2023 | Faculty of Sciences

UHASSELT
KNOWLEDGE IN ACTION

University of Antwerp

Doctoral dissertation submitted to obtain the degrees of
- Doctor of Sciences: Statistics | UHasselt
- Doctor of Social Sciences: Sociology | UAntwerpen

Signe Møgelmose

DOCTORAL DISSERTATION

Population dynamics and household structures in infectious disease modelling: A demographic perspective

DSI DATA SCIENCE INSTITUTE
UHASSELT

Promoters: Prof. Dr Niel Hens | UHasselt
Prof. Dr Karel Neels | UAntwerpen

D/2023/2451/69

Contents

List of Tables

List of Figures

List of Abbreviations

ASFR Age-specific fertility rate

CBR Crude birth rate

CCBM Cohort-component-based model

CDR Crude death rate

EPM External population model

FPB Belgian federal planning bureau

GAM Generalised additive model

HH Household

HIV Human immunodeficiency virus

HZ Herpes zoster

IBM Individual-based model

IPUMS Integrated Public Use Microdata Series

LE Life expectancy

LTCF Long-term care facility

MSM Men who have sex with men

RSV Respiratory Syncytial Virus

SA Sensitivity analysis

SARS-CoV Severe acute respiratory syndrome coronavirus

SIPP Survey of Income and Program Participation

Statbel Belgian statistical office

SPC Supplementary professional contacts

TFR Total fertility rate

VZV Varicella-zoster virus

Chapter 1

Introduction

Infectious diseases do not spread randomly in a population nor affect people equally upon infection. Some population groups are more likely to acquire an infection or to experience a severe outcome in case of disease because of risk factors that may not be randomly distributed in the population. In infectious disease epidemiology, models are used to understand which factors influence the mechanisms by which infectious diseases spread and afflict a population in order to predict future outbreaks and to evaluate prevention and control strategies.

Some of the factors influencing the spread and burden of infectious diseases are related to demography and demographic events (e.g. age, births, migration), which typically are included in epidemiological models, though often in a highly simplified manner. Demographic structures and events in a population, however, result from complex demographic processes that tend to change over time. In the context of infectious disease epidemiology, it is not well understood how these underlying processes shape current and future population structures with relevance for the transmission and burden of infectious diseases.

For that reason, the aim of this dissertation is to explore and improve infectious disease models with dynamic host populations with the purpose of investigating the impact of demographic structures and changes on the transmission and burden of infectious diseases transmitted through close contact. Next, we provide some background information, including a brief overview of infectious diseases and the associated burden in a historical context, a short introduction to infectious disease modelling and the most commonly used methods, as well as the demographic methods applied in infectious

disease modelling. Finally, we briefly describe past and ongoing demographic changes and how these influence the transmission and burden of infectious diseases.

1.1 Infectious diseases

From ancient plagues to the current coronavirus (COVID-19) pandemic, infectious diseases have indisputably had a major impact on humanity. The emergence and spread of infectious diseases increased as civilisation evolved. Intensified animal keeping, more connected communities, extended trading and travelling, rising population density and urbanisation led to a higher risk of outbreaks, epidemics and pandemics [1, 2]. An outbreak constitutes a sudden rise in the number of infections in a relatively small area. It becomes an epidemic when spreading to a larger geographical area and is finally considered a pandemic when the spread is global [3]. While some infections become endemic, meaning they are present in a population at a more or less stable and predictable rate, outbreaks of other infectious diseases have been recurring throughout history.

The Black Death (1346–1353) was a plague pandemic that was immensely destructive and heavily reduced the populations in some regions of Europe and other parts of the World [4, 5]. Until the mid-17$^{\text{th}}$ century, areas of Europe faced several successive plague epidemics, limiting the population growth [1]. In addition to the plague, outbreaks of leprosy, syphilis, smallpox, yellow fever, typhoid fever, cholera and other infectious diseases were the norm [6]. The first cholera pandemic emerged in 1817 as it spread from India to multiple regions of the world and several cholera pandemics followed in the 19$^{\text{th}}$ and 20$^{\text{th}}$ centuries.

The high mortality caused by infectious diseases, particularly in infants and young children, is reflected in estimates of life expectancy (see Figure 1.1). In Belgium in 1841, life expectancy at birth was about 40 years for both sexes combined [7]. In the 19$^{\text{th}}$ century, life expectancy in Belgium and many other countries improved slowly, but was interrupted by different outbreaks. However, quarantine measures, sanitary cordons and hygiene practices led to faster improvements in mortality at the end of the 19$^{\text{th}}$ century and begininng of the 20$^{\text{th}}$ century until an abrupt interruption was caused by the First World War (1914-1918) and the Spanish flu pandemic (1918-1919) [8]. The Spanish flu caused as many as 50 million deaths globally, although such estimates are associated with substantial uncertainty [9].

After the Spanish flu pandemic, mortality from infectious diseases declined substantially in most European and Western countries, partly due to rising living standards and public health interventions in the first half of the 20th century and particularly due to advances in medicine (e.g. antibiotics, mass vaccination) in the second half [8]. Improvements in life expectancy followed accordingly as non-communicable diseases replaced communicable diseases as the main causes of death in many countries. The long-term improvement in life expectancy is one of the most spectacular developments in the past two centuries. In Belgium in 2021, life expectancy at birth was approaching 82 years, a tremendous improvement from the aforementioned 40 years in 1841 [7].

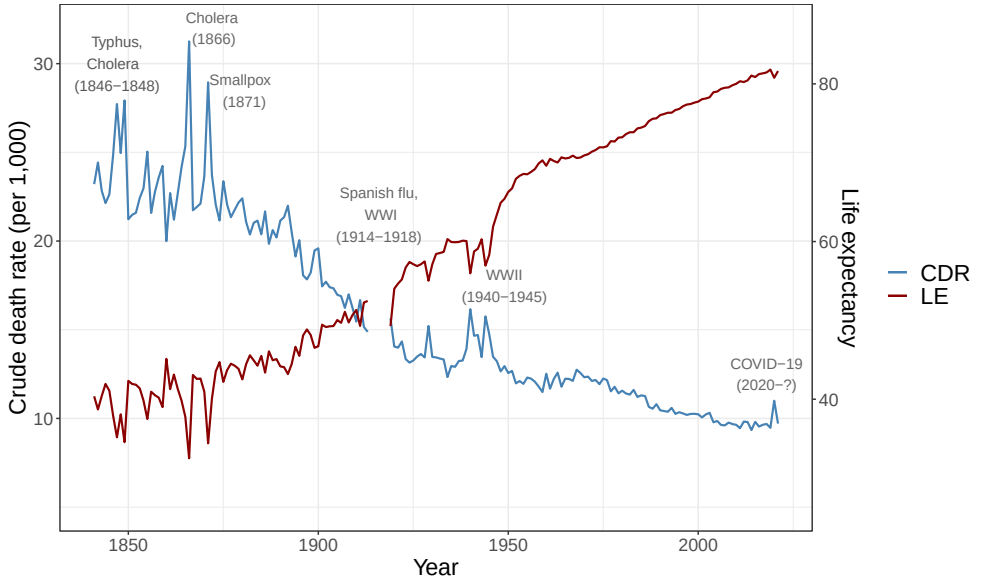

Figure 1.1. Crude death rate per 1,000 people (CDR, left y-axis, blue) and life expectancy at birth for both sexes combined (LE, right y-axis, red) for Belgium, 1841-2021. Missing data 1914-1918. Source: Human Mortality Database and Eggerickx et al. [10].

Nevertheless, infectious diseases remain a public health threat, particularly in low-income countries [11]. Viral descendants of the Spanish flu have caused seasonal influenza epidemics and pandemics in 1957-1958, 1968-1969 and 2009 [12]. Moreover, in recent decades, new threats have arisen from emerging infectious diseases, including human immunodeficiency virus (HIV), Ebola and severe acute respiratory syndrome coronavirus (SARS-CoV). COVID-19 being the most recent example of a devastating pandemic disease [13].

1.2 Modelling of infectious diseases

Mathematical and computational models have become important tools for analysing the spread of infectious diseases and to evaluate control measures, for which controlled trials typically are practically or ethically infeasible [14, 15]. Predictive models can guide policymakers in the assessment of different mitigation strategies, as seen during the COVID-19 pandemic [16]. Infectious disease modelling also provides a framework for investigating transmission dynamics and the role different factors play in the spread of an infection by examining them in isolation. Such models can give an understanding of which elements are important to consider and improve the accuracy of predictive models, hopefully leading to better strategies for prevention and control of infectious diseases [15, 17].

Mathematical modelling in epidemiology goes back to the 18^{th} century, when Daniel Bernoulli presented a paper on smallpox [18], demonstrating how inoculation in the long term could increase life expectancy and the population size. The method, however, was criticised in the scientific community and had little impact at that time [19, 20]. It was not until the beginning of the 20^{th} century that theories explaining the mechanisms by which infectious diseases spread were formulated. Hamer (1906) [21] presented the mass action principle defining the incidence (number of new cases per time unit) of directly transmitted infections as a function of the product of the number of susceptible and infectious individuals. Kermack and McKendrick (1927) [22] formulated additional theories, including the threshold theorem stating that for a given set of rates of transmission and recovery/death, a critical threshold density of susceptibles exists, which, if exceeded, gives rise to an epidemic. Moreover, they introduced the compartmental SIR model, a framework still used today. Since then, the field of mathematical epidemiology has grown substantially [17, 23–25].

1.2.1 Compartmental models

The spread of infectious diseases is often modelled at the population level using the compartmental modelling approach introduced by Kermack and McKendrick (1927) [22, 26]. The population is divided into compartments according to disease state (e.g. susceptible, infectious and recovered) and assumptions define the transitions between compartments. The model structure should reflect the natural history of the given infection. The model structure *Susceptible-Infectious* (SI), for example, is used when a once infected individual remains infected and infectious for life (e.g. HIV). Meanwhile, *Susceptible-Infectious-Recovered* (SIR) model structures would be appropriate

for infections for which individuals develop immunity after having been infected [25]. For other infections, it could be important to include temporary immunity of infants resulting from maternally-derived antibodies (M), or an exposed state (E) for infected but not yet infectious individuals (i.e. latent period) [27].

After the compartments (i.e. disease states) have been defined, the transitions from one compartment to another are described. Considering a simple SIR model (see example in Figure 1.2), the transition from *Susceptible* to *Infectious* results from disease transmission. Depending on the infection, different routes can lead to transmission between human or animal hosts, including direct contact (e.g. leprosy), respiratory route (e.g. influenza, tuberculoses, SARS-CoV-2), faecal-oral route (e.g. typhoid, dysentery), sexual contact (e.g. HIV) or through insect vectors (e.g. malaria, dengue). The interaction between hosts allowing transmission of an infection if one host is infectious and the other is susceptible will be referred to as *contact* [25].

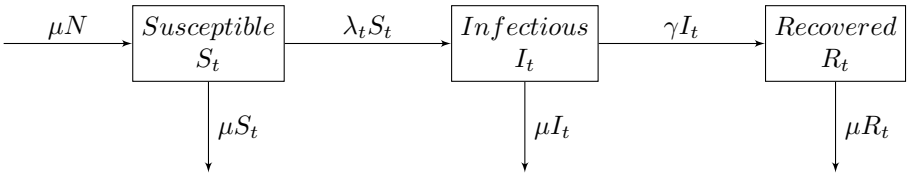

Figure 1.2. Example of SIR model. $\lambda_t = \beta I_t$, γ = recovery rate, μ = crude death/birth rate, $N = S_t + I_t + R_t$.

Contact patterns are typically expressed in rates and are modified by a transmission probability to take into account that some contacts between susceptible and infectious individuals do not lead to transmission [17]. The product of the contact rates and the transmission probability is denoted by β, the rate at which two specific individuals come into *effective contact* per unit time [25, 28]. The progression from *Susceptible* to *Infectious* can be formulated as $\lambda_t S_t$ (mass action principle), where

$$\lambda_t = \beta I_t \tag{1.1}$$

Hence, λ_t represents the risk of a susceptible individual becoming infected [25]. The size of each compartment can be defined in absolute numbers or as a proportion of the population, often depending on whether density-dependent or frequency-dependent transmission is assumed (see chapter 2 in Keeling and Rohani [17] for further details). The progression from *Infectious* to *Recovered* depends on the length of the infec-

tious period, which typically is estimated from clinical data. Finally, in the example presented in Figure 1.2, newborns enter the population (μN) and dead individuals are removed (μS_t, μI_t, μR_t) using the simplifying assumption of equal crude death and birth rates, meaning the population size is constant. The whole system can be represented by difference equations [25].

Besides disease state, the host population is often stratified by additional characteristics, describing differential behaviour influencing the risk of contracting and transmitting infection (e.g. age, sex, sexual behaviour) [17]. However, the complexity of the system of compartments and the corresponding number of equations increases inevitably as additional population heterogeneity is incorporated. Moreover, some heterogeneities are difficult to implement in compartmental models and population-level models in general, for example spatial structures, life trajectories, local interactions, network structures and adaptive behaviour. If such elements are important to take into account considering the given infection and the purpose of the model, other modelling approaches are typically applied, such as individual-based models (IBMs). We use the abbreviation IBM to refer to all infectious disease models at the individual level, including agent-based models [26].

1.2.2 Individual-based models

While population-level models, including compartmental models, can be described as top-down models due to the imposed aggregate parameters, IBMs follow a bottom-up approach. In an IBM, each individual in the population is explicitly represented and can be assigned an extensive set of relevant attributes and behaviours, for example disease state, age, sex, health trajectory, membership of subgroups, spatial location, sexual and social behaviour. Moreover, individuals can interact with each other and their environment based on a set of defined rules [29]. All events are tied to the individual, meaning that the life-course and health trajectory of each person is tracked [14]. Consequently, outbreaks at the population level emerge from the interactions between the individuals and their environment.

In the compartmental SIR model in Figure 1.2, the risk of infection was assumed to be proportional to the number of infectious individuals in the population (Equation 1.1). This equation, however, tends to overestimate the actual risk of infection in small populations, as a susceptible individual may have contact with multiple infectious individuals, but only one of those contacts will lead to disease transmission. Thus, in IBMs, Equation 1.1 is often replaced with the Reed-Frost formula [28]:

$$\lambda_t = 1 - (1 - p)^{I_t} \tag{1.2}$$

where p is the probability of an effective contact between two specific individuals in each time step. If each individual's membership of subgroups (e.g. household, school, workplace) is considered in the model, Equation 1.2 is typically extended to distinguish between infection resulting from contact within the subgroup and in the community, since contact patterns, and by that the risk of infection, often vary considerably by setting [25, 30, 31].

IBMs provide a framework for exploring how subgroups, networks and individual variation influence disease transmission at the population level, with importance for intervention strategies. For example, non-pharmaceutical interventions targeted at specific population groups, such as diagnostic testing in schools, school closure, control of within- and between-household mixing and limitation of spatial mobility, have been investigated using IBMs [32–34]. This is rather complex to carry out in a population-level model, but can still be achieved [35]. Moreover, IBMs have proven particularly useful when an individual's history of exposure influences the risk of infection or another outcome of interest (e.g. [36, 37]). Nevertheless, IBMs can be difficult to set up and are computationally intensive. Detailed models require many input parameters and it can be difficult to disentangle the contribution of each input parameter on the overall outcome [25]. Moreover, detailed data may be limited, making it necessary to make strong assumptions. However, improvements in hardware performance have reduced these disadvantages [38].

1.3 Demographic modelling of host population

Demographic structures often help to explain heterogeneity in the spread of infectious diseases. Age in particular is an important factor in the epidemiology of many infectious diseases. The proportion susceptible to an immunising infection (in the absence of vaccination) is likely to decrease with age, since it indicates the years of exposure [25]. Moreover, the number of social contacts vary considerably by age and people generally have a high proportion of their contacts with people of similar age (i.e. assortative mixing) [30]. Thus, social contact patterns are likely to affect the age-specific exposure to directly transmitted and airborne pathogens. The risks associated with an infection may also vary across ages or increase if coinciding with age-related events. For example, the morbidity and mortality of respiratory virus

infections is pronounced in the elderly population [39], adults with mumps or measles face an increased risk of encephalitis and rubella infection during pregnancy is associated with increased health risks for the child. Interventions such as vaccination programmes are typically also designed to target specific age groups [25]. Finally, epidemiological data is often collected by age groups.

Other demographic characteristics that may be relevant to include in the host population include sex, spatial structures, household membership, social structures, school enrolment, workplace, sexual orientation or health status (noncommuncable diseases). The relevance of each component depends on the infection and the purpose of the model. For sexually transmitted diseases, sex and characteristics describing sexual behaviour are typically taken into account [40]. Meanwhile, households have been given particular attention when modelling infectious diseases spread through close contact, such as influenza and COVID-19 [35, 41–47], because household members tend to have more frequent and intimate contacts of a longer duration [31, 48, 49]. Moreover, household members often belong to different age groups (e.g. parents, children, grandparents) with distinct health trajectories and are part of different formal and informal groups outside the household (e.g. schools, workplaces, sports clubs, social groups). Households may thus facilitate the spread of an infection between population groups that are not directly connected (i.e. bridging function) [50].

The demographic method used to construct the host population, however, limits the feasibility of incorporating demographic characteristics beyond age and sex. The population heterogeneity required considering the infection, setting and research question should thus be taken into account when choosing the type of method, as well as the necessary degree of resemblance between the host population and a real population.

1.3.1 Static, stationary and stable populations

Many demographic characteristics change throughout an individual's life. People are ageing, moving between households, forming new relationships, developing chronic illnesses etc. Moreover, newborn children and immigrants enter the population, while others leave the population due to death or emigration. Thus, populations evolve over time. The demographic structures of the host population in infectious disease models, however, has traditionally been assumed to be static, stationary or stable [27]. In a static population, births, deaths, migration and ageing are disregarded completely. If population dynamics are incorporated in the form of a stationary population, the

number of people entering and leaving each age group in the population are assumed to be equal, implying no population growth and a constant age distribution [17]. Stationary populations are a special case of a stable population. Both have a constant relative age distribution, but the latter may be growing or shrinking as the assumption of equal flows in and out of the population is relaxed. An actual population, however, is rarely perfectly stable, but if subjected to constant vital rates for an extended period of time, it will gradually converge to stability. The resulting age structure of the stable population is entirely determined by the vital rates and may differ markedly from that of the initial population [51].

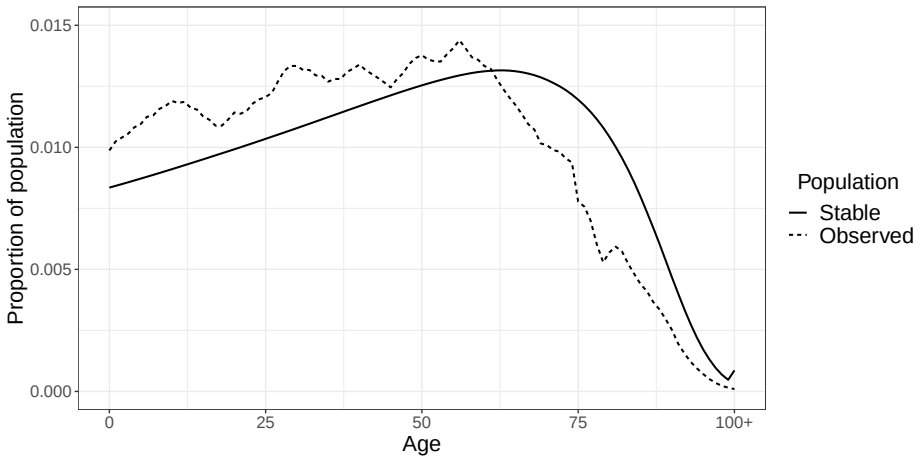

Figure 1.3. The relative age distribution of Belgium in 2021 (dashed line) and that of the corresponding stable population (solid line) assuming vital rates of 2021 and no migration. Source: Statbel and own construction of stable population.

Stable and static populations may be justified when considering epidemics occurring over a time frame that is sufficiently short to limit the effect of demographic processes (e.g. weeks, months). However, as the time frame extends or when modelling endemic diseases, the outcome may be very sensitive to assumptions about fertility, mortality and migration [25]. Population-level predictions of epidemiological outcomes may also be affected by differences between the composition of a stable host population and that of the actual population. The age distribution observed for Belgium in 2021, for example, deviates markedly from that of the stable population implied by the fertility and mortality schedules for Belgium in 2021, assuming no migration (see Figure 1.3). Finally, the epidemiology of an infectious disease may be influenced by demographic change in the host population.

1.3.2 Representative population structures

Host populations with a representative demographic composition, in the sense that they resemble an observed population, are increasingly being incorporated in models of disease transmission. Generally, two different demographic methods are used to model host populations with realistic structures, the cohort component method and microsimulation [52]. The cohort component method is used to create population-level models, which have many similarities with compartmental models. The population is divided into subgroups with differential risks of demographic events, typically limited to fertility, mortality and migration, and the changes in each subgroup is computed separately over time. In most cohort component models, the population is only segregated by age and sex [25]. The cohort component population model is combined with a compartmental disease model by breaking each disease compartment down by the demographic subgroups.

Microsimulation is a commonly used term to describe IBMs in demographic research. Similar to IBMs for disease modelling, microsimulation constitutes a useful tool for incorporating a higher degree of heterogeneity in the host population and the processes underlying its demographic structures. The set of demographic attributes can expand in a flexible manner, individuals can interact (e.g. form a union), social structures and networks can be implemented (e.g. households, kinship networks) and the life history of an individual can be taken into account when assessing the probability of future events [53, 54]. A demographic microsimulation can easily be combined with an IBM for disease transmission by adding the disease state to each individual's attributes and other components relevant for disease transmission. In practice, the demographic model and the infectious disease model are typically developed jointly.

Regardless of the method, demographic processes need to be taken into account to obtain a somewhat realistic evolution in the composition of the host population. If the vital rates remain constant, demographic changes will be observed at first as the effect of past changes in fertility, mortality and migration emerge. However, the population will eventually converge to the implied stable age composition. Consequently, some studies advance to nonstable host populations with dynamic demographic processes, especially if the period under consideration is long [36, 44, 47, 55–60]. We refer to these as dynamic populations.

1.3.3 Demographic change and disease transmission

A population's history of mortality, fertility and migration is reflected in its age distribution [25]. This is seen in Figure 1.3, where the age distribution of the observed population is rather uneven (dashed line) as a result of past changes in the vital rates, while the stable age distribution is smooth (solid line). The historical declines in mortality described earlier (see Figure 1.1) were in most high-income countries of today followed by substantial fertility declines starting in the late 19th century or early 20th century. Over decades, the populations transitioned to a regime of low mortality and fertility, a process also referred to as the *demographic transition* [61]. During the transition, the population grows rapidly, but the stages prior to and after the completion of the transition are associated with low population growth or even declines [62]. By now, the demographic transition has reached all world regions, although some populations are still in the early stages [63].

The changing mortality and fertility levels induce changes in the population age composition. Populations at the early stages of the demographic transition have a younger age structure than those at the later stages, while populations that have completed the demographic transition will be characterised by an increasingly older age structure, if fertility and mortality remain low [62]. In the course of the demographic transition, the median age of the population will start increasing, a phenomenon referred to as *population ageing*. All countries in the world are currently experiencing population ageing, and some low- and middle-income countries are ageing at an unprecedented speed [64]. However, the proportion of the population aged 65 years and older is still substantially higher in upper-middle income and high-income countries, including Belgium (19.4% in 2021) [65].

Like many other high-income countries of today, Belgium faced an increase in fertility in the mid-20th century, also labelled the *baby boom*. This was followed by a decline to below-replacement fertility in the 1970s, corresponding to a total fertility rate (TFR)[1] below 2.1. Since then, the TFR in Belgium has remained well below the replacement level, with consequences for the population age composition [65]. The large size of the generations born in Belgium in the mid-20th century is seen from the wide base of the population pyramid for 1970 (grey bars) in Figure 1.4. Considerably smaller generations followed (blue line), and currently the generation of the baby boom is moving into the older age groups, causing a temporary acceleration of population ageing, which is expected to continue in the next few decades (red line) [66].

[1] "The average number of children a woman would bear if she survived through the end of the reproductive age span and experienced at each age a particular set of age-specific fertility rates" [25].

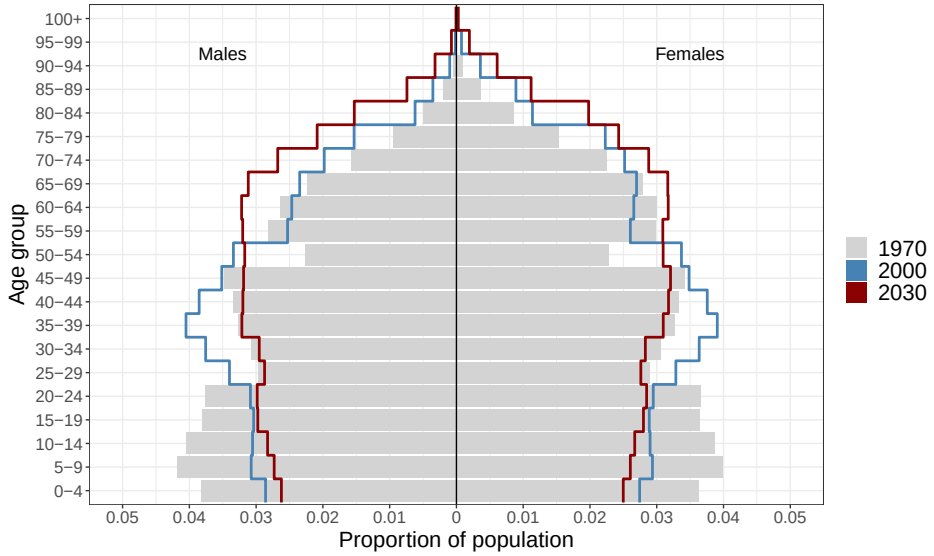

Figure 1.4. Age-sex groups as proportions of the Belgian population in 1970, 2000 and projected in 2030. Males on the left side, females on the right side. Source: World Population Prospects 2019, United Nations Population Division.

These historical declines in fertility and mortality and the induced changes in the population age composition are associated with a decreasing incidence of infections traditionally considered to be childhood diseases, such as measles [42, 67–71]. Considering measles prior to vaccination, a reduction in births imply that the proportion of susceptible individuals (young children) relative to immune individuals (older children and adults) is decreasing, hence measles circulation drops and the mean age at infection increases [72]. Moreover, the shifting age composition in ageing populations imply that the burden of infections with a higher incidence and severity in older adults is expected to increase [73–75]. Older adults face an increased susceptibility to many infections and a higher risk of a severe outcome in case of disease due to the progressive deterioration of immune functions with age (immunosenescene) [76]. The COVID-19 pandemic, for example, had a disproportionate impact on the older adult population [77, 78].

1.4 Motivation and aim

The impact of demographic change on the epidemiology of infectious diseases has mainly been investigated by considering changes in the population age composition. This has undoubtedly improved the understanding of long-term transmission dynamics at the population level, with importance for the evaluation of immunisation pro-

grammes. However, demographic changes with relevance for disease transmission go beyond changes in the age structure and are not always noticeable at the population level. As mentioned earlier, age- and household-structured host populations have proven valuable for modelling infections transmitted through close contact [43, 45], which are the focus of this dissertation. However, only few studies consider a dynamic evolution in the population age structure as well as the household structure (e.g. [41, 42, 47, 58, 79–81]), although both are affected by changing patterns of fertility, mortality and migration.

The household size and composition of nuclear families, for example, are highly influenced by the level and timing of fertility. The postponement of parenthood implies that parents gradually will be older when the last child leaves the household. Changing trends in divorce rates, re-partnering and custody arrangements mean that an increasing share of children are members of more than one household and reconstituted families. These phenomena have been observed for decades in many high-income countries and are described in demographic research as the *second demographic transition* [82, 83]. Changing patterns in fertility and family formation may alter the transmission dynamics within the affected families and households. This could have potential spillover effects in the rest of the population due to the high number of contacts of children and the special bridging function of their households.

Additionally, population ageing implies that the population household structure will be increasingly influenced by the living arrangements of older adults, which tend to differ from those of younger age groups. In Belgium, for example, the large majority of older adults live alone, with a partner or in a long-term care facility (LTCF), which typically has a large number of residents [84]. While small households are associated with a smaller risk of acquiring an infection that spreads through close contact, large households are associated with an increased risk [43, 85]. Consequently residents in LTCFs accounted for a substantial share of the morbidity and mortality associated with COVID-19 [86–88], since it is a high-risk population in a high-transmission environment. Thus, relatively small population groups may be disproportionately affected by some infectious diseases due to differential demographics, particularly if the demographics of the group is associated with an unfavourable health trajectory, as in the case of LTCF residents.

The aim of the dissertation is to explore and improve infectious disease models with dynamic host populations with the purpose of investigating the impact of demographic structures and changes on the transmission and burden of infectious diseases transmit-

ted through close contact. To achieve this goal, we need an understanding of several relationships, which leads to the research questions: (i) what is the relationship between age and household structures at the individual and population level and how are these shaped by demographic processes and changes therein? (ii) to what degree are transmission dynamics at the individual, household and population level shaped by age and household structures? (iii) how are population age and household structures expected to evolve over time and how are the demographic changes expected to affect the transmission dynamics and disease burden? Addressing these research questions requires interdisciplinary research combining demographic modelling and infectious disease modelling, which will be presented in this doctoral dissertation.

1.5 Overview of the dissertation

This dissertation consists of five scientific investigations in which we explore and further develop existing methods used to incorporate demographic change in infectious disease modelling and present multiple applications in an attempt to unravel the relationship between population structures and the transmission and burden of infectious diseases.

In Chapter 2, we present a systematic review of the demographic methods and data used to incorporate dynamic host populations in models of infectious disease transmission. We provide an overview of the demographic methods and techniques used to model the initial population structure and the demographic processes (fertility, mortality, migration and household transitions) and discuss their advantages and limitations. Finally, we discuss the potential implications certain demographic methods and assumptions have for the population composition and potentially for epidemiological outcomes.

With the advantages and limitations of the existing infectious disease models incorporating demographic change in mind, we develop a demographic microsimulation which is presented in Chapter 3. The age- and household-structured model is informed with Belgian census data and population register data, making it possible to obtain a high level of detail in the initial population as well as the demographic processes. We specifically simulate the Belgian population from 2011 to 2050 and consider fertility, mortality, migration and household transitions. The microsimulation is developed to be easily combined with a disease transmission model and we present two applications in Chapter 4 and 5. In the first study, we explore the relationship between age and household structures at the individual and population level with the purpose of in-

vestigating how these shape the disease transmission dynamics of an epidemic. In the second study, we model outbreaks of diseases resembling COVID-19 and pandemics influenza and explore how the transmission dynamics and disease burden are affected by population ageing. We specifically consider the changing living arrangements in the older adult population.

In Chapter 6, we present ongoing work investigating the impact of demographic change on the epidemiology of varicella-zoster virus (VZV) and herpes zoster (HZ) in the US population from 1960 to 2020. This involves an individual-level modelling framework similar to the microsimulation developed for Belgium. Finally, in Chapter 7 we highlight the main findings of the research presented in this dissertation and discuss limitations. We also discuss how the presented studies can provide a foundation for further research.

Chapter 2

Incorporating human dynamic populations in models of infectious disease transmission: a systematic review

This chapter is based on published work: "Møgelmose, S., Neels, K. & Hens, N. Incorporating human dynamic populations in models of infectious disease transmission: a systematic review. BMC Infect Dis 22, 862 (2022)."

2.1 Background

In response to infectious disease threats, mathematical and computational models have proven to be invaluable tools in understanding the spread of infectious diseases in human populations and in quantifying possible disease control strategies as well as evaluating public health interventions, particularly in situations where a controlled trial is ethically or practically unfeasible [14].

The host population studied in an infectious disease model is typically assigned demographic characteristics to account for heterogeneity that may influence the spread of an infection. Population age structure, for example, is commonly included as epidemiological parameters often vary by age, such as the proportion susceptible to immunising infections, which typically decreases with age. Furthermore, contact patterns relevant for the spread of close-contact infections are highly assortative with age, which may affect the exposure to infection. Susceptibility to infection may also

vary across ages, as well as the risks associated with an infection [15, 25, 30]. Other demographic characteristics and subgroups (e.g. sex, households, schools and spatial structures) may also play an important role in the transmission process of an infectious disease [48, 50, 89]. This often implies that the burden of an infectious disease in a population may also be influenced by the relative size of those demographic groups, also referred to as the population composition.

The composition of a population tends to change over time as a result of changes in the underlying demographic processes, which include ageing, births, deaths and population movements. Nevertheless, demographic change is a slow process and is often not incorporated in models of infectious disease transmission, since the time period under consideration tends to be short. Moreover, it is often useful to disregard demographic change when focusing on how epidemiological factors alone influence different outcomes [67].

For some infections, settings and research questions, however, the realism of the population composition and how it evolves play an important role. This often applies to analyses of disease transmission dynamics and public health interventions over a longer time frame, where demographic changes are to be expected. Fertility declines, for example, have in some cases been linked to increases in the average age at infection of diseases traditionally considered to be childhood diseases [42, 67–69, 71]. This may affect the disease burden of infections associated with increased morbidity and mortality in certain age groups or during age-related events such as pregnancy [25]. The burden of infections with a higher incidence and severity among the elderly is also expected to increase as a population undergoes ageing, as has been seen with herpes zoster [74, 75]. Such relationships can only be investigated by allowing for demographic change in the host population.

Demographic change can be introduced in models of infectious disease transmission in various ways. The population can be subjected to constant fertility and mortality rates for an extended period of time, where demographic change will result from the gradual convergence of the population to the implied stable population with a constant relative age distribution and a fixed growth rate [51]. In many cases, this provides a useful approach for investigating disease transmission dynamics in a population with a changing composition induced by preceding trends in fertility, mortality and migration. However, as the time period expands, the assumption of constant demographic rates becomes implausible. Thus, for the evaluation of long-term effects, it may be important to consider demographic change in the host population by explic-

itly considering and including dynamic demographic processes. We will refer to this as a dynamic population which is the main focus of this paper. Dynamic populations are incorporated in an increasing number of mathematical and computational models for infectious disease transmission. These models have shown an important impact of demographic change on the long-term dynamics of infectious diseases, as well as for the effectiveness of immunization programmes. For example, long-term demographic changes have been found to have a considerable effect on the epidemiology of varicella and herpes zoster, implying that the demographic assumptions have an impact on the predicted burden of disease [75, 90].

The demographic methods used to incorporate a dynamic host population in models of infectious disease transmission vary considerably. The methods range from adjusting the age distribution over time according to population projections to complex models with dynamic demographic processes and subgroups such as households (e.g. [42, 75]). This includes population-level models and individual-based models (IBMs). We use the term *IBM* to refer to all models at the individual level, including microsimulations and agent-based models [26]. IBMs are increasingly used to model disease transmission, however, it is unclear whether their flexibility also enhances the level of detail incorporated in the demographic modelling. Finally, various demographic assumptions are applied in models of disease transmission, such as no migration, but the implications thereof are not always explained.

With this systematic review, we provide an overview of the methods and techniques used to model dynamic population structures in the context of infectious disease modelling, which to our knowledge has not previously been attempted. We discuss the advantages and limitations of various modelling techniques in order to improve the understanding needed to evaluate their suitability in a given study. Moreover, we discuss the potential implications certain demographic methods and assumptions have for the population composition and potentially for the epidemiological outcomes. As previously mentioned, dynamic host populations are in many cases not incorporated in models of infectious disease transmission and typically for good reasons. Thus, the aim is to identify the smaller group of infectious disease models where dynamic population structures have been a major point of attention. To obtain this, while taking the feasibility of the search into account, we limit our review to infectious disease models with a focus on demographic change. We differentiate in terms of the method that is used to model the host population, the different demographic processes, as well as the data and techniques used to model each demographic process.

2.2 Methods

We carried out a systematic review in accordance with the guidelines of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses [91]. The methods and procedures are described in a protocol (see Appendix A section A.1).

2.2.1 Search

We searched PubMed and Web of Science Core Collection for articles published up to August 25th 2020 without language or time restrictions using the following search string in titles and abstracts:

> *(demography OR "demographic transition" OR "demographic change*" OR "population change*" OR "household structure*" OR "household composition*" OR "population ageing" OR "population aging" OR "aging population" OR "ageing population") AND (infect* OR vaccin* OR epidemic* OR communicable) AND (model* OR simulat*) NOT (animal* OR plant*).*

The asterisk in some search terms represents any group of characters, including no character (e.g. infect*: infected, infection, infectious etc.). The search string includes terms related to demography since we mainly expect dynamic host populations to be incorporated in papers with a focus on demographic change. Broader search terms (e.g. demograph*) and the search of full-text and supplementary material would provide a more thorough search but would result in an unfeasible amount of hits. A detailed overview of the result of each search term and the overall search strategy is shown in Table A.1 in Appendix A. The results of the search strategy were managed in Endnote X9.

2.2.2 Eligibility criteria

The eligibility criteria were defined by two researchers (SM and NH) prior to screening. We included research papers on mathematical and computational models for infectious disease transmission in a human population. The host population should at least be divided into five age groups. Moreover, the population should result from a model including at least fertility and (all-cause) mortality as dynamic processes. No requirements are made for disease-specific mortality, if included in the model. The demographic model can be included explicitly or population structures from another source can be used as input to the disease transmission model, as long as this population is the result of a demographic model explicitly considering dynamic trends for

fertility and mortality. This implies that models assuming constant fertility or mortality rates throughout the entire study period are excluded. However, models with constant rates in a limited part of the study period are still included. Articles only describing the technicalities behind a method or a software tool without applying it to any population are also excluded. Finally, models limited to high-risk groups (MSM community, injecting drug users etc.) are excluded. The screening and selection processes are presented in Figure 2.1. Titles, abstracts and full-texts were screened by a single reviewer (SM) and discussed with the last author (NH) in case of doubt. We also identified articles by screening reference lists of the included articles.

2.2.3 Data extraction and analysis

For all eligible articles, we retrieved and classified information as follows: (1) Setting and population characteristics: country/region, population, demographic characteristics and time horizon; (2) Model specifications and data: model type, demographic processes considered and source of demographic data; (3) Modelling of demographic processes: starting population, fertility, mortality, migration, household networks and sensitivity analysis of demographic assumptions; (4) Specifications of disease transmission model and analyses: disease, vaccination, social mixing and cost-effectiveness analyses. Models from the same article were included if they were eligible and differed from each other in setting, model specifications or demographic processes. To make it clear that models originated from the same article, letters were added to the article number in figures and tables.

2.3 Results

We identified 881 articles (after removing duplicates) by searching the databases PubMed and Web of Science Core Collection with the search string mentioned under Methods. Based on the defined eligibility criteria, we screened titles and abstracts and excluded 724 articles. For the remaining articles, a full-text analysis was carried out in case fulfilment of the eligibility criteria was uncertain. Most articles excluded at this stage were assuming constant fertility and/or mortality rates. We identified 13 articles through snowball sampling. Finally, 46 articles, containing 53 different models, were included in the qualitative analysis. The data retrieved from each article can be found in Appendix A in tables A.2 to A.9. In this section, the term *study* is used to refer to all models within one article.

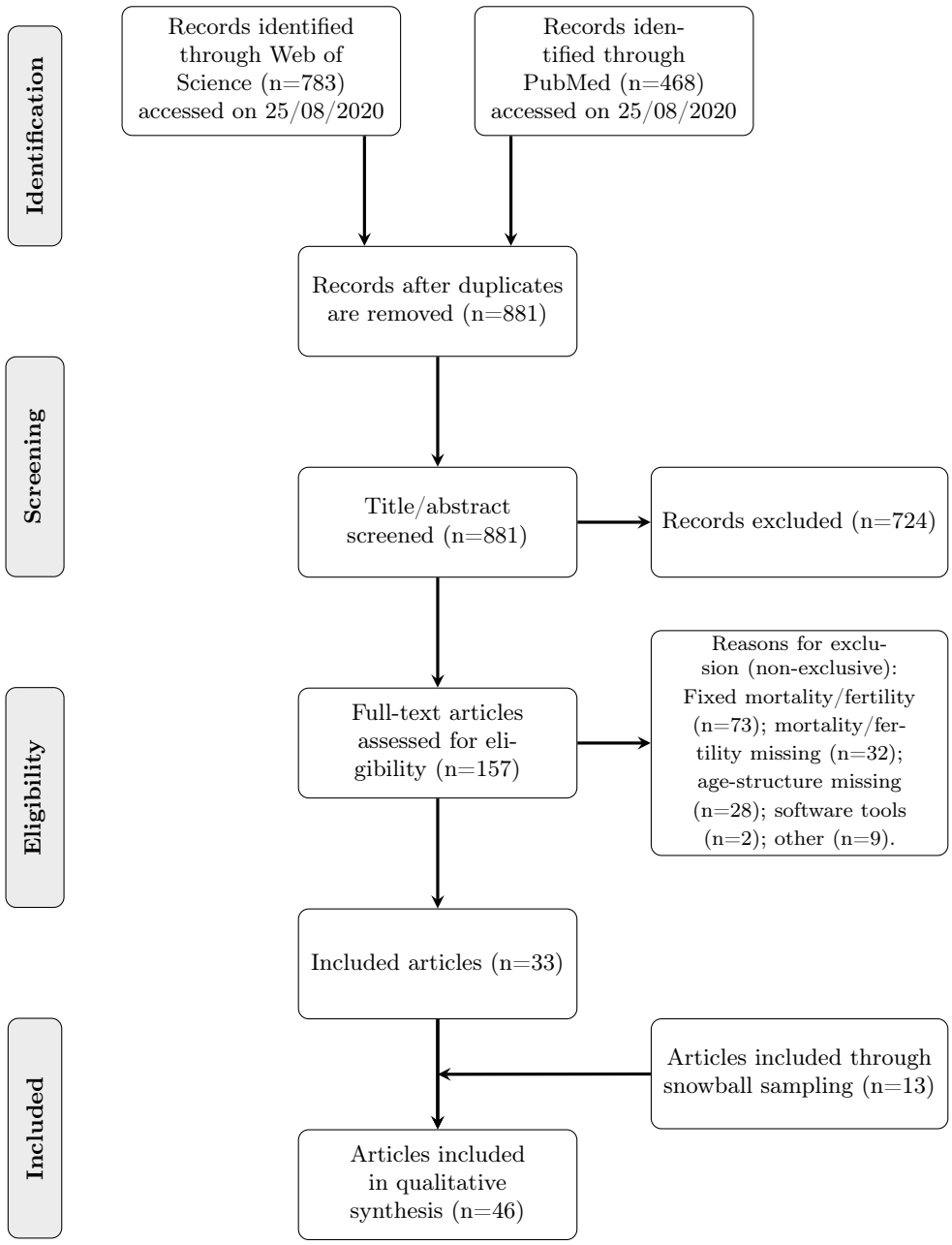

Figure 2.1. PRISMA flow diagram of the article selection process.

2.3.1 Setting and time period

In the included studies, populations were modelled for countries, regions and cities in Europe (20), Asia (16), Africa (14), North America (12), Oceania (9) and South America (6). Twelve studies covered multiple populations. In most of the studies, past as well as future time periods were modelled, six studies only covered a historical period while five studies only looked at projections of the future. The length of the modelling period varied between 10 and 250 years as seen in Figure 2.2.

2.3.2 Model type

We divide the demographic models in the included studies into three types: (i) disease models that use demographic population prospects as an external input (EPMs: external population models), (ii) cohort-component-based models that use cohort-component projections to model demographic change (CCBMs) and (iii) individual-based models that model demographic events at the level of individual life courses (IBMs). First, EPMs draw the annual population composition from an external source and use it as an input for the disease transmission model instead of modelling the demographic processes explicitly. Given assumptions for the different components of demographic change, statistical agencies often generate projection sets that provide annual information on population composition, typically by age and sex. In this approach, the population composition is allowed to vary over time, but population dynamics cannot be attributed to changes in fertility, mortality or migration separately, because only the resulting population composition is used.

Second, in CCBMs, the population is divided into subgroups to which group-specific rates for fertility and mortality are applied in each projection step to work out population change over time. In most cases, CCBMs do not consider household or family dynamics. Depending on the assumptions made, emigrants and immigrants are subtracted and added, respectively, by age and sex [51]. As is the case for the EPMs discussed earlier, CCBMs are typically integrated into a compartmental disease transmission model by adding the demographic sub-groups (e.g. age groups) to each compartment. As a result, both disease transmission and population dynamics are modelled at the population level.

Third, in IBMs, the unit of analysis is the individual. In IBMs, all individuals are assigned a set of attributes (e.g. age, sex, marital status) and are in every time interval subject to covariate-specific risks of demographic events, such as union formation or dissolution, fertility, mortality and migration [54]. However, the number of covariates

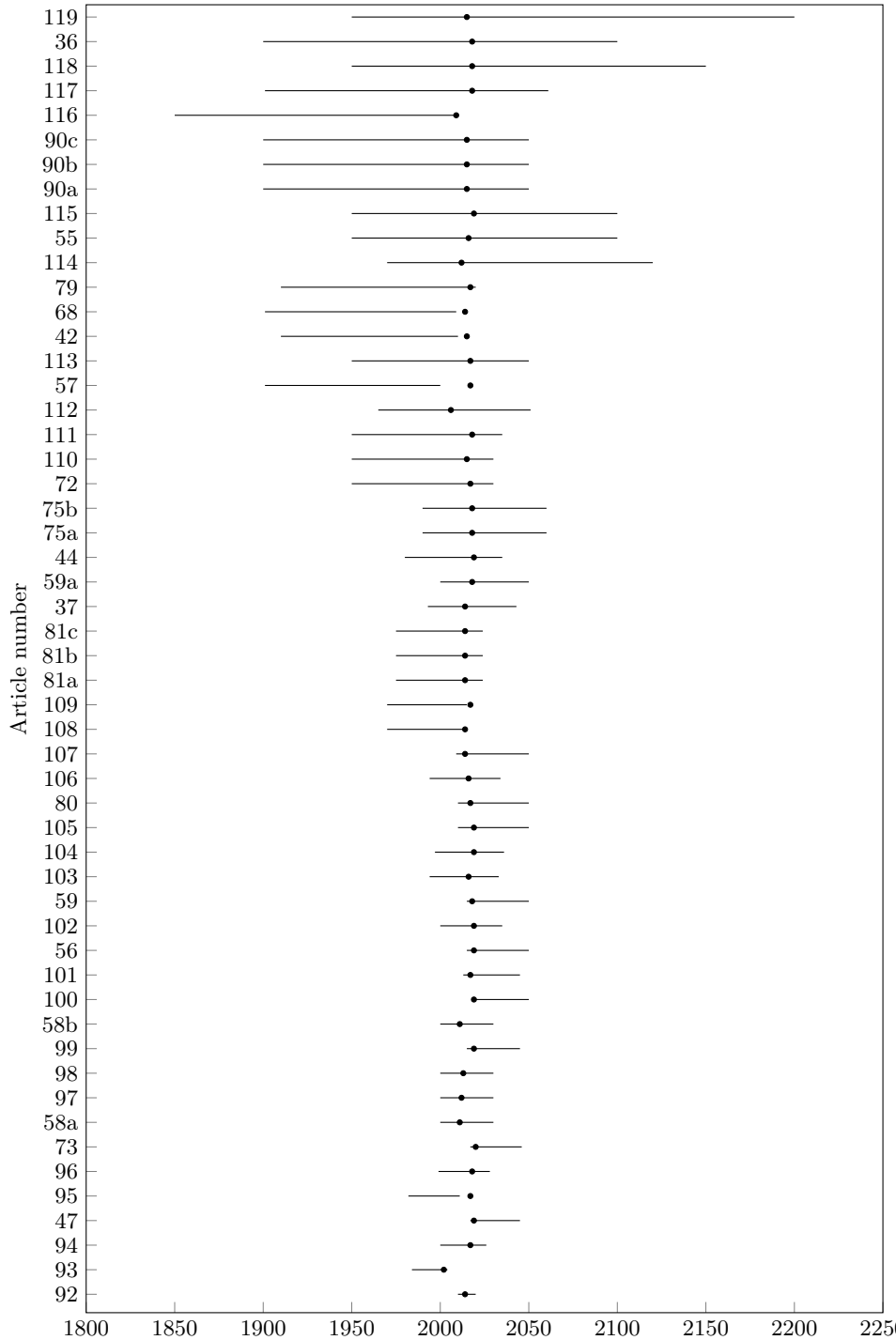

Figure 2.2. Modelling period by article number in the reference list (points indicate years of publication)

included in each demographic process varies substantially between the included IBM studies. Given the predicted probabilities, a random number generator is used to determine whether an individual experiences the event and the individual's attributes are updated accordingly [120]. This makes it possible to track the life course of each individual. In order to simulate disease spread in a demographic IBM, a disease state is added to the individual attributes. Moreover, interactions between individuals as well as subgroups (e.g. households, schools) and network structures (e.g. mobility networks) can be included when relevant for disease transmission [14], which was done in several studies.

In 14 studies, existing population prospects were used as external input in the disease model (see Figure 2.3 and Figure 2.4). CCBMs and IBMs were applied in 24 and 7 studies, respectively, while one study applied both approaches. This implies that the majority of the included studies use population-level models, but fertility, mortality, and in some cases also migration, were in most cases modelled explicitly.

2.3.3 Starting population

Most studies based the starting population on the observed population composition in the first year of the modelling period or on a population sample. In 18 studies, however, the starting population was generated by simulating demographic events and disease transmission in an initial population using a set of demographic rates for a defined period of time (see Appendix A tables A.2 and A.3 for more detail). In this way, an epidemiological equilibrium can be obtained in the starting population, while respecting any specified demographic constraints (e.g. demographic generation intervals, birth intervals). In most of these studies, the fertility and mortality rates remained constant in the initialisation period, which eventually leads to a stable population.

The relative age distribution of a stable population is not influenced by the initial age distribution, but is entirely determined by the fertility and mortality rates assumed [51]. Consequently, the age composition in the initial non-stable population and the resulting stable starting population may differ considerably. Some studies compensated for this by adjusting the demographic rates used to generate the starting population. Household membership was included in the starting populations of the eight models incorporating households. Individual-level data on household position and composition were lacking in most studies and marginal distributions of household size and age compositions were applied instead. Different algorithms and

constraints were applied to obtain somewhat realistic age differences between household members in the starting population. In several studies, household members were assigned different positions (e.g. adult in a union, single adult, child) based on data or defined rules.

2.3.4 Demographic processes

The 32 studies using IBMs and CCBMs included at least dynamic trends for fertility and mortality. In most of these studies, covariates were included in the fertility and mortality processes (e.g. age and sex), as seen in Figure 2.3 and Figure 2.4 and further described below. Migration was included in 22 studies and households were only incorporated in eight models. Finally, demographic sensitivity analyses were performed in nine models. In some EPMs [37, 103, 104, 106], it was assumed that the number of births in a given year equals the size of the youngest age group, while a decrease or increase over time in all other age groups was ascribed to mortality and immigration, respectively. These details are not included in Figure 2.4 because the changing size of an age group cannot be ascribed to one demographic process alone. The external population prospects typically result from a set of assumptions for fertility, mortality and migration, but since these are not modelled explicitly in the infectious disease model, EPMs are not considered in the further discussion of the demographic processes. The subgroups by which the population in an EPM is decomposed are described in Figure 2.4 instead.

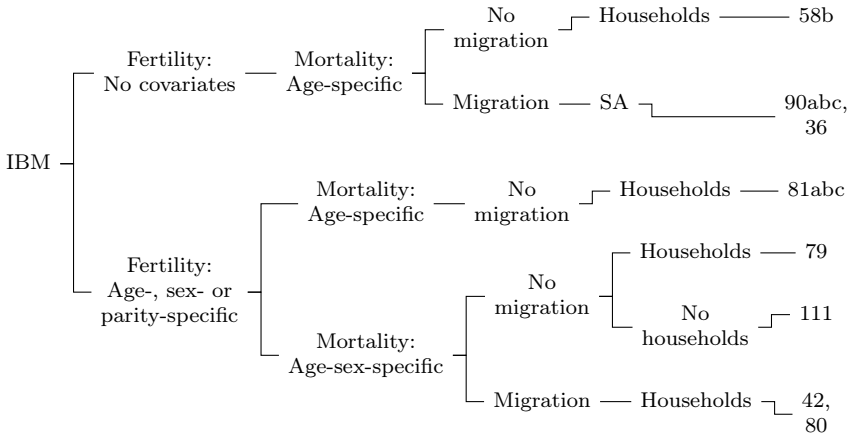

Figure 2.3. Branching diagram of IBMs, demographic processes and covariates with article numbers (SA: Sensitivity analysis).

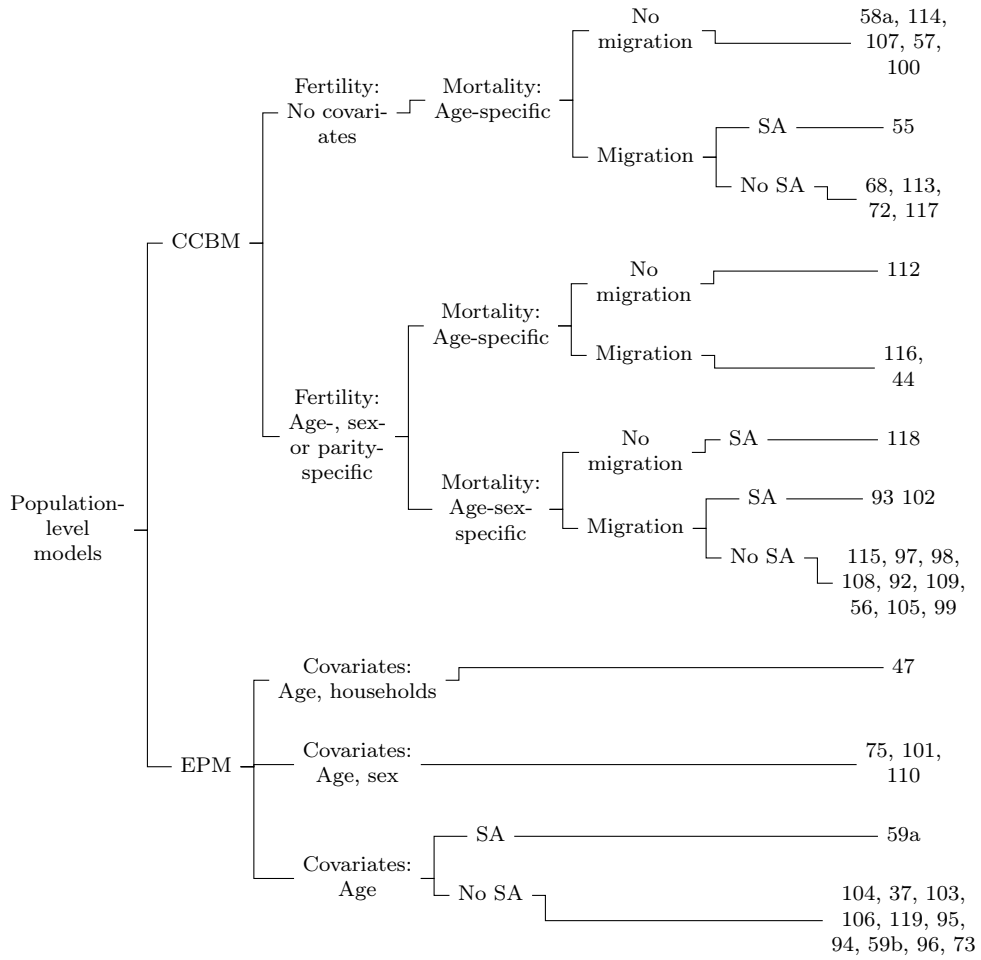

Figure 2.4. Branching diagram of population-level models, demographic processes and covariates with article numbers (SA: Sensitivity analysis).

Fertility

In order to model past fertility patterns, the large majority of models used observed fertility rates, probabilities or birth numbers, as shown in Figure A.1 and Table A.4 in Appendix A, which in most cases were obtained from official statistical agencies. In case annual estimates of vital statistics were not available, interpolation, averages over multiple years or step-wise functions were used. Only three models applied a scenario-based approach with an assumed trend. A larger variety was seen in the methods for projecting future fertility trends. Official fertility projections were applied in 15 models, which were mainly taken from national statistical agencies

or the United Nations World Population Prospects. Eight models carried the last observation forward, meaning that future fertility trends were assumed to remain at the last observed level. One study combined the previous two approaches by applying official projections for the years available, and for the remaining projection period, the last projected value was carried forward. Authors developed their own scenarios for future fertility trends in five models, either expressed as a yearly percentage change or as policy scenarios, while extrapolation was applied in one model. Finally, the applied approach was unclear in one model (the author was contacted but does no longer have access to the information).

About the same proportion of IBMs and CCBMs, which amounted to 15 models in total, included no covariates when modelling fertility. In most of these cases, crude birth rates (CBR) were used. Age was included in the fertility process in one model and 16 models included sex as well as age, meaning that age-specific fertility rates (ASFR) were applied to females in their fecund age range (typically 15–49 years of age). Five models, which were all IBMs, took birth parity (number of children ever born to a female) into account, in addition to age and sex, and two of these also considered birth interval by assuming a minimum amount of time between subsequent births for a given female.

Mortality

Past (all-cause) mortality patterns were modelled using observed mortality rates or numbers of deaths in most models as shown in Figure A.2 and Table A.5 in Appendix A. Interpolation, averages over multiple years or step-wise functions were applied in cases where yearly estimates were not available. Future mortality patterns were modelled using official projections in 16 models and 10 models used the most recent observation for the whole projection period. Other methods, including extrapolation and the scenario-based approach, were applied in a smaller number of models. Mortality was age-dependent in 21 models and age-sex-dependent in 16 models. Furthermore, several models included disease-related mortality or certain risk factors (not included in Figure A.2).

Migration

Migration was included in 24 models in the form of net migration and mainly obtained from official estimates and projections (see Table A.6 in Appendix A). Three of these models assumed that the composition of the migrant population was similar to that of the native population, but it was acknowledged in the studies, that in reality these

tend to differ markedly [42, 113, 115]. The age and age-sex distributions of the net-migrant population were considered in 9 and 10 models, respectively. In most of these cases, a lack of data made it necessary to assume that the composition of the net-migrant population remains constant over time. Internal migration was modelled in a broad sense in three models, namely as migration between rural and urban areas and without considering any covariates. Migration was incorporated in the majority of CCBMs and in about half of the IBMs.

Households

Households were incorporated in eight models, of which seven were IBMs and one was an EPM [47] (see Table A.7 in Appendix A). The households evolved dynamically over time in all models. In the IBMs, individuals move between households or create new households (e.g. child leaving parental household). In most cases, the probabilities for household transitions were fixed over time and equal across all ages eligible for a given transition. Households were also dynamic in the EPM, however, a new population of individuals was generated and assigned to households at the beginning of each simulation year according to an algorithm using the observed or projected age composition and the last observed age distribution by household size. This implies that changing household structures could not be traced back to individual household transitions.

