 1.4 km during the early phase of the pandemic and increased to 5.3 km later on, comparable to 5.8 km estimated by using commuting flows. A possible explanation for this change lies in the adoption of protective behaviors emerged as a response to the pandemic threat, which may have been triggered by an initial overestimation of the risk. Our findings provide quantitative insights on the impact of human behavior on epidemic dynamics and shed light on the existing gaps in the knowledge of spreading mechanisms of pandemic influenza. Since preparedness plans for pandemic influenza are currently based on mobility patterns not accounting for human response, we believe that our findings could support the design of more effective control strategies.

Finally, we analyzed the current epidemiology of measles in Italy to provide some insights into its epidemiological features in a low circulation setting (Chapter 4). First, we provide updated estimates of the measles seroprevalence by age in all Italian regions. Results obtained show that the regional differences in measles vaccination coverage over the last years have led to a fragmented epidemiological framework, e.g. the proportion of susceptibles of school-age ranges from 9% to 27% depending on the region. Second, we estimate a serial interval for measles of 11.4 days, by analyzing the clusters of two cases reported to the Italian National Surveillance System between 2013-2015, and we use this estimate to reconstruct the missing epidemiological links. Finally, we use a computational model to estimate where secondary infections occur and the attack rates in different regions. Results obtained suggest that in the Italian low-circulation setting most of secondary measles cases are linked to contacts in the "general community", about 67%, rather than in households or at school. The estimated overall probability of an imported case to generate outbreaks with more than 50 cases is about 8%. However, given the heterogeneities in the regional serological profile by age, the probability of experiencing this occurrence varies from 3%

to 38% depending on the region considered and, once the epidemic starts, we can expect regional attack rates in the range 0.3% to 3.4%. Results obtained provide insights into some key epidemiological features of measles in the low-circulation Italian setting. In particular, our findings highlight that the transmission dynamics of measles in Italy are strongly region-dependent and suggest that, in order to achieve measles elimination, intervention strategies tailored to the different regions should be considered, rather than nationwide.

List of Figures

Bibliography

[1] Austin Alchon S (2003) A pest in the land: new world epidemics in a global perspective. (University of New Mexico Press).

[2] World Health Organization (WHO) (2016) Measles fact sheet (`http://www.who.int/mediacentre/factsheets/fs286/en/`).

[3] World Health Organization (WHO) (2004) The world health report 2004–changing history. Geneva. (`http://www.who.int/whr/2004/en/`).

[4] Durrheim DN (2016) Measles elimination — using outbreaks to identify and close immunity gaps. New England Journal of Medicine 375(14):1392–1393. PMID: 27705259.

[5] Kermack WO, McKendrick AG (1927) A contribution to the mathematical theory of epidemics. (The Royal Society), Vol. 115, pp. 700–721.

[6] Garnett GP, Cousens S, Hallett TB, Steketee R, Walker N (2011) Mathematical models in the evaluation of health programmes. The Lancet 378(9790):515–525.

[7] Lunelli A et al. (2013) Understanding the dynamics of seasonal influenza in italy: incidence, transmissibility and population susceptibility in a 9-year period. Influenza and Other Respiratory Viruses 7(3):286–295.

[8] Christensen H, Hickman M, Edmunds WJ, Trotter CL (2013) Introducing vaccination against serogroup B meningococcal disease: an economic and mathematical modelling study of potential impacts. Vaccine 31(23):2638–2646.

[9] Trotter CL, Edmunds WJ (2002) Modelling cost effectiveness of meningococcal serogroup C conjugate vaccination campaign in England and Wales. BMJ 324(7341):809.

[10] Melegaro A, Edmunds W (2004) Cost-effectiveness analysis of pneumococcal conjugate vaccination in england and wales. Vaccine 22(31):4203–4214.

[11] Turner D et al. (2006) The cost-effectiveness of influenza vaccination of healthy adults 50–64 years of age. Vaccine 24(7):1035–1043.

[12] Nichol K, Margolis K, Wuorenma J, Von Sternberg T (1994) The efficacy and cost effectiveness of vaccination against influenza among elderly persons living in the community. New England journal of medicine 331(12):778–784.

[13] Lieu TA et al. (1994) Cost-effectiveness of a routine varicella vaccination program for us children. Jama 271(5):375–381.

