REAL-TIME NOWCASTING AND FORECASTING OF COVID-19 IN THE UK: THE FIRST WAVE?

During the first wave of the COVID-19 pandemic in the UK, the recurrent message has been that policy was being based on scientific evidence. But how did the interaction between policy-makers and science work and what type of contribution did scientists make to the decision process?
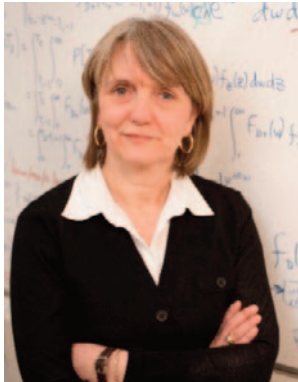
Professor Daniela De Angelis
Professor of Statistical Science for Health
Deputy Director and Programme Leader
MRC Biostatistics Unit
University of Cambridge

As of the 28th of June, over 300,000 individuals have been confirmed to have acquired infection with the SARS-CoV-2 virus. Amongst these over 40,000 have died. Over the last months we have learnt that the UK COVID-19 pandemic has three contexts: community, hospitals and care homes. A lock-down was introduced on the 23rd of March and we are gradually easing it, desperately trying to go back to normal life. Behind this knowledge and its input to policy, lies the work of many individuals:

epidemiologists, mathematicians, statisticians and behavioural scientists who became members of relevant advisory groups. Figure 1 explains the structure of this interaction involving the Civil Contingencies Committee (COBR), the Scientific Advisory Group for Emergencies (SAGE), the New and Emerging Respiratory Virus Threats Advisory Group (NERVTAG), and the Scientific Pandemic Influenza Groups on Modelling (SPI-M-O) and Behaviours (SPI-B).

SPI-M-O is the committee providing quantitative evidence to SAGE, particularly on the dynamics of SARS-CoV-2 transmission. Constituted during the H1N1 Pandemic in 2009, it has been maintained over the years to ensure preparedness in the event of an influenza pandemic and recently extended to deal with other emerging diseases. Members come from a number of research institutions who freely accepted the invitation to contribute to the work. Regular commissions are received from Cabinet Office →
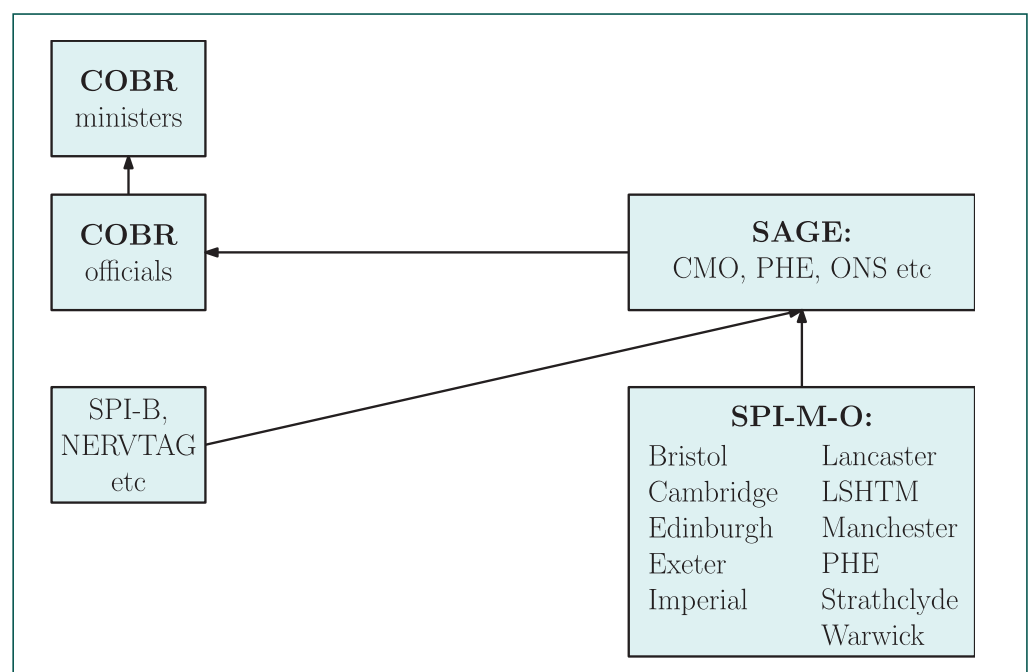

Figure 1

and answers are provided swiftly (24/48 hours!) from the various groups. Results are discussed at a SPI-M-O meeting and a consensus is reached. This consensus is communicated to SAGE and further discussed. The evidence provided may or may be considered in the final deliberations.

The questions posed throughout the pandemic have varied. During the containment phase, when the plan was to use tracing of contacts of symptomatic confirmed infected individuals, the need was to understand the potential of the pandemic. Then the questions asked were: how many people will get infected? When is the peak occurring? What is the likely duration? What age groups/geographical locations will be worst affected? In the mitigation phase, when the need was to mitigate the impact on the NHS by 'flattening the curve' there was a need to understand what non-pharmaceutical interventions would be most effective (e.g. Closing schools? Isolating symptomatic individuals? Banning big gatherings?). Once the lockdown was introduced, in the suppression phase, the questions were: how is the pandemic progressing after the lock-down? When would it be safe to re-instate social networks?

