Assessing the effects of non-pharmaceutical interventions on SARS-CoV-2 spread in Belgium by means of a compartmental, age-stratified, extended SEIQRD model and public mobility data

Tijs W. Alleman[1,*], Jenna Vergeynst[1,2], Lander De Visscher[2], Michiel Rollier[2], Daniel Illana[1], Elena Torfs[1], Ingmar Nopens[1], and Jan M. Baetens[2]

[1]*BIOMATH, Department of Data Analysis and Mathematical Modelling, Ghent University, Coupure links 653, 9000 Gent, Belgium*
[2]*KERMIT, Department of Data Analysis and Mathematical Modelling, Ghent University, Coupure links 653, 9000 Gent, Belgium*
[*]*Corresponding author: Tijs W. Alleman, tijs.alleman@ugent.be*

March 29, 2021

Abstract

We present a compartmental, age-stratified, nation-level, extended SEIQRD model to model SARS-CoV-2 spread in Belgium. We demonstrate the robustness of the calibration procedure by calibrating the model using incrementally larger datasets and dissect the model results by computing the effective reproduction number at home, in workplaces, in schools and during leisure activities. We find that schools are a major contributor to SARS-CoV-2 spread, with the potential to increase the basic reproduction number from $R_e = 0.66 \pm 0.04$ to $R_e = 1.09 \pm 0.05$ under lockdown measures. The model accounts for the main characteristics of SARS-CoV-2 transmission and COVID-19 disease and features a detailed representation of hospitals with parameters derived from two Ghent (Belgium) hospitals. Social contact during the pandemic is modelled by scaling pre-pandemic contact matrices with publically available mobility data from Google and with effectivity parameters inferred from hospitalization data. The combination of the deterministic epidemiological model, which incorporates rigid a-priori knowledge on disease dynamics, and the calibrated effectivity parameters in the social contact model allow us to combine the ease of long-term extrapolation and scenario-analysis of compartmental models with the flexibility of a data-driven model.

keywords: SARS-CoV-2, compartmental epidemiological model, non-pharmaceutical interventions, Google mobility data, effective reproduction number, model calibration

1 Introduction

After an initial outbreak in early 2020 in Wuhan, China, *Severe acute respiratory syndrome coronavirus 2* (SARS-CoV-2) has spread globally [21]. SARS-CoV-2 is capable of sustained human-to-human transmission [31], and may cause severe disease and death, especially in older individuals. The SARS-CoV-2 pandemic has, in general, shown a remarkedly low incidence among children and young adults [8, 27, 36]. Furthermore, presymptomatic transmission is a major contributor to SARS-CoV-2 spread [25, 39]. Both on March 15th, 2020 and on October 19th, 2020, the Belgian governments imposed social restrictions after testing & tracing methods had failed to prevent large-scale spread of SARS-CoV-2. Recently, pharmaceutical interventions under the form of vaccinations have become available. If natural immunity wanes or if SARS-CoV-2 mutates, it is expected that SARS-CoV-2 will become endemic [32]. Hence, there is a need for well-informed models and knowledge build-up to assist policymakers in choosing the best non-pharmaceutical and pharmaceutical interventions during future SARS-CoV-2 outbreaks.

Currently, four other models exist to inform policymakers Belgium. The agent-based model (ABM) of Willem et al. [42], the data-driven model by Barbe et al. [4] and the nation-level, age-stratified compartmental models of Abrams et al. [3] and Franco [12]. The latter feature similar disease dynamics as our own model but rely on different assumptions to model social contact. The different model outputs are currently combined into an ensemble to inform policy makers [41]. In the ensemble, each model fulfills a niche, for instance, the ABM of Willem et al. [42] is good for studying microscopic social behaviour, and was used to inform the optimal *household bubble* size. The model of Barbé excels at short-term forecasts while our own model, together with the compartmental models of Abrams et al. [3] and Franco [12], are well-fit to study long-term scenarios.

In this work, we built a compartmental, age-stratified, nation-level model which accounts for the main characteristics of SARS-CoV-2 disease. The model features a detailed representation of hospitals with parameters derived from two Ghent (Belgium) hospitals. We built a social contact model which scales pre-pandemic contact matrices from a study by Willem et al. [43] with publically available mobility data from Google and with effectivity parameters derived from hospitalization data using an *Markov Chain Monte Carlo* (MCMC) method [13]. Tardiness in compliance to social restrictions is included using a delayed-ramp model. We find that the combination of the deterministic epidemiological model, which incorporates rigid a-priori knowledge on disease dynamics, and the calibrated effectivity parameters in the social contact model allow us to combine the ease of long-term extrapolation and scenario-analysis of compartmental models with the flexibility of a data-driven model. Further, the model does not require *ad-hoc* tweaking and is computationally cheap, making it ideal to perform optimizations which require thousands of model evaluations.

Using a hospitalization dataset of 348 *coronavirus disease 19* (COVID-19) patients in two Ghent (Belgium) hospitals, we computed hospitalization parameters and used them in our model as proxies for Belgium. Using the obtained parameters, we found the model was able to predict the total number of patients in the Belgian hospitals well. We calibrated the model to high-level hospitalization data made publically available by the Belgian *Scientific Instute of Public Health* (Sciensano) and demonstrated its robustness. We computed the basic reproduction numbers (R_0) and the time to reach com-

pliance to lockdown measures during both *coronavirus disease 2019* (COVID-19) waves in Belgium. Using the calibrated model, we computed the relative share of contacts and the effective reproduction numbers and found these to be in line with estimates from other authors at home, at school, at work and during leisure activities to asses their effect on SARS-CoV-2 spread during both 2020 COVID-19 waves. We observed a strong correlation between schools re-opening and increases in SARS-CoV-2 transmission. More precisely, schools have the potential to increase the basic reproduction number from $R_e = 0.67 \pm 0.04$ to $R_e = 1.09 \pm 0.05$ under lockdown measures.

Throughout the work, Belgium is used as a case but the scope of the work is extendable to other countries. Since February 2021, the effects of new SARS-CoV-2 strains and pharamaceutical interventions (vaccines) need to be accounted for. For this purpose, a model extension was developed and is currently used in the aforementioned model ensemble. However, due to the longetivity of this work, we chose to limit the scope of this study to the effects of non-pharamaceutical interventions.

2 Materials and methods

2.1 The extended SEIQRD-model ($\mathbf{SEI_pI_aQRD}$)

2.1.1 Disease dynamics

The SEIR(D) model [19] is a compartmental model that subdivides the human population into four groups: 1. susceptible individuals (S), 2. exposed individuals in the latent phase (E), 3. infectious individuals capable of transmitting the disease (I) and 4. individuals removed from the population either through immunization or death (R/D). Despite being a simple and idealized reality, the SEIR(D) dynamics are used extensively to predict the outbreak of infectious diseases and this was no different during the SARS-CoV-2 outbreak earlier this year [8, 21, 45].

In this work, we extended the SEIRD model to incorporate more expert knowledge on SARS-CoV-2 disease dynamics. For that purpose, the infectious compartment was split into four parts. The first is a period of presymptomatic infectiousness because several studies have shown that presymptomatic transmission is a dominant transmission mechanism of SARS-CoV-2 [25, 39]. After the presymptomatic period, three possible infectious outcomes are modeled: (1) Asymptomatic outcome, for individuals who show no symptoms at all, (2) Mild outcome, for individuals with mild symptoms who recover at home, and (3) Hospitalization, when mild symptoms worsen. Children and young adults have a high propensity to experience an asymptomatic or mild outcome, while older individuals have a higher propensity to be hospitalized [25, 39]. Belgian hospitals generally have two wards for COVID-19 patients: 1) *cohort*, where patients are not monitored continuously and 2) *Intensive care units* (ICUs), for patients with the most severe symptoms. Intensive care includes permanent monitoring, the use of ventilators or the use of extracorporeal membrane oxygenation (ECMO). Patients generally spend limited time in the emergency room and/or in a buffer ward before going to cohort. After spending limited time in cohort, some patients are transferred to ICU. Patients can perish in both wards, but mortalities are generally lower in cohort. After a stay in an ICU, patients return to cohort for recovery in the hospital. During the recovery stay, mortality is generally limited. We assume that mildly infected individuals and hospitalized patients cannot infect susceptibles (Q). The

model dynamics are depicted in Figure 1.

2.1.2 Model structure and equations

In this work, we implemented the extended SEIQRD dynamics shown in Figure 1 using ordinary differential equations, without spatial stratification and with age-stratification. This was accomplished by defining a system of $K.N$ ordinary differential equations, one for every of the $K = 11$ model compartments, each of which is further split into $N = 9$ age-stratified metapopulations. The age-groups have different contact rates with other age-groups and the disease progresses differently for each age-group, making the model behave in a realistic way. Our model consists of 9 age classes, i.e., [0,10[, [10,20[, [20,30[, [30,40[, [40,50[, [50,60[, [60,70[, [70,80[, [80, ∞[. The advantage of this approach are the limited computational resources required to explore scenarios and perform optimizations that require thousands of function evaluations. The disadvantage is the assumption of homogeneous mixing of the entire Belgian population, i.e. every contact between individuals in the Belgian population is assumed to be completely *random*. More realistic approaches are spatial patch models, network-based models, agent-based models or combinations thereof. However, these come at a substantial significant computational cost. Because Belgium is a small and heavily urbanised country, a spatially explicit model becomes relevant at very low SARS-CoV-2 prelevances. The macroscopic coarse-graining of homogeneous mixing works well to describe major COVID-19 waves, but is less fit for monitoring the disease at low prelevances. The model dynamics are translated into the following system of coupled ordinary differential equations,

$$\dot{S}_i = -\beta S_i \sum_{j=1}^{N} N_{c,ij}\Big(\frac{I_j + A_j}{T_j}\Big), \tag{1}$$

$$\dot{E}_i = \beta S_i \sum_{j=1}^{N} N_{c,ij}\Big(\frac{I_j + A_j}{T_j}\Big) - (1/\sigma)\cdot E_i, \tag{2}$$

$$\dot{I}_i = (1/\sigma)E_i - (1/\omega)I_i, \tag{3}$$

$$\dot{A}_i = (\mathrm{a}_i/\omega)I_i - (1/d_{\mathrm{a}})A_i, \tag{4}$$

$$\dot{M}_i = ((1-\mathrm{a}_i)/\omega)I_i - ((1-h_i)/d_m + h_i/d_{\mathrm{hosp}})M_i, \tag{5}$$

$$\dot{ER}_i = (h_i/d_{\mathrm{hosp}})M_i - (1/d_{\mathrm{ER}})ER_i, \tag{6}$$

$$\dot{C}_i = c_i(1/d_{\mathrm{ER}})ER_i - (m_{C,i}/d_{c,D})C_i - ((1-m_{C,i})/d_{c,R})C_i, \tag{7}$$

$$\dot{ICU}_i = (1-c_i)(1/d_{\mathrm{ER}})ER_i - (m_{ICU,i}/d_{\mathrm{ICU},D})ICU_i \tag{8}$$

$$-((1-m_{ICU,i})/d_{\mathrm{ICU},R})ICU_i, \tag{9}$$

$$\dot{C}_{\mathrm{ICU,rec,i}} = ((1-m_{ICU,i})/d_{\mathrm{ICU},R})ICU_i - (1/d_{\mathrm{ICU,rec}})C_{\mathrm{ICU,rec,i}}, \tag{10}$$

$$\dot{D}_i = (m_{ICU,i}/d_{\mathrm{ICU},D})ICU_i + (m_{C,i}/d_{\mathrm{c},D})C_i, \tag{11}$$

$$\dot{R}_i = (1/d_a)A_i + ((1-h_i)/d_m)M_i + ((1-m_{C,i})/d_{c,R})C_i \tag{12}$$

$$+(1/d_{\mathrm{ICU,rec}})C_{\mathrm{ICU,rec,i}}, \tag{13}$$

for $i = 1, 2, \ldots, 9$. Here, T stands for total population (Table 1), S stands for susceptible, E for exposed, I for presymptomatic and infectious, A for asymptomatic and infectious, M for mildly symptomatic and infectious, ER for emergency room and/or buffer ward, C for cohort, $C_{\mathrm{ICU,rec}}$ for a recovery stay in cohort coming from Intensive Care, ICU for Intensive Care Unit, H for hospitalised,

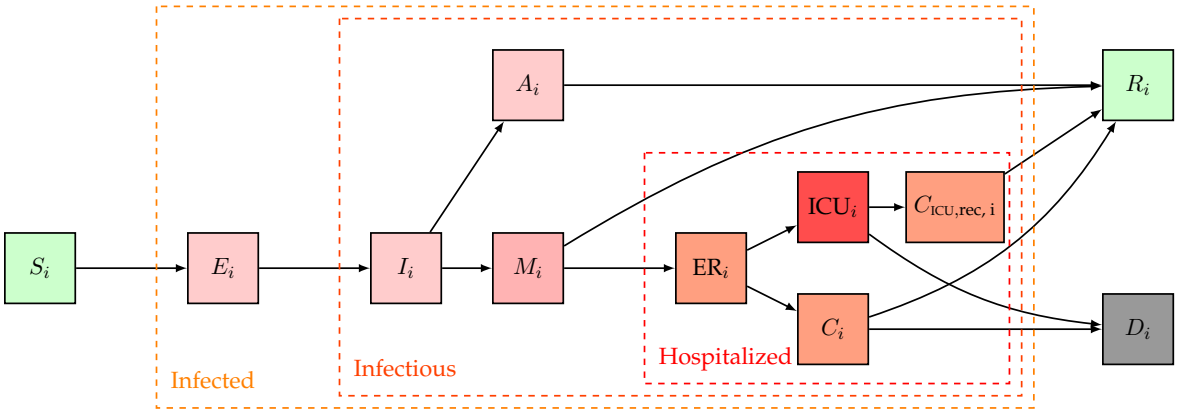

Figure 1: Extended SEIQRD dynamics used in this study. Here, S stands for susceptible, E for exposed, I for presymptomatic and infectious, A for asymptomatic and infectious, M for mildly symptomatic and infectious, ER for emergency room and/or buffer ward, C for cohort, $C_{\text{ICU,rec}}$ for a recovery stay in cohort coming from IC, ICU for Intensive Care Unit, D for dead and R for recovered. A subscript i is used to denote the ith age strate of the model, the model has a total of nine age strata. An overview of the model parameters can be found in table 3.

D for dead and R for recovered. A subscript to these variables is used to refer to one of the nine age strata in the model. Using the above notation, all model states are 9-dimensional vectors,

$$\boldsymbol{S} = [S_1(t)\ S_2(t)\ \ldots\ S_i(t)],$$

where $S_i(t)$ denotes the number of susceptibles in age-class i at time t after the introduction of SARS-CoV-2 in the population. As initial condition, the whole population is assumed susceptible to SARS-CoV-2 and one exposed patient in every age class is assumed, so $E_i(0) = 1$ for all $i = 1, 2, ..., 9$. The time between the start of the simulation and the start of data collection must then be estimated when calibrating the model. An overview of all model parameters, their values and their meaning can be found in table 3. In what follows, the most important model parameters and their chosen values are motivated.