[14] Beutels P, Clara R, Tormans G, Van Doorslaer E, Van Damme P (1996) Costs and benefits of routine varicella vaccination in german children. Journal of Infectious Diseases 174(Supplement 3):S335–S341.

[15] Brisson M, Edmunds W (2002) The cost-effectiveness of varicella vaccination in canada. Vaccine 20(7):1113–1125.

[16] van Hoek AJ, Melegaro A, Gay N, Bilcke J, Edmunds WJ (2012) The cost-effectiveness of varicella and combined varicella and herpes zoster vaccination programmes in the united kingdom. Vaccine 30(6):1225–1234.

[17] European Centre for Disease Prevention and Control (ECDC) (2015) Current practices in immunisation policymaking in European countries. (`http://ecdc.europa.eu/en/healthtopics/documents/090613_influenza_ah1n1_situation_report_1700hrs.pdf`).

[18] Meltzer MI, Santibanez S, Fischer L, et al. (2016) Modeling in Real Time During the Ebola Response. MMWR 65(Suppl–3):85–89.

[19] Mossong J et al. (2008) Social contacts and mixing patterns relevant to the spread of infectious diseases. PLoS Med 5(3):e74.

[20] Colizza V, Barrat A, Barthelemy M, Vespignani A (2006) The role of the airline transportation network in the prediction and predictability of global epidemics. Proceedings of the National Academy of Sciences 103(7):2015–2020.

[21] Tizzoni M et al. (2014) On the use of human mobility proxies for modeling epidemics. PLoS Comput Biol 10(7):1–15.

[22] Merler S, Ajelli M (2010) The role of population heterogeneity and human mobility in the spread of pandemic influenza. Proceedings of the Royal Society B: Biological Sciences 277(1681):557–565.

[23] Merler S, Ajelli M (2014) Deciphering the relative weights of demographic transition and vaccination in the decrease of measles incidence in Italy. Proceedings of the Royal Society of London B: Biological Sciences 281(1777):20132676.

[24] Manfredi P, Williams JR (2004) Realistic population dynamics in epidemiological models: the impact of population decline on the dynamics of childhood infectious diseases: Measles in Italy as an example. Mathematical Biosciences 192(2):153–175.

[25] Ferguson N (2007) Capturing human behaviour. Nature 446(7137):733–733.

[26] Poletti P, Ajelli M, Merler S (2011) The Effect of Risk Perception on the 2009 H1n1 Pandemic Influenza Dynamics. PLoS ONE 6(2):e16460.

[27] He D, Dushoff J, Day T, Ma J, Earn DJ (2013) Inferring the causes of the three waves of the 1918 influenza pandemic in england and wales. Proceedings of the Royal Society of London B: Biological Sciences 280(1766):20131345.

[28] Bauch CT, Bhattacharyya S (2012) Evolutionary game theory and social learning can determine how vaccine scares unfold. PLoS Comput Biol 8(4):e1002452.

[29] Funk S, Salathé M, Jansen VAA (2010) Modelling the influence of human behaviour on the spread of infectious diseases: a review. Journal of The Royal Society Interface.

[30] Donahue JG, Choo PW, Manson JE, Platt R (1995) The incidence of Herpes Zoster. Archives of Internal Medicine 155(15):1605–1609.

[31] Oxman MN (2009) Herpes Zoster pathogenesis and cell-mediated immunity and immunosenescence. The Journal of the American Osteopathic Association 109(6_suppl_2):S13–S17.

[32] Hope-Simpson RE (1965) The nature of Herpes Zoster: a long-term study and a new hypothesis. Journal of the Royal Society of Medicine 58(1):9–20.

[33] Oxman MN (1995) Immunization to reduce the frequency and severity of Herpes Zoster and its complications. Neurology 45(12 Suppl 8):S41–S46.

[34] Burke BL et al. (1982) Immune responses to varicella-zoster in the aged. Archives of Internal Medicine 142(2):291–293.

[35] Terada K, Kawano S, Yoshihiro K, Morita T (1993) Proliferative response to varicella-zoster virus is inversely related to development of high levels of varicella-zoster virus specific igg antibodies. Scandinavian Journal of Infectious Diseases 25(6):775–778.

[36] Vossen MT et al. (2004) Development of virus-specific cd4+ t cells on reexposure to varicella-zoster virus. Journal of Infectious Diseases 190(1):72–82.

