

Social Mixing Matrix Literature Review

April 23, 2020

Standard epidemic models, such as SIR, SIS and their extensions SIRS, MSEIRS, SEIR are well established mathematical methods that simulate the spread of an infectious disease [4, 8, 7]. One limitation of these models is that they base their prediction on the a-priori assumption that the rate of contact in a population is constant across environments and age groups. As a result, these methods neglect the vital component of how populations interact. Therefore, researchers who solely rely on the SIR/SIS family of models often formulate contact rates (σ) and reproduction numbers (R_0) based on make non-data driven hypotheses. As a reminder, R_0 is the "The threshold for many epidemiology models is the basic reproduction number R_0 , which is defined as the average number of secondary infections produced when one infected individual is introduced into a host population where everyone is susceptible" [8, p.21]. In general, $R_0 > \sigma$ but in many models $R_0 = \sigma$.

Social Mixing Matrices (SMMs) are considered important tools to establish an as much accurate as possible R_0 and σ and therefore in building effective defence against respiratory pathogens. They capture social distancing patterns, based on direct observations and surveys, and as a result they provide crucial information to both scientists and public officials on how a directly transmitted infectious disease is spreading. Because of the granularity of the data that SMMs offer their use has the potential to increase model robustness and prediction accuracy compared to models that rely solely on indirect data [9, 1].

On their study, [2] explored the construction of the called next generation matrices (NGMs). NGMs are a vital part of the SMMs construction and utilisation process. Many papers that have used SMMs, one way or the other, they had to transform their initial population interaction matrix to be NGM compliant in order to adequately compute the integral R_0 and σ [9, 3, 5]. In

other words, NGMs show the number of novel infections from each category (age groups, gender groups, etc) in consecutive generations [2, p.874]. From the NGM, we can map the R_0 to the dominant eigenvalue of the matrix. The ODE models that are used in epidemiology are linearised and form a Jacobian matrix [4, Section 2.1.2], then [2] showed how to decompose this matrix into a transition component Σ (change in state) and a transmission part T (new infections).

To illustrate the work of [2], we will use the epidemiological model from [6]. The model will be adapted to COVID-19 cases and to include social mixing matrix data instead of the sexual contact matrix that the authors used. The compartmental transmission model has a form as shown in Figure 1 [6, p.5]. From the Figure 1, the studied population is divided into the following non-overlapping compartments: S is for individuals who are at risk of Covid-19 infection, I is for infected individuals, G is for infected who developed serious symptoms and required ICU admission, P is those who have recovered, and are seropositive and immune, N is for individuals who are recovered, immune but seronegative. The indices g , a and s indicate that every compartment is stratified by gender, age and level of social activity. Movements between compartments are occurring at per capita rates specified by the following parameters: λ is the **force of infection** dependent on the proportion of individuals in I , **PSC** is the probability of becoming seropositive, **WIP** is the Covid-19 incubation period, **DAI** is the duration of asymptomatic (i.e. without Covid-19) infection symptoms, **DWT** is the duration of treatment for Covid-19, and **DI** is the duration of immunity. Subscripts denote stratification of parameters: for example, **DWT** _{g} means that in our model this parameter is gender-dependent.

The system of ODEs is then formulated as follow

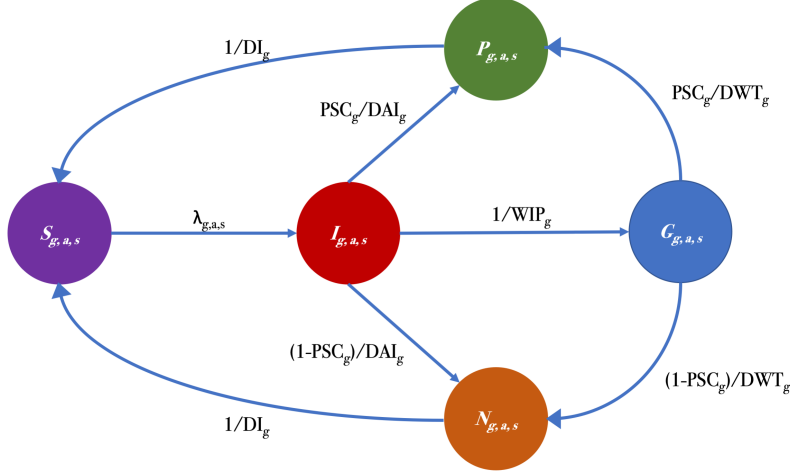


Figure 1: A compartmental Covid-19 transmission model.