2.1.3 Model parameters

Transmission rate and social contact data The transmission rate of the disease depends on the product of four contributions (Equation 1). The first contribution, $(I_j + A_j)/T_j$, is the fraction of contagious individuals in age group j. We thus assume presymptomatic and asymptomatic individuals spread the disease, while mildly infected are assumed to self-quarantine and hospitalized individuals cannot infect health care workers. The second contribution, $N_{c,ij}$, is the average number of human-to-human interactions of an individual in age group i, with an individual in age group j per day. The sum of the first two contributions over all age groups j, $\sum_{j=1}^{N} N_{c,ij}(I_j + A_j)/T_j$, is the number of contacts of an individual in age group i that can result in SARS-CoV-2 transmission. This is multiplied with the number of susceptibles in age group i (S_i), and with β, the probability of contracting COVID-19 when encountering a contagious individual, to compute the number of effective contacts

at every timestep. We assume that the per contact transmission probability β is independent of age and we infer its distribution by calibrating the model to national Belgian hospitalization data. We assume individuals of all ages have an equal susceptibility to and transmissability to SARS-CoV-2 infection. In a model based inference-based study by Davies et al. [8], it was deduced that children were less susceptible to SARS-CoV-2 disease. However, this is most likely due to underreporting of cases in children. Viner et al. [37] analyzed 32 studies who reported on the susceptiblity of children and found that data were insufficient to conclude that transmissability of SARS-CoV-2 by children is lower than by adults. The number of (pre-pandemic) human-human interactions, $\boldsymbol{N}_c$, are both place and age-dependent. These matrices assume the form of a 9x9 *interaction matrix* where an entry i, j denotes the number of social contacts age group i has with age group j per day. These matrices are available for homes, schools, workplaces, in public transport, and leisure activities, from a study by Willem et al. [43]. The total number of prepandemic social interactions must be translated into an appropriate weighted sum of the contributions in different places, adequately describing pandemic social behaviour (Section 2.3). The basic reproduction number R_0, defined as the expected number of secondary cases directly generated by one case in a population where all individuals are susceptible to infection, is computed using the next generation matrix (NGM) approach introducted by Diekmann et al. [10, 11]. For our model, the basic reproduction number of age group i is,

$$R_{0,i} = (a_i d_a + \omega)\beta \sum_{j=1}^{N} N_{c,ij} \tag{14}$$

and the population basic reproduction number is calculated as the weighted average over all age groups using the demographic data in Table 1. The detailed algebra underlying the computation Equation 14 is presented in the supplementary materials (Section 7.4).

Latent period, Pre-symptomatic infectiousness Estimates for the incubation period range from 3.6 days to 6.4 days, with most estimates close to 5 days. For the average serial interval, the period between the onset of symptoms in the primary case and onset of symptoms in the secondary case, estimates range from 4.0 to 7.5 days [30]. These estimates are roughly the same, which indicates that presymptomatic transmission is contributing to SARS-Cov-2 spread. If the transmission takes place during the symptomatic period of the primary case, the serial interval is longer than the incubation period. However, this relationship can be reversed when presymptomatic transmission takes place. The secondary case may even experience illness onset before onset in their infector, resulting in a negative serial interval. Liu et al. [25] estimated that 23 % of transmissions in Shenzen may have originated from presymptomatic infections, He et al. [17] concluded that 44 % of secondary cases were infected during the presymptomatic stage. Wei et al. [39] investigated all 243 cases of COVID-19 reported in Singapore from January 23–March 16 and identified seven clusters of cases in which presymptomatic transmission is the most likely explanation for the occurrence of secondary cases. They determined that presymptomatic transmission exposure occurred 1-3 days before the source patient developed symptoms. In Equation 3, σ denotes the length of the latent, non-infectious period and in Equation 4, ω is the length of the presymptomatic infectious period. In this work, we assume the incubation period, equal to $\omega + \sigma$, lasts 5.2 days [25]. We then infer the length of the presymptomatic infectious period from hospitalization data using an MCMC method (Section 2.4).

Duration of the infectious period The duration of infectiousness is determined by the number of days patients are able to spread viral particles. He et al. [17] analyzed both data from 77 infector–infectee transmission pairs and 414 throat swabs collected from 94 patients from symptom onset up to 32 days after onset. The highest viral load was observed in throat swabs at the time of symptom onset. Other studies reported similar PCR results, with patients have the highest viral load of the coronavirus at the time they are diagnosed, with viral loads declining gradually over time [20, 26, 35, 47]. The inferred infectiousness profile of He et al. [17] is an approximate normal distribution, with the peak infectivity roughly at the time of symptom onset. Infectiousness was estimated to decline quickly within 7 days after symptom onset. Zou et al. [47] found a steep decline in viral load between 6 and 12 days after symptom onset. A comparison of viral load between symptomatic and one asymptomatic case revealed similar viral loads, an indicator that asymptomatic individuals can be as infectious as symptomatic patients. It was further concluded that viral load alone is not a clear predictor of disease outcome. Given the evidence presented above, we assume the average duration of infectiousness is 7 days for mildly symptomatic cases. The average duration of asymptomatic infectiousness will be inferred from hospitalization data using an MCMC method (Section 2.4).

Disease severity and hospitalizations The model parameter a_i (Equation 4) is the probability of an individual in age group i having a subclinical infection. Several authors have attempted to estimate the fraction of asymptomatic infections. Li et al. [22] estimated that 86 % of coronavirus infections in China were "undocumented" in the weeks before their government instituted stringent quarantines . However, this Figure includes an unknown number of mildly symptomatic cases and is thus an overestimation of the asymptomatic fraction. In Iceland, citizens were invited for testing regardless of symptoms. Of all people with positive test results, 43 % were asymptomatic [15]. As previously mentioned, there is a strong relationship between the age of the patient and the severity of the disease. Wu et al. [44] estimated the relative fraction of asymptomatic cases per age group from publically available Wuhan data. In this study, the subclinical fractions per age group estimated by Davies et al. [8] are used (table 1). For Belgium, this results in a population average subclinical fraction of 57 %. In Equation 5, h_i is the fraction of mild cases that require hospitalization in age group i and in Equation 7, c_i is the fraction of the hospitalized which remain in cohort in age group i. In this study, we use the age-stratified hospitalization probabilities computed by Verity et al. [36] (table 1). The distributions between cohort and ICU are computed using data from 370 patients treated in two Ghent (Belgium) hospitals (Section 2.2).

In Equation 5, d_{hosp} is the average time between first symptoms and hospitalization, which was previously estimated as 5-9 days by Linton et al. [23] and as 4 days by To et al. [35]. In Equation 6, d_{ER} is the time a patient spends in the emergency room and/or buffer wards before going to cohort or ICU. In Equations 7, 8 and 9, $d_{c,R}$, $d_{c,D}$, $d_{\text{ICU},R}$ and $d_{\text{ICU},D}$ are the average lengths of a hospital stay in cohort and in an ICU. The subscript R denotes the duration if the patient recovers, while subscript D denotes the duration if the patient perishes. $\boldsymbol{m_C}$ and $\boldsymbol{m}_{\text{ICU}}$ are the mortalities of patients in cohort and in ICU. In Equation 10, $d_{\text{ICU,rec}}$ is the length of a recovery and observation stay in cohort after being in ICU. Because detailed hospitalization data has unafortunately not yet been made publicly available by the Belgian Scientific Institute of Public Health (Sciensano), all hospitalization parameters are inferred from data of 370 patients, treated in two Ghent (Belgium) hospitals. These parameters are used as temporary proxies to fit the model to the total number of patients in Belgian hospitals and

ICU until the complete Belgian dataset is made publically available. The methodology of the analysis is presented in Section 2.2, the results of the analysis are presented in Section 3.1.

Testing, tracing and quarantine, re-susceptibility Testing and tracing are not explicitly implemented for this study. However, the average effect of the testing and tracing bodies setup during the pandemic will be included in an indirect way after the parameter inference. Liu et al. [24] found that after SARS-CoV-2 infection, it is unlikely that long-lasting protective antibodies are produced. It is thus deemed likely that re-susceptibility will play an important role in future modeling work.

Table 1: Structure of the Belgian population per 10 year intervals [34]. Subclinical fractions per age group as reported by Davies et al. [8]. Hospitalization probability for symptomatic infections per age group in China [36].

Age group (years)	number of individuals	Subclinical fraction (%)	Hospitalization (%)
$[0, 10[$	1 305 219	71	0.1
$[10, 20[$	1 298 970	79	0.3
$[20, 30[$	1 395 385	73	1.2
$[30, 40[$	1 498 535	67	3.2
$[40, 50[$	1 524 152	60	4.9
$[50, 60[$	1 601 891	51	10.2
$[60, 70[$	1 347 696	37	16.6
$[70, 80[$	908 725	31	24.3
$[80, \infty[$	658 753	31	27.3
Population	11.539.326	57	8.1

2.2 Analysis of hospital data

The raw datasets consisted of 396 patients in two Ghent (Belgium) hospitals, Ghent University hospital and AZ Maria Middelares. For every patient the following data were provided: 1) age, 2) sex, 3) hospital admission date, 4) hospital discharge date, 5) date of ICU transfer, 6) date of ICU discharge, 7) outcome (recovered or deceased). Since only age and gender of the patient were provided, all data were anonymous and could not be traced back to individual patients. The date at which patients first reported having symptoms was only recorded by Ghent University hospital, while the time spent in the emergency room, buffer ward or triage was only registered by AZ Maria Middelares. Data from 26 patients had to be excluded from the analysis because one or more entries were missing. Of the remaining 370 patients, 22 patients were discharged from the emergency room and could not be used in the analysis. Thus, in total, data from the remaining 348 patients were used to compute the following hospitalisation parameters: $\boldsymbol{c}$, $d_{c,R}$, $d_{c,D}$, $d_{\text{ICU},R}$ and $d_{\text{ICU},D}$, $d_{\text{ICU,rec}}$, $\boldsymbol{m_C}$ and $\boldsymbol{m}_{\text{ICU}}$. The length of stay in the emergency room or a buffer ward (d_{ER}) was computed solely using data from AZ Maria Middelares and the average time from symptom onset to hospitalisation (d_{hosp}) was computed using

the Ghent University hospital data. To determine if the duration of a cohort or ICU stay differed significantly if the patient recovers or dies, a non-parametric Mann-Whitney U-test was used.

2.3 Social contact model

Mobility reductions As previously mentioned, the social behaviour of the Belgian population must be translated into a linear combination of the aforementioned pre-pandemic interaction matrices. Mathematically, we must find tangible coefficients so that the linear combination of pre-pandemic interaction matrices, i.e.,

$$\boldsymbol{N}_{\text{c}} = \alpha \boldsymbol{N}_{\text{c, home}} + \beta \boldsymbol{N}_{\text{c, schools}} + \gamma \boldsymbol{N}_{\text{c, work}} + \delta \boldsymbol{N}_{\text{c, transport}} + \epsilon \boldsymbol{N}_{\text{c, leisure}} + \phi \boldsymbol{N}_{\text{c, others}} \, , \tag{15}$$

is a good representation of macroscopic social behaviour during the pandemic. Instead of using pre-pandemic contact matrices, modellers would ideally use pandemic contact matrices to build disease models as these are expected to better represent mixing behaviour under lockdown measures. Although these new contact studies under social restrictions will be valuable during future pandemics, such matrices were not available at the start of the pandemic. Hence, our model builds upon pre-pandemic knowledge of social behaviour to make a prediction on pandemic social behaviour.

First, the *Google Community Mobility Reports* for Workplaces, Transit stations, Retail & recreation and Groceries & pharmacy are used as proxies to scale the work, transport, leisure and other social contact matrices (Figure 2). Two surges in COVID-19 cases were observed in Belgium, resulting in two lockdowns. The first lockdown was imposed on March 15th, 2020 and lasted until May 4th, 2020 and involved the closure of schools, bars, clubs, restaurants, all non-essential shops and a closure of border to non-essential travel (Table 4). From May 4th, 2020 until July 1st, 2020 the lockdown was gradually lifted. During the first lockdown, schools remained fully closed until May 18th, 2020 and were only re-opened to a very limited extent before the end of the school year on July 1st, 2020. The second lockdown was imposed on October 19th, 2020 and is still ongoing at the time of writing. Schools were closed on November 2nd, 2020 and re-opened on November 16th, 2020. Further, schools have been closed during Christmas holidays from December 18th, 2020 until January 4th, 2021. Universities have remained fully closed since October 19th, 2020. Briefly summarized, the first 2020 COVID-19 wave consisted of 1) a rapid surge in cases, 2) a lockdown, and 3) a release of lockdown measures. The second 2020 COVID-19 wave consisted of 1) a rapid surge in cases, 2) a lockdown with schools closed, 3) a lockdown with varying school policies. A more detailed overview of all key events in Belgium during the COVID-19 pandemic is provided in the supplementary materials (Section 7.3).

During both lockdowns, residential mobility increases were observed (Figure 2). Although the mobility figures indicate people spent more time at home, this does not mean people have more contacts at home (especially under stay-at-home orders). Further, the coefficients of the interaction matrices are also influenced by *network* effects and may thus not scale linearly with the decrease in the observed mobility. During a strict lockdown where social *bubbles* are the norm, an intra-household contact does not have the same weight as a random, extra-household contact as long as the virus is outside of the social *bubble*. However, amplifying the fraction of household contacts under lockdown measures will increase intergenerational mixing of the population under lockdown, which is unrealistic and will lead to overestimations of the hospitalizations. The inability to capture such dynamics is an inherent

downside of compartmental epidemiological models. We thus used a coefficient of 1.0 for the home interaction matrix $\boldsymbol{N}_{\text{c,home}}$.

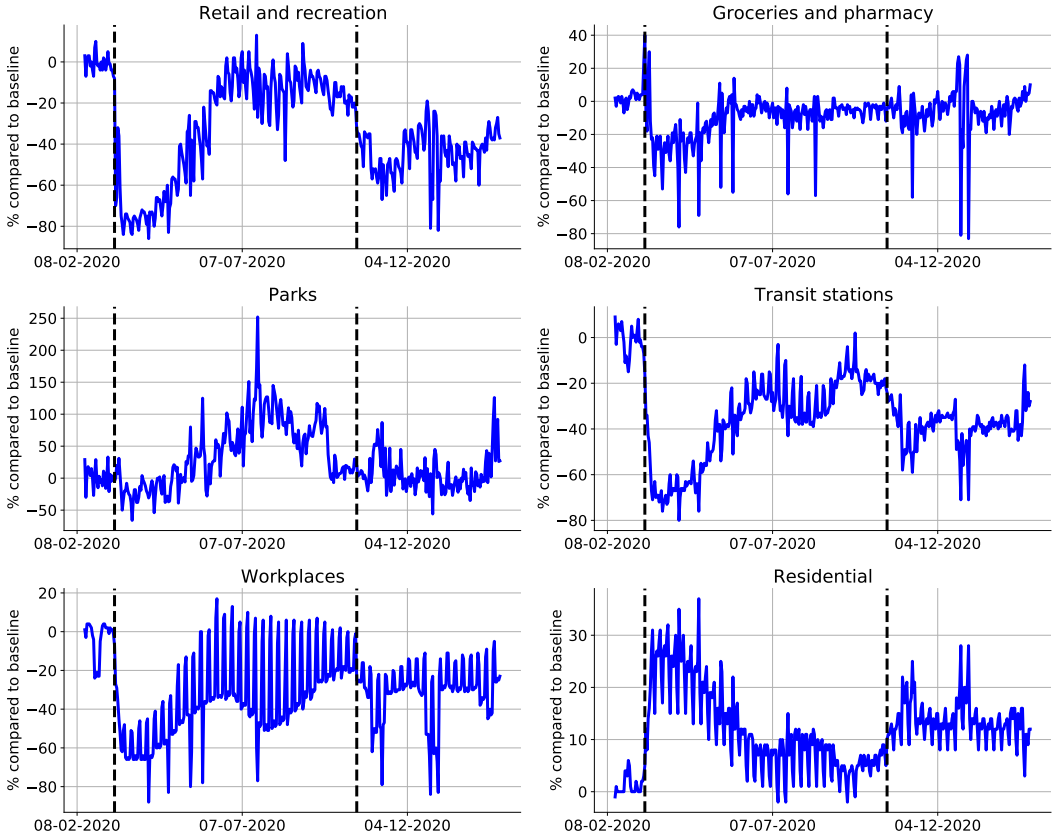

Figure 2: Mobility data extracted from the *Google Community Mobility Reports*. Dashed lines indicates the start of the first lockdown on Friday March 13th, 2020 and the start of the second lockdown on Monday October 19th, 2020. The mobility reduction in *workplaces* is used to scale the work interactions matrix, the *retail & recreation* reduction is used to scale the leisure interaction matrix, the *groceries & pharamacy* reduction is used to scale the other interaction matrix, the *transit stations* reduction is used to scale the public transport mobility matrix.