[37] Ogunjimi B et al. (2011) Exploring the impact of exposure to primary varicella in children on varicella-zoster virus immunity of parents. Viral Immunology 24(2):151–157.

[38] Ogunjimi B et al. (2014) Influence of frequent infectious exposures on general and varicella-zoster virus-specific immune responses in pediatricians. Clinical and Vaccine Immunology 21(3):417–426.

[39] Thomas SL, Wheeler JG, Hall AJ (2002) Contacts with varicella or with children and protection against Herpes Zoster in adults: a case-control study. The Lancet 360(9334):678–682.

[40] Brisson M, Gay N, Edmunds W, Andrews N (2002) Exposure to varicella boosts immunity to Herpes-Zoster: implications for mass vaccination against chickenpox. Vaccine 20(19):2500–2507.

[41] Karhunen M et al. (2010) Modelling the impact of varicella vaccination on varicella and zoster. Epidemiology and Infection 138(4):469.

[42] Poletti P et al. (2013) Perspectives on the impact of varicella immunization on Herpes Zoster. a model-based evaluation from three european countries. PloS one 8(4):e60732.

[43] Guzzetta G et al. (2013) Hope-Simpson's progressive immunity hypothesis as a possible explanation for Herpes Zoster incidence data. American Journal of Epidemiology 177(10):1134–1142.

[44] Rasch G, Hellenbrand W (2004) Germany adds varicella vaccine to the national vaccination programme. Eurosurveillance weekly 8:31.

[45] Seward JF et al. (2002) Varicella disease after introduction of varicella vaccine in the United States, 1995-2000. JAMA 287(5):606–611.

[46] Stefanoff P et al. (2010) Varicella and Herpes Zoster surveillance and vaccination recommendations 2010-2011, (Venice II Consortium), Technical report.

[47] Brisson M et al. (2010) Modeling the impact of one-and two-dose varicella vaccination on the epidemiology of varicella and zoster. Vaccine 28(19):3385–3397.

[48] van Hoek AJ, Melegaro A, Zagheni E, Edmunds WJ, Gay N (2011) Modelling the impact of a combined varicella and zoster vaccination programme on the epidemiology of varicella zoster virus in england. Vaccine 29(13):2411–2420.

[49] Jumaan AO et al. (2005) Incidence of Herpes Zoster, before and after varicella-vaccination-associated decreases in the incidence of varicella, 1992–2002. Journal of Infectious Diseases 191(12):2002–2007.

[50] Reynolds MA, Chaves SS, Harpaz R, Lopez AS, Seward JF (2008) The impact of the varicella vaccination program on Herpes Zoster epidemiology in the United States: a review. Journal of Infectious Diseases 197(Supplement 2):S224–S227.

[51] Law B, Chateau D, Walld R, Roos L (2004) Temporal trends in the annual population-based incidence of Herpes Zoster by age and gender: Manitoba, 1979–1998. Can J Infect Dis Med Microbiol 15:357–8.

[52] Russell M, Schopflocher D, Svenson L, Virani S (2007) Secular trends in the epidemiology of shingles in Alberta. Epidemiology and Infection 135(06):908–913.

[53] Ragozzino M, Melton 3rd L, Kurland L, Chu C, Perry H (1982) Population-based study of Herpes Zoster and its sequelae. Medicine 61(5):310–316.

[54] Brisson M et al. (2001) Epidemiology of varicella zoster virus infection in Canada and the United Kingdom. Epidemiology and Infection 127(2):305–314.

[55] Pérez-Farinós N et al. (2007) Varicella and Herpes Zoster in Madrid, based on the Sentinel General Practitioner Network: 1997–2004. BMC Infectious Diseases 7(1):1.

[56] Livi-Bacci M (2012) A concise history of world population. (Cambridge, MA: Wiley-Blackwell).

[57] Kohler HP, Billari FC, Ortega JA (2002) The emergence of lowest-low fertility in Europe during the 1990s. Population and development review 28(4):641–680.

[58] Instituto Nacional de Estadística. (2014) INEbase (`http://www.ine.es/inebmenu/mnu_cifraspob.htm`). Accessed: 2016-10-19.

[59] University of California Berkeley (USA) and Max Planck Institut for demographic research (Germany). (2014) 2014 Human Mortality Database (`www.mortality.org`).

