University of Oxford

Department of Statistics

A thesis submitted for the degree of Doctor of Philosophy

Epidemiological modelling and analysis of COVID-19

Author:

Ruth McCabe

Supervised by:

Professor Christl Donnelly, University of Oxford

Dr. Mike Beadsworth, Liverpool School of Tropical Medicine

Dr. Jake Dunning, University of Oxford

Dr. Maria Van Kerkhove, World Health Organization

Dr. Roberto Vivancos, UK Health Security Agency

Dr. Oliver J. Watson, Imperial College London

Michaelmas Term, 2023

Acknowledgements

I am very grateful to the many people who have encouraged and supported me over the past three years. If I have missed anyone here, please take this as my struggling end-of-thesis brain rather than me not appreciating the time that we have spent together.

To Christl, you are truly the most incredible supervisor and mentor. I still can't believe that you offered me this opportunity, but I will always be grateful that you did. Every success that I have had during my DPhil is down to your guidance and support. Thank you for always empowering me, for giving me pep talks when needed (often before I had even realised that they were needed - sometimes I feel that you know me better than I do!), and for encouraging me to embrace my (borderline-neurotic) need to be organised as an asset and not an annoyance. I hope that the last three years are just the beginning of a long and fruitful working relationship, because it is a true privilege to work with and learn from you.

To OJ, the biggest thank you in the world for letting me adopt you as a supervisor - I cannot imagine my DPhil without your guidance, wisdom and support. It has been a pleasure and privilege to learn from you. I know for a fact that, in particular, my technical and data viz skills have improved substantially directly because of us working together. You have given me so much of your time despite never being obligated to and I will be eternally grateful for that. I am also super grateful for you letting me pat Rhubarb whenever possible.

To Maria and to Sophie, thank you so much for letting me join your team(s) at WHO, something that was a dream of mine for a long time (and something that I distinctly recall telling Christl that I wanted to do on the very first day of my DPhil). You have given me opportunities I could only dream of being given at this stage of my career which have already shaped and, I am confident, will continue to shape my future career. I very much look forward to continuing to work together on MERS and other global health threats. Maria, an additional thank you for being a shining example of effective leadership in a crisis - you inspire so many people (including myself) every day and we are so grateful to have you there to guide us through difficult times.

Thank you to Jasmina Panovska-Griffiths, Josie Park and Gabriel Danelian for welcoming me into the Epi-Ensemble team at UKHSA and for the many interesting and thought-provoking discussions we had during my time there.

Thank you (again) to Christl for inviting me to work with the Statistics Expert Group (SEG) in producing their evidence for the Infected Blood Inquiry. Yet again, this an opportunity I never imagined I could have at this stage of my career and was such an incredible learning experience for me. In particular, it was a genuine honour to work with and learn from two of my all-time favourite scientists (who are also wonderful people) – Professors Sir David Spiegelhalter and Sheila Bird.

Without the personal support and mentorship I received from Professors Thomas Churcher and Katharina Hauck, I know that I would not have received this DPhil opportunity. You are exceptional PIs and I am extremely grateful for your supervision and the opportunities you both gave me.

Additionally, thank you to my Higher maths teacher Mrs Hale. Her enthusiasm for maths was truly second to none - she is one of the main reasons why I pursued a maths degree and I am so grateful to her for her entirely infectious passion for the subject.

Thank you to the HPRU in Emerging and Zoonotic Infections for funding my DPhil and my placement at WHO, and a huge thank you to the incredible, heroic support staff in the Department of Statistics who have not only provided me with administrative support but also friendship and laughter over the past three years.

Thank you to all of Team Donnelly for welcoming me with open arms, for celebrating my successes and for allowing me to join in celebrating yours. In particular I must thank Sarah, Matt, Nic, Cathal and Tarek for being amazing friends over the years. You are all incredible scientists - I hugely admire each of you and I am excited to see the brilliant things that you will all achieve in the course of your careers. Thank you for emotionally counselling me through every IT meltdown (far too frequent an occurrence) and most importantly, thank you for the many, many laughs in our chaotic Friday group meetings.

I am so proud to have been a member of the Department of Statistics at Oxford. I am completely in awe of my peers who are among the smartest people I have ever met but also, importantly for me, such lovely and fun people to be around. Thank you in particular to Vik, Chris, Anna, Sahra, Amitis, Kamelia, Tiggy, Xi for always having my back and for finding laughter in even the greyest of days. One of my fondest memories from my time here will be the copious amounts of scheming and tea-drinking as part of the Tea Drinking Trio with the amazing Mariagrazia and Maria. Thank you also to everyone from the department that I've had the pleasure of enjoying lunch or a pint with, and for collectively never allowing a social occasion, no matter how big or small, pass without multiple stats jokes.

A huge thank you to Alison, Mike, Jackie, Robbo, Mish, Heather, Leah, Qassim, Ettie, Tori, Alexandra, Alessandra, Nora, Ashleigh, Collins, Shreya, Rabikah, Sumali, Joe, Rebecca, Pablo and Richard for your friendship, laughter and emotional support. I am so fortunate to have such amazing friends without whose love and support I would not have made it to the end of this journey.

I am so lucky that I have a family who I would want to be friends with even if we weren't related. Thank

you to the whole 'Warwickshire BBQ' team for being the best humans around - I am always the happiest person when I have a weekend with you lined up. Thank you to Liz and Tom for being my surrogate parents for the last 10 years. I really hope you know how much I appreciate the huge amount much you have done for me since I moved down South. Thank you to Maz, Al, Claire and Chash for being the big sisters I never had. But the biggest thank you of all must go to Joe, Raine, Luke, Micah and Archie for the most amazing (and chaotic) afternoons in Barford. You guys have shown me I can do whatever I put my mind to (for example, without you all I wouldn't know that I can fit in every corner of the kid's soft play at Hatton as an adult).

The biggest anti-thank you (/f*** you) must go to cancer for inflicting far too much physical and emotional suffering on too many of my loved ones, especially in the past three years. You have book-ended my DPhil in the most unwelcome of ways.

I don't quite know how to properly thank my Mum, Dad and Michael for their support, not only for this thesis, but throughout my whole life. You guys mean the entire world to me. Michael, thank you for being a rock the last few months and thank you also for always keeping me grounded in reality (by calling me a neeeeeerd every so often among other things). Mum and Dad, you have given me everything you have your entire life and I wouldn't have achieved anything without your unrelenting support. Thank you for instilling your incredible work ethic in me, for encouraging me that all I ever have to do to make you proud is "my best", and for never limiting, in fact, for always encouraging, my ambitions and hare-brained schemes (this degree being one of the latter). I will keep working hard in everything I do to continue making you proud.

But last and by no means least, the work in this thesis is dedicated to my Auntie Evelyn - the kindest, most selfless and bravest person that I know. The world would be a much better place with more people like you in it.

Contents

List of Figures

List of Tables

List of Abbreviations

Abbreviation	Meaning
AFRO	African Regional Office
anti-RBD	total antibody to SARS-CoV-2 receptor binding domain
BR	binding ratio
CFR	case fatality ratio
CI	confidence interval
COVAX	Access to COVID-19 tools Accelerator's Vaccines Pillar
COVID-19	coronavirus disease 2019
DABA	double antigen binding assay
DSTL	Defence Science and Technology Laboratory
ELISA	enzyme-linked immunosorbent assay
EMRO	Eastern Mediterranean Regional Office
FMOH	Federal Ministry of Health
GBD	global burden of disease
GS	Google Scholar
HIC	high-income country
HIV	human immunodeficiency virus
IBI	Infected Blood Inquiry
ICU	intensive care unit
IBM	individual-based model
IFR	infection fatality ratio
IgG	immunoglobulin G
IgM	immunoglobulin M
IMD	index of multiple deprivation
JCVI	Joint Committee on Vaccination and Immunisation
LFD	lateral flow device
LIC	low-income country
m	million
MCMC	Markov Chain Monte Carlo
MERS	Middle East respiratory syndrome
MERS-CoV	Middle East respiratory syndrome coronavirus
NP	nucleotide protein
NPI	non-pharmaceutical intervention
NW	North-West
OD	optical density
ONS	Office for National Statistics
ONS CIS	Office for National Statistics COVID-19 Infection Survey
p	p-value
PCR	polymerase chain reaction
PHEIC	public health emergency of international concern
PHSM	public health and social measure
PIS	participant information sheet
R_0	basic reproduction number
$R(t)$	effective reproduction number
$r(t)$	instantaneous growth rate

RDT	rapid diagnostic test
REACT	REal-time Assessment of Community Transmission
RECOVERY	Randomised Evaluation of COVID-19 Therapy
S1	spike protein S1 subunit
SAGE	Scientific Advisory Group for Emergencies
SARS	severe acute respiratory syndrome
SARS-CoV-1	severe acute respiratory syndrome coronavirus
SARS-CoV-2	severe acute respiratory syndrome coronavirus 2
sd	standard deviation
SEIR	susceptible-exposed-infected-recovered
SIR	susceptible-infected-recovered
SPI-M	Scientific Pandemic Influenza Group on Modelling
SSE	sum of squared errors
UK	United Kingdom of Great Britain and Northern Ireland
UKAS	United Kingdom Accreditation Service
UKHSA	United Kingdom Health Security Agency
UNOCHA	United Nations Office for the Coordination of Humanitarian Affairs
VOC	variant of concern
VOI	variant of interest
WCIS	winter COVID-19 study
WHO	World Health Organization
WoS	Web of Science

Chapter 1

Introduction

1.1 Background

The transmission of an infectious disease through a susceptible population can cause substantial and long-lasting disruption to the normal functioning of society. This has been exemplified clearly by the novel virus severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) which, upon its discovery in late 2019, left no corner of the globe unaffected in some way. The highly transmissible SARS-CoV-2 virus, which causes coronavirus disease 2019 (COVID-19), risked the lives of millions of people [1, 2] and threatened the stability of healthcare systems around the globe [3, 4]. In turn, this influenced governments to take unprecedented actions to control the spread of the virus, in the form of stringent public health and social measures (PHSMs) (including, but not limited to, "lockdown" measures, physical distancing and isolation of infectious cases).

In 2022, the World Health Organization (WHO) estimated that the COVID-19 pandemic was associated with almost 15 million (m) excess deaths globally in 2020 and 2021 alone [5]. While this figure in itself represents enormous suffering, it does not capture the toll of the loss of livelihood from the implementation of PHSMs [6]; the exhaustion of healthcare workers having worked through several waves of SARS-CoV-2 while also providing treatment for other conditions [7]; potentially millions of people left with post-COVID-19 syndrome ("long COVID") [8]; and the substantial increase in mental health disorders worldwide [9], to name only a few such problems. Such examples demonstrate that the ramifications of a pandemic range far beyond the boundaries of public health alone.

While COVID-19 is one of the most recent, unfortunately it is not the only example of an outbreak of an infectious disease to cause wide-ranging and long-lasting devastation [10–16]. Leaders in the field

of public health had long warned about the high likelihood of a pandemic caused by the emergence of a novel or influenza-like virus [17–20]. In 2018, the WHO updated its list of priority pathogens most likely to cause a "public health emergency", which includes Middle East respiratory syndrome (MERS) and severe acute respiratory syndrome (SARS), to also include an unknown "Disease X" [21]. The discovery of the (previously) novel SARS-CoV-2 is indeed an example of an emergence of a "Disease X", although this pathogen now holds its own place on this list [22].

Undoubtedly, the response to the COVID-19 pandemic has resulted in some outstanding scientific advances, including: the development, testing and roll-out of a vaccine within 12 months of the discovery of SARS-CoV-2 (the Oxford-AstraZeneca vaccine) [23, 24]; the implementation of a nationwide randomised clinical trial for COVID-19 treatments within 3 months of its discovery (the Randomised Evaluation of COVID-19 Therapy (RECOVERY) trial) [25]; global scaling up of infectious disease surveillance capabilities, particularly genomic sequencing [26]; to name only a few. However, many of these advances were possible in such a short time frame due to the years of research and preparatory work undertaken in response to other pathogens. For example, the basis of the vaccines against SARS-CoV-2 came from the almost decade of work undertaken by experts across the world to develop a vaccine for MERS coronavirus (MERS-CoV) [24, 27]. This was captured by Professor Sir Peter Horby, who has noted that SARS-CoV-2 is from a well-understood family of viruses (coronaviruses) for which decades of research had already been completed. This meant that the global community already had a strong basis upon which to respond. What concerns Professor Horby greatly is the emergence of a new virus from a less-well-understood family, for which our foundations upon which to respond will be less strong, or a virus (like HIV [human immunodeficiency virus]) which is less amenable to vaccine development [28].

Although the COVID-19 pandemic has arguably been the most severe and far-ranging public health event of the last four years, it is certainly not the only outbreak to have occurred in this period. Since 2020, the world has dealt with: a global mpox (formerly monkeypox) outbreak [29]; children presenting with novel hepatitis [30, 31]; an outbreak of Ebola virus disease in Uganda [32]; outbreaks of Marburg virus disease in Equatorial Guinea [33] and Tanzania [34]; cholera outbreaks reported in 30 countries in 2022 [35]; global reports of avian influenza (H5N1) [36]; a zoonotic spillover of Influenza A (H1N2) in the United Kingdom of Northern Ireland and Great Britain (UK) [37]; a sharp increase in respiratory illnesses in children in China [38]; and many more [39]. This demonstrates that the world remains highly susceptible to the spread of an infectious disease, whether this is a known pathogen or an unknown "Disease X". As noted by Dr. Tedros Ghebreyesus, Director General of the WHO, "the next pandemic is a case of *when*, not *if*" [40].

Many scientists have a critical role to play in preparing for and responding to a pandemic, from the development of therapeutics and diagnostics to building robust national health and surveillance systems. Statisticians and mathematical modellers are no exception [41–43]. Models and analyses make substantial contributions to decision-making of public health officials during outbreaks by shedding light on questions such as: How many people infected with a virus will die? [44, 45] Might unmitigated spread of a virus overwhelm healthcare systems? [1, 46, 47] However, the transmission of infections through a population is intrinsically complex and not directly observable. Statisticians and modellers have the responsibility to harness all available information and use this, for example, to understand the mechanisms driving the transmission of a virus, to estimate key epidemiological quantities, and often to bring findings together to project what could happen in the future under different conditions.

Four years after the emergence of SARS-CoV-2, the emergency phase of the COVID-19 pandemic is over [48]. However, the virus remains in circulation across the world, continuing to evolve and to pose threats to global public health [49]. However, our collective knowledge of how to best respond to COVID-19 and outbreaks of infectious disease more generally also evolves, as we build upon the strengths of previous responses and learn from the elements that were less successful. As demonstrated particularly well by MERS-CoV, the work undertaken to respond to a particular virus is not only essential in the context of that pathogen alone [27]. All of the work in this thesis has been written, and should be viewed, with this philosophy in mind: not only do the efforts that we put into understanding the transmission dynamics of SARS-CoV-2 help us to respond to the current threats posed by the ongoing circulation of this pathogen, but they will help inform responses to future outbreaks.

1.2 DPhil objectives

This DPhil has two main objectives, each of which are addressed by three pieces of work (Table 1.1). Firstly, the use of modelling in informing public health policy is investigated. The work pertaining to this objective includes primary data collected from key stakeholders, namely scientific advisors and the public, reflecting upon their experiences of modelling during the COVID-19 pandemic, and discusses the communication of uncertainty to non-technical audiences. The reliability of and uncertainty surrounding model outputs are highly dependent on the data used to fit them. Often, these data are officially-reported cases and deaths, but as discussed below, these only capture a subset of the true disease burden. Consequently, the second thesis objective is to investigate the suitability of alternative data sets to track the transmission dynamics of SARS-CoV-2.

Table 1.1: An overview of the six papers which make up the main body of this thesis.

Objective	Paper	Chapter	Title	Status	Reference	Appendix
1. To investigate the role of modelling in informing public health policy	I	3	Disease transmission and control modelling at the science-policy interface	Published in *Journal of the Royal Society Interface Focus*	[50]	A
	II	4	Communicating uncertainty in epidemic models	Published in *Epidemics*	[51]	B
	III	5	Public awareness and opinions on the use of mathematical transmission modelling to inform public health policy in the United Kingdom	Published in *Journal of the Royal Society Interface*	[52]	C
2. To investigate the suitability of alternative sources to track transmission dynamics	IV	6	Tracking the incidence and risk factors for SARS-CoV-2 infection using historical maternal booking samples	Published in *PLoS One*	[53]	D
	V	7	Alternative epidemic indicators for COVID-19 in settings with incomplete death registration systems	Published in *Science Advances*	[54]	E
	VI	8	Inferring community transmission of SARS-CoV-2 in the United Kingdom using the ONS COVID-19 Infection Survey	Accepted for publication in *Infectious Disease Modelling*	[55]	F

1.2.1 Objective 1: To investigate the role of modelling in informing public health policy

Mathematical modelling is an important component of scientific evidence presented to decision-makers during an outbreak of an infectious disease [43]. The implementation of restrictive PHSMs has drawn increased attention towards the mechanisms by which scientific advice contributes to policy decisions. In particular the influence of modelling regarding the COVID-19 response has been discussed by multiple witnesses during their testimonies to the UK COVID-19 Inquiry in late 2023 [56]. Much of the research conducted in the field of infectious disease epidemiology can have implications for policy, with this often being an explicit goal of researchers within this field. In order for science to inform policy there must be effective communication mechanisms between key stakeholders, but these associations are rarely considered formally. Papers I - III address these issues, as described briefly below.

1.2.1.1 Paper I: Disease transmission and control modelling at the science-policy interface

Paper I [50] explores the complex relationships between models, decision-making, the media and the public during the COVID-19 pandemic in the UK. Doing so not only provides important historical context of COVID-19 modelling and how it has shaped the UK response but, looking towards future pandemic preparedness, understanding these relationships and how they might be improved is critical. As such, we have synthesised information gathered via three methods: a survey to publicly listed attendees of relevant

government advisory committees, interviews with science communication experts and former scientific advisors, and reviewed some of the key COVID-19 modelling literature from 2020.

1.2.1.2 Paper II: Communicating uncertainty in epidemic models

Paper II [51] addresses the issue of the communication of uncertainty arising in epidemic models to non-technical audiences, as noted by respondents to our research in Paper I. The paper includes a discussion of potential sources of uncertainty, and provides an alternative visualisation of the modelling conducted as part of [57], with the aim of aiding the communication of this to non-technical audiences.

1.2.1.3 Paper III: Public awareness and opinions on the use of mathematical transmission modelling to inform public health policy in the United Kingdom

Paper III [52] responds directly to another of the key findings of Paper I in which a participant in our research noted that communication with the public during the COVID-19 pandemic had been "difficult". As such, we surveyed the public on their opinions on and awareness of the use of models in informing public health policy, asking the same questions with regard to the period prior to and during the COVID-19 pandemic to elicit how opinions could have changed during this period. To overcome the difficulties of obtaining a representative sample of the public, we generated two samples of the population: a "representative" sample using the platform *Prolific Academic* and a "non-probability" sample using *Twitter*, and focus on the comparison between them.

1.2.2 Objective 2: To investigate the suitability of alternative data sources to track transmission dynamics

Officially-reported epidemiological data, such as the number of cases and deaths, are crucial for understanding the transmission and severity of a specific virus. It is known that these quantities are often subsets of the true burden of infections [45] and deaths [58], respectively, as their existence depends on infrastructure such as national testing systems and robust vital registration systems, the availability of which is not just dependent on location but also on the phase of the epidemic. For example, there is often a delay between the discovery of an emerging infectious disease and national testing programmes to detect cases among those not requiring medical attention. Incomplete and/or unrepresentative data contribute directly to the quality and robustness of model outputs. With this in mind, this thesis explores questions such as: How representative are officially-reported data of the true burden of an infectious disease? How does this influence our understanding of the transmission dynamics of SARS-CoV-2? Are there alternative sources of data that could be used to supplement or even improve this information? In the absence of COVID-19 surveillance programmes, how can we track the spread of SARS-CoV-2? Papers IV - VI

each present a different case study on alternative data sources to track the spread of SARS-CoV-2 in different locations and at different time points, and are introduced below.

1.2.2.1 Paper IV: Tracking the incidence and risk factors for SARS-CoV-2 infection using historical maternal booking samples

Paper IV [53] presents the results of a retrospective serosurvey conducted on maternal booking samples in London across September 2019 - September 2020. In this early phase of the pandemic (early 2020), widespread testing was not yet available and, as such, there is limited information available into the onset of community transmission in the UK. We use our estimates of the prevalence of seroreactivity to derive estimates of the incidence of seroreactivity, which we take as a proxy for incidence of infection. We compare the trends in our estimated incidence to those observed in officially-reported hospitalisations and deaths over the same period, when available, to discuss the viability of these routinely-collected samples as a potential data source for tracking the community transmission of a virus.

1.2.2.2 Paper V: Alternative epidemic indicators for COVID-19 in settings with incomplete death registration systems

Excess mortality has been used frequently as an alternative means by which to measure the burden of the COVID-19 pandemic in settings without robust COVID-19 surveillance infrastructure. Paper V [54] directly addresses one of the main issues cited in the WHO global world mortality estimates [5], which is that all-cause mortality data are often missing in low-income countries (LICs) due to lack of complete death registration systems, thus estimates of excess mortality are derived from pooling information from countries with similar demographics. In this paper, we undertake a more data-driven approach by using mathematical modelling and the infection fatality ratio (IFR) to consider the suitability of alternative epidemic indicators to track disease dynamics, namely burial site worker reports in Addis Ababa, satellite imagery of cemeteries in Aden and social-media-conducted surveys of infection in Khartoum. Each setting also has independently-conducted serosurveys which we compare to the seroprevalence implied by our model fit to the alternative sources, and argue that similarities of these values provide evidence of the ability of the alternative sources to track the epidemics in each setting in the absence of complete, officially-reported mortality data.

1.2.2.3 Paper VI: Inferring community transmission of SARS-CoV-2 in the United Kingdom using the ONS COVID-19 Infection Survey

Paper VI [55] addresses a challenge faced by the *Epi-Ensemble Modelling team*, responsible for producing estimates of the effective reproduction number ($R(t)$) and instantaneous growth rate ($r(t)$), at the United

Kingdom Health Security Agency (UKHSA) during the period in early 2022 when the four nations of the UK began to transition from an "emergency" to "endemic" state.[1] For approximately two years, government-published estimates of $R(t)$ and $r(t)$ were estimated from an ensemble of up to 14 models, incorporating many different data sources available throughout the UK. However, the majority of the surveillance systems in place were largely scaled down from April 2022 as part of the devolved nations (England, Scotland, Wales and Northern Ireland) governments' "living with COVID-19" policy [59–62], which made estimation of these important quantities substantially more difficult. This paper presents a method to estimate $R(t)$ and $r(t)$ directly from the percentage of individuals testing positive under the Office for National Statistics (ONS) COVID-19 Infection Survey (CIS) [63], which was one of the main surveillance streams to continue after the cessation of community testing. We validate our estimates through comparison with the government-published estimates of these quantities [64, 65], estimating the proportion of variation in the government-based estimates that can be explained by the ONS-based estimates. This method is presented as a possible solution to the estimation of $R(t)$ and $r(t)$ in the absence of more traditional epidemiological data, such as case incidence, but is not intended as an alternative to ensemble modelling, which is advantageous in real-time epidemic modelling.

1.3 Thesis structure

This thesis is structured as follows (Table 1.1):

- Chapter 2 presents a background of the emergence of SARS-CoV-2 and the subsequent COVID-19 pandemic, and introduces the concepts and literature relevant to the work in this thesis.
- Chapters 3 - 5 present Papers I - III, respectively, which relate to Objective 1.
- Chapters 6 - 8 present Papers IV - VI, respectively, which relate to Objective 2.
- Chapter 9 summarises and reflects upon the work of this thesis.
- Appendices A - F present the supplementary material corresponding to Papers I - VI, respectively.

Please note that as each main results chapter (Chapters 3 - 8) has a specific introduction and discussion, Chapters 2 and 9 should be considered high-level overviews.

This is an integrated thesis, with five out of the six chapters published in peer-reviewed journals and the remaining one accepted for publication in a peer-reviewed journal. For consistency, in this thesis each chapter is presented in its published/accepted format without adaptation, with the exception of small formatting changes (for example, unifying different journal's notations for p-values to p; use of UK

[1] Across January - April 2022 I worked as a mathematical modeller within this team.

English; capitalising of Figure and Table; use of double rather than single quotation marks etc.).

All information presented was correct at time of writing. However, due to the retrospective reporting and updating of cases and deaths, there may be inconsistencies in specific dates of the first recorded cases. For example, in Paper IV (Chapter 6, written in Autumn 2021), the first reported case in the UK is documented as occurring on 2 February 2020, but official figures now report this (December 2023) as occurring on 30 January 2020, which is what is reported in Chapter 2.

1.4 Work completed during DPhil outside of this thesis

Papers I - VI comprise the main work undertaken during the period of this DPhil. However, other work was undertaken during this time (October 2020 - December 2023), with corresponding publications set out in Appendix G. Additionally, three internships were undertaken in policy-related public health settings (without suspension of studies): at the UKHSA as a mathematical modeller (January - April 2022); at the Infected Blood Inquiry (IBI) as a statistician (May - September 2022); and at the WHO as a statistical epidemiologist working on MERS-CoV (May - July 2023; subsequently hired as a WHO part-time consultant from August - December 2023). In particular, the work for IBI constituted complex statistical modelling which was presented as evidence to the Inquiry in September 2022 [66, 67].

Chapter 2

Background and literature review

This chapter provides background information on the COVID-19 pandemic, key epidemiological quantities, statistical and modelling methodologies commonly deployed in outbreak analysis, and a review of literature relevant to this thesis. It should be noted that each paper (presented in Chapters 3 - 8) has a specific introduction, and so this chapter provides a high-level overview of relevant topics and the history of the COVID-19 pandemic.

2.1 The COVID-19 pandemic

The novel coronavirus severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) was first identified in Wuhan, China in December 2019. The unfortunate combination of the high transmissibility and high severity of this emerging pathogen, compounded by the absence of vaccines and therapeutics, posed a major threat to life and to the stability of healthcare systems [68–74]. By late January 2020, Chinese government officials imposed strict public health and social measures (PHSMs) in many badly affected provinces, and the re-opening of schools after the Lunar New Year holiday was delayed nationwide [75]. These measures also included the implementation of the first, and arguably among the strictest, coronavirus disease 2019 (COVID-19) "lockdown" in Wuhan on 23 January 2020, a policy response that has since been described as "unprecedented" [76–78] and captured the attention of the international media [79–84]. On 30 January, the outbreak was designated a "public health emergency of international concern (PHEIC)" by the World Health Organization (WHO) [85] under the 2005 International Health Regulations [86].

Although Wuhan was the initial epicentre of the epidemic, SARS-CoV-2 spread rapidly around the globe. The first case outside of China was reported in Thailand on 13 January 2020, followed by Japan and the Republic of Korea on 14 and 19 January 2020, respectively [87]. By the end of January, 22

countries, including some in Europe and North America, had reported at least one case of COVID-19, with this figure rising to 55 by the end of February, and then included Latin American and African countries [87]. The WHO declared the outbreak a pandemic on 11 March 2020 [88]. Within weeks, the "unprecedented" scenes observed in Wuhan tragically became less of an anomaly as governments around the world also faced the serious dual threats of the overwhelming of healthcare systems and substantial loss of life [1, 2, 89, 90].

Across 2020 – 2022, countries oscillated in and out of restrictive measures in a bid to balance control of the virus with the restriction of civil liberties. "Lockdowns" were implemented in many countries alongside other physical distancing measures, but it is important to recall the nuance in the meaning and implementation of these types of measures under different national or regional policymakers [91, 92]. For example, in January 2021, in some regions of Italy, hospitality was permitted to open for take away under "lockdown", but this was not the case as part of the Netherland's "lockdown" [92]. On the other hand, Sweden never implemented a full, legally-binding "lockdown" and relied on citizens to adapt their behaviour accordingly, a decision which has been debated widely by scientists and policymakers [93, 94]. Even when "lockdowns" were relaxed (in countries which implemented them), other PHSMs, such as isolation of cases, often remained in place to some degree. The Blavatnik COVID-19 Government Response tracker is an exceptional resource which documented the various measures in place across the world in real-time [89], possible due to the assistance of $1,500$ volunteers.

The approval of safe and effective vaccines against severe COVID-19 in late 2020 proved critical in allowing many societies to "reopen" from restrictive measures, without risking the substantial number of deaths following a wave of infections that had been observed previously [95, 96]. It has been estimated that the vaccines prevented up to 20 million (m) deaths from COVID-19 globally in the first year of the roll-out alone [97]. By mid-2021 many European countries had relaxed the most severe restrictions due to vaccine-induced immunity in the population [98] despite having a growing number of COVID-19 cases [87]. However, the inequitable distribution of vaccinations against SARS-CoV-2 globally meant that not all countries were in such a privileged position [99, 100]. In mid-2021, of all vaccines given globally, only approximately 15.9% of vaccines were administered in lower-middle-income countries, and, shockingly, only 0.25% in low-income countries (LICs) [101, 102]. In particular, less than 2% of all doses were administered in African countries [101], despite countries in this region facing growing epidemics across April - July 2021 [103, 104]. In July 2021, Dr. Mike Ryan, Executive Director of the WHO Health Emergencies Programme, summarised the severity of the global situation, stating that "if we're really lucky, we can have it [SARS-CoV-2] under control next year [2022]" [105]. Despite calls from the WHO and the

development of global initiatives such as COVAX, the Access to COVID-19 tools (ACT)-Accelerator's Vaccines Pillar [106], vaccine inequity has persisted. No LIC was expected to attain the G20 target of 70% coverage by mid-2022 [100], and only two (out of 27) LICs were expected to attain the G20 target of at least 40% vaccination coverage by the end of 2021. Watson et al. [97] estimated that were this latter target reached, one out of every five deaths in LICs would have been averted. In September 2023, the WHO reported that 8% of Africa's population over 60 years of age had received a booster dose [107] while, at time of writing (December 2023), Our World in Data reported that LICs had now received 1.65% of all administered vaccines [101].

Many countries, particularly in the Asia-Pacific region, attempted initially to eliminate SARS-CoV-2 transmission in full, often referred to as the "zero COVID" policy which necessitated legally-binding, stringent and long-lasting PHSMs. By Autumn 2021, Malaysia, Thailand, Vietnam and New Zealand pivoted from this strategy [108, 109] due to a combination of factors including anti-SARS-CoV-2 vaccinations, the emergence of the Delta variant[1], so-called "pandemic fatigue", wherein a population become tired of the restrictions placed upon daily life [110], and the need for economic recovery from such measures. In contrast, China only announced the end of its "zero COVID" policy in December 2022 following three years of among the strictest control policies observed globally [111].

Throughout 2022, countries slowly dropped testing and isolation requirements [89]. At the beginning of 2022, around 20m cases per week were being reported to WHO (globally), but this was around 3m by December of that year [87]. A large jump in reported cases in the last week of December 2022/January 2023 (to over 40m at the peak) is attributable to the substantial change in policy in China around that time. However, between March and December 2023, weekly reported cases at the global level have fallen to and remained consistently at fewer than 1m cases per week [87].

The full impact of the pandemic will take decades to fully present itself, and it is likely that we will be unable to fully quantify the losses that we have suffered. However, what we can quantify paints a bleak picture. To date (December 2023), over 772m cases and almost 7m deaths have been officially reported to WHO [87], although as previously stated it is known that this is an underestimate of the true number of COVID-19 infections and deaths worldwide. Estimates suggest that the death toll in 2020 and 2021 could be almost 3 times as great as the officially-reported figure (5.5m reported vs. 15m estimated excess) [5, 87]. Furthermore, estimates suggest the pandemic has resulted in close to US$14 trillion in economic losses, and approximately 70m more people living in extreme poverty than there were pre-pandemic [100].

[1] Please see below for discussion of SARS-CoV-2 variants.

While the pandemic has subsided from the emergency phase, SARS-CoV-2 is now endemic globally, and will require continuous monitoring and research going forwards, as all endemic pathogens do.

2.1.1 The COVID-19 pandemic in the United Kingdom

The United Kingdom of Great Britain and Northern Ireland (UK) reported its first official case of COVID-19 on the 30 January 2020, with the first official death following on 2 March 2020 [112]. However, there is evidence to suggest more widespread circulation of SARS-CoV-2 in the UK in early 2020 than implied by the official data [1, 113–116]. In addition to the devastating scenes reported from China and then northern Italy, mathematical modelling studies suggested that without sufficient countermeasures, there could be as many as 500,000 deaths by Autumn 2020 [1], or that demand for intensive care unit (ICU) beds could exceed 13 - 80 times the available capacity [46]. The stark warnings informed decisions that led to the announcement of the first "lockdown" on 23 March 2020 [117]. Some form of restrictions or testing requirements remained in place until April 2022.

Although the UK is one country, it is made up of four devolved nations (England, Scotland, Wales and Northern Ireland). In each nation, healthcare decision-making is the responsibility of the devolved governments, and so from approximately the loosening of the first "lockdown" in May 2020, different policies were implemented at different time points in each nation and, as described above with the example of Italy and the Netherlands [92], were nuanced in their exact requirements. The devolved nations each moved to more localised control of the epidemic, via the use of a tiered alert system per local authority [117–121], in a bid to minimise the number of people and length of time spent under restrictions. Nonetheless, most of the UK population were under some form of restrictive measures for the rest of 2020 and, although plans were in place to allow for a longer period of mixing over the Christmas period of that year, the emergence of the Alpha SARS-CoV-2 variant resulted in this policy being revoked just days before it was due to come into place [122–126].

The Alpha variant induced the re-implementation of strict "lockdown" measures lasting the first quarter of 2021 in each nation. This coincided with national SARS-CoV-2 vaccination programmes, which underpinned government decision-making regarding the loosening of severe restrictions planned for later in 2021 [96, 127]. The protection provided by vaccination was evidenced clearly in Spring 2021, when the emergence of the Delta variant caused a steep increase in the number of SARS-CoV-2 infections, but not in COVID-19-related deaths [112, 128].[2] By the end of Summer 2021, most of the severe restrictions were removed across the four nations [131–134] and in the first quarter of 2022, the four nations of the

[2]The same cannot be said for India, where the Delta variant was first reported, which observed a devastating wave of infections and mortality from this variant due, at least in part, to lack of access to vaccinations [129, 130].

UK formally transitioned from an "emergency" response to treating COVID-19 as "endemic" [59–62]. This involved removing testing requirements and the scaling down most of the "exceptionally good" and "world-beating" surveillance systems in place to track the spread of the virus [135]. The Office for National Statistics (ONS) COVID-19 Infection Survey (CIS) [63] was one exception to this and showed SARS-CoV-2 continuing to circulate at high levels in early 2023, until the initial "pause" of the study in March 2023 [136]. However, the new Winter COVID-19 Infection Study [137], announced by the United Kingdom Health Security Agency (UKHSA) and ONS in October 2023 to track the prevalence of SARS-CoV-2 throughout the winter period 2023/24, indicated a prevalence of around 1% in England and Scotland in late November 2023 [138].

As of 19 May 2022, the last date for which these values are provided for the UK as a whole on the government dashboard [112], the UK had reported over 22m cases, of which 1.2m were reinfections. To the end of August 2023, there were almost $230,000$ deaths with COVID-19 on the death certificate in the UK [112]. The UK government announced a public inquiry into the "response and impact of the pandemic" in the UK, with the aim to "learn lessons for the future" [139]. Although this was announced in April 2021, hearings began in Spring 2023 with witnesses expected to continue giving evidence until 2026 [140], suggesting that it will be several years until the final findings and recommendations of the inquiry are published.

2.2 The epidemiology of SARS-CoV-2

SARS-CoV-2 is primarily spread from human to human through infectious droplets or aerosols (which, for example, were exhaled by an infected individual) [141–143]. Not all infected with the virus will develop symptoms; estimates suggest that one third of infections among unvaccinated individuals are asymptomatic [144, 145], but there is substantial variation when accounting for factors such as geographic location and age [146, 147]. Transmission can occur from presymptomatic and asymptomatic individuals, which contribute to the difficulty of controlling the spread of this virus [148–154]. Clinical presentation most typically includes respiratory symptoms, developing into pneumonia in cases of serious illness, as well as fever, headaches, and sore throats (a non-exhaustive list), but there is substantial individual-level heterogeneity in clinical presentation [155–160]. Estimates from early in the pandemic (pre-vaccination) suggested up to 20% of those infected over 80 years of age were likely to require hospitalisation [69, 161]. As knowledge of the virus developed, the understanding of the clinical profile of COVID-19 was updated. For example, the UK government added loss of taste and smell to the official list of symptoms in May 2020, after the first wave of infections [162]. Similarly, the emergence of new variants also induced changes in clinical presentation [160]. Although symptomatic individuals recover

between $6-19$ days on average [163], the WHO reports that $10\%-20\%$ of SARS-CoV-2 infections develop into post-COVID-19 syndrome [164]. Estimates of key epidemiological quantities for SARS-CoV-2 are presented throughout Section 2.4.

The risk of death from COVID-19 is strongly correlated with age [44, 159, 165], socioeconomic status [165, 166], ethnicity [165, 167] and the presence of underlying health conditions [159, 168]. Health and care workers have a substantially higher chance of becoming infected [169], with the global death toll among this group estimated at around $115,000$ in the first 18 months of the pandemic alone [170]. A more detailed discussion of the COVID-19 infection fatality ratio (IFR) can be found in Section 2.4.2.

Since the emergence of wildtype SARS-CoV-2, the virus has continued to evolve and new variants have emerged. On each occasion, public health officials sought rapidly to understand any changes to the virus, primarily in terms of three key characteristics: transmissibility, severity, and vaccine effectiveness [171] (although these are not the only characteristics of the virus epidemiology which could be impacted e.g. [172]). In particular, vaccine effectiveness can be measured in different ways, for example, protection against infection compared to protection against severe disease or death upon infection. However, over time it becomes increasingly difficult to disentangle the characteristics of variants (e.g. severity) from artefacts of the varied population immunity profile, arising from a combination of previous infection and vaccination. Substantial changes to at least one of the three aforementioned criteria "at a degree of global public health significance" led to the variant being designated a Variant of Concern (VOC) by the WHO, which has occurred five times since the beginning of the pandemic (Alpha, Beta, Gamma, Delta and Omicron) [173]. A further seven variants were designated Variants of Interest (VOIs), in which the aforementioned characteristics changed to the degree of "an emerging risk to global public health", but did not cross the threshold to become VOCs [173].

Towards the end of 2020, three VOCs, Alpha, Beta and Gamma, were declared following (initial) detection in the UK, South Africa and Brazil, respectively, reflecting the widespread global transmission of SARS-CoV-2 [174, 175]. Genomic investigation was instigated following rapidly increasing case numbers in each setting [175–177], and retrospective analyses showed the emergence of these variants in each location prior to their detection (in September, August and November 2020 for Alpha [178], Beta [175] and Gamma [177], respectively). Cases of each variant were subsequently detected internationally (with respect to where they were first detected) [176, 179–181]. Analyses showed that Alpha, Beta and Gamma were either intrinsically more transmissible or had the ability to escape immunity compared to wildtype SARS-CoV-2 [175, 177, 182–186], but fortunately vaccine effectiveness against both infection and severe

disease generally remained high for each variant [187–189] (with an exception of symptomatic infection against the Beta variant which was notably lower [187]). The impact on severity was less clear-cut: while Alpha was shown to have increased severity over wildtype SARS-CoV-2 [186, 190]; Beta was shown to have increased severity over Alpha [179, 191]; and it has been suggested that Gamma also had increased severity, although the evidence is less definitive [192]. Each of these three VOCs became dominant at national levels in different parts of the world, replacing the existing wildtype strain and often inducing the reintroduction of PHSMs, but none became dominant at the global level.

The Delta variant, first detected in India from as early as October 2020 [193], was designated a VOC in May 2021. Estimates suggested that this variant was $40\% - 60\%$ more transmissible than the Alpha variant [129, 194] and could evade immunity [195, 196]. Consequently, large waves of infections followed in many parts of the world, and unlike previous variants, Delta out-competed all other strains of the virus to become the dominant variant globally [87, 197]. In comparison to the Alpha variant, vaccine effectiveness against infection was estimated to have reduced [187, 198] but remained high against death [187, 199]. There were mixed results on severity compared to previous variants [186, 200]. In November 2021, the Delta variant was replaced by the Omicron variant [174]. First detected in South Africa, unlike previous VOCs, this variant was not retrospectively detected months before formal designation as a VOC [173]. Following the patterns described previously, Omicron was estimated as being more transmissible than Delta and having a lower level of vaccine effectiveness against infection but less so against severe disease [186, 188, 189, 196, 201–203]. Such findings resulted in the Joint Committee on Vaccination and Immunisation (JCVI) in the UK recommending rapid expansion of the COVID-19 vaccination booster campaign to include all adults over the age of 18 [204]. These properties also allowed Omicron to rapidly out-compete Delta, with the WHO reporting that this occurred within just four weeks [205]. However, Omicron was found to be less severe than the Delta variant [186, 196, 206–208].

As of December 2023, no new VOC has emerged to out-compete Omicron. Omicron sub-lineages continue to comprise the majority of sequenced SARS-CoV-2 cases reported to WHO, with four related subvariants designated as being of "interest" and a further seven under monitoring [209].

2.3 Epidemiological data

2.3.1 Officially-reported data

Officially-reported data, for example, the incidence of cases and deaths, are traditionally the primary source by which to track the spread and burden of an emerging infection. While these data are critical

to outbreak analysis and response, they are subjected to biases that may risk their representativeness of the true toll of the outbreak [5, 45, 210–217]. Both metrics rely on the availability and accessibility of testing infrastructure; but deaths also require established robust vital registration systems to be captured officially. The availability of these components is dependent on time and location: at the beginning of an outbreak testing infrastructure is unlikely to be widely available and thus biased to the most severe cases [45]. This is exemplified by Imai et al. [218], who estimated between $427 - 4,471$ infections in Wuhan in early January compared to just 41 reported cases. Similarly, Jit et al. [115] estimated that only 1% of infections were captured in official data in the UK up to March 2020. Regarding mortality data, in 2020 the WHO estimated that 40% of the worlds deaths from all causes go unregistered [219], predominantly in LICs. Although deaths have traditionally been considered as more reliable in capturing the burden of an infectious disease, there are examples from this pandemic which challenge this assumption [58, 214, 220, 221]. Even if these surveillance issues were overcome, differing definitions of a COVID-19 case and death across countries would continue to make national-level comparisons unreliable [222–224].

The UK has been recognised globally as having robust SARS-CoV-2 testing surveillance systems [225, 226], providing a wealth of data by which to track the spread of COVID-19 in each of the four nations. Notably, these data have been made publicly available in real-time from early in the pandemic via the UK Government COVID-19 dashboard [112], which has drawn widespread praise [226, 227]. Different surveillance systems inherently have different sampling mechanisms underpinning them, thus capture a different subset of the infected population. Within the UK, community-level surveillance systems include the use of both polymerase chain reaction (PCR) and lateral flow device (LFD) tests, and are split into pillars, with Pillar 1 capturing those with clinical need and healthcare workers, and Pillar 2 focussing on the community in general.[3] Although cases resulting from such non-random surveillance systems typically capture a biased subset of all infections, they remain an important indicator of the circulation of a virus through a population. In particular, these data are often used to estimate key epidemiological parameters, including the effective reproduction number ($R(t)$) and instantaneous growth rate ($r(t)$), [229–231] which in turn were used by governments to inform public health policy throughout the COVID-19 pandemic [65, 232–234].

2.3.2 Population-level studies of infection and seroprevalence

Surveillance studies deploying random sampling methods and large sample sizes, such as the ONS CIS [63] and the REal-time Assessment of Community Transmission (REACT) [235], are regarded as "gold-standard" in terms of estimating the prevalence of infection or antibodies to a virus within a specific

[3] There are also Pillars 3 and 4, referring to antibody testing and national-level surveillance studies for both prevalence of infection or antibodies, such as the ONS CIS, respectively [228].

population. Both types of surveys seek to sample randomly from a population of interest, regardless of factors such as symptoms or behaviour. Infection surveys use swab-tests (e.g. PCR or LFD) to detect active infection, thus obtaining estimates of the prevalence of the infection while serological surveys (sero-surveys) test for antibodies within blood, thus estimating seroprevalence [236]. Such studies are often, depending on sampling strategy, subject to fewer of the biases typically suffered by community-level testing surveillance systems described above and have been a crucial resource in estimating the burden and understanding the epidemiology of SARS-CoV-2 globally [216, 237–240]. Papers IV - VI (Objective 2; Table 1.1) each depend on studies of infection and/or seroprevalence in order to understand further the transmission dynamics in different settings and at different stages of respective epidemics.

The ONS CIS [63] was the only regular COVID-19 testing study encompassing all four nations of the UK, and remained the primary way in which the circulation of SARS-CoV-2 could be tracked after the cessation of community testing in April 2022. Although the study was initially "paused" in March 2023, the Winter COVID-19 Infection Study will track the prevalence of SARS-CoV-2 throughout the winter period 2023/24 [136, 137]. In addition to tracking the state of the epidemic [241], data from the ONS CIS have also been used to examine risk factors for severe infection [242, 243], the burden of long COVID by socioeconomic status [244], within-household transmission dynamics [245] and the evolutionary dynamics of SARS-CoV-2 variants over time [246]. Further information on this study can be found in Paper VI (Chapter 8).

While infection surveys are key to tracking the current burden of infection, there is only a relatively short window during which infected people will test positive (approximately $10 - 14$ days for SARS-CoV-2) [247]. Due to the longer duration of antibodies, particularly immunoglobulin G (IgG), which on average only begins to wane after several months [44, 248], estimates of seroprevalence are an important epidemic indicator. They provide means by which to assess the total number of people infected over time [249–251], particularly in the early phase of an outbreak, but even after the implementation of vaccination, seroprevalence remains important to track levels of immunity in the population [240, 252, 253]. There have been more than $4,400$ SARS-CoV-2 seroprevalence studies conducted across the world since the beginning of the pandemic, as documented by *Serotracker* [254]. In some settings with low reported COVID-19 death tolls, serosurveys have revealed extensive SARS-CoV-2 transmission [255–258] inconsistent with the number of infections expected based on officially reported deaths and estimates of the IFR [44, 259]. Further discussion is presented in Paper V (Chapter 7).

2.3.3 Excess mortality

Given under-reporting in COVID-19 deaths [5, 58], excess mortality has been used as an alternative means by which to track the impact of the COVID-19 pandemic. Excess mortality is defined as the difference between observed and expected deaths, with a positive value indicating a greater number of deaths than expected and vice versa. There are different methods by which to estimate excess mortality, but the principal idea relies on the estimation of a baseline level of mortality that is expected based on that which has been observed in years prior, adjusting for contributing factors such as population growth, humanitarian incidents and long-term mortality trends [224].

Karlinsky and Kobak [224] have shown that trends observed in excess mortality follow those observed in officially-reported COVID-19 mortality [224]. Although it is not possible to confirm that all excess deaths are attributable directly as a cause of being infected with SARS-CoV-2, Ghafari et al. [260] have demonstrated agreement between estimates of attack rates under models fit to excess mortality data and observed seroprevalence studies. In 2021, the Peruvian government revised its official COVID-19 death toll upwards by a factor of close to three (from $69,342$ to $185,380$), to align these figures with observed excess mortality and to overcome the issue of a positive test around the time of death [261]. The use of excess mortality has been adopted by the WHO to estimate the true toll of the pandemic [5] for similar reasons. However, estimation of excess mortality relies on the availability of all-cause excess mortality data, which in turn relies on vital registration systems. Consequently, many of the estimates of excess mortality in LICs have relied solely on model-based inference by pooling information from countries with similar socioeconomic and demographic characteristics, rather than being driven by data from that particular setting [5]. Further discussion of this topic can be found in Paper V (Chapter 7).

2.4 Key epidemiological quantities

This section presents an overview of five key epidemiological quantities of interest in outbreak control: the reproduction number, the IFR, the incubation period, the serial interval and the generation time.

2.4.1 The reproduction number

The basic reproduction number, R_0, is defined as the average number of secondary infections caused by an infected individual in a wholly susceptible population [262–264]. This is a key quantity used to define the transmissibility of a virus [265]. Estimation of this was a priority of many research groups at the beginning of the outbreak, suggesting an R_0 of wildtype SARS-CoV-2 of $2-3$ [70, 71, 266, 267].

A number of factors can influence the average number of secondary infections caused by an infected individual, for example, the implementation and lifting of PHSMs, the population immunity profile or the emergence of a new variant. Such effects are captured via $R(t)$, which estimates average secondary infections by an infected individual at time t which can in turn be used to assess whether an outbreak is growing ($R(t) > 1$) or in decline ($R(t) < 1$). After the introduction of control measures, attention turned to focus on estimating $R(t)$. Kucharski et al. [268] demonstrated a decline in $R(t)$ in Wuhan from 2.35 to close to 1 by late January 2020 after the introduction of PHSMs, while Knock et al. [269] showed the power of lockdowns in bringing $R(t)$ below 1 for each region in England. Further examples of the effects of the introduction of various PHSMs on $R(t)$ across Europe in the first half of 2020 can be seen in [270].

Multiple methods exist to estimate $R(t)$ [229, 271–274], with some also being developed since the emergence of SARS-CoV-2 [230, 231, 275]. Wallinga and Teunis [271] developed a likelihood-based framework using case data with dates of symptom onset and the distribution of the generation time. Cori et al. [229] built upon this method by instead using a Bayesian framework with case incidence at time t modelled as a Poisson process. However, both methods are hindered by censoring and can be unreliable during periods of low incidence (or lack of reporting, for example, at the beginning of an outbreak). As such, Parag [230] presented an alternative method for the estimation of $R(t)$ to overcome the aforementioned issues in which this quantity is modelled as a hidden Markov process and estimated via Bayesian recursive filtering and smoothing. $R(t)$ is also often recoverable from mathematical compartmental models.

In the UK, the government published estimates of $R(t)$ (and also $r(t)$) for each of the four nations since mid-2020 (approximately, dependent on nation). Estimation relied on an ensemble of up to 14 different models, each with different assumptions and data inputs as set out in Park et al. [65]. The diversity of surveillance data available in the UK across the pandemic makes a clear case for an ensemble modelling approach. Among the ensemble were models using the methods produced by Cori et al. [229] and Scott et al. [231] fitting to COVID-19 case data, and a model by Vohringer et al. [276] additionally incorporating genomic sequencing data. More complex mechanistic models, featuring two individual-based models, each being calibrated to multiple data sets such as hospitalisations, deaths, symptomatic reporting from NHS 111 and serology [277–282], were also included. Two models fitted to data from the ONS CIS [275, 283, 284].

During the COVID-19 pandemic, many political leaders, including UK Prime Minister Boris Johnson and German Chancellor Angela Merkel, frequently referred to the (effective) reproduction number (now colloquially known as the "R number") during press conferences [285]. In particular, Johnson has often

referred to the role that estimates of $R(t)$ have played in informing his decisions to lock down and re-open society [286, 287], and was formally listed as an alert criterion in the "UK COVID-19 alert system" [288].

2.4.2 Infection fatality ratio

While the case fatality ratio (CFR) is defined as the ratio of confirmed or suspected cases to the total number of deaths among these cases, the IFR is defined as the ratio of infections to the deaths caused by the infection, and indicates the severity of the virus. Although arguably a more straightforward concept than R_0, the total number of infections is not directly observable and must be estimated.

Statistical methods and mathematical models can be used to estimate the total number of infections [45, 58, 69, 218]. In early 2020, the aforementioned study by Imai et al. [218] inferred the potential number of SARS-CoV-2 infections in Wuhan using the probability of overseas travel and the average time taken for a case to be detected. Verity et al. [69] combined data on the number of infected travellers on repatriation flights with official case and death data within a more sophisticated Bayesian framework to produce age-stratified estimates of the IFR. The early estimates from Verity et al. [69] have proven to be remarkably accurate when compared to studies conducted later in the pandemic (e.g. [44]) at which point more data were available and our understanding of the epidemiology of SARS-CoV-2 had vastly improved. Both Imai et al. [218] and Verity et al. [69] concluded that the true number of infections exceeded the number of reported cases, with the difference described as "substantial" by Imai et al. [218].

Serological data provide an additional means by which to both estimate the IFR as well as to validate alternative methods, such as those outlined above. For example, Brazeau et al. [44] used 10 such studies of SARS-CoV-2 antibody prevalence within a Bayesian statistical model, which additionally accounts for delays between infection and antibodies being detectable and infection and death, the uncertainty in the serostudies and antibody test sensitivity and specificity, to produce age-stratified IFR estimates. They estimated the IFR at less than 0.2% among children under 19 years old, but this rose to above 2% for those aged $70-74$, and reached 16.4% among those over 90 years old (although the sample size of the latter group is small) [44]. In the same study, Brazeau et al. [44] also demonstrated large differences in the IFR by setting, with estimates of 1.1% in high-income countries (HICs) compared to 0.22% in LICs. O'Driscoll et al. [259] also used a Bayesian model to produce their estimates, which included fitting to seroprevalence data from 22 seroprevalence studies. These estimates are similar to those of Brazeau et al. [44], showing a strong log-linear relationship between age and mortality.

As with $R(t)$, the IFR also varies over time. For example, it has been well-documented that vacci-

nation has successfully reduced the IFR [97, 199, 289, 290]. In addition to vaccination, therapeutics, improved clinical care and/or the emergence of new variants also impacts the IFR. Early results from the Randomised Evaluation of COVID-19 Therapy (RECOVERY) trial showed that dexamethasone was associated with a significant reduction in mortality among COVID-19 patients [291]. Moreover, while Knock et al. [269] showed a 20% reduction in the IFR in England by the end of 2020 attributable to improved clinical care, Perez-Guzman et al. [186] illustrated the stark difference in the intrinsic IFR of the Alpha variant ($2.7\% - 3.2\%$) compared to the Omicron variant ($0.6\% - 0.8\%$) in England.

2.4.3 The incubation period distribution

The incubation period is defined as the time between an individual becoming infected and developing symptoms [292]. This often varies widely within individuals hence is presented frequently as a probability distribution rather than as a single estimate. The incubation period distribution informs self-isolation and quarantine policies for individuals suspected to have been exposed to a virus [70, 292–294]. Estimation of the incubation period distribution requires dates of individual symptom onset and estimates of the date of infection, of which the latter is not typically directly observable, and therefore also has to be estimated. This can involve epidemiological investigation on a case-by-case basis, for example through tracing and discussions with close contacts (potential infectors) of the infectee [295]. Statistical methods have been developed in order to estimate the incubation period while accounting for these uncertainties. Reich et al. [296] derived likelihoods for this parameter by firstly considering the times between infection and developing symptoms as doubly interval-censored data in comparison to reducing these data to a single observation, referred to as interval-reduced data.

The sophistication of techniques used to estimate the incubation period distribution across pathogens vary substantially. Virlogeux et al. [293] treated the exposure and symptom onset times as random variables and used Markov Chain Monte Carlo methods to estimate the mean of alternative parametric distributions for MERS-CoV. They then compared these estimated parametric distributions to those obtained using a non-parametric estimator developed by Turnbull [297]. However, for SARS-CoV-2 estimation has primarily relied on fitting parametric probability distributions to the observed data, with exposure times determined via epidemiological investigations, as presented in [70, 298–300]. Linton et al. [292] also adjusted for right-censoring in their estimation procedure. Variation in definitions and self-reporting of both exposure and symptom onset likely contribute to the variability in estimates [293].

The incubation period of wildtype SARS-CoV-2 is estimated to range between $2 - 18$ days [70, 292, 301, 302]. Galmiche et al. [303] demonstrated variation in the incubation period by variant, with the

mean for the Omicron variant estimated to be one day shorter than that of the wildtype strain (3.6 vs 4.6 days, respectively). A similar effect of a shortening incubation period for new VOCs was found by Wu et al. [304].

2.4.4 Generation time and serial interval distributions

While the incubation period is calculated based on individual cases, the generation time and serial interval measure the timings of disease progression between infector and infectee. The generation time is defined as the period between the dates of infection and the serial interval as the period between the onset of symptoms between infector-infectee pairs [305]. Similar to the incubation period, these quantities are typically described by probability distributions to account for variability across pairs. It follows that the generation time and serial interval distributions are identical when there is a fixed time between infection and symptom onset. However, this is not the case for SARS-CoV-2 as transmission can occur before an infected individual develops symptoms [294, 306]. These quantities are important for inferring the transmissibility of a virus from the incidence time series, as demonstrated by their use in the estimation of $R(t)$ in [229, 230].

As the onset of symptoms can be observed and typically the infection event cannot, the generation time is often approximated using the serial interval. Britton and Tomba [307] demonstrated that while the distributions of these parameters should always have the same mean, the variance of the serial interval will likely be greater than that of the generation time, thus rendering the former a biased estimator of the latter.

Regardless of whether estimating the generation time or serial interval, in addition to dates of individual symptom onset and estimates of the date of infection, these quantities require knowledge of probable transmission pairs, which presents a further challenge for epidemiologists [308]. For example, at the beginning of the COVID-19 pandemic, some researchers estimated the (distribution or mean/median and standard deviation of) serial interval of SARS-CoV-2 using small numbers of transmission pairs [70, 294, 309–311]. Again the sophistication of methods deployed varies. Some research groups [70, 295] fit parametric distributions to assumed transmission pairs, determined via contact-tracing and interviews, whereas others, such as [309, 312] opted for a Bayesian approach, building upon the methods set out by Reich et al. [296]. In order to estimate the mean generation time of SARS-CoV-2, Ganyani et al. [294] used similar methods to [296] but additionally conditioned on previously estimated values of the incubation period.

Estimates of the mean serial interval of wildtype SARS-CoV-2 range from 3 to 8 days [70, 294, 302, 309–

311]. As with the other parameters discussed in this section, the distributions of these parameters have also been shown to change over time, primarily due to the emergence of new variants [172, 313, 314]. Hart et al. [172] demonstrated a shorter generation time for the Delta variant (mean 3.2 days) compared to the Alpha variant (mean 4.5 days), which the authors state could contribute to the Delta variant transmitting more quickly in households than the Alpha variant. Moreover, in 2020 Ali et al. [315] demonstrated that PHSMs were associated with a substantial reduction in the serial interval in the early months of the COVID-19 pandemic in China (from an average of 7.8 to 2.6 days).

2.5 Mathematical modelling of disease transmission, progression and control

The transmission of infectious diseases is a complex process [316]. Mathematical models are invaluable tools which simplify complex real-world processes and synthesise multiple pieces of information relating to the transmission, progression and control of an infectious disease over time. This typically includes the incorporation of key epidemiological parameters such as those set out in the previous section (Section 2.4). Through this coherent framework, which captures the multifaceted, time-dependent relationships between transmission and disease progression, disease transmission processes can be quantified and the results used to understand the dynamics and potential impacts of infectious disease outbreaks [262, 317, 318].

Mathematical modelling [262, 316, 319–324] has long supported decision-making in public health; the COVID-19 pandemic was no exception [43, 325]. There are multiple questions that these tools can help to answer: what could be the impact of an unmitigated epidemic on healthcare and mortality? [1, 46] What impact could control measures have on transmission? [269, 326–328] What factors drive transmission? [329, 330] What might have happened if different decisions were taken? [97, 269] Such models also allow for the validation of statistical and mathematical methodologies used in outbreak analysis [271].

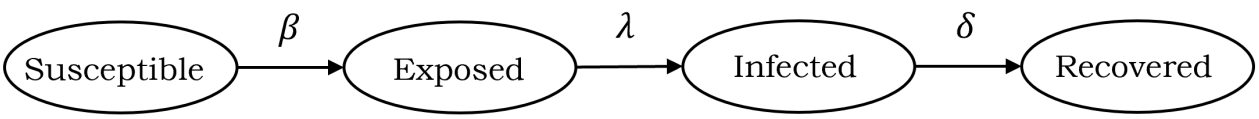

Figure 2.1: The basic structure of a Susceptible-Exposed-Infected-Recovered (SEIR) model.

Generally, in a mathematical model the population is split into distinct compartments from pre-infection (susceptible) to recovery or death, with flow through the model over time governed by epidemiological

parameters estimated from a variety of data sources [318]. Models typically follow susceptible-infected-recovered (SIR) or susceptible-exposed-infected-recovered (SEIR) structures, a basic example of which is provided in Figure 2.1, and are formally defined using ordinary differential equations. Denoting by $S(t)$, $E(t)$, $I(t)$ and $R(t)$[4] the number of people in the susceptible, exposed, infected and recovered compartments, respectively, at time t, the following set of first-order differential equations describe the model presented in Figure 2.1:

$$\frac{dS(t)}{dt} = -\beta S(t)E(t)I(t) \tag{2.1}$$

$$\frac{dE(t)}{dt} = \beta S(t)E(t)I(t) - \lambda E(t) \tag{2.2}$$

$$\frac{dI(t)}{dt} = \lambda E(t) - \delta I(t) \tag{2.3}$$

$$\frac{dR(t)}{dt} = \delta I(t) \tag{2.4}$$

The basic pathogen-agnostic SIR and SEIR frameworks can then be adapted to meet the needs of specific investigations. For example, due to the demonstrated risk of sustained nosocomial transmission of MERS-CoV, there are multiple examples of models for this pathogen with a particular focus on this transmission route [329, 331]. In the context of SARS-CoV-2, it became evident that the inclusion of compartments for hospitalisations and deaths was extremely important, in particular to decision-makers [1, 2, 46, 161, 186, 269]. However, there are also multiple examples of more bespoke adaptations for this pathogen. For example, in the UK, an important feature of the epidemic was the transmission of SARS-CoV-2 in care homes, as modelled by Knock et al. [269] and Hall et al. [332].

Mathematical models typically consider transition between states at the population level. Particularly in the case of COVID-19, the population is often stratified into age groups to account for the heterogeneity risks of transmission and progression within these groups, for example in [2, 161, 333–335]. Agent- or individual-based models (IBMs) continue to model flow through transmission and disease progression compartments but specifically capture individual-level behaviour and transmission risks in particular contexts [1, 320, 336, 337]. This could be particularly advantageous when there are complex immunity dynamics within the target population following infection and/or vaccination, as these two sources of immunity may have different levels of effectiveness and wane according to different time scales. Population-level dynamics can still be recovered by aggregating individuals of interest at each time point.

Mathematical models can be deterministic or stochastic. The choice depends on the questions under

[4]Note that $R(t)$ is standard notation for two different things in epidemiological modelling: the effective reproduction number at time t and the number of people in the recovered compartment of a mathematical model at time t. Throughout this thesis, $R(t)$ will predominantly refer to the effective reproduction number, but I will ensure that it is clear if this is not the case.

consideration and, often, the time stage of the epidemic. Chowell et al. [331] noted that stochastic models are useful in situations where there is low incidence, which is exemplified by the use of such models at the beginning of the COVID-19 pandemic [268, 333]. Stochastic models incorporate the effects of random chance. While this typically results in greater uncertainty and computational complexity, some have argued that stochasticity is necessary to reflect the complexity of transmission processes [338].

2.5.1 Model assumptions

Although extremely useful, mathematical models are necessarily imperfect representations of reality. Assumptions are inherently required in the development and parameterisation of these models as part of the quantification process and can be incorporated in many ways. The structure of compartments and the paths between them, as discussed above, is one such example (for example in [269, 331, 332]). A further example is through the modelling of control measures, which, undoubtedly, has been of great interest throughout the pandemic [1, 2] (see Paper I (Chapter 3) for further discussion).

The combination of the necessity of assumptions and the uncertainty surrounding parameter values (see Section 2.5.2 and Paper II (Chapter 4)) means that researchers often undertake scenario-based analyses, particularly when using models to project future epidemic trajectories [1, 57, 326]. In a scenario-based analysis, models are run under different assumptions of epidemic spread. For example, Ferguson et al. [1] considered the implementation of different mitigation and suppression strategies at different time points and in terms of their ability to reduce transmission of SARS-CoV-2 in March 2020 through to September 2021 in the UK and the United States of America. Similar approaches were taken by Davies et al. [46], who focused their analysis solely on the UK, and Walker et al. [2] who performed their analysis at a global scale, stratifying countries according to income status.

2.5.2 Parameter estimation

The estimates presented in Section 2.4 are among those that can be used to parameterise mathematical models. While some use estimates from previous analyses focused solely on parameter estimation, for example, in [44, 293–295], many researchers perform their own parameter estimation in conjunction with transmission modelling. Although different methods to do so were set out in Section 2.4, in modelling studies this is frequently conducted using MCMC methods [270, 326, 333, 339–341], which Endo et al. [342] described as "relatively easy" when the observed data can be appropriately described using a known, parametric probability distribution.

Regardless of the estimation method, typically, within a model, parameters are often fixed to one specific

value, such as the median estimate. However, as outlined in Section 2.4, there is often considerable uncertainty in estimates of key epidemiological quantities. Therefore, it may be that a range of values is scanned across (examples of both strategies, in the same models, can be found in [333]). For example, estimates of quantities such as the incubation period and serial interval distributions can be obtained by sampling from their respective probability distributions. Analogous to scenario-based analyses above, the parameter space is often explored by producing multiple realisations of the model using different parameter sets to capture this uncertainty.

2.5.3 Model estimation

A variety of methods exist to fit mathematical transmission models to data. The choice depends on the model assumptions, available data and computational complexity.

Numerical methods, such as those based on Runge-Kutta iterative approximation methods, are commonly used to simulate from models [343, 344], such as in [345, 346]. An alternative approach for (stochastic) model estimation is the Gillespie Stochastic Simulation algorithm [347], initially developed to model chemical reactions, in which transitions between states are governed by propensity functions defining the probability of a transition occurring in an infinitesimally-small time interval.

Endo et al. [342] noted that particle MCMC, in which both the dynamic process and measurement errors of the observed data are modelled, is a powerful but under-used method in infectious disease modelling. In this instance, observed time series of epidemiological data, for example daily incidence of cases and deaths, are modelled as partially-observed Markov chains. Such methodologies allow for easier exploration of the parameter and state space, and for easier incorporation of uncertainty [342]. An example of a model fitted using this methodology can be found in [57].

Dankwa et al. [348] underlined the importance of the concept of identifiability in ensuring the reliability of inferences drawn from (all) models, but noted that it is often not considered in the field of infectious disease epidemiology. Parameters are said to be "identifiable" if it is possible to define a unique set of parameter estimates for a particular model structure and dataset [348, 349]. For example, were a model not to be identifiable, different input parameter sets could produce identical outputs; consequently, answering questions surrounding, for example, the drivers of transmission or the impact of countermeasures becomes substantially more difficult. Roosa et al. [350] demonstrated that the risk of issues with identifiability is positively related to model complexity.

2.5.4 Alternatives to dynamical mathematical modelling of transmission

Although the utility and widespread use of dynamical mathematical models for epidemics has been demonstrated throughout this section, these models require assumptions in addition to incidence data and parameter estimates which can present challenges for implementation.

Many alternative methods have been developed. Some groups have opted for straightforward time series models for variables such as cases and deaths, for example in [351, 352]. While this overcomes many (but not all) of the assumption and data requirements, these relatively simple models also do not capture the intricacies and interconnections of transmission dynamics and epidemiological outcomes, with separate models required to explore metrics of interest. Alternatively, Chowell et al. [353] proposed producing short-term forecasts of case numbers using generalised logistic growth models fit to sub-epidemics, defined as epidemics within distinct groups of the population. This methodology includes three parameters, the growth rate, scaling of the growth rate and the final epidemic size, which require either estimation or exploration. Roosa et al. [354] used this framework to produce real-time forecasts of COVID-19 incidence in early 2020. Branching processes have also been proposed to model the spread of an infection [355]. This allows for the incorporation of the serial interval, for example, as demonstrated by Abbott et al. [356]. However, similarly to time series models, such frameworks typically focus on only one set of estimates at once, in comparison to mathematical models which can simultaneously output multiple quantities (e.g. infections, hospitalisations and deaths).

Chapter 3

Paper I: Disease transmission and control modelling at the science-policy interface

R McCabe and CA Donnelly. Disease transmission and control modelling at the science–policy interface. *Interface Focus*, 11(6):20210013, 2021. doi: https://doi.org/10.1098/rsfs.2021.0013.

Abstract

The coronavirus disease 2019 (COVID-19) pandemic has disrupted the lives of billions across the world. Mathematical modelling has been a key tool deployed throughout the pandemic to explore the potential public health impact of an unmitigated epidemic. The results of such studies have informed government's decisions to implement non-pharmaceutical interventions to control the spread of the virus. In this article we explore the complex relationships between models, decision-making, the media and the public during the COVID-19 pandemic in the United Kingdom of Great Britain and Northern Ireland (UK). Doing so not only provides important historical context of COVID-19 modelling and how it has shaped the UK response, but as the pandemic continues and looking towards future pandemic preparedness, understanding these relationships and how they might be improved is critical. As such, we have synthesised information gathered via three methods: a survey to publicly listed attendees of SAGE, SPI-M and other comparable advisory bodies, interviews with science communication experts and former scientific advisors, and reviewing some of the key COVID-19 modelling literature from 2020. Our research highlights the desire for increased bidirectional communication between modellers, decision-makers and the public, as well as the need to convey uncertainty inherent in transmission models in a clear manner. These aspects should

be considered carefully ahead of the next emergency response.

Author contributions

RM conducted the literature review, drafted the questionnaire, conducted the interviews, analysed the questionnaire responses and the content of interviews and drafted the manuscript. CAD conceived the study and edited the questionnaire and manuscript.

Ethics

This study has approval from the Medical Sciences Interdivisional Research Ethics Committee at the University of Oxford with Ethics Approval Reference R74566/RE001.

Reproducibility

Appendix A provides the quantitative survey results. Written answers to survey questions and transcripts of interviews are not provided to protect the confidentiality of participants.

Acknowledgements

We would like to acknowledge survey respondents and interviewees for their participation in this research, particularly in light of all the demands placed on their time throughout the pandemic.

Appendix

Further information relating to this paper can be found in Appendix A.

3.1 Introduction

Since the discovery of the novel coronavirus SARS-CoV-2 in late 2019, the lives of billions of people across the planet have been substantially disrupted. This highly transmissible virus, which causes coronavirus disease 2019 (COVID-19), has threatened the stability of healthcare systems globally with the official death count sitting at over 3.8 million people as of 15 June 2021 [357].

Mathematical models of transmission dynamics are commonly used to understand the dynamics and potential impacts of infectious disease outbreaks. As the initial epidemic unfolded in Wuhan in early 2020, scientists sought to understand rapidly the key epidemiological characteristics of the virus, such as the serial interval distribution and infection fatality ratio, in order to parameterise such models [69, 298, 309]. This then allowed the exploration of urgent public health questions, for example: how many people have been infected with the virus? Might the unmitigated spread of the virus overwhelm healthcare systems? [1, 46] The results of such studies are one piece of evidence used to inform governments' decisions to take unprecedented actions to control the spread of the virus in the form of stringent non-pharmaceutical interventions (NPIs). As well as reducing transmission of the virus, these measures have drastically disrupted the normal functioning of society with no section of the population left unaffected by the pandemic and its control.

In the light of the unprecedented pandemic situation that we have collectively experienced, we have sought to explore the complex relationships between models, decision-making, the media and the public during the COVID-19 pandemic in the United Kingdom of Great Britain and Northern Ireland (UK). Doing so not only provides an important historical context of COVID-19 modelling and how it has shaped the UK response, but as the pandemic continues and looking towards future pandemic preparedness, understanding these relationships and how they might be improved is critical. As such, we have synthesised information gathered via three methods, encompassing both original, primary data gathered by the authors in addition to some key modelling literature. Combining these complementary sources gives a rich and holistic insight into the utility, diversity, development and influence of modelling in policy in a more detailed way than considering one of these sources in isolation. First, we considered some of the most prominent modelling studies internationally and then focus on specific modelling studies conducted in the UK, the results of which were presented to scientific advisory groups before the introduction of the first national 'lockdown' in March 2020. Second, we surveyed individuals publicly listed as members, or having attended a meeting, of the Scientific Advisory Group for Emergencies (SAGE), the Scientific Pandemic Influenza Group on Modelling (SPI-M) and other comparable advisory bodies about the development and role of modelling in COVID-19 decision-making. Third, we conducted interviews with individuals

including science communication experts and former scientific advisors in order to understand further the interplay between these communities, particularly in comparison to previous health emergencies.

3.2 Methods

3.2.1 Literature review of COVID-19 modelling studies in 2020

Since the beginning of the pandemic, an extremely large number of studies involving mathematical modelling, restricted in this review to compartmental or individual-based models, of the transmission and control of SARS-CoV-2 have been published. While many make important contributions to the scientific understanding of the spread and control of the virus, we have sought to identify and synthesise a subset of key influential studies published in 2020. Therefore, this review does not synthesise all of the COVID-19 modelling literature.

We searched Web of Science (WoS) for (COVID-19 OR SARS-CoV-2 OR 'coronavirus disease') AND (model* OR simulat* OR impact) AND (transm* OR 'non-pharmaceutical intervention') on 15 December 2020. The results of this search were ordered according to WoS citation count and the 100 most cited papers were downloaded. Due to the rapidly evolving public health situation, many scientists published influential research as preprints to ensure this important information was publicly available without delay. Therefore, we also searched Google Scholar (GS) for (model COVID-19 OR SARS-CoV-2 OR 'coronavirus disease' OR 'non-pharmaceutical intervention' simulat OR impact OR transm). The restrictive nature of GS did not allow for an identical search to WoS, and we found that changing the order of search terms affected the search results. Searches in GS returned $1,000$ results and are presented according to a bespoke algorithm using undeclared relevancy criteria. These $1,000$ results and their corresponding GS citation counts were extracted on 15 December 2020.

WoS citation counts encompass only those citations from other papers within the WoS database. Therefore, the GS citation count of a peer-reviewed paper often differs to WoS citation count due to the broader inclusion criteria. To combine the results of both searches, GS citation counts for each of the WoS papers were extracted on 15 December 2020. The two groups of studies were then combined, duplicates were removed, and the studies were ordered according to their GS citation count. Using this list, the abstracts were sequentially screened until all 100 WoS studies had been considered. Systematic reviews and publications without mathematical models of SARS-CoV-2 transmission and disease progression as defined above were excluded from the review. Figure 3.1 summarises this process.

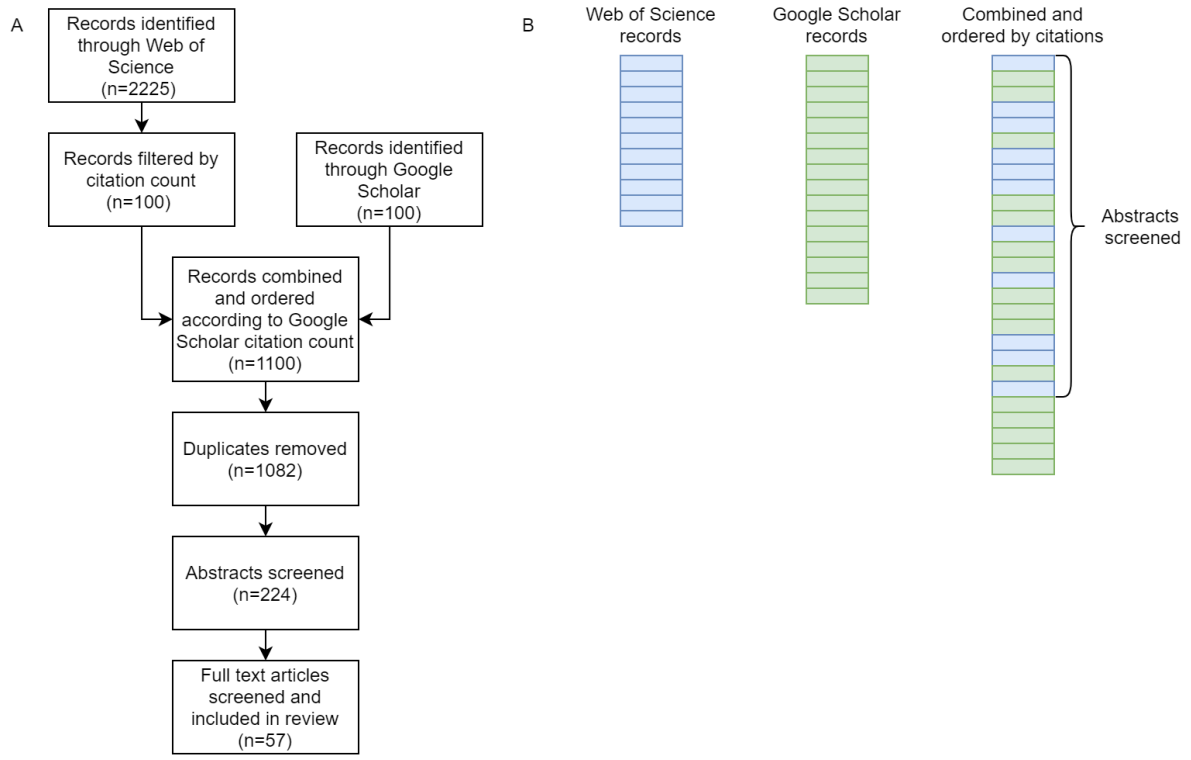

Figure 3.1: Overview of literature review search methodology. (A) Schematic diagram illustrating the process in which articles were identified and screened. (B) Illustration of how records from WoS and GS were combined and the number of abstracts that were screened was determined.

3.2.2 Survey to attendees of the Scientific Advisory Group for Emergencies, the Scientific Pandemic Influenza Group on Modelling and other comparable advisory committees

In order to further understand the development and role of modelling in COVID-19 decision-making, a short survey of eight multiple-choice and two longer-answer questions was sent to individuals publicly listed as being a member of, or having attended a meeting of, the SAGE, SPI-M or the Scottish Government COVID-19 Advisory Group during the pandemic [358–360]. In addition, a small number of individuals involved in similar committees overseas were also included. This study has approval from the Medical Sciences Interdivisional Research Ethics Committee at the University of Oxford with Ethics Approval Reference R74566/RE001. The participant information sheet (PIS) is included in the Appendix (Section A.2).

The survey was produced on Microsoft Forms and contained the questions set out in Table A.1. After each multiple-choice question, participants were invited to provide any further details which they consider relevant to the question and/or their answer to said question via an optional open-ended question box. Participants were sent a link over email along with the PIS. Responses were gathered between

25 January 2021 and 30 March 2021. Answers to all multiple-choice questions remain anonymous. Participants were explicitly informed that any written answers provided, either via the open-ended questions or follow-up to multiple-choice questions, may be quoted in the article and attributed to them under one of three levels of anonymity: full anonymity, job title only, full name and job title (Table A.1, Question 18). Participants that have been quoted were notified of and consented to the exact wording used in the manuscript ahead of publication.

3.2.3 Interviews with science communication experts and former scientific advisers

To further examine the interplay between scientists, decision-makers, the media and the public with respect to modelling throughout the pandemic, individuals including science communication experts, former scientific advisers and policy engagement officers were invited to participate in interviews. The selection was based on the individual's expertise in the interface between science, policy and media and for which an email address could be obtained. This study has approval from the Medical Sciences Interdivisional Research Ethics Committee at the University of Oxford with Ethics Approval Reference R74566/RE001. The PIS can be found in the Appendix (Section A.4).

The interviewer prepared discussion points and questions ahead of the interview, which were shared with the interviewee upon request. Discussion points broadly covered: the responsibilities and opportunities of an individual in their role in the pandemic; their experience of the utility, or lack of, mathematical modelling throughout the pandemic; how communication between communities of interest could be improved and how the COVID-19 pandemic compared to previous health emergencies they have been involved in. The interviewer allowed for a natural flow of conversation wherever possible and encouraged interviewees to shape the conversation in the way that they desired in order to avoid influencing points raised by interviewees.

All interviews except for one were conducted virtually, with interviews lasting between 20 minutes and 1 hour depending on the availability of the interviewee.

3.3 Results

3.3.1 Primary data collection

A total of 189 individuals were identified as either being a member of or having attended a meeting of SAGE, SPI-M or other comparable advisory groups as of January 2021. Of these, we found contact

information online for 174 individuals who were subsequently invited to participate in the survey. We received a total of 46 responses (26% of those emailed; 24% of those identified), with most respondents participating anonymously. Figure 3.2 and Table A.1 present the results of the multiple-choice questions, the implications of which are discussed throughout the remainder of the article.

A total of 13 individuals were identified as suitable for interview based on the selection criteria. We conducted 11 interviews (85% of those contacted), the details of which are summarised in Table A.2.

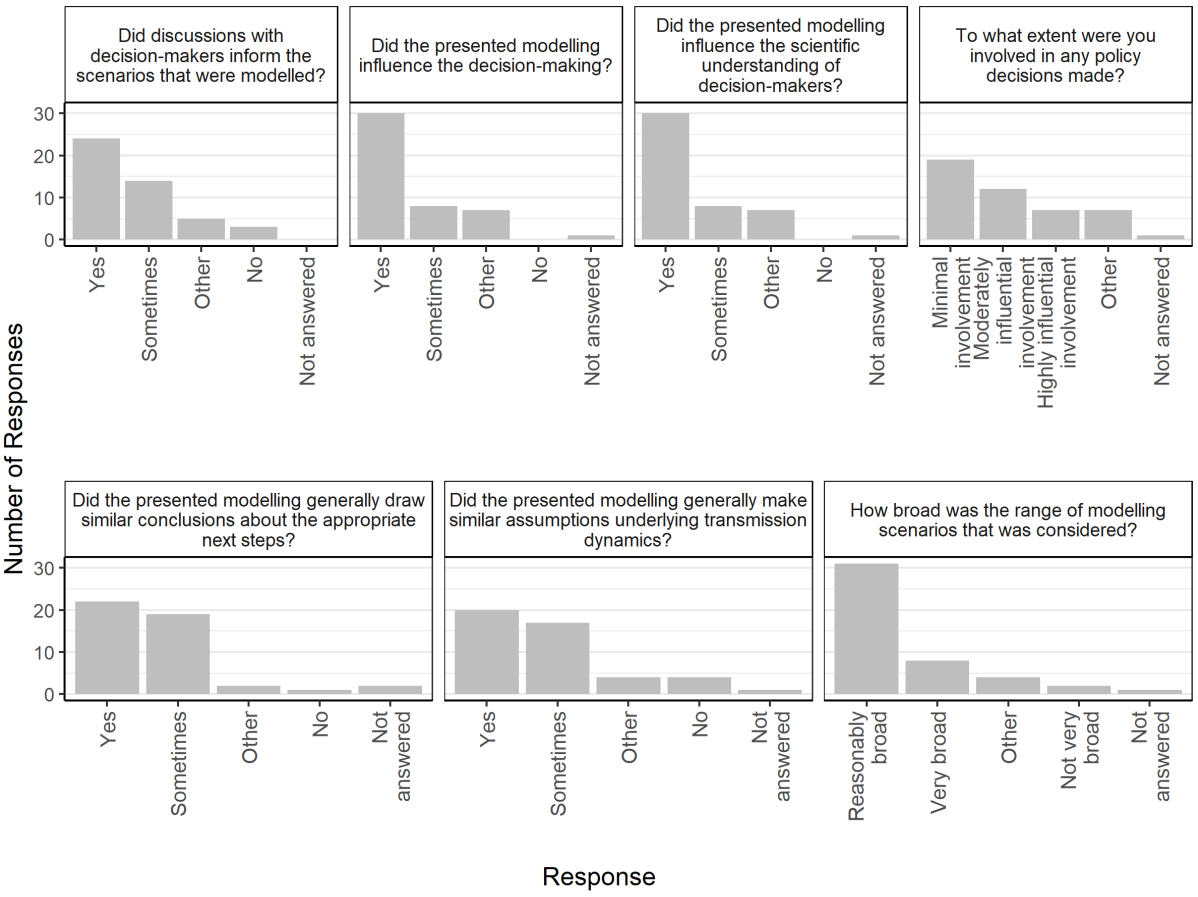

Figure 3.2: Results of the multiple-choice questions in the survey to attendees of SAGE, SPI-M and other comparable advisory bodies.

3.3.2 Mathematical models for outbreak response

The transmission of infectious diseases is a complex process, the dynamics of which are driven by a combination of biological, sociological and environmental factors [316]. Scientists often use mathematical models to simplify complex real-world transmission and disease progression processes by splitting the population into distinct compartments from pre-infection to recovery or death, with flow through the model governed by epidemiological parameters estimated from a variety of data sources [318]. These

models inherently require assumptions so that the process can be quantified and used to explore various questions (which may or not be directly of relevance to decision-makers).

While many models used in decision-making appear to be focussed primarily on providing epidemic projections based on certain assumptions or scenarios, models can be used far beyond this purpose: to understand a previous outbreak, to test the reliability of assumptions and to unravel the drivers of transmission. Additionally, studies measuring key parameters, such as the household attack rate, can inform decision-makers directly as well as through their estimates being used in transmission models.

Professor Sir Charles Godfray, Director of the Oxford Martin School, noted that the exploration of a situation as complex as an epidemic in which the future trajectory is unknown requires modelling: "There is no alternative". This is reinforced by Professor Jason Leitch, National Clinical Director for the Scottish Government, who added that modelling "is crucial".

3.3.3 The diversity of COVID-19 modelling

The WoS search returned 2,225 papers in total of which the 100 most cited were extracted. The citation counts of these papers were (with one exception) greater in GS, with the order of most cited to least cited generally preserved. Nineteen duplicates were removed. From the remaining 1,081 papers, the abstracts of the 223 most cited (thus encompassing the 100 most cited WoS studies) were screened. Of these, 53 met the inclusion criteria thus were included in the review (Figure 3.1) [1, 46, 161, 266, 268, 270, 326–328, 330, 333–336, 339–341, 345, 346, 349, 361–393].

Dr Sabine van Elsland, External Relationships & Communications Manager at the Medical Research Council Centre for Global Infectious Disease Analysis (MRC GIDA) at Imperial College London, noted her pride in the way the academic community came together to understand the characteristics of SARS-CoV-2. This is exemplified by the breadth of studies included in our review, many of which were published within just weeks of the beginning of the outbreak in Wuhan.

Models included in the review broadly followed susceptible–infected–recovered (SIR) or susceptible–exposed–infected–recovered (SEIR) structures, but are diverse in their applications and adapted as required to meet the needs of specific investigations. A common example of this among the reviewed papers was the inclusion of compartments for individuals requiring hospitalisation (15 studies [1, 46, 161, 266, 326, 327, 340, 346, 361, 365, 368, 371, 375, 379, 382]). Some researchers opted for more bespoke adaptations: Kucharski et al. [268] duplicated model compartments for the population in Wuhan and international

travellers while Eikenberry et al. [368] split the population into compartments according to mask-wearing habits. The diversity spanned beyond model structure. For example, studies in our review examined outbreaks in different locations [46, 161, 270, 328, 334], the impact of factors such as age or asymptomatic infection on transmission [268, 330, 333, 335], the use of NPIs as a control measure [1, 339, 368, 375], or, often, a combination of multiple factors. In isolation, each study contributes an interesting analysis, but together they build a rich and wide-ranging scientific understanding of SARS-CoV-2 and options for its control.

Different data sources and the way model parameters were estimated varied more than basic model structure. Modellers examining the early outbreak in Wuhan had limited data with which to parameterise models, which often centred on relatively small numbers confirmed COVID-19 cases and even used biological characteristics of the severe acute respiratory syndrome coronavirus (SARS-CoV-1) [266, 268, 298, 309, 330, 336, 345, 349, 361, 370]. More informative data fed into models as they became available, for example as the epidemic spread to other locations and more affected countries published their epidemiological data. This was explained by Professor Leitch, who noted the improvement of mathematical modelling over the course of the pandemic as more information had become available as well as people's understanding of it: "The modelling has gotten better because 14 months ago [March 2020] we didn't know what this virus was, it didn't have a name. So now we know a lot more about what it [the virus] does and how it behaves and so we end up with better modelling and better translation of modelling". Nonetheless, 32 studies in our review reported data quality, uncertainty in epidemiological parameters and/or rapidly changing insights generated from data in the unfolding epidemics, as caveats of their analysis [1, 46, 268, 270, 328, 333, 335, 336, 340, 345, 346, 349, 361, 363–367, 370–373, 377–382, 385, 388, 389, 392, 393]. The final point was explicitly noted by SAGE with their published evidence [359] and was emphasised by many survey respondents and interviewees.

Throughout 2020, control of the virus centred on the implementation of NPIs, such as social distancing and school closures. As such, it is unsurprising that modelling their impact on metrics of interest, such as peak infections, peak hospitalisations and cumulative COVID-19-related deaths, was a central focus of most studies in this review (48 studies) [1, 46, 161, 266, 268, 270, 326–328, 330, 333–336, 339–341, 346, 349, 361–376, 379–381, 383, 384, 386–393]. Some also presented metrics relating to the effect of the intensity or timing of NPI implementation on reducing these metrics [1, 46, 266, 340, 364, 375]. The specific interventions modelled varied substantially between studies: social distancing, quarantine and isolation of cases were among the most common, with 18 studies considering multiple interventions [1, 46, 270, 326–328, 330, 335, 339, 340, 346, 349, 363, 365, 366, 369, 375, 391] and 11 studies considering generic reductions

in transmission rather than a particular intervention [266, 341, 362, 367, 373, 374, 380, 381, 386, 390, 392]. These were modelled structurally in different ways, from altering transmission rate parameters or contact patterns, and implemented at different times, for example at fixed time points or triggered based on disease burden. This heterogeneity means that even within the same setting it is difficult to compare study results directly, but all studies in this review considering NPIs illustrated the potentially severe consequences of allowing a COVID-19 epidemic to progress without mitigation.

The authors of four studies self-reported their role in informing decision-making either in the UK or locally in the United States [1, 46, 382, 383]. Weissman et al. [382] presented estimates of the demand for hospital resources from COVID-19 patients in Philadelphia under different assumptions of the epidemic doubling time. In contrast with the other three papers, Paltiel et al.'s [383] study did not focus on healthcare rather on the role of testing in preventing outbreaks on US college campuses. In addition to estimating the number of cases under different testing policies, they provided an analysis of the cost effectiveness of each scenario in relation to the number of tests and cases averted. Other studies in our review are likely to have been used in decision-making as well, but the authors have not declared this explicitly.

3.3.4 The COVID-19 pandemic and science advisory mechanisms in the United Kingdom of Great Britain and Northern Ireland

Within one year of recording the first confirmed case of COVID-19 on 30 January 2020, the UK had registered almost four million confirmed cases and nearly 110,000 deaths due to the virus [394, 395]. This was in addition to the substantial economic toll, with a reduction in gross domestic product growth and an increase in unemployment [396, 397]. At the time of revision (June 2021), since the government implemented the first national lockdown on 23 March 2020, most of the population have remained under some form of restrictions aimed at controlling the spread of the virus [398].

SAGE is the central advisory body responsible for providing "scientific and technical advice to support government decision-makers during emergencies" [399]. Professor Leitch noted that if decision-makers were not being asked to make such difficult and complex decisions, there would not be a need for specialist advisers. An anonymous survey respondent underlined their role as an adviser: "Policy decisions are made by policymakers. Our job is to ensure they have the evidence.". This was corroborated by multiple survey respondents, with 41% stating that they had "minimal involvement" in any policy decisions, compared to just 15% who considered their involvement as "highly influential" (Figure 3.2). An overview of the science advisory mechanisms in the UK is given in [43].

Scientific participants of SAGE are politically independent experts from a range of fields, which during the COVID-19 pandemic has included medicine, virology and epidemiology [359]. SAGE synthesises scientific evidence pertinent to the threat posed by SARS-CoV-2 from a variety of sources, of which modelling is only a subset as noted by multiple survey respondents and Dr Claire Craig, former Director of the Government Office for Science. Mathematical modelling of transmission dynamics and disease progression is the primary focus of SPI-M, a specialist advisory group consisting of politically independent modelling experts which reports to SAGE. Professor Godfray noted that he has been impressed by the way in which the modelling community in the UK has "stepped up to support decision-making" throughout the pandemic. As stated by an anonymous survey respondent, "SAGE meets to consider the results of the various models and then to develop a consensus statement (for decision-makers) that summarises what can be inferred from these". Figure 3.3A shows an example of combined model outputs from different SPI-M modelling groups to forecast intensive care unit (ICU) occupancy in early April 2020, while Figure 3.3B presents the individual model outputs that this combined output is comprised [400]. Although some survey respondents commented that they felt SAGE and SPI-M structures were "too large in their current form", in a modelling-themed issue of Philosophical Transactions of Royal Society B, Dr Ellen Brooks-Pollock, Dr Leon Danon, Dr Thibaut Jombart and Dr Lorenzo Pellis note that "The way SPI-M-O operates evolved during 2020—starting with a small number of modellers and expanding to around 50 modellers regularly attending the weekly meetings. This plurality of opinion was key to generating robust and reliable advice." [43].

3.3.5 The use of transmission modelling in the UK COVID-19 pandemic response

Professor Sir John Beddington, former government Chief Scientific Adviser 2008 – 2013, drew on his experience of modelling during the H1N1 influenza pandemic in 2009. He explained that, in light of many unknown parameter values, scenario-based modelling is the most appropriate and informative analysis for scientists to conduct. However, he recognised that during this pandemic, the modelling community are fortunate to have nationwide surveys to estimate COVID-19 prevalence over time, which can act as an independent verification of the results generated via mathematical models. This was further emphasised by Professor Linda Bauld, Bruce and John Usher Chair of Public Health at the University of Edinburgh, who noted that "Like all science, modelling needs to be triangulated with other data sources and understood in the wider context".

Two papers included in our review contained results of scenario-based analyses that were presented

to SAGE prior to the first national lockdown in March 2020 [1, 46]. Both explored the potential future impact of broadly similar NPIs on measures of epidemic magnitude such as the number of infections, the number of deaths and healthcare capacity, another key metric of COVID-19 control for the UK government [401]. In both cases, estimates of peak hospital bed demand from COVID-19 patients were presented alongside the corresponding length of time spent under NPI restrictions for a range of scenarios, highlighting the trade-offs inherent in such interventions. The nuances of the individual models, both in the differing assumptions and stochasticity arising in the model fitting process, mean that specific estimates from these studies are not directly comparable. Nonetheless, both studies fully aligned on their overall conclusions that without sufficient mitigation critical care resources would be overwhelmed at several orders of magnitudes above baseline capacity. As explained by Dr Marc Baguelin, a researcher at Imperial College London and London School of Hygiene and Tropical Medicine (LSHTM), "How fast models reach a conclusion differs but a clear majority of models are usually in favour of a specific decision/outcome", a statement corroborated by other survey respondents.

At the time of revision (June 2021), the aforementioned themed issue of Philosophical Transactions of the Royal Society B entitled "Modelling that shaped the early COVID-19 pandemic response in the UK" was published featuring 20 articles from various modelling groups regularly contributing to SPI-M during the pandemic [43]. In addition to further highlighting the diversity and applications of modelling in informing policy, the editors provided a timeline (January–July 2020) of when the work in each article was undertaken which illustrates the evolution of the focus of and data available to the contributing scientists over the first wave of the virus as mentioned in Section 3.3.3. For example, the first article used a mathematical model to estimate the basic reproduction number, R_0, in Wuhan (early 2020) [402]. Shortly after, scientists used parameters estimated using the outbreak in Wuhan to parameterise models to understand the potential spread of SARS-CoV-2 in England and Wales [113]. As the pandemic progressed, scientific understanding of the virus increased and more data became available, modelling was applied to consider more specific issues, for example, the risk and impact of outbreaks in care homes [332] and the implications of waning population immunity on transmission in the UK [403].

While the power of mathematical models is well understood, their influence on policy can be difficult to quantify [404]. Nonetheless, Scott Heald, Head of Profession for Statistics at Public Health Scotland, as well as Professor Leitch, highlighted the utility of models in planning for a range of policy options within the Scottish Government. For example, Heald noted that estimates of the number of cases have been used to inform the resources required for the test-and-trace programme in addition to hospital resources. It has been reported that "the Scottish Government uses the publicly available Imperial model as reported

in their Report 13 [270] to help understand the COVID-19 epidemic in Scotland over the longer term and what the reproductive rate at a point in time (R_t) is for Scotland.". Above all, Heald underlined, similarly to Professor Beddington and Professor Leitch, that the rapidly changing situation meant that while no model could reliably predict exactly what will happen in the future, the results can be reliably used to demonstrate what will happen if sufficient action is not taken. As noted by Professor Leitch: "Modelling doesn't lend itself to binary choice, but it is part of the framework of information that we had to take into account". These statements were further corroborated by several survey respondents, with one noting that models give "real decision advantage".

3.3.6 Development of modelling for use in decision-making

Modelling is only a useful tool to decision-makers if it considers "realistic" policy options, as noted by an anonymous survey respondent. Most survey respondents (85%) agreed that discussions with decision-makers at least sometimes informed the scenarios that were modelled and presented to SAGE and/or SPI-M (Figure 3.2). One such respondent noted that this very much depends on what is being asked, adding that "Sometimes it is very constrained by scenario other times people are free to bring interesting output and analyses to the group". Dr Craig noted that the availability of evidence and data can often frame policy discussions, likening this to lamp-posts illuminating distinct areas on, but not the entirety of, the ground (issues pertaining to the policy discussion). Nonetheless, when asked about scenarios covered at these committees, 82% of respondents categorised the range considered as at least "reasonably" or "very" broad (Figure 3.2). This is illustrated by Dr Mike Tildesley, a researcher at the University of Warwick, who noted: "There have been a significant number of different modelling scenarios considered by SPI-M, from short- and long-term forecasting, to the effect of school re-opening, to the effect of Christmas relaxation, the impact of bubbles, transmission on university campuses, the impact of a circuit breaker lockdown etc.", and echoed by Professor John Edmunds, a researcher at LSHTM, who stated "We have tried to cover the main policy options". One respondent noted that while "it was always possible to be broader", it is "important that scenarios were reasonable".

Respondents also generally noted a degree of flexibility given to modellers in the underlying assumptions in their models. Dr Ellen Brooks-Pollock, a researcher at the University of Bristol, noted "SPI-M contributors have been encouraged to make their own and best assumptions based on the evidence available at the time", while Billy Quilty, a researcher at LSHTM, added that "Requests were generally broad, with the specific scenarios and assumptions determined through conversations within our modelling group and with SPI-M/SAGE members". Despite this breadth and flexibility, the majority of survey respondents (79%) believed that different models were generally based on similar assumptions (Figure 3.2).

This is exemplified through a comparison of the studies by Ferguson et al. [1] and Davies et al. [46].

3.3.7 Communication between modellers and decision-makers

When asked how communication between modellers and decision-makers could be improved, many noted the success of the system at present. According to Professor Edmunds: "I actually think that the structures, such as SPI-M and SAGE have worked very well", with another, Dr Lorenzo Pellis, a researcher at the University of Manchester, adding "Obviously, the current state of things has evolved since the start of 2020, so if future health emergencies were starting from the current level of communication, it would be great". Many respondents stated that while modelling scenarios were informed by decision-makers, these conversations were often conducted by a third party, i.e. the SPI-M secretariat, rather than between the two parties directly. Furthermore, many suggested that communication flowed more often from modellers to decision-makers than in the opposite direction. Another respondent would have preferred to "Have an actual conversation between the two parties, rather than a one-way road. Key priorities and questions should be discussed between policymakers and modellers as there is often a trade-off between what decision-makers want to know, and what is achievable given available data". Dr Craig echoed that discussions between the two communities are important, particularly regarding real-time work as during the pandemic: "most of the time what you really want is discussions such that the questions scientists ask and answer are informed by the policymaking, but the policymaking questions is informed by an understanding of what the science is telling them and also what it's capable of doing". The potential mutual benefits of two-way communication were echoed by many respondents, with some suggesting that decision-makers should attend SAGE and SPI-M meetings to "hear the nuances of the debates" and "to enable the modelling teams to communicate the impact of uncertainty upon modelling outcomes".

Models of disease transmission are inherently complex and understanding the intricacies of how they work requires a substantial investment of time. William Pryor, Head of Policy Engagement at the University of Oxford, noted the mutual benefit to be derived from the more open exchange of evidence and expertise between decision-making and research communities. He also noted, however, that both needed to develop the skills required to combine the breadth of the former's knowledge and depth of the latter's. As noted by Professor Graham Medley, chair of SPI-M, modellers have the responsibility of ensuring that they develop as high-quality models as possible, but how these are used by decision-makers is out of their control. Another respondent touched upon this: "There are many more considerations in policy than just the science", a sentiment echoed by Dr Craig and Professor Godfray in our interviews, as well as Sir Patrick Vallance, government Chief Scientific Adviser, in [405]. These differences in perspective can make communication challenging, especially during an emergency such as the COVID-19 pandemic,

as explained by Pryor and also Allen et al. in their report for the International Network for Government Science Advice [406].

Communicating the uncertainty inherent in mathematical models to decision-makers was highlighted by Lord John Krebs, former Chairman of the UK Food Standards Agency, and Professor Beddington as a key area requiring improvement and confirmed by many survey respondents. One noted: "I think outputs appear to influence the thinking of decision-makers, though they are often too focused on a single scenario, rather than adequately incorporating the uncertainty inherent in such tasks". As highlighted by our literature review and discussed in [41, 407–409], the uncertainty in models ranges beyond measures of traditional statistical uncertainty: model parameterisation, the choice and structure of modelling frameworks and further assumptions about the drivers of transmission all introduce additional layers of uncertainty into models which are essential to communicate in the results. While these issues existed long before the discovery SARS-CoV-2, Professor Beddington noted that the issue of unknown parameter values is more prominent during the COVID-19 pandemic due to the novelty of the virus than it was during the H1N1 influenza pandemic which behaved similarly to other strains of influenza. In a seminar hosted by Turing-RSS Laboratory and Joint Biosecurity Centre in April 2021 [410], Professor Marc Lipsitch commented that equal emphasis should be placed on highlighting the assumptions underpinning the results as is given to the results themselves through the use of "if X then Y" statements rather than focussing solely on the "then Y" aspect.

Many survey respondents noted that both communities have a responsibility to use times between emergencies to improve preparedness. Professor Christine Middlemiss, UK Chief Veterinary Officer, drew on her experience of how a previous outbreak informed emergency responses since then: "Learn from lessons already learnt in the broader One Health space. (Foot and mouth disease) FMD 2001 provided a wealth of understanding on just how important and how to improve this communication, and we have invested significantly in it and evolved it". In terms of modelling, Professor Medley noted: "During a pandemic there is no time to develop novel models and to inform policy—the timelines are very short. We really need a next generation of models to be developed in the coming years to better suit the next pandemic". Professor Sir Jeremy Farrar, Director of the Wellcome Trust, highlighted the need to "Increase levels of knowledge, understanding, scientific literacy before any emergency. The work done in the "non-crisis" times determines the outcomes in a crisis. Don't try and forge those links, those insights in a highly dynamic situation, you will always be behind the curve".