Effectivity parameters During the first lockdown, we estimated that the overall effectiveness of the contacts (Ω) was approximately one third of what would be expected based on the *Google Community Mobility* reductions and the pre-pandemic contacts. Over the course of the first lockdown, work mobility decreased by 56 %, the public transport mobility decreased by 65 %, leisure mobility decreased by 72 % and grocery (others) mobility decreased by 26 % (Table 4). Mathematically,

$$\boldsymbol{N}_{\text{c}} = \underbrace{\Omega}_{\approx 0.30} \Big[\boldsymbol{N}_{\text{c, home}} + (1-0.56)\boldsymbol{N}_{\text{c, work}} + (1-0.65)\boldsymbol{N}_{\text{c, transport}} + (1-0.72)\boldsymbol{N}_{\text{c, leisure}} + (1-0.26)\boldsymbol{N}_{\text{c, others}} \Big], \tag{16}$$

Intuitively, the effectivity of a contacts may not scale linearily with the observed mobility reductions. The net effectivity of the contacts under lockdown measures depends on a combination of the pre-pandemic physical proximity and duration of the contact, the effectivity of preventive measures and on behavioural changes. As an example, the effects of alcohol gel and face masks might be large in the workplace and in grocery stores, but not at home or during leisure activities. To account for different effectivities of contacts in different places, we could introduce one additional parameter per contact matrix, bound between zero and one, and infer its distribution from the available hospitalization data. However, estimating six effectivity parameters was unfeasible because of identifiability issues. We determined that the effectivity parameters of public transport and other places could not be identified. This is most likely because very little contacts are made in those places [28]. Consequently, the effectivity parameters of public tranport, other places and leisure contacts were lumped to reduce the number of effectivity parameters from six to four. Finally, the linear combination of interaction matrices used to represent social contact under lockdown measures is,

$$\begin{aligned} \boldsymbol{N}_{\mathrm{c}}(t) = \Omega_{\mathrm{home}}\boldsymbol{N}_{\mathrm{c,\,home}} + \Omega_{\mathrm{schools}}H_{\mathrm{schools}}(t)\boldsymbol{N}_{\mathrm{c,\,schools}} + \Omega_{\mathrm{work}}G_{\mathrm{work}}(t)\boldsymbol{N}_{\mathrm{c,\,work}} + \\ \Omega_{\mathrm{rest}}\Big[G_{\mathrm{transit}}(t)\boldsymbol{N}_{\mathrm{c,\,transport}} + G_{\mathrm{retail\ \&\ recreation}}(t)\boldsymbol{N}_{\mathrm{c,\,leisure}} + G_{\mathrm{grocery\ \&\ pharmacy}}(t)\boldsymbol{N}_{\mathrm{c,\,others}}\Big]. \end{aligned} \tag{17}$$

Here, $\boldsymbol{N}_{\mathrm{c,\,home}}$, $\boldsymbol{N}_{\mathrm{c,\,schools}}$, $\boldsymbol{N}_{\mathrm{c,\,work}}$, $\boldsymbol{N}_{\mathrm{c,\,transport}}$, $\boldsymbol{N}_{\mathrm{c,\,leisure}}$ and $\boldsymbol{N}_{\mathrm{c,\,others}}$ denote the pre-pandemic contact matrices at home, in schools, in workplaces, on public transport, during leisure activities and during other activities [43]. G_{work}, G_{transit}, $G_{\mathrm{retail\ \&\ recreation}}$ and $G_{\mathrm{grocery\ \&\ pharmacy}}$ denote the *Google Community Mobility Reports* mobility reductions in the respective categories. H_{schools} is a Heaviside step function to open or close schools, as school opening cannot be deduced from the Google community mobility reports. In spite of their limited re-opening on May 18th, 2020, schools are assumed to be closed during the first lockdown. Ω_{home}, $\Omega_{\mathrm{schools}}$, Ω_{work}, Ω_{rest} are the effectivity parameters at home, in schools, at work and during leisure, public transport and other activities. Note that these parameters are not time-varying, there is thus no link between the effectivity parameters and the mobility reduction. However, when relaxing measures, an increase in mobility will likely be accompanied by an increase of the effectiveness of the contacts. This is due to mentality changes upon relaxation, as measures will gradually be ingnored more.

Obedience to measures In reality, compliance to social restrictions is gradual and cannot be modeled using a step-wise change of the social interaction matrix $\boldsymbol{N}_{\boldsymbol{c}}$ (Section 2.1.3). This can be seen upon close inspection of the *Google community mobility report* after lockdown measures were taken (Figure 2). Because Google mobility data are updated daily in the model, the effect of gradual mobility changes are inherently included. However, the added value of a social compliance model is to gradually introduce the effects of the effectivity parameters in the model. Further, since the compliance model parameters will be estimated from hospitalization data, the added degrees of freedom aid in obtaining a better model fit to the peak hospitalizations. In our model, we use a delayed ramp to model compliance, i.e.,

$$\boldsymbol{N}_{\boldsymbol{c}}(t - t_0) = \boldsymbol{N}_{\mathrm{c,\,old}} + f(t - t_0, \tau, l)(\boldsymbol{N}_{\mathrm{c,\,new}} - \boldsymbol{N}_{\mathrm{c,\,old}}) \tag{18}$$

where,

$$f(t - t_0, \tau, l) = \begin{cases} 0.0, & \text{if } t - t_0 \leq \tau \\ \frac{t - t_0}{l} - \frac{\tau}{l}, & \text{if } \tau < t - t_0 \leq \tau + l \\ 1.0, & \text{otherwise} \end{cases}$$

where τ is the number of days before measures start having an effect and l is the number of additional days after the time delay until full compliance is reached. Both parameters are calibrated to the daily number of hospitalizations in Belgium (Section 2.4). The difference $t - t_0$ denotes the number of days since a change in social policy.

2.4 Parameter identification and model predictions

Aim of the calibration procedure To demonstrate the robustness of the social contact model and calibration method, for each of the 2020 COVID-19 waves, we calibrate the model to a minimal dataset and then increase the amount of data used in the calibration procedure to see if the simulation results and posterior parameter distributions convergence. For the first COVID-19 epidemic, we first calibrate the model using data until April 4th, 2020, and then extend the data range used in the calibration in two week increments until July 1st, 2020. During the second wave, we first calibrate the model on November 7th, 2020 and then extend the calibration to the date of schools re-opening on November 16th, 2020, the date of schools closing for Christmas holidays on December 18th, 2020 and we finally calibrate on February 1st, 2021. By February 1st, 2021, the full impact of school closure and decrease in work mobility during the holiday period is visible in the new hospitalizations. Extending the calibration beyond February 1st, 2021 is part of future research, as the emergence of more contagious strains and the national vaccination campaign need to be included from this point onward (Table 4). As previously mentioned, the effectivity of contacts in schools cannot be studied during the first COVID-19 wave because schools were only opened to a very limited extend before their final closure on July 1st, 2020.

Parameters The model parameters R_0, l, τ, Ω_{home}, Ω_{schools}, Ω_{work} and Ω_{rest} must be calibrated to the available hospitalization data. From Equation 14, the basic reproduction number depends on four model parameters, β, ω, d_a and a_i. We calibrate β, ω, d_a to hospitalization data, Since a_i is an age-stratified parameter, it consists of nine values rendering its calibration computationally unfeasable. The calibration of β, ω and d_a should allow sufficient degrees of freedom to obtain a robust estimate of the basic reproduction number without increasing the demand for computational resources too much. In total, nine parameters must be calibrated to hospitalization data.

Data The aim of the calibration procedure is to obtain a parameter set which leads to a good agreement between the model predictions and the observed data. We calibrate the nine aforementioned parameters to the timeseries of daily new hospitalizations (H_{in}), which are available for download at `https://epistat.sciensano.be/Data`. The model is calibrated seperately to the first and second COVID-19 waves in Belgium.

Statistical model Rather than using the method of least squares (Gaussian model), we aim to explicitly take into account the underlying uncertainty in the data. Since daily hospitalizations are discrete counts, the simplest statistical model describing the model output is the Poisson distribution. An attractive advantage of this distribution is that the variance of a Poisson-distributed random variable is the same as the expected value. The method of least squares inherently assumes the variance of the data are constant. In reality, beause the magnitude of the hospitalizations during a COVID-19

wave varies significantly over time, so does the variance. We thus assume the data are independent and identically distributed (i.d.d.) sequences of poisson variables. The resulting log-likelihood function is,

$$\log L(\boldsymbol{y} \mid \boldsymbol{x}, \boldsymbol{\theta}) = -\sum_{i=1}^{N} \left[y_i(\boldsymbol{\theta}) - x_i \log(y_i(\boldsymbol{\theta})) \right], \tag{19}$$

where the vector of parameters, $\boldsymbol{\theta}$, that maximizes the log-likelihood function must be found. In Equation 19, $\boldsymbol{y}$ denotes the vector of model predictions, $\boldsymbol{x}$ denotes the timeseries of data and N represents the number of datapoints.

Calibration procedure The fitting procedure is performed in two steps. Maximising the result of Equation 19 is computationally demanding and suffers from the presence of local maxima. We thus need an efficient way to scan through the nine-dimensional parameter space $\boldsymbol{\theta} = \{\beta, \omega, ..., \Omega_{\text{rest}}\}$. A good technique to initially broadly identify the region where the global maximum is situated is *Particle Swarm Optimisation* (PSO) [18]. When a region of interest has been identified, we use the maximum-likelihood estimates as initial values for the ensemble sampler for Markov Chain Monte Carlo (MCMC) proposed by Goodman and Weare [13]. For all parameters, uniform prior distributions were used.

2.5 Effects of non-pharamaceutical interventions

To better compare the effects of mobility changes on the daily number of new hospitalizations, we compute the relative share of contacts and the effective reproduction number (R_e) at home, in schools, in workplaces and for the combination of leisure, public transport and other contacts. The number of effective contacts in the aforementioned places at time t are equal to,

$$\begin{aligned}
\boldsymbol{N}^*_{\text{c, home}}(t) &= \Omega_{\text{home}} \boldsymbol{N}_{\text{c, home}}, && (20)\\
\boldsymbol{N}^*_{\text{c, schools}}(t) &= \Omega_{\text{schools}} H_{\text{schools}}(t) \boldsymbol{N}_{\text{c, schools}}, && (21)\\
\boldsymbol{N}^*_{\text{c, work}}(t) &= \Omega_{\text{work}} G_{\text{work}}(t) \boldsymbol{N}_{\text{c, work}}, && (22)\\
\boldsymbol{N}^*_{\text{c, rest}}(t) &= \Omega_{\text{rest}} \Big[G_{\text{transit}}(t) \boldsymbol{N}_{\text{c, transport}} + G_{\text{r \& r}}(t) \boldsymbol{N}_{\text{c, leisure}} + G_{\text{g \& p}}(t) \boldsymbol{N}_{\text{c, others}} \Big], && (23)
\end{aligned}$$

where $\boldsymbol{N}^*_{\text{c, home}}$, $\boldsymbol{N}^*_{\text{c, schools}}$, $\boldsymbol{N}^*_{\text{c, work}}$, $\boldsymbol{N}^*_{\text{c, rest}}$ denote the number of effective contacts at home, in schools, at work or for the sum of leisure, public transport and other contacts. The relative share of contacts in location x is computed as,

$$r_x(t) = \frac{\boldsymbol{N}^*_{\text{c, x}}(t)}{\boldsymbol{N}^*_{\text{c, home}}(t) + \boldsymbol{N}^*_{\text{c, schools}}(t) + \boldsymbol{N}^*_{\text{c, work}}(t) + \boldsymbol{N}^*_{\text{c, rest}}(t)}. \tag{24}$$

The effective reproduction number for age goup i, in place x and at time t is computed as,

$$R_{e,x,i}(t) = \frac{S_i(t)}{S_i(0)} (a_i d_a + \omega) \beta \sum_{j=1}^{N} N^*_{c,x,ij}(t), \tag{25}$$

Finally, the population average effective reproduction number in place x is computed as the weighted average over all age groups using the demographics listed in Table 1.

3 Results

3.1 Hospitalization parameters

The average time from symptom onset to hospitalisation was 7.5 days (IQR 3.0 - 10.0 days). The average stay in the emergency room and/or buffer was 1.1 days (IQR: 0.2 - 1.2 days) and the average time between admision in cohort and admission in ICU was 1.3 days (IQR: 0.0 - 2.0 days). Of the 348 hospitalised patients, 86 patients (24.7 %) required intensive care at some point during their stay and 262 (75.3 %) remained in cohort. The overall mortality in the hospital is 19.5 %, the mortality in cohort was slightly lower at 16.8 % and the mortality in ICU was higher (27.9 %). The average length of the stay in a cohort ward was 7.7 days (IQR: 4.0 - 10.0 days) and the average length of an ICU stay was 12.5 days (2.9 - 17.9 days) ($p = 0.05$). The average cohort stay was 8.0 days (IQR: 4.2 - 10.9 days) if the patient had recovered and 6.1 days (IQR: 3.0 - 7.6 days) if the patient had died ($p = 0.007$). The average ICU stay was 9.9 days (IQR: 2.0 - 14.0 days) if the patient had recovered and 19.3 days (IQR: 7.9 - 23.8 days) if the patient had died ($p = 0.001$). Out of 62 surviving patients in ICU, 46 stayed in a cohort ward after their stay in ICU, with an average length of stay of 11.7 days (IQR: 6.0 - 13.3 days). A violin representation of the residence time distributions can be found in Figure 9. The computed age-stratified hospitalization parameters $\boldsymbol{c}$, $\boldsymbol{m}_{\mathrm{C}}$ and $\boldsymbol{m}_{\mathrm{ICU}}$, are listed in Table 2. From the sample sizes it can be deduced that older individuals have a much higher chance of hospitalization than younger individuals. In general, both the chance of needing intensive care and the mortalities increase with the patient's age. In the dataset, no indivuals under 30 years old were in need of intensive care and no individuals under 40 year old have died in the hospital from a SARS-CoV-2 infection. Hence, the mortalities in cohort and ICU, $\boldsymbol{m}_{\mathrm{C}}$ and $\boldsymbol{m}_{\mathrm{ICU}}$, are equal to 0.0 % for individuals under 40 years old. This will not impact the goodness-of-fit to the Belgian hospitalization data, but care is warranted when using these data to perform health economic evalutions such as computing *quality adjusted life years* lost to COVID-19. The age-stratified mortalities, age-stratified distributions between cohort and ICU and the residence time distributions derived from the hospital dataset were propagated in the model, resulting in a good prediction of the total number of patients in Belgian hospitals during both 2020 COVID-19 waves (Figure 10).