Sensitivity analyses

Sensitivity analyses of population projections were included in a small number of models (see Table A.8 in Appendix A). Alternative scenarios for the overall age distribution or for fertility, mortality and/or migration separately were applied in seven models, two models quantified the uncertainty of the demographic parameters with confidence and prediction intervals, respectively, while one model compared step-wise functions with interpolation between five-year estimates.

2.4 Discussion

We identified 46 studies, which contained 53 models for infectious disease transmission in populations with dynamic demographic processes. The dynamic population models in the included studies varied in the methodology, the demographic characteristics and processes included in the model and overall complexity. Population- level models, EPMs and CCMBs, were most common, while individual-based models were least

common. This was to be expected, as the use of IBMs in infectious disease modelling is relatively new, but has been expanding in recent decades [26]. Moreover, fertility, mortality and migration were modelled explicitly in most models (CCBMs and IBMs), while EPMs using population prospects from a statistical agency as external input made up a smaller share of the models. One demographic model cannot generally be considered superior to another because its suitability must be evaluated jointly with the infectious disease model and the aim of the given study. However, the demographic methods and assumptions serve different purposes and are associated with different limitations and possibilities for extensions.

In EPMs, the population composition changes over time as individuals enter and leave the population, but the demographic change cannot be traced back to fertility, mortality and migration separately, since these processes are not modelled explicitly in the infectious disease model. The underlying demographic assumptions are often available from the statistical agencies developing the population prospects. The population heterogeneity in EPMs is determined by the level of detail provided in the population prospects, which often is limited to the age or age-sex composition. In infectious disease modelling, however, demographic variables additional to age and sex are often not required. Thus, EMPs provide a straightforward and relatively simple implementation of a dynamic host population in cases where it is not of interest to consider fertility, mortality and migration separately. However, it is worthwhile to explicitly state and discuss the demographic methods and assumptions underlying the population projection, even when they are not modelled explicitly, as different demographic assumptions may give rise to quite different population dynamics, which in turn may affect the epidemiological outcomes.

In CCBMs, the processes that generate changes in the population composition are explicitly incorporated as assumptions or sub-models. Thus, a strength of CCBM is the possibility to assess variation in epidemiological outcomes given alternative scenarios for each of the demographic components. This is particularly relevant when considering time periods far into the future. Moreover, characteristics beyond age and sex can be included in a CCBM, but the system of demographic subgroups and disease state compartments becomes quickly very complex. Consequently, none of the included CCBMs incorporated demographic subgroups beyond age and sex. The highest degree of flexibility is provided by IBMs, which make it possible to include more heterogeneity in both the population and the disease transmission process. The majority of the IBMs (60%) included households as well as other demographic subgroups (e.g. schools, workplaces). Demographic subgroups and networks are important for

the transmission process of many infectious diseases. Households especially play a central role due to the higher frequency and intimacy of social contacts among people living together [43, 48]. Individual-based modelling may also make it easier to track the life course and/or health trajectory of individuals. Thus, past events can be taken into account, when determining the probability of future events. Five IBMs, for example, included birth intervals and/or parity in the fertility process, which would be more difficult to accomplish in a population-level model.

Nevertheless, similar methods were used to model fertility, mortality and migration in the majority of models explicitly incorporating demographic processes (CCBMs and IBMs). In about 40% of these models, the number of births was modelled proportionally to the size of the total population (crude birth rate), disregarding the age structure, and changes therein, of the female population [121]. As a result, the number of births in the dynamic population will only be correct as long as the age-sex composition remains similar to that of the population on which the crude rate was based. Age-specific fertility rates (ASFRs), which are directly standardized for age-sex composition, are preferable in that respect and were used in most of the remaining studies. To the extent that characteristics of household members, parents or siblings (e.g. age, vaccination status) or other kinship-related factors play a role in the disease transmission process, it is necessary to incorporate these characteristics in the fertility process as well. This was seen in the aforementioned IBMs, which included birth interval and parity in the fertility process in order to obtain appropriate generation intervals and household compositions.

The relationship between (all-cause) mortality and age was acknowledged in all IBMs and CCBMs and about half also considered the impact of sex. In most settings, however, mortality in the middle and older ages has been shown to vary by a considerably larger set of factors, including household composition, living arrangement and marital status, especially among males [122]. This could be particularly relevant to take into account when modelling ageing populations, where household structures are increasingly influenced by the developments in the elderly population. However, this would require very detailed data which often is not accessible. Thus, characteristics beyond age and sex were unsurprisingly not considered in the mortality process of any of the included studies.

In addition to fertility and mortality, about two-thirds of the CCBMs and IBMs modelled migration. A typical age pattern is often observed in migration, with a peak in young adulthood and in childhood, because of children joining their parents in migration. In high-income countries, a smaller peak is also observed around the retirement ages [51]. This was incorporated in the majority of the studies modelling migration. Migration patterns are highly complex and typically associated with a high degree of uncertainty [123, 124], which could partly explain why migration was not taken into account in about a third of the models. However, in most countries, migration flows cannot be considered negligible and changes in population size and composition may be biased if migration is ignored [125]. Statistical agencies typically provide estimates of past as well as projected net migration.

Households and other subgroups were incorporated in eight models, which all were IBMs. Such structures are more cumbersome to implement in population-level models, thus individual-level modelling seems to be preferred if demographic subgroups are considered important for the disease transmission process, setting or research question at hand. Despite the flexibility of IBMs, it remains complex to model household structures. Detailed data on household characteristics, in particular historical data, is very limited, which can make it necessary to make strong assumptions regarding household structures. For example, several studies assigned individuals to households according to marginal distributions of household size and composition rather than individual or household-level data, which often is unavailable.

However, households did evolve dynamically over time in all models, which is important for assessing the impact of demographic change on disease transmission dynamics. Declining household sizes and changing compositions are resulting from population ageing due to decreasing fertility rates and rising life expectancy [126]. Consequently, the number of household contacts decline and the age structure in the household contacts changes [42]. This implies that boosting of immunity through household transmission becomes less likely [127]. Most studies, however, focused on the most common household types and/or positions and left out the rest completely, or gathered them in one category. In many cases this approach may be warranted given that the required level of demographic precision is, in this context, determined by its relevance for disease transmission, the setting and research question. Nevertheless, some household types may be important for the disease spread and burden, even if they represent a relatively small proportion of the population. For example, nursing homes and other special care facilities for the elderly make up a small share of the households in most countries but provide an optimal environment for the spread of

many infections [128]. Moreover, the enhanced age and underlying chronic illnesses place this population group at increased risk of many infections [129, 130]. Thus, in some contexts, less common household types can be important to consider if the data is available.

Future demographic trends are associated with a large degree of uncertainty, which is an important aspect in projecting populations [123]. Thus, an assessment of this uncertainty and its impact on epidemiological outcomes by the means of sensitivity analyses is highly relevant. Statistical agencies often provide different scenarios or prediction intervals for future fertility, mortality and migration levels, which can be used for this purpose. A smaller number of included models (10) performed sensitivity analyses by assessing the impact of variation in demographic trends on epidemiological outcomes.

The findings of this systematic review should be considered in light of several limitations. More studies may have been relevant to include but were not captured in the search due to the requirement of a reference to certain demographic terms in the title or abstract. However, the number of hits would be unfeasible to handle if the demographic search terms were omitted and if searching full-text and supplementary materials. Moreover, static and stable host populations are most common in infectious disease modelling and the incorporation of dynamic demographic processes involves a degree of complexity that we would expect most researchers to omit unless the study has a specific focus on the impact of demography or demographic change. Note that, to minimise any potential bias, we also conducted snowball sampling.

2.5 Conclusions

We systematically reviewed the literature on infectious disease modelling with a dynamic host population. We found that population-level modelling (EPMs and CCBMs) was more common than individual-based modelling. EMPs provide a straightforward and relatively simple implementation of a dynamic host population, while CCBMs are a bit more complex but make it possible to consider each demographic process separately and to test different demographic assumptions. Demographic characteristics beyond age and sex were only included in IBMs, including birth interval and parity, as well as households and other demographic subgroups. However, we found that the majority of IBMs modelled fertility, mortality and migration in a similar manner to the CCBMs, namely by the use of crude rates or age- (sex)-specific rates. We recommend avoiding the use of crude rates, if possible, as they disregard the pop-

ulation age structure and changes therein. In addition to fertility and mortality, we recommend including migration in the demographic model, since most countries face substantial migration flows and changes in population size and composition may be biased if migration is ignored. The approach used to model each demographic process implies certain assumptions, and the implications these may have for the population composition should be given careful consideration, and above all be stated clearly. Finally, the inherent uncertainty in demographic trends and their potential impact on epidemiological outcomes is ideally addressed using sensitivity analyses.

Chapter 3

Demographic microsimulation

3.1 Introduction

In a demographic microsimulation, each individual in the population is represented explicitly and assigned a set of characteristics of interest (e.g. age, sex, marital status, household membership). The initial population structure typically resembles that of a real population [54, 131]. The population evolves over time as a result of individual demographic events (e.g. death, migration, marriage). The occurrence, timing and sequence of events that an individual experiences are determined stochastically using probability models (e.g. hazard models, competing risk models), which relate the likelihood of an event at a given time to certain demographic characteristics and/or the duration since a reference event (e.g. duration since previous birth) [132]. The initial population and the estimated parameters in the demographic probability models are typically based on empirical data, such as vital statistics, surveys, census data or population registers [132].

The events that are determined to occur are executed according to algorithms of varying complexity (e.g. union formation, removal of deceased individuals) and may include some constraints to avoid highly unlikely or illegal events (e.g. incestuous relationships) [54]. The scheduling of demographic events depends on whether time is treated as a continuous or discrete variable. In the first case, the timing and sequence of events are determined, while in the latter case, it is not determined when an event takes place, but whether it takes place in a defined time interval [133, 134]. As the time interval shortens, the result of a discrete-time model resembles that of

a continuous-time model more and more [54]. Microsimulation is not to replace the traditional population-level models in demographic research, but constitutes a valuable and more appropriate tool in certain settings. If population heterogeneity is of importance, microsimulation is often to prefer, since the state space (the number of demographic variables and their possible values) can expand more easily. Moreover, life histories can be tracked in a microsimulation, making it possible to consider processes where events that have occurred in the past are determining for future outcomes. Finally, interactions between variables as well as individuals can be incorporated in a microsimulation in a flexible manner, while this remains highly complex in population-level models. Interactions between individuals is often of particular importance as many demographic events involve several individuals (e.g. marriage, divorce, leaving household) [53, 54, 133].

These features of individual-level modelling have also proven to be valuable in the field of infectious disease modelling. Like in demographic research, population heterogeneity beyond age and sex is often of importance, which can be cumbersome to implement in the traditional population-level compartmental models. Moreover, individual-level modelling allows for an explicit representation of network structures (e.g. mobility, households, sex partners) and interactions between individuals potentially leading to disease transmission. Finally, past infections or other events can be tracked in the life history of each individual if considered relevant for future susceptibility to an infection [14, 26].

Several individual-level models and software packages have been developed in the fields of demography and infectious disease modelling. We do not indent to provide an exhaustive list, but mention SOCSIM [135, 136], MIC-CORE [137] and MicSim [138] as examples of demographic microsimulation models. SOCSIM is often used in the field of family demography and kinship networks [139–143]. Microsimulation models have also been developed to generate population projections and facilitate government planning, including LifePaths from Statistics Canada [144] and MOSART from Statistics Norway [145]. Individual-based software tools have also been developed for simulating the spread of infectious diseases, including STDSIM for sexually transmitted diseases [146], the Spectrum and Epidemic Projection Package of UNAIDS [147] and 4Flu for seasonal influenza [37]. Moreover, some researchers make the source code of their IBMs publicly available, allowing others to adapt them to different infectious diseases and settings, for example [42, 60, 79, 148].

Infectious disease modelling typically involves some degree of demographic modelling to define the host population and changes therein. However, the flexibility of individual-level modelling is typically mainly exploited in the disease transmission process and less often used to advance the demographic modelling [52]. Our aim is to adopt methods and techniques from demographic microsimulation to construct an age- and household-structured population with dynamic demographic processes, that can be combined with an IBM of disease transmission in a flexible manner.

Many existing individual-level models, particularly within infectious disease modelling, are limited by the availability of detailed demographic data. However, access to individual-level longitudinal data from Belgian census and populations registers resolves this issue. The Belgian population observed on January 1st 2011 serves as the initial population in our microsimulation. The individual characteristics include age, sex, household membership, household position and/or the household type (e.g. single parent, resident in LTCF). Moreover, kinship networks and fertility trajectories are re-created in the initial population using procedures described in section 3.3.

The population is updated in discrete time steps of one year by simulating demographic events in the form of births, deaths, migration and household transitions. Models for each demographic process relate the probability of an event in a given time step to individual characteristics, including age, sex and household position. Next, we provide further detail about the model structure, the creation of the initial population, the methods used to model each demographic process and the set of assumptions and rules used to execute the demographic events in the simulation.

3.2 Model structure

The model consist of two parts, the model initialisation and the simulation (see Figure 3.1). During model initialisation, the census data and additional demographic data from the Belgian statistical office (Statbel) and the Belgian Federal Planning Bureau (FPB) are cleaned, re-structured and organised to make them suitable for creating the input files: the initial population, demographic rates and probability models. While the initial population is the starting point of the simulation, the demographic rates and probability models serve as input in the demographic processes in the simulation. Since the simulation is in discrete time steps (i.e. years), we need to predetermine the order in which the demographic events take place. We assume that births take place first and are followed by household transitions, migration and finally deaths.

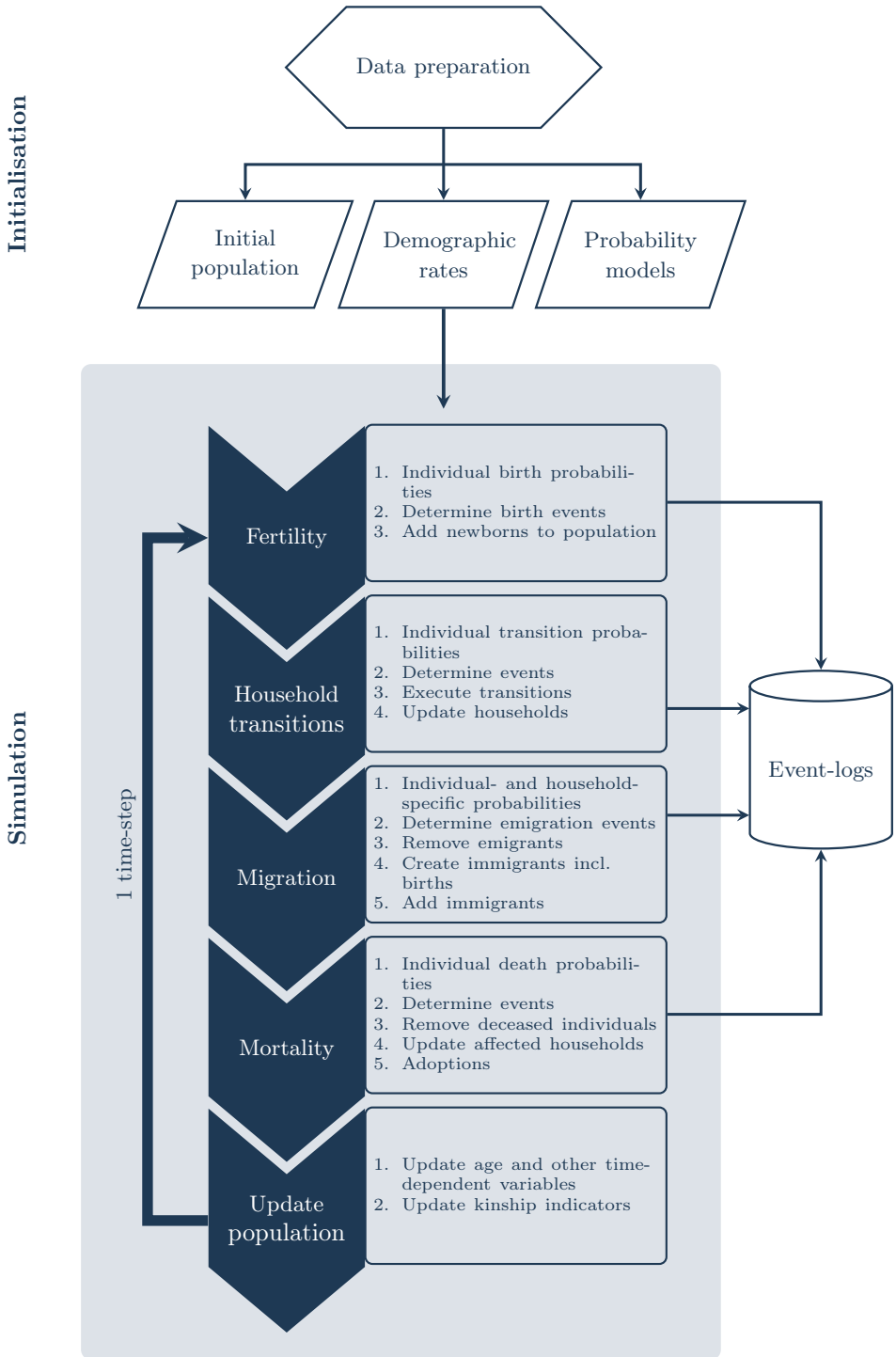

Figure 3.1. Flowchart of microsimulation.

Each event is assigned a date within the time step and we allow individuals to experience multiple events in the same year as long as the order is preserved. If an event involves multiple individuals (e.g. union formation), the individuals involved are assigned the same event date. The characteristics of the individuals experiencing an event are updated immediately. For example, the household position of an individual is updated before it is determined whether the same individual emigrates or dies within the same time step. Since the processes of emigration and mortality are conditional on household position, the given individual's probabilities of those events taking place may have changed after the household transition. After all events have been executed, kinship indicators and time-dependent variables are updated to prepare the population for the following time step. All demographic events are recorded in an event log file, including information on coded ID, event type, event date, coded household ID and household position before and after the event. This makes it possible to recreate the population in smaller time steps (i.e. days), which typically is necessary in infectious disease modelling. Finally, we use the event log file for model validation, as we compare the simulation results to aggregate measures, such as past and projected vital rates, and adjust the demographic input rates and probability models if necessary. More detailed measures for model validation are currently not available for the time period considered in the simulation.

3.3 Initial population

3.3.1 Data

The initial population is based on Belgian census and population registers on January 1st 2011. Additionally, time-dependent variables are also collected for January 1st 2012. For each individual in the population, the following variables are extracted:

- Coded ID
- Date of birth
- Sex
- Coded ID of household of residence (2011 and 2012)
- LIPRO household position (2011 and 2012)
- Coded ID of parents

We generate additional variables (household position, partner matching, kinship and birth trajectory) based on the original data and a set of assumptions.

3.3.2 Household position

All individuals in the population have a LIPRO household position for 2011 and 2012, which describes the relation an individual has to other household members and/or the type of household [149]. The LIPRO household positions only allow one family nucleus per household and all other household members are assigned LIPRO positions in relation to that family nucleus. Consequently, in a multigenerational household consisting of three related generations, the grandparents would typically have the LIPRO position married/cohabitating union, their child would be child in family with cohabiting/married parents and the partner of their child and their grandchildren would be non-family related.

Table 3.1. Categories of the variable household position.

Position	Description
Child	Individual (regardless of age) living in their parental household without own children or partner
Union	Individual living together with their partner but without child(ren)
Union+	Individual living together with their partner and child(ren)
Single	Individual living in a one-person household
NFRA	Individual living without own family nucleus but living in same household as unrelated family
Other	Individual living together with other unrelated individuals
Collective	Individual living in collective household (prison, special care facility, nursing homes,...)
Single+	Individual living together with their child but without a partner in the household
*Single+**	Single+ living in their own parental household (only used in cases where the distinction is necessary)
Union in multigenerational household (multi_U)	Individual living together with their partner, child and family of child (i.e. grandchild and/or partner of child)
Single in multigenerational household (multi_S)	Individual without partner living with their child and family of child (i.e. grandchild and/or partner of child)

In order to allow for more than one family nucleus and at the same time disregard the distinction between marriage and cohabitation, the LIPRO household positions are modified in the new variable *household position*. The categories of this variable are described in Table 3.1, while subsection B.1.1 in Appendix B contains the list of the original LIPRO household position and the specific procedures and assumptions used to create each new category.

3.3.3 Partner matching

Individuals with household position union(+) are matched if they live in the same household, since the original LIPRO household positions only contain one union per household. A smaller number of errors are detected (e.g. union between parent and child), which are changed to more appropriate categories in the variable household position (see Appendix B subsection B.1.2 for further details). Unions are also detected through the parental IDs. Individuals who are parents to the same child and live in the same household are assumed to be a couple and assigned the household position union(+). Additionally, individuals living together in 2011 and 2012, but only with household position union(+) in 2012 are assumed to already be in a union in 2011. For reasons of simplicity, same-sex unions are change to opposite-sex unions with the female being the youngest. This implies that we assume demographic events for individuals in a same-sex union to be similar to those of opposite-sex unions. It should be noted that unions formed by individuals living in different households are not detected.

3.3.4 Parent-child matching

The parental IDs are used to create the variables ID mother and ID father by looking up the sex of the respective individuals. For individuals with household position *child*, a parent ID is considered to be incorrect and set to NA if there are less than 12 years between parent and child. Moreover, the parent IDs are modified if both parents to a child younger than 16 years do not live in the same household as their child (e.g. foster care) or both IDs are missing. In those cases, the parental role is assigned to a household member who is at least 14 years older than the child (see Appendix B subsection B.1.3 for further details).

3.3.5 Birth trajectory

We use the parent IDs to re-create birth trajectories of the female population. A frequency table containing the IDs of mothers is used to compute the number of births, which is then added to the information of the given mother. Moreover, the birth date of each child is added to the information of the mother and the date of the most recent birth is used to create the variable *index birth*. There will be some discrepancies between the computed and the actual birth trajectories because some females have given birth to individuals that are no longer in the population (stillbirths, emigrants). Moreover, some parent IDs are missing.

3.3.6 Population sample

We create a representative sample of the census population to reduce computation time in the microsimulation. A sample of 500,000 households was drawn from the census conditional on the household size distribution. Collective households with more than 230 residents were excluded, as this is a technical limitation in the household network model of Krivitsky et al. [150], which is implemented in the disease modelling application. Nevertheless, households of that size are very rare. All members of the sampled households, about 10% of the census population, make up the sample population. The household size distribution and age distribution for the sample population and full census are compared in Figure 3.2. Moreover, age distributions by household size are shown in Figure B.1 in Appendix B.

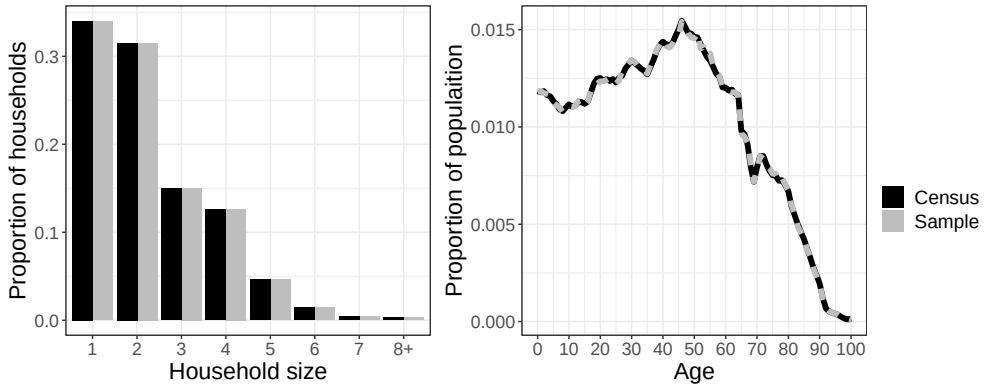

Figure 3.2. Household size distribution (left) and age distribution (right) in Belgian census 2011 and sample population.

The parent(s) of sampled individuals may not be included in the sample population if an (adult) child and its parent(s) live in different households. In those cases, the

(adult) child is assigned 'replacement' parents. Females in the sample population of the same age and household position as the actual mother in the full census make up a pool of candidate 'replacement' mothers and one is randomly drawn and assigned to the individual. The partner of the assigned mother is assigned as 'replacement' father. For individuals with only a father in the full census, the same procedure is carried out but based on the characteristics of the father. The birth trajectories of the 'replacement' parents remain unchanged.

The sample population constitutes the initial population in the microsimulation. An overview of the original and generated demographic variables in the initial population are provided in Table 3.2. These variables will be updated, if applicable, as an individual experiences demographic events in the microsimulation.

Table 3.2. Demographic variables in initial population

Variable	**Description**	**Possible values**
Individual ID		1, 2, 3, ...
Sex		Male, female
Birth date		Date
Household ID		1, 2, 3, ...
Household position	Generated from LIPRO	Child, single, single+, union, union+, NFRA, collective, multi_U, multi_S, other
ID of partner	ID of partner living in same household	1, 2, 3, ..., NA
In union	ID of partner is applicable	0=No, 1=Yes
ID of mother		1, 2, 3, ..., NA
ID of father		1, 2, 3, ..., NA
Parent indicator	Parent(s) present in population and not living in collective household	0=No, 1=Yes
Child-family indicator	Child with own family present in population	0=No, 1=Yes, NA
Birth parity	Number of previous births	0, 1, 2, ...
Index birth	Most recent birth	Date, NA
Time since index birth	Days	0, 1, 2, 3, ..., NA

3.4 Fertility

3.4.1 Modelling

We use the birth trajectories of the 14- to 50-year-old female population up to December 31st 2011 (described in subsection 3.3.5) to model fertility in 2011. We calculate the age at the index birth (most recent birth before January 1st 2011) and the duration since the index birth, if applicable. For females giving birth in 2011, the duration between that birth and the index birth is used. In case a female had two births in 2011 resulting from subsequent pregnancies, the first birth in 2011 is considered the index birth. The household positions of the females on January 1st 2011 are also included in the fertility data.

Generalised additive models (GAMs) [151, 152], are fitted to the fertility data. In a GAM, like in a generalised linear model, a link function describes the relationship between the linear predictor and the expected value of the response, but in a GAM, the linear predictor involves a sum of smooth functions of covariates. The model structure is generally described by Wood (2017, eq. 4.1) [153] as

$$g(\mu_i) = A_i\theta + f_1(x_{1i}) + f_2(x_{2i}) + f_3(x_{3i}, x_{4i}) + \cdots \tag{3.1}$$

where $g(.)$ is the link function, $\mu_i \equiv E(Y_i)$ and the response variable Y_i follows an exponential family distribution with mean μ_i and scale parameter ϕ. A_i and θ represent a row of the model matrix for strictly parametric model components and the corresponding parameter vector, respectively. Finally, f_j are smooth functions of the covariates x_k. A smooth function is the sum of a number of basis functions, $b_j(x)$, weighted by the corresponding regression coefficient, β_j (Wood, 2017, eq. 4.3) [153]:

$$f(x) = \sum_{j=1}^{k} b_j(x)\beta_j \tag{3.2}$$

We use penalised cubic regression splines with shrinkage (specified as bs="cs" in R package mgcv) and vary the number and location of knots, which defines the number and interval of the basis functions, in order to obtain an optimal fit. The GAMs are fitted separately for first, second, third and higher order births. The dependent variable is binary (1=birth, 0=no birth) and we use a logit link function. For females

at risk of a first birth in 2011, the following model was fitted:

$$\log\left(\frac{\mathrm{P}(1st\ birth)_i}{1-\mathrm{P}(1st\ birth)_i}\right) = \alpha + f_{h(i)}(age_i) + \epsilon_i \tag{3.3}$$

The term f represents the smooth function, the variable age_i is the age in the middle of the year of individual i and $h(i)$ denotes its household position on January 1st. Thus, separate smooth functions are fitted for each household position. In Figure 3.3, the predicted and estimated probabilities of a first birth are shown by household position. Overall, the predicted and estimated probabilities are very similar except at the very young ages where births are uncommon.

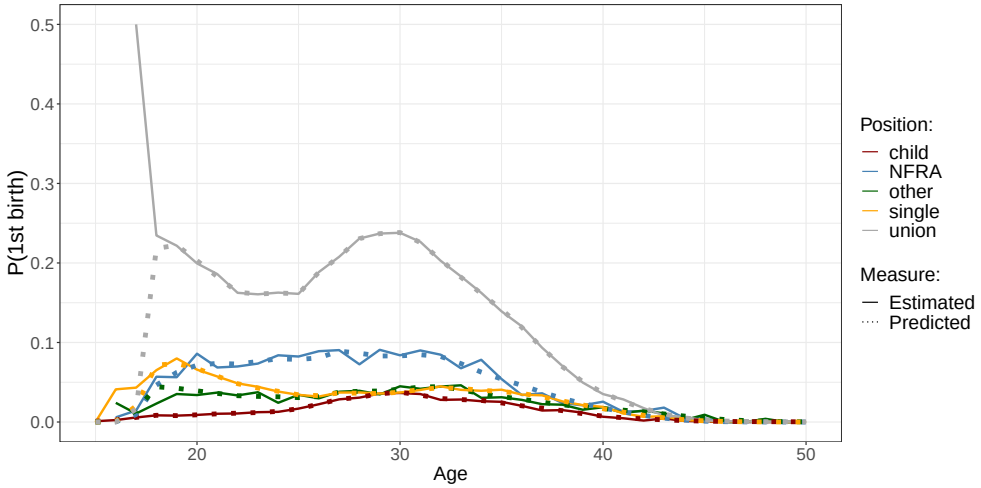

Figure 3.3. Estimated and predicted probability of 1st birth in 2011 by household position. Data source: Belgian census 2011 and population registers.

For second and higher order births, the following model was fitted separately for each combination of parity and household position group (union vs. non-union):

$$log\left(\frac{\mathrm{P}(j_{th}\ birth)_i}{1-\mathrm{P}(j_{th}\ birth)_i}\right) = \alpha_1 + \alpha_2 \cdot parity_i \cdot z_i + f_{12}(age_i^{index}, duration_i^{index}) + \epsilon_i \tag{3.4}$$

where $j = parity_i + 1$ with $parity_i$ denoting the birth parity of individual i and z_i is a dummy variable which is 1 for parity three and higher and zero for all other parities. The covariate age_i^{index} is the age of the female at the index birth and $duration_i^{index}$ is the duration since index birth. These two variables are fitted with a tensor product

smooth, f_{12}, which can be formulated as follows [153]:

$$f_{12}(x_1, x_2) = \sum_{(i=1)}^{I} \sum_{(j=1)}^{J} \delta_{ij} b_{1i}(x_1) b_{2j}(x_2) \tag{3.5}$$

where b_1 and b_2 are basis functions, I and J are the corresponding basis dimensions and δ is a vector of coefficients. The predicted and estimated birth probabilities are shown in Figure 3.4 and Figure B.2 in Appendix B for different ages at index birth.

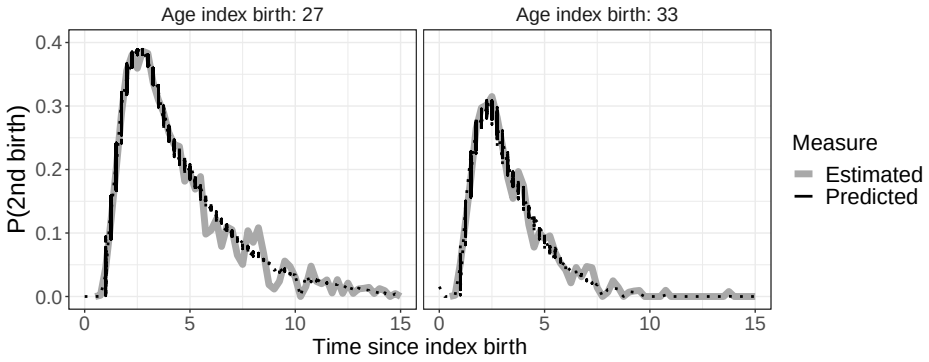

Figure 3.4. Estimated and predicted probability of 2nd birth in 2011 for females in a union and by age at index birth (columns) and time since index birth (years). Data source: Belgian census 2011 and population registers.

3.4.2 Simulation

In the microsimulation, females of age 14-50 years with the household positions *union, union+, single, single+, child, NFRA* or *other* are considered eligible for giving birth. This means that females in collective households and the oldest generation in multi-generational households are excluded. Furthermore, females with a duration since index birth of less than 12 months (pregnancy of nine months and period of lactational amenorrhoea of three months) are also excluded from giving birth in the given time step.

The eligible females are assigned a probability of giving birth in the given time step by applying the fertility model corresponding to their parity and household position and converting the log odds to probabilities. The fertility models are independent of time, however, we have incorporated the option to adjust the resulting birth probabilities by a time-specific factor. Specifically, we chose a vector of time-specific factors which

result in a TFR that resembles the observed and projected rates by Statbel and FPB. If a female's probability of giving birth is higher than or equal to a randomly drawn number between zero and one, she is assigned a newborn child and her birth trajectory is updated accordingly. A number of females are assumed to give birth to twins corresponding to the estimated proportion in 2011. The birth events are assigned a random date within the time step. The newborns are added to the population and the household of the mother. If the mother has a partner, the partner is assigned as the father of the child.

3.5 Household transitions

3.5.1 Transition probabilities

The household transition probabilities are, in contrast to the other demographic processes, assumed to remain constant over time. This is a rather unlikely assumption, but detailed data covering a longer time period is lacking. We thus calculate household transition probabilities based on the individual household positions observed on January 1st 2011 and January 1st 2012. Some individuals are not present in the population in 2012 and the household position is for that reason missing (NA). We also set indirect transitions resulting from other demographic events to NA because those transitions are incorporated in the other demographic processes. This includes changes in the household position *child* due to the death or emigration of the parents, given that the child is 16 years or older. Union dissolutions resulting from the death of the partner are also considered to be indirect transitions, as well as transitions from *single+* and *union+* resulting from the last child leaving the parental household. The individuals with missing household positions are assumed to have left the population and/or risk-set in the middle of 2011.

We calculate the transition probabilities conditional on a set of demographic characteristics: age group, sex, birth event in 2011 (yes/no), parent indicator, child-family indicator (described in Table 3.2). Z denotes the combination of characteristics, for example, females in age group 20-21 with no birth in 2011, with a parent indicator of 1 and child-family indicator of 0. The parent and child-family indicators are included because some transitions are conditional on kinship networks between households. For example, an individual cannot move back to the parental household if the parents are no longer in the population. We compute two types of transition probabilities; the probability that an individual with demographic characteristics Z and household

position j experience any household transition (overall probability), and the probability that the individual moves from household position j to k given that a transition takes place (directed probability). The overall transition probabilities for household position j are calculated as follows:

$$T_j^K(Z) = \frac{N_j^K(Z)}{N_j(Z) - \left(\frac{N_j^m(Z)}{2}\right)} \tag{3.6}$$

where subscripts and superscripts indicate the household position in 2011 and 2012, respectively. $j \notin K$, thus the set K contains all household positions different from j, which are considered eligible transitions (described in detail in subsection 3.5.2). The risk-set, i.e. the number of individuals with household position j in 2011, is denoted by N_j, while N_j^K is the number of individuals from the risk set with household positions different from j in 2012. Finally, N_j^m is the number of individuals from the risk set with a missing household position in 2012 (left the population or indirect transition). The directed household transition probabilities are calculated by limiting the risk-set to those with a transition, N_j^K, and specifying the household position in 2012 as $k \in K$, hence $k \neq j$:

$$T_j^k(Z) = \frac{N_j^k(Z)}{N_j^K(Z)} \tag{3.7}$$

Examples of estimated household transitions probabilities are shown in Figure B.3 and Figure B.4 in Appendix B.

3.5.2 Transitions

In the microsimulation, the occurrence of household transitions are determined in a step-wise procedure. First, each individual is assigned the overall probability of a transition conditional on their set of demographic characteristics, $T_j^K(Z)$. Second, a subset is created of individuals determined to experience a transition. Third, each individual in the subset draws a new household position from a vector containing the directed transition probabilities, $T_j^k(Z)$, for all eligible values of k given j and Z. Some transitions are not considered eligible because they are rare or for the sake of simplicity. Individuals younger than 16 years of age are not eligible for any transitions, but may move to a new household as a result of household transitions of their parents. Moreover, only a smaller set of transitions are eligible for parents to

a child born within the given year. The white and grey cells in Table 3.3 indicate eligible transitions for individuals with and without a birth event in the given year, respectively. Black cells are non-applicable transitions for everyone. Cells with I represent indirect transitions resulting from other demographic events (e.g. child leaving parental household, oldest generation in multi-generational household moves out). For the individuals with an event, the transition from one household position, and typically also household, to another is executed using different procedures. Each procedure involves a set of assumptions and rules, which are described next, and may involve individuals that were not assigned a transition in the step-wise procedure (i.e. indirect transitions).

Union formation

Household transitions to *union(+)* are disregarded if the individual is male. Instead, the process of union formation starts when a female experiences a transition to *union(+)* and the males are indirectly assigned as partners. For each female entering a union, we search for a match in a pool of male candidates. A male is only considered a candidate if he is not already in a union and is 16 years of age or older. The age of the future partner of a given female is drawn from an age distribution based on the union formations observed in 2011. The age distribution of the male partners was computed by the age group of the female partners in 2011 (see Figure B.5 in Appendix B). The pool of male candidates is then narrowed down to those with an age within the drawn age group. Each remaining candidate is assigned a match probability corresponding to their probability of a household transition to a union, $T_j^k(Z)$ where k is *union(+)*, given their demographic characteristics. A partner is then assigned to the female by drawing from the pool of male candidates conditional on the match probabilities. The newly formed union is assigned to a new household as well as children previously living with any of the partners.

Union dissolution

Union dissolution takes place if a female with household position *union(+)* is assigned a transition to *child*, *single(+)*, *other*, *NFRA* or *collective* and if a female with household position *multi_U* is assigned a transition to *single* or *collective*. Union dissolution initiated by a household transition of the male partner is disregarded, like for union formation. However, when a female leaves the union and thereby the household according to our assumptions, a new household position is also drawn for the male partner based on the corresponding household transition rates. In case of a dissolu-

Table 3.3. Transitions for household positions. White: Applicable for individuals without a newborn. Grey: Applicable for individuals with a newborn. Black: Non-applicable/no transition. I: indirect transition/other demographic events)

From \ **To**	*Child*	*Union*	*Union+*	*Single*	*Single+*	*Single+**	*Other*	*NFRA*	*Collec.*	*Multi_ U*	*Multi_ S*
Child											
Union			I								
Union+		I				I					
Single											
Single+				I		I					
*Single+**	I				I						
Other											
NFRA											
Collec.											
Multi_ U			I								
Multi_ S					I						

tion of *union+*, the children are assigned a main parent, which they follow to their new household. The probability of becoming the main parent is assumed to be 0.25 and 0.75 for the father and mother, respectively. The eligible household transitions from the position *union+* are conditional on whether the individual is assigned the role of main parent. If the main parent is moving to their own parental household, the new household position is *single+** and not *child*. Moreover, a transition to *single*, *other*, *NFRA* or *collective* is changed to *single+* if the individual is assigned the role of main parent. Finally, for reasons of simplicity it is not possible to move directly from one union to another. A union dissolution is also triggered by a transition from *union* to *collective*, because we only consider unions between individuals living in the same household. However, in most cases the relation probably continues.

Transition to parental household

Individuals can move back to the household of their parent(s) if one or both are present in the population and the parent is not living in a collective household (parent indicator of 1, see Table 3.2). The household position changes to *child*, when moving back to the parental household, also for adult individuals. This transition is not eligible for the oldest generation in multi-generational households for whom a parental household rarely is present in the population. Single parents and individuals in a union cannot move into their parental household, unless it is in combination with a union dissolution. However, the parents of an individual with household position *union(+)* or *single+* can move into their child's household.

Transition to multi-generational household

Multi-generational households can be created by two different types of transitions. As already mentioned, individuals can move from a union and back to their parental household together with the children, for whom they are the main parent, but without the former partner. In that case the individual's household position changes to *single+**. Moreover, the household position of the oldest generation in the household (the grandparents) indirectly changes to $multi\_U$ or $multi\_S$, depending on whether they are in a union or single.

Individuals can also experience a direct transition to the position $multi\_U/multi\_S$, if they have a child with household position *union(+)* or *single+* (i.e. child-family indicator of 1, see Table 3.2). In that case, the older generation moves into the household of their child's family. This type of transition does not affect the household position of the younger generation, except for single parents, which transition from

single+ to *single+**. The oldest generation in multi-generational households follow their child in case of future household transitions (e.g. union formation or dissolution), but can also leave the household and change household position to *union* or *single* depending on whether it involves a union dissolution.

Transition to household positions *NFRA* and *other*

Transitions to *other* and *NFRA* are only eligible if the individual is not a single parent or assigned the role of main parent after union dissolution. Moreover, the oldest generation in multi-generational households are also not eligible for such household transitions. For individuals with a transition to *NFRA*, a target household size is drawn from the household size distribution of non-family related adults observed in 2011. Candidate households are found by identifying individuals with household position *single+* and *union(+)*, which are considered to be in a family nucleus. A household ID is drawn from the pool of candidate households with the corresponding target size. The individual is assigned the household position *NFRA* and the new household ID. The household positions of the new household members are not affected by the transition.

In a similar manner, a target household size is drawn for each individual with a transition to *other* conditional on the observed household size distribution of *other* households in 2011. The group of individuals assigned to a target size of two are matched in pairs based on their age and assigned a new household ID. For an individual with a target household size larger than two, an existing *other* household matching the target size of the individual is drawn at random. The household ID of the individual is updated accordingly and the household position changes to *other*. The household positions of the new household members are not affected by the transition.

Transition to collective household

All collective households are assigned a maximum capacity. For households with less than 10 residents, the maximum capacity is set to the initial household size, while households with 10 or more residents are assigned an extra capacity corresponding to 15% of the initial size. The median age of the household members in each collective household is computed in each time step. Transitions to a collective household are only eligible if the individual is not a single parent or assigned the role of main parent after union dissolution. For individuals with an eligible transition to collective household, a target household size is drawn from the age-specific household size distribution for collective households observed in 2011. The individuals are assigned to existing

collective households with a median age of ±15 years of their own age and with a size corresponding to their target household size. When the household capacity is met, no more individuals are assigned. In case some individuals have not been assigned a household, new collective households are created with a maximum capacity corresponding to their target household size. Age is taken into account in the allocation of people in collective households in an attempt to distinguish LTCFs mainly occupied by elderly people from prisons and other institutions mainly containing younger people.

Transition to *single(+)*

All individuals aged 16 years are eligible for a transition to single (parent). The final household position depends on whether the individual has the role as main parent. In both cases, a new household is created.

Birth-related household transitions

The individuals with a newborn are eligible for a birth-related household transition in the same year conditional on their household position before the birth (see grey cells in Table 3.3). Some transitions are not considered eligible in the year of the birth because they are rare or for simplicity (e.g. adoption immediately after birth, union dissolution). Females with household position *child* can stay in their parental household after giving birth, but the household position changes to *single+** and the household becomes a multi-generational household. However, birth events among females with household position *child* can also involve a move to a new household either as *single+* or *union+*.

The transition to *union+* involves a partner search based on the matching process described earlier. Females with position *single(+)* either stay in their household and keep the household position after giving birth or move to *union+* with a matched partner. Births to females with household position *NFRA* or *other* always involve a move to a new household and a change to household position *single+* or *union+*. Females in a union and with a birth event stay in the union and change to household position *union+*. The household transitions are assigned the same event date as the birth. Moreover, household IDs and household positions are also updated for the individuals indirectly affected by birth-related household transitions (e.g. new partner, children, grandparents in multi-generational households).

Indirect transitions

Some of the described procedures explicitly consider the individuals indirectly affected by a specific household transition (e.g. males in the union formation procedure, children moving with their main parent after union dissolution), and their household IDs and positions are updated accordingly. Other indirect transitions are captured in a check-up procedure. For example, household positions *union+* and *single+* are changed if the last child has left the parental household.

3.6 Migration

Observed (2011-2020) and projected (2021-2050) immigration and emigration rates are retrieved from Statbel [154]. Distributions of age, sex and household positions of immigrants as well as emigrants are computed based on the observed migration in 2011 (see Figure 3.5).

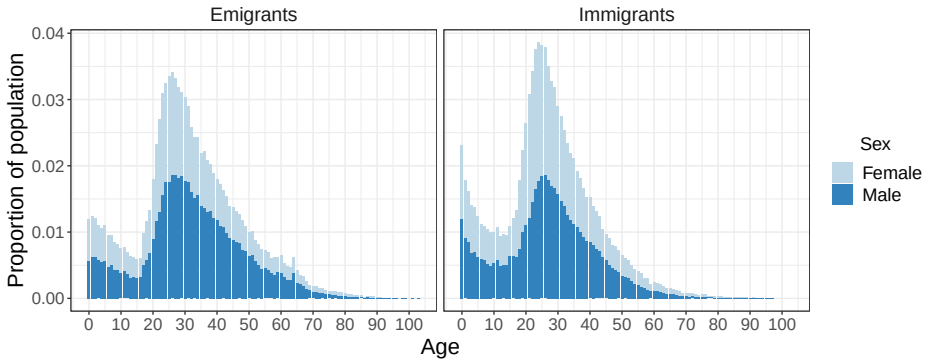

Figure 3.5. Age-sex distribution of emigrants (left) leaving Belgium in 2011 and immigrants (right) entering Belgium in 2011.

3.6.1 Immigration

The number of immigrants in a given time step is computed according to the observed or projected immigration rate for the given simulation year. Age, sex and household position of the immigrants are drawn from the distribution observed in 2011. Females assigned household position *union(+)* are matched with male immigrants using the same union formation procedure as for the native population described earlier. Family reunification between natives and immigrants is thus not modelled explicitly. Female immigrants with household position *union(+)* and *single+* are assigned as mothers to

children based on an allocation procedure. For a given child, candidate mothers are identified by computing the age difference between the female and the child. Females remain in the pool of candidates if the age difference is between 14 and 45 years. The female candidates are assigned a probability of being the mother of the given child, which is based on the observed distribution in the age difference between mothers and children immigrating in 2011 by the age of the child (age group 0-9, 10-19, 20+). A female is then drawn from the pool of candidates conditional on her age and that of the given child and assigned as the mother.

In case the pool of candidates is empty, females who already have been assigned as mothers to other children are added to the pool again and can in that way be mothers to multiple children. In case the pool of candidates remains empty, the upper limit for the age difference between mother and child is removed. The children are added to the household of the mother and her partner, if any, which is assigned as the father of the child. Some immigrants with household position *child* are assigned native parents according to the age-specific proportions observed in the immigrant population in 2011 using the same allocation procedure as described above.

The remaining immigrants without a household position, are assigned the positions *single*, *NFRA* and *other* according to the distribution of household positions observed for immigrants in 2011. Immigrants with household position *NFRA* are assigned to a household in the native population using the same procedure as for the native population. All immigrants are assigned a date of entering the population which is the same for all members of a given household. Immigrants are not at risk of a household transition in the time step of entry into the population, but females of fertile age are at risk of a birth.

3.6.2 Emigration

The number of emigrants in a given time step is computed according to the observed or projected emigration rate for the given simulation year. Emigrants are split into two groups, individuals emigrating alone and whole households emigrating. Individuals are eligible for emigrating alone if they are 16 years or older and have the household position *child*, *single*, *other* or *NFRA*. Eligible individuals are chosen for emigration conditional on the household position and age-sex distribution observed for emigrants in 2011 (see Figure 3.5). Whole households are eligible for emigration with the exception of collective households. Eligible households are assigned an age-composition-specific probability of emigrating, which is based on the observed com-

position of households that emigrated in 2011. Conditional on the household-specific probabilities, households are sampled until the estimated number of emigrants in the given time step has been met. Emigrants are assigned an event date within the time step for leaving the population, which can only take place after any other demographic events of the given individual. Household members emigrating together are assigned the same event date. Emigrants are removed from the population and cannot re-enter the population.

3.7 Mortality

By breaking the central rate of mortality (m_x) calculated for 2011 further down by the household positions *union(+)*, *non-union* (*single/child/NFRA/other*) and *collective*, rate ratios can be computed by dividing the age-, sex- and household-specific mortality rates by the overall mortality rate for the given age and sex. The rate ratios are computed for the age groups 0-49, 50-59, 60-69, 70-74, 75-79, 80-84, 85-90 and 91+ (see Figure 3.6).

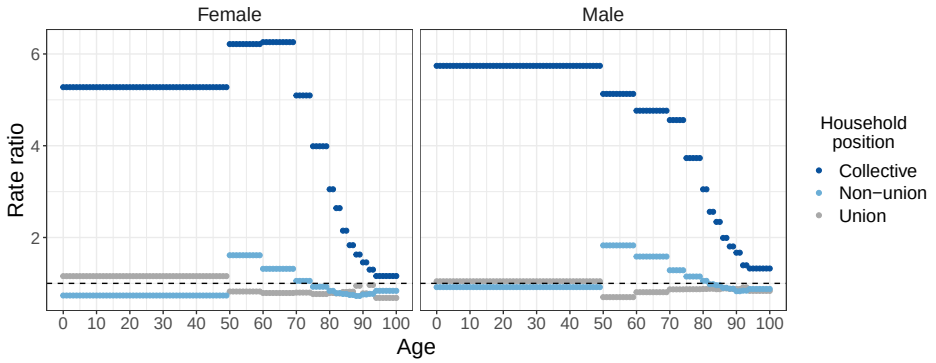

Figure 3.6. Mortality rate ratios by sex, household position and age group. Data source: Belgian census 2011 and population registers.

In each time step, all individuals have a probability of dying based on age, sex- and household-position-specific mortality probabilities (qx). These probabilities are computed by first converting the observed or projected age-sex-specific mortality probabilities from Statbel for the given simulation year to mortality rates, which are then multiplied by the household-specific rate ratios. Finally, the mortality rates are converted back to mortality probabilities and assigned to the individuals according to age, sex and household position. The mortality probability of each individual is compared to a random number between one and zero to determine whether they die within the

time step. In the case of death, a date is assigned, which will always be after any other demographic events of the given individual. The dead individuals are removed from the population. The partner of a dead individual is at risk of a household transition. Moreover, children of a dead individual and with household position *child* follow the other parent, if applicable. In case both parents are dead and the child is younger than 16 years, it is adopted by another family.

3.8 Event log file

All demographic events in the microsimulation are recorded in an event log file (see example in Table 3.4). This includes the following variables: ID, household ID (initial and target), household position (initial and target), type of event and date of event. In this way, the evolution of the population can easily be re-created and used for other applications, e.g. as input in models for infectious disease transmission. Moreover, the population can be re-created in smaller time steps, for example days or weeks, which often is required when modelling the spread of an infectious disease.

Table 3.4. Example of recordings in event log file. HH: Household.

ID	Sex	Birth date	Event type	Event date	HH ID		HH position	
					Initial	New	Initial	New
1234	F	May 1st 1989	Household transition	Aug 15th 2013	13	972	Child	Union
7856	M	Nov 20th 1987	Household transition	Aug 15th 2013	667	972	Single	Union
8764	F	Apr 10th 2016	Birth	Apr 10th 2016	NA	972	NA	Child

Event logs for fictive individuals are shown in Table 3.4, where a female with ID 1234 leaves the parental household on August 15th 2013 to form a union with a male with ID 7856 who was living in a single-person household prior to the union formation. On April 10th 2016, the female gives birth to a child that is assigned ID 8764. This part of the life history of the female with ID 1234 is also visualised in Figure 3.7. The changes in the characteristics of the female are highlighted and the arrows indicate time in the form of the date (black) and her age (grey).

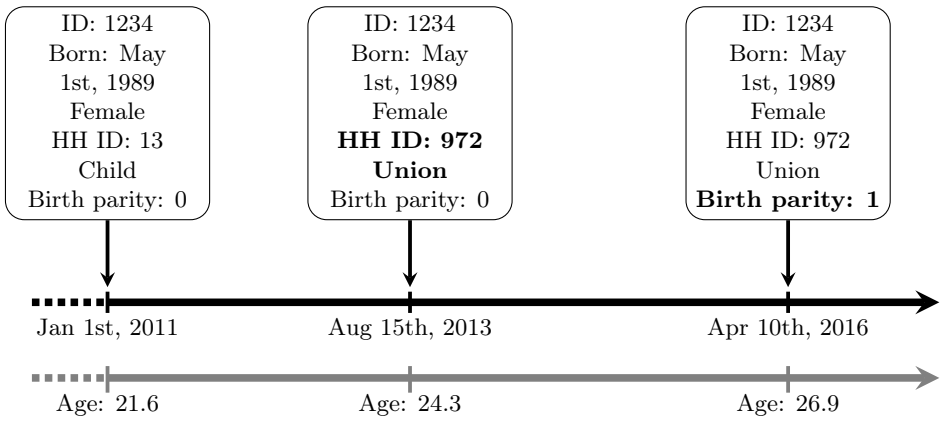

Figure 3.7. Example of individual life history along time (black arrow) and age (grey arrow). Individual characteristics in text box with events in bold. HH ID: Household ID.

3.9 Limitations

The microsimulation model faces several limitations with regard to event scheduling, modelling of the demographic processes and the execution of demographic events. We simulate the population in discrete time steps and assumptions regarding the ordering and timing of events are imposed exogenously. In reality, however, the events in the life course of an individual occur in continuous time and may be interdependent [132]. Nevertheless, we assign a date to each event and thereby make it possible to allow for multiple events within one time step. Moreover, some of the demographic processes are conditional on past events and the duration since reference events. A female's probability of a birth, for example, is conditional on the duration since her previous birth (if applicable), and her probability of a household transition in the same time step is conditional on whether the birth takes place or not. Nonetheless, we are not taking interdependence in event sequences into account. A continuous-time framework would resolve some of these limitations, but requires highly detailed data. Moreover, it remains problematic to estimate interdependent processes and the implementation of interdependence of transitions for several individuals (e.g. household members) is still complex in continuous-time models [155].

Additional limitations of the microsimulation pertain to household positions and transitions. We do not distinguish between married and cohabitating couples, although their demographic processes (i.e. births, union dissolution) tend to differ [156, 157]. Another simplification pertains to same-sex unions, which we assume to have the

same demographic processes as opposite-sex unions, although it is unlikely to hold [158]. Moreover, unions between non-household members are not considered. Neither is the possibility of being a member of several households (e.g. children of divorced parents). However, we do distinguish between unions with and without children in the household, as their probabilities of union dissolution vary considerably.

We made these assumptions while keeping in mind that the simulated population will serve as an input in transmission models of close-contact infectious diseases, where household transmission is substantial. Consequently, we considered it more important to specify transitions for families with children than to distinguish marriage from cohabitation. However, the model framework and structure provide opportunities to incorporate more detail. A higher level of detail may be worthwhile to consider in the household transition process, if permitted by the data. While the household transitions are conditional on a range of characteristics, the time spent in the household position is not taken into account. Consequently, the timing and number of transitions in an individual's life course may not reflect reality. Moreover, the transition probabilities remain constant over time. These aspects could be incorporated using longitudinal data on household position and/or household membership covering an extended period of time.

In microsimulation, different types of matching procedures can be used for union formations. We have developed a so-called closed mating model with a female-dominant matching algorithm [131]. This matching procedure ensures that every (female) individual finds a partner, while avoiding too many matches with unlikely age differences. Since we apply a procedure where the female's as well as the male's probability of entering a union is conditional on the female's age, among other variables, the age pattern of union formation for males deviates to some degree from the estimated transition probabilities. This issue could be resolved by determining which males experience a transition to union in the same manner as for females, i.e. using household transition probabilities only. A sequential matching procedure based on a compatibility measure (e.g. age difference) would then follow. However, this procedure would be more computationally intensive (see chapter 10 in Zinn [131] for further details).

Other possible improvements of the union formation and dissolution procedures relate to migration. We assume that immigrants only form unions among each other in the year of arrival. A substantial share of immigration into Belgium and other European countries, however, is related to family formation and reunification [159].

Moreover, we do not consider union dissolution as a result of emigration nor do we take into account that emigration tends to be more likely among recently arrived immigrants [160]. We are thus not explicitly incorporating the mechanisms driving migration, which are highly complex. Migration background and migration history, however, could be incorporated in the microsimulation using the current data and model structure.

Since the process of fertility incorporates the females' birth trajectories, the timing of births in the simulation reflects the observed pattern in 2011 (the year of the fertility data) rather well. In the fertility simulation, we allow for a specification of a time-dependent factor by which all birth probabilities are adjusted in the corresponding simulation year. This implies equal relative changes in birth probabilities over time regardless of age, household position and birth trajectory, which does not reflect reality. In Belgium, for example, the postponement of childbearing to older ages has continued since 2011 [161], which is not captured when applying an aggregate time-dependent factor. Official projections, however, are typically only made for the total fertility rate or age-specific fertility rates, limiting the possible granularity of a time-dependent factor in the fertility model. A preferable alternative is to expand the fertility process by incorporating the factors driving the changes in fertility. This includes education level and migration background, by which fertility patterns vary considerably [162–164]. Fertility postponement observed in Belgium, for example, has to a large degree been explained by rising educational enrolment [162].

Similarly, the mortality process in the microsimulation could be improved by including a variable describing the health of each individual (e.g. chronic conditions, frailty) instead of the current use of household position as a proxy. This would also make it possible to model household transitions conditionally on health state, with particular relevance for transitions to collective households (i.e. LTCFs). Moreover, an indication of health state or presence of non-communicable diseases would be valuable in the context of infectious disease modelling, as these factors typically influence the morbidity and mortality associated with an infection. [165, 166].

In conclusion, the microsimulation model provides a flexibility that allows us to incorporate more detail in the demographic processes and to explicitly model the mechanisms driving demographic change. In many cases, however, this would require access to additional data with high granularity and/or additional, potentially time-consuming, computations. The decision to expand the model thus depends on the feasibility and gain thereof considering the study at hand.

Chapter 4

Population age and household structures shape transmission dynamics of emerging infectious diseases: a longitudinal microsimulation approach

This chapter is based on submitted work: "Møgelmose, S., Vijnck, L., Neven, F., Neels, K., Beutels, P. & Hens, N. Population age and household structures shape transmission dynamics of emerging infectious diseases: a longitudinal microsimulation approach. medRxiv 2023.06.05.23290874 [Preprint]."

4.1 Background

Host population demographics and patterns of host-to-host interactions are important drivers of heterogeneity in infectious disease transmission. To describe the dynamics of infections transmitted via close contact interactions, particular attention has been given to social mixing patterns, which can be captured by demographic structures. The frequency and pattern of social mixing with relevance for the spread of close-contact infectious diseases are highly dependent on age. Children, teenagers and young adults have more contacts and are disproportionately more likely to mix with people of their own age than with adults older than 25 years. Adults also display age-assortative mixing, but their average contact frequency is lower and their contacts are less concentrated in their own age group than those of youngsters [30, 167].

Consequently, age patterns are seen in susceptibility and exposure to many pathogens [20]. Additionally, changes in the immune system throughout the life course can add to these age-specific differences. Children tend to be more susceptible to infections given that their immune system is still maturing, while in older adults, the ageing of the immune system (immunosenescence) may increase their susceptibility to infection and to more severe disease upon infection [76]. Likewise, the infectiousness of infected individuals may also vary by age [168, 169].