3.3.8 Communication of modelling and decision-making to the public

The NPIs introduced to control the spread of SARS-CoV-2 have profoundly affected public life. According to an anonymous respondent: "Communication with policymakers hasn't really been a problem. Communication with the public a little more so". Professor Bauld noted that when (frequently) asked about the data informing decision-making during media appearances, she must "always refer to the modelling, because in the basket of indicators it is not just [say] incidence, prevalence, test positivity, it is also modelling forward what hospital capacity could be".

Like most decision-makers, the public are not experts in infectious disease modelling. Striking the balance between conveying the overarching results of a modelling study and the assumptions underpinning these results is challenging, as described by Robert Cuffe, Head of Statistics at the BBC, and echoed by Lord Krebs, Heald and Professor Leitch. Robert Cuffe also noted a transition away from reporting on individual modelling studies as occasionally observed at the beginning of the pandemic, with emphasis placed on measures such as case numbers instead. Although crucial to SAGE and SPI-M, Cuffe noted that "having multiple versions of the same number" in itself is not of interest to the public. Professor Leitch and Professor Bauld corroborated this and noted that they communicate modelling to the public in general terms. For example, Professor Leitch would say that "cases will increase over the next week", as opposed to "there will be [say] 127 new cases by next week".

Several science communication experts, including Fiona Fox, Director of the Science Media Centre, Dr van Elsland, Professor Leitch and Professor Bauld, have been striving to make modelling studies more accessible to and better understood by the public, particularly throughout the COVID-19 pandemic. In particular, Professor Bauld spoke about the personal responsibility she has felt to use her 25 years' experience in public health to support the communication of scientific evidence used in decision-making in an understandable way to the public. Referring to modelling, she noted that her role is "to understand the main messages, communicate the uncertainty and then to know when to stop commenting and pass over to someone who understands the methods more than I do".

Fox, Dr van Elsland and Professor Leitch noted a strong commitment from most journalists to ensuring that modelling results are communicated clearly and accurately. Dr van Elsland identified "transparency" as among the most critical factors when communicating modelling results to the public, so that the caveats of the analysis are understood while ensuring that the results are taken seriously, as corroborated by [411]. She noted the importance of "being clear and open about what the models do and do not know", a sentiment echoed by Fox, Lord Krebs, Dr Craig, Professor Godfray and Heald.

Professor Leitch stated that when communicating with the public during the pandemic, he is "not just trying to educate, but trying to get people to change their behaviour" in line with government regulations, which, as an adviser, is an additional challenge. Since March 2020, he has had a regular slot on football show "Off the Ball" on BBC Scotland [412], the "most listened to radio show in Scotland" and with an audience that "doesn't listen to other radio shows", in which he answers COVID-19-related questions posed by listeners. Professor Leitch explained that: "[listeners] often see the pandemic through their own lens of, for example, their wedding or their football match and you have to try and make your answers relevant to that. What you try and do is the translate an individual situation into a message or lesson for a broader population because there's probably a hundred thousand people with the same challenge".

Fox emphasised that modelling results should be communicated to the media and the public directly from the scientists themselves. This reinforces the distinction between the scientists (including modellers) and decision-makers, something which many survey respondents believe has been blurred during the current pandemic. Many scientists [407, 413, 414] have also raised this issue when criticising the UK government's "following the science" catchphrase during the pandemic. Professor Leitch stated that advisers "try to draw a distinction between the politician who is leading the briefing who has a particular job and the clinical advisers who have a different job and it is important to continuously redraw that distinction", referring to the Scottish Government COVID-19 briefings. Fox and Lord Krebs both noted that this distinction is a key component of building public trust in both modelling and decision-making. Dr Craig added that even if there is a distinction between these communities, if the public does not perceive this to be the case then this perception could equally weaken their trust in the decision-making process.

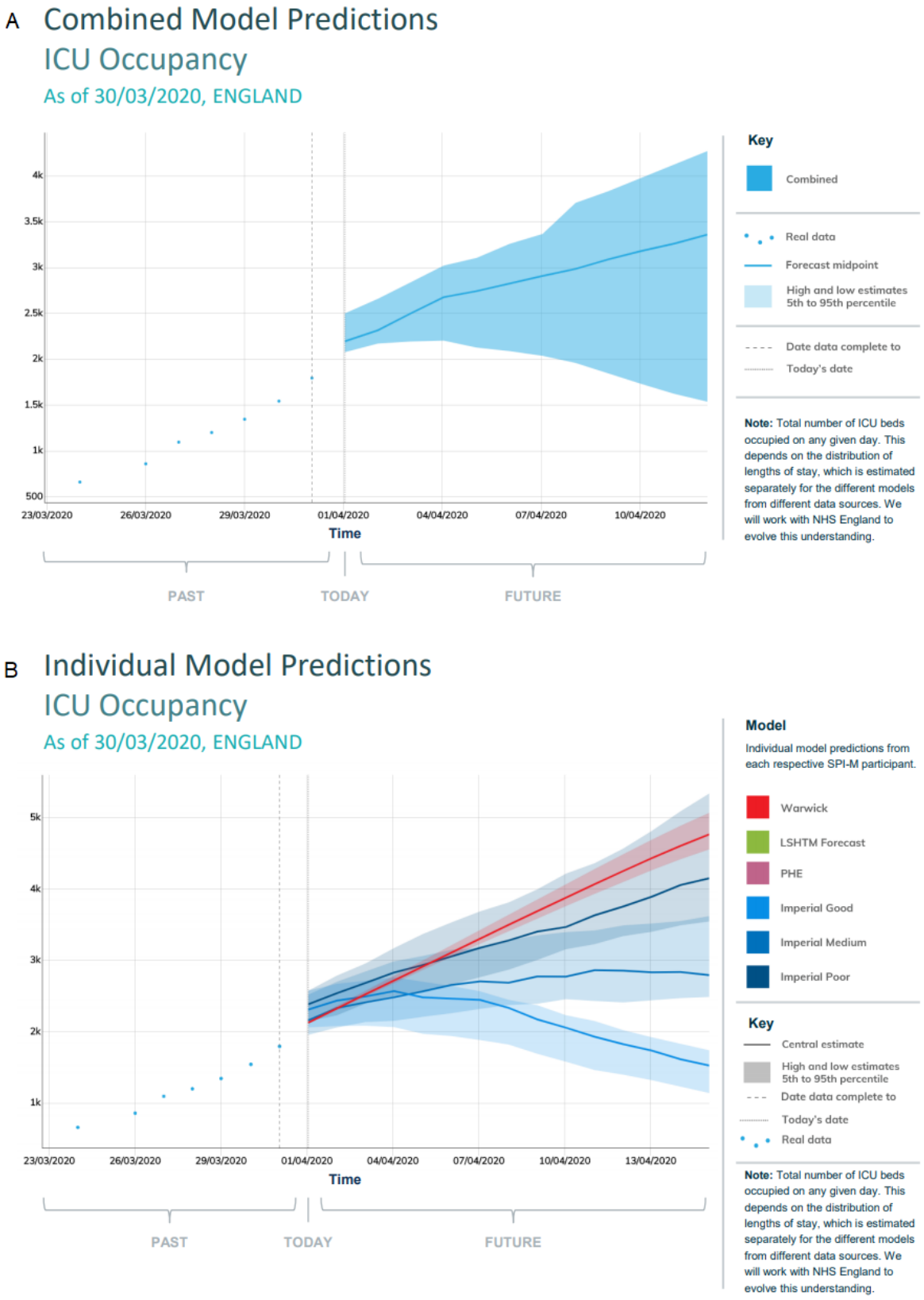

Figure 3.3: An example of modelling produced by SPI-M modelling groups in late March 2020 to provide short-term predictions of ICU occupancy in England in April 2020 [400]. (A) The individual modelling outputs, undertaken independently by different modelling groups, are combined to provide a wide range of possible trajectories for the epidemic in the following fortnight. (B) The outputs from the six individual model analyses which make up that presented in (A) are presented.

3.4 Discussion

This article has sought to provide a broader perspective on the scope and use of mathematical modelling throughout the COVID-19 pandemic, particularly in informing policy in the UK. Our literature review, although not exhaustive, highlighted both the diverse applications of modelling in providing evidence on a wide range of scientific and policy-focussed issues and the pace at which researchers have worked since the emergence of SARS-CoV-2 in early 2020. The respondents to our survey and our interviewees confirmed that modelling informed decision-making at key stages of the unfolding COVID-19 epidemics in the UK, although it was a subset of the evidence considered by SAGE and decision-makers. Similarly, our interviewees highlighted a largely positive interest from the British media and public in understanding the mechanisms of mathematical modelling and how to interpret the results.

Not unexpectedly given the pressures on everyone, respondents and interviewees indicated that communication between communities could be improved, but multiple competing factors must be balanced. Many survey respondents suggested that more direct communication with decision-makers would benefit both parties, with some stating that decision-makers should attend SAGE meetings to hear the evidence debated. However, many survey respondents and interviewees sought to make the distinction between the roles of scientists and decision-makers clear. Throughout the pandemic, the degree of separation between these roles has varied [415]. For example, the news that unelected UK government officials had attended SAGE meetings proved controversial, with some questioning the independence of scientific advice [416, 417]. On the contrary, minutes from the 58th SAGE meeting on 21 September 2020 [418] presented five options, including a circuit-breaker lockdown, for the "package of interventions [that] will need to be adopted to reverse this exponential rise in cases", which was not actioned by decision-makers at that time. Ultimately, understanding this distinction is crucial for the public, to whom the decision-makers are accountable.

Communication of uncertainty to non-modellers was also a key area identified by survey respondents and interviewees as requiring improvement. While uncertainty is inherent in all mathematical and statistical models, how to convey this clearly to non-technical audiences remains an active area of research [409, 419], particularly in epidemiology [408]. Decision-makers need to judge the likelihood of different scenarios and their potential consequences ahead of implementing policy, especially policies with as wide-ranging impacts as national lockdowns. The rapidly evolving pandemic often necessitated fast decisions and communicating the intricacies underpinning the uncertainty in model results can be time-consuming. Methods for conveying uncertainty fall outside the scope of this article. However, some respondents alluded to improving the general scientific understanding of decision-makers, with Professor

Charles Bangham of Imperial College London, noting the example of the Vice-President of Taiwan being a professor of epidemiology. Additionally, Professor Gabriel Leung, senior author of [364], was previously Under Secretary for Food and Health in Hong Kong [420]. However, such examples could further cloud the distinction between scientists and decision-makers. The communication of uncertainty to the public requires a careful balancing of transparency and interpretability. As noted by several survey respondents, this should be developed between emergencies to facilitate effective communication of uncertainty during emergency situations.

Vast quantities of new information about COVID-19 become available daily across the world. While this improves our collective understanding of the virus, this also puts unprecedented pressure on scientists and decision-makers. Professor Bauld noted a "huge pressure to get it [science communication] right in the face of changing evidence". In a dynamic situation such as this one, it is critical that both parties acknowledge the implications of new evidence and use it to inform their actions. This also presents a challenge for the media in communicating this to the public in an understandable and coherent manner, but it is critical to any emergency response that society fosters an open-minded attitude to updated scientific evidence over time. The importance of doing so was underlined by Dr Pellis, Dr Ian Hall and Professor Medley in the BBC Documentary "Lockdown 1.0 — Following the science?" [421], in which it is stated that revised estimates of the epidemic doubling time in March 2020 and the consequences this would have on hospital capacity were presented to SPI-M and subsequently passed onto SAGE, leading to "the cascade of full lockdown". Although our knowledge of the virus is continuously improving, there are still a great many unknowns about the transmission of SARS-CoV-2. Professor Beddington highlighted the equal importance of culture in which scientists underline that a particular fact is unknown in addition to updating scientific evidence over time, something which he believes Professor Chris Whitty, Chief Medical Officer and Sir Patrick Vallance have done particularly well throughout the pandemic. The former was echoed by Fox, Tim Fielden and Professor Godfray in a discussion hosted by the Oxford Martin School in February 2021 [422].

Our literature review was only able to encompass a small subset of modelling studies released since the beginning of the pandemic, given the extremely large number published over the course of 2020. However, our review informed our wider discussion on modelling rather than constituting a systematic modelling review. In this special case of the COVID-19 pandemic, we included studies which are not peer reviewed. Many scientists have published preprints during the COVID-19 pandemic to ensure fast and unrestricted access to important new information in a rapidly evolving situation. Fox noted that while the "justification for this practice is clear", she hopes that the scientific community return to the peer

review system after the emergency phase of the pandemic is over.

Many of findings in this article are based on the experiences and opinions of relevant parties (based primarily in the UK), gathered by the authors. Although almost all those invited for the interview took part, most of those invited to take part in the survey did not. This is frequently the case in research studies, but caution is needed to avoid assuming that those who responded are representative of those who did not, or that these findings are applicable to settings outside of the UK. While we intended for our survey questions to be as open-ended as possible, we appreciate that some wording may have lacked clarity and thus the interpretation of results may be subjective. Similarly, interviewees were not asked identical questions, as each interview was shaped by the interviewee. Furthermore, some respondents of the survey believed that our research was being conducted prematurely. One respondent wrote: "I write about the COVID SAGE in the present tense because this is not over by a long way and I just think it's far too early to be doing stuff like this", with another adding that "I'm feeling quite fatigued by the whole process at the moment and would need some more temporal distance from the whole process to be more objective about the interactions between modelling and policymaking". We acknowledge the extreme pressure that scientists and advisers have been under since the beginning of the outbreak in Wuhan and understand that the views given at this stage of the pandemic may change in the future.

The effects of the COVID-19 pandemic are likely to define many years to come. The loss of life and destruction of livelihoods across the world, either directly or indirectly attributable to COVID-19, is substantial and must never be omitted from any discussion of the pandemic. Scientists and decision-makers have been placed in an extraordinarily challenging situation. As the pandemic continues, we are reminded of the responsibilities of both parties with respect to mathematical modelling and public health. Scientists have to ensure that their models are of the highest scientific quality while also acknowledging their inherent limitations and the urgent need for results, and decision-makers have to understand the nuances underlying results and consider outputs in conjunction with all other evidence. We hope that our article contributes to an important discussion of the interplay between these parties and highlights matters to consider ahead of the next emergency.

Chapter 4

Paper II: Communicating uncertainty in epidemic models

R McCabe, MD Kont, N Schmit, C Whittaker, A Løchen, PGT Walker, AC Ghani, NM Ferguson, PJ White, CA Donnelly and OJ Watson. Communicating uncertainty in epidemic models. *Epidemics*, 37:100520, 2021. doi: https://doi.org/10.1016/j.epidem.2021.100520.

Abstract

While mathematical models of disease transmission are widely used to inform public health decision-makers globally, the uncertainty inherent in results are often poorly communicated. We outline some potential sources of uncertainty in epidemic models, present traditional methods used to illustrate uncertainty and discuss alternative presentation formats used by modelling groups throughout the COVID-19 pandemic. Then, by drawing on the experience of our own recent modelling, we seek to contribute to the ongoing discussion of how to improve upon traditional methods used to visualise uncertainty by providing a suggestion of how this can be presented in a clear and simple manner.

Author contributions

- **Ruth McCabe: Formal analysis, Writing – original draft, Writing – review & editing.**
- Mara D. Kont: Writing – review & editing.
- Nora Schmit: Writing – review & editing.
- Charles Whittaker: Writing – review & editing.
- Alessandra Løchen: Writing – review & editing.

- Patrick G.T. Walker: Writing – review & editing.
- Azra C. Ghani: Writing – review & editing.
- Neil M. Ferguson: Writing – review & editing.
- Peter J. White: Writing – review & editing.
- Christl A. Donnelly: Writing – review & editing.
- Oliver J. Watson: Conceptualization, Formal analysis, Writing – original draft, Writing – review & editing.

Reproducibility

Model fitting and analysis is performed in R. Data and code to reproduce figures are available at `https://github.com/j-idea/europe-icu-capacity`.

Appendix

Further information relating to this paper can be found in Appendix B.

Communicating uncertainty in epidemic models

During a public health crisis, such as an outbreak of a novel pathogen, decision-makers are often faced with making difficult and rapid decisions in the face of uncertainty [409]. Mathematical models of disease transmission are one piece of evidence widely used to inform public health decision-makers globally, but, like all models, their outputs are uncertain [423, 424]. Communicating this uncertainty during the COVID-19 pandemic in a clear and understandable manner to decision-makers is vital, but uncertainty has been often poorly reflected in the visualisations presented [408]. The most important considerations when deciding on a data visualisation is knowing who the audiences are and ensuring that key messages can be easily and quickly absorbed [409]. The pandemic has produced numerous examples of where statistics have been misunderstood and efforts to inform have ultimately confused audiences [425]. One lesson that must be learnt from the pandemic is more effective ways to communicate quantitative findings and their uncertainty. Drawing on the experience of our own recent COVID-19 modelling, we seek to contribute to the ongoing discussion of how to improve upon traditional methods used to visualise uncertainty by providing suggestions of how this can be presented in a clear and simple manner.

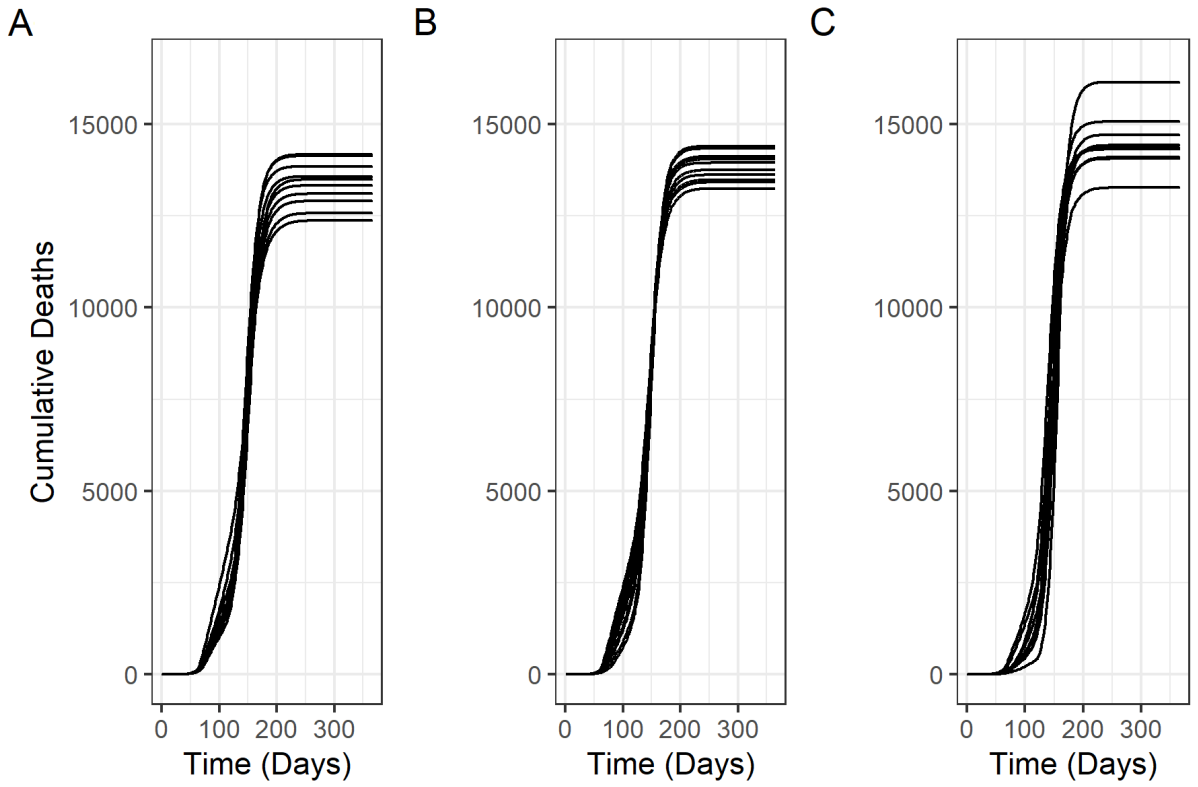

Figure 4.1: Uncertainty in epidemic models arising from different sources. Lines correspond to ten individual simulations of daily deaths over a one-year period using (A) a deterministic model with different sets of parameters per simulation; (B) a stochastic model with fixed parameters across simulations; and (C) a stochastic model with different sets of parameters per simulation.

The transmission of infections through a population is intrinsically complex and not directly observable. Transmission is often observed indirectly through recording cases of illness or death, or testing individuals for antibodies indicating history of previous infection. Mathematical models represent these processes by classifying the population into distinct states of infection and disease progression, with transitions between states governed by various epidemiological parameters. There are multiple parameter estimates and each is subject to uncertainty, so many sets of plausible combinations of parameter values exist. This uncertainty can be explored by producing multiple realisations of the model using different parameter sets. Uncertainty can also be introduced into simulations via the use of stochastic models, which, unlike deterministic models, incorporate the effects of random chance and are inherently "noisy". The choice of model structure and parameterisation depends on the questions under consideration and often on the stage of the epidemic. For example, stochastic models were commonly deployed at the beginning of the COVID-19 pandemic, at which point there was limited understanding of the transmission dynamics of SARS-CoV-2 [268, 333].

Figure 4.1 presents an illustration of ten simulation trajectories of daily deaths under three adaptations of a previously published model [2] in which uncertainty arises from different sources. Figure 4.1A displays output from a deterministic model in which the uncertainty is driven by varying model parameters, whereas Figure 4.1B presents the results of a stochastic model in which the parameters are held constant for each realisation. Finally, Figure 4.1C presents a stochastic model in which model parameters are also varied. The combination of two sources of uncertainty lead to much greater variation in trajectories than that which is observed under the models with a single source of uncertainty. Regardless of the source and magnitude of the uncertainty in epidemic models, it is critical that uncertainty is always considered carefully in the presentation of any results. Additionally, depending on the audience and hypotheses under investigation, it could also aid the interpretation of results if the uncertainties in model inputs are also presented alongside model outputs.

Results of epidemic models are most frequently summarised using simple summary statistics, such as the median and interquartile range. Drawing on our recent COVID-19 modelling work investigating intensive care unit (ICU) spare capacity in Europe [57], we present an example of such an illustration in Figure 4.2. Data are simulated daily hospital and ICU demand from COVID-19 patients and COVID-19 deaths using a stochastic age-dependent susceptible-exposed-infected-recovered (SEIR) model of SARS-CoV-2 transmission. The model was fit to COVID-19 deaths and ICU demand, with 100 simulation trajectories presented. Simulation realisations were sampled from the posterior parameter space from model fitting and used to provide scenario projections under two hypothetical intervention policies: 1)

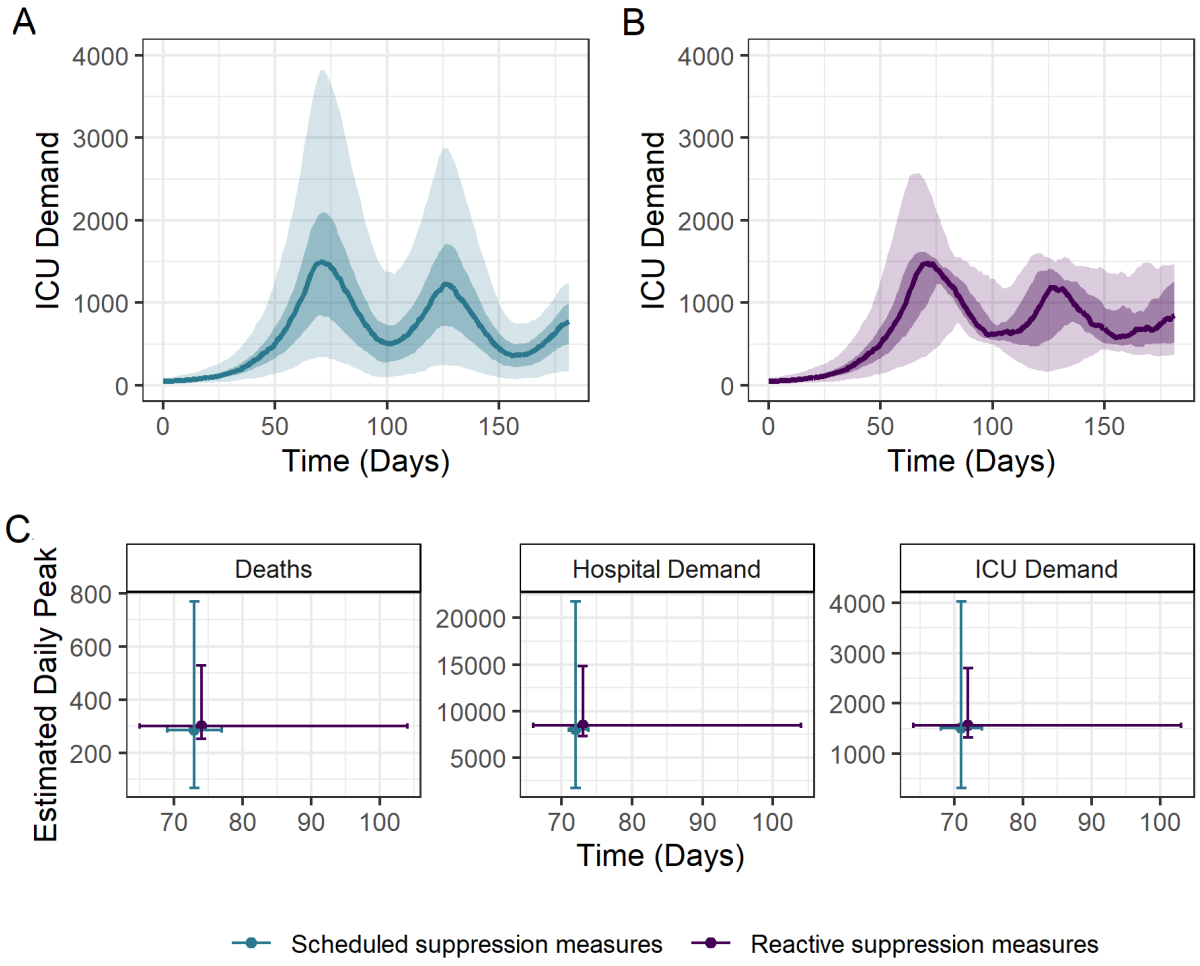

Figure 4.2: Traditional presentation of epidemic projections under two suppression strategies. The median (line) with 50% credible intervals (dark shading) and 95% credible intervals (light shading) of 100 simulations of intensive care unit (ICU) demand from COVID-19 patients is shown for strategies in which suppression measures are (A) scheduled at a fixed time point and (B) triggered reactively. The estimated median daily peak in deaths, hospital demand and ICU demand and the estimated time of each peak with 95% quantile ranges are shown under the two suppression scenarios in (C).

suppression measures are introduced at scheduled points in time and 2) suppression measures are triggered based on the number of COVID-19 patients in ICUs. Trajectories are summarised using the median with 50% and 95% credible intervals (calculated as 25th and 75th and 2.5th and 97.5th centiles, respectively) at each time point of the simulation. Additionally, we have provided two metrics of importance to decision-makers: the timing and the magnitude of peak ICU bed demand per simulation, presented as point estimates and 95% credible intervals. The median trajectories are very similar in both scenarios (Figure 4.2A and 4.2B; Figure B.1), as are the point estimates of the magnitude and timing of each peak (Figure 4.2C). However, under the "scheduled" strategy there is greater uncertainty in the magnitude than the timing of the first peak whereas, the alternative "reactive" strategy has an additional layer of uncertainty, with there being heterogeneity in both the magnitude and timing of peaks (Figure 4.2C).

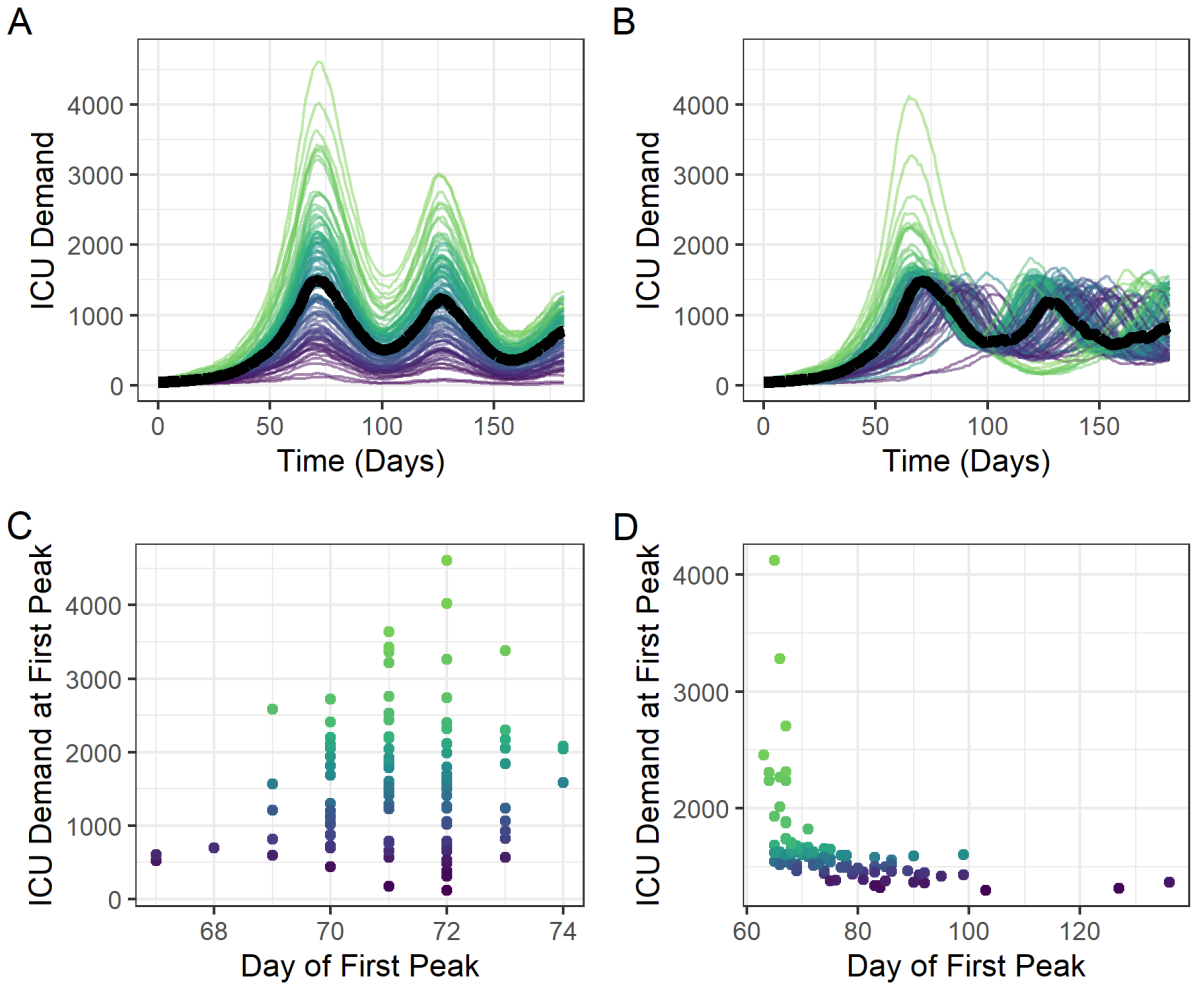

Figure 4.3: Communicating uncertainty in epidemic projections of ICU demand under two suppression strategies. Individual simulations of intensive care unit (ICU) demand from COVID-19 patients is shown for strategies in which suppression measures are (A) scheduled at a fixed time point and (B) triggered reactively, with corresponding relationships between the magnitude and timing of peak ICU demand for each realisation shown in (C) and (D), respectively. In (A) and (B) the median trajectory is shown with a black line. The colour of each simulation realisation depicts the ranking of ICU demand at first peak [high in green to low in purple] and allows for the metrics in (C) and (D) to be more easily linked to the trajectories in (A) and (B) respectively.

This allows decision-makers at a national-level to consider the possible burden different policies would place on the healthcare system, while also providing information to decision-makers at a hospital-level about the potential level of additional resources required to meet the projected surge in demand, conditional on the policy choice.

However, such aggregation often hides the nuances of individual trajectories as well as important features of the epidemic. For example, the typical peak number of infections across the set of individual trajectories may be drastically underestimated if results for multiple different, asynchronous epidemic trajectories are crudely summarised. We have encountered such issues in our work [57] (Figure 4.3) and hypothesise

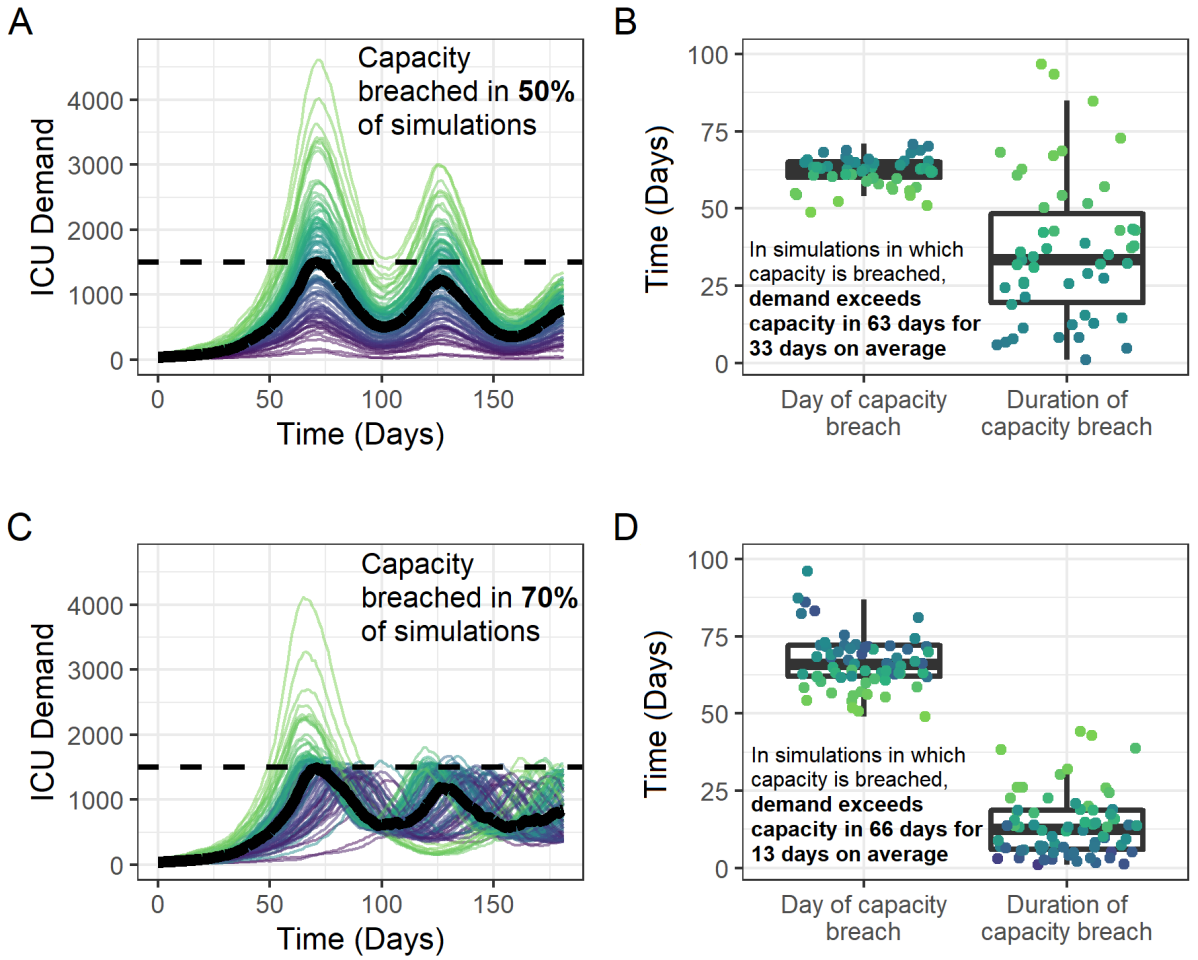

Figure 4.4: Linking epidemic projections of ICU demand under two suppression strategies to potential capacity breaches. Individual simulations of intensive care unit (ICU) demand from COVID-19 patients compared to an assumed ICU capacity of $1,500$ available beds is shown for strategies in which suppression measures are (A) scheduled at a fixed time point and (C) triggered reactively. In simulations in which capacity is breached, corresponding metrics of the day on which capacity is breached and the duration of the capacity breach shown in (B) and (D), respectively. In (A) and (C) the median trajectory is shown with a black line. The colour of each simulation realisation depicts the ranking of ICU demand at first peak [high in green to low in purple] and allows for the metrics in (B) and (D) to be more easily linked to the trajectories in (A) and (C) respectively.

that this is also true of other research groups who have chosen similar aggregation methods given the underlying dynamics of compartmental transmission models [161, 336, 382]. We have observed that models in which acquisition of immunity has a substantial impact on transmission dynamics or in which interventions are triggered when a threshold burden of infection is reached are particularly prone to producing asynchronous trajectories that are difficult to summarise in ways that do not obscure relevant properties of individual trajectories. These effects are often amplified when using stochastic models (Figure 4.1).

Alternative means of communicating uncertainty have been explored throughout the COVID-19 pan-

demic. For example Juul et al. [408], use central ranking systems which consider entire epidemic curves in relation to one another to produce curve boxplots and compare the likelihood of individual simulations in relation to key metrics of interest (i.e. daily number of new cases requiring hospitalisations). On the other hand Davies et al. [333], plot a subset of "representative" curves drawn from each decile corresponding to the total number of simulated new cases. Additionally Koo et al. [363], show each realisation of their simulations and use darker colours to indicate those which are closer to the median. While these approaches enable variation in epidemic trajectories to be communicated, we believe that sometimes the interpretability of the results to non-technical audiences is hindered. We argue that focussing directly on the metrics that inform policy decisions and making judicious choices of what quantities are plotted allows results, both in terms of central estimates and uncertainty to be communicated accurately, whilst ensuring these results are accessible to non-technical audiences.

In Figure 4.3, we instead present the individual trajectories of ICU demand over the simulation period alongside individual metrics of magnitudes and peak timings. In an effort to engage the audience more with the uncertainty, the colour of each epidemic trajectory reflects the ranked magnitude of the first peak allowing the key metrics presented in Figure 4.3C and 4.3D to be more easily linked to the trajectory's epidemic dynamics in Figure 4.3A and 4.3B. Figure 4.3 paints a different picture of the potential unfolding epidemics than that provided via the simple summary statistics in Figure 4.2. This is particularly applicable to the "reactive" strategy, which observes a high degree of variability across trajectories which is not clearly conveyed using summary statistics calculated at each time point. The individual metrics of magnitude and peak timing allow for a clearer interpretation of the necessary support that national-level decision-makers ought to allow hospital-level decision-makers to prepare for a surge in demand, depending on the NPI strategy they implement. For example, the "scheduled" strategy suggests public health action should focus on providing additional hospital beds (Figure 4.3C) whereas the "reactive" strategy suggests that officials may also need to consider how quickly they need to procure additional hospital beds (Figure 4.3D).

To further aid public health officials, in Figure 4.4 we present our simulations in conjunction with measures of ICU capacity. Under the assumption of $1,500$ available beds, 50% of the simulations from the "scheduled" strategy estimate that capacity will be breached in comparison to 70% of simulations from "reactive" strategy (Figure 4.4A and 4.4C). In Figures 4.4B and 4.4D, the first day and duration of capacity breaches are presented individually for those simulations in which capacity is exceeded, again using colour to link to the individual trajectories as in Figure 4.3. In each plot, we have used text to provide brief, key messages to the audience without obscuring the full range of uncertainty underlying

each message. Together, this provides a more detailed insight into the range of scenarios that ICUs may be faced with than that which could be gained from solely looking at the individual trajectories and may be particularly useful to national-level decision-makers in deciding which NPI strategy to implement. Similar to the timing of the first peak (Figures 4.3C and 4.3D), there is greater variability in the first day of a possible capacity breach under the "reactive" strategy compared to the "scheduled" strategy. However, although more simulations breach capacity under the "reactive" strategy, the duration of the breach is anticipated to be shorter compared to the "scheduled" strategy given the model assumptions. Figures B.2 and B.3 are analogous to Figures 4.3 and Figure 4.4 but instead consider general hospital demand.

We believe these simple visualisations allow for the key metrics to be more easily understood while being transparent about the uncertainty in our epidemic trajectories. Although we have chosen in this instance to focus on the dynamics of ICU demand at the first peak, understanding the audience is still paramount and it is imperative that visualisations are tailored as such. For example, decision-makers may be more interested in the outcomes of specific interventions and contrasting alternative scenarios. This is where interactive tools become increasingly valuable and allow users to explore scenarios specific to their needs, such as surrounding COVID-19 vaccine allocation strategies COVID-19 Scenario Analysis Tool [426]. However, greater flexibility introduces further complexities in striking the correct balance between informing and not overwhelming audiences in chosen data visualisations [427].

Although the clear communication of uncertainty was important long before the COVID-19 pandemic, the severity and rapidly-evolving nature of this emergency has underlined its importance. Brooks-Pollock et al. [43] note that "communication to non-specialists became an overnight skill required of disease modellers". Some scientific advisors in the UK have noted the improvement in both advisor's and decision-maker's understanding of mathematical modelling over the course of the pandemic, but as noted by other former advisors and current members of the UK Scientific Advisory Group for Emergencies (SAGE), the communication of uncertainty inherent in these models requires improvement [50]. The responsibility of communication has traditionally fallen solely on scientists, but it has been noted that developing the general scientific understanding of decision-makers could also ease the intense pressure placed upon the scientific advisors when responding to an emergency [50]. Additionally, it may be more appropriate to train intermediaries for specialist science communication and data visualisation roles in order to share the substantial burden that such activities place on both scientists' and decision-makers' time.

As the pandemic continues, mathematical models of transmission to inform policy tend to become increas-

ingly complex to account for vaccination and the emergence of variants of SARS-CoV-2 with different phenotypes while still modelling the impact of population behaviour, for example through the implementation of non-pharmaceutical interventions. Clear communication of uncertainty will remain important now and also in future health emergencies. We hope here to have added to a much-needed discussion about effective visualisation approaches for conveying uncertainty.

Chapter 5

Paper III: Public awareness of and opinions on the use of mathematical transmission modelling to inform public health policy in the United Kingdom

R McCabe and CA Donnelly. Public awareness of and opinions on the use of mathematical transmission modelling to inform public health policy in the United Kingdom. *Journal of the Royal Society Interface*, 20:20230456, 2023. doi: https://doi.org/10.1098/rsif.2023.0456.

Abstract

Mathematical modelling is used to inform public health policy, particularly so during the COVID-19 pandemic. As the public are key stakeholders, understanding the public perceptions of these tools is vital. To complement our previous study on the science-policy interface, novel survey data were collected via an online panel ("representative" sample) and via social media ("non-probability" sample). Many questions were asked twice for the period "prior to" and again in reference to "during" the COVID-19 pandemic.

Respondents reported being increasingly aware of modelling in informing policy during the pandemic, with

higher levels of awareness among social media respondents. Modelling informing policy was perceived as more reliable during the pandemic than in reference to the pre-pandemic period in both samples. Trust in government public health advice remained high within both samples but was lower during the pandemic in comparison to the (retrospective) pre-pandemic period. The decay in trust was greater among social media respondents. Many respondents explicitly made the distinction that their trust was reserved for "scientists" and not "politicians". Almost all respondents believed governments have responsibility for communicating modelling to the public. These results provide a reminder of the skewed conclusions that could be drawn from non-representative samples.

Author contributions

RM and CAD designed the study. **RM** conducted the analysis and wrote the first draft of the manuscript. CAD edited the manuscript.

Ethics

This study has approval from the Medical Sciences Interdivisional Research Ethics Committee at the University of Oxford with Ethics Approval Reference R76166/RE001. Each participant provided consent to take part in the research via a mandatory question at the beginning of the survey at the end of the Participant Information Sheet, thus providing written consent.

Reproducibility

In line with the ethics approval obtained, raw research data is accessible to the Medical Sciences Interdivisional Research Ethics Committee at the University of Oxford, upon their request. Anonymised summary data are provided in Table 5.1. Code for R analyses are provided at: `https://zenodo.org/records/10201376` (doi: 10.5281/zenodo.10201375) [428].

Acknowledgements

The authors would like to thank a number of individuals who supported this analysis. First, to Professor Sally Sheard, Dr Hayley Mableson and Dr Michael Humann for their helpful comments when drafting the survey. Second, to all those who shared our survey via social media. Third, to Cathal Mills at the University of Oxford for his assistance with data visualisation. Last, but by no means least, the authors are incredibly grateful to all those who took the time to complete the study

Appendix

Further information relating to this paper can be found in Appendix C.

5.1 Introduction

Scientific advisory mechanisms in the United Kingdom of Great Britain and Northern Ireland (UK) were brought into the full view of the public eye during the COVID-19 pandemic. Restrictive physical distancing measures, such as lockdowns, affected all within society and, rightfully, sparked a widespread interest in the evidence underpinning such decision-making by the government. As society reflects on the events of the COVID-19 pandemic, particularly in anticipation of the outcome of the public inquiry into the UK government's handling of the pandemic [139], the use, or not, of scientific evidence to inform policy will be examined closely.

Mathematical transmission modelling is a key component of scientific evidence presented to the government during outbreaks of infectious diseases: the COVID-19 pandemic was no exception [50, 429]. Already, there is a growing body of work focusing on the use of this set of tools during the recent pandemic [43, 325, 430], including one by the authors of this study [50]. In our previously published study, we examined the interplay between mathematical transmission modelling, scientists, the media and the public throughout the first year of the pandemic [50]. Participants in our research, namely scientific advisors and public communication experts, indicated that direct communication with the public had been challenging. As the success of public health policy often depends upon the cooperation of the public, it is important that the public perception of scientific advice, in this case specifically mathematical modelling, is understood. Much of the research into public opinions throughout the COVID-19 pandemic has related to vaccine hesitancy [431–433] or adherence to public health and social measures [434–437]. Although Marshall et al. [438] interviewed members of the public regarding Test, Trace and Isolate policies with a view to improving assumptions underpinning mathematical modelling, respondents were not explicitly asked about their perceptions of modelling.

We sought to complement our initial study [50] by gathering and analysing data on the public's awareness of and opinions on the use of mathematical transmission modelling to inform public health policy in the UK via an online questionnaire covering topics such as means of awareness and responsibility of the communication of modelling, feelings of reliability of modelling in informing policy, and trust in government public health policy (Table 5.1). Even with sufficient financial resources, obtaining a "representative" sample of the population is notoriously difficult and thus the interpretation of the resulting data must be cautious regarding the opinions of the adult population in the UK. Consequently, we collected data via two sampling methods: via an online panel (using the platform Prolific Academic [439] and considered a "representative" sample in the context of this paper) and via social media (using Twitter [440] and considered a "non-probability" sample). As such, the goals of this study are two-fold: to better under-

stand public awareness of and opinions on the use of modelling to inform public health policy but also, critically, to explore and interpret any systematic differences between the two samples obtained (online panel and social media).

5.2 Materials and Methods

5.2.1 Survey design and distribution

This study has approval from the Medical Sciences Interdivisional Research Ethics Committee at the University of Oxford with Ethics Approval Reference R76166/RE001. The Participant Information Sheet (PIS), provided in Appendix C, sets out the instructions given to participants ahead of taking the survey.

We designed an online survey of 24 questions, comprising of 16 multiple-choice and 8 open-ended questions (Table 5.1). The survey was split into four sections of unequal length.

To begin, participants were asked to describe a transmission model and state the level of confidence that they have in their description. These were the only two compulsory questions in the survey. Once participants answered this question, a description of transmission models in the context of this survey was given so that answers given in the rest of the survey were relevant to the research questions at hand.

The middle two sections contained most questions which form the basis of the analysis in this paper. Participants were asked a series of questions about their awareness of and opinions on transmission modelling to inform policy before the COVID-19 pandemic. The next section then repeated these questions, also alongside some additional questions, but asked participants to answer in relation to the period during the COVID-19 pandemic. As questions about time periods were both responded to simultaneously (summer 2021), it is important to be aware of possible recall bias of participants in reference to the period prior to the pandemic.

Finally, participants were asked to provide demographic information, namely age group, gender, the sector that they work in and their COVID-19 vaccination status.

The survey was designed in Microsoft Forms. The questions are presented in full in Table 5.1.

Participants were gathered via two methods: via an online panel and via social media. Only individuals aged over 18 and residing in the UK were permitted to partake in the survey.

The online panel was surveyed via Prolific Academic, an online platform with a large, diverse database of individuals who have consented to participate in research on a multitude of different topics. Recruitment to Prolific Academic primarily relies on "word of mouth" [441]. The company actively advertises that "we currently offer representative samples of two national populations: the United Kingdom and the United States" [442], but also note that "no sample can ever be 100% representative, so using a representative sample does not guarantee that your results will be perfectly generalisable to the population (although it does help to improve generalisability).". Our survey was sent to a sample of approximately 500 people signed up to the platform, who were selected to target "representation of the UK population" across a range of demographic variables, including "Age", "Sex", "Country of Birth" and "Employment Status". As such, this sample is considered a "representative" sample within the context of this paper. Responses were gathered across 7 – 9 July 2021.

Alongside this, we shared our survey on the social media platform Twitter by tweeting a link to it, and asked collaborators and accounts associated with the departments to which the authors are affiliated to also do so. An overview of the accounts that tweeted or quote tweeted the survey is presented in Table C.1. Undoubtedly, a network effect will have influenced who saw the survey in their timeline and who will have selected to participate. Consequently, this sample is considered a non-probability sample. A priori, it was expected that this method of participant recruitment could substantially skew the demographics and answers of the participants within this sample. The survey was first shared on Twitter on 7 July 2021. Although responses were received until 3 September 2021, the vast majority were obtained between 7 – 20 July 2021.

5.2.2 Statistical analysis

Data were filtered so as to only include observations with all four demographic variables answered (age group; gender; sector; vaccination status). Throughout, percentages of respondents have been rounded to the nearest integer.

5.2.2.1 Classification of open-ended responses

Responses to three open-ended questions, namely "How would you describe a transmission model?"; "Before the COVID-19 pandemic, where do you think those who developed and used transmission models worked?" and "During the COVID-19, where do you think those who developed and used transmission models worked?", were categorised to aid the interpretation of the results. Categories were not predetermined and were refined having considered the entirety of the data.

Descriptions of transmission modelling were deemed to be more relevant to the context of our research in relation to infectious disease epidemiology and/or to mathematics and statistics more broadly. Each response was considered individually to assess the sentiment of the description. For example, the use of words such as "virus", "disease", "infection", "illness", "epidemic" or those akin to "COVID-19" (e.g. "coronavirus", "covid"), or others such as "statistics", "mathematical", "data", "prediction" or "simulation", were all interpreted as capturing the essence of "disease transmission and control modelling" and thus were considered as "more relevant", compared to "less relevant", to the topic at hand.

Answers to the open-ended question about where transmission modellers work (prior to and during the COVID-19 pandemic) were also assessed on an individual basis. If respondents stated multiple places of work, each were categorised individually. While "academia" and references to "university" were classified as "Research (Academia)", any other reference to "research" or "science" or "laboratories" were classed as "Research (Unspecified affiliation)". "Public health bodies" encompassed either direct reference to the term "public health" or specific organisations such as the World Health Organization or Public Health England (the United Kingdom Health Security Agency now combines elements of Public Health England and NHS Test and Trace [443]). Responses were classified as "Government" if either that word was used, a specific department was named, such as the "Department of Health" or reference was made to the "civil service". Any reference to "journalists" or "the press" or "the media" were clustered together under the heading "Media", but not those stating "Social media" (of which there were extremely few). "Healthcare services" encompassed reference to "health bodies" without direct mention of "public health". For example, this includes terms such as "hospitals", "the NHS" and "healthcare providers". References of "advisory" to government, or "advisers", including the use of "SAGE", was classified as "Advisory roles". All other responses were placed into "Other". In most instances, answers were classified in more than one category as they were not mutually exclusive.

Responses to the remaining open-ended questions were not explicitly categorised but, as above, were considered on an individual basis.

5.2.2.2 Answers to multiple-choice questions

Chi-squared tests were used to test for differences in the proportions of respondents selecting specific answers to multiple-choice questions across samples. Within respondents, Wilcoxon Signed Rank tests were used to assess the differences between time periods within respondents. Questions in which there was a discrete scale from one extreme to the other, for example: "Not at all confident" to "Very confident";

or "No trust whatsoever" to "High level of trust", were enumerated on a scale where 0 represented the middle strength option (e.g., "Moderately confident" or "Moderate level of trust"), with 1 assigned to the highest level (e.g. "Very confident" or "High level of trust") and -1 assigned to the lowest level (e.g. "Not at all confident" or "No trust whatsoever") for each time point. Cases with a binary outcome, for example, whether or not a respondent selected a specific means of being aware of transmission modelling (e.g. "Newspaper (online or print)"), were enumerated as 1 if this option was selected and 0 otherwise. Specific details for each question, alongside results tables, are presented in Appendix C.

5.2.2.3 Reliability scores

Answers to the question "On a scale from 1–10 with 1 being "extremely unreliable" and 10 being "extremely reliable", how do you feel about the use of transmission models in informing public health policy?", both "Prior to the COVID-19 pandemic" and "During the COVID-19 pandemic" are referred to as "reliability scores" throughout. This is the only question in which participants were asked for a numeric answer. Mann-Whitney U tests were used to assess differences in scores between samples within time periods, while Wilcoxon Signed Rank tests were used to assess differences in scores between time periods within respondents.

5.2.2.4 Regression analysis

Linear regression was used throughout the study to quantify the relationships between the responses to each multiple-choice question (outcome) and other variables (explanatory, primarily demographic variables, but also occasionally answers to other multiple-choice questions). For example, we regressed responses to the question "Were you aware of the use of transmission models in informing public health policy?", with the answer scale enumerated as described above, on age and sex (Supplementary Table 3). Due to the number of multiple-choice questions, many linear regression models were analysed with the specific details and results presented in full in Appendix C.

Throughout, identical regression models were fitted separately on data from the online panel and social media samples. Chi-squared tests were used to assess the overall significance of categorical variables in each model. In the main text, results of these tests are reported if they show statistical significance at the 0.01 level.

When assessing the relationship between demographic and other variables, due to the large number of employment categories and limited number of respondents selecting that they are unvaccinated, only age group and gender were included within regression models. The youngest age group, 18–25 years, and

females, the most frequently reported gender, were designated as the reference levels.

5.3 Results

Participants totalled 509 and 207 across the online panel and social media samples, respectively (total participants 716). Of these, 504 (99%) and 202 (98%) had "complete" demographic data and thus were included in the analysis (total participants 706). Table 5.1 includes the results of the multiple-choice questions by sample. Response rates to multiple-choice question were high in both samples, with at least 98% of respondents with "complete" demographic data answering each question.

Aside from the demographics of the respondents, the results are discussed with respect to four main themes: awareness, reliability, communication and trust in government advice. Extensive further analyses are presented in Appendix C.

5.3.1 Respondent demographics

Gender had similar distributions across samples (Table 5.1; Figure 5.1A), with females holding a slight majority (52% and 51% for the online panel and social media, respectively).

There was heterogeneity across the samples in terms of age group and employment sector (Table 5.1; Figure 5.1C; Figure 5.1D). Age groups of the online panel respondents were approximately evenly distributed except for the oldest age group (66+ years) in which there were substantially fewer participants. In contrast, the number in each age group of the social media sample steadily increased up to 56–65 years, which had a notably larger percentage of participants than other groups (40%), before also falling in the 66+ category. While half of the employment sector options had fairly consistent percentages of participants working in them across samples, (namely: "Agriculture", "Business, finance and technology"; "Health and social work"; "Production (energy, manufacturing, mining)"; "Public"), there were fairly large differences elsewhere. For example, social media respondents were significantly more likely than their online panel counterparts to work in "Education" (chi-squared test: $\chi^2_{df=1} = 13.39$; $n = 706$; $p < 0.001$) or "Research" (chi-squared test: $\chi^2_{df=1} = 12.62$; $n = 706$; $p < 0.001$). This is likely to be attributable to the fact that it was predominantly researchers or research institutions sharing the survey (Table C.1).

Social media respondents were also significantly more likely to be vaccinated against COVID-19 than online panel respondents (97% and 87%, respectively; Figure 5.1; chi-squared test: $\chi^2_{df=1} = 13.00$; $n = 706$; $p < 0.001$). However, similar percentage of participants in both samples recorded having no intention of

being vaccinated (5% and 3%, respectively; chi-squared test: $\chi^2_{df=1} = 0.74$; $n = 706$; $p = 0.390$), with the remaining differences explained by those indicating their intention to be vaccinated in the future.

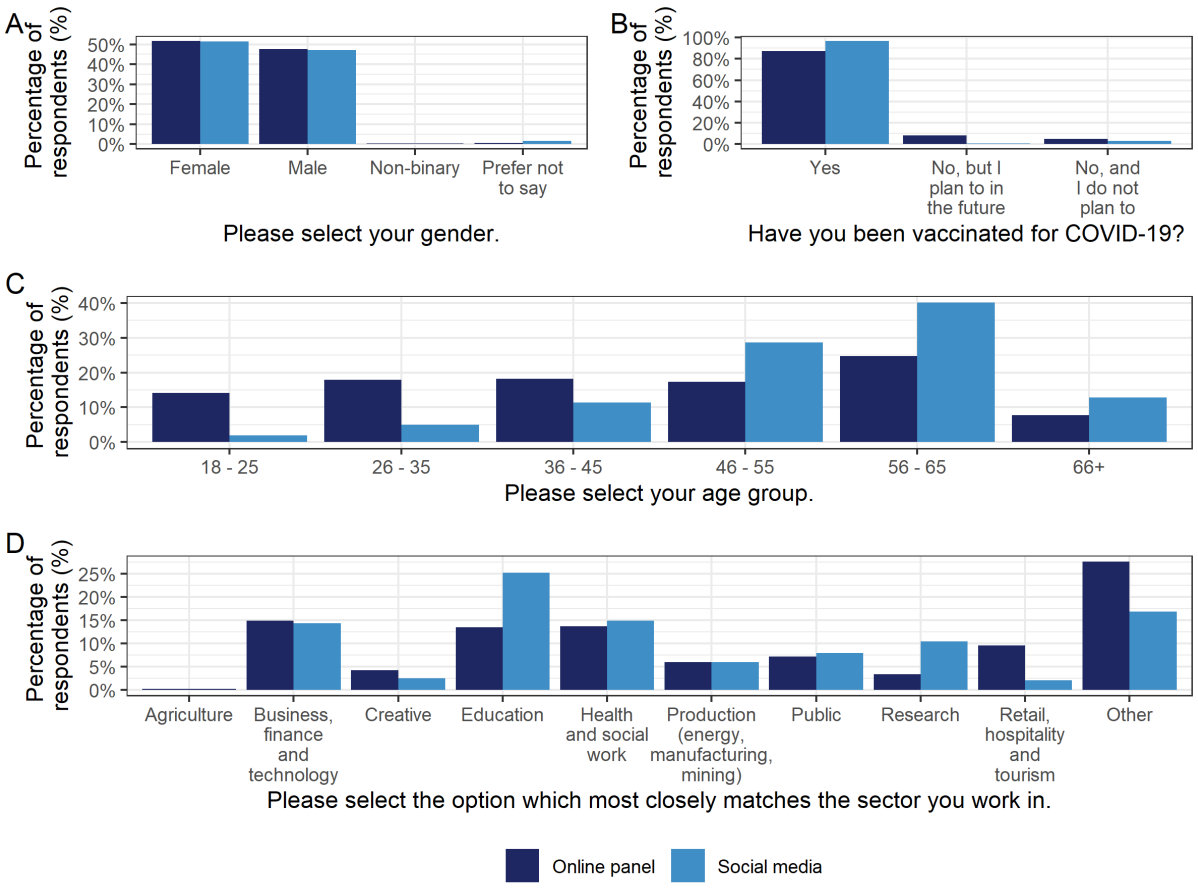

Figure 5.1: Demographics of survey participants stratified by sample. (A) Gender. (B) COVID-19 vaccination status. (C) Age group. (D) Sector participants work in. Underlying data are presented in Table 5.1.

5.3.2 Awareness of mathematical transmission modelling

Within samples and time periods, levels of awareness of mathematical transmission modelling (herein referred to as "modelling") generally were indistinguishable from the levels of awareness of modelling being used in policy specifically (Table C.2). An exception was online panel respondents reporting being significantly less aware of modelling being used for policy than modelling in general in reference to the pre-pandemic period (Wilcoxon signed rank test: $p < 0.001$).

The majority of participants across both samples were aware of modelling being used to inform public health policy in the period of the pandemic (Table 5.1; Figure 5.2), with these levels being significantly higher than those reported for the pre-pandemic period (Online panel: 27% pre-pandemic, 74% during; Wilcoxon signed rank test $p < 0.001$; Social media: 57% pre-pandemic, 97% during; Wilcoxon signed

rank test $p < 0.001$). However, social media respondents were significantly more likely to be aware of modelling informing policy than online panel respondents regarding each time period (chi-squared test: prior: $\chi^2_{df=1} = 56.05$; $n = 706$; $p < 0.001$; during: $\chi^2_{df=1} = 44.55$; $n = 706$; $p < 0.001$).

Online panel respondents were primarily aware of modelling through "News show (TV or online)" or "Newspapers (online or print)" across both time periods (Figure C.1; Table C.4). However, social media respondents were primarily aware of modelling through "Academic reports and papers" prior to the COVID-19 pandemic, but "News show (TV or online)", "Newspapers (online or print)" and "Social media" were the most popular selections during the pandemic. For both time periods, social media respondents were significantly more likely than the online panel to select "Academic reports and papers" (Table C.5; chi-squared test: prior: $p < 0.001$; during: $p < 0.001$), which could be driven in part by the significantly greater percentage of respondents from this sample working in the "Education" and "Research" sectors (Table 5.1; Table C.1). In reference to the period prior to the pandemic, respondents from both samples were similarly likely to select "Social media", but in the period of the pandemic social media respondents were significantly more likely to select this option than their online panel counterparts (Table C.5; chi-squared test: $p < 0.001$).

Most participants stated that they had an "About right" level of information of modelling during the pandemic with respondents from both samples being similarly likely to select this response (51% and 60% of the online panel and social media respondents, respectively; chi-squared test: $\chi^2_{df=1} = 4.76$; $n = 706$; $p = 0.029$). Some of those selecting the option having an "About right" amount of information also selected that they were unaware of modelling, demonstrating that not all have interest in this topic, in both samples (Figure C.3; Table C.7).

5.3.3 Reliability of mathematical transmission modelling in informing policy

Online panel and social media respondents reported similar mean reliability scores of 6.3 (standard deviation (sd) 1.8) and 6.4 (sd 1.9), respectively, in reference to the pre-pandemic period (Mann-Whitney U test: $p = 0.493$; Figure C.4). During the pandemic, mean reliability scores rose significantly to 6.9 (sd 2.0) among online panel respondents (Wilcoxon Signed Rank test: $p < 0.001$), but rose similarly among the social media respondents (mean 6.7 (sd 2.3); although this increase was not significant, Wilcoxon Signed Rank test: $p = 0.204$, largely due to the more limited sample size ($n = 202$ for social media compared to $n = 504$ in the online panel)). There was no evidence to suggest significant differences between mean reliability scores across the two samples during the pandemic (Mann-Whitney U test: $p = 0.613$).

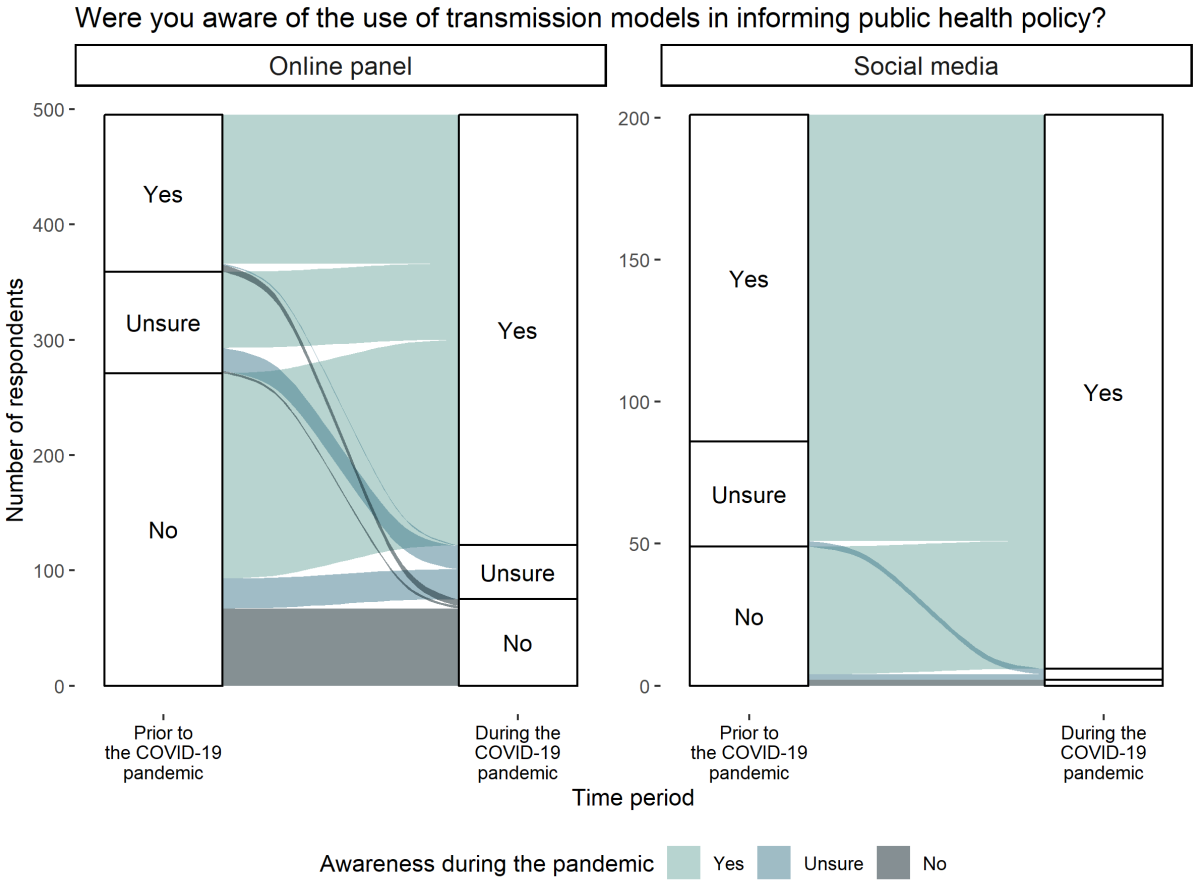

Figure 5.2: Awareness of modelling informing public health policy. Responses to the question "Were you aware of the use of transmission models in informing public health policy?" both "Prior to" (reported retrospectively) and "During" the COVID-19 pandemic for the online panel and social media samples. Observations corresponding to a respondent who did not answer the question for at least one time point were removed from the figure to aid visual presentation (n=8 online panel; n=1 social media). Due to the vast majority (97%) of social media respondents reporting being aware of transmission modelling during the COVID-19 pandemic, the options of "Unsure" and "No" are unable to be labelled but appear in the same order as in the three preceding response pillars with "Unsure" above "No". Underlying data are presented in Table 5.1.

A majority of respondents changed their reliability score in reference to different time periods, with respondents from each sample being similarly likely to do so (Figure 5.3; Table C.8; online panel 55%; social media 65%; chi-squared test: $\chi^2_{df=1} = 4.96$; $n = 706$; $p = 0.026$). Respondents were similarly more likely to perceive greater reliability during the pandemic compared to before, regardless of sample (40% and 39% for online panel and social media, respectively; $\chi^2_{df=1} = 0.04$; $n = 706$; $p = 0.840$). However, social media respondents were around 70% more likely to have a decreased feeling of reliability compared to their online panel counterparts (26% and 15% for social media and online panel, respectively; $\chi^2_{df=1} = 10.34$; $n = 706$; $p = 0.001$).

During the COVID-19 pandemic, being male was significantly associated with lower reliability scores

compared to females, within the social media sample only (Table C.9; chi-squared test: $p = 0.001$). Awareness was not associated with reliability score within this sample during the pandemic ($p = 0.817$). In contrast, within the online panel in reference to both time points, increased awareness of the use of modelling in informing policy was associated with a significant increase in feeling of reliability rather than the demographic variables (Table C.10; online panel and pre-pandemic: $p < 0.001$; online panel and during: $p < 0.001$).

5.3.4 Responsibility of communicating mathematical transmission modelling to the public

Almost all respondents believed that governments have the responsibility of ensuring that the public are informed about the use of modelling in informing policy decisions (92% of online panel respondents; 91% of social media respondents; Table 5.1; Figure C.10; chi-squared test: $\chi^2_{df=1} = 0.10$; $n = 706$; $p = 0.756$). Social media respondents were significantly more likely to select those who "develop" or "use" transmission models, "The media", and "Myself, as a member of the public" as having the responsibility than online panel respondents (Table C.13). Except for "Myself, as a member of the public", the three aforementioned responses were equally popular within each sample (Table C.13).

5.3.5 Trust in government advice

Most respondents reported having "moderate" trust in the government regarding public health issues, regardless of sample or time period (Table 5.1; Figure 5.4A). Particularly in reference to the pre-pandemic period, many respondents stated their belief that the government had the best interest of the public at heart, with one stating "I didn't have [a] reason not to" (online panel). Within both samples, respondents noted that they had a high level of trust in scientific evidence rather than the government: "I trusted government advice because this is provided by scientific advisors" (online panel) and "I trusted the scientist[s] not the politicians" (social media), acknowledging that decisions taken by government can be influenced by more than just science: "It can be difficult to disentangle the science from politics" (social media). The perception of the government not trusting scientific advice was frequently mentioned as a reason for a lower level of trust, for example: "They didn't always listen to scientists, which deteriorated my trust" (online panel). These views were reflected in the reliability scores with a higher level of trust in government advice being associated with higher median reliability scores for both samples. However, there was substantially greater variance in reliability scores among those with lower trust levels. (Figure C.12; Table C.14).

Most social media respondents reported a different level of trust in government public health advice

regarding the period of the pandemic compared to the pre-pandemic period (54%), significantly more than the minority of online panel respondents also doing so (35%) (chi-squared test: $\chi^2_{df=1} = 22.51$; $n = 706$; $p < 0.001$) (Table C.15). In particular, the percentage of respondents selecting having "No trust whatsoever" rose significantly during the pandemic from 10% (both samples) to 25% and 38% for online panel and social media respondents, respectively (Wilcoxon signed rank test: Online panel: $p < 0.001$; Social media: $p < 0.001$). 26% of social media respondents and 18% of online panel respondents went from a "High level of trust" to "No trust whatsoever" during the pandemic (chi-squared test: $\chi^2_{df=1} = 1.48$; $n = 706$; $p = 0.223$). Reasons given for this included: "Felt that they [the government] were more driven by political and economic priorities rather than public health." (online panel) and "Not convinced they followed the science on every occasion as they stated they were doing." (social media). Separately, another respondent noted: "My trust has reduced but not yet to the position of no trust" (online panel). Together, this demonstrates an erosion of trust among respondents, but substantially more so among the social media sample.

One respondent who went from "High" to "No" trust gave the following reason for their opinion: "Government kept changing the advice given during the pandemic." (online panel). However, when asked directly about trust in relation to the government changing their advice based on scientific evidence, a minority of respondents from both samples stated consequently having "Less trust" (Table 5.1; Figure 5.4B; 13% and 10% for online panel and social media, respectively; chi-squared test: $\chi^2_{df=1} = 1.69$; $n = 706$; $p = 0.194$). Among those stating they find changing advice "Less trustworthy", respondents noted: "It makes me feel as though the government make up things to give the public false hope until there is scientific research available to support" (online panel) and "Gets too confusing [with] mixed messages" (social media). Others among the online panel sample suggested that it was the rapid pace at which advice changed which influenced their level of trust: "the advice changed so frequently it was hard to know what to believe was the right thing to do".

Social media respondents (56%) were (borderline) "More trusting" of changing advice based on new evidence than their online panel counterparts (44%) (chi-squared test: $\chi^2_{df=1} = 6.70$; $n = 706$; $p = 0.010$), although this was still the most common response among both samples. Respondents noted: "It shows a willingness to address a changing situation" (online panel) and "During a pandemic there will be need to amend advice as more data becomes available and more [is] known about the virus" (social media). Across both samples, it was clear that respondents placed a high value on understanding the source of the new information ("More trust as long as sources are quoted" (online panel)) and the method in which it was explained to the public ("I understand that evidence as it gathered can lead to new understand-

ing and therefore improved modelling. However, it can be dependent on how well it is explained and communicated." (social media)).

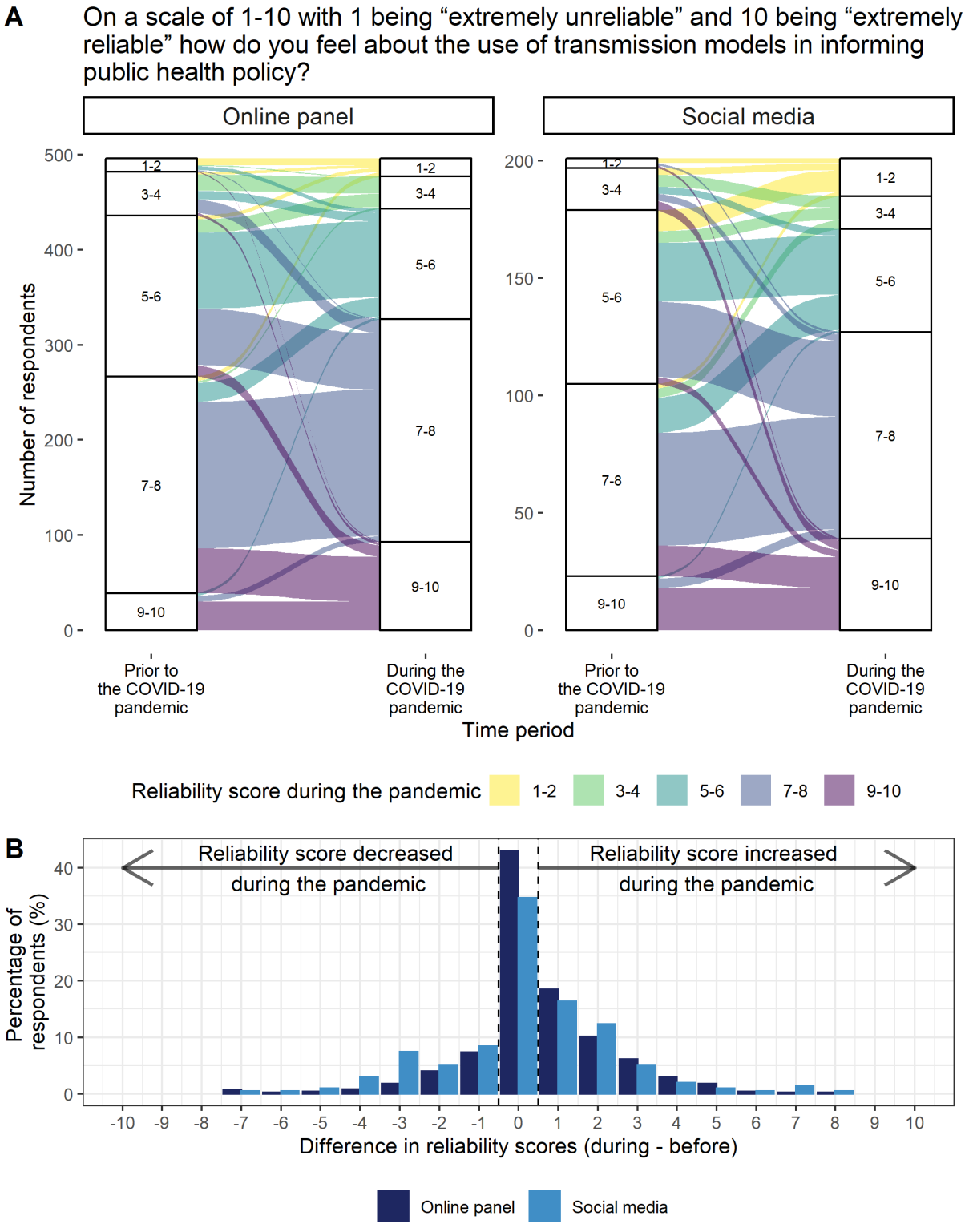

Figure 5.3: Reliability of using transmission modelling to inform public health policy. (A) Responses to the question "On a scale of 1 – 10 with 1 being "extremely unreliable" and 10 being "extremely reliable" how do you feel about the use of transmission modelling in informing public health policy?" both "Prior to" (reported retrospectively) and "During" the COVID-19 pandemic, for both the online panel and social media samples. Responses were categorised into groups of 2 (from 1–2) to (9–10) and observations corresponding to a respondent which did not answer the question for at least one time point were removed from the figure to aid visual presentation ($n = 8$ online panel; $n = 1$ social media). Underlying data are presented in Table 5.1. (B) Distribution of the within-respondent differences in reliability scores from during the pandemic compared to prior for both the online panel and social media samples. Difference in scores is defined as the reliability score during the pandemic minus reliability score prior to the pandemic.

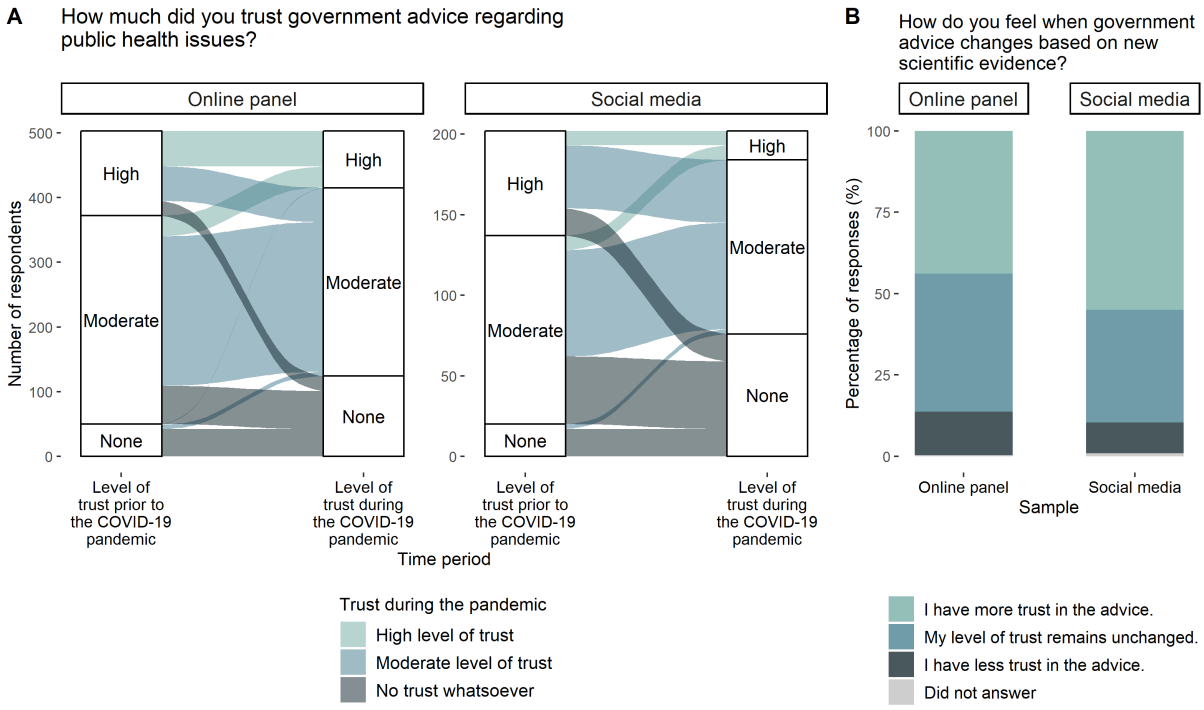

Figure 5.4: Trust in government scientific advice. (A) The percentage of observations from both the online panel and social media samples according to time period and answer to the question "How much did you trust government advice regarding public health issues?" both "Prior to" (reported retrospectively) and "During" the COVID-19 pandemic. Observations corresponding to respondents who did not answer the question for at least one time point were removed from the figure to aid visual presentation ($n = 1$ respondent from the online panel). Underlying data are presented in Table 5.1. (B) Answers to the question "How do you feel when government advice changes based on new scientific evidence?" for both the online panel and social media samples. Underlying data are presented in Table 5.1.

5.4 Discussion

This study was undertaken to complement the findings of our previously published work on the relationship between modelling and policy [50] by providing insights into the UK public's awareness and opinion of transmission modelling in informing policy. As such, novel data were gathered under two different sampling mechanisms, an online panel ("representative" sample) and social media ("non-probability" sample) and in respect of two time periods, "Prior to" and "During" the COVID-19 pandemic.

Almost all respondents reported being increasingly aware of the use of modelling in informing policy during the pandemic, with significantly higher levels of awareness among social media respondents than online panel respondents. Awareness generally stemmed from the news media and social media during the pandemic. The mean reliability score increased during the pandemic compared to the (retrospective) pre-pandemic period under both samples, with awareness being positively associated with reliability only within the online panel. Trust in government public health advice remained high across samples and time periods overall but was lower in the period of the pandemic compared to the pre-pandemic period. The decay in trust was notably greater among social media respondents. Many respondents from both samples explicitly made the distinction that their trust was reserved for "scientists" and not "politicians". Almost all respondents, regardless of sample, believed governments have responsibility for the communication of modelling to the public.

Our results are not, nor claim to be, a definitive representation of all public opinion on this topic. Instead, we have focussed primarily on the differences in the responses between our two samples. Specifically, our results validate our a priori expectation of a skewed response among social media respondents thus providing an important reminder of the potentially biased conclusions that could be drawn from non-representative samples [444]. Throughout, it is important to recall that social media respondents were likely to have accessed the survey directly from a Twitter account they follow related to public health or mathematical modelling, thus demonstrating an active interest in and/or familiarity with the subject matter (Table C.1). In contrast, online panel respondents saw this survey without any context. Furthermore, social media respondents were significantly more likely to work in "Research" or "Education" than online panel respondents, which may account for a substantial amount of the differences observed in the responses received, given that modelling for public health policy was primarily undertaken by academics [325].

While the general sentiment of the responses was consistent across samples, often social media respondents were more likely to have stronger or polarised views. Examples of this include trust in the government

decreasing within both samples but at a greater level for social media respondents compared to the online panel, and almost all social media respondents stating awareness of the use of modelling in informing policy compared to three quarters of online panel respondents. In general, age and gender were only found to be significantly associated with or approaching significant association for responses among the social media sample. Nonetheless, caution is still needed when interpreting the results from the online panel as the suitability of this sample as a representation of the UK population is unclear. However, 87% of online panel respondents (aged over 18 years old) stated having had at least one dose of a vaccine for SARS-CoV-2 compared to 76.9% of the population over 12 years old in the UK on 9 July 2021 (when online panel survey responses were gathered) [445]. As vaccination of those aged under 18 years was not recommended in the UK until 15 July 2021 [446], it is reasonable to consider 76.9% a lower bound of vaccination coverage of the UK adult population thus suggesting reasonable representativeness of the online panel with regards to this variable.

Given the role that modelling has played in the pandemic response in the UK [50, 325, 429, 430], and the corresponding coverage in the media [447–450], it is unsurprising that respondents reported an increased awareness of the use of modelling as a tool to inform public health policy. However, it is perhaps unexpected that this translated to most participants in both samples. The popularity of the news media as a means of awareness underlines the important role that journalists play as science communicators, and emphasises the value of organisations such as the Science Media Centre [451]. A number of the experts interviewed as part of our aforementioned study praised the commitment that many journalists showed to clear and accurate communication of modelling studies [50]. Professor Sir Chris Whitty also spoke about communicating modelling to the public during the pandemic during his testimony to the UK COVID-19 Inquiry in November 2023, noting that this has to be done with "great care" and in a way that is different to presenting raw data [452]. The rise in popularity of social media as a means for being aware of modelling is unsurprising among social media respondents, but the significant four-fold increase among online panel users in the period of the pandemic was particularly notable. Akin to the case of the news media, this finding further demonstrates the role that social media can play in science communication, particularly in the COVID-19 pandemic, and, consequently, the importance of ensuring that studies are properly represented to users of this platform by those sharing them.

One of the major findings of the authors' initial science-policy paper was the importance of distinguishing between scientists and decision-makers, with only the latter being directly accountable to the public [50]. Many respondents, from both samples, noted this distinction explicitly when commenting on the question regarding trust in government public health policy, and within both samples during the pan-

demic respondents were significantly more likely to state that modellers work within "Academia" rather than "Government" (see Appendix C). While this suggests a reasonable understanding of the difference between these two roles, further and continued emphasis of the difference in roles may be warranted ahead of future emergencies, public health-related or otherwise, as "Government" became a significantly more popular response during the pandemic compared to the pre-pandemic period among the online panel. Nearly all respondents from both samples indicated that the government have the responsibility of ensuring that the public are informed of scientific evidence underpinning their decisions. This remains compatible with another key finding of our initial study that science must be communicated by scientists [50], for example by having scientists present at Government-organised press conferences.

There will often be tensions between scientific advice and decision-making. This was particularly evident during the COVID-19 pandemic when our understanding of SARS-CoV-2 was rapidly evolving as new data became available, with public health guidance changing almost equally rapidly to reflect new findings. This was highlighted by science communication experts in our previous study as a key challenge of communicating with the public and was encapsulated well by a social media respondent: "Politicians don't like u-turns while scientists are keen to constantly evolve thinking and theories. These outlooks are not particular[ly] compatible when it comes to communicating subtle changes in public health messaging.". However, most respondents from both samples noted that changing advice either did not change or increase their trust in said advice, although this was significantly more common among social media respondents. Our results indicate a key factor of trust is transparency or evidence, also highlighted in our initial study [50] rather than rapidly changing advice.

The fact that respondents reported "No awareness" of modelling in informing policy and that for some (online panel $n = 10$; social media $n = 1$) this was an "About right" level of knowledge provides an important reminder that this topic is, understandably, not of interest to all. Furthermore, while awareness was associated with increased feelings of reliability of modelling, the fact that some "Aware" respondents selected a reliability score in the lower half of the scale during the pandemic highlights the complexities of this relationship. Nonetheless, some respondents seemed to show a genuine interest and appreciation of modelling: "I have found a new respect for those who do modelling, to be honest I find it fascinating" (online panel).

It is important to note the limitations of this study. All results rely on the opinions and experiences of our participants. Although we distributed the survey via two platforms, as discussed we are unable to ascertain the representativeness of the online panel in terms of the UK population, as is cautioned

by Prolific Academic [442]. Furthermore, although respondents were asked questions in relation to the period "Prior to" and "During" the COVID-19 pandemic separately, sampling only took place during the latter period. Therefore, responses for the pre-pandemic period may have been influenced by recall bias and should be interpreted with caution. Due to the nature of the sampling platforms, all online panel responses were obtained in a much shorter period (days) than social media responses (weeks). Around this time, all public health restrictions in England were lifted after a one-month delay [453], which could have influenced participants' responses. We also note that the limited sample sizes obtained under the two sampling mechanisms ($n = 504$ for the online panel and $n = 202$ for social media) necessitates caution when interpreting differences between the two samples. Finally, although respondents were given multiple opportunities to add open-ended comments, multiple-choice answers may lose some of the nuance of participants' opinions.

As noted by one respondent: "I think we have all learned something about science and how it impacts on our daily lives throughout the pandemic". Ultimately, the public are key stakeholders of public health policy, and if mathematical transmission models are used to inform important decisions then there is a responsibility to assess people's understanding of and feelings towards them. We hope that the findings of this study will be useful in informing science communication strategies in future emergencies. For example, by being more transparent about the evidence underpinning the implementation of or rapid change to policy or by better utilising the news media and/or social media as means of communicating with the public.

Table 5.1: Questions, and multiple-choice answers where relevant, for the survey to the public. The two compulsory questions are highlighted with an asterisk (*). Totals and percentages are based on responses with "complete" demographic data, namely age group, gender, sector and vaccination status within each sample.

Question	**Question type**	**Response options (if (multiple-choice)**	**Number of online panel respondents (n (%))**		**Number of social media respondents (n (%))**	
How would you describe a transmission model?*	Open-ended		504		202	
What level of confidence do you have in your description?*	Multiple-choice	Not at all confident	200	(40%)	70	(35%)
		Moderately confident	262	(52%)	111	(55%)
		Very confident	42	(8%)	21	(10%)
Before the COVID-19 pandemic						
For the purposes of this research, a transmission model is defined as a mathematical model of the spread of a virus through a population in order to simulate quantities such as, but not limited to, the daily number of deaths, new infections and hospitalisations due to the virus.						
Prior to the COVID-19 pandemic, how were you aware of transmission modelling? Please select all that apply.	Multiple-choice	I was not aware	314	(62%)	74	(37%)
		Newspaper (Online or print)	69	(14%)	34	(17%)
		News show (TV or online)	68	(13%)	27	(13%)
		Social media	34	(7%)	16	(8%)
		Internet search	49	(10%)	22	(11%)
		Academic reports and papers	53	(11%)	48	(24%)
		Other	36	(7%)	45	(22%)
		Did not answer	0	(0%)	0	(0%)
Before the COVID-19 pandemic, where do you think those who developed and used transmission models work?	Open-ended					
Prior to the COVID-19 pandemic, were you aware of the use of transmission models in informing public health policy (e.g. to model the spread of pandemic influenza)?	Multiple-choice	Yes	137	(27%)	116	(57%)
		No	274	(54%)	49	(24%)
		Unsure	91	(18%)	37	(18%)
		Did not answer	2	(0%)	0	(0%)
Prior to the COVID-19 pandemic, on a scale of 1 - 10 with 1 being "extremely unreliable" and 10 being "extremely reliable", how did you feel about the use of transmission models in informing public health policy? Please select a number below.	Multiple-choice	1	11	(2%)	4	(2%)
		2	3	(1%)	0	(0%)
		3	25	(5%)	13	(6%)
		4	24	(5%)	5	(2%)
		5	112	(22%)	47	(23%)
		6	60	(12%)	27	(13%)
		7	118	(23%)	35	(17%)
		8	110	(22%)	47	(23%)
		9	28	(6%)	19	(9%)
		10	11	(2%)	4	(2%)
		Did not answer	4	(1%)	1	(0%)
Prior to the COVID-19 pandemic, how much did you trust government advice regarding public health issues?	Multiple-choice	High level of trust	131	(26%)	65	(32%)
		Moderate level of trust	322	(64%)	117	(58%)
		No trust whatsoever	51	(10%)	20	(10%)
		Did not answer	0	(0%)	0	(0%)
Please provide a brief explanation regarding	Open-ended					

your answer to the previous question						
During the COVID-19 pandemic						
For the purposes of this research, a transmission model is defined as a mathematical model of the spread of a virus through a population in order to simulate quantities such as, but not limited to, the daily number of deaths, new infections and hospitalisations due to the virus.						
During the COVID-19 pandemic, how were you aware of transmission modelling? Please select all that apply.	Multiple-choice	I was not aware	114	(23%)	6	(3%)
		Newspaper (Online or print)	204	(40%)	123	(61%)
		News show (TV or online)	300	(60%)	142	(70%)
		Social media	147	(29%)	125	(62%)
		Internet search	122	(24%)	72	(36%)
		Academic reports and papers	86	(17%)	97	(48%)
		Other	19	(4%)	23	(11%)
		Did not answer	0	(0%)	0	(0%)
During the COVID-19 pandemic, where do you think those who developed and used transmission models work?	Open-ended					
During to the COVID-19 pandemic, were you aware of the use of transmission models in informing public health policy (e.g. the implementation of so-called "lockdown" measures)?	Multiple-choice	Yes	374	(74%)	195	(97%)
		No	75	(15%)	2	(1%)
		Unsure	47	(9%)	4	(2%)
		Did not answer	8	(2%)	1	(0%)
During the COVID-19 pandemic, on a scale of 1 - 10 with 1 being "extremely unreliable" and 10 being "extremely reliable", how did you feel about the use of transmission models in informing public health policy? Please select a number below.	Multiple-choice	1	11	(2%)	10	(5%)
		2	9	(2%)	6	(3%)
		3	13	(3%)	9	(4%)
		4	21	(4%)	5	(2%)
		5	62	(12%)	24	(12%)
		6	56	(11%)	20	(10%)
		7	105	(21%)	36	(18%)
		8	129	(26%)	53	(26%)
		9	68	(13%)	30	(15%)
		10	26	(5%)	9	(4%)
		Did not answer	4	(1%)	0	(0%)
During the COVID-19 pandemic, how much did you trust government advice regarding public health issues?	Multiple-choice	High level of trust	88	(17%)	18	(9%)
		Moderate level of trust	281	(58%)	108	(53%)
		No trust whatsoever	124	(25%)	76	(38%)
		Did not answer	1	(0%)	0	(0%)
Please provide a brief explanation regarding your answer to the previous question	Open-ended					
How much do you know about how transmission modelling has been used throughout the COVID-19 pandemic?	Multiple-choice	Too little	233	(46%)	60	(30%)
		About right	257	(51%)	122	(60%)
		Too much	10	(2%)	17	(8%)
		Did not answer	4	(1%)	3	(1%)
Who do you think has the responsibility of ensuring that the public are informed about the use	Multiple-choice	The government	462	(92%)	183	(91%)
		Those who develop transmission models	198	(39%)	130	(64%)
		Those who use	217	(43%)	129	(64%)

modelling in policy decisions, particularly in COVID-19 pandemic?		transmission models				
		The media	221	(44%)	121	(60%)
		Myself, as a member	108	(21%)	68	(34%)
Please select all that apply.		of the public				
		None of the above	1	(0%)	2	(1%)
		Did not answer	1	(0%)	0	(0%)
If answered "None of the above", please write who you think has the responsibility of ensuring that the public are informed about the use of modelling in policy decisions.	Open-ended					
How do you feel when government advice changes based on new scientific evidence?	Multiple-choice	I have more trust in the advice	221	(44%)	111	(56%)
		I have less trust in the advice	67	(13%)	19	(10%)
		My level of trust remains unchanged	214	(42%)	70	(35%)
		Did not answer	2	(0%)	2	(1%)
Please provide a brief explanation regarding your answer to the previous question.	Open-ended					
Please provide any further comments about any of the previous questions.	Open-ended					
Demographic information						
Please note that you will not be able to be identified based on any information you provide in this section.						
Please select your age group.	Multiple-choice	18 - 25	71	(14%)	4	(2%)
		26 - 35	90	(18%)	10	(5%)
		36 - 45	92	(18%)	23	(11%)
		46 - 55	87	(17%)	58	(29%)
		56 - 65	125	(25%)	81	(40%)
		66+	39	(8%)	26	(13%)
Please select your gender.	Multiple-choice	Female	260	(52%)	104	(51%)
		Male	240	(48%)	95	(47%)
		Non-binary	1	(0%)	0	(0%)
		Prefer not to say	3	(1%)	3	(1%)
Please select the option below which most closely matches the sector you work in.	Multiple-choice	Agriculture	1	(0%)	0	(0%)
		Business, finance and technology	75	(15%)	29	(14%)
		Creative	21	(4%)	5	(2%)
		Education	68	(13%)	51	(25%)
		Health and social work	69	(14%)	30	(15%)
		Production (energy, manufacturing, mining)	30	(6%)	12	(6%)
			36	(7%)	16	(8%)
		Public	36	(7%)	16	(8%)
		Research	17	(3%)	21	(10%)
		Retail, hospitality and tourism	48	(10%)	4	(2%)
		Other	139	(28%)	34	(17%)
COVID-19 vaccination						
Have you been	Multiple-choice	Yes	439	(87%)	195	(97%)

vaccinated for COVID-19?		No, but I plan to in the future	41	(8%)	1	(0%)
		No, and I do not plan to	24	(5%)	6	(3%)

Chapter 6

Paper IV: Tracking the incidence and risk factors for SARS-CoV-2 infection using historical maternal booking samples

E Mullins*, R McCabe*, SM Bird, P Randell, MJ Pond, L Regan, E Parker, M McClure and CA Donnelly. Tracking the incidence and risk factors for SARS-CoV-2 infection using historical maternal booking serum samples. *PLOS ONE*, 17(9):1–15, 09 2022. doi: `https://doi.org/10.1371/journal.pone.0273966`.

Abstract

The early transmission dynamics of SARS-CoV-2 in the UK are unknown but their investigation is critical to aid future pandemic planning. We tested over 11,000 anonymised, stored historic antenatal serum samples, given at two north-west London NHS trusts in 2019 and 2020, for total antibody to SARS-CoV-2 receptor binding domain (anti-RBD). Estimated prevalence of seroreactivity increased from 1% prior to mid-February 2020 to 17% in September 2020. Our results show higher prevalence of seroreactivity to SARS-CoV-2 in younger, non-white ethnicity, and more deprived groups. We found no significant interaction between the effects of ethnicity and deprivation. Derived from prevalence, the estimated incidence of seroreactivity reflects the trends observed in daily hospitalisations and deaths in London that followed 10 and 13 days later, respectively. We quantified community transmission of SARS-CoV-2 in London, which peaked in late March/early April 2020 with no evidence of community transmission until after January 2020. Our study was not able to determine the date of introduction of the SARS-CoV-

2 virus but demonstrates the value of stored antenatal serum samples as a resource for serosurveillance during future outbreaks.

Author contributions

- Edward Mullins: Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Writing – original draft, Writing – review & editing.
- **Ruth McCabe: Formal analysis, Investigation, Writing – original draft, Writing – review & editing.**
- Sheila M. Bird: Formal analysis, Funding acquisition, Investigation, Writing – review & editing.
- Paul Randell: Funding acquisition, Investigation, Methodology, Resources, Writing – review & editing.
- Marcus J. Pond: Investigation, Project administration, Resources, Writing – review & editing.
- Lesley Regan: Funding acquisition, Methodology, Writing – review & editing.
- Eleanor Parker: Investigation, Project administration, Resources, Writing – review & editing.
- Myra McClure: Investigation, Methodology, Writing – review & editing.
- Christl A. Donnelly: Formal analysis, Funding acquisition, Investigation, Methodology, Writing – review & editing.

EM and RM contributed equally to this work. RM conducted all of the analysis and wrote most of the initial draft of the manuscript.

Ethics

Consent was not obtained. Ethical approval for a non-consent, anonymised study was gained, REC reference 20/NI/0107, because strong safeguards against deductive disclosure of the identities of individuals with SARS-CoV-2 antibodies were inbuilt.

Reproducibility

Data and code to reproduce these analyses can be found at: `https://github.com/ruthmccabe/historic_antenatal_serostudy`.

Acknowledgements

We wish to thank several individuals who helped with this study. Infrastructure support was provided by the NIHR Imperial College Biomedical Research Centre, Ben Glampson at Imperial College Healthcare Trust (ICHT) and Stephen Cole at Chelsea and Westminster Healthcare Foundation Trust (CWHFT). Tina Cotzias and Natasha Mohammed for acted as PIs for ChelWest Foundation NHS Trust. Jonathan Dennis supported with initial literature searches and prepared the study for ethical review. Sarah Littell at NW London Pathology identified and aliquoted samples. Victoria Male provided helpful discussions about immune responses in pregnancy. We thank the REal-time Assessment of Community Transmission (REACT) Study investigators for sharing antibody prevalence data. We are grateful to Samreen Ijaz at UKHSA for conducting validation testing on 2019 seroreactive samples. We are grateful to Wendy Barclay, Peter Horby and Peter Openshaw for their comments on this manuscript.

Appendix

Further information relating to this paper can be found in Appendix D.

6.1 Background

The highly transmissible severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) causes coronavirus disease 2019 (COVID-19), which has led to widespread suffering and loss of life since its discovery in Wuhan, China in late 2019. In the UK, the first two polymerase chain reaction (PCR)-confirmed COVID-19 cases were confirmed on 2 February 2020 and the first official COVID-19 death occurred on 2 March 2020 [112]. However, infection by SARS-CoV-2 does not invariably lead to symptomatic COVID-19: mild or asymptomatic infection is also common [454].

For several reasons, epidemiological surveillance systems based on reported clinical disease typically capture only a subset of the true number of infections and deaths attributable to a specific pathogen. Reasons include, but are not limited to, the occurrence of asymptomatic infections, limited testing capacity and delays in registration of deaths or mis-attributed cause [45]. For example, in England there are earlier deaths attributable to COVID-19 by autopsy, the (known) first of which occurred on 30 January 2020 [116].

Anonymised serological studies are an important mechanism to infer population-level prevalence and incidence of infection and avoid under-reporting inherent in the reliance upon the incidence of clinical infection [236]. The insights gained from such data allow for better understanding of the transmission dynamics of, and demographic groups most at risk from, a particular virus. Linked to self-completed behavioural risks questionnaire, nationwide surveillance surveys to detect SARS-CoV-2 antibodies in England did not start until after the peak of the first wave [63, 235].

Ascertaining estimates of the prevalence and incidence of infection in the early weeks of SARS-CoV-2 transmission are critical learning to inform preparations for future outbreaks. Nonetheless, such studies are challenging to conduct retrospectively because they rely on the availability of routinely stored blood or other biological samples from well-defined sub-populations, such as blood donors [455–457] or pregnant women [458], but typically lack non-demographic risk-factor information.

In the UK, serum samples routinely given by women at booking for antenatal care, and known as "booking samples", are stored for at least two years post-collection as the samples may be informative in monitoring the pregnancy or diagnosing illness in the infant during the first year of life [459, 460]. Archives of these booking samples present a unique opportunity to test retrospectively for total antibody to SARS-CoV-2 receptor binding domain (anti-RBD), and, by extension, shed light on early transmission dynamics, within a young (18–44 years old), broadly healthy, and ethnically diverse female population

across all deciles the Index of Multiple Deprivation (IMD) [461].

In this study, we present a retrospective analysis of antenatal serum samples from two north-west London trusts across October 2019 – September 2020 which were subsequently tested for SARS-CoV-2 seroreactivity. We used the results to estimate the prevalence of seroreactivity over time and in association with age, ethnicity, and deprivation and to examine the trends in incidence of seroreactivity over the study period.

6.2 Methods

6.2.1 Study design

Consent was not obtained. Ethical approval for a non-consent, anonymised study was gained, REC reference 20/NI/0107, because strong safeguards against deductive disclosure of the identities of individuals with SARS-CoV-2 antibodies were inbuilt.