3.2 Model calibration

The population average basic reproduction number was computed as $R_0 = 4.16$ (IQR: 3.90 - 4.39) for the first 2020 COVID-19 wave and as $R_0 = 3.69$ (IQR: 3.64 - 3.75) for the second 2020 COVID-19 wave. Compliance to social measures were similar for both 2020 COVID-19 waves, with an average delay of 0.22 (IQR: 0.07-0.31) and 0.39 (IQR: 0.20 - 0.52) days, and a time to reach full compliance to measures of 9.17 (IQR: 8.89 - 9.50) and 6.94 (IQR: 6.71 - 7.18) days respectively.

Figure 4 summarizes the results of six model calibrations using hospitalization datasets starting on March 15th, 2020 until April 4th, 2020 and subsequently increasing in two week increments. Here, Figure 4 (a) represents the minimal dataset, where the data range used for the calibration was equal to March 15th, 2020 until April 4th, 2020. Opposed is Figure 4f, which uses the maximal dataset, using hospitalization data from March 15th, 2020 until July 1st, 2020. Using the mimimal dataset

Table 2: Computed fraction of hospitalized remaining in cohort and not transferring to ICU (c), mortality in cohort (m_{C}) and mortality in ICU (m_{ICU}) per age group for two Belgian hospitals ($n = 370$).

Age group (years)	sample size	c (%)	m_{C} (%)	m_{ICU} (%)
$[0, 10[$	2	100.0	0.0	0.0
$[10, 20[$	7	100.0	0.0	0.0
$[20, 30[$	9	80.0	0.0	0.0
$[30, 40[$	9	87.5	0.0	0.0
$[40, 50[$	33	78.1	8.0	14.3
$[50, 60[$	67	74.2	2.2	6.3
$[60, 70[$	62	61.3	0.0	16.7
$[70, 80[$	74	64.3	17.8	44.0
$[80, \infty[$	107	88.2	36.7	58.3
Average	370	75.3	16.8	27.9

(Figure 4a), the posterior distributions are uninformative and model prediction uncertainty is large. Using additional data from April 15th, 2020 (Figure 4b) onwards, the model captures the observed downward trend in the hospitalization data. Before the release of social restrictions on May 4th, 2020 (Figure 4a-4c), the posterior distributions seem to converge to distributions different from the ones found using the maximal dataset (Figure 4f). However, during the gradual lifting of lockdown restrictions (Figure 4d-4f), the posterior distributions monotonically converge to their final distributions.

In a similar fashion, four calibrations on hospitalization datasets of increasing length during the second COVID-19 wave were performed and the results are summarized in Figure 5. Once more, the minimal dataset (Figure 5a), which uses data from September 1st, 2020 until November 7th, 2020 does not result in informative posterior distributions of the effectivity parameters. Uncertainty on the model prediction is large, but the mean model prediction is ballpark accurate. As soon as schools are opened on November 16th, 2020, the daily hospitalizations evolve to a plateau. In spite of large uncertainty on the model prediction, the emergence of the hospitalization plateau is captured in the uncertainty band, and the model thus provides a ballpark estimate using the minimal dataset. Although model accuracy has risen, a similar conclusion can be drawn for the calibration using data until schools re-opening on November 16th, 2020. When including data in the hospitalization dataset until schools closure for the Christmas holidays on December 18th, 2020 (Figure 5c), the model correctly attributes the increased transmission to the opening of schools. In Figure 5c, it can be seen that the the effectivity parameter for schools is almost equal to the maximum value of one. Although the posteriors of the effectivity parameters still differ significantly from their final distributions, the model provides an accurate prediction for the future evolution of the new hospitalizations during Christmas holidays and until schools re-opening on January 4th, 2021. From the inference using the maximal dataset (Figure 5d), it is clear that the model attributes high effectivities for contacts at home and in schools.

3.3 Effects of non-pharamaceutical interventions

To better compare the effects of non-pharamceutical interventions between both 2020 COVID-19 waves, we computed the relative share of contacts and the effective reproduction number at home, in schools, in workplaces and for the sum of leisure, public transport and other contacts (Figure 6). In this way, we are able to dissect the force of infection in our model, allowing us to asses the relative impact of contacts made at different locations on SARS-CoV-2 transmission. In prepandemic times, leisure and work contacts account for the bulk of total contacts, while under strict lockdown measures (March 15th, 2020 - May 4th, 2020 and October 19th, 2020 - November 16th, 2020), the contacts at home are the main driver of SARS-CoV-2 spread. The effective reproduction number under strict lockdown measures was equal to $R_e = 0.67$ (IQR: 0.48 - 0.76) for the first COVID-19 epidemic and was equal to $Re = 0.66$ (IQR: 0.61 - 0.69) for the second COVID-19 epidemic. Aside from the interactions at home, leisure contacts had the second most impact during the first COVID-19 wave, with roughly twice the impact of work contacts. When lifting social restrictions from May 4th, 2020 onwards, the relative contribution of home contacts gradually declines, while the contributions of work and leisure become more important. The effective reproduction number gradually increases and approaches the critical value of $R_e = 1$ by the beginning of summer (average of June, 2020 $R_e = 0.91$, IQR: 0.77 - 1.00).

As soon as schools are re-opened on November 16th, 2020, a plateau in the daily number of hospitalizations emerges (Figure 6). There were no other major policy changes around this time, except schools re-opening. Our model deduces this correlation by inferring posterior values of the effectivity of contacts in schools close to one, meaning school contacts were highly effective for SARS-CoV-2 transmission. Schools have an impact similar to the home interactions, with both contributing roughly 40 % to the total number of effective contacts during the second COVID-19 wave. It is clear that the opening of schools under lockdown can tip the scale, and push the effective reproduction number just above the critical value of $R_e = 1$. When shools are opened, the effective reproduction number increases from $R_e = 0.66 \pm 0.04$ to $R_e = 1.09 \pm 0.05$, causing a stagnation in the decline of the daily hospitalizations. To further validate this result, we extracted the number of laboratory confirmed cases in youths $[0, 20[$, the working population $[20, 60[$ and the senior population $[60, \infty[$ from the *Belgian Scientific Institute of Public Health* (Sciensano). The timeseries were normalized with the number of cases on November 21st, 2020[1] to allow a better comparison. The number of laboratory confirmed cases amongst youths start increasing as soon as schools are openend on November 16th, 2020 (Figure 3). A similar pattern is observed during school closure and re-opening for Christmas holidays, altough it should be noted the relationship is less clear. This is most likely the effect of Christmas and New Year celebrations and returning travellers. The use of a time-lagged cross-correlation revealed a significant lead-relationship between the number of cases in youths and the working population by 9 days, and a leading relationship between the number of cases amongst youths and the senior population by 13 days (Section 7.5).

[1] Date of school reopening 2021-11-16 plus one five day incubation period.

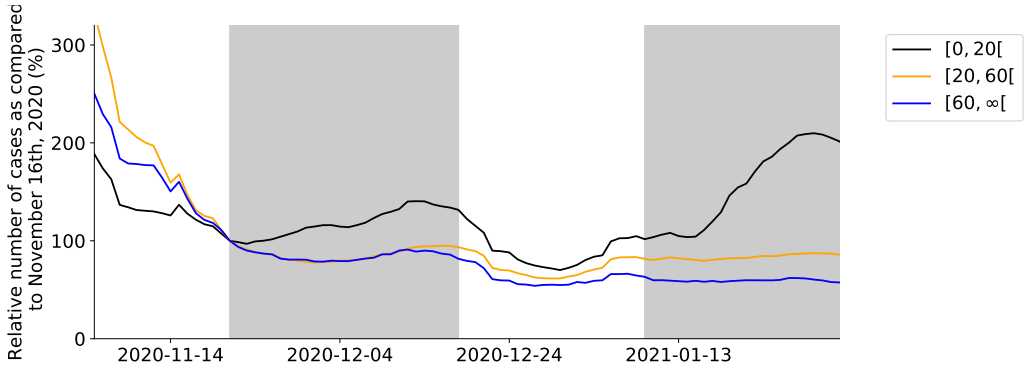

Figure 3: Relative number of confirmed cases in youths, the working population and the senior population during the period November 2nd, 2020 until February 1st, 2021, as compared to the number of confirmed cases in each group on November 16th, 2020. The grey shade is used to indicate schools were open.

4 Discussion

We computed hospitalization parameters using data from two hospitals in Ghent. The ICU admission probabilities and mortalities in cohort and in ICU indicate that COVID-19 has a much higher severity in older individuals, which is in line with estimates from authors in the USA and China [5, 36]. The average time from symptom onset to hospitalisation was 7.5 days, which is in line with estimates from Linton et al. [23]. Of the 348 hospitalized patients, 86 patients (24.7 %) required intensive care at some point during their stay and 262 (75.3 %) remained in cohort. The result is in line with the estimate of Wu and McGoogan [46], who estimated that one quarter of all hospitalized patients requires intensive care. The inferred hospitalization parameters are valuable proxies for Belgium and resulted in an adequate goodness-of-fit to the total number of patients in Belgian hospitals.

We obtained an average basic reproduction number of $R_0 = 4.16$ (IQR: 3.90 - 4.39) for the first 2020 COVID-19 wave and of $R_0 = 3.69$ (IQR: 3.64 - 3.75) for the second 2020 COVID-19 wave, which is in line with the global consensus range of $R_0 = [2, 4]$. The estimate for the second COVID-19 wave is slightly lower, and this is most likely because this estimate implicitly includes the effects of preventive measures and mentality changes that were gradually adopted during the first 2020 COVID-19 wave. Large differences in the basic reproduction number exist between the different age groups (Figure 7). It is clear that the youths and working-aged population drive the pandemic while people of ages 70 or above can hardly sustain a SARS-CoV-2 pandemic amongst themselves, this is mainly because elderly individuals have limited social interactions (Figure 7). Still, these individuals make up roughly 35 % of all hospitalizations. The biggest risk group are the individuals aged 50 to 70, which make up roughly 50 % of the expected hospitalizations. The high expected fraction of hospitalizations in this age group is due to a trade-off between social contact and hospitalization risk. These individuals have of plenty social contact and at the same time, have a high propensity to hospitalization. The compliance to social measures were similar between the both 2020 COVID-19 waves, little lag was observed (0.22 vs. 0.39 days) and the time to reach full compliance were of the same magnitude

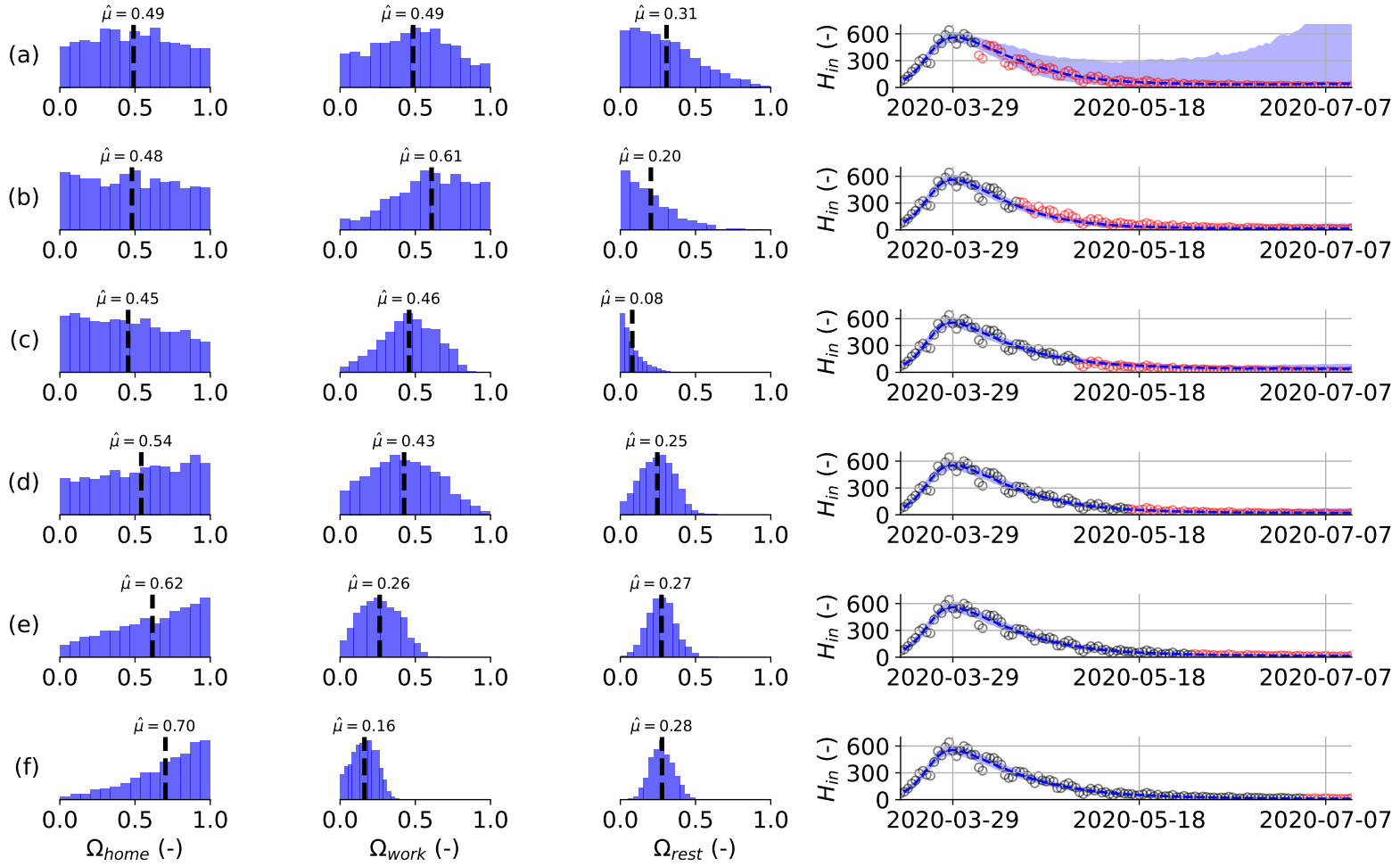

Figure 4: (left) Estimated posterior distributions for the effectivity of a contact at home (Ω_{home}), in the workplace (Ω_{work}) and for the sum of leisure activities, other activities and public transport (Ω_{rest}), (right) together with the resulting model prediction for the daily hospitalizations from March 15th, 2020 until July 14th, 2020 (right). The effectivity of school contacts could not be deduced during the first 2020 COVID-19 wave because schools were only re-opened very limited before their final closure on July 1st, 2020. Calibration performed using the daily hospitalizations in Belgium until: (a) 2020-04-04, (b) 2020-04-15, (c) 2020-05-01, (d) 2020-05-01, (e) 2020-06-01 and (f) 2020-07-01. Calibration data in black, validation data in red. Model predictions are accurate in all but the minimal calibration dataset (a). Monotonic convergence of the effectivity parameter posteriors is reached quickly after lockdown release on May 4th, 2020 (d-f).