Population structures beyond age further add to the heterogeneity in social mixing patterns and in disease transmission dynamics. Due to the high frequency, long duration and intimacy of within-household contacts, household transmission constitutes a substantial risk factor in infectious disease dynamics [31, 48]. Moreover, households often contain people from different generations (e.g. parents and children) belonging to different subpopulations outside the household, which, for example, can facilitate the spread of an infection from schools to workplaces [50]. Consequently, age- and household-structured models of infectious disease transmission with social mixing have proven highly valuable for modelling the transmission of close-contact infectious diseases [32, 41, 42, 45, 170].

Still, it remains challenging to model an age- and household-structured population, and in particular the changes therein, in a well-founded and feasible manner. Detailed household data is usually unavailable, which often makes it necessary to recreate households by relying on probability matching and/or to make simplifying assumptions, like limiting to specific household sizes or types (e.g. nuclear families). Less common living arrangements such as long-term care facilities (LTCFs) or multigenerational households are often disregarded, although they may be important for disease transmission [52]. Moreover, only few studies incorporate evolving age and household structures in the host population or consider multiple populations with different compositions, for example [41, 47, 79, 81, 171].

While demographic change can be reasonable to disregard when the period under consideration is short, it may be necessary to allow for changing population structures when investigating disease transmission dynamics and control strategies in different populations or over a longer time frame (i.e. years or decades depending on the population, infection and research question), where demographic changes become more apparent. Moreover, a thorough understanding of the relationship between disease transmission dynamics and host population structures, as well as the demographic processes underlying these structures, may be important for assessing how future

demographic changes potentially could affect transmission dynamics.

Demographic change has in many countries led to an increasing median age of the population (population ageing), which has become a global phenomenon [172]. Many of the most developed countries have been ageing for decades as a result of declining fertility rates and increasing life expectancy, and the ageing of the large generations born in the mid-twentieth century is currently causing a temporary acceleration of population ageing [64].

We investigate how emerging infections are spreading in a relatively old (i.e. high median age), and still ageing, host population. We use longitudinal microdata drawn from Belgian census and population registers, including individual-level information on age, sex, household membership and kinship links. The data feeds into a demographic microsimulation, which includes dynamic demographic processes for fertility, mortality, migration and household transitions, making it possible to model the Belgian population over time with evolving age and household structures.

We subsequently combine the demographic microsimulation with a two-level mixing model, which distinguishes between exposure to infection in the household and exposure in the community at large. We base contact networks within households on empirical data, rather than making the common assumption of random mixing. We simulate the spread of an emerging close-contact infectious disease in 2020, 2030, 2040 and 2050, which allows sufficient time for noticeable changes in age and household structures to emerge. Furthermore, we vary the age profiles of infectiousness and susceptibility to reflect specific infections, including influenza and SARS-CoV-2.

We aim to explore how the relationship between age and household structures affects disease transmission dynamics of an epidemic at the individual and the population level. Moreover, we show how demographic processes alter the population structures in an ageing population and investigate how this affects the transmission dynamics across population groups. The paper is organised as follows: In section 4.2, we give specific details on the demographic microsimulation, including the demographic data and processes considered. Similarly, we describe the disease transmission model with two levels of mixing along with the model parameters. Section 4.3 presents the population structures and changes therein and documents the disease incidence by age and household composition. Furthermore, the impact of epidemiological heterogeneities within the population is visualized in a scenario analysis. Finally, in section 4.4, we discuss the results as well as the strengths and limitations of the study.

4.2 Methods

We model the host population and the spread of an infection at the individual level using microsimulation, also referred to as individual-based modelling or agent-based modelling [26]. Each individual in the population is represented explicitly and assigned a set of relevant attributes (e.g. age, sex, marital status, household membership, disease state). The population evolves over time as a result of individual events and events emerging from interactions between individuals (e.g. marriage, death, social contact, disease transmission). All events are tied to the individual, meaning that the life course and health trajectory of each person is tracked [14]. Consequently, outbreaks at the population level emerge from the interactions between the individuals. Next, we describe the initial population and the processes used to determine the occurrence of demographic events and disease transmission events.

4.2.1 Demographic microsimulation

We developed a demographic microsimulation to simulate the Belgian population from 2011 to 2051. The initial population in the microsimulation is based on the Belgian census from January 1st 2011, from which we drew a household-based sample corresponding to about 10% of the total population. For each individual, we have information on their date of birth, sex, coded ID of parents, birth trajectory (parity and date of most recent birth if applicable), household ID and household position (e.g. in union, child, single parent). Thus, individuals can be linked to each other through household membership and kinship.

In each time step (i.e. day), individuals can enter and leave the population as a result of births, deaths and migration. Moreover, the household position of an individual may change and transitions between households or the creation of a new household can take place. Household transitions include children leaving the parental household, union formation or dissolution and older adults moving to LTCFs. The household transition and other demographic events of one individual may thus also affect the household position of other individuals. A single parent, for example, changes household position to single after the last child moves out and an individual in a union becomes single (parent) after their partner dies.

The demographic events take place by comparing an individual's probability of a given event to a random number between zero and one. The event is executed if the probability is larger than or equal to the random number. The probability of a demographic event taking place varies by individual characteristics, including age, sex

and household position, and changes over time except for the household transition rates. Finally, ageing takes place at the end of the time step and the population is updated accordingly.

We assume that mortality, fertility and migration levels in the microsimulation resemble the observed and projected rates by the Belgian Statistical Office (Statbel) and the Belgian Federal Planning Bureau (FPB). This implies below-replacement level fertility (i.e. a total fertility rate (TFR) below 2.1 [51]), as the TFR in the simulation is decreasing from approximately 1.8 in 2011 to 1.5 in 2020 followed by a slow increase to about 1.62 by 2050. Moreover, we assume continuous improvements in longevity, especially for males, for which the life expectancy at birth is increasing from 77.7 years in 2011 to 85.4 years in 2050 [173, 174]. Consequently, the population will continue ageing, with implications for the household structures.

We keep track of all demographic events in an event log file, which makes it possible to re-create the population and the changes therein. The host population in the disease transmission model thus evolves in a deterministic manner, making it possible to solely ascribe differential outcomes in a given simulation year to the disease transmission parameters. The demographic data, model and source code are described in detail in Chapter 3 and Appendix C.1.

4.2.2 Disease transmission model

Although we allow demographic events to take place on a daily basis, substantial effects of population ageing will only emerge after several years, as demographic change is a slow process. We have therefore chosen to simulate disease outbreaks every ten years. In 2020, 2030, 2040 and 2050, ten randomly chosen individuals become infected, in an otherwise fully susceptible population, on January 1st of each respective year. We consider an infectious disease transmitted via close contact, which can be represented by a *Susceptible-Infectious-Recovered* (SIR) model and consider several scenarios for age-specific susceptibility and infectiousness. The probability of becoming infected, and thus moving from the susceptible to infectious state, is calculated using a two-level mixing model, where an individual can acquire infection as a result of interactions with an infected household member (local contact) or an infected individual in the general population (global contact) [175]. For each combination of parameter settings in the two-level mixing model, we run 30 simulations, but in the analysis, we disregard those where an outbreak never takes place (i.e. total attack rate of less than 0.5%).

Within- and between-household interactions

We use a contact matrix to estimate social interactions (i.e. a proxy for an at-risk event at which infection can be transmitted) between non-household members in the general population. The contact matrix is based on social contact data collected in a survey in Belgium in 2010-2011 [167] and made available through the SOCRATES data tool [176] (see Figure C.1). Contacts between household members were excluded, as these are captured by the household level of the model, but contacts taking place in the household with non-household members were included. Additionally, supplementary professional contacts (SPC) were excluded.

To model interactions among household members, we construct a household contact network for each household in the population. Specifically, each household member is represented by a vertex, and a link between two vertices indicates a contact between those two household members. The links are constructed using an exponential random graph model developed by Krivitsky et al. [150], which was fitted to data from two contact surveys conducted in Belgium in 2010-2011 [50, 167]. The household contacts are limited to those involving skin-to-skin touching. The model accounts, amongst other things, for the type of household and the age-sex composition. The probability of a contact between two household members, however, is independent of past contacts between them. In each time step (i.e. day), we apply the fitted model from Krivitsky et al. [150] to the households in the simulated population and simulate who comes into contact with whom within each household. The mean network density (the number of links in a household relative to the number of possible links [177]) by household size and type are shown in Figure C.2. Contacts between household members are often repetitive because of the high contact density in the households. In the general population, repetitive contacts are less likely because of the large population size. Nevertheless, they may still be important, but data on the share of repetitive contacts is lacking.

Risk of infection

Each susceptible individual i acquires infection at time t with probability $p_i(t)$:

$$p_i(t) = 1 - \prod_{\substack{j \neq i \\ j \in h_i}} (1 - \beta_h s_i z_j a_{ij}(t) I_j(t)) \cdot \prod_{j \notin h_i} (1 - \beta_p s_i z_j c_{ij}(t) I_j(t)) \,, \tag{4.1}$$

where h_i denotes the household of individual i and the parameters β_h and β_p rep-

resent the probabilities of disease transmission given contact between a susceptible and infectious individual within the household and in the general population, respectively. We vary the transmission parameters to reflect different settings (e.g. high vs. low household transmission) and are thus not calibrating the parameters to data for specific infections. The relative susceptibility and infectiousness given the ages of individual i and j are represented by s_i and z_j, respectively, while $I_j(t)$ takes the value one if individual j is infected at time t and zero otherwise. The contact network in household h_i is represented by an adjacency matrix $\mathbf{A}$ of which the element $a_{ij}(t)$ equals one if household members i and j come into contact with each other at time t, and zero otherwise. A new adjacency matrix is generated in each time step conditional on the household composition.

The social contact matrix from Figure C.1 contains the mean number of contacts per day in the general population between each age group, m_{ij}, thus we compute the probability by which individual i and j come into contact with each other on a given day (time t is discretised in days) based on the age groups to which i and j belong, $c_{ij}(t)$, as follows:

$$c_{ij}(t) = \frac{m_{ij}}{N_j(t)}, \tag{4.2}$$

We assume disease transmission in the general population to be frequency-dependent, meaning that the number of effective contacts made by each person remains unchanged as the population grows. Thus, to keep the age-specific contacts constant over time, the element m_{ij} is divided by $N_j(t)$, the size of the age group of j at time t. In each time step, the probability of infection based on Equation 4.1 is computed for all susceptible individuals in the population and their disease state is updated accordingly. We are not taking factors like seasonality, weekends and school holidays into account, as we are focussing on the role of population heterogeneity in the spread of an infection. The interplay between these factors remain a topic for further research.

Infectious period

We assume that the infectious period follows a gamma distribution with a mean of 3.8 days and a standard deviation of 2 days, reflecting the infectious period for influenza [31, 178, 179]. For each newly infected individual, a value is drawn from the distribution and rounded to the nearest integer. An infected individual recovers and obtains immunity when the infectious period has passed.

Age-specific susceptibility and infectiousness

We consider different scenarios for age-specific susceptibility and infectiousness. Whereas this is primarily an explorative set of scenarios, some of these age-specific susceptibility and infectiousness profiles were motivated by specific infections, including influenza and Respiratory Syncytial Virus (RSV) (e.g. scenario S4 [180]), SARS-CoV-2 (e.g. scenario S1 or S3 [181–183]). Other considered elements includes the general phenomenon of immunosenescence depicting higher susceptibility at older age (e.g. scenario S2). Moreover, we use the scenarios to assess the role different population groups play in the spread of an infection. The age-specific susceptibility and infectiousness scenarios are shown in Figure 4.1 in relative terms, meaning that a value below one corresponds to reduced susceptibility or infectiousness and a value larger than one implies increased susceptibility or infectiousness for individuals in the given age group. Susceptibility is age-dependent (different from one) in scenario S1-S4, while infectiousness is age-dependent in scenario I1-I4. We compare the different scenarios to a baseline case where infectiousness and susceptibility are equal across all ages.

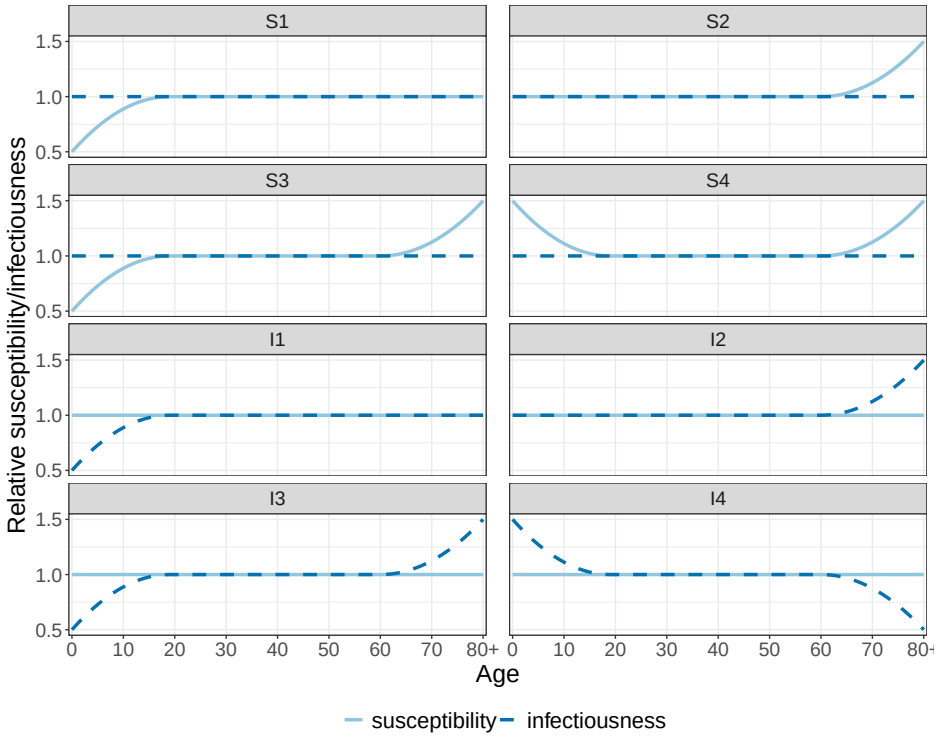

Figure 4.1. Age-specific susceptibility and infectiousness scenarios.

4.3 Results

4.3.1 Population dynamics

Age and household size are closely connected as seen in Figure 4.2. Children and adolescents most often live with their parent(s) and siblings, meaning that households of size three and larger are most common at young ages, which implies a similar pattern in the parental age groups (e.g. ages 30-55). The average household size starts to decrease in late adolescence, as children leave the parental household, and increases again from the late twenties with the entry into parenthood. Again, a similar pattern is seen in the parental generation (e.g. age 50+), but with a continuous decrease in mean household size until the mid-seventies, when widowhood and the need for LTCFs become more prevalent. Consequently, single-person households and very large households (i.e. LTCFs) become more frequent in the oldest age groups.

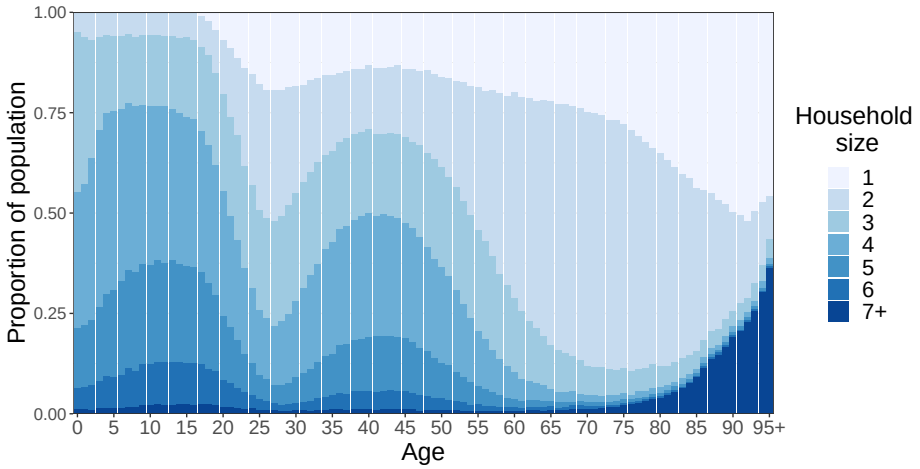

Figure 4.2. Household size distribution by age group of simulated population in 2020.

The population structures do not change dramatically over time, as the Belgian population in 2020 already has an old age structure and we are not assuming extreme changes in the vital rates in the simulation period of 30 years. Nevertheless, the simulated population continues ageing between 2020 and 2050 as seen in Figure 4.3 (left panel). The share of the population older than 60 years increases by 22% between 2020 and 2050, while the share in the oldest age group alone (i.e. 81+) increases by more than 75%. This is the result of past long-term trends of declining fertility and increasing longevity, which continue to a certain degree in the simulation pe-

riod. Moreover, the ageing of the large generations born in the mid-twentieth century causes a temporary acceleration of the population ageing.

The changes in the age distribution are also reflected in the household size distribution as seen in Figure 4.3 (right panel). The proportion of the population living in a single-person household increases as the population ages, since the proportion of people living alone is higher in the older age groups. The proportion living in households of size two is increasing from 2020 to 2040, which is mainly due to the increased survival of elderly males in a union. While an increasing share of the population lives in small households of size one to three, the proportion of larger households of size four to six is decreasing. Households of those sizes are to a large extent occupied by nuclear families in the age range of 0-50 years (see Figure C.3 and C.4 in Appendix C), which makes up a decreasing proportion of the simulated population from 2020 to 2050. The proportion of the population living in households of size seven and larger remains quite stable during the simulation period, but it is the result of two opposite trends. The proportion of parents with a large number of children, and therefore with a large household, is decreasing, while people living in LTCFs (i.e. elderly population), make up an increasing proportion of the population.

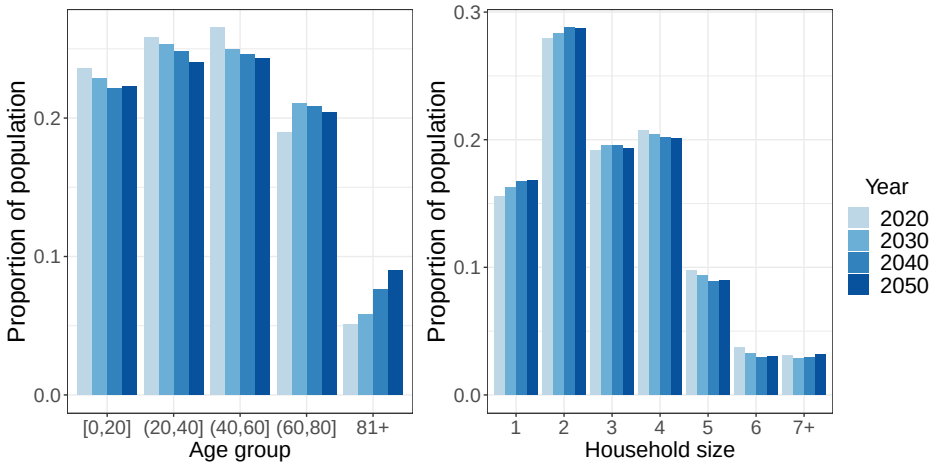

Figure 4.3. Age distribution (left) and household size distribution (right) of the simulated population in 2020, 2030, 2040 and 2050.

The age-specific household size distributions slightly change over time (see Figure C.3 and C.4 in Appendix C). The average household size for children and their parents (i.e. younger than 50 years) is decreasing. This is a consequence of single parent families becoming more prevalent and a decreasing TFR prior to 2020, followed by a slow,

but not full, recovery (see Figure C.5). Meanwhile, the average household size in age group 50-70 is increasing as the likelihood that their household accommodates (adult) children increases, due to past changes in the timing of childbearing. A change in the household size distribution is also seen in the elderly population as a result of improvements in longevity. Consequently, single-person households and collective households (i.e. LTCFs) are increasingly being replaced by two-person households.

4.3.2 Disease transmission dynamics

Attack rate

The proportion of the population that contracts the infection during an outbreak (the attack rate) responds to changes in the transmission parameters (β_h and β_p in Equation 4.1) in a non-linear pattern (see Figure C.6 in Appendix C). The relative increase in the attack rate diminishes as the transmission probabilities increase, especially when the transmission probability given contact within the household increases. Potential household infections are limited by the size of the household and more than 15% of the simulated population live in single-person households. Thus, as the household transmission probability keeps increasing, everyone with an infected household member will eventually become infected as well. The attack rates in each susceptibility and infectiousness scenario deviate from the baseline scenario since some population groups face an increased or decreased risk of acquiring or transmitting the infection given the specific scenario. However, the differences between the scenarios diminish as the transmission probabilities increase.

Age- and household-specific transmission dynamics: Baseline scenario

The age-specific attack rates in the baseline scenario are shown in Figure 4.4 for varying probabilities of household (closed vs. open circle) and community transmission (upper vs. lower panel) and at the different time points. We see that some population groups are more likely to get infected than others, also when discarding age-specific differences in susceptibility and infectiousness. The proportion of children and adolescents getting infected is larger than that of any other age group. The adult population also faces relatively large attack rates, which decrease from age 50 onwards. This reflects the age pattern in social contacts outside the household (see Figure C.1 in Appendix C). Nevertheless, social mixing in the general population alone cannot explain the age distribution in the attack rates. The age group 20-29, for example, has more contacts in the general population than the age group 30-49,

yet lower attack rates. This is due to the difference in household composition of these age groups. Individuals in their 20s are much more likely to live in households of size one or two, and thereby have a lower number of possible household contacts than people of age 40-49, who often live in larger households as seen in Figure 4.2. Moreover, the household members of the two groups tend to differ in case of a larger household. People aged 30-49 living in a household of size three or larger often have young children living with them, while the 20- to 29-year-olds are more likely to live together with their parents or unrelated adults in a house-sharing arrangement (see Figure C.4 in Appendix C). The mean density of the contact network is higher in the first household constellation than in the latter because of the presence of young children (see Figure C.2 in Appendix C). Finally, children and teenagers are more likely to bring an infection into the household given their high number of contacts in the general community, including schools, thus putting the parents at an increased risk (assuming that susceptibility and infectiousness are independent of age).

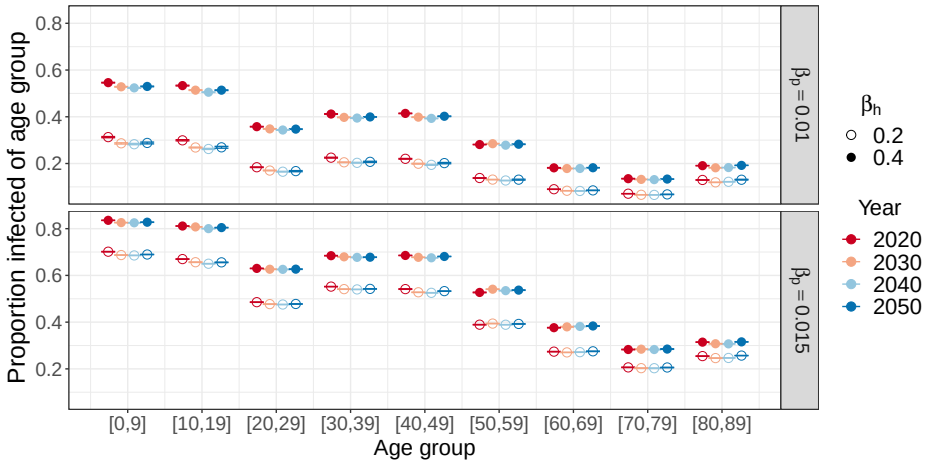

Figure 4.4. Mean age-specific attack rate with 95% confidence interval for varying transmission parameters (β_h: filled vs. open circle, β_p: upper vs. lower panel) in the baseline scenario (susceptibility and infectiousness are equal across age).

The attack rates in the oldest age groups can also only be explained by considering household structure. The attack rate is higher in age group 90+ than in 80-89, despite both age groups having identical social contact rates outside the household. From the age of 80 onwards, small households of size one and two are increasingly replaced by very large households (see Figure 4.2), such as LTCFs.

Household-specific attack rates (i.e. the proportion of household members acquiring infection) are shown in Figure 4.5 by household size (x-axis) and household transmission probability (open vs. closed circle). Moreover, we distinguish between a risk set containing all households (upper panel) and a risk set only containing households with at least one infected household member during the outbreak (lower panel). Generally, the household-specific attack rate increases by household size (Figure 4.5 upper panel). This result is a combination of how likely an infection is to enter the household and how easily it spreads within that household. The number of individuals that can bring an infection into the household increases with the household size, however, the likelihood of it happening also depends on the social contact patterns of each household member. The infection is more likely to spread to households with at least one child younger than 13 years than to households of the same size without children, because children have a relatively large number of contacts in the community (see Figure C.7 in Appendix C).

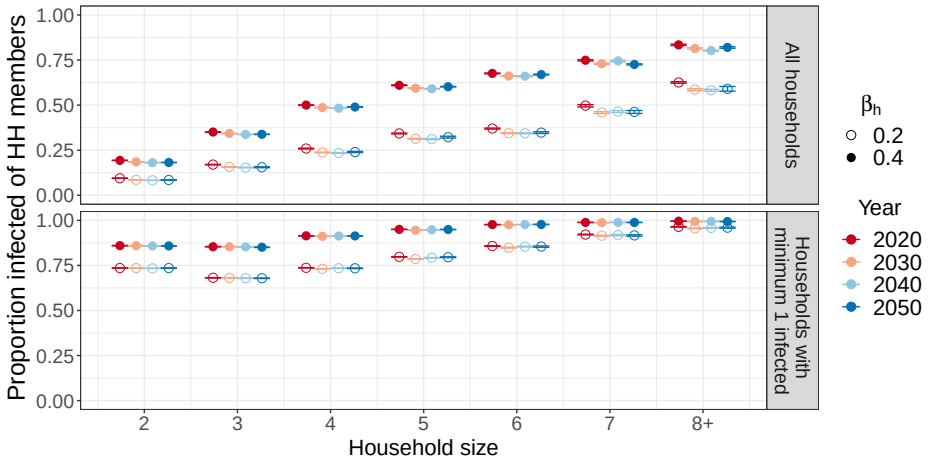

Figure 4.5. Mean proportion of household members getting infected by household size in baseline scenario. Upper panel: Estimate based on all households. Lower panel: Estimate based on households with minimum one infected individual. $\beta_p = 0.01$.

After the infection has entered a household, the further spread is affected by the household size and composition. The within-household transmission is visualised in the lower panel of Figure 4.5, as the estimates are limited to households with at least one infected individual. In that case, the differences across household sizes are substantially smaller and the mean proportion of household members getting infected even decreases from household size two to three. The decrease, however, is only observed for households without a young child. The presence of a child in

the household affects the within-household transmission across all household sizes of less than eight, as young children tend to have more contacts with their household members (i.e. parents, siblings) than teenagers and young adults do with theirs (see Figure C.8 in Appendix C).

Effect of demographic change on transmission dynamics: Baseline scenario

The proportion of the population acquiring the infection during an outbreak is decreasing from 2020 to 2040 in all scenarios but the trend stabilises or reverses between 2040 and 2050 (see Figure C.6 in Appendix C). This is the result of changing household structures and population ageing. The elderly population, which over time makes up an increasing proportion of the population, has relatively few contacts on average since the majority lives in small households and has few contacts in the general population. Consequently, the elderly population has a lower risk of acquiring and transmitting an infection than younger age groups.

Additionally, the changing household compositions in the population younger than 50 years of age resulting from low fertility levels and an increasing proportion of single parent families, decreases their risk of infection over time, with implications for the overall incidence. Meanwhile, the proportion infected of age group 50-70 remains more or less stable, despite an increasing proportion living in households larger than size two. Finally, improved longevity implies that the elderly population (especially females) becomes more likely to live with their partner than alone or in LTCFs, which affects the incidence in the oldest age group (90+). The relationship between risk of infection and household size persists as the population is ageing, but the proportion of infected household members is decreasing over time across all household sizes (Figure 4.5 upper panel). Meanwhile, the within-household transmission remains stable over time (Figure 4.5 lower panel).

Effect of age-specific infectiousness and susceptibility

We further investigate the role different age groups play in the spread of an infectious disease. In Figure 4.6, we compare the age-specific attack rates in 2020 across the susceptibility and infectiousness scenarios to those of the baseline scenario (i.e. corresponding to Figure 4.4). This is visualised for varying population transmission probabilities (upper vs. lower panel). Differences from the baseline attack rate are not only seen in the age groups with modified susceptibility or infectiousness, but also in the rest of the population to varying degrees. The susceptibility and infectiousness of children affect all population groups, and the parental generation in particular (e.g.

age group 30-39), to a markedly larger degree than changes in the susceptibility and infectiousness of the elderly population. In scenario S1 (I1), children have a relatively low susceptibility (infectiousness) which affects all other age groups substantially, while the relatively high susceptibility (infectiousness) from age 65 onwards in scenario S2 (I2) has a much smaller effect on the incidence in other age groups. This is also seen by comparing scenarios S1 and S3 (I1 and I3), where the attack rates below the age of 60 do not differ substantially, despite the increased susceptibility (infectiousness) of the elderly population in the latter scenario. Even the elderly population itself is only somewhat affected by changes in their infectiousness. However, it applies to children as well as the elderly, that a change in their infectiousness only has a slightly larger impact on the incidence in the rest of the population compared to the same change in susceptibility.

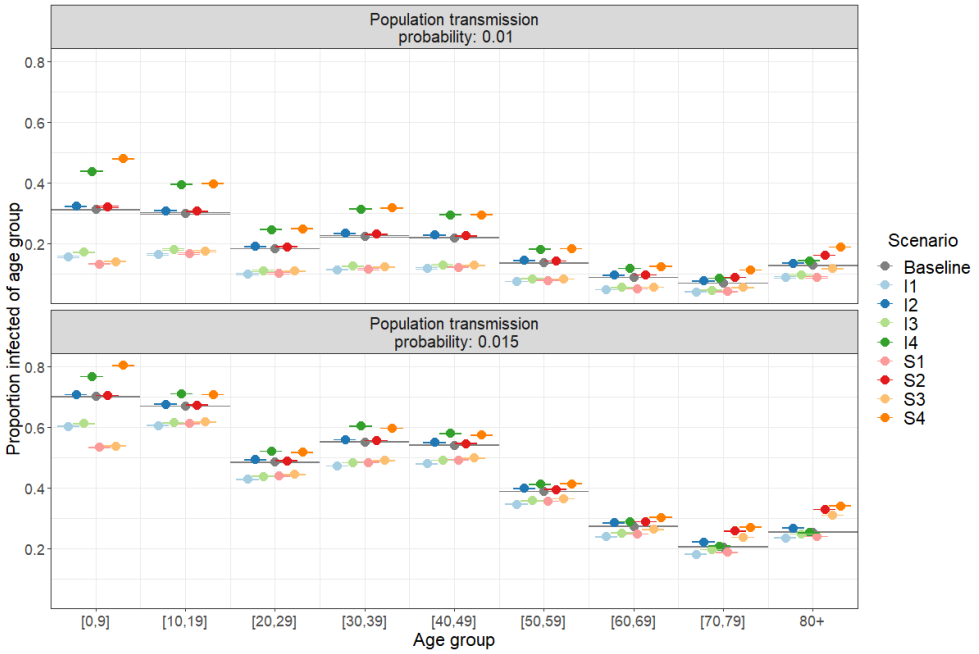

Figure 4.6. Mean and 95% bootstrap confidence interval for the mean age-specific attack rate in baseline scenario (grey) and each susceptibility/infectiousness scenario across varying population transmission probabilities (upper vs. lower panel) and household transmission probability of 0.2. Simulation year 2020

As the population transmission probability (upper vs. lower panel in Figure 4.6) increases, the absolute difference between each scenario and the baseline attack rate diminishes, except for the age group 70 and older when subject to increased susceptibility as in scenario S2, S3 and S4. This is particularly pronounced in scenario

S3, where children have relatively low susceptibility while the old age groups have high susceptibility. As the transmission probabilities increase, the attack rates in age groups 70-79 and 80+ in S3 shift from being below to being above those of the baseline scenario. The attack rates change slightly over time in all scenarios, but the position of each scenario relative to the other scenarios and the baseline remains unchanged (see Figure C.9 and Figure C.10 in Appendix C).

4.4 Discussion

An understanding of demographic structures in the host population and how these influence disease transmission can be important for determining which population subgroups are most at risk and most effective to target in an intervention [17]. Moreover, an understanding of the demographic processes underlying the population structures may be important for assessing how future demographic changes potentially could affect transmission dynamics.

Using longitudinal microdata drawn from Belgian census and population registers, we model a host population with evolving age and household structures using microsimulation and illustrate how population ageing and changing household dynamics may further unfold in the next decades. We combine the demographic microsimulation with an epidemic model with two levels of mixing and illustrate a strong link between age and household structures and the implications thereof for the risk of infection for different population subgroups during an epidemic. Additionally, we quantify the potential impact of changing age and household structures on disease transmission.

The age structures in the social contact patterns are mirrored in the age-specific attack rates as the youngest age groups, with the most community contacts, have the highest risk of infection. The attack rates in the adult population, however, are to a larger degree explained by the differences in household compositions between young adults, middle-aged adults and older adults, which are related to the timing of demographic events. Young adults in their 20s face a relatively low risk of infection on average, but it increases with the entry into parenthood, when assuming that susceptibility and infectiousness are independent of age. The child will eventually have a relatively high number of contacts outside the household (e.g. in day-care, school) as well as frequent contact with the parent(s) within the household, making the risk of infection high compared to other population groups. These relationships change to some degree in the different scenarios for age-specific infectiousness and susceptibility.

The incidence in families with children and/or adolescents decreased during the simulation period as a consequence of changing household compositions. In the decade prior to the first outbreak simulations, 2010-2019, the total fertility rate decreased followed by a slow, but not full, recovery during the remaining simulation period, similarly to observed and projected rates by Statbel and FPB [174]. Changing fertility levels combined with an increase in single parent families affected the household compositions and indirectly disease transmission. This also affects disease transmission in other population groups since families with (school-age) children are important drivers in an epidemic.

As the children grow up and, in most cases, eventually leave the parental household, the number of household contacts of the now middle-aged parental generation is decreasing again, often in combination with decreasing community contacts, leading to lower attack rates. However, we found that the proportion of middle-aged people with (adult) children living in their household is increasing during the simulation. This change is related to the postponement of parenthood since the probability of leaving the parental household is assumed to be constant over time. Parents are increasingly older when the last child leaves the parental household because the average age at childbirth was increasing prior to 2011 when our simulation begins [161]. Several years later, these past fertility trends affect the household composition of the middle-aged population and indirectly their risk of infection. The increasing household size was expected to increase the risk of infection in the middle-aged population, but the effect is more or less counterbalanced by the decrease in the overall incidence induced by the changing household compositions in the younger age groups.

It should be noted that the relative distribution of births by the age of the mother is only slightly changing in the first decades of the simulation and eventually stabilises. However, the average age at childbirth in Belgium has increased since 2011 and this is expected to continue in the future, to some degree [174]. Hence, the average age at childbirth, and indirectly the age at which children leave the parental household, is likely to increase more than in our microsimulation.

The risk of infection in the elderly population was also found to be highly dependent on their living arrangement. Community contacts are decreasing with age and a large proportion of old people live alone or only with their partner, which minimises the number of occasions where transmission of a close-contact infection can take place. However, from the age of 80 onwards, the proportion of the population living in LTCFs, which tend to be very large households, increases considerably and so does

the risk of infection. Consequently, the age pattern in the attack rates in the elderly population is shaped by the proportion living in collective households.

During the simulation period, collective households and single-person households in the population older than 80 years of age are increasingly being replaced by two-person households, as a result of improved longevity, particularly of males. We assume that the sex differential in mortality decreases during the simulation period as a result of larger improvements in men's mortality than in that of women. This implies that an increasing proportion of elderly women are living with a partner instead of alone, which intuitively should increase their risk of infection. However, the probability of moving to LTCFs, which are very large households associated with a high risk of infection, is substantially higher for elderly people living alone than for those living with a partner. Consequently, gradually fewer elderly people move to LTCFs and therefore incur a substantially lower risk of infection.

The future living arrangements and mortality of the elderly population are, like all other demographic processes, associated with uncertainty [184]. However, the sex differential in mortality has been decreasing and this is considered likely to continue, to some extent, in the future. The resulting changes in the household structures of the elderly population seen in our microsimulation are in agreement with existing studies of past and future trends in the living arrangements and mortality of older adults [185, 186].

In addition to social contact patterns, age and household structures, we also explored how epidemiological heterogeneities within the population may influence the spread of an infection. We incorporated different scenarios for susceptibility to infection when exposed and infectiousness when infected according to age. As anticipated given the social contact patterns and household structures, the susceptibility and infectiousness of children and adolescents were highly influential for the disease transmission in the whole population and in the parental generation in particular. Changes to these epidemiological parameters in the elderly population clearly affected that age group, but exerted much less influence on other age groups.

The elderly population, however, is affected differently by changes in the transmission parameters than the rest of the population, when subject to elevated susceptibility. As the probability of transmission given an effective contact increases, the underestimation of the attack rate in the elderly population when omitting age-dependent susceptibility (baseline scenario) increases, while it decreases in the rest of the population, and in all other scenarios. Generally, many older adults escape infection

due to their limited number of contacts within and outside the household. However, as the transmission probabilities increase, a proportionately larger share of households with elderly people are reached by the infection and more individuals within the households become infected if their susceptibility is elevated. Thus, the impact of epidemiological heterogeneities (e.g. age-specific susceptibility) is not only dependent on the transmission parameters but also on other heterogeneities in the population, including social contact patterns and household structures.

Overall, we find that the strong relationship between age and household structures at the individual and population level, in combination with social mixing patterns and epidemiological parameters, shape the spread of an emerging infection. Disease transmission in the adult population in particular is influenced by differential household compositions. Moreover, we highlight how demographic processes alter population structures with differential impact on the disease transmission dynamics across population groups. Nevertheless, our study faces several limitations with regard to demographic modelling as well as infectious disease modelling.

We recognise that the demographic processes in the microsimulation are simplifications of highly complex processes and that the inherent uncertainty in population projections preferably is described in the form of probability distributions [123, 187]. Moreover, the sensitivity in the association between demographic and epidemiological outcomes could have been explored. However, expanding our demographic microsimulation is not considered necessary to fulfil the aim of this paper, which is to document the impact of population structures and the changes therein on the spread of an emerging infection. For future research, however, it would be relevant to compare the microsimulation and two-level mixing model to other epidemic models with household-structured host population (e.g. [188]).

In our model of infectious disease transmission, we assume a fully susceptible population, restricting our study to emerging infections. Expanding the study to endemic infectious diseases requires not only information on age-specific patterns of prior immunity but also information on how immunity is distributed in households. Alternatively, a population at an endemic disease equilibrium can be generated, for example by simulating disease transmission over a long period of time before the actual analysis [41, 42]. However, this would require a rather complex technique and/or detailed (historical) demographic data. Moreover, if the fertility and mortality schedules remain constant while generating an endemic disease equilibrium, the population eventually acquires the age distribution of the stable population associated with those

underlying schedules of vital rates, which may not resemble the population of interest [51].

Another limitation in our model of disease transmission concerns the social contact patterns. If SPCs had been included in the social contact matrix, the incidence in the population of working age would have been slightly higher, however, the relationships found between population groups would remain. Additionally, we assume that the contact patterns in a household of a given composition are constant over time and that community transmission is frequency-dependent with constant age-specific contacts. Methods for adjusting social contact matrices to evolving demographic structures have been proposed, but these are not based on empirical evidence for how contact patterns behave over longer time frames as population structures evolve [189, 190]. A comparison of two social contact surveys in Belgium five years apart suggests that the contact rates can be assumed stable, but the demographic change in this period is of course limited [167].

Chapter 5

Exploring the impact of population ageing on the spread of emerging respiratory infections and the associated burden of mortality

This chapter is based on submitted work: "Møgelmose, S., Neels, K., Beutels, P. & Hens, N. Exploring the impact of population ageing on the spread of emerging respiratory infections and the associated burden of mortality."

5.1 Background

The population age structures in most high-income countries have for decades been shifting towards older ages (i.e. population ageing) as a result of increasing life expectancy and persistent below-replacement fertility levels. Currently, a temporary acceleration of population ageing is seen in many countries due to the ageing of the large generations born in the mid-twentieth century [191]. Moreover, population ageing has become a global phenomenon and the proportion of older adults in many low- and middle-income countries is increasing at an unprecedented speed [64, 172, 192, 193].

The rising burden of non-communicable diseases induced by population ageing has rightfully been given a lot of attention [194–196]. However, morbidity and mortality due to infectious diseases, respiratory infections in particular, remain substantial in the elderly population [39]. The progressive deterioration of immune functions with

age (immunosenescene) increases older adults' susceptibility to infection and their risk of a severe outcome in case of disease [76]. The COVID-19 pandemic, for example, has had a disproportionate impact on the older adult population and on those living in long-term care facilities in particular (LTCFs) [77, 78, 86–88]. Several aspects of LTCFs (e.g. communal meals, group activities, staff rotation) make them an optimal environment for rapid spread of many infectious diseases [85, 128]. Additionally, many LTCF residents have underlying chronic illnesses, which, in addition to their old age, may increase the severity of an infection [129, 130]. Nevertheless, LTCF residents only make up a minority of the older adult population in most high-income countries. The majority of old people typically live alone or with a partner. Moreover, social contact surveys from several European countries have shown that people aged 65 and older have the lowest mean number of contacts [167], and thus fewer interactions that potentially could lead to disease transmission. Consequently, the incidence of infections transmitted via close contact may be relatively low in the oldest age groups, yet the disease burden is typically substantial [197–201].

This implies that high-income countries with ageing populations may face a decreasing overall incidence of an infectious disease (e.g. influenza), but it could coincide with an increasing burden. However, the future burden of infectious diseases in older adults may, among other things, be influenced by the future health and living arrangements at old age. It remains unclear whether the increases in life expectancy are accompanied by a proportionate increase in healthy life expectancy [202–205]. Health at old age and living arrangements are naturally connected, with relevance for infections transmitted via close contact. The proportion of elderly people living with a partner is expected to increase due to improved longevity, particularly of males [186]. Nevertheless, the proportion living in LTCFs is also likely to increase as the proportion of the oldest people (i.e. 85+) increases [185, 206].

Several studies have investigated the impact of population ageing on the spread and burden of different infectious diseases, including measles, influenza, pneumonia and herpes zoster (e.g. [36, 37, 42, 73, 75, 97, 118, 207, 208]). Nevertheless, only few studies consider a household-structured population and to our knowledge none of them incorporate LTCFs. We aim to improve the understanding of how changing age and household structures affect the future transmission dynamics and mortality burden of respiratory infections in an ageing population, and explicitly explore the role of living arrangements in the older adult population. Specifically, we consider the Belgian population, which, like most other high-income countries, has a relatively old age structure and is still ageing. We use the demographic microsimulation presented in

Chapter 3, which is based on longitudinal microdata drawn from Belgian census and population registers. The microsimulation includes dynamic demographic processes for fertility, mortality, migration and household transitions, making it possible to model the Belgian population over time with evolving age and household structures. In addition to private households, collective households (e.g. LTCFs) are represented in the microsimulation. Due to the uncertainty surrounding the future health and living arrangements of older adults, we consider three demographic scenarios with respect to the proportion of LTCF residents. We refer to the scenarios as *low, medium* and *high* to describe the proportion of older adults living in LTCFs relatively to the other scenarios.

We subsequently combine the demographic microsimulation with a disease transmission model representing the spread of SARS-CoV-2 and a novel influenza A virus. The model is a modification of the two-level mixing model presented in Chapter 4, which distinguishes between exposure to infection in the household and exposure in the community at large. Additionally, the model implements contact networks within households which are based on empirical data, rather than making the common assumption of random mixing [150]. We simulate outbreaks of SARS-CoV-2 and influenza in a fully susceptible population in 2020, 2030, 2040 and 2050, which allows sufficient time for demographic change to emerge.

We first illustrate how the age and household structures are altered in an ageing population. Secondly, we explore how the changing population structures affect the spread of the two respiratory infections (i.e. incidence) and the burden of mortality in the form of disease-related deaths and quality-adjusted life-years (QALYs) lost. In health economics, QALY expectations (gains or losses) represent a commonly used summary measure of longevity adjusted for the combined impact of death and morbidity [209]. Finally, we investigate to what extent our findings at the individual and population level are affected by changes in the living arrangements in the older adult population.

5.2 Methods

5.2.1 Demographic microsimulation

We simulate the Belgian population using the microsimulation presented in Chapter 3. The initial population is based on a sample from the Belgian census in 2011. For each individual, we have information on their date of birth, sex, ID of parents, birth

trajectory (parity and date of most recent birth if applicable), household ID and household position (e.g. in union, child, single parent). Thus, individuals can be linked to each other through household membership and kinship. The population evolves over time as individuals can enter and leave the population as a result of births, deaths and migration. Moreover, individuals can move between or create new households, for example as a part of union formation or dissolution. Finally, all individuals are ageing over time and the population is updated accordingly.

The probability of a demographic event taking place varies by individual characteristics, including age, sex and household position, and changes over time except for the household transition rates. We assume that mortality, fertility and migration levels in the microsimulation resemble the observed and projected rates by the Belgian Statistical Office (Statbel) and the Belgian Federal Planning Bureau (FPB). This implies below-replacement fertility (a total fertility rate below 2.1 [44]) and continuous improvements in longevity, especially for males [45,46] (see Figure D.1 and Figure D.2 in Appendix D). Consequently, the population will continue ageing, with implications for the household structures. Fertility trajectory and/or household position are included as covariates in the sub-models for fertility and mortality, as they have been shown to affect the probability of having a(nother) child and dying, respectively [210, 211]

We consider three demographic scenarios pertaining to the household structures in the older adult population (i.e. people aged 75 and older). The large majority of LTCF residents live in a single-person household prior to moving to the LTCF, thus we created three scenarios by varying the probability of leaving a single-person household for people aged 75 and older. The cut-off is made at the age of 75 years since only a small proportion of the population reside in LTCFs at younger ages (see Figure D.13 in Appendix D). We refer to the scenarios as *low, medium* and *high*, as an indication of the proportion of the older adults living in LTCFs. The demographic data, model and source code are described in detail in Chapter 3 and Appendix D.1.

5.2.2 Disease transmission model

In addition to the demographic attributes, all individuals are assigned a disease state. Disease outbreaks take place in the simulated population in 2020, 2030, 2040 and 2050 as ten randomly chosen individuals become infected, in an otherwise fully susceptible population, on January 1st of each respective year. The outbreaks are ten years apart to give sufficient time for demographic changes to emerge. We use an

SEIR-like (*Susceptible-Exposed-Infectious-Recovered*) model to describe the spread of respiratory diseases transmitted via close-contact interactions with the examples of COVID-19 and influenza. The probability of becoming infected, and thus moving from the susceptible to exposed state, is calculated using a two-level mixing model, where an individual can become infected as a result of disease transmission within the household or in the general population [175].

Within- and between-household interactions

We use the same techniques as described in Chapter 4 to model social interactions, which serve as a proxy for an at-risk event at which infection can be transmitted. Contacts between non-household members in the general population are estimated using social contact data collected in a survey in Belgium in 2010-2011 [167] and made available as a contact matrix through the SOCRATES data tool [176] (see Figure D.3 in Appendix D). Contacts between household members were excluded, as these are captured by the household level of the model, but contacts taking place in the household with non-household members were included. Additionally, supplementary professional contacts (SPC) were excluded. SPC is a category for individuals with more than 20 professional contacts per day (e.g. bus drivers). These are likely to be less important than other types of contacts when it comes to the transmission of close-contact infectious diseases [212].

For each household, we construct a contact network to model interactions among household members. Contacts are determined to take place using an exponential-family random graph model developed by Krivitsky et al. [150], which was fitted to data from the social contact survey mentioned above [167] and a household contact survey [50], both conducted in Belgium in 2010-2011. Household contacts are limited to those involving skin-to-skin touching. The household contact network is, amongst other things, conditional on the type of household and the age-sex composition. In each time step (i.e. day), we apply the fitted model from Krivitsky et al. [150] to generate a contact network for each household in the simulated population. The household contact networks may thus change every day. The mean network density (i.e. the number of links in a household relative to the number of possible links [177]) by household size and type are shown in Figure D.4 in Appendix D.

Influenza

We formulate an SEIR model to describe the spread of a novel influenza virus such as the influenza A (H1N1)pdm09 virus that emerged in 2009. When acquiring the infection, the individual is not infectious at first (i.e. exposed or latent period), but becomes infectious as the latent period ends and eventually recovers as the infectious period comes to an end. Disease-related mortality is not considered explicitly in the model, but estimated after the simulation. Each susceptible individual i acquires the infection at time t with the probability:

$$p_i(t) = 1 - \prod_{\substack{j \neq i \\ j \in h_i}} (1 - \beta_h a_{ij}(t) I_j(t)) \cdot \prod_{j \notin h_i} (1 - \beta_p c_{ij}(t) I_j(t)) \,, \tag{5.1}$$

where h_i denotes the household of individual i and the parameters β_h and β_p represent the probability of disease transmission given contact between a susceptible and infectious individual within the household and in the general population, respectively. We select transmission parameters, β_h and β_p, that result in a group-to-group reproduction number (R_* [175]) of about 1.5, which resembles the basic reproduction number estimated for influenza A(H1N1)pdm09 [213–215]. This is further described in section D.5 and section D.6 in Appendix D. $I_j(t)$ takes the value one if individual j is infectious at time t and is otherwise zero. The contact network in household h_i is represented by an adjacency matrix A and the element $a_{ij}(t)$ equals one if household members i and j come into contact with each other at time t and is otherwise zero. A new adjacency matrix is generated in each time step.

The social contact matrix from Figure C.1 contains the mean number of contacts per day in the general population between each age group, m_{ij}, thus we compute the rate at which individual i and j come into contact with each other at time t given their age groups, $c_{ij}(t)$, as follows:

$$c_{ij}(t) = \frac{m_{ij}}{N_j(t)}, \tag{5.2}$$

The element m_{ij} is divided by $N_j(t)$, the size of the age group of j at time t, to keep the age-specific contacts constant over time. This implies that we assume disease transmission in the general population to be frequency-dependent, meaning that the number of effective contacts made by each person remains unchanged as the population grows. In each time step, the probability of infection is computed for all suscep-

tible individuals in the population and their disease state is updated accordingly. The latent period is drawn from a uniform distribution with 1 day as minimum and 5 days as maximum. We assume that the infectious period follows a gamma distribution with a mean of 3.8 days and standard deviation of 2 days [31, 178, 179]. For each newly infected individual, a value is drawn from the distribution. An infected individual recovers and obtains immunity when the infectious period has passed.

COVID-19

In order to model the spread of SARS-CoV-2, we use a model similar to that of Willem et al. [32], which involves an extension of the SEIR model. Infectious individuals are initially pre-symptomatic and some develop symptoms while others remain asymptomatic (see Figure 5.1).

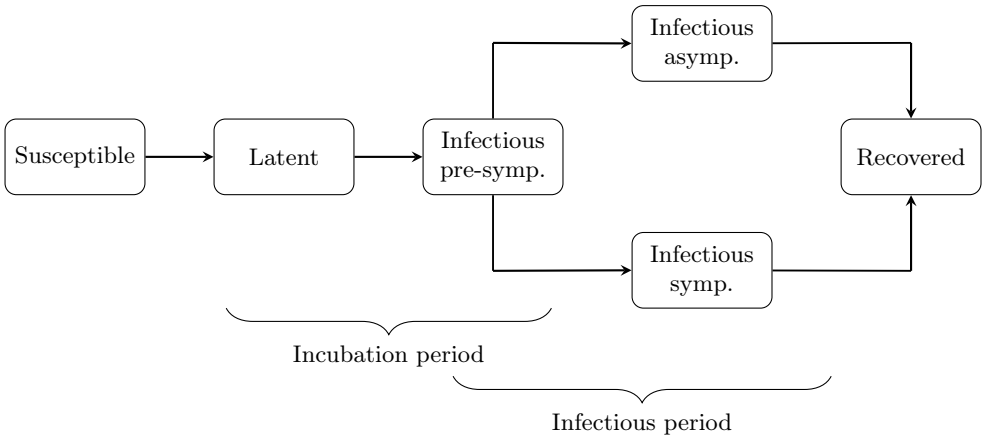

Figure 5.1. Disease transmission process for COVID-19. Symp.: Symptomatic.

Each susceptible individual i acquires infection at time t with probability:

$$\begin{aligned} p_i(t) = 1 - &\prod_{\substack{j \neq i \\ j \in h_i}} (1 - \beta_{h,a} a_{ij}(t) I_{j,a}(t) s_i) \cdot \prod_{\substack{j \neq i \\ j \in h_i}} (1 - \beta_{h,s} a_{ij}(t) I_{j,s}(t) s_i) \\ &\cdot \prod_{j \notin h_i} (1 - \beta_{p,a} c_{ij}(t) I_{j,a}(t) s_i) \cdot \prod_{j \notin h_i} (1 - \beta_{p,s} c_{ij}(t) I_{j,s}(t) s_i) \,. \end{aligned} \tag{5.3}$$

The same notation is used as for the influenza model, but the subscripts indicate whether the infectious individual is symptomatic (s) or asymptomatic (a). Infected individuals without symptoms are assumed to be half as infectious compared to those with symptoms, however, we acknowledge that this parameter is associated with uncertainty [216, 217]. We select transmission parameters, β_h and β_p, that result in a group-to-group reproduction number of about 3, to reflect the estimated basic reproduction number in Belgium prior to lockdown [32, 218, 219]. This if further described in section D.5 and section D.6 in Appendix D

$I_{j,a}(t)$ ($I_{j,s}(t)$) takes the value one if individual j is infectious and asymptomatic (symptomatic) at time t and is otherwise zero. The parameter s_i represents age-specific susceptibility and is 0.5 if individual i is younger than 18 years of age and is otherwise one, as we assume that children and teenagers are half as susceptible as adults [183]. The incubation period contains a latent period and a pre-symptomatic period. In the latent period, the individual is infected but not yet infectious, whereas the individual is infectious in the pre-symptomatic period, but shows no symptoms (yet). The incubation period is based on findings of Li et al. [216] and He et al. [220] and is assumed to follow a log-normal distribution with mean and standard deviation on the log scale of 1.43 and 0.66, respectively (see Figure D.7). The incubation period spans over at least two days since we assume that infectiousness starts one day prior to symptom onset at the latest and one day after infection at the earliest.

Based on infectiousness profiles from He et al. [220], a discrete distribution for the pre-symptomatic infectious period was estimated by Willem et al. [32] (see Figure D.8). For each newly infected individual, we draw from the distributions for the incubation period and the pre-symptomatic period. The length of the latent period for a given individual is obtained by subtracting the sampled value for the pre-symptomatic period from that of the incubation period (after rounding to a discrete number of days).

The distribution of the infectious period (including pre-symptomatic period) is assumed to follow a normal distribution with a mean of six days and a standard deviation of one (see Figure D.9). For each infected individual, the length of the infectious period is drawn from the distribution and the pre-symptomatic period is subtracted in order to obtain the remaining days of infectiousness. It is determined whether the individual shows symptoms during this period according to age-specific probabilities estimated by Willem et al. [32] (see Figure D.10). The probability of being symptomatic is based on the age-specific relative susceptibility to symptomatic infection

reported by Wu et al. [221] assuming 50% of the overall cases in the population to be symptomatic. An infected individual recovers and obtains immunity when the infectious period has passed. We run 50 simulations using the COVID-19 and influenza models, respectively, but limit our analysis to those simulations where an outbreak takes place (i.e. total attack rate of 0.5% or more).

Disease-related mortality

We estimate influenza-related deaths by applying the infection fatality rates (IFRs) for the influenza A(H1N1)pdm09 pandemic estimated by Riley et al. [222] based on a serological survey of a cohort of households in Hong Kong (see Figure D.11). The rates are by age group, but no estimates are available for children younger than three years. Consequently, we apply the IFR of the age group 3-19 to all ages younger than 19, although this is likely to underestimate the fatalities in the youngest children. For COVID-19, we use IFRs estimated by Molenberghs et al. [77] for Belgium in the period March 8th to May 9th 2020. The IFRs are broken down by age and household type (see Figure D.12). The considered household types are LTCFs and non-LTCFs. For ages younger than 60 years, there is no distinction between the household types, likely due to the small number of LTCF residents of that age. LTCF residents are not directly identifiable in the demographic microsimulation. Therefore, we use the household position *collective* as a proxy. This household position covers residents in different types of institutions, including prisons and LTCFs, but we expect the large majority of older adults with that household position to actually be living in LTCFs.

Quality-adjusted life years

To provide an estimate of the potential years of life lost due to premature death and the health-related quality of those years of life lost, we estimate the QALY losses attributable to COVID-19 and influenza fatalities using the method presented by Briggs et al. [223]. Pre-existing comorbidities are associated with an increased risk of a fatal outcome upon infection with SARS-CoV-2 or influenza [182, 224, 225], which is taken into account when estimating the QALY losses (see further details in section D.9 in Appendix D). However, we do not consider QALY losses from morbidity due to non-fatal infections.

5.3 Results

5.3.1 Population ageing

The population is ageing during the whole simulation period. Between 2020 and 2030, it is primarily due to an increasing proportion aged 65-79 years, while the age group 80 years and older increases fastest in the remaining decades (see Figure 5.2 left panel). This reflects the ageing of the large generations born in the mid-twentieth century. Population ageing induces changes in the household size distribution (see Figure 5.2 right panel). The elderly population primarily lives in small households (size 1-2) or very large households in the form of LTCFs (see Figure D.13). Consequently, an increasing proportion of the population lives in households of these sizes (1, 2 and 8+) as the population is ageing. Additionally, households of nuclear families are decreasing in size due to low fertility and an increase in single-parent families. It should be noted that the group size 8+ primarily is made up by LTCFs, which tend to have 25-100 residents in the simulated population.

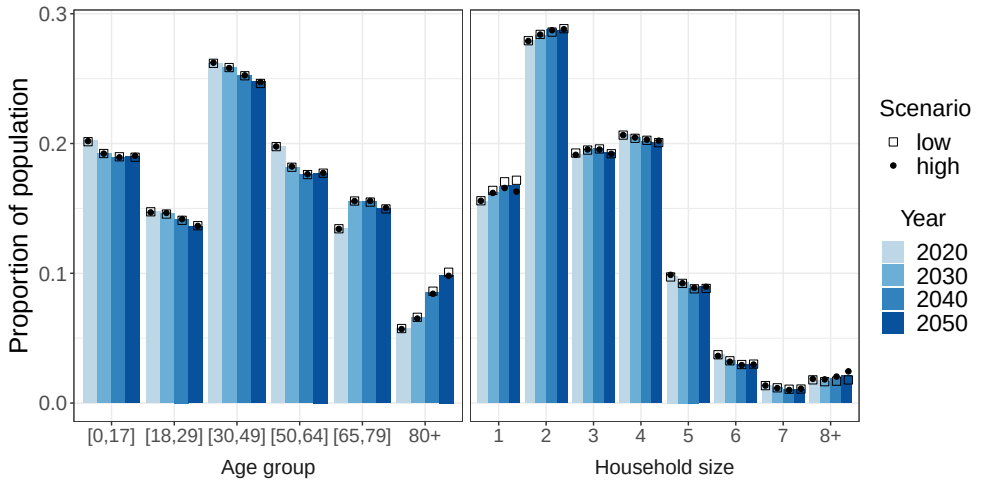

Figure 5.2. Age and household-size distribution by simulation year and scenario (medium: bar, low: square, high: circle).