[60] Statistical Office of the European Commission (Eurostat) (2011) Migrants in Europe. A statistical portrait of the first and second generation. Technical report (`http://ec.europa.eu/eurostat/documents/3217494/5727749/KS-31-10-539-EN.PDF/bcf27a60-7016-4fec-98c5-e8488491ebbd`).

[61] World Bank (2014) Net Migration (`http://data.worldbank.org/indicator/sm.pop.netm`).

[62] Bover O, Velilla P (1999) Migration in Spain: historical background and current trends. (Bonn, Germany: Institute for the study of Labor).

[63] United Nations Department of Economic and Social Affairs. (2012) UN World Population Prospects: the 2012 revision (`http://https://esa.un.org/unpd/wpp/`). Accessed: 2014-12-29.

[64] Melegaro A, Jit M, Gay N, Zagheni E, Edmunds WJ (2011) What types of contacts are important for the spread of infections? Using contact survey data to explore European mixing patterns. Epidemics 3(3):143–151.

[65] Fumanelli L, Ajelli M, Manfredi P, Vespignani A, Merler S (2012) Inferring the structure of social contacts from demographic data in the analysis of infectious diseases spread. PLoS Comput Biol 8(9):e1002673.

[66] Nardone A et al. (2007) The comparative sero-epidemiology of varicella zoster virus in 11 countries in the European region. Vaccine 25(45):7866–7872.

[67] García Cenoz M et al. (2008) Incidencia de la varicela y el herpes zoster antes de la introducción de la vacunación sistemática infantil en navarra, 2005-2006. Anales del sistema sanitario de Navarra 31(1):71–80.

[68] Livi-Bacci M (1999) The population of Europe: a history. (Oxford, UK: Blackwell).

[69] Brisson M, Edmunds W, Gay N, Law B, De Serres G (2000) Analysis of varicella vaccine breakthrough rates: implications for the effectiveness of immunisation programmes. Vaccine 18(25):2775–2778.

[70] Civen R et al. (2009) The incidence and clinical characteristics of Herpes Zoster among children and adolescents after implementation of varicella vaccination. The Pediatric infectious disease journal 28(11):954–959.

[71] Gaillat J et al. (2011) Does monastic life predispose to the risk of Saint Anthony's fire (Herpes Zoster)? Clinical Infectious Diseases 53(5):405–410.

[72] Gialloreti LE et al. (2010) Epidemiology and economic burden of herpes zoster and post-herpetic neuralgia in Italy: a retrospective, population-based study. BMC Infectious Diseases 10(1):230.

[73] Iozzi F et al. (2010) Little italy: An agent-based approach to the estimation of contact patterns- fitting predicted matrices to serological data. PLoS Comput Biol 6(12):1–10.

[74] Kawai K, Gebremeskel BG, Acosta CJ (2014) Systematic review of incidence and complications of herpes zoster: towards a global perspective. BMJ open 4(6):e004833.

[75] Réseau Sentinelles, INSERM/UPMC (2015) Temporal series of varicella and HZ incidence (`http://www.sentiweb.fr`).

[76] Gauthier AH (1999) Historical trends in state support for families in europe (post-1945). Children and Youth Services Review 21(11):937–965.

[77] Yih WK et al. (2005) The incidence of varicella and herpes zoster in Massachusetts as measured by the Behavioral Risk Factor Surveillance System (BRFSS) during a period of increasing varicella vaccine coverage, 1998–2003. BMC Public Health 5(1):68.

[78] Mc Lean AR, Anderson RM (1988) Measles in developing countries. Part I. Epidemiological parameters and patterns. Epidemiol Infect 100:111.

[79] De Jong M, Diekmann O, Heesterbeek H (1995) Epidemic Models, ed. Mollison D. (Cambridge University), p. 84.

[80] Seiler HE (1949) A study of Herpes Zoster particularly in its relationship to chicken-pox. The Journal of Hygiene 47(3):253–262.

[81] Gilks W (2005) Markov Chain Monte Carlo. (Wiley Online Library).

[82] National Institute of Statistics and Economic Studies - France (2016) Estimations de population et statistiques de l'état civil: Solde migratorie (`http://www.insee.fr/`).

[83] Béraud G et al. (2015) The French Connection: The First Large Population-Based Contact Survey in France Relevant for the Spread of Infectious Diseases. PloS ONE 10(7):e0133203.