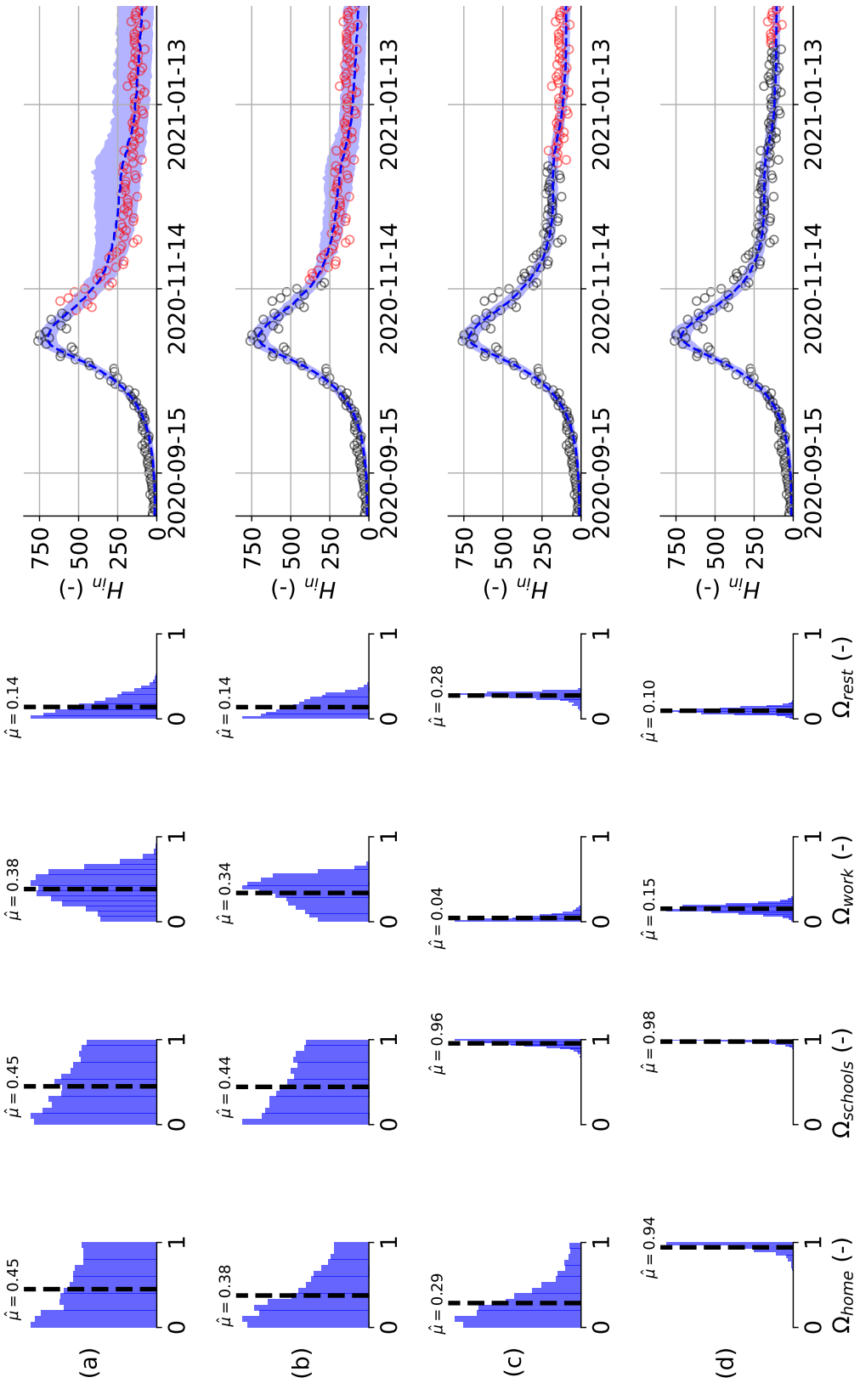

Figure 5: Estimated posterior distributions for the effectivity of a contact at home (Ω_{home}), at school (Ω_{schools}), in the workplace (Ω_{work}) and for the sum of leisure activities, other activities and public transport (left), together with the resulting model prediction for the daily hospitalizations from September 1st, 2020 until February 14th, 2021 (right). Calibration performed using the daily hospitalizations in Belgium until: (a) 2020-11-07, (b) 2020-11-16, (c) 2020-12-18, (d) 2021-02-01. Calibration data in black, validation data in red. Model predictions are accurate for all calibration datasets. Monotonic convergence of the schools effectivity parameter is reached a-posteriori schools re-opening (c).

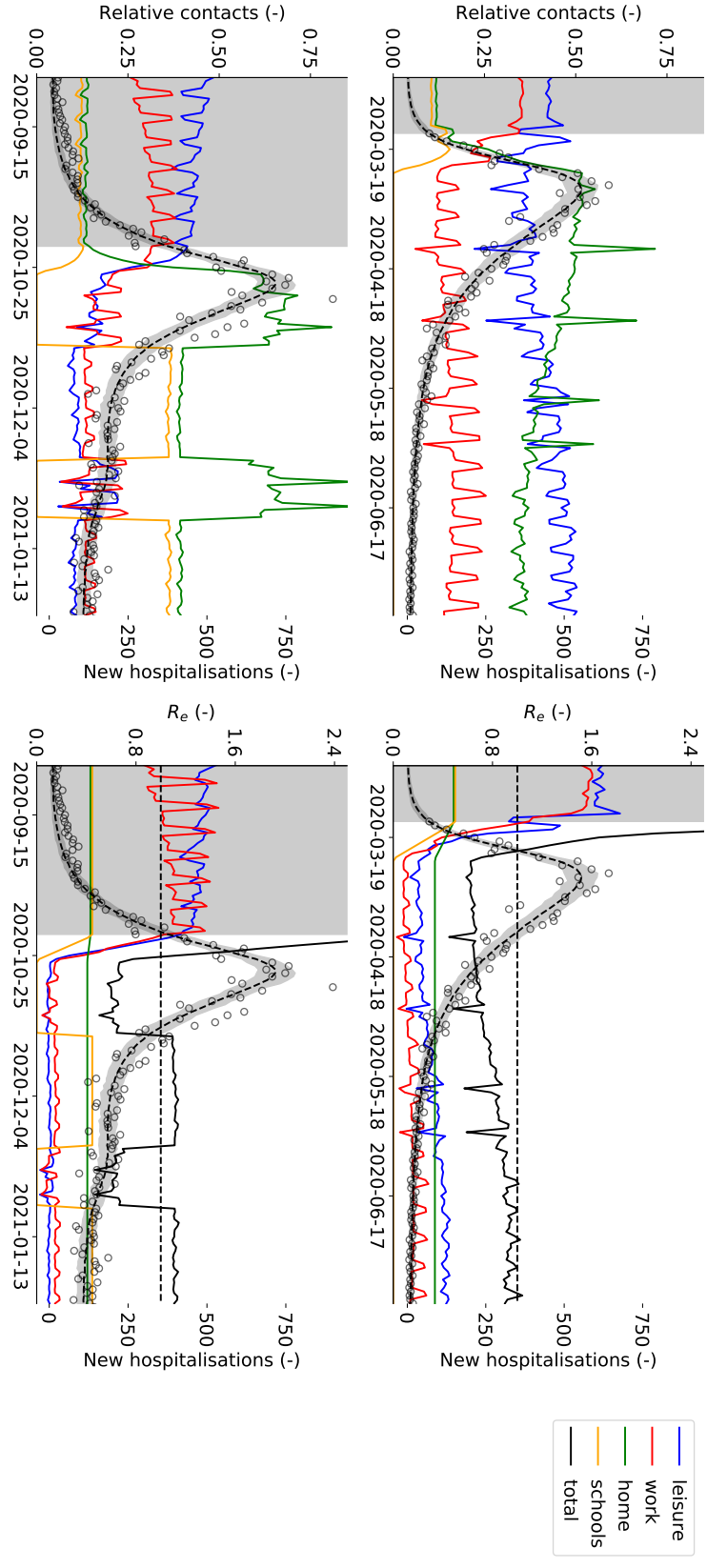

Figure 6: (First column) Relative share of contacts at home, in the workplace, in schools and for the sum of leisure activities, (Second column) effective reproduction number (R_e) at home, in the workplace, in schools and for the sum of leisure activities, other activities and public transport. The right axis denotes the predicted number of daily Belgian hospitalizations. The first row depicts the first COVID-19 wave in Belgium, from March 15th, 2020 until July 14th, 2020, while the second row depicts the second COVID-19 wave in Belgium, from September 1st, 2020 until February 1st, 2020. Mean and 95 % confidence interval of 1000 model realisations. During both lockdowns, home interactions have the largest share of effective contacts. During lockdown release, the relative importance of work and leisure contacts start increasing. Schools opening and closing has a large impact on the effective reproduction number, and can end a decreasing trend in hospitalizations.

(9.17 vs 6.94 days). Thus, compliance to lockdown restrictions can be modeled using a ramp function without lag, eliminating one of the model's parameters, namely τ (Equation 18).

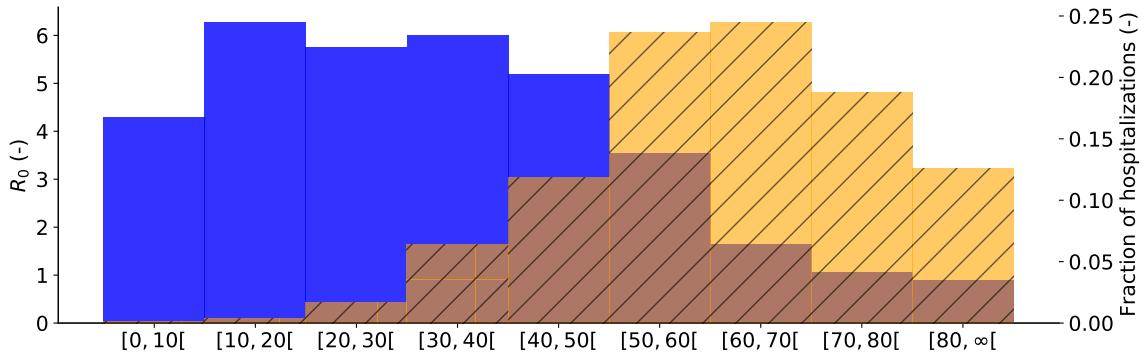

Figure 7: Basic reproduction number per age group ($R_{0,i}$), for Belgium (blue). Expected fraction of the total Belgian hospitalizations during the first COVID-19 wave, as predicted by the model, from March 15th, 2020 until July 1st, 2020 in age group i (orange, striped). Youths and working-aged population drive the pandemic, while the senior population is mostly in need of hospital care.

We calibrated the model's effectivity parameters (Ω_{home}, Ω_{schools}, Ω_{work}, Ω_{rest}) on incrementally larger hospitalization datasets and found that the model provides accurate forecasts under the observed mobility changes, even when the posteriors still depend on the extent of the dataset. However, *correct*[2] effectivity parameters could only be deduced a posteriori events. This is because *information* on the effectiveness of contacts can only be obtained by observing the hospitalizations under changing policies. Examples are the effects of leisure and work relaxations during the first COVID-19 wave and the effect of schools re-opening during the second COVID-19 wave. From April 15th, 2020 onwards (Figure 4, panel b) the ever decreasing trend in the daily hospitalizations is nicely captured even with posteriors seemingly converging to distributions different than those of the maximal dataset (panel f). In spite, on May 1st 2020 (panel c), the model could have been used to accurately inform policymakers on the effects of lifting work and leisure restrictions just four days later. As soon as restrictions are lifted, the posteriors quicly converge to their final distributions. A similar observation is made with regard to the schools effectivity parameter. From November 7th, 2020 onwards (Figure 5, panel a) the effect of schools re-opening is captured in the model uncertainty, in spite of deviant posterior distributions. From Decenber 18th, 2020 onwards (panel c) the effect of schools re-opening is captured both in the model predictions and the effectivity parameters. Because accurate posteriors can only be inferred a posteriori, the modeler must asses if policy changes have been sufficient to deduce accurate effectivity posteriors. This is important when performing scenario analysis, as incomplete knowledge of the effectivity posterior can significantly alter the results. Our calibration method allows to extract the most out of pre-pandemic contact studies and publically available mobility data, and has proven usefull to inform policy makers during the acute crisis.

[2] Assuming the inferred posterior distributions of the maximal dataset are correct.

Finally, we would like to further discuss the importance of schools in the SARS-CoV-2 pandemic. As previously mentioned in section 3.3, there seems to be a strong correlation between school re-opening, the rise of laboratory confirmed cases amongst youths, the rise of the number of clusters in schools, and the emergence of plateaus in the daily hospitalizations (Figures 3 and 6). Our model incorporates this correlation as high effectivities of school contacts. An increase in the effective reproduction number, from $R_e = 0.66 \pm 0.04$ to $R_e = 1.09 \pm 0.05$, is observed when schools are re-opened. Although the present evicence is circumstantial, and correlation does not imply causation, schools seem to play a critical role in SARS-CoV-2 spread. Thus, schools closure is an effective way of countering an epidemic SARS-CoV-2 trend. However, given the negative consequences of school closure, we recommend policymakers to invest in ways to ensure qualitative online education during future SARS-CoV-2 outbreaks, but even more generally, in case of a future pandemic.

5 Conclusions

- We obtained an average basic reproduction number of $R_0 = 4.16$ (IQR: 3.90 - 4.39) and $R_0 = 3.69$ (IQR: 3.64 - 3.75) for both 2020 COVID-19 waves in Belgium. We found that SARS-CoV-2 strongly discriminates between inidividuals of different age groups, with youths and the working-aged population driving the pandemic, and the senior population needing hospital care. These results are in line with the established consensuses and highlight the model's validity.
- The combination of the deterministic epidemiological model, which incorporates a-priori knowledge on disease dynamics, and the social contact model whose infectivity parameters were inferred allow us to make the most out of the available pre-pandemic data. The method is computationally cheap and does not require ad-hoc tweaking to obtain a good fit to the observed data. A disadvantage is that the effectivity parameter distributions only monotonically converge to their *correct* posterior distributions a posteriori policy changes. Still, the model is able to make accurate predictions on the daily number of hospitalizations under the observed mobility, but care by the modeller is warranted when performing scenario analysis.
- As soon as schools were re-opened on November 16th, 2020, the number of confirmed cases amongst youths starts increasing. A significant lead relationship between the number of cases amongst youths and the working population, and youths and the senior population was found. Our model incorporates this correlation as high effectivities of school contacts. When schools were re-opened under lockdown policies, the model indicates the effective reproduction number increased from $R_e = 0.66 \pm 0.04$ to $R_e = 1.09 \pm 0.05$. Thus, schools closure is an effective measure to counter an epidemic SARS-CoV-2 trend.

6 Future research

- The calibration procedure should be repeated using pandemic social contact matrices, which are currently being gathered for Belgium by Coletti et al. [6]. Further, the effects of integrating the contacts with their duration should be explored. A comparison between the different results can then be made.

- The effective reproduction number in the different places should be compared to data on SARS-CoV-2 clusters to further validate the model.
- It is expected that lockdown measures in Belgium will be lifted soon. The impact of releasing measures on the daily hospitalizations should be studied to find a link between the effectivity parameters and the mobility reductions.
- If schools are a major contributor to SARS-CoV-2 spread, administering a vaccine with high transmission-blocking potential to youths is expected to have a similar effect as schools closure. Due to their localized nature, vaccination for SARS-CoV-2 in schools is logistically easier than vaccinating the general population.