The differences in the demographic scenarios only slightly affect the age distribution. The proportion aged 80 years and older is marginally larger in the scenario *low* (square) than in *medium* (bar) and *high* (dot), because of the smaller proportion of LTCF residents, which have a higher all-cause mortality (further information on all-cause mortality in the microsimulation is available in Chapter 3). The demographic scenarios have a more profound impact on the household size distribution. The pro-

portion in household size 1 and 8+ in scenario *low* (square) and *high* (dot) gradually diverge from the medium scenario (bar), but in opposite directions. Consequently, the proportion living in LTCFs (i.e. size 8+) relative to the proportion living alone is highest in the scenario *high* and lowest in scenario *low*, while the *medium* scenario is in between. The other household sizes are more or less unaffected. Household size distributions by age groups, scenario and simulation year can be seen in Figure D.14 and Figure D.15 in Appendix D.

5.3.2 Transmission dynamics

As expected, the proportion of the population that becomes infected during an outbreak (attack rate) in the COVID-19 model is substantially larger than for influenza, due to the differences in the transmission parameters. The attack rate is decreasing over time in both models, but after 2040, a slight increase is seen for influenza (see Figure D.17 in Appendix D). Older adults, which are increasingly replacing the younger population, have relatively fewer contacts on average since the majority live in small households (see Figure D.13) and have fewer contacts in the general population (see Figure D.3). Consequently, the older adult population has a lower risk of acquiring and transmitting an infection than younger age groups.

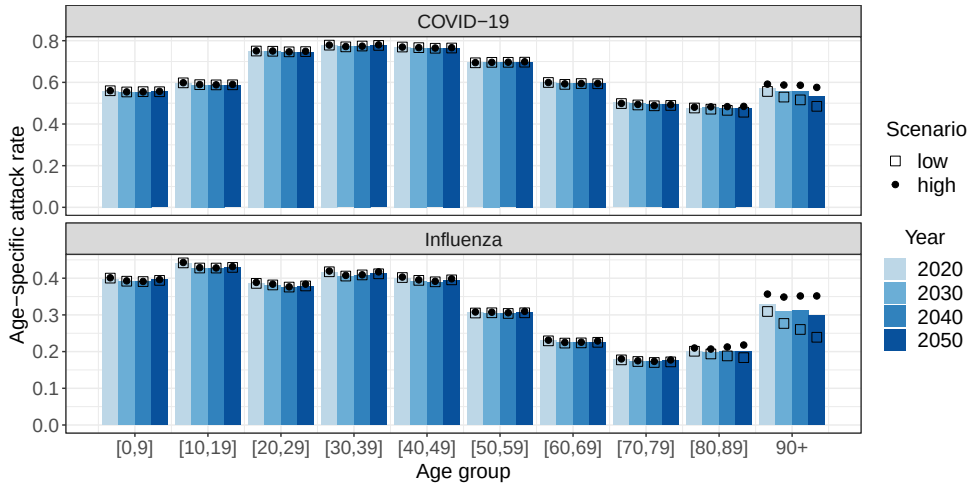

Figure 5.3. Mean age-specific attack rates by simulation year, model and demographic scenario. Note the different scales on y-axes.

The age-specific attack rates in the COVID-19 and influenza models naturally differ in magnitude, but other patterns are also seen (see Figure 5.3, note different scales on y-axes). For COVID-19 (Figure 5.3, upper panel), the attack rate is largest in the

adult population, which reflects the lower susceptibility of children and the increasing probability of being symptomatic and thereby more contagious with age. Meanwhile, influenza (Figure 5.3, lower panel) is more prevalent in children and adolescents compared to adults, which is resulting from the age-specific differences in contact patterns within and outside the household. This indirectly affects the parental generation (i.e. ages 30-49), which has the highest incidence in the adult population. In both models, the attack rate in the elderly population is shaped by the proportion of the age groups living in LTCFs, as the risk of infection increases with household size (see Figure D.19 in Appendix D).

These age patterns in the transmission processes imply that the impact of population ageing on the spread of COVID-19 and influenza differ. Since children and adolescents are the main drivers in influenza transmission, the attack rate declines substantially as the nuclear families decrease in size (i.e. low fertility and increase in single-parent families) and are increasingly being replaced by elderly people with a relatively low risk of infection (see Figure D.17). Meanwhile, the decrease in the attack rate of COVID-19 is less pronounced, and barely observed for the 20- to 79-year-olds (see Figure D.18). The risk of community transmission in the young and middle-aged adult population remains substantial (i.e. at work-places) due to the increased probability of being symptomatic and the assumption of frequency-dependent transmission. Moreover, old people (i.e. 70+) account for a larger share of infections with COVID-19 than with influenza, thus the decrease in the overall attack rate of COVID-19 induced by population ageing is less pronounced.

Since the population is ageing, the age composition of the infected people in the population is also shifting, but not necessarily to the same degree. In Figure 5.4, we compare the relative change in the age distribution (black bars) to the relative change in the age distribution of infected people (blue bars), both as a proportion of the total population size in simulation year 2030, 2040 and 2050 and using the corresponding values for 2020 as the reference. Infected people younger than 65 years of age make up a decreasing proportion of the population, while the proportion of infected people aged 65 and older is increasing, as to be expected considering the changes in the age distribution. The proportion of infected children and adolescents (i.e. younger than 18 years) in the population is decreasing more than the overall proportion of the age group across all demographic scenarios and models. This results from the decreasing household size of nuclear families (see Figure D.14), which is associated with a lower risk of infection for children and their parents as described earlier. Nevertheless, in the COVID-19 model, young and middle-aged adults experience a more or less equal

relative change in the proportions since community transmission is more pronounced. This is also the case for age group 50-64 in the influenza model, but due to an increasing mean household size as the share living together with their adult children is increasing. The growth in the proportion of infected 65 to 79-year-olds in the population is slower than the overall growth in the age group, as the share living in single-person households is increasing, especially for people in their seventies.

The relationship between the growth in the proportion of infected people aged 80 and older in the population and the general growth in the age group vary by demographic scenario. In the medium scenario, the growth rates are very similar, meaning that the age group is barely benefitting from the lower transmission in the young population. First of all, older adults have few contacts with children and adolescents. Second of all, the share of the age group 80+ living with a partner instead of alone is increasing. Finally, the age group is increasingly made up by people aged 90 and older, which have a higher attack rate.

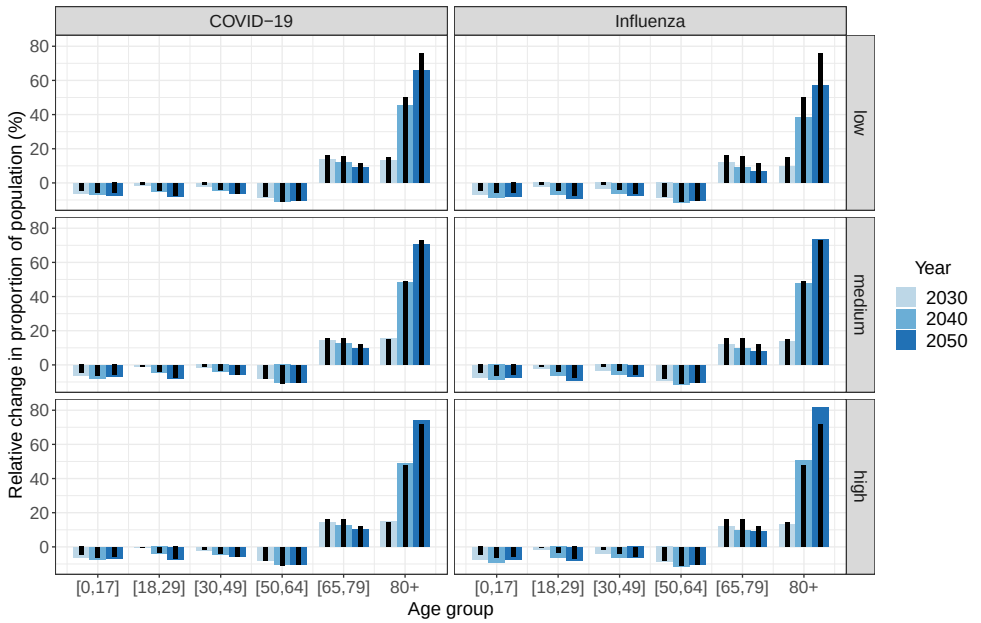

Figure 5.4. Mean relative change in size of age group (black bars) and number of infected people in age group (blue bars) as proportion of total population compared to 2020. Demographic scenarios by row and models by column.

In the scenario with a relatively low and decreasing share of the population living in LTCFs (first row in Figure 5.4), the proportion of infected people aged 80 years and older in the population is growing at a slower rate than the age group overall,

while this relationship is reversed in the scenario with a high and slightly increasing share living in LTCFs (third row in Figure 5.4). The differences between the scenarios *low* and *high* for age group 80+ are generally larger in the influenza model than in the COVID-19 model. The risk of infection for an elderly person living in a small household compared to someone living in an LTCF differs substantially more in the influenza model than in the COVID-19 model, thus the response to the scenarios is more pronounced in the first case (see Figure D.20). This is again related to the age-specific susceptibility and infectiousness in the COVID-19 model.

5.3.3 Burden of disease-related mortality

Although the overall attack rate is decreasing over time, the number of deaths per 1,000 people in the population is increasing, due to the shift in the age distribution of the infected population (see Figure D.21). Since the applied fatality rates are associated with substantial uncertainty, we limit the analysis of disease-related deaths to a comparison across time, age and demographic scenarios. Deaths attributable to COVID-19 are highly concentrated in the older adult population (see Figure 5.5, upper panel). Influenza-related deaths are also more pronounced in the older adult population, however, differences within the older age groups only reflect differential attack rates since the same IFR is applied to everyone aged 60 and older.

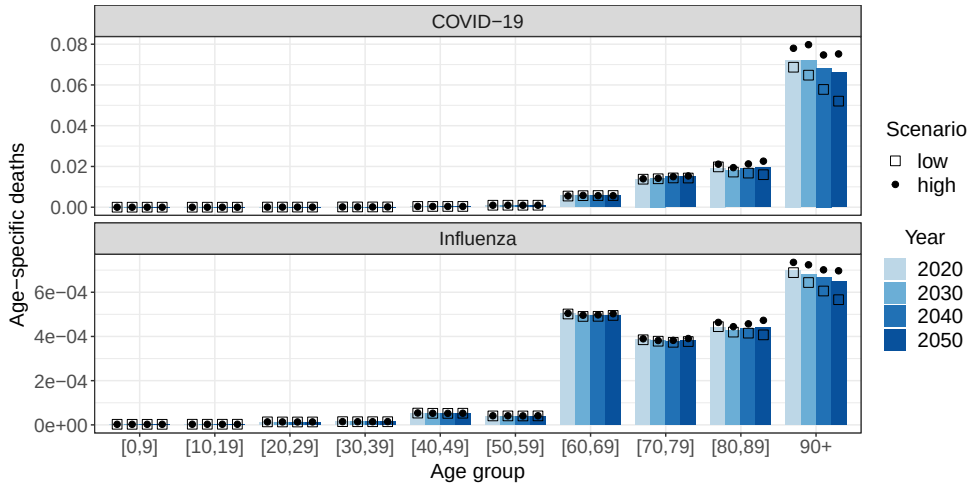

Figure 5.5. Mean age-specific disease-related death rate by year, model and demographic scenario. Note the different scales on y-axes.

The differences in fatalities between the demographic scenarios are induced by the aforementioned relationship between the proportion of infected older adults and the proportion of those living in LTCFs. Moreover, the applied IFRs for COVID-19 from Molenberghs et al. [77] are broken down by household type (i.e. LTCF vs. non-LTCF), hence the number of COVID-19 deaths in our simulation is more sensitive to changes in the population living in LTCFs.

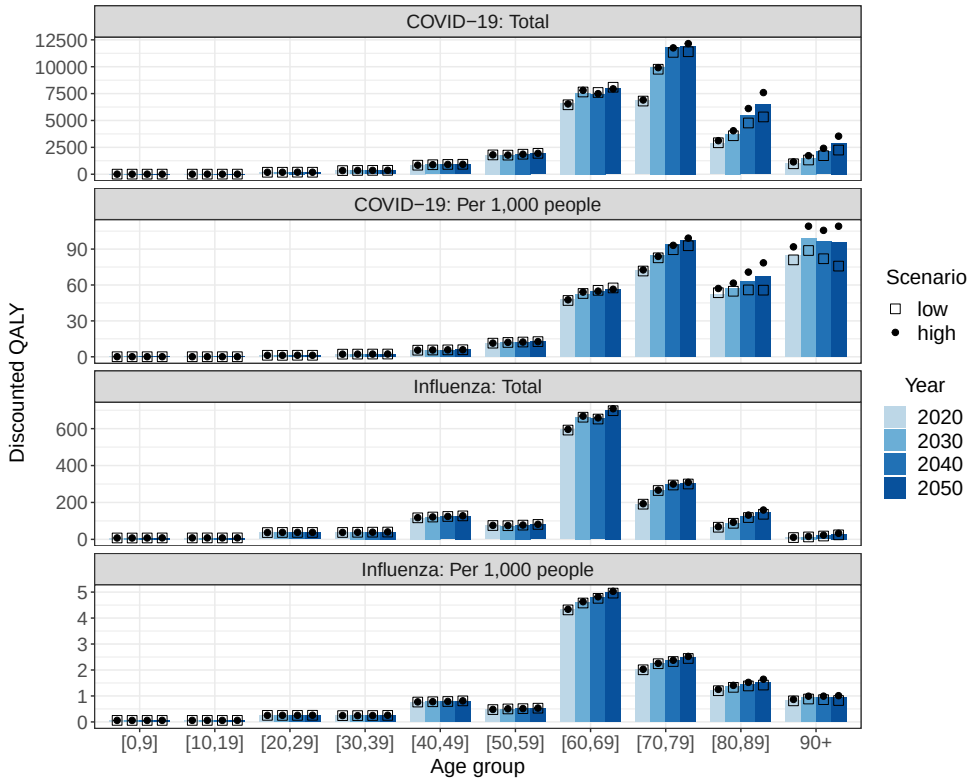

Figure 5.6. Mean age-specific QALY losses in total and per 1,000 people by year and demographic scenario.

Clearly, the average number and quality of years of life lost due to a premature death decrease with age. The middle-aged adults thus account for a larger share of the total QALY losses than of the fatalities (Figure 5.6 row one and three, note different Y-axis scales). However, the largest QALY losses in absolute values are seen in the 60-79 year olds for both COVID-19 and influenza. When taking the age distribution into account, the QALY losses become more pronounced in the oldest age groups (see Figure 5.6, row two and four). In both models, the total QALY losses are increasing over time and at a rate slightly higher than that of the increase in deaths (see Figure D.21).

Moreover, the age-specific QALY losses per 1,000 people are increasing in several age groups despite a stable or even decreasing disease-related death rate (see Figure 5.6), because life expectancy is increasing. In 2020, for example, the average life expectancy of a 75 year old is about 12 years, while it is expected to increase to 15 years by 2050 (see Figure D.22). However, the COVID-19 related QALY losses per 1,000 people in age group 90+ is decreasing between 2030 and 2050 for the demographic scenarios with a low or medium proportion of older adults living in LTCFs. The increasing life expectancy in this age group cannot compensate for the decrease in deaths associated with the changing living arrangements.

5.4 Discussion

Increasing life expectancy and persistently low fertility levels have led to old population age structures in most high-income countries, and population ageing is ongoing as the large generations born in the mid-twentieth century move into the older age categories [191]. Population ageing has potential implications for the burden of infectious diseases as the morbidity and mortality of many infections are concentrated in the older adult population [76, 226], as seen in the COVID-19 pandemic. The demographic microsimulation and two-level mixing model applied in our study allow to investigate the potential impact of population ageing on the transmission dynamics and burden of COVID-19 and influenza, while explicitly considering changes in the household structures, particularly among the older adults. Our focus on the future living arrangements in the older adult population is motivated by the disproportionate burden of the COVID-19 pandemic among LTCF residents [77, 86–88].

Our results suggest that population ageing on the one hand is associated with smaller total attack rates in COVID-19 and influenza epidemics, but on the other hand is causing a substantially larger disease burden of mortality, even if the proportion of older adults living in LTCFs were to decrease. Moreover, we find that not only the shift in the age distribution, but also the induced changes in the household structures are important to consider when assessing the potential impact of population ageing on the transmission and burden of respiratory infections.

Respiratory infections are predominantly caught by close contact with an infectious individual and transmission often takes place between household members [17, 43]. Since older adults in Belgium have few contacts and the majority live alone or with their partner, the number of occasions where they could acquire a respiratory disease is relatively low. A decrease in the overall attack rate of COVID-19 and influenza

is thus a logical consequence of population ageing. This relationship, however, is modified by the age pattern in the attack rates, which in turn is influenced by the susceptibility to and infectiousness upon infection as well as contact patterns.

In our simulation, COVID-19 attack rates were highest in the young and middle-aged adult population, while influenza incidence was highest in children and adolescents, which is in line with serological studies [227, 228]. Moreover, older adults made up a larger share of the infected population in the COVID-19 model than in the influenza model. Consequently, the increasing proportion of older people in the population led to a greater relative decline in the overall number of influenza infections, which was amplified by the decreasing size of nuclear families, as they play an important role in influenza transmission. The changing composition of nuclear families resulted from a decreasing fertility rate prior to 2020, which recovered slowly, but not fully, in the remaining simulation period, similar to observed and projected rates by Statbel and FPB [174]. Additionally, single-parent families became more prevalent. As a result of this, the average influenza incidence in children and the parental generation declined. This was less pronounced in the COVID-19 model, where the attack rate in the young and middle-aged adults were more or less unchanged due to substantial community transmission (i.e. at work-places).

Since older adults have the lowest number of community contacts [30], their risk of infection is highly dependent on their living arrangements. Older people in Belgium tend to either live in very small households (alone or with a partner) or in very large LTCFs. Several typical aspects of LTCFs, such as the size, shared meals, group activities, staff rotation, visitors, makes it easy for an infection to enter and spread rapidly [85, 128]. The variability in older adults' risks of infection is thus considerably larger than in any other age group. This implies that the future incidence in the older adult population is closely connected to changes in their household structures.

In the microsimulation, single-person households (e.g. a widow) in the age group 80+ are increasingly being replaced by two-person households (e.g. a couple), as the sex differential in mortality diminishes due to larger improvements in the life expectancy of males than in that of females. On the one hand, living with a partner instead of living alone increases the risk of infection as household transmission becomes a possibility. On the other hand, the probability of moving to a LTCF, which is associated with a substantially higher risk of infection, is markedly lower for elderly people living with a partner than for those living alone. However, the future mortality, health and living arrangements in the older adult population are associated with a large degree of

uncertainty [184, 186, 202–205]. Therefore, we presented three demographic scenarios pertaining to the proportion of older adults living in LTCFs relative to the proportion living alone.

The attack rates in the old age groups follow the proportion living in LTCFs and therefore differ substantially between the demographic scenarios. However, the sensitivity of the attack rates to household structures among older adults was larger for influenza than for COVID-19. This is again related to the different age patterns in disease transmission. The risk of acquiring COVID-19 remains relatively high for older non-LTCF residents because their susceptibility is high and the few contacts they do have will typically be with other old people, which are most likely to be symptomatic in case of infection and thereby more contagious. The influenza attack rate in older non-LTCF residents is markedly lower than other population groups, as children and adolescents are the main drivers of the spread and rarely live together with old people and generally have few contacts with them. Meanwhile, the attack rate for LTCF residents of COVID-19 as well as influenza are the highest in the population. Thus, the larger differential in the influenza attack rates between LTCF and non-LTCF residents imply that the overall attack rate of the older adults responds stronger, in relative terms, to changes in the living arrangements.

Although population ageing is associated with a decreasing proportion of infected people in the total population, disease-related deaths and QALY losses are increasing substantially. The lost QALYs increase faster than the deaths because a projected increase in life expectancy is accounted for in the QALY estimations. The speed at which the burden increases is influenced by the living arrangements among older adults, which can be considered a proxy for the health at old age. However, even in a scenario with a diminishing proportion of LTCF residents, the burden of disease-related mortality increases considerably in the whole simulation period.

We emphasise that our study is an investigation of the effects of population ageing on transmission dynamics and burden of disease-related mortality, and the results should not be interpreted as predictions. Moreover, our findings should be seen in the light of several limitations. First, we restricted our study to emerging infections by assuming that the initial population is fully susceptible and do not consider behavioural changes (e.g. changing contact patterns) during the outbreak, which would reduce the size of the simulated outbreaks. Additionally, age patterns of prior immunity or mitigation strategies in certain population groups (e.g. LTCFs, schools) may shift the age distribution of the infected population and thereby modify the impact of

population ageing. For example, some degree of pre-existing immunity to influenza A (H1N1)pdm09 was found in older adults, which may have resulted from exposure to H1N1 viruses earlier in life [227]. Nevertheless, we disregard these elements in order to obtain a clear understanding of the effects of population ageing alone.

Second, we assume that social contact patterns in the general population remain constant over time, since little is known about how contact patterns are affected by changing population structures. Nevertheless, the household contact patterns change along with the household composition as we generate new household networks in each time-step (i.e. day). The extrapolation of household contact networks for private households to LTCFs may be questionable due to the different structures, compositions and relations within the households. However, the large outbreaks among LTCF residents in our simulation reflect estimations of COVID-19 cases and the spread in LTCFs in Belgium prior to the implementation of mitigation measures [77, 229].

Finally, our estimates of disease burden are based on adjusting the QALYs lost due to deaths attributable to COVID-19 and influenza, but do not include QALY losses from morbidity due to non-fatal COVID-19 or influenza. Furthermore, our QALY estimates are associated with a considerable degree of uncertainty pertaining to the applied IFRs and the parameter settings in the QALY estimations. Moreover, we applied constant IFRs and parameters in the QALY estimations. The alternative parameter settings suggested in Briggs et al. [223] for the estimation of QALYs did not change the relationships we obtained. Nevertheless, the age-specific morbidity and mortality associated with respiratory diseases may change over time because of medical innovations and/or improved health at old age. Developments in healthy life expectancy, however, remain unclear [202–205]. We partially addressed this uncertainty with the demographic scenarios in our analysis.

5.5 Conclusions

Population ageing is associated with smaller outbreaks of emerging respiratory infections such as SARS-CoV-2 and novel influenza A virus. Nevertheless, the burden of mortality increases substantially, even if the population living in LTCFs, which face a high risk of infection and a fatal outcome, were to decrease. The variability in older adults' risks of infection is considerably larger than in any other age group, which is related to their living arrangements. Not only the shift in the age distribution, but also the induced changes in the household structures are important to consider when assessing the potential impact of population ageing on the transmission and burden

of emerging respiratory infections. Age patterns in epidemiological parameters may exacerbate or alleviate these relationships.

Chapter 6

The impact of demographic change on the epidemiology of varicella and herpes zoster: US population 1960-2020

This chapter is based on ongoing work: "Møgelmose, S., Hens, N. & Lewnard, J. The impact of demographic change on the epidemiology of varicella and herpes zoster: US population 1960-2020. [Working paper]."

6.1 Background

Infection with varicella zoster virus (VZV) can lead to varicella or chickenpox, a highly contagious and widespread childhood disease. Based on national seroprevalence studies, it has been estimated that prior to the introduction of a varicella vaccine in 1995, more than 95% of the US population acquired varicella before the age of 20 years [230]. VZV may reactivate later in life, typically in older adults or immunosuppressed individuals, and can cause herpes zoster (HZ), also known as shingles [74]. Hope-Simpson presented the hypothesis that the reactivation of VZV could be inhibited by re-exposure to VZV [231], also referred to as the *exogenous boosting hypothesis*. Consequently, the introduction of a varicella vaccine is expected to induce an increase in HZ cases due to reduced VZV circulation. The magnitude of exogenous boosting, however, is associated with a great deal of uncertainty [232, 233]. Furthermore, an increasing incidence of HZ has been observed in several countries prior to the introduction of vaccination against varicella [234, 235]. Modelling studies suggest

that reduced varicella circulation resulting from decreasing birth rates and population ageing has contributed to this [36, 57, 75]. However, little is known about the role population structures beyond age, such as households, played in this. In this study, we aim to explore the impact of demographic change on VZV transmission dynamics and the incidence of HZ. Specifically, we develop an age- and household-structured population reflecting that of the US population between 1960 and 2020. The model includes a disease process for the spread and reactivation of VZV similar to that of Melegaro et al. [36], which is informed with historical epidemiological data. We intend to investigate the past dynamics of varicella and HZ in the US, prior to and after the introduction of vaccination, at the individual, household and population level.

6.2 Methods

We apply a framework similar to that of the microsimulation presented in Chapter 3 to model the US population between 1960 and 2020. We base the initial population on a population sample from 1960. Over time, demographic events (fertility, mortality, migration and household transitions) and ageing take place and the population changes accordingly. The events take place in discrete time steps of one year, but each event is assigned a date within the given year. We thus allow an individual to experience multiple events within one year, but assume that birth events (own birth or giving birth) take place before household transitions, followed by migration events and finally death. Since we are modelling an observed time period, we calibrate the applied demographic rates by comparing the simulated population to sample populations from IPUMS. We specifically evaluate deviations in the age-specific household size distributions and positions and adjust the household transition probabilities accordingly. We extend the demographic model with a disease transmission process for VZV similar to that of Melegaro et al. [36].

6.2.1 Demographic data

We access cross-sectional samples of the US population between 1960 and 2020 in IPUMS USA and IPUMS CPS [236], which are databases containing harmonized microdata from the U.S. census and the U.S. labor force survey, respectively. Households are sampled in the surveys and individuals are sampled as parts of the households. IPUMS samples are stratified to some degree (e.g. geography, socio-economic factors) and sample weights are provided to ensure a representative sample population. The sample for 1960, which is used as the initial population, is from IPUMS USA

and contains 1% of the total population. Samples for 1970 to 2020, which are used for model calibration, are from IPUMS CPS, a monthly survey of about 100,000-150,000 individuals. We adjust the IPUMS CPS samples by a household-specific weight (*HWTFINL*) to obtain a representative sample population at the household level. We exclude institutional group quarters from all samples (e.g. correctional facilities and nursing homes). The age and household size distributions of the sample populations are compared to the observed distributions in Figure E.1 and Figure E.2 in Appendix E.

For each individual in a sample, we include the variables age, sex, household ID, ID of mother (*MOMLOC*), ID of father (*POPLOC*) and ID of partner (*SPLOC*). A birth date is assigned to each individual within the boundaries given by their age indicated on the census date. Unions are detected using the partner IDs, but unions between individuals living in different households are not detected. Parent-child links are based on the parental IDs and checked for compatibility (i.e. age difference between parent and child of at least 12 years). In case of incompatibility or missing parental IDs of individuals younger than 16 years of age, another adult household member is assigned as the parent. If one parental ID is missing but the known parent is in a union, the partner is assigned as the other parent of the child. Parent-child links between individuals in different households are not detected. Households only containing children younger than 16 years are removed. We use the partner and parent-child links to create a new variable; *household position*. This describes the position of an individual in a household in relation to the other household members. We define the household positions as follows:

Table 6.1. Categories of the variable household position.

Position	**Description**
Union	Individual living together with their partner (and potentially others, e.g. children)
Single	Living alone
Child	Individual (regardless of age) living in their parental household
Multigeneration	Oldest generation (union or single) in a multi-generational household, defined as households containing three or more generations
Single parent	Individual living together with their child(ren) (regardless of age), but without a partner
Other	Individual living in other household constellations

We also use the parent-child links to re-create the fertility trajectory of the female population of fertile age (i.e. 14-50). The frequency of the mother IDs is added to the corresponding female as the number of previous births (i.e. parity) and the birth dates of the children are used to estimate the birth intervals. The fertility trajectories will be biased to some degree as information on children that are not present in the household at the time of the census is missing.

6.2.2 Demographic modelling

As mentioned earlier, the initial population is based on the population sample for 1960. The population changes over time as individuals enter and leave the population as a result of fertility, mortality and migration. Additionally, individuals move between households or create new households. Each event is assigned a date within the time step (i.e. year) at random. In case an individual experiences multiple events within one time step, the dates are adjusted so that the order is respected (e.g. household transition before death). All demographic events are tracked in an event-log making it possible to re-create the population in smaller time steps.

Fertility

We create a fertility data set using the re-created birth trajectories of females of fertile age from each sample (parity and birth interval). Females are considered to have a birth event if they gave birth in the year prior to the census, as the birth trajectories are obtained retrospectively. Thus, the variable *year* in the fertility data is the year prior to the census. The household positions on the census date are also included in the data. Using the fertility data, we compute the probability of a birth by age, household position (simplified to union vs. non-union) and fertility trajectory (parity and length of birth interval). We apply linear interpolation for the years without a population sample.

The number of births in a time step is based on the observed age-specific fertility rates for the corresponding year [65]. For a given age group, the number of births is calculated by multiplying the age-specific fertility rate by the size of the female population of the given age in the middle of the time step. Each female in the risk set is assigned the birth probability corresponding to her age, household position and fertility trajectory. Conditional on the birth probabilities, we draw from the pool of females until the age-specific number of births has been reached. The newborns are added to the population and specifically to the household of the mother.

Migration

Due to the lack of historical as well as recent data on age and household profiles of emigrants and immigrants, we use the R package *DemoTools* to estimate the age-sex profile of the net-migrant population based on the Rogers-Castro migration model [237]. Observed net-migration rates are obtained from the UN Statistics Division [65]. The number of migrants thus changes over time, but the age-sex distribution remains constant as shown in Figure 6.1.

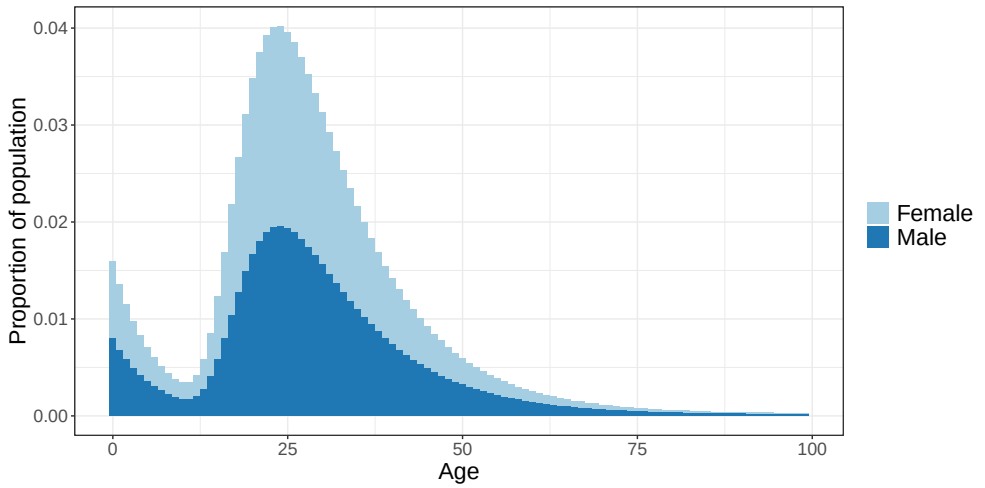

Figure 6.1. Age-sex distribution of net-migrant population.

The net-migration rates between 1960 and 2020 are positive, meaning that the number of immigrants is higher than the number of emigrants. Consequently, migrants are created according to the assumed age-sex distribution and added to the existing population. The migrant population is divided into households. Individuals younger than 18 years are considered to be children and are assigned a mother. The mother-child matching is based on the age-specific fertility rates observed in the US in the year the child was born. Consequently, we assume the fertility trajectories in the migrant population to be similar to that of the native population, despite being unlikely, since detailed data is lacking. The mother is matched with a male individual in the migrant population according to the age difference observed for couples marrying in the US in the given year. Two-thirds of female migrants without a child are also assigned a partner. This implies that the households in the native population are not affected by migration. However, this is a simplification as migration often is associated with family reunification and union formation.

Mortality

We apply the observed age- and sex-specific mortality rates for 1960-2020 obtained from the United Nations Statistics Division [65] and convert these to death probabilities. Individuals with a death probability higher than a random number drawn between zero and one are assumed to die within the time step and are removed.

Household transitions

Fertility, mortality and migration affect the composition of households in the population. Additionally, individuals can move between households or create new households independently of other demographic events. The applicable household transitions depend on the household position. Some transitions are considered non-applicable because they are uncommon (e.g. union to other) or to limit model complexity (e.g. transitions between unions is not applicable). Moreover, some transitions are modelled as a population-level process (i.e. union formation). The household transitions are shown in Table 6.2, where white cells are applicable transitions, black cells are non-applicable and grey cells indicate processes carried out at the population level. Finally, cells with *I* indicate indirect transitions resulting from other demographic events (e.g. single to single parent).

Table 6.2. Household transitions. Black=not applicable, white=applicable, I=indirect transitions, grey= population-level transition

From \ **To**	Union	Child	Single	Single parent	Multigen.	Other
Union						
Child				*I*		
Single				*I*		
Single parent			*I*			
Multigen.						
Other				*I*		

Union dissolution between 1960 and 1979 is based on estimated age-specific divorce rates of females [238, 239]. For 1983 to 2020, we estimate the age-specific dissolution rates for females in the Survey of Income and Program Participation (SIPP) [240]. SIPP is a nationally representative longitudinal household-based survey providing comprehensive information at the individual- and household-level, including information on relations between household members. Before 1996, only married couples are detected as unions in the SIPP data, while cohabiting couples are included from 1996 onwards. The rise in cohabitation in the US since the mid-20$^{\text{th}}$ Century [241] implies

that we are increasingly underestimating union dissolution up to 1996. In case of no available data for a given year, data from the closest time period is applied. In case of union dissolution in a household with children present, the children stay with the mother, and the father moves to a single-person household.

After the procedure of union dissolution has been carried out, we calculate the age-sex specific proportion living in a union and compare it to the observed proportions according to the census data from IPUMS USA and IPUMS CPS for the given year. In case the age-sex-specific proportion of unions in the simulation is higher than the one observed in the sample population of the given year, union dissolutions take place and the proportion living in a union is re-calculated. This re-calculation is necessary because some union dissolution may imply too few unions in the age group of the opposite sex. Finally, the number of males and females entering a union is calculated and individuals with the corresponding age and with household position *single, child, single parent* or *other* are drawn at random. The candidates are matched based on the age differences observed in the census in the given year. In case of an unequal number of females and males, no union formation takes place for the remaining individuals.

A transition to *child* does not mean that the individual moves into the actual parental household. Instead, we draw an age group (five-year age groups) from the age-specific fertility rates observed in the birth year of the given individual and assign a female with the appropriate age difference the role as mother and her partner, if applicable, as the father. In this way, we do not have to keep track of kinship links between households, which is computationally time-consuming and complicates sampling procedures in the disease transmission model.

Multi-generational households are created by moving individuals (union, single or single parent) to a household containing one or more children (i.e. household position not age). The individuals moving in to the new household should be the oldest generation. We therefore use the same matching procedure as for individuals moving to a parental household to secure an appropriate age difference between the generations. The transition to the oldest generation in a multi-generational household is only applicable for individuals in households of size one or two.

New households are created for the individuals moving to a single-person household and for a proportion of the individuals moving to the household position *other*. The rest move to existing *other* households according to the household size distribution observed in the samples.

6.2.3 Disease transmission model

We model VZV similarly to Melegaro et al. [36]. Maternal immunity (M) is assumed to last for six months after which individuals become susceptible to varicella (S). A proportion of individuals vaccinated against varicella undergo vaccine failure and remain susceptible to breakthrough varicella (S^*). Susceptible individuals (S and S^*) are exposed to a force of infection with two levels of mixing; in the household and in the community:

$$\lambda(a,t) = \beta_c \sum_{a'} C_{c,aa'}(t)(y_{c,a'}(t) + \alpha y^*_{c,a'}(t)) + \beta_h(y_h(t) + \alpha y^*_h(t)) \tag{6.1}$$

The subscripts reflect community (c) and household (h). The transmission parameters, β_h and β_c, differ for contacts between household members and contacts in the general community between non-household members. We assume that all household members come into contact with each other daily, while contacts in the community (i.e. all contacts outside the household, including at work, school and leisure activities) are based on a contact matrix ($C_{c,aa'}$), estimated for the US population [189]. The contact matrix contains the average number of daily contacts an individual of age a has with individuals of age a'. It is adjusted to changes in the demographic composition in accordance with method two in Arregui et al. [190]. The contact matrices for 1960 and 2010 are shown in Figure 6.2. The number of contacts is multiplied by the proportion of people in the given age group with natural ($y_{c,a'}$) and breakthrough varicella ($y^*_{c,a'}$). Individuals with breakthrough varicella are assumed to be half as contagious as those with natural varicella, hence $\alpha = 0.5$.

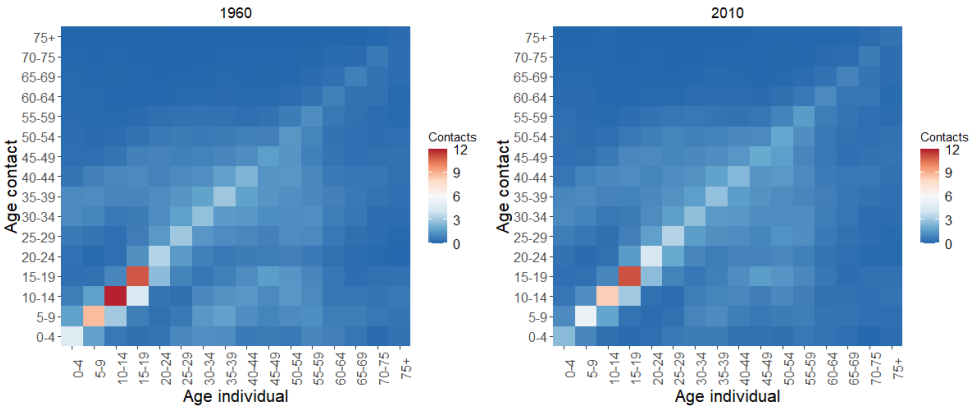

Figure 6.2. Social contact matrices adjusted to US population composition in 1960 (left) and 2010 (right) [189, 190].

Individuals infected with varicella (I and I^*) recover and gain immunity after three weeks and are then susceptible to HZ (SZ and SZ^*). We apply the model PI from Melegaro et al. [36] which assumes progressive immunity, thus susceptible individuals are assumed to develop HZ at a rate ρ, which depends on their age, the number of VZV boosting events (i) and the time elapsed since the last VZV exposure (τ). The VZV reactivation rate is calculated as follows [36]:

$$\rho_i(a,\tau) = \rho_0 q^{(i-1)^2} e^{\theta_a(a-a_0)} e^{\theta_\tau \tau} \tag{6.2}$$

where ρ_0 is the risk of developing HZ immediately after recovering from the varicella infection, the parameter q shapes the reduction in the risk of HZ due to re-exposure to VZV, while θ_a and θ_τ shape the increasing risk of HZ with age (a, a_0=45 years) and time since last exposure to VZV (τ), respectively. Individuals successfully immunised against varicella as well as individuals becoming susceptible to HZ after experiencing breakthrough varicella are assumed to develop HZ at the reduced rate $\chi\rho_i(a,\tau)$. The boosting events (i.e. re-exposure to VZV of individuals susceptible to HZ) take place at the rate $z\lambda(a,t)$, where $z \in [0,1]$. We assume a HZ vaccination coverage of c_{HZ}, and a HZ vaccine efficacy of p_{HZ}. After an individual has developed HZ or has been successfully immunised against HZ, life-long immunity to VZV reactivation is assumed.

6.2.4 Endemicity and re-sampling

Since we are lacking epidemiological data for VZV in the 1960s, we apply a simulation technique to obtain a population in which varicella is endemic. We first obtain a stable population by running the demographic model for 100 years with fixed demographic rates corresponding to those of 1960. In the stable population, we initiate VZV transmission by changing the disease state of ten randomly chosen individuals to $infected$. The demographic rates remain constant and the disease transmission continues until varicella is endemic in the population. At this stage, the population composition no longer reflects the initial population of 1960 because the demographic rates have been fixed for a long period of time (see Figure E.3 in Appendix E). For that reason, we return to the nonstable sample population of 1960 and assign each individual a disease state and history of exposure by sampling from the stable population with endemic varicella conditionally on the individual's age and household composition. The resulting population is used as a starting point in the simulation of VZV transmission in the dynamic population resembling that of the US from 1960 to

2020. In the dynamic simulation, we explore different transmission parameters and compare the simulation results to epidemiological data from the 1970s, 1980s and 1990s [242, 243]

6.3 Preliminary results

6.3.1 Population dynamics

The age distribution changes dramatically in the simulation period. In 1960, the population contains a large proportion of children as seen in Figure 6.3, but as fertility declines in the following decades (see Figure E.4 in Appendix E), the subsequent generations are substantially smaller, and by 2020, an old population age structure has emerged.

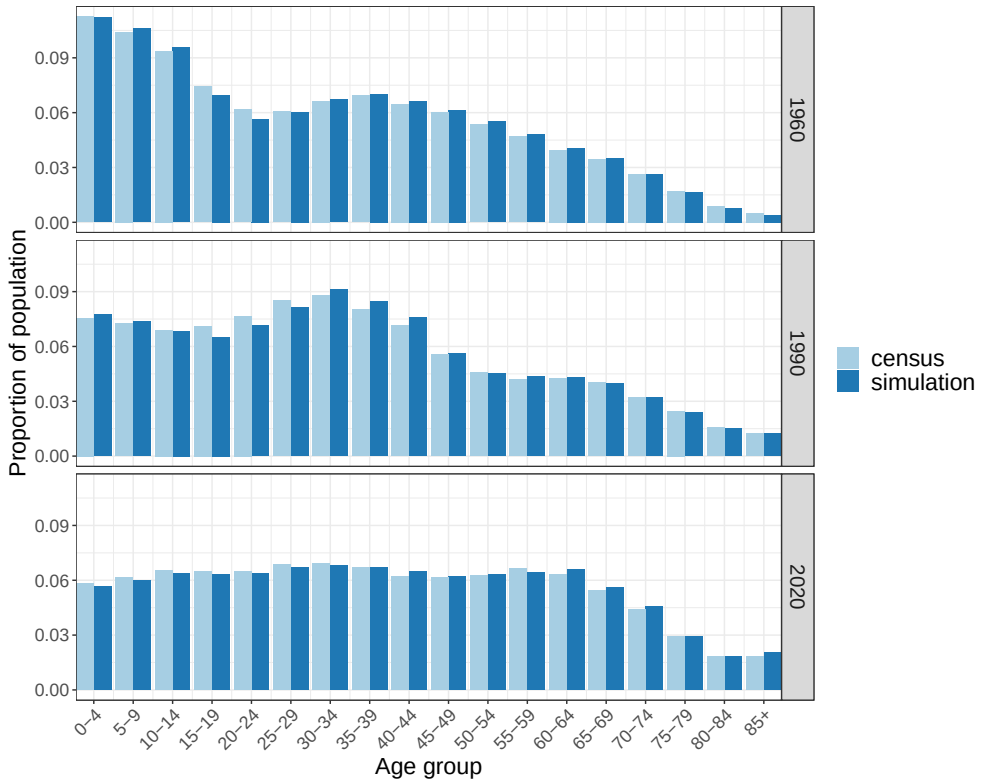

Figure 6.3. Age distribution in 1960, 1990 and 2020 in the simulated population (dark bars) and according to the US Census Bureau (light bars).

The household size distribution also changes substantially as a result of declining fertility and mortality rates (see Figure 6.4). As fewer children are born and life expectancy increases, large households are increasingly being replaced by small households. The age and household structures in the simulated population (dark bars) only deviate slightly from those reported by the US Census Bureau (light bars). The deviations remain reasonable when disaggregating the household size distributions by age groups (see Figure E.5-E.8 in Appendix E).

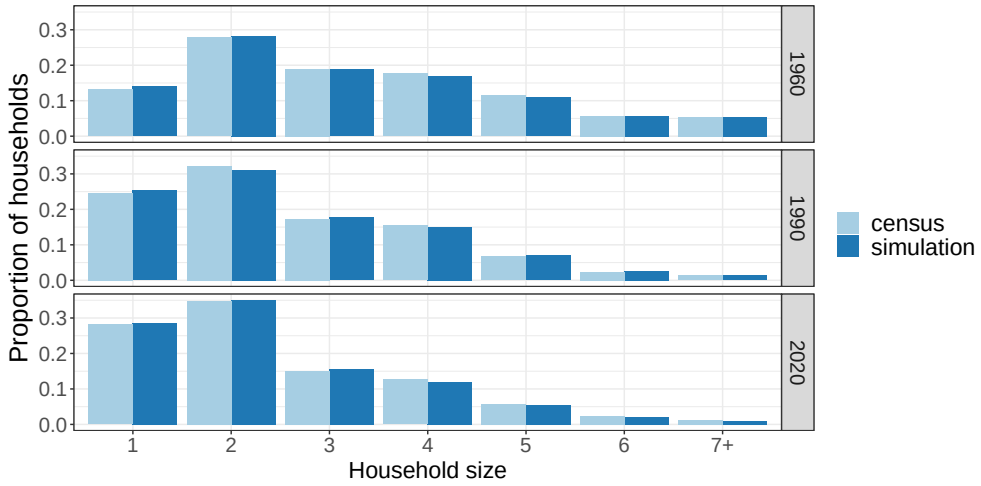

Figure 6.4. Household size distribution in 1960, 1990 and 2020 in the simulated population (dark bars) and according to the US Census Bureau (light bars).

6.4 Next steps

At this moment, we are exploring a range of transmission parameters shaping the disease transmission within and between households in the dynamic population, with the purpose of obtaining simulation results reflecting observed measures, including age-specific incidence and seroprevalence [242, 243]. After the calibration of the simulation model, epidemiological outcomes will be computed and analysed, including attack rates by age and household composition. The contribution of the changing population structures will be considered explicitly in the analysis.

Chapter 7

Discussion

In recent decades, severe outbreaks of emerging and re-emerging infectious diseases (e.g. HIV, MERS, SARS, COVID-19, monkeypox), aided by increasing global connectivity, have demonstrated that infectious diseases remain a public health threat, also to populations in high-income countries. Moreover, outbreaks of vaccine-preventable diseases such as measles and mumps are re-emerging due to suboptimal vaccine uptake, waning vaccine-induced immunity and unvaccinated clusters in populations with otherwise high vaccine coverage [244–248].

Infectious disease modelling provides an important tool for analysing and predicting the spread of infectious diseases and to evaluate prevention and control measures. Population structures contribute to the heterogeneity in the spread and burden of many infectious diseases. Age and household structures have been given particular attention when modelling infectious diseases spread through close contact [35, 41–47]. Social contact patterns influencing the exposure to an infection vary considerable by age, as well as many epidemiological parameters [25, 30]. Moreover, household members have distinct contact patterns and often belong to different generations and subgroups outside the household, making it a unique entity for disease transmission [31, 48–50].

A population's age and household structures evolve over time as they are influenced by past trends in fertility, mortality and migration. Populations in many high-income countries have an old age structure with a low mean household size, resulting from sustained below-replacement fertility and rising life expectancy. Moreover, population ageing is currently accelerating in several countries as the large generations born in the

mid-20$^{\text{th}}$ century become older [66]. Nevertheless, the relationship between population structures, the changes therein and the transmission and burden of infectious diseases has not yet been fully unravelled.

The overarching aim of this dissertation was to explore and further improve infectious disease models incorporating a dynamic host population with the purpose of investigating the impact of demographic structures and changes on the transmission and burden of infectious diseases transmitted through close contact. Specifically, we aimed to address the following research questions: (i) what is the relationship between age and household structures at the individual and population level and how are these shaped by demographic processes and changes therein? (ii) to what degree are transmission dynamics at the individual, household and population level shaped by age and household structures? (iii) how are population age and household structures expected to evolve over time and how are the demographic changes expected to affect the transmission dynamics and disease burden?

We carried out interdisciplinary research combining demographic modelling with infectious diseases modelling, which was presented in Chapters 2 to 6. Next, we provide an overview of the main findings and consider them in a broader perspective. This is followed by a discussion of the limitations of the dissertation, prospects for future research and finally a conclusion.

7.1 Main findings

Efforts have already been made to assess the impact of demographic change on the epidemiology and burden of infectious diseases (e.g. [36, 44, 47, 55–60]). Various demographic methods and assumptions were applied in the existing literature, but often with an insufficient discussion of the choice of method and the implications thereof for the realism of the population structures. Consequently, we systematically reviewed the demographic methods, assumptions and data used to incorporate a dynamic host population in models of disease transmission, which was presented in Chapter 2.

Population-level modelling was more common than individual-level modelling, as to be expected since the general use of computer simulations like IBMs only have become feasible in recent decades with improvements in hardware performance. Moreover, a host population with a realistic and evolving age-sex composition is typically less cumbersome to model at the population level than with an IBM. The advantages

of individual-level modelling emerge when additional heterogeneity is required in the population or transmission process, for example household structures, schools, local interactions or health trajectories [26]. Consequently, several IBMs incorporated additional population structures, including households. However, the flexibility of IBMs was rarely exploited to improve the level of detail in the demographic processes, with potential implications for the population structures, the composition of households in particular.

Households are not only a unique entity from a disease transmission perspective, but also from a demographic perspective. Many demographic, social and economic processes involve decisions that are often made at the household level, such as childbearing, living arrangements, labour force participation and migration [249]. Consequently, household compositions are affected by changes in these processes, which may affect the disease transmission dynamics.

With the objective of enhancing the demographic modelling of age- and household-structured host populations, we presented a demographic microsimulation in Chapter 3, developed for applications in infectious disease modelling. Access to individual-level longitudinal data from Belgian census and population registers made it possible to define relationships between household members and to re-create kinship networks and fertility trajectories. Moreover, less common household types and constellations were represented, including LTCFs. Finally, past individual events, household positions and/or household types were incorporated in the demographic processes. The model was used to simulate the Belgian population from 2011 to 2050.

The microsimulation was developed in a flexible manner to allow for sensitivity or scenario analyses, as future developments in demographic processes are associated with substantial uncertainty [123]. Nevertheless, the changing population structures in the simulation are partly a result of fertility, mortality and migration schedules observed in Belgium prior to the census in 2011. For example, past trends in fertility postponement implied that parents in the simulation were increasingly older when the last child left the parental household.

The demographic events in the simulated population were tracked, making it easy to re-create the population in disease modelling applications. The first application was presented in Chapter 4, where we investigated how the spread of an emerging infectious disease was affected by age and household structures in the host population and how these relationships were influenced by demographic change. We illustrated a strong relationship between age and household structures at the individual and

population level explained through an investigation of the underlying demographic processes. The thorough understanding of these relationships and processes provided the basis for explaining the demographic changes in the simulated population. The age and household structures had an impact on the disease transmission dynamics, but the magnitude of the relationship depended on epidemiological heterogeneity in the population. Moreover, not only the size of households but also their compositions, which are influenced by the timing of demographic events, were crucial for explaining how the infection spread at the individual, household and population level.

Drawing on our understanding of how population structures and changes therein shape the spread of an emerging infectious disease, we presented a second application in Chapter 5, with the objective of investigating how population ageing affects the burden of respiratory infections. Older adults face an increased susceptibility to many infections and a higher risk of a severe outcome in case of disease due to the progressive deterioration of immune functions with age [76]. This was seen in the COVID-19 pandemic, which had a disproportionate impact on the older adult population and those living in LTCFs in particular [77, 78].

We developed a model for disease transmission resembling the spread of SARS-CoV-2 and pandemic influenza, which was combined with the demographic microsimulation. We focused on the living arrangements in the elderly population and how they evolve in an ageing population, as well as the implications thereof for the spread and burden of infectious diseases. Specifically, we considered different scenarios for the proportion of older adults living in LTCFs. Unique features of LTCFs make them an optimal environment for rapid spread of infectious diseases, and the health trajectory of the residents often make them more vulnerable to infection [85, 128–130]. Consequently, this relatively small population group, which is often disregarded in disease modelling, faced a markedly higher risk of infection in our simulations and accounted for a substantial share of the disease burden, reflecting observations made in the COVID-19 pandemic [86–88]. The burden of future epidemics in our simulation increased with the ageing of the population, even in case of a decreasing proportion of LTCF residents reflecting a scenario of improving health at older ages.

In Chapter 4 and 5, we only considered epidemic settings, as the emergence of SARS-CoV-2 and the related COVID-19 pandemic emphasised the need for an improved understanding of the role of population structures in the spread and burden of emerging infectious diseases. However, a dynamic microsimulation of an age- and household-structured host population is also useful for evaluating the long-term dynamics of

endemic infectious diseases and the effectiveness of immunisation programmes. In Chapter 6, we presented ongoing work involving a demographic microsimulation for the US population from 1960 to 2020, which will feed into an IBM for VZV and HZ. The demographic methods were similar to those used to develop the microsimulation in Chapter 3, but with several modifications as the data and its granularity differed. Moreover, we considered the US population in a time period in the past for which observed demographic data was available. This facilitated the use of re-sampling techniques to secure resemblance between the simulated and observed populations.

Existing studies have looked into the impact of demographic change on the epidemiology of VZV and HZ [36, 57, 75, 90, 250], however, population structures beyond age, such as households, are rarely considered [171]. Declining fertility rates since the 1960s has led to a changing age composition of the US population, as well as a declining mean household size. These trends are reproduced in the microsimulation, allowing for an investigating of the implications of the demographic changes for the epidemiology of VZV and HZ at the individual, household and population level.

7.2 Limitations and future work

With the research presented in this dissertation, we attempted to improve the understanding of highly complex demographic and epidemiological phenomena by simplifying them using statistical and computational modelling. The formulation of a model involves the trade-off between accuracy, transparency and flexibility, for which the appropriate balance is determined by the research question and purpose of the model, as well as the feasibility (e.g. data availability) [17]. With this in mind, we take a critical look at our choices of modelling, discuss the limitations they impose and consider directions for future interdisciplinary research in demography and infectious disease epidemiology.

Based on the systematic review, we found individual-level modelling to be the best framework for a dynamic host population with age and household structures. The accuracy of our demographic microsimulation for Belgium, however, is still limited by the amount of heterogeneity incorporated in the population and in the demographic processes. Fertility, mortality and migration schedules are influenced by a wide range of demographic, social, cultural and environmental factors that are disregarded in the microsimulation. The flexibility of individual-level modelling makes it possible to include such factors, if the granularity of the demographic data allows it. The mechanisms driving demographic change can thus be modelled explicitly. The model

transparency, however, decreases with the complexity. Moreover, the developed microsimulation is already rather advanced considering the context of infectious disease modelling, and several demographic and epidemiological relationships are still to be explored before further model expansions would be relevant. This includes investigations of the role of kinship networks in the spread of an infection, as these may add to the clustering of contacts. Moreover, more detailed analyses of disease transmission in LTCFs remain to be carried out using the current microsimulation model.

Nonetheless, a description of general health (e.g. chronic conditions, frailty) in the population, particularly at old age, would be a significant improvement of the microsimulation. So far, we considered living arrangements (i.e. LTCF, union, non-union) to be a proxy for the health of older adults, but it obviously does not provide us with the complete picture. Chronic health conditions have been found to be associated with increased risks of morbidity and mortality due to COVID-19 and other infectious diseases [165, 166, 251–253]. Moreover, population ageing is accompanied by increased rates of chronic diseases, emphasising the need for further research of the interface between demographic change, chronic diseases and infectious disease transmission [254–256].

Another potential improvement of the microsimulation pertains to the uncertainty of projections. Since future population structures are associated with uncertainty, it is relevant to consider how sensitive epidemiological outcomes are to variability in the demographic processes. We partly addressed this in the microsimulation by the use of demographic scenarios. Nevertheless, it would be worthwhile to consider a wider range of scenarios involving all the included processes. The scenarios are preferably theoretically founded or motivated by official population forecast (scenario-based or probabilistic) to ensure realistic parameter ranges. Alternatively, fully probabilistic projections for fertility, mortality and migration could be incorporated [257, 258], which have the advantage over scenario-based forecasts of attaching a probability to the range of possible outcomes [259].

Official population forecasts, however, are typically modelled at the population level, providing no information on how aggregate trends emerge from changes in individual-level characteristics, behaviours and interactions. By incorporating the factors and mechanisms driving the changes in mortality, fertility and migration (e.g. health, education level, family reunification patterns), assumptions about future aggregate trends in the demographic processes (e.g. TFR) become obsolete. Instead, theoretically and/or empirically founded relationships can be used to describe individual

behaviour (e.g. educational participation and childbearing). Assumptions will still be needed to describe how these relationships are expected to evolve over time, but the uncertainty can now be ascribed to the mechanisms underlying the demographic process. Nevertheless, it would require a substantial expansion of our microsimulation to achieve this, and the gains thereof would for many applications in infectious disease modelling probably not be considered sufficient.

Not only the demographic microsimulation, but also the presented disease transmission models face several limitations and possible improvements. In the transmission process, we make a distinction between household mixing and community mixing. While the first is rather detailed as we draw on work by Krivitsky et al. [150], allowing us to base the household contact networks on empirical data, the latter has potential for more nuance since we consider the community as a whole. Moreover, it is not evident whether the generated household contact networks should be extrapolated to unique living arrangements like LTCFs, which also vary considerably in size, demographics of residents, level of care etc. [84, 260]. Further research is needed to explore social contact patterns in less common settings like LTCFs. Additionally, we are lacking an understanding of how social contact patterns are influenced by demographic change. For ageing populations, for example, it remains unknown whether contact with older adults becomes more likely. In the presented studies, we assume that contact patterns within households as well as in the community remain constant over time. Methods for adjusting social contact matrices to changing demographics have been proposed, but these are not based on empirical evidence [189, 190]. A comparison of two social contact surveys carried out in Belgium suggests stable age-specific contact rates, but the surveys were only five years apart, implying limited demographic change [167]. Future contact surveys would therefore be valuable.

In the epidemic settings, we did not consider any behavioural changes (e.g. reduced social contact) during the outbreaks, which is an unlikely assumption. Nevertheless, we deemed it a necessary and useful assumption as it allowed us to investigate the role of demographic change in isolation. Moreover, our aims did not involve the accuracy of epidemic predictions or evaluations of intervention strategies that require implementation of behavioural change. Nevertheless, the disease transmission process in the epidemic settings were constructed to allow for temporal changes in the community contacts. Hence, this can be explored if deemed necessary.

In future work, we will make use of demographic microsimulation to investigate the impact of demographic change on the long-term transmission dynamics of endemic infectious diseases. Specifically, we presented ongoing work on VZV and HZ in the US population from 1960 to 2020. The main limitations of the study pertains to the availability of historical epidemiological and demographic data. Moreover, the procedure for obtaining an endemic equilibrium involves re-sampling of households, which may cause a lack of resemblance between the observed and modelled populations. The microsimulation developed for the Belgian population has not yet been combined with transmission models for endemic infections. However, this could be obtained using the methods from the ongoing VZV study for the US. In that way, the relationship between demography and infectious disease epidemiology can be studied for a wider range of pathogens.