All these questions can be addressed through the use of models of SARS-CoV-2 transmission. These are mathematical constructs built to approximate the unobserved process of epidemic spread, i.e. the process of interaction of infected individuals into a (totally) susceptible population, spreading infection through
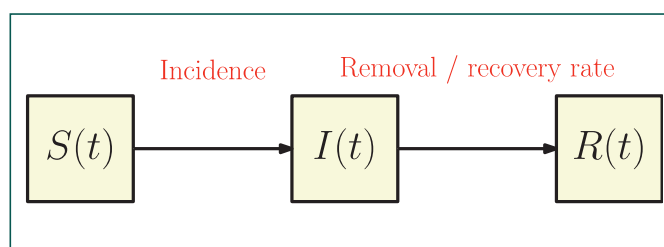

Figure 2

contact with these susceptible individuals. The infection spread will depend on how infectious the infected individuals are and the level of contacts they have with those susceptible. Figure 2 provides the simplest example of such models, referred to as the Susceptible, Infected, Recovered (SIR) model. As time progresses, more individuals become infected if no intervention is introduced, until the pool of susceptible individuals is depleted. People might then die or recover and no longer take part in the generation of infections, unless, after recovery, their immunity wanes and they become susceptible again. The quantity of interest here is the movement from the susceptible compartment to the infectious compartment, i.e. the number of new infections. These movements will depend on unknown quantities (parameters) reflecting the infectiousness of the virus and the likelihood of an infection given a contact. These unknown quantities involve the basic reproduction number, R_0, the average number of infections generated by a typical infected individual throughout their infectious period in a totally susceptible population. A value of R_0 higher than 1 indicates ongoing transmission.

The models adopted by the groups contributing to SPI-M-O have a much more complex structure than the simple SIR

model in Figure 2. They typically include a higher number of compartments and further stratify the population by age groups, geography and contexts (e.g. workplace, schools etc..). The interaction between susceptible and infected individuals in the different population strata, expressed as the average numbers of daily contacts, drives disease transmission and generates new infections until there are no longer susceptible individuals, or an intervention that reduces this interaction is introduced. In this more realistic context, the aim is to monitor the pandemic evolution over time by estimating the level of transmission, the number of new infections and predict future burden, in different age groups and regions.

In the last four months, the Medical Research Council Biostatistics Unit (MRC-BSU) at Cambridge University, has contributed, in collaboration with Public Health England (PHE), to SPI-M-O by providing regular updates on the state of the pandemic in England from a transmission model. The population is subdivided into four categories: susceptible, infected but not infectious, infected and infectious, recovered. We further stratify by age groups (< 1, 1–4, 5–14, 15–24, 25–44, 45–64, 65–74, 75+ years old) and by the seven National Health Service regions. The transmission between