By initial design, we planned to analyse the results of $11,000$ antenatal booking samples given at Imperial College Healthcare NHS Trust and Chelsea and Westminster NHS Foundation Trust from December 2019 to May 2020 to examine how the prevalence of seroreactivity to SARS-CoV-2 varied over time in this population. Samples were retrieved from the $-20°$C storage freezer at North-West (NW) London Pathology, held in boxes of 100 samples in loose chronological order. Samples were thawed prior to vortexing each sample for 10–15 seconds. Where there was sufficient sample stored, 150μL of serum was aliquotted into a single well of a 96 well plate using the DS2 instrument (Dynex Technologies–USA). Once aliquotted, the plates were stored in a $-80°$C freezer until they were transported on dry ice to the Molecular Diagnostics Unit (Imperial College London) for testing.

Prior to anonymization, samples were cross-classified by fortnight, age group (18–29 years, 30–34 years, 35–44 years), ethnicity (all white, all Asian, all Black, Other), and deprivation (most-deprived two IMD deciles 1–2, intermediate four deciles 3–6, least-deprived four deciles 7–10). During the aliquoting process, a small number of pairs of samples was identified to have come from the same person; the second of such samples to be aliquoted was marked as a duplicate and excluded from the formal analysis. We gave an undertaking not to report on any cross-classification for which the observed count was under 25. This undertaking was met by pooling across adjacent fortnights.

Testing of $5,350$ samples initially confirmed that sufficiently narrow confidence intervals for naïve preva-

lence estimates were obtained by testing around 500 samples per fortnight across the initial sampling period. Since samples prior to mid-February were reactive for anti-SARS-CoV-2 receptor binding domain (RBD) antibodies, a decision was taken in April 2021 to test around 500 samples per fortnight (rather than approximately all 1,000 available samples held by NW London Pathology for the two north-west London trusts per fortnight) and focus the conserved test-resources on samples from fortnights earlier than late December 2019. The testing period was also extended to fortnights through September 2020 to allow comparison with national serosurveys running during this period. Additionally, due to the anticipated stabilised high prevalence in fortnights after May 2020, the number tested in those later fortnights was reduced to around 200.

In total, an aliquot of the stored sample was retrieved from 12,348 antenatal serum samples, collected over fortnightly intervals from fortnight 22 in 2019 (22 October – 4 November 2019) to fortnight 19 in 2020 (9 – 22 September 2020) with accompanying anonymised demographic information where available, as detailed above.

Following the initial rounds of testing, the continued presence of seroreactive samples across all tested fortnights in 2019 led us to seek additional funding to conduct further testing of booking samples from earlier in 2019. The earliest maternal booking samples available at the time were from June 2019; 1,000 of these were aliquoted and tested. Further information on this validatory testing is provided below. As these data would be used only to aid interpretation of our main findings, it was not necessary (nor possible) to cross-classify samples by age group, ethnicity, and deprivation. Ethical approval for the original study was expanded to include the testing of these June 2019 samples, with the amendments to the study period classed as "non-substantial".

6.2.2 Testing of samples

Samples were assigned a testing bin number, based on sample year, fortnight, age group, ethnicity group, and deprivation group (except for June 2019 samples which were only assigned a sample year and month). Testing was thereby conducted anonymously.

For the detection of total antibodies to SARS-CoV-2 receptor binding domain (anti-RBD) we used an in-house hybrid double antigen binding assay (DABA), the Imperial College London Hybrid DABA, which is a two-step sequential enzyme-linked immunosorbent assay (ELISA) and is accredited by the United Kingdom Accreditation Service (UKAS). The Imperial Hybrid DABA detects and quantifies total antibodies to SARS-CoV-2 anti-RBD [462]. This assay has a specificity greater than 99.6% (95% CI

99.6%–100%), defined by testing 825 serum samples that predated the COVID-19 pandemic, 100 of which were historic antenatal booking serum samples from 2018; and a sensitivity of 98.9% (95% confidence interval (CI) 96.8%–99.8%) when evaluating 276 serum samples from individuals with RT-PCR-confirmed SARS-CoV-2 infection [463]. Further details are provided in Rosadas et al. [464].

6.2.3 Defining seroreactivity

As set out in Rosadas et al. [464], the cutoff for seroreactivity was established by adding 0.1 to the average of optical density (OD) obtained for three negative controls assayed in each run. The signal-to-cutoff value, known as the binding ratio (BR), for each sample was determined by dividing the sample OD by the cutoff OD. A sample was generally considered seroreactive for SARS-CoV-2 anti-RBD when the BR was greater than or equal to 1. BRs between 0.8 and 1.2 were considered to display a weak signal (borderline non-seroreactive; borderline seroreactive).

6.2.4 Statistical analysis

Data from October 2019 to September 2020 were cleaned to contain only "complete" observations in terms of BR, sampling fortnight, and categorisation into each of the age, ethnicity, and deprivation groups outlined previously.

Assuming that the number of reactive samples is a binomially-distributed random variable, prevalence of seroreactivity was estimated using the maximum likelihood estimate (MLE) as follows:

$$\hat{Prevalence\ of\ seroreactivity} = \frac{Number\ of\ seropositive\ observations}{Number\ of\ total\ observations}$$

The prevalence of seroreactivity was estimated per fortnight as well as for each age, ethnicity and deprivation groups over the entire sampling period, and is presented as the MLE with 95% exact binomial confidence intervals. Prevalence of seroreactivity for the month of June 2019 was also estimated using this method.

Two-sided Fisher's exact tests and chi-squared tests were used to examine statistically significant changes in the proportions of seroreactive observations across fortnights.

Binomial logistic regression was used to quantify the relationship between seroreactivity and sampling fortnight, age, ethnicity and deprivation. The oldest age group, all-white ethnicity (ethnicity group with lowest prevalence), and most-deprived deprivation group were selected as the baseline groups. Per-

protocol, the baseline sampling fortnight was the fortnight prior to lockdown. To test for interaction between ethnicity and deprivation, a further model with an interaction term between ethnicity and deprivation was considered, with the overall significance of this term determined by a chi-squared test.

Estimates of the incidence of seroreactivity per fortnight were derived from estimated prevalence of seroreactivity using a bootstrapping procedure. For each bootstrap sample, $11,256$ bootstrap-observations in total were sampled with replacement from the available data (i.e. $11,256$ actual observations), according to fortnightly groupings, and used to estimate prevalence of seroreactivity as detailed above. A shape-constrained P-spline was then fitted to the bootstrapped prevalence estimates to characterise the temporal trends in the data (Figure D.1). The incidence per fortnight was simply the difference in the current and previous fortnight's modelled prevalence estimates. The incidence among susceptible persons per fortnight was then estimated by dividing the difference in the current and previous fortnight's modelled prevalence estimates by the proportion of population estimated to have been susceptible at the previous fortnight (1 minus the model estimate of the proportion seroreactive). Bootstrap sampling and analysis were repeated $1,000$ times. The reported central estimate is the median with 95% confidence interval constructed by taking the 2.5th and 97.5th quantiles of the bootstrap samples, respectively.

Estimated incidence of seroreactivity over the study period was compared to subsequent seven-day average daily hospital admission and death rates in London [112], akin to Riley et al. [465]. Hospital admissions and deaths were shifted by a lag of L_H and L_D days, respectively, to capture the delay between infection, the development of seroreactivity, hospitalisation, and death, and scaled by a value of β_H and β_D to capture the different scales of these time series. L_H and L_D were allowed to take integer values from 0 to 40 (days). For each integer value from 0 to 40 days, individual linear regression models estimated incidence per fortnight on the appropriately-lagged daily hospital admissions and death rates, weighted proportional to the inverse of the variance of the incidence estimates. The best-fitting lag was selected according to the model which minimised the weighted sum of squared errors (SSE). Conditional on the best-fitting lag, the scaling parameter was the slope estimated in the regression model.

6.2.5 Validation

To validate our testing by DABA in samples from June 2019, all seroreactive samples were re-tested blind with 2 consecutive, negative controls for each positive sample, in the United Kingdom Health Security Agency (UKHSA) Colindale laboratory using their nucleotide protein (NP) and spike protein S1 subunit (S1) capture assays, with and without blocking for seasonal coronaviruses.

Table 6.1: Demographics of the study sample. The number of total observations and the number of seroreactive observations per group within each demography variable across the primary study period (October 2019 –September 2020).

Demography variable	**Group**	**Total observations (n)**	**Positive observations (x)**	**Estimated seroprevalence ($\hat{s}$)**
Age (years)	18 - 29	2,930	226	7.7% (6.8% - 8.7%)
	30 - 34	4,135	256	6.2% (5.5% - 7.0%)
	35 - 44	4,226	218	5.2% (4.5% - 5.9%)
Ethnicity	All-black	839	87	10.4% (8.4% - 12.6%)
	All-Asian	2,016	160	8.0% (6.8% - 9.2%)
	All-white	4,613	208	4.5% (3.9% - 5.2%)
	Other	3823	245	6.4% (5.7% - 7.2%)
IMD Decile	Deciles 1 - 2	1,546	120	7.8% (6.5% - 9.2%)
	Deciles 3 - 6	6,492	427	6.6% (6.0% - 7.2%)
	Deciles 7 - 10	3,253	153	4.7% (4.0% - 5.5%)

6.3 Results

Fifty-five duplicates were removed from the 12,348 retrieved samples. Of the 12,293 non-duplicated retrieved samples from October 2019 to September 2020, 11,256 (92%) had binding ratio value, sampling fortnight and were categorised into each of the relevant age, ethnicity and IMD groups and so were included in the analysis.

Seven hundred of the 11,256 samples were seroreactive (defined by a $BR \geq 1$), resulting in an overall prevalence of seroreactivity estimate of 6.2% (95% CI $5.8\% - 6.7\%$). The vast majority of seroreactive samples (665; 95%) were categorised as "clearly seroreactive" ($BR \geq 1.2$) (Table D.1).

The median BRs among seroreactive observations per fortnight remained at low levels across October 2019 – February 2020 (Figure 6.1). From late February 2020, there is a sharp increase in the median BRs per fortnight through late April 2020. Thereafter there is a decrease, although BR values do not fall to the level observed at the beginning of the sampling period. Among non-seroreactive observations, median BRs remained fairly constant across fortnights throughout the main sampling period (Figure 6.1).

6.3.1 Prevalence by demographic variables and by fortnight

There was heterogeneity in the sample size across ethnicity and IMD groups, even more than anticipated, with fewest observations seen for the youngest age-group (18–29 years: 25.9%), black ethnicity (7.4%) and the most deprived IMD-group (IMD 1–2: 13.7%). Overall, the prevalence of seroreactivity decreased with age; was more than twice as common in pregnant black women compared to pregnant white women; and increased as deprivation increased (Table 6.1).

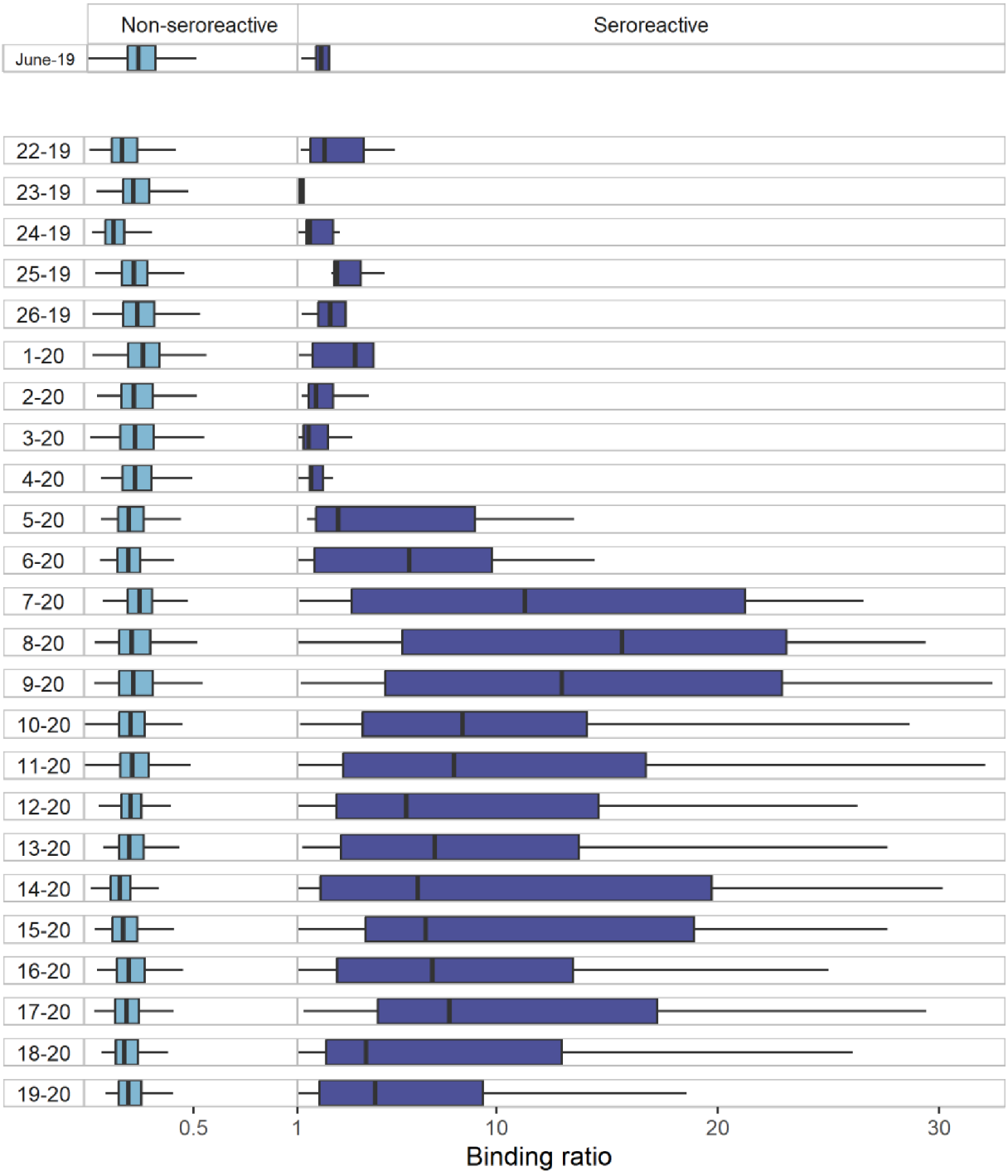

Figure 6.1: Binding ratio values among non-seroreactive (left) and seroreactive (right) observations per fortnight (October 2019-September 2020) and per month for June 2019. Note that the assay approximately 22–23 non-linear. From left to right, lines on the boxplot indicate: 1.5 times the interquartile range (IQR) less than the first quartile, first quartile, first quartile, median third quartile and 1.5 times the IQR greater than the third quartile. Outliers, defined as any point outwith the lower and upper bounds described have been removed so as not to obscure or distort the presentation of the other results.

Seroreactive samples were found in the earliest fortnight of the primary study period (October 2019 – September 2020): fortnight 22 of 2019 (22 October – 4 November 2019) (Figure 6.2; Table 6.2). Each

age, ethnicity (except the all-black group) and deprivation group was represented among the earliest seroreactive samples.

Estimated prevalence of seroreactivity increased substantially across February – May 2020 (Figure 6.2; Table 6.2) coincident with an increase in national confirmed case numbers, hospitalisations and deaths attributable to COVID-19 [112]. Despite reasonably large sample sizes, prevalence of seroreactivity estimates in 2019 were low and somewhat volatile over time, which four-weekly summaries counteracted (Table 6.2).

While there was an overall increase in the prevalence of seroreactivity across June – September 2020, some of the fortnightly heterogeneity in prevalence estimates in this period was reflective of the comparatively smaller sample size (Figure 6.2; Table 6.2). This was also overcome using four-weekly summaries. For example, prevalence of seroreactivity significantly increased from 12.4% (95% CI $10.1\% - 14.9\%$) in June 2020 (fortnights 12–13) to 17.1% (95% CI $13.5\% - 21.2\%$) in September 2020 (fortnights 18–19) (chi-squared test: $\chi^2 = 4.52$; $n = 1175$; $p = 0.034$).

Eight out of $1,000$ samples from the additional June 2019 samples were also found to be seroreactive. The median BR of these eight samples was similar to those observed across the initial fortnights of the primary study period (Figure 6.1). Two-sided Fisher's exact tests indicated no statistically significant changes in the odds of observing a seroreactive sample across each fortnight between fortnight 22 of 2019 and fortnight 5 of 2020 in comparison to June 2019, suggesting no evidence of community transmission of SARS-CoV-2 within our sample until after January 2020 (Table D.2).

6.3.1.1 Multifactorial logistic regression

Women in the youngest age category were significantly more likely to be seroreactive than women in the oldest age category ($p = 0.004$). Minority ethnicity groups were associated with significantly increased seroreactivity compared to the all-white group ($p < 0.001$ for all-black, all-Asian and other groups each compared to all-white) (Table 6.3). Relative deprivation was associated with seroreactivity, but significantly so only for the most deprived IMD sub-grouping in comparison to the least deprived ($p = 0.020$).

The 10 fortnights between 22 October 2019 and 10 March 2020 (which precede the reference fortnight of 11 to 24 March) had significantly reduced levels of seroreactivity. On the other hand, all fortnights after the reference fortnight (13 fortnights between 25 March 2020 and 22 September 2020) had significantly increased levels of seroreactivity, compared to the reference fortnight.

Table 6.2: Sample size and naïve prevalence of seroreactivity over time. The total number of observations and the number of seroreactive observations per fortnight and per four-week period using a seroreactive threshold of BR value ≥ 1 in the primary study period (and per month in June 2019). This is used to calculate naïve estimates of the prevalence of seroreactivity with exact binomial 95% confidence intervals. Note fortnight 26 of 2019 spans across 15 rather than 14 days.

Collection fortnight (fortnight-year)	**Calendar date**	**Fortnightly**			**4-week period**		
		Total observations (n)	**Seroreactive observations (x)**	**Estimated prevalence of seroreactivity ($\hat{s}$)**	**Total observations (n)**	**Seroreactive observations (x)**	**Estimated prevalence of seroreactivity ($\hat{s}$)**
June-2019	01/06/2019-30/06/2019				1000	8	0.8% (0.3% - 1.6%)
22-2019	22/10/2019-04/11/2019	608	8	1.3% (0.6% - 2.6%)	1218	13	1.1% (0.6% - 1.8%)
23-2019	05/11/2019-18/11/2019	610	5	0.8% (0.3% - 2.6%)			
24-2019	19/11/2019-02/12/2019	546	6	1.1% (0.4% - 2.4%)	1060	9	0.8% (0.4% - 1.6%)
25-2019	03/12/2019-16/12/2019	514	3	0.6% (0.1% - 1.7%)			
26-2019	17/12/2019-31/12/2019	554	5	0.9% (0.3% - 2.1%)	1021	10	1.0% (0.5% - 1.8%)
1-2020	01/01/2020-16/12/2019	467	5	1.1% (0.4% - 2.5%)			
2-2020	15/01/2020-28/01/2020	572	6	1.1% (0.4% - 2.3%)	1241	13	1.0% (0.6% - 1.8%)
3-2020	29/01/2020-11/02/2020	669	7	1.1% (0.4% - 2.1%)			
4-2020	12/02/2020-25/02/2020	530	9	1.7% (0.8% - 3.2%)	1158	16	1.4% (0.8% - 2.2%)
5-2020	26/02/2020-10/03/2020	628	7	1.1% (0.5% - 2.3%)			
6-2020	11/03/2020-24/03/2020	659	18	2.7% (1.6% - 4.3%)	1051	52	4.9% (3.7% - 6.4%)
7-2020	25/03/2020-07/04/2020	392	34	8.7% (6.1% - 11.9%)			
8-2020	08/04/2020-21/04/2020	622	73	11.7% (9.3% - 14.5%)	1259	150	11.9% (10.2% - 13.8%)
9-2020	22/04/2020-05/05/2020	637	77	12.1% (9.7% - 14.9%)			
10-2020	06/05/2020-19/05/2020	584	83	14.2% (11.5% - 17.3%)	1066	136	12.8% (10.8% - 14.9%)
11-2020	20/05/2020-02/06/2020	482	53	11.0% (8.4% - 14.1%)			
12-2020	03/06/2020-16/06/2020	499	58	11.6% (9.0% - 14.8%)	777	96	12.4% (10.1% - 14.9%)
13-2020	17/06/2020-30/06/2020	278	38	13.7% (9.9% - 18.3%)			
14-2020	01/07/2020-14/07/2020	254	37	14.6% (10.5% - 19.5%)	542	72	13.3% (10.5% - 16.4%)
15-2020	15/07/2020-28/07/2020	288	35	12.2% (8.6% - 16.5%)			
16-2020	29/07/2020-11/08/2020	267	40	15.0% (10.9% - 19.8%)	465	65	14.0% (11.0% - 17.5%)
17-2020	12/08/2020-25/08/2020	198	25	12.6% (8.3% - 18.1%)			
18-2020	26/08/2020-08/09/2020	233	35	15.0% (10.7% - 20.3%)	398	68	17.1% (13.5% - 21.2%)
19-2020	09/09/2020-22/09/2020	165	33	20.0% (14.2% - 26.9%)			

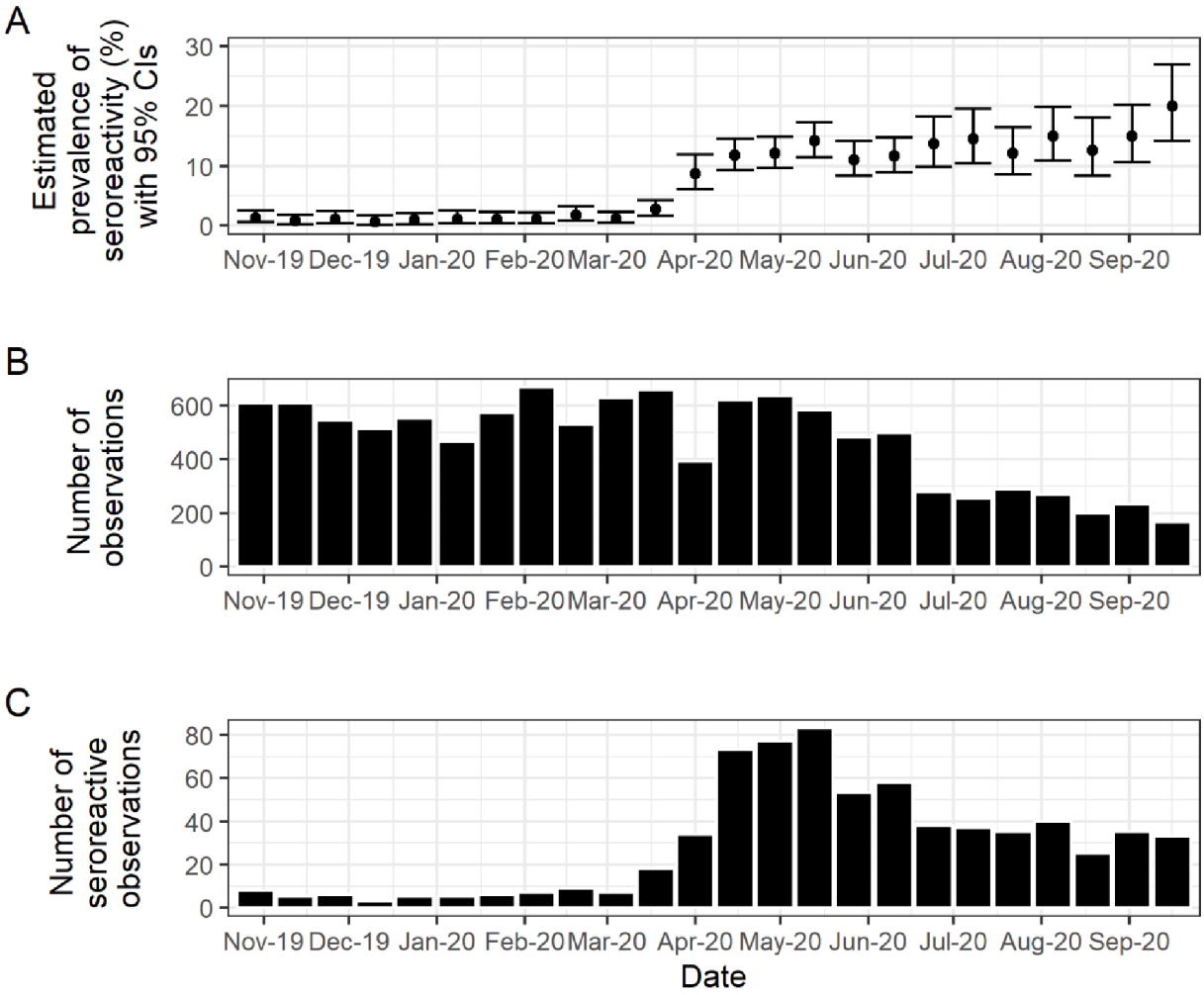

Figure 6.2: Prevalence of seroreactivity over time. (A) Estimated prevalence of seroreactivity over time with exact binomial 95% confidence intervals. (B) Total number of observations per fortnight. (C) Number of seroreactive observations per fortnight.

The interaction term between ethnicity and deprivation groups was not statistically significant (chi-squared test: $\chi^2 = 4.80$; degrees of freedom $= 6$; overall $p = 0.570$).

6.3.2 Incidence of seroreactivity over time

The estimated incidence of seroreactivity per fortnight remained around zero until February 2020 (Figure 6.3). After this time, there was a notable increase in estimated incidence per fortnight. This reached a peak between 11 March and 7 April 2020, corresponding to fortnights 6 and 7 of 2020, which encompasses the implementation of the first national lockdown on 23 March 2020. Incidence was then estimated to fall until around mid-August 2020 after which point it began to increase again. The shape of the incidence curve derived from estimated prevalence of seroreactivity reflects the trends displayed in daily hospital admission and death rates of COVID-19-confirmed cases in London 10 and 13 days later, respectively

Table 6.3: Results of logistic regression for seroreactivity using demographic variables. Coefficient estimates, standard errors and p-values for the logistic regression model with age group, ethnicity, IMD and fortnight of sample (across the primary study period October 2019 - September 2020) as predictors of seroreactivity. The dashed line indicates the timing of the reference fortnight.

Variable	**Interpretation**	**Estimate**	**Standard error**	**Adjusted odds ratio**	p
(Intercept)	35 - 44 All-white IMD decile group 1 - 2 11/03/2020 - 24/03/2020	-3.93	0.27		
Age	18 - 29	0.30	0.10	1.35	0.004
	30 - 34	0.12	0.10	1.13	0.223
Ethnicity	All-black	0.91	0.14	2.47	< 0.001
	All-Asian	0.57	0.11	1.77	< 0.001
	Other	0.40	0.10	1.49	< 0.001
Deprivation	IMD decile 3 - 6	-0.08	0.11	0.92	0.461
	IMD decile 7 - 10	-0.32	0.14	0.73	0.020
Fortnight	22/10/2019 - 04/11/2019	-0.74	0.43	0.48	0.084
	05/11/2019 - 18/11/2019	-1.23	0.51	0.29	0.016
	19/11/2019 - 02/12/2019	-0.92	0.48	0.38	0.041
	03/12/2019 - 16/12/2019	-1.61	0.63	0.20	0.010
	17/12/2019 - 31/12/2019	-1.17	0.51	0.31	0.022
	01/01/2020 - 14/01/2020	-0.95	0.51	0.39	0.061
	15/01/2020 - 28/01/2020	-0.97	0.48	0.38	0.041
	29/01/2020 - 11/02/2020	-0.98	0.45	0.38	0.030
	12/02/2020 - 25/02/2020	-0.43	0.41	0.65	0.295
	26/02/2020 - 10/03/2020	-0.89	0.45	0.41	0.048
	25/03/2020 - 07/04/2020	1.24	0.30	3.47	< 0.001
	08/04/2020 - 21/04/2020	1.57	0.27	4.79	< 0.001
	22/04/2020 - 05/05/2020	1.59	0.27	4.88	< 0.001
	06/05/2020 - 19/05/2020	1.78	0.27	5.93	< 0.001
	20/05/2020 - 02/06/2020	1.52	0.28	4.73	< 0.001
	03/06/2020 - 16/06/2020	1.55	0.28	4.58	< 0.001
	17/06/2020 - 30/06/2020	1.75	0.30	5.78	< 0.001
	01/07/2020 - 14/07/2020	1.84	0.30	6.32	< 0.001
	15/07/2020 - 28/07/2020	1.62	0.30	5.07	< 0.001
	29/07/2020 - 11/08/2020	1.87	0.30	6.47	< 0.001
	12/08/2020 - 25/08/2020	1.67	0.32	5.31	< 0.001
	26/08/2020 - 08/09/2020	1.85	0.30	6.39	< 0.001
	09/09/2020 - 22/09/2020	2.26	0.31	9.56	< 0.001

(Figure 6.3). Estimates of the incidence among susceptible persons is presented in Appendix D.

6.3.3 Validation of seroreactivity

Of the 8 samples from June 2019 which were seroreactive on DABA, none were found to be reactive on UKHSA NP or S1 capture assays. Of the 16 non-seroreactive samples on DABA (corresponding negative controls from June 2019), 2 were seroreactive on NP Immunoglobulin G (IgG) assay but were non-seroreactive after blocking for seasonal coronaviruses.

Internal validation found 1 of the 8 samples from June 2019 which were seroreactive on DABA was positive on Immunoglobulin M (IgM) S1 capture assay (Table D.3).

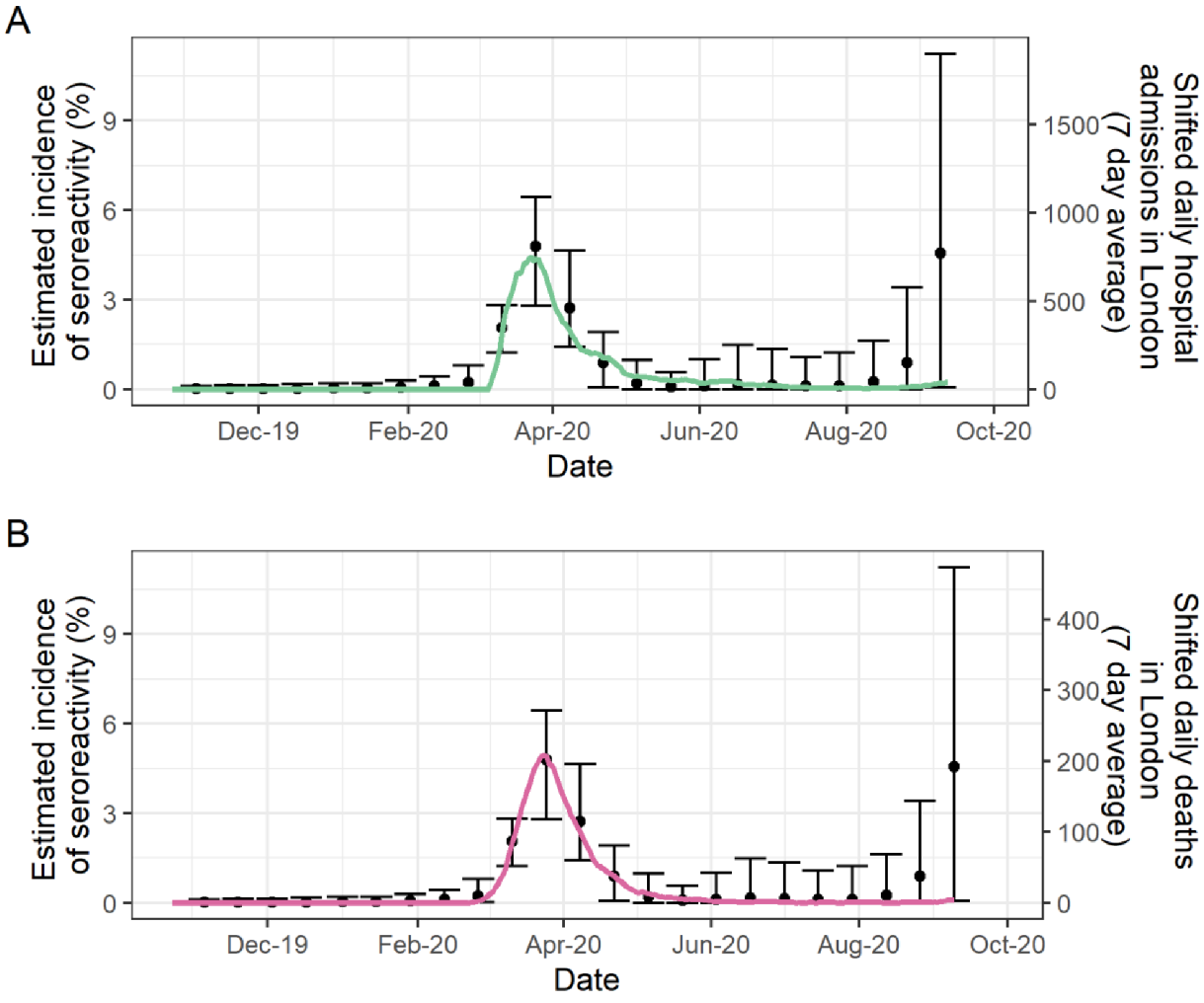

Figure 6.3: Estimated incidence of seroreactivity over time with 95% bootstrapped confidence intervals. (A) Seven-day averages of daily hospital admissions in London shifted by 10 days and scaled by 0.006. (B) Seven-day averages of daily deaths in London shifted by 13 days and scaled by 0.024. See Section 6.2.4 for further details.

6.4 Discussion

Our study presents a novel analysis of historic antenatal serum samples from an ethnically and socially diverse population in north-west London from late October 2019 to September 2020 in London, during which time SARS-CoV-2 was first detected. There were three direct flights per week from Wuhan to London Heathrow airport, in north-west London, until air travel from China to the UK was suspended on 29 January 2020 [466].

We found reactive samples throughout our testing period, but no evidence of sustained community transmission until after January 2020, assuming a seroconversion period of approximately 2–3 weeks. The temporal trends displayed in our estimated incidence of seroreactivity prevalence matched those observed in daily hospital admission and death rates in London that followed 10 and 13 days later, respectively [112].

Our results showed higher prevalence of seroreactivity to SARS-CoV-2 in non-white ethnicity groups, in the most deprived group and in younger age groups, in accordance with findings of other research groups [467–469]. The absence of significant interaction between ethnicity group and IMD deprivation group indicated no evidence in the data that the ethnicity-risk was exacerbated by relative deprivation [470]. Ethnicity itself was a key factor in SARS-CoV-2 infection rates in our study population.

Since the assay used for this study has a high specificity of 100% (95% CI 99.6%–100%) and sensitivity of 98.9% (95% CI 96.8%–99.8%) [462] our estimates for the prevalence of seroreactivity were not adjusted for sensitivity and specificity due to the negligible impact that this had on the results.

Our results did not identify the introduction into the United Kingdom of the highly transmissible SARS-CoV-2 variant, first detected in Wuhan. The external testing of our June 2019 seroreactive samples from 2019 indicate that the seroreactivity in the early samples was likely attributable to cross-reactivity with antibody to related viruses in these antenatal sera, as described in other studies [471, 472]. Incidence estimates depend on changes in changes in prevalence and so are unaffected by a low background level of cross-reactivity.

The end of our study period briefly coincided (for four fortnights) with the beginning of the England-wide Real-time Assessment of Community Transmission (REACT-2) study which had an initial response rate of 34% [469]. When comparing our estimates to a subset of the REACT-2 adult respondents matching our demographic variables, our estimates were consistent with the proportion antibody positive in REACT-2 (among female participants from London and aged 18 to 44 years) in their in June – July

2020 surveillance but not thereafter when the REACT-2 proportions declined (Figure D.3). However, REACT-2 used a different assay. Antibody responses wane with elapsed time since infection, but may do so differently between assays [469]. Adjusting for sensitivity and specificity of the different assay, using the method by Rogan et al. [473], did not explain the observed differences. Nonetheless, our estimates clearly captured the increase in seroreactivity prevalence in the first half of 2020, as reflected also in the first REACT-2 survey (June – July 2020).

Other serosurveillance studies have not provided robust insights into community transmission in late 2019 and early 2020. Lumley et al. [458] reported on $1,000$ antenatal samples from April and May 2020, of which 53 were reactive. Testing of samples from blood donors across the UK has also been undertaken. Thompson et al. [456] tested $3,500$ blood donor samples in Scotland across March to May 2020, of which 111 were reactive but none of the 500 obtained on 17 March 2020; and Public Health England tested $1,000$ blood donor samples per region per week, beginning from late March 2020 [457]. Additionally, Dickson et al. [455] analysed $4,751$ residual blood samples from Scottish regional biochemistry laboratories estimating a seroprevalence of 4.3% across April – June 2020. We note that the population of blood donors is typically less socially and ethnically diverse than pregnant women, two risk factors which have been shown to be associated with risk of infection of SARS-CoV-2.

Other published historical seroprevalence studies for detection of SARS-CoV-2 have often focussed on the timing of introduction rather than using their data to derive estimates of incidence. An historical seroprevalence study tested stored samples held on 959 patients in a lung cancer treatment trial from sites across Italy, detecting anti-RBD IgG and IgM antibodies to SARS-CoV-2 in September 2019 in 14% of samples [474]. Surprisingly, seroprevalence in this cohort fell to 2.8% in January 2020, rising to 20% in February 2020. The authors did not describe assay performance and background detection rates. Furthermore, their results were unable to be confirmed in an independent laboratory, at the request of the World Health Organization, thus reducing confidence in their observations [475].

The COVID-19 pandemic has underlined the importance of robust surveillance systems to detect community transmission as early as possible. Alternative, less-traditional means by which to detect public health impact have been explored throughout various epidemics. For example, during the COVID-19 pandemic all-cause excess mortality has been used to examine the possibility of wider transmission than those captured in official statistics as well as deaths caused indirectly by COVID-19's impacts (for example, through overwhelmed healthcare systems). One such study in England and Wales showed an increase in excess deaths from March 2020, which matches our finding that widespread community transmission

was not observed until after January 2020 [476].

Antenatal booking samples from at least two years ago (March 2020, at time of writing) are still currently stored throughout the UK (and potentially overseas [477]) and could allow mapping of the introduction and spread of the virus in this population over time with demographic characterisation. However, storage of these samples is not routinely required beyond 2 years. Pregnant women's increased susceptibility for some respiratory viruses make them a crucial group for surveillance, however the evidence for increased susceptibility to SARS-CoV-2 in pregnancy is mixed [478]. Advice for pregnant women to shield may affect the generalisability of our data to other populations, however these samples were taken at around 10 – 12 weeks gestation when the majority of women would have only recently have become aware of pregnancy.

In future epidemics, stored maternal booking serum samples from areas centred on UK travel hubs could be used, once reliable assays have been developed, to give retrospective estimates of early incidence rates; to identify populations at risk during the earliest stages of transmission and hence to track the emergence of a new virus. Such estimates would complement any blood donor serosurveillance undertaken. Surveillance for historical SARS-CoV-2 seroreactivity in blood donors could, even now, usefully compare $5,000$ to $10,000$ stored samples from each of June and December in 2018 versus 2019 to allow for donors' lower-risk while robustly investigating seasonality and cross-reactivity.

The data we report show the value of historical maternal seroprevalence studies in exploring the dynamics of the SARS-CoV-2 pandemic and pre-pandemic periods. The maternal serum samples from north-west London from October 2019 onwards remain in storage and could, in combination with sera stored in other regions of the UK and around the world, offer fascinating insights into the spread of the SARS-CoV-2 virus in different regional populations. As an anonymised tool for the future, these samples are a comprehensive rolling serum bank which stands ready to be activated to understand new and emerging viral threats. This could be an important component of ongoing surveillance at a time when some SARS-CoV-2 infection-surveillance programmes are being wound down [135, 479].

Chapter 7

Paper V: Alternative epidemic indicators for COVID-19 in settings with incomplete death registration systems

R McCabe*, C Whittaker, RJ Sheppard, N Abdelmagid, A Ahmed, IZ Alabdeen, NF Brazeau, AEA Abd Elhameed, AS Bin-Gouth, A Hamlet, R AbuKoura, G Barnsley, JA Hay, M Alhaffar, EK Besson, SM Saje, BG Sisay, SH Gebreyesus, AP Sikamo, A Worku, YS Ahmed, DH Mariam, MM Sisay, F Checchi, M Dahab, BS Endris, AC Ghani, PGT Walker, CA Donnelly and OJ Watson*. Alternative epidemic indicators for COVID-19 in three settings with incomplete death registration systems. *Science Advances* 9, eadg7676 (2023). doi: `https://doi.org/10.1126/sciadv.adg7676`.

Abstract

Not all COVID-19 deaths are officially reported, and particularly in low-income and humanitarian settings, the magnitude of reporting gaps remains sparsely characterized. Alternative data sources, including burial site worker reports, satellite imagery of cemeteries, and social media–conducted surveys of infection may offer solutions. By merging these data with independently conducted, representative serological studies within a mathematical modeling framework, we aim to better understand the range of underreporting using examples from three major cities: Addis Ababa (Ethiopia), Aden (Yemen), and Khartoum (Sudan) during 2020. We estimate that 69 to 100%, 0.8 to 8.0%, and 3.0 to 6.0% of COVID-19 deaths were reported in each setting, respectively. In future epidemics, and in settings where vital registration

systems are limited, using multiple alternative data sources could provide critically needed, improved estimates of epidemic impact. However, ultimately, these systems are needed to ensure that, in contrast to COVID-19, the impact of future pandemics or other drivers of mortality is reported and understood worldwide.

Author contributions

OJW conceived the idea for the study and designed the methodology used. **RM**, CW and OJW conducted the analysis with contributions to software and model fitting by RS, NB, GB, AH and PGTW. NA, AA, AEAAE, RAK, IZA and MD contributed data, model parameters and situational awareness for the analysis in Khartoum. BA MA, ASBG, EKB and FC contributed data, model parameters and situational awareness for the analysis for the Aden study. SMJ, BGS, SHG, AP, AW, YS, DHM, MMS, AK and BSE contributed data, model parameters and situational awareness for the analysis for the Addis Ababa study. ACG and CAD provided supervision throughout. **RM** and OJW wrote the first draft of the manuscript. All authors contributed to redrafting.

RM and OJW contributed equally to this work. **RM** conducted the analysis for Addis Ababa, with input from CW, and Aden and OJW conducted the analysis for Khartoum.

Reproducibility

All data and code needed to evaluate the conclusions in the paper are provided at `https://zenodo.org/record/7774954` [480]. The corresponding GitHub repository can be found at `https://github.com/mrc-ide/covid-alternative-mortality`.

Appendix

Further information relating to this paper can be found in Appendix E.

7.1 Introduction

Accurate ascertainment of infections, cases and deaths of an emerging infectious disease are vital in order to implement an effective public health response. However, these quantities depend on both testing capacity and robust vital registration systems. Consequently, only a subset of the true burden of a disease is captured within official statistics. During the coronavirus disease 2019 (COVID-19) pandemic, official COVID-19 data obscured the perception of the global burden of the disease [213]. For example, in some settings with low reported COVID-19 death tolls, serological surveys (serosurveys) have revealed extensive severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) transmission [255, 257, 258]. These levels of transmission are inconsistent with the number of infections expected based on officially reported deaths and estimates of the infection fatality ratio (IFR) [44, 259], with the most parsimonious explanation that COVID-19 deaths go unreported [481].

Due to the underreporting of COVID-19 deaths [58], excess mortality has been used frequently as an alternative means by which to assess the impact of the COVID-19 pandemic [482]. On 5 May 2022, the World Health Organization (WHO) published estimates of the "full" global death toll of the pandemic based on all-cause excess mortality [483]. Although not subject to the same biases as COVID-19 deaths, which typically rely on a proven SARS-CoV-2 infection around the time of the death, estimates of excess mortality still require complete data on the total number of deaths from all causes, which in turn requires mechanisms by which to record these deaths. Unfortunately, robust vital registration systems do not exist in many parts of the world, with the WHO estimating in 2020 that 40% of the world's deaths from all causes occur unregistered [219].

In settings without official all-cause mortality data, estimation of excess mortality during the pandemic has relied solely on model-based inference by pooling information from countries with similar socioeconomic and demographic characteristics. An alternative approach is to leverage alternative mortality data sources and epidemic indicators that can rapidly provide a more data-driven understanding of epidemic dynamics, such as social-media-shared obituary notifications [58]. Characterising the potential biases in these novel data sources is crucial to understanding their suitability for tracking the spread of SARS-CoV-2, as well as more broadly for mortality monitoring in the absence of complete vital registration. In response, we consider here three alternative data sources generated during the pandemic to understand the transmission of SARS-CoV-2 in two cities in sub-Saharan Africa and one in the Middle East. These include burial site worker reports in Addis Ababa [484], satellite imagery of cemeteries in Aden [220] and social-media-conducted surveys of symptomatic infection in Khartoum [485] (Table 7.1).

Table 7.1: Sources of alternative data collected at study locations and seroprevalence estimates. See Appendix E for further details of each study and the specifics of the assays used.

Setting	Alternative data source		Seroprevalence		
	Overview	**Source**	**Overview**	**Reported estimate**	**Source**
Addis Ababa, Ethiopia	Burial site worker cemetery reports 01/01/2015 - 26/01/2021	[484]	Random sample of 956 households 22/07/2020 - 10/08/2020	Immunoglobulin (IgG): 1.9% (95% CI 0.4% – 3.7%) Combined IgG/IgM: 3.5% (95% CI 1.7% – 5.4%)	[486]
Aden, Yemen	Satellite imagery of cemetery burials 21/07/2016 - 19/09/2020	[220]	Cross-sectional household study of 2,001 people 28/11/2020 - 13/12/2022	IgG: 25.0% (95% CI 23.2% – 26.9%) IgM: 0.2% (95% CI 0.1% – 0.4%) Combined IgG/IgM: 27.4% (95% CI 25.6% – 29.3%)	[256]
Khartoum, Sudan	Survey of historic symptomatic infections conducted through social media channels 26/05/2020 - 03/06/2020	[485]	Cross-sectional household study of 2,375 people 01/03/2021 - 10/04/2021	Combined IgG/IgM: 54.6% (95% CI 51.4% – 57.8%)	[487]

In each setting, representative serosurveys have also been conducted independently at different time-points before the availability of vaccines. We incorporate these within a previously developed mathematical modelling framework [58] to quantify how informative each data source is for explaining population seroprevalence under different assumptions for the IFR in each setting. In brief, the model uses a Susceptible-Exposed-Infected-Recovered structure, which explicitly captures pathways through hospital care and the impact this has on mortality (see Section 7.6 for further information). Our default assumption is that the relationship between IFR and age is consistent with the log-linear relationship estimated in high-income settings as estimated in multiple modelling studies [44, 259] and meta-analyses [481], although there is evidence that the IFR may be higher in lower-income settings [481]. We test this hypothesis first by fitting to the alternative sources for mortality (Addis Ababa and Aden) and case incidence (Khartoum) data using the relationship between IFR and age as estimated in Brazeau et al. [44]. From these model fits, we compare the model predicted seroprevalence against the reported seroprevalence estimates using a chi-squared test, arguing that this is good evidence that the alternative datasets are reliable for tracking the transmission of SARS-CoV-2 and the assumed IFR is suitable should these values not differ significantly. Additionally, in settings where the alternative sources and the serosurveys do not agree it presents an opportunity to better understand the biases in these alternative datasets as well as explore alternative hypotheses for the severity of COVID-19 in these settings. Full methods are presented in Section 7.6 and in Appendix E.

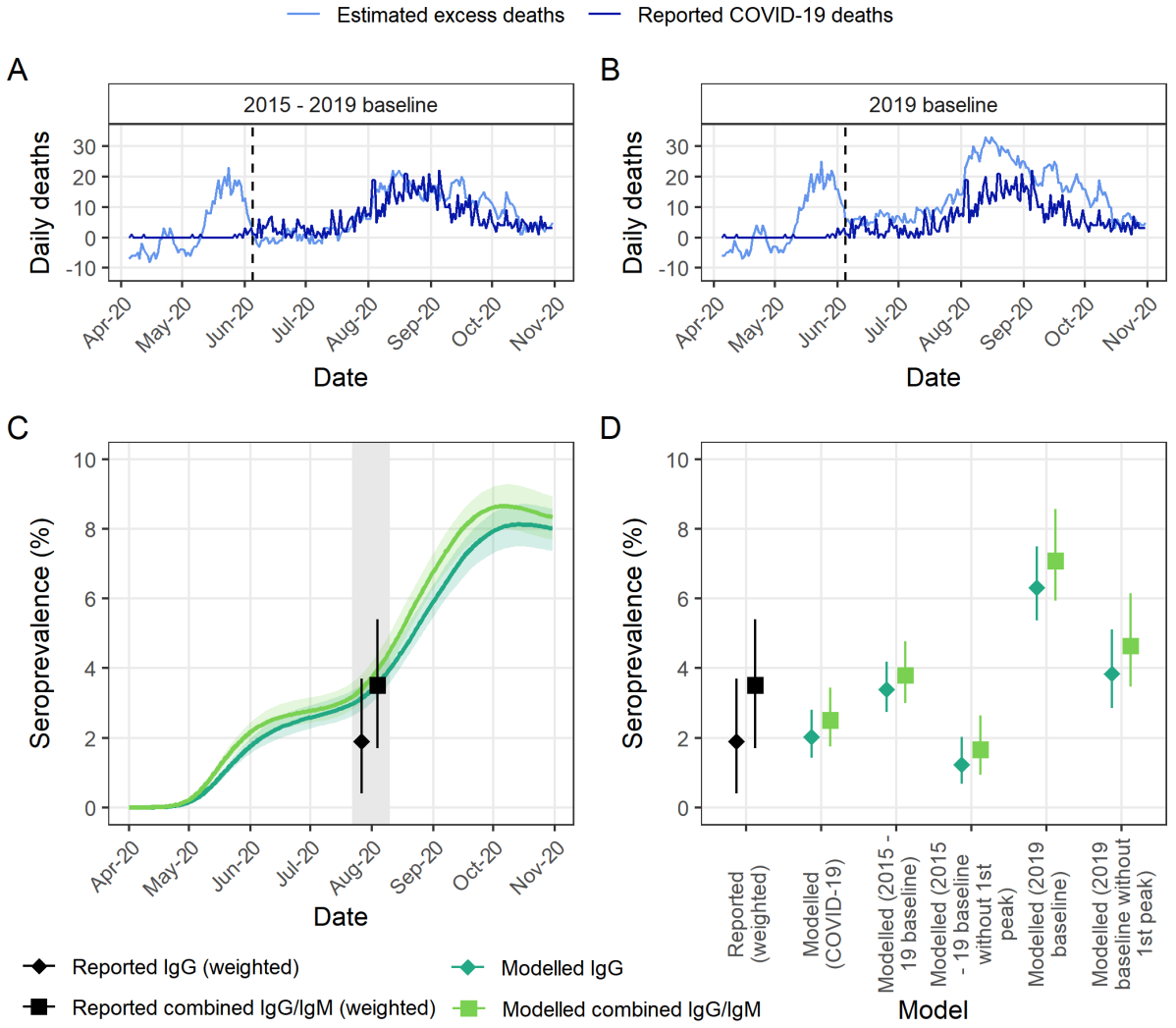

Figure 7.1: Mortality reporting and seroprevalence in Addis Ababa, Ethiopia in 2020. (A) Reported COVID-19 deaths compiled from the Ethiopian Public Health Institute and estimated excess mortality using cemetery surveillance from [484] using data from 2015 - 2019 to derive the baseline. Dashed line indicates the end of the early peak in excess mortality featured in the sensitivity analysis. (B) As in (A), but with excess mortality derived from 2019 data only. (C) Estimated seroprevalence from model fit to excess mortality under 2015 - 2019 baseline (median and 95% credible intervals) with the first peak compared to reported values by Abdella et al. [486]. Grey shaded area highlights the sampling period of the serosurvey, with weighted reported estimates both corresponding to this entire period. (D) Seroprevalence estimated under models fit to either reported COVID-19 or different estimates of excess mortality with different baselines (median and 95% credible intervals) compared to reported seroprevalence in [486].

7.2 Addis Ababa

In Addis Ababa, excess mortality was derived from cemetery surveillance data collected via the Addis Ababa Mortality Surveillance Program across January 2015 - January 2021, as detailed by Endris et al. [484]. These data comprise the number of burials per day at all cemeteries in the city according to official records kept by workers at each site. As cremation is not practised in Addis Ababa, this

surveillance programme is expected to capture the vast majority of deaths from all causes in the city [484].

Reported COVID-19 deaths [488] largely matched estimated excess mortality from June 2020 onwards, though this was in part dependent on how baseline mortality was estimated (Figures 7.1A and 7.1B). While the annual number of burials in each year from 2015 to 2018 was largely consistent (average $12,862$; sd 389), there was a significant reduction in the number of burials observed in 2019 compared to the mean from 2015 - 2018 (total $11,256$; chi-squared test $p < 0.001$, Figure E.3). Therefore, we explored two methods for estimating baseline mortality in 2020, in which we used mortality data either from 2015 to 2019 or exclusively from 2019 (Spearman correlation between excess mortality and COVID-19 deaths over our study period: $r = 0.59$ (using 2015 to 2019 as baseline) and $r = 0.69$ (using just 2019 as baseline)). Regardless of the baseline, we observed a peak in excess mortality observed in May 2020 that was not associated with a peak in reported COVID-19 deaths (Figures 7.1A and 7.1B). We explored whether this peak could be due to COVID-19 through an additional sensitivity analysis by starting the model from 5 June 2020 instead of 6 April 2020 (the date of the first COVID-19 death) (Figure 7.1D) (Spearman correlation between excess mortality and COVID-19 deaths from June 2020 onwards: $r = 0.74$ and $r = 0.75$ for 2015-to-2019 and 2019-only baselines, respectively). Therefore, we considered four excess mortality scenarios in total for Addis Ababa, which we assumed to be equally plausible a priori.

We fitted the mathematical model separately to reported COVID-19 deaths and the four estimated excess mortality time series and compared the number of infections estimated by the model to those at the time of a serosurvey conducted by Abdella et al. [486] in July - August 2020. The study used random sampling of 956 households in the city and reported a weighted prevalence of Immunoglobulin G (IgG) and combined IgG/IgM antibodies of 1.9% ($0.4\% - 3.7\%$) and 3.5% ($1.7\% - 5.4\%$), respectively, having post-stratified by age and sex to account for potential biases in their sample (see Appendix E for further information).

The model fit to COVID-19 deaths produced estimates of seroprevalence not statistically different from the values reported by Abdella et al. [486] ($p = 0.884$ and $p = 0.334$ for IgG and combined IgG/IgM, respectively), suggesting that officially reported deaths are representative of the true COVID-19 death toll in this setting. Similarly, we observed that the majority of the seroprevalence estimates based on model fits to the excess mortality inferred from cemetery burial data were not statistically significantly different from the reported seroprevalence in Abdella et al. [486] for both antibody types (Table E.2). However, this is expected given the similarity in the COVID-19 and excess mortality time series (Figures 7.1A and 7.1B). The exception to this was the model fit to excess mortality inferred using 2019 only as

the baseline and including the first peak ($p < 0.001$ and $p = 0.002$ for IgG and combined IgG/IgM antibodies, respectively), with the resulting model-estimated seroprevalence three times and twice as great as the reported prevalence of IgG and combined IgG/IgM antibodies of Abdella et al. (Figure 7.1D). This suggests that excess mortality under this baseline scenario overestimated COVID-19 mortality and is unsuitable for informing the size of the COVID-19 epidemic. Consequently, we approximate that, depending on the selection of mortality baseline, between 68.7% (1,064 of 1,549, reflecting the 2015 - 2019 scenario with the first peak in May included) - 100% of COVID-19 deaths were reported across April - November 2020 in Addis Ababa. The upper estimate of 100% (complete) reporting reflects both the near agreement between the cemetery-inferred excess mortality and reported COVID-19 deaths, and the non-significant difference between the seroprevalence inferred based on model fits to reported COVID-19 deaths and the reported seroprevalence (Figure 7.1C, Table E.5).

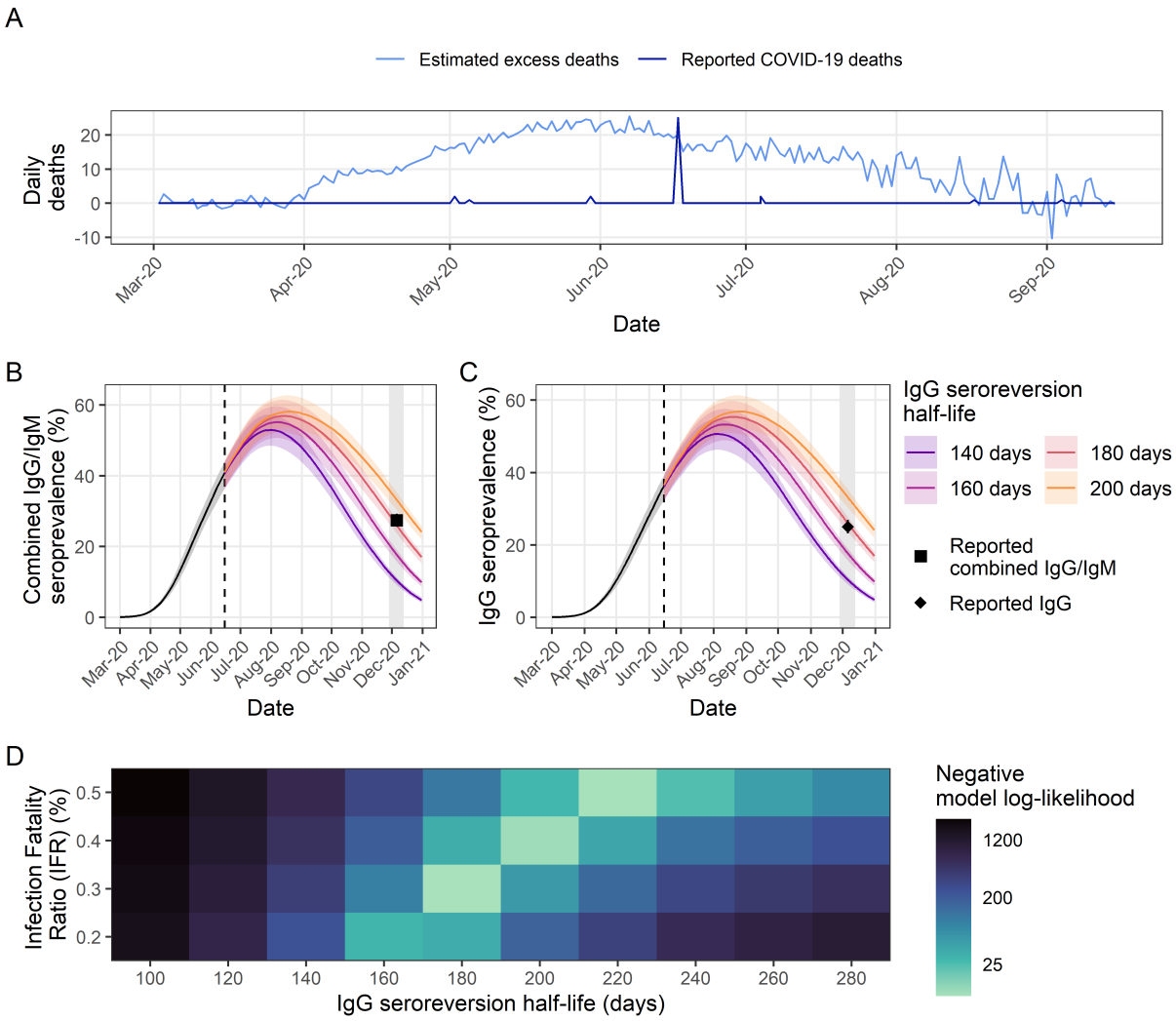

Figure 7.2: Mortality reporting and seroprevalence in Aden, Yemen in 2020. (A) Reported COVID-19 deaths and estimated excess mortality from satellite surveillance of burials from [220]. (B) Estimated seroprevalence of combined IgG/IgM (positive for either IgG and/or IgM) from the model fitted to excess mortality using the default Infection Fatality Ratio (IFR = 0.3%) under different IgG seroreversion half-lives compared to observed values from [256](reported seroprevalence: 27.4% (95% CI 25.6% – 29.3%)). Dashed vertical line indicates point in time when seroprevalence dynamics divert due to difference in IgG half life modelled. (C) As in (B) but for IgG antibodies (reported seroprevalence: 25.0% (95% CI 23.2% – 26.9%)). (D) Combined negative log-likelihood of models of IgG and combined IgG/IgM antibodies under varying assumptions of the IFR and IgG seroreversion half-life. In (B) and (D) the IgM seroreversion half-life is held constant at 50 days. Values associated with the highest log-likelihoods (shown here in light-blue) indicate the best fit of the parameter values to the observed data.

7.3 Aden

In Aden, excess mortality was estimated from satellite imagery surveillance of all cemeteries in the city to estimate daily burials between July 2016 - September 2020, by quantifying expansions to cemetery surface area and validating this with civil death registrations [220]. From these data, a wave of excess mortality was estimated to have occurred between April - September 2020, peaking in early June. Reported COVID-19 deaths were only available via scraping death counts tweeted daily by the Yemen Supreme

National Emergency Committee for COVID-19 [489]. Excess mortality estimates suggested $2,120$ (95% CI: $424 - 4,137$) excess deaths occurred across 1st April - 19 September 2020, compared to just 34 officially reported COVID-19 deaths during this same period [489] implying substantial under-ascertainment of COVID-19 mortality (Figure 7.2A).

In December 2020, Bin-Ghouth et al. [256] conducted a cross-sectional household serosurvey of $2,001$ people and estimated that 25.0% (95% CI: $23.2\% - 26.9\%$) and 0.2% (95% CI: $0.1\% - 0.4\%$) of the population had either IgG or IgM antibodies to SARS-CoV-2, respectively, with a total of 27.4% (95% CI: $25.6\% - 29.3\%$) estimated to have IgG and/or IgM antibodies (combined IgG/IgM). Because of the long delay of approximately five months between the peak in excess mortality and the implementation of the serosurvey in this setting, it was necessary to consider the decay of antibodies after infection over time. Antibody kinetics vary substantially among individuals resulting in high uncertainty around the duration of seropositivity after infection [248]. To the best of our knowledge, these parameters had not been quantified for the assay used in this study and so we conducted sensitivity analyses to explore the impact of different assumptions for IgG and IgM seroreversion. IgM antibodies endure for a substantially shorter duration than IgG antibodies. As such, we found that varying the IgM half-life had a minimal impact in the context of this analysis and thus assumed a central value of 50 days as in [44] (statistically insignificant differences with reported values, Figure E.11, Table E.3), but that the IgG half-life was an important determinant of our estimated seroprevalence (Figure 7.2, Table C.3).

Seroprevalence inferred from the model fit to reported COVID-19 deaths was statistically significantly different from that reported for all antibody types and under all seroreversion half-lives considered ($p < 0.001$ for IgG and combined IgG/IgM, $p < 0.009$ for IgM), suggesting that official reported COVID-19 deaths substantially underestimate the true death toll. However, in model fits to the satellite-derived excess mortality, we found no statistically significant differences between reported and modelled seroprevalence of IgG and combined IgG/IgM antibodies ($p = 0.651$ and $p = 0.563$, respectively) with an assumed IgG seroreversion half-life of 180 days. This finding suggests that these data are a more accurate representation of the first COVID-19 outbreak in Aden than officially reported deaths.

Although the default IFR (IFR = 0.3%) and IgG seroreversion half-life (180 days) maximised the combined model log-likelihood of IgG and combined IgG/IgM antibodies, we were able to find other parameter combinations which could also explain the reported seroprevalence (Figure 7.2). We found that a slightly higher IFR of 0.4% or 0.5% combined with a slightly longer seroreversion half-life of 200 or 220 days, respectively, also produced non-statistically significant differences between reported and modelled sero-

prevalence of both antibody types ($p = 0.289$ and $p = 0.649$ for IgG and $p = 0.944$ and $p = 0.385$ for combined IgG/IgM antibodies, respectively) and produced similarly high log-likelihoods (Figure 7.2, Table E.3). Although a lower IFR of 0.2% was able to also produce non-statistically significant differences when using shorter seroreversion half-lives (Table E.3), the model was unable to recreate the excess deaths from the satellite data, suggesting that IFR in Aden was greater than 0.2% (Figure E.10).

Overall, the broad agreement between the two independent data sources under plausible assumptions of antibody seroreversion and the IFR strongly supports COVID-19 to be the driver of the observed satellite-inferred excess mortality. Consequently, we estimate that just 1.6% (34/2,120) of COVID-19 deaths in Aden in summer 2020 were captured in official statistics (Table E.5). Based on the uncertainty in the satellite-inferred excess mortality (95% CI: $424 - 4,137$), we report our uncertainty in the reporting fraction to be between $0.8\% - 8.0\%$.

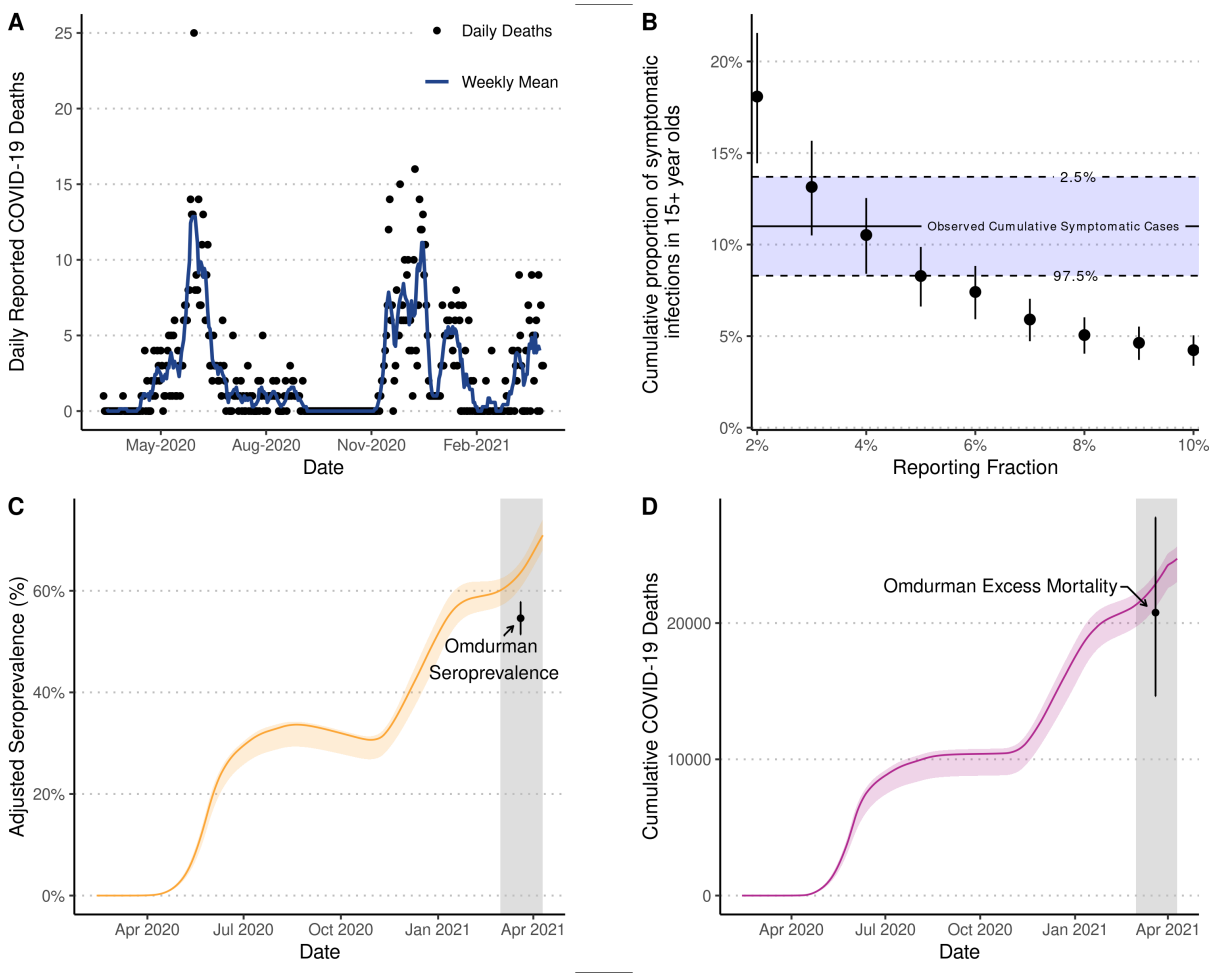

Figure 7.3: Estimates of under-ascertainment of deaths in Khartoum. (A) Daily and weekly mean reported COVID-19 deaths in Khartoum. We fit models to the reported COVID-19 deaths in (A) under different assumptions for what proportion of the true number of COVID-19 deaths these represent (reporting fractions). In (B) the resultant model fits were used to estimate the proportion of individuals aged over 15 years that would have experienced a symptomatic COVID-19 infection by 3 June 2020. The points and vertical bars show the median and 95% CI for each model fit and are compared against the observed cumulative number of symptomatic cases in Khartoum estimated from a social-media-distributed survey [485], suggesting that 4% of COVID-19 deaths were detected. A cross-sectional, household mortality and serosurvey conducted in Omdurman in March - April 2021 estimated seroprevalence to be 54.6% (95% CI: $51.4\% - 57.8\%$) after adjusting for specific test performance [487]. This estimate is depicted in (C) by point and whiskers and is compared against the adjusted seroprevalence (median and 95% CI shown in orange) predicted by a model fit with an assumed mortality reporting fraction of 4%. Based on the same survey, $20,766$ (95% CI: $14,641 - 27,750$) excess deaths are estimated to have occurred across Khartoum state. This estimate is depicted in (D) by point and whiskers and is compared against the cumulative number of COVID-19 deaths (median and 95% CI shown in purple) predicted by a model fit with an assumed mortality reporting fraction of 4%. In (C) and (D) the grey shaded area highlights the sampling period of the serosurvey and mortality survey in Omdurman.

7.4 Khartoum

In Khartoum, we leveraged an online survey distributed through social media channels [485]. The survey presented respondents with a list of symptoms known to be associated with SARS-CoV-2 infection and used to triage patients in Sudan. Respondents were asked which of these symptoms they had experienced

since the start of the pandemic. Respondents were also asked both if they had received i) a SARS-CoV-2 diagnostic test and ii) the outcome of any test taken. From this survey, the SARS-CoV-2 infection status of survey participants who had not received a COVID-19 test was inferred based on their reported symptom(s), resulting in an estimate of the symptomatic attack rate in the general population. The survey collected responses opportunistically, resulting in $5,018$ responses from individuals over 15 years old throughout Khartoum State between 26 May and 3 June 2020. From these respondents, 11.0% of individuals were estimated to have experienced a symptomatic infection by 3 June, with lower and upper estimates (based on different methods for inferring symptomatic infection) of 8.3% and 13.7%.

By the end of March 2021 (prior to a pause in publication of COVID-19 reports by the Sudan Federal Ministry of Health), 938 COVID-19 deaths had been reported in Khartoum (Figure 7.3A). We fit the transmission model to the observed COVID-19 deaths making the assumption that the reported COVID-19 deaths reflected a fixed proportion of the true total number of COVID-19 deaths over time (Figures E.12 and E.13). From these, our central estimate is that 4% of COVID-19 deaths were reported based on comparison against the 11.0% of individuals estimated to have experienced a symptomatic infection by 3 June (Figure 7.3B). We report our uncertainty in the reporting fraction to be between $3\% - 6\%$, with the modelled 95% confidence interval for the cumulative proportion of symptomatic infections overlapping with the lower or upper bounds reported of 8.3% and 13.7% (Figure 7.3B).

Between March 1 and April 10 2021, a cross-sectional household mortality and seroprevalence survey was conducted in Omdurman - a city in Khartoum state estimated to encompass $\sim 35\%$ of the population of Khartoum state [487]. In this survey, 54.6% (95% CI: $51.4\% - 57.8\%$) of the population were seropositive and $7,113$ excess deaths (95% CI: $5,015 - 9,505$) were estimated to have occurred, which when scaled linearly to reflect the population of Khartoum state would be $20,766$ (95% CI: $14,641 - 27,750$) excess deaths [487]. By comparison, with 4% of COVID-19 deaths assumed to have been detected, we estimate that 63.6% (95% CI: $60.6\% - 65.7\%$) of the population would have been seropositive by March 20 2021 (Figure 7.3C), which is significantly different to the reported seroprevalence of 54.6% (chi-squared test $p < 0.001$, Table E.4), and that $22,870$ (95% CI: $21,150 - 23,640$) deaths would have occurred (Figure 7.3D). Consequently, while the inferred reporting fraction of 4% based on the symptomatic survey yields estimates of the epidemic size that are similar to those observed in Omdurman almost a year later, estimates of the epidemic size are most accurately recreated with an assumed reporting fraction of 4.5% (Figure E.14, Table E.5).

7.5 Discussion

In this study, we demonstrate the validity of using alternative data sources, namely burial site worker reports in Addis Ababa [484], satellite imagery of cemeteries in Aden [220] and social-media-conducted surveys of symptomatic infection in Khartoum [485], to track the dynamics and burden of the COVID-19 epidemic in each of these settings. By incorporating these data within a previously-published mathematical modelling framework [58], we could infer the number of infections implied by each source and validate the accuracy of these results by comparing them with independently-conducted serosurveys. Consequently, we estimate that between $67\% - 100\%$ of COVID-19 deaths were reported in Addis Ababa, while in Aden and Khartoum, we demonstrated that there are likely to have been large unobserved epidemics with $0.8\% - 8.0\%$ and $3.0\% - 6.0\%$ of deaths reported, respectively (Table E.5).

Understanding the impact and burden of an infectious disease is integral to facilitating an effective public health response. Our results challenge the perception of the global burden of the disease as reported in official figures which, as shown here in the cases of Aden and Khartoum, can suffer from under-reporting in resource-poor settings. In each setting, we found statistically insignificant differences between our model-estimated seroprevalence based on the alternative data source and those reported in serosurvey, each of which tested at least 950 participants. This provides reasonable confidence in the accuracy of these sources to describe and understand the respective epidemic dynamics (Tables E.2 - E.4). Hence, we consider that these alternative data sources represent a viable, and likely cost-effective, route to estimating disease burden and impact in settings where high-quality civil registration system mortality data are not currently available.

Our framework utilises the IFR when comparing mortality inferred from alternative data sources to reported seroprevalence. Although this parameter has not been estimated directly within the three settings in this study, we have demonstrated that the default IFR accounting for the demography in Addis Ababa (IFR = 0.22%) estimated by Brazeau et al. [44] accurately captured the reported seroprevalence. However, in Aden, our results suggested that the IFR is at least 0.3%, which is higher than the IFR estimated by Brazeau et al. [44] after accounting for the demography in Aden (IFR = 0.17%). Similarly, in Khartoum, an IFR of approximately 0.38% accurately captured the reported seroprevalence, which is notably higher than the default IFR estimated by Brazeau et al. [44] after accounting for the demography in Khartoum (IFR = 0.20%). These studies should not be viewed as an accurate assessment of the IFR, with the mortality inferred from these alternative data unlikely to perfectly reflect the true COVID-19 mortality in each setting. Nonetheless, in addition to demonstrating these data are indeed informative for describing and understanding epidemic dynamics, our results provide further evidence that the re-

lationship between age and IFR that has been consistently observed in high-income countries [44, 259] also exists in these low-income settings. It is further possible that the aforementioned relationship with age could lead to an underestimation of the IFR in some settings, as suggested by a recent meta-analysis of data from developing countries [481] and supported by our presented analyses of data from Aden and Khartoum.

Our results could be further corroborated via sources other than those directly used in our analyses. In Aden, multiple reports of hospital capacity being reached coincided with the epidemic peak estimated via satellite image-derived excess mortality [220, 490]. The saturation of healthcare facilities could also explain the higher IFR indicated in this setting. A similar situation may also have occurred in Khartoum, with over 70% of health centres closed in Khartoum as a COVID-19 containment measure during the first wave [491]. Additional serosurveys have also been conducted in each setting. In Khartoum, a seroprevalence survey conducted in 22 neighbourhoods and relying on voluntary enrolment organised through resistance committees estimated a seroprevalence of 18.3% (95% CI: $16.0\% - 20.9\%$) using rapid antibody immunochromatography tests between 22 May - 5 July 2020 [492]. Polymerase chain reaction (PCR)-confirmed prevalence of infection in the same survey was estimated at 35.0% (95% CI: $32.1\% - 38.0\%$), raising concerns that the sampling scheme had resulted in an upward bias. However, the seroprevalence inferred from our model fits with an assumed reporting fraction of 4% was consistent with this survey ($p = 0.493$), confirming previous viral kinetics modelling of this study [493]. In Addis Ababa, higher seroprevalence of combined IgG/IgM antibodies has been estimated in other surveys [494], which estimated a value of 10.9% in August 2020, with this value rising to 53.7% in December 2020. However, these estimates were derived from sampling focused on healthcare workers, who are known to have a higher risk of exposure to SARS-CoV-2 than the general population [169]. Similarly, samples collected from Médecins Sans Frontières staff in Aden between September - November 2020 and tested using rapid serology lateral flow tests resulted in an estimated seroprevalence of 19.4% (95% CI: 17.9%–20.8%) [495]. However, a follow-up survey of Médecins Sans Frontières staff in Aden during January 2021 (after a period with very few COVID-19 cases reported in Aden [489]) but analysed using an electrochemiluminescence immunoassay resulted in an estimated seroprevalence of 59.0% (95% CI: 52.2%–65.9%) [495]. This study both shows high exposure in healthcare staff but also demonstrates the importance of accounting for waning rapid test sensitivity. Lastly, excess mortality published by the WHO [483] yielded similar magnitudes to those estimated by our analyses in Aden and Khartoum. However, we estimate a much greater reporting fraction in Addis Ababa than the corresponding national estimate (Table E.5, Table E.6 and Figure E.15), highlighting how death registration measured in individual cities or locations often will not reflect the realities of the whole countries, especially during periods of crises.

There are a number of limitations of this analysis. First, our estimated reporting fractions correspond to the entire study period and do not account for time-varying trends in detection of COVID-19 deaths. Second, there is uncertainty as to whether all excess deaths can be attributed directly due to COVID-19 [5, 224, 496], which is why we have conducted extensive sensitivity analyses to assess the robustness of our analysis in all three settings. Third, while we have demonstrated strong agreement between the alternative sources and seroprevalence estimates, it is still possible that the mortality inferred from the alternative source does not capture all deaths in each setting. For example, the Global Burden of Disease (GBD) [497] model suggests higher annual mortality in Addis Ababa than that presented in the cemetery surveillance data used in this study. While the GBD is a model-based estimate and the Addis Ababa Mortality Surveillance Program has shown to be representative of historic censuses [498], it is still likely that at least some deaths are missed by the cemetery data. If the true excess death toll in Addis Ababa is much higher than found in the Addis Ababa Mortality Surveillance Program, this would suggest that the IFR in Addis Ababa would need to be higher than in this analysis to be in agreement with the observed seroprevalence. Fourth, SeroTracker [254] suggested all three serosurveys may be subject to "moderate" bias, with corresponding estimates defined as "likely correct for the target population", as opposed to "very likely correct" for studies with "low" bias, based on nine criteria including sampling technique, sample size and consistency of reporting results [254, 499]. Nonetheless, throughout our study we have considered the full range of uncertainty estimated in the serosurveys and also estimated from our model fits, with these intervals shown to overlap under certain assumptions in each study. Analogously to the mortality data, we have focussed on comparing the range of uncertainty intervals as opposed to point estimates alone. This is also applicable to the social media-based survey in Khartoum, which may suffer from self-reporting biases. Fifth, SARS-CoV-2 assays are often developed on high-income populations and the applicability to other settings has been questioned [500], which could influence the results. Sixth, we were unable to obtain specific seroconversion and seroreversion rate estimates for the assays used in these studies and had to rely on a sensitivity analysis around estimates derived from studies of other assays. Similarly, our modelling framework operates under the assumption of independence between the likelihood of the detection of IgG and combined IgG/IgM antibodies, due to a lack of data to prove otherwise. These assumptions may explain why we were unable to recreate the relatively large difference between IgG and combined IgG/IgM antibodies observed in the serosurvey in Addis Ababa, despite adjusting for different seroreversion half-lives of these two antibody measurements. Crucially, our qualitative conclusions about the significant degree of underreporting and the usefulness of these alternative data sources are robust to the limitations described above. Consequently, while there are likely biases in the alternative data sources leveraged, we argue that these biases are substantially smaller than those in

official reported COVID-19 statistics and are thus vitally useful as proxies for COVID-19 mortality and epidemic dynamics.

The COVID-19 pandemic has caused a substantial loss of life, but reported deaths are likely to only capture a small proportion of the true death toll in settings without robust vital registration. In the absence of all-cause mortality data, we have validated the suitability of alternative data sources, namely burial site worker reports, satellite imagery of cemeteries and social media-conducted surveys of symptomatic infection, in tracking epidemics in Addis Ababa, Aden and Khartoum across 2020. These sources provide a critical insight into the dynamics of SARS-CoV-2 in these settings and contradict the hypothesis that low-income countries were spared the worst of the pandemic. Our flexible, data-driven modelling framework is readily adaptable to other settings in which there are alternative sources of mortality and estimates of seroprevalence, such as Lusaka [501] and Haiti [502]. However, while the modelling framework can be adapted to other settings and further alternative data sources, not all data sources will be suitable in other locations. For example, cemetery burial data are unlikely to be representative in settings that practice cremation and satellite imagery approaches may be less reliable in settings with extensive cloud coverage.

The global community must prioritise providing support and investment in the development of vital registration systems to capture all-cause mortality across the globe. Better vital registration is not only essential for future pandemic preparedness plans but also is foundational for evaluating progress made across multiple areas of public health and is essential for the Sustainable Development Goals' mission to "leave no-one behind" [503].

7.6 Materials and Methods

We explain the general methodology, applicable to each of the three settings. Information on the specific datasets and sources for each modeled city are presented in Appendix E. A schematic diagram of the methods is shown in Figure E.1. All data and code are provided at `https://github.com/mrc-ide/covid-alternative-mortality/` [480].

7.6.1 Mathematical model of SARS-CoV-2 transmission and disease progression

We use a previously published age-structured SARS-CoV-2 transmission model [2] and fitting framework [58] to fit the weekly estimated excess deaths in each setting in our study. In overview, the model is

a Susceptible-Exposed-Infected-Recovered-Susceptible compartmental model, which is population-based and age-structured. The model explicitly represents disease severity and resultant passage through different healthcare levels, with an assumed elevated severity when healthcare capacity is exceeded as defined in Walker et al. [2]. The model is capable of modelling vaccinations (see [504]), but in all settings considered in this study, vaccination campaigns had not started.

Model fitting was carried out within a Bayesian framework using a Metropolis-Hastings Markov Chain Monte Carlo (MCMC) based sampling scheme, which estimates the epidemic start date, R_0, and the time varying reproduction number, $R(t)$, using a series of pseudo-random walk parameters (ρ_n), which alter transmission every 2-weeks, given by:

$$R(t) = R_0.f(-\rho_1 - \rho_2 - ... - \rho_n)$$

where $f(x) = \frac{2.exp(x)}{1+exp(x)}$. Each random walk parameter is introduced two weeks after the previous parameter, serving to capture changes in transmission every two weeks. The last change in transmission, n, is maintained for the last 4 weeks prior to the current day to reflect our inability to estimate the effect size of this parameter due to the approximate 21-day delay between infection and death. Each model fit is tailored to each setting, incorporating demographics, with the population size in 5-year age bands, and the effective number of general hospital beds and intensive care beds for each city.

7.6.2 Estimation of seroprevalence

Seroprevalence over time was derived from the total number of infections estimated by the model, adjusted for rates of seroconversion, seroreversion and serological assay sensitivity.

In each setting, we assumed that seroconversion followed an exponential distribution with mean time to seroconversion of 13.3 days for IgG antibodies [505]. Seroconversion for IgM antibodies was assumed to occur more quickly, with mean time to seroconversion of 12.3 days [505]. Similarly, in each setting seroreversion was assumed to follow a Weibull distribution with shape parameter 3.7 [44] and with the scale parameter adjusted to enforce a specific half-life for IgG or IgM antibodies, respectively. As default, we assumed a half-life of 50 days and 140 days for IgM and IgG, respectively, as estimated in Brazeau et al. [44]. In Addis Ababa, a negligible amount of time had passed between the start of the epidemic inferred from the alternative mortality source and the serosurvey period and consequently the assumed duration of seroreversion does not impact our findings. However, because in Aden there was a substantial amount of time between the excess mortality and seroprevalence studies, we instead conducted a sensitivity analysis by varying the half-life of IgG antibodies between 100 and 280 days and of IgM antibodies

between 30 and 70 days to capture the uncertainty surrounding the true value of these parameters. Last, in Khartoum, the observed seroprevalence estimated by Moser et al. [487] for Omdurman has already been adjusted to account for decreasing diagnostic test sensitivity over time as a result of waning antibody titres, which resulted in an increase from 34.3% crude seroprevalence to an adjusted seroprevalence of 54.6% [487]. Consequently, we estimate seroprevalence from our model using a different approach. We continue to model the lag from infection to seroconversion, with a mean of 13.3 days. However, we consider individuals to only become seronegative again once they have moved from the recovered infection compartment to the susceptible compartment (i.e. by definition they should not possess antibodies to detect), which has a mean duration of 365 days and is described by an Erlang-2 distribution, i.e. is the sum of two independent exponential distributions each with a mean duration of 365/2 days.

Seroreversion and seroconversion distributions for combined IgG and IgM antibodies was estimated by treating the IgG and IgM distributions as independent. Although this assumption is unlikely to hold in reality, we validated our approach using Monte Carlo simulation which produced an essentially identical probability distribution of seropositivity when the two antibody types were dependent as to that when they were treated independently (Figure E.2).

7.6.3 Testing for differences between observed and expected seroprevalence

Chi-squared tests were used to determine whether or not the difference between observed and expected seroprevalence was statistically significant. Let θ_1 denote reported seroprevalence and let θ_2 denote our modelled seroprevalence. Our test statistic is:

$$T = \frac{(\theta_1 - \theta_2)^2}{var(\theta_1) + var(\theta_2)}$$

where

$$var(\theta_1) \approx (\frac{\text{upper confidence limit} - \text{lower confidence limit}}{2 \times 1.96})^2$$

with 95% confidence interval limits those reported in the serosurveys.

Under the null hypothesis $T \sim \chi^2_1$, where χ^2_1 denotes a Chi-squared distribution with 1 degree of freedom, we assume that there are no statistically significant differences between θ_1 and θ_2. A p-value (p) greater than 0.05 indicates that there is not enough evidence to reject the null hypothesis, and we conclude that there are no statistically significant differences between the reported and modelled values.

7.6.4 Estimation of reporting fractions

Reporting fractions are defined as total COVID-19 deaths over the period divided by total positive excess deaths over the same period (Table E.5).

Chapter 8

Paper VI: Inferring community transmission of SARS-CoV-2 in the United Kingdom using the ONS COVID-19 Infection Survey

R McCabe, G Danelian, J Panovska-Griffiths and CA Donnelly. Inferring community transmission of SARS-CoV-2 in the United Kingdom using the ONS COVID-19 Infection Survey. *Infectious Disease Modelling*, 2024. doi: `https://doi.org/10.1016/j.idm.2024.01.011`. [*In press.*]

Abstract

Key epidemiological parameters, including the effective reproduction number, $R(t)$, and the instantaneous growth rate, $r(t)$, generated from an ensemble of models, have been informing public health policy throughout the COVID-19 pandemic in the four nations of the United Kingdom of Great Britain and Northern Ireland (UK). However, estimation of these quantities became challenging with the scaling down of surveillance systems as part of the transition from the "emergency" to "endemic" phase of the pandemic.

The Office for National Statistics (ONS) COVID-19 Infection Survey (CIS) provided an opportunity to continue estimating these parameters in the absence of other data streams. We used a penalised spline model fitted to the publicly-available ONS CIS test positivity estimates to produce a smoothed estimate of the prevalence of SARS-CoV-2 positivity over time. The resulting fitted curve was used to estimate the "ONS-based" $R(t)$ and $r(t)$ across the four nations of the UK. Estimates produced under this model

are compared to government-published estimates with particular consideration given to the contribution that this single data stream can offer in the estimation of these parameters.

Depending on the nation and parameter, we found that up to 77% of the variance in the government-published estimates can be explained by the ONS-based estimates, demonstrating the value of this singular data stream to track the epidemic in each of the four nations. We additionally find that the ONS-based estimates uncover epidemic trends earlier than the corresponding government-published estimates.

Our work shows that the ONS CIS can be used to generate key COVID-19 epidemiological parameters across the four UK nations, further underlining the enormous value of such population-level studies of infection. This is not intended as an alternative to ensemble modelling, rather it is intended as a potential solution to the aforementioned challenge faced by public health officials in the UK in early 2022.

Author contributions

RM conceived the idea for the study following discussions with GD, JPG and CAD. **RM** conducted the analysis and drafted the manuscript, with all authors contributing to redrafting.

Reproducibility

All data and code used are available from `https://github.com/ruthmccabe/ons-test-positivity-model`.

Acknowledgements

RM was supported by the UK Health Security Agency (UKHSA) and Isaac Newton Institute (INI) Knowledge Transfer Network (KTN) in funding and coordinating a 3-month placement at the UKHSA, respectively.

Appendix

Further information relating to this paper can be found in Appendix F.

8.1 Introduction

Key epidemiological parameters, including the effective reproduction number, $R(t)$, and the instantaneous growth rate, $r(t)$, have been used to inform public health policy throughout the COVID-19 pandemic [232–234, 506]. Estimation of these quantities by public health officials in the United Kingdom (UK) has relied on an ensemble of models which encompass a range of data sources and assumptions [65]. These parameters are traditionally estimated using case data as a proxy for the infection incidence curve [229, 230], but methods have also been developed to estimate these parameters from other sources, such as hospitalisations [278] and genomic data [276]. The four nations of the UK (England, Scotland, Wales and Northern Ireland, herein ordered by population size) were recognised globally as having comprehensive SARS-CoV-2 testing surveillance systems [135, 225, 226], comprising of widescale community testing [112], nationwide surveys of infection [63, 235], genomic data [507, 508], and wastewater surveillance [509], but these were largely scaled down from their peak capacity as part of the transition from an "emergency" to "endemic" state in the first half of 2022 [59–62]. Consequently, estimation of $R(t)$ and $r(t)$, particularly using an ensemble model approach, became more challenging due to the reduction in available data streams.

The Office for National Statistics (ONS) COVID-19 Infection Survey (CIS) [63] was a primary means by which to understand, quantify the impacts of and track SARS-CoV-2 transmission within the UK [241, 242, 244, 246]. This COVID-19 testing study invited members of randomly selected private households across the UK to complete polymerase chain reaction (PCR) tests, regardless of symptoms or behaviour. The ONS CIS continued for one year after the cessation of community testing, until its "pause" in March 2023 [136], and in October 2023 it was announced that a similar study would be undertaken for the upcoming 2023/24 winter period [137]. Although the survey can estimate incidence of infection, this lags the main metric of interest, the percentage of the population testing positive for SARS-CoV-2 infection, which is used as a proxy for the prevalence of the virus in each nation of the UK. These data, widely regarded as "gold-standard" due to the random sampling methods underpinning them, provided an opportunity to continue estimating $R(t)$ and $r(t)$ after the scaling down of many of the surveillance systems described above.

By adapting the methods deployed by another community surveillance survey with randomly-selected participants, the REal-time Assessment of Community Transmission (REACT) study [235, 510–513], this paper presents a model to estimate $R(t)$ and $r(t)$ directly from publicly-available ONS CIS estimates of test positivity in each nation of the UK. The estimates produced under this model are compared to government-published ensemble estimates, to assess the validity of this method to track the spread of

SARS-CoV-2 in the absence of other surveillance data. In particular, we consider the level of contribution that this single data stream can offer in the estimation of these parameters. This methodology is then used to provide estimates of $R(t)$ and $r(t)$ in each nation of the UK until the pause of the ONS CIS in the first quarter of 2023.

8.2 Methods

8.2.1 Estimating $R(t)$ and $r(t)$ using the ONS COVID-19 Infection Survey (CIS)

8.2.1.1 The ONS CIS

The ONS CIS was the only long-term SARS-CoV-2 testing study in randomly selected households encompassing all four nations of the UK. In brief, private households were randomly selected across each nation, irrespective of factors such as members displaying symptoms or having contact with a known case, and household members aged over 2 years were invited to complete multiple PCR tests over time. The random sampling produced estimates that were unaffected by test-seeking behaviour and public availability of diagnostic tests. Further details regarding sampling can be found in [514].

The primary outcome of interest was the estimated percentage of people testing positive for SARS-CoV-2, derived from the number of positive tests out of the total tests completed and post-stratified by key variables, such as gender and ethnicity. Until 29 July 2022, the raw numbers of positive and total tests were not publicly available, although they have since been reported retrospectively for the entire study period [515–518]. The publication of corresponding estimates of incidence of infection, based on the primary outcome (prevalence of test positivity), was lagged by a couple of weeks, as this is substantially more complex to estimate, but ceased to be published from June 2022 [519]. Initially, the CIS had overlapping reporting windows of approximately 10–14 days but reporting settled into approximately weekly windows.

In this study, we focus exclusively on the publicly-available estimates of test positivity provided as point estimates with 95% credible intervals, derived from the post-stratification model [514], due to these being consistently available to the public in real-time for the majority of the ONS CIS study period. Furthermore, these estimates are unaffected by issues concerning deductive disclosure, in which the number of positive tests cannot be disclosed due to their small number, as experienced by both Wales and Northern Ireland at the beginning of the study in 2020.

The survey commenced on different dates in each nation (Table 8.1) and there is some heterogeneity

Table 8.1: Availability of estimates used in this study: test positivity from the ONS CIS survey and official estimates of the effective reproduction number and growth rate as published by the UK government. The study period for each nation is determined as the beginning date of the ONS CIS until the "pause" date of the ONS CIS (March 2023), while the period of overlap states the period for which both ONS-based estimates and government-published estimates are available for comparison.

Nation	**ONS CIS coverage**	**Parameter**	**Government-published estimate coverage**	**Period of overlap**
England	27/04/20 - 13/03/23	Reproduction number	29/05/20 - 23/12/22	29/05/20 - 23/12/22
		Growth rate	12/06/20 - 23/12/22	12/06/20 - 23/12/22
Scotland	03/10/20 - 13/03/23	Reproduction number	28/05/20 - 08/12/22	03/10/20 - 08/12/22
		Growth rate	18/06/20 - 08/12/22	03/10/20 - 08/12/22
Wales	27/07/20 - 13/03/23	Reproduction number	28/05/20 - 09/12/22	27/07/20 - 09/12/22
		Growth rate	12/06/20 - 27/12/21	27/07/20 - 27/12/21
Northern Ireland	11/09/20 - 07/03/23	Reproduction number	26/05/20 - 31/05/22	11/09/20 - 31/05/22
		Growth rate	09/06/20 - 17/11/20	11/09/20 - 17/11/20

in the reporting windows across nations, but was "paused" on 13 March 2023 in England, Scotland and Wales and on 7 March 2023 in Northern Ireland (with the last publication on 24 March 2023 [520]). In all settings, the midpoint of the reporting window is taken as the "date" of the observation, for the purposes of model fitting.

In October 2023, the UKHSA and the ONS announced a "Winter COVID-19 study" (WCIS) running from November 2023 – March 2024, with similar aims to the ONS CIS of ascertaining prevalence of SARS-CoV-2 infection in communities across the UK [137]. However, as these data are not available at the time of writing (October 2023) and their exact format is currently unknown, our study only covers the first, continuous period of the ONS CIS from 2020 to the end of March 2023. Nonetheless, the implications of our methodology for the newly announced WCIS are outlined in the Discussion.

8.2.1.2 Modelling $R(t)$ and $r(t)$

A two-step approach is used to estimate $R(t)$ and $r(t)$ from the ONS CIS, with these estimates denoted by $\hat{R}(t)$ and $\hat{r}(t)$, respectively, and referred to as "ONS-based" estimates. The methods are primarily adapted from those presented in Eales et al. [510] and akin to those in [276, 469, 521]. Specifically, a penalised spline model is fit to the publicly-available ONS CIS test positivity estimates to produce a smoothed estimate of the prevalence of SARS-CoV-2 positivity over time before the epidemiological parameters of interest are estimated directly from the resulting curve. Each step is described in turn below.

8.2.1.3 Step 1: fitting a spline to the ONS CIS test positivity data

The model used by the REACT study [510] was taken as the basis for spline-fitting in this study and is explained here briefly. The REACT-1 data consist of the number of positive (Y_t) and total (N_t) tests on each day t of the study period, allowing test positivity of SARS-CoV-2 to be naively estimated as $p_t = \frac{Y_t}{N_t}$. Smoothed test positivity, defined as $\hat{p}(t)$, is estimated via a linear combination of K B-splines up to the 2nd degree, defined by a sequence of equidistant knots distributed throughout the period under consideration. Specifically:

$$\hat{p}(t) = \sum_{k=1}^{K} a_k B_{k,3}(t)$$

where $B_{k,3}(t)$ represents the kth 3rd order B-spline (see Appendix F) and a_k are the model coefficients, defined by a second-order random-walk prior distribution. The binomial likelihood is then parameterised by the observed number of total tests, N_t, and the estimated test positivity, $\hat{p}(t)$, at each time point.

Although the principal idea is the same, several adjustments were required to adapt the REACT spline model to ONS test positivity estimates. First, an additional sampling step was required to capture the uncertainty in the ONS test positivity estimates, which are provided in a different format to the REACT-1 data. Let μ_t denote the central ONS estimate of SARS-CoV-2 test positivity at time t, with l_t and u_t denoting the corresponding lower and upper limits of the credible interval, respectively. The variance of the estimate can be estimated as $\sigma_t^2 \approx \left(\frac{u_t - l_t}{3.92}\right)^2$. For each time t, π_t, the test positivity, is considered a random variable following a Beta distribution parameterised using μ_t and σ_t^2 as shown:

$$\begin{aligned}
\pi_t &\sim Beta(\alpha_t, \beta_t) \\
\alpha_t &= \mu_t\Big(\frac{\mu_t(1-\mu_t)}{\sigma_t^2} - 1\Big) > 0 \\
\beta_t &= \sigma_t\Big(\frac{1-\mu_t}{\mu_t}\Big) > 0.
\end{aligned}$$

The derivation of the expressions for α_t and β_t is shown in Appendix F. For every iteration, i, our algorithm begins by sampling $\pi_t^{(i)}$ for every time point t ahead of the construction of the B-splines. Essentially, $\pi_t^{(i)}$ is then taken as the "data point" at time t being fitted to in iteration i; without this additional parametric bootstrapping sampling we would fit exclusively to the point estimate and discard the information about uncertainty contained within the 95% credible intervals. The sampled value $\pi_t^{(i)}$ is then transformed onto the unconstrained *logit* scale:

$$logit(\pi_t) = log\Big(\frac{\pi_t}{1-\pi_t}\Big)$$

Second, the weekly timescale of the ONS CIS data, compared to the daily timescale of the REACT-1 study data, also necessitated adjustments to the model. In this instance, a first-order random walk was deemed a more appropriate prior distribution for the model coefficients due to the rapidly changing epidemic dynamics in these weekly data. Furthermore, knots could not be placed at the REACT-1-chosen value of every five days. Rather, a sensitivity analysis of the placing of knots was undertaken by considering scenarios in which the number of equidistant knots was equal to a percentage of the total data points, specifically 20%, 30% and 40%, balancing capturing sufficient information but without overfitting.

The final adaptation was to use a Normal likelihood function, again due to the different format of the data, based on the results of a simulation study, with details set out in Appendix F. Bringing everything together:

$$\begin{aligned}
\pi_t &\sim Beta(\alpha_t, \beta_t) \\
a_1 &\sim Normal(0, 1) \\
a_k &\sim Normal(a_{k-1}, \tau^2) \\
\tau^2 &\sim InverseGamma(0.0001, 0.0001) \\
logit(\hat{p}(t)) &= \sum_{k=1}^{K} a_k B_{k,3}(t) \\
logit(\pi_t) &\sim Normal(logit(\hat{p}(t)), \gamma^2) \\
\gamma^2 &\sim InverseGamma(0.0001, 0.0001)
\end{aligned}$$

This model was fitted using a No-U-Turn sampler [522], implemented in STAN, with 4 chains, each with $20,000$ iterations and a burn-in of $2,000$ iterations. To ensure approximately-continuous estimates of $R(t)$ and $r(t)$, the coefficients from the model are used to estimate $\hat{p}(t)$ at a granular time step (hundredths of one day).

8.2.1.4 Step 2a: Estimating $R(t)$

The effective reproduction number, $R(t)$, is estimated using the renewal equation [229, 271]:

$$\hat{R}(t) = \frac{\hat{p}(t)}{\int_{\tau=0}^{\infty} \hat{p}(t-\tau) g(\tau) d\tau}$$

where $g(\cdot)$ denotes the distribution of the generation time, defined as the time between the infections of infector-infectee pairs. The integral is approximated by summation.

The generation time distribution has been shown to change over time with the emergence of new variants

Table 8.2: Parameterisations of the generation time distribution used in the analysis. Variant time periods are defined by the earliest date at which 50% of the daily tests were attributable to the emerging variant according to Our World in Data [523]. Uncertainty was not characterised for the estimates obtained for the Omicron wave.

Variant	**Time period**	**Parametric distribution**	**Mean**	**Standard deviation**	**Source**
Wildtype	01/04/2020 - 25/12/2020	Lognormal	4.2 (3.3 - 5.3)	4.9 (3.0 - 8.3)	Hart et al. [313]
Alpha	26/12/2020 - 21/05/2021	Gamma	5.5 (4.7 - 6.5)	4.0 (3.3 - 5.3)	Hart et al. [172]
Delta	22/05/2021 - 21/11/2021	Gamma	4.7 (4.1 - 5.6)	3.3 (3.0 - 4.3)	Hart et al. [172]
Omicron	22/12/2021 - 13/03/2023	Gamma	3.3	3.5	Abbott et al. [314]

[172]. As such, estimates of $g(\cdot)$ used in this study were extracted from the literature for four distinct time periods defined by epidemic waves: Wildtype, Alpha, Delta and Omicron. Each study distribution was parameterised using data from the UK and thus is assumed to be applicable to this setting. The transition dates between time periods were determined by the earliest date at which 50% of the daily tests were attributable to the emerging variant according to Our World in Data [523]. The distributions are summarised in Table 8.2.

8.2.1.5 Step 2b: Estimating $r(t)$

The instantaneous growth rate, $r(t)$, is approximated as follows:

$$\hat{r}(t) = \frac{\frac{d\hat{p}(t)}{dt}}{\hat{p}(t)}$$

The quantities of interest, $\hat{R}(t)$ and $\hat{r}(t)$, are estimated for all posterior realisations of the spline model, with results presented as the median and the 2.5% and 97.5% quantiles throughout.