Acknowledgements

This work is the result of a team effort. I want to thank Bram De Jaegher, Daan Van Hauwermeiren, Stijn Van Hoey and Joris Van den Bossche for their help in maintaining the GitHub repo, coding the visualizations and for teaching me the basics of object-oriented programming in Python. I would like to thank Mieke Descheppere, statistician at the Ghent University hospital and Wim Verbeke, MD from AZ Delta Roeselare, for sharing their insights on hospital dynamics. I would like to thank prof. Steven Callens, MD, prof. Pascal Coorevits, MD and prof. Ernst Nietzschel, MD of the Ghent university hospital and Leen Van Hoeymissen from AZ Maria Middelares for sharing their data. We thank VZW 100 km Dodentocht Kadee for their financial support through the organisation of the 2020 100 km COVID-Challenge.

Competing interests

The authors have no competing interests to declare.

Role of the funding source

This work was supported by the UGent *Special Research Fund* and by the *Research Foundation Flanders* (FWO), project number G0G2920N. Further, the computational resources and services used in this work were also provided by the VSC (*Flemish Supercomputer Center*), funded by FWO and the Flemish Government. The funding sources played no role in study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication.

CRediT author statement

Tijs W. Alleman: Conceptualization, Software, Methodology, Investigation, Data Curation, Writing - Original Draft. **Jenna Vergeynst:** Conceptualization, Software, Writing - Review. & Editing, Project administration. **Lander De Visscher:** Methodology. **Michiel Rollier:** Methodology, Writing - Review. **Elena Torfs:** Conceptualization, Funding acquisition. **Daniel Illana:** Visualization. **Ingmar**

Nopens: Conceptualization, Funding acquisition, Project administration, Writing - Review. & Editing **Jan Baetens:** Conceptualization, Funding acquisition, Project administration, Writing - Review & Editing.

7 Supplementary materials

7.1 Overview of model parameters

Table 3: Overview of simulation parameters used in the age-stratified extended SEIQRD model.

Symbol	Parameter	Value	Unit	Reference
$\boldsymbol{a}$	subclinical fraction per age group	Table 1	$(-)$	Davies et al. [8]
$\boldsymbol{h}$	fraction of mildly infected individuals requiring hospitalisation	Table 1	$(-)$	Verity et al. [36]
$\boldsymbol{c}$	fraction of hospitalisations requiring ICU transfer	Table 2	$(-)$	Hospital dataset
d_a	duration of subclinical infection	4.1	days	Inferred
d_m	duration of mild infection	7	days	To et al. [35]
d_{hosp}	average time from symptom onset to hospitalization	7.5	days	Hospital dataset
d_{ER}	time spent in the emergency room and/or in a buffer ward	2.5	days	Hospital dataset
$d_{c,R}$	length of cohort stay if recovered	8.0	days	Hospital dataset
$d_{c,D}$	length of cohort stay if deceased	6.1	days	Hospital dataset
$d_{\text{ICU},R}$	length of ICU stay if recovered	9.9	days	Hospital dataset
$d_{\text{ICU},D}$	length of ICU stay if deceased	19.3	days	Hospital dataset
$d_{\text{ICU,rec}}$	length of recovery and observation stay in cohort after ICU stay	6.5	days	Hospital dataset
$\boldsymbol{m_C}$	mortality in cohort	Table 2	$(-)$	Hospital dataset
$\boldsymbol{m_{ICU}}$	mortality in ICU	Table 2	$(-)$	Hospital dataset
σ	length of latent period	4.9	days	Computed
ω	length of presymptomatic infectious period	0.3	days	Inferred
$\sigma + \omega$	length of incubation period	5.2	days	Liu et al. [25]
β	probability of infection upon contact with an individual capable of transmitting SARS-CoV-2 under the assumption that the infectee is 100 % susceptible to SARS-CoV-2 infection	0.073	$(-)$	Inferred
$\boldsymbol{T_0}$	total population	Table 1	people	StatBEL [34]
N_c	contact matrix	9x9 matrix	day^{-1}	Willem et al. [43]

7.2 Overview of model assumptions and limitations

The following assumptions were made with regard to the SEIQRD dynamics:

1. All individuals experience a brief presymptomatic, infectious period.
2. All individuals are equally susceptible to SARS-CoV-2 infection.
3. Asymptomatic and mild cases automatically lead to recovery and in no case to death.
4. Mildly infected and hospitalized individuals cannot infect susceptibles (= *quarantined*). Thus, the model cannot be used to model the effect of transmission to healthcare workers.
5. All deaths come from hospitals, meaning no patients died outside a hospital [14].
6. Both 2020 COVID-19 waves were fitted seperately, assuming a fully susceptible population.

The following assumptions to the hospital dynamics were made:

1. All patients spend time in the emergency room and/or buffer ward before going to either a cohort ward or an ICU.
2. After the emergency room and/or buffer ward, patients immediately branch into ICU or cohort. In real life, a patient may first spend some time in a cohort ward before going to an ICU and this is not accounted for.
3. Residence times in cohort and in ICU differ depending on the outcome of the infection (recovered or deceased).
4. All recovered ICU patients spend some additional time in cohort (recovery and observation stay).

The following assumptions were made in the social contact model:

1. Prepandemic contact matrices by Willem et al. [43] are scaled with mobility reductions extracted from the *Google Community Mobility Reports* and an effectivity parameter inferred from hospitalization data using a *Markov-Chain Monte-Carlo* method in order to mimic pandemic social behaviour.
2. The pre-pandemic contact matrices used included all contacts, both physical and non-physical, and the duration of contacts is not integrated.
3. The effectivity of the contacts (Ω_x) are bound between zero and one. This implies that if work mobility is reduced to 40 % of its pre-pandemic value, the work contacts can account for no more than 40 % of their pre-pandemic value.
4. There is no link between the effectivity parameters and the mobility reduction. However, when relaxing measures, an increase in mobility will likely be accompanied by an increase of the effectiveness of the contacts. This is due to mentality changes upon relaxation, as measures will gradually be ingnored more.

7.3 Key events

The first lockdown, which started on March 15th, 2020 and lasted until May 4th, 2020 involved the closure of schools, bars, clubs, restaurants, all non-essential shops and a closure of border to non-essential travel (Table 4). The *Community Mobility Reports* show a 56 % reduction in workplace mobility (Figure 2 and Table 4). Based on surveys from the Belgian National Bank, 28.6 % of all employees was able to work from home, 29.9 % remained in the workplace and 4.4 % worked both from home and in the workplace. 32.4 % were temporary unemployed and 4.8 % were absent [40]. Public transport mobility decreased by 65 %, leisure mobility decreased by 72 % and grocery & pharmacy mobility was reduced by 26 %. From March 15th, 2020 until May 4th, 2020, mobility remained practically constant at the aforementioned reductions. On May 4th, 2020 the lockdown was gradually lifted by re-opening all non-essential shops and lifting telework restrictions. The effect can be seen in the *Google Community Mobility Reports* (Figure 2), by the end of April, workplace and retail & recreation mobility gradually start increasing. By July 1st, 2020, almost all social measures had been lifted. During the first lockdown, schools remained fully closed until May 18th, 2020 and were only re-opened to a very limited extent before the end of the school year on July 1st, 2020. For this reason, schools are assumed to remain closed during the first COVID-19 wave. During July, there were little social restrictions, and this resulted in new, localized infection clust. During most of August 2020, a lockdown with curfew was imposed in Belgium's Antwerp province. We do not attempt to model the hospitalizations during July and August 2020, as modelling localized infection clusters with a nation-level epidemiological model can only be accomplished by sever ad-hoc tweaks in the social contact model. A spatial model extension was developed to account for such localized phenomena.

During the second lockdown from October 19th 2020 until present day (26/02/2021), workplace mobility has been reduced by approximately 25 %. During Autumn break and Christmas holidays, workplace mobility further declined to approximately 45 %. Public transport mobility decreased by 30 % and by 50 % during holidays, leisure mobility decreased by 40-50 % and grocery & pharmacy mobility have decreased by approximtely 5-10 %. Primary and secundary schools were closed between October 19th, 2020 and re-opened on November 16th, 2020. Further, schools have been closed during Christmas holidays from December 18th, 2020 until January 4th, 2021 and were closed during spring break from February 15th, 2021 until February 21th, 2021. Universities have remained fully closed since October 19th, 2020.

A 30 % and 15 % increase in residential mobility were observed during the lockdowns. However, the figure indicates people spent on average 15 % -30 % more time at home every day, which does not necessarily mean having 15 % - 30 % more contacts at home. Further, the coefficients of the interaction matrices are also influenced by *network* effects and may thus not scale linearly with the decrease in the observed mobility. During a strict lockdown where social *bubbles* are the norm, an intra-household contact does not have the same weight as a random, extra-household contact as long as the virus is outside of the social *bubble*. However, amplifying the fraction of household contacts under lockdown measures will increase intergenerational mixing of the population under lockdown, which is unrealistic and will lead to overestimations of the hospitalizations. We thus used a coefficient of 1.0 for the home interaction matrix $\boldsymbol{N}_{\text{c,home}}$.

Table 4: Dates of key events during the first and second lockdown in Belgium. Google mobility reduction (see Figure 2), computed as the average reduction between one key event and the next.

Date	Key event	Details	G_{work}	G_{transit}	$G_{\text{r\& r}}$	$G_{\text{g \& p}}$	H_{schools}
First COVID-19 wave (March - July 2020)							
15/03/2020	Lockdown	Closure of schools, bars, clubs and restaurants; Closure of all non-essential shops; Non-essential travel forbidden. [9]	- 56 %	-65 %	-72 %	-26 %	- 100 %
04/05/2020	Lockdown release phase Ia	Re-opening of industry and B2B services. Re-opening of non-essential retail. Merging of two *social bubbles* allowed [29].	- 44 %	-54 %	-57 %	-18 %	- 100 %
11/05/2020	Lockdown release phase Ib	Re-opening of all businesses and shops. Working at home remains the norm where possible.	- 38 %	-45 %	-46 %	-12 %	- 100 %
18/05/2020	Lockdown release phase IIa	Re-opening of businesses that involve the most human-human contact (f.i. hairdressers). Re-opening of schools for graduating classes in elementary and secondary education [38].	- 38 %	-39 %	-39 %	-8 %	- 100 %
04/06/2020	Lockdown release phase III	Re-opening of bars and restaurants. Gatherings up to 10 persons are allowed.	-22 %	-27 %	-15 %	-4 %	- 100 %
01/07/2020	Lockdown release phase IV	Closure of schools for summer holidays. Gatherings of up to 15 persons are allowed.	-32 %	-27 %	-11 %	-8 %	- 100 %
01/08/2020	Antwerp Lockdown	The number of infections starts increasing in Antwerp province, where a second lockdown with curfew is imposed [7].	-28 %	-33 %	-32 %	-6 %	- 100 %
Second COVID-19 wave (September 2020 - present)							
01/09/2020	End of summer holidays	Opening of elementary and secondary schools.	-18 %	-17 %	-14 %	-5 %	- 0 %
19/10/2020	Lockdown	Closure of bars and restaurants; Curfew; Strict social restrictions. [1]	-26 %	-31 %	-39 %	-3 %	- 0 %
02/11/2020	Lockdown	Closure of non-essential stores; Closure of all schools. [2]	-43 %	-48 %	-55 %	-13 %	- 100 %
16/11/2020	Schools reopen	Elementary and secondary schools reopen. Universities remain closed.	-27 %	-37 %	-44 %	-5 %	- 0 %
12/18/2020 - 04/01/2021	Christmas holidays	Elementary and secondary schools close. Decrease in work related mobility.	-45 %	-47 %	-42 %	-4 %	- 100 %
04/01/2021 - 15/02/2021	Period between holidays	Elementary and secondary schools reopen. British variant (501Y.V1) starts spreading [9]. Vaccination campaign in elderly homes starts [33].	-27 %	-38 %	-43 %	-6 %	- 0 %

7.4 Basic reproduction number

Since the system of differential equations (eq. 1- eq. 13), is autonomous, the eigenvalues of the Jacobian matrix evaluated at its hyperbolic equilibrium point can be used to determine the nature of that equilibrium [16]. The basic reproduction number (R_0) is computed as the spectral radius of the Jacobian matrix at the disease-free equilibrium [10]. Our model has eight infected states: $E, I, A, M, ER, C, C_{\text{ICU,rec}}$ and ICU (Figure 1). At the disease-free equillibrium, the whole population is susceptible to the infectious disease, $S_i = T_i$,

$$\boldsymbol{u^*} = (T_i, 0, 0, 0, 0, 0, 0, 0, 0, 0). \tag{26}$$

The Jacobian $\boldsymbol{J}$ is defined as,

$$\boldsymbol{J} = \begin{bmatrix} \left.\frac{\partial f_1}{\partial x_1}\right|_{\boldsymbol{u^*}} & \cdots & \left.\frac{\partial f_1}{\partial x_n}\right|_{\boldsymbol{u^*}} \\ \vdots & \ddots & \vdots \\ \left.\frac{\partial f_m}{\partial x_1}\right|_{\boldsymbol{u^*}} & \cdots & \left.\frac{\partial f_m}{\partial x_n}\right|_{\boldsymbol{u^*}} \end{bmatrix}, \tag{27}$$

where n and m are equal to the number of infected compartments. Next, the Jacobian is decomposed in the following form,

$$\boldsymbol{J}^* = (\boldsymbol{T} + \boldsymbol{\Sigma})\boldsymbol{J}\ . \tag{28}$$

The matrix $\boldsymbol{T}$ contains all terms that lead to *transmissions* of SARS-CoV-2, while $\boldsymbol{\Sigma}$ contains all terms that lead to *transitions*. For our model,

$$\boldsymbol{T} = \begin{bmatrix} 0 & \beta \sum_{j=1}^{N} N_{c,ij} & \beta \sum_{j=1}^{N} N_{c,ij} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \tag{29}$$

where an entry $T_{i,j}$ is the rate at which individuals in infected state j gives rise to individuals in infected state i. And,

$$
\boldsymbol{\Sigma} = \begin{bmatrix}
-1/\sigma & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1/\sigma & -1/\omega & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & a_i/\omega & -1/d_a & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & (1-a_i)/\omega & 0 & -\left(\frac{1-h_i}{d_m}+\frac{h_i}{d_{\text{hosp}}}\right) & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \frac{h_i}{d_{\text{hosp}}} & -1/d_{\text{ER}} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{c_i}{d_{\text{ER}}} & -\left(\frac{m_{C,i}}{d_{c,D}}+\frac{1-m_{C,i}}{d_{c,R}}\right) & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{1-c_i}{d_{\text{ER}}} & 0 & -\left(\frac{m_{\text{ICU},i}}{d_{\text{ICU},D}}+\frac{1-m_{\text{ICU},i}}{d_{\text{ICU},R}}\right) & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \frac{1-m_{\text{ICU},i}}{d_{\text{ICU},R}} & -\frac{1}{d_{\text{ICU,rec}}} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}, \tag{30}
$$

where an element $\mathbf{\Sigma}^{-1}_{i,j}$ is the expected time that an individual who presently has state j will spend in state i during its entire epidemiological *life*. The next generation matrix (NGM) is then calculated as,

$$\text{NGM} = -\boldsymbol{T}\boldsymbol{\Sigma}^{-1} \, . \tag{31}$$

The basic reproduction number R_0 is defined as the spectral radius[3] $\boldsymbol{\rho}$ of this matrix [10],

$$R_0 = \boldsymbol{\rho}(-\boldsymbol{T}\boldsymbol{\Sigma}^{-1}) \, , \tag{32}$$

which becomes for our model,

$$R_{0,i} = (a_i d_a + \omega)\beta \sum_{j=1}^{N} N_{c,ij} \, . \tag{33}$$

A linear relationship between the reproduction number and the chance of infection upon contact (β), the number of contacts ($\boldsymbol{N_c}$) and the sum of the durations of infectiousness for those compartments able to infect susceptibles makes sense.