groups is informed by the POLYMOD study (Mossong et al, 2008), giving the mean number of contacts between the different age groups and contexts, updated throughout the pandemic to account for changes in behaviour (E. van Leeuwen, F. Sandmann, 2020); literature from the Chinese pandemic provides information on the natural history of SARS-CoV-2 (Li et al, 2020; Verity et al, 2020) ; and sequential serological surveys from testing blood donors over time and in different regions provide information on the proportion of the population that has already been infected (https://www.gov.uk/governmen t/publications/national-covid-19-surveillance-reports/sero-surveill ance-of-covid-19). All these data sources are then combined with data on daily numbers of age and region-specific deaths in individuals with laboratory confirmed SARS-CoV-2 to reconstruct the underlying number of infections and characterise transmission nationally and in the different regions. Death data, although a lagged signal of infection, provide information on the shape of the epidemic curve; and data from serological studies, giving information on the proportion of the population in different age group with antibodies, inform the magnitude of the pandemic. The typical output, produced regularly (https://www.mrc-bsu.cam.ac.uk/now-casting/), included: the reconstructed number of infections over time; trends in R_t, the *effective* reproduction number (now referring to a population not totally susceptible), an indicator of ongoing transmission; and
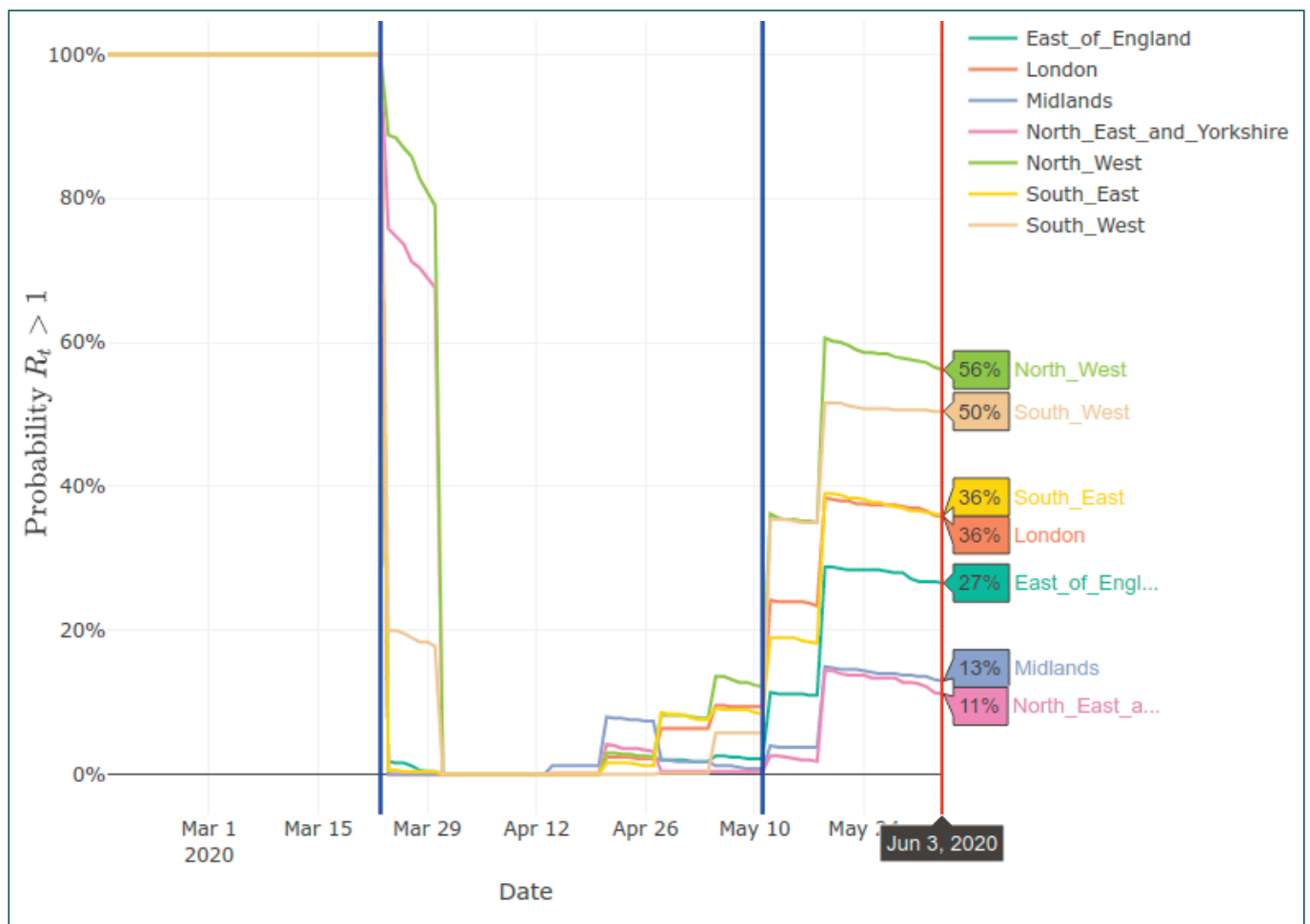
Figure 3

short-term predictions for the number of deaths. Figure 3 gives an example of these outputs, portraying the probability of R_t being above the value of 1 in the different regions at the beginning of June.

It this type of quantification, provided by the different groups, that gave evidence to policy makers to monitor levels of transmission and inform decisions.

Currently, transmission appears to be slowing down in all regions of England and the number of daily infections is decreasing over time. The estimated proportions of already infected individuals in the various age groups and regions are, however, low, alerting to the possibility of a second wave of infection. So, continued

monitoring remains essential. The approach taken until now will need to be complemented by more granular surveillance tools aimed at identifying and managing local outbreaks.

These interesting times offer plenty of exciting professional experience to a statistician studying disease transmission like me!

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, W. J. Edmunds, Social contacts and mixing patterns relevant to the spread of infectious diseases. PLOS Medicine. 5, e74 (2008).

E. van Leeuwen, F. Sandmann, Augmenting contact matrices with time-use data for fine-grained intervention modelling of disease dynamics: a modelling analysis (2020), doi:10.1101/2020.06.03.20067793.

R. Verity, L. C. Okell, I. Dorigatti, P. Winskill, C. Whittaker, N. Imai, G. Cuomo-Dannenburg, H. Thompson, P. G. T. Walker, H. Fu, A. Dighe, J. T. Griffin, M. Baguelin, S. Bhatia, A. Boonyasiri, A. Cori, Z. Cucunubá, R. FitzJohn, K. Gaythorpe, W. Green, A. Hamlet, W. Hinsley, D. Laydon, G. Nedjati-Gilani, S. Riley, S. van Elsland, E. Volz, H. Wang, Y. Wang, X. Xi, C. A. Donnelly, A. C. Ghani, N. M. Ferguson, Estimates of the severity of coronavirus disease 2019: a model-based analysis. The Lancet Infectious Diseases. 20, 669–677 (2020). □