8.2.2 Analysis of $\hat{R}(t)$ and $\hat{r}(t)$

8.2.2.1 Government-published estimates

Estimates of $R(t)$ and $r(t)$ for SARS-CoV-2 in each of the four nations were produced for and published by the government from early in the pandemic for either weekly or biweekly time periods [65, 112]. The estimates were derived from an ensemble of (up to 14) independently run models as part of a cross-government and academic modelling hub that comprised the UK Health Security Agency (UKHSA) Epidemiological Ensemble team and Scientific Pandemic Influenza Group on Modelling, Operational sub-group (SPI-M-O) [65, 524]. These models encompassed a range of different assumptions and data streams as set out

in [65]. Among the data streams used, officially reported cases and hospital admissions were the most common, but there were also two models which fit to the ONS CIS data [275, 283, 284]. Consequently, there is a degree of inbuilt dependency between the government-published estimates and the estimates obtained in this study, as demonstrated below.

Ensemble model outputs were combined into a single estimate with associated uncertainty, using an established method derived by the Defence Science and Technology Laboratory (DSTL) [525]. These estimates are publicly available for download and are presented as 90% confidence intervals for the period, without any central estimate [112]. We apply the weekly or biweekly estimates to each day in the corresponding time period, and approximate a central estimate by taking the mid-point of the upper and lower bounds, herein denoted by $\tilde{R}(t)$ and $\tilde{r}(t)$ for the effective reproduction number and instantaneous growth rate, respectively. These estimates are herein referred to as "government-published" estimates.

There is heterogeneity in the availability of the government-published estimates, depending on the parameter and nation under consideration (Table 8.1). For example, official estimates of the growth rate in Northern Ireland are only available from June to November 2020. All estimates for each nation ceased to be published by 23 December 2022 [64].

8.2.2.2 Comparison of ONS-based and government-published estimates

The ONS-based estimates $\hat{R}(t)$ and $\hat{r}(t)$ are initially presented from the beginning of the survey period in each nation until the end of 2022, as this is when the government-published estimates ceased to be publicly available for both parameters and for all four nations (Table 8.1). The periods for which the ONS-based estimates can be compared to government-published estimates for each parameter and country combination are set out in Table 8.1. Due to the short period of government published estimates for the growth rate in Northern Ireland, ONS-based estimates of this parameter cannot be compared in any meaningful way to the corresponding official estimates and are thus omitted from our analysis.

We use three metrics to consider the relationship between the (median) ONS-based [$\hat{R}(t)$ and $\hat{r}(t)$] and government-published [$\tilde{R}(t)$ and $\tilde{r}(t)$] estimates of the parameters of interest, each of which are now presented. These are herein referred to as the "comparison metrics".

First, we use linear regression models of the form:

$$\tilde{R}(t) = \beta_0 + \beta_1 \hat{R}(t) + \epsilon$$

and

$$\tilde{r}(t) = \gamma_0 + \gamma_1 \hat{r}(t) + \epsilon$$

to assess the level of variation within the government-published estimates explained by the ONS-based estimates. This is captured by:

$$R^2 = 1 - \frac{\text{Residual sum of squares}}{\text{Total sum of squares}}$$

Second, we consider the Spearman rank correlation, selected due to the lack of assumptions regarding the distributions of the data.

Finally, the parameters of interest in this paper are sometimes interpreted dichotomously by policymakers, for example, to assess whether the epidemic is either growing or shrinking. As such, the proportion of point estimates over the study period for which the modelled and official estimates are on the same side of the "growth thresholds", 0 for the growth rate and 1 for the effective reproduction number, over time is also presented as the third and final metric for comparison. This quantity is herein referred to as the "agreement proportion".

A sensitivity analysis was used to examine the potential of a time-lag between estimates, assessed via the three comparison metrics. Specifically, lags between plus and minus 20 days from the default values were considered. A positive lag of L days indicates that the ONS-based estimates of day D, $\hat{R}(D)$ and $\hat{r}(D)$, are compared with the government-published estimates at day $D+L$, $\tilde{R}(D+L)$ and $\tilde{r}(D+L)$, and vice versa. In the main plots, we present our ONS-based estimates alongside the government-published estimates for the time lag and knot value which under which R^2 is maximised. A comparison of estimates without any time lags are presented in Appendix F.

8.2.3 $\hat{R}(t)$ and $\hat{r}(t)$ in the first quarter of 2023

Finally, $\hat{R}(t)$ and $\hat{r}(t)$ are produced for the first quarter of 2023 by fitting the spline model to ONS CIS estimates from 1 November 2022 until the "pause" of the survey on 13 March 2023 in England, Scotland and Wales and on 7 March 2023 in Northern Ireland. These are presented for information but without comparison, due to the lack of availability of the government-published estimates in this period.

All data and code used are available from: `https://github.com/ruthmccabe/ons-test-positivity-model`.

8.3 Results

Figure 8.1 to Figure 8.4 present the results for England, Scotland, Wales and Northern Ireland, respectively. Across the board, we can demonstrate strong agreement between our ONS-based estimates and the government-published estimates for both parameters. Depending on nation and parameter, up to 77% of the variance in the government-published estimates can be explained by the ONS-based estimates, providing evidence of the suitability of this singular data stream to track the epidemic in each of the four nations (Table 8.3). Similarly, we observed high maximum values of the Spearman rank correlation, ranging between 0.72 and 0.87 (Table 8.3).

We found that the model in England produced similar values of the three comparison metrics for all numbers of knots considered (Figure 8.1D-F). However, this was not the case in Scotland and Wales, where the number of knots played a more important role (Figure 8.2D-F; Figure 8.3D-F): for example, the maximum observed R^2 for $R(t)$ in Scotland rose by more than 40% from 0.53 to 0.74 when doubling the number of knots from 20% to 40% of total data points. In addition to requiring a greater number of knots, the spline fits in Scotland and Wales, and additionally in Northern Ireland, all have substantially greater uncertainty which is propagated through to $\hat{R}(t)$ and $\hat{r}(t)$. This is likely driven by the increased uncertainty arising from smaller sample sizes in the ONS CIS, which are proportional to the smaller population sizes in these countries compared to England.

Our sensitivity analysis has highlighted a time delay between the ONS-based estimates compared to the government-based estimates. In England, Scotland and Wales, the ONS-based estimates are correlated with later government-published estimates, suggesting that the ONS-based estimates can capture epidemic trends more quickly. (This effect can be seen clearly in Figure F.11 to Figure F.13). The models in England and Scotland indicated a similar time delay of around 8 days for $R(t)$ and 10 days for $r(t)$. In Scotland, R^2 rose by almost 50% from 0.49 with no time delay to its maximum value (0.74) under a delay of 8 days for the $R(t)$ under the model with 40% knots, emphasising the importance of considering such delays. The time delay which maximised the metrics of interest was much greater in Wales, sitting at around 2 weeks for both parameters.

For $R(t)$ and $r(t)$ in England and Scotland, and $R(t)$ in Wales, the agreement proportion sits close to 0.80 depending on the time delay and number of knots used (range 0.79–0.82). The majority of instances in which there is not agreement on whether the epidemic is growing or shrinking occurs for values close to the threshold (e.g. one estimate being slightly over the threshold while the other is slightly under and vice versa), rather than there being large differences between the two estimates. The is what drives

Table 8.3: Maximum values of the three comparison metrics used to assess the relationship between the ONS-based and government-published estimates of $R(t)$ and $r(t)$ for each nation. The value of the metric is provided alongside the lag* (in days) and the percentage of knots** out of the total data points for which this maximum value is produced. For any instance in which there were multiple combinations producing the same maximum value, the lag closest to 0 and lowest percentage of knots was selected for presentation here.

		Effective reproduction number			**Instantaneous growth rate**		
		R^2	**Spearman correlation**	**Agreement proportion**	R^2	**Spearman correlation**	**Agreement proportion**
England	Value	0.71	0.85	0.81	0.77	0.87	0.82
	Lag*	-8	-6	-10	-9	-7	-7
	Knots**	30%	30%	40%	30%	30%	40%
Scotland	Value	0.74	0.86	0.79	0.69	0.83	0.80
	Lag*	-8	-8	-4	-11	-10	-7
	Knots**	40%	40%	40%	40%	40%	40%
Wales	Value	0.66	0.81	0.79	0.67	0.82	0.42
	Lag*	-12	-10	-10	-15	-14	-13
	Knots**	40%	40%	40%	40%	40%	40%
Northern Ireland	Value	0.52	0.72	0.52			
	Lag*	9	9	10			
	Knots**	40%	40%	40%			

the low agreement proportion of $r(t)$ in Wales, despite the corresponding R^2 and Spearman correlation values being relatively high.

Our ONS-based estimates for Northern Ireland have the weakest relationship with the government-published estimates. $\tilde{R}(t)$ fluctuates substantially more so than $\hat{R}(t)$ throughout the period, in particular from November 2021 (around the time of the emergence of the Omicron variant). This relationship is reflected in all three comparison metrics, with a maximum of only 50% of the variance in the government-published estimates being explained by the ONS-based estimates. Furthermore, the ONS-based estimates are (weakly) correlated with an earlier government-based, in contrast to England, Scotland and Wales, meaning that in Northern Ireland the ONS-based estimates are slower to track the epidemic trends in this setting.

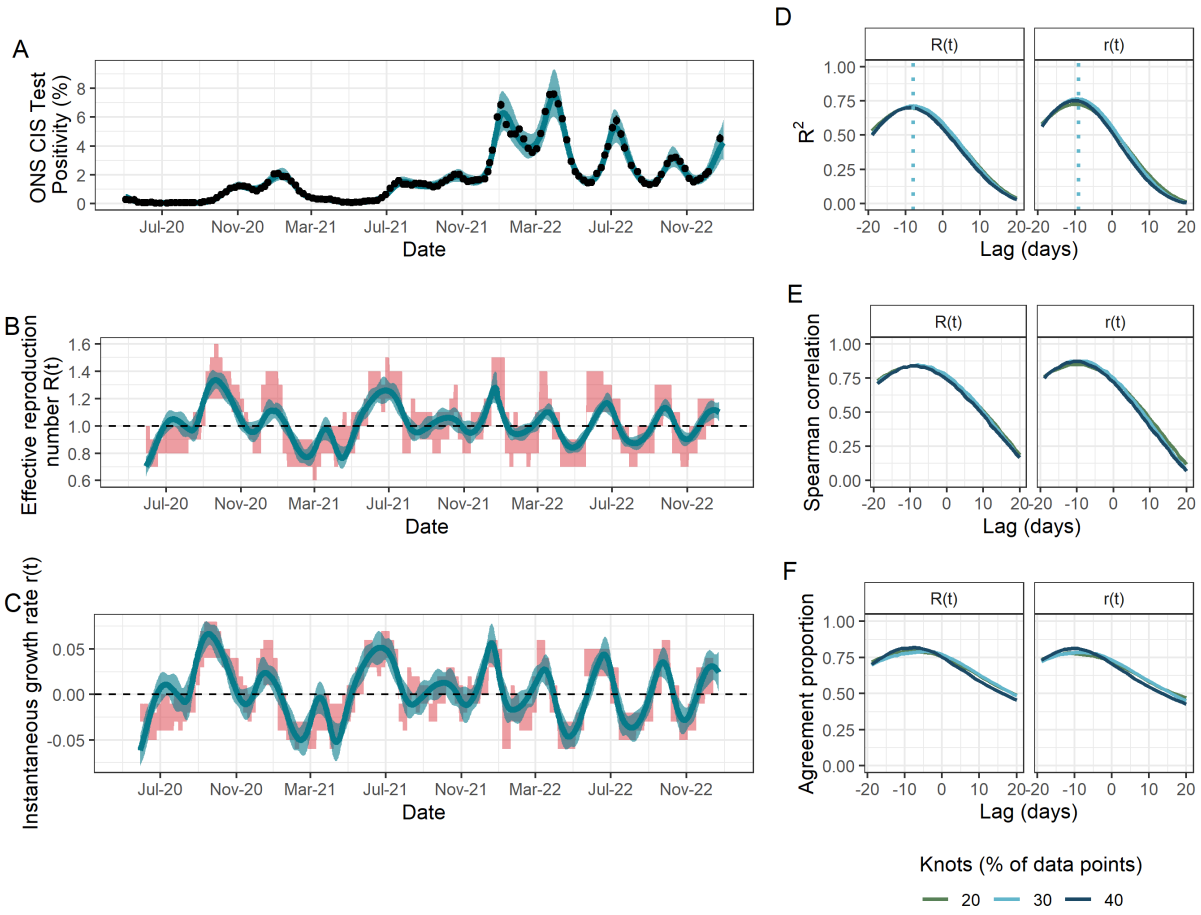

Figure 8.1: Fit to ONS CIS data, resulting estimates of $R(t)$ and $r(t)$, and comparison metrics for England. (A) Spline model fit ($\hat{p}(t)$) (blue; median line with 95% credible intervals) to data (black points; ONS point estimate with 95% credible intervals) with the number of knots totalling 30% of the total data points. (B) The ONS-based estimated effective reproduction number ($\hat{R}(t)$) (blue; median line with 95% credible intervals) alongside the government-published estimates ($\tilde{R}(t)$) (red; 90% confidence intervals), lagged by -8 days. The black dashed line highlights the epidemic growth threshold of 1. (C) The ONS-based estimated instantaneous growth rate ($\hat{r}(t)$) (blue; median line with 95% credible intervals) alongside the government-published estimates ($\tilde{r}(t)$) (red; 90% confidence intervals), lagged by -9 days. The dashed line highlights the epidemic growth threshold of 0. (D) – (F) Results of the three metrics used to assess the agreement between ONS-based and government-published estimates under different lags and for models fitted with a different number of knots, taken as a percentage of the total data points. (D) R^2. Dotted lines indicate the lag for which this metric is maximised for each parameter and is thus what is applied to the government-based estimates in panels (B) and (C). (E) Spearman rank correlation. (F) Agreement proportion.

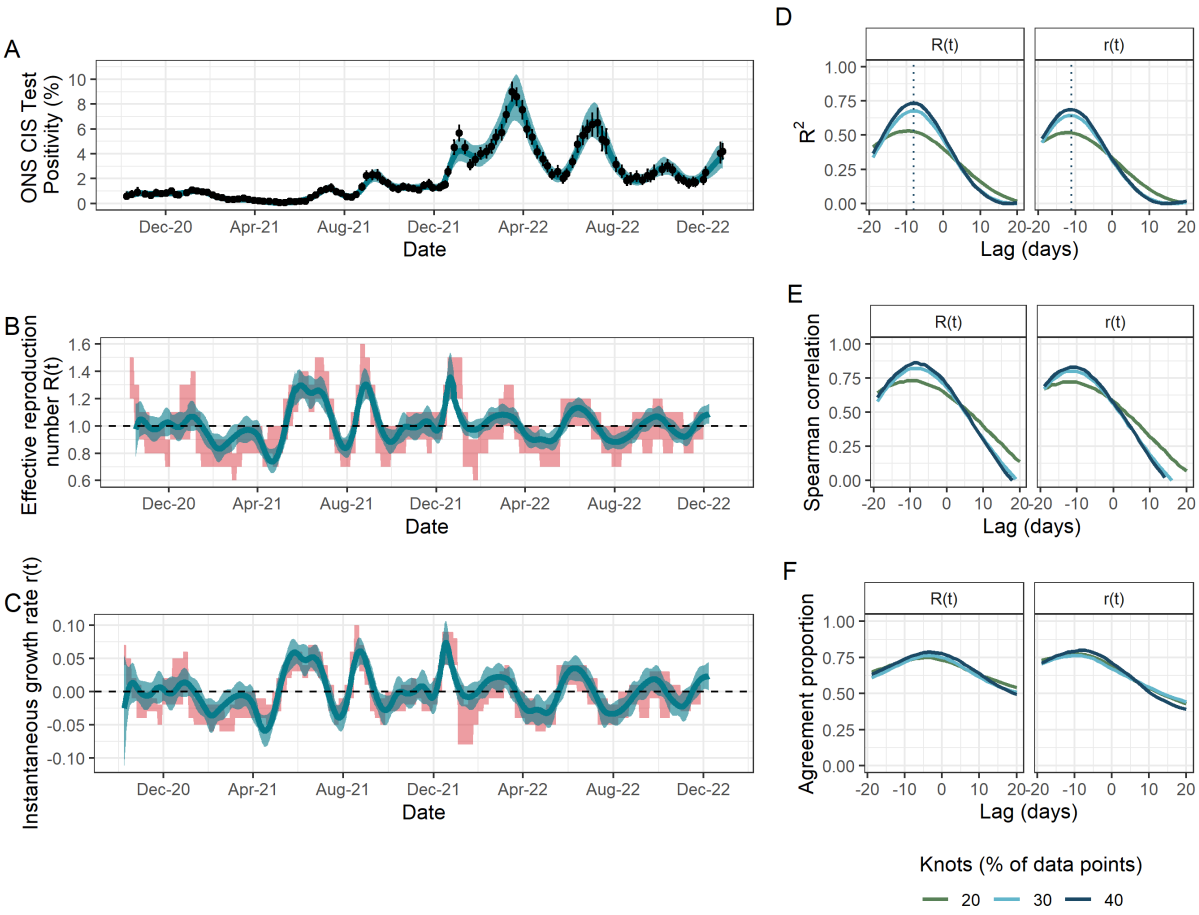

Figure 8.2: Fit to ONS CIS data, resulting estimates of $R(t)$ and $r(t)$, and comparison metrics for Scotland. (A) Spline model fit ($\hat{p}(t)$) (blue; median line with 95% credible intervals) to data (black points; ONS point estimate with 95% credible intervals) with the number of knots totalling 40% of the total data points. (B) The ONS-based estimated effective reproduction number ($\hat{R}(t)$) (blue; median line with 95% credible intervals) alongside the government-published estimates ($\tilde{R}(t)$) (red; 90% confidence intervals), lagged by -8 days. The black dashed line highlights the epidemic growth threshold of 1. (C) The ONS-based estimated instantaneous growth rate ($\hat{r}(t)$) (blue; median line with 95% credible intervals) alongside the government-published estimates ($\tilde{r}(t)$) (red; 90% confidence intervals), lagged by -11 days. The dashed line highlights the epidemic growth threshold of 0. (D) – (F) Results of the three metrics used to assess the agreement between ONS-based and government-published estimates under different lags and for models fitted with a different number of knots, taken as a percentage of the total data points. (D) R^2. Dotted lines indicate the lag for which this metric is maximised for each parameter and is thus what is applied to the government-based estimates in panels (B) and (C). (E) Spearman rank correlation. (F) Agreement proportion.

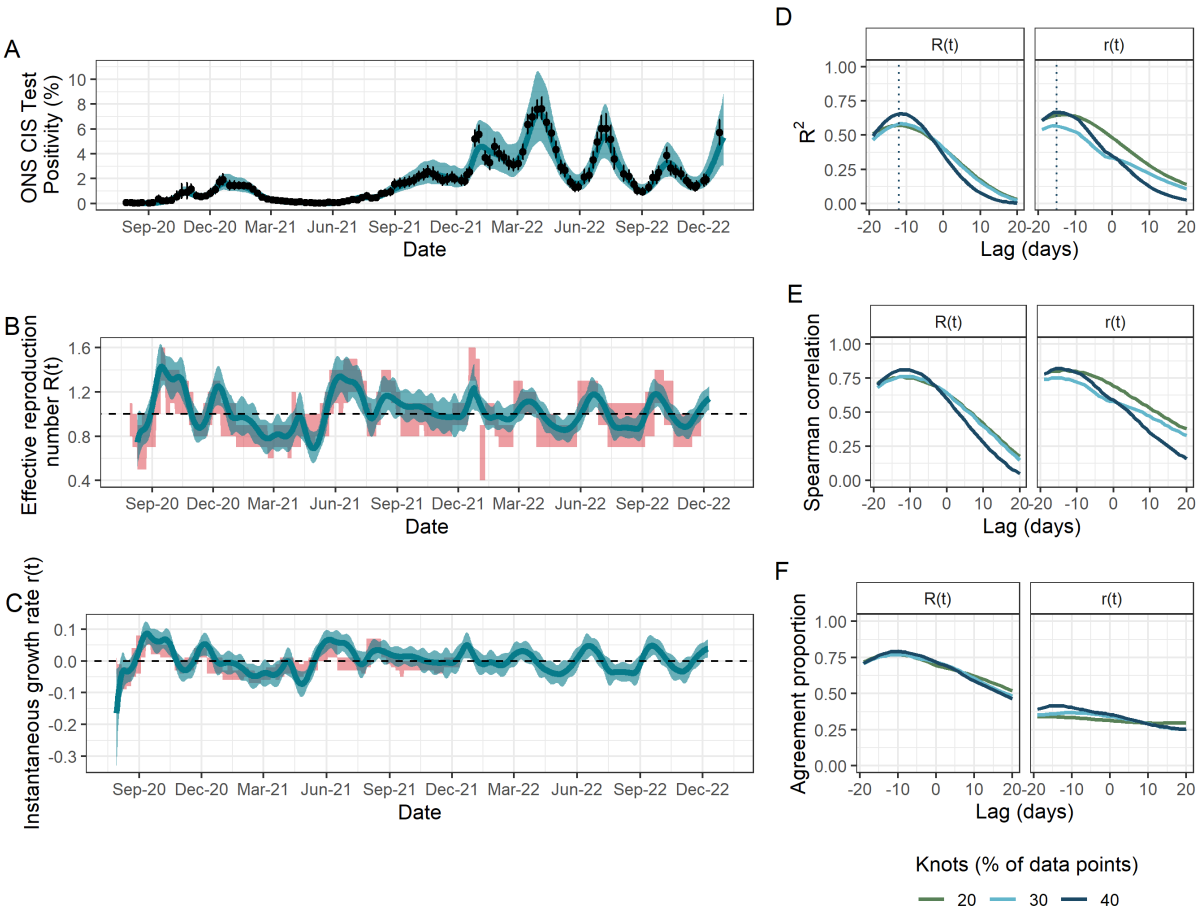

Figure 8.3: Fit to ONS CIS data, resulting estimates of $R(t)$ and $r(t)$, and comparison metrics for Wales. (A) Spline model fit ($\hat{p}(t)$) (blue; median line with 95% credible intervals) to data (black points; ONS point estimate with 95% credible intervals) with the number of knots totalling 40% of the total data points. (B) The ONS-based estimated effective reproduction number ($\hat{R}(t)$) (blue; median line with 95% credible intervals) alongside the government-published estimates ($\tilde{R}(t)$) (red; 90% confidence intervals), lagged by -12 days. The black dashed line highlights the epidemic growth threshold of 1. (C) The ONS-based estimated instantaneous growth rate ($\hat{r}(t)$) (blue; median line with 95% credible intervals) alongside the government-published estimates ($\tilde{r}(t)$) (red; 90% confidence intervals), lagged by -15 days. The dashed line highlights the epidemic growth threshold of 0. (D) – (F) Results of the three metrics used to assess the agreement between ONS-based and government-published estimates under different lags and for models fitted with a different number of knots, taken as a percentage of the total data points. (D) R^2. Dotted lines indicate the lag for which this metric is maximised for each parameter and is thus what is applied to the government-based estimates in panels (B) and (C). (E) Spearman rank correlation. (F) Agreement proportion.

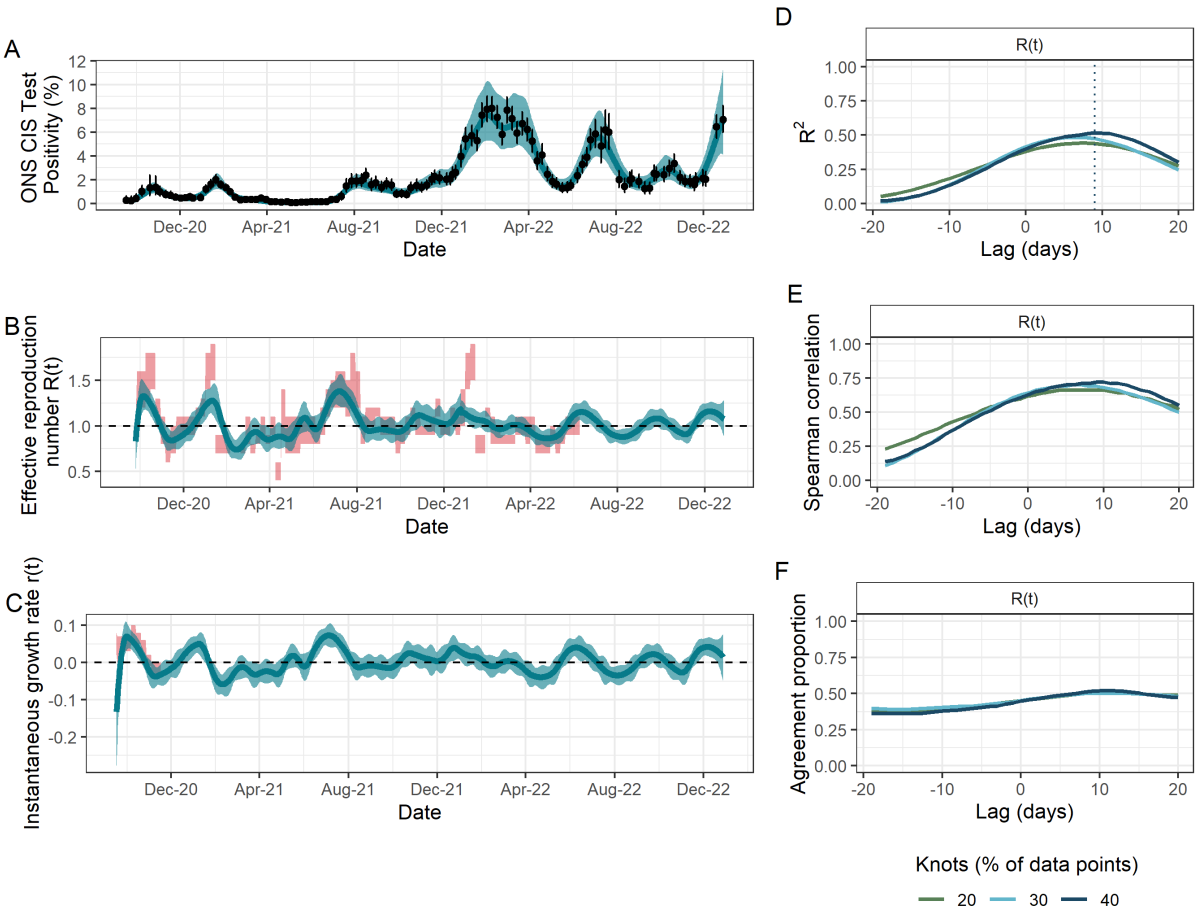

Figure 8.4: Fit to ONS CIS data, resulting estimates of $R(t)$ and $r(t)$, and comparison metrics for Northern Ireland. (A) Spline model fit ($\hat{p}(t)$) (blue; median line with 95% credible intervals) to data (black points; ONS point estimate with 95% credible intervals) with the number of knots totalling 40% of the total data points. (B) The ONS-based estimated effective reproduction number ($\hat{R}(t)$) (blue; median line with 95% credible intervals) alongside the government-published estimates ($\tilde{R}(t)$) (red; 90% confidence intervals), lagged by -8 days. The black dashed line highlights the epidemic growth threshold of 1. (C) The ONS-based estimated instantaneous growth rate ($\hat{r}(t)$) (blue; median line with 95% credible intervals) alongside the government-published estimates ($\tilde{r}(t)$) (red; 90% confidence intervals), lagged by -11 days. The dashed line highlights the epidemic growth threshold of 0. (D) – (F) Results of the three metrics used to assess the agreement between ONS-based and government-published estimates of $R(t)$ under different lags and for models fitted with a different number of knots, taken as a percentage of the total data points. (D) R^2. Dotted lines indicate the lag for which this metric is maximised for each parameter and is thus what is applied to the government-based estimates in panels (B) and (C). (E) Spearman rank correlation. (F) Agreement proportion.

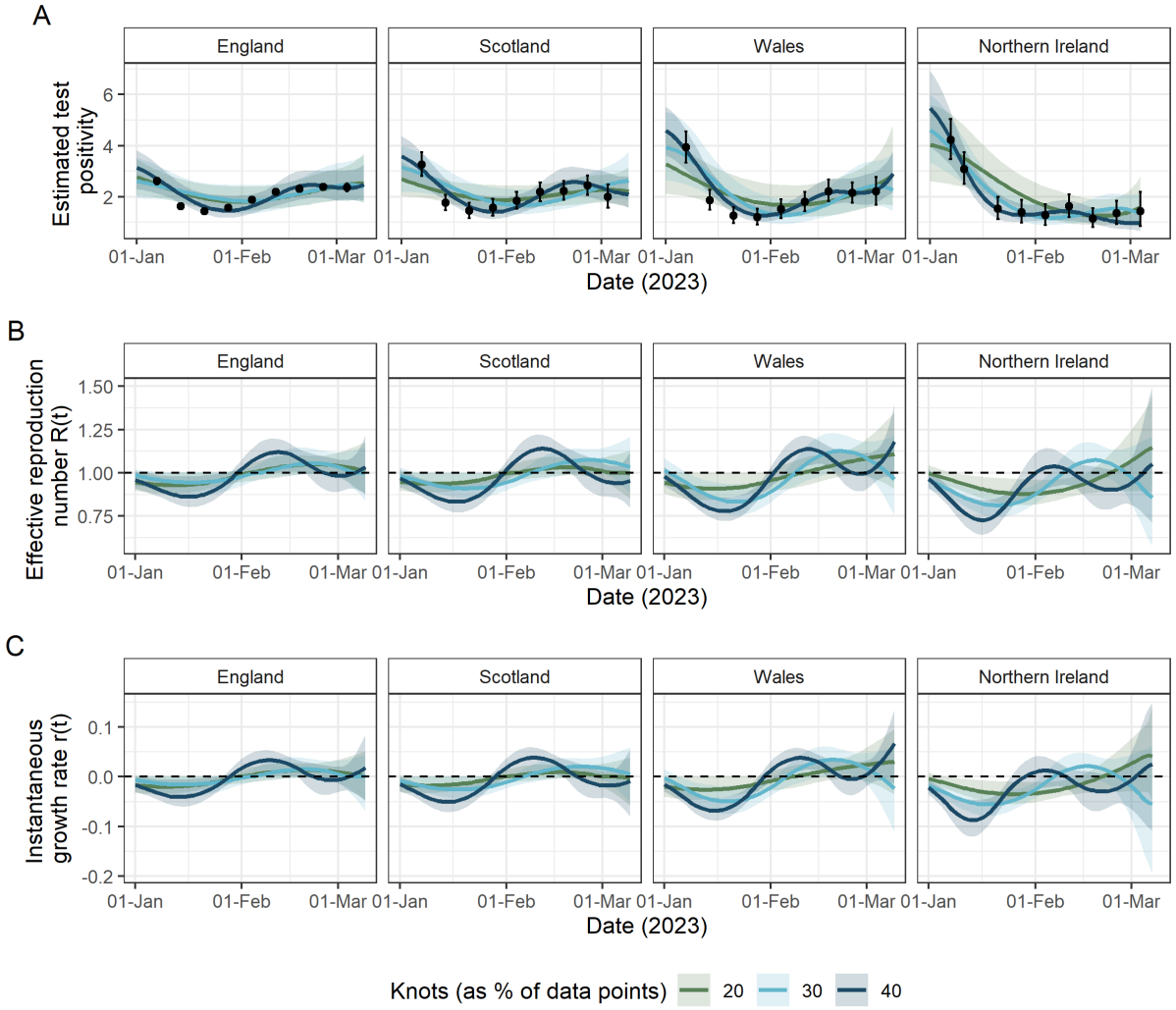

Figure 8.5: Fit to ONS CIS data and resulting estimates of $R(t)$ and $r(t)$ for England, Scotland, Wales and Northern Ireland from January 2023 - March 2023, at which point the ONS CIS was "paused". Throughout, lines represent median values and shaded areas are the corresponding 95% credible intervals (A) Spline model fits ($\hat{p}(t)$) to data (black points; ONS point estimate with 95% credible intervals) with the number of knots totalling 20% (green), 30% (light blue) and 40% (purple) of the total data points. (B) The estimated effective reproduction number ($\hat{R}(t)$) resulting from the spline fits in (A). The dashed line highlights the epidemic growth threshold of 1. (C) The estimated instantaneous growth rate ($\hat{r}(t)$) resulting from the spline fits in (A). The dashed line highlights the epidemic growth threshold of 0.

8.3.1 Estimates for 2023

Figure 8.5 presents the spline fits and estimates of $\hat{R}(t)$ and $\hat{r}(t)$ from January – March 2023 in each nation, under varying numbers of knots in the spline model.

In each nation, ONS test positivity decreases to mid-January, which results in the corresponding estimates falling below the epidemic growth thresholds: $\hat{R}(t) < 1$ and $\hat{r}(t) < 0$. However, the degree to which the estimates indicate a shrinking epidemic is dependent on the number of knots used, with a lower number of knots resulting in the estimates being closer to the threshold and vice versa.

From mid-January, test positivity increases in all nations, peaking in mid-February in England, Scotland and Wales, and in early February in Northern Ireland. This results in epidemic growth, $\hat{R}(t) > 1$ and $\hat{r}(t) > 0$, for most of the month of February. In England and Scotland, by the end of the period, both $\hat{R}(t)$ and $\hat{r}(t)$ approach the epidemic threshold showing a stable epidemic, but in Wales and Northern Ireland the estimates imply that the epidemic was continuing to grow.

8.4 Discussion

We have demonstrated and validated a method to estimate the $R(t)$ and $r(t)$ of SARS-CoV-2 using the publicly available ONS CIS, which became a primary source characterising the ongoing epidemics in the four nations of the UK after the scaling down of community testing in Spring 2022, until the survey's "pause" in March 2023. We have shown strong agreement between our ONS-based and government-published estimates across mid-2020 until the end of 2022, showing the suitability of this model applied to these data to track the trends in these key epidemic parameters. Specifically, we demonstrated that up to 77% of the variation in the government-published estimates could be explained by our ONS-based estimates, depending on the nation and parameter under consideration, which was complemented by high values of the Spearman rank correlation and agreement proportion. We have also found that most estimates under this model, except for Northern Ireland, lead government-reported estimates by up to 2 weeks, a potentially advantageous gap in terms of producing timely real-time modelling estimates. These results are important for the WCIS announced in October 2023 [137], as the methodology deployed here could potentially be used on the data generated by this study to track community transmission in the UK over winter 2023/24. Furthermore, our study demonstrates the epidemiological value of population-level studies using random sampling, such as the ONS CIS and REACT, in terms of estimating key epidemiological parameters in addition to, for example, examining risk factors for severe infection [241, 242], the burden of long COVID-19 [244] and the evolutionary dynamics of SARS-CoV-2 variants over time [246].

Our work is not intended as an alternative to ensemble modelling, which is advantageous in its ability to synthesise information from multiple sources [526]. Although deployed in multiple different fields [527–529], this is particularly important in the context of modelling the SARS-CoV-2 epidemic in the UK, due to the diversity of surveillance data streams available. As discussed by Park et al. [65], the ensemble modelling approach has many strengths including increased prediction ability and greater robustness. We see our work completing this approach, by presenting a potential solution to the challenge faced by public health officials in the UK in early 2022 given the large scaling-down of the surveillance systems at the time. With less data available, some ensemble models could become less reliable, while the approach presented here would not be affected given the continuation of the ONS CIS beyond this period.

One of the strengths of this study is its application to all nations of the UK separately, which is uncommon in the literature [46, 113, 282, 283, 530]. This work underlines the importance of setting-specific analysis of an epidemic and the differences that can arise even when these settings are geographically close and socioeconomically similar. While we were able to produce estimates which matched government-reported estimates closely, we found that both the smoothness of the spline and the time delay between the ONS-based and government-published estimates were dependent on the nation of the UK, with smaller ONS CIS sample sizes (proportional to population size) resulting in less smooth fits. As previously mentioned, this is likely attributable to the increased variability observed in the ONS CIS data among countries with smaller sample sizes. Moreover, each of the four nations have devolved governments which implemented slightly different public health and social measures at varying time points. This nuance would be difficult to capture accurately in a UK-level model.

The importance of setting-specific modelling was underlined in Northern Ireland. We found that, here, $\hat{R}(t)$ has the weakest relationship with $\tilde{R}(t)$ of all of the nations, and this relationship could not even be considered for $r(t)$ due to the lack of government-published estimates in this country. In addition to the weaker relationship, the temporal relationship between the two sets of estimates were the opposite of that observed in the other three nations, with $\hat{R}(t)$ trailing $\tilde{R}(t)$. Northern Ireland has the smallest population size of the four nations of the UK (approximately 1.9 million [531]), which inevitably increases the volatility of all surveillance data observed in this setting, and indeed this was demonstrated through the volatility of $\tilde{R}(t)$. In particular, this volatility may have contributed to the lack of published estimates $\tilde{r}(t)$, occurring when a consensus estimate from the ensemble models could not be generated due to the small number of reliable model outputs available.

Our model builds upon the methodology set out in [510] by adapting the model to fit to the publicly-available ONS CIS estimates of test positivity. Of course, other methods also exist to estimate $R(t)$ and $r(t)$ using the ONS CIS outputs, for example by fitting the model of Eales et al. [510] to the (retrospectively-reported) numbers of positive and total tests. While theoretically it would be possible to use ONS-estimated incidence of infection within previously published frameworks such as [229–231], these figures were published up to three weeks later than the corresponding test positivity estimates and ceased to be published in June 2022, thus making these data unsuitable for ongoing real-time modelling of the epidemic. As an alternative, Abbott and Funk [275] deconvolve ONS test positivity estimates into incidence, which they then model using a Gaussian process. However, in addition to the assumption of the generation time distribution, which is a necessity of the renewal model, the probability density function of the time from infection until PCR positivity is also required. Similarly, the ONS CIS is among the data streams that Birrell et al. [283, 284] fit to as part of their age- and NHS-England-region-stratified Susceptible-Exposed-Infected-Recovered (SEIR) model. However, this model is specific to England, and also requires additional parameters to calibrate the complex mechanistic model to the observed data, as is the case with mathematical models of this structure. By contrast, our method only has two key parameters which require tuning to the setting of interest: firstly, the generation time distribution, which is a common assumption in models with an element of mechanistic transmission [229–231] and secondly, the smoothness of the spline, controlled by the number of knots. For the latter, we conducted a sensitivity analysis to assess the impact of the number of knots and show that there are often viable multiple options, although we found that data sets with smaller sample sizes were better fitted by models with more knots. Another important strength of our model is that we have presented a method to fit directly to the publicly-available estimates of test positivity, which was the format in which the survey results were presented for the vast majority of the study period (until 29 July 2022). Furthermore, by fitting exclusively to these publicly-available estimates, anyone may reproduce our results and/or adapt the model to similar data in other contexts, for example in different countries or for other pathogens, or even to other data streams presented in this format if appropriate. Moreover, the generation time distribution can be updated easily over time, as was required due to the emergence of new variants [172].

It is important to be aware of the limitations of this work. First, our estimates suffer from boundary effects attributable to the ONS CIS survey starting mid-way through a wave of infection. This is particularly noticeable in Wales and Northern Ireland (also with smaller sample sizes) and may contribute to the lower levels of agreement observed between the ONS-based and government-published estimates for each comparison metric. Second, as discussed, there is not a standard way to select the optimal level of smoothing and thus was determined in this instance via sensitivity analysis. Eales et al. [510] have

taken a similar approach in their selection the number of knots. Third, as $R(t)$ and $r(t)$ are not directly observable quantities and thus there are no "true" values by which to validate our model. Instead, we have focussed on comparison with the government-published estimates and considered the amount of information that the ONS-based estimates alone could yield. This implicitly treats the government-published estimates as reliable and accurate estimates of the true values of $R(t)$ and $r(t)$. Fourth, although our results suggest that our ONS-based estimates provide more timely insights into epidemic trends than the corresponding government-based estimates, this does not consider the time delays between the collection and publication of the ONS CIS data. Although this delay would not change the date on which trends occur, it would change the date when the trends would be able to be estimated and so real-time modelling gains may not be fully realised. This could be overcome by more timely access to the ONS CIS data, but this may compromise the data being publicly available and could also make estimation more logistically challenging. Although not a limitation of our work per se, it is also important to highlight that the ONS CIS was expensive to run (approximately £945m to the end of December 2022 [532]), which undoubtedly raises questions regarding the long-term implementation of similar schemes in the future.

8.5 Conclusions

In this study, we have demonstrated the utility and validity of this model for estimating $R(t)$ and $r(t)$ using the ONS CIS test positivity data, which was a primary source measuring the ongoing epidemics in the four nations of the UK. Our model provides a reliable means by which to track the ongoing epidemics in each of the four nations of the UK after the scaling down of SARS-CoV-2 surveillance, which was largely reduced due to the transition from "emergency" to "endemic" state in Spring 2022. Our study highlights the critical role that studies such as the ONS CIS play as part of an effective and data-driven epidemic response. Although the ONS CIS was "paused" in March 2023, the WCIS (announced in October 2023) should provide a new surveillance data stream, similar to the ONS CIS data, for the winter period 2023/24. Consequently, this model is uniquely placed to provide a reliable method for tracking the spread of the UK epidemic throughout this potentially challenging winter period.

Chapter 9

Discussion and conclusions

This final chapter brings together the main findings of the work in this thesis to discuss the strengths, limitations and conclusions of the research. It should be noted that each paper (presented in Chapters 3 - 8) has a specific discussion, and so this chapter provides a high-level synthesis of this information.

9.1 Summary of main findings

This thesis presents six papers, comprising of three papers to address each of the two main objectives (Table 1.1):

Objective 1: To investigate the role of modelling in informing public health policy (Papers I - III);
Objective 2: To investigate the suitability of alternative data sources to track transmission dynamics (Papers IV - VI).

The main findings and limitations of the work relating to each objective will be discussed in turn in the following two subsections.

9.1.1 Objective 1: To investigate the role of modelling in informing public health policy

Papers I - III explored the use of and communication between key stakeholders regarding mathematical modelling of epidemics in informing public health policy, particularly during the coronavirus disease 2019 (COVID-19) pandemic. The research conducted in these papers has encompassed multiple components, including the design and implementation of surveys and interviews to collect novel data (Papers I and III); synthesis of both quantitative and qualitative data (Papers I and III); synthesis of relevant litera-

ture (Paper I); analysis and presentation of mathematical modelling of an epidemic under different public health and social measure (PHSM) scenarios (Paper II); statistical analysis of quantitative survey results drawn via two sampling methods (Paper III).

In Paper I [50], we confirmed that mathematical modelling indeed was among the key evidence contributing to pandemic decision-making in the COVID-19 pandemic in the United Kingdom of Great Britain and Northern Ireland (UK) over the course of 2020 (the time period of our research) and demonstrated the diversity in the applications of mathematical modelling to address policy-related issues, with this work often completed under intense time pressure. We then focused attention on the perspective of (current, at the time of undertaking the research, and former) scientific advisors and science communicators, posing questions regarding how the process of providing scientific advice with regard to mathematical modelling worked in the first year of the pandemic in the UK and how it could have been improved. Our primary data strongly highlighted the need to convey the uncertainty inherent in transmission models in a clear manner, acknowledging that many key stakeholders (for example, politicians and the public) are not always familiar with this difficult, technical concept. We also found most participants in our research indicated that communication between all key stakeholders could be improved, again acknowledging that the intense time pressures faced by all which made this more challenging. However, there was a divergence of opinions regarding the form that this should take: while some respondents wanted a two-way dialogue between advisors and decision-makers, others cautioned against this as it may blur the important distinction between scientists and decision-makers, which must be understood by the public. Nevertheless, the science communication experts that we interviewed noted that direct communication with the public had been challenging, for example, conveying the complexity of the model and its underpinning assumptions in a way that is understandable to a non-technical audience.

Paper II [51] directly addressed the issue of clear communication of uncertainty to non-technical audiences underlined by respondents to Paper I, drawing upon the experience of our own mathematical modelling of the impact of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) pandemic under different PHSMs on intensive care unit (ICU) capacity. We demonstrated that the use of summary statistics such as the median and 95% quantiles, a frequent means by which to present model outputs, can obscure the full range of possible outcomes indicated by a model. We illustrated this using our ICU modelling, which suggested an almost identical outcome under both PHSM scenarios when considering the associated summary statistics, but upon closer inspection of the individual outputs, important differences in terms of the volatility of different model realisations under one of the strategies emerged. We suggested that a reasonable alternative is to present all trajectories, without crude summarisation,

alongside policy-actionable metrics such as the time and magnitude of possible capacity breaches, which are not immediately obvious from looking at the trajectories alone. We highlighted the importance of such an approach particularly under more complex models, for example, those including stochasticity and/or those capturing multifaceted immunity dynamics.

The research questions for Paper III [52] were developed in direct response to the comments made by respondents to Paper I regarding the communication of mathematical modelling and decision-making to the public, whom public health policies, such as "lockdowns", affect most. We again collected novel data via responses to a survey which we distributed via two mechanisms: an online panel ("representative" sample in the context of this paper) and via social media ("non-probability" sample). We asked questions relating to the means of awareness and responsibility of the communication of modelling, feelings of reliability of modelling in informing policy, and trust in government public health policy in reference to two time points: before the COVID-19 pandemic (asked retrospectively) and during the COVID-19 pandemic. We focused on comparing differences between the two samples, which also demonstrated the importance of varied conclusions that could be drawn based on the sampling mechanism used. We found all respondents were increasingly aware of the use of modelling in informing policy during the pandemic, with significantly higher levels of awareness among social media respondents than online panel respondents. Awareness generally stemmed from the news media and social media during the pandemic. Transmission modelling informing public health policy was perceived as more reliable during the pandemic compared to the pre-pandemic period in both samples, with awareness being positively associated with reliability within both samples and time points, except for social media during the pandemic. Trust in government public health advice remained high across samples and time periods overall but was lower in the period of the pandemic compared to the pre-pandemic period. The decay in trust was notably greater among social media respondents. Many respondents from both samples explicitly made the distinction that their trust was reserved for "scientists" and not "politicians". Almost all respondents, regardless of sample, believed governments have responsibility for the communication of modelling to the public.

9.1.1.1 Strengths and limitations of work relating to Objective 1

The work relating to Objective 1 not only provides important historical context of COVID-19 modelling, how it has shaped the UK response, and the public's perception of this, but looking towards future pandemic preparedness, understanding these relationships and how they may be improved is critical. We believe that our work contributes to these important conversations not only through the collection of novel data from many of the key stakeholders such as scientific advisors, science communicators and the public (Papers I and III), but also by providing tangible suggestions to address some of the key

issues raised, for example, the clear communication of uncertainty (Paper II). While we acknowledge the limitations of primary research below, we are proud to have been able to collect and share data with the scientific community from a unique time point in our history (the COVID-19 era throughout early-2020 - mid-2021, which cannot be added to without increasing risks of recall bias etc.).

The main limitation of the work in Papers I and III is that it is predominantly based on our participants' experiences and opinions, leaving our data vulnerable to, for example, personal opinion and/or recall bias. While it would not be possible to explore the questions set out in this objective sufficiently without collecting data from key stakeholders as we have done here, we must still be cautious in the conclusions and generalisations that we can draw from these data. This issue is compounded further by sample size. In Paper I, although almost all those invited for the interview took part ($n = 11$ out of 13 invited), most of those invited to take part in the survey did not ($n = 46$ out of 174 invited) and it is unclear whether those who responded are representative of the whole population. In Paper III, both sample sizes are moderate in size ($n = 504$ for the online panel; $n = 202$ for social media), and we are unable to ascertain how representativeness of the online panel is in terms of the UK population. In this paper, we focus on comparison with levels of vaccination, but could expand this to include age and sex also. However, as mentioned previously, one of the key aims of Paper III was to the assess the differences in results obtained under the two sampling methods, with one of our main findings relating to this very topic. Throughout both papers, we have been clear that we do not consider our work to be definitive in terms of the mechanisms of the science-policy interface in the UK during the COVID-19 pandemic, rather we hope that it makes important contributions to this topic, on which there are a multitude of opinions and experiences to consider, as demonstrated particularly well through the ongoing UK COVID-19 Inquiry [139].

Another common limitation of collecting primary data thus applicable to our work is striking the difficult but critical balance between open-ended questions so as not to influence participant's responses but retaining enough clarity to ensure the results are relevant and comparable. We sought to balance these competing issues for both the surveys and interview questions in Papers I and III in the design process, but acknowledge that this is subjective and there is likely room for improvement in the wording of our questions. While we did consult colleagues with experience in this field, in future we would also seek to conduct a small pilot of a survey ahead of formal data collection. Finally, we also acknowledge that our literature review in Paper I was only able to encompass a small subset of modelling studies released over the course of 2020. However, the purpose of our review was to inform our wider discussion on modelling rather than constituting a systematic modelling review.

9.1.2 Objective 2: To investigate the suitability of alternative data sources to track transmission dynamics

Epidemiological analysis is typically underpinned by traditional data sources such as the numbers of officially-reported cases and deaths pertaining to a particular pathogen. However, in Papers IV - VI, we discuss the biases inherent in these data, and present and validate alternative sources by which to understand the transmission dynamics of SARS-CoV-2 in different settings and at different time points. As with Objective 1, the research conducted in these papers has drawn on different methodologies and skillsets: statistical analysis of a seroprevalence study (Paper IV); mathematical modelling of epidemics in low-income settings (Paper V); and statistical modelling of the effective reproduction number ($R(t)$) and instantaneous growth rate ($r(t)$) in the UK using a population-level prevalence study (Paper VI).

In Paper IV [53], we retrospectively tested 11,000 stored antenatal serum samples from two north-west London NHS trusts across late 2019 to late 2020 for antibodies to SARS-CoV-2. While the principle is identical to testing stored samples from blood donors, different ethnicities and levels of deprivation present in the UK population are not typically well-represented within this group, which is problematic as these two characteristics are associated with increased risk of infection with SARS-CoV-2 and severe illness from COVID-19 [165–167]. Pregnant women, although typically restricted to a younger subset of the population and all biologically female, do not have the same representation problems and thus these stored antenatal samples present an opportunity by which to understand further the spread and burden of SARS-CoV-2. In this paper, we estimated the prevalence and incidence (derived from prevalence) of seroreactivity over time and examined the association between age, ethnicity and deprivation on prevalence of seroreactivity. We found that prevalence of seroreactivity increased from 1% prior to mid-February 2020 to 17% in September 2020 and showed higher prevalence of seroreactivity to SARS-CoV-2 in younger, non-white ethnicity, and more deprived groups, in accordance with findings of other research groups. The temporal trends displayed in our estimated incidence of seroreactivity reflected those observed in daily hospitalisations and deaths in London that followed 10 and 13 days later, respectively. We quantified community transmission of SARS-CoV-2 in London, which peaked in late March/early April 2020 with no evidence of community transmission until after January 2020, as expected. We were unable to estimate the date of introduction of the SARS-CoV-2 virus into north-west London. However, in future epidemics, stored maternal booking serum samples could be used, once reliable assays have been developed, to give retrospective estimates of early incidence rates, to identify populations at risk during the earliest stages of transmission and hence to track the emergence of a new virus.

In Paper V [54], we simultaneously estimated levels of under-reporting of COVID-19 mortality in three

low-income settings (Addis Ababa, Ethiopia; Aden, Yemen; Khartoum, Sudan) and addressed an important limitation of the World Health Organization (WHO)-commissioned estimates of excess mortality, that estimation in settings without all-cause mortality data had relied solely on model-based inference by pooling information from countries with similar socioeconomic and demographic characteristics. In this paper, we presented a local-data-driven alternative to this approach, drawing on three case studies from settings which had both an alternative source of data, namely burial site worker reports in Addis Ababa; satellite imagery of cemeteries in Aden; and social media-conducted surveys of infection in Khartoum, and an independently-conducted, representative seroprevalence study. Specifically, we calibrated a previously-published model of SARS-CoV-2 transmission and disease progression to these alternative sources and utilised the infection fatality ratio (IFR) to estimate the number of infections implied by the alternative source. We then validated these estimates using the independently-conducted seroprevalence study, finding no statistically significant differences between our model-estimated seroprevalence based on the alternative data source and those reported in serosurvey in each setting. Consequently, we estimated that 69 to 100%, 0.8 to 8.0%, and 3.0 to 6.0% of COVID-19 deaths were reported in Addis Ababa, Aden and Khartoum, respectively. Our results challenge the perception of the global burden of the disease as reported in official figures which, as shown here in the cases of Aden and Khartoum, can suffer from under-reporting in low-income settings. However, should the trends that we have observed in Addis Ababa hold at a national-level, this would suggest that national-level excess mortality, and consequently mortality under-reporting, may have been overestimated for Ethiopia as a whole. Our work underlined the need for setting-specific epidemic data, and we concluded that these alternative data sources represent a viable, and likely cost-effective, route to estimating disease burden and impact in settings where high-quality civil registration system mortality data are not currently available.

Finally, Paper VI [55] addressed directly a problem faced by public health officials at the United Kingdom Health Security Agency (UKHSA) in the first quarter of 2022 as to how to continue estimation of $R(t)$ and $r(t)$ in each nation of the UK, as requested by the respective governments, given the scale-down of SARS-CoV-2 surveillance, particularly community testing. During the emergency phase of the pandemic, estimation of these quantities relied on an ensemble of models but this became much more difficult to implement after these substantial changes to surveillance as part of the transition to treatment of this pathogen as "endemic". We leveraged the use of one of the remaining data streams, the Office for National Statistics (ONS) COVID-19 Infection Survey (CIS), to estimate $R(t)$ and $r(t)$ in the home nations by drawing on the methods deployed by the REal-time Assessment of Community Transmission (REACT) study. We validated our results using the government-published ensemble model estimates for the periods of overlap, finding that up to 77% of the variance in the government-published estimates

can be explained by our ONS-based estimates. We additionally found that the ONS-based estimates uncover epidemic trends earlier than the corresponding government-published estimates. This work also demonstrated the importance of setting-specific epidemic modelling, even within the geographically, demographically and socioeconomically similar four nations of the UK. Our work outlined the value of this singular data stream to track the epidemic in each of the four nations, indicating a potential solution to the aforementioned challenge faced by public health officials in the UK in early 2022 rather than as an alternative to ensemble modelling.

9.1.2.1 Strengths and limitations of work relating to Objective 2

The work pertaining to Objective 2 provides practical solutions to the real, common data challenges encountered in the analysis of infectious diseases as discussed previously. Data issues early in the pandemic were frequently cited as limitations of the modelling included in the literature review and noted by scientific advisors surveyed in Paper I. In an ideal scenario, traditional epidemiological data such as officially-reported cases and deaths would be complete and not subjected to biases, or given imperfect reporting, we would always have multiple sources of data to call upon and to triangulate our findings. However, we know that both of these scenarios are unrealistic and unsustainable in the long-term, for any pathogen, as demonstrated particularly well throughout the recent pandemic and highlighted in Papers IV - VI. The papers relating to this objective creatively repurpose existing resources, not produced initially for the uses presented here, to track the transmission dynamics of the SARS-CoV-2 epidemic across different locations (low- and high-income) and time periods (early-outbreak and post-emergency phase) for which there were few alternative means to do so. Our practical and data-driven approaches provide tangible means by which to understand the spread and burden of a pathogen when official data are sparse or missing altogether. However, as highlighted throughout these papers, these resources should not be seen as replacements for high-quality, complete and timely epidemiological data, which would always be preferable. Although the methods deployed here have been exclusively applied to SARS-CoV-2 in this thesis, they are pathogen-agnostic and thus could be utilised in future outbreaks of a known or unknown pathogen to overcome some of the inevitable data challenges that the community will be faced with at that time. We also note that a considerable strength of the work in these papers is that they provide further evidence of the importance of setting-specific modelling. For example, in Paper V via the case of Addis Ababa in Ethiopia and in Paper VI with respect to the four nations of the UK, we demonstrate the potential pitfalls of assuming common trends in countries which are geographically close as well as demographically and socioeconomically similar.

While we have validated the use of these resources, a clear limitation of this work is that these data

may not always be available in every setting and at each time point. For example, while the modelling framework in Paper V could be adapted to other settings and further alternative data sources, not all data sources will be suitable in other locations. Specifically, cemetery burial data are unlikely to be representative in settings that practice cremation and satellite imagery approaches may be less reliable in settings with extensive cloud coverage. Furthermore, while the collection and storage of antenatal serum samples used in Paper IV is implemented across the UK, and would potentially be implementable in other high-income countries (HICs), it is unlikely to be a realistic option in countries in which the collection and subsequent storage of these samples is not common practice. A similar argument applies to programmes such as the ONS CIS which, while extremely valuable epidemiologically, cost close to £1bn up to the end of 2022 [532].

A further limitation of this work comes from the fact that these analyses were conducted retrospectively and so we have not been able to fully validate their suitability for real-time epidemic analysis. For example, were there to be substantial time delays between the date of observation and the date of the availability of said observation, the use of these resources in real-time may not be possible (although, of course, the date on which the trends are recorded within the data would remain unaffected for retrospective analysis). This caveat is applicable to all of the alternative resources utilised in Papers IV - VI. Additionally, some of our analyses rely on serological data, for which there are two main caveats. Firstly, in the case of Paper V, SeroTracker suggested all three serosurveys used for validation may be subject to "moderate" bias, with corresponding estimates defined as "likely correct for the target population", as opposed to "very likely correct" for studies with "low" bias. Additionally, SARS-CoV-2 assays are often developed on high-income populations and the applicability to other settings is not fully understood. Secondly, in the case of Paper IV, the susceptibility of the assay used to cross-reactivity (uncovered during rather than prior to our study) meant that we could not estimate the date of introduction of SARS-CoV-2 into our sample population, rather we focussed on the trends in incidence of seroprevalence which is unaffected by this issue given a constant level of cross-reactivity.

Finally, as with all models, we acknowledge that they have to be parameterised, which we have achieved using the most up-to-date and appropriate literature available at the time and for the setting under consideration. Nonetheless, we appreciate that our results are therefore contingent, to a degree, on our choice of parameterisation. Wherever possible, we have undertaken sensitivity analyses to test and explore the impact of these choices (for example, time to seroreversion and the IFR in Paper V; number of knots in the spline model in Paper VI), but there remains a degree of uncertainty as to the most appropriate selection.

9.2 Future research questions

While the specific work presented here in Papers I - VI has come to a practical conclusion upon both their publication and submission of this thesis, there are limitations which could be addressed and further questions that could be explored based on the findings and experiences of conducting this research.

Regarding the work in Papers I and III, given further opportunity it would be important to consider explicitly the views of decision-makers themselves. This group are omitted directly from the conversation due to the difficulty in being able to contact them, especially during the emergency phase of the pandemic. However, it would likely still be difficult to facilitate even now due to the politicisation of the pandemic, and particularly due to the ongoing COVID-19 Inquiry [139] to which many are giving evidence. As Paper III was developed in response to the findings of Paper I, we present these papers as connected but ultimately distinct pieces of work. Going forwards, it would be advantageous to have all parties involved in one coherent discussion, perhaps through the use of focus groups, rather than considering each in relative isolation.

Within Paper IV, the biggest limitation to overcome is the use of an assay to test for SARS-CoV-2 antibodies which is highly sensitive but less susceptible to cross-reactivity; doing so may have allowed for estimation of the introduction of this pathogen within our sample population. This study focussed on north-west London having been proposed initially as a pilot to demonstrate the utility of this resource in epidemic analysis, with the plan to then extend this England-wide. We hypothesised that we could have used these data to provide unique insights into the initial introduction and subsequent spatial spread of SARS-CoV-2 nationally, for which there are so few means to examine such questions. Unfortunately, we were unable to secure funding to undertake this larger-scale work, and due to the two-year retention policy of samples this would no longer be possible for SARS-CoV-2 as the samples from the period of the emergence of this virus are no longer available. Nonetheless, we have still demonstrated the utility of these samples for epidemic analysis and would seek to use them again upon the emergence of another emerging virus in the UK.

While the main goal of Paper VI was to provide a practical solution to a problem faced by UKHSA as part of the transition from treatment of SARS-CoV-2 as “endemic” rather than an “emergency”, it also sparks an important discussion about the contributions of individual streams of data to estimate important epidemiological quantities, such as, but not exclusively, $R(t)$. A natural next step of the work in this thesis would be to look at estimates of $R(t)$ obtained under the different but overlapping surveillance data streams, for example, Pillar 1, Pillar 2, polymerase chain reaction (PCR) and lateral flow

device (LFD) testing, in the UK using previously developed methods where possible, and then undertaking comparisons with the government-published estimates. Key questions of interest could include: Do these surveillance data, in isolation, suggest similar epidemiological trends to that observed by the government-published ensemble estimates? How do estimates vary, for example in terms of precision and uncertainty, under different data streams and does this change over time? How much of the variation in the government-published ensemble estimates can be explained by an individual data source? How much is gained, purely from a modelling perspective, from having full-scale community testing using LFDs compared to just PCR testing for those with a clinical need? Such questions may shed light on what resources could be prioritised during a future outbreak in the UK.

A further potential research question comes from the modelling literature review in Paper I, in which researchers frequently cited high levels of uncertainty surrounding estimates of key epidemiological parameters of SARS-CoV-2 as limitations of their work. We discussed similar issues with uncertainty in parameter values during our modelling work in Paper V, particularly regarding time to seroreversion. Therefore, it would be interesting to explore initial estimates of epidemiological parameters, how they have changed over time, and to unpick the drivers of those changes (for example, are changes in estimates an artefact of more reliable data or due to genuine epidemiological changes, perhaps due to the emergence of a new variant). Once these data were collected, different estimates of the same parameter could be input into an identical mathematical model structure and the results used to assess the influence of this parameter. Through exploration of these questions for SARS-CoV-2, it could be possible to consider the extent to which this frequently cited limitation is a cause for concern, which could guide initial modelling responses to future outbreaks.

9.3 Conclusions

As stated in Chapter 1, not only do the efforts that we put into understanding the transmission dynamics of SARS-CoV-2 help us to respond to the current threats posed by the ongoing circulation of this pathogen, but they will help inform responses to future outbreaks. With this in mind, and bringing this body of work together, the following points can be considered as the key conclusions of this thesis:

- Mathematical modelling plays a crucial role, though it is not the exclusive source of evidence guiding public health policy during an outbreak (Paper I).
- The science-policy interface is highly complex and, in times of emergency, effective communication between relevant stakeholders can be challenging. Whatever form communication takes, the public must be clearly aware of the distinction between those producing the evidence and those whose job

it is to make decisions based on all of the relevant evidence. Our findings suggest this distinction is already well understood by the public (Papers I and III).

- The communication of uncertainty inherent in epidemic models to non-technical audiences requires a careful balancing of transparency and interpretability (Papers I and II).
- The public want transparency regarding the evidence underpinning the implementation of and/or rapid change to public health policies, such as "lockdowns", and often find such information via the news media and/or social media (Paper III).
- Skewed conclusions could be drawn from non-representative samples of the target population, for example, if gathered through social media (Paper III).
- There is no evidence of community transmission of SARS-CoV-2 in London until after January 2020, with this peaking in late March/early April for the first wave of the pandemic in 2020. Prevalence of seroreactivity to SARS-CoV-2 was associated with younger, non-white ethnicity, and more deprived groups (Paper IV).
- Antenatal serum samples are a useful, routinely-collected resource in the UK which could be used to track the spatiotemporal spread of a novel pathogen while surveillance systems are being established (Paper IV).
- Alternative epidemic indicators, namely burial site worker reports, satellite imagery of cemeteries and social-media-conducted surveys of infection, provide a data-driven alternative to estimates of excess mortality in settings without robust all-cause mortality data. From these sources and with validation from independently-conducted serological studies, we estimate that $69-100\%$, $0.8-8.0\%$ and $3.0-6.0\%$ of COVID-19 deaths were reported in Addis Ababa, Aden and Khartoum, respectively, in the first-wave of the pandemic in these settings (Paper V).
- Estimates of $R(t)$ and $r(t)$ derived from the ONS CIS can explain up to 77% of the variation observed in the government-published ensemble estimates of these quantities, and can uncover these key epidemic trends earlier than these estimates (Paper VI).
- Tailoring epidemic modelling to specific settings is essential because even countries that seem similar in terms of geography, demographics, and/or socioeconomic factors can exhibit distinct trends and require different modelling approaches (Papers V and VI).
- It is always preferable to have multiple sources of data for analysis of an epidemic, ideally in the form of robust and complete epidemiological data. However, given that this is broadly unrealistic, there are validated alternative sources by which to reliably track spread and burden of existing or novel pathogens (Papers IV - VI).

Bibliography

[1] N Ferguson, D Laydon, G Nedjati Gilani, N Imai, K Ainslie, M Baguelin, et al. Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand. Imperial College London, 2020. Available from: `https://doi.org/10.25561/77482`.

[2] PGT Walker, C Whittaker, OJ Watson, M Baguelin, P Winskill, A Hamlet, et al. The impact of COVID-19 and strategies for mitigation and suppression in low- and middle-income countries. *Science*, 369(6502):413–422, 2020. doi: 10.1126/science.abc0035.

[3] P Christen, JC D'Aeth, A Løchen, R McCabe, D Rizmie, N Schmit, et al. The J-IDEA Pandemic Planner: A Framework for Implementing Hospital Provision Interventions During the COVID-19 Pandemic. *Medical Care*, 59(5):371–378, 2021. doi: 10.1097/MLR.0000000000001502.

[4] R McCabe, N Schmit, P Christen, JC D'Aeth, A Løchen, D Rizmie, et al. Adapting hospital capacity to meet changing demands during the COVID-19 pandemic. *BMC Medicine*, 369(413), 2020. doi: 10.1186/s12916-020-01781-w.

[5] W Msemburi, A Karlinsky, V Knutson, S Aleshin-Guendel, S Chatterji, and J Wakefield. The WHO estimates of excess mortality associated with the COVID-19 pandemic. *Nature*, 613, 2022. doi: 10.1038/s41586-022-05522-2.

[6] Impact of COVID-19 on people's livelihoods, their health and our food systems. Joint statement by ILO, FAO, IFAD and WHO, 2020. Available from: `https://www.who.int/news/item/13-10-2020-impact-of-covid-19-on-people's-livelihoods-their-health-and-our-food-systems`.

[7] RF Mollica and GL Fircchione. Mental and physical exhaustion of health-care practitioners. *The Lancet*, 398(10318), 2021. doi: 10.1016/S0140-6736(21)02663-5.

[8] AV Ballering, SKR van Zon, TC olde Hartman, and JGM Rosmalen. Persistence of somatic symptoms after COVID-19 in the Netherlands: an observational cohort study. *The Lancet*, 400(10350): 452–461, 2022. doi: 10.1016/S0140-6736(22)01214-4.

[9] DF Santomauro, AM Mantilla Herrera, J Shadid, P Zheng, C Ashbaugh, DM Pigott, et al. Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the COVID-19 pandemic. *The Lancet*, 398(10312), 2021. doi: 10.1016/S0140-6736(21) 02143-7.

[10] P Piot, M Bartos, PD Ghys, N Walker, and B Schwartländer. The global impact of HIV/AIDS. *Nature*, 410:968–973, 2001. doi: 10.1038/35073639.

[11] M Brzezinski. The impact of past pandemics on economic and gender inequalities. *Economics & Human Biology*, 43:101039, 2021. doi: 10.1016/j.ehb.2021.101039.

[12] M Salamatbakhsh, K Mobaraki, S Sadeghimohammadi, and J Ahmadzadeh. The global burden of premature mortality due to the Middle East respiratory syndrome (MERS) using standard expected years of life lost, 2012 to 2019. *BMC Public Health*, 19, 2019. doi: 10.1186/s12889-019-7899-2.

[13] JWT Elston, C Cartwright, P Ndumbi, and J Wright. The health impact of the 2014–15 Ebola outbreak. *Public Health*, 143:60–70, 2017. doi: 10.1016/j.puhe.2016.10.020.

[14] KM Adams Waldorf, EM Olson, BR Nelson, MTE Little, and L Rajagopal. The Aftermath of Zika: Need for Long-Term Monitoring of Exposed Children. *Trends in Microbiology*, 26(9):729–732, 2018. doi: 10.1016/j.tim.2018.05.011.

[15] NAM Molinari, IR Ortega-Sanchez, ML Messonnier, WW Thompson, PM Wortley, E Weintraub, and CB Bridges. The annual impact of seasonal influenza in the US: Measuring disease burden and costs. *Vaccine*, 25(27):5086–5096, 2007. doi: 10.1016/j.vaccine.2007.03.046.

[16] G Yamey, M Schäferhoff, OK Aars, B Bloom, D Carroll, M Chawla, et al. Financing of international collective action for epidemic and pandemic preparedness. *The Lancet Public Health*, 5, 2017. doi: 10.1016/S2214-109X(17)30203-6.

[17] PA Patriarca and NJ Cox. Influenza pandemic preparedness plan for the United States. *Journal of Infectious Diseases*, 176 Suppl 1:S4–7, 1997. doi: 10.1086/514174.

[18] AS Fauci. Pandemic influenza threat and preparedness. *Emerging Infectious Diseases*, 12:73–77, 2006. doi: 10.3201/eid1201.050983.

[19] HV Fineberg. Pandemic Preparedness and Response — Lessons from the H1N1 Influenza of 2009. *New England Journal of Medicine*, 370(14):1335–1342, 2014. doi: 10.1056/NEJMra1208802.

[20] BH Bird and JAK Mazet. Detection of Emerging Zoonotic Pathogens: An Integrated One Health Approach. *Annual Review of Animal Biosciences*, 6(1):121–139, 2018. doi: 10.1146/ annurev-animal-030117-014628.

[21] 2018 Annual review of diseases prioritized under the Research and Development Blueprint. The World Health Organization, 2018. Available from: `http://tinyurl.com/28usyxue`.

[22] Prioritizing diseases for research and development in emergency contexts. The World Health Organization, 2023. Available from: `https://www.who.int/activities/prioritizing-diseases-for-research-and-development-in-emergency-contexts`.

[23] Oxford University breakthrough on global COVID-19 vaccine. University of Oxford, 2021. Available from: `https://www.research.ox.ac.uk/article/2020-11-23-oxford-university-breakthrough-on-global-covid-19-vaccine`.

[24] Professor Dame Sarah Gilbert delivers 44th Dimbleby Lecture. University of Oxford, 2021. Available from: `https://www.ox.ac.uk/news/2021-12-07-professor-dame-sarah-gilbert-delivers-44th-dimbleby-lecture`.

[25] RECOVERY Trial rolled out across the UK. RECOVERY, 2020. Available from: `https://www.recoverytrial.net/news/update`.

[26] M Allan, L Lièvre, H Laurenson-Schafer, S de Barros, Y Jinnai, S Andrews, et al. The World Health Organization COVID-19 surveillance database. *International Journal for Equity in Health*, 21 (Supplement 3), 2022. doi: 10.1186/s12939-022-01767-5.

[27] EPI-WIN webinar: MERS-CoV, a circulating coronavirus with epidemic and pandemic potential - Pandemic preparedness, prevention and response with a One Health approach. World Health Organization, 2023. Available from: `http://tinyurl.com/5n7khfdj`.

[28] P Horby. Personal communication, 9 September 2023.

[29] H Laurenson-Schafer, N Sklenovská, A Hoxha, SM Kerr, P Ndumbi, J Fitzner, et al. Description of the first global outbreak of mpox: an analysis of global surveillance data. *The Lancet Global Health*, 11:e1012 – e1023, 2023. doi: 10.1016/S2214-109X(23)00198-5.

[30] M Cevik, AL Rasmussen, II Bogoch, and J Kindrachuk. Acute hepatitis of unknown origin in children. *BMJ*, 377, 2022. doi: 10.1136/bmj.o1197.

[31] S Mandal, R Simmons, G Ireland, A Charlett, M Desai, L Coughlan, et al. Paediatric acute hepatitis of unknown aetiology: a national investigation and adenoviraemia case-control study in the UK. *The Lancet Child & Adolescent Health*, 2023. doi: 10.1016/S2352-4642(23)00215-8.

[32] Ebola outbreak 2022 - Uganda. World Health Organization, 2022. Available from: `https://www.who.int/emergencies/situations/ebola-uganda-2022`.

[33] Equatorial Guinea confirms first-ever Marburg virus disease outbreak. World Health Organization, 2023. Available from: `https://www.afro.who.int/countries/equatorial-guinea/news/equatorial-guinea-confirms-first-ever-marburg-virus-disease-outbreak`.

[34] Tanzania confirms first-ever outbreak of Marburg Virus Disease. World Health Organization, 2023. Available from: `https://www.afro.who.int/countries/united-republic-of-tanzania/news/tanzania-confirms-first-ever-outbreak-marburg-virus-disease`.

[35] Cholera – Global situation. World Health Organization, 2023. Available from: `https://www.who.int/emergencies/disease-outbreak-news/item/2023-DON437`.

[36] Ongoing avian influenza outbreaks in animals pose risk to humans. World Health Organization, 2023. Available from: `https://www.who.int/news/item/12-07-2023-ongoing-avian-influenza-outbreaks-in-animals-pose-risk-to-humans`.

[37] Influenza A(H1N2) variant virus infection - United Kingdom of Great Britain and Northern Ireland. World Health Organization, 2023. Avaiable from: `https://www.who.int/emergencies/disease-outbreak-news/item/2023-DON496`.

[38] Upsurge of respiratory illnesses among children-Northern China. World Health Organization, 2023. Available from: `https://www.who.int/emergencies/disease-outbreak-news/item/2023-DON494`.

[39] Disease Outbreak News. World Health Organization, 2023. Available from: `https://www.who.int/emergencies/disease-outbreak-news/1`.

[40] WHO press conference on global health issues – 14 September 2023. World Health Organization, 2023. Available from: `http://tinyurl.com/4ur6zrx5`.

[41] C Viboud, K Sun, R Gaffey, M Ajelli, L Fumanelli, S Merler, et al. The RAPIDD ebola forecasting challenge: Synthesis and lessons learnt. *Epidemics*, 22:13–21, 3 2018. doi: 10.1016/j.epidem.2017.08.002.

[42] MD Van Kerkhove and N Ferguson. Epidemic and intervention modelling–a scientific rationale for policy decisions? Lessons from the 2009 influenza pandemic. *Bulletin of the World Health Organization*, 90:306–310, 2012. doi: 10.2471/BLT.11.097949.

[43] E Brooks-Pollock, L Danon, T Jombart, and L Pellis. Modelling that shaped the early COVID-19 pandemic response in the UK. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 376(1829):20210001, 7 2021. doi: 10.1098/rstb.2021.0001.

[44] NF Brazeau, R Verity, S Jenks, H Fu, C Whittaker, P Winskill, et al. Estimating the COVID-19 infection fatality ratio accounting for seroreversion using a statistical model. *Communications Medicine*, 2(1), 2022. doi: 10.1038/s43856-022-00106-7.

[45] J Lessler, H Salje, MD Van Kerkhove, NM Ferguson, S Cauchemez, I Rodriquez-Barraquer, et al. Estimating the Severity and Subclinical Burden of Middle East Respiratory Syndrome Coronavirus Infection in the Kingdom of Saudi Arabia. *American Journal of Epidemiology*, 183(7):657–663, 2016. doi: 10.1093/aje/kwv452.

[46] NG Davies, AJ Kucharski, RM Eggo, A Gimma, and WJ Edmunds. Effects of non-pharmaceutical interventions on COVID-19 cases, deaths, and demand for hospital services in the UK: a modelling study. *The Lancet Public Health*, 5(7):e375–e385, 7 2020. doi: 10.1016/S2468-2667(20)30133-X.

[47] SY Cho, JM Kang, YE Ha, GE Park, JY Lee, JH Ko, et al. MERS-CoV outbreak following a single patient exposure in an emergency room in South Korea: an epidemiological outbreak study. *The Lancet*, 388:994–1001, 2016. doi: 10.1016/S0140-6736(16)30623-7.

[48] Statement on the fifteenth meeting of the IHR (2005) Emergency Committee on the COVID-19 pandemic. World Health Organization, 2023. Available from: `http://tinyurl.com/yc3n6fbs`.

[49] WHO press conference on global health issues – 25 August 2023. World Health Organization, 2023. Available from: `https://www.who.int/multi-media/details/who-press-conference-on-global-health-issues-25-august-2023`.

[50] R McCabe and CA Donnelly. Disease transmission and control modelling at the science–policy interface. *Interface Focus*, 11(6):20210013, 2021. doi: 10.1098/rsfs.2021.0013.

[51] R McCabe, MD Kont, N Schmit, C Whittaker, A Løchen, PGT Walker, et al. Communicating uncertainty in epidemic models. *Epidemics*, 37:100520, 2021. ISSN 1755-4365. doi: 10.1016/j.epidem.2021.100520.

[52] R McCabe and CA Donnelly. Public awareness of and opinions on the use of mathematical transmission modelling to inform public health policy in the United Kingdom. *Journal of the Royal Society Interface*, 20:20230456, 2023. doi: 10.1098/rsif.2023.0456.

[53] E Mullins, R McCabe, SM Bird, P Randell, MJ Pond, L Regan, et al. Tracking the incidence and risk factors for SARS-CoV-2 infection using historical maternal booking serum samples. *PLOS ONE*, 17(9):1–15, 2022. doi: 10.1371/journal.pone.0273966.

[54] R McCabe, C Whittaker, RJ Sheppard, N Abdelmagid, A Ahmed, IZ Alabdeen, et al. Alternative

epidemic indicators for COVID-19 in three settings with incomplete death registration systems. *Science Advances*, 9(23):eadg7676, 2023. doi: 10.1126/sciadv.adg7676.

[55] R McCabe, G Danelian, J Panovska-Griffiths, and CA Donnelly. Inferring community transmission of SARS-CoV-2 in the United Kingdom using the ONS COVID-19 Infection Survey. *Infectious Disease Modelling*, 2024. doi: 10.1016/j.idm.2024.01.011. In press.

[56] Core UK Decision-making and Political Governance (Module 2) – Public Hearings, 2023. Available from: https://covid19.public-inquiry.uk/hearings/core-uk-decision-making-and-political-governance-module-2-public-hearings/.

[57] R McCabe, MD Kont, N Schmit, C Whittaker, A Løchen, M Baguelin, et al. Modelling intensive care unit capacity under different epidemiological scenarios of the COVID-19 pandemic in three Western European countries. *International Journal of Epidemiology*, 50(3):753–767, 2021. doi: 10.1093/ije/dyab034.

[58] OJ Watson, M Alhaffar, Z Mehchy, C Whittaker, Z Akil, NF Brazeau, et al. Leveraging community mortality indicators to infer COVID-19 mortality and transmission dynamics in Damascus, Syria. *Nature Communications*, 12(2394), 2021. doi: 10.1038/s41467-021-22474-9.

[59] COVID-19 Response: Living with COVID-19. Cabinet Office, 2022. Available from: https://www.gov.uk/government/publications/covid-19-response-living-with-covid-19/covid-19-response-living-with-covid-19.

[60] Coronavirus (COVID-19): Test and Protect - transition plan. Scottish Government, 2022. Available from: https://www.gov.scot/publications/test-protect-transition-plan/.

[61] Together for a safer future: Wales' long-term Covid-19 transition from pandemic to endemic. Welsh Government, 2022. Available from: https://gov.wales/sites/default/files/publications/2022-05/wales-long-term-covid-19-transition-from-pandemic-to-endemic.pdf.

[62] COVID-19 Test, Trace and Protect Transition Plan. Department of Health, Northern Ireland, 2022. Available from: https://www.health-ni.gov.uk/sites/default/files/publications/health/Test-Trace-Transition-Plan.pdf.

[63] COVID-19 Infection Survey. Office for National Statistics, 2023. Available from: https://www.ons.gov.uk/surveys/informationforhouseholdsandindividuals/householdandindividualsurveys/covid19infectionsurvey.

[64] The R value and growth rate. gov.uk, 2022. Available from: https://www.gov.uk/guidance/the-r-value-and-growth-rate.

[65] J Park, L Bevan, A Sanchez-Marroquin, G Danelian, T Bayley, H Manley, et al. Combining models to generate a consensus effective reproduction number R for the COVID-19 epidemic status in England. *medRxiv*, 2023. doi: 10.1101/2023.02.27.23286501.

[66] Expert Report to the Infected Blood Inquiry: Statistics. Infected Blood Inquiry, 2022. Available from: `https://www.infectedbloodinquiry.org.uk/sites/default/files/documents/Expert%20Report%20to%20the%20Infected%20Blood%20Inquiry%20-%20Statistics.pdf`.

[67] S Hayes, R McCabe, D De Angelis, CA Donnelly, SJW Evans, GF Medley, et al. Modelling hepatitis C infection acquired from blood transfusions in the UK between 1970 and 1991 for the Infected Blood Inquiry. *medRxiv*, 2023. doi: 10.1101/2023.11.03.23298029.

[68] B Du, C Wang, and M Singer. Learning for the next pandemic: the Wuhan experience of managing critically ill people. *BMJ*, 375, 2021. doi: 10.1136/bmj-2021-066090.

[69] R Verity, LC Okell, I Dorigatti, P Winskill, C Whittaker, N Imai, et al. Estimates of the severity of coronavirus disease 2019: a model-based analysis. *The Lancet Infectious Diseases*, 20(6):669–677, 2020. doi: 10.1016/S1473-3099(20)30243-7.

[70] Q Li, X Guan, P Wu, X Wang, L Zhou, Y Tong, et al. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia. *New England Journal of Medicine*, 382(13): 1199–1207, 2020. doi: 10.1056/NEJMoa2001316.

[71] N Imai, A Cori, I Dorigatti, M Baguelin, C Donnelly, S Riley, et al. Report 3: Transmissibility of 2019-nCoV. Imperial College London, 2020. Available from: https://doi.org/10.25561/77148.

[72] Z Du, Wang L, Cauchemez S, Xu X, Cowling BJ, and Meyers LA. Risk for Transportation of Coronavirus Disease from Wuhan to Other Cities in China. *Emerging Infectious Diseases*, 26, 2020. doi: 10.3201/eid2605.200146.

[73] JT Wu, K Leung, M Bushman, N Kishore, R Niehus, PM de Salazar, et al. Estimating clinical severity of COVID-19 from the transmission dynamics in Wuhan, China. *Nature Medicine*, 26, 2020. doi: 10.1038/s41591-020-0822-7.

[74] AL Phelan, R Katz, and LO Gostin. The Novel Coronavirus Originating in Wuhan, China: Challenges for Global Health Governance. *JAMA*, 323(8):709–710, 2020. doi: 10.1001/jama.2020.1097.

[75] H Fu, H Wang, X Xi, A Boonyasiri, Y Wang, W Hinsley, et al. Database of epidemic trends and control measures during the first wave of COVID-19 in mainland China. *International Journal of Infectious Diseases*, 102:463–471, 2021. ISSN 1201-9712. doi: 10.1016/j.ijid.2020.10.075.

[76] T Burki. The Wuhan Lockdown. *The Lancet Respiratory Medicine*, 10(7), 2022. doi: 10.1016/S2213-2600(22)00177-1.

[77] D Fu, Y He, and Y Guo. Wuhan lockdown after COVID-19 outbreak: a doctor's perspective. *The BMJ Opinion*, 2020. Available from: `https://blogs.bmj.com/bmj/2020/04/24/wuhan-lockdown-after-covid-19-outbreak-a-doctors-perspective/`.

[78] Wuhan lockdown 'unprecedented', shows commitment to contain virus: WHO representative in China. Reuters, 2020. Available from: `https://www.reuters.com/article/us-china-health-who-idUSKBN1ZM1G9`.

[79] L Kuo. Coronavirus: panic and anger in Wuhan as China orders city into lockdown. The Guardian, 2020. Available from: `https://www.theguardian.com/world/2020/jan/23/coronavirus-panic-and-anger-in-wuhan-as-china-orders-city-into-lockdown`.

[80] R Inocencio. Wuhan quarantines 11 million people as coronavirus spreads. CBS News, 2020. Available from: `https://www.cbsnews.com/news/coronavirus-wuhan-implements-unprecedented-lockdown-slow-spread-2020-01-22/`.

[81] J Griffiths, N Gan, T John, and A Vera. Wuhan coronavirus death toll rises, as city imposes transport lockdown. CNN, 2020. Available from: `https://edition.cnn.com/2020/01/22/asia/china-wuhan-coronavirus-deadly-intl-hnk/index.html`.

[82] J Goldstein and JE Singer. New York Braces for Coronavirus: 'It's Inevitable'. The New York Times, 2020. Available from: `https://www.nytimes.com/2020/01/27/nyregion/new-york-city-coronavirus.html`.

[83] In Pictures: Coronavirus outbreak in China. Al Jazeera, 2020. Available from: `https://www.aljazeera.com/gallery/2020/1/28/in-pictures-coronavirus-outbreak-in-china`.

[84] From Wuhan to Australia: A timeline of key events in the spread of the deadly coronavirus. ABC News, 2020. Available from: `https://www.abc.net.au/news/2020-01-29/coronavirus-timeline-from-wuhan-china-to-global-crisis/11903298`.

[85] Statement on the second meeting of the International Health Regulations (2005) Emergency Committee regarding the outbreak of novel coronavirus (2019-nCoV). World Health Organization, 2020. Available from: `http://tinyurl.com/3wjcd4cb`.

[86] International Health Regulations (2005) – Third edition. World Health Organization, 2016. Available from: `https://www.who.int/publications/i/item/9789241580496`.

[87] WHO Coronavirus (COVID-19) Dashboard. World Health Organization, 2023. Available from: `https://covid19.who.int/`.

[88] Rolling updates on coronavirus disease (COVID-19). World Health Organization, 2020. Available from: `https://www.who.int/emergencies/diseases/novel-coronavirus-2019/events-as-they-happen`.

[89] COVID-19 Government Response Tracker. Blavatnik School of Government, 2023. Available from: `https://www.bsg.ox.ac.uk/research/covid-19-government-response-tracker`.

[90] M Day. Covid-19: surge in cases in Italy and South Korea makes pandemic look more likely. *BMJ*, 368, 2020. doi: 10.1136/bmj.m751.

[91] COVID-19: examining the effectiveness of non-pharmaceutical interventions. The Royal Society, 2023. Available from: `https://tinyurl.com/mr2suvk9`.

[92] From lockdowns to pool parties: how Covid rules vary around the world. The Guardian, 2021. Available from: `https://www.theguardian.com/world/2021/jan/30/from-lockdowns-to-pool-parties-how-covid-rules-vary-around-the-world`.

[93] M Paterlini. What now for Sweden and covid-19? *BMJ*, 375, 2021. doi: 10.1136/bmj.n3081.

[94] H Habib. Has Sweden's controversial covid-19 strategy been successful? *BMJ*, 369, 2020. doi: 10.1136/bmj.m2376.

[95] U Agrawal, S Bedston, C McCowan, J Oke, L Patterson, and C Robertson. Severe COVID-19 outcomes after full vaccination of primary schedule and initial boosters: pooled analysis of national prospective cohort studies of 30 million individuals in England, Northern Ireland, Scotland, and Wales. *The Lancet*, 400, 2022. doi: 10.1016/S0140-6736(22)01656-7.

[96] SPI-M-O: Summary of further modelling of easing restrictions – Roadmap Step 2, 31 March 2021. gov.uk, 2021. Available from: `http://tinyurl.com/3uujbsan`.

[97] OJ Watson, G Barnsley, J Toor, AB Hogan, P Winskill, and AC Ghani. Global impact of the first year of COVID-19 vaccination: a mathematical modelling study. *The Lancet Infectious Diseases*, 22(9):1293–1302, 2022. doi: 10.1016/S1473-3099(22)00320-6.

[98] COVID: How is Europe lifting lockdown restrictions? BBC, 2021. Available from: `https://www.bbc.co.uk/news/explainers-53640249`.

[99] E Mathieu, H Ritchie, E Ortiz-Ospina, M Roser, J Hasell, C Appel, et al. A global database of COVID-19 vaccinations. *Nature Human Behaviour*, 5(7), 2021. doi: 10.1038/s41562-021-01122-8.

[100] Accelerating COVID-19 Vaccine Deployment. World Health Organization, 2022. Available from: https://www.who.int/publications/m/item/accelerating-covid-19-vaccine-deployment.

[101] Coronavirus (COVID-19) Vaccinations. Our World in Data, 2023. Available from: https://ourworldindata.org/covid-vaccinations.

[102] World Bank Country and Lending Groups. The World Bank, 2023. Available from: https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups.

[103] TV Padma. COVID vaccines to reach poorest countries in 2023 — despite recent pledges. *Nature*, 595:342–343, 2021. doi: 10.1038/d41586-021-01762-w.

[104] Weekly epidemiological update on COVID-19 - 20 July 2021. World Health Organization, 2021. Available from: https://www.who.int/publications/m/item/weekly-epidemiological-update-on-covid-19---20-july-2021.

[105] LIVE Q&A on #COVID19 with @DrMikeRyan & Dr @mvankerkhove. #AskWHO. World Health Organization, 2021. Available from: https://www.pscp.tv/w/1djGXqPoAQzJZ.

[106] COVAX. World Health Organization, 2023. Available from: https://www.act-a.org/vaccines-covax.

[107] EPI-WIN webinar: Global Update on COVID-19: managing the risks from new variants. World Health Organization, 2023. Available from: https://tinyurl.com/yfxmjchm.

[108] J Yeung and B Westcott. Malaysia, Thailand and Vietnam are leaving their zero-Covid policies behind, but they aren't ready to open up, experts warn. CNN, 2021. Available from: https://edition.cnn.com/2021/09/21/asia/covid-zero-asia-vietnam-thailand-malaysia-intl-hnk/index.html.

[109] B Westcott. New Zealand to abandon zero-Covid strategy as Delta variant proves hard to shake. CNN, 2021. Available from: https://edition.cnn.com/2021/10/05/asia/new-zealand-ardern-covid-zero-intl-hnk/index.html.

[110] F Delussu, M Tizzoni, and Gauvin. Evidence of pandemic fatigue associated with stricter tiered COVID-19 restrictions. *PLOS Digital Health*, 1(5):1–14, 2022. doi: 10.1371/journal.pdig.0000035.

[111] S Mallapaty. China is relaxing its zero-COVID policy — here's what scientists think. *Nature*, 612, 2022. doi: 10.1038/d41586-022-04382-0.

[112] Coronavirus (COVID-19) in the UK. gov.uk, 2023. Available from: `https://coronavirus.data.gov.uk/`.

[113] L Danon, E Brooks-Pollock, M Bailey, and M Keeling. A spatial model of COVID-19 transmission in England and Wales: early spread, peak timing and the impact of seasonality. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 376(1829):20200272, 7 2021. doi: 10.1098/rstb.2020.0272.

[114] L du Plessis, JT McCrone, AE Zarebski, V Hill, C Ruis, B Gutierrez, et al. Establishment and lineage dynamics of the SARS-CoV-2 epidemic in the UK. *Science*, 371(6530):708–712, 2021. doi: 10.1126/science.abf2946.

[115] M Jit, T Jombart, ES Nightingale, A Endo, S Abbott, et al. Estimating number of cases and spread of coronavirus disease (COVID-19) using critical care admissions, United Kingdom, February to March 2020. *Eurosurveillance*, 25(18):2000632, 2020. doi: 10.2807/1560-7917.ES.2020.25.18.2000632.

[116] M Weaver. Daughter calls for Kent man to be recognised as UK's first Covid victim. The Guardian, 2021. Available from: `https://www.theguardian.com/uk-news/2021/mar/04/daughter-calls-for-kent-man-to-be-recognised-as-uks-first-covid-victim`.

[117] Timeline of UK government coronavirus lockdowns and restrictions. Institute for Government, 2023. Available from: `https://www.instituteforgovernment.org.uk/data-visualisation/timeline-coronavirus-lockdowns`.

[118] Full list of local restriction tiers by area. gov.uk, 2020. Available from: `https://www.gov.uk/guidance/full-list-of-local-restriction-tiers-by-area`.

[119] Coronavirus (COVID-19): protection levels - reviews and evidence. Scottish Government, 2020. Available from: `https://www.gov.scot/collections/coronavirus-covid-19-protection-levels-reviews-and-evidence/`.

[120] Coronavirus Control Plan: Alert Levels in Wales. Welsh Government, 2021. Available from: `https://www.gov.wales/sites/default/files/publications/2021-02/coronavirus-control-plan-alert-levels-in-wales.pdf`.

[121] Covid-19: NI schools to close and three-tier system begins. BBC News, 2020. Available from: `https://www.bbc.co.uk/news/uk-54527565`.

[122] Guidance for the Christmas period. gov.uk, 2020. Available from: `https://www.gov.uk/guidance/guidance-for-the-christmas-period`.

[123] New guidance issued for the festive period. Scottish Government, 2020. Available from: https://www.gov.scot/news/new-guidance-issued-for-the-festive-period/.

[124] Written Statement: Alert level four restrictions. Welsh Government, 2020. Available from: https://www.gov.wales/written-statement-alert-level-four-restrictions.

[125] Coronavirus: NI facing six-week lockdown from 26 December. BBC News, 2020. Available from: https://www.bbc.co.uk/news/uk-northern-ireland-55349545.

[126] Christmas rules 2020: What are the new rules on mixing? BBC News, 2020. Available from: https://www.bbc.co.uk/news/explainers-55056375.

[127] UK COVID-19 vaccines delivery plan. gov.uk, 2021. Available from: https://www.gov.uk/government/publications/uk-covid-19-vaccines-delivery-plan/uk-covid-19-vaccines-delivery-plan.

[128] JT McCrone, V Hill, S Bajaj, RE Pena, BC Lambert, R Inward, et al. Context-specific emergence and growth of the SARS-CoV-2 Delta variant. *Nature*, 610, 2022. doi: 10.1038/s41586-022-05200-3.

[129] W Yang and J Shaman. COVID-19 pandemic dynamics in India, the SARS-CoV-2 Delta variant and implications for vaccination. *Journal of The Royal Society Interface*, 19(191):20210900, 2022. doi: 10.1098/rsif.2021.0900.

[130] M Pons-Salort, J John, OJ Watson, NF Brazeau, R Verity, G Kang, et al. Reassessing Reported Deaths and Estimated Infection Attack Rate during the First 6 Months of the COVID-19 Epidemic, Delhi, India. *Emerging Infectious Diseases*, 28:759–766, 2022. doi: 10.3201/eid2804.210879.

[131] COVID-19 Response - Spring 2021 (Summary). gov.uk, 2021. Available from: https://www.gov.uk/government/publications/covid-19-response-spring-2021/covid-19-response-spring-2021-summary.

[132] Coronavirus (COVID-19): Scotland's route map through and out of the crisis. Scottish Government, 2020. Available from: http://tinyurl.com/4msuts9u.

[133] Coronavirus timeline: the response in Wales. Welsh Government, 2023. Available from: https://research.senedd.wales/research-articles/coronavirus-timeline-the-response-in-wales/.

[134] Covid-19: What is the current picture shaping the roadmap? BBC News, 2021. Available from: https://www.bbc.co.uk/news/uk-northern-ireland-56301394.

[135] J Tapper. Dismay as funding for UK's 'world-beating' Covid trackers is axed. The Guardian, 2022. Available from: https://www.theguardian.com/world/2022/mar/12/dismay-as-funding-for-uks-world-beating-covid-trackers-is-axed.

[136] COVID-19 Infection Survey participants thanked for 'huge contribution' to pandemic response. UK Health Security Agency, 2023. Available from: http://tinyurl.com/ys49czh2.

[137] UKHSA and ONS launch new Winter COVID-19 Infection Study. UKHSA, 2023. Available from: https://ukhsa-newsroom.prgloo.com/news/ukhsa-and-ons-launch-new-winter-covid-19-infection-study.

[138] Winter Coronavirus (COVID-19) Infection Study, England and Scotland. Office for National Statistics, 2023. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/datasets/wintercoronaviruscovid19infectionstudyenglandandscotland.

[139] UK Covid-19 Inquiry, 2023. Available from: https://covid19.public-inquiry.uk/.

[140] Clare Dyer. Covid-19: UK's pandemic inquiry prepares to hear from witnesses. *BMJ*, 381, 2023. doi: 10.1136/bmj.p1314.

[141] JW Tang, LC Marr, Y Li, and I Eames. The role of SARS-CoV-2 aerosol transmission during the COVID-19 pandemic. *Interface Focus*, 12(2):20220003, 2022. doi: 10.1098/rsfs.2022.0003.

[142] H Rutter, S Parker, W Stahl-Timmins, C Noakes, A Smyth, R Macbeth, S Fitzgerald, and ALJ Freeman. Visualising SARS-CoV-2 transmission routes and mitigations. *BMJ*, 375, 2021. doi: 10.1136/bmj-2021-065312.

[143] R Zhang, Y Li, AL Zhang, and MJ Molina. Identifying airborne transmission as the dominant route for the spread of COVID-19. *PNAS*, 117, 2020. doi: 10.1073/pnas.2009637117.

[144] Key points summary: asymptomatic testing for SARS-CoV-2 using antigen-detecting lateral flow devices (evidence from performance data October 2020 to May 2021). gov.uk, 2023. Available from: https://tinyurl.com/3cr2mmd2.

[145] W Shang, L Kang, G Cao, Y Wang, p Gao, J Liu, et al. Percentage of Asymptomatic Infections among SARS-CoV-2 Omicron Variant-Positive Individuals: A Systematic Review and Meta-Analysis. *Vaccines*, 10, 2022. doi: 10.3390/vaccines10071049.

[146] Q Ma, J Liu, Q Liu, L Kang, R Liu, W Jing, et al. Global Percentage of Asymptomatic SARS-CoV-2 Infections Among the Tested Population and Individuals With Confirmed COVID-19 Diagnosis:

A Systematic Review and Meta-analysis. *JAMA Network Open*, 4(12):e2137257–e2137257, 2021. doi: 10.1001/jamanetworkopen.2021.37257.

[147] X Chen, Z Huang, J Wang, S Zhao, M Chi-Sang Wong, K Chun Chong, et al. Ratio of asymptomatic COVID-19 cases among ascertained SARS-CoV-2 infections in different regions and population groups in 2020: a systematic review and meta-analysis including 130 123 infections from 241 studies. *BMJ Open*, 11(12), 2021. doi: 10.1136/bmjopen-2021-049752.

[148] AL Rasmussen and SV Popescu. SARS-CoV-2 transmission without symptoms. *Science*, 371(6535): 1206–1207, 2021. doi: 10.1126/science.abf9569.

[149] DP Oran and EJ Topol. Prevalence of Asymptomatic SARS-CoV-2 Infection. *Annals of Internal Medicine*, 173(5):362–367, 2020. doi: 10.7326/M20-3012.

[150] Y Song and E Shim. Proportion of Pre-Symptomatic Transmission Events Associated with COVID-19 in South Korea. *Journal of Clinical Medicine*, 11, 2022. doi: 10.3390/jcm11143925.

[151] M Casey-Bryars, J Griffin, C McAloon, A Byrne, J Madden, D McEvoy, et al. Presymptomatic transmission of SARS-CoV-2 infection: a secondary analysis using published data. *BMJ Open*, 11 (6), 2021. doi: 10.1136/bmjopen-2020-041240.

[152] W Gao, J Lv, Y Pang, and L Li. Role of asymptomatic and pre-symptomatic infections in covid-19 pandemic. *BMJ*, 375, 2021. doi: 10.1136/bmj.n2342.

[153] MA Mugglestone, NV Ratnaraja, A Bak, J Islam, JE Wilson, J Bostock, et al. Presymptomatic, asymptomatic and post-symptomatic transmission of SARS-CoV-2: joint British Infection Association (BIA), Healthcare Infection Society (HIS), Infection Prevention Society (IPS) and Royal College of Pathologists (RCPath) guidance. *BMC Infectious Disease*, 22, 2022. doi: 10.1186/s12879-022-07440-0.

[154] COVID-19 Pandemic Planning Scenarios. Centers for Disease Control and Prevention, 2021. Available from: `https://www.cdc.gov/coronavirus/2019-ncov/hcp/planning-scenarios.html`.

[155] A Parasher. COVID-19: Current understanding of its Pathophysiology, Clinical presentation and Treatment. *Postgraduate Medical Journal*, 97(1147):312–320, 2020. doi: 10.1136/postgradmedj-2020-138577.

[156] Clinical Spectrum of SARS-CoV-2 Infection. National Institute for Health, 2023. Available from: `https://www.covid19treatmentguidelines.nih.gov/overview/clinical-spectrum/`.

[157] M Lipman, RC Chambers, M Singer, and JS Brown. SARS-CoV-2 pandemic: clinical picture of COVID-19 and implications for research. *Thorax*, 75(8):614–616, 2020. doi: 10.1136/thoraxjnl-2020-215024.

[158] Symptoms of COVID-19. US Center for Disease Control, 2022. Available from: `https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/symptoms.html`.

[159] PN Perez-Guzman, A Daunt, S Mukherjee, P Crook, R Forlano, MD Kont, et al. Clinical Characteristics and Predictors of Outcomes of Hospitalized Patients With Coronavirus Disease 2019 in a Multiethnic London National Health Service Trust: A Retrospective Cohort Study. *Clinical Infectious Diseases*, 73(11):e4047–e4057, 2020. doi: 10.1093/cid/ciaa1091.

[160] M-K Looi. How are covid-19 symptoms changing? *BMJ*, 380, 2023. doi: 10.1136/bmj.p3.

[161] H Salje, C Tran Kiem, N Lefrancq, N Courtejoie, P Bosetti, J Paireau, et al. Estimating the burden of SARS-CoV-2 in France. *Science*, 369(6500):208–211, 2020. doi: 10.1126/science.abc3517.

[162] Statement from the UK Chief Medical Officers on an update to coronavirus symptoms: 18 May 2020. Department of Health and Social Care, 2020. Available from: `https://tinyurl.com/4n7f4xws`.

[163] CH Sudre, B Murray, T Varsavsky, MS Graham, RS Penfold, RC Bowyer, et al. Attributes and predictors of long COVID. *Nature Medicine*, 27, 2021. doi: 10.1038/s41591-021-01292-y.

[164] Coronavirus disease (COVID-19): Post COVID-19 condition. World Health Organization, 2023. Available from: `https://www.who.int/news-room/questions-and-answers/item/coronavirus-disease-(covid-19)-post-covid-19-condition`.

[165] T Beaney, A L Neves, A Alboksmaty, H Ashrafian, K Flott, A Fowler, et al. Trends and associated factors for Covid-19 hospitalisation and fatality risk in 2.3 million adults in England. *Nature Communications*, 13, 2022. doi: 10.1038/s41467-022-29880-7.

[166] Deaths involving COVID-19 by local area and socioeconomic deprivation: deaths occurring between 1 March and 31 July 2020. Office for National Statistics, 2020. Available from: `https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/bulletins/deathsinvolvingcovid19bylocalareasanddeprivation/deathsoccurringbetween1marchand31july2020`.