[84] Khoshnood B et al. (2006) Seroprevalence of varicella in the French population. The Pediatric Infectious Disease Journal 25(1):41–44.

[85] Mick G, Hans G (2013) Postherpetic neuralgia in europe: The scale of the problem and outlook for the future. Journal of Clinical Gerontology and Geriatrics 4(4):102 – 108.

[86] Goldman G, King P (2013) Review of the United States universal varicella vaccination program: Herpes zoster incidence rates, cost-effectiveness, and vaccine efficacy based primarily on the Antelope Valley Varicella Active Surveillance Project data. Vaccine 31(13):1680–1694.

[87] Brisson M, Edmunds W, Gay N (2003) Varicella vaccination: Impact of vaccine efficacy on the epidemiology of VZV. Journal of Medical Virology 70(S1):S31–S37.

[88] Oxman M et al. (2005) A Vaccine to Prevent Herpes Zoster and Postherpetic Neuralgia in Older Adults. New England Journal of Medicine 352(22):2271–2284. PMID: 15930418.

[89] Damm O et al. (2015) Systematic review of models assessing the economic value of routine varicella and herpes zoster vaccination in high-income countries. BMC Public Health 15(1):533.

[90] Marziano V et al. (2015) The impact of demographic changes on the epidemiology of herpes zoster: Spain as a case study. Proceedings of the Royal Society of London B: Biological Sciences 282(1804).

[91] ISTAT (2015) Database of the Italian Institute of Statistics. (`http://demo.istat.it`). Accessed: 2015-6.

[92] United Nations Department of Economic and Social Affairs. (2015) UN World Population Prospects: the 2015 revision (`http://https://esa.un.org/unpd/wpp/`).

[93] Seward JF, Zhang JX, Maupin TJ, Mascola L, Jumaan AO (2004) Contagiousness of varicella in vaccinated cases: a household contact study. JAMA 292(6):704–708.

[94] Consiglio Superiore di Sanità (2015) Piano Nazionale Prevenzione Vaccinale 2016-2018, PNPV (2016-2018) (`http://www.quotidianosanita.it/allegati/allegato1955037.pdf`). Accessed: 2016-10.

[95] European Centre for Disease Prevention and Control (ECDC) (2015) Varicella vaccination in the European Union. Stockholm (`http://ecdc.europa.eu/en/publications/Publications/Varicella-Guidance-2015.pdf`).

[96] Bilcke J et al. (2012) The health and economic burden of chickenpox and herpes zoster in Belgium. Epidemiology and Infection 140(11):2096–2109.

[97] Public Health England (PHE) (2016) Shingles: guidance and vaccination programme (`https://www.gov.uk/government/collections/shingles-vaccination-programme`).

[98] European Centre for Disease Prevention and Control (ECDC) (2016) Vaccine schedule. recommended immunisations for Herpes Zoster (`http://vaccine-schedule.ecdc.europa.eu/Pages/Scheduler.aspx`).

[99] Vázquez M et al. (2001) The effectiveness of the varicella vaccine in clinical practice. New England Journal of Medicine 344(13):955–960.

[100] Thiry N et al. (2004) An economic evaluation of varicella vaccination in italian adolescents. Vaccine 22(27):3546–3562.

[101] Coudeville L, Brunot A, Giaquinto C, Lucioni C, Dervaux B (2004) Varicella vaccination in italy. Pharmacoeconomics 22(13):839–855.

[102] Van Hoek A, Gay N, Melegaro A, Opstelten W, Edmunds W (2009) Estimating the cost-effectiveness of vaccination against herpes zoster in england and wales. Vaccine 27(9):1454–1467.

[103] Fornaro P et al. (1999) Epidemiology and cost analysis of varicella in italy: results of a sentinel study in the pediatric practice. The Pediatric infectious disease journal 18(5):414–419.

[104] Zotti C, Maggiorotto G, Migliardi A (2002) [costs of varicella]. Annali di igiene : medicina preventiva e di comunita 14(4 Suppl 6):29—33.

[105] Giaquinto C et al. (2002) Epidemiologia ed esiti della varicella in italia: risultati di uno studio prospettico sui bambini (0–14 anni) seguiti dai pediatri di libera scelta (studio pedianet). Ann Ig 14(Suppl 6):21–7.

[106] Fraser C et al. (2009) Pandemic Potential of a Strain of Influenza A (H1N1): Early Findings. Science 324(5934):1557–1561.