7.3 Conclusion

Population age and household structures are closely connected. The composition of a household evolves as one or more of the household members experience a demographic event (e.g. death, union dissolution), which often has an age-related likelihood of occurrence. Thus, past fertility, mortality and migration schedules not only determine the population age structure, but are also some of the factors influencing the size and composition of households in the population. Individual-level modelling provides a flexible and useful framework to consider the relationship between population age and household structures, as well as how changes therein emerge from demographic processes.

Households comprise a unique entity in demography but also in infectious disease transmission. We investigated how the relationship between age and household structures shape the spread of an emerging infectious disease in a population with an old age structure. In addition to household size, differential household compositions helped explaining the disease transmission dynamics at the individual and household level, in the adult population in particular.

The Belgian population is ageing, which is expected to continue in the coming decades, as a result of persistently low fertility levels and rising life expectancy. On the one hand, population ageing is associated with smaller outbreaks of emerging infectious diseases like COVID-19 and pandemic influenza, but on the other hand, it is causing a substantially larger disease burden. Future developments in the health, living arrangements and social contact patterns in the older adult population may exacerbate

or mitigate this relationship.

Summary

Some population groups are more likely to acquire an infection or to experience a severe outcome in case of disease due to risk factors that may not be randomly distributed in the population. Some of these factors are related to demographic characteristics and structures (e.g. age, sex, household composition), which typically are incorporated in the host population in models of infectious disease transmission, though often in a highly simplified manner. Demographic structures, however, result from complex demographic processes that tend to change over time. In the context of infectious disease epidemiology, it is not well understood how these underlying processes shape current and future population structures with relevance for the transmission and burden of infectious diseases. For that reason, the aim of this dissertation was to explore and improve infectious disease models with dynamic host populations with the purpose of investigating the impact of demographic structures and changes on the transmission and burden of infectious diseases transmitted through close contact.

To create an overview of the existing literature, we first carried out a systematic review of the demographic methods and data used to incorporate dynamic host populations in models of infectious disease transmission. We found that population-level modelling was more common than individual-based modelling. The advantages of IBMs emerge when heterogeneity beyond age and sex, for example household structure, is required in the population or transmission process. The flexibility of IBMs, however, was rarely used to advance the demographic modelling of the host population. With the advantages and limitations of the existing literature in mind, we developed a demographic microsimulation for an age- and household-structured population, tailored for applications in infectious disease modelling. We specifically simulated the Belgian

population from 2011 to 2050 and considered the demographic processes of fertility, mortality, migration and household transitions. The microsimulation was extended by a disease transmission model to investigate how the spread of an emerging infectious disease was affected by age and household structures in the host population and how these relationships were influenced by demographic change. The age and household structures had an impact on the disease transmission dynamics, but the magnitude of the relationship depended on epidemiological heterogeneity in the population. Moreover, the size and composition of households were crucial for explaining how the infection spread at the individual, household and population level.

In a second application of the microsimulation, we investigated how population ageing affects the mortality burden of respiratory infections. The disease transmission model was modified to resemble the spread of SARS-CoV-2 and novel influenza A virus. We focused on the living arrangements in the older adult population, as the COVID-19 pandemic, for example, has had a disproportionate impact on those living in LTCFs. Similarly, we found that this relatively small population group, which is often disregarded in infectious disease modelling, faced a markedly higher risk of infection in our simulations and accounted for a substantial share of the burden of mortality associated with the respiratory infections. The burden of future epidemics increased with the ageing of the population, but the magnitude of this relationship depended on the living arrangements and general health in the older adult population.

Dynamic microsimulation of an age- and household-structured host population is also useful for evaluating the long-term dynamics of endemic infectious diseases and the effectiveness of immunisation programmes. We therefore presented ongoing work involving a demographic microsimulation for the US population from 1960 to 2020, which was extended by a disease transmission model for VZV and HZ. The demographic methods were similar to those used to develop the microsimulation for Belgium, but with several modifications, as the data and its granularity differed. The host population, however, was modelled with age and household structures, as household transmission rarely has been explored explicitly in the existing literature. Declining fertility rates since the 1960s has led to a changing age composition of the US population, as well as a declining mean household size. These trends are reproduced in the microsimulation, allowing for an investigating of the implications of the demographic changes for the epidemiology of VZV and HZ at the individual, household and population level.

Samenvatting

Sommige bevolkingsgroepen hebben een grotere kans om een infectie op te lopen of om een ernstige afloop te ervaren in geval van ziekte als gevolg van risicofactoren die mogelijk niet willekeurig over de bevolking verdeeld zijn. Sommige van deze factoren houden verband met demografische kenmerken en structuren (bv. leeftijd, geslacht, samenstelling van het huishouden), die typisch worden opgenomen in modellen van infectieziekten, zij het vaak op een sterk vereenvoudigde manier. Demografische structuren zijn echter het resultaat van complexe demografische processen die in de loop van de tijd veranderen. Er is onvoldoende begrip over de manier warop deze onderliggende processen vorm geven aan huidige en toekomstige bevolkingsstructuren die relevant zijn voor de transmissie en de last van infectieziekten. Het doel van dit proefschrift is daarom tweeledig. De eerste doelstelling is het verkennen en verbeteren van modellen van infectieziekten met dynamische gastheerpopulaties. Daarnaast wordt onderzocht wat de invloed van demografische structuren en veranderingen is op de transmissie en last van infectieziekten die via nauw contact worden overgedragen.

Om een overzicht te krijgen van de bestaande literatuur, hebben we eerst een systematische review uitgevoerd van de demografische methoden en gegevens die gebruikt worden om dynamische gastheerpopulaties op te nemen in modellen van infectieziekten. We ontdekten dat modellering op populatieniveau gebruikelijker was dan modellering op individueel niveau. De voordelen van IBM's komen naar voren wanneer heterogeniteit naast leeftijd en geslacht vereist is in de populatie of het transmissieproces. Dit is bijvoorbeeld het geval bij huishoudens. Deze flexibiliteit van IBM's werd echter zelden gebruikt om de demografische modellering van de gastpopulatie te bevorderen. Met de voordelen en beperkingen van de bestaande literatuur in gedachten, ontwikkelden we een demografische microsimulatie voor een leeftijds- en huishoudensgestructureerde

populatie, op maat gemaakt voor toepassingen in de modellering van infectieziekten. We simuleren specifiek de Belgische bevolking van 2011 tot 2050 en beschouwen de demografische processen van vruchtbaarheid, sterfte, migratie en huishoudenstransities. De microsimulatie werd uitgebreid met een transmissiemodel om te onderzoeken hoe de verspreiding van een opkomende infectieziekte werd beïnvloed door leeftijds- en huishoudenstructuren in de gastpopulatie en hoe deze relaties werden beïnvloed door demografische veranderingen. De leeftijds- en huishoudenstructuren hebben een impact op de transmissiedynamiek, maar de grootte van de relatie hangt af van epidemiologische heterogeniteit in de populatie. Bovendien waren de grootte en samenstelling van huishoudens van cruciaal belang om te verklaren hoe de infectie zich verspreidde op individueel, huishoud- en populatieniveau.

In een tweede toepassing van de microsimulatie onderzochten we hoe de vergrijzing van de bevolking de mortaliteitslast van respiratoire infecties beïnvloedt. Het transmissiemodel werd aangepast om verspreiding van SARS-CoV-2 en een nieuwe variant van het influenza A-virus te simuleren. We richtten ons op de samenstelling van huishoudens in de populatie van oudere volwassenen, omdat de COVID-19 pandemie bijvoorbeeld een onevenredig grote impact had op degenen die in zorginstellingen wonen. Ook ontdekten we dat deze relatief kleine bevolkingsgroep, die vaak buiten beschouwing wordt gelaten in modellen voor infectieziekten, een duidelijk hoger infectierisico liep in onze simulaties en verantwoordelijk was voor een aanzienlijk deel van de mortaliteitslast gerelateerd aan respiratoire infecties. De last van toekomstige epidemieën nam toe met de vergrijzing van de bevolking, maar de omvang van deze samenhang hing af van de samenstelling van huishoudens in de oudere volwassen bevolking.

Dynamische microsimulatie van een gastpopulatie met leeftijds- en huishoudensstructuur is ook nuttig voor het evalueren van de langetermijndynamiek van endemische infectieziekten en de effectiviteit van immunisatieprogramma's. Daarom presenteerden we een demografische microsimulatie voor de Amerikaanse bevolking van 1960 tot 2020, die werd uitgebreid met een transmissiemodel voor VZV en HZ. De demografische methoden waren vergelijkbaar met degenen die gebruikt werden om de microsimulatie voor België te ontwikkelen, maar met verschillende aanpassingen omdat de gegevens en de granulariteit verschilden. Aangezien huishoudenstransmissie zelden expliciet is onderzocht in de bestaande literatuur, werd de gastpopulatie evenwel gemodelleerd met leeftijds- en huishoudenstructuren. Dalende vruchtbaarheidscijfers sinds de jaren 1960 hebben geleid tot een veranderende leeftijdsopbouw van de Amerikaanse bevolking en een afnemende gemiddelde grootte van huishoudens.

Deze trends werden gereproduceerd in de microsimulatie. Hierdoor zullen de implicaties van de demografische veranderingen voor de epidemiologie van VZV en HZ op individueel, huishoud- en populatieniveau kunnen worden onderzocht.

Scientific output

Publications in peer-reviewed scientific journals

- **Møgelmose, S.**, Neels, K. & Hens, N. Incorporating human dynamic populations in models of infectious disease transmission: a systematic review. BMC Infect Dis 22, 862 (2022).
- Willem, L., Abrams, S., Libin, P. J. K., Coletti, P., Kuylen, E., Petrof, O., **Møgelmose, S.**, Wambua, J., Herzog, S. A., Faes, C., Beutels, P., & Hens, N. (2021). The impact of contact tracing and household bubbles on deconfinement strategies for COVID-19. Nature Communications, 12(1).

Submitted manuscripts

- **Møgelmose, S.**, Vijnck, L., Neven, F., Neels, K., Beutels, P. & Hens, N. Population age and household structures shape transmission dynamics of emerging infectious diseases: a longitudinal microsimulation approach.
- **Møgelmose, S.**, Neels, K., Beutels, P. & Hens, N. Exploring the impact of population ageing on the spread of emerging respiratory infections and the associated burden of mortality.
- Vanderlocht, J., **Møgelmose, S.**, Van Kerckhove, K., Beutels, P. & Hens, N. Chronic disease patients have fewer social contacts: a pilot survey with implications for transmission dynamics.

Working papers

- **Møgelmose, S.**, Hens, N. & Lewnard, J. The impact of demographic change on the epidemiology of varicella and herpes zoster: US population 1960-2020.

Acknowledgements

I would like to express my deepest gratitude to the exceptional people who supported and guided me throughout my PhD. First and foremost, I am incredibly grateful to my promoters, Niel and Karel. Thank you for giving me the opportunity to pursue a PhD, for your guidance, continuous encouragement and patience. I have especially enjoyed the interdisciplinary aspect of my PhD, and I highly appreciate that you have supported and seen opportunities in this collaboration. I would also like to thank the jury members for their time and great feedback. Additionally, I want to give my deepest appreciation to Joseph Lewnard for hosting and welcoming me in Berkeley. I would like to thank my colleagues from CPFH, especially Naomi, Tair, Julie, Layla and Leen, as well as my colleagues from DSI, CHERMID and the SIMID group. You have all contributed to a supportive, ambitious and fun environment, where my questions and doubts always have been met with engagement and encouragement. A special thanks to Evelyn and Julie, who started as my colleagues and have become my dear friends. I am deeply grateful to my family in Denmark for their support, frequent visits and for always taking a great interest in my life abroad. Thanks to my friends in and outside Belgium for providing me with laughter, energy and a fresh perspective whenever the PhD has caused me frustrations. A special thanks to my Belgian in-laws for the warm welcome to the family. Finally, I want to thank Vincent for your unwavering support and belief in me, for your belief in us and for embarking on a new adventure with me in Denmark.

Signe Møgelmose
4 September 2023
Antwerp

Bibliography

[1] J. Piret and G. Boivin. Pandemics Throughout History. *Frontiers in Microbiology*, 11(631736), 2021. doi: 10.3389/fmicb.2020.631736.

[2] J. F. Lindahl, D. Grace, and T. Strand. The consequences of human actions on risks for infectious diseases: a review. *Infection Ecology and Epidemiology*, 5(1), 2015. doi: 10.3402/IEE.V5.30048.

[3] D. Grennan. What is a pandemic? *JAMA*, 321(9), 2019. doi: 10.1001/jama.2019. 0700.

[4] B. Bramanti, N. C. Stenseth, L. Walløe, and X. Lei. Plague: A disease which changed the path of human civilization. In R. Yang and A. Anisimov, editors, *Yersinia pestis: Retrospective and Perspective*, volume 918. Springer, 2016. ISBN 9789402408904. doi: 10.1007/978-94-024-0890-4.

[5] A. Izdebski, P. Guzowski, R. Poniat, L. Masci, J. Palli, C. Vignola, M. Bauch, C. Cocozza, R. Fernandes, F. C. Ljungqvist, T. Newfield, A. Seim, D. Abel-Schaad, F. Alba-Sánchez, L. Björkman, A. Brauer, A. Brown, S. Czerwiński, A. Ejarque, M. Fiłoc, . . . , and A. Masi. Palaeoecological data indicates land-use changes across Europe linked to spatial heterogeneity in mortality during the Black Death pandemic. *Nature Ecology and Evolution*, 6(3):297–306, 2022. doi: 10.1038/s41559-021-01652-4.

[6] P. S. Brachman. Infectious diseases - Past, present, and future. *International Journal of Epidemiology*, 32(5):684–686, 2003. doi: 10.1093/ije/dyg282.

[7] University of California (Berkeley) and Max Planck Institute for Demographic

Research. Human Mortality Database. Accessed 2023-03-20. URL https://www.mortality.org.

[8] L. Shaw-Taylor. An introduction to the history of infectious diseases, epidemics and the early phases of the long-run decline in mortality. *Economic History Review*, 73(3), 8 2020. doi: 10.1111/ehr.13019.

[9] N. P. Johnson and J. Mueller. Updating the accounts: global mortality of the 1918-1920 "Spanish" influenza pandemic. *Bulletin of the history of medicine*, 76(1):105–115, 2002. doi: 10.1353/bhm.2002.0022.

[10] T. Eggerickx, J.-P. Sanderson, and C. Vandeschrick. Mortality in Belgium from nineteenth century to today. Variations according to age, sex, and social and spatial contexts. *Quetelet Journal*, 8(2):7–59, 5 2020. doi: 10.14428/rqj2020.08.02.01.

[11] World Health Organization. Top 10 causes of deaths, 2020. URL https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death.

[12] D. M. Morens and A. S. Fauci. The 1918 influenza pandemic: Insights for the 21st century. *Journal of Infectious Diseases*, 195(7):1018–1028, 2007. doi: 10.1086/511989.

[13] D. M. Morens and A. S. Fauci. Emerging Pandemic Diseases: How We Got to COVID-19. *Cell*, 182(3), 2020. doi: 10.1016/j.cell.2020.08.021.

[14] G. P. Garnett, S. Cousens, T. B. Hallett, R. Steketee, and N. Walker. Mathematical Models in the Evaluation of Health Programmes. *The Lancet*, 378: 515–525, 2011. doi: 10.1016/S0140-6736(10)61505-X.

[15] H. W. Hethcote. The Mathematics of Infectious Diseases. *SIAM Review*, 42(4): 599–653, 2000. doi: 10.1137/S0036144500371907.

[16] L. P. James, J. A. Salomon, C. O. Buckee, and N. A. Menzies. The Use and Misuse of Mathematical Modeling for Infectious Disease Policymaking: Lessons for the COVID-19 Pandemic. *Medical Decision Making*, 41(4):379–385, 2021. doi: 10.1177/0272989X21990391.

[17] M. J. Keeling and P. Rohani. *Modeling Infectious Diseases in Humans and Animals*. Princeton University Press, Princeton, 2007. doi: 10.2307/j.ctvcm4gk0.

[18] D. Bernoulli. Essai d'une nouvelle analyse de la mortalité causée par la petite vérole et des advantages de línoculation pour la prévenir. *Mem. Math. Phys. Acad. Roy. Sci.*, 1760.

[19] D. J. Daley and J. M. Gani. *Epidemic Modelling: An Introduction*. Cambridge University Press, 2001. ISBN 9780511608834. doi: 10.1017/CBO9780511608834.

[20] L. Held, N. Hens, P. O'Neill, and J. Wallinga, editors. *Handbook of Infectious Disease Data Analysis*. Chapman and Hall/CRC, 1st edition, 2019. doi: 10.1201/9781315222912.

[21] W. H. Hamer. Epidemic Disease in England - The Evidence of Variablity and of Persistency of Type. *The Lancet*, pages 733–739, 1906.

[22] W. O. Kermack and A. G. McKendrick. A Contribution to the Mathematical Theory of Epidemics. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 115(772):700–721, 1927. doi: 10.1098/rspa.1927.0118.

[23] N. Bailey. *The Mathematical Theory of Epidemics*. Griffin, London, 1957.

[24] R. M. Anderson and R. M. May. *Infectious Diseases of Humans. Dynamics and Control*. Oxford University Press, 1992.

[25] E. Vynnycky and R. G. White. *An Introduction to Infectious Disease Modelling*. Oxford University Press, Oxford, 2010.

[26] L. Willem, F. Verelst, J. Bilcke, N. Hens, and P. Beutels. Lessons from a decade of individual-based models for infectious disease transmission: a systematic review (2006-2015). *BMC Infectious Diseases*, 17(1):612, 2017. doi: 10.1186/s12879-017-2699-8.

[27] A. M. John. Endemic Disease in Host Populations with Fully Specified Demography. *Theoretical Population Biology*, 37(3):455–471, 1990. doi: 10.1016/0040-5809(90)90048-Z.

[28] H. Abbey. An Examination of the Reed-Frost Theory of Epidemics. *Human Biology*, 24(3):201–233, 1952.

[29] D. L. DeAngelis and V. Grimm. Individual-Based Models in Ecology After Four Decades. *F1000Prime Reports*, 6(39), 2014. doi: 10.12703/P6-39.

[30] J. Mossong, N. Hens, M. Jit, P. Beutels, K. Auranen, R. Mikolajczyk, M. Massari, S. Salmaso, G. S. Tomba, J. Wallinga, J. Heijne, M. Sadkowska-Todys, M. Rosinska, and W. J. Edmunds. Social Contacts and Mixing Patterns Relevant to the Spread of Infectious Diseases. *PLoS Medicine*, 5(3), 2008. ISSN 15491277. doi: 10.1371/journal.pmed.0050074.

[31] S. Cauchemez, F. Carrat, C. Viboud, A. J. Valleron, and P. Y. Boëlle. A Bayesian MCMC approach to study transmission of influenza: Application to household longitudinal data. *Statistics in Medicine*, 23(22):3469–3487, 2004. ISSN 02776715. doi: 10.1002/sim.1912.

[32] L. Willem, S. Abrams, P. J. Libin, P. Coletti, E. Kuylen, O. Petrof, S. Møgelmose, J. Wambua, S. A. Herzog, C. Faes, P. Beutels, and N. Hens. The impact of contact tracing and household bubbles on deconfinement strategies for COVID-19. *Nature Communications*, 12(1):1–9, 2021. ISSN 20411723. doi: 10.1038/s41467-021-21747-7.

[33] A. Torneri, L. Willem, V. Colizza, C. Kremer, C. Meuris, G. Darcis, N. Hens, and P. J. K. Libin. Controlling SARS- CoV-2 in schools using repetitive testing strategies. *eLife*, 11(e75593), 2022. doi: 10.7554/eLife.75593.

[34] F. Spooner, J. F. Abrams, K. Morrissey, G. Shaddick, M. Batty, R. Milton, A. Dennett, N. Lomax, N. Malleson, and N. Nelissen. A dynamic microsimulation model for epidemics. *Social Science & Medicine*, 291(114461), 2021. doi: 10.1016/j.socscimed.2021.114461.

[35] J. Hilton, H. Riley, L. Pellis, R. Aziza, S. P. Brand, I. K. Kombe, J. Ojal, A. Parisi, M. J. Keeling, D. James Nokes, R. Manson-Sawko, and T. House. A computational framework for modelling infectious disease policy based on age and household structure with applications to the COVID-19 pandemic. *PLoS Computational Biology*, 18(9):1–38, 2022. ISSN 15537358. doi: 10.1371/journal.pcbi.1010390.

[36] A. Melegaro, V. Marziano, E. Del Fava, P. Poletti, M. Tirani, C. Rizzo, and S. Merler. The impact of demographic changes, exogenous boosting and new vaccination policies on varicella and herpes zoster in Italy: A modelling and cost-effectiveness study. *BMC Medicine*, 16(117), 2018. ISSN 17417015. doi: 10.1186/s12916-018-1094-7.

[37] M. Eichner, M. Schwehm, J. Hain, H. Uphoff, B. Salzberger, M. Knuf, and R. Schmidt-Ott. 4Flu - an individual based simulation tool to study the effects

of quadrivalent vaccination on seasonal influenza in Germany. *BMC Infectious Diseases*, 14(1), 2014. ISSN 14712334. doi: 10.1186/1471-2334-14-365.

[38] L. Willem, K. V. Kerckhove, D. L. Chao, N. Hens, and P. Beutels. A Nice Day for an Infection? Weather Conditions and Social Contact Patterns Relevant to Influenza Transmission. *PLoS ONE*, 7(11), 2012. doi: 10.1371/journal.pone.0048695.

[39] N. J. Kassebaum. Global, regional, and national burden of diseases and injuries for adults 70 years and older: Systematic analysis for the Global Burden of Disease 2019 Study. *The BMJ*, 376, 2022. ISSN 17561833. doi: 10.1136/bmj-2021-068208.

[40] G. P. Garnett. An introduction to mathematical models in sexually transmitted disease epidemiology. *Sexually Transmitted Infections*, 78(1):7–12, 2002. ISSN 13684973. doi: 10.1136/sti.78.1.7.

[41] J. Hilton and M. J. Keeling. Incorporating household structure and demography into models of endemic disease. *Journal of the Royal Society Interface*, 16(157), 2019. ISSN 17425662. doi: 10.1098/rsif.2019.0317.

[42] N. Geard, K. Glass, J. M. McCaw, E. S. McBryde, K. B. Korb, M. J. Keeling, and J. McVernon. The Effects of Demographic Change on Disease Transmission and Vaccine Impact in a Household Structured Population. *Epidemics*, 13:56–64, 2015. ISSN 18780067. doi: 10.1016/j.epidem.2015.08.002.

[43] T. House and M. J. Keeling. Household Structure and Infectious Disease Transmission. *Epidemiology and Infection*, 137:654–661, 2009. ISSN 0950-2688. doi: 10.1017/S0950268808001416.

[44] W. Mahikul, L. J. White, K. Poovorawan, N. Soonthornworasiri, P. Sukontamarn, P. Chanthavilay, G. F. Medley, and W. Pan-ngum. Modeling household dynamics on Respiratory Syncytial Virus (RSV). *Plos One*, 14(7):e0219323, 2019. doi: 10.1371/journal.pone.0219323.

[45] L. Pellis, S. Cauchemez, N. M. Ferguson, and C. Fraser. Systematic selection between age and household structure for models aimed at emerging epidemic predictions. *Nature Communications*, 11(1), 2020. ISSN 20411723. doi: 10.1038/s41467-019-14229-4.

[46] A. Endo, M. Uchida, A. J. Kucharski, and S. Funk. Fine-scale family structure shapes influenza transmission risk in households: Insights from primary schools in Matsumoto city, 2014/15. *PLoS Computational Biology*, 15(12):1–18, 2019. ISSN 15537358. doi: 10.1371/journal.pcbi.1007589.

[47] V. Marziano, P. Poletti, F. Trentini, A. Melegaro, M. Ajelli, and S. Merler. Parental vaccination to reduce measles immunity gaps in Italy. *eLife*, 8, 2019. doi: 10.7554/elife.44942.

[48] I. Longini, J. Koopman, A. Monto, and J. Fox. Estimating Household and Community Transmission Parameters for Influenza. *American Journal of Epidemiology*, 115(5):736–751, 1982.

[49] I. K. Kombe, P. K. Munywoki, M. Baguelin, D. J. Nokes, and G. F. Medley. Model-based estimates of transmission of respiratory syncytial virus within households. *Epidemics*, 27:1–11, 2019. ISSN 18780067. doi: 10.1016/j.epidem .2018.12.001.

[50] N. Goeyvaerts, E. Santermans, G. Potter, A. Torneri, K. Van Kerckhove, L. Willem, M. Aerts, P. Beutels, and N. Hens. Household members do not contact each other at random: Implications for infectious disease modelling. *Proceedings of the Royal Society B: Biological Sciences*, 285(1893), 2018. ISSN 14712954. doi: 10.1098/rspb.2018.2201.

[51] S. H. Preston, P. Heuveline, and M. Guillot. *Demography. Measuring and Modeling Population Processes*. Blackwell Publishers, Oxford, 2001.

[52] S. Møgelmose, K. Neels, and N. Hens. Incorporating human dynamic populations in models of infectious disease transmission: A systematic review. *BMC Infectious Diseases*, 22(862), 2022. doi: 10.1186/s12879-022-07842-0.

[53] M. Spielauer. What is social science microsimulation? *Social Science Computer Review*, 29(1):9–20, 2011. ISSN 08944393. doi: 10.1177/0894439310370085.

[54] E. Zagheni. Microsimulation in Demographic Research. *International Encyclopedia of the Social & Behavioral Sciences: Second Edition*, 2:343–346, 2015. doi: 10.1016/B978-0-08-097086-8.31018-2.

[55] T. Van Effelterre, C. Marano, and K. H. Jacobsen. Modeling the hepatitis A epidemiological transition in Thailand. *Vaccine*, 34(4):555–562, 2016. ISSN 18732518. doi: 10.1016/j.vaccine.2015.11.052.

[56] M. Haacker, T. Bärnighausen, and R. Atun. HIV and the growing health burden from non-communicable diseases in Botswana: modelling study. *J Glob Health*, 9(1):10428, 2019. doi: 10.7189/jogh.09.010428.

[57] V. Costantino, H. F. Gidding, and J. G. Wood. Projections of zoster incidence in Australia based on demographic and transmission models of varicella-zoster virus infection. *Vaccine*, 35:6737–6742, 2017. doi: 10.1016/j.vaccine.2017.09.090.

[58] G. Guzzetta, M. Ajelli, Z. Yang, S. Merler, C. Furlanello, and D. Kirschner. Modeling Socio-Demography to Capture Tuberculosis Transmission Dynamics in a Low Burden Setting. *Journal of Theoretical Biology*, 289:197–295, 2011. ISSN 00092665. doi: 10.1016/j.jtbi.2011.08.032.

[59] S. Arregui, M. J. Iglesias, S. Samper, D. Marinova, C. Martin, J. Sanz, and Y. Moreno. Data-driven model for the assessment of mycobacterium tuberculosis transmission in evolving demographic structures. *Proceedings of the National Academy of Sciences*, 115(14), 2018. ISSN 10916490. doi: 10.1073/pnas.1720606115.

[60] N. Geard, J. M. McCaw, A. Dorin, K. B. Korb, and J. McVernon. Synthetic Population Dynamics : A Model of Household Demography. *Journal of Artificial Societies and Social Simulation*, 16(1), 2013.

[61] K. Davis. The World Demographic Transition. *The Annals of the American Academy of Political and Social Science*, 237(1):1–11, 1945. ISSN 0002-7162. doi: 10.1177/000271624523700102.

[62] J. Bongaarts. Human Population Growth and the Demographic Transition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364: 2985–2990, 2009. ISSN 0962-8436. doi: 10.1098/rstb.2009.0137.

[63] D. Kirk. Demographic Transition Theory. *Population Studies*, 50(3):361–387, 2015.

[64] United Nations. Department of Economic and Social Affairs. World Population Ageing 2019. Technical report, Department of Economic and Social Affairs, Population Division, 2019.

[65] United Nations. Department of Economic and Social Affairs. Population Division. World Population Prospects 2022. Technical report, 2022.

[66] W. C. Sanderson, S. Scherbov, P. Gerland, W. C. Sanderson, S. Scherbov, and P. Gerland. The end of population aging in high-income countries. *Vienna Yearbook of Population Research*, 16:163–176, 2018.

[67] P. Manfredi and J. R. Williams. 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*, 192:153–175, 2004. ISSN 00255564. doi: 10.1016/j.mbs.2004.11.006.

[68] S. Merler and M. Ajelli. Deciphering the Relative Weights of Demographic Transition and Vaccination in the Decrease of Measles Incidence in Italy. *Proceedings of the Royal Society, Biological Sciences*, 281, 2014. ISSN 1471-2954. doi: 10.1098/rspb.2013.2676.

[69] M. J. Ferrari, B. T. Grenfell, and P. M. Strebel. Think Globally, Act Locally: The Role of Local Demographics and Vaccination Coverage in the Dynamic Response of Measles Infection to Control. *Philosophical transactions of the Royal Society, Biological sciences*, 368(1623):20120141, 2013. ISSN 1471-2970. doi: 10.1098/rstb.2012.0141.

[70] R. Reves. Declining fertility in England and Wales as a major cause of the twentieth century decline in mortality. The role of changing family size and age structure in infectious disease mortality in infancy. *American Journal of Epidemiology*, 122(1):112–126, 1985. ISSN 0002-9262. doi: 10.1093/oxfordjournals.aje.a114070.

[71] D. A. T. Cummings, S. Iamsirithaworn, J. T. Lessler, A. Mcdermott, A. Nisalak, R. G. Jarman, D. S. Burke, and R. V. Gibbons. The Impact of the Demographic Transition on Dengue in Thailand : Insights from a Statistical Analysis and Mathematical Modeling. 6(9), 2009. doi: 10.1371/journal.pmed.1000139.

[72] F. Trentini, P. Poletti, S. Merler, and A. Melegaro. Measles immunity gaps and the progress towards elimination : a multi-country modelling analysis. *The Lancet Infectious Diseases*, 3099(17), 2017. ISSN 1473-3099. doi: 10.1016/S1473-3099(17)30421-8.

[73] S. E. Talbird, E. M. La, J. Carrico, S. Poston, J. E. Poirrier, J. K. DeMartino, and C. S. Hogea. Impact of population aging on the burden of vaccine-preventable diseases among older adults in the United States. *Human Vaccines and Immunotherapeutics*, 2020. doi: 10.1080/21645515.2020.1780847.

[74] D. van Oorschot, H. Vroling, E. Bunge, J. Diaz-Decaro, D. Curran, and B. Yawn. A systematic literature review of herpes zoster incidence worldwide. *Human Vaccines & Immunotherapeutics*, 17(6):1714–1732, 2021. ISSN 2164-5515. doi: 10.1080/21645515.2020.1847582.

[75] J. Horn, O. Damm, W. Greiner, H. Hengel, M. E. Kretzschmar, A. Siedler, B. Ultsch, F. Weidemann, O. Wichmann, A. Karch, and R. T. Mikolajczyk. Influence of demographic changes on the impact of vaccination against varicella and herpes zoster in Germany - a mathematical modelling study. *BMC Medicine*, 16(3), 2018. ISSN 17417015. doi: 10.1186/s12916-017-0983-5.

[76] L. Haynes. Aging of the Immune System: Research Challenges to Enhance the Health Span of Older Adults. *Frontiers in Aging*, 1(October):1–4, 2020. doi: 10.3389/fragi.2020.602108.

[77] G. Molenberghs, C. Faes, J. Verbeeck, P. Deboosere, S. Abrams, L. Willem, J. Aerts, H. Theeten, B. Devleesschauwer, N. B. Sierra, F. Renard, S. Herzog, P. Lusyne, J. Van Der Heyden, H. Van Oyen, P. Van Damme, and N. Hens. COVID-19 mortality, excess mortality, deaths per million and infection fatality ratio, Belgium, 9 March 2020 to 28 June 2020. *Eurosurveillance*, 27(7):1–10, 2022. ISSN 15607917. doi: 10.2807/1560-7917.ES.2022.27.7.2002060.

[78] N. J. Henry, A. Elagali, M. Nguyen, M. G. Chipeta, and C. E. Moore. Variation in excess all-cause mortality by age, sex, and province during the first wave of the COVID-19 pandemic in Italy. *Scientific Reports*, 12(1):1–12, 2022. ISSN 20452322. doi: 10.1038/s41598-022-04993-7.

[79] P. T. Campbell, J. McVernon, and N. Geard. Determining the Best Strategies for Maternally Targeted Pertussis Vaccination Using an Individual-Based Model. *Am J Epidemiol*, 186(1):109–117, 2017. doi: 10.1093/aje/kwx002.

[80] Z. Xu, K. Glass, C. L. Lau, N. Geard, P. Graves, and A. Clements. A Synthetic Population for Modelling the Dynamics of Infectious Disease Transmission in American Samoa. *Scientific Reports*, 7, 2017. ISSN 20452322. doi: 10.1038/s41598-017-17093-8.

[81] F. Liu, W. T. A. Enanoria, K. J. Ray, M. P. Coffee, A. Gordon, T. J. Aragón, G. Yu, B. J. Cowling, and T. C. Porco. Effect of the One-Child Policy on Influenza Transmission in China: A Stochastic Transmission Model. *PLoS ONE*, 9(2), 2014. ISSN 19326203. doi: 10.1371/journal.pone.0084961.

[82] D. van de Kaa. Europe's Second Demographic Transition. *Population Bulletin*, 41(1):1–59, 1987.

[83] R. Lesthaeghe. The Unfolding Story of the Second Demographic Transition. *Population and Development Review*, 36(2):211–251, 2010. ISSN 0098-7921. doi: 10.1111/j.1728-4457.2010.00328.x.

[84] K. Van den Bosch, P. Willemé, J. Geerts, J. Breda, S. Peeters, S. Van De Sande, F. Vrijens, C. Van de Voorde, and S. Stordeur. Residential care for older persons in Belgium: Projections 2011 – 2025. Technical report, 2011.

[85] W. Gardner, D. States, and N. Bagley. The Coronavirus and the Risks to the Elderly in Long-Term Care. *Journal of Aging and Social Policy*, 32(4-5): 310–315, 2020. ISSN 15450821. doi: 10.1080/08959420.2020.1750543.

[86] K. Modig, M. Lambe, A. Ahlbom, and M. Ebeling. Excess mortality for men and women above age 70 according to level of care during the first wave of COVID-19 pandemic in Sweden: A population-based study. *The Lancet Regional Health - Europe*, 4:100072, 2021. ISSN 26667762. doi: 10.1016/j.lanepe.2021.100072.

[87] D. N. Fisman, I. Bogoch, L. Lapointe-Shaw, J. McCready, and A. R. Tuite. Risk Factors Associated with Mortality among Residents with Coronavirus Disease 2019 (COVID-19) in Long-term Care Facilities in Ontario, Canada. *JAMA Network Open*, 3(7):1–7, 2020. ISSN 25743805. doi: 10.1001/jamanetworkopen.2020.15957.

[88] C. J. Cronin and W. N. Evans. Nursing home quality, COVID-19 deaths, and excess mortality. *Journal of Health Economics*, 82, 2022.

[89] S. Merler and M. Ajelli. 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, 2010. ISSN 0962-8452. doi: 10.1098/rspb.2009.1605.

[90] V. Marziano, P. Poletti, G. Guzzetta, M. Ajelli, P. Manfredi, and S. Merler. The Impact of Demographic Changes on the Epidemiology of Herpes Zoster: Spain as a Case Study. *Proceedings. Biological sciences*, 282, 2015. ISSN 1471-2954. doi: 10.1098/rspb.2014.2509.

[91] D. Moher, A. Liberati, J. Tetzlaff, D. G. Altman, D. Altman, G. Antes, D. Atkins, V. Barbour, N. Barrowman, J. A. Berlin, J. Clark, M. Clarke,

D. Cook, R. D'Amico, J. J. Deeks, P. J. Devereaux, K. Dickersin, M. Egger, E. Ernst, P. C. Gøtzsche, ..., and P. Tugwell. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, 6(7), 2009. ISSN 15491277. doi: 10.1371/journal.pmed.1000097.

[92] M. Penazzato, V. Bendaud, L. Nelson, J. Stover, and M. Mahy. Estimating future trends in paediatric HIV. *AIDS*, 28(Suppl 4):445–51, 2014. doi: 10.1097/QAD.0000000000000481.

[93] Y. Mekonnen, R. Jegou, R. A. Coutinho, J. Nokes, and A. Fontanet. Demographic impact of AIDS in a low-fertility urban African setting: projection for Addis Ababa, Ethiopia. *J Health Popul Nutr*, 20(2):120–129, 2002.

[94] M. Eichner, M. Schwehm, L. Eichner, and L. Gerlier. Direct and indirect effects of influenza vaccination. *BMC Infectious Diseases*, 17(1):1–8, 2017. ISSN 14712334. doi: 10.1186/s12879-017-2399-4.

[95] S. Li, C. Ma, L. Hao, Q. Su, Z. An, F. Ma, S. Xie, A. Xu, Y. Zhang, Z. Ding, H. Li, L. Cairns, H. Wang, H. Luo, N. Wang, L. Li, and M. J. Ferrari. Demographic transition and the dynamics of measles in six provinces in China: A modeling study. *PLoS Medicine*, 14(4):1–18, 2017. ISSN 15491676. doi: 10.1371/journal.pmed.1002255.

[96] P. Turgeon, V. Ng, R. Murray, and A. Nesbitt. Forecasting the incidence of salmonellosis in seniors in Canada: A trend analysis and the potential impact of the demographic shift. *PLoS One*, 13(11):e0208124, 2018. doi: 10.1371/journal.pone.0208124.

[97] S. A. McDonald, A. van Lier, D. Plass, and M. E. Kretzschmar. The Impact of Demographic Change on the Estimated Future Burden of Infectious Diseases: Examples from Hepatitis B and Seasonal Influenza in the Netherlands. *BMC Public Health*, 12(1046), 2012. ISSN 1471-2458. doi: 10.1186/1471-2458-12-1046.

[98] S. A. McDonald, M.-J. J. Mangen, A. Suijkerbuijk, E. Colzani, and M. E. E. Kretzschmar. Effects of an ageing population and the replacement of immune birth cohorts on the burden of hepatitis A in the Netherlands. *BMC Infect Dis*, 13:120, 2013. ISSN 1471-2334. doi: 10.1186/1471-2334-13-120.

[99] E. Yerushalmi, P. Hunt, S. Hoorens, C. Sauboin, and R. Smith. Exploring the Use of a General Equilibrium Method to Assess the Value of a Malaria

Vaccine: An Application to Ghana. *MDM Policy and Practice*, 4(2), 2019. doi: 10.1177/2381468319894345.

[100] F. Trentini, P. Poletti, A. Melegaro, and S. Merler. The introduction of 'No jab, No school' policy and the refinement of measles immunisation strategies in high-income countries. *BMC Med*, 17(1):86, 2019. doi: 10.1186/s12916-019-1318-5.

[101] J. E. Hood, M. R. Golden, J. P. Hughes, S. M. Goodreau, A. E. Siddiqi, S. E. Buskin, and S. E. Hawes. Projected demographic composition of the United States population of people living with diagnosed HIV. *AIDS Care*, 29(12): 1543–1550, 2017. doi: 10.1080/09540121.2017.1308466.

[102] C. C. Ku and P. J. Dodd. Forecasting the impact of population ageing on tuberculosis incidence. *PLoS ONE*, 14(9), 2019. ISSN 19326203. doi: 10.1371/journal.pone.0222937.

[103] R. Schmidt-Ott, M. Schwehm, and M. Eichner. Influence of social contact patterns and demographic factors on influenza simulation results. *BMC Infectious Diseases*, 16(1):1–11, 2016. ISSN 14712334. doi: 10.1186/s12879-016-1981-5.

[104] R. Schmidt-Ott, D. Molnar, A. Anastassopoulou, E. Yanni, C. Krumm, R. Bekkat-Berkani, G. Dos Santos, P. Henneke, M. Knuf, M. Schwehm, and M. Eichner. Assessing direct and indirect effects of pediatric influenza vaccination in Germany by individual-based simulations. *Human Vaccines and Immunotherapeutics*, 00(00):1–10, 2019. ISSN 2164554X. doi: 10.1080/21645515.2019.1682843.

[105] A. Khalifa, J. Stover, M. Mahy, P. Idele, T. Porth, and C. Lwamba. Demographic change and HIV epidemic projections to 2050 for adolescents and young people aged 15-24. *Glob Health Action*, 12(1):1662685, 2019. doi: 10.1080/16549716.2019.1662685.

[106] C. Dolk, M. Eichner, R. Welte, A. Anastassopoulou, L. A. Van Bellinghen, B. Poulsen Nautrup, I. Van Vlaenderen, R. Schmidt-Ott, M. Schwehm, and M. Postma. Cost-Utility of Quadrivalent Versus Trivalent Influenza Vaccine in Germany, Using an Individual-Based Dynamic Transmission Model. *PharmacoEconomics*, 34(12):1299–1308, 2016. ISSN 11792027. doi: 10.1007/s40273-016-0443-7.

[107] G. M. Knight, U. K. Griffiths, T. Sumner, Y. V. Laurence, A. Gheorghe, A. Vassall, P. Glaziou, and R. G. White. Impact and cost-effectiveness

of new tuberculosis vaccines in low- and middle-income countries. *Proceedings of the National Academy of Sciences*, 111(43):15520–15525, 2014. doi: 10.1073/pnas.1404386111.

[108] M. Mahy, M. Nzima, M. K. Ogungbemi, D. A. Ogbang, M. C. Morka, and J. Stover. Redefining the HIV epidemic in Nigeria: from national to state level. *AIDS*, 28(4), 2014. doi: 10.1097/QAD.0000000000000459.

[109] M. Mahy, M. Penazzato, A. Ciaranello, L. Mofenson, C. T. Yianoutsos, M. A. Davies, and J. Stover. Improving estimates of children living with HIV from the Spectrum AIDS Impact Model. *AIDS*, 31:S13–S22, 2017. doi: 10.1097/QAD.0000000000001306.

[110] A. Sibley, K. H. Han, A. Abourached, L. A. Lesmana, M. Makara, W. Jafri, R. Salupere, A. M. Assiri, A. Goldis, F. Abaalkhail, Z. Abbas, A. Abdou, F. Al Braiki, F. Al Hosani, K. Al Jaberi, M. Al Khatry, M. A. Al Mulla, H. Al Quraishi, A. Al Rifai, ..., and J. D. Schmelzer. The present and future disease burden of hepatitis C virus infections with today's treatment paradigm - Volume 3. *Journal of Viral Hepatitis*, 22:21–41, 2015. ISSN 13652893. doi: 10.1111/jvh.12476.

[111] M. Smit, J. Olney, N. P. Ford, M. Vitoria, S. Gregson, A. Vassall, and T. B. Hallett. The growing burden of noncommunicable disease among persons living with HIV in Zimbabwe. *Aids*, 32(6):773–782, 2018. ISSN 14735571. doi: 10.1097/QAD.0000000000001754.

[112] L. Gao and H. Hethcote. Simulations of Rubella Vaccination Strategies in China. *Mathematical Biosciences*, 202:371–385, 2006. ISSN 00255564. doi: 10.1016/j.jmmm.2005.07.010.

[113] T. Van Effelterre, A. Guignard, C. Marano, R. Rojas, and K. H. Jacobsen. Modeling the hepatitis A epidemiological transition in Brazil and Mexico. *Human Vaccines and Immunotherapeutics*, 13(8):1942–1951, 2017. ISSN 2164554X. doi: 10.1080/21645515.2017.1323158.

[114] T. Van Effelterre, R. De Antonio-Suarez, A. Cassidy, L. Romano-Mazzotti, and C. Marano. Model-based projections of the population-level impact of hepatitis A vaccination in Mexico. *Hum Vaccin Immunother*, 8(8):1099–1108, 2012. doi: 10.4161/hv.20549.

[115] A. Heffernan, G. S. Cooke, S. Nayagam, M. Thursz, and T. B. Hallett. Scaling up prevention and treatment towards the elimination of hepatitis C: a global mathematical model. *The Lancet*, 393(10178):1319–1329, 2019. ISSN 1474547X. doi: 10.1016/S0140-6736(18)32277-3.

[116] J. P. Aparicio and C. Castillo-Chavez. Mathematical modelling of tuberculosis epidemics. *Mathematical Biosciences and Engineering*, 6(2):209–237, 2009. doi: 10.3934/mbe.2009.6.209.

[117] D. Jayasundara, B. B. Hui, D. G. Regan, A. E. Heywood, C. R. MacIntyre, and J. G. Wood. Modelling the decline and future of hepatitis A transmission in Australia. *Journal of Viral Hepatitis*, 26(1):199–207, 2019. doi: 10.1111/jvh.13018.

[118] J. R. Williams, P. Manfredi, and A. Melegaro. The potential impact of the demographic transition in the Senegal-Gambia region of sub-Saharan Africa on the burden of infectious disease and its potential synergies with control programmes: The case of hepatitis B. *BMC Medicine*, 16(1):1–13, 2018. ISSN 17417015. doi: 10.1186/s12916-018-1100-0.

[119] A. van Lier, A. Lugner, W. Opstelten, P. Jochemsen, J. Wallinga, F. Schellevis, E. Sanders, H. de Melker, and M. van Boven. Distribution of Health Effects and Cost-effectiveness of Varicella Vaccination are Shaped by the Impact on Herpes Zoster. *EBioMedicine*, 2(10):1494–1499, 2015. doi: 10.1016/j.ebiom.2015.08.017.

[120] E. Van Imhoff and W. Post. Microsimulation Methods for Population Projection. *Population: An English Selection*, 10(1):97–138, 1998.

[121] C. Newell. *Methods and Models in Demography*. Belhaven Press, London, 1988.

[122] P. Zueras, R. Rutigliano, and S. Trias-Llimós. Marital status, living arrangements, and mortality in middle and older age in Europe. *International Journal of Public Health*, 65(5):627–636, 2020. ISSN 1420911X. doi: 10.1007/s00038-020-01371-w.

[123] W. Lutz and J. R. Goldstein. How to deal with uncertainty in population forecasting? *International Statistical Review*, 72(1):1–4, 2004. ISSN 1751-5823. doi: 10.1111/j.1751-5823.2004.tb00219.x.

[124] J. Bijak. *Forecasting international migration in Europe: A Bayesian view.* Springer Netherlands, 1st edition, 2010. ISBN 978-90-481-8897-0. doi: 10.1007/978-90-481-8897-0.

[125] IOM. World Migration Report 2020. Technical report, UN, New York, 2019.

[126] United Nations. Department of Economic and Social Affairs. Population Division. Patterns and trends in household size and composition: Evidence from a United Nations dataset. Technical report, 2019.

[127] K. Glass, J. McCaw, and J. McVernon. Incorporating Population Dynamics into Household Models of Infectious Disease Transmission. *Epidemics*, 3:152–158, 2011.

[128] L. J. Strausbaugh, S. R. Sukumar, and C. L. Joseph. Infectious disease outbreaks in nursing homes: An unappreciated hazard for frail elderly persons. *Clinical Infectious Diseases*, 36(7):870–876, 2003. ISSN 10584838. doi: 10.1086/368197.

[129] P. Gaspard, A. Mosnier, L. Simon, Olivia Ali-Brandmeyer, C. Rabaud, S. Larocca, B. Heck, S. Aho-Glele, P. Pothier, and K. Ambert-Balay. Gastroenteritis and respiratory infection outbreaks in French nursing homes from 2007 to 2018: Morbidity and all-cause lethality according to the individual characteristics of residents. *PLoS ONE*, 14(9):1–15, 2019. ISSN 19326203. doi: 10.1371/journal.pone.0222321.

[130] R. A. Garibaldi. Residential care and the elderly: the burden of infection. *Journal of Hospital Infection*, 43, 1999.

[131] S. Zinn. *A Continuous-Time Microsimulation and First Steps Towards a Multi-Level Approach in Demography.* PhD thesis, 2011.

[132] F. Willekens. Biographic Forecasting: Bridging the Micro- Macro Gap in Population Forecasting. *New Zealand Population Review*, 31(1), 2005.

[133] E. Van Imhoff and W. Post. Microsimulation methods for population projections. *Population*, 52(4):889–932, 1997. ISSN 00324663. doi: 10.2307/1534618.

[134] A. Harding and F. Willekens. Continuous-time Microsimulation in Longitudinal Analysis. *New Frontiers in Microsimulation Modelling*, pages 413–436, 2018. doi: 10.4324/9781315248066-16.

[135] E. Hammel, C. Mason, and K. Wachter. *SOCSIM II a sociodemographic microsimulation program rev. 1.0 operating manual.* University of California, Berkeley, 1990.

[136] E. A. Hammel, D. W. Hutchinson, K. W. Wachter, R. T. Lundy, and R. Z. Deuel. *The socsim demographic-sociological microsimulation program : operating manual.* Institute of International Studies University of California, Berkeley, 1976.

[137] S. Zinn, J. Gampe, J. Himmelspach, and A. M. Uhrmacher. MIC-CORE: A tool for microsimulation. *Proceedings - Winter Simulation Conference*, (December): 992–1102, 2009. ISSN 08917736. doi: 10.1109/WSC.2009.5429424.

[138] S. Zinn. The MicSim Package of R: An Entry-Level Toolkit for Continuous-Time Microsimulation. *International Journal of Microsimulation*, 7(73):3–32, 2014.

[139] E. Zagheni. The impact of the HIV/AIDS epidemic on kinship resources for orphans in Zimbabwe. *Population and Development Review*, 37(4):761–783, 2011. ISSN 00987921. doi: 10.1111/j.1728-4457.2011.00456.x.

[140] E. A. Hammel. Demographic dynamics and kinship in anthropological populations. *Proceedings of the National Academy of Sciences of the United States of America*, 102(6):2248–2253, 2005. ISSN 00278424. doi: 10.1073/pnas.0409762102.

[141] K. W. Wachter, J. E. Knodel, and M. Vanlandingham. AIDS and the elderly of Thailand: Projecting familial impacts. *Demography*, 39(1):25–41, 2002. ISSN 00703370. doi: 10.2307/3088362.

[142] R. Margolis and A. M. . Verdery. A Cohort Perspective on the Demography of Grandparenthood : Past, Present, and Future Changes in Race and Sex Disparities in the United States. *Demography*, 56(4):1495–1518, 2019.

[143] M. Murphy. Long-Term Effects of the Demographic Transition on Family and Kinship Networks in Britain. *Population and Development Review*, 37:55–80, 2011. ISSN 00987921. doi: 10.1111/j.1728-4457.2011.00378.x.

[144] M. Spielauer, C. Hicks, S. Gribble, G. Rowe, X. Lin, K. Moore, L. Plager, and H. Nguyen. The LifePaths Microsimulation Model: An Overview. Technical report, Statistics Canada. Modelling Division., 2013.

[145] L. Andreassen, D. Fredriksen, H. M. Gjefsen, E. Halvorsen, and N. M. Stølen. The dynamic cross-sectional microsimulation model MOSART. *International Journal of Microsimulation*, 13(1):92–113, 2020. ISSN 17475864. doi: 10.34196/ijm.00214.

[146] C. P. Van Der Ploeg, C. Van Vliet, S. J. De Vlas, J. O. Ndinya-Achola, L. Fransen, G. J. Van Oortmarssen, and J. D. F. Habbema. STDSIM: A microsimulation model for decision support in STD control. *Interfaces*, 28(3): 84–100, 1998. ISSN 00922102. doi: 10.1287/inte.28.3.84.

[147] K. K. Case, T. B. Hallett, S. Gregson, K. Porter, and P. D. Ghys. Development and future directions for the Joint United Nations Programme on HIV/AIDS estimates. *Aids*, 28(September):S411–S414, 2014. ISSN 14735571. doi: 10.1097/QAD.0000000000000487.

[148] P. T. Campbell, J. McVernon, P. McIntyre, and N. Geard. Influence of population demography and immunization history on the impact of an antenatal pertussis program. *Clinical Infectious Diseases*, 63, 2016. ISSN 15376591. doi: 10.1093/cid/ciw520.

[149] E. van Imhoff and N. Keilman. *LIPRO 2.0: an application of a dynamic demographic projection model to household structure in the Netherlands.* Swets & Zeitlinger, Amsterdam/Lisse, 1991.

[150] P. N. Krivitsky, P. Coletti, and N. Hens. A Tale of Two Datasets: Representativeness and Generalisability of Inference for Samples of Networks. *arXiv: 2202.03685*, 2022.

[151] T. Hastie and R. Tibshirani. Generalized Additive Models. *Statistical science*, 1(3):297–318, 1986.

[152] T. Hastie and R. Tibshirani. *Generalized Additive Models.* Chapman and Hall, London, 1990.

[153] S. N. Wood. *Generalized Additive Models. An introduction with R.* CRC Press. Taylor & Francis Group, second edi edition, 2017. ISBN 978-1-4987-2833-1.

[154] Statbel. Total international migration. URL `https://statbel.fgov.be/en/themes/population/population-movement/migration`.

[155] J. Li and C. O'Donoghue. A survey of dynamic microsimulation models: uses, model structure and methodology. *International Journal of Microsimulation*, 6:3–55, 2013. ISSN 1871-9872. doi: 10.1093/jae/ejm029.

[156] B. Perelli-Harris. How Similar are Cohabiting and Married Parents? Second Conception Risks by Union Type in the United States and Across Europe. *European Journal of Population*, 30(4):437–464, 10 2014. ISSN 15729885. doi: 10.1007/s10680-014-9320-2.

[157] M. Kreidl and Z. Zilinčíková. How Does Cohabitation Change People's Attitudes toward Family Dissolution? *European Sociological Review*, 37(4):541–554, 8 2021. ISSN 14682672. doi: 10.1093/esr/jcaa073.

[158] M. Kolk and G. Andersson. Two Decades of Same-Sex Marriage in Sweden: A Demographic Account of Developments in Marriage, Childbearing, and Divorce. *Demography*, 57(1):147–169, 2 2020. ISSN 15337790. doi: 10.1007/s13524-019-00847-6.

[159] C. Van Mol and H. de Valk. Migration and Immigrants in Europe: A Historical and Demographic Perspective. In B. Garcés-Mascareñas and R. Penninx, editors, *Integration Processes and Policies in Europe*, IMISCOE Research Series, chapter 3. Springer, 2016. ISBN 9783319216744. doi: 10.1007/978-3-319-21674.

[160] J. Van Hook and W. Zhang. Who Stays? Who Goes? Selective Emigration Among the Foreign-Born. *Popul Res Policy Rev*, 30(1):1–24, 2011. doi: 10.1007/s11113-010-9183-0.

[161] Statistics Belgium (Statbel). Births and fertility. URL `https://statbel.fgov.be/en/themes/population/births-and-fertility#figures`.

[162] K. Neels, M. Murphy, M. Ní Bhrolcháin, and Beaujouan. Rising Educational Participation and the Trend to Later Childbearing. *Population and Development Review*, 43(4):667–693, 12 2017. ISSN 17284457. doi: 10.1111/padr.12112.

[163] K. Neels and D. De Wachter. Postponement and recuperation of Belgian fertility: How are they related to rising female educational attainment? *Vienna Yearbook of Population Research*, (1):77–106, 2010. ISSN 17284414. doi: 10.1553/populationyearbook2010s77.

[164] L. Van Landschoot, J. Van Bavel, and H. A. de Valk. Estimating the contribution of mothers of foreign origin to total fertility: The recent recovery of period

fertility in the Belgian region of Flander. *Demographic Research*, 30(1):361–376, 2014. ISSN 14359871. doi: 10.4054/DemRes.2014.30.12.

[165] L. Semenzato, J. Botton, J. Drouin, F. Cuenot, R. Dray-Spira, A. Weill, and M. Zureik. Chronic diseases, health conditions and risk of COVID-19-related hospitalization and in-hospital mortality during the first wave of the epidemic in France: a cohort study of 66 million people. *The Lancet Regional Health - Europe*, 8, 2021. ISSN 26667762. doi: 10.1016/j.lanepe.2021.100158.

[166] M. Drozd, M. Pujades-Rodriguez, P. J. Lillie, S. Straw, A. W. Morgan, M. T. Kearney, K. K. Witte, and R. M. Cubbon. Non-communicable disease, sociodemographic factors, and risk of death from infection: a UK Biobank observational cohort study. *The Lancet Infectious Diseases*, 21(8):1184–1191, 8 2021. ISSN 14744457. doi: 10.1016/S1473-3099(20)30978-6.

[167] T. V. Hoang, P. Coletti, Y. W. Kifle, K. V. Kerckhove, S. Vercruysse, L. Willem, P. Beutels, and N. Hens. Close contact infection dynamics over time: insights from a second large-scale social contact survey in Flanders, Belgium, in 2010-2011. *BMC Infectious Diseases*, 21(1):1–15, 2021. ISSN 14712334. doi: 10.118 6/s12879-021-05949-4.

[168] E. Goldstein, M. Lipsitch, and M. Cevik. On the effect of age on the transmission of SARS-CoV-2 in households, schools, and the community. *Journal of Infectious Diseases*, 223(3):362–369, 2021. ISSN 15376613. doi: 10.1093/infdis /jiaa691.

[169] L. A. Paul, N. Daneman, K. L. Schwartz, M. Science, K. A. Brown, M. Whelan, E. Chan, and S. A. Buchan. Association of Age and Pediatric Household Transmission of SARS-CoV-2 Infection. *JAMA Pediatrics*, 175(11):1151–1158, 2021. ISSN 21686211. doi: 10.1001/jamapediatrics.2021.2770.

[170] E. Kuylen, L. Willem, J. Broeckhove, P. Beutels, and N. Hens. Clustering of susceptible individuals within households can drive measles outbreaks: an individual-based model exploration. *Scientific Reports*, 10(1):1–12, 2020. ISSN 20452322. doi: 10.1038/s41598-020-76746-3.

[171] R. Silhol and P.-Y. Boëlle. Modelling the Effects of Population Structure on Childhood Disease: The Case of Varicella. *PLoS computational biology*, 7(7), 2011. ISSN 1553-7358. doi: 10.1371/journal.pcbi.1002105.

[172] S. H. Preston and A. Stokes. Sources of Population Aging in More and Less Developed Countries. *Population and Development Review*, 38(2):221–236, 2012. doi: 10.1111/j.1728-4457.2012.00490.x.

[173] Federaal Planbureau. Prospectieve sterftequotiënten. Technical Report December, 2009.

[174] M. Vandresse. Federal Planning Bureau. Modelling fertility for national population projections. The case of Belgium. Technical report, Federal Planning Bureau (Belgium), 2020.

[175] F. Ball, D. Mollison, and G. Scalia-Tomba. Epidemics with Two Levels of Mixing. *The Annals of Applied Probability*, 7(1):46–89, 1997.