[167] CE Rodriguez-Diaz, V Guilamo-Ramos, L Mena, E Hall, B Honermann, JS Crowley, et al. Risk for COVID-19 infection and death among Latinos in the United States: examining heterogeneity in transmission dynamics. *Annals of Epidemiology*, 52:46–53.e2, 2020. doi: 10.1016/j.annepidem.2020.07.007.

[168] JY Ko, ML Danielson, M Town, G Derado, KJ Greenlund, PD Kirley, et al. Risk Factors for Coronavirus Disease 2019 (COVID-19)–Associated Hospitalization: COVID-19–Associated Hospitalization Surveillance Network and Behavioral Risk Factor Surveillance System. *Clinical Infectious Diseases*, 72(11):e695–e703, 2020. doi: 10.1093/cid/ciaa1419.

[169] LH Nguyen, DA Drew, MS Graham, AD Joshi, C-G Guo, W Ma, et al. Risk of COVID-19 among front-line health-care workers and the general community: a prospective cohort study. *The Lancet Public Health*, 5(9):e475–e483, 9 2020. doi: 10.1016/S2468-2667(20)30164-X.

[170] Director-General's opening remarks at the World Health Assembly - 24 May 2021. World Health Organization, 2021. Available from: `https://www.who.int/director-general/speeches/detail/director-general-s-opening-remarks-at-the-world-health-assembly---24-may-2021`.

[171] Updated working definitions and primary actions for SARS-CoV-2 variants. World Health Organization, 2023. Available from: `https://www.who.int/publications/m/item/updated-working-definitions-and-primary-actions-for--sars-cov-2-variants`.

[172] WS Hart, E Miller, NJ Andrews, P Waight, PK Maini, S Funk, et al. Generation time of the alpha and delta SARS-CoV-2 variants: an epidemiological analysis. *The Lancet Infectious Diseases*, 22 (5), 2022. doi: 10.1016/S1473-3099(22)00001-9.

[173] Historical working definitions and primary actions for SARS-CoV-2 variants. World Health Organization, 2023. Available from: `https://www.who.int/publications/m/item/historical-working-definitions-and-primary-actions-for-sars-cov-2-variants`.

[174] SARS-CoV-2 variants of concern as of 21 September 2023. European Centre for Disease Prevention and Control, 2023. Available from: `https://www.ecdc.europa.eu/en/covid-19/variants-concern`.

[175] H Tegally, E Wilkinson, M Giovanetti, A Iranzadeh, V Fonseca, J Giandhairi, et al. Detection of a SARS-CoV-2 variant of concern in South Africa. *Nature*, 592, 2021. doi: 10.1038/s41586-021-03402-9.

[176] MUG Kraemer, V Hill, C Ruis, S Dellicour, S Bajaj, JT McCrone, et al. Spatiotemporal invasion dynamics of SARS-CoV-2 lineage B.1.1.7 emergence. *Science*, 373(6557):889–895, 2021. doi: 10.1126/science.abj0113.

[177] NR Faria, TA Mellan, C Whittaker, IM Claro, DDS Candido, S Mishra, et al. Genomics and epidemiology of the P.1 SARS-CoV-2 lineage in Manaus, Brazil. *Science*, 372(6544):815–821, 2021. doi: 10.1126/science.abh2644.

[178] AS Walker, KD Vihta, O Gethings, E Pritchard, J Jones, T House, et al. Tracking the Emergence of SARS-CoV-2 Alpha Variant in the United Kingdom. *New England Journal of Medicine*, 385 (27):2582–2585, 2021. doi: 10.1056/NEJMc2103227.

[179] T Funk, A Pharris, G Spiteri, N Bundle, A Melidou, M Carr, and Gothers. Characteristics of SARS-CoV-2 variants of concern B.1.1.7, B.1.351 or P.1: data from seven EU/EEA countries, weeks 38/2020 to 10/2021. *Eurosurveillance*, 26(16), 2021. doi: 10.2807/1560-7917.ES.2021.26.16.2100348.

[180] A O'Toole, V Hill, OG Pybus, A Watts, II Bogoch, K Khan, et al. Tracking the international spread of SARS-CoV-2 lineages B.1.1.7 and B.1.351/501Y-V2 with grinch. *Wellcome Open Research*, 6, 2021. doi: 10.12688/wellcomeopenres.16661.2.

[181] Risk related to spread of new SARS-CoV-2 variants of concern in the EU-/EEA. European Centre for Disease Prevention and Control, 2020. Available from: `https://www.ecdc.europa.eu/sites/default/files/documents/COVID-19-risk-related-to-spread-of-new-SARS-CoV-2-variants-EU-EEA.pdf`.

[182] E Volz, S Mishra, M Chand, JC Barrett, R Johnson, and L Geidelberg. Assessing transmissibility of SARS-CoV-2 lineage B.1.1.7 in England. *Nature*, 593, 2021. doi: 10.1038/s41586-021-03470-x.

[183] CAB Pearson, TW Russell, NG Davies, AJ Kucharski, CMMID COVID-19 Working Group, WJ Edmunds, et al. Estimates of severity and transmissibility of novel South Africa SARS-CoV-2 variant 501Y.V2. Centre for Mathematical Modelling of Infectious Diseases, 2021. Available from `https://cmmid.github.io/topics/covid19/reports/sa-novel-variant/2021_01_11_Transmissibility_and_severity_of_501Y_V2_in_SA.pdf` [Preprint].

[184] FG Naveca, V Nascimento, VC de Souza, AL Corado, F Nascimento, G Silva, et al. COVID-19 in Amazonas, Brazil, was driven by the persistence of endemic lineages and P.1 emergence. *Nature Medicine*, 27, 2021. doi: 10.1038/s41591-021-01378-7.

[185] EC Sabino, LF Buss, MPS Carvalho, CA Prete, MAE Crispim, NA Fraiji, et al. Resurgence of COVID-19 in Manaus, Brazil, despite high seroprevalence. *The Lancet*, 397, 2021. doi: 10.1016/S0140-6736(21)00183-5.

[186] PN Perez-Guzman, E Knock, N Imai, T Rawson, Y Elmaci, J Alcada, et al. Epidemiological drivers of transmissibility and severity of SARS-CoV-2 in England. *Nature Communications*, 14, 2023. doi: 10.1038/s41467-023-39661-5.

[187] B Zeng, L Gao, Q Zhou, K Yu, and F Sun. Effectiveness of COVID-19 vaccines against SARS-CoV-2 variants of concern: a systematic review and meta-analysis. *BMC Medicine*, 20, 2022. doi: 10.1186/s12916-022-02397-y.

[188] MA Gram, H-D Emborg, AB Schelde, NU Friis, KF Nielsen, IR Moustsen-Helms, et al. Vaccine effectiveness against SARS-CoV-2 infection or COVID-19 hospitalization with the Alpha, Delta, or Omicron SARS-CoV-2 variant: A nationwide Danish cohort study. *PLOS Medicine*, 19(9):1–18, 09 2022. doi: 10.1371/journal.pmed.1003992.

[189] T Braeye, L Catteau, R Brondeel, JAF. van Loenhout, K Proesmans, L Cornelissen, et al. Vaccine effectiveness against transmission of alpha, delta and omicron SARS-COV-2-infection, Belgian contact tracing, 2021–2022. *Vaccine*, 41(20):3292–3300, 2023. doi: 10.1016/j.vaccine.2023.03.069.

[190] NG Davies, CI Jarvis, K van Zandvoort, S Clifford, FY Sun, S Funk, et al. Increased mortality in community-tested cases of SARS-CoV-2 lineage B.1.1.7. *Nature*, 593, 2021. doi: 10.1038/s41586-021-03426-1.

[191] LJ Abu-Raddad, H Chemaitelly, HH Ayoub, HM Yassine, FM Benslimane, HA Al Khatib, et al. Severity, Criticality, and Fatality of the Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Beta Variant. *Clinical Infectious Diseases*, 75(1):e1188–e1191, 2021. doi: 10.1093/cid/ciab909.

[192] ARR Freitas, OA Beckedorff, LP de Góes Cavalcanti, AM Siqueira, DB de Castro, CF da Costa, et al. The emergence of novel SARS-CoV-2 variant P.1 in Amazonas (Brazil) was temporally associated with a change in the age and sex profile of COVID-19 mortality: A population based ecological study. *The Lancet Regional Health - Americas*, 1:100021, 2021. doi: 10.1016/j.lana.2021.100021.

[193] D Tian, Y Sun, J Zhou, and Q Ye. The Global Epidemic of the SARS-CoV-2 Delta Variant, Key Spike Mutations and Immune Escape. *Frontiers in Immunology*, 12, 2021. doi: 10.3389/fimmu.2021.751778.

[194] SPI-M-O: Consensus Statement on COVID-19. gov.uk, 2021. Available from: `https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/993321/S1267_SPI-M-O_Consensus_Statement.pdf`.

[195] P Mlcochova, SA Kemp, MS Dhar, G Papa, B Meng, IATM Ferreira, et al. SARS-CoV-2 B.1.617.2 Delta variant replication and immune evasion. *Nature*, 599, 2021. doi: 10.1038/s41586-021-03944-y.

[196] W Yang and JL Shaman. COVID-19 pandemic dynamics in South Africa and epidemiological characteristics of three variants of concern (Beta, Delta, and Omicron). *eLife*, 11:e78933, 2022. doi: 10.7554/eLife.78933.

[197] WHO Director-General's opening remarks at the 8th meeting of the IHR Emergency Committee on COVID-19 – 14 July 2021. World Health Organization, 2021. Available from: `https://tinyurl.com/y4pv79c7`.

[198] J Lopez Bernal, N Andrews, C Gower, E Gallagher, R Simmons, S Thelwall, et al. Effectiveness of Covid-19 Vaccines against the B.1.617.2 (Delta) Variant. *New England Journal of Medicine*, 385 (7):585–594, 2021. doi: 10.1056/NEJMoa2108891.

[199] A Sheikh, C Robertson, and B Taylor. BNT162b2 and ChAdOx1 nCoV-19 Vaccine Effectiveness against Death from the Delta Variant. *New England Journal of Medicine*, 385(23):2195–2197, 2021. doi: 10.1056/NEJMc2113864.

[200] A Sheikh, J McMenamin, B Taylor, and C Robertson. SARS-CoV-2 Delta VOC in Scotland: demographics, risk of hospital admission, and vaccine effectiveness. *The Lancet*, 397, 2021. doi: 10.1016/S0140-6736(21)01358-1.

[201] AB Hogan, SL Wu, P Doohan, OJ Watson, P Winskill, G Charles, et al. Report 48: The value of vaccine booster doses to mitigate the global impact of the Omicron SARS-CoV-2 variant. Imperial College London, 2021. Available from: `https://doi.org/10.25561/93034`.

[202] N Ferguson, A Ghani, A Cori, A Hogan, W Hinsley, E Volz, et al. Report 49: Growth, population distribution and immune escape of Omicron in England. Imperial College London, 2021. Available from: `https://doi.org/10.25561/93038`.

[203] TK Burki. Omicron variant and booster COVID-19 vaccines. *The Lancet Respiratory Medicine*, 10 (2):e17, 2022. ISSN 2213-2600. doi: 10.1016/S2213-2600(21)00559-2.

[204] JCVI advice on the UK vaccine response to the Omicron variant. Department for Health and Social Care, 2021. Available from: `https://www.gov.uk/government/publications/uk-vaccine-response-to-the-omicron-variant-jcvi-advice/jcvi-advice-on-the-uk-vaccine-response-to-the-omicron-variant`.

[205] One year since the emergence of COVID-19 virus variant Omicron. World Health Organization, 2022. Available from: `https://www.who.int/news-room/feature-stories/detail/one-year-since-the-emergence-of-omicron`.

[206] N Ferguson, A Ghani, W Hinsley, E Volz, et al. Report 50 - Hospitalisation risk for Omicron cases in England. Imperial College London, 2021. Available from: `https://doi.org/10.25561/93035`.

[207] P Relan, NV Motaze, K Kothari, L Askie, O Le Polain de Waroux, MD Van Kerkhove, et al. Severity and outcomes of Omicron variant of SARS-CoV-2 compared to Delta variant and severity of Omicron sublineages: a systematic review and metanalysis. *BMJ Global Health*, 8(7), 2023. doi: 10.1136/bmjgh-2023-012328.

[208] C Hyams, R Challen, R Marlow, E Nguyen, E Begier ad J Southern, et al. Severity of Omicron (B.1.1.529) and Delta (B.1.617.2) SARS-CoV-2 infection among hospitalised adults: a prospective cohort study in Bristol, United Kingdom. *The Lancet Regional Health – Europe*, 25, 2023. doi: 10.1016/j.lanepe.2022.100556.

[209] COVID-19 Epidemiological Update - 24 November 2023. World Health Organization, 2023. Available from: `https://www.who.int/publications/m/item/covid-19-epidemiological-update---24-november-2023`.

[210] CL Gibbons, MJJ Mangen, D Plass, AH Havelaar, RJ Brooke, P Kramarz, et al. Measuring underreporting and under-ascertainment in infectious disease datasets: a comparison of methods. *BMC Public Health*, 14(1), 2014. doi: 10.1186/1471-2458-14-147.

[211] V Albani, J Loria, E Massad, and J Zubelli. COVID-19 underreporting and its impact on vaccination strategies. *BMC Infectious Diseases*, 21, 2021. doi: 10.1186/s12879-021-06780-7.

[212] TJ Doyle, MK Glynn, and SL Groseclose. Completeness of Notifiable Infectious Disease Reporting in the United States: An Analytical Literature Review. *American Journal of Epidemiology*, 155 (9):866–874, 2002. doi: 10.1093/aje/155.9.866.

[213] D Ghosh, JA Bernstein, and TB Mersha. COVID-19 pandemic: The African paradox. *Journal of Global Health*, 10, 2020. doi: 10.7189/jogh.10.020348.

[214] C Whittaker, PGT Walker, M Alhaffar, A Hamlet, BA Djaafara, A Ghani, et al. Under-reporting of deaths limits our understanding of true burden of covid-19. *BMJ*, 375, 2021. doi: 10.1136/bmj.n2239.

[215] S Anand, M Montez-Rath, J Han, J Bozeman, R Kerschmann, P Beyer, et al. Prevalence of SARS-CoV-2 antibodies in a large nationwide sample of patients on dialysis in the USA: a cross-sectional study. *The Lancet*, 396(10259):1335–1344, 2020. doi: 10.1016/S0140-6736(20)32009-2.

[216] S Riley, KEC Ainslie, O Eales, CE Walters, H Wang, C Atchison, et al. Resurgence of SARS-CoV-2:

Detection by community viral surveillance. *Science*, 372(6545):990–995, 2021. doi: 10.1126/science.abf0874.

[217] G Nicholson, B Lehmann, T Padellini, KB Pouwels, R Jersakova, J Lomax, et al. Improving local prevalence estimates of SARS-CoV-2 infections using a causal debiasing framework. *Nature Microbiology*, 7(1), 2022. doi: 10.1038/s41564-021-01029-0.

[218] N Imai, I Dorigatti, A Cori, S Riley, and N Ferguson. Report 1 - Estimating the potential total number of novel Coronavirus (2019-nCoV) cases in Wuhan City, China. Imperial College London, 2020. Available from: `https://doi.org/10.25561/77149`.

[219] SCORE global report 2020. World Health Organization, 2020. Available from: `https://www.who.int/data/stories/score-global-report-2020---a-visual-summary`.

[220] ES Koum Besson, A Norris, AS Bin-Ghouth, T Freemantle, M Alhaffar, Y Vazquez, et al. Excess mortality during the COVID-19 pandemic: a geospatial and statistical analysis in Aden governorate, Yemen. *BMJ Global Health*, 6(3), 2021. doi: 10.1136/bmjgh-2020-004564.

[221] CJ Gill, L Mwananyanda, WB MacLeod, G Kwenda, RC Pieciak, L Etter, et al. What is the prevalence of COVID-19 detection by PCR among deceased individuals in Lusaka, Zambia? A postmortem surveillance study. *BMJ Open*, 12(12), 2022. doi: 10.1136/bmjopen-2022-066763.

[222] C Morris and A Reuben. Coronavirus: Why are international comparisons difficult? BBC, 2020. Available from: `https://www.bbc.co.uk/news/52311014`.

[223] T Riffe, E Acosta, and the COVerAGE-DB team. Data Resource Profile: COVerAGE-DB: a global demographic database of COVID-19 cases and deaths. *International Journal of Epidemiology*, 50 (2):390–390f, 2021. doi: 10.1093/ije/dyab027.

[224] A Karlinsky and D Kobak. Tracking excess mortality across countries during the COVID-19 pandemic with the World Mortality Dataset. *eLife*, 10:e69336, 2021. doi: 10.7554/eLife.69336.

[225] N Dean. Tracking COVID-19 infections: time for change. *Nature*, 602(7896), 2022. doi: 10.1038/d41586-022-00336-8.

[226] J Clarke, T Beaney, and A Majeed. UK scales back routine covid-19 surveillance. *BMJ*, 376, 2022. doi: 10.1136/bmj.o562.

[227] Winner of statistical 'trustworthiness, quality value' award announced. Royal Statistical Society, 2022. Available from: `https://rss.org.uk/news-publication/news-publications/2022/general-news/winner-of-statistical-trustworthiness,-quality-val/`.

[228] COVID-19 testing data: methodology note. gov.uk, 2022. Available from: `https://www.gov.uk/government/publications/coronavirus-covid-19-testing-data-methodology/covid-19-testing-data-methodology-note`.

[229] A Cori, NM Ferguson, C Fraser, and S Cauchemez. A New Framework and Software to Estimate Time-Varying Reproduction Numbers During Epidemics. *American Journal of Epidemiology*, 178 (9):1505–1512, 2013.

[230] KV Parag. Improved estimation of time-varying reproduction numbers at low case incidence and between epidemic waves. *PLOS Computational Biology*, 17(9):1–23, 2021. doi: 10.1371/journal.pcbi.1009347.

[231] JA Scott, A Gandy, S Mishra, S Bhatt, S Flaxman, HJT Unwin, et al. Epidemia: An R Package for Semi-Mechanistic Bayesian Modelling of Infectious Diseases using Point Processes. *arXiv*, 2021. doi: 10.48550/arXiv.2110.12461.

[232] Slides to accompany coronavirus press conference: 11 May 2020. gov.uk, 2020. Available from: `https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/884352/slides_-_11_05_2020.pdf`.

[233] Slides and datasets to accompany coronavirus press conference: 22 October 2020. gov.uk, 2020. Available from: `https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/928812/Press_conference_slides_-_Thursday_22_October_.pdf`.

[234] Government publishes latest R number. gov.uk, 2020. Available from: `https://www.gov.uk/government/news/government-publishes-latest-r-number`.

[235] REal-time Assessment of Community Transmission (REACT) Study. Imperial College London, 2021. Available from: `https://www.imperial.ac.uk/medicine/research-and-impact/groups/react-study/`.

[236] CJE Metcalf, J Farrar, FT Cutts, NE Basta, AL Graham, J Lessler, et al. Use of serological surveys to generate key insights into the changing global landscape of infectious disease. *The Lancet*, 388: 728–730, 2016. doi: 10.1016/S0140-6736(16)30164-7.

[237] AJ Siegler, PS Sullivan, T Sanchez, B Lopman, M Fahimi, C Sailey, et al. Protocol for a national probability survey using home specimen collection methods to assess prevalence and incidence of SARS-CoV-2 infection and antibody response. *Annals of Epidemiology*, 49:50–60, 2020. doi: 10.1016/j.annepidem.2020.07.015.

[238] R Subramanian, Q He, and M Pascual. Quantifying asymptomatic infection and transmission of COVID-19 in New York City using observed cases, serology, and testing capacity. *Proceedings of the National Academy of Sciences*, 118(9):e2019716118, 2021. doi: 10.1073/pnas.2019716118.

[239] KD Vihta, KB Pouwels, TEA Peto, E Pritchard, DW Eyre, T House, et al. Symptoms and Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Positivity in the General Population in the United Kingdom. *Clinical Infectious Diseases*, 75(1):e329–e337, 2021. doi: 10.1093/cid/ciab945.

[240] A Frei, M Kaufmann, R Amati, A Butty Dettwiler, V von Wyl, AM Annoni, et al. Development of hybrid immunity during a period of high incidence of Omicron infections. *International Journal of Epidemiology*, page dyad098, 2023. doi: 10.1093/ije/dyad098.

[241] KB Pouwels, T House, E Pritchard, JV Robotham, PJ Birrell, A Gelman, et al. Community prevalence of SARS-CoV-2 in England from April to November, 2020: results from the ONS Coronavirus Infection Survey. *The Lancet Public Health*, 6, 2021. doi: 10.1016/S2468-2667(20)30282-6.

[242] S Rhodes, J Wilkinson, N Pearce, W Mueller, M Cherrie, K Stocking, et al. Occupational differences in SARS-CoV-2 infection: analysis of the UK ONS COVID-19 infection survey. *Journal of Epidemiology & Community Health*, 76(10):841–846, 2022. doi: 10.1136/jech-2022-219101.

[243] D Ayoubkhani, C Bermingham, KB Pouwels, M Glickman, V Nafilyan, F Zaccardi, et al. Trajectory of long COVID symptoms after COVID-19 vaccination: community based cohort study. *BMJ*, 377, 2022. doi: 10.1136/bmj-2021-069676.

[244] S Shabnam, C Razieh, H Dambha-Miller, T Yates, C Gillies, YV Chudasama, et al. Socioeconomic inequalities of Long COVID: a retrospective population-based cohort study in the United Kingdom. *Journal of the Royal Society of Medicine*, 116(8):263–273, 2023. doi: 10.1177/01410768231168377.

[245] T House, H Riley, L Pellis, KB Pouwels, S Bacon, A Eidukas, et al. Inferring risks of coronavirus transmission from community household data. *Statistical Methods in Medical Research*, 31(9): 1738–1756, 2022. doi: 10.1177/09622802211055853.

[246] KA Lythgoe, T Golubchik, M Hall, T House, R Cahuantzi, G MacIntyre-Cockett, et al. Lineage replacement and evolution captured by 3 years of the United Kingdom Coronavirus (COVID-19) Infection Survey. *Proceedings of the Royal Society B: Biological Sciences*, 290(2009):20231284, 2023. doi: 10.1098/rspb.2023.1284.

[247] J Hellewell, TW Russell, R Matthews, A Severn, S Adam, L Enfield, et al. Estimating the effectiveness of routine asymptomatic PCR testing at different frequencies for the detection of SARS-CoV-2 infections. *BMC Medicine*, 19, 2021. doi: 10.1186/s12916-021-01982-x.

[248] AS Iyer, FK Jones, A Nodoushani, M Kelly, M Becker, D Slater, et al. Persistence and decay of human antibody responses to the receptor binding domain of SARS-CoV-2 spike protein in COVID-19 patients. *Science Immunology*, 5(52):eabe0367, 2020. doi: 10.1126/sciimmunol.abe0367.

[249] Y Fei, H Xu, X Zhang, SS Musa, S Zhao, and D He. Seroprevalence and infection attack rate of COVID-19 in Indian cities. *Infectious Disease Modelling*, 7(2):25–32, 2022. doi: 10.1016/j.idm.2022.03.001.

[250] O Byambasuren, CC Dobler, K Bell, DP Rojas, J Clark, ML McLaws, et al. Comparison of seroprevalence of SARS-CoV-2 infections with cumulative and imputed COVID-19 cases: Systematic review. *PLOS ONE*, 16(4):1–14, 2021. doi: 10.1371/journal.pone.0248946.

[251] I Bergeri, MG. Whelan, H Ware, L Subissi, A Nardone, HC Lewis, et al. Global SARS-CoV-2 seroprevalence from January 2020 to April 2022: A systematic review and meta-analysis of standardized population-based studies. *PLOS Medicine*, 19(11):1–24, 2022. doi: 10.1371/journal.pmed.1004107.

[252] Coronavirus (COVID-19) latest insights: Antibodies. Office for National Statistics, 2023. Available from: `https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/articles/coronaviruscovid19latestinsights/antibodies`.

[253] H Soeorg, P Jõgi, P Naaber, A Ottas, K Toompere, and I Lutsar. Seroprevalence and levels of IgG antibodies after COVID-19 infection or vaccination. *Infectious Diseases*, 54(1):63–71, 2022. doi: 10.1080/23744235.2021.1974540.

[254] RK Arora, A Joseph, J Van Wyk, S Rocco, A Atmaja, E May, et al. SeroTracker: a global SARS-CoV-2 seroprevalence dashboard. *The Lancet Infectious Diseases*, 21(4):e75–e76, 2021. doi: 10.1016/S1473-3099(20)30631-9.

[255] S Uyoga, IMO. Adetifa, HK Karanja, J Nyagwange, J Tuju, P Wanjiku, et al. Seroprevalence of anti-SARS-CoV-2 IgG antibodies in Kenyan blood donors. *Science*, 371(6524):79–82, 2021. doi: 10.1126/science.abe1916.

[256] AS Bin-Ghouth, S Al-Shoteri, N Mahmoud, A Musani, NM Baoom, AA Al-Waleedi, et al. SARS-CoV-2 seroprevalence in Aden, Yemen: a population-based study. *International Journal of Infectious Diseases*, 115:239–244, 2022. ISSN 1201-9712. doi: 10.1016/j.ijid.2021.12.330.

[257] KE Wiens, P Mawien, J Rumunu, D Slater, S Jones, A Caflisch, et al. Seroprevalence of Severe Acute Respiratory Syndrome Coronavirus 2 IgG in Juba, South Sudan. 27:1598–1606. doi: 10.3201/eid2706.210568.

[258] HC Lewis, H Ware, M Whelan, L Subissi, Z Li, X Ma, et al. SARS-CoV-2 infection in Africa: a systematic review and meta-analysis of standardised seroprevalence studies, from January 2020 to December 2021. *BMJ Global Health*, 7(8), 2022. doi: 10.1136/bmjgh-2022-008793.

[259] M O'Driscoll, G Riberio Dos Santos, L Wang, DAT Cummings, AS Azman, J Paireau, et al. Age-specific mortality and immunity patterns of SARS-CoV-2. 590, 2021. doi: 10.1038/s41586-020-2918-0.

[260] M Ghafari, OJ Watson, A Karlinsky, L Ferretti, and A Katzourakis. A framework for reconstructing SARS-CoV-2 transmission dynamics using excess mortality data. *Nature Communications*, 13, 2022. doi: 10.1038/s41467-022-30711-y.

[261] Owen Dyer. Covid-19: Peru's official death toll triples to become world's highest. *BMJ*, 373, 2021. doi: 10.1136/bmj.n1442.

[262] RM Anderson and RM May. *Infectious diseases of humans: dynamics and control*. Oxford, United Kingdom: Oxford University Press, 1991.

[263] K Dietz. The estimation of the basic reproduction number for infectious diseases. *Statistical Methods in Medical Research*, 2(1):23–41, 1993.

[264] JA Jacquez and P O'Neill. Reproduction numbers and thresholds in stochastic epidemic models I. Homogeneous populations. *Mathematical Biosciences*, 107(2):161–186, 1991.

[265] P van den Driessche and J Watmough. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Mathematical Biosciences*, 180(1):29–48, 2002.

[266] S Sanche, YT Lin, C Xu, E Romero-Severson, N Hengartner, and R Ke. High Contagiousness and Rapid Spread of Severe Acute Respiratory Syndrome Coronavirus 2. *Emerging Infectious Diseases*, 26(7):1470–1477, 2020.

[267] JT Wu, K Leung, and GM Leung. 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, 2020.

[268] AJ Kucharski, TW Russell, C Diamond, Y Liu, J Edmunds, S Funk, et al. Early dynamics of transmission and control of COVID-19: a mathematical modelling study. *The Lancet Infectious Diseases*, 20(5):553–558, 5 2020.

[269] ES Knock, LK Whittles, JA Lees, PN Perez-Guzman, R Verity, RG FitzJohn, et al. Key epidemiological drivers and impact of interventions in the 2020 SARS-CoV-2 epidemic in England. *Science Translational Medicine*, 13(602):eabg4262, 2021. doi: 10.1126/scitranslmed.abg4262.

[270] S Flaxman, S Mishra, A Gandy, HJT Unwin, TA Mellan, H Coupland, et al. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature*, 584(7820):257–261, 8 2020.

[271] J Wallinga and P Teunis. Different Epidemic Curves for Severe Acute Respiratory Syndrome Reveal Similar Impacts of Control Measures. *American Journal of Epidemiology*, 160(6):509–516, 2004. doi: 10.1093/aje/kwh255.

[272] LMA Bettencourt and RM Ribeiro. Real time Bayesian estimation of the epidemic potential of emerging infectious diseases. *PLoS One*, 3(5), 2008. doi: 10.1371/journal.pone.0002185.

[273] S Riley, C Fraser, CA Donnelly, AC Ghani, LJ Abu-Raddad, AJ Hedley, et al. Transmission dynamics of the etiological agent of SARS in Hong Kong: impact of public health interventions. *Science*, 300:1961–1966, 2003. doi: 10.1126/science.1086478.

[274] EJ Amundsen, H Stigum, JA Røttingen, and OO Aalen. Definition and estimation of an actual reproduction number describing past infectious disease transmission: application to HIV epidemics among homosexual men in Denmark, Norway and Sweden. *Epidemiology and Infection*, 132(6): 1139–1149, 2004. doi: 10.1017/s0950268804002997.

[275] Sam Abbott and Sebastian Funk. Estimating epidemiological quantities from repeated cross-sectional prevalence measurements. *medRxiv*, 2022. doi: 10.1101/2022.03.29.22273101.

[276] HS Vöhringer, T Sanderson, M Sinnott, N De Maio, T Nguyen, R Goater, et al. Genomic reconstruction of the SARS-CoV-2 epidemic in England. *Nature*, 600(7889), 2021. doi: 10.1038/s41586-021-04069-y.

[277] CE Overton, L Pellis, B Stage, F Scarabel, J Burton, C Fraser, et al. EpiBeds: Data informed modelling of the COVID-19 hospital burden in England. *PLOS Computational Biology*, 18(9):1–20, 2022. doi: 10.1371/journal.pcbi.1010406.

[278] RE Moore, C Rosato, and S Maskell. Refining epidemiological forecasts with simple scoring rules. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 380(2233):20210305, 2022. doi: 10.1098/rsta.2021.0305.

[279] M Baguelin, E Knock, L Whittles, R FitzJohn, J Lees, A Cori, et al. sircovid. Imperial College London, 2023. Available from: `https://mrc-ide.github.io/sircovid/`.

[280] CC Kerr, RM Stuart, D Mistry, RG Abeysuriya, K Rosenfeld, GR Hart, et al. Covasim: An agent-based model of COVID-19 dynamics and interventions. *PLOS Computational Biology*, 17(7):1–32, 2021. doi: 10.1371/journal.pcbi.1009149.

[281] R Hinch, WJM Probert, A Nurtay, M Kendall, C Wymant, M Hall, et al. OpenABM-Covid19—An agent-based model for non-pharmaceutical interventions against COVID-19 including contact tracing. *PLOS Computational Biology*, 17(7):1–26, 2021. doi: 10.1371/journal.pcbi.1009146.

[282] MJ Keeling, L Dyson, G Guyver-Fletcher, A Holmes, MG Semple, ISARIC4C Investigators, et al. Fitting to the UK COVID-19 outbreak, short-term forecasts and estimating the reproductive number. *Statistical Methods in Medical Research*, 31(9):1716–1737, 2022. doi: 10.1177/09622802211070257.

[283] P Birrell, J Blake, E van Leeuwen, N Gent, and D De Angelis. Real-time nowcasting and forecasting of COVID-19 dynamics in England: the first wave. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 376(1829):20200279, 2021. doi: 10.1098/rstb.2020.0279.

[284] P Birrell, J Blake, E van Leeuwen, MRC Biostatistics Unit COVID-19 Working Group, and D De Angelis. COVID-19: nowcast and forecast, 2022. Available from: `https://joshuablake.co.uk/public-RTM-reports/iframe.html`.

[285] D Adam. A guide to R — the pandemic's misunderstood metric. *Nature*, 583:346–348, 2020. doi: 10.1038/d41586-020-02009-w.

[286] Prime Minister's statement on coronavirus (COVID-19): 10 May 2020. gov.uk, 2020. Available from: `https://www.gov.uk/government/speeches/pm-address-to-the-nation-on-coronavirus-10-may-2020`.

[287] Prime Minister's statement on coronavirus (COVID-19): 3 July 2020. gov.uk, 2020. Available from: `https://www.gov.uk/government/speeches/prime-ministers-statement-on-coronavirus-covid-19-3-july-2020`.

[288] UK COVID-19 alert level methodology: an overview. gov.uk, 2021. Available from: `https://www.gov.uk/government/publications/uk-covid-19-alert-level-methodology-an-overview/uk-covid-19-alert-level-methodology-an-overview`.

[289] AB Suthar, J Wang, V Seffren, RE Wiegand, S Griffing, and E Zell. Public health impact of covid-19 vaccines in the US: observational study. *BMJ*, 377, 2022. doi: 10.1136/bmj-2021-069317.

[290] N Andrews, J Stowe, F Kirsebom, S Toffa, R Sachdeva, C Gower, et al. Effectiveness of COVID-19 booster vaccines against COVID-19-related symptoms, hospitalization and death in England. *Nature Medicine*, 28, 2022. doi: 10.1038/s41591-022-01699-1.

[291] The RECOVERY Collaborative Group. Dexamethasone in Hospitalized Patients with COVID-19. *New England Journal of Medicine*, 384(8):693–704, 2021. doi: 10.1056/NEJMoa2021436.

[292] NM Linton, T Kobayashi, Y Yang, K Hayashi, AR Akhmetzhanov, S Jung, et al. 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), 2020. doi: 10.3390/jcm9020538.

[293] V Virlogeux, V Fang, M Park, JT Wu, and BJ Cowling. Comparison of incubation period distribution of human infections with MERS-CoV in South Korea and Saudi Arabia. *Scientific Reports*, 6(35389), 2016. doi: 10.1038/srep35839.

[294] T Ganyani, C Kremer, D Chen, A Torneri, C Faes, J Wallinga, et al. Estimating the generation interval for coronavirus disease (COVID-19) based on symptom onset data, March 2020. *Eurosurveillance*, 25(17), 2020. doi: 10.2807/1560-7917.ES.2020.25.17.2000257.

[295] Q Bi, Y Wu, S Mei, C Ye, X Zou, Z Zhang, et al. Epidemiology and transmission of COVID-19 in 391 cases and 1286 of their close contacts in Shenzhen, China: a retrospective cohort study. *The Lancet Infectious Diseases*, 20:911–919, 2020. doi: 10.1016/S1473-3099(20)30287-5.

[296] NG Reich, J Lessler, DAT. Cummings, and R Brookmeyer. Estimating incubation period distributions with coarse data. *Statistics in Medicine*, 28(22):2769–2784, 2009. doi: 10.1002/sim.3659.

[297] BW Turnbull. The empirical distribution function with arbitrarily grouped, censored and truncated data. *Journal of the Royal Statistical Society: Series B (Methodological)*, 38(3):290–295, 1976. doi: 10.1111/j.2517-6161.1976.tb01597.x.

[298] SA Lauer, KH Grantz, Q Bi, FK Jones, Q Zheng, HR Meredith, et al. The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. *Annals of internal medicine*, 172(9):577–582, 2020. doi: 10.7326/M20-0504.

[299] JA Backer, D Klinkenberg, and J Wallinga. Incubation period of 2019 novel coronavirus (2019-nCoV) infections among travellers from Wuhan, China, 20–28 January 2020. *Eurosurveillance*, 25 (5), 2020. doi: 10.2807/1560-7917.ES.2020.25.5.2000062.

[300] JA Quesada, A López-Pineda, VF Gil-Guillén, JM Arriero-Marín, F Gutiérrez, and C Carratala-Munuera. Período de incubación de la COVID-19: revisión sistemática y metaanálisis. *Revista Clínica Española*, 221(2):109–117, 2021. doi: 10.1016/j.rce.2020.08.005.

[301] Y Wu, L Kang, Z Guo, J Liu, M Liu, and W Liang. Incubation Period of COVID-19 Caused by Unique SARS-CoV-2 Strains: A Systematic Review and Meta-analysis. *JAMA Network Open*, 5 (8):e2228008–e2228008, 2022. ISSN 2574-3805. doi: 10.1001/jamanetworkopen.2022.28008.

[302] M Alene, L Yismaw, MA Assemie, DB Kektema, W Gietaneh, and TY Birham. Serial interval and incubation period of COVID-19: a systematic review and meta-analysis. *BMC Infectious Diseases*, 21(1), 2021. doi: 10.1186/s12879-021-05950-x.

[303] S Galmiche, T Cortier, T Charmet, L Scaheffer, O Chény, C von Platen, et al. SARS-CoV-2 incubation period across variants of concern, individual factors, and circumstances of infection in France: a case series analysis from the ComCor study. *The Lancet Microbe*, 4(6), 2023. doi: 10.1016/S2666-5247(23)00005-8.

[304] X Xu, Y Wu, AG Kummer, Y Zhao, Z Hu, Y Wang, et al. Assessing changes in incubation period, serial interval, and generation time of SARS-CoV-2 variants of concern: a systematic review and meta-analysis. *BMC Medicine*, 21, 2023. doi: 10.1186/s12916-023-03070-8.

[305] E Kenah, M Lipsitch, and JM Robins. Generation interval contraction and epidemic data analysis. *Mathematical Biosciences*, 213(1):71–79, 2008. doi: 10.1016/j.mbs.2008.02.007.

[306] MA Johansson, TM Quandelacy, S Kada, PV Prasad, M Steele, JT Brooks, et al. SARS-CoV-2 Transmission From People Without COVID-19 Symptoms. *JAMA Network Open*, 4(1):e2035057–e2035057, 2021. doi: 10.1001/jamanetworkopen.2020.35057.

[307] T Britton and GS Tomba. Estimation in emerging epidemics: biases and remedies. *Journal of the Royal Society Interface*, 16, 2019. doi: 10.1098/rsif.2018.0670.

[308] B Rai, A Shukla, and LK Dwivedi. Estimates of serial interval for COVID-19: A systematic review and meta-analysis. *Clinical Epidemiology and Global Health*, 9:157–161, 2021. doi: 10.1016/j.cegh.2020.08.007.

[309] H Nishiura, NM Linton, and AR Akhmetzhanov. Serial interval of novel coronavirus (COVID-19) infections. *International Journal of Infectious Diseases*, 93:284–286, 2020. doi: 10.1016/j.ijid.2020.02.060.

[310] M Ki and Task force for 2019-nCov. Epidemiologic characteristics of early cases with 2019 novel coronavirus (2019-nCoV) disease in Korea. *Epidemiology and Health*, 42:e202007, 2020. doi: 10.4178/epih.e2020007.

[311] S Zhao, D Gao, Z Zhuang, MKC Chong, Y Cai, J Ran, et al. Estimating the Serial Interval of the Novel Coronavirus Disease (COVID-19): A Statistical Analysis Using the Public Data in Hong Kong From January 16 to February 15, 2020. *Frontiers in Physics*, 8:347, 2020. doi: 10.3389/fphy.2020.00347.

[312] M Li, K Liu, Y Song, M Wang, and J Wu. Serial Interval and Generation Interval for Imported and Local Infectors, Respectively, Estimated Using Reported Contact-Tracing Data of COVID-19 in China. *Frontiers in Public Health*, 8:942, 2021. doi: 10.3389/fpubh.2020.577431.

[313] WS Hart, S Abbott, A Endo, J Hellewell, E Miller, N Andrews, et al. Inference of the SARS-CoV-2 generation time using UK household data. *eLife*, 11:e70767, 2022. doi: 10.7554/eLife.70767.

[314] S Abbott, K Sherratt, M Gerstung, and S Funk. Estimation of the test to test distribution as a proxy for generation interval distribution for the omicron variant in england. *medRxiv*, 2022. doi: 10.1101/2022.01.08.22268920.

[315] ST Ali, L Wang, EHY Lau, X Xu, Z Du, Y Wu, et al. Serial interval of SARS-CoV-2 was shortened over time by nonpharmaceutical interventions. *Science*, 369(6507):1106–1109, 2020. doi: 10.1126/science.abc9004.

[316] NG Reich, LC Brooks, SJ Fox, S Kandula, CJ McGowan, E Moore, et al. A collaborative multiyear, multimodel assessment of seasonal influenza forecasting in the United States. *Proceedings of the National Academy of Sciences*, 116(8):3146–3154, 2019. doi: 10.1073/pnas.1812594116.

[317] ME Halloran. Concept of Infectious Disease Epidemiology. In KJ Rotham and S Greenland, editors, *Modern Epidemiology*, chapter 10. Lippincott-Raven, Philadelphia, 1998.

[318] WO Kermack and AG McKendrick. Contributions to the mathematical theory of epidemics—I. *Bulletin of Mathematical Biology*, 53(1):33–55, 1991.

[319] MI Meltzer, CY Atkins, S Santibanez, B Knust, BW Petersen, ED Ervin, et al. Estimating the future number of cases in the Ebola epidemic – Liberia and Sierra Leone, 2014–2015. Center for Disease Prevention and Control, 2014. Available from: `https://stacks.cdc.gov/view/cdc/24901`.

[320] NM Ferguson, DAT Cummings, S Cauchemez, C Fraser, S Riley, A Meeyai, et al. Strategies for containing an emerging influenza pandemic in Southeast Asia. *Nature*, 437:209–214, 2005. doi: 10.1038/nature04017.

[321] HM Topazian, N Schmit, I Gerard-Ursin, GD Charles, H Thompson, AC Ghani, et al. Modelling the relative cost-effectiveness of the RTS,S/AS01 malaria vaccine compared to investment in vector control or chemoprophylaxis. *Vaccine*, 41(20):3215–3223, 2023. doi: 10.1016/j.vaccine.2023.04.011.

[322] NM Ferguson, ZM Cucunubá, I Dorigatti, GL Nedjati-Gilani, CA Donnelly, M-G Basáñez, et al. Countering the Zika epidemic in Latin America. *Science*, 353(6297):353–354, 2016. doi: 10.1126/science.aag0219.

[323] C Fraser, S Riley, RM Anderson, and NM Ferguson. Factors that make an infectious disease outbreak controllable. *Proceedings of the National Academy of Sciences*, 101(16):6146–6151, 2004. doi: 10.1073/pnas.0307506101.

[324] SB Kim, M Yoon, NS Ku, MH Kim, JE Song, JY Ahn, et al. Mathematical Modeling of HIV Prevention Measures Including Pre-Exposure Prophylaxis on HIV Incidence in South Korea. *PLOS ONE*, 9(3):1–9, 2014. doi: 10.1371/journal.pone.0090080.

[325] GF Medley. A consensus of evidence: The role of SPI-M-O in the UK COVID-19 response. *Advances in Biological Regulation*, 86:100918, 2022. doi: 10.1016/j.jbior.2022.100918. COVID-19 models and expectations.

[326] AB Hogan, BL Jewell, E Sherrard-Smith, JF Vesga, OJ Watson, C Whittaker, et al. Potential impact of the COVID-19 pandemic on HIV, tuberculosis, and malaria in low-income and middle-income countries: a modelling study. *The Lancet Global Health*, 8(9):e1132–e1141, 9 2020. doi: 10.1016/S2214-109X(20)30288-6.

[327] JS Weitz, SJ Beckett, AR Coenen, D Demory, M Dominguez-Mirazo, J Dushoff, et al. Modeling shield immunity to reduce COVID-19 epidemic spread. *Nature Medicine*, 26(6):849–854, 6 2020. doi: 10.1038/s41591-020-0895-3.

[328] J Rocklöv, H Sjödin, and A Wilder-Smith. COVID-19 outbreak on the Diamond Princess cruise ship: estimating the epidemic potential and effectiveness of public health countermeasures. *Journal of Travel Medicine*, 27(3), 2020. doi: 10.1093/jtm/taaa030.

[329] T Oraby, MG Tyshenko, HH Balkhy, Y Tasnif, A Quiroz-Gaspar, Z Mohamed, et al. Analysis of the Healthcare MERS-CoV Outbreak in King Abdulaziz Medical Center, Riyadh, Saudi Arabia, June–August 2015 Using a SEIR Ward Transmission Model. *International Journal of Environmental Research and Public Health*, 17(8), 2020. doi: 10.3390/ijerph17082936.

[330] R Li, S Pei, B Chen, Y Song, T Zhang, W Yang, et al. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). *Science*, 368(6490):489–493, 2020. doi: 10.1126/science.abb3221.

[331] G Chowell, S Blumberg, L Simonsen, MA Miller, and C Viboud. Synthesizing data and models for the spread of MERS-CoV, 2013: Key role of index cases and hospital transmission. *Epidemics*, 9: 40–51, 2014. doi: 10.1016/j.epidem.2014.09.011.

[332] I Hall, H Lewkowicz, L Webb, T House, L Pellis, J Sedgwick, et al. Outbreaks in care homes may lead to substantial disease burden if not mitigated. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 376(1829):20200269, 7 2021. doi: 10.1098/rstb.2020.0269.

[333] NG Davies, P Klepac, Y Liu, K Prem, M Jit, and RM Eggo. Age-dependent effects in the transmission and control of COVID-19 epidemics. *Nature Medicine*, 26(8):1205–1211, 8 2020. doi: 10.1038/s41591-020-0962-9.

[334] K Prem, Y Liu, TW Russell, AJ Kucharski, RM Eggo, N Davies, et al. The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: a modelling study. *The Lancet Public health*, 5(5):e261–e270, 2020. doi: 10.1016/S2468-2667(20)30073-6.

[335] J Zhang, M Litvinova, Y Liang, Y Wang, W Wang, S Zhao, et al. Changes in contact patterns shape the dynamics of the COVID-19 outbreak in China. *Science*, 368(6498):1481–1486, 6 2020. doi: 10.1126/science.abb8001.

[336] M Chinazzi, JT Davis, M Ajelli, C Gioannini, M Litvinova, S Merler, et al. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science*, 368(6489):395–400, 2020. doi: 10.1126/science.aba9757.

[337] WS Sun and RISEWIDs Team. Individual-based simulation model for COVID-19 transmission in Daegu, Korea. *Epidemiology and Health*, 42(0):e2020042–0, 2020. doi: 10.4178/epih.e2020042.

[338] M Roberts, V Andreasen, A Lloyd, and L Pellis. Nine challenges for deterministic epidemic models. *Epidemics*, 10:49–53, 2015. doi: 10.1016/j.epidem.2014.09.006.

[339] H Tian, Y Liu, Y Li, C-H Wu, B Chen, MUG Kraemer, et al. An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. *Science*, 368(6491):638–642, 5 2020. doi: 10.1126/science.abb6105.

[340] A Aleta, D Martín-Corral, A Pastore Y Piontti, M Ajelli, M Litvinova, M Chinazzi, et al. Modelling the impact of testing, contact tracing and household quarantine on second waves of COVID-19. *Nature Human Behaviour*, 4(9):964–971, 2020. doi: 10.1038/s41562-020-0931-9.

[341] G Sebastiani, M Massa, and E Riboli. COVID-19 epidemic in Italy: evolution, projections and impact of government measures. *European Journal of Epidemiology*, 35(4):341–345, 2020. doi: 10.1007/s10654-020-00631-6.

[342] A Endo, E van Leeuwen, and M Baguelin. Introduction to particle Markov-chain Monte Carlo for disease dynamics modellers. *Epidemics*, 29:100363, 2019. doi: 10.1016/j.epidem.2019.100363.

[343] LF Shampine and MW Reichelt. The MATLAB ODE Suite. *SIAM Journal on Scientific Computing*, 18(1):1–22, 1997. doi: 10.1137/S1064827594276424.

[344] K Soetaert, T Petzoldt, and RW Setzer. Solving Differential Equations in R: Package deSolve. *Journal of Statistical Software, Articles*, 33(9):1–25, 2010. doi: 10.18637/jss.v033.i09.

[345] TM Chen, J Rui, Q-P Wang, Z-Y Zhao, J-A Cui, and L Yin. A mathematical model for simulating the phase-based transmissibility of a novel coronavirus. *Infectious Diseases of Poverty*, 9(1):24, 2020. doi: 10.1186/s40249-020-00640-3.

[346] B Ivorra, MR Ferrández, M Vela-Pérez, and AM Ramos. Mathematical modeling of the spread of the coronavirus disease 2019 (COVID-19) taking into account the undetected infections. The case of China. *Communications in Nonlinear Science & Numerical Simulation*, 88:105303, 2020. doi: 10.1016/j.cnsns.2020.105303.

[347] DT Gillespie. Stochastic simulation of chemical kinetics. *Annual Review of Physical Chemistry*, 58: 35–55, 2007. doi: 10.1146/annurev.physchem.58.032806.104637.

[348] EA Dankwa, AF Brouwer, and CA Donnelly. Structural identifiability of compartmental models for infectious disease transmission is influenced by data type. *Epidemics*, 41:100643, 2022. ISSN 1755-4365. doi: 10.1016/j.epidem.2022.100643.

[349] WC Roda, MB Varughese, D Han, and MY Li. Why is it difficult to accurately predict the COVID-19 epidemic? *Infectious Disease Modelling*, 5:271–281, 2020.

[350] K Roosa and G Chowell. Assessing parameter identifiability in compartmental dynamic models using a computational approach: application to infectious disease transmission models. *Theoretical Biology and Medical Modelling*, 16, 2019. doi: 10.1186/s12976-018-0097-6.

[351] E Gecili, A Ziady, and RD Szczesniak. Forecasting COVID-19 confirmed cases, deaths and recoveries: Revisiting established time series modeling through novel applications for the USA and Italy. *PloS One*, 16(1):1–11, 2021. doi: 10.1371/journal.pone.0244173.

[352] OB Da'ar and AE Ahmed. Underlying trend, seasonality, prediction, forecasting and the contribution of risk factors: an analysis of globally reported cases of Middle East Respiratory Syndrome Coronavirus. *Epidemiology and Infection*, 146(11):1343–1349, 2018. doi: 10.1017/S0950268818001541.

[353] G Chowell, A Tariq, and JM Hyman. A novel sub-epidemic modeling framework for short-term forecasting epidemic waves. *BMC Medicine*, 17(164), 2019. doi: 10.1186/s12916-019-1406-6.

[354] K Roosa, Y Lee, R Luo, A Kirpich, R Rothenberg, JM Hyman, et al. Real-time forecasts of the COVID-19 epidemic in China from February 5th to February 24th, 2020. *Infectious Disease Modelling*, 5:256–263, 2020. doi: 10.1016/j.idm.2020.02.002.

[355] CP Farrington, MN Kanaan, and NJ Gay. Branching process models for surveillance of infectious diseases controlled by mass vaccination. *Biostatistics*, 4(2), 2003. doi: 10.1093/biostatistics/4.2.279.

[356] S Abbott, J Hellewell, J Munday, CMMID nCoV working group, and S Funk. The transmissibility of novel Coronavirus in the early stages of the 2019-20 outbreak in Wuhan: Exploring initial point-source exposure sizes and durations using scenario analysis. *Wellcome Open Research*, 5(17), 2020.

[357] Coronavirus disease (COVID-19) pandemic. World Health Organization, 2021. Available from: `https://www.who.int/emergencies/diseases/novel-coronavirus-2019`.

[358] List of participants of SAGE and related sub-groups. gov.uk, 2021. Available from: `https://tinyurl.com/2bk76k3a`.

[359] Scientific evidence supporting the government response to coronavirus (COVID-19). gov.uk, 2021. Available from: `https://www.gov.uk/government/collections/scientific-evidence-supporting-the-government-response-to-coronavirus-covid-19#meeting-minutes-and-supporting-papers`.

[360] Scottish Government COVID-19 Advisory Group. Scottish Government, 2021. Available from: `https://www.gov.scot/groups/scottish-government-covid-19-advisory-group/`.

[361] SM Kissler, C Tedijanto, E Goldstein, YH Grad, and M Lipsitch. Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. *Science*, 368(6493):860–868, 5 2020. doi: 10.1126/science.abb5793.

[362] A Atkeson. What Will Be the Economic Impact of COVID-19 in the US? Rough Estimates of Disease Scenarios. National Bureau of Economic Research, 2020. Available from: `https://www.nber.org/papers/w26867`.

[363] JR Koo, AR Cook, M Park, Y Sun, H Sun, JT Lim, et al. Interventions to mitigate early spread of SARS-CoV-2 in Singapore: a modelling study. *The Lancet Infectious Diseases*, 20(6):678–688, 6 2020. doi: 10.1016/S1473-3099(20)30162-6.

[364] K Leung, JT Wu, D Liu, and GM Leung. First-wave COVID-19 transmissibility and severity in China outside Hubei after control measures, and second-wave scenario planning: a modelling impact assessment. *The Lancet*, 395(10233):1382–1393, 4 2020. doi: 10.1016/S0140-6736(20)30746-7.

[365] PGT Walker, C Whittaker, OJ Watson, M Baguelin, KEC Ainslie, S Bhatia, et al. Report 12 - The global impact of COVID-19 and strategies for mitigation and suppression. 2020. Available from: `https://doi.org/10.25561/77735`.

[366] S Hsiang, D Allen, S Annan-Phan, K Bell, I Bolliger, T Chong, et al. The effect of large-scale anti-contagion policies on the COVID-19 pandemic. *Nature*, 584(7820):262–267, 8 2020. doi: 10.1038/s41586-020-2404-8.

[367] M Gatto, E Bertuzzo, L Mari, S Miccoli, L Carraro, R Casagrandi, et al. Spread and dynamics of the COVID-19 epidemic in Italy: Effects of emergency containment measures. *Proceedings of the National Academy of Sciences of the United States of America*, 117(19):10484–10491, 5 2020. doi: 10.1073/pnas.2004978117.

[368] SE Eikenberry, M Mancuso, E Iboi, T Phan, K Eikenberry, Y Kuang, et al. To mask or not to mask: Modeling the potential for face mask use by the general public to curtail the COVID-19 pandemic. *Infectious Disease Modelling*, 5:293–308, 2020. doi: 10.1016/j.idm.2020.04.001.

[369] H Wang, Z Wang, Y Dong, R Chang, C Xu, X Yu, et al. Phase-adjusted estimation of the number of Coronavirus Disease 2019 cases in Wuhan, China. *Cell Discovery*, 6:10, 2020. doi: 10.1038/s41421-020-0148-0.

[370] Y Fang, Y Nie, and M Penny. Transmission dynamics of the COVID-19 outbreak and effectiveness of government interventions: A data-driven analysis. *Journal of Medical Virology*, 92(6):645–659, 6 2020. doi: 10.1002/jmv.25750.

[371] SM Moghadas, A Shoukat, MC Fitzpatrick, CR Wells, P Sah, A Pandey, et al. Projecting hospital utilization during the COVID-19 outbreaks in the United States. *Proceedings of the National Academy of Sciences of the United States of America*, 117(16):9122–9126, 4 2020. doi: 10.1073/pnas.2004064117.

[372] R Singh and R Adhikari. Age-structured impact of social distancing on the COVID-19 epidemic in India. *arXiv*, 2020. Available from: `https://doi.org/10.48550/arXiv.2003.12055`.

[373] P Boldog, T Tekeli, Z Vizi, A Dénes, FA Bartha, and G Röst. Risk Assessment of Novel Coronavirus COVID-19 Outbreaks Outside China. *Journal of Clinical Medicine*, 9(2), 2 2020. doi: 10.3390/jcm9020571.

[374] RA Neher, R Dyrdak, V Druelle, EB Hodcroft, and J Albert. Potential impact of seasonal forcing on a SARS-CoV-2 pandemic. *Swiss Medical Weekly*, 150:w20224, 3 2020. doi: 10.4414/smw.2020.20224.

[375] CN Ngonghala, E Iboi, S Eikenberry, M Scotch, CR MacIntyre, MH Bonds, et al. Mathematical assessment of the impact of non-pharmaceutical interventions on curtailing the 2019 novel Coronavirus. *Mathematical Biosciences*, 325:108364, 7 2020. doi: 10.1016/j.mbs.2020.108364.

[376] H Sjödin, A Wilder-Smith, S Osman, Z Farooq, and J Rocklöv. Only strict quarantine measures can curb the coronavirus disease (COVID-19) outbreak in Italy, 2020. *Eurosurveillance*, 25(13), 2020. doi: 10.2807/1560-7917.ES.2020.25.13.2000280.

[377] CY Yang and J Wang. A mathematical model for the novel coronavirus epidemic in Wuhan, China. *Mathematical Biosciences and Engineering*, 17(3):2708–2724, 3 2020. doi: 10.3934/mbe.2020148.

[378] F Ndaïrou, I Area, JJ Nieto, and DFM Torres. Mathematical modeling of COVID-19 transmission dynamics with a case study of Wuhan. *Chaos, Solitons, and Fractals*, 135:109846, 6 2020. doi: 10.1016/j.chaos.2020.109846.

[379] K Chatterjee, K Chatterjee, A Kumar, and S Shankar. Healthcare impact of COVID-19 epidemic in India: A stochastic mathematical model. *Medical Journal, Armed Forces India*, 76(2):147–155, 4 2020. doi: 10.1016/j.mjafi.2020.03.022.

[380] AA Toda. Susceptible-Infected-Recovered (SIR) Dynamics of COVID-19 and Economic Impact. *arXiv*, 2020. Available from: `https://arxiv.org/abs/2003.11221`.

[381] S Choi and M Ki. Estimating the reproductive number and the outbreak size of COVID-19 in Korea. *Epidemiology and health*, 42:e2020011, 2020. doi: 10.4178/epih.e2020011.

[382] GE Weissman, A Crane-Droesch, C Chivers, T Luong, A Hanish, MZ Levy, et al. Locally Informed Simulation to Predict Hospital Capacity Needs During the COVID-19 Pandemic. *Annals of Internal Medicine*, 173(1):21–28, 7 2020. doi: 10.7326/M20-1260.

[383] AD Paltiel, A Zheng, and RP Walensky. Assessment of SARS-CoV-2 Screening Strategies to Permit the Safe Reopening of College Campuses in the United States. *JAMA Network Open*, 3(7):e2016818, 7 2020. doi: 10.1001/jamanetworkopen.2020.16818.

[384] DB Larremore, B Wilder, E Lester, S Shehata, JM Burke, JA Hay, et al. Test sensitivity is secondary to frequency and turnaround time for COVID-19 surveillance. *Science Advances*, 7(1), 2021. doi: 10.1126/sciadv.abd5393.

[385] P Shi, Y Dong, H Yan, C Zhao, X Li, W Liu, et al. Impact of temperature on the dynamics of the COVID-19 outbreak in China. *Science of The Total Environment*, 728:138890, 2020. doi: 10.1016/j.scitotenv.2020.138890.

[386] J Fernández-Villaverde and CI Jones. Estimating and simulating a SIRD model of COVID-19 for many countries, states and cities. National Bureau of Economic Research, 2020. Available from: `https://www.nber.org/papers/w27128`.

[387] AL Bertozzi, E Franco, G Mohler, MB Short, and D Sledge. The challenges of modeling and forecasting the spread of COVID-19. *Proceedings of the National Academy of Sciences of the United States of America*, 117(29):16732–16738, 7 2020. doi: 10.1073/pnas.2006520117.

[388] YC Chen, PE Lu, CS Chang, and TH Liu. A Time-Dependent SIR Model for COVID-19 With Undetectable Infected Persons. *IEEE Transactions on Network Science and Engineering*, 7(4): 3279–3294, 10 2020. doi: 10.1109/TNSE.2020.3024723.

[389] L López and X Rodó. The end of social confinement and COVID-19 re-emergence risk. *Nature Human Behaviour*, 4(7):746–755, 7 2020. doi: 10.1038/s41562-020-0908-8.

[390] A Atkeson. How Deadly Is COVID-19? Understanding The Difficulties With Estimation Of Its Fatality Rate. National Bureau of Economic Research, 2020. Available from: `https://www.nber.org/papers/w26965`.

[391] J Jia, J Ding, S Liu, G Liao, J Li, B Duan, et al. Modeling the Control of COVID-19: Impact of Policy Interventions and Meteorological Factors. *Electronic Journal of Differential Equations*, 2020 (3):1–24, 2020.

[392] J Wangping, H Ke, S Yang, C Wenzhe, W Shengshu, Y Shanshan, et al. Extended SIR Prediction of the Epidemics Trend of COVID-19 in Italy and Compared With Hunan, China. *Frontiers in Medicine*, 7:169, 2020. doi: 10.3389/fmed.2020.00169.

[393] Y Li, B Wang, R Peng, C Zhou, Y Zhan, Z Liu, et al. Mathematical Modeling and Epidemic Prediction of COVID-19 and Its Significance to Epidemic Prevention and Control Measures. *Annals of Infectious Disease and Epidemiology*, 5(1), 2020. Available from: `http://tinyurl.com/mrau6pdr`.

[394] Cases in United Kingdom. gov.uk, 2021. Available from: `https://coronavirus.data.gov.uk/details/cases`.

[395] Deaths in United Kingdom. gov.uk, 2021. Available from: `https://coronavirus.data.gov.uk/details/deaths`.

[396] Coronavirus and the impact on output in the UK economy: December 2020. Office for National Statistics, 2021. Available from: `https://www.ons.gov.uk/economy/grossdomesticproductgdp/articles/coronavirusandtheimpactonout\putintheukeconomy/december2020`.

[397] Employment in the UK: January 2021. Office for National Statistics, 2021. Available from: `https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/\bulletins/employmentintheuk/january2021`.

[398] Coronavirus Legislation. gov.uk, 2020. Available from: `https://www.legislation.gov.uk/coronavirus`.

[399] Scientific Advisory Group for Emergencies: About us. gov.uk, 2021. Available from: `https://www.gov.uk/government/organisations/scientific-advisory-group-for-emergencies/about`.

[400] SPI-M-O: COVID-19 combined pilot model predictions, 30 March 2020. gov.uk, 2020. Available from: `https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data\/file/917243/SAGE22_20200401_SPI-M_Nowcast_S0093_for_release.pdf`.

[401] Prime Minister's statement on coronavirus (COVID-19): 23 March 2020. gov.uk, 2020. Available from: `https://www.gov.uk/government/speeches/pm-address-to-the-nation-on-coronavirus-23-march-2020`.

[402] JM Read, JRE Bridgen, DAT Cummings, A Ho, and CP Jewell. Novel coronavirus 2019-nCoV (COVID-19): early estimation of epidemiological parameters and epidemic size estimates. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 376(1829):20200265, 7 2021. doi: 10.1098/rstb.2020.0265.

[403] T Crellen, L Pi, EL Davis, TM Pollington, TCD Lucas, D Ayabina, et al. Dynamics of SARS-CoV-2 with waning immunity in the UK population. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 376(1829):20200274, 7 2021. doi: 10.1098/rstb.2020.0274.

[404] The Science Behind COVID-19: Interpreting Epidemiological Trends and Models. Swiss Tropical and Public Health Institute Virtual Event Series, 2021. Available from: `https://www.swisstph.ch/en/about/events/ves/`.

[405] Chief Scientific Adviser's Sunday Telegraph article: 31 May 2020. Government Office for Science, 2020. Available from: `https://www.gov.uk/government/speeches/chief-scientific-advisers-sunday-telegraph-article-31-may-2020`.

[406] K Allen, T Buklijas, A Chen, N Simon-Kumar, L Cowen, J Wilsdon, et al. Tracking global evidence-to-policy pathways in the coronavirus crisis: A preliminary report. International Network for Government Scientific Advice, 2020. Available from: `https://www.ingsa.org/wp-content/uploads/2020/09/INGSA-Evidence-to-Policy-Tracker_Report-1_FINAL_17Sept.pdf`.

[407] A Saltelli, G Bammer, I Bruno, E Charters, M Di Fiore, E Didier, et al. Five ways to ensure that models serve society: a manifesto. *Nature*, 582:482 – 484, 2020. doi: 10.1038/d41586-020-01812-9.

[408] JL Juul, K Græsbøll, LE Christiansen, and S Lehmann. Fixed-time descriptive statistics underestimate extremes of epidemic curve ensembles. *Nature Physics*, 17(1):5–8, 2021. doi: 10.1038/s41567-020-01121-y.

[409] C Craig. *How Does Government Listen to Scientists?* Palgrave Macmillan, Cham, 2019. ISBN 978-3-319-96086-9.

[410] Methodological challenges and approaches in COVID-19 epidemiology. The JBC-Turing-RSS Laboratory., 2021. Available from: `https://www.turing.ac.uk/events/methodological-challenges-and-approaches-covid-19-epidemiology`.

[411] CM Barton, M Alberti, D Ames, J-A Atkinson, J Bales, E Burke, et al. Call for transparency of COVID-19 models. 368(6490):482–483, 5 2020. doi: 10.1126/science.abb8637.

[412] Off the Ball. BBC Scotland, 2021. Available from: `https://www.bbc.co.uk/programmes/b0079mcc`.

[413] A Stevens. Governments cannot just 'follow the science' on COVID-19. *Nature Human Behaviour*, 4(6):560, 2020. doi: 10.1038/s41562-020-0894-x.

[414] K Abbasi. Covid-19: politicisation, "corruption," and suppression of science. *BMJ*, 371:m4425, 2020. doi: 10.1136/bmj.m4425.

[415] R Pielke. Five Modes of Science Engagement, 2015. Available from: `http://rogerpielkejr.blogspot.com/2015/01/five-modes-of-science-engagement.html`.

[416] Addendum to thirteenth SAGE meeting on Covid-19, 5th March 2020, 2020. Available from: `https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file\/888781/S0381_Thirteenth_meeting_on_Wuhan_Coronavirus__Covid-19__.pdf`.

[417] F Lawrence, S Carrell, and D Pegg. Attendees of Sage meetings worried by presence of Cummings. The Guardian, 2020. Available from: `https://www.theguardian.com/world/2020/apr/26/attendees-of-sage-coronavirus-meetings-worried-by-presence-of-dominic-cummings`.

[418] Fifty-eighth SAGE meeting on Covid-19, 21st September 2020. Technical report, 2020. Available from: `https://assets.publishing.service.gov.uk/media/5f8432e18fa8f5045d71574a/S0768_Fifty-eighth_SAGE_meeting_on_Covid-19.pdf`.

[419] L Maxim and J Van der Sluijs. Quality of epidemiological studies: Procedural rules for uncertain science for policy, a case study on bisphenol-A. *Environmental Science & Policy*, 84:80–87, 2018. doi: 10.1016/j.envsci.2018.03.010.

[420] Professor Gabriel M Leung. Hong Kong University, 2021. Available from: `http://www.med.hku.hk/en/About-HKUMed/Leadership/Professor-Gabriel-M-Leung`.

[421] R Cookson. Lockdown 1.0 - Following the science? BBC Two, 2020. Available from: `https://www.bbc.co.uk/iplayer/episode/m000pjr1/lockdown-10-following-the-science`.

[422] Fiona Fox, Tom Feilden & Prof Sir Charles Godfray in conversation: "Science communication in times of emergency". Oxford Martin School, University of Oxford, 2021. Available from: `https://www.oxfordmartin.ox.ac.uk/videos/science-communication/`.

[423] J Zelner, J Riou, R Etzioni, and A Gelman. Accounting for uncertainty during a pandemic. *Patterns*, 2(8):100310, 2021. doi: 10.1016/j.patter.2021.100310.

[424] K Shea, MC Runge, D Pannell, WJM Probert, S Li, M Tildesley, et al. Harnessing multiple models for outbreak management. *Science*, 368(6491):577–579, 2020. doi: 10.1126/science.abb9934.

[425] J Cheshire. Next slide please: data visualisation expert on what's wrong with the UK government's coronavirus charts. The Conversation, 2020. Available from: `https://tinyurl.com/3h6pt78m`.

[426] COVID-19 Scenario Analysis Tool. MRC Centre for Global Infectious Disease Analysis, Imperial College London, 2021. Available from: `www.covidsim.org`.

[427] P Fox and J Hendler. Changing the Equation on Scientific Data Visualization. *Science*, 331(6018): 705–708, 2011. doi: 10.1126/science.1197654.

[428] R McCabe and CA Donnelly. ruthmccabe/public-survey-modelling, 2023. Available from: `https://zenodo.org/records/10201376`.

[429] Technical report on the COVID-19 pandemic in the UK. Department of Health and Social Care, 2022. Available from: `https://www.gov.uk/government/publications/technical-report-on-the-covid-19-pandemic-in-the-uk`.

[430] C Pagel and CA Yates. Role of mathematical modelling in future pandemic response policy. *BMJ*, 378, 2022. doi: 10.1136/bmj-2022-070615.

[431] MH Kalayou and SM Awol. Myth and Misinformation on COVID-19 Vaccine: The Possible Impact on Vaccination Refusal Among People of Northeast Ethiopia: A Community-Based Research. *Risk Management and Healthcare Policy*, 15:1859–1868, 2022. doi: 10.2147/RMHP.S366730.

[432] T Santavicca, C Ngov, R Frounfelker, D Miconi, A Levinsson, and C Rousseau. COVID-19 vaccine hesitancy among young adults in Canada. *Canadian Journal of Public Health*, 114, 2023. doi: 10.17269/s41997-022-00693-x.

[433] RR Marzo, R Shrestha, B Sapkota, S Acharya, N Shrestha, M Pokharel, et al. Perception towards vaccine effectiveness in controlling COVID-19 spread in rural and urban communities: A global survey. *Frontiers in Public Health*, 10, 2022. ISSN 2296-2565. doi: 10.3389/fpubh.2022.958668.

[434] A Elia, DJ Negara, Ferdinand, S Neneng, Meitiana, T Anden, et al. State versus action orientation and compliance during the COVID-19 pandemic in Indonesia. *Heliyon*, 8(10):e10725, 2022. ISSN 2405-8440. doi: 10.1016/j.heliyon.2022.e10725.

[435] A Al Mamun, N Hayat, HTM Dieu, NR Zainol, and AA Salameh. COVID-19 preventive behavior among university students in Southeast Asia: Effects of knowledge, concern, awareness, and perceived risk. *Frontiers in Public Health*, 10, 2022. doi: 10.3389/fpubh.2022.958021.

[436] EY Roga, GG Bekele, and DN Gonfa. Compliance level toward COVID-19 preventive measures and associated factors among the Ambo University community, 2021. *Frontiers in Public Health*, 10, 2022. doi: 10.3389/fpubh.2022.958270.

[437] M Chambon, J Dalege, D Borsboom, LJ Waldorp, HLJ van der Maas, and F van Harreveld. How compliance with behavioural measures during the initial phase of a pandemic develops over time: A longitudinal COVID-19 study. *British Journal of Social Psychology*, 62(1):302–321, 2023. doi: 10.1111/bjso.12572.

[438] GC Marshall, R Skeva, C Jay, MEP Silva, M Fyles, T House, EL Davis, L Pi, GF Medley, BJ Quilty, L Dyson, L Yardley, and E Fearon. Public perceptions and interactions with UK COVID-19 Test, Trace and Isolate policies, and implications for pandemic infectious disease modelling [version 1; peer review: awaiting peer review]. *F1000Research*, 11(1005), 2022. doi: 10.12688/f1000research.124627.1.

[439] Prolific. Prolific, 2023. `https://www.prolific.co/`.

[440] About. Twitter, 2023. `https://about.twitter.com/en`.

[441] Prolific's participants. Prolific, 2023. `https://tinyurl.com/3u2a495w`.

[442] Representative samples. Prolific, 2023. `https://researcher-help.prolific.com/hc/en-gb/articles/360019236753-Representative-samples`.

[443] Letter from Lord Bethell to Dr Jenny Harries, UKHSA chief executive. gov.uk, 2021. Available from: `https://www.gov.uk/government/publications/ukhsa-priorities-in-2021-to-2022/letter-from-lord-bethell-to-dr-jenny-harries-ukhsa-chief-executive`.

[444] VC Bradley, S Kuriwaki, M Isakov, D Sejdinovic, and S Meng, X Flaxman. Unrepresentative big surveys significantly overestimated US vaccine uptake. *Nature*, 600:695 – 700, 2021. doi: 10.1038/s41586-021-04198-4.

[445] Vaccinations in the United Kingdom. gov.uk, 2023. Available from: `https://coronavirus.data.gov.uk/details/vaccinations`.

[446] JCVI statement on COVID-19 vaccination of children and young people aged 12 to 17 years: 15 July 2021. gov.uk, 2021. Available from: `https://tinyurl.com/2kzxa6bx`.

[447] D Adam. Special report: The simulations driving the world's response to COVID-19. *Nature*, 580: 316 – 318, 2020. doi: 10.1038/d41586-020-01003-6.

[448] L Koch. Scientific modelling is steering our response to coronavirus. But what is scientific modelling? The Conversation, 2020. Available from: `https://tinyurl.com/3hsnynre`.

[449] I Sample. Coronavirus exposes the problems and pitfalls of modelling. The Guardian, 2020. Available from: `https://www.theguardian.com/science/2020/mar/25/coronavirus-exposes-the-problems-and-pitfalls-of-modelling`.

[450] L Andrews and E Craig. It's easier to believe in Father Christmas than SAGE's 6,000 Covid deaths-a-day projection: Experts say doomsday scenario is 'fiction' as chief modeller admits they don't do optimistic estimates because 'that doesn't get decisions made'. The Daily Mail, 2021. Available from: `https://www.dailymail.co.uk/news/article-10328333/Why-easier-believe-Father-Christmas-SAGEs-6k-Covid-deaths-day-projection.html`.

[451] About us. Science Media Centre, 2023. Available from: `https://www.sciencemediacentre.org/`.

[452] UK COVID-19 Inquiry 21 November 2023 (Testimony of Professor Sir Chris Whitty). UK Covid-19 Inquiry, 2023. Available from: `https://covid19.public-inquiry.uk/wp-content/uploads/2023/11/21174834/C-19-Inquiry-21-November-2023-Module-2-Day-23.pdf`.

[453] Moving to step 4 of the roadmap. gov.uk, 2021. Available from: `https://www.gov.uk/government/publications/covid-19-response-summer-2021-roadmap/moving-to-step-4-of-the-roadmap`.

[454] D Buitrago-Garcia, D Egli-Gany, MJ Counotte, S Hossmann, H Imeri, AM Ipekci, et al. Occurrence and transmission potential of asymptomatic and presymptomatic SARS-CoV-2 infections: A living systematic review and meta-analysis. *PLoS Medicine*, 17(9):1–25, 09 2020. doi: 10.1371/journal.pmed.1003346.

[455] E Dickson, NE Palmateer, J Murray, C Robertson, C Waugh, LA Wallace, et al. Enhanced surveillance of COVID-19 in Scotland: population-based seroprevalence surveillance for SARS-CoV-2 during the first wave of the epidemic. *Public Health*, 190:132–134, 2021. doi: 10.1016/j.puhe.2020.11.014.

[456] CP Thompson, NE Grayson, RS Paton, JS Bolton, J Lourenço, BS Penman, et al. Detection of neutralising antibodies to SARS-CoV-2 to determine population exposure in Scottish blood donors between March and May 2020. *Eurosurveillance*, 25(42), 2020. doi: 10.2807/1560-7917.ES.2020.25.42.2000685.

[457] Sero-surveillance of COVID-19. Public Health England, 2021. Available from: `https://www.gov.uk/government/publications/national-covid-19-surveillance-reports/sero-surveillance-of-covid-19`.

[458] SF Lumley, DW Eyre, AL McNaughton, A Howarth, S Hoosdally, SB Hatch, et al. SARS-CoV-2 antibody prevalence, titres and neutralising activity in an antenatal cohort, United Kingdom, 14 April to 15 June 2020. *Eurosurveillance*, 25(42), 2020. doi: 10.2807/1560-7917.ES.2020.25.41.2001721.

[459] Infectious diseases in pregnancy screening (IDPS): programme overview. Public Health England, 2021. Available from: `https://www.gov.uk/guidance/infectious-diseases-in-pregnancy-screening-programme-overview#screening-test`.

[460] Maternity Services Monthly Statistics June 2019, experimental statistics. NHS Digital, 2019. Available from: `https://digital.nhs.uk/data-and-information/publications/statistical/maternity-services-monthly-statistics/june-2019/analysis`.

[461] Health of women before and during pregnancy: health behaviours, risk factors and inequalities. Public Health England, 2019. Available from: `https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/844210/Health_of_women_before_and_during_pregnancy_2019.pdf`.

[462] C Rosadas, P Randell, M Khan, MO McClure, and RS Tedder. Testing for responses to the wrong SARS-CoV-2 antigen? *The Lancet*, 396, 2020. doi: 10.1016/S0140-6736(20)31830-4.

[463] K El Bouzidi, T Pirani, C Rosadas, S Ijaz, M Pearn, S Chaudhry, et al. Severe Acute Respiratory Syndrome Coronavirus-2 Infections in Critical Care Staff: Beware the Risks Beyond the Bedside. *Critical Care Medicine*, 49, 2021. doi: 10.1097/CCM.0000000000004878.

[464] C Rosadas, M Khan, E Parker, F Marchesin, K Katsanovskaja, M Sureda-Vives, et al. Detection and quantification of antibody to SARS CoV 2 receptor binding domain provides enhanced sensitivity,

specificity and utility. *Journal of Virological Methods*, 302:114475, 2022. doi: 10.1016/j.jviromet.2022.114475.

[465] S Riley, H Wang, O Eales, D Haw, CE Walters, KEC Ainslie, et al. REACT-1 round 9 final report: Continued but slowing decline of prevalence of SARS-CoV-2 during national lockdown in England in February 2021. *medRxiv*, 2021. Available from: `https://doi.org/10.1101/2021.03.03.21252856`.

[466] E Mahase. Coronavirus: UK screens direct flights from Wuhan after US case. *BMJ*, 368, 2020. doi: 10.1136/bmj.m265.

[467] MS Razai, HKN Kankam, A Majeed, A Esmail, and DR Williams. Mitigating ethnic disparities in covid-19 and beyond. *BMJ*, 372, 2021. doi: 10.1136/bmj.m4921.

[468] R Mathur, CT Rentsch, CE Morton, WJ Hulme, A Schultze, B MacKenna, et al. Ethnic differences in SARS-CoV-2 infection and COVID-19-related hospitalisation, intensive care unit admission, and death in 17 million adults in England: an observational cohort study using the OpenSAFELY platform. *The Lancet*, 397(10286):1711–1724, 2021. doi: 10.1016/S0140-6736(21)00634-6.

[469] H Ward, C Atchison, M Whitaker, KEC Ainslie, J Elliott, L Okell, et al. SARS-CoV-2 antibody prevalence in England following the first peak of the pandemic. *Nature Communications*, 12, 2021. doi: 10.1038/s41467-021-21237-w.

[470] English Indices of Deprivation. Ministry of Housing Communities & Local Government, 2019. Available from: `https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/833973/File_2_-_IoD2019_Domains_of_Deprivation.xlsx`.

[471] E Fraley, C LeMaster, D Banerjee, S Khanal, R Selvarangan, and T Bradley. Cross-reactive antibody immunity against SARS-CoV-2 in children and adults. *Cellular & Molecular Immunology*, 18, 2021. doi: 10.1038/s41423-021-00700-0.

[472] TA Bates, JB Weinstein, S Farley, HC Leier, WB Messer, and FG Tafesse. Cross-reactivity of SARS-CoV structural protein antibodies against SARS-CoV-2. *Cell Reports*, 34, 2021. doi: 10.1016/j.celrep.2021.108737.

[473] WJ Rogan and B Gladen. Estimating prevalence from the results of a screening test. *American Journal of Epidemiology*, 107(1):71–76, 1978. doi: 10.1093/oxfordjournals.aje.a112510.

[474] G Apolone, E Montomoli, A Manenti, M Boeri, F Sabia, I Hyseni, et al. Unexpected detection of SARS-CoV-2 antibodies in the prepandemic period in Italy. *Tumori Journal*, 107(5), 2021. doi: 10.1177/0300891620974755.

[475] E Parodi. WHO asks for re-checks of research on when coronavirus first surfaced in Italy. Reuters, 2021. Available from: `https://www.reuters.com/business/healthcare-pharmaceuticals/who-re-checks-research-when-coronavirus-first-surfaced-italy-2021-06-01/`.

[476] E Kontopantelis, MA Mamas, J Deanfield, M Asaria, and T Doran. Excess mortality in England and Wales during the first wave of the COVID-19 pandemic. *Journal of Epidemiology & Community Health*, 75(3):213–223, 2021. doi: 10.1136/jech-2020-214764.

[477] Requirements for the retention of laboratory records and diagnostic material (Seventh Edition 2018). National Pathology Accreditation Advisory Council, 2018. Available from: `https://www.safetyandquality.gov.au/publications-and-resources/resource-library/requirements-retention-laboratory-records-and-diagnostic-material`.

[478] DJ Jamieson and SA Rasmussen. An update on COVID-19 and pregnancy. *American Journal of Obstetrics & Gynecology*, 226, 2022. doi: 10.1016/j.ajog.2021.08.054.

[479] J Daly. COVID sewage surveillance scaled back, but experts say data is crucial. ABC News, 2022. Available from: `https://www.abc.net.au/news/2022-01-31/sewage-covid-surveillance-scaled-back-experts-say-data-crucial/100784648`.

[480] R McCabe and OJ Watson. mrc-ide/covid-alternative-mortality: v.0.1.0, 2022. Available from: `https://zenodo.org/record/7185171`.

[481] AT Levin, N Owusu-Boaitey, S Pugh, BK Fosdick, AB Zwi, A Malani, et al. Assessing the burden of COVID-19 in developing countries: systematic review, meta-analysis and public policy implications. *BMJ Global Health*, 7(5), 2022. doi: 10.1136/bmjgh-2022-008477.

[482] DM Weinberger, J Chen, T Cohen, FW Crawford, F Mostashari, D Olson, et al. Estimation of Excess Deaths Associated With the COVID-19 Pandemic in the United States, March to May 2020. *JAMA Internal Medicine*, 180(10):1336–1344, 10 2020. doi: 10.1001/jamainternmed.2020.3391.

[483] Estimates of Excess Mortality Associated With COVID-19 Pandemic (as of 25 March 2022). World Health Organization, 2022. Available from: `https://www.who.int/data/stories/global-excess-deaths-associated-with-covid-19-january-2020-december-2021`.

[484] BS Endris, SM Saje, ZT Metaferia, BG Sisay, T Afework, YG Mengistu, et al. Excess mortality in the face of COVID-19: Evidence from Addis Ababa Mortality Surveillance Program. *SSRN*, 2021. doi: 10.2139/ssrn.3787447.

[485] F Mohamed, S Satti, M Ali Amhed, A Ahmed, M Yagob, Fakhreldeen, et al. Estimation of Coronavirus (COVID-19) Infections in Khartoum State (Sudan). 2020.

Available from: https://www.researchgate.net/publication/342122988_Estimation_of_Coronavirus_COVID-19_Infections_in_Khartoum_State_Sudan.

[486] S Abdella, S Riou, M Tessema, A Assefa, A Seifu, A Blachman, et al. Prevalence of SARS-CoV-2 in urban and rural Ethiopia: Randomized household serosurveys reveal level of spread during the first wave of the pandemic. *eClinicalMedicine*, 35, 2021. doi: 10.1016/j.eclinm.2021.100880.

[487] W Moser, M Fahal, E Abualas, S Bedri, M T Elsir, MFERO Mohamed, et al. SARS-CoV-2 Antibody Prevalence and Population-Based Death Rates, Greater Omdurman, Sudan. *Emerging Infectious Diseases*, 28:1026–1030, 2022. doi: 10.3201/eid2805.211951.

[488] Ethiopian Public Health Institute Public Health Emergency Operations Center COVID-19. Ethiopian Public Health Institute. Available from: https://ephi.gov.et/download/ephi-pheoc-covid-19/.

[489] Official account of Yemen Supreme National Emergency Committee for COVID-19. Twitter. Available from: https://twitter.com/ysneccovid19.

[490] Yemen cemetery struggles to dig enough graves as coronavirus spreads. Reuters, 2020. Available from: https://www.reuters.com/article/uk-health-coronavirus-yemen-death-idUKKBN24A0XK.

[491] UN Office for the Coordination of Humanitarian Affairs. Impact of COVID-19 on continuity of health services. Situation Reports, 2020. Available from: https://reports.unocha.org/en/country/sudan/card/49cURoYa8t/.

[492] W McCollum. Sudan FETP Conducts Targeted Testing for COVID-19 in Khartoum State. Training Programs in Epidemiology and Public Health Interventions Network, 2020. Available from: https://www.tephinet.org/sudan-fetp-conducts-targeted-testing-for-covid-19-in-khartoum-state.

[493] OJ Watson, N Abdelmagid, A Aljaile, AA Elhameed, C Whittaker, and N Brazeau. Report 39: Characterising COVID-19 epidemic dynamics and mortality under-ascertainment in Khartoum, Sudan. Imperial College London, 2020. Available from: https://doi.org/10.1038/s41467-021-22474-9.

[494] EK Gudina, S Ali, E Girma, A Gize, B Tegene, GB Hundie, et al. Seroepidemiology and model-based prediction of SARS-CoV-2 in Ethiopia: longitudinal cohort study among front-line hospital workers and communities. *The Lancet Global Health*, 9, 2021. doi: 10.1016/S2214-109X(21)00386-7.

[495] R Malaeb, N Yousef, O Al-Nagdah, QH Ali, MAS Saeed, A Haider, et al. High seroprevalence of antibodies against SARS-CoV-2 among healthcare workers 8 months after the first wave in Aden, Yemen. *PLoS Global Public Health*, 2(11):1–11, 11 2022. doi: 10.1371/journal.pgph.0000767.

[496] H Pifarré i Arolas, E Acosta, G López-Casasnovas, A Lo, C Nicodemo, T Riffe, et al. Years of life lost to COVID-19 in 81 countries. *Scientific Reports*, 11(1), 2021. doi: 10.1038/s41598-021-83040-3.

[497] A Misganaw, M Naghavi, A Walker, AH Mirkuzie, AZ Giref, TM Berheto, et al. Progress in health among regions of Ethiopia, 1990-2019: a subnational country analysis for the Global Burden of Disease Study 2019. *The Lancet*, 399(10332), 2022. doi: 10.1016/S0140-6736(21)02868-3.

[498] A Misganaw, DH Mariam, and T Araya. Association of socioeconomic and behavioral factors with adult mortality: analysis of data from verbal autopsy in Addis Ababa, Ethiopia. *BMC Public Health*, 13(1), 2013. doi: 10.1186/1471-2458-13-634.

[499] Z Munn, S Moola, K Lisy, D Riitano, and C Tufanaru. Methodological guidance for systematic reviews of observational epidemiological studies reporting prevalence and cumulative incidence data. *International Journal of Evidence-Based Healthcare*, 13:147–153. doi: 10.1097/XEB.0000000000000054.

[500] A Nkuba Ndaye, A Hoxha, J Madinga, J Mariën, M Peeters, FH Leendertz, et al. Challenges in interpreting SARS-CoV-2 serological results in African countries. *The Lancet Global Health*, 9, 2021. doi: 10.1016/S2214-109X(21)00060-7.

[501] L Mwananyanda, CJ Gill, W MacLeod, G Kwenda, R Pieciak, Z Mupila, et al. Covid-19 deaths in Africa: prospective systematic postmortem surveillance study. *BMJ*, 372, 2021. doi: 10.1136/bmj.n334.

[502] V DeGennaro Jr, T Schwartz, R Henderson, and MC Elie. A cross-sectional cohort study of prevalence of antibodies to COVID-19 in Port-au-Prince, Haiti. *medRxiv*, 2021. Available from: `https://doi.org/10.1101/2021.06.30.21259815`.

[503] United Nations General Assembly. Transforming our world: the 2030 Agenda for Sustainable Development, 2015. Available from: `https://sdgs.un.org/2030agenda`.

[504] AB Hogan, P Winskill, OJ Watson, PGT Walker, C Whittaker, M Baguelin, et al. Within-country age-based prioritisation, global allocation, and public health impact of a vaccine against SARS-CoV-2: A mathematical modelling analysis. *Vaccine*, 39(22):2995–3006, 2021. doi: 10.1016/j.vaccine.2021.04.002.

[505] B Borremans, A Gamble, KC Prager, SK Helman, AM McClain, C Cox, et al. Quantifying antibody kinetics and RNA detection during early-phase SARS-CoV-2 infection by time since symptom onset. *eLife*, 9:e60122, 2020. doi: 10.7554/eLife.60122.

[506] R Anderson, C Donnelly, D Hollingsworth, M Keeling, C Vegvari, R Baggaley, et al. Reproduction number (R) and growth rate (r) of the COVID-19 epidemic in the UK: methods of estimation, data sources, causes of heterogeneity, and use as a guide in policy formulation. The Royal Society, 2020. Available from: `https://royalsociety.org/-/media/policy/projects/set-c/set-covid-19-R-estimates.pdf?la=en-GB&hash=FDFFC11968E5D247D8FF641930680BD6`.

[507] The COVID-19 Genomics UK (COG-UK) consortium. An integrated national scale SARS-CoV-2 genomic surveillance network. *The Lancet Microbe*, 1(3), 2020. doi: 10.1016/S2666-5247(20)30054-9.

[508] COVID–19 Genomic Surveillance. Wellcome Sanger Institute, 2023. Available from: `https://covid19.sanger.ac.uk/lineages/raw`.

[509] M Morvan, AL Jacomo, C Souque, ML Wade, T Hoffmann, K Pouwels, et al. An analysis of 45 large-scale wastewater sites in England to estimate SARS-CoV-2 community prevalence. *Nature Communications*, 13(1), 2022. doi: 10.1038/s41467-022-31753-y.

[510] O Eales, KEC Ainslie, CE Walters, H Wang, C Atchison, D Ashby, et al. Appropriately smoothing prevalence data to inform estimates of growth rate and reproduction number. *Epidemics*, 40:100604, 2022. doi: 10.1016/j.epidem.2022.100604.

[511] O Eales, H Wang, D Haw, KEC Ainslie, CE Walters, C Atchison, et al. Trends in SARS-CoV-2 infection prevalence during England's roadmap out of lockdown, January to July 2021. *PLOS Computational Biology*, 18(11):1–16, 2022. doi: 10.1371/journal.pcbi.1010724.

[512] P Elliott, O Eales, N Steyn, D Tang, B Bodinier, H Wang, et al. Twin peaks: The Omicron SARS-CoV-2 BA.1 and BA.2 epidemics in England. *Science*, 376(6600):eabq4411, 2022. doi: 10.1126/science.abq4411.

[513] P Elliott, D Haw, H Wang, O Eales, CE Walters, KEC Ainslie, et al. Exponential growth, high prevalence of SARS-CoV-2, and vaccine effectiveness associated with the Delta variant. *Science*, 374(6574):eabl9551, 2021. doi: 10.1126/science.abl9551.

[514] Coronavirus (COVID-19) Infection Survey: methods and further information. Office for National Statistics, 2023. Available from: `https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/methodologies/covid19infectionsurveypilotmethodsandfurtherinformation`.

[515] Coronavirus (COVID-19) Infection Survey: England: Previous versions, 2023. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/datasets/coronaviruscovid19infectionsurveydata/2022.

[516] Coronavirus (COVID-19) Infection Survey: Scotland: Previous versions. Office for National Statistics, 2023. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/datasets/covid19infectionsurveyscotland/2022.

[517] Coronavirus (COVID-19) Infection Survey: Wales: Previous versions. Office for National Statistics, 2023. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/datasets/covid19infectionsurveywales/2022.

[518] Coronavirus (COVID-19) Infection Survey: Northern Ireland: Previous versions. Office for National Statistics, 2023. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/datasets/covid19infectionsurveynorthernireland/2022.

[519] Coronavirus (COVID-19) Infection Survey, UK: 15 July 2022. Office for National Statistics, 2022. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/coronaviruscovid19infectionsurveypilot/15july2022.

[520] Coronavirus (COVID-19) Infection Survey, UK: 24 March 2023. Office for National Statistics, 2023. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/coronaviruscovid19infectionsurveypilot/24march2023.

[521] J Wallinga and M Lipsitch. How generation intervals shape the relationship between growth rates and reproductive numbers. *Proceedings of the Royal Society B: Biological Sciences*, 274(1609): 599–604, 2007. doi: 10.1098/rspb.2006.3754.

[522] MD Homan and A Gelman. The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo. *The Journal of Machine Learning Research*, 15:1593 – 1623, 2014. Available from: https://www.jmlr.org/papers/volume15/hoffman14a/hoffman14a.pdf.

[523] SARS-CoV-2 variants in analyzed sequences, United Kingdom. Our World in Data, 2023. Available from: https://ourworldindata.org/grapher/covid-variants-area?country=~GBR.

[524] Reproduction number (R) and growth rate: methodology. gov.uk, 2021. Available from: `https://www.gov.uk/government/publications/reproduction-number-r-and-growth-rate-methodology/reproduction-number-r-and-growth-rate-methodology`.

[525] T Maishman, S Schaap, DS Silk, SJ Nevitt, DC Woods, and VE Bowman. Statistical methods used to combine the effective reproduction number, R(t), and other related measures of COVID-19 in the UK. *Statistical Methods in Medical Research*, 31(9):1757–1777, 2022. doi: 10.1177/09622802221109506.

[526] SV Barai and Y Reich. Ensemble modelling or selecting the best model: Many could be better than one. *AI EDAM*, 13(5):377–386, 1999. doi: 10.1017/S0890060499135029.

[527] M Di Napoli, F Carotenuto, A Cevasco, P Confuorto, D Di Martire, M Firpo, et al. Machine learning ensemble modelling as a tool to improve landslide susceptibility mapping reliability. *Landslides*, 17 (8), 2020. doi: 10.1007/s10346-020-01392-9.

[528] G Grenouillet, L Buisson, N Casajus, and S Lek. Ensemble modelling of species distribution: the effects of geographical and environmental ranges. *Ecography*, 34(1):9–17, 2011. doi: 10.1111/j.1600-0587.2010.06152.x.

[529] L Zhou, KK Lai, and L Yu. Least squares support vector machines ensemble models for credit scoring. *Expert Systems with Applications*, 37(1):127–133, 2010. doi: 10.1016/j.eswa.2009.05.024.

[530] CR Simpson, C Robertson, E Vasileiou, E Moore, C McCowan, U Agrawal, et al. Temporal trends and forecasting of COVID-19 hospitalisations and deaths in Scotland using a national real-time patient-level data platform: a statistical modelling study. *The Lancet Digital Health*, 3(8), 2021. doi: 10.1016/S2589-7500(21)00105-9.

[531] Northern Ireland population mid-year estimate. Office for National Statistics, 2022. Available from: `https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/timeseries/nipop/pop`.

[532] Covid-19 infection survey cost. Office for National Statistics, 2023. Available from: `https://www.ons.gov.uk/aboutus/transparencyandgovernance/freedomofinformationfoi/covid19infectionsurveycost`.

[533] Ethopia: Country confirms first COVID-19 related death April 5/update 7. Crisis24, 2020. Available from: `https://crisis24.garda.com/alerts/2020/04/ethiopia-country-confirms-first-covid-19-related-death-april-5-update-7`.

[534] S H Watare, Z A Edea, H T Desta, Fantu Lombamo, Negussie Yohannes, and Mulugeta Asefa. The First Hundred Days of Novel Corona-virus Disease 2019 in Ethiopia, 13 March to 21 June 2020: Retrospective Analysis. Research Square, 2020.

[535] G Timotewos, A Solomon, A Mathewos, A Addissie, S Bogale, T Wondemagegnehu, et al. First data from a population based cancer registry in Ethiopia. *Cancer Epidemiology*, 53:93–98, 2018. ISSN 1877-7821. doi: 10.1016/j.canep.2018.01.008.

[536] United Nations Department of Economic and Social Affairs. World Urbanization Prospects: The 2018 Revision. United Nations, 2018. Available from: `https://population.un.org/wup/publications/Files/WUP2018-Report.pdf`.

[537] AD Laytin, M Sultan, F Debebe, Y Walelign, G Fisseha, and A Gebreyesus. Critical care capacity in Addis Ababa, Ethiopia: A citywide survey of public hospitals. *Journal of Critical Care*, 63:1–7, 2021. ISSN 0883-9441. doi: 10.1016/j.jcrc.2021.01.010.

[538] S Murthy, A Leligdowicz, and NKJ Adhikari. Intensive Care Unit Capacity in Low-Income Countries: A Systematic Review. *PLoS ONE*, 10(1):1–12, 01 2015. doi: 10.1371/journal.pone.0116949.

[539] AA Bawazir. Cancer incidence in Yemen from 1997 to 2011: a report from the Aden cancer registry. *BMC Cancer*, 18, 2018. doi: 10.1186/s12885-018-4411-9.

[540] United Nations Department of Economic and Social Affairs. World Population Prospects 2019 - Volume II: Demographic Profiles. United Nations, 2020. Available from: `https://doi.org/10.18356/7707d011-en`.

[541] S El Sirgany and S Kiley. Coronavirus death rates in Yemen's Aden could exceed its wartime fatalities. CNN, 2020. Available from: `https://edition.cnn.com/videos/tv/2020/06/10/ctw-yemen-covid-coronavirus.cnn`.

[542] COVID-19 Pandemic [title translated into English]. Sudan Health Observatory, 2022. Available from: `http://sho.gov.sd/corona/index.php`.

[543] COVID-19 infections and deaths during the second wave of the pandemic in Sudan. Federal Ministry of Health Sudan Facebook Page, 2020. Available from: `https://m.facebook.com/story.php?story_fbid=2764654530474174&id=182728912754405`.

[544] Sudan: Population density and potential COVID-19 hotspots. United Nations Office for the Coordination of Humanitarian Affairs, 2020. Available from: `https://reports.unocha.org/en/country/sudan/card/ZXsJqLs913/`.

[545] Statistics on hospital sector in Sudan. Sudan NextGen, 2019. Available from: https://sudannextgen.com/wp-content/uploads/2019/10/Statistics-on-Hospital-sector-in-Sudan-2.pdf.

[546] H Sulieman, W El-Mahdi, M Awadelkareem, and L Nazer. Characteristics of Critically-Ill Patients at Two Tertiary Care Hospitals in Sudan. *Sultan Qaboos University Medical Journal*, 18, 2018. doi: 10.18295/squmj.2018.18.02.011.

[547] Coronavirus (COVID-19) Infection Survey, UK: 6 January 2023. Office for National Statistics, 2023. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/coronaviruscovid19infectionsurveypilot/6january2023.

Appendix A

Appendix: Paper I: Disease transmission and control modelling at the science-policy interface

Supplementary material for Paper I (Chapter 3)

R McCabe and CA Donnelly. Disease transmission and control modelling at the science–policy interface. *Interface Focus*, 11(6):20210013, 2021. doi: https://doi.org/10.1098/rsfs.2021.0013.

A.1 Questions in survey to attendees of SAGE, SPI-M and other comparable advisory committees

Table A.1 presents the survey questions sent to attendees of SAGE, SPI-M and other comparable advisory committees alongside number of respondents per option.

Table A.1: Questions, and multiple-choice (MC) answers where relevant, for the survey sent to attendees of SAGE, SPI-M and other comparable advisory committees.

Question		**Question type**	**Responses (if MC)**	**Number of respondents** ($(n(\%))$	
Questions on modelling					
1	Have you ever attended any SAGE, SPI-M or similar (for example, Scottish Government COVID-19 Advisory Group) meeting related to COVID-19?	Multiple-choice	Yes	46	(100%)
			No	0	(0%)
2	Did discussions with decision-makers inform the scenarios that were modelled?	Multiple-choice	Yes	24	(52%)
			No	3	(7%)
			Sometimes	14	(30%)
			Other	5	(11%)
			[Not answered]	0	(0%)
3	[Optional] Please provide any further comments you feel are relevant to Question 2.	Open-ended		33	(72%)
4	Did the presented modelling influence the scientific understanding of decision-makers?	Multiple choice	Yes	30	(65%)
			No	0	(0%)
			Sometimes	8	(17%)
			Other	7	(15%)
			[Not answered]	1	(2%)
5	[Optional] Please provide any further comments you feel are relevant to Question 4.	Open-ended		26	(56%)
6	Did the presented modelling influence the decision-making?	Multiple-choice	Yes	30	(65%)
			No	0	(0%)
			Sometimes	8	(17%)
			Other	7	(15%)
			[Not answered]	1	(2%)
7	[Optional] Please provide any further comments you feel are relevant to Question 6.	Open-ended		26	(56%)
8	To what extent were you involved in any policy decisions made?	Multiple-choice	Highly influential involvement	7	(15%)
			Moderately influential involvement	12	(26%)
			Minimal involvement	19	(41%)
			Other	7	(15%)
			[Not answered]	1	(2%)
9	[Optional] Please provide any further comments you feel are relevant to Question 8.	Open-ended		25	(54%)
10	Did the presented modelling generally make similar assumptions underlying transmission dynamics?	Multiple-choice	Yes	20	(43%)
			No	4	(9%)
			Sometimes	17	(37%)
			Other	4	(9%)
			[Not answered]	1	(2%)
11	[Optional] Please provide any further comments you feel are relevant to Question 10.	Open-ended		25	(54%)
12	Did the presented modelling generally draw similar conclusions about the appropriate next steps?	Multiple-choice	Yes	22	(48%)
			No	1	(2%)
			Sometimes	19	(41%)
			Other	2	(4%)

			[Not answered]	2	(4%)
13	[Optional] Please provide any further comments you feel are relevant to Question 12.	Open-ended		21	(46%)
14	How broad was the range of modelling scenarios that was considered?	Multiple-choice	Very broad	8	(17%)
			Reasonably broad	31	(67%)
			Not very broad	2	(4%)
			Other	4	(9%)
			[Not answered]	1	(2%)
15	[Optional] Please provide any further comments you feel are relevant to Question 14.	Open-ended		18	(39%)
16	Based on your experience of COVID-19, how could communication between policymakers and modellers be improved when responding to a future health emergency?	Open-ended		38	(83%)
17	Please provide any further comments about your experiences of COVID-19 modelling and/or policymaking here.	Open-ended		27	(59%)
Personal information					
18	Please select one of the following options.	Multiple-choice	[1]		
			[2]		
			[3]		
19	[Optional] Name	Open-ended			
20	[Optional] Job title and institution	Open-ended			

A.2 Survey Participant Information Sheet

The Participant Information Sheet for survey respondents is included on the following two pages.

[1] I give permission for written answers in this form to be used as quotes in the article which will be attributed to me using my full name and job title as stated below.

[2] I give permission for written answers in this form to be used as quotes in the article but without my full name. Instead, I can be referred to anonymously using only the description of my job title below.

[3] I give permission for written answers in this form to be used as quotes in the article only under complete anonymity.