[107] Health Protection Agency (HPA) (2010) Epidemiological report of pandemic (H1n1) 2009 in the UK (http://webarchive.nationalarchives.gov.uk/20140714084352/http:/www.hpa.org.uk/webc/HPAwebFile/HPAweb_C/1284475321350).

[108] Chowell G et al. (2011) Characterizing the Epidemiology of the 2009 Influenza A/H1N1 Pandemic in Mexico. PLoS Med 8(5):1–13.

[109] Nicoll A, Coulombier D (2009) Europe's initial experience with pandemic (H1N1) 2009 - mitigation and delaying policies and practices. Euro Surveillance 14(29).

[110] WHO (2010) Situation update in the European Region: overview of influenza surveillance data week 40/2009 to week 07/2010. (http://www.euro.who.int/__data/assets/pdf_file/0003/91839/E93581.pdf).

[111] Merler S, Ajelli M, Pugliese A, Ferguson NM (2011) Determinants of the spatiotemporal dynamics of the 2009 h1n1 pandemic in europe: Implications for real-time modelling. PLoS Comput Biol 7(9):e1002205.

[112] European Centre for Disease Prevention and Control (ECDC) (2010) The 2009 A(H1n1) pandemic in Europe. A review of the experience. (http://ecdc.europa.eu/en/publications/Publications/101108_SPR_pandemic_experience.pdf).

[113] Poggensee G et al. (2010) The first wave of pandemic influenza (H1N1) 2009 in Germany: from initiation to acceleration. BMC Infect. Dis. 10:155.

[114] Calatayud L et al. (2010) Pandemic (H1n1) 2009 virus outbreak in a school in London, April-May 2009: an observational study. Epidemiol. Infect. 138(2):183–191.

[115] Health Protection Agency West Midlands H1N1v Investigation Team (2009) Preliminary descriptive epidemiology of a large school outbreak of influenza A(H1n1)v in the West Midlands, United Kingdom, May 2009. Euro Surveill. 14(27).

[116] Miller E et al. (2010) Incidence of 2009 pandemic influenza a {H1N1} infection in england: a cross-sectional serological study. The Lancet 375(9720):1100 – 1108.

[117] Hardelid P, Andrews N, Hoschler K, Stanford E, Baguelin M (2010) Assessment of baseline age-specific antibody prevalence and incidence of infection to novel influenza a h1n1 2009. Health Technology Assessment 14(55).

[118] Merler S et al. (2015) Spatiotemporal spread of the 2014 outbreak of Ebola virus disease in Liberia and the effectiveness of non-pharmaceutical interventions: a computational modelling analysis. Lancet Infect Dis 15(2):204–211.

[119] Ajelli M et al. (2015) The 2014 Ebola virus disease outbreak in Pujehun, Sierra Leone: epidemiology and impact of interventions. BMC Medicine 13:281.

[120] Chretien JP, Riley S, George DB (2015) Mathematical modeling of the West Africa Ebola epidemic. eLife 4:e09186.

[121] World Health Organization (WHO) (2015) | One year into the Ebola epidemic: a deadly, tenacious and unforgiving virus (`http://www.who.int/csr/disease/ebola/one-year-report/introduction/en/`).

[122] Ferguson NM et al. (2005) Strategies for containing an emerging influenza pandemic in Southeast Asia. Nature 437(7056):209–214.

[123] Poletti P, Caprile B, Ajelli M, Pugliese A, Merler S (2009) Spontaneous behavioural changes in response to epidemics. Journal of Theoretical Biology 260(1):31–40.

[124] Reluga TC (2010) Game Theory of Social Distancing in Response to an Epidemic. PLOS Comput Biol 6(5):e1000793.

[125] Perra N, Balcan D, GonÃ§alves B, Vespignani A (2011) Towards a Characterization of Behavior-Disease Models. PLOS ONE 6(8):e23084.

[126] Poletti P, Ajelli M, Merler S (2012) Risk perception and effectiveness of uncoordinated behavioral responses in an emerging epidemic. Mathematical Biosciences 238(2):80–89.

[127] Merler S, Ajelli M, Fumanelli L, Vespignani A (2013) Containing the accidental laboratory escape of potential pandemic influenza viruses. BMC Med 11:252.