[176] L. Willem, T. Van Hoang, S. Funk, P. Coletti, P. Beutels, P. Beutels, N. Hens, and N. Hens. SOCRATES: An online tool leveraging a social contact data sharing initiative to assess mitigation strategies for COVID-19. *BMC Research Notes*, 13(1):1–8, 2020. ISSN 17560500. doi: 10.1186/s13104-020-05136-9.

[177] A. J. O'Malley and P. V. Marsden. The analysis of social networks. *Health Services and Outcomes Research Methodology*, 8(4):222–269, 2008. ISSN 13873741. doi: 10.1007/s10742-008-0041-z.

[178] J. Glasser, D. Taneri, Z. Feng, J. H. Chuang, P. Tüll, W. Thompson, M. M. McCauley, and J. Alexander. Evaluation of targeted influenza vaccination strategies via population modeling. *PLoS ONE*, 5(9):1–8, 2010. ISSN 19326203. doi: 10.1371/journal.pone.0012777.

[179] F. Carrat, E. Vergu, N. M. Ferguson, M. Lemaitre, S. Cauchemez, S. Leach, and A. J. Valleron. Time lines of infection and disease in human influenza: A review of volunteer challenge studies. *American Journal of Epidemiology*, 167 (7):775–785, 2008. ISSN 00029262. doi: 10.1093/aje/kwm375.

[180] P. Arevalo, H. Q. McLean, E. A. Belongia, and S. Cobey. Earliest infections predict the age distribution of seasonal influenza a cases. *eLife*, 9:1–30, jul 2020. ISSN 2050084X. doi: 10.7554/eLife.50060.

[181] N. G. Davies, P. Klepac, Y. Liu, K. Prem, M. Jit, C. Covid, and R. M. Eggo. Age-dependent effects in the transmission and control of COVID-19 epidemics. 26(August), 2020. doi: 10.1038/s41591-020-0962-9.

[182] V. Bajaj, N. Gadi, A. P. Spihlman, S. C. Wu, C. H. Choi, and V. R. Moulton. Aging, Immunity, and COVID-19: How Age Influences the Host Immune Response to Coronavirus Infections? *Frontiers in Physiology*, 11:1–23, 2021. ISSN 1664042X. doi: 10.3389/fphys.2020.571416.

[183] N. Franco, P. Coletti, L. Willem, L. Angeli, A. Lajot, S. Abrams, P. Beutels, C. Faes, and N. Hens. Inferring age-specific differences in susceptibility to and infectiousness upon SARS-CoV-2 infection based on Belgian social contact data. *PLoS Computational Biology*, 18(3):1–17, 2022. ISSN 15537358. doi: 10.1371/journal.pcbi.1009965.

[184] N. Keilman. Erroneous Population Forecasts. In T. Bengtsson and N. Keilman, editors, *Old and New Perspectives on Mortality Forecasting*. Springer, Cham, Switzerland, 2019. ISBN 9783030050740. doi: 10.1007/978-3-030-05075-7{_}21.

[185] M. Poulain, L. Dal, and A. Herm. Trends in living arrangements and their impact on the mortality of older adults: Belgium 1991-2012. *Demographic Research*, 43:401–430, 2020. ISSN 14359871. doi: 10.4054/DEMRES.2020.43.15.

[186] N. Keilman and S. Christiansen. Norwegian Elderly Less Likely to Live Alone in the Future. *European Journal of Population*, 26(1):47–72, 2010. ISSN 0168-6577. doi: 10.1007/s10680-009-9195-9.

[187] A. Wi??niowski, P. W. F. Smith, J. Bijak, J. Raymer, and J. J. Forster. Bayesian Population Forecasting: Extending the Lee-Carter Method. *Demography*, 52(3): 1035–1059, 2015. ISSN 15337790. doi: 10.1007/s13524-015-0389-y.

[188] F. Ball and P. Neal. A General Model for Stochastic SIR Epidemics with Two Levels of Mixing. *Mathematical Biosciences*, 180:73–102, 2002. ISSN 00255564. doi: 10.1016/S0025-5564(02)00125-6.

[189] K. Prem, A. R. Cook, and M. Jit. Projecting social contact matrices in 152 countries using contact surveys and demographic data. *PLoS Computational Biology*, (3):0381–0391, 2017. ISSN 15491277. doi: 10.1371/journal.pmed.0050074.

[190] S. Arregui, A. Aleta, J. Sanz, and Y. Moreno. Projecting social contact matrices to different demographic structures. *PLoS Computational Biology*, 14(12):1–18, 2018. ISSN 15537358. doi: 10.1371/journal.pcbi.1006638.

[191] G. Lanzieri. The greying of the baby boomers. A century-long view of ageing in European populations. Technical report, 2011.

[192] G. Gavazzi, F. Herrmann, and K. H. Krause. Aging and infectious diseases in the developing world. *Clinical Infectious Diseases*, 39(1):83–91, 2004. ISSN 10584838. doi: 10.1086/421559.

[193] P. Shetty. Grey matter: Ageing in developing countries. *The Lancet*, 379(9823): 1285–1287, 2012. ISSN 1474547X. doi: 10.1016/S0140-6736(12)60541-8.

[194] L. Chaker, A. Falla, S. J. van der Lee, T. Muka, D. Imo, L. Jaspers, V. Colpani, S. Mendis, R. Chowdhury, W. M. Bramer, R. Pazoki, and O. H. Franco. The global impact of non-communicable diseases on macro-economic productivity: a systematic review. *European Journal of Epidemiology*, 30(5):357–395, 2015. ISSN 15737284. doi: 10.1007/s10654-015-0026-5.

[195] F. Kämpfen, N. Wijemunige, and B. Evangelista. Aging, non-communicable diseases, and old-age disability in low- and middle-income countries: a challenge for global health. *International Journal of Public Health*, 63(9):1011–1012, 2018. ISSN 1420911X. doi: 10.1007/s00038-018-1137-z.

[196] J. Y. Xi, X. Lin, and Y. T. Hao. Measurement and projection of the burden of disease attributable to population aging in 188 countries, 1990-2050: A population-based study. *Journal of Global Health*, 12, 2022. ISSN 20472986. doi: 10.7189/jogh.12.04093.

[197] J. Bartoszko and M. Loeb. The burden of influenza in older adults: meeting the challenge. *Aging Clinical and Experimental Research*, 33(3):711–717, 2021. ISSN 17208319. doi: 10.1007/s40520-019-01279-3.

[198] J. R. Glynn and P. A. Moss. Systematic analysis of infectious disease outcomes by age shows lowest severity in school-age children. *Scientific Data*, 7(1):1–13, 2020. ISSN 20524463. doi: 10.1038/s41597-020-00668-y.

[199] D. Mertz, T. H. Kim, J. Johnstone, P. P. Lam, M. Science, S. P. Kuster, S. A. Fadel, D. Tran, E. Fernandez, N. Bhatnagar, and M. Loeb. Populations at risk for severe or complicated influenza illness: Systematic review and meta-analysis. *BMJ (Online)*, 347(7923):1–15, 2013. ISSN 17561833. doi: 10.1136/bmj.f5061.

[200] A. M. Moa, R. I. Menzies, J. K. Yin, and C. R. MacIntyre. Modelling the influenza disease burden in people aged 50–64 and 65 years in Australia. *In-*

fluenza and other Respiratory Viruses, 16(1):132–141, 2022. ISSN 17502659. doi: 10.1111/irv.12902.

[201] K. Wing, D. J. Grint, R. Mathur, H. P. Gibbs, G. Hickman, E. Nightingale, A. Schultze, H. Forbes, V. Nafilyan, K. Bhaskaran, E. Williamson, T. House, L. Pellis, E. Herrett, N. Gautam, H. J. Curtis, C. T. Rentsch, A. Y. S. Wong, B. MacKenna, A. Mehrkar, S. Bacon, I. J. Douglas, S. J. W. Evans, L. Tomlinson, B. Goldacre, R. M. Eggo, and the OpenSAFELY platform. Association between household composition and severe COVID-19 outcomes in older people by ethnicity: an observational cohort study using the OpenSAFELY platform. *International Journal of Epidemiology*, 51(6), 2022. doi: 10.1093/ije/dyad041.

[202] H. Beltrán-Sánchez, S. Soneji, and E. M. Crimmins. Past, Present, and Future of Healthy Life Expectancy. *Cold Spring Harb Perspect Med.*, 5, 2015. ISSN 15577740.

[203] K. Christensen, G. Doblhammer, R. Rau, and J. W. Vaupel. Ageing populations: the challenges ahead. 374(9696):1196–1208, 2010. doi: 10.1016/S0140-6736(09)61460-4.Ageing.

[204] J. A. Salomon, H. Wang, M. K. Freeman, T. Vos, A. D. Flaxman, A. D. Lopez, and C. J. L. Murray. Healthy life expectancy for 187 countries, 1990-2010: a systematic analysis for the Global Burden Disease Study 2010. *The Lancet*, 380 (9859):2144–2162, 12 2012. ISSN 0140-6736. doi: 10.1016/S0140-6736(12)61690-0.

[205] J. W. Vaupel, F. Villavicencio, and M. P. Bergeron-Boucher. Demographic perspectives on the rise of longevity. *Proceedings of the National Academy of Sciences of the United States of America*, 118(9):1–10, 2021. ISSN 10916490. doi: 10.1073/pnas.2019536118.

[206] C. J. Evans, Y. Ho, B. A. Daveson, S. Hall, I. J. Higginson, and W. Gao. Place and Cause of Death in Centenarians: A Population-Based Observational Study in England, 2001 to 2010. *PLoS Medicine*, 11(6):1–13, 2014. ISSN 15491676. doi: 10.1371/journal.pmed.1001653.

[207] J. R. Williams and P. Manfredi. Ageing Populations and Childhood infections: The Potential Impact on Epidemic patterns and Morbidity. *International Journal of Epidemiology*, 33:566–572, 2004. ISSN 03005771. doi: 10.1093/ije/dyh098.

[208] P. C. Wroe, J. A. Finkelstein, G. T. Ray, J. A. Linder, K. M. Johnson, S. Rifas-Shiman, M. R. Moore, and S. S. Huang. Aging population and future burden of pneumococcal pneumonia in the United States. *J Infect Dis*, 205(10):1589–1592, 2012. doi: 10.1093/infdis/jis240.

[209] M. R. Gold, D. Stevenson, and D. G. Fryback. HALYs and QALYs and DALYs, Oh My: Similarities and Differences in Summary Measures of Population Health. *Annual Review of Public Health*, 23(1):115–134, 2002. ISSN 0163-7525. doi: 10.1146/annurev.publhealth.23.100901.140513.

[210] S. Gadeyne. *The ultimate inequality: socio-economic differences in all-cause and cause-specific mortality in Belgium in the first part of the 1990s.* Brussel : Centrum voor Bevolkings- en Gezinsstudiën, 2006. ISBN 9040302456.

[211] K. Neels. *Reproductive strategies in Belgian fertility, 1960-1990.* Brussel : CBGS, 2006. ISBN 9040302448.

[212] G. Béraud, S. Kazmercziak, P. Beutels, D. Levy-Bruhl, X. Lenne, N. Mielcarek, Y. Yazdanpanah, P. Y. Boëlle, N. Hens, and B. Dervaux. The French connection: The first large population-based contact survey in France relevant for the spread of infectious diseases. *PLoS ONE*, 10(7):1–22, 2015. ISSN 19326203. doi: 10.1371/journal.pone.0133203.

[213] M. Biggerstaff, S. Cauchemez, C. Reed, M. Gambhir, and L. Finelli. Estimates of the reproduction number for seasonal, pandemic, and zoonotic influenza: A systematic review of the literature. *BMC Infectious Diseases*, 14(1):1–20, 2014. ISSN 14712334. doi: 10.1186/1471-2334-14-480.

[214] B. J. Cowling, M. S. Lau, L. M. Ho, S. K. Chuang, T. Tsang, S. H. Liu, P. Y. Leung, S. V. Lo, and E. H. Lau. The effective reproduction number of pandemic influenza: Prospective estimation. *Epidemiology*, 21(6):842–846, 2010. ISSN 10443983. doi: 10.1097/EDE.0b013e3181f20977.

[215] E. Petersen, M. Koopmans, U. Go, D. H. Hamer, N. Petrosillo, F. Castelli, M. Storgaard, S. Al Khalili, and L. Simonsen. Comparing SARS-CoV-2 with SARS-CoV and influenza pandemics. *The Lancet Infectious Diseases*, 20(9): e238–e244, 2020. ISSN 14744457. doi: 10.1016/S1473-3099(20)30484-9.

[216] Q. Li, X. Guan, P. Wu, X. Wang, L. Zhou, Y. Tong, R. Ren, K. S. Leung, E. H. Lau, J. Y. Wong, X. Xing, N. Xiang, Y. Wu, C. Li, Q. Chen, D. Li, T. Liu, J. Zhao, M. Liu, ..., and Z. Feng. Early Transmission Dynamics in Wuhan,

China, of Novel Coronavirus–Infected Pneumonia. *New England Journal of Medicine*, 382(13):1199–1207, 2020. ISSN 0028-4793. doi: 10.1056/nejmoa2001316.

[217] D. McEvoy, C. McAloon, A. Collins, K. Hunt, F. Butler, A. Byrne, M. Casey-Bryars, A. Barber, J. Griffin, E. A. Lane, P. Wall, and S. J. More. Relative infectiousness of asymptomatic SARS-CoV-2 infected persons compared with symptomatic individuals: A rapid scoping review. *BMJ Open*, 11(5):1–8, 2021. ISSN 20446055. doi: 10.1136/bmjopen-2020-042354.

[218] S. Abrams, J. Wambua, E. Santermans, L. Willem, E. Kuylen, P. Coletti, P. Libin, C. Faes, O. Petrof, S. A. Herzog, P. Beutels, and N. Hens. Modelling the early phase of the Belgian COVID-19 epidemic using a stochastic compartmental model and studying its implied future trajectories. *Epidemics*, 35:100449, 2021. ISSN 18780067. doi: 10.1016/j.epidem.2021.100449.

[219] P. Coletti, P. Libin, O. Petrof, L. Willem, S. Abrams, S. A. Herzog, C. Faes, E. Kuylen, J. Wambua, P. Beutels, and N. Hens. A data-driven metapopulation model for the Belgian COVID-19 epidemic: assessing the impact of lockdown and exit strategies. *BMC Infectious Diseases*, 21(1):1–12, 2021. ISSN 14712334. doi: 10.1186/s12879-021-06092-w.

[220] X. He, E. H. Lau, P. Wu, X. Deng, J. Wang, X. Hao, Y. C. Lau, J. Y. Wong, Y. Guan, X. Tan, X. Mo, Y. Chen, B. Liao, W. Chen, F. Hu, Q. Zhang, M. Zhong, Y. Wu, L. Zhao, F. Zhang, B. J. Cowling, F. Li, and G. M. Leung. Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nature Medicine*, 26(5):672–675, 2020. ISSN 1546170X. doi: 10.1038/s41591-020-0869-5.

[221] J. T. Wu, K. Leung, M. Bushman, N. Kishore, R. Niehus, P. M. de Salazar, B. J. Cowling, M. Lipsitch, and G. M. Leung. Estimating clinical severity of COVID-19 from the transmission dynamics in Wuhan, China. *Nature Medicine*, 26(4):506–510, 2020. ISSN 1546170X. doi: 10.1038/s41591-020-0822-7.

[222] S. Riley, K. O. Kwok, K. M. Wu, D. Y. Ning, B. J. Cowling, J. T. Wu, L. M. Ho, T. Tsang, S. V. Lo, D. K. Chu, E. S. Ma, and J. S. Peiris. Epidemiological characteristics of 2009 (H1N1) pandemic influenza based on paired sera from a longitudinal community cohort study. *PLoS Medicine*, 8(6), 2011. ISSN 15491676. doi: 10.1371/journal.pmed.1000442.

[223] A. H. Briggs, R. Meacock, D. A. Goldstein, E. Kirwin, and T. Wisløff. Estimating (quality-adjusted) life-year losses associated with deaths: With application to COVID-19. *Health Economics*, 30:699–707, 2022. doi: 10.1002/hec.4208.

[224] P. Adab, S. Haroon, M. E. O'Hara, and R. E. Jordan. Comorbidities and covid-19. *The BMJ*, pages 19–20, 2022. ISSN 17561833. doi: 10.1136/bmj.o1431.

[225] E. Pérez-Flores, J. C. Izquierdo-Puente, J. J. Castillo-Pérez, G. Ramírez-Rosales, I. Grijalva-Otero, C. López-Macías, R. A. García-Ramírez, C. Grajales-Muñiz, and J. M. Mejía-Aranguré. Quantifying the mortality caused by the H1N1 influenza virus during the 2009 pandemic in Mexico. *Journal of Infection in Developing Countries*, 8(6):742–748, 2014. ISSN 19722680. doi: 10.3855/jidc.3622.

[226] N. J. Kassebaum. Global, regional, and national burden of diseases and injuries for adults 70 years and older: Systematic analysis for the Global Burden of Disease 2019 Study. *The BMJ*, 376, 2022. ISSN 17561833. doi: 10.1136/bmj-2021-068208.

[227] C. Reed, J. M. Katz, K. Hancock, A. Balish, and A. M. Fry. Prevalence of Seropositivity to Pandemic Influenza A/H1N1 Virus in the United States following the 2009 Pandemic. *PLoS ONE*, 7(10), 2012. ISSN 19326203. doi: 10.1371/journal.pone.0048187.

[228] S. Stringhini, A. Wisniak, G. Piumatti, A. S. Azman, S. A. Lauer, H. Baysson, D. D. Ridder, D. Petrovic, S. Schrempft, K. Marcus, S. Yerly, I. A. Vernez, O. Keiser, S. Hurst, K. M. Posfay-Barbe, D. Trono, D. Pittet, L. Gétaz, F. Chappuis, I. Eckerle, N. Vuilleumier, B. Meyer, A. Flahault, L. Kaiser, and I. Guessous. Seroprevalence of anti-SARS-CoV-2 IgG antibodies in Geneva, Switzerland (SEROCoV-POP): a population-based study. *Lancet*, 396, 2020. doi: 10.1016/ S0140-6736(20)31304-0.

[229] J. Crèvecoeur, N. Hens, T. Neyens, Y. Larivière, B. Verhasselt, and H. Masson. Change in COVID19 outbreak pattern following vaccination in long-term care facilities in Flanders, Belgium. *Vaccine*, 40:6218–6224, 2022.

[230] P. W. Kelley, B. P. Petruccelli, P. Stehr-Green, R. L. Erickson, and C. J. Mason. The Susceptibility of Young Adult Americans to Vaccine-Preventable Infections: A National Serosurvey of US Army Recruits. *JAMA*, 266(19):2724–2729, 11 1991. ISSN 0098-7484. doi: 10.1001/jama.1991.03470190072032.

[231] R. E. Hope-Simpson. The Nature of Herpes Zoster: A Long-Term Study and a New Hypothesis. *Journal of the Royal Society of Medicine*, 58(1):9–20, 1965. ISSN 01410768. doi: 10.1177/003591576505800106.

[232] B. Ogunjimi, P. Van Damme, and P. Beutels. Herpes Zoster Risk Reduction through Exposure to Chickenpox Patients: A Systematic Multidisciplinary Review. *PLoS ONE*, 8(6), 2013. ISSN 19326203. doi: 10.1371/journal.pone.0066485.

[233] H. Forbes, I. Douglas, A. Finn, J. Breuer, K. Bhaskaran, L. Smeeth, S. Packer, S. M. Langan, K. E. Mansfield, R. Marlow, H. Whitaker, and C. Warren-Gash. Risk of herpes zoster after exposure to varicella to explore the exogenous boosting hypothesis: Self controlled case series study using UK electronic healthcare data. *The BMJ*, 368, 2020. ISSN 17561833. doi: 10.1136/bmj.l6987.

[234] K. Kawai, B. P. Yawn, P. Wollan, and R. Harpaz. Increasing incidence of herpes zoster over a 60-year period from a population-based study. *Clinical Infectious Diseases*, 63:221–226, 2016. ISSN 15376591. doi: 10.1093/cid/ciw296.

[235] K. Kawai, B. G. Gebremeskel, and C. J. Acosta. Systematic review of incidence and complications of herpes zoster: towards a global perspective. *BMJ Open*, 4(11), 2014. ISSN 20446055. doi: 10.1136/bmjopen-2014-004833.

[236] IPUMS. URL `https://www.ipums.org/ accessed 2023-06-22`.

[237] A. Rogers and L. J. Castro. Model migration schedules. Technical report, 1981.

[238] R. T. Michael. The Rise in Divorce Rates, 1960-1974: Age-Specific Components. *Demography*, 15(2):177–182, 1978.

[239] A. P. Krishnan and A. K. Kayani. Estimates of Age Specific Divorce Rates for Females in the United States , 1960-1969. *Journal of Marriage and Family*, 36 (1):72–76, 1974.

[240] Survey of Income and Program Participation (SIPP). URL `https://www.census.gov/programs-surveys/sipp.html accessed on 2023-06-26`.

[241] P. J. Smock. Cohabitation in the United States: An appraisal of research themes, findings, and implications. *Annual Review of Sociology*, 26:1–20, 2000. ISSN 03600572. doi: 10.1146/annurev.soc.26.1.1.

[242] S. R. Preblud and L. J. D'Angelo. From the center for disease control chickenpox in the United States, 1972-1977. *Journal of Infectious Diseases*, 140(2):257–260, 1979. ISSN 15376613. doi: 10.1093/infdis/140.2.257.

[243] M. A. Reynolds, D. Kruszon-Moran, A. Jumaan, D. S. Schmid, and G. M. McQuillan. Varicella seroprevalence in the U.S.: Data from the National Health and Nutrition Examination Survey, 1999-2004. *Public Health Reports*, 125(6): 860–869, 2010. ISSN 14682877. doi: 10.1177/003335491012500613.

[244] J. Schenk, S. Abrams, H. Theeten, P. V. Damme, P. Beutels, and N. Hens. Immunogenicity and persistence of trivalent measles, mumps, and rubella vaccines: a systematic review and meta-analysis Julie. *Lancet Infectious Diseases*, 21:286–295, 2021.

[245] C. Cohen, J. M. White, E. J. Savage, J. R. Glynn, Y. Choi, N. Andrews, D. Brown, and M. E. Ramsay. Vaccine effectiveness estimates, 2004-2005 mumps outbreak, England. *Emerging Infectious Diseases*, 13(1):12–17, 2007. ISSN 10806059. doi: 10.3201/eid1301.060649.

[246] M. J. Knol, A. T. Urbanus, E. M. Swart, L. Mollema, W. L. Ruijs, R. S. van Binnendijk, M. J. Te Wierik, H. E. de Melker, A. Timen, and S. J. Hahné. Large ongoing measles outbreak in a religious community in the Netherlands since may 2013. *Eurosurveillance*, 18(36):20580, 2013. ISSN 15607917. doi: 10.2807/1560-7917.es2013.18.36.20580.

[247] D. Antona, D. Lévy-Bruhl, C. Baudon, F. Freymuth, M. Lamy, C. Maine, D. Floret, and I. P. du Chatelet. Measles Elimination Efforts and 2008-2011 Outbreak, France. *Romanian Journal of Infectious Diseases*, 16(1):16–24, 2013. ISSN 20696051. doi: 10.3201/eid1903.121360.

[248] A. Siani. Measles outbreaks in Italy: A paradigm of the re-emergence of vaccine-preventable diseases in developed countries. *Preventive Medicine*, 121(January): 99–104, 2019. ISSN 10960260. doi: 10.1016/j.ypmed.2019.02.011.

[249] United Nations Population Division. Population facts: Household size and composition around the world. Technical report, United Nations. Department of Economic and Social Affairs. Population Division, 2017. URL `http://www.un.org/en/development/desa/population/publications/pdf/popfacts/PopFacts_2017-2.pdf`.

[250] V. Marziano, P. Poletti, G. Béraud, P. Y. Boëlle, S. Merler, and V. Colizza. Modeling the impact of changes in day-care contact patterns on the dynamics of varicella transmission in France between 1991 and 2015. *PLoS Computational Biology*, 14(8):1–13, 2018. ISSN 15537358. doi: 10.1371/journal.pcbi.1006334.

[251] B. Bigdelou, M. R. Sepand, S. Najafikhoshnoo, J. A. T. Negrete, M. Sharaf, J. Q. Ho, I. Sullivan, P. Chauhan, M. Etter, T. Shekarian, O. Liang, G. Hutter, R. Esfandiarpour, and S. Zanganeh. COVID-19 and Preexisting Comorbidities: Risks, Synergies, and Clinical Outcomes. *Frontiers in Immunology*, 13(May): 1–16, 2022. ISSN 16643224. doi: 10.3389/fimmu.2022.890517.

[252] J. L. Atkins, J. A. Masoli, J. Delgado, L. C. Pilling, C. L. Kuo, G. A. Kuchel, and D. Melzer. Preexisting Comorbidities Predicting COVID-19 and Mortality in the UK Biobank Community Cohort. *Journals of Gerontology - Series A Biological Sciences and Medical Sciences*, 75(11):2224–2230, 2020. ISSN 1758535X. doi: 10.1093/gerona/glaa183.

[253] E. Ge, Y. Li, S. Wu, E. Candido, and X. Wei. Association of pre-existing comorbidities with mortality and disease severity among 167,500 individuals with COVID-19 in Canada: A population-based cohort study. *PLoS ONE*, 16 (10 October):1–18, 2021. ISSN 19326203. doi: 10.1371/journal.pone.0258154.

[254] M. J. Divo, C. H. Martinez, and D. M. Mannino. Ageing and the epidemiology of multimorbidity. *European Respiratory Journal*, 44(4):1055–1068, 2014. ISSN 13993003. doi: 10.1183/09031936.00059814.

[255] B. C. Choi, H. Morrison, T. Wong, J. Wu, and Y. P. Yan. Bringing chronic disease epidemiology and infectious disease epidemiology back together. *Journal of Epidemiology and Community Health*, 61(9):832, 2007. ISSN 0143005X. doi: 10.1136/jech.2006.057752.

[256] A. Badawi, M. Drebot, and N. H. Ogden. Convergence of chronic and infectious diseases: a new direction in public health policy. *Canadian Journal of Public Health*, 110(4):523–524, 2019. ISSN 19207476. doi: 10.17269/s41997-019-00228-x.

[257] A. E. Raftery and H. Ševčíková. Probabilistic population forecasting: Short to very long-term. *International Journal of Forecasting*, 39(1):73–97, 2023. ISSN 01692070. doi: 10.1016/j.ijforecast.2021.09.001.

[258] L. Alkema, P. Gerland, A. Raftery, and J. Wilmoth. The United Nations Probabilistic Population Projections: An Introduction to Demographic Forecasting with Uncertainty. *Foresight (Colch)*, 2015(37):19–24, 2015. ISSN 1555-9068. URL http://www.ncbi.nlm.nih.gov/pubmed/26617476%0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4662414.

[259] N. Keilman, D. Q. Pham, and A. Hetland. Why population forecasts should be probabilistic - Illustrated by the case of Norway. *Demographic Research*, 6: 409–453, 2002. ISSN 14359871. doi: 10.4054/DemRes.2002.6.15.

[260] R. T. Konetzka, E. M. White, A. Pralea, D. C. Grabowski, and V. Mor. A systematic review of long-term care facility characteristics associated with COVID-19 outcomes. *Journal of the American Geriatrics Society*, 69(10):2766–2777, 2021. ISSN 15325415. doi: 10.1111/jgs.17434.

[261] I. Casado, I. Martínez-Baz, R. Burgui, F. Irisarri, M. Arriazu, F. Elía, A. Navascués, C. Ezpeleta, P. Aldaz, J. Castilla, I. Abad, J. Agreda, E. Álvarez, J. J. Arana, I. Arceiz, E. Arina, M. D. Artajo, A. Arza, K. Ayerdi, ..., and J. Castilla. Household transmission of influenza A(H1N1)pdm09 in the pandemic and post-pandemic seasons. *PLoS ONE*, 9(9), 2014. ISSN 19326203. doi: 10.1371/journal.pone.0108485.

[262] Z. J. Madewell, Y. Yang, I. M. Longini, M. E. Halloran, and N. E. Dean. Household Secondary Attack Rates of SARS-CoV-2 by Variant and Vaccination Status: An Updated Systematic Review and Meta-analysis. *JAMA Network Open*, 5(4):E229317, 2022. ISSN 25743805. doi: 10.1001/jamanetworkopen.2022.9317.

[263] O. Mondiale. New influenza A (H1N1) virus : global epidemiological situation, June 2009 = Nouveau virus grippal A (H1N1) : situation épidémiologique mondiale, juin 2009. *Weekly Epidemiological Record = Relevé épidémiologique hebdomadaire*, 84(25):249–257, 2009. URL https://apps.who.int/iris/handle/10665/241366.

[264] L. Van Wilder, R. Charafeddine, P. Beutels, R. Bruyndonckx, I. Cleemput, S. Demarest, D. De Smedt, N. Hens, A. Scohy, N. Speybroeck, J. Van der Heyden, R. T. Yokota, H. Van Oyen, J. Bilcke, and B. Devleesschauwer. Belgian population norms for the EQ-5D-5L, 2018. *Quality of Life Research*, 31(2): 527–537, 2022. ISSN 15732649. doi: 10.1007/s11136-021-02971-6. URL https://doi.org/10.1007/s11136-021-02971-6.

Appendix A

Incorporating human dynamic populations in models of infectious disease transmission: a systematic review - Appendix

A.1 Protocol for the systematic review of incorporating human dynamic populations in models of infectious disease transmission

A.1.1 Background

The host population in models of infectious disease transmission is typically based on rather strong assumptions regarding the demographic composition and how it changes over time. Stable populations, meaning that the relative age composition remains constant over time, are often used [15, 67]. Such assumptions have proven to be useful to gain epidemiological insights and may be justified for disease outbreaks taking place over shorter time frames during which no considerable demographic change is expected [27]. In reality, however, populations never reach stability as fertility and mortality levels are subject to continuous change. Furthermore, the demographic processes and their changes over time may influence social contact patterns with importance for disease transmission. This is acknowledged in an increasing number of mathematical and computational models for infectious diseases, which incorporate changes in the population structures over time. These models have shown an important impact of

demographic change on the dynamics of infectious diseases, as well as for the effectiveness of immunization programmes (e.g. [42, 47, 75]). Nevertheless, the included population dynamics and demographic modelling approach vary highly from model to model. The methods used to incorporate dynamic population structures in infectious disease models has to our knowledge never been systematically summarised.

A.1.2 Objectives

The objective of this systematic review is to summarise and discuss the methods that have been used to incorporate dynamic population structures into models for infectious disease transmission. This includes the methods used to model the host population, the different demographic processes considered, as well as the data and techniques used to model each demographic process. The questions to be answered in the systematic review are:

- Which methods are used to model a dynamic host population?
- Which demographic processes are explicitly incorporated in infectious disease models with a dynamic population?
- How are the demographic processes modelled and which data is used?

A.1.3 Methods

Eligibility criteria

We will search for publications of mathematical and computational models for infectious disease transmission in a human population that results from a model including at least fertility and mortality as dynamic processes. The demographic model can be included explicitly or a population from another source can be used as input to the disease transmission model as long as this population is the result of a demographic model explicitly considering dynamic trends for fertility and mortality. This implies that models assuming constant fertility or mortality rates throughout the entire study period are excluded. However, models with constant rates in a limited part of the study period are still included. The inclusion and exclusion criteria are listed below.

Inclusion criteria:

- Mathematical and computational models for infectious disease transmission among humans

- A dynamic population (observed or synthetic)
- Fertility and mortality are modelled as dynamic processes
- A population modelled with dynamic fertility and mortality trends by another source

Exclusion criteria:

- Less than five age groups
- Use of stable populations
- Fixed mortality or fertility rates throughout the entire study period
- Population modelled with intervals in time of 10 or more years (e.g. 2020, 2030)
- Models limited to high-risk groups (MSM community, injecting drug users)
- Non-communicable disease
- Technical papers or software tools without any application
- Reviews (unless it is a review of incorporation of demography in infectious disease models)

Information sources

We will search the electronic databases Web of Science and PubMed from the earliest date of the database to 25.08.2020. Additionally, we will carry out a manual search by screening reference lists of included papers.

Search strategy

The following search string is used to search in titles and abstracts:

- demography OR "demographic transition" OR "demographic change*" OR "population change*" OR "household structure*" OR "household composition*" OR "population ageing" OR "population aging" OR "aging population" OR "ageing population"
- AND (infect* OR vaccin* OR epidemic* OR communicable)
- AND (model* OR simulat*)
- NOT (animal* OR plant*)

A.1.4 Study records

Data management

We will make use of the reference software EndNote X9 to manage the identified publications.

Selection process

Titles and abstracts will be screened by using the inclusion and exclusion criteria stated above. Articles will be reviewed in full-text in case of doubt.

Data collection process

The reviewer will extract the data using a standardised form. Papers might be excluded at the data collection stage if it becomes apparent that inclusion criteria are not met or if there is not enough information in the paper to extract the required data.

Data items

The extracted data will include the following:

- Setting and population characteristics
  - Country/region/city
  - Population
  - Demographic characteristics (age, sex, etc.)
  - Time horizon
- Model specifications and data
  - Model type
  - Demographic processes considered
  - Source of demographic data
- Modelling of demographic processes
  - Starting population
  - Fertility

  - Mortality
  - Migration
  - Households and networks
  - Demographic sensitivity analyses
- Specifications of disease transmission model and analyses
  - Disease(s)
  - Vaccination
  - Social mixing
  - Cost-effectiveness analyses

A.1.5 Data analysis

The qualitative data analysis will include:

- A flowchart describing included and excluded articles
- Tables presenting the characteristics of the included articles with information regarding
  - Setting and population characteristics
  - Model specifications and data
  - Modelling of demographic processes included in study
  - Components of disease transmission model
- Figures visualising
  - Model types
  - Demographic processes included in model
  - Demographic data

A.2 Figures and tables

Table A.1. Search strategy and hits

#	**Search**	**Web of science: Abstract/title**	**PubMed: Abstract/title**
1	"demographic transition"	2.353	1.237
2	demography	27.772	10.98
3	"demographic change*"	6.641	3.117
4	"population change*"	5.31	2.137
5	"household structure*"	915	344
6	"household composition*"	1.037	519
7	("population ageing" OR "population aging")	4.561	2.663
8	("aging population" OR "ageing population")	11.861	10.28
9	infect*	1.899.379	1.784.209
10	vaccin*	349.222	318.058
11	epidemic*	117.329	108.101
12	communicable	12.809	18.047
13	model*	9.303.490	2.931.098
14	simulat*	3.292.060	535.675
15	1 OR 2 OR 3 OR 4 OR 5 OR 6 OR 7 OR 8	57.934	30.033
16	9 OR 10 OR 11 OR 12	2.175.981	2.055.072
17	13 OR 14	10.933.783	3.228.393
18	15 AND 16 AND 17	869	652
19	18 NOT (animal* OR plant*)	783	468

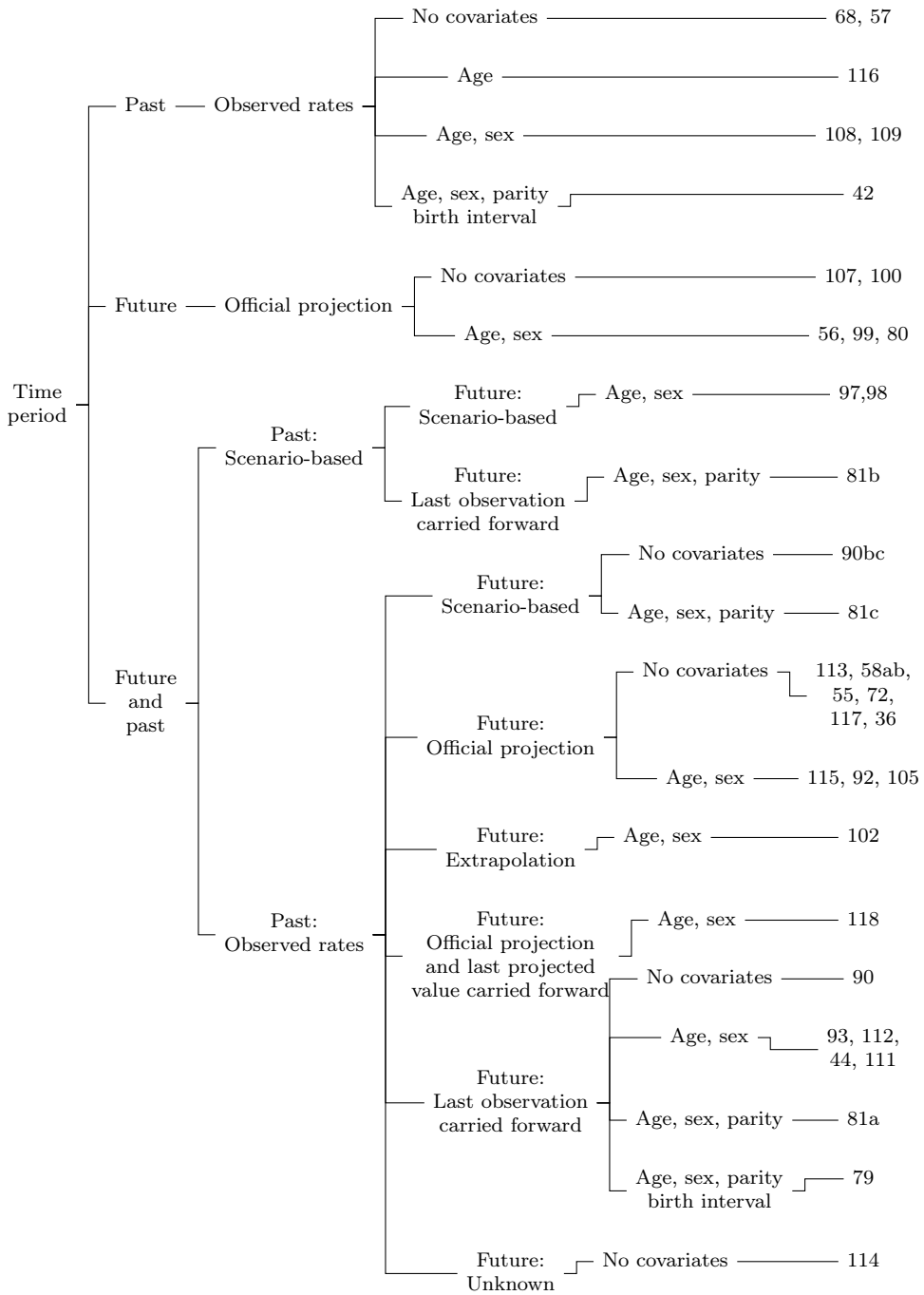

Figure A.1. Branching diagram of fertility modelling with article numbers.

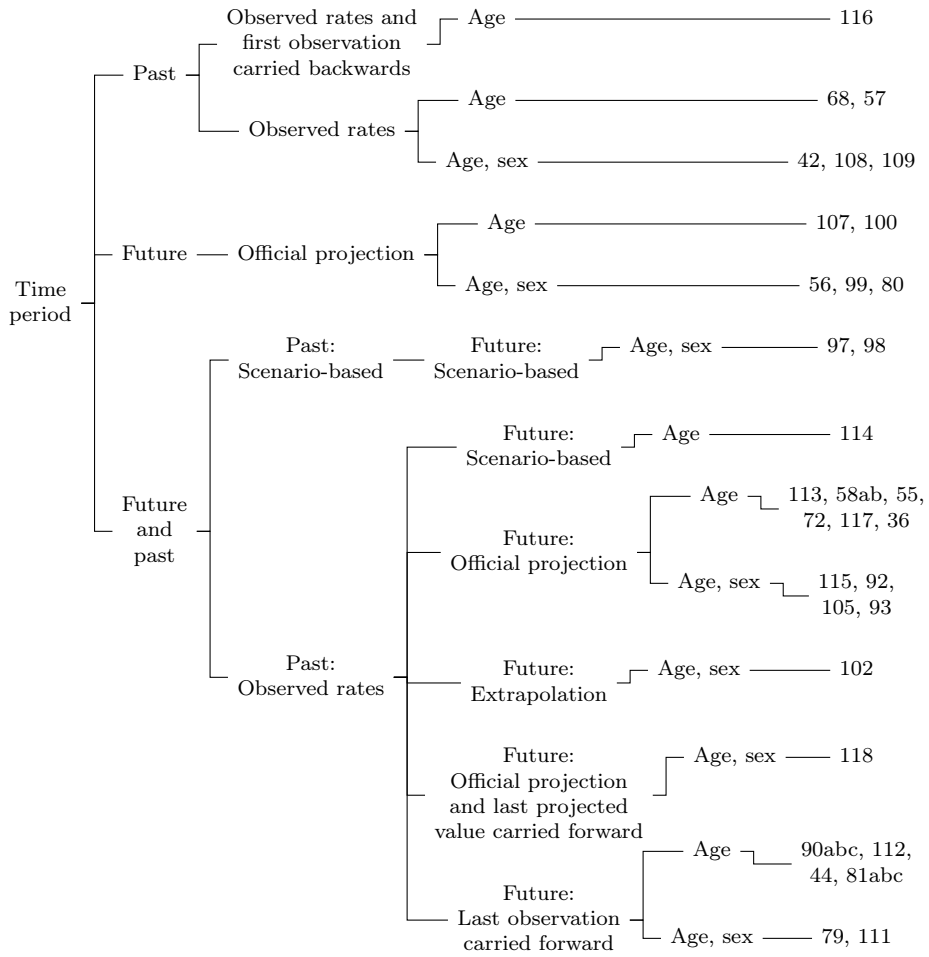

Figure A.2. Branching diagram of mortality modelling with article number.

Table A.2. IBMs and CCBMs included in systematic review (NA: Not applicable, obs.: Observed)

Article	Setting	Time horizon	Demographic characteristics	Demographic data source	Model type	Demographic processes	Initialisation period: rates
Mekonnen et al. [93]	Addis Ababa	1984-2004	Age, sex	Central Statistical Authority(Ethiopia), UN WPP	CCBM	Fertility, mortality, migration	NA
Gao and Hethcote [112]	China	1965-2051	Age, sex	National Bureau of Statistics (China), Population Information and Research Center (China)	CCBM	Fertility, mortality	NA
Aparicio and Castillo-Chavez [116]	USA (urban population)	1850-2009	Age	US Census Bureau	CCBM	Fertility, mortality, migration	Observation of 1850 (fixed)
Guzzetta et al. [58]a	Arkansas	2000-2030	Age	US Bureau of Census, US Bureau of Labor Statistics, US National Center for Education Statistics, CDC National Vital Statistics	CCBM	Fertility, mortality	NA
McDonald et al. [97]	Netherlands	2000-2030	Age, sex	Statistics Netherlands	CCBM	Fertility, mortality, migration	NA
Van Effelterre et al. [114]	Mexico	1970-2120	Age	US Census Bureau WHO	CCBM	Fertility, mortality	Adjusted to obtain a stable population
McDonald et al. [98]	Netherlands	2000-2030	Age, sex	Statistics Netherlands	CCBM	Fertility, mortality, migration	NA
Merler and Ajelli [68]	Italy	1901-2009	Age	ISTAT	CCBM	Fertility, mortality, migration	Observation of 1901 (fixed)
Mahy et al. [108]	Nigeria	1970-2013	Age, sex, state, rural/urban	UN WPP	CCBM	Fertility, mortality, migration	NA
Knight et al. [107]	91 LMIC	2009-2050	Age	UN WPP	CCBM	Fertility, mortality	Observation of 2009 (fixed)
Penazzato et al. [92]	21 African countries	2010-2020	Age, sex	UN WPP	CCBM	Fertility, mortality, migration	NA
Van Effelterre et al. [55]	Thailand	1950-2100	Age	UN WPP, WHO	CCBM	Fertility, mortality, internal migration	Adjusted to obtain a stable population
Costantino et al. [57]	Australia	1901-2000	Age	Australian Bureau of Statistics	CCBM	Fertility, mortality	Observation of 1901 (fixed)
Trentini et al. [72]	Australia, Ethiopia, Kenya, Ireland, Italy, South Korea, Singapore, UK, USA	1950-2030	Age	UN WPP	CCBM	Fertility, mortality, migration	Observation of 1950 (fixed)
Van Effelterre et al. [113]	Brazil, Mexico	1950-2050	Age	UN WPP, WHO	CCBM	Fertility, mortality, internal migration	Adjusted to obtain a stable population
Mahy et al. [109]	160 countries	1970-2015	Age, sex	UN WPP	CCBM	Fertility, mortality, migration	NA

Table A.2 IBMs and CCBMs included in systematic review (NA: Not applicable, obs.: Observed) *continued*

Article	Setting	Time horizon	Demographic characteristics	Demographic data source	Model type	Demographic processes	Initialisation period: Rates
Williams et al. [118]	Senegal and Gambia as one population	1950-2150	Age, sex	UN WPP	CCBM	Fertility, mortality	Fertility obs. 1950/55, female life table with net reproductive rate of 1
Jayasundara et al. [117]	Australia	1901-2061	Age	Australian Bureau of Statistics	CCBM	Fertility, mortality, migration	Observations of 1901 (fixed)
Mahikul et al. [44]	Thailand	1980-2035	Age, sex, internal migration group	Population and Housing Census, Ministry of Public Health	CCBM	Fertility, mortality, migration	NA
Haacker et al. [56]	Botswana	2015-2050	Age, sex	UN WPP	CCBM	Fertility, mortality, migration	NA
Heffernan et al. [115]	190 countries	1950-2100	Age, sex	UN WPP	CCBM	Fertility, mortality, migration	Observation of 1950 (fixed)
Khalifa et al. [105]	148 countries	2010-2050	Age, sex	UN WPP, National census data	CCBM	Fertility, mortality, migration	NA
Ku and Dodd [102]	Taiwan	2000-2035	Age, sex	Department of Statistics (Taiwan)	CCBM	Fertility, mortality, migration	NA
Trentini et al. [100]	USA, South Korea, Singapore, Australia, Italy, UK, Ireland	2018-2050	Age	UN WPP	CCBM	Fertility, mortality	NA
Yerushalmi et al. [99]	Ghana	2015-2045	Age, sex	UN WPP, DHS	CCBM	Fertility, mortality, migration	NA
Guzzetta et al. [58]b	Arkansas	2000-2030	Age, household membership, school, workplace	US Bureau of Census, US Bureau of Labor Statistics, US National Center for Education Statistics, CDC National Vital Statistics	IBM	Fertility, mortality, household transitions, school/workplace transitions	NA
Liu et al. [81]a	China	1975-2024	Age, sex, household membership, school membership, mother link	Census data, Yearbooks for China	IBM	Fertility, mortality, household transitions, school transitions	NA
Liu et al. [81]b	China	1975-2024	Age, sex, household membership, school membership, mother link	Census data, Yearbooks for China	IBM	Fertility, mortality, household transitions, school transitions	NA
Liu et al. [81]c	China	1975-2024	Age, sex, household membership, school membership, mother link	Census data, Yearbooks for China	IBM	Fertility, mortality, household transitions, school transitions	NA

Table A.2 IBMs and CCBMs included in systematic review (NA: Not applicable, obs.: Observed) *continued*

Article	Setting	Time horizon	Demographic characteristics	Demographic data source	Model type	Demographic processes	Initialisation period: Rates
Marziano et al. [90]a	Spain	1900-2050	Age	UN WPP, HMD INEbase, Eurostat, World Bank	IBM	Fertility, mortality, migration	Observation of 1900 (fixed)
Marziano et al. [90]b	Spain	1900-2050	Age	UN WPP, HMD INEbase, Eurostat, World Bank	IBM	Fertility, mortality, migration	Observation of 1900 (fixed)
Marziano et al. [90]c	Spain	1900-2050	Age	UN WPP, HMD INEbase, Eurostat, World Bank	IBM	Fertility, mortality, migration	Observation of 1900 (fixed)
Geard et al. [42]	Australia	1910-2010	Age, sex, household membership	Australian Bureau of Statistics, Survey data	IBM	Fertility, mortality, migration, household transitions	Observations of 1910 (fixed)
Xu et al. [80]	American Samoa	2010-2050	Age, sex, household membership, residence, household position	US Census Bureau American Samoa Statistical Yearbook 2015	IBM	Fertility, mortality, migration, household transitions	NA
Campbell et al. [79]	Australia	1910-2020	Age, sex, household membership	Australian Bureau of Statistics, Survey data	IBM	Fertility, mortality, household transitions	Observations of 1910 (fixed)
Melegaro et al. [36]	Italy	1900-2100	Age	HMD, MPIRDR, ISTAT	IBM	Fertility, mortality, migration	Observation of 1900 (fixed)
Smit et al. [111]	Zimbabwe	1950-2035	Age, sex	UN WPP	IBM	Fertility, mortality	NA

Table A.3. EPMs included in systematic review (NA: Not applicable, obs.: Observed)

Article	Setting	Time horizon	Demographic characteristics	Demographic data source	Initialisation period: Rates	Population composition
Eichner et al. [37]	Germany	1993-2043	Age	DESTATIS	Adjusted to keep age distribution constant	Obs., official projection
van Lier et al. [119]	The Netherlands	1950-2200	Age	Statistics Netherlands		Obs., official projection, official projection carried forward
Sibley et al. [110]	15 countries (Europe, Asia)	1950-2030	Age, sex	UN WPP, National databases	NA	Obs. and official projection
Schmidt-Ott et al. [103]	Belgium, Finland, Germany, Great Britain, Italy, Luxembourg, Netherlands, Poland	1994-2033	Age	DESTATIS, Eurostat	Adjusted to keep age distribution constant	Obs., official projection
Dolk et al. [106]	Germany	1994-2034	Age	DESTATIS	NA	Obs., official projection
Li et al. [95]	Six provinces in China	1982-2011	Age	National census	NA	Obs. (1982, 1990, 2000, 2010) with linear interpolation
Eichner et al. [94]	Germany	2000-2026	Age	DESTATIS	NA	Obs., official projection (main scenario)
Hood et al. [101]	USA	2013-2045	Age, sex, ethnicity/race	US Census Bureau	NA	Official projection
Horn et al. [75]a	Germany	1990-2060	Age, sex	DESTATIS	NA	Obs., official projection (medium scenario)
Horn et al. [75]b	Germany	1990-2060	Age, sex	DESTATIS	NA	Obs., official projection (high migration 2016-2025)
Arregui et al. [59]a	12 countries (Africa, Asia)	2000-2050	Age	UN WPP	NA	Obs., official projection
Arregui et al. [59]b	12 countries (Africa, Asia)	2015-2050	Age	UN WPP	NA	Obs., synthetic (young, static, ageing)
Turgeon et al. [96]	Provinces/territories in Canada	1999-2028	Age	Statistics Canada	NA	Obs., official projection (medium growth scenario)
Schmidt-Ott et al. [104]	Germany	1997-2036	Age	DESTATIS	Adjusted to keep age distribution constant	Obs., official projection
Marziano et al. [47]	Italy	2017-2045	Age, household membership	ISTAT	NA	Official projection of age distribution, household size and age composition of 2017
Talbird et al. [73]	USA (age 50+)	2017-2046	Age	US Census Bureau	NA	Official projection

Table A.4. Modelling of fertility (TFR: total fertility rate, ASFR: age-specific fertility rate, CBR: crude birth rate, obs.: observation, NA: not applicable)

Article	Measure	Past time period	Future time period	Covariates
Mekonnen et al. [93]	TFR, age distribution (model schedule), disease state	Obs. (5-year average)	Last obs. carried forward	Age, sex, disease state
Gao and Hethcote [112]	ASFR	Obs. with interpolation	Last obs. carried forward	Age, sex
Aparicio and Castillo-Chavez [116]	Age-specific per capita birth rate	Obs. and linear interpolation	NA	Age
Guzzetta et al. [58]a	CBR	Obs.	Official projection	None
McDonald et al. [97]	ASFR	Obs. for 2000 with a fixed annual increase for ages >27	Obs. for 2000 with a fixed annual increase for ages >27	Age, sex
Van Effelterre et al. [114]	Number of births	Obs.	Uncertain. Author no longer has access to information.	None
McDonald et al. [98]	ASFR	Obs. for 2000 with a fixed annual increase for ages >27	Obs. for 2000 with a fixed annual increase for ages >27	Age, sex
Merler and Ajelli [68]	CBR	Obs. and linear interpolation	NA	None
Mahy et al. [108]	TFR, age distribution, state	Obs.	NA	Age, sex, state
Knight et al. [107]	Number of births	NA	Official projection	None
Penazzato et al. [92]	TFR, age distribution (model schedule), disease state	Obs.	Official projection	Age, sex, disease state
Van Effelterre et al. [55]	Number of births	Obs. (5-year average)	Official projection (5-year average)	None
Costantino et al. [57]	CBR	Obs.	NA	None
Trentini et al. [72]	CBR	Obs.	Official projection	None
Van Effelterre et al. [113]	Number of births	Obs. with smoothing	Official projection with smoothing	None
Mahy et al. [109]	TFR, age distribution (model schedule)	Obs.	NA	Age, sex
Williams et al. [118]	ASFR	Obs. averaged over countries	Official projections averaged over countries, last projected value carried forward	Age, sex
Jayasundara et al. [117]	CBR	Obs.	Official projection	None
Mahikul et al. [44]	ASFR	Obs.	Last obs. carried forward	Age, sex
Haacker et al. [56]	TFR, age distribution	NA	Official projection (5-year average)	Age, sex
Heffernan et al. [115]	ASFR	Obs. (5-year average)	Official projection (5-year average)	Age, sex
Khalifa et al. [105]	TFR, age distribution	Obs. (5-year average)	Official projection (5-year average)	Age, sex
Ku and Dodd [102]	ASFR	Obs.	Projection with Lee-Carter model	Age, sex
Trentini et al. [100]	CBR	NA	Official projection	None
Yerushalmi et al. [99]	TFR, age distribution	NA	Official projection	Age, sex
Guzzetta et al. [58]b	CBR	Obs.	Official projection	None

Table A.4 Modelling of fertility (TFR: total fertility rate, ASFR: age-specific fertility rate, CBR: crude birth rate, obs.: observation, NA: not applicable) *continued*

Article	Measure	Past time period	Future time period	Covariates
Liu et al. [81]a	Age-, parity- and policy -specific fertility rate	Obs.	Last obs. carried forward	Age, sex, parity
Liu et al. [81]b	Age-, parity- and policy -specific fertility rate	Obs. excl. 2nd order births	Last obs. carried forward	Age, sex, parity
Liu et al. [81]c	Age-, parity- and policy -specific fertility rate	Obs.	Last obs. for 1st order births carried forward, 2nd order birth rate increases to 1st order birth rate	Age, sex, parity
Marziano et al. [90]a	CBR	Obs.	Last obs. carried forward	None
Marziano et al. [90]b	CBR	Obs.	Linear decrease to 0	None
Marziano et al. [90]c	CBR	Obs.	Linear increase to doubling of last obs.	None
Geard et al. [42]	Number of births, age distribution, birth interval of minimum 270 days	Obs.	NA	Age, sex, parity, birth interval
Xu et al. [80]	ASFR	NA	Official projection	Age, sex
Campbell et al. [79]	Number of births, age distribution, birth interval of minimum 270 days	Obs.	Last obs. carried forward	Age, sex, parity, birth interval
Melegaro et al. [36]	CBR	Obs.	Official projection	None
Smit et al. [111]	ASFR	Obs. (5-year average)	Last obs. carried forward	Age, sex

Table A.5. Modelling of mortality (ASMR: age-specific mortality rate, ASSMR: age-sex-specific mortality rate, obs.: observation, NA: not applicable)

Article	Measure	Past time period	Future time period	Covariates
Mekonnen et al. [40]	Life expectancy at birth, model life table, disease-related mortality	Obs.	Official projection	Age, sex, disease state
Gao et al. [41]	ASMR	Obs. with interpolation	Last obs. carried forward	Age
Aparicio et al. [42]	Crude mortality rate, life table, disease-related mortality	First obs. carried backwards (1850-1899), function fitted to obs.	NA	Age, disease state
Guzzetta et al. [43]a	ASMR, disease-related mortality	Obs.	Official projection	Age, disease state
McDonald et al. [44]	ASSMR	Obs. 2000 with a fixed annual age-specific decrease	Obs. 2000 with a fixed annual age-specific decrease	Age, sex
Van Effelterre et al. [45]	ASMR	Obs. with linear interpolation	Last obs. with annual age-specific adjustment	Age
McDonald et al. [46]	ASSMR	Obs. 2000 with a fixed annual age-specific decrease	Obs. 2000 with a fixed annual age-specific decrease	Age, sex
Merler et al. [10]	ASMR	Obs. and linear interpolation	NA	Age
Mahy et al. [47]	Life expectancy at birth by sex, model life table, disease-related mortality, state	Obs.	NA	Age, sex, state, disease state
Knight et al. [48]	ASMR	NA	Official projection	Age, disease state
Penazzato et al. [49]	Life expectancy at birth, model life table, disease-related mortality	Obs.	Official projection	Age, sex, disease state
Van Effelterre et al. [50]	ASMR	Obs.	Exponentially decreasing function fitted to official projections	Age
Costantino et al. [51]	ASMR	Obs.	NA	Age
Trentini et al. [52]	ASMR	Obs. with linear interpolation	Official projection with linear interpolation	Age
Van Effelterre et al. [25]	ASMR	Obs.	Official projection	Age
Mahy et al. [53]	Life expectancy at birth, model life table, disease-related mortality	Obs.	NA	Age, sex, disease state
Williams et al. [54]	ASSMR, disease-related mortality	Obs. averaged over countries	Official projections averaged over countries, last projected value carried forward	Age, sex, disease state
Jayasundara et al. [55]	ASMR	Obs. (averages over multiple years) with interpolation	Official projection	Age
Mahikul et al. [56]	ASMR, disease-related mortality	Obs.	Last obs. carried forward	Age, disease state

Table A.5 Modelling of mortality (ASMR: age-specific mortality rate, ASSMR: age-sex-specific mortality rate, obs.: observation, NA: not applicable) *continued*

Article	Measure	Past time period	Future time period	Covariates
Haacker et al. [57]	Life expectancy at birth by sex, model life table, disease-related mortality	NA	Official projection (5-year average)	Age, sex, disease state
Heffernan et al. [26]	ASSMR, disease-related mortality, risk group mortality	Obs. (5-year average)	Official projection (5-year average)	Age, sex, disease state, risk group
Khalifa et al. [58]	Life expectancy at birth by sex, model life table, disease-related mortality	Obs. (5-year average)	Official projection (5-year average)	Age, sex, disease state
Ku et al. [59]	ASMR	Obs.	Projection with Lee-Carter and Coale-Kisker method	Age, sex
Trentini et al. [60]	ASMR	NA	Official projection	Age
Yerushalmi et al. [61]	ASSMR, disease-related mortality, regional variation	NA	Official projection	Age, sex, disease state, region
Guzzetta et al. [43]b	ASMR, disease-related mortality	Obs.	Official projection	Age, disease state
Liu et al. [62]a	ASMR	Obs.	Last obs. carried forward	Age
Liu et al. [62]b	ASMR	Obs.	Last obs. carried forward	Age
Liu et al. [62]c	ASMR	Obs.	Last obs. carried forward	Age
Marziano et al. [16]a	ASMR	Obs.	Last obs. carried forward	Age
Marziano et al. [16]b	ASMR	Obs.	Last obs. carried forward	Age
Marziano et al. [16]c	ASMR	Obs.	Last obs. carried forward	Age
Geard et al. [9]	ASSMR	Obs.	NA	Age, sex
Xu et al. [63]	Number of deaths, life table by sex (fixed)	NA	Official projection	Age, sex
Campbell et al. [64]	ASSMR	Obs.	Last obs. carried forward	Age, sex
Melegaro et al. [65]	ASMR	Obs.	Official projection	Age
Smit et al. [66]	ASSMR, disease-related mortality	Obs. (5-year average)	Last obs. carried forward	Age, sex, disease state