7.5 Time-lagged cross correlation

We extracted the number of laboratory confirmed cases in youths $[0, 20[$, the working population $[20, 60[$ and the senior population $[60, \infty[$ from the *Belgian Scientific Institute of Public Health* (`https://epistat.sciensano.be/Data`) from November 2nd, 2020 to February 1st 2020. We then normalized the timeseries with the number of cases on November 21st, 2020 and visualized the result in Figure 3. Using the Python module *pandas*, the dataseries were shifted with k days and the cross correlation was computed. The procedure was performed for $k \in [-15, 5]$ days, the resulting *cross correlation function* is shown in Figure 8 and the results of the analysis are summarized in Table 5. Next, we constructed a statiscal test to check if the covariance between two series x and y, shifted with the number of days resulting in the maximum covariance, k_max, varied significantly from zero. Thus, the null hypothesis is,

$$H_0 : \rho_{xy}(k_{\text{max}}) = 0.0 \, . \tag{34}$$

If the cross correlation of lag k_{max} is zero, then, for a fairly large timeseries consisting of n datapoints, the covariance $\rho_{xy}(k_{\text{max}})$ will be approximately normally distributed, with mean zero and standard deviation $\sigma = \frac{1}{\sqrt{n-|k|}}$. Since approximately 95% of a normal population is within 2 standard deviations of the mean, a test will reject the hypothesis that the cross correlation of lag k equals zero when,

$$|\rho(k)| \geq \frac{2}{\sqrt{n-|k|}} \, . \tag{35}$$

The null hypothesis was rejected for all timeseries.

[3] Largest absolute eigenvalue.

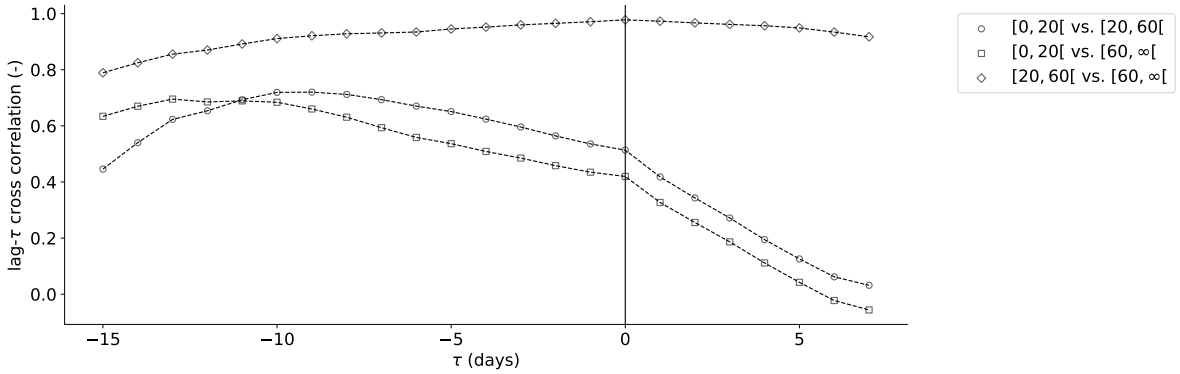

Figure 8: Cross correlation between the number of cases in Belgium in the age groups $[0-20[$, $[20-60[$ and $[60-\infty[$, from November 2nd, 2020 until February 1st 2020 in function of the number of days the timeseries are shifted relative to each other (τ). The maximum cross correlation is obtained when the series $[0-20[$ and $[20-60[$ are shifted -9 days, the maximum cross correlation is obtained when the series $[0-20[$ and $[60-\infty[$ are shifted -13 days, and the maximum cross correlation is obtained when the series $[20-60[$ and $[60-\infty[$ are not shifted.

Table 5: Results of the time-lagged cross-correlation between the number of cases in the age groups $[0-20[$, $[20-60[$ and $[60-\infty[$. Data from November 2nd, 2020 until February 1st 2020 were used in the analysis, which is equal to the daterange range shown in Figure 3.

Age group (years)	**Time-lag (days)**	**Covariance (-)**
$[0-20[$ vs. $[20-60[$	-9	0.72
$[0-20[$ vs. $[60-\infty[$	-13	0.70
$[20-60[$ vs. $[60-\infty[$	0	0.98

7.6 Supplementary data and figures

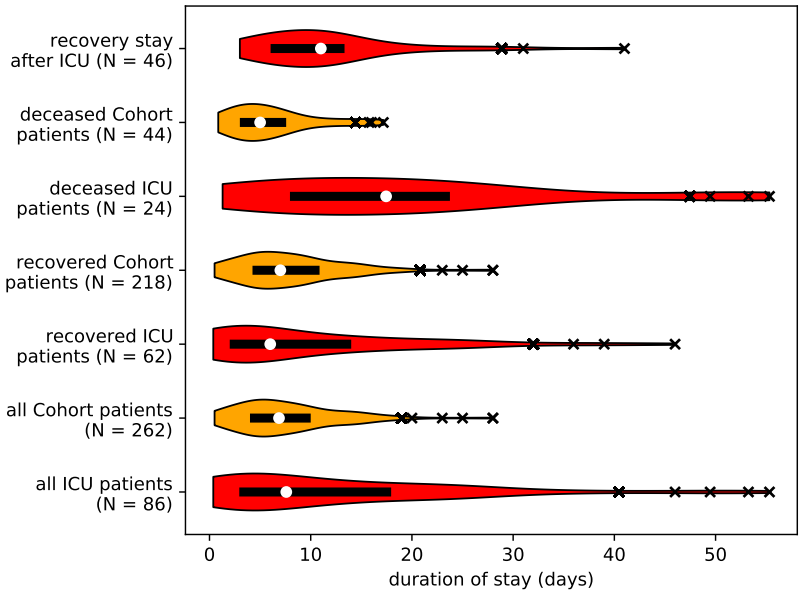

Figure 9: Observations of the length of a hospital stay for patients in cohort and ICU wards. Overall, if recovered, if deceased and for a recovery stay in cohort after ICU. Interquartile range, median observation and suspected outliers (more than 1.5 times above the third quartile) indicated.

References

[1] Abbeloos, J.-F. (2020a). Alle munitie uit de kast om dijkbreuk te vermijden. *De Standaard*.

[2] Abbeloos, J.-F. (2020b). 'dit zijn de maatregelen van de laatste kans'. *De Standaard*.

[3] Abrams, S., Wambua, J., Santermans, E., Willem, L., Kuylen, E., Coletti, P., Libin, P., Faes, C., Petrof, O., Herzog, S. A., Beutels, P., and Hens, N. (2021). Modeling the early phase of the belgian covid-19 epidemic using a stochastic compartmental model and studying its implied future trajectories. *medRxiv*.

[4] Barbe, K., Blotwijk, S., and Cools, W. (2020). Data-driven epidemiological model to monitor the sustainability of hospital care. Technical Report ICDS300420, Vrije Universiteit Brussel.

[5] CDC COVID-19 Response Team (2020). Severe Outcomes Among Patients with Coronavirus Disease 2019 (COVID-19) - United States, February 12-March 16, 2020. *Morbidity and Mortality Weekly Report*, 69:343–346.

[6] Coletti, P., Wambua, J., Gimma, A., Willem, L., Vercruysse, S., Vanhoutte, B., Jarvis, C. I., Van Zandvoort, K., Edmunds, J., Beutels, P., and Hens, N. (2020). CoMix: comparing mixing patterns in the Belgian population during and after lockdown. *Scientific Reports*, 10(1):21885.

[7] Cools, S. (2020). Antwerpen voert avondklok in. *De Standaard*.

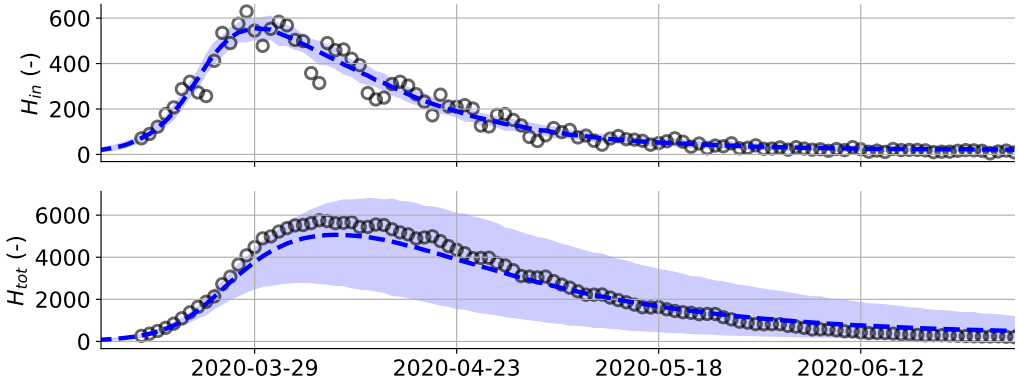

Figure 10: Model predictions for the daily Belgian hospitalizations (top) and the total number of patients in Belgium hospitals (bottom) during the first COVID-19 wave in Belgium, from March 15th, 2020 until July 1st, 2020. Mean and 95 % confidence interval of 1000 model realisations. The model is calibrated to the daily Belgian hospitalizations (H_{in}), the prediction for the total number of patients (H_{tot}) is obtained after propagating the age-stratified mortalities, age-stratified distributions between cohort and ICU and the residence time distributions derived from the hospital dataset in the model. Propagated uncertainty is large but the mean model prediction is accurate. This is most likely due to the broad residence time distributions (see Figure 9).

[8] Davies, N. G., Klepac, P., Liu, Y., Prem, K., Jit, M., Pearson, C. A. B., Quilty, B. J., Kucharski, A. J., Gibbs, H., Clifford, S., Gimma, A., van Zandvoort, K., Munday, J. D., Diamond, C., Edmunds, W. J., Houben, R. M. G. J., Hellewell, J., Russell, T. W., Abbott, S., Funk, S., Bosse, N. I., Sun, Y. F., Flasche, S., Rosello, A., Jarvis, C. I., Eggo, R. M., and working Group, C. C.-. (2020). Age-dependent effects in the transmission and control of COVID-19 epidemics. *Nature Medicine*.

[9] De Smet, D. (2021). De britse variant is niet meer te stuiten, de britse golf wel. *De Standaard*.

[10] Diekmann, O., Heesterbeek, J., and J.A.J., M. (1990). On the definition and the computation of the basic reproduction ratio R0 in models for infectious diseases in heterogeneous populations. *Journal of mathematical biology*, 28(4):365–382.

[11] Diekmann, O., Heesterbeek, J., and M.G., R. (2010). The construction of next-generation matrices for compartmental epidemic models. *Journal of the Royal Society, Interface*, 7(47):873–885.

[12] Franco, N. (2020). Covid-19 belgium: Extended seir-qd model with nursing homes and long-term scenarios-based forecasts. *medRxiv*.

[13] Goodman, J. and Weare, J. (2010). Ensemble samplers with affine invariance. *Communications in Applied Mathematics and Computational Science*, 5(1):65–80.

[14] Governement, B. F. (2020). De trends van de afgelopen dagen zetten zich voort.

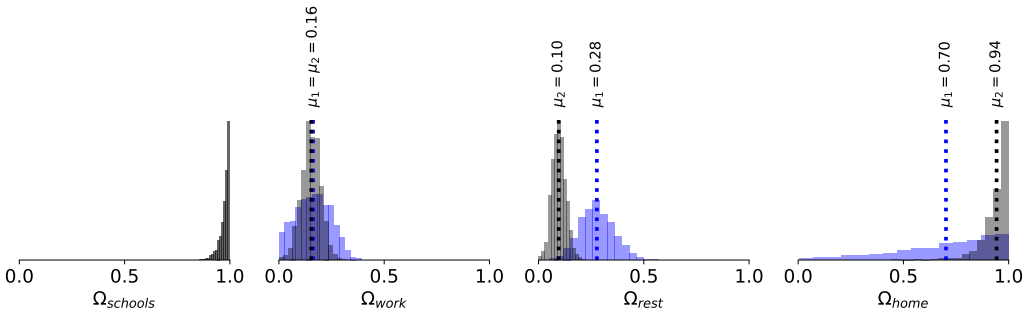

Figure 11: Inferred effectivity parameters at home (Ω_{home}), in the workplace (Ω_{work}), in schools (Ω_{schools}) and for the sum of leisure activities, other activities and public transport (Ω_{rest}), for the first COVID-19 wave (blue) and for the second COVID-19 wave (black). The effectivity of contacts in schools could not be deduced during the first COVID-19 wave because schools remained practically closed until July 1st, 2020. However, a high effectivity of contacts in schools could be deduced during the second COVID-19 wave. The effectivity of work contacts was roughly the same during both 2020 COVID-19 waves. The effectivity of leisure contacts was estimated to be lower during the second COVID-19 wave, however, leisure policies were not varied (yet) during the second COVID-19 wave, so the estimate must be taken with a grain of salt. Home contacts were deemed more effective by the model during the second COVID-19 wave.

[15] Gudbjartsson, D. F., Helgason, A., Jonsson, H., Magnusson, O. T., Melsted, P., Norddahl, G. L., Saemundsdottir, J., Sigurdsson, A., Sulem, P., Agustsdottir, A. B., Eiriksdottir, B., Fridriksdottir, R., Gardarsdottir, E. E., Georgsson, G., Gretarsdottir, O. S., Gudmundsson, K. R., Gunnarsdottir, T. R., Gylfason, A., Holm, H., Jensson, B. O., Jonasdottir, A., Jonsson, F., Josefsdottir, K. S., Kristjansson, T., Magnusdottir, D. N., le Roux, L., Sigmundsdottir, G., Sveinbjornsson, G., Sveinsdottir, K. E., Sveinsdottir, M., Thorarensen, E. A., Thorbjornsson, B., Löve, A., Masson, G., Jonsdottir, I., Möller, A. D., Gudnason, T., Kristinsson, K. G., Thorsteinsdottir, U., and Stefansson, K. (2020). Spread of sars-cov-2 in the icelandic population. *New England Journal of Medicine*, 0(0):null.

[16] Hartman, P. (1960). A lemma in the theory of structural stability of differential equations. volume 11, pages 610–620.

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

[18] Kennedy, J. and Eberhart, R. (1995). Particle swarm optimization. In *Proceedings of ICNN'95 - International Conference on Neural Networks*, volume 4, pages 1942–1948 vol.4.

[19] Kermack, W. O., McKendrick, A. G., and Walker, G. T. (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 115(772):700–721.

[20] Lescure, F.-X., Bouadma, L., Nguyen, D., Parisey, M., Wicky, P.-H., Behillil, S., Gaymard, A., Bouscambert-Duchamp, M., Donati, F., Hingrat, Q. L., Enouf, V., Houhou-Fidouh, N., Valette, M., Mailles, A., Lucet, J.-C., Mentre, F., Duval, X., Descamps, D., Malvy, D., Timsit, J.-F., Lina, B., Van-der Werf, S., and Yazdanpanah, Y. (2020). Clinical and virological data of the first cases of COVID-19 in Europe: a case series. *The Lancet Infectious Diseases*, 0(0).