[128] Education, Audiovisual and Culture Executive Agency, European Commission (2009) Key Data on Education in Europe (`http://ecdc.europa.eu/en/healthtopics/documents/090613_influenza_ah1n1_situation_report_1700hrs.pdf`).

[129] Simini F, Gonzalez MC, Maritan A, Barabasi AL (2012) A universal model for mobility and migration patterns. Nature 484(7392):96–100.

[130] Department of Transports (2015) National Travel Survey 2013. (`https://www.gov.uk/government/statistics/national-travel-survey-2013`).

[131] European Centre for Disease Prevention and Control (ECDC) (2009) ECDC Situation report on new influenza A(H1N1) infection 13 June 2009 (`http://ecdc.europa.eu/en/healthtopics/documents/090613_influenza_ah1n1_situation_report_1700hrs.pdf`).

[132] Civil Aviation Authority (2015) Airport data 1990 onwards (http://www.caa.co.uk/Data-and-analysis/UK-aviation-market/Airports/Datasets/UK-Airport-data/Airport-data-1990-onwards/).

[133] Cauchemez S et al. (2009) Household transmission of 2009 pandemic influenza a (h1n1) virus in the united states. New England Journal of Medicine 361(27):2619–2627. PMID: 20042753.

[134] Merler S et al. (2013) Pandemic Influenza A/H1n1pdm in Italy: Age, Risk and Population Susceptibility. PLoS ONE 8(10):e74785.

[135] Birrell PJ et al. (2011) Bayesian modeling to unmask and predict influenza A/H1n1pdm dynamics in London. Proc. Natl. Acad. Sci. U.S.A. 108(45):18238–18243.

[136] Biggerstaff M, Cauchemez S, Reed C, Gambhir M, Finelli L (2014) Estimates of the reproduction number for seasonal, pandemic, and zoonotic influenza: a systematic review of the literature. BMC Infect. Dis. 14:480.

[137] Cauchemez S, Valleron AJ, Boelle PY, Flahault A, Ferguson NM (2008) Estimating the impact of school closure on influenza transmission from sentinel data. Nature 452(7188):750–754.

[138] Ajelli M, Poletti P, Melegaro A, Merler S (2014) The role of different social contexts in shaping influenza transmission during the 2009 pandemic. Scientific Reports 4:7218.

[139] Ferguson NM et al. (2006) Strategies for mitigating an influenza pandemic. Nature 442(7101):448–452.

[140] Springborn M, Chowell G, MacLachlan M, Fenichel EP (2015) Accounting for behavioral responses during a flu epidemic using home television viewing. BMC Inf Dis 15(1):1–14.

[141] Health Protection Agency (HPA) (2009) Pandemic (H1n1) 2009 in England: an overview of initial epidemiological findings and implications for the second wave (http://www.checktheevidence.com/pdf/Swine%20Flu%20-%20NHS%20-%20Epidemiology.pdf).

[142] Dorigatti I, Cauchemez S, Ferguson NM (2013) Increased transmissibility explains the third wave of infection by the 2009 H1n1 pandemic virus in England. Proc. Natl. Acad. Sci. U.S.A. 110(33):13422–13427.

[143] Fumanelli L, Ajelli M, Merler S, Ferguson NM, Cauchemez S (2016) Model-Based Comprehensive Analysis of School Closure Policies for Mitigating Influenza Epidemics and Pandemics. PLOS Comput. Biol. 12(1):e1004681.

[144] Funk S, Gilad E, Jansen VAA (2010) Endemic disease, awareness, and local behavioural response. Journal of Theoretical Biology 264(2):501–509.

[145] Sadique MZ et al. (2007) Precautionary behavior in response to perceived threat of pandemic influenza. Emerging Infect. Dis. 13(9):1307–1313.

[146] SteelFisher GK, Blendon RJ, Bekheit MM, Lubell K (2010) The public's response to the 2009 H1n1 influenza pandemic. N. Engl. J. Med. 362(22):e65.

[147] Birrell PJ, Zhang XS, Pebody RG, Gay NJ, De Angelis D (2016) Reconstructing a spatially heterogeneous epidemic: Characterising the geographic spread of 2009 A/H1N1pdm infection in England. Sci Rep 6:29004.

[148] Charaudeau S, Pakdaman K, Boëlle PY (2014) Commuter mobility and the spread of infectious diseases: Application to influenza in france. PLoS ONE 9(1):1–9.