Table A.6. Modelling of migration (obs.: observation, , NA: not applicable)

Article	Measure	Past time period	Future time period	Covariates
Mekonnen et al. [93]	Net-migration rate, model age-sex schedule	Obs. of 1984 (fixed)	Last obs. carried forward	Age, sex
Aparicio and Castillo-Chavez [116]	Net-migration rates (urban population)	Function fitted to obs.	NA	None
McDonald et al. [97]	Age-sex-specific net-migration rate	Average obs. 2000-2009	Average obs. carried forward	Age, sex
McDonald et al. [98]	Age-sex-specific net-migration rate	Average obs. 2000-2009	Average obs. carried forward	Age, sex
Merler and Ajelli [68]	Immigration and emigration rates, age distribution (fixed)	Obs.	NA	Age
Mahy et al. [108]	Net-migration rate, model age-sex schedule	Obs.	NA	Age, sex
Penazzato et al. [92]	Net-migration rate, model age-sex schedule	Obs.	Official projection	Age, sex
Van Effelterre et al. [55]	Internal migration rate (urban/rural)	Obs. (5-year average)	Official projection (5-year average)	None
Trentini et al. [72]	Net-migration rate, age distribution (3 time points only)	Obs. (5-year averages)	Official projection (5-year averages)	Age
Van Effelterre et al. [113]	Internal migration rate (urban/rural), net international migration rate (Mexico)	Obs. with smoothing	Official projection with smoothing	None
Mahy et al. [109]	Net-migration rate, model age-sex schedule	Obs.	NA	Age, sex
Jayasundara et al. [117]	Net-migration rate, age distribution (average 1976-2015)	Obs.	Official projection	Age
Mahikul et al. [44]	Net international migration rate, internal migration rate	Obs.	Last obs. carried forward	Age
Haacker et al. [56]	Net-migration rate, model age-sex schedule	NA	Official projection (5-year average)	Age, sex
Heffernan et al. [115]	Net-migration rate	Obs. (5-year average)	Official projection (5-year average)	None
Khalifa et al. [105]	Net-migration rate, model age-sex schedule	Obs. (5-year average)	Official projection (5-year average)	Age, sex
Ku and Dodd [102]	Age-sex-specific net-migration rate	Obs. with residual method	Mean of obs. carried forward	Age, sex
Yerushalmi et al. [99]	Net-migration rate, Model age-sex schedule	NA	Official projection	Age, sex
Marziano et al. [90]a	Net-migration rate, age distribution (fixed)	Obs.	Last obs. carried forward	Age
Marziano et al. [90]b	Net-migration rate, age distribution (fixed)	Obs.	Last obs. carried forward	Age
Marziano et al. [90]c	Net-migration rate, age distribution (fixed)	Obs.	Last obs. carried forward	Age
Geard et al. [42]	Net-migration rate	Fixed rate from 1950	NA	None

Table A.6 Modelling of migration (obs.: observation, , NA: not applicable) *continued*

Article	Measure	Past time period	Future time period	Covariates
Xu et al. [80]	Net-migration rate, age distribution (fixed)	NA	Official projection	Age
Melegaro et al. [36]	Net-migration rate	Obs.	Official projection	Age

Table A.7. Households and networks (NA: not applicable)

Article	Household types	Transitions	Transition rates	Other networks
Guzzetta et al. [58]b	Single person household, married couple with/without child(ren), single(s) with/without child(ren), other households	Marriage,divorce, leaving current household to create a new	Dynamic	Schools, workplaces membership, spatial location of households, schools and workplaces
Liu et al. [81]a	Union with/without child(ren), single adult with/without child(ren), non-family related individuals with/without nuclear family	Union formation, leaving parental household	Fixed over time and equal across eligible ages	School links
Liu et al. [81]b	Union with/without child(ren), single adult with/without child(ren), non-family related individuals with/without nuclear family	Union formation, leaving parental household	Fixed over time and equal across eligible ages	School links
Liu et al. [81]c	Union with/without child(ren), single adult with/without child(ren), non-family related individuals with/without nuclear family	Union formation, leaving parental household	Fixed over time and equal across eligible ages	School links
Geard et al. [42]	Union with/without child(ren), single parent with child(ren), single	Union formation and dissolution, child leaving parental household	Fixed and equal across eligible ages	NA
Xu et al. [80]	Couple with/without child(ren), single with/without child(ren), non-related adult(s), non-related adult(s) living with nuclear family, multi-generational household, institutions	Union formation and dissolution, senior individuals moving to household of adult child, single person household to institution	Fixed and equal across ages	Household are assigned a location while ensuring plausible living space per capita
Campbell et al. [79]	Union with/without child(ren), single parent with child(ren), single	Union formation and dissolution, child leaving parental household	Time-dependent but equal across eligible ages	NA
Marziano et al. [47]	Couple with/without child(ren), single adult with child(ren), household size 1-7	New households are generated each year according to the household size and age composition of 2017.	NA	NA

Table A.8. Sensitivity analysis

Article	Demographic processes	Time period	Method
Mekonnen et al. [93]	Fertility	2000-2024	Scenario-based (low/high)
Van Effelterre et al. [55]	Fertility and internal migration	1950-2100	Interpolation between 5-year averages
Williams et al. [118]	Fertility	2015-2100	Scenario-based
Ku and Dodd [102]	Fertility, mortality, migration	2018-2035	Prediction intervals
Marziano et al. [90]a	Fertility, mortality, migration	2010-2050	Scenario-based (UN medium variant)
Marziano et al. [90]b	Fertility, mortality, migration	2010-2050	Scenario-based (UN medium variant)
Marziano et al. [90]c	Fertility, mortality, migration	2010-2050	Scenario-based (UN medium variant)
Melegaro et al. [36]	Fertility	2016-2100	Scenario-based
Arregui et al. [59]a	Age distribution	2015-2050	Confidence interval, alternative age distributions (young/old population)

Table A.9. Infectious disease modelling (CEA: cost-effectiveness analysis, NA: not applicable)

Article	Disease	Vaccination	Social mixing	CEA
Mekonnen et al. [93]	HIV/AIDS	No	No	No
Gao and Hethcote [112]	Rubella	Yes	Proportionate mixing	NA
Aparicio and Castillo-Chavez [116]	Tuberculosis	NA	NA	NA
Guzzetta et al. [58]a	Tuberculosis	NA	Homogeneous mixing	NA
McDonald et al. [97]	Hepatitis B, influenza	Yes	NA	NA
Van Effelterre et al. [114]	Hepatitis A	Yes	WAIFW matrix	NA
McDonald et al. [98]	Hepatitis A	Yes	NA	NA
Merler and Ajelli [68]	Measles	Yes	Homogeneous mixing and sensitivity analysis (POLYMOD)	NA
Mahy et al. [108]	HIV/AIDS	No	No	No
Knight et al. [107]	Tuberculosis	Yes	No	Yes
Penazzato et al. [92]	HIV/AIDS	No	No	No
Van Effelterre et al. [55]	Hepatitis A	NA	Homogeneous mixing	NA
Costantino et al. [57]	varicella zoster virus	NA	POLYMOD	NA
Trentini et al. [72]	Measles	Yes	Homogeneous mixing	NA
Van Effelterre et al. [113]	Hepatitis A	NA	Homogeneous mixing	NA
Mahy et al. [109]	HIV/AIDS	No	No	No
Williams et al. [118]	Hepatitis B	Yes	WAIFW matrix	NA
Jayasundara et al. [117]	Hepatitis A	Yes	POLYMOD	NA
Mahikul et al. [44]	Melioidosis	NA	NA	NA
Haacker et al. [56]	HIV and non-communicable diseases	NA	NA	NA
Heffernan et al. [115]	Hepatitis C	NA	NA	NA
Khalifa et al. [105]	HIV	NA	NA	NA
Ku and Dodd [102]	Tuberculosis	NA	NA	NA
Trentini et al. [100]	Measles	Yes	Homogeneous mixing	NA
Yerushalmi et al. [99]	Malaria	Yes	NA	Yes
Guzzetta et al. [58]b	Tuberculosis	No	Heterogeneous mixing	NA
Liu et al. [81]a	Influenza	No	Household contacts, school contacts and social contacts.	NA
Liu et al. [81]b	Influenza	No	Household contacts, school contacts and social contacts.	NA
Liu et al. [81]c	Influenza	No	Household contacts, school contacts and social contacts.	NA
Marziano et al. [90]a	varicella zoster virus	Yes	Synthetic social contact matrices	No

Table A.9 Infectious disease modelling (CEA: cost-effectiveness analysis, NA: not applicable) *continued*

Article	Disease	Vaccination	Social mixing	CEA
Marziano et al. [90]b	varicella zoster virus	Yes	Synthetic social contact matrices	No
Marziano et al. [90]]c	varicella zoster virus	Yes	Synthetic social contact matrices	No
Geard et al. [42]	"measles-like" illness	Yes	Dynamic contact matrix based on POLYMOD	NA
Xu et al. [80]	Lymphatic filariasis (not modelled)	No	NA	NA
Campbell et al. [79]	Pertussis	Yes	Dynamic contact matrix based on POLYMOD	NA
Melegaro et al. [36]	Varicella and herpes zoster	Yes	Synthetic social contact matrices	Yes
Smit et al. [111]	HIV, non-communicable diseases	No	NA	NA
Eichner et al. [37]	Influenza	Yes	Dynamic contact matrix based on POLYMOD	NA
van Lier et al. [119]	Varicella and herpes zoster	Yes	Yes (POLYMOD)	Yes
Sibley et al. [110]	Hepatitis C	No	NA	No
Schmidt-Ott et al. [103]	Influenza	Yes	Dynamic contact matrix based on POLYMOD	NA
Dolk et al. [106]	Influenza	Yes	Dynamic contact matrix based on POLYMOD	Yes
Li et al. [95]	Measles	Yes	No	No
Eichner et al. [94]	Influenza	Yes	Yes (POLYMOD)	No
Hood et al. [101]	HIV	No	No	No
Horn et al. [75]a	Varicella zoster virus	Yes	POLYMOD	NA
Horn et al. [75]b	Varicella zoster virus	Yes	POLYMOD	NA
Arregui et al. [59]a	Tuberculosis	No	Heterogeneous mixing	No
Arregui et al. [59]b	Tuberculosis	No	Heterogeneous mixing	No
Turgeon et al. [96]	Salmonella	No	No	No
Schmidt-Ott et al. [104]	Influenza	Yes	Dynamic contact matrix based on POLYMOD	NA
Marziano et al. [47]	Measles	Yes	POLYMOD	NA
Talbird et al. [73]	Influenza, pertussis, herpes zoster, and pneumococcal disease	Yes	No	Yes

Appendix B

Demographic microsimulation - Appendix

B.1 Initial population

B.1.1 Household position

All individuals in the population have a LIPRO household position for 2011 and 2012, which describes the relation an individual has to other household members and/or the type of household [149]. The LIPRO typology contains the following categories:

- CMAR: Child in family with married parents
- CUMR: Child in family with cohabiting parents
- C1PA: Child in one-parent family
- SING: Single (one-person household)
- MAR0: Married, living with spouse but without children
- MAR+: Married, living with spouse and one or more children
- UNM0: Cohabiting, no children present
- UNM+: Cohabiting with one or more children
- H1PA: Head of one-parent family
- NFRA: Non-family related adult (adult living with MAR0/MAR+/UNMO/ UNM+/H1PA)

- OTHR: Other (e.g. multiple single adults living together) Collective: Member of collective household

The original LIPRO household positions are modified in the new variable household position. Each category of this variable is described below as well as the included LIPRO positions. For some categories, LIPRO positions are only included in a given household position if certain conditions involving other variables are fulfilled.

- Child *(child)*
  - Individual living in parental household without own children or partner
  - LIPRO positions: CMAR, CUNM, C1PA
  - LIPRO positions with conditions:
    - NFRA/OTHR if individuals is younger than 16
- Union without child *(union)*
  - Individual living together with partner and without children
  - LIPRO positions: MAR0, UNM0
  - LIPRO positions with conditions:
    - CMAR/CUNM/C1PA/NFRA if individual is in a union in 2012 (2011) with a partner that also was a household member in 2011 (2012)
    - CMAR/CUNM/C1PA/NFRA/OTHR/H1PA if two individuals living in the same household are parents to the same child (no longer in household)
- Union with child *(union+)*
  - Individual living together with partner and child
  - LIPRO positions: MAR+, UNM+
  - LIPRO positions with conditions:
    - CMAR/CUNM/C1PA/NFRA if individual is in union in 2012 (2011) with a partner that also was a household member in 2011 (2012) and a (step-)child is present in the household (further description in subsection 3.3.3 in Chapter 3)

        * CMAR/CUNM/C1PA/NFRA/OTHR/H1PA if two individuals living in the same household are parents to the same child and a (step-)child is present in the household (further description in subsection 3.3.3 in Chapter 3)

- Single-person household *(single)*
  - Individual living in a one-person household
  - LIPRO position: SING
  - LIPRO positions with conditions:
    * All LIPRO positions if household of size 1
- Non-family related adult *(NFRA)*
  - Individual living without own family nucleus but living in same household as unrelated family nucleus
  - LIPRO position: NFRA
- Other *(Other)*
  - Individual living together with other unrelated individuals
  - LIPRO position: OTHR
- Collective household *(collective)*
  - Individual living in collective household (prison, special care facility, nursing homes, student accommodation,...)
  - LIPRO position: Collective
  - LIPRO positions with conditions:
    * All LIPRO positions if household members have category 'collective'
- Single parent *(single+)*
  - Individual living together with their child but without a partner in the household
  - LIPRO position: H1PA
  - LIPRO positions with conditions:

        * NFRA/OTHR/CMAR/CUNM/C1PA/MAR+/UNM+ if individual without a partner or with a missing partner match (further description in subsection B.1.2) lives in the same household as their child
    - Single parents living in their own parental household are referred to as single+* in cases where the distinction is necessary
- Union in multigenerational household *(multi_U)*
    - Individual living together with partner, child and family nucleus of child (e.g. grandchild and/or partner of child)
- Single in multigenerational household *(multi_S)*
    - Individual without partner living with child and family nucleus of child (e.g. grandchild and/or partner of child)

In Table B.1, the categories in the new variable household position are broken down by LIPRO household position for the census population. Differences between the two variables are especially seen for the LIPRO positions *NFRA* and *OTHR*.

Table B.1. Household position vs. LIPRO household position in Belgian census 2011.

		LIPRO household position											
		SING	MAR0	MAR1	CMAR	UNM0	UNM1	CUMR	H1PA	C1PA	NFRA	OTHR	Collective
Household position	Child	28	2	0	2065579	406	83	460206	1	693979	62715	18581	656
	Collective	0	0	0	0	0	0	0	0	10	0	39	127021
	Multi_S	0	0	2	32	6	32	22	15128	72	22396	130	3
	Multi_U	0	112	41438	23	26	4638	20	17	52	7363	48	0
	NFRA	0	0	0	18	0	0	13	0	8	69709	0	0
	Other	0	6	0	0	120	15	0	0	0	0	125519	0
	Single	1589161	0	0	0	35	24	3	4	3	42	68	1510
	Single+	0	2	6	9916	405	199	1888	440286	8534	4434	12442	3
	Union	0	1915841	579	2472	488299	287	55	2	1462	5424	366	0
	Union+	0	6571	2161670	9338	1069	544300	593	939	5464	17169	4127	0

B.1.2 Partner matching

Errors cause a smaller number of unions between parent and child (521), which are changed to more appropriate categories in the variable household position. Moreover, no partner is initially found for 1,700 individuals because no one else in the household has position *union(+)*, but 1,396 of these individuals are matched with a household member with a missing value as LIPRO household position, if they are of opposite sex and the age difference is less than 15 years. In case of multiple matches, the union with the smallest age difference is chosen. The remaining individuals were changed to *single+* (157) and *other* (147).

B.1.3 Parent-child matching

In cases where both parents do not live in the same household as their child (e.g. foster care) or both IDs are missing, the parental role is assigned to a household member who is at least 14 years older than the child. In case multiple household members fulfil that requirement, the one with an age difference to the child closest to 29 years (mean age of Belgian females at first birth in 2016[1]) is chosen. The parent IDs were adjusted for 63,075 individuals. For a small number of children (1,438), no suitable parent was found in the household and the individuals were instead assigned to randomly chosen households of females in a union of at least 30 years of age.

[1] https://ec.europa.eu/eurostat/documents/2995521/8774296/3-28032018-AP-EN.pdf/fdf8ebdf-a6a4-4153-9ee9-2f05652d8ee0

B.1.4 Sample population

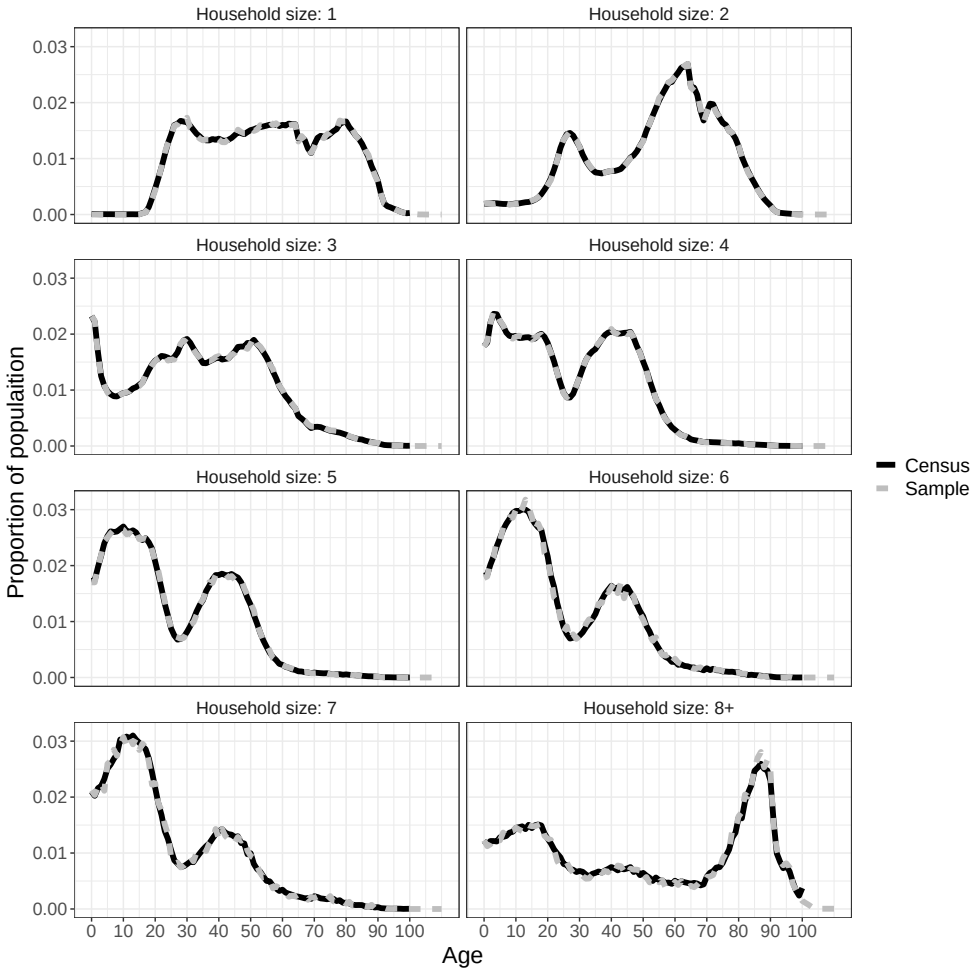

Figure B.1. Age distribution by household size in Belgian census 2011 and sample population.

B.2 Fertility

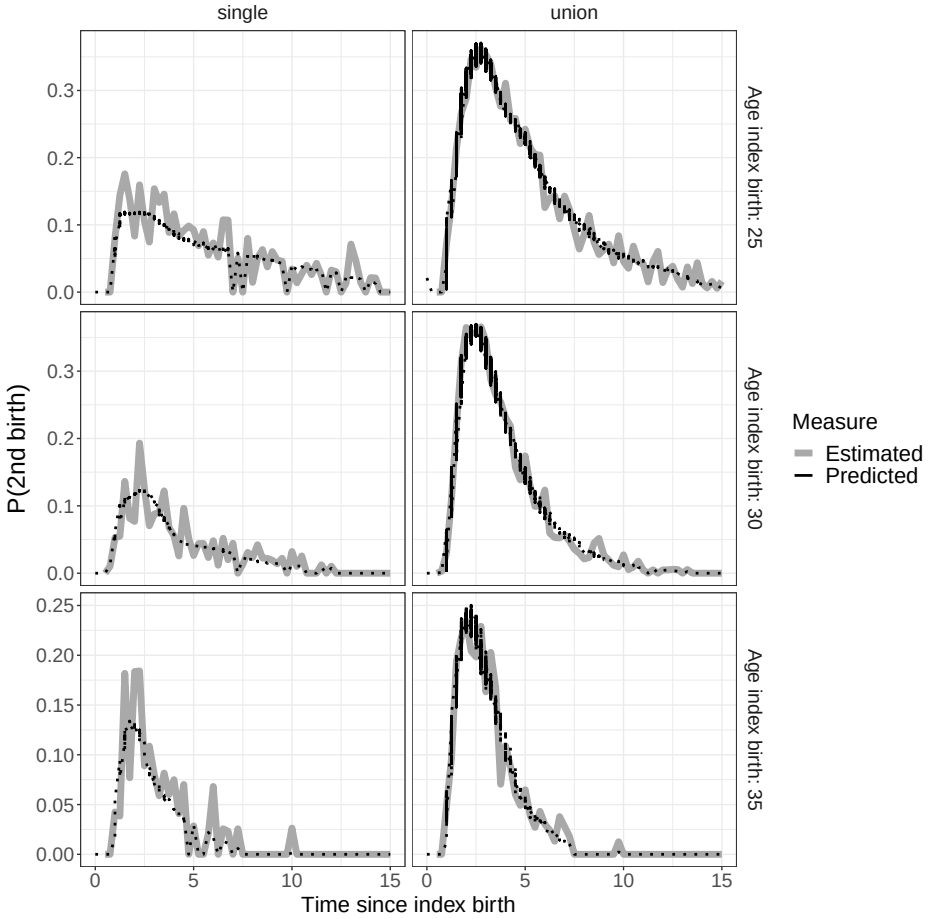

Figure B.2. Predicted and estimated probability of 2nd birth in 2011 based on Belgian census by time since index birth (x-axis), household position (columns) and index age (rows).

B.3 Household transitions

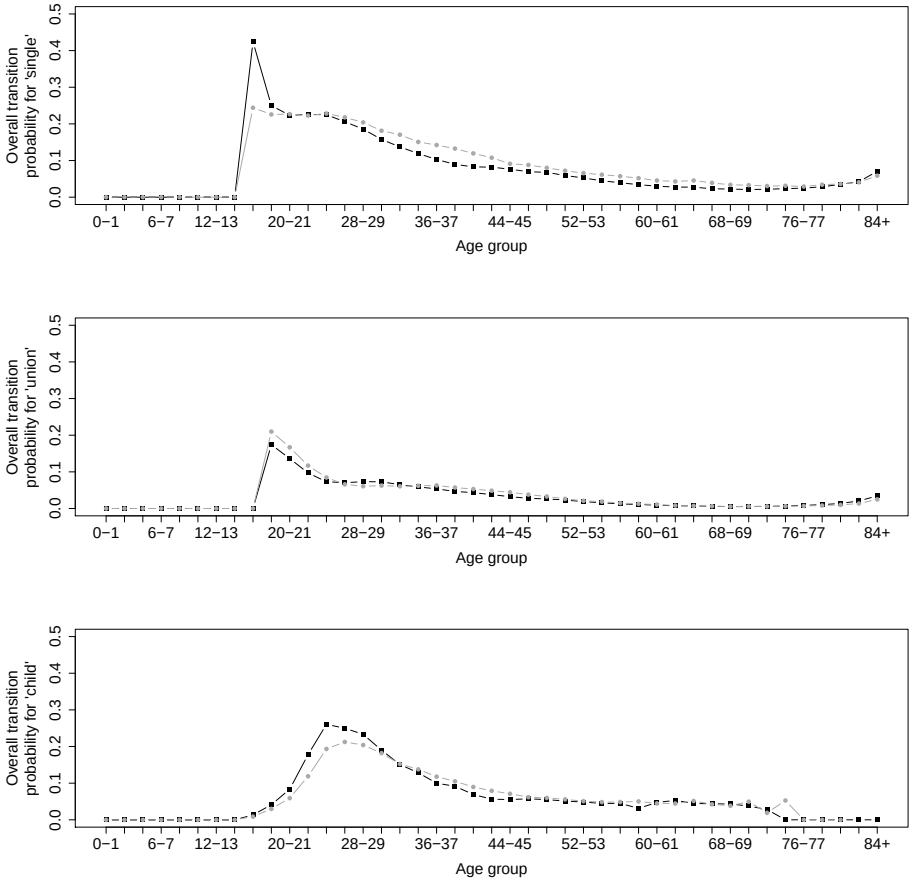

Figure B.3. Overall household transition probabilities by age and sex (male=black, female=grey) for household positions single, union and child in case of no birth event in the same year and parent indicator of 1.

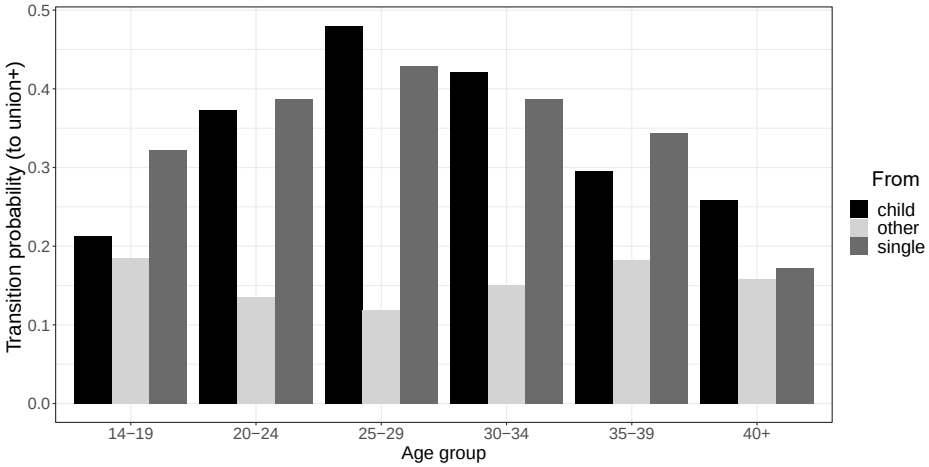

Figure B.4. Age-specific probabilities of household transition from child, other or single to union+ for females with a birth event in the same year.

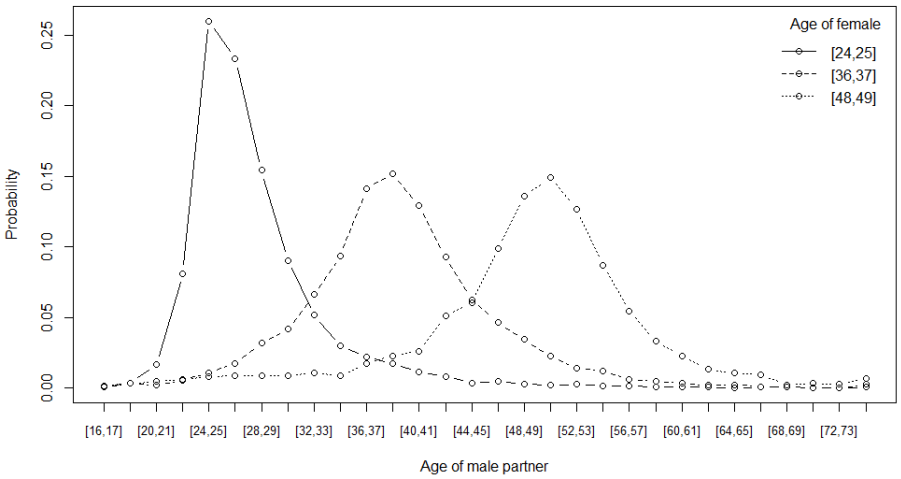

Figure B.5. Age-specific probability of male entering union conditional on the age of female partner.

Appendix C

Population age and household structures shape transmission dynamics of emerging infectious diseases: a longitudinal microsimulation approach - Appendix

C.1 Demographic microsimulation

A detailed description of the demographic data and model implementation as well as source code are available from the GitHub repository:
`https://github.com/signemoegelmose/demographic_microsimulation_EXTERNAL`

We do not have permission to share the demographic input data from the Belgian population registers. New input files thus need to be created to run the code.

C.2 Social contact matrix

The social contact matrix applied in the disease transmission model is visualised in Figure C.1, where the colour indicates the average daily number of contacts between age groups.

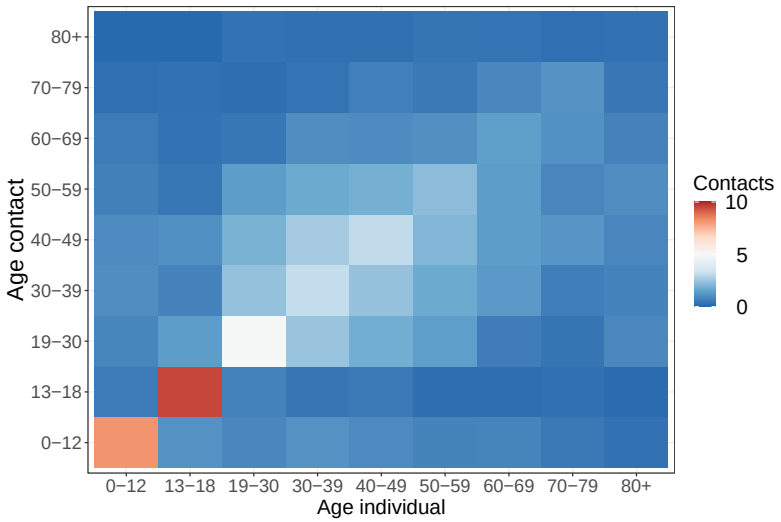

Figure C.1. Age-specific social contacts in Belgium excluding contacts with household members and excluding supplementary professional contacts [167, 176].

C.3 Household network density

The distribution of the household network densities (i.e. the number of links in a household relative to the number of possible links) by household size and type are shown in Figure C.2, where the blue line indicates the overall mean, while the red and green dashed lines indicate the mean density for households with and without at least one child younger than 13 years, respectively.

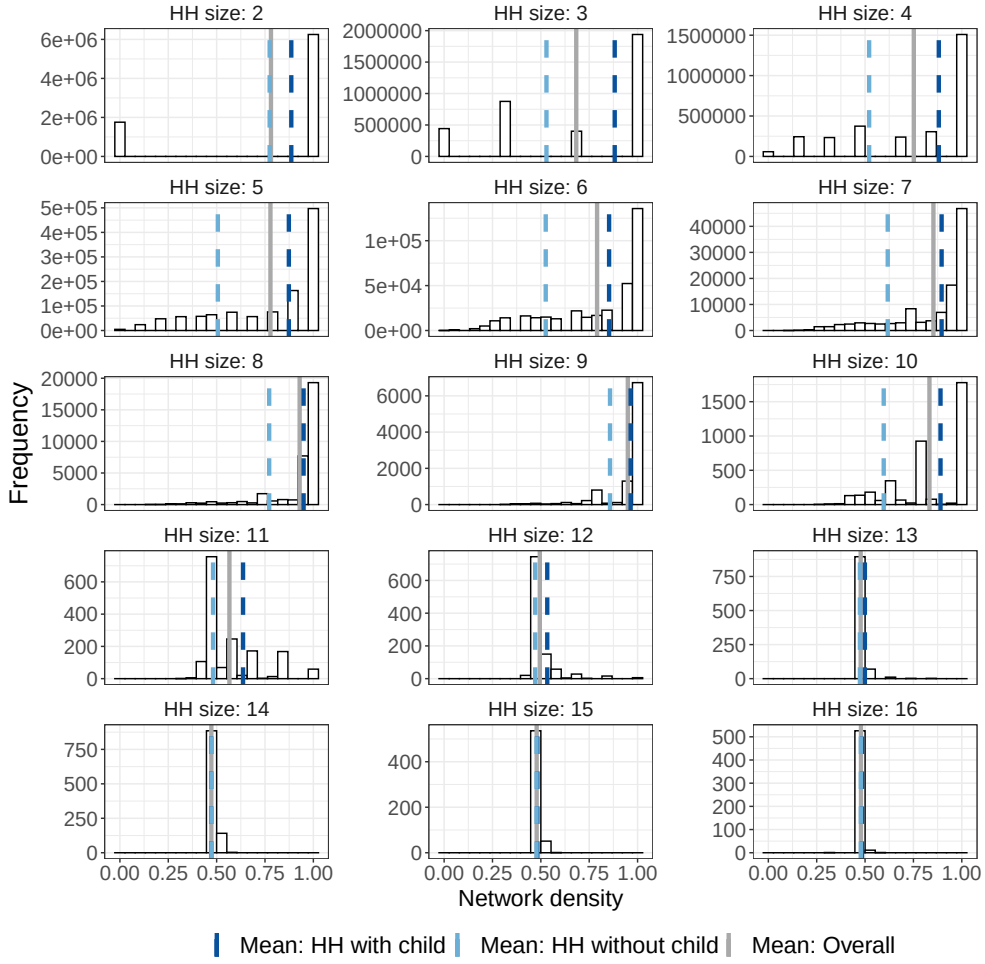

Figure C.2. Histograms of household network densities by household size (HH: household, child: age < 13).

C.4 Household size by age groups

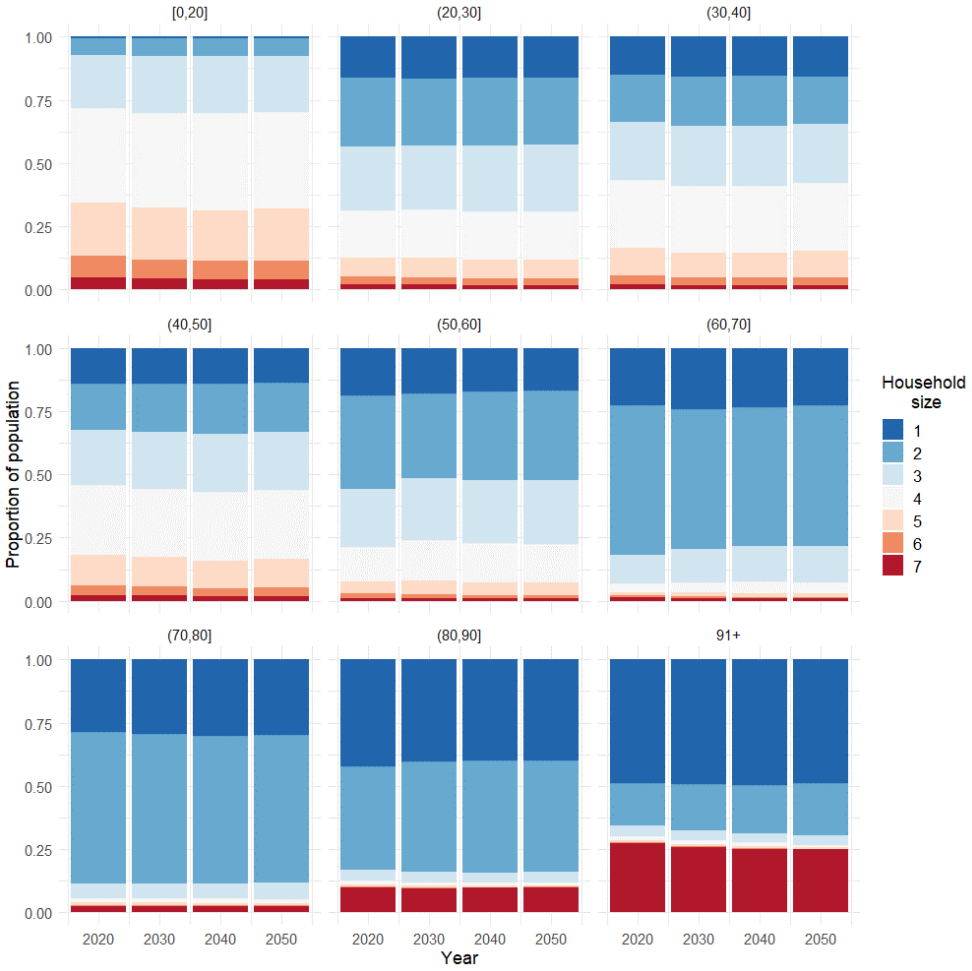

Figure C.3. Household size distribution by age group and year (7=7+).

C.5 Household position by age groups

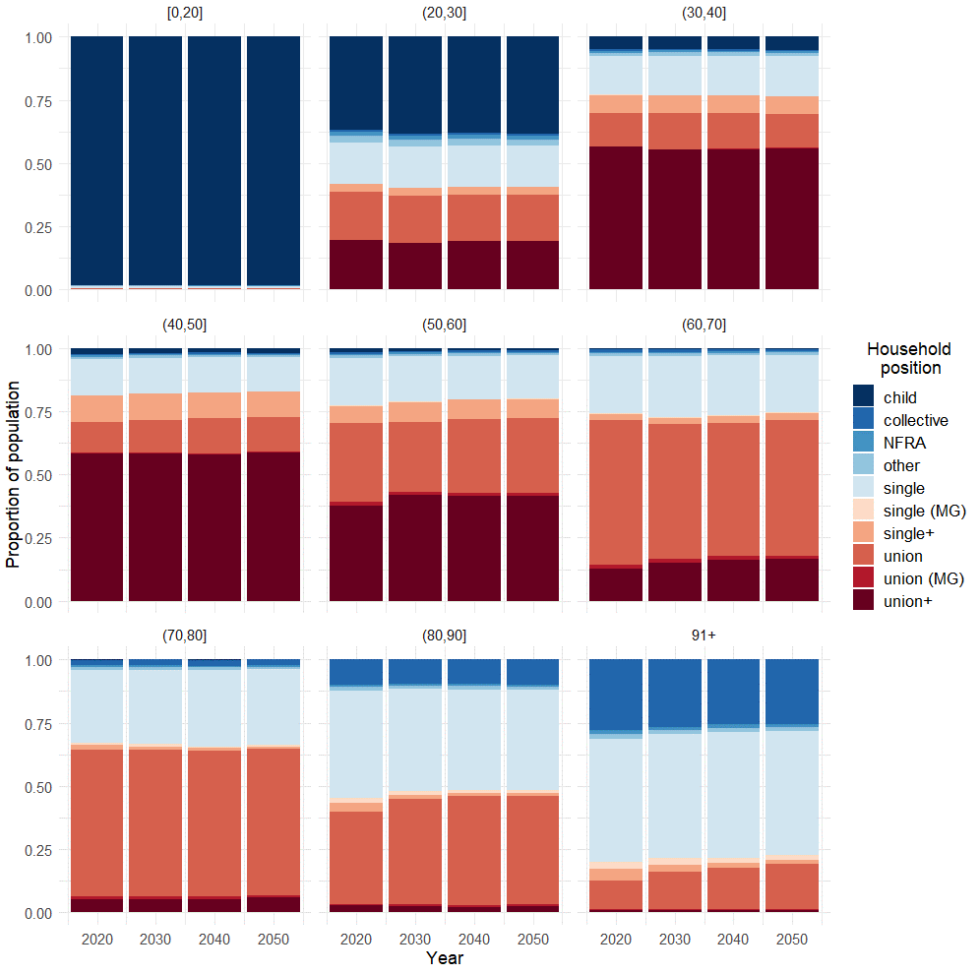

Figure C.4. Household position distribution by age group and year. NFRA: non-family related adult, MG: oldest generation in multi-generational household, single+: single parent, union+: union with child(ren) in household.

C.6 Total fertility rate

Figure C.5 displays the total fertility rate observed in Belgium from 1960 to 2020 (grey) and the total fertility rate in the microsimulation (black) from 2011 to 2050.

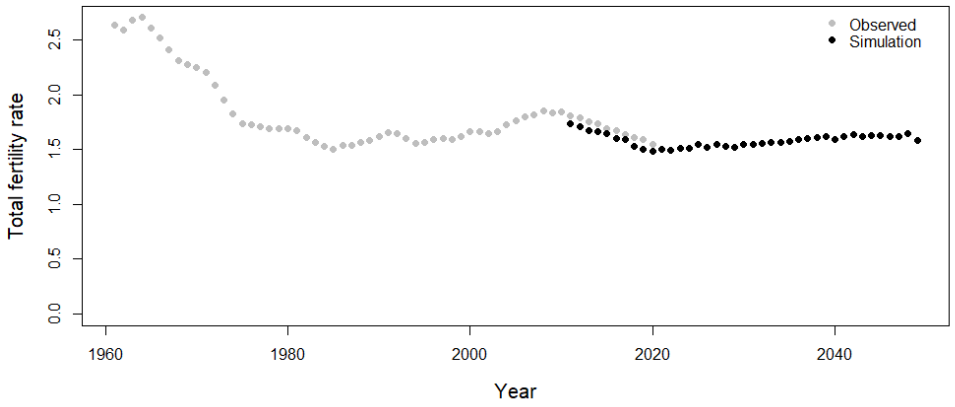

Figure C.5. Total fertility rate. Observed (Statistics Belgium): 1960-2020. Simulation: 2011-2050.

C.7 Incidence by transmission parameters and scenarios

Figure C.6 contains the overall attack rate (i.e. the proportion infected of the total population) by transmission probabilities (panels and colours), year (shape) and scenario (x-axis).

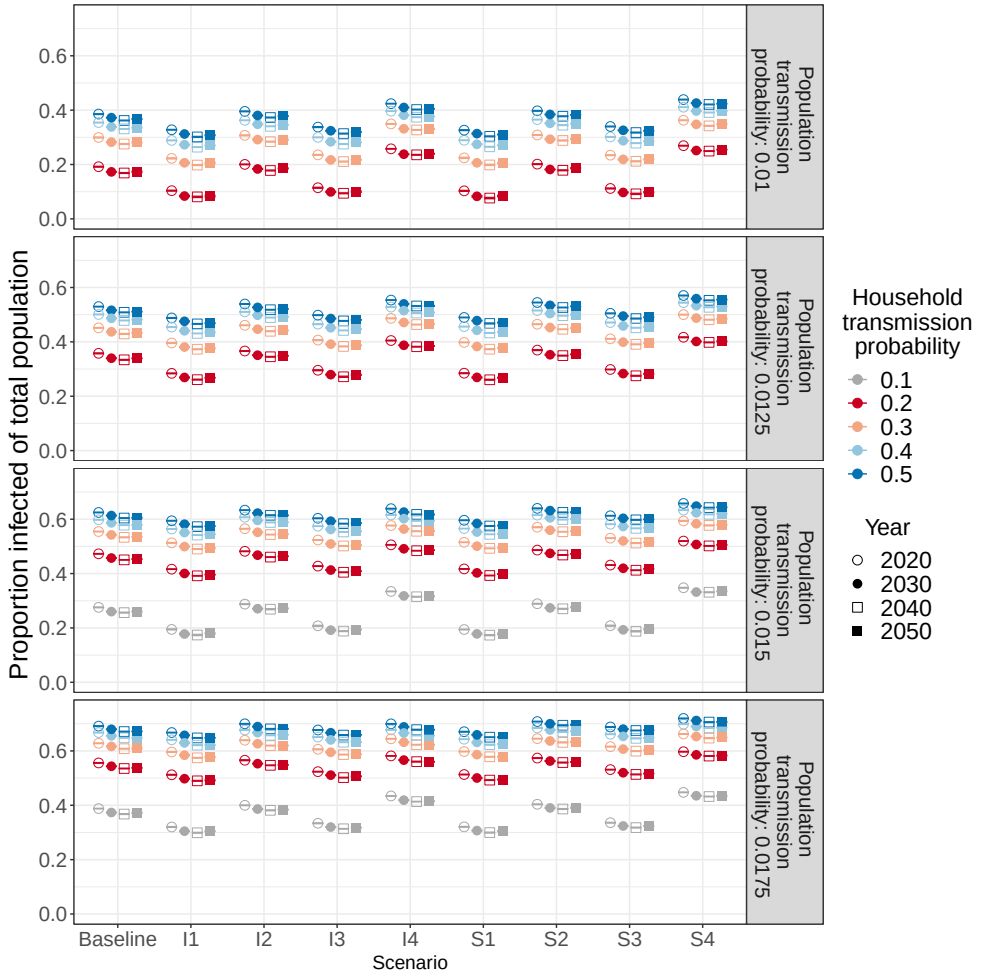

Figure C.6. Mean attack rate in total population with 95% confidence interval for varying transmission parameters (β_h, β_p), susceptibility and infectiousness scenarios and simulation year.

C.8 Incidence by household size and type

Figure C.7 and Figure C.8 contain the average proportion of households with at least one infected household member and the average proportion of infected household members in households with at least one infected household member, respectively, in the baseline scenario and by household size (x-axis), type (shape), year (colour) and transmission probability (panels).

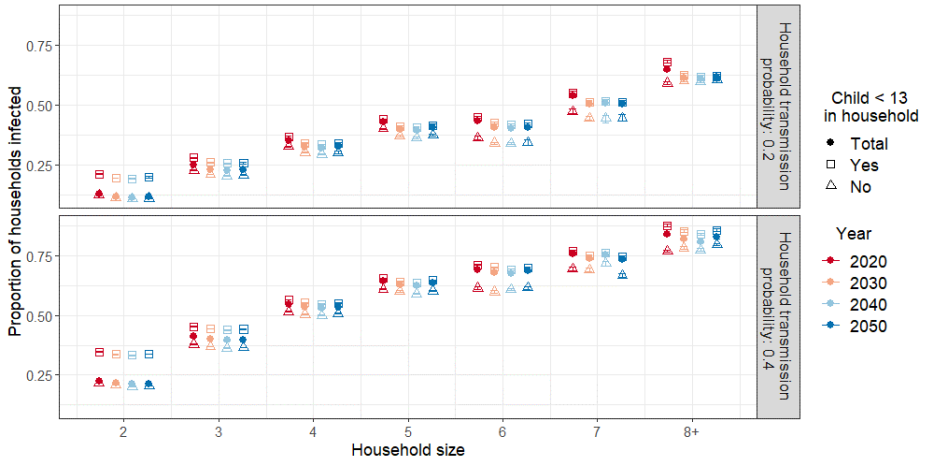

Figure C.7. Mean proportion of households with minimum one infected household member by household size and type in baseline scenario. Population transmission probability: 0.01.

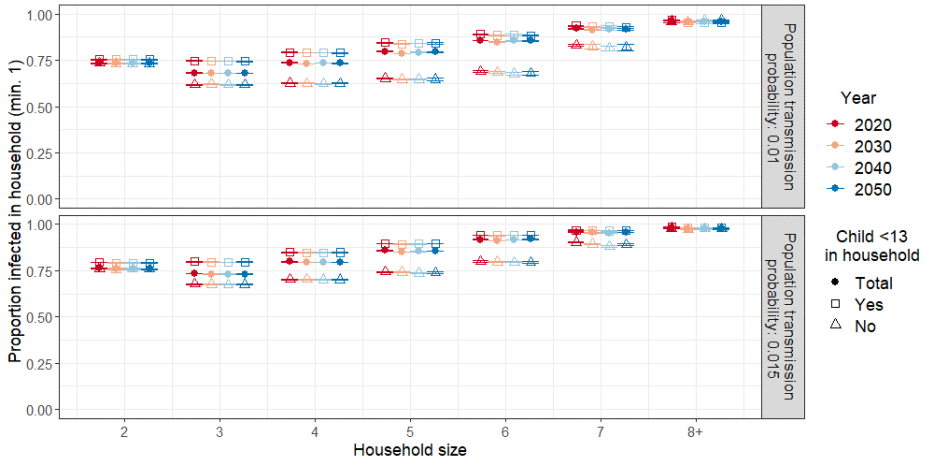

Figure C.8. Mean proportion of household members getting infected by household size and type in baseline scenario. Estimate based on households with minimum one infected individual. Household transmission probability: 0.2.

C.9 Age-specific attack rate by scenario and over time

Figure C.9 and Figure C.10 contain the age-specific attack rates by scenario (colour), year (panel) and transmission probability.

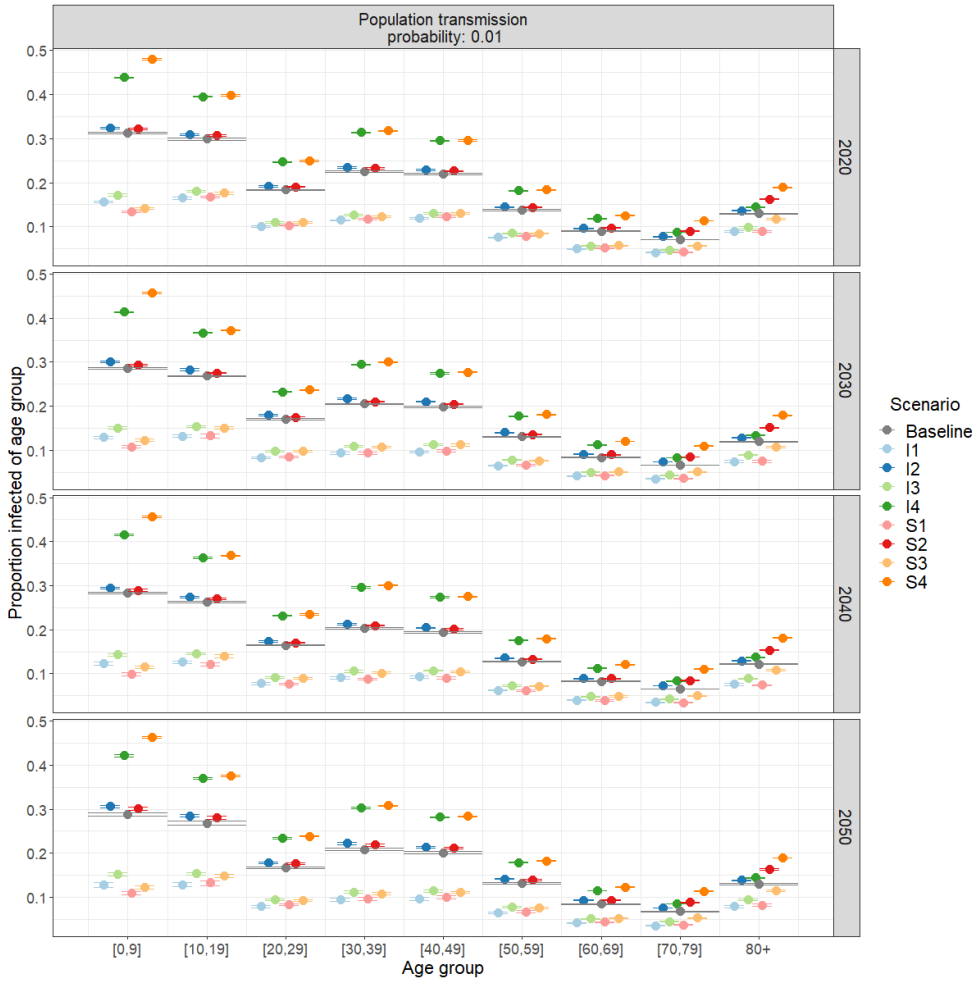

Figure C.9. Mean and 95% bootstrap confidence interval for the mean age-specific attack rate in baseline scenario (grey) and each susceptibility/infectiousness scenario across simulation years. Household transmission probability: 0.2, population transmission probability: 0.01

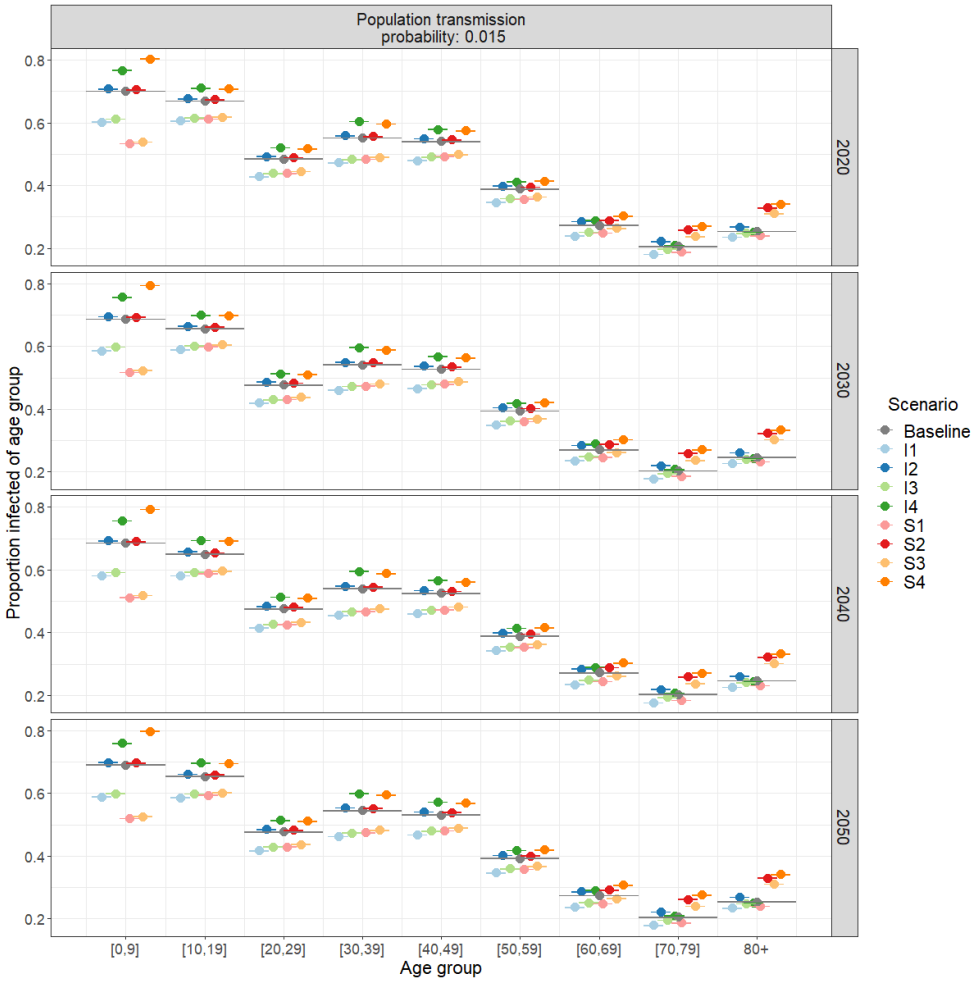

Figure C.10. Mean and 95% bootstrap confidence interval for the mean age-specific attack rate in baseline scenario (grey) and each susceptibility/infectiousness scenario across simulation years. Household transmission probability: 0.2, population transmission probability: 0.015.

Appendix D

Exploring the impact of population ageing on the spread of emerging respiratory infections and the associated burden of mortality - Appendix

D.1 Demographic microsimulation

A detailed description of the demographic data and model implementation as well as source code are available from the GitHub repository:
`https://github.com/signemoegelmose/demographic_microsimulation_EXTERNAL`

We do not have permission to share the demographic input data from the Belgian population registers. New input files thus need to be created to run the code.

D.2 Trends in fertility and life expectancy (Statbel)

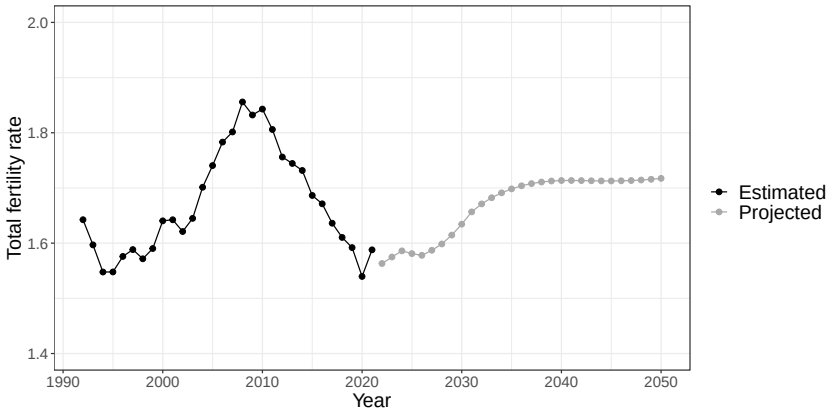

Figure D.1. Estimated and projected total fertility rate for Belgium, 1992-2050. Source: Statbel.

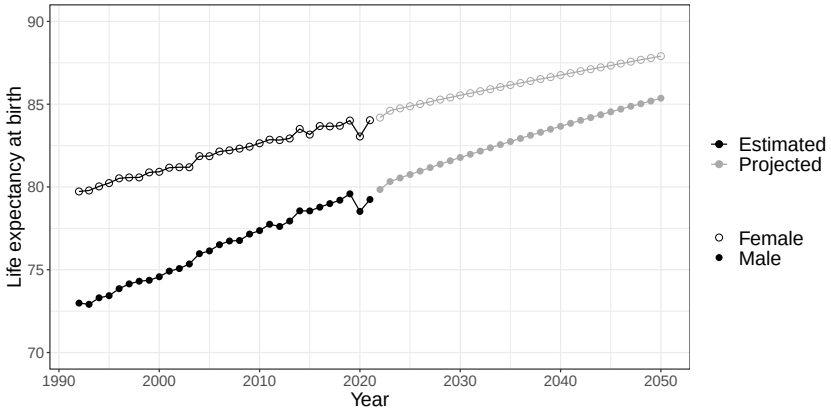

Figure D.2. Estimated and projected life expectancy at birth for Belgium, 1992-2050. Source: Statbel.

D.3 Social contact matrix

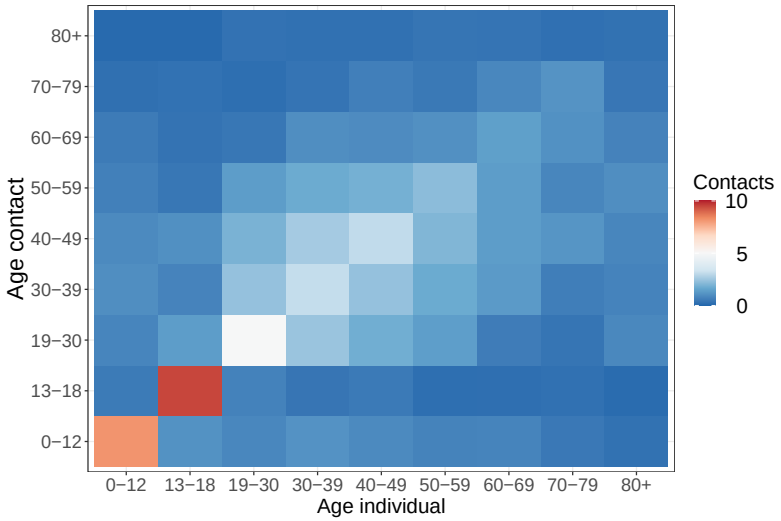

Figure D.3. Age-specific social contacts in Belgium excluding contacts with household members and excluding supplementary professional contacts [167, 176].