[21] Li, Q., Guan, X., Wu, P., Wang, X., Zhou, L., Tong, Y., Ren, R., Leung, K. S., Lau, E. H., Wong, J. Y., Xing, X., Xiang, N., Wu, Y., Li, C., Chen, Q., Li, D., Liu, T., Zhao, J., Liu, M., Tu, W., Chen, C., Jin, L., Yang, R., Wang, Q., Zhou, S., Wang, R., Liu, H., Luo, Y., Liu, Y., Shao, G., Li, H., Tao, Z., Yang, Y., Deng, Z., Liu, B., Ma, Z., Zhang, Y., Shi, G., Lam, T. T., Wu, J. T., Gao, G. F., Cowling, B. J., Yang, B., Leung, G. M., and Feng, Z. (2020a). Early transmission dynamics in wuhan, china, of novel coronavirus–infected pneumonia. *New England Journal of Medicine*, 0(0):null.

[22] Li, R., Pei, S., Chen, B., Song, Y., Zhang, T., Yang, W., and Shaman, J. (2020b). Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (sars-cov2). *Science*.

[23] Linton, N. M., Kobayashi, T., Yang, Y., Hayashi, K., Akhmetzhanov, A. R., Jung, S.-M., Yuan, B., Kinoshita, R., and Nishiura, H. (2020). Incubation Period and Other Epidemiological Characteristics of 2019 Novel Coronavirus Infections with Right Truncation: A Statistical Analysis of Publicly Available Case Data. *Journal of clinical medicine*, 9(2).

[24] Liu, T., Wu, S., Tao, H., Zeng, G., Zhou, F., Guo, F., and Wang, X. (2020a). Prevalence of igg antibodies to sars-cov-2 in wuhan - implications for the ability to produce long-lasting protective antibodies against sars-cov-2. *medRxiv*.

[25] Liu, Y., null, n., Funk, S., and Flasche, S. (2020b). The contribution of pre-symptomatic infection to the transmission dynamics of covid-2019 [version 1; peer review: 1 approved]. *Wellcome Open Research*, 5(58).

[26] Liu, Y., Yan, L.-M., Wan, L., Xiang, T.-X., Le, A., Liu, J.-M., Peiris, M., Poon, L. L. M., and Zhang, W. (2020c). Viral dynamics in mild and severe cases of COVID-19. *The Lancet Infectious Diseases*, 0(0).

[27] Molenberghs, G., Faes, C., Aerts, J., Theeten, H., Devleesschauwer, B., Bustos Sierra, N., Braeye, T., Renard, F., Herzog, S., Lusyne, P., Van der Heyden, J., Van Oyen, H., Van Damme, P., and Hens, N. (2020). Belgian covid-19 mortality, excess deaths, number of deaths per million, and infection fatality rates (8 march - 9 may 2020). *medRxiv*.

[28] Mossong, J., Hens, N., Jit, M., Beutels, P., Auranen, K., Mikolajczyk, R., Massari, M., Salmaso, S., Tomba, G. S., Wallinga, J., Heijne, J., Sadkowska-Todys, M., Rosinska, M., and Edmunds, W. J. (2008). Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLOS Medicine*, 5(3):1–1.

[29] Open VLD (2020). Coronavirus : België heeft z'n exitstrategie vastgelegd.

[30] Park, M., Cook, A. R., Lim, J. T., Sun, Y., and Dickens, B. L. (2020). A Systematic Review of COVID-19 Epidemiology Based on Current Evidence. *Journal of clinical medicine*, 9(4).

[31] Riou, J. and Althaus, C. L. (2020). Pattern of early human-to-human transmission of wuhan 2019 novel coronavirus (2019-ncov), december 2019 to january 2020. *Eurosurveillance*, 25(4).

[32] Shaman, J. and Galanti, M. (2020). Will sars-cov-2 become endemic? *Science*, 370(6516):527–529.

[33] Snoekx, K. (2021). Vaccinaties in rusthuizen komen in stroomversnelling. *De Standaard*.

[34] StatBEL (2020). Structure of the Population.

[35] To, K. K.-W., Tsang, O. T.-Y., Leung, W.-S., Tam, A. R., Wu, T.-C., Lung, D. C., Yip, C. C.-Y., Cai, J.-P., Chan, J. M.-C., Chik, T. S.-H., Lau, D. P.-L., Choi, C. Y.-C., Chen, L.-L., Chan, W.-M., Chan, K.-H., Ip, J. D., Ng, A. C.-K., Poon, R. W.-S., Luo, C.-T., Cheng, V. C.-C., Chan, J. F.-W., Hung, I. F.-N., Chen, Z., Chen, H., and Yuen, K.-Y. (2020). Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by SARS-CoV-2: an observational cohort study. *The Lancet Infectious Diseases*, 20(5):565–574.

[36] Verity, R., Okell, L. C., Dorigatti, I., Winskill, P., Whittaker, C., Imai, N., Cuomo-Dannenburg, G., Thompson, H., Walker, P. G. T., Fu, H., Dighe, A., Griffin, J. T., Baguelin, M., Bhatia, S., Boonyasiri, A., Cori, A., Cucunubá, Z., FitzJohn, R., Gaythorpe, K., Green, W., Hamlet, A., Hinsley, W., Laydon, D., Nedjati-Gilani, G., Riley, S., van Elsland, S., Volz, E., Wang, H., Wang, Y., Xi, X., Donnelly, C. A., Ghani, A. C., and Ferguson, N. M. (2020). Estimates of the severity of coronavirus disease 2019: a model-based analysis. *The Lancet Infectious Diseases*.

[37] Viner, R. M., Mytton, O. T., Bonell, C., Melendez-Torres, G., Ward, J., Hudson, L., Waddington, C., Thomas, J., Russell, S., van der Klis, F., Koirala, A., Ladhani, S., Panovska-Griffiths, J., Davies, N. G., Booy, R., and Eggo, R. M. (2020). Susceptibility to sars-cov-2 infection amongst children and adolescents compared with adults: a systematic review and meta-analysis. *medRxiv*.

[38] Vlaamse regering (2020). Heropstart van de lessen op school: wie, waarom en hoe.

[39] Wei, W. E., Li, Z., Chiew, C. J., Yong, S. E., Toh, M. P., and Lee, V. J. (2020). Presymptomatic Transmission of SARS-CoV-2 — Singapore, January 23–March 16, 2020. *Morbidity and Mortality Weekly Report*, 69:411–415.

[40] Whatty, S. (2020). Tijdelijke werkloosheid bijna gehalveerd. *Het Laatste Nieuws*.

[41] Willem, L. (2021). Restore.

[42] Willem, L., Abrams, S., Petrof, O., Coletti, P., Kuylen, E., Libin, P., Mogelmose, S., Wambua, J., Herzog, S. A., Faes, C., Beutels, P., and Hens, N. (2020). The impact of contact tracing and household bubbles on deconfinement strategies for covid-19: an individual-based modelling study.

[43] Willem, L., Van Kerckhove, K., Chao, D. L., Hens, N., and Beutels, P. (2012). A nice day for an infection? weather conditions and social contact patterns relevant to influenza transmission. *PLOS ONE*, 7(11):1–7.

[44] Wu, J. T., Leung, K., Bushman, M., Kishore, N., Niehus, R., de Salazar, P. M., Cowling, B. J., Lipsitch, M., and Leung, G. M. (2020a). Estimating clinical severity of COVID-19 from the transmission dynamics in Wuhan, China. *Nature Medicine*, 26(4):506–510.

[45] Wu, J. T., Leung, K., and Leung, G. M. (2020b). Nowcasting and forecasting the potential domestic and international spread of the 2019-ncov outbreak originating in wuhan, china: a modelling study. *The Lancet*, 395(10225):689 – 697.

[46] Wu, Z. and McGoogan, J. M. (2020). Characteristics of and important lessons from the coronavirus disease 2019 (covid-19) outbreak in china: Summary of a report of 72314 cases from the chinese center for disease control and prevention. *JAMA*.

[47] Zou, L., Ruan, F., Huang, M., Liang, L., Huang, H., Hong, Z., Yu, J., Kang, M., Song, Y., Xia, J., Guo, Q., Song, T., He, J., Yen, H.-L., Peiris, M., and Wu, J. (2020). Sars-cov-2 viral load in upper respiratory specimens of infected patients. *New England Journal of Medicine*, 382(12):1177–1179. PMID: 32074444.

ETHICS QUESTIONNAIRE

848 Name of the promotor:

INGMAR NOPENS, JAN BAETENS

Title of the proposal/project:

THE USE OF A COMPARTMENTAL, AGE-STRATIFIED, EXTENDED SEIRQD MODEL AND PUBLIC MOBILITY DATA TO ASSESS THE EFFECTS OF NON-PHARMACEUTICAL INTERVENTIONS ON SARS-CoV-2 SPREAD IN BELGIUM.

A. INTRODUCTION

Research should be conducted according to ethical standards. Some are imposed by law, others are generally accepted in (international) scientific practice.

Researchers must reflect on the ethical aspects of their proposal and identify any ethical issues in the ethics questionnaire (**part B**). The actual questionnaire is followed by the declaration on honour (**part C**).

Sometimes there is a legal obligation to submit the project proposal to an ethics committee. This is made clear in the questionnaire. However, researchers may also submit their proposal to an ethics committee for an ethical approval if there is no legal obligation, e.g., when a journal requests an ethical approval. For further information, check our website in Dutch or English. A dedicated research tip informs you on what is requested when specific ethics issues apply, and provides further background information.

Even if no ethical issues apply to your proposal, you must always complete **part C** of the questionnaire (**Declaration on honour**).

2. Humans

	YES?
1. Does your research involve human participants?	☐
a. Are they volunteers for social or human sciences research?	☐
b. Are they persons unable to give informed consent (including children / minors)?	☐
c. Are they vulnerable individuals or groups?	☐
d. Are they children / minors?	☐
e. Are they patients?	☐
f. Are they healthy volunteers for medical studies?	☐
2. Does your research involve physical interventions on the study participants?	☐
a. Does it involve invasive techniques?	☐
b. Does it involve collection of biological samples?	☐

If you checked box 1.a, 1.b, 1.c or 1.d, please note that not every research involving human participants triggers the obligation to request an ethical approval. However, the journal in which you want to publish the results of your research might ask you to submit an ethical approval. For this reason, it might be advisable to request ethical approval anyway before the start of the project.

If you checked box 1.e, 1.f or 2, you must submit your proposal to the Committee for Medical Ethics, as soon as your application has been approved for funding. The project can only start when the committee has formally given an ethical approval of the project.

4. Personal data

	YES?
Does your research involve collecting and/or processing of personal data?	☐

If you checked the box, the EU General Data Protection Regulation (GDPR) requires that all personal data processing activities are registered in the GDPR register of each institution where the processing takes place before the start of the processing. At UGent, this registration should be done in dmponline.be. Check our website for more information on how to register your personal data processing activities.

6. Access and Benefit-Sharing and the Nagoya Protocol

851

	YES?
Does your research involve genetic (biological) resources and/or traditional knowledge associated with genetic resources, that are in scope of the Nagoya Protocol and/or the related EU Regulation 511/2014? → Name of the country/ies of origin:	☐

Usually, you must obtain a 'Prior Informed Consent' (PIC) from the Competent National Authority in the country of origin (provider country) prior to the access and utilization of the genetic resources or traditional knowledge. The conditions for utilization, and benefit sharing, must be negotiated and registered in 'Mutually Agreed Terms' (MAT). Check our website in Dutch or English for more information.

8. Dual-use and military applications

852

	YES?
1. Does your research have the potential for military applications?	☐
2. Does your research involve dual-use items in the sense of EU Regulation 428/2009, or other items for which an export license is required?	☐

If you checked any of the boxes, be aware that research proposals with potential military applications or involving dual-use items must comply with Ghent University's dual-use research policy and must be reported to Ghent University's Dual-Use Contact Point. You might need an export license if you wish to export dual-use items. Check our website in Dutch or English for more information.

Not every dual-use research triggers the obligation to request an ethical approval. However, you must inform the Dual Use Contact Point of your dual-use research using the Dual-Use Notification Form. The journal in which you want to publish the results of your research might ask you to submit an ethical approval. For this reason, it might be advisable to request an ethical approval from the Committee on Human Rights Policy and Dual-Use Research anyway before the start of the project.

10. Environment and health & safety

853

	YES?
1. Does your research involve the use of elements (chemical, physical, sound, …) that may cause harm to the environment (water, air, soil, …), or to animals or plants?	☐
2. Does your research involve the use of elements (chemical, physical, sound, …) that may cause harm to humans, including research staff and their co-workers?	☐
3. Is (part of) your research carried out within protected areas?	☐
4. Do the proposed experiments make use of GMOs or pathogens?	☐
5. Do the proposed experiments make use of activities, installations or products that need to be covered by permits (narcotic drugs and precursors, hormonal substances, explosives and precursors, cyanides, ozone-depleting substances, ionizing radiation, radioactive substances, soils/animals/animal parts and by-products/plants from third countries, …)?	☐

If you checked the box at questions 4 and/or 5, please ensure that you comply with the applicable legislations and guidelines regarding the environment, health and safety.

For question 4, an attestation concerning biosafety is required. See our Dutch webpage and milieu@ugent.be.

For question 5, different types of permits or attestations or a compulsory notification may be required:

- Narcotic drugs and precursors, hormonal substances, explosive compounds, ozone-depleting substances, soils/animals/animal parts and by-products/plants from third countries: see our Dutch webpage and milieu@ugent.be.
- Cyanides and prohibited substances: see our Dutch webpage and veiligheid@ugent.be.
- Ionizing radiation and radioactive substances: see our Dutch webpage and straling@ugent.be.

C. DECLARATION ON HONOUR

854 Please confirm which of the following situations apply to your research (multiple answers are possible). It is required to at least answer either question 1 or 2.

1. I declare on honour that **none of the ethical issues** in part B of the questionnaire **apply to my proposal**.	☒
2. I declare on honour that **one or more of the ethical issues** in part B of the questionnaire do **apply to my proposal** and I commit to adhere to all relevant legislation, regulations and institutional guidelines and policies.	☐
3. I declare that for one or more ethical issues in part B of the questionnaire **an ethical approval is required**. I declare on honour that **I will submit my project proposal** in due time to the competent research **ethics committee(s)**. Please indicate which ethics committee(s) will deal with your application: ☐ Medical Ethics Committee ☐ Federal Commission for Embryos ☐ Ethics Committee for Animal Testing of the Faculty of Veterinary Sciences, also competent for the Faculty of Bioscience Engineering ☐ Ethics Committee for Animal Testing of the Faculty of Medicine and Health Sciences, also competent for the Faculty of Pharmaceutical Sciences ☐ Ethics Committee for Animal Testing of the Faculty of Sciences and the VIB ☐ Committee on Human Rights Policy and Dual-use Research ☐ Ethics Committee of the Faculty of Arts and Philosophy ☐ Ethics Committee of the Faculty of Engineering and Architecture ☐ Ethics Committee of the Faculty of Law and Criminology ☐ Ethics Committee of the Faculty of Economics and Business Administration ☐ Ethics Committee of the Faculty of Psychology and Educational Sciences ☐ Ethics Committee of the Faculty of Political and Social Sciences	☐