[149] Bedford T et al. (2015) Global circulation patterns of seasonal influenza viruses vary with antigenic drift. Nature 523(7559):217–220.

[150] Kucharski AJ, Conlan AJK, Eames KTD (2015) School's Out: Seasonal Variation in the Movement Patterns of School Children. PLoS ONE 10(6):e0128070.

[151] Shaman J, Pitzer VE, Viboud C, Grenfell BT, Lipsitch M (2010) Absolute Humidity and the Seasonal Onset of Influenza in the Continental United States. PLoS Biology 8(2):e1000316.

[152] Truscott J, Ferguson NM (2012) Evaluating the adequacy of gravity models as a description of human mobility for epidemic modelling. PLoS Comput. Biol. 8(10):e1002699.

[153] Chowell G, Hengartner N, Castillo-Chavez C, Fenimore P, Hyman J (2004) The basic reproductive number of ebola and the effects of public health measures: the cases of congo and uganda. Journal of Theoretical Biology 229(1):119 – 126.

[154] Pan American Health Organization (PAHO) (2016) Americas declared free of measles. (`www.sciencedaily.com/releases/2016/09/160928093510.htm`).

[155] World Health Organization (WHO) (2010) Renewed commitment to measles and rubella elimination and prevention of congenital rubella syndrome in the WHO European Region by 2015. Moscow, Russia, WHO Regional Officile for Europe (`http://www.euro.who.int/__data/assets/pdf_file/0008/119546/RC60_edoc15.pdf`).

[156] World Health Organization (WHO) (2012) Global measles and rubella strategic plan, 2012–2020. Geneva, Switzerland: World Health Organization; 2012. (`http://www.who.int/immunization/newsroom/Measles_Rubella_StrategicPlan_2012_2020.pdf`).

[157] Consiglio Superiore di Sanità (Ministero della Salute) (2011) Piano Nazionale Prevenzione Vaccinale PNPV (2012-2014) (http://www.salute.gov.it/imgs/c_17_pubblicazioni_1721_allegato.pdf).

[158] Ministero della Salute. (2000-2015) Vaccine coverages in Italy (http://www.epicentro.iss.it/temi/vaccinazioni/dati_Ita.asp).

[159] Rota MC et al. (2008) Measles serological survey in the italian population: interpretation of results using mixture model. Vaccine 26(34):4403–4409.

[160] Venables WN, Ripley BD (2002) Modern Applied Statistics with S. (Springer, New York), Fourth edition. ISBN 0-387-95457-0.

[161] Vink MA, Bootsma MCJ, Wallinga J (2014) Serial intervals of respiratory infectious diseases: a systematic review and analysis. American Journal of Epidemiology 180(9):865–875.

[162] Schenzle D (1984) An age-structured model of pre-and post-vaccination measles transmission. Mathematical Medicine and Biology 1(2):169–191.

[163] Bolker B, Grenfell B (1995) Space, persistence and dynamics of measles epidemics. Philosophical Transactions of the Royal Society B: Biological Sciences 348(1325):309–320.

[164] Jansen VA et al. (2003) Measles outbreaks in a population with declining vaccine uptake. Science 301(5634):804–804.

[165] Finkenstädt BF, Bjørnstad ON, Grenfell BT (2002) A stochastic model for extinction and recurrence of epidemics: estimation and inference for measles outbreaks. Biostatistics 3(4):493–510.

[166] Manfredi P, Ciofi degli Atti M, Mandolini D, Salmaso S (2003) Measles in the Italian regions: estimate of infection parameters. Epidemiology and prevention 27(6):340–347.

[167] Manfredi P, Williams JR, Atti MLCD, Salmaso S (2005) Measles Elimination in Italy: Projected Impact of the National Elimination Plan. Epidemiology and Infection 133(1):87–97.

[168] Manfredi P, Cleur EM, Williams JR, Salmaso S, Atti MCd (2005) The pre-vaccination regional epidemiological landscape of measles in italy: contact patterns, effort needed for eradication, and comparison with other regions of europe. Population Health Metrics 3(1):1.

[169] Team WER (2014) Ebola virus disease in west africa—the first 9 months of the epidemic and forward projections. N Engl J Med 2014(371):1481–1495.

[170] Earn DJ, Rohani P, Bolker BM, Grenfell BT (2000) A simple model for complex dynamical transitions in epidemics. Science 287(5453):667–670.