Department of Statistics

24-29 St Giles'
Oxford
OX1 3LB

Professor Christl Donnelly
University Direct Line: 01865 ******
University e-mail: christl.donnelly@stats.ox.ac.uk

Transmission modelling at the science-policy interface

CUREC Approval Reference: R74566/RE001

General Information

The aim of this study is to support us in writing an article for a special edition of the *Journal of the Royal Society Interface Focus* themed on COVID-19 in society. We are looking at this through the lens of mathematical modelling. As well as reviewing some of the key literature, we want to further understand the relationship between modelling and policymaking as well as how this is reported in the media. As the pandemic continues, understanding these relationships and how they can be improved is critical. Moreover, this research will provide historical context of COVID-19 modelling and how this has shaped the UK response in the coming years.

We appreciate your interest in participating in this online survey. You have been invited to participate as you are publicly listed as being a member of or having attended a meeting of SAGE, SPI-M or other comparable advisory body.

You may ask any questions by contacting the researcher (details below).

The Principal Researcher is Ruth McCabe, who is attached to the Department of Statistics at the University of Oxford. This project is being completed under the supervision of Professor Christl Donnelly.

The link above provides access to the survey. This includes 8 multiple choice, with option to expand answers, and 2 open-ended questions. The final questions ask for explicit consent from you under a chosen level of anonymity. There is an option to enter a name and job title, but if this information is not provided there is no way in which you can be identified. This should take about 10 minutes to complete.

The data gathered will be used to form the discussion within the article. Answers to multiple-choice questions remain anonymous while answers to open-ended questions may be quoted within the article under the chosen level of anonymity. Only the Principal Researcher and Supervisor will have access to the raw data, although summaries and, conditional on consent, some individual quotes will be used in the article and the corresponding DPhil thesis chapter.

Do I have to take part?

No. Please note that participation is voluntary. If you do decide to take part, you may withdraw at any point, for any reason, before submitting your answers by closing the browser.

If you have already taken part, but now wish to withdraw, please let us know by emailing us. However, if you have not provided a name or job title we will be unable to identify your response.

How will my data be used?

The data we will collect that could identify you will be your name and job title, but only if you voluntarily enter this information.

Identifiable information will be deleted as soon as it is no longer required for the research. This is expected

to be in approximately 4.5 years' time to allow for completion of the DPhil thesis. Research data (including consent records) will be stored for three years after publication or public release.

Your IP address will not be stored. We will take all reasonable measures to ensure that data remain confidential.

The responses you provide will be used in an academic publication under the specified level of anonymity determined in the final questions of the survey. If you explicitly provide consent to use your name and job title, or just job title, then this will be written in the article next to the quote provided in the survey, if it is chosen for inclusion, and thus may be attributable to you. Otherwise, any quotes you provide (and we use) will be entirely unattributable to you.

Who will have access to my data?

The University of Oxford is the data controller with respect to your personal data, and as such will determine how your personal data is used in the study. The University will process your personal data for the purpose of the research outlined above. Research is a task that we perform in the public interest. Further information about your rights with respect to your personal data is available from https://compliance.admin.ox.ac.uk/individual-rights.

The data you provide will not be shared with any other party other than the principal researcher and their supervisor.

This survey will be written up for a DPhil degree.

Who has reviewed this study?

This project has been reviewed by, and received ethics clearance through, the University of Oxford Central University Research Ethics Committee R74566/RE001.

Who do I contact if I have a concern or I wish to complain?

If you have a concern about any aspect of this study, please contact Ruth McCabe at ruth.mccabe@linacre.ox.ac.uk or her supervisor Professor Christl Donnelly at christl.donnelly@stats.ox.ac.uk, and we will do our best to answer your query. We will acknowledge your concern within 10 working days and give you an indication of how it will be dealt with. If you remain unhappy or wish to make a formal complaint, please contact the Chair of the Medical Sciences Interdivisional Research Ethics Committee at the University of Oxford who will seek to resolve the matter as soon as possible:

Email: ethics@medsci.ox.ac.uk; Address: Research Services, University of Oxford, Wellington Square, Oxford OX1 2JD

If you have read the information above, have already taken part and are still happy for your data to be used as described above, then no further action is needed. We thank you for your participation.

If you have read the information above, have already taken part and now wish to withdraw from the study then please contact Professor Christl Donnelly (christl.donnelly@stats.ox.ac.uk).

If you have read the information above, agree to participate with the understanding that the data (including any personal data) you submit will be processed accordingly and have not already taken part, you can take part now by clicking this link.

A.3 Interview Details

Table A.2 provides details of the interviewees and interviews.

Table A.2: The names of interview participants, alongside the date of interview and platform on which it was conducted.

Interview participant	**Job Title**	**Date**	**Platform**
Professor Linda Bauld	Bruce and John Usher Chair of Public Health at the University of Edinburgh	10 June 2021	Zoom
Professor Sir John Beddington	Former Government Chief Scientific Adviser 2008 - 2013	17 February 2021	Zoom
Dr. Claire Craig	Former Director of the Government Office for Science	15 April 2021	In person at the Queen's College, Oxford
Robert Cuffe	Head of Statistics at the British Broadcasting Corporation (BBC)	29 January 2021	Telephone
Fiona Fox	Director of the Science Media Centre	25 January 2021	Microsoft Teams
Professor Sir Charles Godfray	Director of the Oxford Martin School	24 February 2021	Microsoft Teams
Scott Heald	Head of Profession for Statistics at Public Health Scotland	10 February 2021	Microsoft Teams
Lord John Krebs	Former Chairman of the UK Food Standard Agency	28 January 2021	Telephone
Professor Jason Leitch	National Clinical Director for the Scottish government	18 May 2021	Microsoft Teams
William Pryor	Head of Policy Engagement at the University of Oxford	2 February 2021	Microsoft Teams
Dr. Sabine van Elsland	External Relationships and Communications Manager at the Medical Research Council Centre for Global Infectious Disease Analysis (MRC GIDA) at Imperial College London	21 January 2021	Microsoft Teams

A.4 Interview Participant Information Sheet

The Participant Information Sheet for interviewees is included on the following two pages.

Department of Statistics

24-29 St Giles'
Oxford
OX1 3LB

Professor Christl Donnelly
University Direct Line: 01865 ******
University e-mail: christl.donnelly@stats.ox.ac.uk

Transmission modelling at the science-policy interface

CUREC Approval Reference: R74566/RE001

General Information

The aim of this study is to support us in writing an article for a special edition of the *Journal of the Royal Society Interface Focus* themed on COVID-19 in society. We are looking at this through the lens of mathematical modelling. As well as reviewing some of the key literature, we want to further understand the relationship between modelling and policymaking as well as how this is reported in the media. As the pandemic continues, understanding these relationships and how they can be improved is critical. Moreover, this research will provide historical context of COVID-19 modelling and how this has shaped the UK response in the coming years.

We appreciate your participation in this study. You were invited to participate due to your expertise in the interface between science, policy and media.

You may ask any questions by contacting the researcher (details below).

The Principal Researcher is Ruth McCabe, who is attached to the Department of Statistics at the University of Oxford. This project is being completed under the supervision of Professor Christl Donnelly.

The interview took place virtually, on a platform of your choice. You were asked broad questions about the relationships between modelling, decision-making, the media and the public but were strongly encouraged to shape the conversation in the way that you feel most comfortable with. You were asked for permission for the meeting to be recorded. The interview was scheduled to take approximately half an hour.

The data gathered is being used to form the discussion within the article. Only the Principal Researcher and Supervisor will have access to the raw data, although summaries and, conditional on consent, some individual quotes will be used in the article and the corresponding DPhil thesis chapter.

Do I have to take part?

No. Please note that participation is voluntary. If you have taken part, you may withdraw at any point, for any reason by contacting Professor Christl Donnelly.

How will my data be used?

The recording of the interview will be transcribed and the original recording deleted.

Transcripts and identifiable information will be deleted as soon as it is no longer required for the research. This is expected to be in approximately 4.5 years' time to allow for conferment of the DPhil thesis. Research data (including consent records) will be stored for three years after publication or public release.

We will take all reasonable measures to ensure that data remain confidential.

The responses you provided will be used in an academic publication under the specified level of anonymity determined by you. If you explicitly provide consent to use your name and job title, or just job title, then this

will be written in the article next to quotes provided in the interview, if it is chosen for inclusion, and thus may be attributable to you. You will be asked to approve the use of your quotations before this identifiable information is provided.

Who will have access to my data?

The University of Oxford is the data controller with respect to your personal data, and as such will determine how your personal data is used in the study. The University will process your personal data for the purpose of the research outlined above. Research is a task that we perform in the public interest. Further information about your rights with respect to your personal data is available from https://compliance.admin.ox.ac.uk/individual-rights.

The data you provide will not be shared with any other party other than the principal researcher and their supervisor.

This interview will be written up for a DPhil degree.

Who has reviewed this study?

This project has been reviewed by, and received ethics clearance through, the University of Oxford Central University Research Ethics Committee R74566/RE001.

Who do I contact if I have a concern or I wish to complain?

If you have a concern about any aspect of this study, please contact Ruth McCabe at ruth.mccabe@linacre.ox.ac.uk or her supervisor Professor Christl Donnelly at christl.donnelly@stats.ox.ac.uk, and we will do our best to answer your query. We will acknowledge your concern within 10 working days and give you an indication of how it will be dealt with. If you remain unhappy or wish to make a formal complaint, please contact the Chair of the Medical Sciences Interdivisional Research Ethics Committee at the University of Oxford who will seek to resolve the matter as soon as possible:

Email: ethics@medsci.ox.ac.uk; Address: Research Services, University of Oxford, Wellington Square, Oxford OX1 2JD

If you have read the information above, have already taken part and are still happy for your data to be used as described above, then no further action is needed. We thank you for your participation.

If you have read the information above, have already taken part and now wish to withdraw from the study then please contact Professor Christl Donnelly (christl.donnelly@stats.ox.ac.uk).

Appendix B

Appendix: Paper II: Communicating uncertainty in epidemic models

Supplementary material for Paper II (Chapter 4)

R McCabe, MD Kont, N Schmit, C Whittaker, A Løchen, PGT Walker, AC Ghani, NM Fergusion, PJ White, CA Donnelly and OJ Watson. Communicating uncertainty in epidemic models. *Epidemics*, 37:100520, 2021. doi: `https://doi.org/10.1016/j.epidem.2021.100520`.

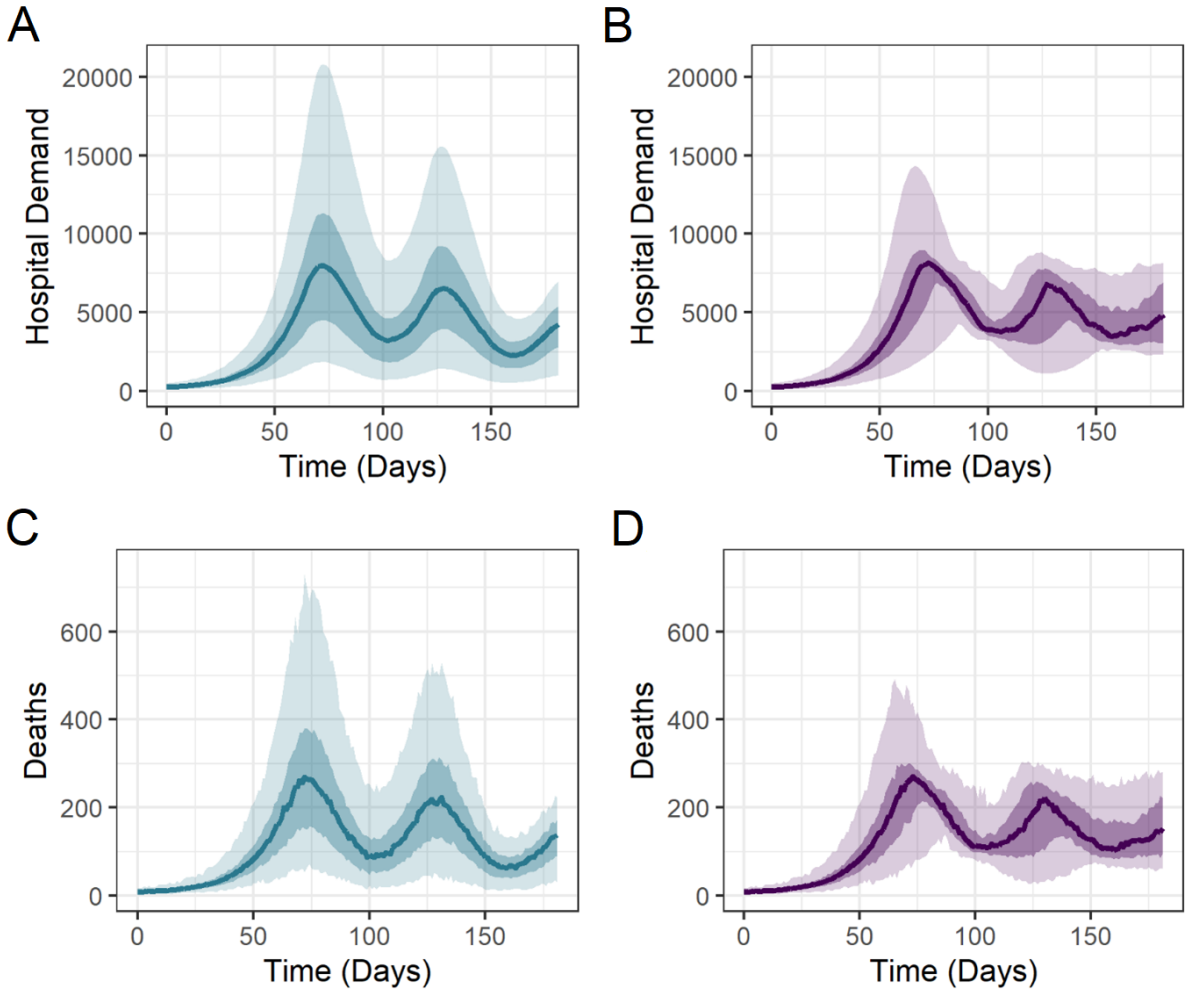

Figure B.1: Traditional presentation of epidemic forecasts under two suppression strategies. The median (line) with 50% credible intervals (dark shading) and 95% credible intervals (light shading) of 100 simulations of hospital demand from COVID-19 patients and COVID-19 deaths are shown for strategies in which suppression measures are (A) and (C) scheduled at a fixed time point and (B) and (D) triggered reactively.

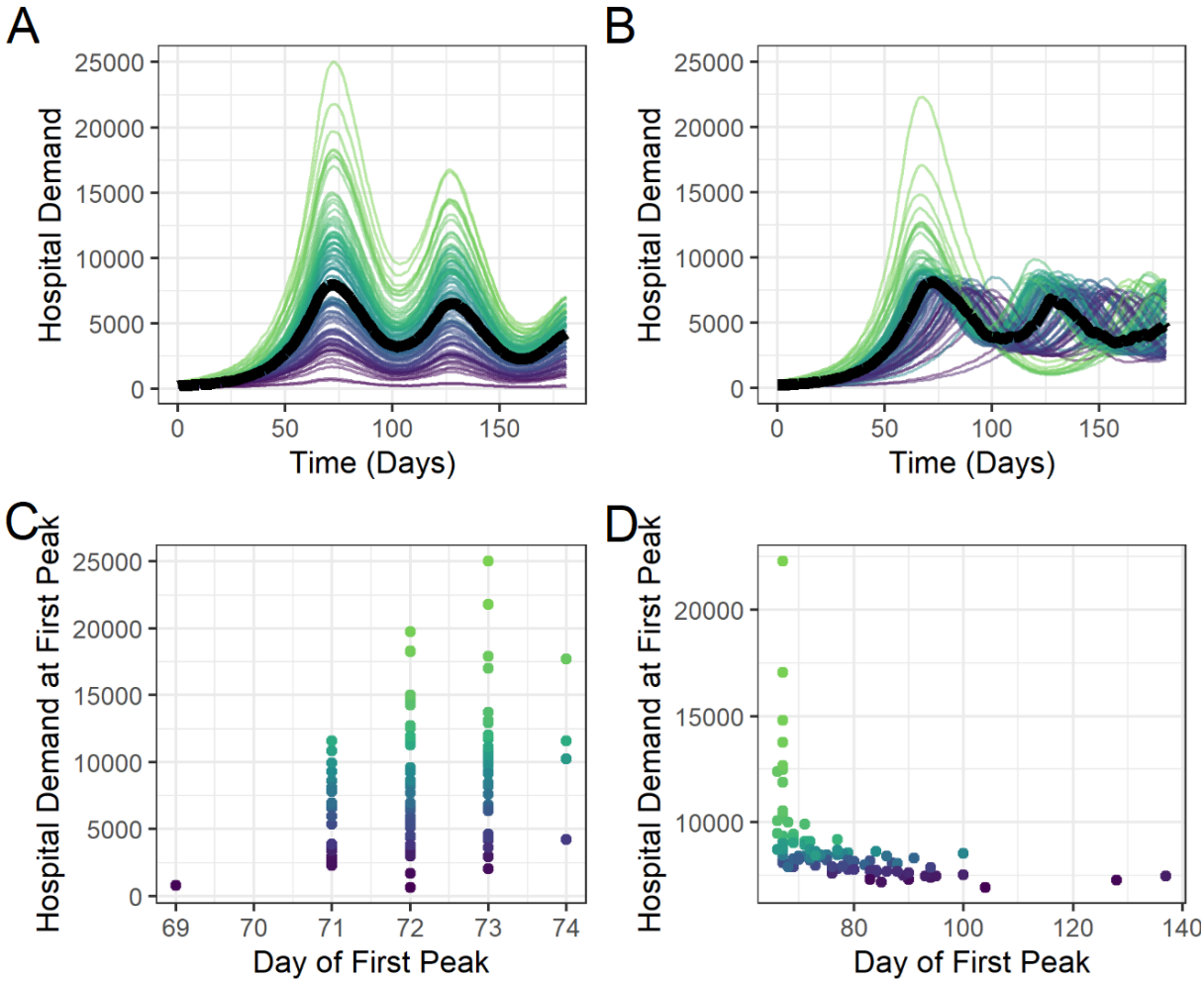

Figure B.2: Communicating uncertainty in epidemic forecasts of hospital demand under two suppression strategies. Individual simulations of hospital demand from COVID-19 patients is shown for strategies in which suppression measures are (A) scheduled at a fixed time point and (B) triggered reactively, with corresponding relationships between the magnitude and timing of peak hospital demand for each realisation shown in (C) and (D), respectively. In (A) and (B) the median trajectory is shown with a black line. The colour of each simulation realisation depicts the ranking of hospital demand at first peak [high in green to low in purple] and allows for the metrics in (C) and (D) to be more easily linked to the trajectories in (A) and (B) respectively.

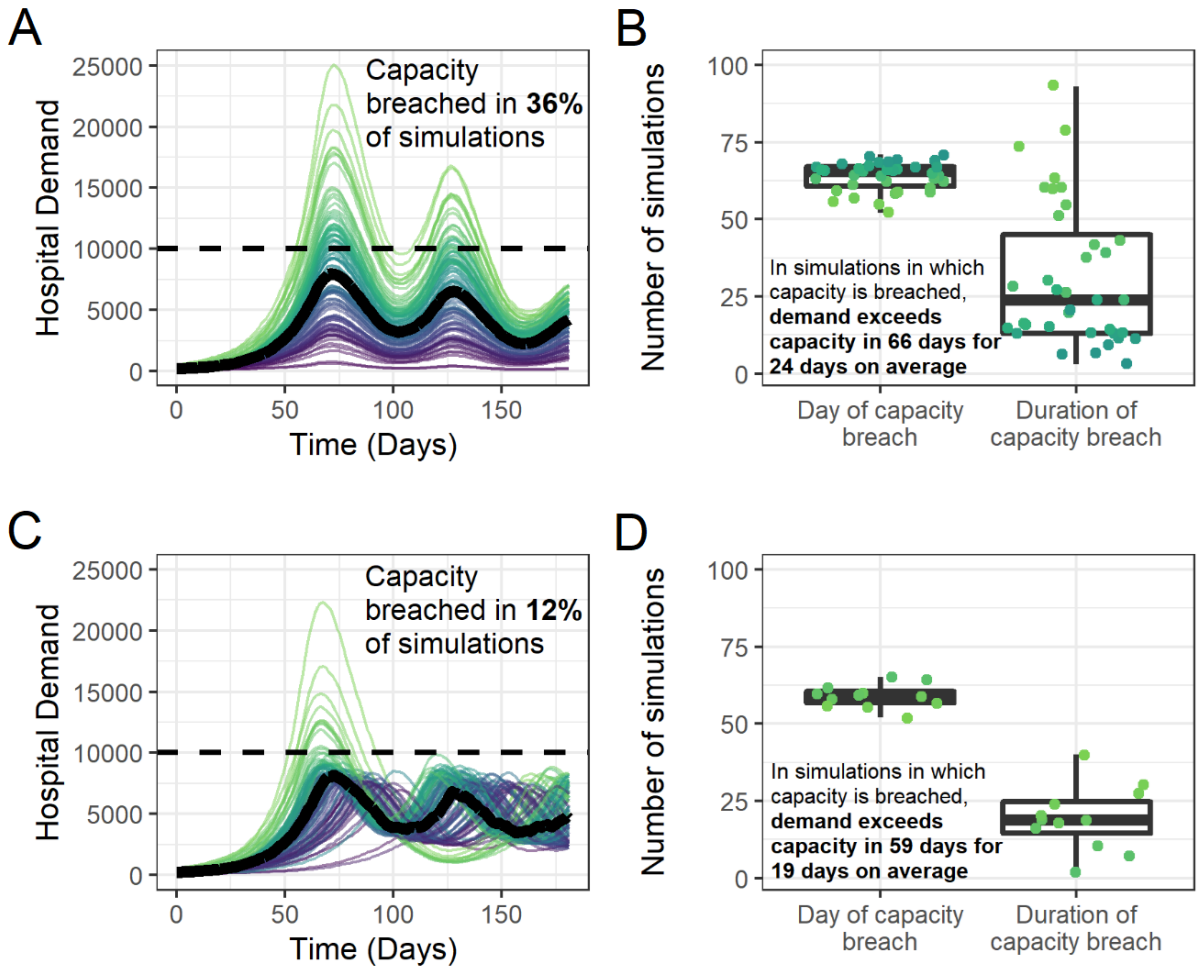

Figure B.3: Linking epidemic forecasts of hospital demand under two suppression strategies to potential capacity breaches. Individual simulations of hospital demand from COVID-19 patients compared to an assumed hospital capacity of 10000 available beds is shown for strategies in which suppression measures are (A) scheduled at a fixed time point and (C) triggered reactively. In simulations in which capacity is breached, corresponding metrics of the day on which capacity is breached and the duration of the capacity breach shown in (B) and (D), respectively. In (A) and (C) the median trajectory is shown with a black line. The colour of each simulation realisation depicts the ranking of hospital demand at first peak [high in green to low in purple] and allows for the metrics in (B) and (D) to be more easily linked to the trajectories in (A) and (C) respectively.

Appendix C

Appendix: Paper III: Public awareness and opinions on the use of mathematical transmission modelling to inform public health policy in the United Kingdom

Supplementary material for Paper III (Chapter 5)

R McCabe and CA Donnelly. Public awareness of and opinions on the use of mathematical transmission modelling to inform public health policy in the United Kingdom. *Journal of the Royal Society Interface*, 20:20230456, 2023. doi: `https://doi.org/10.1098/rsif.2023.0456`.

C.1 Participant Information Sheet

The Participant Information Sheet for survey respondents is included on the following three pages.

Department of Statistics
24-29 St Giles'
Oxford
OX1 3LB

Professor Christl Donnelly
University Direct Line: 01865 ******
University e-mail: christl.donnelly@stats.ox.ac.uk

Awareness and opinions on the use of transmission modelling in informing policy

CUREC Approval Reference: R76166/RE001

General Information

We recently wrote an article for a special edition of the *Journal of the Royal Society Interface Focus* themed on COVID-19 in society. We looked at this through the lens of mathematical modelling. We surveyed attendees of the Scientific Advisory Group for Emergencies (SAGE) or the Scientific Pandemic Influenza Group on Modelling (SPI-M), meetings to elicit their experiences of the development and use of transmission modelling in policy throughout the COVID-19 pandemic. We now want to build upon this work by surveying the public on their awareness and opinions of the use of modelling in informing policy throughout the pandemic. Please note that participants are not required to have any prior knowledge of transmission modelling to participate.

We appreciate your interest in participating in this online survey. You have been invited to participate as you are an adult over 18 years of age and residing within the United Kingdom of Great Britain and Northern Ireland.

You may ask any questions by contacting the researcher (details below).

The Principal Researcher is Ruth McCabe, who is attached to the Department of Statistics at the University of Oxford. This project is being completed under the supervision of Professor Christl Donnelly.

You are invited to participate in a survey consisting of 13 multiple choice questions and 8 open-ended questions. At the end of the survey, basic demographic information covering age, gender and occupation are asked for. Participants cannot be identified from this information. This should take about 10 minutes.

The data gathered will be used to write a short commentary in follow up to our original article and will also be incorporated into the corresponding doctoral thesis in statistics. Answers to all questions are anonymous. Only the Principal Researcher and Supervisor will have access to the raw, anonymised data.

Do I have to take part?

No. Please note that participation is voluntary. If you do decide to take part, you may withdraw at any point for any reason before submitting your answers by closing the browser. However, once submitted your data cannot be removed due to their being no identifiable information collected.

How will my data be used?

We will not collect any information that could be used to identify you.

Anonymous research data (including consent records) will be stored for three years after publication or public release.

Your IP address will not be stored. We will take all reasonable measures to ensure that data remain confidential.

The responses you provide will be used in an academic publication.

Who will have access to my data?

The researchers named above will have access to the raw data that you provide, which is anonymised and from which you cannot be identified. The University will process the data you provide for the purpose of the research outlined above. Research is a task that we perform in the public interest. Further information about your rights with respect to your personal data is available from https://compliance.admin.ox.ac.uk/individual-rights.

The data you provide may be shared with third parties (researchers at partner universities) for analysis and publication purposes.

This survey will be written up for a doctoral thesis in statistics.

Who has reviewed this study?

This project has been reviewed by, and received ethics clearance through, the Medical Sciences Interdivisional Research Ethics Committee, a subcommittee of the University of Oxford Central University Research Ethics Committee R76166/RE001.

Who do I contact if I have a concern or I wish to complain?

If you have a concern about any aspect of this study, please write to Ruth McCabe at ruth.mccabe@stats.ox.ac.uk or her supervisor Professor Christl Donnelly at christl.donnelly@stats.ox.ac.uk, and we will do our best to answer your query. We will acknowledge your concern within 10 working days and give you an indication of how it will be dealt with. If you remain unhappy or wish to make a formal complaint, please contact the Chair of the Medical Sciences Interdivisional Research Ethics Committee at the University of Oxford who will seek to resolve the matter as soon as possible:

Email: ethics@medsci.ox.ac.uk; Address: Research Services, University of Oxford, Wellington Square, Oxford OX1 2JD

The following two statements form the first two questions of the survey, which you must complete in order to proceed to the rest of the questions.

Please note that you may only participate in this survey if you are 18 years of age or over and residing in the United Kingdom of Great Britain and Northern Ireland (UK).

☐ I certify that I am 18 years of age or over and reside in the UK.

If you have read the information above and agree to participate with the understanding that the data you submit will be processed accordingly, please check the relevant box below to get started.

☐ Yes, I agree to take part

C.2 Methods

Table C.1 presents the details of the accounts who shared the link to the survey, either by tweeting it directly or quote tweeting the link, as of 18 July 2023.

Table C.1: Details of accounts who directly tweeted or quote tweeted (QT) the link to the survey.

Type of tweet	**Account name**	**Twitter handle**	**Date tweeted**	**Link**	**No. QTs**	**No. retweets**	**No. likes**	**No. replies**
Direct	HPRU EZI	@hpruezi	7 July 2021	[1]	3	6	5	4
Direct	Oxford Statistics	@oxfordstats	July 2021	[2]	Although we have the link to this Tweet, it is unfortunately no longer available and so specific information on reach is unavailable.			
QT	Liam McCabe	@liamrun	7 July 2021	[3]	0	0	1	0
QT	Ruth McCabe	@ruth_mccabe	16 July 2021	[4]	3	6	6	1
QT	Jameel Institute	@Imperial_Jameel	16 July 2021	[5]	0	1	6	0
QT	Dr David Telford	@Davie_T	16 July 2021	[6]	0	1	4	0
QT	Marie	@marieclarkg12	16 July 2021	[7]	0	1	3	2
QT	Katharina Hauck	@kdhauck	16 July 2021	[8]	0	0	6	0
QT	Oonagh Gil	@OonaghGil	17 July 2021	[9]	0	1	3	0
QT	MRC Centre for Global Infectious Disease Analysis	@MRC_Outbreak	20 July 2021	[10]	0	2	3	0
QT	David Spiegelhalter	@d_spiegel	16 July 2021	[11]	0	9	24	2

[1] https://twitter.com/HPRUezi/status/1412793471895748608
[2] https://twitter.com/OxfordStats/status/1413446810111135753
[3] https://twitter.com/liamrun/status/1412819957696638976
[4] https://twitter.com/ruth_mccabe/status/1415943038900264962
[5] https://twitter.com/Imperial_Jameel/status/1415984380045152256
[6] https://twitter.com/Davie_T/status/1415958167687733251
[7] https://twitter.com/marieclarkg12/status/1415988394447020033
[8] https://twitter.com/kdhauck/status/1416000017597358080
[9] https://twitter.com/OonaghGil/status/1416347577730146306
[10] https://twitter.com/MRC_Outbreak/status/1417485287932956673
[11] https://twitter.com/d_spiegel/status/1422190746787106817

C.3 Results: Figures, tables and additional analyses

C.3.1 Awareness of mathematical transmission modelling

C.3.1.1 Were you aware of the use of transmission models in informing public health policy?

Table C.2 presents the results of Wilcoxon signed rank tests for differences between awareness of modelling generally and awareness specifically in relation to use in policy, for both samples and time periods. Responses in which a respondent stating having no awareness or being unsure were enumerated as 0, with responses stating awareness enumerated as 1.

Table C.3 presents the results of the linear model regressing responses to the question "Were you aware of the use of transmission models in informing public health policy?" both "Prior to" and "During" the COVID-19 pandemic on age group and gender for each sample. The three responses provided as multiple-choice options, "No", "Unsure" and "Yes", were enumerated as -1, 0, 1, respectively.

Table C.2: Differences between awareness of transmission modelling generally and specifically in relation to use in policy for both time periods and samples according to Wilcoxon signed rank tests. $p < 0.01$ is considered statistically significant.

		Number of respondents (n (%))		**Wilcoxon signed rank test statistic**	n	p
		Not aware generally	**Not aware or unsure regarding policy**			
Prior to the COVID-19 pandemic	Online panel	314 (62%)	365 (72%)	3927	504	< 0.001
	Social media	74 (37%)	86 (42%)	581	202	0.065
During the COVID-19 pandemic	Online panel	114 (23%)	122 (24%)	1326	504	0.180
	Social media	6 (3%)	6 (3%)	18	202	1.000

Table C.3: Coefficient estimates, standard errors and p-values for the linear model regressing responses to the question "Were you aware of the use of transmission models in informing public health policy?" both "Prior to" and "During" the COVID-19 pandemic on age group and gender for each sample. Chi-squared tests for the overall significance of age and gender are also presented. The three responses provided as multiple-choice options, "No", "Unsure" and "Yes", were enumerated as -1, 0, 1, respectively. $p < 0.01$ is considered statistically significant. *Baseline comprised of "Female" and "18 - 25" years.

Sample	**Response**	**Variable**	**Interpretation**	**Estimate**	**Standard error**	p	**Chi-squared test**	
							χ^2	p
Online panel	Prior to the COVID-19 pandemic, were you aware of the use of transmission models in informing public health policy?	Intercept	Baseline*	-0.42	0.11	< 0.001		
		Age (years)	26 - 35	0.17	0.14	0.227	3.63 (df=5)	0.432
			36 - 45	0.05	0.14	0.725		
			46 - 55	0.20	0.14	0.140		
			56 - 65	0.06	0.13	0.650		
			66+	-0.13	0.17	0.465		
		Gender	Male	0.14	0.08	0.072	3.56 (df=3)	0.189
			Non-binary	-0.75	0.87	0.387		
			Prefer not to say	-0.34	0.50	0.495		
Social media	Prior to the COVID-19 pandemic, were you aware of the use of transmission models in informing public health policy?	Intercept	Baseline*	0.29	0.43	0.501		
		Age (years)	26 - 35	0.14	0.49	0.774	2.01 (df=5)	0.718
			36 - 45	0.07	0.45	0.882		
			46 - 55	-0.19	0.43	0.669		
			56 - 65	-0.13	0.43	0.771		
			66+	-0.10	0.45	0.823		
		Gender	Male	0.28	0.12	0.019	5.30 (df=2)	0.024
			Prefer not to say	0.84	0.50	0.019		
Online panel	During the COVID-19 pandemic, were you aware of the use of transmission models in informing public health policy?	Intercept	Baseline*	0.51	0.01	< 0.001		
		Age (years)	26 - 35	0.01	0.12	0.927	2.11 (df=5)	0.718
			36 - 45	0.03	0.12	0.792		
			46 - 55	0.18	0.12	0.140		
			56 - 65	0.13	0.11	0.223		
			66+	0.11	0.15	0.458		
		Gender	Male	0.03	0.07	0.694	0.83 (df=3)	0.679
			Non-binary	0.48	0.74	0.517		
			Prefer not to say	0.43	0.43	0.313		
Social media	During the COVID-19 pandemic, were you aware of the use of transmission models in informing public health policy?	Intercept	Baseline*	0.94	0.12	< 0.001		
		Age (years)	26 - 35	-0.19	0.14	0.185	0.38 (df=5)	0.247
			36 - 45	-0.06	0.13	0.623		
			46 - 55	-0.01	0.12	0.940		
			56 - 65	-0.01	0.12	0.962		
			66+	-0.01	0.13	0.921		
		Gender	Male	0.08	0.03	0.023	0.30 (df=2)	0.247
			Prefer not to say	0.05	0.14	0.708		

C.3.1.2 How were you aware of mathematical transmission modelling?

Figure C.1 presents responses to the question "How were you aware of transmission modelling?" for both samples and for both time points.

Table C.4 presents the results of Wilcoxon signed rank tests to assess differences in the number of

respondents selecting responses to the aforementioned questions across the two time periods. For each multiple-choice answer, data were enumerated to a binary scale with 1 indicating that this answer was selected and 0 indicating that it was not selected.

Table C.5 present the results of chi-squared tests for differences between the proportions of respondents selecting responses to the aforementioned question across the two samples.

Although not among the most common responses, the percentage of respondents reporting "Internet search" as their source of information on modelling almost doubled and more than tripled in the period during the pandemic compared to the period before, for the online panel (Wilcoxon signed rank test $p < 0.001$) and social media (Wilcoxon signed rank test $p < 0.001$) respectively, indicating an increased appetite for information on mathematical modelling among both samples. This was verified by individuals commenting that their trust is often highly dependent on transparency of evidence, which was also highlighted in our initial study.

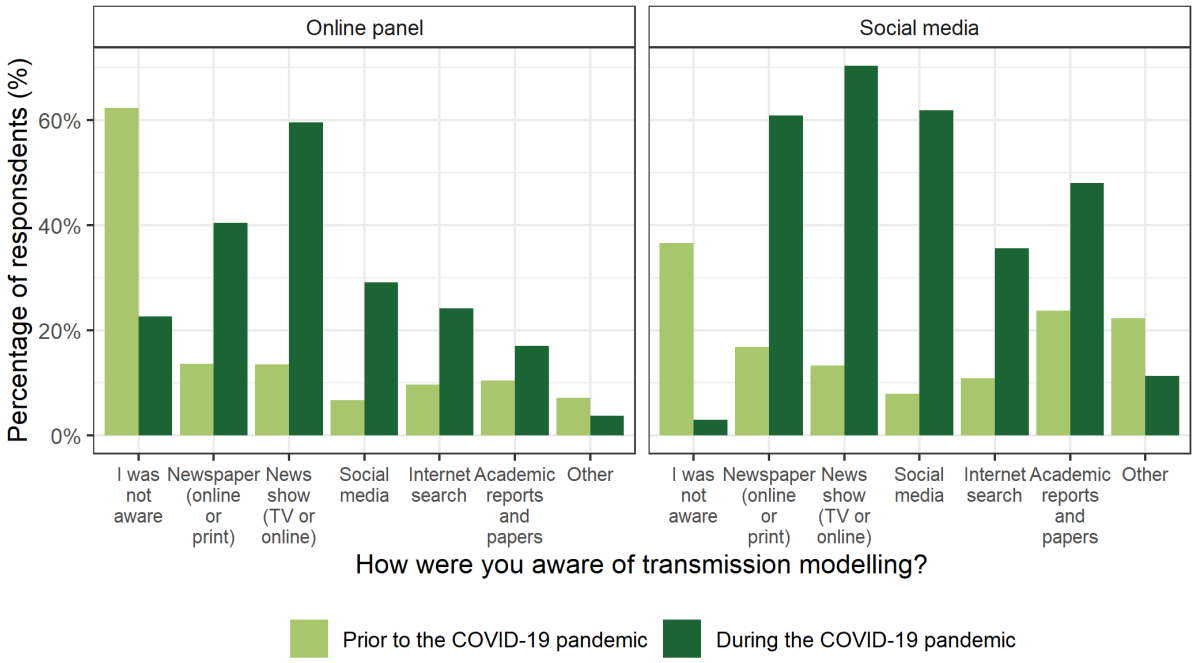

Figure C.1: Responses to the question "How were you aware of transmission modelling?" either "Prior to" or "During" the COVID-19 pandemic for both the online panel and social media samples. Underlying data are presented in Table 5.1.

Table C.4: Differences between the popularity of responses to the question "How were you aware of transmission modelling?" across time periods within samples, according to Wilcoxon signed rank tests. $p < 0.01$ is considered statistically significant.

		Number of respondents (n (%))		**Wilcoxon signed rank test statistic**	n	p
		Prior to the COVID-19 pandemic	**During the COVID-19 pandemic**			
Online panel	I was not aware	314 (62%)	114 (23%)	21318	504	< 0.001
	Newspaper (online or print)	69 (14%)	204 (40%)	288		< 0.001
	News show (TV or online)	68 (13%)	300 (60%)	966		< 0.001
	Social media	34 (7%)	147 (29%)	180		< 0.001
	Internet search	49 (10%)	122 (24%)	588		< 0.001
	Academic reports and papers	53 (11%)	86 (17%)	684		< 0.001
	Other	36 (7%)	19 (4%)	513		0.005
Social media	I was not aware	74 (37%)	6 (3%)	2450	202	< 0.001
	Newspaper (online or print)	34 (17%)	123 (61%)	144		< 0.001
	News show (TV or online)	27 (13%)	142 (70%)	183		< 0.001
	Social media	16 (8%)	125 (62%)	56		< 0.001
	Internet search	22 (11%)	72 (36%)	189		< 0.001
	Academic reports and papers	48 (24%)	97 (48%)	244		< 0.001
	Other	45 (22%)	23 (11%)	446		< 0.001

C.3.1.3 How much did you know about how transmission modelling has been used throughout the COVID-19 pandemic?

Figure C.2 presents responses to the question "How much did you know about how transmission modelling has been used throughout the COVID-19 pandemic?" under both the online panel and social media samples.

Table C.6 presents the results of the linear regression model regressing the answers to this question on age group and gender. The three responses provided as multiple-choice options, "Too little", "About right" and "Too much", were enumerated as -1, 0, 1, respectively.

Large numbers of respondents from both samples (46% from the online panel and 30% from social media) reported having "Too little" information about the use of transmission modelling during the pandemic, although this was significantly greater among the online panel (chi-squared test: $\chi^2_{df=1} = 15.55$; $n = 706$; $p < 0.001$). By comparison, social media respondents (8%) were significantly more likely to report having "Too much" information than the online panel (2%) ($\chi^2_{df=1} = 14.52$; n=706; $p < 0.001$). However, it

Table C.5: Differences between the popularity of responses to the question "How were you aware of transmission modelling?" across samples within time periods, according to chi-squared tests. $p < 0.01$ is considered statistically significant.

		Number of respondents (n (%))		$\chi^2_{df=1}$	n	p
		Online panel	**Social media**			
Prior to the COVID-19 pandemic	I was not aware	314 (62%)	74 (37%)	37.35	706	< 0.001
	Newspaper (online or print)	69 (14%)	34 (17%)	0.90		0.342
	News show (TV or online)	68 (13%)	27 (13%)	< 0.001		1.000
	Social media	34 (7%)	16 (8%)	0.15		0.698
	Internet search	49 (10%)	22 (11%)	0.11		0.743
	Academic reports and papers	53 (11%)	48 (24%)	19.57		< 0.001
	Other	36 (7%)	45 (22%)	31.05		< 0.001
During the COVID-19 pandemic	I was not aware	114 (23%)	6 (3%)	38.09		< 0.001
	Newspaper (online or print)	204 (40%)	123 (61%)	23.36		< 0.001
	News show (TV or online)	300 (60%)	142 (70%)	6.70		0.010
	Social media	147 (29%)	125 (62%)	63.79		< 0.001
	Internet search	122 (24%)	72 (36%)	8.90		0.003
	Academic reports and papers	86 (17%)	97 (48%)	70.37		< 0.001
	Other	19 (4%)	23 (11%)	13.62		< 0.001

is important to remember that most of the social media sample recorded being aware of transmission modelling, either generally or specifically in policy, during the pandemic period. As mentioned in the main text, these differences could be driven by a combination of a significantly greater percentage of social respondents working in the "Education" and "Research" sectors and by their relationship to the accounts used to share the survey (Table C.1).

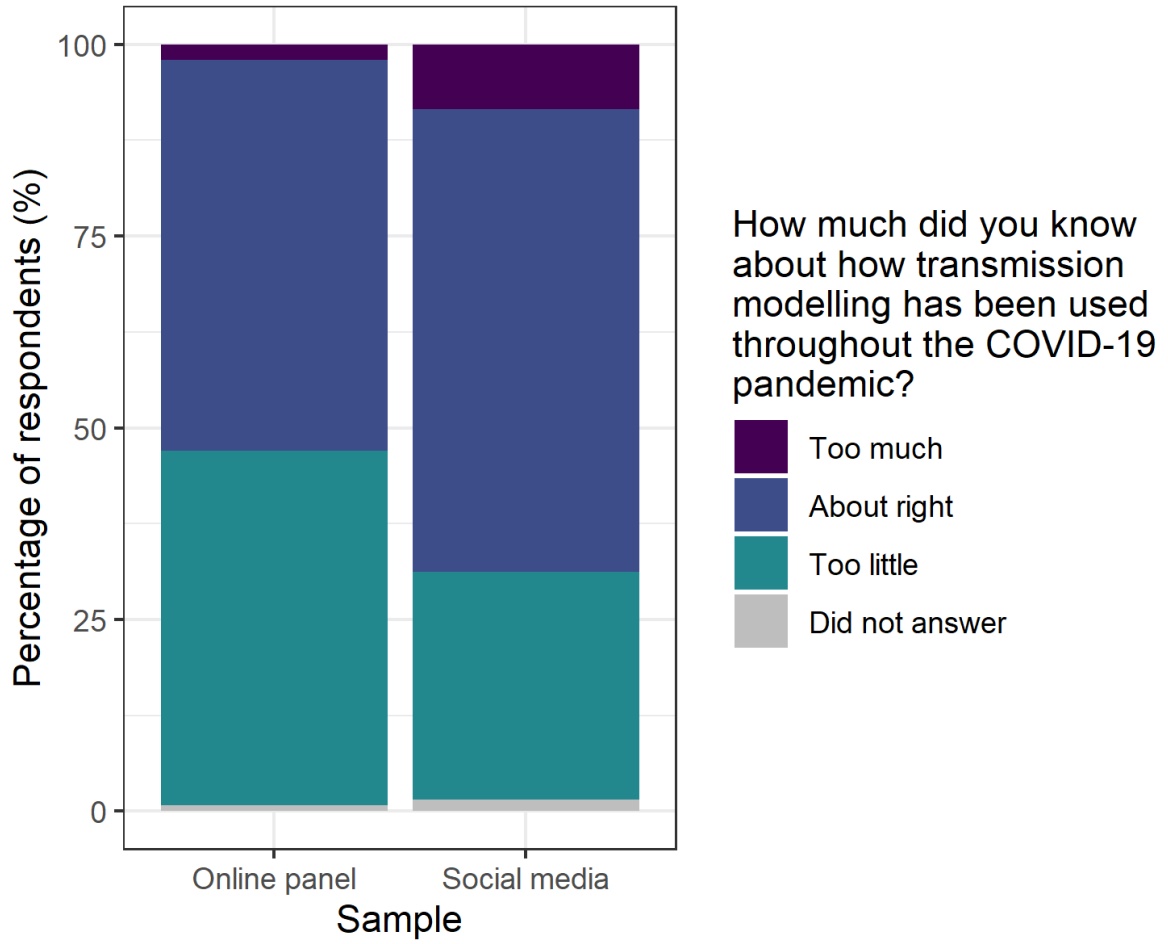

Figure C.2: Responses to the question "How much did you know about how transmission modelling has been used throughout the COVID-19 pandemic?" for the online panel and social media samples. Underlying data are presented in Table 5.1.

C.3.1.4 Level and means of awareness

Figure C.3 presents the results of "During the COVID-19 pandemic, how were you aware of transmission modelling?" stratified by responses to the question "How much do you know about how transmission modelling has been used throughout the COVID-19 pandemic?". The underlying data are presented in Table C.7.

Almost all social media respondents selecting having "Too much" information on modelling during the pandemic, selected "Newspapers (online or print)" and "Academic reports and papers" as means of their awareness, with "Social media" also being a common response (Table C.7; Figure C.3). Within the online panel, those with "Too much" information were primarily aware through "News shows" (Table C.7; Figure C.3).

Table C.6: Coefficient estimates, standard errors and p-values for the linear model regressing awareness of the use of transmission modelling during COVID-19 pandemic on age group (years) and gender within each model for each sample. Chi-squared tests for the overall significance of age and gender are also presented. The three responses provided as multiple-choice options, "Too little", "About right" and "Too much", were enumerated as -1, 0, 1, respectively. $p < 0.01$ is considered statistically significant. *Baseline comprised of "Female" and "18 - 25" years.

Sample	**Response**	**Variable**	**Interpretation**	**Estimate**	**Standard error**	p	**Chi-squared test**	
							χ^2	p
Online panel	How much do you know about how transmission modelling has been used throughout the COVID-19 pandemic?	Intercept	Baseline*	-0.52	0.07	< 0.001		
		Age (years)	26 - 35	-0.00	0.09	0.999	1.20 (df=5)	0.530
			36 - 45	0.06	0.09	0.508		
			46 - 55	0.15	0.09	0.094		
			56 - 65	0.04	0.08	0.655		
			66+	0.08	0.11	0.444		
		Gender	Male	0.05	0.05	0.340	0.63 (df=3)	0.535
			Non-binary	0.52	0.54	0.338		
			Prefer not to say	-0.18	0.31	0.567		
Social media	How much do you know about how transmission modelling has been used throughout the COVID-19 pandemic?	Intercept	Baseline*	-0.04	0.29	0.901		
		Age (years)	26 - 35	0.41	0.34	0.232	4.63 (df=5)	0.018
			36 - 45	-0.25	0.31	0.433		
			46 - 55	-0.17	0.30	0.560		
			56 - 65	-0.28	0.30	0.341		
			66+	-0.25	0.31	0.414		
		Gender	Male	0.05	0.08	0.555	0.12 (df=2)	0.834
			Prefer not to say	-0.01	0.34	0.966		

C.3.2 Reliability of transmission modelling in informing public health policy

C.3.2.1 On a scale of 1 - 10, with 1 being "extremely unreliable" and 10 being "extremely reliable" how did you feel about the use of transmission modelling in informing public health policy?

Figure C.4 presents histograms of the reliability scores within each sample across both time periods.

Figure C.5 is analogous to Figure 5.3A but presents each of the 10 possible reliability scores individually rather than collapsing into 5 categories of 2.

Table C.8 presents the number and percentage of respondents within each sample changing reliability score in the period of the pandemic compared to the period prior (reported retrospectively).

Table C.9 presents the results of the linear models regressing reliability score on age and gender, per sample and in reference to each time period.

Table C.10 presents the results of the linear models regressing reliability score on age, gender, and

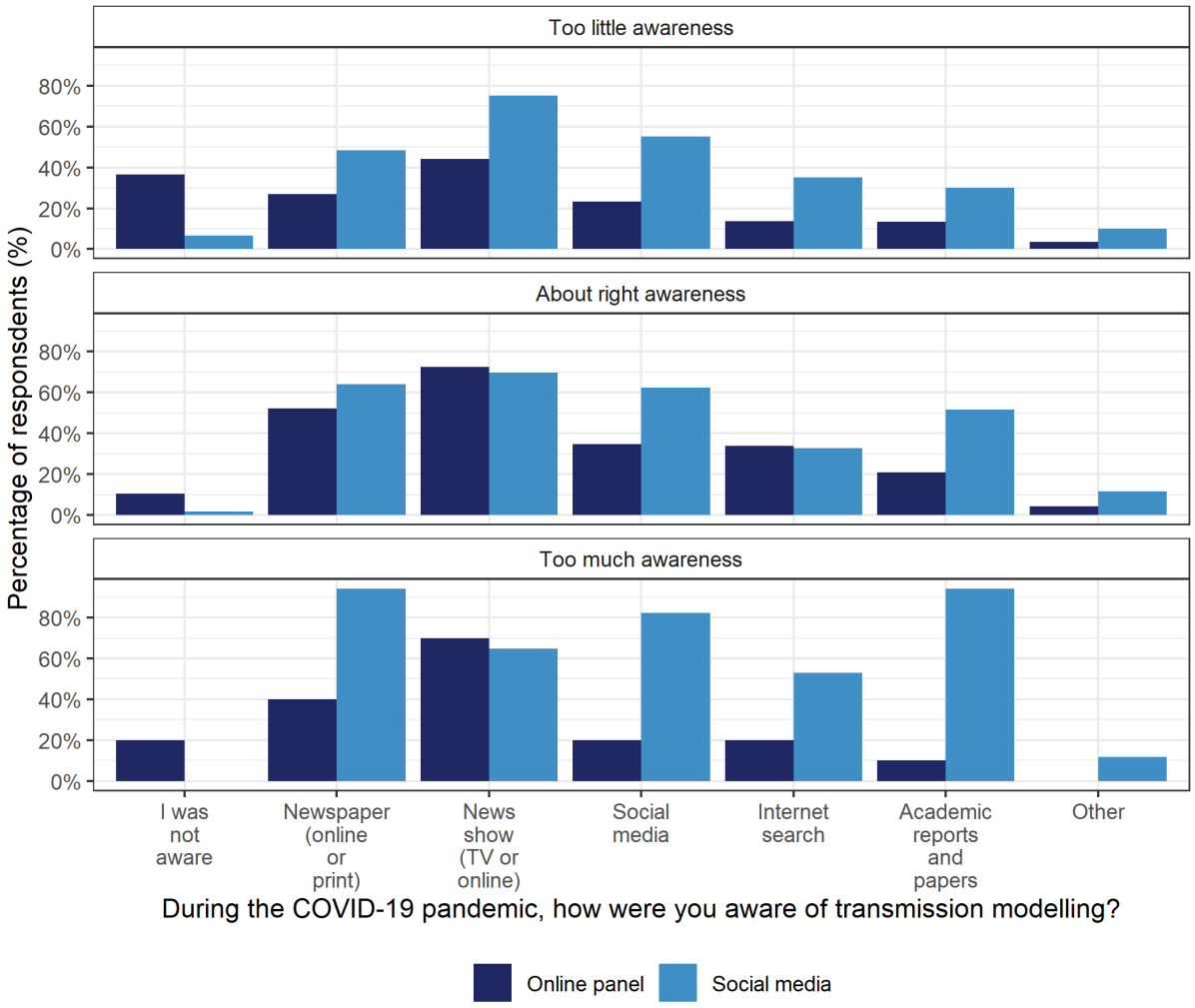

Figure C.3: Responses to the question "During the COVID-19 pandemic, how were you aware of transmission modelling?" stratified by responses to the question "How much did you know about how transmission modelling has been used throughout the COVID-19 pandemic?" for both the online panel and social media samples. Percentages are taken with respect to sample platform and level of awareness. Underlying data are presented in Table C.7.

awareness of the use of modelling in informing public health policy per sample and in reference to each time period.

Figure C.6 and Table C.11 present the relationship between reliability scores prior to (reported retrospectively) and during the COVID-19 pandemic under each sample.

Table C.12 presents a linear model regressing the difference in reliability scores on age group and gender. Difference in reliability score is defined as the score given during the pandemic minus the score given prior.

Table C.7: Number and percentage of responses under each category of the question "During the COVID-19 pandemic, how were you aware of transmission modelling?" stratified by responses to the question "How much did you know about how transmission modelling has been used throughout the COVID-19 pandemic?" for both the online panel and social media samples. As this question was open-ended, participants often provided more than one answer and so percentages do not represent mutually exclusive responses. Percentages taken with respect to the number selecting each awareness category within each sample as in Table 5.1.

	Online panel			**Social media**		
	Too little awareness	**About right awareness**	**Too much awareness**	**Too little awareness**	**About right awareness**	**Too much awareness**
I was not aware	85 (36%)	27 (11%)	2 (20%)	4 (7%)	2 (2%)	0 (0%)
Newspaper (online or print)	63 (27%)	134 (52%)	4 (40%)	29 (48%)	78 (64%)	16 (94%)
News show (TV or online)	103 (44%)	186 (72%)	7 (70%)	45 (75%)	85 (70%)	11 (65%)
Social media	54 (23%)	89 (35%)	2 (20%)	33 (55%)	76 (62%)	14 (82%)
Internet search	32 (14%)	87 (34%)	2 (20%)	21 (35%)	40 (33%)	9 (53%)
Academic reports and papers	31 (13%)	54 (21%)	1 (10%)	18 (30%)	63 (52%)	16 (94%)
Other	8 (3%)	11 (4%)	0 (0%)	6 (10%)	14 (11%)	2 (12%)

Table C.8: Number and percentage of respondents within each sample classified according to differences in reliability scores prior to and during the pandemic.

		Online panel	**Social media**
Number of respondents (n (%))	Reliability score increased	203 (40%)	79 (39%)
	Reliability score unchanged	217 (43%)	70 (35%)
	Reliability score decreased	76 (15%)	52 (26%)
	At least one score was missing	8 (2%)	1 (0%)

C.3.2.2 Reliability scores stratified by awareness

Figure C.7 presents reliability scores for both samples and time periods, stratified by responses to the question "Were you aware of transmission modelling being used in informing public health policy?". There were statistically significant differences in reliability scores among those with awareness compared to those with no awareness among online panel respondents, holding time period constant, (Mann-Whitney U test: online panel and prior $p < 0.001$; online panel and during $p < 0.001$), but not among social media respondents (Mann-Whitney U test: social media and prior $p = 0.197$; social media and during $p = 0.148$).

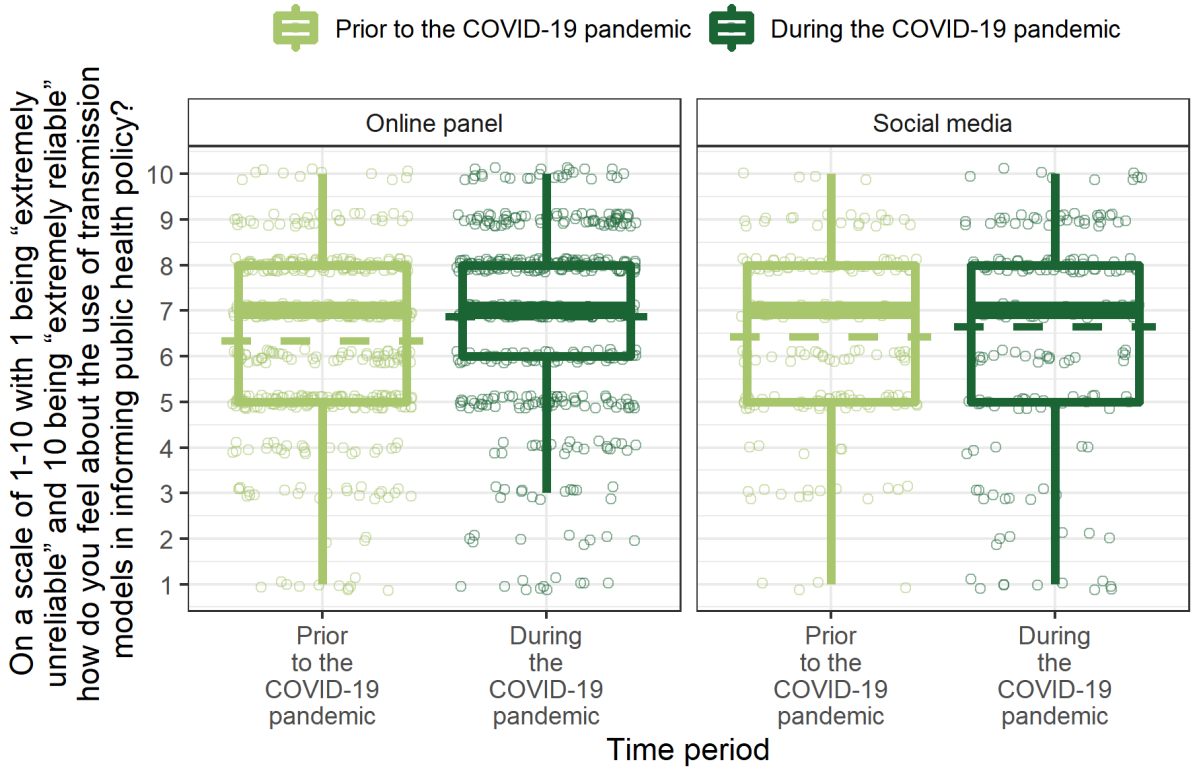

Figure C.4: Reliability of using transmission modelling to inform public health policy. (A) Responses to the question "On a scale of 1 – 10 with 1 being "extremely unreliable" and 10 being "extremely reliable" how do you feel about the use of transmission modelling in informing public health policy?" both "Prior to" and "During" the COVID19 pandemic, for both the online panel and social media samples. From bottom to top, the solid lines on the boxplot indicate: 1.5 times the interquartile range (IQR) less than the first quartile, first quartile, median, third quartile and 1.5 times the IQR greater than the third quartile. The dashed line corresponds to the mean. All responses are shown by the points, and so outliers, defined as any point outside the lower and upper bounds described, have been removed from the boxplots as they are shown in the presentation of the data. Points represent each reliability score and have been jittered to aid visual presentation. Underlying data are presented in Table 5.1.

Figure C.8 presents reliability scores according to responses to the question "How much do you know about how transmission modelling has been used throughout the COVID-19 pandemic?" for both samples. Online panel respondents who recorded an "About right" level of knowledge on the use of transmission modelling during the pandemic had significantly higher reliability scores than those stating they had "Too little" knowledge (Mann-Whitney U test: $p < 0.001$), but this was not a significant factor among social media respondents (Mann-Whitney U test: $p = 0.396$) (Figure C.7). The few respondents recording they had "Too much" knowledge had significantly lower reliability scores than those with "About right" knowledge within both samples (Mann-Whitney U test: Online panel $p = 0.002$; social media $p < 0.001$).

Figure C.9 presents reliability scores according to responses to the question "How were you aware of transmission modelling?" both "Prior to" and "During" the COVID-19 pandemic, for both samples. There was little distinction in reliability scores according to the means by which respondents were aware

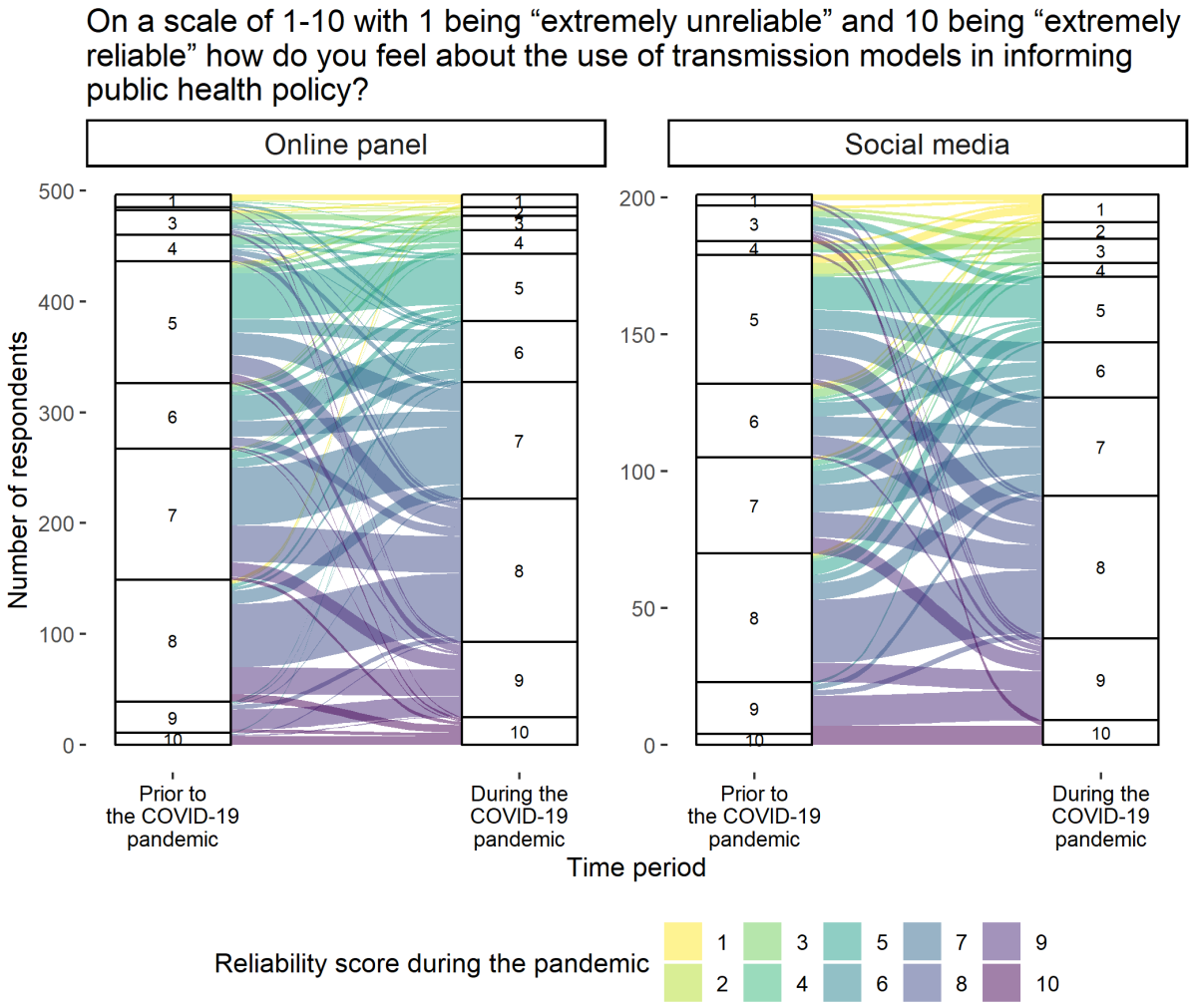

Figure C.5: Responses to the question “On a scale of 1 – 10 with 1 being “extremely unreliable” and 10 being “extremely reliable” how do you feel about the use of transmission modelling in informing public health policy?” both “Prior to” (reported retrospectively) and “During” the COVID-19 pandemic, for both the online panel and social media samples. Observations corresponding to a respondent which did not answer the question for at least one time point were removed from the figure to aid visual presentation ($n = 8$ online panel; $n = 1$ social media). Due to small numbers of respondents from the online panel selecting a score of “2”, this is unable to be labelled, but appears in numeric order. Underlying data are presented in Table 1.

of modelling, with scores distributed across the entire scale under each option within each sample.

Table C.9: Coefficient estimates, standard errors and p-values for the linear regression models with age group and gender as predictors of reliability scores during each period and for each sample. Chi-squared tests for the overall significance of age and gender are also presented. $p < 0.01$ is considered statistically significant. *Baseline comprised of "Female" and "18 - 25" years.

Sample	**Response**	**Variable**	**Interpretation**	**Estimate**	**Standard error**	p	**Chi-squared test**	
							χ^2	p
Online panel	Reliability score prior to the COVID-19 pandemic	Intercept	Baseline*	6.66	0.24	< 0.001		
		Age (years)	26 - 35	-0.07	0.29	0.815	12.99 (df=5)	0.567
			36 - 45	-0.27	0.29	0.351		
			46 - 55	-0.48	0.29	0.106		
			56 - 65	-0.33	0.27	0.225		
			66+	-0.08	0.36	0.816		
		Gender	Male	-0.20	0.17	0.234	10.19 (df=3)	0.386
			Non-binary	0.41	1.84	0.824		
			Prefer not to say	1.23	1.06	0.247		
Social media	Reliability the score prior to the COVID-19 pandemic	Intercept	Baseline*	6.95	0.96	< 0.001		
		Age (years)	26 - 35	-0.89	1.12	0.426	4.66 (df=5)	0.933
			36 - 45	-0.36	1.03	0.728		
			46 - 55	-0.39	0.98	0.692		
			56 - 65	-0.24	0.97	0.807		
			66+	-0.60	1.02	0.554		
		Gender	Male	-0.27	0.27	0.319	14.09 (df=2)	0.141
			Prefer not to say	-2.05	1.12	0.069		
Online panel	Reliability score during the COVID-19 pandemic	Intercept	Baseline*	7.12	0.25	< 0.001		
		Age (years)	26 - 35	-0.13	0.32	0.674	24.22 (df=5)	0.290
			36 - 45	-0.26	0.31	0.400		
			46 - 55	-0.63	0.32	0.049		
			56 - 65	-0.00	0.30	0.994		
			66+	-0.18	0.39	0.643		
		Gender	Male	-0.11	0.18	0.526	3.18 (df=3)	0.846
			Non-binary	0.01	1.99	0.994		
			Prefer not to say	0.68	1.15	0.554		
Social media	Reliability score during the COVID-19 pandemic	Intercept	Baseline*	7.84	1.11	< 0.001		
		Age (years)	26 - 35	-2.87	1.29	0.027	59.66 (df=5)	0.031
			36 - 45	-1.28	1.18	0.281		
			46 - 55	-0.53	1.13	0.637		
			56 - 65	-0.58	1.12	0.607		
			66+	-0.62	1.17	0.598		
		Gender	Male	-0.78	0.31	0.013	65.70 (df=2)	0.001
			Prefer not to say	-3.93	1.29	0.003		

C.3.3 Responsibility of communicating modelling to the public

Figure C.10 presents the responses to the question "Who do you think has the responsibility of ensuring that the public are informed about the use of modelling in policy decisions, particularly in the COVID-19 pandemic?" for both samples.

Table C.13 presents the results of chi-squared tests for statistically significant differences in the proportion of respondents selecting multiple-choice answers to the question "Who do you think has the responsibility of ensuring that the public are informed about the use of modelling in policy decisions,

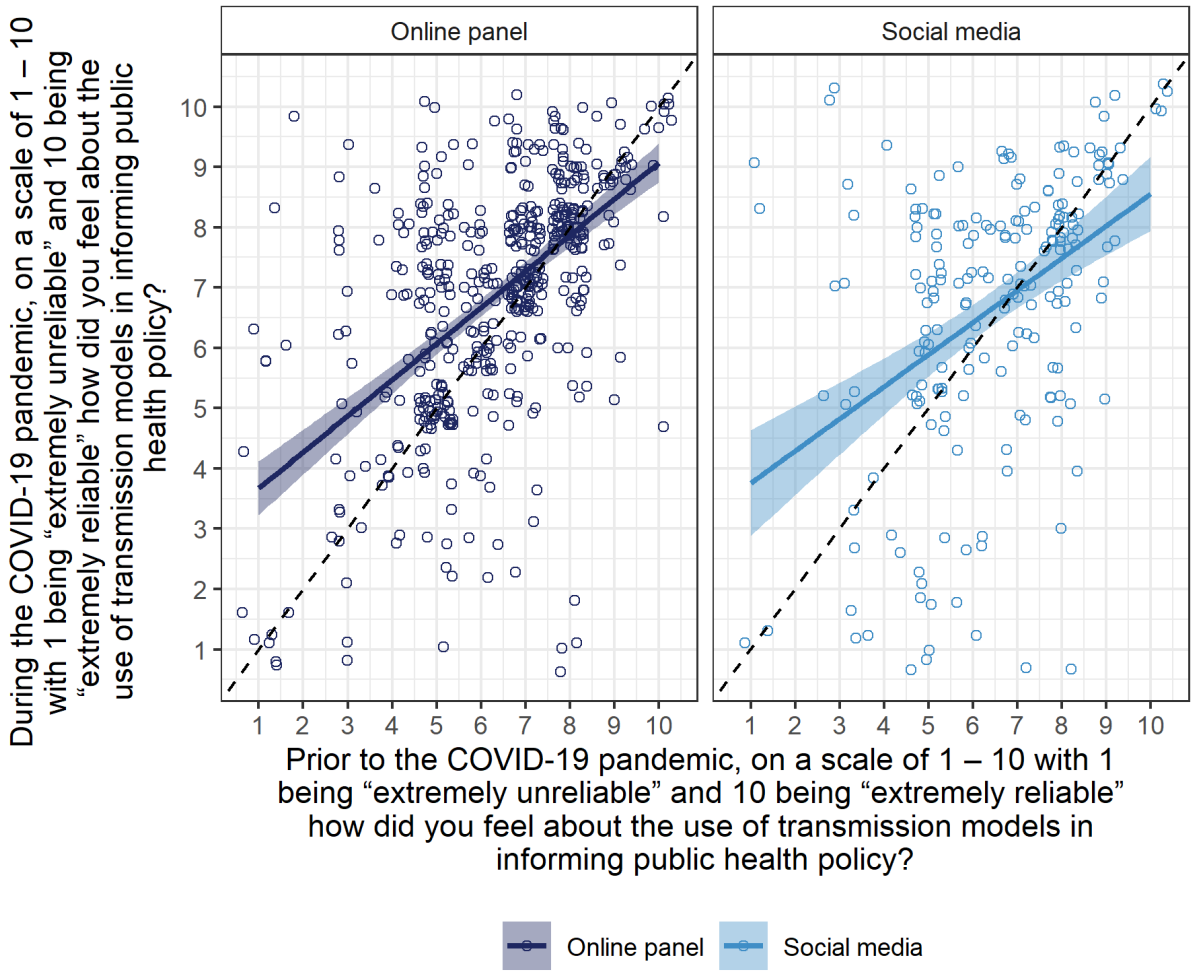

Figure C.6: : Relationship between reliability scores given "Prior to" and "During" the COVID-19 pandemic for both the online panel and social media samples. Reliability scores obtained via the question: "On a scale of 1 – 10 with 1 being "extremely unreliable" and 10 being "extremely reliable" how did you feel about the use of transmission models in informing public health policy?". Points represent each reliability score and have been jittered to aid visual presentation. Coloured lines show the linear model presented in Table C.11, with shaded areas representing 95% confidence intervals. Black dashed line indicates where reliability scores during the pandemic equal those prior to the pandemic.

particularly in the COVID-19 pandemic?" between the two samples.

Figure C.11 presents the results to the aforementioned question stratified by answers to the question "How much did you know about how transmission modelling has been used throughout the COVID-19 pandemic?". The robustness of these results to awareness suggests that this is not an influencing factor.

Of the few who selected "None of the above" to the aforementioned question, of which there were one from the online panel and two from the social media sample, respondents were either unsure (online panel), believed in a "collaborative" (social media) approach or the "NHS" (social media).

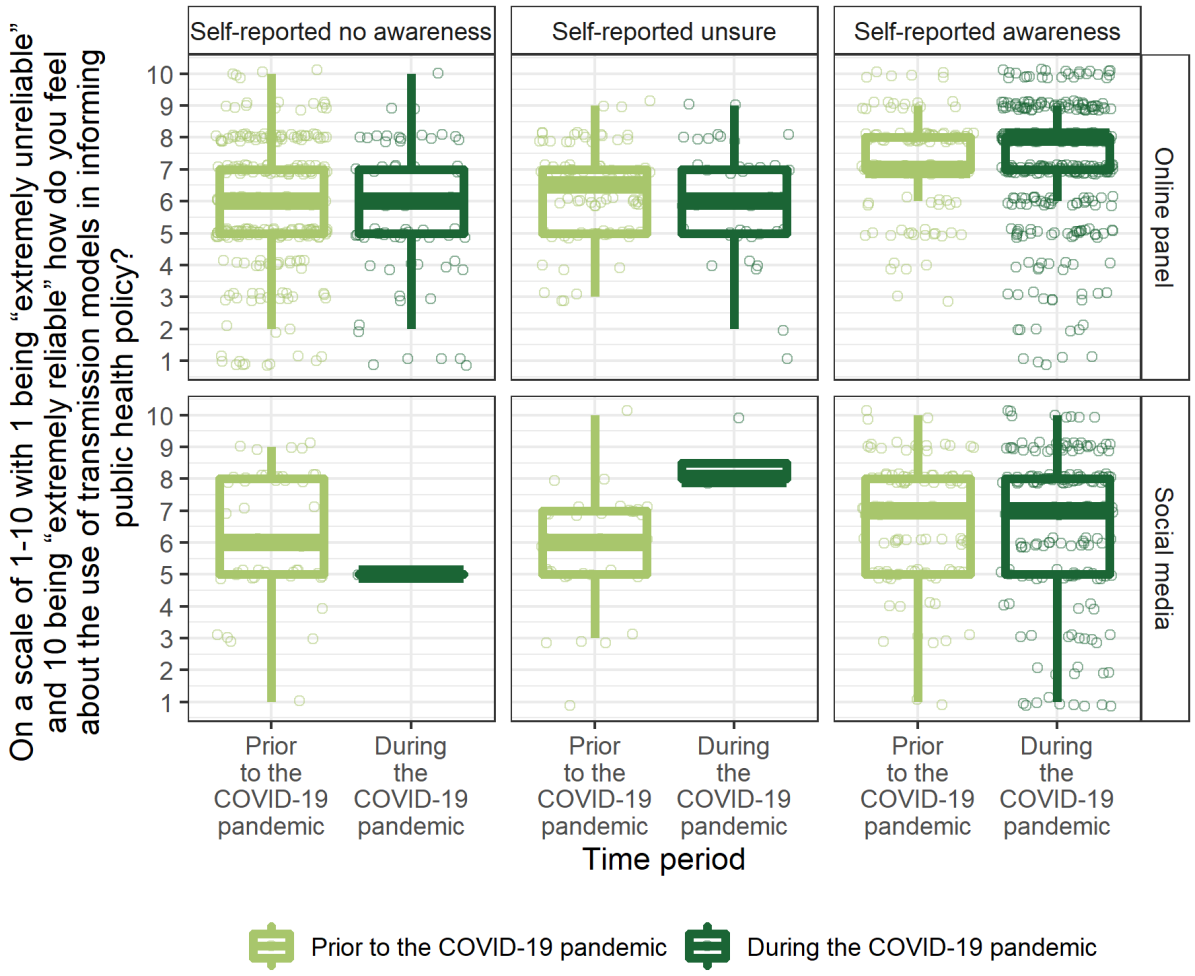

Figure C.7: Responses to the question "On a scale of 1 – 10 with 1 being "extremely unreliable" and 10 being "extremely reliable" how do you feel about the use of transmission modelling in informing public health policy?" both "Prior to" and "During" the COVID-19 pandemic, for both the online panel and social media samples, stratified by self-reported level of awareness according to the question "Were you aware of the use of transmission models in informing public health policy?" both "Prior to" and "During" the COVID-19 pandemic. From bottom to top, the solid lines on the boxplot indicate: 1.5 times the interquartile range (IQR) less than the first quartile, first quartile, median, third quartile and 1.5 times the IQR greater than the third quartile. All responses are shown by the points, and so outliers, defined as any point outside the lower and upper bounds described, have been removed from the boxplots as they are shown in the presentation of the data. Points represent each reliability score and have been jittered to aid visual presentation.

C.3.4 Trust in government advice

C.3.4.1 How much did you trust government advice regarding public health issues?

Figure C.12 presents reliability scores stratified by answers to question "How much did you trust government advice regarding public health issues?" both "Prior to" and "During" the COVID-19 pandemic for both samples. Table C.14 present the median and variances of the reliability scores for the same stratifications.

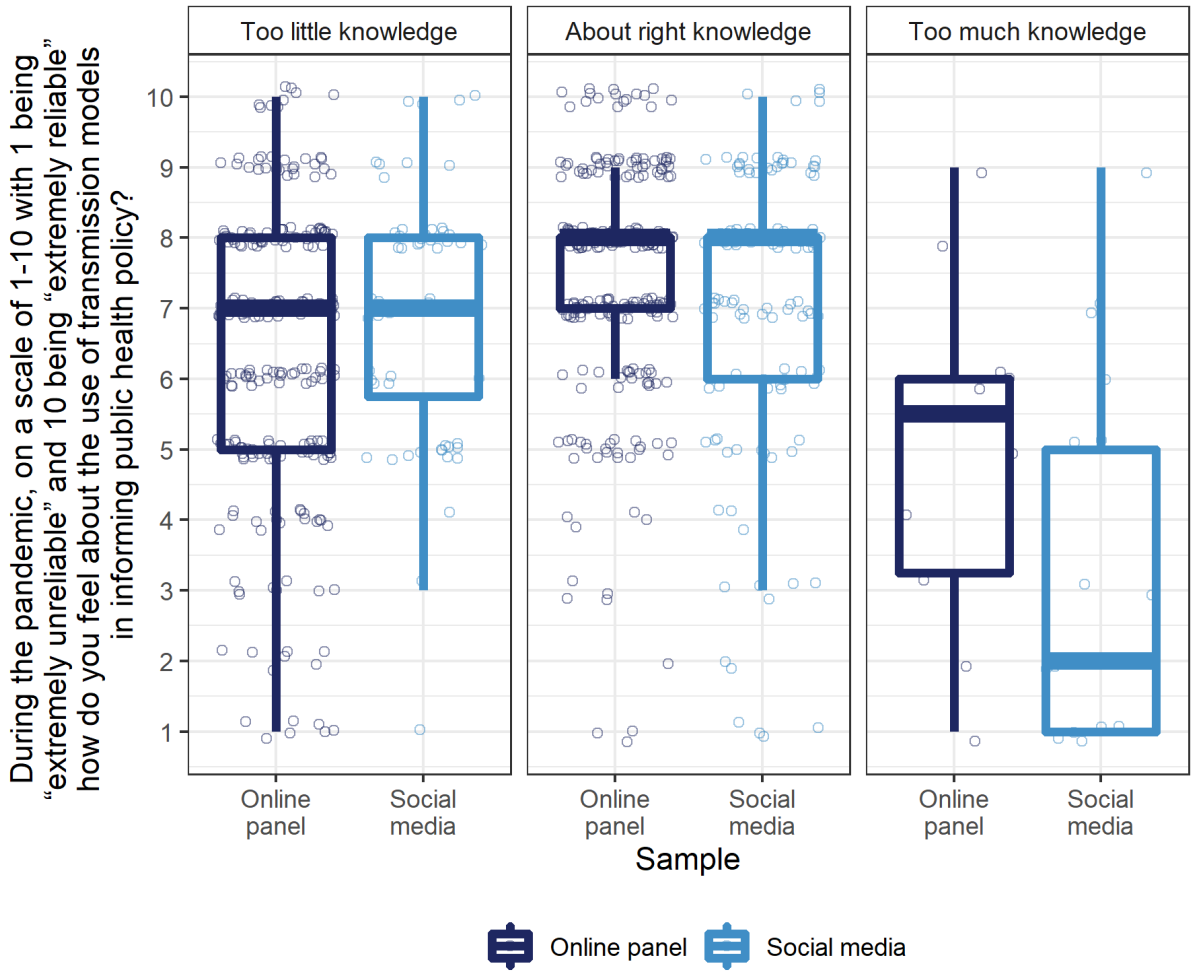

Figure C.8: Responses to the question "During the COVID-19 pandemic, on a scale of 1 – 10 with 1 being "extremely unreliable" and 10 being "extremely reliable" how do you feel about the use of transmission modelling in informing public health policy?" for both the online panel and social media samples, stratified by self-reported level of awareness according to the question "How much do you know about how transmission modelling has been used throughout the COVID-19 pandemic?". From bottom to top, the solid lines on the boxplot indicate: 1.5 times the interquartile range (IQR) less than the first quartile, first quartile, median, third quartile and 1.5 times the IQR greater than the third quartile. The dashed line corresponds to the mean. All responses are shown by the points, and so outliers, defined as any point outside the lower and upper bounds described, have been removed from the boxplots as they are shown in the presentation of the data. Points represent each reliability score and have been jittered to aid visual presentation.

Table C.15 presents the number and percentage of respondents within each sample classified according to differences in responses to the aforementioned question across time periods.

Table C.16 presents the results of the corresponding linear model regressing the responses to the question "How much did you trust government advice regarding public health issues?" both "Prior to" and "During" the COVID-19 pandemic on age group and gender of respondents within each sample. The three levels of trust provided in the multiple-choice options "No level of trust whatsoever", "Moderate level of

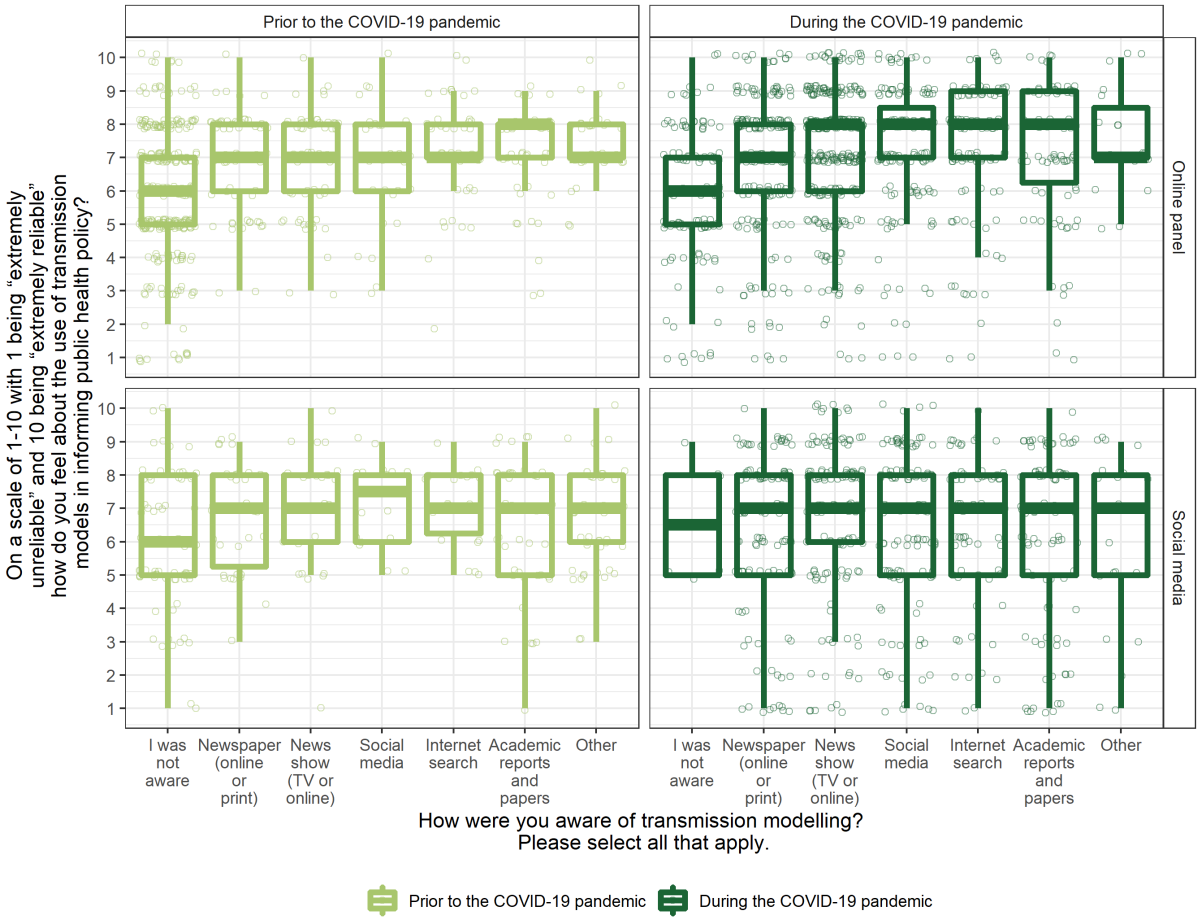

Figure C.9: Responses to the question "On a scale of 1 – 10 with 1 being "extremely unreliable" and 10 being "extremely reliable" how do you feel about the use of transmission modelling in informing public health policy?" both "Prior to" and "During" the COVID-19 pandemic, for both the online panel and social media samples, stratified by responses to the question "How were you aware of transmission modelling?" both "Prior to" and "During" the COVID-19 pandemic. From bottom to top, the solid lines on the boxplot indicate: 1.5 times the interquartile range (IQR) less than the first quartile, first quartile, median, third quartile and 1.5 times the IQR greater than the third quartile. The dashed line corresponds to the mean. All responses are shown by the points, and so outliers, defined as any point outside the lower and upper bounds described, have been removed from the boxplots as they are shown in the presentation of the data. Points represent each reliability score and have been jittered to aid visual presentation.

trust" and "High level of trust" were enumerated as -1, 0, 1 respectively.

Table C.17 presents the results of the linear model regressing changes in responses to the question "How much did you trust government advice regarding public health issues?" in the period of the pandemic compared to the period prior, on age group and gender within each sample. The three responses provided in the multiple-choice options, "No trust whatsoever", "Moderate level of trust" and "High level of trust", were enumerated as -1, 0, 1, respectively, and so the differences (response) could take values of -2, -1, 0, 1, 2.

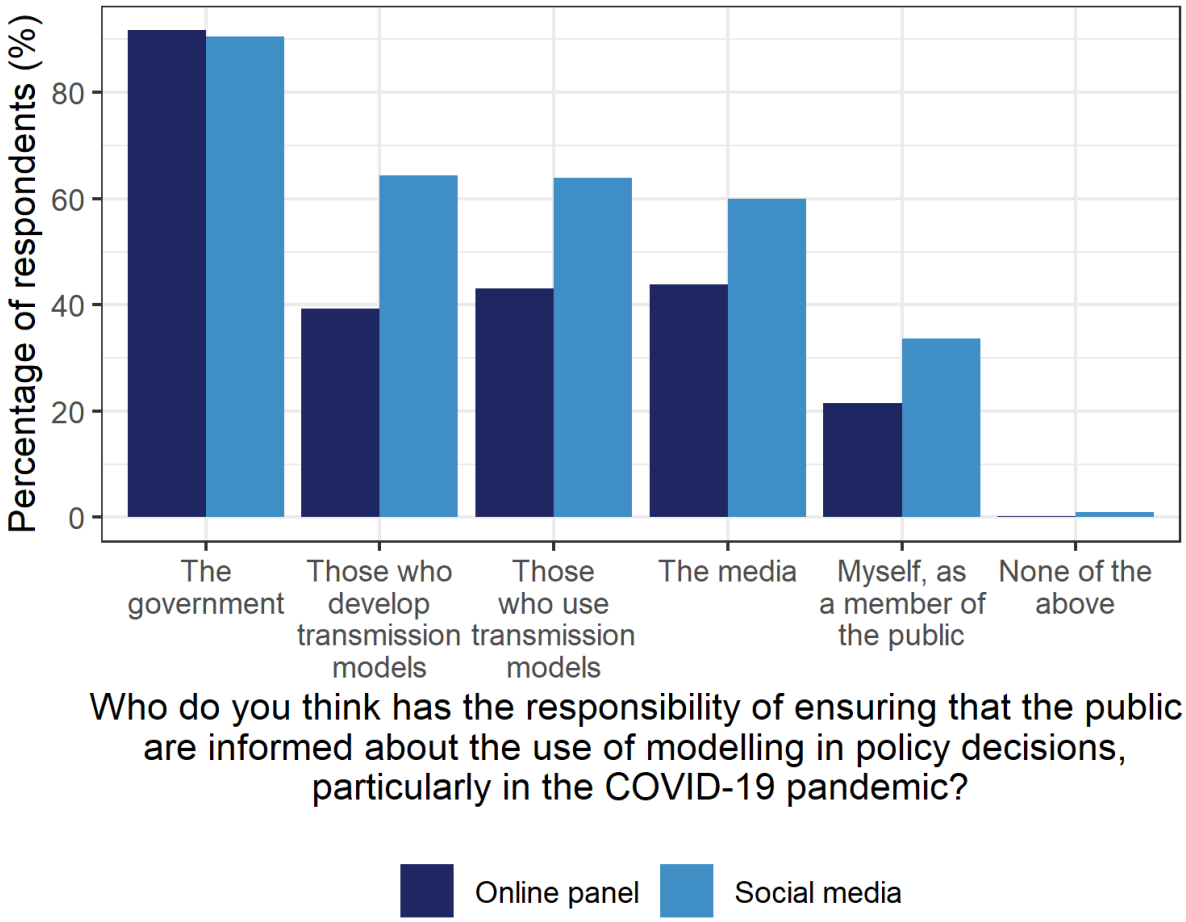

Figure C.10: Responses to the question "Who do you think has the responsibility of ensuring that the public are informed about the use of modelling in policy decisions, particularly in the COVID-19 pandemic? Please select all that apply." for both samples. Underlying data are presented in Table 5.1.

C.3.4.2 Trust in government advice and awareness

Table C.18 presents the results of the linear model regressing responses to the question "How much did you trust government advice regarding public health issues?" on age group, gender and awareness of the use of modelling in policy during each period and for each sample. The three responses provided in the multiple-choice options for the response, "No trust whatsoever", "Moderate level of trust" and "High level of trust", were enumerated as -1, 0, 1, respectively.

Similarly, the responses to the question "Were you aware of transmission modelling in informing public health policy?" used as an explanatory variable, "No", "Unsure" and "Yes" were enumerated as -1, 0, 1, respectively. Data to the response question stratified by awareness are presented in Figure C.13, with the percentages underlying this figure set out in Table C.19.

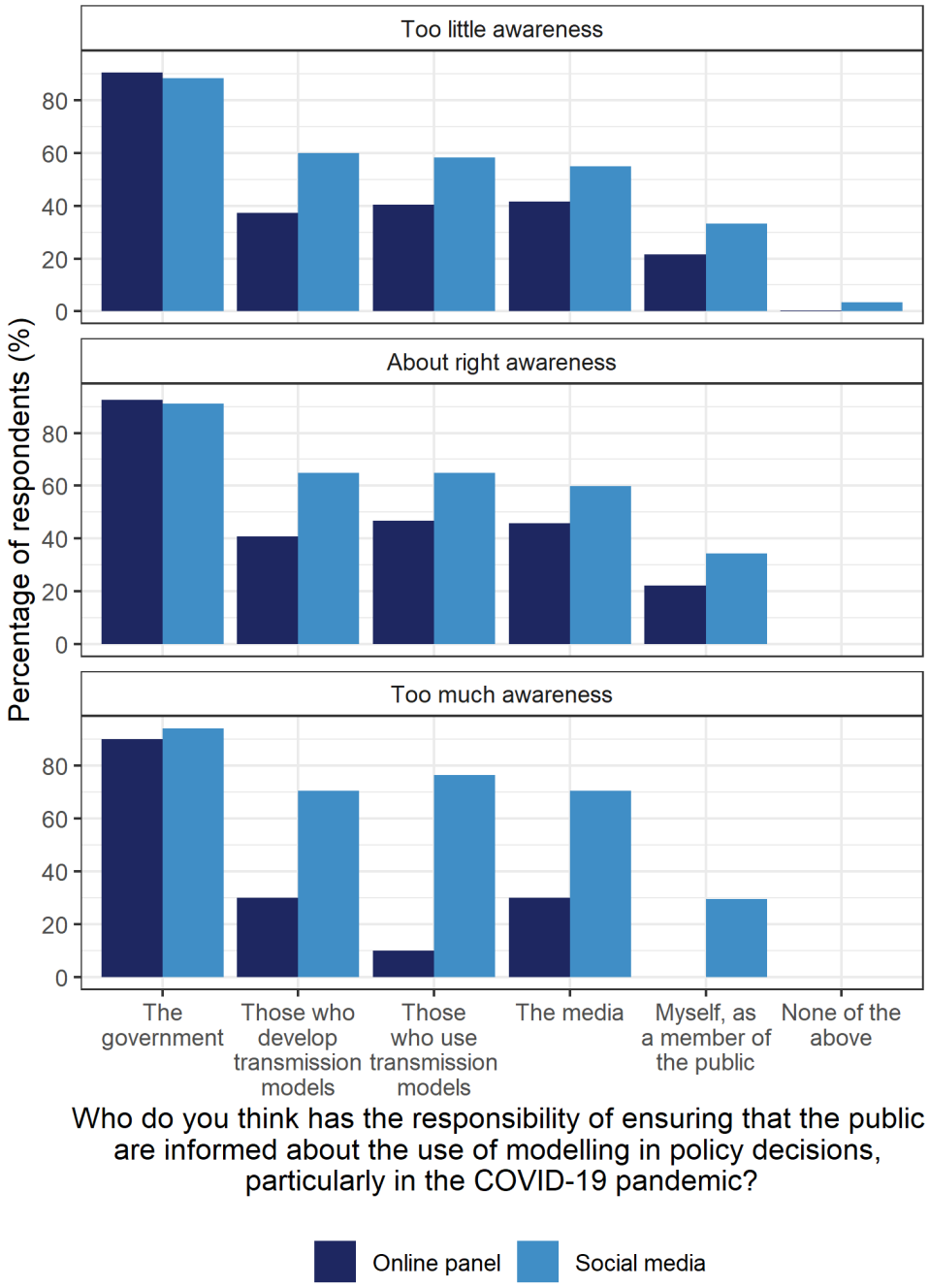

Figure C.11: Responses to the question "Who do you think has the responsibility of ensuring that the public are informed about the use of modelling in policy decisions, particularly in the COVID-19 pandemic? Please select all that apply." for both the online panel and social media samples, stratified by answers to the question "How much do you know about how transmission modelling has been used throughout the COVID-19 pandemic?".

C.3.4.3 How do you feel when government advice changes based on new scientific evidence?

Table C.20 presents the results of the corresponding linear model regressing the responses to the question "How do you feel when government advice changes based on new scientific evidence?" on gender and age group of respondents within each sample. The three levels of trust provided in the multiple-choice

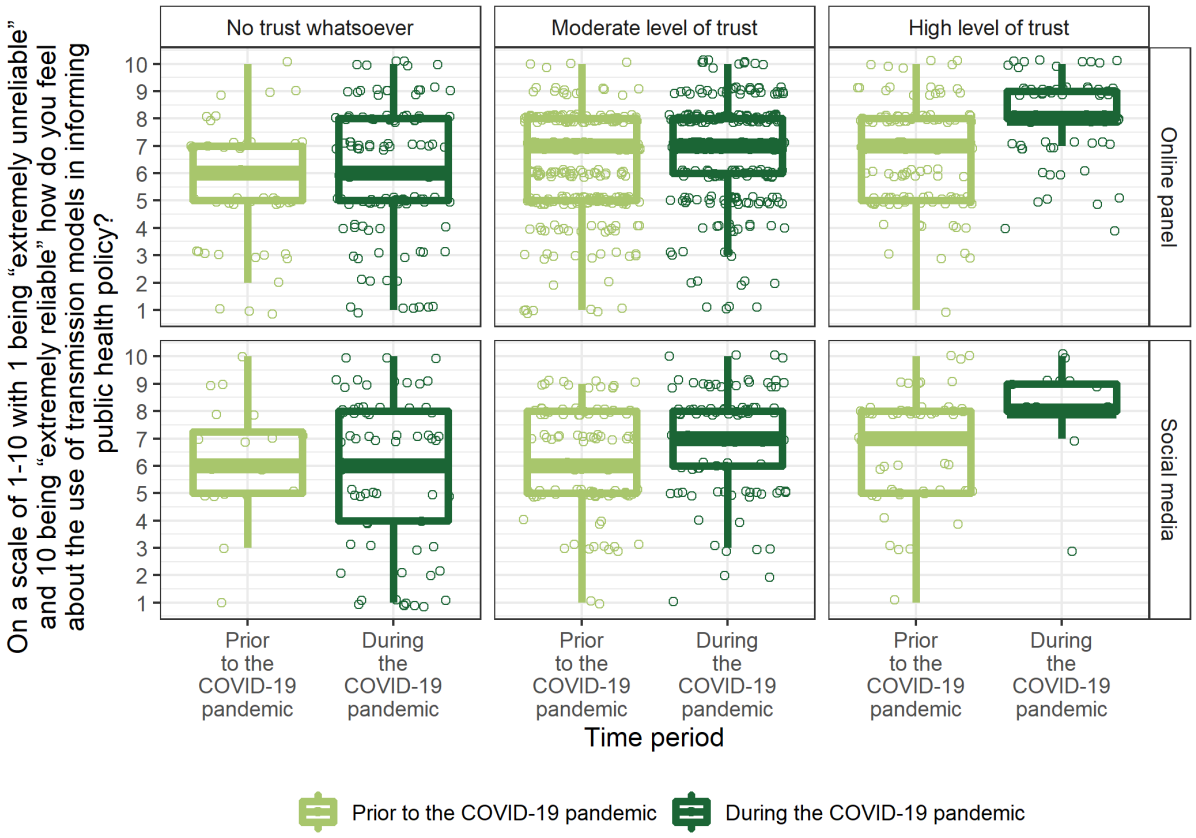

Figure C.12: Responses to the question "On a scale of 1 – 10 with 1 being "extremely unreliable" and 10 being "extremely reliable" how do you feel about the use of transmission modelling in informing public health policy?" both "Prior to" and "During" the COVID-19 pandemic, for both the online panel and social media samples, stratified by responses to the question "How much did you trust government advice regarding public health issues?" both "Prior to" and "During" the COVID-19 pandemic. From bottom to top, the solid lines on the boxplot indicate: 1.5 times the interquartile range (IQR) less than the first quartile, first quartile, median, third quartile and 1.5 times the IQR greater than the third quartile. The dashed line corresponds to the mean. All responses are shown by the points, and so outliers, defined as any point outside the lower and upper bounds described, have been removed from the boxplots as they are shown in the presentation of the data. Points represent each reliability score and have been jittered to aid visual presentation.

options "I have less trust in the advice", "My level of trust remains unchanged" and "I have more trust in the advice" were enumerated as -1, 0, and 1, respectively.

Table C.21 expands the regression presented in Table C.20 but also controls for responses to the question "Were you aware of the use of transmission models in informing public health policy?" both "Prior to" and "During" the COVID-19 pandemic. Responses to this question, "Yes", "Unsure" and "No" were enumerated as 1, 0, and -1, respectively.

Analogously to Table C.21, Table C.22 expands the regression in Table C.20 but also controls for responses to the question "How much did you trust government scientific advice regarding public health

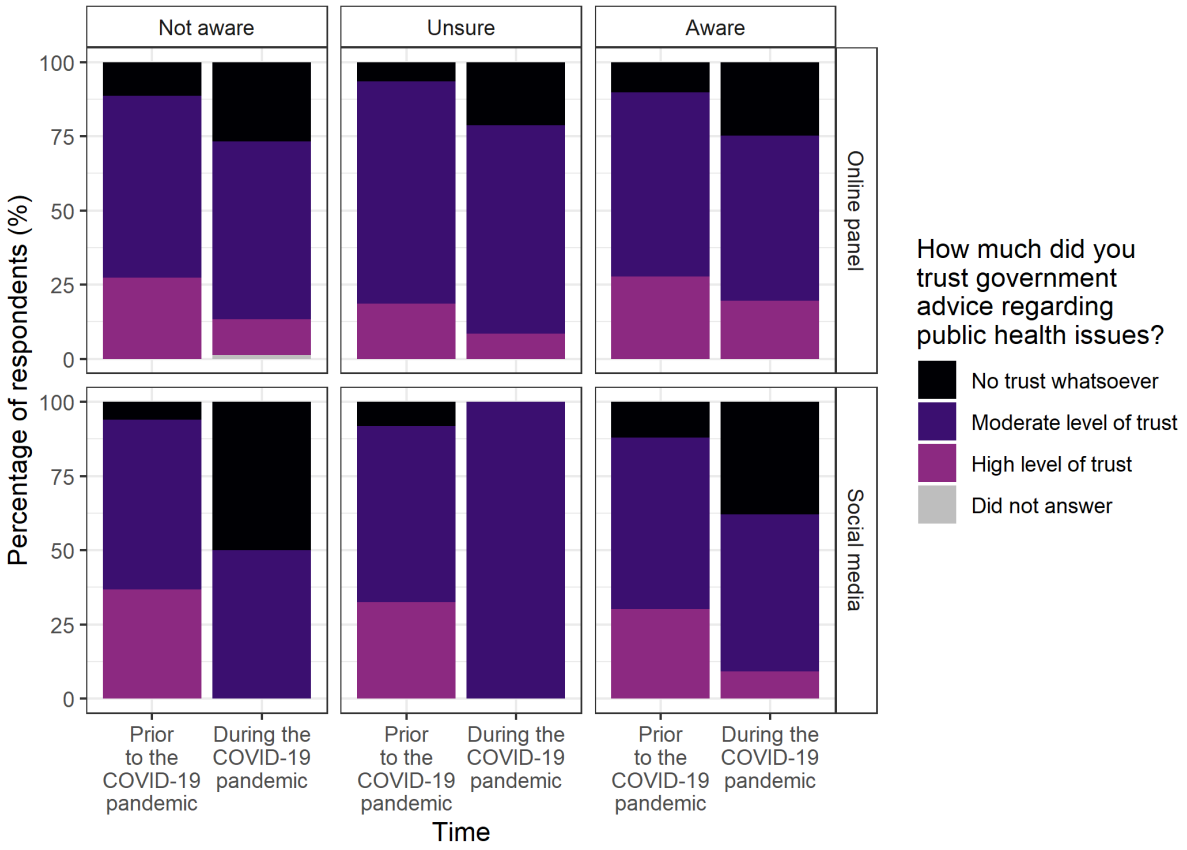

Figure C.13: The percentage of observations from both the online panel and social samples according to time period and answer to the question "How much did you trust government advice regarding public health issues?" both "Prior to" and "During" the COVID-19 pandemic, stratified by responses to the question "Were you aware of the use of transmission models in informing public health policy?" both "Prior to" and "During" the COVID-19 pandemic. Underlying data are presented in Table C.19.

policy?" both "Prior to" and "During" the COVID-19 pandemic. Responses to this question, "High level of trust", "Moderate level of trust" and "No trust whatsoever" were enumerated as 1, 0, and -1, respectively.