$$\begin{aligned}
\dot{S}_{g,s,a} &= -\lambda_{g,s,a}(t)S_{g,s,a} + (P_{g,s,a} + N_{g,s,a})/DI_g + \frac{1}{r}S_{g,s,a-1} - \frac{1}{r}S_{g,s,a} + \\
&\quad \frac{1}{R} \sum_{g,s} (S_{g,s,20} + I_{g,s,20} + G_{g,s,20} + P_{g,s,20} + N_{g,s,20}) \times \\
&\quad \delta_1(a)(\pi_1\delta_1(s) + \pi_2\delta_2(s) + \pi_3\delta_3(s) + \pi_4\delta_4(s)) \\
\dot{I}_{g,s,a} &= \lambda_{g,s,a}(t)S_{g,s,a} - (1/WIP_g + 1/DAI_g)I_{g,s,a} + \frac{1}{r}I_{g,s,a-1} - \frac{1}{r}I_{g,s,a} \\
\dot{G}_{g,s,a} &= I_{g,s,a}/WIP_g - G_{g,s,a}/DWT_g + \frac{1}{r}G_{g,s,a-1} - \frac{1}{r}G_{g,s,a} \\
\dot{P}_{g,s,a} &= PSC_g(I_{g,s,a}/DAI_g + G_{g,s,a}/DWT_g) - P_{g,s,a}/DI_g + \frac{1}{r}P_{g,s,a-1} - \frac{1}{r}P_{g,s,a} \\
\dot{N}_{g,s,a} &= (1 - PSC_g)(I_{g,s,a}/DAI_g + G_{g,s,a}/DWT_g) - N_{g,s,a}/DI_g + \frac{1}{r}N_{g,s,a-1} - \frac{1}{r}N_{g,s,a}
\end{aligned}$$

The stochastic compartmental system has in total 5 states of which one is infected state I and three uninfected states S , P and N . The total population remains constant by adding an ageing term in the susceptible compartment,

$$\frac{1}{R} \sum(\dots) \times \delta\pi.$$

A landmark study SMM based study was POLYMOD [9] that includes $n = 7,290$ participants from 8 European countries between years 2005-2006. This is the first study of such inter-state scale and since then it has been widely used amongst publications in the field. POLYMOD, contains information for 97,904 contacts grouped by age, gender, type of contact (physical/conversational), place and frequency. The study expanded the pool of potential information that researchers for epidemiological diseases could access.

One limitation with the POLYMOD study is the population bias that exhibits on children and young adult ages. The reason is that the focus of the study has been deliberately on these age groups as the infectious diseases that the authors studied were most likely to be transmitted through young ages. For example, 38% of the participants are under 20 years old of age and 12% were over 60 years old. At the time of survey, the European average for these 8 countries for 20 y and under was 17% and the same average for 60 y and over was 21%. As a reference, the EU average for 20 y and under was 16% and 23% respectively. There is also a small change over the years, in 2018 (the same year as the BBC survey in UK) the population in the 8 countries for 20 y and under was 15% and for 60y and over was 25%. The inclusion of developing Eastern European countries has not managed to reverse the ageing population that developed countries suffer from. The subsequent rapid development of the new European Union countries has affected them with the same problem of an ageing population ¹.

In their results, the authors showed some interesting outcomes. First, using countries as dummy variables do seem to make a difference in the social mixing matrix. For instance, German participants showed per daily average $\mu = 8$ fewer number of contacts compared to their Italian counterparts $\mu = 20$.

Work and transport involve the fewer average number of physical contact; it might look counter-intuitive for the latter but transport could potentially entail

¹Information sourced from Eurostat data [demo_pjanbroad]

References

- [1] M. Baguelin, S. Flasche, A. Camacho, N. Demiris, E. Miller, and W. J. Edmunds. Assessing optimal target populations for influenza vaccination programmes: an evidence synthesis and modelling study. *PLoS Med*, 10(10):e1001527, 2013.
- [2] O. Diekmann, J. A. P. Heesterbeek, and M. G. Roberts. The construction of next-generation matrices for compartmental epidemic models. *J R Soc Interface*, 7(47):873–885, 2010.
- [3] L. Fumanelli, M. Ajelli, P. Manfredi, A. Vespignani, and S. Merler. Inferring the structure of social contacts from demographic data in the analysis of infectious diseases spread. *PLOS Computational Biology*, 8(9):1–10, 09 2012.
- [4] M. Keeling, P. Rohani, and P. U. Press. *Modeling Infectious Diseases in Humans and Animals*. Princeton University Press, 2008.
- [5] P. Klepac, A. J. Kucharski, A. J. Conlan, S. Kissler, M. Tang, H. Fry, and J. R. Gog. Contacts in context: large-scale setting-specific social mixing matrices from the bbc pandemic project. *medRxiv*, 2020.
- [6] I. A. Korostil, G. W. Peters, J. Cornebise, and D. G. Regan. Adaptive markov chain monte carlo forward projection for statistical analysis in epidemic modelling of human papillomavirus. *Statistics in Medicine*, 32(11):1917–1953, 2013.
- [7] M. Li. *An Introduction to Mathematical Modeling of Infectious Diseases*. Mathematics of Planet Earth. Springer International Publishing, 2018.
- [8] S. Ma and Y. Xia. *Mathematical Understanding of Infectious Disease Dynamics*. Lecture Notes Series, Institute for Mathematical Sciences. World Scientific, 2009.
- [9] J. Mossong, N. Hens, M. Jit, P. Beutels, K. Auranen, R. Mikolajczyk, M. Massari, S. Salmaso, G. S. Tomba, J. Wallinga, J. Heijne, M. Sadkowska-Todys, M. Rosinska, and W. J. Edmunds. Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLOS Medicine*, 5(3):1–1, 2008.