D.4 Household network density

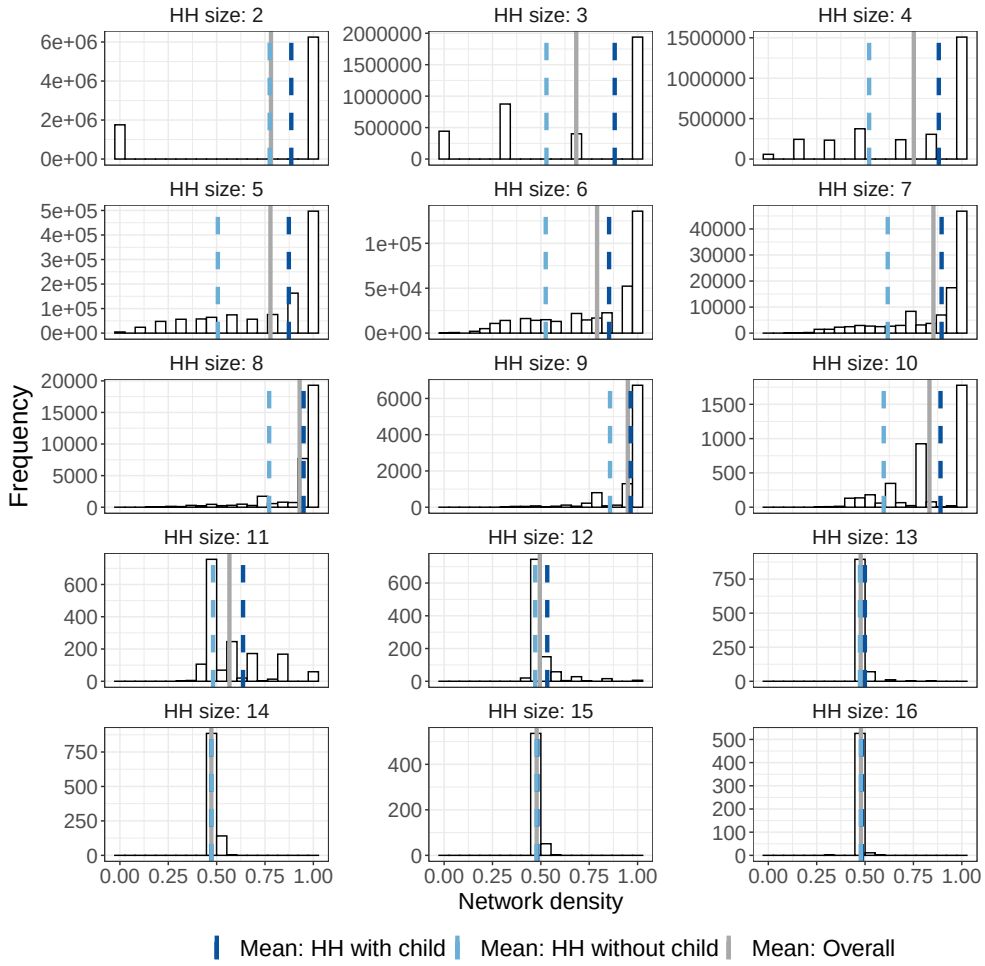

Figure D.4. Histograms of household network densities by household size (HH: household, child: age < 13).

D.5 Threshold parameter R_*

We compute the threshold parameter R_* (group-to-group reproduction number) based on Ball et al. [175]. The basic reproduction number R_0 is not used because it requires large group sizes, which is not the case for the households in our two-level mixing model. The computation of R_* is based on equation (3.31) in Ball et al. [175]:

$$R_* = \lambda_G E[T_I] \mu_h^{-1} \sum_{n=1}^{\infty} (1 + \mu_{n-1,1})\, n h_n \tag{D.1}$$

$$R_* = \mu R_G \tag{D.2}$$

where n corresponds to household size, h_n is the proportion of households of size n and μ_h is the mean household size. We compute the average final size in households of size n, $(1 + \mu_{n-1,1})$, by starting with one randomly chosen infected individual in each household, which then can pass on the infection to household members, which also can transmit the infection within the household. Meanwhile transmission in the general population is disregarded. Finally, the average final size by household size is calculated and used to compute $\mu = \mu_h^{-1} \sum_{n=1}^{\infty} (1 + \mu_{n-1,1})\, n h_n$ (i.e. the average number of household infections). The basic reproduction number in the general population when disregarding household transmission, $R_G = \lambda_G E[T_I]$, is computed by initially infecting one randomly chosen individual in the population. The individual can transmit the infection to others in the population, but the newly infected individuals cannot pass on the infection. The average number of secondary cases is then calculated.

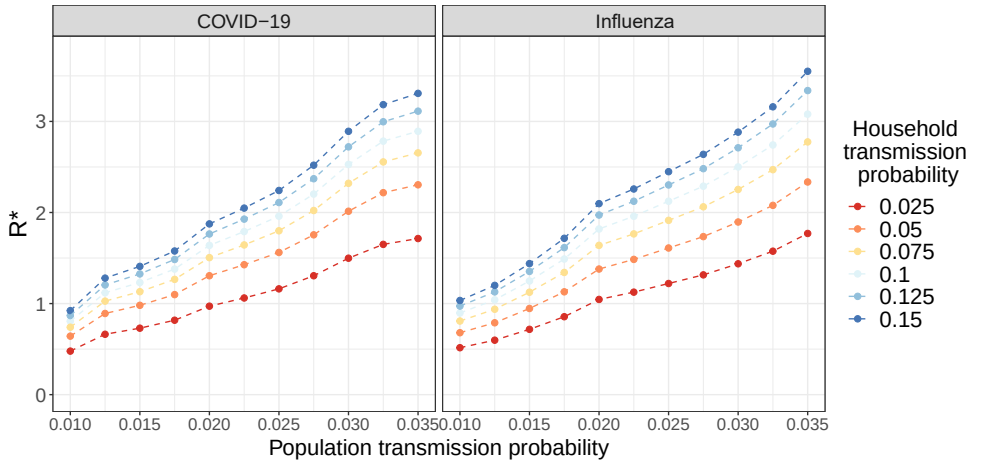

Figure D.5. R_* for medium demographic scenario in 2020.

D.6 Transmission parameters: β_h and β_p

We select the transmission parameters, β_h and β_p, based on the R_* in Figure D.5 and the household secondary attack rates in Figure D.6. In the model for COVID-19, we assume an R_* of approximately 3 to reflect SARS-CoV-2 transmission in Belgium prior to the implementation of mitigation measures [32, 218, 219]. In the ILI model, we assume an R_* of approximately 1.5, which resembles the basic reproduction number estimated for the 2009 influenza A/H1N1 pandemic [213–215]. We choose the parameters $\beta_{h,s} = 0.125$ and $\beta_{p,s} = 0.0325$ for COVID-19 and $\beta_h = 0.075$ and $\beta_p = 0.02$ for ILI. With these parameters we obtain household secondary attack rates of approximately 0.34 and 0.19 for COVID-19 and ILI. This reflects estimated household secondary attack rates from several studies [261–263], however, these are associated with substantial variability.

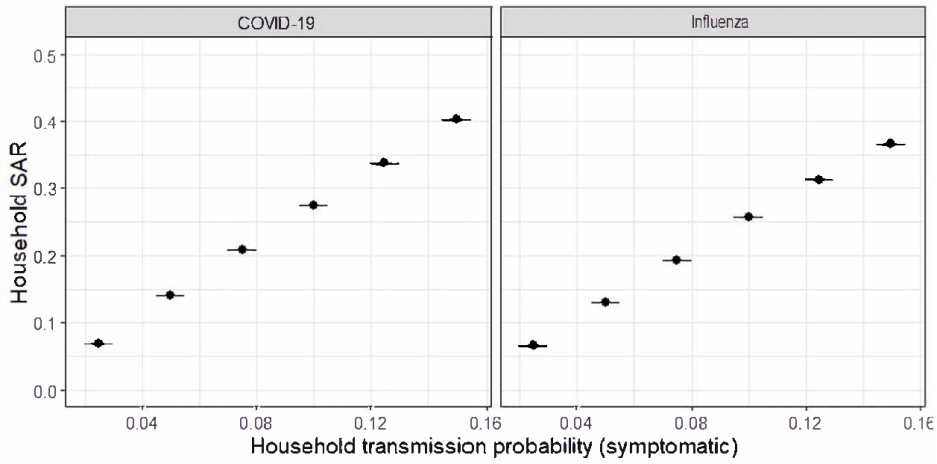

Figure D.6. Household secondary attack rate (household transmission only) for medium demographic scenario in 2020.

D.7 Transmission parameters: COVID-19

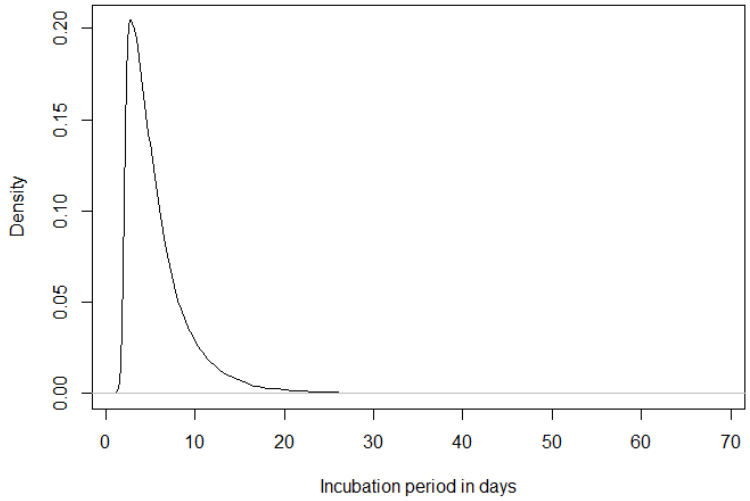

Figure D.7. Incubation period from He et al. [220]

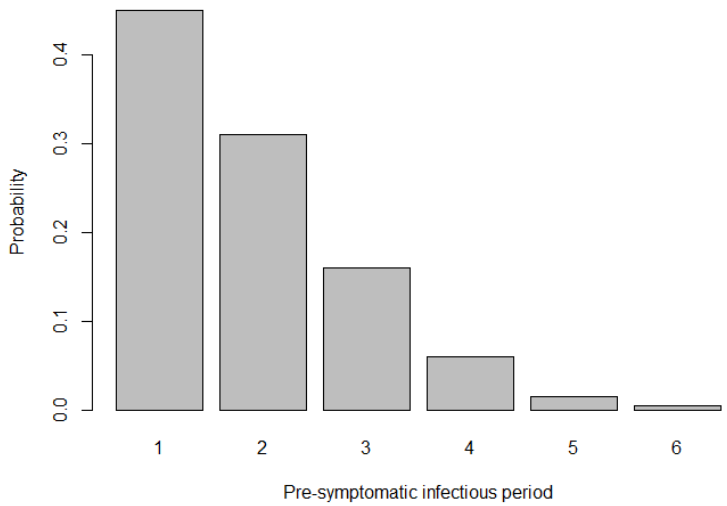

Figure D.8. Distribution for pre-symptomatic period in days [32].

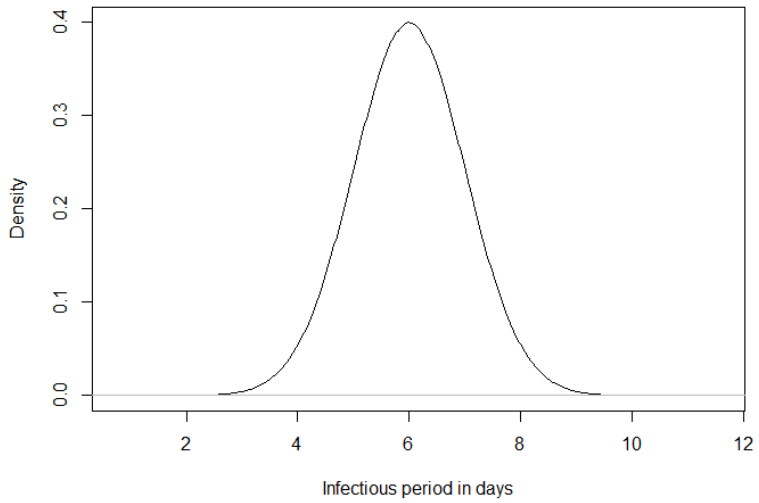

Figure D.9. Distribution for infectious period [32].

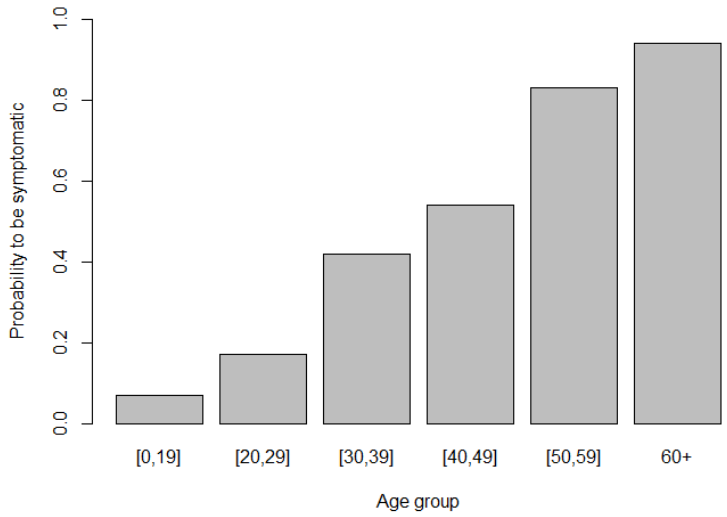

Figure D.10. Probability to be symptomatic by age group (Willem at al., 2020) [32]

D.8 Disease-related mortality

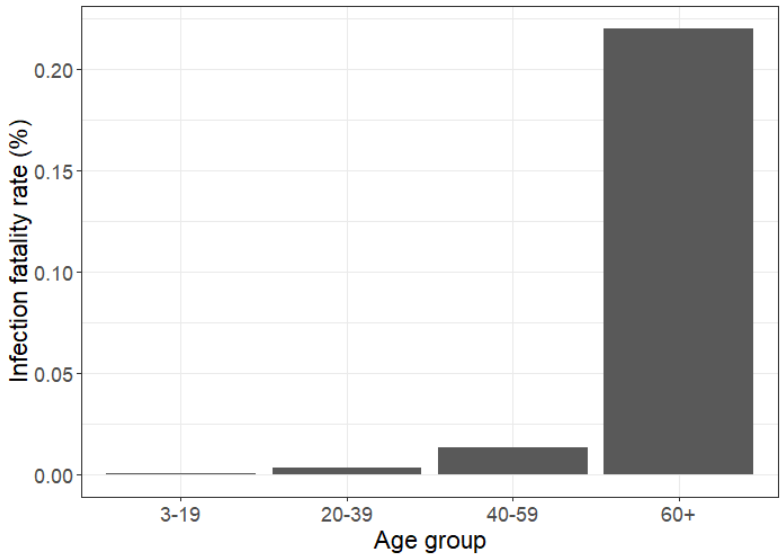

Figure D.11. Infection fatality rates for 2009 (H1N1) pandemic influenza in Hong Kong by age group [222].

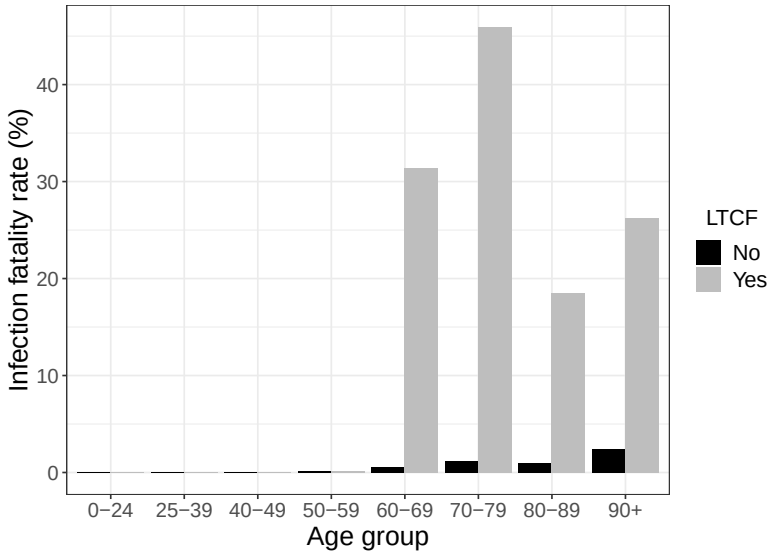

Figure D.12. Infection fatality rates for COVID-19 in Belgium by age group and household type (LTCF or non-LTCF) [77].

D.9 Estimation of QALY losses

We estimate the QALYs lost due to premature death at age x, $1 \leq x < \omega$, as follows [223]:

$$dQALY(x) = \frac{\sum_{u=x}^{\omega} L_s(u) \cdot Q(u) \cdot qCM \cdot (1+r)^{-(u-x)}}{l_s(x)}, \tag{D.3}$$

where

$$l_s(x) = 100,000 \cdot \prod_{a=1}^{x} e^{-d(a) \cdot SMR} \tag{D.4}$$

with $L_s(x)$ being the average of $l_s(x)$ and $l_s(x+1)$ and $Q(x)$ denotes the population average quality of life tariff at age x for Belgium in 2018 [264]. The parameter qCM adjust the quality of life for the impact of pre-existing comorbidity, while the impact of comorbidity on the risk of dying is summarised in the standardised mortality ratio (SMR). Finally, the instantaneous death rate is denoted $d(x)$ and r is the discount rate applied to incorporate the assumption that current health benefits are valued higher than future health benefits. As in Briggs et al. [223], we assume that $SMR = 1.5$ and $qCM = 90\%$.

D.10 Population structures

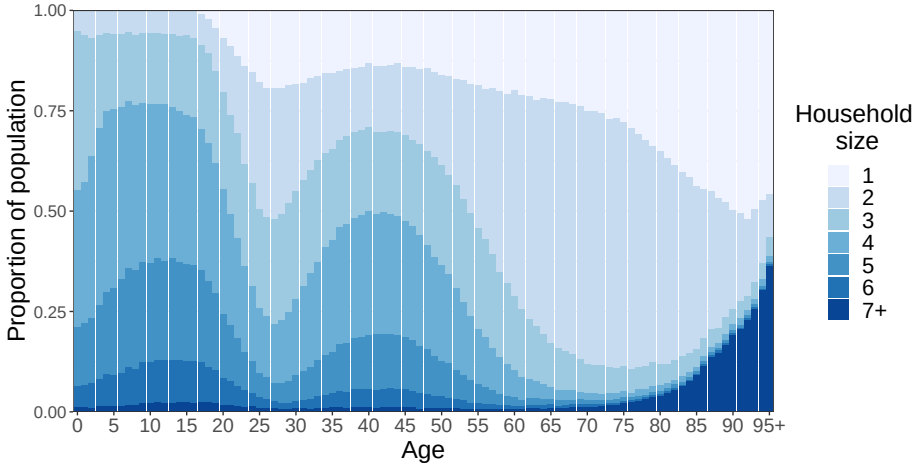

Figure D.13. Household size distribution by age group of simulated population in 2020. Medium scenario.

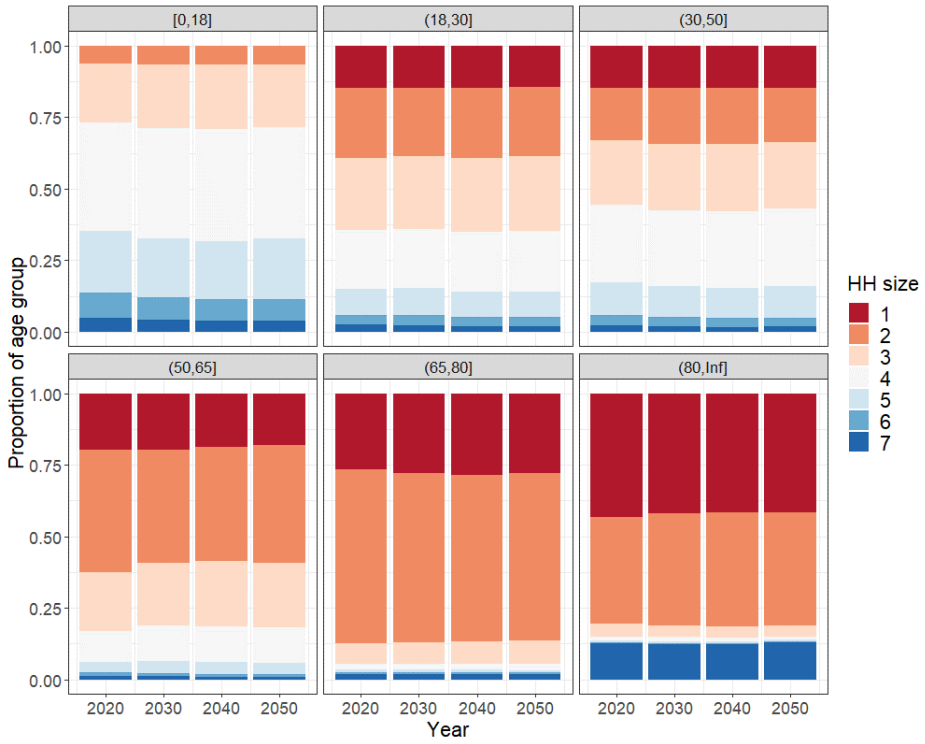

Figure D.14. Household size distribution by age group and year in medium scenario. 7=7+.

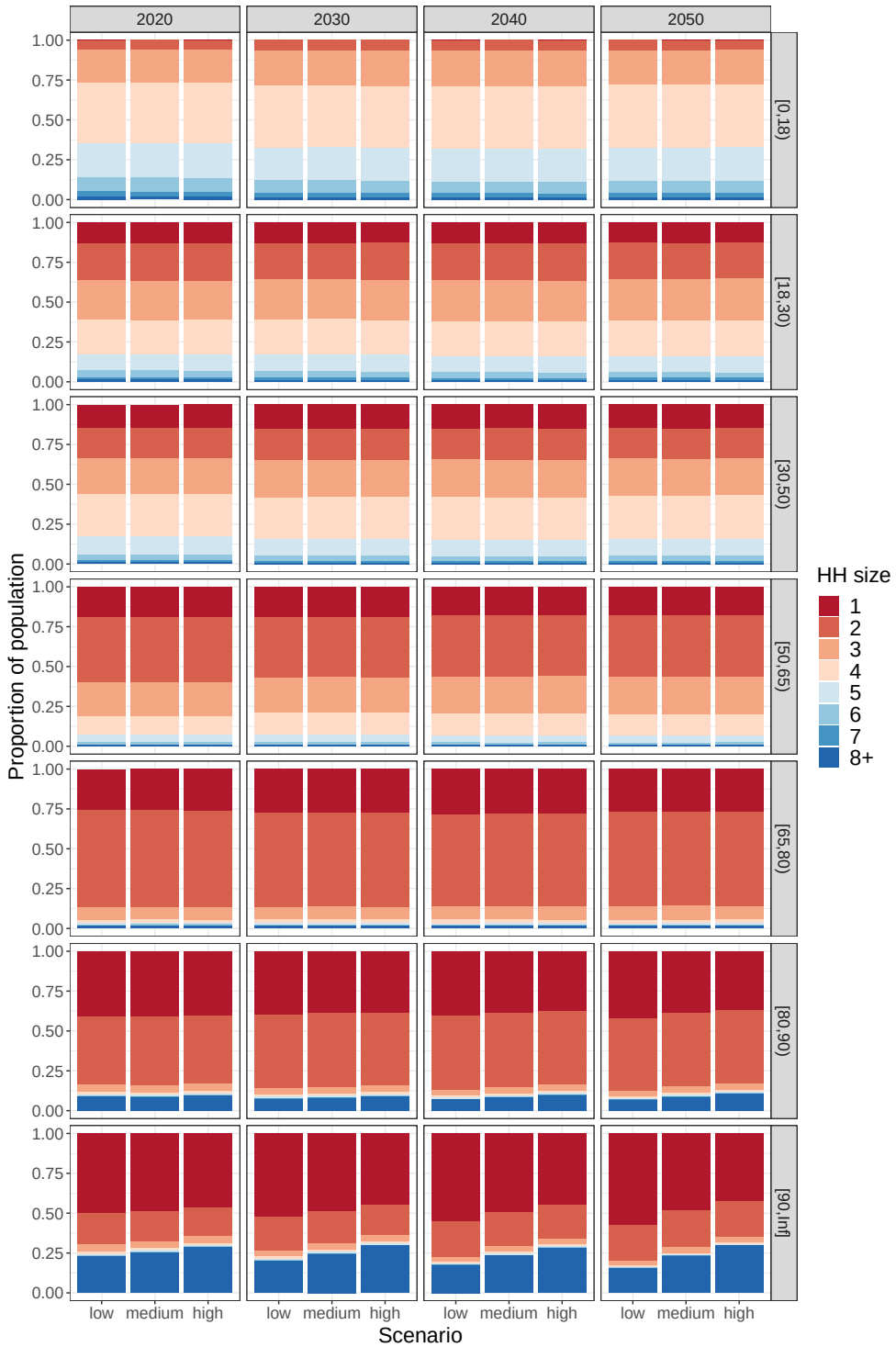

Figure D.15. Household size distribution by age (right y-axis), year and scenario.

D.11 Additional results

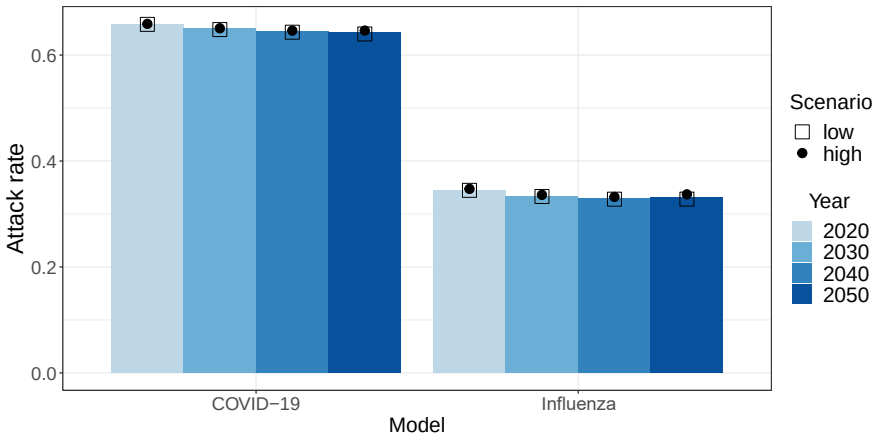

Figure D.16. Mean attack rate in total population by simulation year, model and demographic scenario.

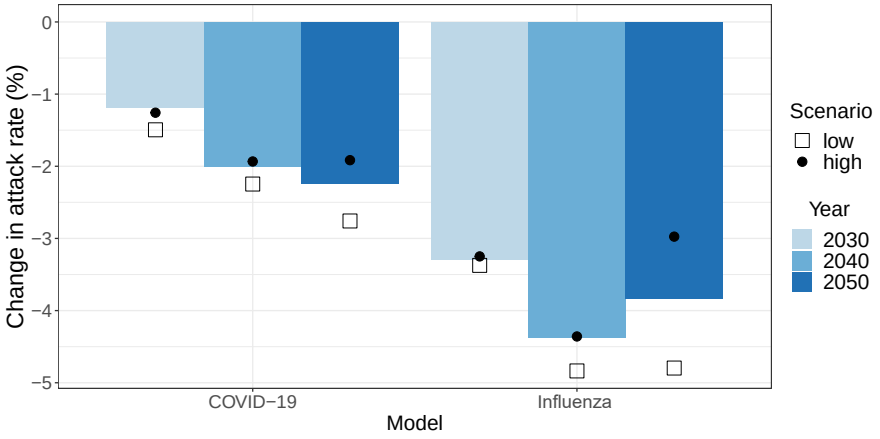

Figure D.17. Change in overall attack rates relative to 2020.

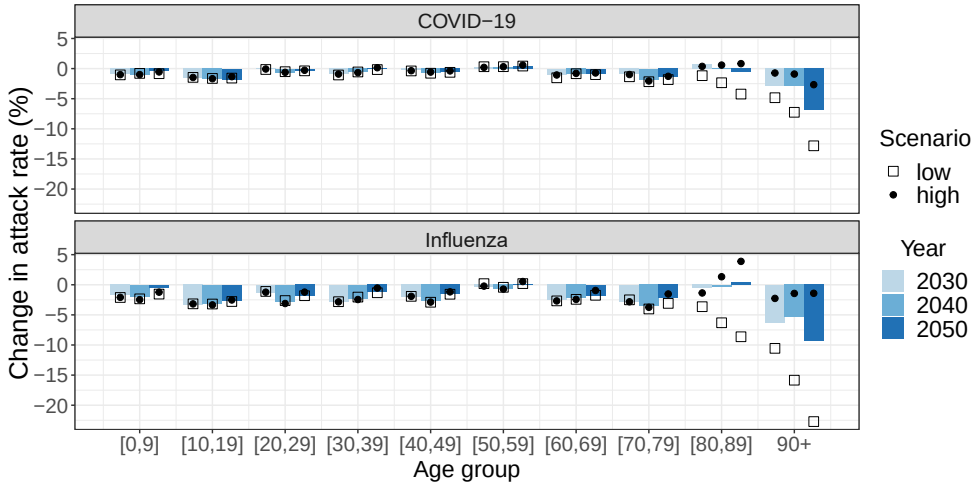

Figure D.18. Change in age-specific attack rates relative to 2020.

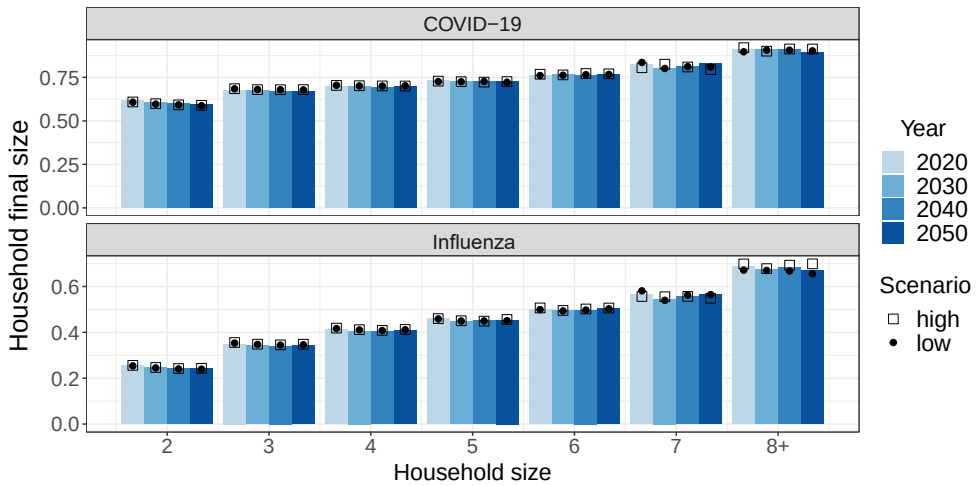

Figure D.19. Mean number of infected household members as proportion of household size by year, model and demographic scenario.

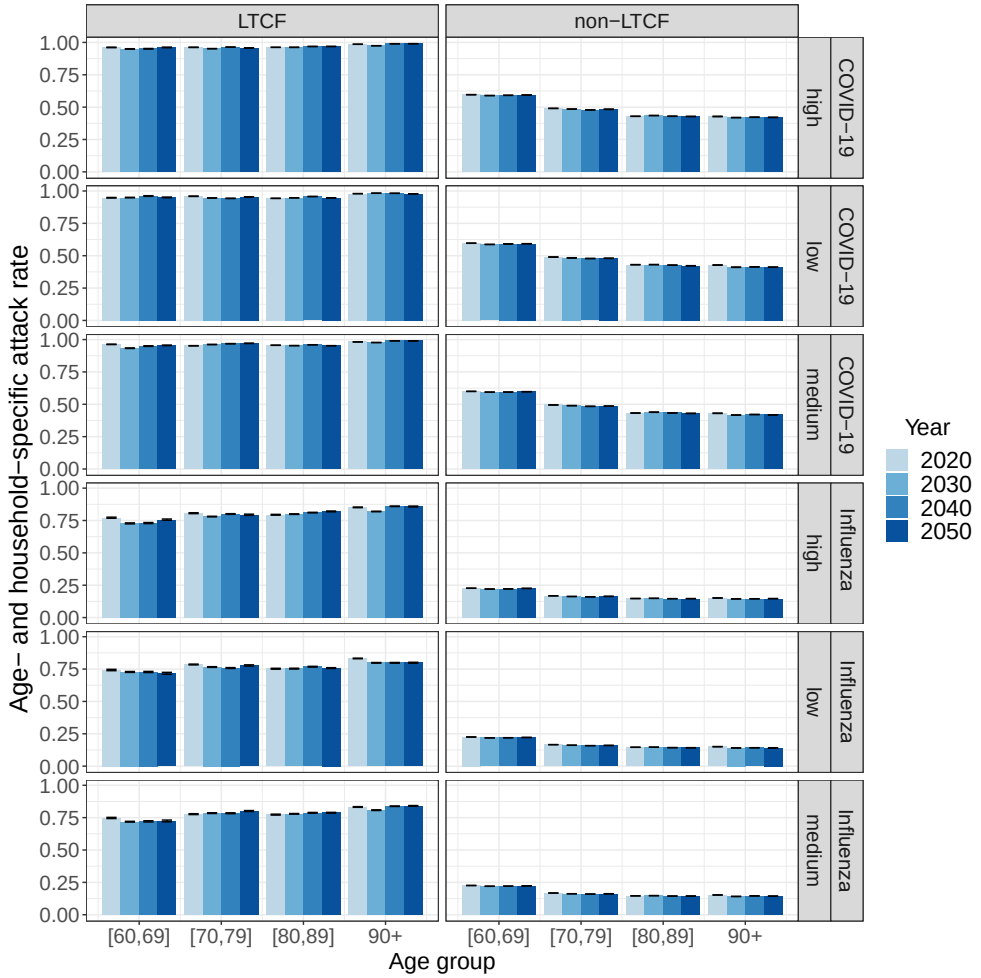

Figure D.20. Age and household-specific attack rate in the elderly population. Columns: Household type. Rows: Model and demographic scenario.

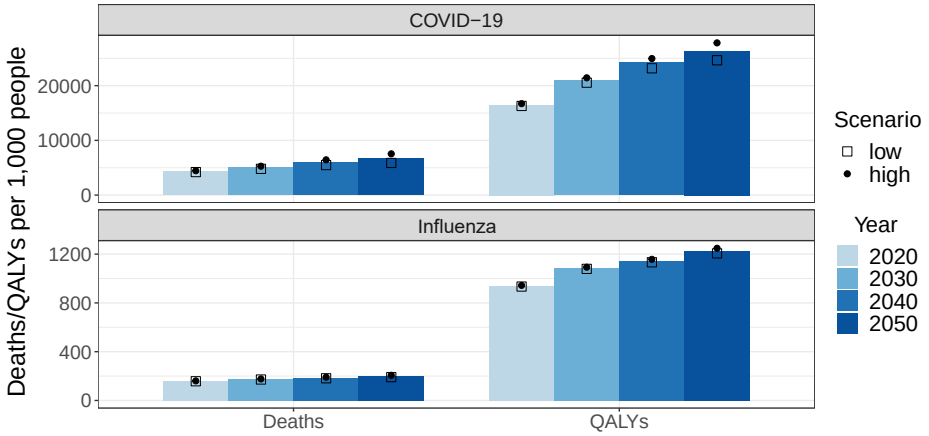

Figure D.21. Disease-related deaths and QALY losses per 1,000 people in the population. Upper panel: COVID-19, lower panel: ILI.

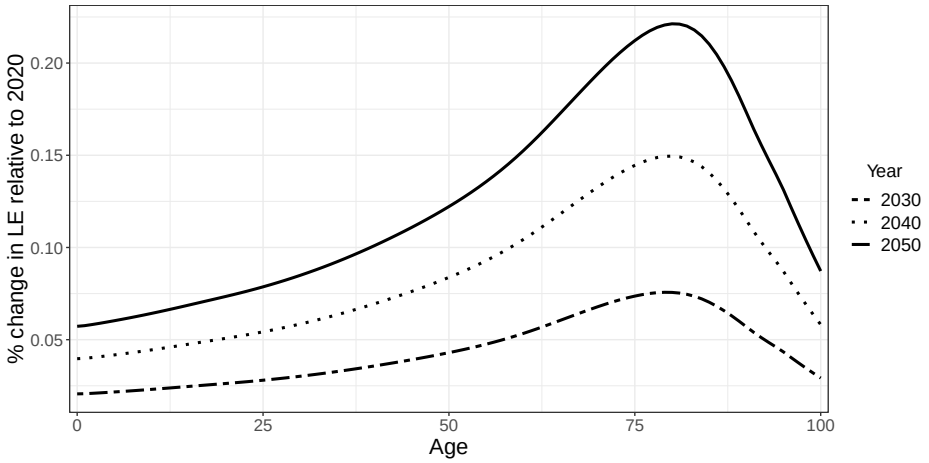

Figure D.22. Relative change in life expectancy across ages compared to 2020 (Statbel).

Appendix E

The impact of demographic change on the epidemiology of varicella and herpes zoster: US population 1960-2020 - Appendix

E.1 IPUMS samples

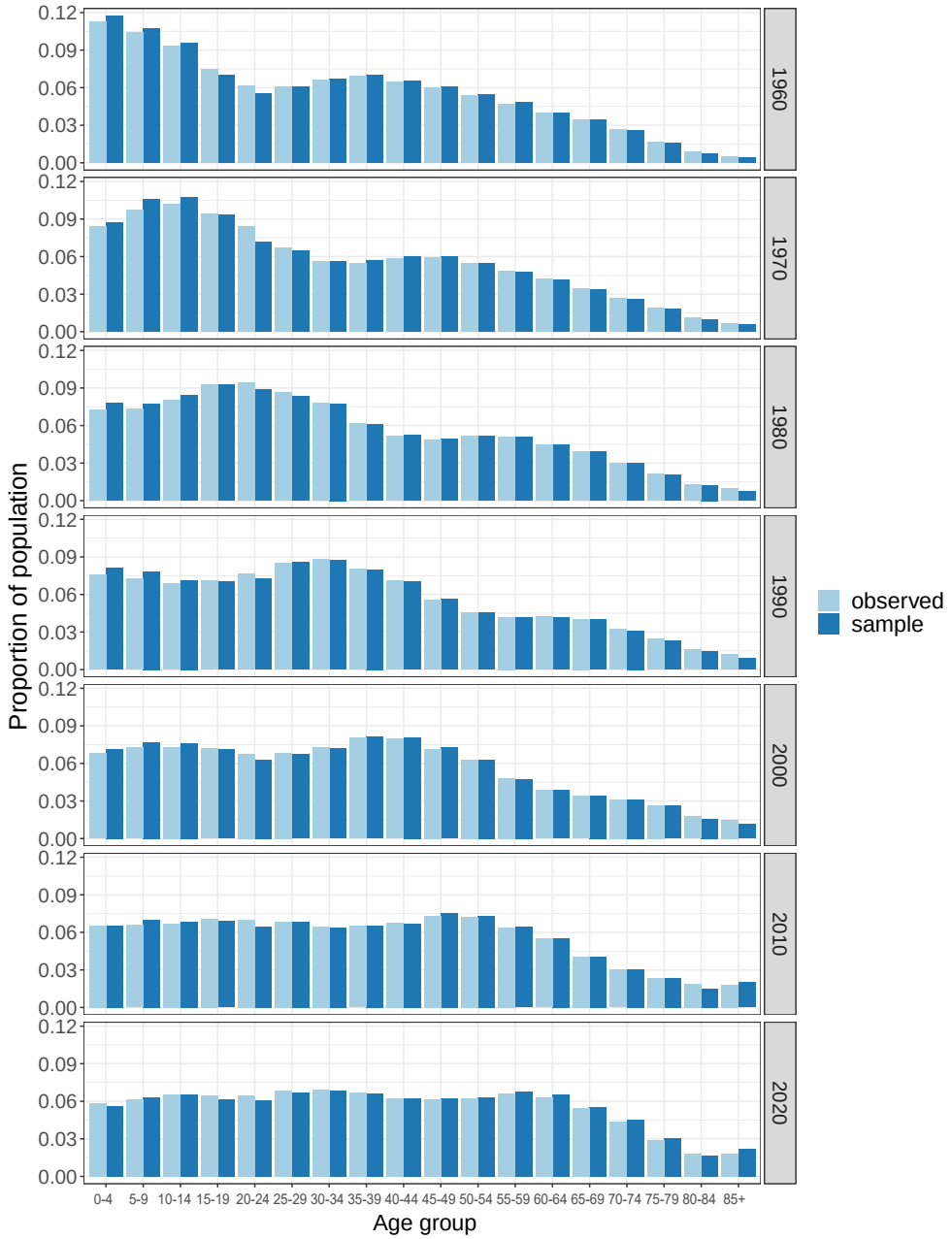

Figure E.1. Age distribution 1960-2020. Sample (1960: IPUMS USA; 1970-2020: IPUMS CPS) vs. observed population (US Census Bureau).

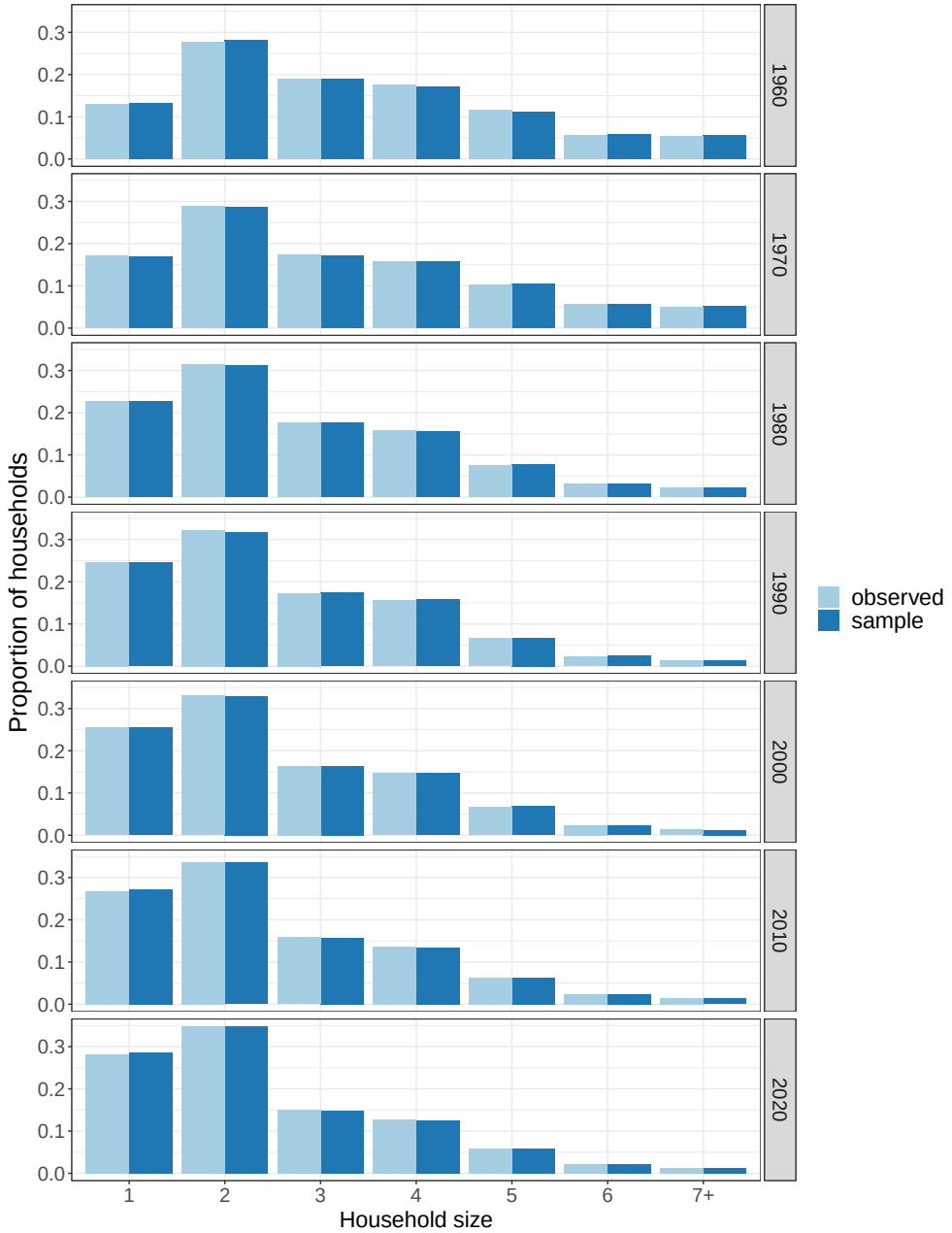

Figure E.2. Household size distribution 1960-2020. Sample (1960: IPUMS USA; 1970-2020: IPUMS CPS) vs. observed population (US Census Bureau).

E.2 Initial and stable age distribution

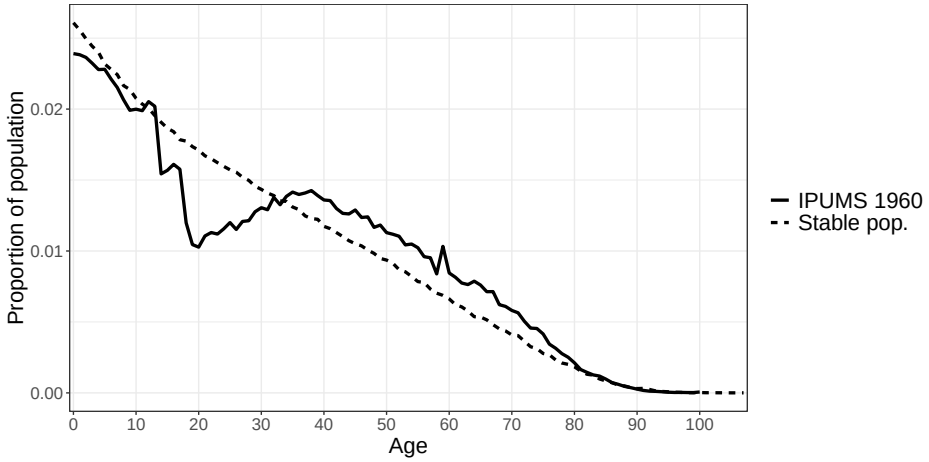

Figure E.3. Age distribution of US population 1960 from IPUMS (solid line) and the corresponding stable population (dashed line) assuming vital rates of 1960 and no migration.

E.3 Age-specific fertility

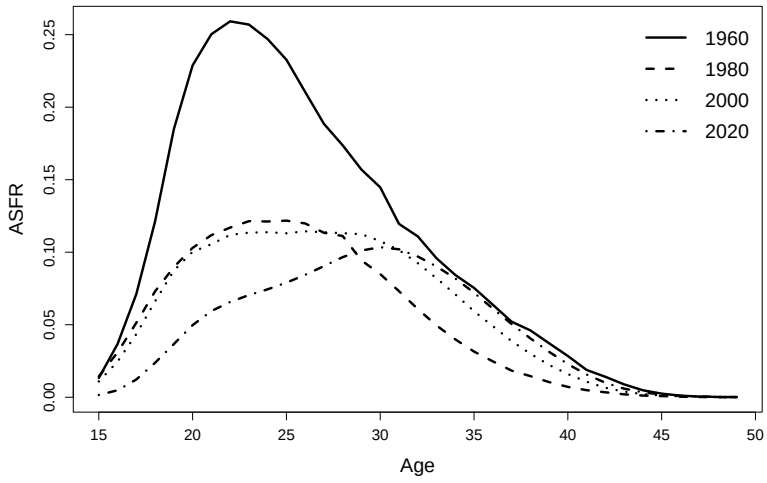

Figure E.4. Age-specific fertility rate estimated for the US in 1960 (solid line), 1980 (dashed line), 2000 (dotted line) and 2020 (dashed-dotted line). Source: UN Statistics Division.

E.4 Age-specific household size distribution

Figure E.5-E.8 show the household size distributions disaggrgated by age group in the simulated population (dark bars) and population samples from IPUMS (light bars). It should be noted that the marginal distributions (i.e. age and household size distribution) in the samples from IPUMS also deviate to some degree from the marginal distributions provided by the US Census Bureau as shown in Figure E.1-E.2.

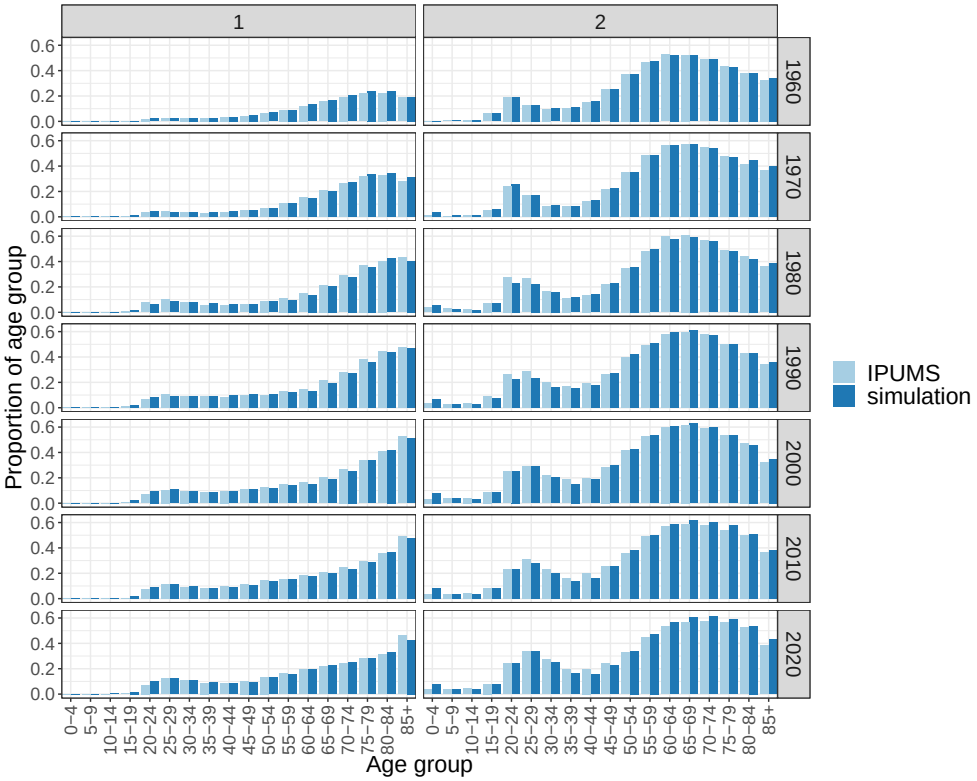

Figure E.5. Proportion of age group (x-axis) living in household size 1 and 2 (columns) by year (row) based on IPUMS (light bars) and simulation (dark bars).

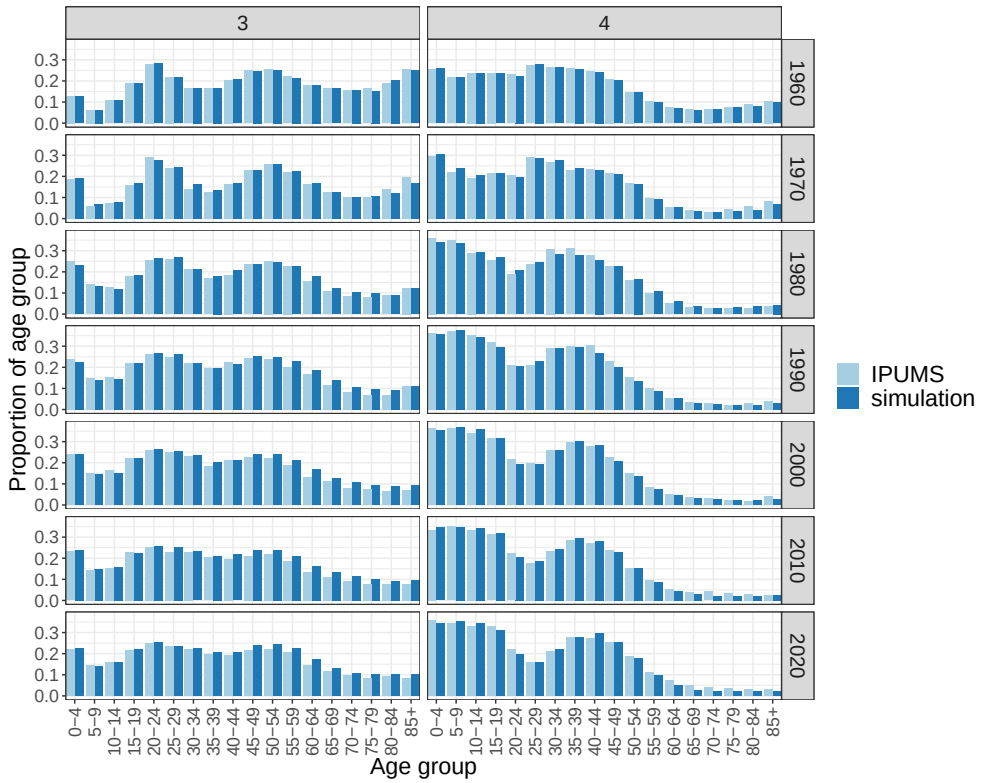

Figure E.6. Proportion of age group (x-axis) living in household size 3 and 4 (columns) by year (row) based on IPUMS (light bars) and simulation (dark bars).

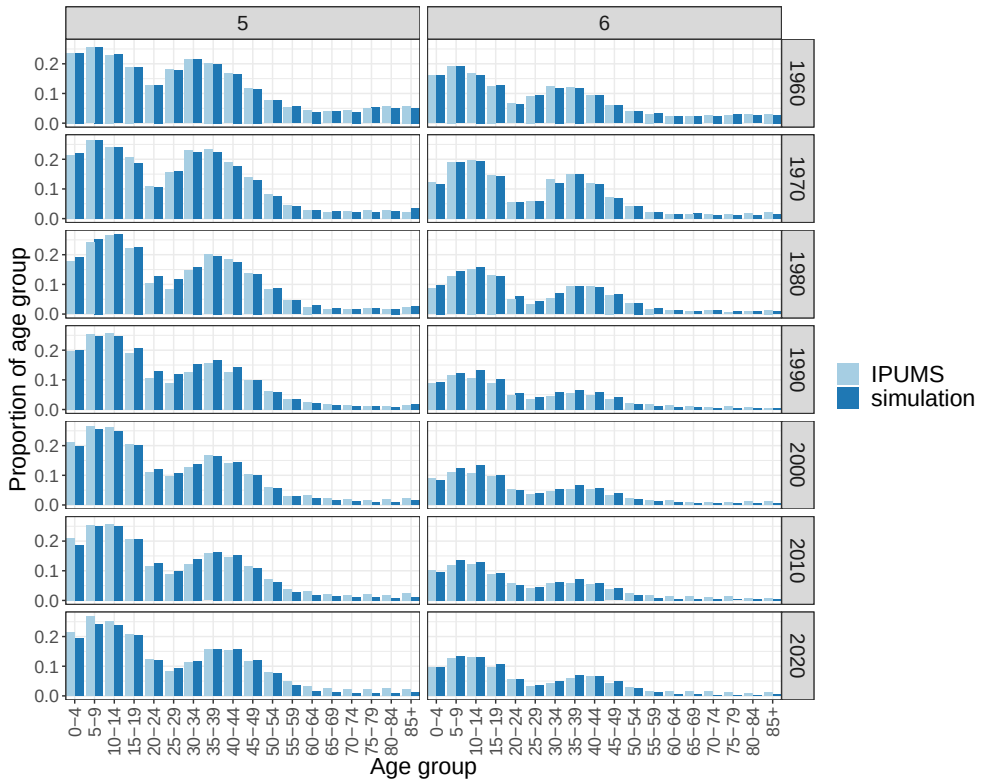

Figure E.7. Proportion of age group (x-axis) living in household size 5 and 6 (columns) by year (row) based on IPUMS (light bars) and simulation (dark bars).

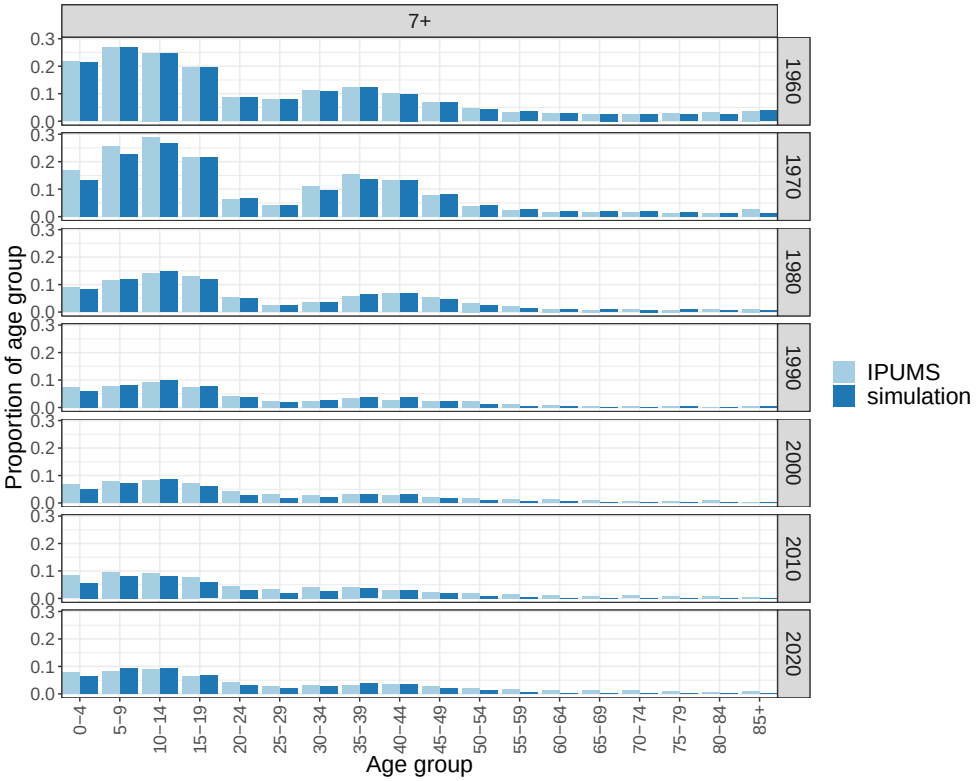

Figure E.8. Proportion of age group (x-axis) living in household size 7 and larger (column) by year (row) based on IPUMS (light bars) and simulation (dark bars).