C.3.4.4 Trust in changing advice and awareness

Figure C.14 presents reliability scores for both samples and time periods stratified by responses to the question "How do you feel when government advice changes based on new scientific evidence?".

Figure C.15 presents the results of the question "How do you feel when government advice changes based on new scientific evidence?" stratified by response to the question "How much did you trust government advice regarding public health issues?" both "Prior to" and "During" the COVID-19 pandemic and under both samples. Table C.23 presents the data underlying this figure.

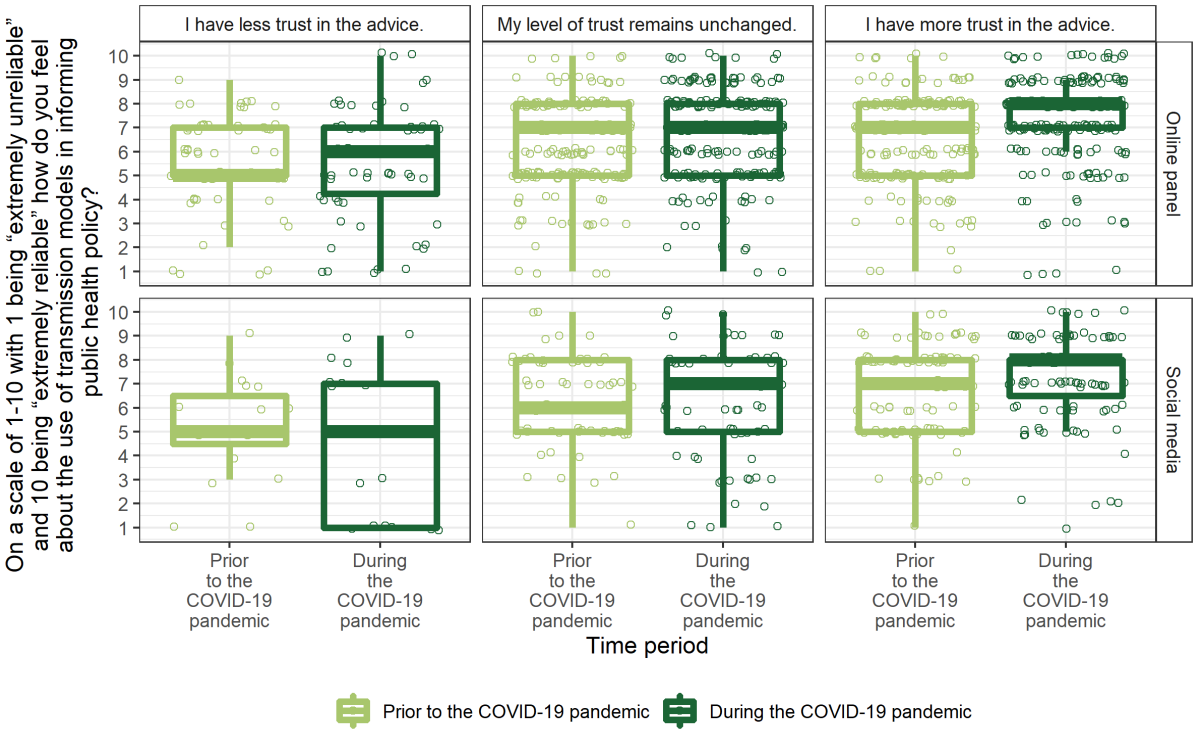

Figure C.14: Responses to the question "On a scale of 1 – 10 with 1 being "extremely unreliable" and 10 being "extremely reliable" how do you feel about the use of transmission modelling in informing public health policy?" both "Prior to" and "During" the COVID-19 pandemic, for both the online panel and social media samples, stratified by responses to the question "How much do you feel when government advice changes based on new scientific evidence?". From bottom to top, the solid lines on the boxplot indicate: 1.5 times the interquartile range (IQR) less than the first quartile, first quartile, median, third quartile and 1.5 times the IQR greater than the third quartile. The dashed line corresponds to the mean. All responses are shown by the points, and so outliers, defined as any point outside the lower and upper bounds described, have been removed from the boxplots as they are shown in the presentation of the data. Points represent each reliability score and have been jittered to aid visual presentation.

C.3.5 Where do those who develop transmission models work?

Figure C.16 presents responses to the question "Where do you think those who developed and used transmission models work?" both "Prior to" and "During" the COVID-19 pandemic for both samples. The

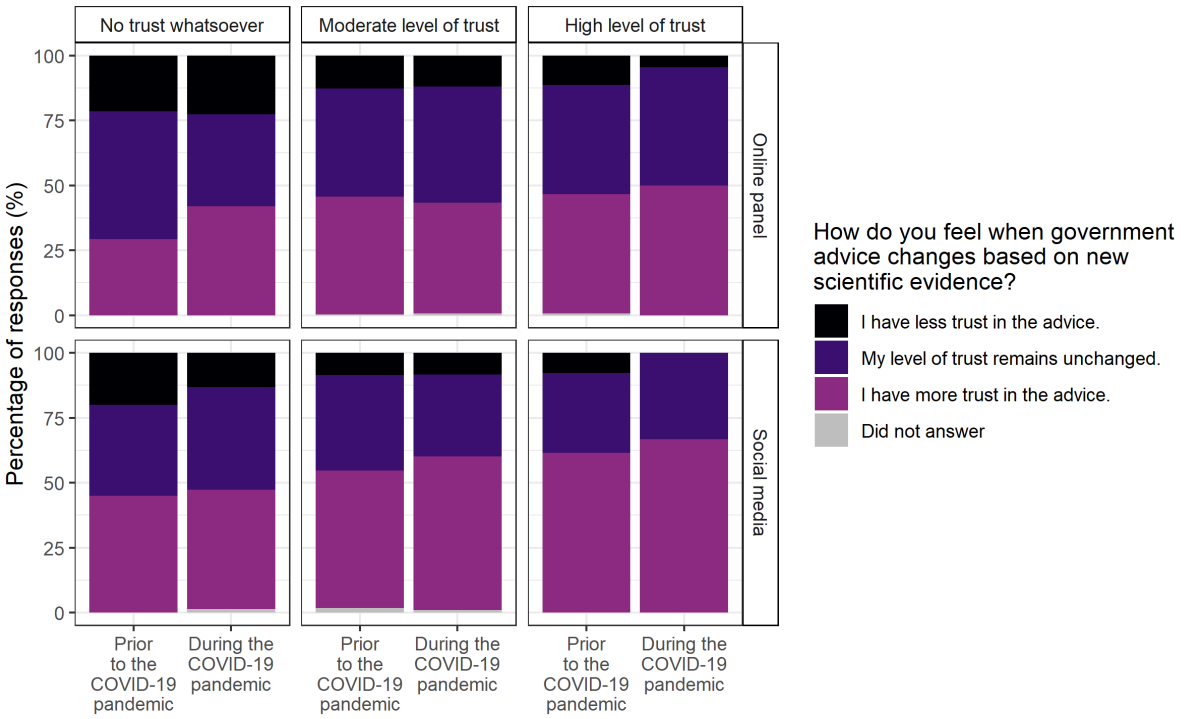

Figure C.15: Responses to the question "How do you feel when government advice changes based on new scientific evidence?" for both the online panel and social media samples, stratified by answers to the question "How much did you trust government advice regarding public health policy?" at the time "Prior to" and "During" the COVID-19 pandemic. Underlying data are presented in Table C.22.

underlying data are presented in Table C.24, with Wilcoxon signed rank tests for significant differences across time presented in Table C.24. For each category, data were enumerated to a binary scale with 1 indicating that this category was referred to and 0 indicating that it was not. Table C.25 presents the results of chi-squared tests for differences in the proportion of individual selecting each option across samples.

Many respondents from the online panel stated that they considered those who developed transmission models worked either within unspecified research institutions (35%) or within academia specifically (33%) prior to the COVID-19 pandemic. "Research"-related answers remained the most common responses in the period of the pandemic. Approximately the same percentage of online panel respondents referred to "Academia" specifically (36%; p-value=0.027; Table C.24) and to research within unspecified institutions (30%; p-value=0.021; Table C.24). The number of respondents who believed that modellers worked within "Government" during the pandemic rose significantly (29% during up from 18% prior; Table C.24), as did "Healthcare services" (14% during up from 8% prior; Table C.24). In particular, direct mention of the NHS doubled during (6%) compared to prior (3%) to the pandemic (Table C.24).

Within the social media sample, there were few categories with statistically significant differences in the percentage of respondents selecting them across the two time periods (Table C.24)). "Research (Academia)" was the most common response across both time periods (70% prior and 73% during; Table C.24)). The second most common response was "Public health bodies" (28% in both time periods), which was similarly popular among this cohort as it was among online panel respondents in each time period. However, "Research (unspecified affiliations)" was a significantly less common response among social media respondents in both time periods compared to the online panel sample (Table C.25). "Healthcare services" was also a comparatively less popular choice among this cohort than online panel respondents among both time periods, as was explicit mention of the NHS in the period of the pandemic but not in the period prior (Table C.25)).

In both samples, respondents referred to "Advisory roles" specifically in the period during the pandemic (4% online panel and 7% social media, respectively; Table C.25)), but not at all in the period prior. Respondents from both samples were equally likely to report this option (Table C.25)). Furthermore, "SAGE" was explicitly mentioned in the period of the pandemic, with approximately equal percentages of respondents selecting this option under both samples.

In reference to both periods, social media respondents were equally likely as online panel respondents to record an answer related to "Government" (Table C.25). Although "Research (Academia)" was significantly more common among both samples than "Government" in the period of the pandemic (Wilcoxon signed rank tests: online panel: $p = 0.016$; social media: $p < 0.001$), the magnitude of the differences was substantially different (online panel: 36% "Research (Academia)" vs. 29% "Government"; social media: 73% "Research (Academia)" vs. 22% "Government").

The percentage of respondents answering "Pharmaceutical companies" was relatively small and remained constant across samples and time periods. This was similar when considering "Media". As observed with other questions, significantly more respondents in the online panel sample answered "Unsure" in both time periods compared to the social sample (online panel: 8% prior and 5% during; social media: 1% prior and 0% during; Figure C.16; Table C.24)).

C.3.6 How would you describe a transmission model?

Descriptions of transmission models varied widely both within and across samples. Approximately 49% of the responses from online panel were deemed as relevant to the context of the research as defined in Methods in the main text, compared to 84% of social media respondents (Figure C.17; chi-squared

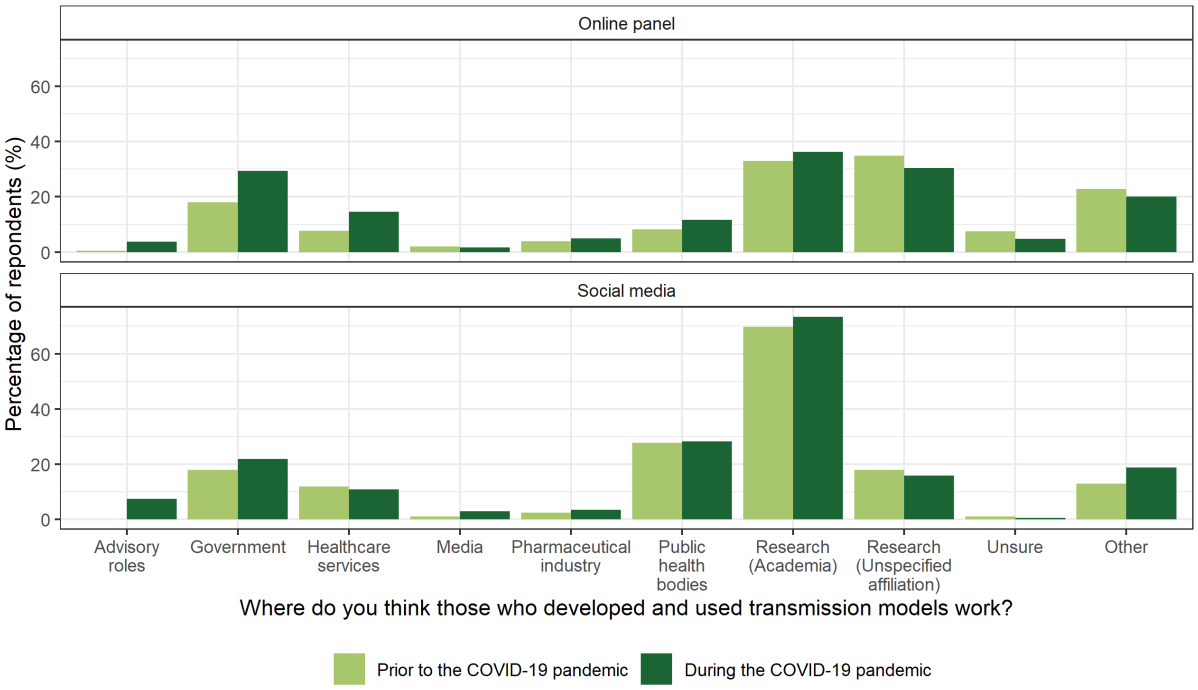

Figure C.16: Responses to the question "Where do you think those who developed and used transmission models work?" either "Prior to" or "During" the COVID-19 pandemic for both the online panel and social media respondents.

test: $\chi^2 = 72.24$; $n = 706$; $p < 0.001$). In general, social media participants were more likely to use technical language such as "confidence interval" or to refer directly to mathematical model structures, such as "Susceptible, Infected, Recovered" than their online panel counterparts, highlighting potential differences in the experience and backgrounds of participants across samples. Nonetheless, the more relevant explanations of individuals were clear, regardless of sample, including:

"Theories that explain how diseases can be transmitted." (online panel respondent)

and

"A mathematical model which replicates the true transmission of the virus in the real world." (social media respondent).

The latter also highlights common occurrence across responses: the use of the term "the virus", or more generally, direct reference to COVID-19, despite this not being mentioned in the survey title or question. There were other interesting and common occurrences across both samples. For example, the "R number" was explicitly mentioned on multiple occasions, perhaps highlighting the continued use of this term at government COVID-19 press conferences and within the media more widely. Many participants also made use of the words akin to "graphic" in their model description, which could similarly be explained by considering that this was a common method of sharing the results of modelling with the public. The word "predict" or "prediction" was also commonly used by participants, which is particularly interesting given that experts have tried to emphasise the difference between scenario-based and predictive models

[325, 429].

Approximately 20% of responses from the online panel were related to communication, for example: "A means to communicate information to someone, people or system.", compared to 3% of responses from social media. However, recall that social media respondents were likely to have accessed the survey directly from an account related to public health or mathematical modelling (Table C.1), and so the topic of the survey is likely to appear less "random" to this group, compared to the online panel respondents who saw this survey without any context. The other most common topics among online panel respondents included transmission modelling in the context of transmitting information (e.g., "A model whereby a message is sent from a source to a transmitter where they are then received and read.") and the sharing of information (e.g., "A guide to on how to disseminate information."). Social media respondents were as likely to give an "opinion" as an answer, such as "production of false statistics" or "guess work erring massively over-cautious", than online panel respondents (3% and 1%, respectively; chi-squared test: $\chi^2 = 3.46$; $n = 706$; $p = 0.063$).

Most respondents from both samples selected being "moderately" or "very" confident in their model description (65% and 60%, respectively; chi-squared test: $\chi^2 = 1.34$; $n = 706$; $p = 0.247$), with "moderately confident" accounting for a majority of responses overall (Table 5.1). In both samples, being male was significantly associated with a higher level of confidence compared to being female, while within the online panel sample only the older age groups were associated with increased levels of confidence compared to the youngest age group (Table C.26).

A different and complex picture of confidence levels emerged when data were stratified by the relevance of the model description (Figure C.17). Although those with "more relevant" descriptions continued to primarily state being "Moderately confident" in their response, more respondents with less relevant model descriptions fell into one of the extremist categories of "Very" or "Not at all" confident (Table C.27). Among the "very confident" respondents, social media respondents were significantly more likely to have written a "more relevant" response than their online panel counterparts (81% and 29%, respectively; chi-squared test: $\chi^2 = 13.43$; $n = 706$; $p < 0.001$).

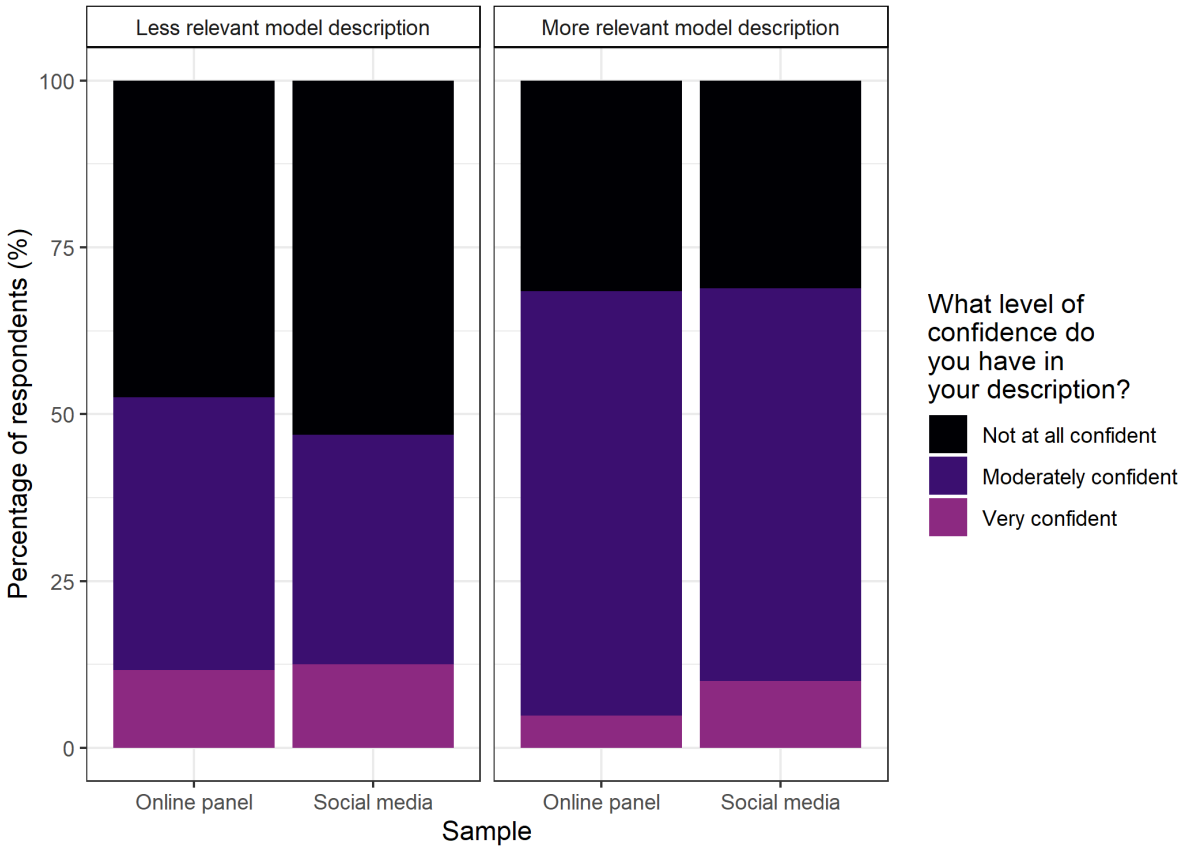

Figure C.17: The percentage of respondents from online panel and social media samples according to relevance of and confidence in answer provided to the question "How would you describe a transmission model?". Underlying data presented in Table 5.1.

Table C.10: Coefficient estimates, standard errors and p-values for the linear regression models with age group, gender and awareness of the use of modelling in policy as predictors of reliability scores during each period and for each sample. The responses to the question "Were you aware of transmission modelling in informing public health policy?" used as an explanatory variable, "No", "Unsure" and "Yes" were enumerated as -1, 0, 1, respectively. Chi-squared tests for the overall significance of age and gender are also presented. $p < 0.01$ are considered statistically significant. *Baseline comprised of "Female" and "18 - 25" years.

Sample	**Response**	**Variable**	**Interpretation**	**Estimate**	**Standard error**	p	**Chi-squared test**	
							χ^2	p
Online panel	Reliability score prior to the COVID-19 pandemic	Intercept	Baseline*	6.95	0.23	< 0.001		
		Age (years)	26 - 35	-0.18	0.28	0.505	12.66 (df=5)	0.520
			36 - 45	-0.29	0.28	0.287		
			46 - 55	-0.62	0.28	0.028		
			56 - 65	-0.36	0.26	0.164		
			66+	0.00	0.35	0.995		
		Gender	Male	-0.31	0.16	0.050	10.65 (df=3)	0.316
			Non-binary	-0.92	1.75	0.597		
			Prefer not to say	1.45	1.01	0.150		
		Awareness	Transmission modelling used in policy	0.69	0.09	< 0.001	174.06 (df=1)	< 0.001
Social media	Reliability the score prior to the COVID-19 pandemic	Intercept	Baseline*	6.85	0.96	< 0.001		
		Age (years)	26 - 35	-0.94	1.10	0.394	4.66 (df=5)	0.930
			36 - 45	-0.38	1.01	0.706		
			46 - 55	-0.32	0.97	0.740		
			56 - 65	-0.19	0.96	0.841		
			66+	-0.59	1.01	0.558		
		Gender	Male	-0.38	0.27	0.163	14.09 (df=2)	0.135
			Prefer not to say	-2.36	1.11	0.036		
		Awareness	Transmission modelling used in policy	0.37	0.16	0.024	17.99 (df=1)	0.024
Online panel	Reliability score during the COVID-19 pandemic	Intercept	Baseline*	6.67	0.25	< 0.001		
		Age (years)	26 - 35	-0.15	0.30	0.608	26.09 (df=5)	0.192
			36 - 45	-0.24	0.30	0.432		
			46 - 55	-0.83	0.30	0.007		
			56 - 65	-0.12	0.28	0.680		
			66+	-0.31	0.38	0.409		
		Gender	Male	-0.14	0.17	0.424	3.07 (df=3)	0.832
			Non-binary	-0.40	1.89	0.832		
			Prefer not to say	0.28	1.09	0.796		
		Awareness	Transmission modelling used in policy	0.88	0.12	< 0.001	205.75 (df=1)	< 0.001
Social media	Reliability score during the COVID-19 pandemic	Intercept	Baseline*	7.93	1.26	< 0.001		
		Age (years)	26 - 35	-2.89	1.28	0.026	55.94 (df=5)	0.039
			36 - 45	-1.05	1.18	0.374		
			46 - 55	-0.52	1.12	0.646		
			56 - 65	-0.56	1.11	0.613		
			66+	-0.60	1.16	0.605		
		Gender	Male	-0.71	0.31	0.025	61.25 (df=2)	0.002
			Prefer not to say	-3.89	1.28	0.003		
		Awareness	Transmission modelling used in policy	-0.15	0.65	0.817	0.25 (df=1)	0.817

Table C.11: Coefficient estimates, standard errors and p-values for the linear regression models with reliability score prior to the COVID-19 pandemic as the predictor of reliability score during the COVID-19 pandemic for each sample.

Sample	**Response**	**Variable**	**Estimate**	**Standard error**	p
Online panel	Reliability score during the COVID-19 pandemic	Intercept	3.07	0.27	< 0.001
		Reliability score prior to the COVID-19 pandemic	0.60	0.04	< 0.001
Social media	Reliability score during the COVID-19 pandemic	Intercept	3.22	0.52	< 0.001
		Reliability score prior to the COVID-19 pandemic	0.53	0.08	< 0.001

Table C.12: Coefficient estimates, standard errors and p-values for the linear regression models with age group and gender as predictors of the difference between reliability scores for each sample. Difference in reliability score is defined as the score given during the pandemic minus the score given prior. Chi-squared tests for the overall significance of age and gender are also presented. $p < 0.01$ is considered statistically significant. *Baseline comprised of "Female" and "18 - 25" years.

Sample	**Response**	**Variable**	**Interpretation**	**Estimate**	**Standard error**	p	**Chi-squared test**	
							χ^2	p
Online panel	Difference in reliability score	Intercept	Baseline*	0.46	0.23	0.049		
		Age (years)	26 - 35	-0.06	0.29	0.824	15.97 (df=5)	0.420
			36 - 45	0.02	0.28	0.936		
			46 - 55	-0.19	0.29	0.520		
			56 - 65	0.33	0.27	0.222		
			66+	-0.10	0.36	0.784		
		Gender	Male	0.09	0.16	0.576	2.24 (df=3)	0.874
			Non-binary	-0.39	1.80	0.828		
			Prefer not to say	-0.55	1.04	0.596		
Social media	Difference in reliability score	Intercept	Baseline*	0.90	1.13	0.426		
		Age (years)	26 - 35	-1.98	1.30	0.130	40.61 (df=5)	0.142
			36 - 45	-0.93	1.20	0.439		
			46 - 55	-0.15	1.14	0.895		
			56 - 65	-0.34	1.13	0.762		
			66+	-0.09	1.19	0.943		
		Gender	Male	-0.53	0.32	0.095	21.16 (df=2)	0.115
			Prefer not to say	-1.89	1.31	0.150		

Table C.13: Results of chi-squared tests for statistically significant differences in the proportion of respondents selecting multiple-choice answers to the question "Who do you think has the responsibility of ensuring that the public are informed about the use of modelling in policy decisions, particularly in the COVID-19 pandemic?" between the online panel and social media samples. Note that respondents could select multiple responses for this question. $p < 0.01$ are considered statistically significant.

	Number of respondents (n (%))		χ^2	n	p
	Online panel	**Social media**			
The government	462 (92%)	183 (91%)	0.10	706	0.756
Those who develop transmission models	198 (39%)	130 (64%)	35.44		< 0.001
Those who use transmission models	217 (43%)	129 (64%)	24.15		< 0.001
The media	221 (44%)	121 (60%)	14.24		< 0.001
Myself, as a member of the public	108 (21%)	68 (34%)	10.89		< 0.001
None of the above	1 (0%)	2 (1%)	0.67		0.411

Table C.14: Median and variances of the responses to the question "On a scale of 1 – 10 with 1 being "extremely unreliable" and 10 being "extremely reliable" how do you feel about the use of transmission modelling in informing public health policy?" for both the online panel and social media samples stratified by responses to the question "How much did you trust government advice regarding public health issues?" both "Prior to" and "During" the COVID-19 pandemic.

		No trust whatsoever		**Moderate level of trust**		**High level of trust**	
		Median	**Variance**	**Median**	**Variance**	**Median**	**Variance**
Online panel	Prior to the COVID-19 pandemic	6.00	4.61	7.00	3.06	7.00	3.21
	During the COVID-19 pandemic	6.00	5.35	7.00	3.21	8.00	1.83
Social media	Prior to the COVID-19 pandemic	6.00	4.41	6.00	3.35	7.00	3.47
	During the COVID-19 pandemic	6.00	7.09	7.00	3.50	8.00	2.26

Table C.15: Number and percentage of respondents within each sample classified according to differences in responses to the question "How much did you trust government advice regarding public health issues?" both "Prior to" and "During" the COVID-19 pandemic.

		Online panel	**Social media**
Number of respondents (n (%))	Trust in government advice increased	203 (40%)	79 (39%)
	Trust in government advice unchanged	328 (65%)	92 (46%)
	Trust in government advice decreased	135 (27%)	98 (49%)
	At least one score was missing	1 (0%)	0 (0%)

Table C.16: Coefficient estimates, standard errors and p-values for the linear model regressing responses to the question "How much did you trust government advice regarding public health issues?" on age group and gender during each period and for each sample. The three responses provided in the multiple-choice options, "No trust whatsoever", "Moderate level of trust" and "High level of trust", were enumerated as -1, 0, 1, respectively. Chi-squared tests for the overall significance of age and gender are also presented. p-values < 0.01 are considered statistically significant. *Baseline comprised of "Female" and "18 - 25" years.

Sample	**Response**	**Variable**	**Interpretation**	**Estimate**	**Standard error**	p	**Chi-squared test**	
							χ^2	p
Online panel	Trust in government prior to the COVID-19 pandemic the	Intercept	Baseline*	0.11	0.07	0.136		
		Age (years)	26 - 35	0.02	0.09	0.831	0.13 (df=5)	0.995
			36 - 45	0.01	0.09	0.953		
			46 - 55	-0.01	0.09	0.933		
			56 - 65	-0.02	0.09	0.795		
			66+	-0.02	0.12	0.886		
		Gender	Male	0.10	0.05	0.047	1.46 (df=3)	0.230
			Non-binary	-0.13	0.59	0.822		
			Prefer not to say	0.22	0.34	0.515		
Social media	Trust in government prior to the COVID-19 pandemic	Intercept	Baseline*	0.23	0.31	0.461		
		Age (years)	26 - 35	0.25	0.36	0.485	2.25 (df=5)	0.312
			36 - 45	-0.07	0.33	0.838		
			46 - 55	-0.00	0.32	0.998		
			56 - 65	0.01	0.31	0.969		
			66+	-0.24	0.33	0.465		
		Gender	Male	0.02	0.09	0.776	0.05 (df=2)	0.939
			Prefer not to say	0.09	0.36	0.805		
Online panel	Trust in government during the COVID-19 pandemic	Intercept	Baseline*	-0.17	0.08	0.039		
		Age (years)	26 - 35	-0.02	0.10	0.864	1.96 (df=5)	0.455
			36 - 45	0.11	0.10	0.279		
			46 - 55	0.05	0.10	0.664		
			56 - 65	0.12	0.10	0.233		
			66+	0.15	0.13	0.238		
		Gender	Male	0.08	0.06	0.167	1.71 (df=3)	0.251
			Non-binary	-0.81	0.65	0.213		
			Prefer not to say	-0.23	0.38	0.539		
Social media	Trust in government during the COVID-19 pandemic	Intercept	Baseline*	-0.48	0.31	0.122		
		Age (years)	26 - 35	-0.00	0.36	0.993	2.90 (df=5)	0.175
			36 - 45	0.02	0.33	0.964		
			46 - 55	0.18	0.32	0.564		
			56 - 65	0.37	0.31	0.243		
			66+	0.15	0.33	0.656		
		Gender	Male	-0.02	0.09	0.811	2.22 (df=2)	0.053
			Prefer not to say	-0.88	0.36	0.016		

Table C.17: Coefficient estimates, standard errors and p-values for the linear model regressing changes in responses to the question "How much did you trust government advice regarding public health issues?" in the period of the pandemic compared to the period prior, on age group and gender within each sample. The three responses provided in the multiple-choice options, "No trust whatsoever", "Moderate level of trust" and "High level of trust", were enumerated as -1, 0, 1, respectively, and so the differences (response) could take integer values between -2 and 2. Chi-squared tests for the overall significance of age and gender are also presented. p-values < 0.01 are considered statistically significant. *Baseline comprised of "Female" and "18 - 25" years.

Sample	**Response**	**Variable**	**Interpretation**	**Estimate**	**Standard error**	p	**Chi-squared test**	
							χ^2	p
Online panel	Change in trust in government regarding public health issues	Intercept	Baseline*	-0.30	0.09	< 0.001		
		Age (years)	26 - 35	-0.02	0.11	0.845	2.81 (df=5)	0.269
			36 - 45	0.12	0.10	0.242		
			46 - 55	0.07	0.11	0.511		
			56 - 65	0.15	0.10	0.118		
			66+	0.19	0.13	0.158		
		Gender	Male	-0.03	0.06	0.634	1.11 (df=3)	0.469
			Non-binary	-0.68	0.67	0.307		
			Prefer not to say	-0.45	0.38	0.239		
Social media	Change in trust in government regarding public health issues	Intercept	Baseline*	-0.72	0.37	0.054		
		Age (years)	26 - 35	-0.26	0.43	0.549	4.56 (df=5)	0.126
			36 - 45	0.08	0.39	0.832		
			46 - 55	0.18	0.37	0.625		
			56 - 65	0.35	0.37	0.341		
			66+	0.39	0.38	0.319		
		Gender	Male	-0.05	0.10	0.657	2.70 (df=2)	0.078
			Prefer not to say	-0.97	0.43	0.024		

Table C.18: Coefficient estimates, standard errors and p-values for the linear model regressing responses to the question "How much did you trust government advice regarding public health issues?" on age group, gender and awareness of the use of modelling in policy during each period and for each sample. The three responses provided in the multiple-choice options for the response, "No trust whatsoever", "Moderate level of trust" and "High level of trust", were enumerated as -1, 0, 1, respectively. Similarly, the responses to the question "Were you aware of transmission modelling in informing public health policy?" used as an explanatory variable, "No", "Unsure" and "Yes" were enumerated as -1, 0, 1, respectively. Chi-squared tests for the overall significance of explanatory variables are also presented. p-values < 0.01 are considered statistically significant. *Baseline comprised of "Female" and "18 - 25" years.

Sample	**Response**	**Variable**	**Interpretation**	**Estimate**	**Standard error**	p	**Chi-squared test**	
							χ^2	p
Online panel	Trust in government prior to the COVID-19 pandemic the	Intercept	Baseline*	0.11	0.08	0.149		
		Age (years)	26 - 35	0.02	0.09	0.828	0.18 (df=5)	0.991
			36 - 45	0.01	0.09	0.942		
			46 - 55	-0.01	0.09	0.937		
			56 - 65	-0.03	0.09	0.734		
			66+	-0.02	0.12	0.886		
		Gender	Male	0.11	0.05	0.042	1.53 (df=3)	0.212
			Non-binary	-0.13	0.59	0.824		
			Prefer not to say	0.22	0.34	0.510		
		Awareness	Transmission modelling used in policy	-0.00	0.03	0.979	0 (df=1)	0.979
Social media	Trust in government prior to the COVID-19 pandemic	Intercept	Baseline*	0.25	0.31	0.420		
		Age (years)	26 - 35	0.26	0.36	0.466	2.45 (df=5)	0.309
			36 - 45	-0.06	0.33	0.849		
			46 - 55	-0.01	0.32	0.964		
			56 - 65	0.00	0.31	0.992		
			66+	-0.25	0.33	0.450		
		Gender	Male	0.05	0.09	0.605	0.05 (df=2)	0.938
			Prefer not to say	0.15	0.36	0.678		
		Awareness	Transmission modelling used in policy	-0.07	0.05	0.159	0.75 (df=1)	0.159
Online panel	Trust in government during the COVID-19 pandemic	Intercept	Baseline*	-0.20	0.09	0.022		
		Age (years)	26 - 35	-0.01	0.10	0.938	1.66 (df=5)	0.551
			36 - 45	0.11	0.10	0.294		
			46 - 55	0.02	0.10	0.829		
			56 - 65	0.11	0.10	0.258		
			66+	0.12	0.13	0.344		
		Gender	Male	0.08	0.06	0.185	1.70 (df=3)	0.254
			Non-binary	-0.84	0.65	0.195		
			Prefer not to say	-0.26	0.38	0.497		
		Awareness	Transmission modelling used in policy	0.05	0.04	0.209	0.66 (df=1)	0.209
Social media	Trust in government during the COVID-19 pandemic	Intercept	Baseline*	-0.41	0.36	0.256		
		Age (years)	26 - 35	-0.02	0.36	0.960	2.67 (df=5)	0.217
			36 - 45	0.03	0.36	0.960		
			46 - 55	0.18	0.32	0.563		
			56 - 65	0.37	0.31	0.242		
			66+	0.15	0.33	0.654		
		Gender	Male	-0.01	0.09	0.929	2.21 (df=2)	0.055
			Prefer not to say	-0.87	0.36	0.017		
		Awareness	Transmission modelling used in policy	-0.09	0.18	0.641	0.08 (df=1)	0.641

Table C.19: The percentage of observations from both the online panel and social media samples according to time period and answer to the question "How much did you trust government advice regarding public health issues?" both "Prior to" and "During" the COVID-19 pandemic, stratified by responses to the question "Were you aware of the use of transmission models in informing public health policy?" both "Prior to" and "During" the COVID-19 pandemic.

Platform	**Time**	**Were you aware of the use of transmission models in informing public health policy?**	**How much did you trust government advice regarding public health issues?**			
			High level of trust	**Moderate level of**	**No trust whatsoever**	**Did not answer**
Online panel	Prior to the COVID-19 pandemic	Yes	38 (28%)	85 (62%)	14 (10%)	0 (0%)
		Unsure	17 (19%)	68 (75%)	6 (7%)	0 (0%)
		No	74 (27%)	168 (61%)	31 (11%)	0 (0%)
	During the COVID-19 pandemic	Yes	73 (20%)	208 (56%)	93 (25%)	0 (0%)
		Unsure	4 (9%)	33 (70%)	10 (21%)	0 (0%)
		No	9 (12%)	45 (60%)	20 (27%)	1 (1%)
Social media	Prior to the COVID-19 pandemic	Yes	35 (30%)	67 (58%)	14 (12%)	0 (0%)
		Unsure	12 (32%)	22 (59%)	3 (8%)	0 (0%)
		No	18 (37%)	28 (57%)	3 (8%)	0 (0%)
	During the COVID-19 pandemic	Yes	18 (9%)	103 (53%)	74 (38%)	0 (0%)
		Unsure	0 (0%)	4 (100%)	0 (0%)	0 (0%)
		No	0 (0%)	1 (50%)	1 (50%)	0 (0%)

Table C.20: Coefficient estimates, standard errors and p-values for the linear model regressing responses to the question "How do you feel when government advice changes based on new scientific evidence?" on age group and gender for each sample. The three responses provided in the multiple-choice options for the response, "I have less trust in the advice", "My level of trust remains unchanged" and "I have more trust in the advice", were enumerated as -1, 0, 1, respectively. Chi-squared tests for the overall significance of age and gender are also presented. $p < 0.01$ is considered statistically significant. *Baseline comprised of "Female" and "18 - 25" years.

Sample	**Response**	**Variable**	**Interpretation**	**Estimate**	**Standard error**	p	**Chi-squared test**	
							χ^2	p
Online panel	Feeling regarding government changing advice based on new scientific evidence	Intercept	Baseline*	0.43	0.09	< 0.001		
		Age (years)	26 - 35	-0.06	0.11	0.560	2.91 (df=5)	0.304
			36 - 45	-0.12	0.11	0.255		
			46 - 55	-0.24	0.11	0.033		
			56 - 65	-0.16	0.10	0.126		
			66+	-0.21	0.14	0.137		
		Gender	Male	0.01	0.06	0.898	0.50 (df=3)	0.790
			Non-binary	-0.37	0.70	0.599		
			Prefer not to say	0.35	0.40	0.386		
Social media	Feeling regarding government changing advice based on new scientific evidence	Intercept	Baseline*	0.60	0.34	0.075		
		Age (years)	26 - 35	0.08	0.39	0.839	1.39 (df=5)	0.669
			36 - 45	-0.18	0.36	0.611		
			46 - 55	0.02	0.34	0.957		
			56 - 65	-0.07	0.34	0.834		
			66+	-0.20	0.36	0.576		
		Gender	Male	-0.14	0.09	0.147	2.75 (df=2)	0.045
			Prefer not to say	-0.87	0.39	0.028		

Table C.21: Coefficient estimates, standard errors and p-values for the generalised linear model regressing responses to the question "How do you feel when government advice changes based on new scientific evidence?" on age group and gender for each sample. The three responses provided in the multiple-choice options for the response, "I have less trust in the advice", "My level of trust remains unchanged" and "I have more trust in the advice", were enumerated as -1, 0, 1, respectively. Responses to the explanatory variable regarding awareness, "Yes", "Unsure" and "No", were enumerated as 1, 0, -1, respectively. Chi-squared tests for the overall significance of explanatory are also presented. $p < 0.01$ is considered statistically significant. *Baseline comprised of "Female" and "18 - 25" years.

Sample	**Response**	**Variable**	**Interpretation**	**Estimate**	**Standard error**	p	**Chi-squared test**	
							χ^2	p
Online panel	Feeling regarding government changing advice the based on new scientific in evidence	Intercept	Baseline*	0.41	0.09	< 0.001		
		Age (years)	26 - 35	-0.08	0.11	0.470	2.55 (df=5)	0.372
			36 - 45	-0.14	0.11	0.208		
			46 - 55	-0.26	0.11	0.020		
			56 - 65	-0.16	0.10	0.118		
			66+	-0.18	0.14	0.202		
		Gender	Male	0.01	0.06	0.932	0.52 (df=3)	0.776
			Non-binary	-0.36	0.70	0.604		
			Prefer not to say	0.34	0.40	0.403		
		Awareness	Transmission modelling used in policy prior to the pandemic	0.05	0.04	0.187	1.85 (df=1)	0.049
			Transmission modelling used in policy during the pandemic	0.08	0.04	0.061	1.68 (df=1)	0.060
Social media	Feeling regarding government changing advice the based on new scientific in evidence	Intercept	Baseline*	0.57	0.39	0.142		
		Age (years)	26 - 35	0.08	0.39	0.840	1.20 (df=5)	0.734
			36 - 45	-0.12	0.36	0.747		
			46 - 55	0.03	0.34	0.929		
			56 - 65	-0.06	0.34	0.853		
			66+	-0.19	0.35	0.593		
		Gender	Male	-0.13	0.10	0.173	2.53 (df=2)	0.056
			Prefer not	-0.89	0.39	0.025		
			to say					
		Awareness	Transmission modelling used in policy prior to the pandemic	0.04	0.06	0.534	0.18 (df=1)	0.515
			Transmission modelling used in policy during the pandemic	0.01	0.20	0.945	0.00 (df=1)	0.945

Table C.22: Coefficient estimates, standard errors and p-values for the generalised linear model regressing responses to the question "How do you feel when government advice changes based on new scientific evidence?" on age group and gender for each sample. The three responses provided in the multiple-choice options for the response, "I have less trust in the advice", "My level of trust remains unchanged" and "I have more trust in the advice", were enumerated as -1, 0, 1, respectively. Responses to the explanatory variable regarding government trust, "High level of trust", "Moderate level of trust" and "No trust whatsoever", were enumerated as 1, 0, -1, respectively. Chi-squared tests for the overall significance of age and gender are also presented. $p < 0.01$ is considered statistically significant. *Baseline comprised of "Female" and "18 - 25" years.

Sample	**Response**	**Variable**	**Interpretation**	**Estimate**	**Standard error**	p	**Chi-squared test**	
							χ^2	p
Online panel	Feeling regarding government changing advice the based on new scientific in evidence	Intercept	Baseline*	0.44	0.09	< 0.001		
		Age (years)	26 - 35	-0.05	0.11	0.624	2.74 (df=5)	0.323
			36 - 45	-0.13	0.11	0.237		
			46 - 55	-0.23	0.11	0.036		
			56 - 65	-0.16	0.10	0.116		
			66+	-0.21	0.14	0.121		
		Gender	Male	-0.01	0.06	0.889	0.50 (df=3)	0.788
			Non-binary	-0.27	0.70	0.702		
			Prefer not to say	0.37	0.40	0.362		
		Government trust	Trust in advice prior to the pandemic	0.05	0.06	0.446	1.70 (df=1)	0.059
			Trust in advice during the pandemic	0.12	0.05	0.024	2.42 (df=1)	0.024
Social media	Feeling regarding government changing advice the based on new scientific in evidence	Intercept	Baseline*	0.65	0.34	0.057		
		Age (years)	26 - 35	0.06	0.39	0.877	1.39 (df=5)	0.662
			36 - 45	-0.18	0.36	0.620		
			46 - 55	-0.01	0.34	0.984		
			56 - 65	-0.12	0.34	0.723		
			66+	-0.20	0.35	0.574		
		Gender	Male	-0.14	0.09	0.152	2.75 (df=2)	0.043
			Prefer not	-0.76	0.40	0.057		
			to say					
		Government trust	Trust in advice prior to the pandemic	0.08	0.08	0.340	0.99 (df=1)	0.131
			Trust in advice during the pandemic	0.13	0.08	0.108	1.12 (df=1)	0.108

Table C.23: The percentage of observations from both the online panel and social media samples according to time period and answer to the question "How do you feel when government advice changes based on new scientific evidence?", stratified by responses to the question "Were you aware of the use of transmission models in informing public health policy?" both "Prior to" and "During" the COVID-19 pandemic.

Platform	**Time**	**How much did you trust government advice regarding public health policy?**	**How do you feel when government advice changes based on new scientific evidence?**			
			I have more trust in the advice	**My level of level of remains unchanged**	**I have less trust in the advice**	**Did not answer**
Online panel	Prior to the COVID-19 pandemic	High level of trust	60 (46%)	55 (42%)	15 (11%)	1 (1%)
		Moderate level of trust	146 (45%)	134 (42%)	41 (13%)	1 (0%)
		No trust whatsoever	15 (29%)	25 (49%)	11 (22%)	0 (0%)
	During the COVID-19 pandemic	High level of trust	44 (50%)	40 (45%)	4 (5%)	0 (0%)
		Moderate level of trust	124 (43%)	130 (45%)	35 (12%)	2 (1%)
		No trust whatsoever	52 (42%)	44 (35%)	28 (23%)	0 (0%)
Social media	Prior to the COVID-19 pandemic	High level of trust	40 (62%)	20 (31%)	5 (8%)	0 (0%)
		Moderate level of trust	62 (53%)	43 (37%)	10 (9%)	2 (2%)
		No trust whatsoever	9 (45%)	7 (35%)	4 (20%)	0 (0%)
	During the COVID-19 pandemic	High level of trust	12 (67%)	6 (33%)	0 (0%)	0 (0%)
		Moderate level of trust	64 (59%)	34 (31%)	9 (8%)	1 (1%)
		No trust whatsoever	35 (46%)	30 (39%)	10 (13%)	1 (1%)

Table C.24: Responses to the question "Where do you think those who developed and used transmission models work?" either "Prior to" or "During" the COVID-19 pandemic for both the online panel and social media samples. Percentages are taken with respect to the sample size of the platform ($n = 504$ for the online panel and $n = 202$ for social media). As this question was open-ended, participants often provided more than one answer and so percentages do not represent mutually exclusive responses. Differences within samples are assessed with Wilcoxon signed rank tests. $p < 0.01$ is considered statistically significant.

		Number of respondents (n (%))		**Wilcoxon signed rank test statistic**	n	p
		Prior to the COVID-19 pandemic	**During the COVID-19 pandemic**			
Online panel	Advisory roles	2 (0%)	19 (4%)	22	504	< 0.001
	(SAGE specifically)	(0 (0%))	(11 (2%))	(0)		(0.001)
	Government	91 (18%)	148 (29%)	1196		< 0.001
	Healthcare services	40 (8%)	73 (14%)	313.5		< 0.001
	(NHS specifically)	(13 (3%))	(32 (6%))	(119)		(0.001)
	Media	10 (2%)	8 (2%)	7.5		0.424
	Pharmaceutical industry	20 (4%)	25 (5%)	88		0.284
	Public health bodies	41 (8%)	59 (12%)	367.5		0.010
	Research (Academia)	166 (33%)	183 (36%)	630		0.027
	Research (Unspecified affiliation)	176 (35%)	153 (30%)	3050		0.021
	Unsure	38 (8%)	24 (5%)	304.5		0.008
	Other	115 (23%)	102 (20%)	2464.5		0.145
Social media	Advisory roles	0 (0%)	15 (7%)	0	202	< 0.001
	(SAGE specifically)	(0 (0%))	(6 (3%))	(0)		(0.020)
	Government	38 (19%)	44 (22%)	138		0.683
	Healthcare services	24 (12%)	22 (11%)	138		0.683
	(NHS specifically)	(11 (5%))	(10 (5%))	(25)		(0.790)
	Media	2 (1%)	6 (3%)	0		0.072
	Pharmaceutical industry	5 (2%)	7 (3%)	22		0.565
	Public health bodies	56 (28%)	57 (28%)	380		0.879
	Research (Academia)	141 (70%)	148 (73%)	320		0.266
	Research (Unspecified affiliation)	36 (18%)	32 (16%)	126		0.383
	Unsure	2 (1%)	1 (0%)	1		1.000
	Other	26 (13%)	38 (19%)	399.5		0.078

Table C.25: Responses to the question "Where do you think those who developed and used transmission models work?" either "Prior to" or "During" the COVID-19 pandemic for both the online panel and social media samples. Percentages are taken with respect to the sample size of the platform ($n = 504$ for the online panel and $n = 202$ for social media). As this question was open-ended, participants often provided more than one answer and so percentages do not represent mutually exclusive responses. Differences with reference to both time periods are assessed with Chi-squared tests. $p < 0.01$ are considered statistically significant.

		Number of respondents (n (%))		**Wilcoxon signed rank test statistic**	n	p
		Online panel	**Social media**			
Prior to the COVID-19 pandemic	Advisory roles	2 (0%)	0 (0%)	0.01	706	0.910
	(SAGE specifically)	(0 (0%))	(0 (0%))	(NA)		(NA)
	Government	91 (18%)	38 (19%)	0.02		0.899
	Healthcare services	40 (8%)	24 (12%)	2.26		0.132
	(NHS specifically)	(13 (3%))	(11 (5%))	(2.88)		(0.095)
	Media	10 (2%)	2 (1%)	0.36		0.548
	Pharmaceutical industry	20 (4%)	5 (2%)	0.55		0.456
	Public health bodies	41 (8%)	56 (28%)	45.05		< 0.001
	Research (Academia)	166 (33%)	141 (70%)	78.25		< 0.001
	Research (Unspecified affiliation)	176 (35%)	36 (18%)	19.26		< 0.001
	Unsure	38 (8%)	2 (1%)	10.38		0.001
	Other	115 (23%)	26 (13%)	8.31		0.004
During the COVID-19 pandemic	Advisory roles	19 (4%)	15 (7%)	3.44		0.063
	(SAGE specifically)	(11 (2%))	(6 (3%))	(0.12)		(0.730)
	Government	148 (29%)	44 (22%)	3.81		0.051
	Healthcare services	73 (14%)	22 (11%)	1.31		0.253
	(NHS specifically)	(32 (6%))	(10 (5%))	(0.29)		(0.593)
	Media	8 (2%)	6 (3%)	0.80		0.372
	Pharmaceutical industry	25 (5%)	7 (3%)	0.44		0.507
	Public health bodies	59 (12%)	57 (28%)	27.44		< 0.001
	Research (Academia)	183 (36%)	148 (73%)	77.62		< 0.001
	Research (Unspecified affiliation)	153 (30%)	32 (16%)	14.97		< 0.001
	Unsure	24 (5%)	1 (0%)	6.49		0.011
	Other	102 (20%)	38 (19%)	0.11		0.745

Table C.26: Coefficient estimates, standard errors and p-values for the linear regression models with gender and age group as predictors of level of confidence within each model for each sample. The three levels of confidence provided in the multiple-choice options, "Not at all confident", "Moderately confident" and "Very confident", were enumerated as -1, 0, 1, respectively. p-values < 0.01 are considered statistically significant. *Baseline comprised of "Female" and "18 - 25" years.

Sample	**Response**	**Variable**	**Interpretation**	**Estimate**	**Standard error**	p	**Chi-squared test**	
							χ^2	p
Online panel	What level of confidence do you have in your description?	Intercept	Baseline*	-0.68	0.08	< 0.001		
		Age (years)	26 - 35	0.15	0.10	0.107	7.29 (df=5)	0.001
			36 - 45	0.37	0.09	< 0.001		
			46 - 55	0.35	0.10	< 0.001		
			56 - 65	0.28	0.09	0.001		
			66+	0.32	0.12	0.009		
		Gender	Male	0.23	0.05	< 0.001		
			Non-binary	0.52	0.60	0.389	(df=3)	
			Prefer not to say	0.41	0.35	0.245		
Social media	What level of confidence do you have in your description?	Intercept	Baseline*	-0.24	0.31	0.445		
		Age (years)	26 - 35	0.25	0.36	0.492	2.58 (df=5)	0.225
			36 - 45	-0.20	0.33	0.536		
			46 - 55	-0.14	0.31	0.656		
			56 - 65	-0.19	0.31	0.534		
			66+	-0.24	0.33	0.459		
		Gender	Male	0.32	0.09	< 0.001	5.10 (df=2)	0.001
			Prefer not to say	0.43	0.36	0.232		

Table C.27: Number and percentage of respondents under each of the three confidence levels depending on the relevance of their description and their sample. Percentages calculated with respect to each column.

What level of confidence do you have in your description?	**Description less relevant**		**Description more relevant**	
	Online panel respondents	**Social media respondents**	**Online panel respondents**	**Social media respondents**
Not at all confident	122 (48%)	17 (53%)	78 (32%)	53 (31%)
Moderately confident	105 (41%)	11 (34%)	157 (64%)	100 (59%)
Very confident	30 (12%)	4 (13%)	12 (5%)	17 (10%)
Total	257 (51%)	32 (16%)	247 (49%)	170 (84%)

Appendix D

Appendix: Paper IV: Tracking the incidence and risk factors for SARS-CoV-2 infection using historical maternal booking samples

Supplementary material for Paper IV (Chapter 6)

E Mullins*, R McCabe*, SM Bird, P Randell, MJ Pond, L Regan, E Parker, M McClure and CA Donnelly. Tracking the incidence and risk factors for SARS-CoV-2 infection using historical maternal booking serum samples. *PLOS ONE*, 17(9):1–15, 09 2022. doi: `https://doi.org/10.1371/journal.pone.0273966`.

D.1 Sample size

Table D.1 presents an overview of the number of samples categorised as 'clearly' and 'borderline' seroreactive or non-seroreactive.

Table D.1: Classification and sample sizes of observations from October 2019 to September 2020 as seroreactive and non-seroreactive according to different thresholds of binding ratio (BR) values.

Assay threshold binding ratios (BRs)	**Interpretation**	**Threshold sample size**	**Cumulative sample size**
$BR \geq 1.2$	Clearly seroreactive	665	665
$1.2 > BR \geq 1.0$	Borderline seroreactive	35	700
$1.0 > BR \geq 0.8$	Borderline non-seroreactive	70	770
$0.8 > BR \geq 0.0$	Clearly non-seroreactive	10,486	11,256

D.2 Estimating prevalence of seroreactivity

Let n be the total number of observations. For each observation i, Y_i denotes whether this observation is seroreactive or non-seroreactive:

$$Y_i = \begin{cases} 1 & \text{if observation is seroreactive} \\ 0 & \text{if observation is non-seroreactive} \end{cases}$$

Each observation can be considered an independent Bernoulli trial for seroreactivity with probability s, where s denotes prevalence of seroreactivity. It follows that the total number of seroreactive observations, $X = x$, is a binomially-distributed random variable with parameters n and s:

$$X = \sum_{i=1}^{n} Y_i \sim Binomial(n, s)$$

Therefore, prevalence of seroreactivity was naively estimated using the maximum likelihood estimator (MLE) of the binomial distribution as follows:

$$\hat{s} = \frac{x}{n}$$

where the number of reactive observations was measured according to clearly and borderline seroreactive observations (binding ratio (BR) value ≥ 1). 95% confidence intervals were generated using the exact binomial method by solving the following equations for s_L and s_U, respectively:

$$\sum_{k=0}^{x} \binom{n}{k} s_U^k (1 - s_U)^{n-k} = \frac{\alpha}{2}$$
$$\sum_{k=0}^{x-1} \binom{n}{k} s_L^k (1 - s_L)^{n-k} = 1 - \frac{\alpha}{2}$$

where $\alpha = 0.05$.

D.3 Fisher's exact test for prevalence in June 2019 compared to main study period

Table D.2 presents the results of the two-sided Fisher's exact tests to test for the significance in the odds of a seroreactive sample between June 2019 compared to each fortnight in the main study.

Table D.2: Results of two-sided Fisher's exact test in the odds of observing a seroreactive sample in June 2019 (1,000 samples, 8 positives) in comparison to each fortnight in the main study period (October 2019 – September 2020).

Collection fortnight (fortnight-year)	**Calendar date**	**Total observations** (n)	**Seroreactive observations** (x)	**Odds ratio (MLE (95% exact confidence intervals))**	p
22-2019	22/10/2019 - 04/11/2019	608	8	0.61 (0.20 - 1.86)	0.313
23-2019	05/11/2019 - 18/11/2019	610	5	0.98 (0.28 - 3.81)	1.000
24-2019	19/11/2019 - 02/12/2019	546	6	0.73 (0.22 - 2.55)	0.581
25-2019	03/12/2019 - 16/12/2019	514	3	1.38 (0.33 - 8.07)	0.759
26-2019	17/12/2019 - 31/12/2019	554	5	0.89 (0.25 - 3.46)	0.781
1-2020	01/01/2020 - 14/01/2020	467	5	0.75 (0.21 - 2.91)	0.565
2-2020	15/01/2020 - 28/01/2020	572	6	0.76 (0.24 - 2.48)	0.607
3-2020	29/01/2020 - 11/02/2020	669	7	0.76 (0.24 - 2.48)	0.607
4-2020	12/02/2020 - 25/02/2020	530	9	0.47 (0.16 - 1.37)	0.127
5-2020	26/02/2020 - 10/03/2020	628	7	0.72 (0.23 - 2.33)	0.597
6-2020	11/03/2020 - 24/03/2020	659	18	0.29 (0.11 - 0.70)	0.004
7-2020	25/03/2020 - 07/04/2020	392	34	0.09 (0.03 - 0.19)	< 0.001
8-2020	08/04/2020 - 21/04/2020	622	73	0.06 (0.03 - 0.13)	< 0.001
9-2020	22/04/2020 - 05/05/2020	637	77	0.06 (0.02 - 0.10)	< 0.001
10-2020	06/05/2020 - 19/05/2020	584	83	0.05 (0.02 - 0.10)	< 0.001
11-2020	20/05/2020 - 02/06/2020	482	53	0.07 (0.03 - 0.14)	< 0.001
12-2020	03/06/2020 - 16/06/2020	499	58	0.06 (0.03 - 0.13)	< 0.001
13-2020	17/06/2020 - 30/06/2020	278	38	0.05 (0.02 - 0.11)	< 0.001
14-2020	01/07/2020 - 14/07/2020	254	37	0.05 (0.02 - 0.11)	< 0.001
15-2020	15/07/2020 - 28/07/2020	288	35	0.06 (0.02 - 0.13)	< 0.001
16-2020	29/07/2020 - 11/08/2020	267	40	0.05 (0.02 - 0.10)	< 0.001
17-2020	12/08/2020 - 25/08/2020	198	25	0.06 (0.02 - 0.13)	< 0.001
18-2020	26/08/2020 - 08/09/2020	233	35	0.05 (0.02 - 0.10)	< 0.001
19-2020	09/09/2020 - 22/09/2020	165	33	0.03 (0.01 - 0.07)	< 0.001

D.4 Logistic regression

The logistic regression model was formally described as:

$$
\begin{aligned}
logit(s) = \beta_0 &+ \beta_1 Age_{18-29} + \beta_2 Age_{30-34} \\
&+ \beta_3 Ethnicity_{All-black} + \beta_4 Ethnicity_{All-Asian} + \beta_5 Ethnicity_{Other} \\
&+ \beta_6 IMD_{deciles3-6} + \beta_7 IMD_{deciles7-10} \\
&+ \beta_8 Fortnight_{2019-22} + \beta_9 Fortnight_{2019-23} + \beta_{10} Fortnight_{2019-24} \\
&+ \beta_{11} Fortnight_{2019-25} + \beta_{12} Fortnight_{2019-26} + \beta_{13} Fortnight_{2020-1} \\
&+ \beta_{14} Fortnight_{2020-2} + \beta_{15} Fortnight_{2020-3} + \beta_{16} Fortnight_{2020-4} \\
&+ \beta_{17} Fortnight_{2020-5} + \beta_{18} Fortnight_{2020-7} + \beta_{19} Fortnight_{2020-8} \\
&+ \beta_{20} Fortnight_{2020-9} + \beta_{21} Fortnight_{2020-10} + \beta_{22} Fortnight_{2020-11} \\
&+ \beta_{23} Fortnight_{2020-12} + \beta_{24} Fortnight_{2020-13} + \beta_{25} Fortnight_{2020-14} \\
&+ \beta_{26} Fortnight_{2020-15} + \beta_{27} Fortnight_{2020-16} + \beta_{28} Fortnight_{2020-17} \\
&+ \beta_{29} Fortnight_{2020-18} + \beta_{30} Fortnight_{2020-19}
\end{aligned}
$$

The additional model with the interaction term between ethnicity and IMD decile was as follows:

$$
\begin{aligned}
logit(s) = \beta_0 &+ \beta_1 Age_{18-29} + \beta_2 Age_{30-34} \\
&+ \beta_3 Ethnicity_{all-black} + \beta_4 Ethnicity_{all-Asian} + \beta_5 Ethnicity_{other} \\
&+ \beta_6 IMD_{deciles3-6} + \beta_7 IMD_{deciles7-10} \\
&+ \beta_8 Fortnight_{2019-22} + \beta_9 Fortnight_{2019-23} + \beta_{10} Fortnight_{2019-24} \\
&+ \beta_{11} Fortnight_{2019-25} + \beta_{12} Fortnight_{2019-26} + \beta_{13} Fortnight_{2020-1} \\
&+ \beta_{14} Fortnight_{2020-2} + \beta_{15} Fortnight_{2020-3} + \beta_{16} Fortnight_{2020-4} \\
&+ \beta_{17} Fortnight_{2020-5} + \beta_{18} Fortnight_{2020-7} + \beta_{19} Fortnight_{2020-8} \\
&+ \beta_{20} Fortnight_{2020-9} + \beta_{21} Fortnight_{2020-10} + \beta_{22} Fortnight_{2020-11} \\
&+ \beta_{23} Fortnight_{2020-12} + \beta_{24} Fortnight_{2020-13} + \beta_{25} Fortnight_{2020-14} \\
&+ \beta_{26} Fortnight_{2020-15} + \beta_{27} Fortnight_{2020-16} + \beta_{28} Fortnight_{2020-17} \\
&+ \beta_{29} Fortnight_{2020-18} + \beta_{30} Fortnight_{2020-19} \\
&+ \beta_{31} Ethnicity_{all-black} \times IMD_{deciles3-6} + \beta_{32} Ethnicity_{all-Asian} \times IMD_{deciles3-6} \\
&+ \beta_{33} Ethnicity_{other} \times IMD_{deciles3-6} + \beta_{34} Ethnicity_{all-black} \times IMD_{deciles7-10} \\
&+ \beta_{35} Ethnicity_{all-Asian} \times IMD_{deciles7-10} + \beta_{36} Ethnicity_{other} \times IMD_{deciles7-10}
\end{aligned}
$$

D.5 Baseline fortnight selection

Fortnight 6 was just before the first UK lockdown in March 2020, was likely to coincide with high rates of infection and was selected as the baseline fortnight.

D.6 Incidence of seroreactivity

Figure D.1 presents the shape constrained P-spline fitted to the estimates of prevalence of seroreactivity over time, in which monotonicity between knots was enforced. In this instance, we specified 24 knots, corresponding to the number of fortnights for which incidence was estimated. The smoothed curve provided the estimates from which incidence was derived.

The bootstrapping procedure is presented below.

Let $f = 1, .., F$ denote the distinct fortnights in the study and let n_f denote the number of obser-

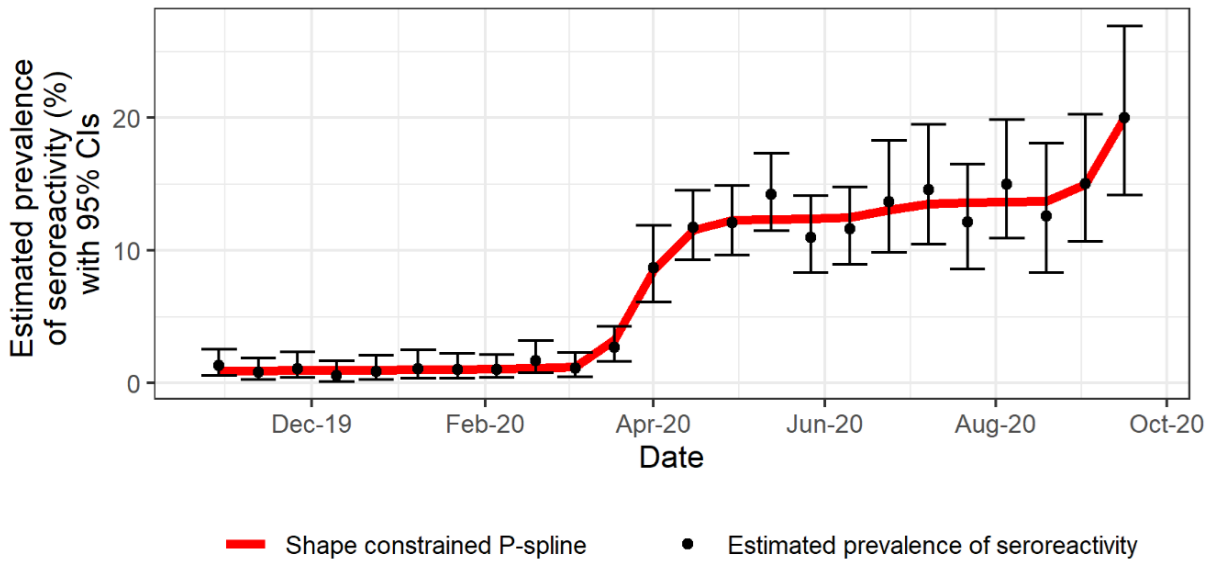

Figure D.1: Shape constrained P-spline (red) fit to prevalence of seroreactivity estimated using entire cleaned data set (black).

vations in fortnight f. Let Z denote the complete data (with $11,256$ observations), consisting of an indicator for seroreactivity (Y_i), the corresponding fortnight of the sample (f_i) and the total number of observations from fortnight f_i (n_{f_i}).

Let $b = 1, ..., 1,000$ denote the bootstrapping sample and let Z^b denote the data obtained by sampling with replacement $B = 11,256$ times from Z under the bth sample.

Prevalence of seroreactivity per fortnight of the bth bootstrapping sample was then calculated as follows:

$$\hat{s}_{b,f} = \frac{\sum_{Y_{b,i} \in f} Y_{b,i}}{n_{b,f}}$$

where $Y_{b,i}$ is the ith observation of Z^b and $n_{b,f}$ is the number of observations in fortnight f under bootstrap sample b.

Denote by $g(\cdot)$ the shape constrained P-splines in which monotonicity between knots is enforced (as detailed above). This function was applied to $\hat{s}_{b,f}$ in order to obtain smoothed estimated of prevalence of seroreactivity, denoted by $\hat{s}_{b,f,smooth} = g(\hat{s}_{b,f})$.

Let $I_{b,f}$ denote the incidence of the bth bootstrapping sample for fortnight f. We therefore have that

$$I_{b,f} = \frac{\hat{s}_{b,f,smooth} - \hat{s}_{b,f-1,smooth}}{1 - \hat{s}_{b,f-1,smooth}}$$

Figure D.2 presents estimates of the incidence of seroreactivity among susceptible persons.

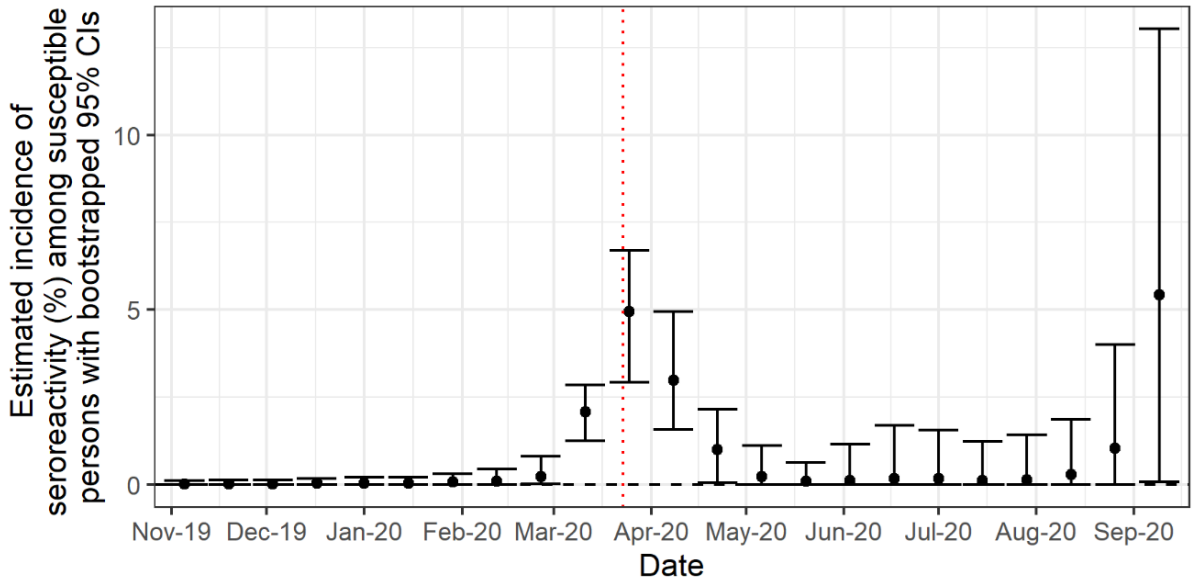

Figure D.2: Estimated incidence of seroreactivity among susceptible persons over time with 95% bootstrapped confidence intervals. The red dotted line indicates the introduction of the first "lockdown" in the UK.

D.7 Comparison of estimates with REACT-2

Figure D.3 presents our estimates of the prevalence of seroreactivity along with the estimates from the REACT-2 study at a national-level and only within London [235].

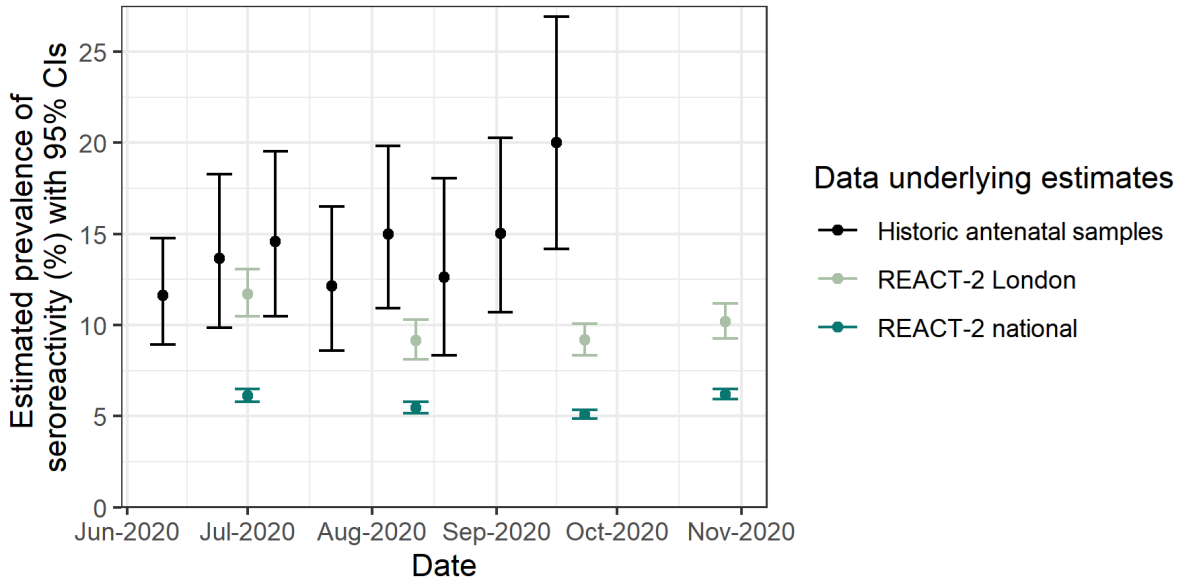

Figure D.3: Estimates of prevalence of seroreactivity using historic antenatal samples in this study (black) compared to 18–44-year-old women across our ethnicity and IMD groups from REACT-2 at a national-level (dark green) and only within London (light green).

D.8 Internal and external validation of June 2019 seroreactive and non-seroreactive samples

Table D.3 presents validation of DABA testing of June 2019 samples, whereby the 8 seroreactive samples in that period and another 16 consecutive negative controls were re-tested blind at UKHSA Colindale with NP capture and S1 capture assays. No positive samples by DABA were positive for NP or S1 capture. After blocking for seasonal coronaviruses, all non-seroreactive samples were confirmed as non-seroreactive.

Positive samples were also re-tested, unblinded, at Imperial College London on S1 capture assays. One of the eight samples seroreactive on DABA was seroreactive on S1 capture IgM assay (Table D.3).

Table D.3: Binding ratios (BR) of validatory testing conducted on June 2019 samples seroreactive under DABA assay along with 16 consecutive negative controls. Samples were re-tested blind at UKHSA (PHE) with NP capture and S1 capture assays. Positive samples were also re-tested, unblinded, at Imperial College London on S1 capture assays. Under all assays, a sample is considered seroreactive if $BR \geq 1$.

Assay	**DABA (BR)**	**PHE NP Capture (no block) (BR)**	**PHE NP Capture (with block) (BR)**	**PHE S1 Capture (BR)**	**Imperial S1 Capture Assay for Immunoglobulin A (BR)**	**Imperial S1 Capture Assay for Immunoglobulin G (BR)**	**Imperial S1 Capture Assay for Immunoglobulin M (BR)**
	2.20	0.17	0.20	0.13	0.10	0.20	0.40
June	11.94	0.17	0.19	0.13	0.20	0.10	1.30
2019	2.43	0.27	0.30	0.13	0.20	0.20	0.70
samples	1.98	0.17	0.21	0.13	0.20	0.20	0.40
tested	1.87	0.21	0.18	0.13	0.30	0.20	0.50
positive	2.49	0.22	0.25	0.18	0.20	0.20	0.30
under	1.81	0.18	0.16	0.13	0.20	0.20	0.20
DABA	1.18	0.17	0.28	0.13	0.20	0.20	0.20
	0.20	0.46	0.23	0.13			
	0.30	0.17	0.19	0.12			
	0.21	0.29	0.25	0.16			
	0.10	0.75	0.21	0.18			
June	0.23	0.19	0.30	0.13			
2019	0.13	0.53	0.41	0.13			
samples	0.10	0.26	0.36	0.13			
tested	0.15	0.19	0.19	0.13			
negative	0.27	0.21	0.17	0.12			
on DABA,	0.32	0.21	0.19	0.13			
consecutive	0.21	0.21	0.25	0.18			
controls	0.16	0.29	0.19	0.13			
	0.09	1.05	0.15	0.13			
	0.60	0.20	0.16	0.13			
	0.21	0.22	0.17	0.11			
	0.18	0.99	0.22	0.18			

Appendix E

Appendix: Paper V: Alternative epidemic indicators for COVID-19 in settings with incomplete death registration systems

Supplementary material for Paper V (Chapter 7)

R McCabe*, C Whittaker, RJ Sheppard, N Abdelmagid, A Ahmed, IZ Alabdeen, NF Brazeau, AEA Abd Elhameed, AS Bin-Gouth, A Hamlet, R AbuKoura, G Barnsley, JA Hay, M Alhaffar, EK Besson, SM Saje, BG Sisay, SH Gebreyesus, AP Sikamo, A Worku, YS Ahmed, DH Mariam, MM Sisay, F Checchi, M Dahab, BS Endris, AC Ghani, PGT Walker, CA Donnelly and OJ Watson*. Alternative epidemic indicators for COVID-19 in three settings with incomplete death registration systems. *Science Advances* 9, eadg7676 (2023). doi: `https://doi.org/DOI:10.1126/sciadv.adg7676`.

E.1 Materials and Methods

We expand on the specific data sets and sources for each modelled city.

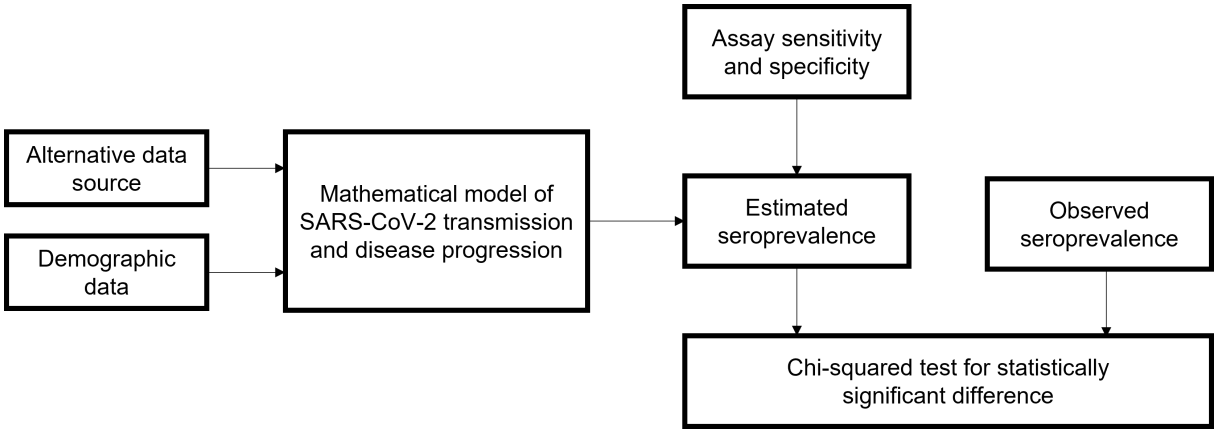

Figure E.1: Schematic diagram of the methodology used in each setting. Within each setting, the SARS-CoV-2 transmission model used in this analysis is calibrated using the alternative data source and demographic data. From the model and assay-specific sensitivity and specificity, we estimate seroprevalence and compare this to that which was observed in independent studies in each setting. Finally, a chi-squared test is used to test whether the difference between estimated and observed seroprevalence is statistically significant.

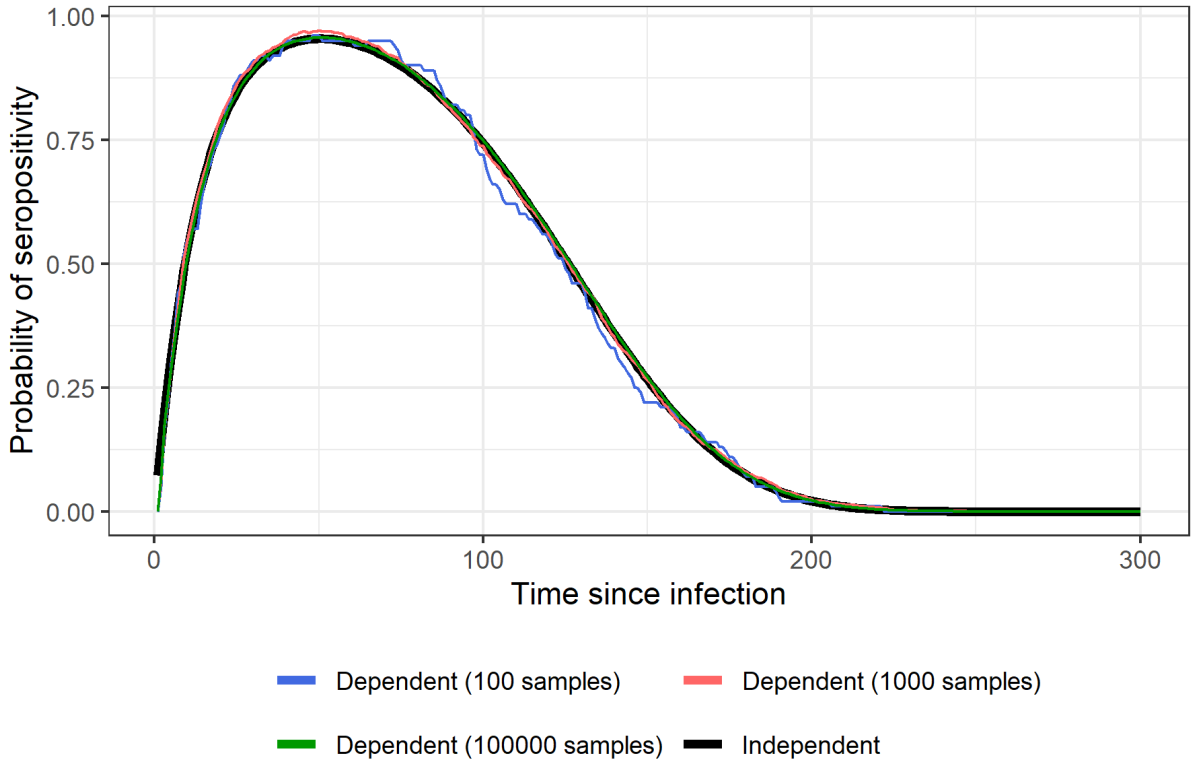

Figure E.2: Probability of seropositivity with and without the assumption of independence between IgG and IgM antibodies.

Table E.1: Further details of seroprevalence assays used. Note that this table omits the alternative data source information (shown in Table 7.1) for the sake of space.

Setting	**Seroprevalence**			
	Overview	**Assay details**	**Reported estimate**	**Source**
Addis Ababa, Ethiopia	Random sample of 956 households 22/07/2020 - 10/08/2020	Core Technology IgM/IgG rapid test IgG: sensitivity 91.4%; specificity 98.6% IgM: sensitivity 89.4%; specificity 98.4%	IgG: 1.9% (0.4% - 3.7%) Combined IgG/IgM: 3.5% (1.7% - 5.4%)	[486]
Aden, Yemen	Cross-sectional household study of 2001 people 28/11/2020 - 13/12/2020	Healgen COVID-19 IgG/IgM rapid test cassette; confirmed via enzyme-linked immunosorbent assay (ELISA) IgG: sensitivity 96.7% (83.3% - 99.4%) IgM: sensitivity 100% (88.7% - 100%) IgG or IgM: sensitivity 100% (88.7% - 100%) specificity 97.5% (91.3% - 99.3%)	IgG: 25.0% (23.2% - 26.9%) IgM: 0.2% (0.1% - 0.4%) Combined IgG/IgM: 27.4% (25.6% - 29.3%)	[256]
Khartoum, Sudan	Cross-sectional household study of 2,375 people 01/03/2021 - 10/04/2021	STANDARD Q COVID-19 IgM/IgG Combo rapid diagnostic test (RDT) (SD-Biosensor) ELISA (Anti-SARS-CoV-2 ELISA [IgG, S1 domain]; Euorimmun) conducted on subsample to estimate study-specific sensitivity and specificity using a meta-analysis of test characteristics with random effects and a Bayesian latent class model. IgG: sensitivity 61.9% (56.8% - 66.9%) IgG: specificity 98.9% (98.2% - 99.5%)	Combined IgG/IgM (corrected for sensitivity/specificity): 54.6% (51.4% - 57.8%)	[487]

E.2 Addis Ababa

E.2.1 COVID-19 deaths in Addis Ababa

At the time of writing, we were unable to obtain a fully complete daily time series of confirmed COVID-19 deaths in Addis Ababa. Instead, we estimated this using information published by the Ethiopian Public Health Institute and trends in national cases in Ethiopia reported to the World Health Organization (WHO) [488]. Based on these data, we estimated daily reported COVID-19 deaths for three distinct time periods before combining into one time series for 2020.

The first confirmed death in Ethiopia was registered on 6 April 2020 and was reported as having occurred in Addis Ababa [533]. By 21 June 2020, Watare et al. [534] reported a total of 63 COVID-19 deaths in Addis Ababa compared to 72 at a national level (87.5%). Therefore, the first time period considered was 6 April - 21 June 2020 and the daily time series was estimated by assuming that each day, the number of deaths in Addis Ababa was 87.5% of the deaths at a national-level rounded to the nearest integer. This produced 66 deaths in total for Addis Ababa for this period, and so three dates were randomly selected from this period and one death was removed from that date's death toll.

The Ethiopian Public Health Institute (EPHI) [488] reported a total of 295 cumulative COVID-19 deaths in Addis Ababa as of 9 August 2020, resulting in 232 deaths in Addis Ababa during this period compared to 318 nationally (73%). Using the same method as above, the second time period for which a daily time series was constructed was 22 June - 9 August 2020. In this instance, there were two fewer deaths than the total and two dates on which national deaths were recorded were randomly selected in which to add one additional death to.

Finally, from 10 August 2020 the EPHI began publishing weekly COVID-19 mortality figures for Addis Ababa (with the exception of the week commencing 19 October, for which the weekly total was assumed to be the median, rounded to the nearest integer, of the two surrounding weeks). The time period from 10 August until 31 October 2020 was the third period under consideration. Within each week, the daily number of COVID-19 deaths was assumed to follow the same proportion of deaths falling on that weekday at national level. The three time series were then joined into a coherent time series from 6 April - 31 October 2020.

E.2.2 Excess deaths in Addis Ababa

Excess mortality was estimated using cemetery surveillance data gathered by Endris et al. [484], which detailed the number of burials per day at all cemeteries ($n = 79$ cemeteries) in Addis Ababa across January 2015 - January 2021. All cemeteries were actively under surveillance for the whole period, except in 2019 when the surveillance programme was reduced with only 10 cemeteries actively under surveillance. These data gaps were filled in 2020 by conducting secondary data collection from the cemeteries, however, there is a significant decrease in deaths in 2019 that may be due to this change in surveillance in 2019. In response, we considered two mortality baselines using data from 2015 - 2019 and exclusively from 2019, respectively. We fit a negative binomial distribution to the total number of daily burials across cemeteries per calendar month. Baseline mortality was then estimated by sampling 100 times from each distribution per date, with excess mortality calculated per sample by taking the difference between the observed and baseline number of deaths per day. Our finalised time series of excess mortality using these data was averaged over each sample and presented as the rolling 7-day mean.

Our decision to use the excess mortality data in this way was due to the absence of a significant trend over time under a negative binomial regression model ($p = 0.408$) for the years 2015 - 2018 (Figure E.3B). When extending the data to also include 2019, time became significant ($p = 0.028$), which is in accordance with the significant reduction in deaths in 2019 compared to the average across 2015 - 2018 as reported in Chapter 7. However, given the noted surveillance bias in 2019 it seemed inaccurate to use data in 2019 as part of a trend analysis to predict baseline mortality in 2020. Therefore, we decided to focus our analysis on the impact of the decrease in all-cause mortality in 2019 and how this related to seroprevalence compared to previous years' mortality, rather than fitting a single model to capture this trend.

E.2.3 Seroprevalence in Addis Ababa

We used estimates of seroprevalence in Addis Ababa from Abdella et al. [486], who conducted a serosurvey of 956 individuals to detect both IgG and combined IgG/IgM antibodies in this area across 22 July - 10 August 2020 using the *Core Technology IgM/IgG rapid test*. This assay has a manufacturer reported sensitivity of 91.4% and 89.4% for IgG and IgM antibodies, respectively. Although confidence intervals are not reported by the manufacturer, Abdella et al. [486] report the results of an independent trial of 200 samples in which IgG and IgM specificity was estimated as 97.0% (95% CI $94.6\% - 99.4\%$) and 98.0% (95% CI $96.0\% - 99.9\%$), respectively. For the purposes of our modelling we used the sensitivity and specificity values estimated by the manufacturer.

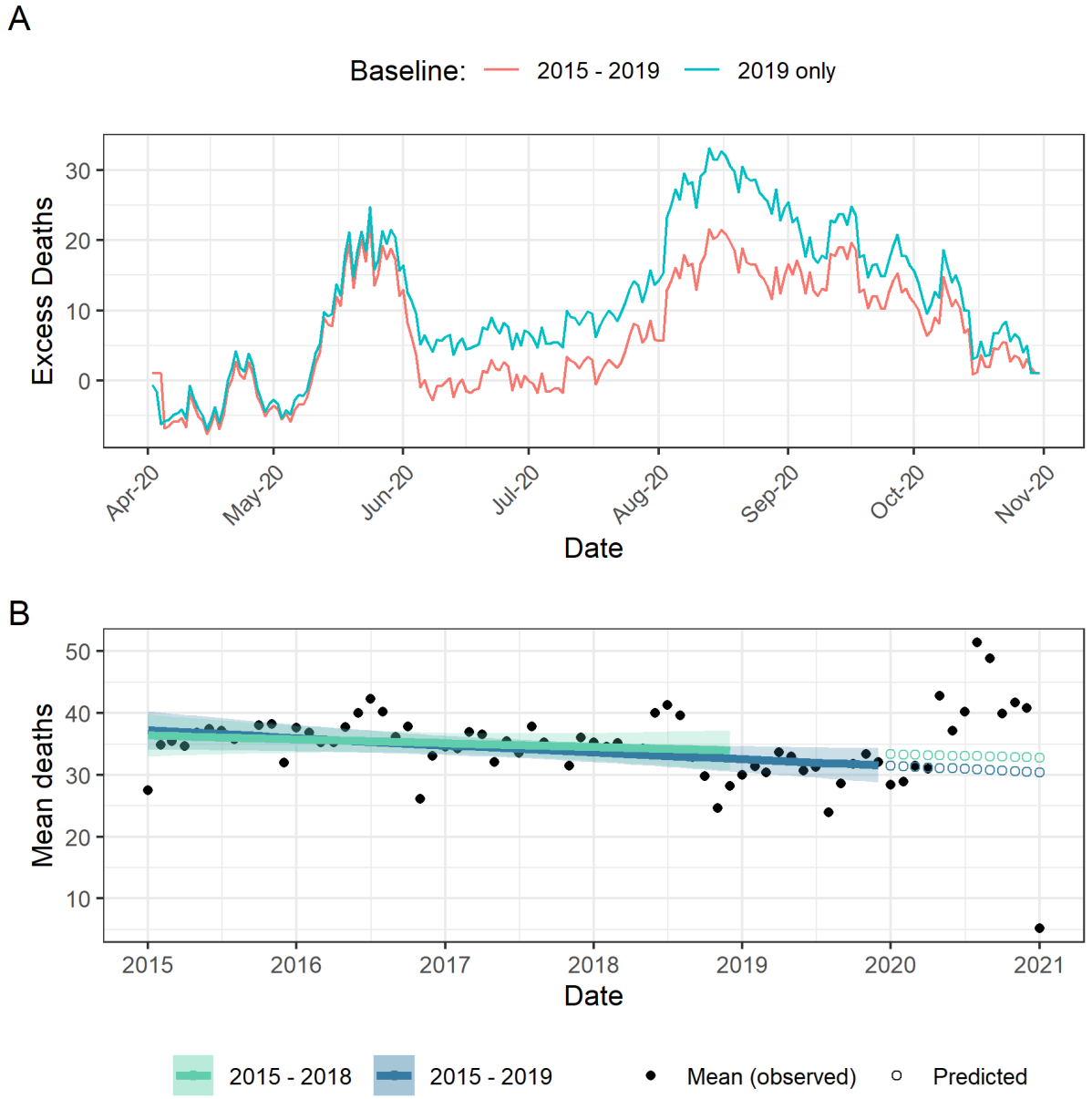

Figure E.3: (A) Excess deaths between April 2020 and November 2020 are shown for two methods (using all years between 2015-2019, or just 2019) of calculating the assumed baseline mortality during 2020 in the absence of the pandemic. (B) Black points show estimated mean monthly deaths (described by monthly fitted negative binomial distributions) 2015 - 2021. Linear trends based on data from 2015 - 2018 (green) and 2015 - 2019 (blue) are shown, revealing the absence of a significant trend over time under a negative binomial regression model ($p = 0.408$) for the years 2015 - 2018.

We were unable to obtain specific seroconversion and seroreversion rates for the assay used in Abdella et al.'s [486] study and therefore assumed our default parameters as set out above. Although we were unable to explain the large difference observed between IgG and combined IgG/IgM antibodies, even when varying the seroreversion half-life of IgM, we hypothesise that this could be attributed to lack of validation of the assay in low-income countries.

E.2.4 Demography and healthcare in Addis Ababa

Age-stratified estimates of the population of Addis Ababa were derived by assuming that the proportions of individuals in each age group were those observed by Timotewas et al. [535] in 2012 - 2013 (summed across male and females), and multiplied by the most recent estimate of the population of Addis Ababa in 2020 (4.8 million) [536].

Laytin et al. [537] estimated 130 intensive care unit (ICU) beds in Addis Ababa in January 2020. Assuming that ICU beds are 1.5% of all hospital beds in low-income countries [538], we estimated there to be $6,867$ acute hospital beds in Addis Ababa.

E.2.5 Model fits and estimated seroprevalence

We fit the model up to 31 October 2020, by which point the serosurvey had taken place and the first wave of reported COVID-19 deaths had plateaued. Figures E.4 - E.6 present the model fits to COVID-19 and excess deaths under different baselines and the associated model-estimated seroprevalence.

Figure E.7 is analogous to Figures 7.1C and 7.1D in the main text, but also includes the unweighted seroprevalence reported by Abdella et al. [486]. The unweighted seroprevalence estimates are greater than their corresponding weighted estimates. The unweighted estimates do not account for the demography of the sampled population. When comparing our modelled seroprevalence to the reported unweighted seroprevalence, there are fewer scenarios that are not statistically different when compared to the weighted seroprevalence. Scenarios that were not statistically different included mortality derived using data from all years to inform the baseline and using 2019 from June 2020 onwards produced estimates of IgG and combined IgG/IgM antibodies consistent with the unweighted results in Abdella et al. [486] ($p = 0.779$; $p = 0.487$; $p = 0.282$; $p = 0.925$, respectively), as did the model fit to reported COVID-19 deaths for IgG antibodies only ($p = 0.112$).

Table E.2: p-values from Chi-squared test (df = 1) for difference between seroprevalence reported by Abdella et al. [486] and estimated seroprevalence in Addis Ababa modelled under different mortality time series for different antibody types. $p > 0.05$ indicates no statistically significant difference at the 0.05 level. Abdella et al. reported both unweighted and weighted estimates of seroprevalence from their study. The unweighted estimates are greater than the corresponding weighted estimates and in general showed a weaker agreement with modelled seroprevalence.

	COVID-19 deaths	Excess mortality scenarios			
		2019 baseline	2019 baseline without 1st peak	2015-2019 baseline	2015-2019 baseline without 1st peak
Comparison to weighted seroprevalence reported in [486]					
IgG	0.884	<0.001	0.072	0.102	0.474
Combined IgG/IgM	0.334	0.002	0.347	0.765	0.074
Comparison to unweighted seroprevalence reported in [486]					
IgG	0.112	<0.001	0.487	0.771	0.007
Combined IgG/IgM	0.007	0.014	0.925	0.282	<0.001

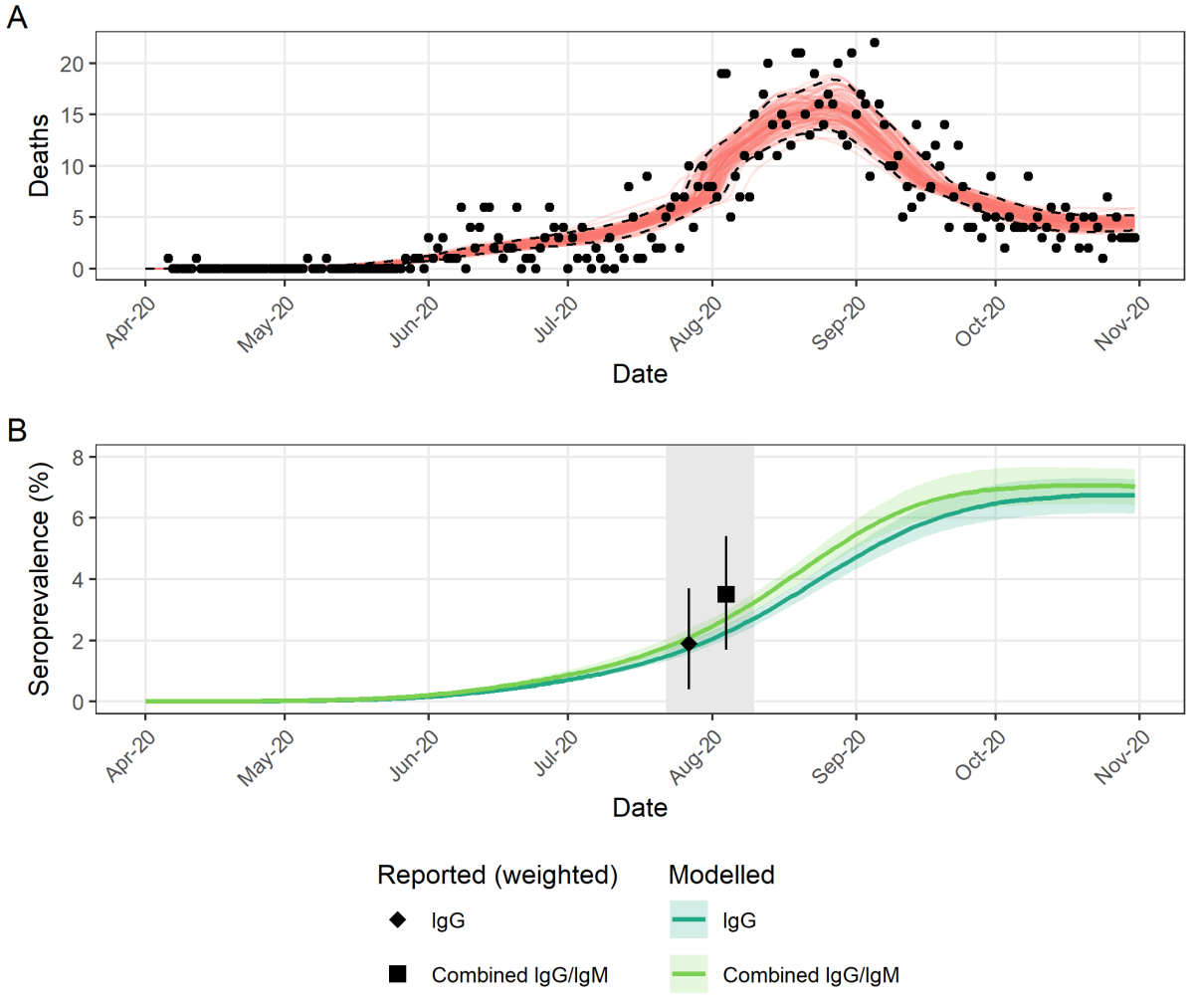

Figure E.4: (A) Model fit (red lines) to reported COVID-19 deaths (black points) in Addis Ababa. (B) Estimated seroprevalence from model fit to COVID-19 deaths in (A) (median and 95% credible intervals) compared to reported values by Abdella et al. [486]. Grey shaded area highlights the sampling period of the serosurvey, with weighted reported estimates both corresponding to this entire period.

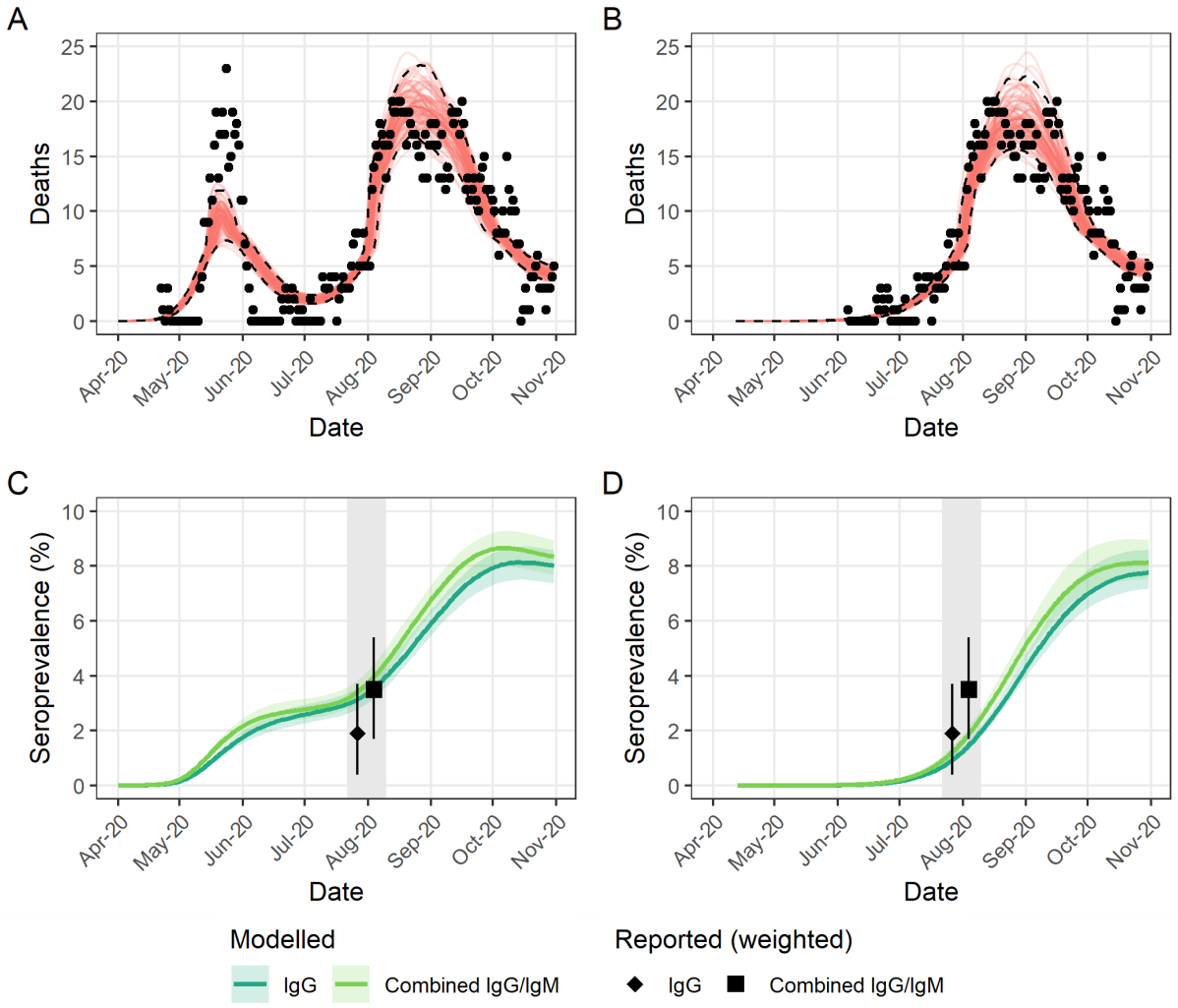

Figure E.5: Model fit and estimated seroprevalence under excess mortality estimated using data from 2015 - 2019 in Addis Ababa. (A) Model fit to excess mortality. (B) Model fit to excess mortality without the first peak observed across May and June 2020. In (A) and (B), red lines represent model realisations and black dots show estimated excess deaths. (C) Estimated seroprevalence (median and 95% credible intervals) of IgG and combined IgG/IgM antibodies under model presented in (A). (D) Estimated seroprevalence (median and 95% credible intervals) of IgG and combined IgG/IgM antibodies under model presented in (B). In (C) and (D), black points indicate weighted seroprevalence as reported by Abdella et al. [486] and grey shaded area highlights the sampling period of the serosurvey, with weighted reported estimates both corresponding to this entire period.

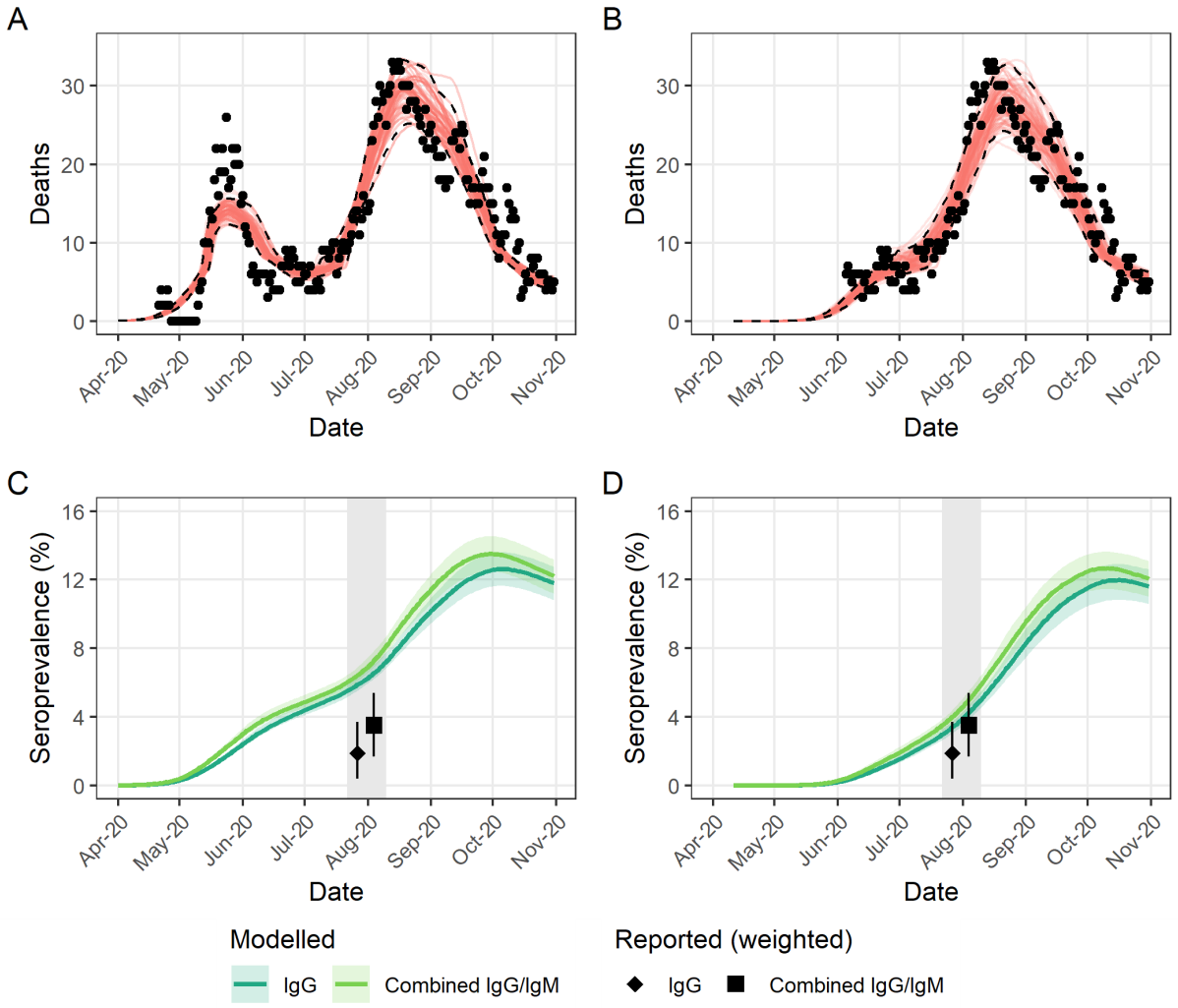

Figure E.6: Model fit and estimated seroprevalence under excess mortality estimated using data from 2019 only in Addis Ababa. (A) Model fit to excess mortality. (B) Model fit to excess mortality without the first peak observed across May and June 2020. In (A) and (B), red lines represent model realisations and black dots show estimated excess deaths. (C) Estimated seroprevalence (median and 95% credible intervals) of IgG and combined IgG/IgM antibodies under model presented in (A). (D) Estimated seroprevalence (median and 95% credible intervals) of IgG and combined IgG/IgM antibodies under model presented in (B). In (C) and (D), black points indicate weighted seroprevalence as reported by Abdella et al. [486] and grey shaded area highlights the sampling period of the serosurvey, with weighted reported estimates both corresponding to this entire period.

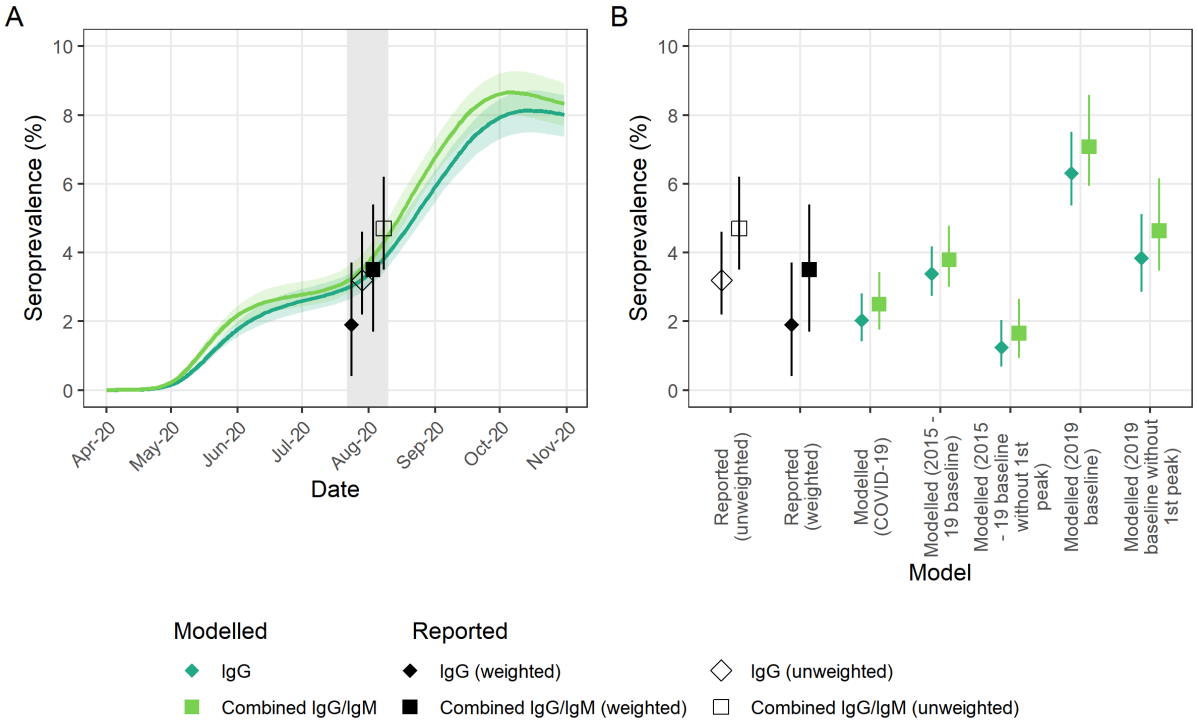

Figure E.7: Seroprevalence in Addis Ababa, Ethiopia in 2020 with reported weighted and unweighted values. (A) Estimated seroprevalence from model fit to excess mortality under 2015 - 2019 baseline with the first peak compared to reported values from Abdella et al. [486]. Grey shaded area highlights the sampling period of the serosurvey, with reported estimates both corresponding to this entire period. (B) Seroprevalence estimated under models fit to different estimates of mortality from COVID-19 and excess mortality with different baselines compared to reported seroprevalence from [486].

E.3 Aden

E.3.1 COVID-19 deaths in Aden

At the time of writing, we were also unable to obtain from a formal ministry of health source the daily time series of confirmed COVID-19 deaths in Aden. Instead, we scraped the cumulative number of deaths from tweets directly referencing Aden from the Twitter account of the Yemen Supreme National Emergency Committee for COVID-19 [489] and converted this into a daily time series for reported COVID-19 deaths.

E.3.2 Excess deaths in Aden

We used excess mortality in Aden as estimated by Koum Besson et al. [220], who derived estimates of excess mortality using satellite images of all active cemeteries in Aden. We converted the cumulative excess data into a daily time series and applied a rolling 7-day mean to the daily estimated number of excess deaths.

E.3.3 Seroprevalence in Aden

Estimates of the seroprevalence of combined IgG/IgM antibodies in Aden across 28 November - 13 December 2020 were sourced from Bin-Ghouth et al. [256], using the *Healgen COVID 19 IgG/IgM Rapid Test Cassette* and confirmed via the *WANTAI SARS-CoV-2 Ab ELISA*. They recruited 2,001 participants to their study. The reported ELISA sensitivity and specificity by the manufacturer are: IgG sensitivity = 96.7% (83.3% − 99.4%), IgM sensitivity = 100% (88.7% − 100%). In the absence of assay-specific estimates of rates of seroconversion, we used the default assumptions as set out above. However, as noted in the main text, due to the long delay between the peak in excess mortality and the implementation of the serosurvey in addition to the lack of assay-specific estimates of rates of seroreversion, we conducted a sensitivity analysis of different IgG and IgM seroreversion half-lives to assess the uncertainty in these parameters.

E.3.4 Demography and healthcare in Aden

We used an analogous methodology to derive age-stratified estimates of the population of Aden as detailed above for Addis Ababa. We used the population pyramid provided by Bawazir [539]. We split the age category "75+" (in years) into "$75-79$" and "$80-84$", as required by our model, by assuming the proportions of individuals within these two categories followed that estimated nationally by the United Nations World Population Prospects 2019 [540]. We multiplied this by the estimated population of Aden (1 million) as assumed in Koum-Besson et al. [220].

Official data on healthcare capacity in Aden were not available at the time of writing. International media reported 18 ICU beds in Aden in June 2020 [541]. We verified this figure by assuming that the reported 520 ICU beds in Yemen were spread across the country proportional to the population of each area. Assuming Aden accounts for 3.35% of the population (total population of Yemen 29.83 million [540]), this implies 17 ICU beds in this area. Under the same assumption of 1.5% of hospital beds are ICU beds as above in Murthy et al. [538], we estimate there to be $1,133$ acute hospital beds in Aden.

E.3.5 Model fits and estimated seroprevalence

Figure E.8 presents the model fit to reported COVID-19 deaths and the resulting estimated seroprevalence of IgM, combined IgG/IgM and IgM antibodies over time. We present estimated seroprevalence under the minimum and maximum seroreversion IgG and IgM half-lives included in our sensitivity analysis, 100 days and 280 days for IgG and 30 days and 70 days for IgM, neither of which can explain the observed seroprevalence.

As our excess mortality time series ended on 16 September 2020 (the end of Koum Bessom et al.'s study [220]), we used our model to project forward the number of daily deaths from this point to the end of 2020 assuming that $R(t)$[1] remained at the level observed on this date. This is illustrated in Figure E.9. Figure E.10 presents multiple model fits to the Aden epidemic under different assumptions of the population average IFR. These figures indicate that an IFR of 0.2% is unable to recreate the wave of excess mortality in Aden, which is due to the depletion of the susceptible population.

Figure E.11 shows the results of a sensitivity analysis of varying the IgM half-life between 30 and 70 days, of which there is minimal impact on the results in comparison to the reported values for IgM and combined IgG/IgM antibodies.

[1] The effective reproduction number.

Table E.3: p-values from Chi-squared test (df = 1) for difference between seroprevalence observed by Bin-Ghouth et al. [256] and estimated seroprevalence in Aden under model fitted to satellite-derived excess mortality, modelled under different assumptions of the IFR and seroreversion half-life. $p > 0.05$ indicates no statistically significant difference at the 0.05 level.

		IFR			
		0.2	**0.3**	**0.4**	**0.5**
IgG					
IgG seroreversion half-life (days)	**100**	<0.001	<0.001	<0.001	<0.001
	120	<0.001	<0.001	<0.001	<0.001
	140	<0.001	<0.001	<0.001	<0.001
	160	0.155	<0.001	<0.001	<0.001
	180	0.007	0.651	0.068	<0.001
	200	<0.001	0.001	0.289	0.089
	220	<0.001	<0.001	0.001	0.649
	240	<0.001	<0.001	<0.001	0.005
	260	<0.001	<0.001	<0.001	<0.001
	280	<0.001	<0.001	<0.001	<0.001
Combined IgG/IgM					
IgG seroreversion half-life (days) with IgM half-life held constant at 50 days	**100**	<0.001	<0.001	<0.001	<0.001
	120	<0.001	<0.001	<0.001	<0.001
	140	<0.001	<0.001	<0.001	<0.001
	160	0.014	<0.001	<0.001	<0.001
	180	0.089	0.563	0.003	<0.001
	200	<0.001	0.015	0.944	0.002
	220	<0.001	<0.001	0.001	0.385
	240	<0.001	<0.001	<0.001	0.130
	260	<0.001	<0.001	<0.001	<0.001
	280	<0.001	<0.001	<0.001	<0.001
IgM					
IgM seroreversion half-life (days)	**30**	0.012	0.019	0.025	0.026
	35	0.016	0.031	<0.001	<0.001
	40	0.023	0.060	0.090	0.089
	45	0.042	0.131	0.191	0.183
	50	0.103	0.307	0.409	0.376
	55	0.249	0.567	0.683	0.623
	60	0.668	0.990	0.932	0.975
	65	0.799	0.662	0.634	0.734
	70	0.288	0.334	0.351	0.439

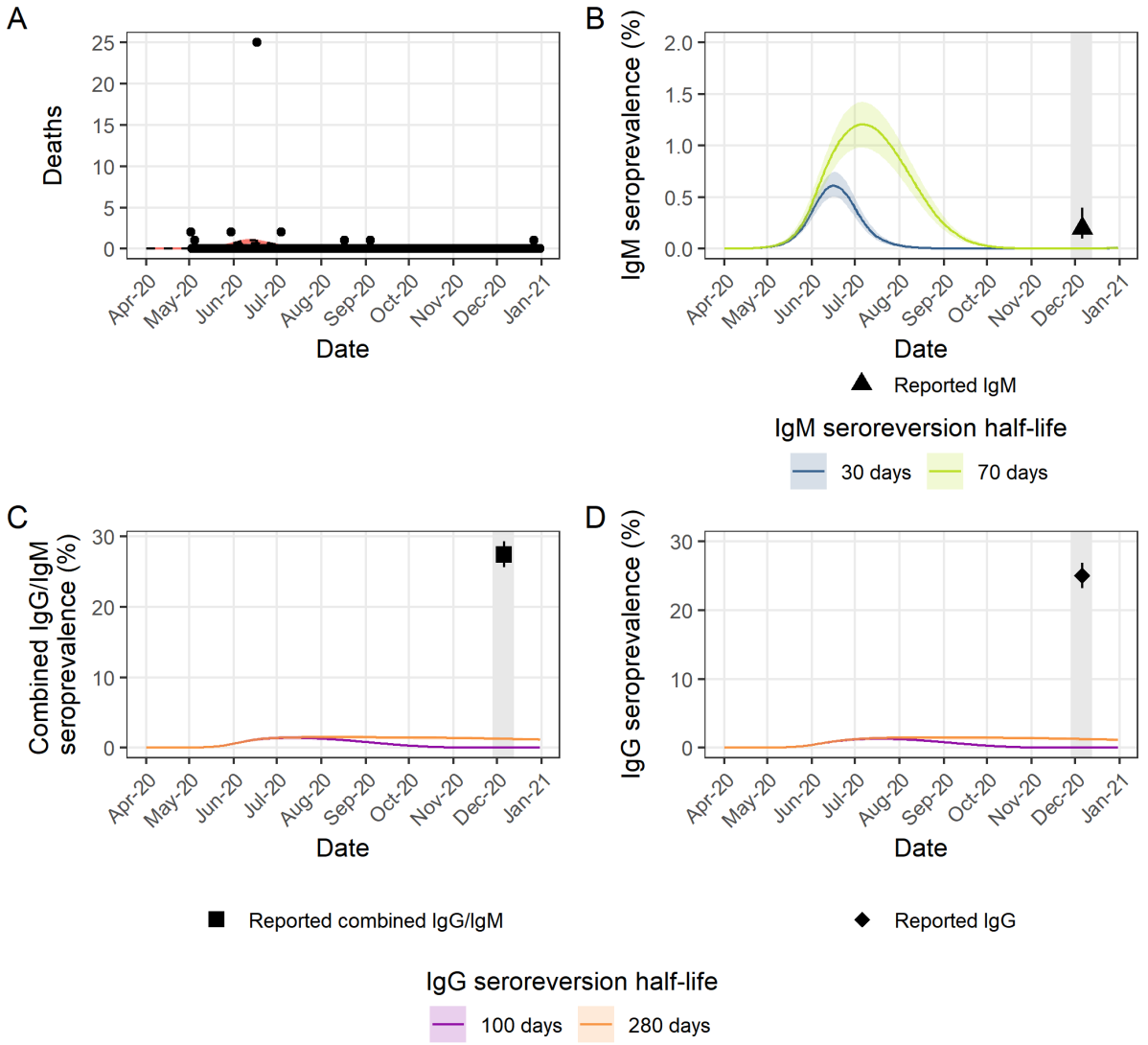

Figure E.8: (A) Model fit to reported COVID-19 deaths in Aden under default infection fatality ratio (IFR=0.3%). Red lines indicate model realisations and black dots show reported COVID-19 deaths. (B) Estimated seroprevalence (median and 95% credible intervals) of IgM antibodies under minimum (30 days) and maximum (70 days) IgM seroreversion half-lives included in our analysis compared to the observed value. Reported IgM by Bin-Gouth et al. [256] is 0.2% (95% CI: 0.1% – 0.4%). (C) Estimated seroprevalence (median and 95% credible intervals) of combined IgG/IgM antibodies under minimum (100 days) and maximum (280 days) seroreversion half-lives included in our sensitivity analysis compared to observed value. IgM half-life held constant at 50 days. Reported value is 27.4% (95% CI: 25.6% – 29.3%) [256]. (D) As in (C) but for IgG antibodies. Reported value is 25.0% (95% CI: 23.2% – 26.9%) (19).

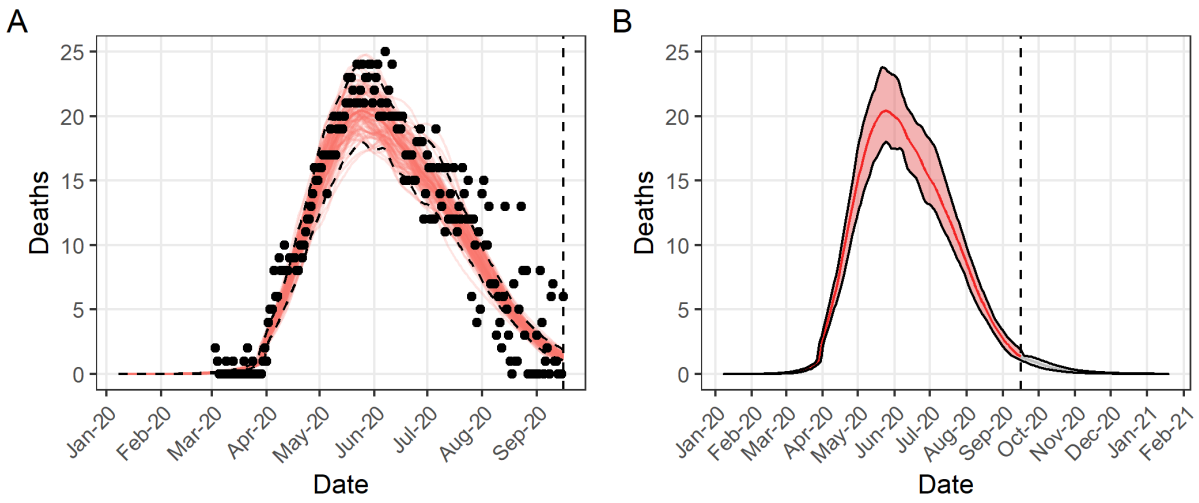

Figure E.9: (A) Model fit to excess mortality in Aden. Red lines indicate model realisations and black dots show excess deaths. (B) Projection from mid-September (median and 95% credible intervals) to end of 2020 (grey) based on fit in (A) (red) assuming $R(t)$ remains at value estimated on the last day of model fit (16 September 2020).

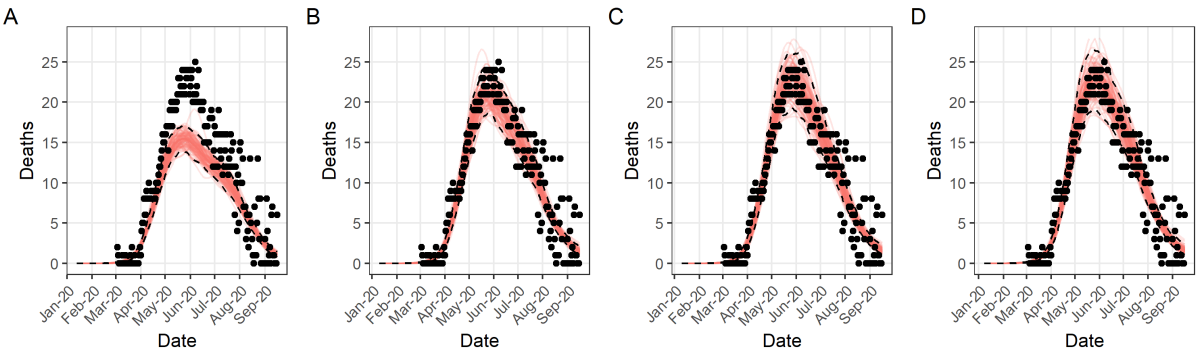

Figure E.10: Model fits to excess mortality in Aden under different assumptions of the Infection Fatality Ratio (IFR). (A) IFR = 0.2; (B) IFR = 0.3; (C) IFR = 0.4; (D) IFR = 0.5. In all panels, red lines indicate model realisations and black dots show excess deaths.

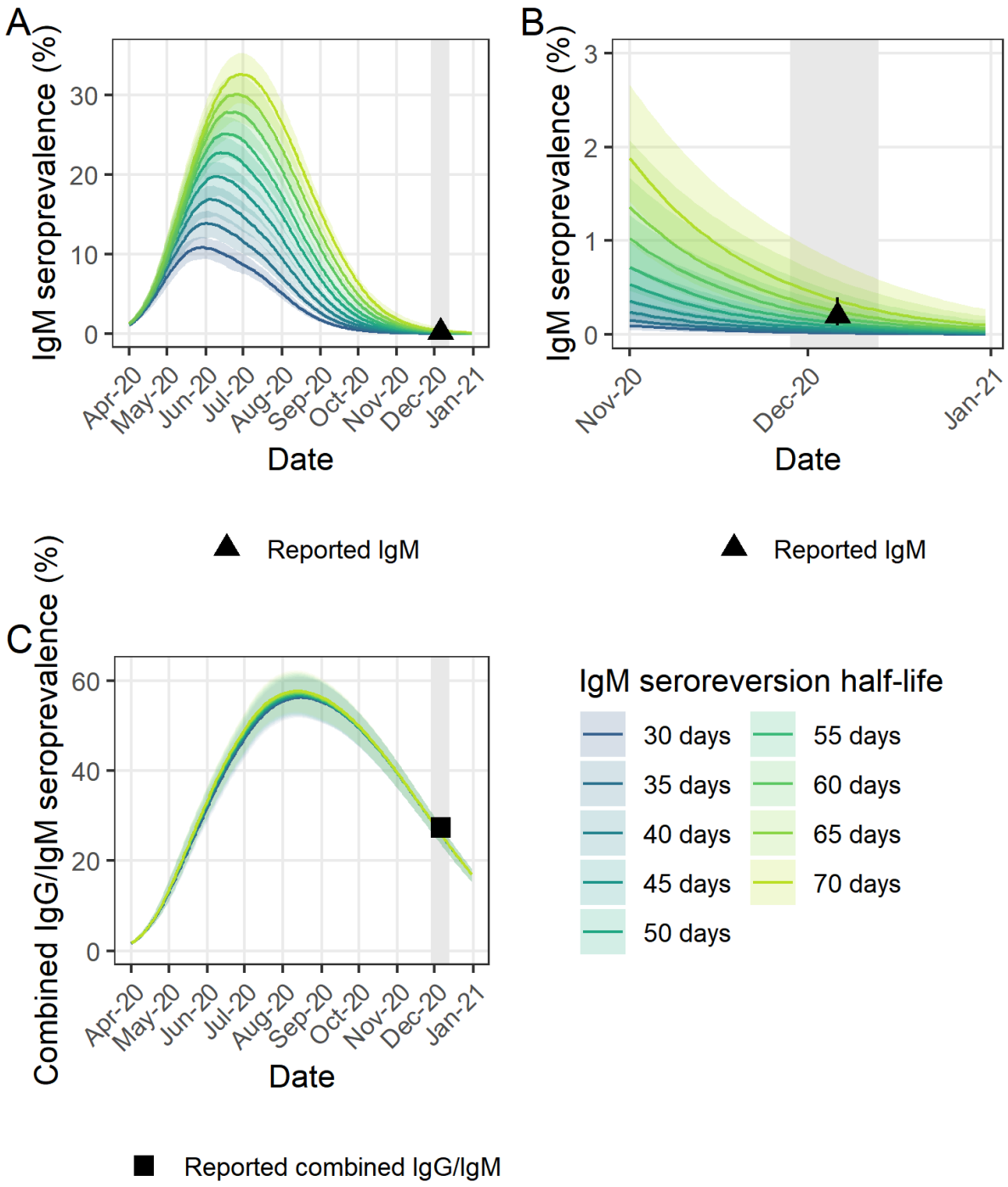

Figure E.11: Estimated seroprevalence in Aden under model fit to excess mortality under default infection fatality ratio (IFR=0.3%). (A) Estimated seroprevalence of IgM antibodies in Aden under different assumptions of the IgM seroreversion half-life compared to that reported by Bin-Gouth et al. [256] (0.2% (95% CI 0.1% - 0.4%)). (B) As in (A) but restricted to November - December 2020 only. (C) Estimated seroprevalence of combined IgG/IgM antibodies under assumptions of the IgM seroreversion half-life compared to that reported in [256] (27.4% (95% CI 25.6% – 29.3%). IgG seroreversion half-life is held constant at 180 days.

E.4 Khartoum

E.4.1 COVID-19 deaths in Khartoum

COVID-19 deaths for Khartoum were scraped from daily epidemiological reports published by the Sudan Health Observatory [542]. On 8 November 2020, the Sudanese Federal Ministry of Health (FMOH) released a periodic revision of COVID-19 data, which included an additional 99 deaths for Khartoum, increasing the cumulative total COVID-19 deaths to 414 by 8 November. These additional 99 deaths were reported to have occurred any time since the beginning of the pandemic. In response, we distributed the additional 99 deaths in proportion to the weekly mean number of COVID-19 deaths reported in the daily epidemiological reports, which ensures the same shape of the epidemic but with an increased magnitude. Lastly, on 17 November 2020, the FMOH released a statement detailing that an additional 46 deaths had been reported from private diagnostic laboratories and had occurred between 8 November 2020 and 17 November 2020 [543]. During this period, only 1 death had been reported in FMOH epidemiological reports, however, it was subsequently followed by a sharp, exponential increase in reported COVID-19 deaths in epidemiological reports. In response, we assumed that the additional 46 deaths were exponentially increasing between 8 November and 17 November 2020, which also aligns with the increase in reported COVID-19 cases in epidemiological reports.

E.4.2 Symptomatic cases in Khartoum

We used estimates of the cumulative number of symptomatic COVID-19 cases in individuals over the age of 15 in Khartoum, which were estimated using an online survey as described in Mohamed et al. [485]. The survey form was distributed through social media channels in Khartoum between 26 May to 3 June 2020 and collected information on whether participants had suffered symptoms associated with COVID-19 since the start of the pandemic. Symptomatic individuals were subsequently asked to report which of any of 13 different symptoms they had suffered and all individuals were asked both i) if they had received a SARS-CoV-2 diagnostic test and ii) the outcome of any test taken. From this survey, the authors created a statistical model to infer the SARS-CoV-2 infection status of survey participants who had not received a SARS-CoV-2 test based on reported symptoms and through this, infer the symptomatic attack rate in the general population by 3 June. We use three estimates of the symptomatic attack rate in 15+ year olds from the study to reflect the uncertainty in the inferred attack rates: 8.3%, 11.0% and 13.7%.

E.4.3 Seroprevalence and mortality survey in Khartoum

We leveraged a cross-sectional, household survey in Omdurman, Sudan, conducted during March - April 2021 by Médecins sans frontières [487]. Omdurman is the largest city in terms of population within Khar-

toum state and represents approximately one-third of the population of Khartoum state. All members of a subset of households, regardless of age were invited to participate in a seroprevalence survey, which recruited 3808 individuals. The survey used the SD-Biosensor STANDARD Q COVID-19 IgM/IgG Combo rapid diagnostic test (RDT), with a subset of dried blood spots also collected for analysis by ELISA (Anti–SARS-CoV-2 ELISA [IgG, S1 domain]; Euroimmun). In the analysis conducted by Médecins sans frontières [487], crude seroprevalence estimates from the RDT and ELISA results were adjusted using published performance estimates for the RDT and ELISA in a meta-analysis with random effects model before being used as inputs for a Bayesian latent class model to yield adjusted seroprevalence estimate of 54.6% (95% CI 51.4%–57.8%).

In the same household survey, a mortality survey was conducted to investigate the death rate for the pre–pandemic (January 1, 2019–February 29, 2020) and pandemic (March 1, 2020 – day of the survey). Data from 27,315 people across 3.716 households collected across the entire recall period showed a 67% (95% CI: 32%–110%) increase in crude death rate between the pre–pandemic (0.12 deaths/10000 people/day [95% CI: 0.10–0.14]) and pandemic (0.20 [95% CI: 0.16–0.23]) periods. This equates to a 7,113 excess deaths across Omdurman (95% CI: 5,015–9,505). To compare this against the predicted COVID-19 deaths from our transmission models fit to data across all Khartoum state, we linearly scale these estimates by 292% to reflect the larger population size of Khartoum state (estimated to be 8,877,146 by the United Nations Office for the Coordination of Humanitarian Affairs (UNOCHA)) compared to Omdurman [487].

E.4.4 Demography and healthcare in Khartoum

We assume the total population size for Khartoum state is 8,877,146, as estimated by UN OCHA [544], with a demographic profile comparable to Sudan as estimated in the United Nations World Population Prospects 2019 [540]. We assume there are 4,640 general hospital beds based on recent per capita estimation by the non-profit organisation NextGen Sudan in October 2019 [545], and that 75 ICU beds are available based on Sulieman et al. [546].

Table E.4: p-values from Chi-squared test (df = 1) Khartoum for reporting fractions explored in Figure 7.3. $p > 0.05$ indicates no statistically significant difference at the 0.05 level.

Reporting fraction	**p-values from Chi-squared test (df=1)**
2%	< 0.001
3%	< 0.001
4%	< 0.001
4.5%	0.492
5%	0.118
6%	< 0.001
7%	< 0.001
8%	< 0.001
9%	< 0.001
10%	< 0.001

Table E.5: Estimated reporting fractions of COVID-19 deaths as a percentage of total estimated excess deaths using alternative sources in each setting. Uncertainty in total estimated excess deaths in Addis Ababa comes from consideration of different mortality baselines, with the 100% upper estimate.

Setting	**Time period (inclusive)**	**Total reported COVID-19 deaths**	**Total estimated COVID-19 deaths (corresponding to reporting fraction range)**	**Estimated reporting fraction**
Addis Ababa[2]	April 2020 - October 2020	1,064	1,064 - 1,549	69% - 100%
Aden[3]	March 2020 - September 2020	34	424 - 4,137	0.8% - 8.0%
Khartoum[4]	March 2020 - April 2021	938	15,630 - 31,270	3.0% - 6.0%

[2]100% upper limit chosen to reflect both the near agreement between the cemetery inferred excess mortality and reported COVID-19 deaths, and the non-significant difference between the seroprevalence inferred based on model fits to reported COVID-19 deaths and the reported seroprevalence. 69% lower limit reflects the excess mortality estimated based on the 2015 - 2019 scenario with the first peak in May included.

[3]Reporting fraction range represents the uncertainty in the satellite-inferred excess mortality (95% CI: $424 - 4,137$).

[4]Reporting fraction range represents the range of reporting fractions for which the corresponding modelled 95% confidence interval for the cumulative proportion of symptomatic infections overlapped with the reported lower or upper bounds of 8.3% and 13.7%.

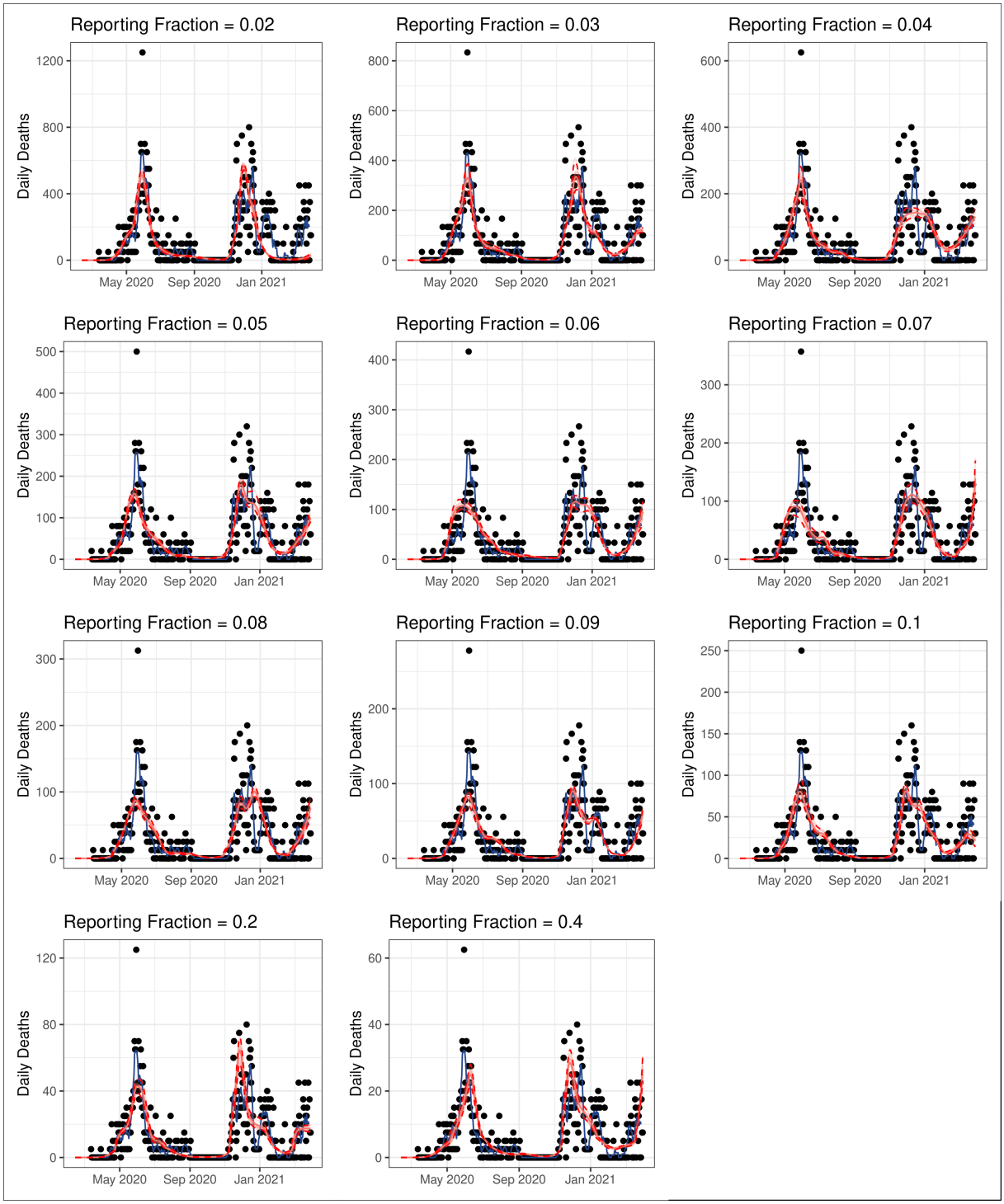

Figure E.12: Model fits in Khartoum based on different reporting fractions. Black dots represent daily reported COVID-19 deaths, with the weekly mean shown with the blue line. Model fits are shown in red, with the 95% CI shown with red dashed lines.

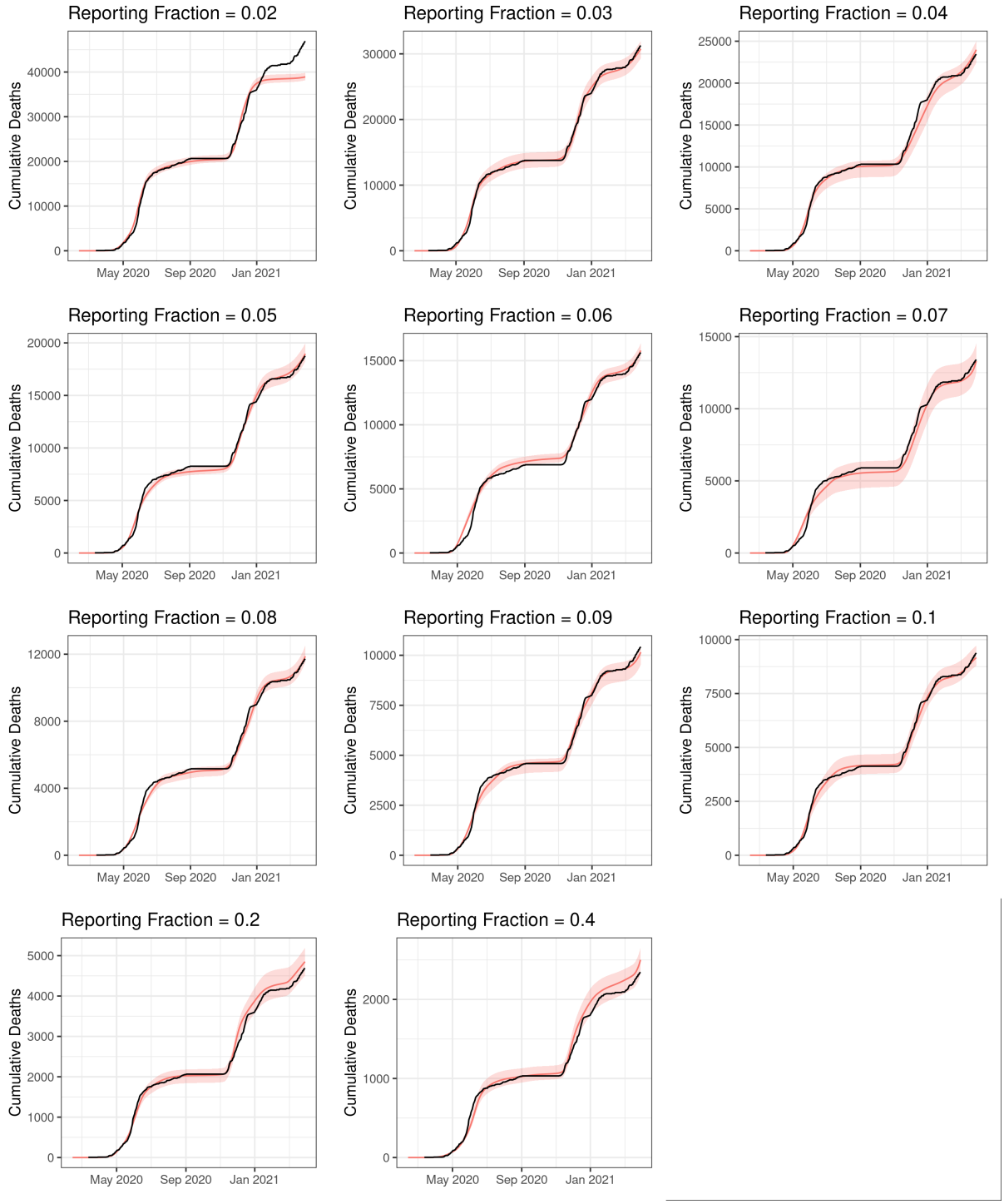

Figure E.13: Model fits in Khartoum based on different reporting fractions. Black lines represent the cumulative reported COVID-19 deaths, with model fits shown in red (median and 95% CI).

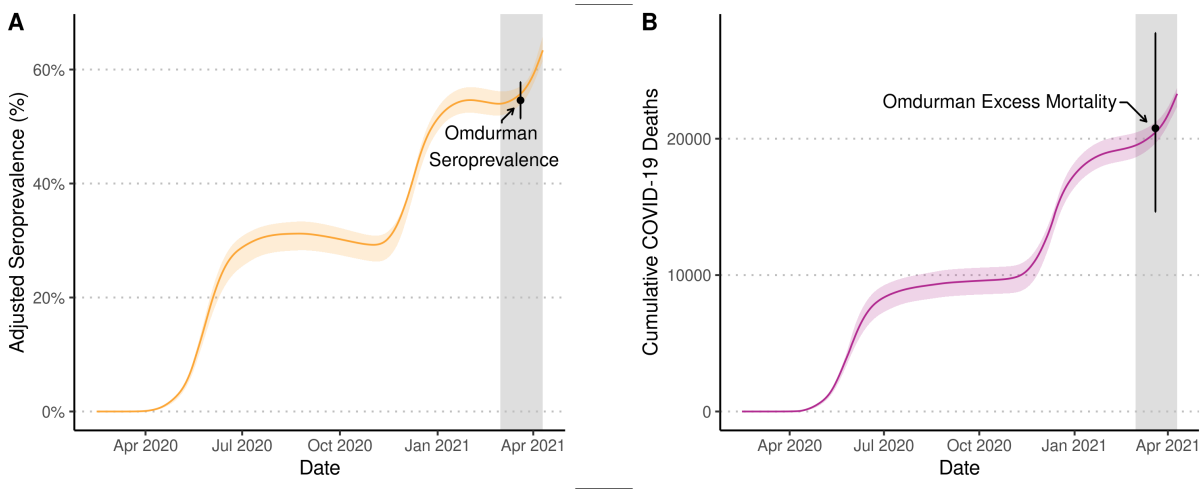

Figure E.14: Best model fit to Omdurman seroprevalence. Model predicted A) adjusted seroprevalence and B) COVID-19 deaths are shown for Khartoum with a reporting fraction equal to 4.5%, which provided the best fit to the Omdurman data, which is shown in black (point estimates, 95% CI shown with vertical bars, and horizontal bars indicate date range for survey). The grey shaded area highlights the sampling period of the serosurvey and mortality survey in Omdurman.

E.5 Comparison of reported and WHO-estimated excess mortality

Using excess mortality estimated by the WHO, we were able to compare reporting fractions of COVID-19 deaths under the alternative sources used in each setting to those at a national level (Figure E.15). These data suggest 8.8% (95% CI: 4.3% − 26.1%), 5.6% (95% CI: 3.5% − 9.3%) and 10.0% (95% CI: 6.4% − 16.4%) of COVID-19 deaths were captured within official statistics in Ethiopia, Yemen and Sudan, respectively (Table E.6). These estimates most likely should not align with our estimated reported fractions, with differences between local and national-level estimates to be expected due to regional heterogeneity, for example, in access to healthcare and death registration services in rural versus urban settings. However, we are encouraged to see similar magnitudes being estimated by our analyses in Aden (parameter uncertainty range: 0.8% − 8.0%) and Khartoum (parameter uncertainty range: 3.0% − 6.0%), which are slightly lower than suggested by the corresponding national-level figures. However, we estimate a much greater reporting fraction in Addis Ababa (parameter uncertainty range: 68.7% − 100%) than the corresponding national estimate, suggesting that national-level excess mortality may have been overestimated if national patterns are comparable to those in Addis Ababa. One plausible explanation for this relates to the WHO methodology for estimating excess mortality. These estimates are modelled based on pooling information from countries with similar characteristics, including WHO region, income status and burden of HIV [483]. However, only four African Region (AMRO) and four Eastern Mediterranean Region (EMRO) countries could provide complete national all-cause mortality data. Consequently, WHO reports very wide confidence intervals for their estimates of excess mortality (Table E.6). These wide ranges demonstrate the value in the alternative data sources we have used here to provide a more nuanced perspective of COVID-19 mortality reporting in low-income settings and add additional data that can be used to refine future estimates of excess mortality.

Table E.6: Estimated reporting fractions of COVID-19 deaths as a percentage of total WHO-estimated excess deaths. WHO excess mortality estimates calculated using the sum of months in which the WHO estimated positive excess mortality.

Setting	**Time period (inclusive)**	**Total reported COVID-19 deaths**	**Total estimated COVID-19 deaths positive excess deaths**	**Estimated reporting fraction**
Ethiopia	April 2020 - October 2020	1,464	16,656 (95% CI: 5,608 - 34,095)	8.8% (95% CI: 4.3% - 26.1%)
Yemen	March 2020 - September 2020	588	10,409 (95% CI: 6,350 - 16,748)	5.6% (95% CI: 3.5% - 9.3%)
Sudan	March 2020 - April 2021	2080	20,723 (95% CI: 12,654 - 32,395)	10.0% (95% CI: 6.4% - 16.4%)

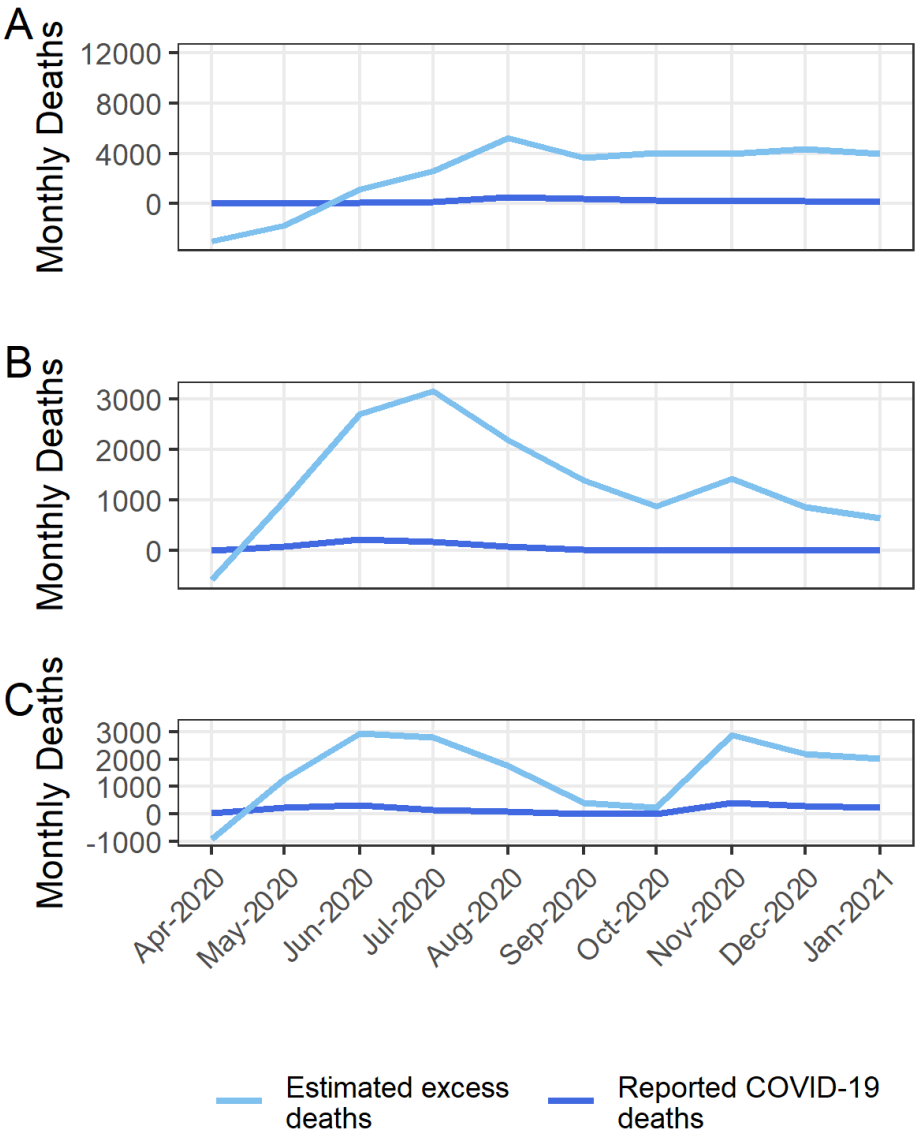

Figure E.15: WHO-estimated excess mortality [483] compared to reported COVID-19 deaths across April 2020 - December 2020. (A) Ethiopia; (B) Sudan; (C) Yemen.

Appendix F

Appendix: Paper VI: Inferring community transmission of SARS-CoV-2 using the ONS COVID-19 Infection Survey

Supplementary material for Paper VI (Chapter 8)

R McCabe, G Danelian, J Panovska-Griffiths and CA Donnelly. Inferring community transmission of SARS-CoV-2 in the United Kingdom using the ONS COVID-19 Infection Survey. *Infectious Disease Modelling*, 2024. doi: `https://doi.org/10.1016/j.idm.2024.01.011`. [*In press.*]

F.1 ONS test positivity spline model

F.1.1 Definition of B-splines

Given a sequence of non-decreasing knots, $t_1, \ldots, t_J$, the kth first-order B-spline is defined as:

$$B_{k,1}(t) = \begin{cases} 1 & t_k < t < t_{k+1} \\ 0 & \text{otherwise} \end{cases}$$

The second and third order kth B-splines are then defined recursively as follows:

$$B_{k,2}(t) = \omega_{k,2}B_{k,1}(t) + \omega_{k,2}B_{k+1,1}(t)$$
$$B_{k,3}(t) = \omega_{k,3}B_{k,2}(t) + \omega_{k,3}B_{k+1,2}(t)$$

where:

$$\omega_{k,2} = \begin{cases} \frac{t-t_k}{t_{k+1}-t_k} & t_k \neq t_{k+1} \\ 0 & \text{otherwise} \end{cases}$$

$$\omega_{k,3} = \begin{cases} \frac{t-t_k}{t_{k+2}-t_k} & t_k \neq t_{k+2} \\ 0 & \text{otherwise} \end{cases}$$

F.1.2 Parameterising the Beta distribution

Recall from Chapter 8 that:

- μ_t denotes the central Office for National Statistics (ONS) estimate at time t;
- l_t and u_t denote the corresponding lower and upper credible intervals, respectively;
- $\sigma_t^2 \approx (\frac{u_t - l_t}{3.92})^2$;
- $\pi_t \sim Beta(\alpha_t, \beta_t)$ is a random variable for the proportion of people testing positive via the ONS survey.

For a random variable distributed according to a Beta distribution, such as X_t, it follows that:

$$\mathbb{E}(X_t) = \frac{\alpha_t}{\alpha_t + \beta_t}$$
$$var(X_t) = \frac{\alpha_t \beta_t}{(\alpha_t + \beta_t)^2(\alpha_t + \beta_t + 1)}.$$

By setting $\mathbb{E}(X_t) = \mu_t$ and $Var(X_t) = \sigma_t^2$, rearranging provides the following expressions for α_t and β_t:

$$\alpha_t = \mu_t\Big(\frac{\mu_t(1-\mu_t)}{\sigma_t^2} - 1\Big) > 0$$
$$\beta_t = \sigma_t\Big(\frac{1-\mu_t}{\mu_t}\Big) > 0.$$

F.1.3 Simulation study to investigate use of Normal likelihood

A simulation study was used in order to examine the appropriateness of using a Normal likelihood distribution in the model. In short, prevalence over time is simulated and by assuming a constant number of tests, the number of positives given the simulated prevalence is determined. The number of positive and total tests can be used to provide a central estimate and 95% confidence interval like the format of the publicly available ONS COVID-19 Infection Survey (CIS) estimates. We then proceed with the first step of the model as described in the main text, by using the data provided to sample from the resulting Beta distribution (see previous section) and compare the sampled values to a Normal distribution.

Let π_t denote the test positivity at time t. An iterative random-walk procedure is used to simulate values of π_t for $t = 1, ..., T$ as follows:

1. Initialise $\pi_t \sim Beta(2, 10)$.
2. To prevent sampling a value of π_t which falls out of the interval [0-1], π_1 is transformed to the logit scale via $logit(\pi_1) = log\left(\frac{\pi_1}{1-\pi_1}\right)$.
3. For $t = 2, ..., T$, $logit(\pi_t) \sim Normal(logit(\pi_{t-1}), \zeta)$, where $\zeta \sim Normal(2, 50)$.
4. For all t, transform $logit(\pi_t)$ back to the interval [0,1] using $\pi_t = \frac{exp(logit(\pi_t))}{exp(logit(\pi_t))+1}$.

Denote by y_t the number of positive tests and by n_t the total number of tests at time t. It is assumed that $n_t = n$ is held constant across the period. For each time point, we determine $y_t = \pi_t \times n$.

Exact 95% binomial confidence intervals are obtained using the number of positives and total tests per time point. To match the format of the ONS data, the maximum likelihood estimate (MLE) $\left(\frac{y_t}{n}\right)$ is taken as the central estimate (μ_t) with the 95% confidence intervals as the uncertainty range (l_t and u_t).

We assume $n = 80,000$ tests are conducted at each time point, in accordance with the number of participants in the ONS CIS in England in December 2022 [547] and we let $T = 150$. This produces data such as that presented in Figure F.1.

Having simulated data in the format of the publicly available ONS estimates, we proceed with mimicking the initial sampling phase of the model. Specifically:

$$\pi_t \sim Beta(\alpha_t, \beta_t)$$

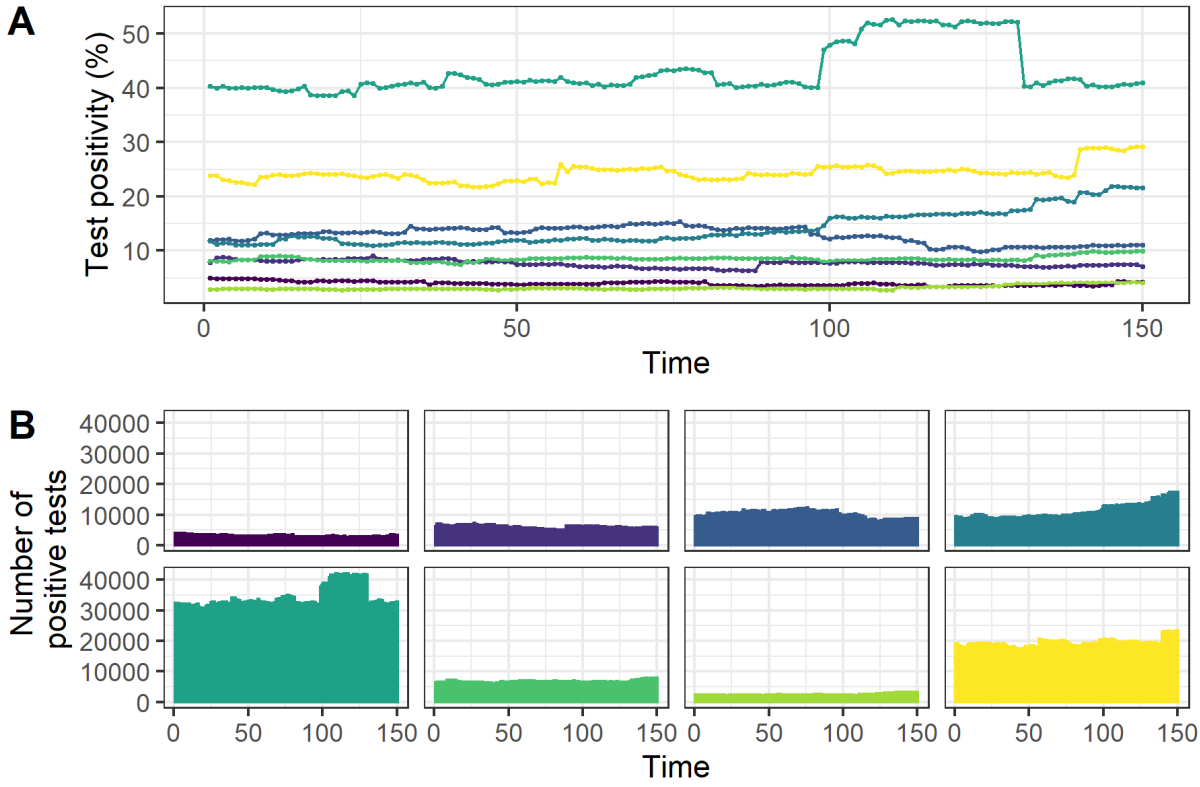

Figure F.1: Examples of data simulated to test the assumption of using a Normal likelihood. Each colour represents a distinct simulation and is used to link the outputs in (A) and (B). At each time point of the simulation, it is assumed that $n = 80,000$ tests are conducted. (A) Test positivity (p_t) simulated from the random walk procedure. (B) The number of positive tests determined via $x_t = \pi_t \times n$.

where:

$$\alpha_t = \mu_t\Big(\frac{\mu_t(1-\mu_t)}{\sigma_t^2} - 1\Big) > 0$$
$$\beta_t = \sigma_t\Big(\frac{1-\mu_t}{\mu_t}\Big) > 0.$$
$$\sigma_t^2 \approx \Big(\frac{u_t - l_t}{3.92}\Big)^2$$

For each t, we sample from the Beta distribution $10,000$ times and take the *logit* of each. Examples of this for one set of simulated data are shown in Figure F.2. The resulting (*logit*) samples are compared to a Normal distribution with mean and standard deviation calculated from the $10,000$ samples using a Kolmogorov-Smirnov test for each $t = 1, \ldots, 150$.

The entire process of simulating data, sampling $10,000$ times from the resulting Beta distribution for each $t = 1, \ldots, 150$ and comparing this to Normal distribution using a Kolmogorov-Smirnov test is repeated $1,000$ times to account for the variability arising from the simulation process. Across the $1,000$ repetitions, most time points were Normally distributed according to Kolmogorov-Smirnov tests with

a critical p-value threshold of 0.05. 33 of the repetitions (3.3%) had 1 time point out of 150 that was not-normally distributed (0.7%) and 2 repetitions (0.2%) had 2 time points out of 150 that were not normally distributed (1.3%). Consequently, the Normal distribution is deemed appropriate the likelihood in the ONS CIS spline model.

F.1.4 Diagnostic traceplots

Figure F.3, Figure F.4, Figure F.5 and Figure F.6 present a selection of traceplots for parameters in the model fit to data from England, Scotland, Wales and Northern Ireland, respectively.

Recall the model specification from Chapter 8:

$$\begin{aligned}
\pi_t &\sim Beta(\alpha_t, \beta_t) \\
a_1 &\sim Normal(0, 1) \\
a_k &\sim Normal(a_{k-1}, \tau^2) \\
\tau^2 &\sim InverseGamma(0.0001, 0.0001) \\
logit(\hat{p}(t)) &= \sum_{k=1}^{K} a_k B_{k,3}(t) \\
logit(\pi_t) &\sim Normal(logit(\hat{p}(t)), \gamma^2) \\
\gamma^2 &\sim InverseGamma(0.0001, 0.0001)
\end{aligned}$$

In each plot, $a[k]$ represents a_k; $Y\_hat[k]$ represents $\hat{p}(k)$ for discrete time points k, which are selected as the first ($k = 1$), the last time, and then equidistant between the first and last time points of the ONS CIS data being fitted to.

tau and gamma represent τ and γ as defined above, respectively.

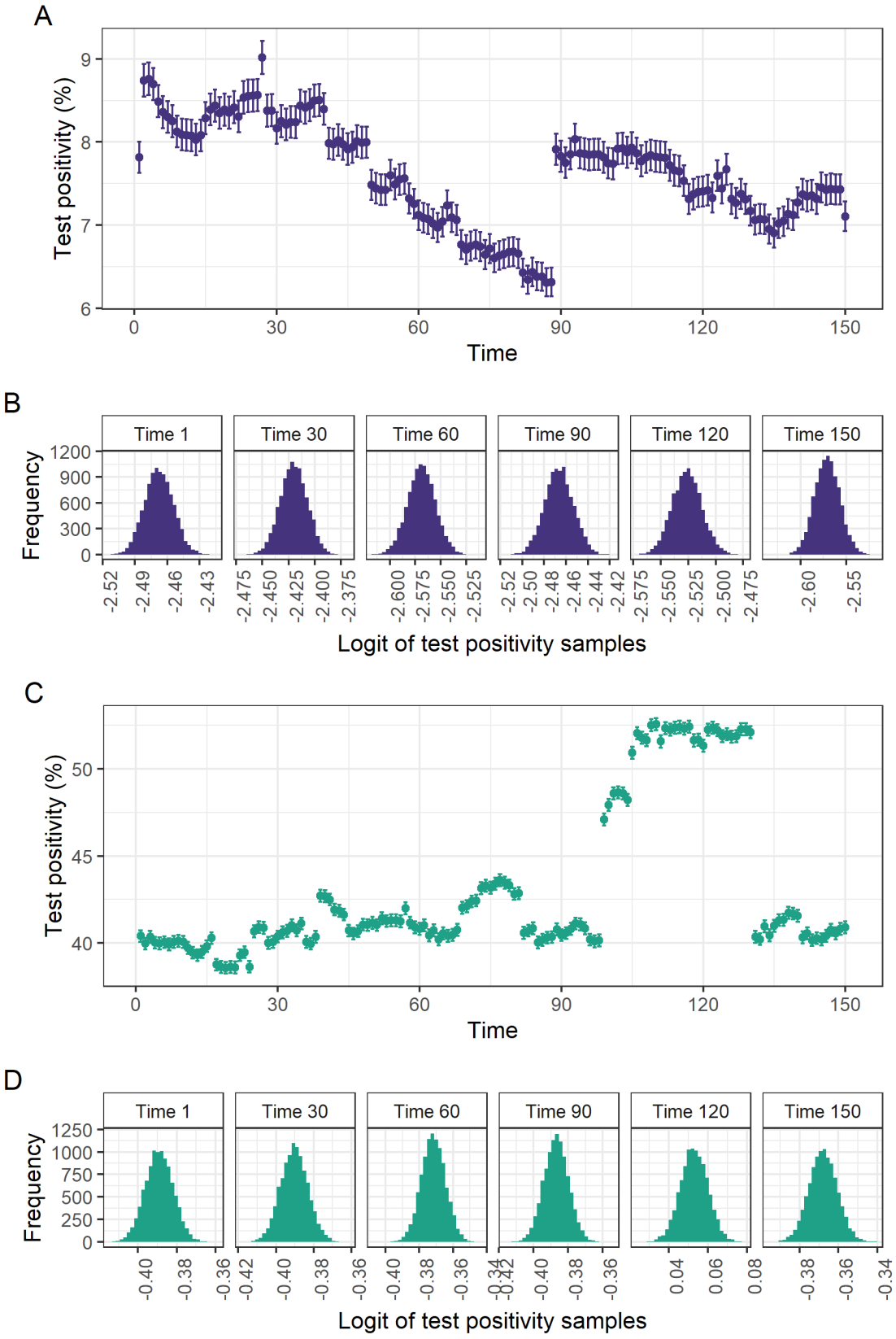

Figure F.2: Examples of simulated data and 10,000 samples from resulting Beta distribution at 6 equidistant time periods in the simulation. Data in the two examples here are also shown in Supplementary Figure 1. (A) and (C) show the simulated test positivity time series formatted in the style of the publicly available ONS CIS estimates (central estimate with 95% confidence intervals). (B) and (D) present histograms of the 10,000 samples transformed to the logit scale at 6 time points corresponding to the data shown in (A) and (C), respectively.

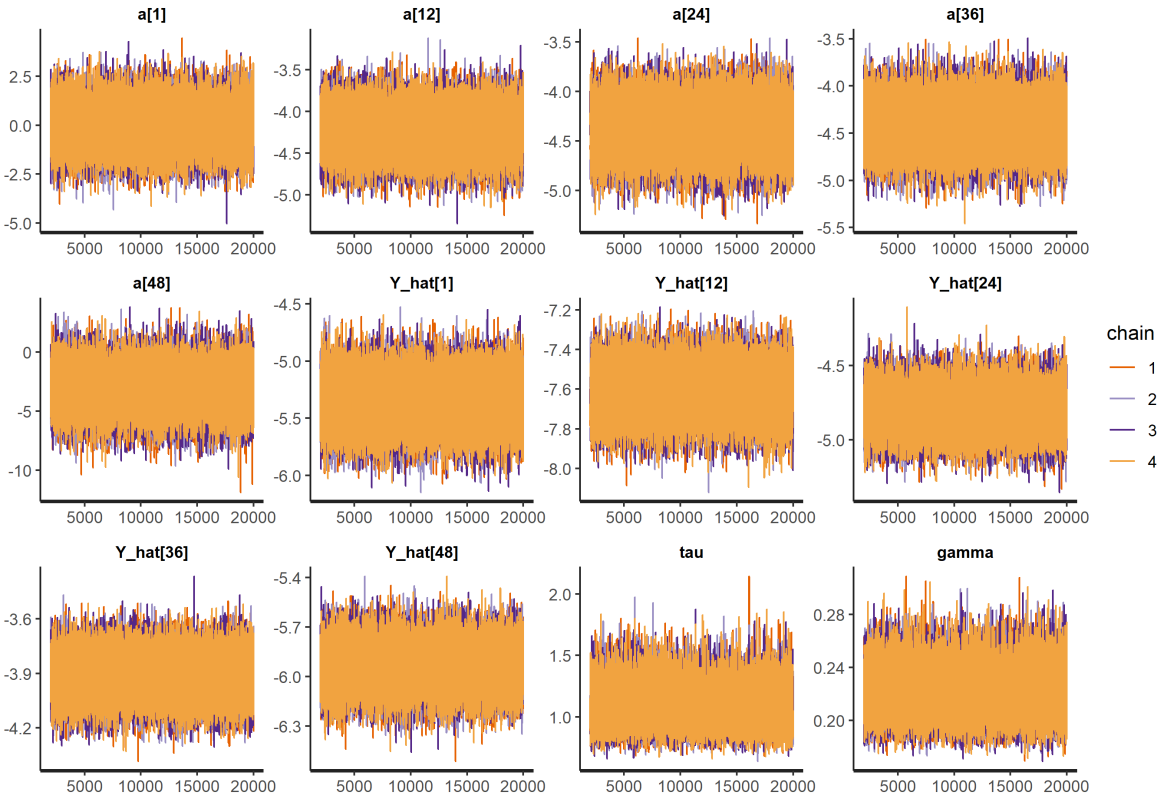

Figure F.3: Traceplots for a subset of parameters in the model fit to data from England.

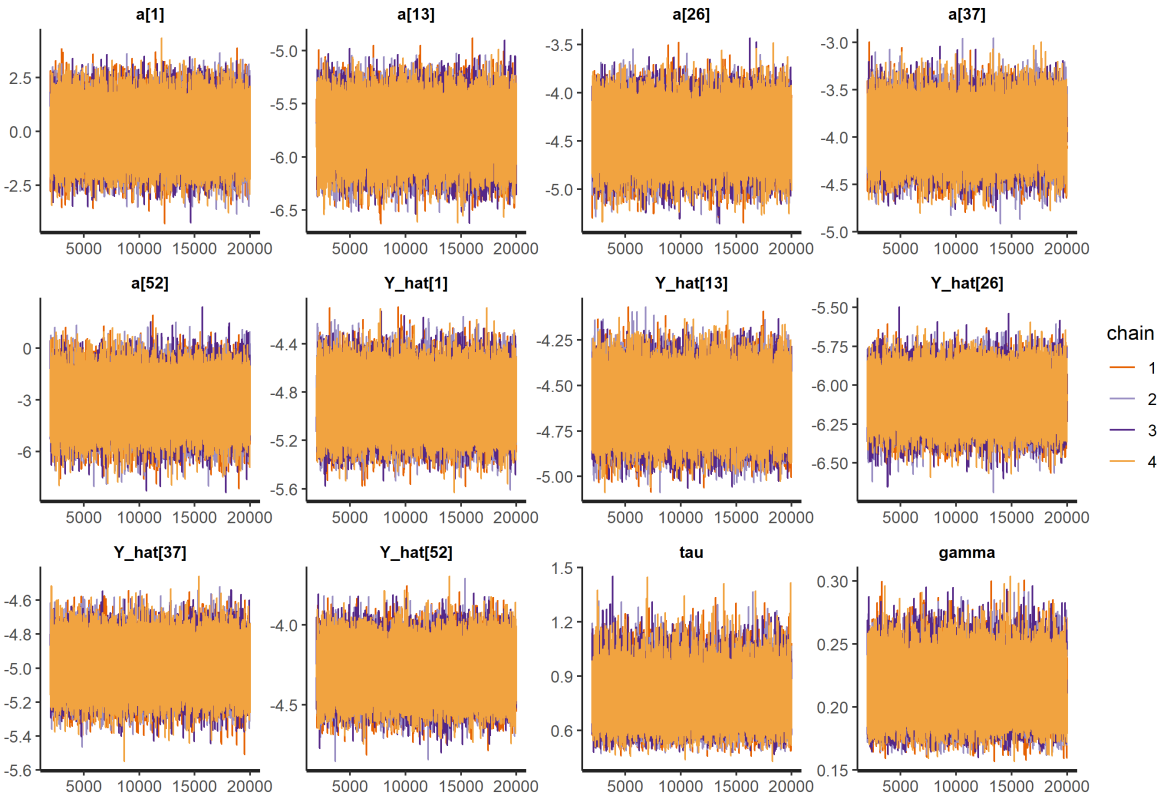

Figure F.4: Traceplots for a subset of parameters in the model fit to data from Scotland.

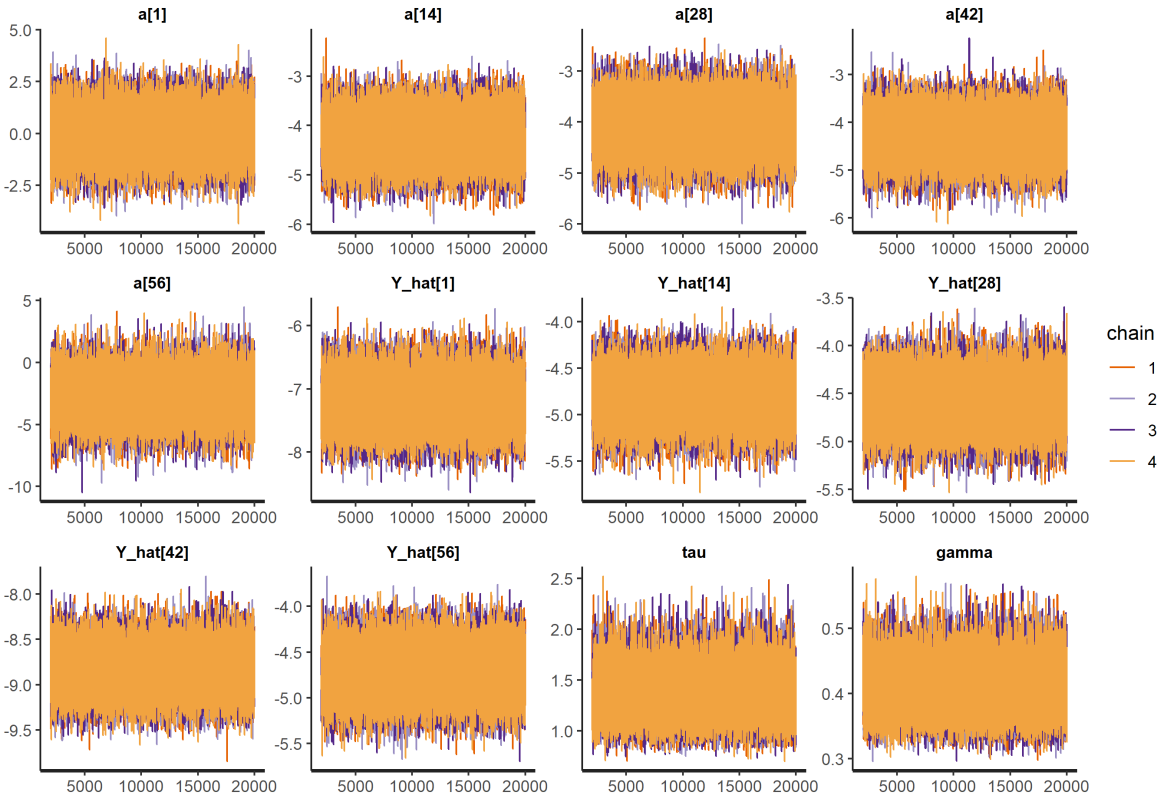

Figure F.5: Traceplots for a subset of parameters in the model fit to data from Wales.

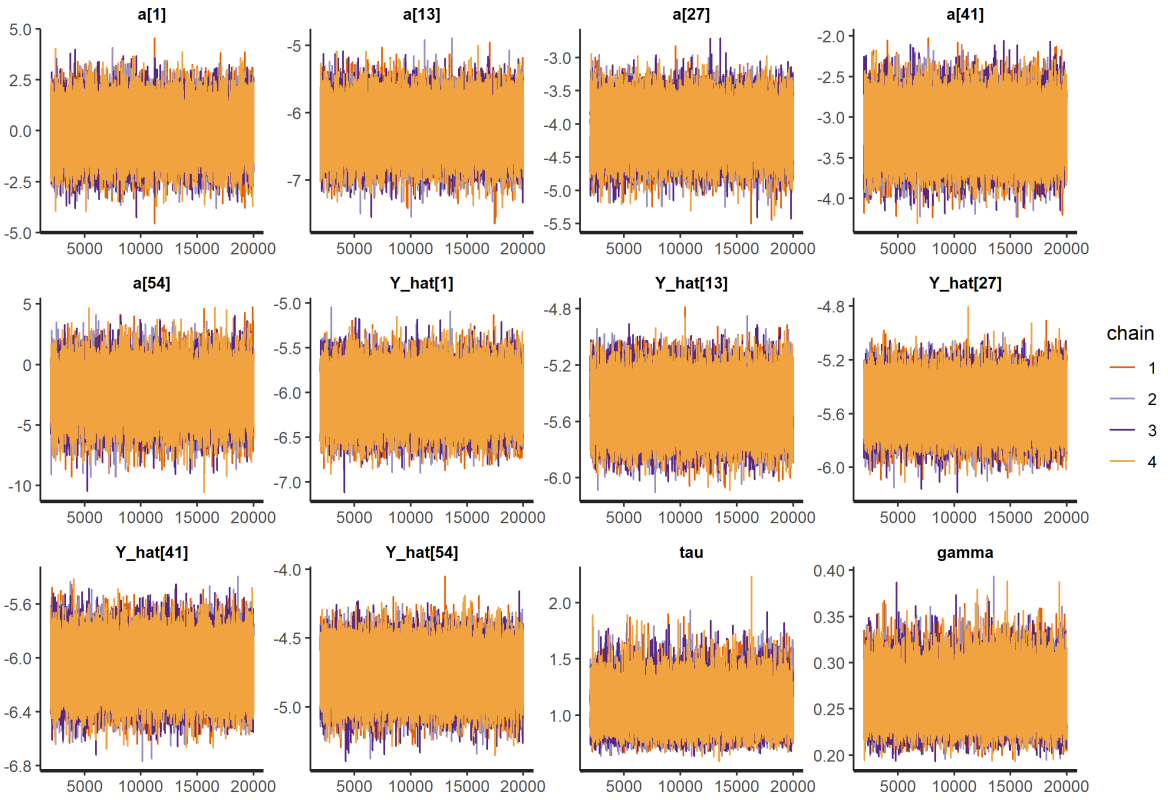

Figure F.6: Traceplots for a subset of parameters in the model fit to data from Northern Ireland.

F.2 Relationship between ONS-based and government-published estimates

F.2.1 Figures including scatterplots between ONS-based and government-published estimates

Figure F.7, Figure F.8, Figure F.9 and Figure F.10 are analogous to Figures 8.1 – 8.4 in Chapter 8, but also include scatter plots highlighting the relationship between the ONS-based and government-published estimates of $R(t)$ and $r(t)$.

F.2.2 Comparison of estimates without lags

Figure F.11, Figure F.12, Figure F.13 and Figure F.14 present the spline fits and resulting estimates $\hat{R}(t)$ and $\hat{r}(t)$ which are compared to $\tilde{R}(t)$ and $\tilde{r}(t)$, respectively, without any time lags.

F.2.3 Regression analysis without inclusion of an intercept

In the main text, we set out the following linear regression models to examine the variation of the government-published estimates explained by the ONS-based estimates:

$$\tilde{R}(t) = \beta_0 + \beta_1 \hat{R}(t) + \epsilon$$

and

$$\tilde{r}(t) = \gamma_0 + \gamma_1 \hat{r}(t) + \epsilon.$$

Here, we consider similar models but without the intercept terms β_0 and γ_0:

$$\tilde{R}(t) = \beta_1 \hat{R}(t) + \epsilon$$

and

$$\tilde{r}(t) = \gamma_1 \hat{r}(t) + \epsilon.$$

In a linear regression model, removing the intercept term forces the model to go through the origin (point $(0, 0)$). In doing so, the assumption is made that the dependent variable, in this instance the government-published estimates, is solely determined by the explanatory variables, the ONS-based estimates, and random noise. Thus, should these explanatory variables be 0, this suggests that the dependent variable

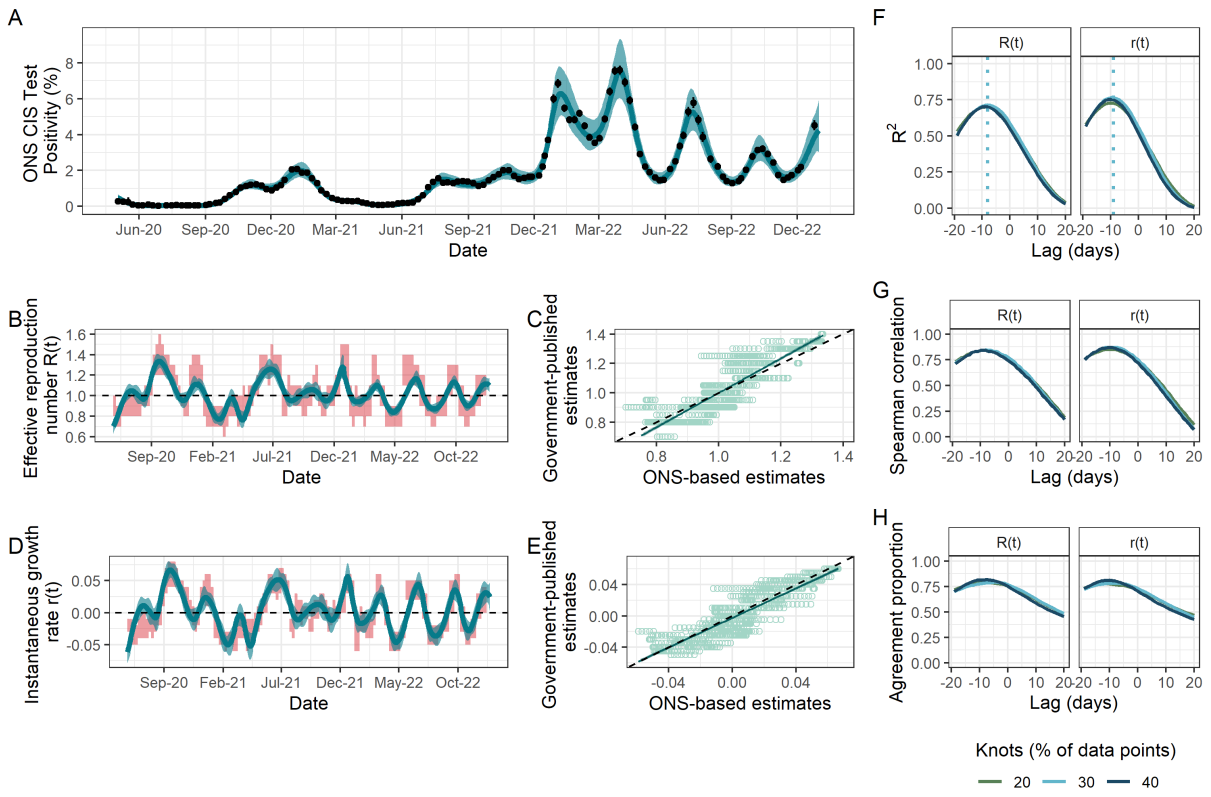

Figure F.7: Fit to ONS CIS data, resulting estimates of $R(t)$ and $r(t)$, and comparison metrics for England. (A) Spline model fit ($\hat{p}(t)$) (blue; median line with 95% credible intervals) to data (black points; ONS point estimate with 95% credible intervals) with the number of knots totalling 30% of the total data points. (B) The ONS-based estimated effective reproduction number ($\hat{R}(t)$) (blue; median line with 95% credible intervals) alongside the government-published estimates ($\tilde{R}(t)$) (red; 90% confidence intervals), lagged by -8 days. The black dashed line highlights the epidemic growth threshold of 1. (C) Scatterplot (light green points) showing the relationship between the ONS-based and government-reported estimates of $R(t)$, with the latter lagged by -8 days. The dark green line shows the trend line from the linear model (with dark green shading corresponding to 95% confidence intervals) regressing the lagged government-published estimates on the ONS-based estimates. Rows of points are due to the limited precision available (due to rounding before publication) for government-published estimates. (D) The ONS-based estimated instantaneous growth rate ($\hat{r}(t)$) (blue; median line with 95% credible intervals) alongside the government-published estimates ($\tilde{r}(t)$) (red; 90% confidence intervals), lagged by -9 days. The dashed line highlights the epidemic growth threshold of 0. (E) Scatterplot (light green points) showing the relationship between the ONS-based and government-reported estimates of $r(t)$, with the latter lagged by -9 days. The dark green line shows the trend line from the linear model (with dark green shading corresponding to 95% confidence intervals) regressing the lagged government-published estimates on the ONS-based estimates. Rows of points are due to the limited precision available (due to rounding before publication) for government-published estimates. (F) – (H) Results of the three metrics used to assess the agreement between ONS-based and government-published estimates under different lags and for models fitted with a different number of knots, taken as a percentage of the total data points. (F) R^2. Dotted lines indicate the lag for which this metric is maximised for each parameter and is thus what is applied to the government-based estimates in panels (B) – (E). (G) Spearman rank correlation. (H) Agreement proportion.

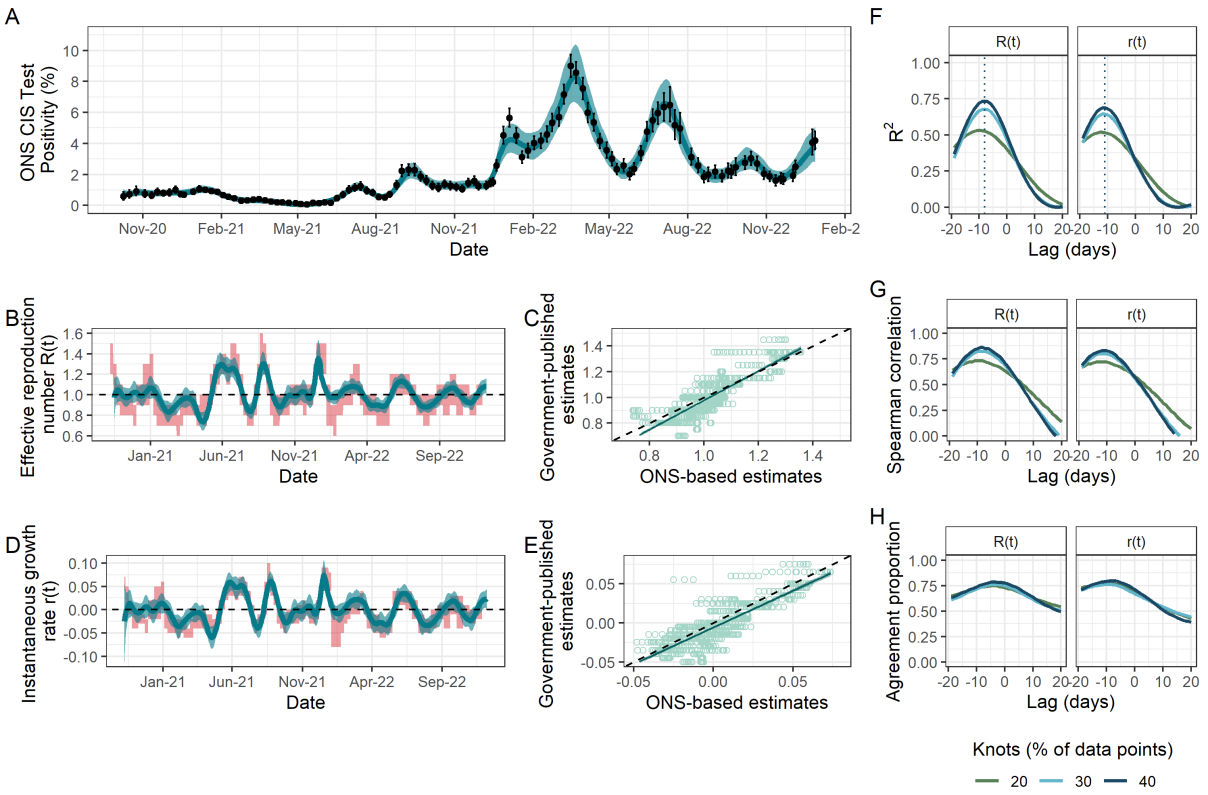

Figure F.8: Fit to ONS CIS data, resulting estimates of $R(t)$ and $r(t)$, and comparison metrics for Scotland. (A) Spline model fit ($\hat{p}(t)$) (blue; median line with 95% credible intervals) to data (black points; ONS point estimate with 95% credible intervals) with the number of knots totalling 40% of the total data points. (B) The ONS-based estimated effective reproduction number ($\hat{R}(t)$) (blue; median line with 95% credible intervals) alongside the government-published estimates ($\tilde{R}(t)$) (red; 90% confidence intervals), lagged by -8 days. The black dashed line highlights the epidemic growth threshold of 1. (C) Scatterplot (light green points) showing the relationship between the ONS-based and government-reported estimates of $R(t)$, with the latter lagged by -8 days. The dark green line shows the trend line from the linear model (with dark green shading corresponding to 95% confidence intervals) regressing the lagged government-published estimates on the ONS-based estimates. Rows of points are due to the limited precision available (due to rounding before publication) for government-published estimates. (D) The ONS-based estimated instantaneous growth rate ($\hat{r}(t)$) (blue; median line with 95% credible intervals) alongside the government-published estimates ($\tilde{r}(t)$) (red; 90% confidence intervals), lagged by -11 days. The dashed line highlights the epidemic growth threshold of 0. (E) Scatterplot (light green points) showing the relationship between the ONS-based and government-reported estimates of $r(t)$, with the latter lagged by -11 days. The dark green line shows the trend line from the linear model (with dark green shading corresponding to 95% confidence intervals) regressing the lagged government-published estimates on the ONS-based estimates. Rows of points are due to the limited precision available (due to rounding before publication) for government-published estimates. (F) – (H) Results of the three metrics used to assess the agreement between ONS-based and government-published estimates under different lags and for models fitted with a different number of knots, taken as a percentage of the total data points. (F) R^2. Dotted lines indicate the lag for which this metric is maximised for each parameter and is thus what is applied to the government-based estimates in panels (B) – (E). (G) Spearman rank correlation. (H) Agreement proportion.

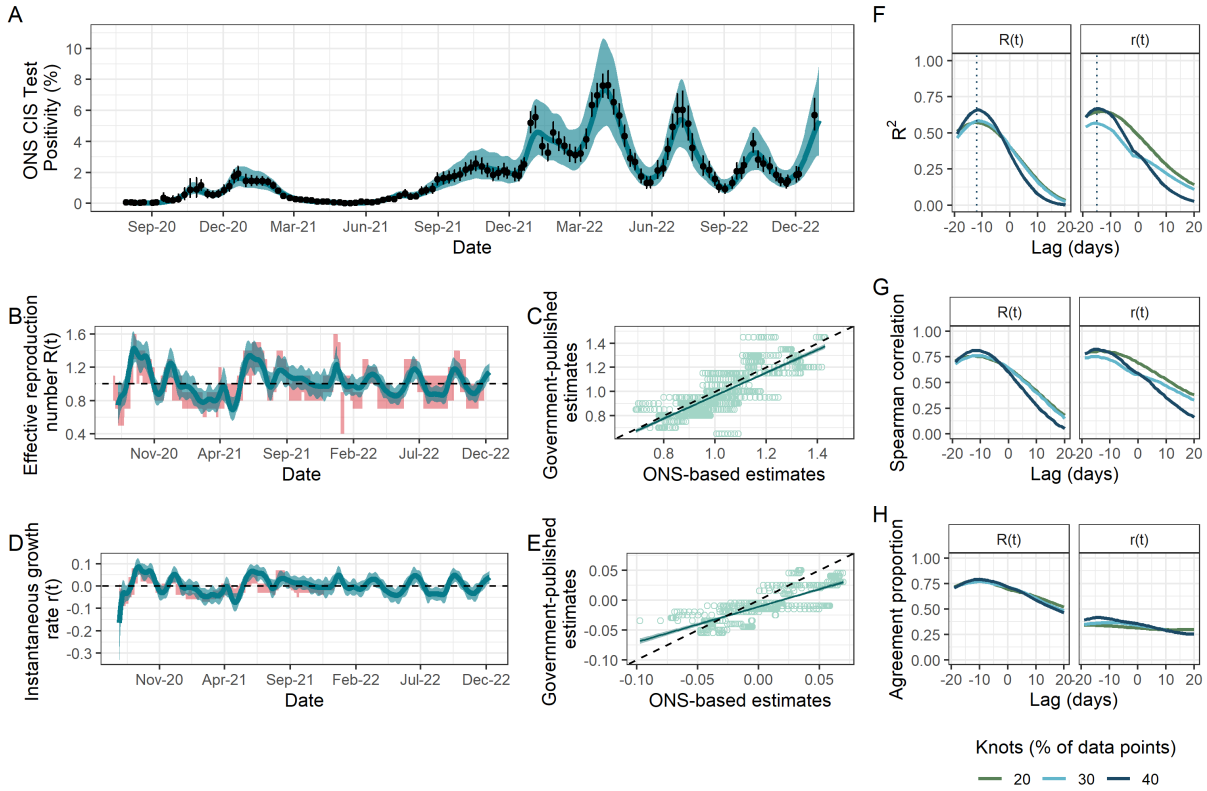

Figure F.9: Fit to ONS CIS data, resulting estimates of $R(t)$ and $r(t)$, and comparison metrics for Wales. (A) Spline model fit ($\hat{p}(t)$) (blue; median line with 95% credible intervals) to data (black points; ONS point estimate with 95% credible intervals) with the number of knots totalling 40% of the total data points. (B) The ONS-based estimated effective reproduction number ($\hat{R}(t)$) (blue; median line with 95% credible intervals) alongside the government-published estimates ($\tilde{R}(t)$) (red; 90% confidence intervals), lagged by -12 days. The black dashed line highlights the epidemic growth threshold of 1. (C) Scatterplot (light green points) showing the relationship between the ONS-based and government-reported estimates of $R(t)$, with the latter lagged by -12 days. The dark green line shows the trend line from the linear model (with dark green shading corresponding to 95% confidence intervals) regressing the lagged government-published estimates on the ONS-based estimates. Rows of points are due to the limited precision available (due to rounding before publication) for government-published estimates. (D) The ONS-based estimated instantaneous growth rate ($\hat{r}(t)$) (blue; median line with 95% credible intervals) alongside the government-published estimates ($\tilde{r}(t)$) (red; 90% confidence intervals), lagged by -15 days. The dashed line highlights the epidemic growth threshold of 0. (E) Scatterplot (light green points) showing the relationship between the ONS-based and government-reported estimates of $r(t)$, with the latter lagged by -15 days. The dark green line shows the trend line from the linear model (with dark green shading corresponding to 95% confidence intervals) regressing the lagged government-published estimates on the ONS-based estimates. Rows of points are due to the limited precision available (due to rounding before publication) for government-published estimates. (F) – (H) Results of the three metrics used to assess the agreement between ONS-based and government-published estimates under different lags and for models fitted with a different number of knots, taken as a percentage of the total data points. (F) R^2. Dotted lines indicate the lag for which this metric is maximised for each parameter and is thus what is applied to the government-based estimates in panels (B) – (E). (G) Spearman rank correlation. (H) Agreement proportion.

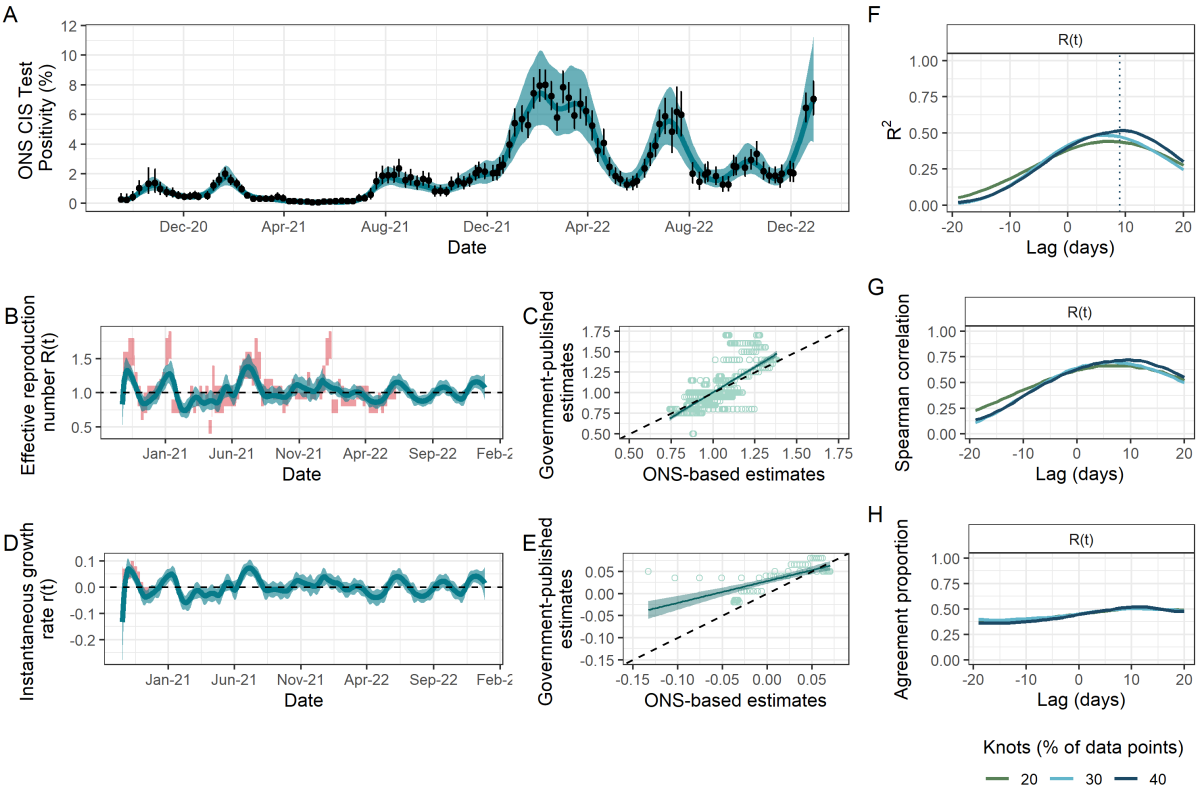

Figure F.10: Fit to ONS CIS data, resulting estimates of $R(t)$ and $r(t)$, and comparison metrics for Northern Ireland. (A) Spline model fit ($\hat{p}(t)$) (blue; median line with 95% credible intervals) to data (black points; ONS point estimate with 95% credible intervals) with the number of knots totalling 40% of the total data points. (B) The ONS-based estimated effective reproduction number ($\hat{R}(t)$) (blue; median line with 95% credible intervals) alongside the government-published estimates ($\tilde{R}(t)$) (red; 90% confidence intervals), lagged by -8 days. The black dashed line highlights the epidemic growth threshold of 1. (C) Scatterplot (light green points) showing the relationship between the ONS-based and government-reported estimates of $R(t)$, with the latter lagged by -8 days. The dark green line shows the trend line from the linear model (with dark green shading corresponding to 95% confidence intervals) regressing the lagged government-published estimates on the ONS-based estimates. Rows of points are due to the limited precision available (due to rounding before publication) for government-published estimates. (D) The ONS-based estimated instantaneous growth rate ($\hat{r}(t)$) (blue; median line with 95% credible intervals) alongside the government-published estimates ($\tilde{r}(t)$) (red; 90% confidence intervals), lagged by -11 days. The dashed line highlights the epidemic growth threshold of 0. (E) Scatterplot (light green points) showing the relationship between the ONS-based and government-reported estimates of $r(t)$, with the latter lagged by -11 days. The dark green line shows the trend line from the linear model (with dark green shading corresponding to 95% confidence intervals) regressing the lagged government-published estimates on the ONS-based estimates. Rows of points are due to the limited precision available (due to rounding before publication) for government-published estimates. (F) – (H) Results of the three metrics used to assess the agreement between ONS-based and government-published estimates of $R(t)$ under different lags and for models fitted with a different number of knots, taken as a percentage of the total data points. (F) R^2. Dotted lines indicate the lag for which this metric is maximised for each parameter and is thus what is applied to the government-based estimates in panels (B) – (E). (G) Spearman rank correlation. (H) Agreement proportion.

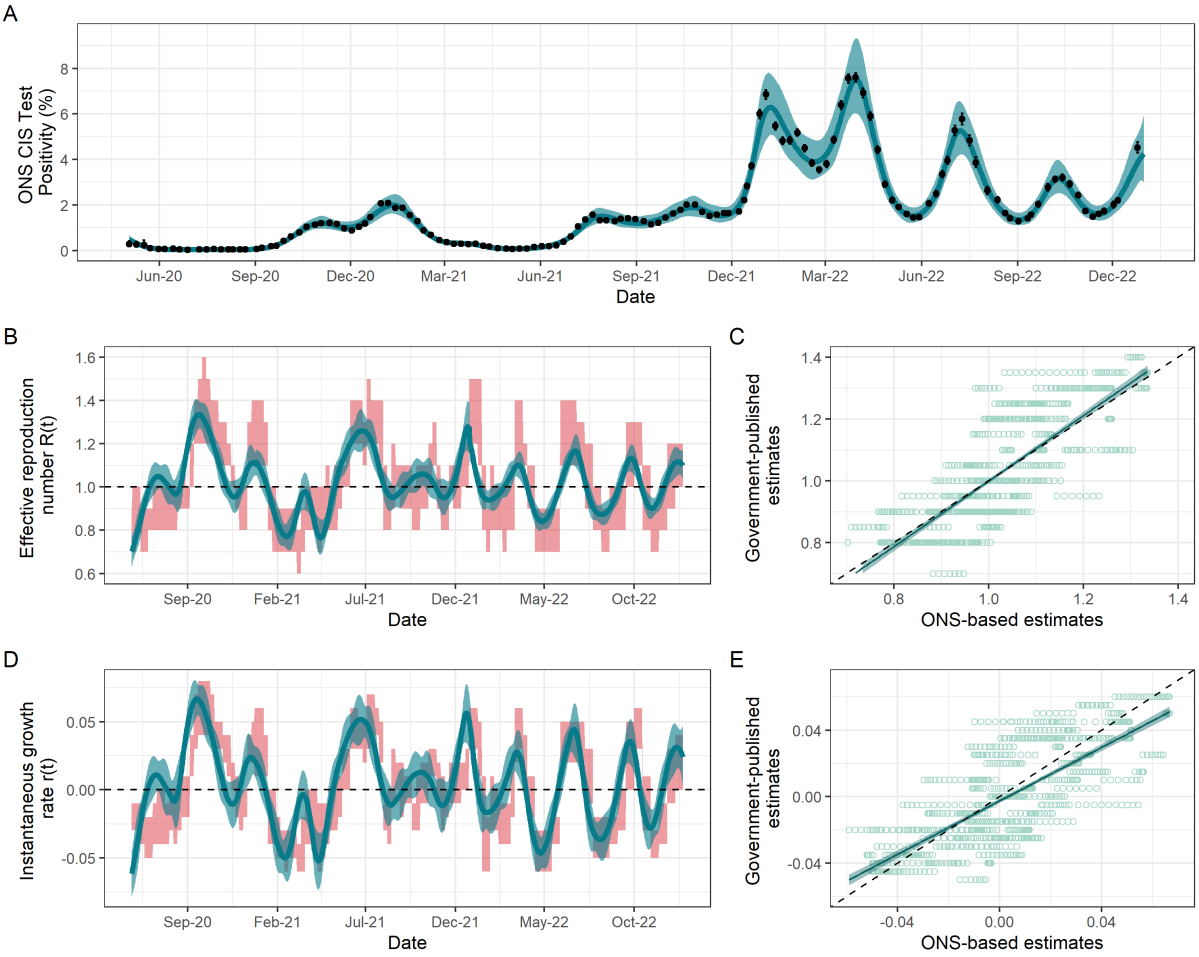

Figure F.11: Fit to ONS CIS data and resulting estimates of $R(t)$ and $r(t)$ for England without any lags. (A) Spline model fit ($\hat{p}(t)$) (blue; median line with 95% credible intervals) to data (black points; ONS point estimate with 95% credible intervals) with the number of knots totalling 30% of the total data points. (B) The ONS-based estimated effective reproduction number ($\hat{R}(t)$) (blue; median line with 95% credible intervals) alongside the government-published estimates ($\tilde{R}(t)$) (red; 90% confidence intervals). The black dashed line highlights the epidemic growth threshold of 1. (C) Scatterplot (light green points) showing the relationship between the ONS-based and government-reported estimates of $R(t)$. The dark green line shows the trend line from the linear model (with dark green shading corresponding to 95% confidence intervals) regressing the government-published estimates on the ONS-based estimates. Rows of points are due to the limited precision available (due to rounding before publication) for government-published estimates. (D) The ONS-based estimated instantaneous growth rate ($\hat{r}(t)$) (blue; median line with 95% credible intervals) alongside the government-published estimates ($\tilde{r}(t)$) (red; 90% confidence intervals). The dashed line highlights the epidemic growth threshold of 0. (E) Scatterplot (light green points) showing the relationship between the ONS-based and government-reported estimates of $r(t)$. The dark green line shows the trend line from the linear model (with dark green shading corresponding to 95% confidence intervals) regressing the lagged government-published estimates on the ONS-based estimates. Rows of points are due to the limited precision available (due to rounding before publication) for government-published estimates.

is determined solely by noise. In this instance, we know that this is an incorrect specification of the relationship between our variables, as the government-published estimates were derived from an ensemble of models. However, we can use the no-intercept model to explore further the relationship between the ONS-based and government-published estimates. If $\hat{\beta}_1$ or $\hat{\gamma}_1$ tend to 1, without the intercept, this implies that

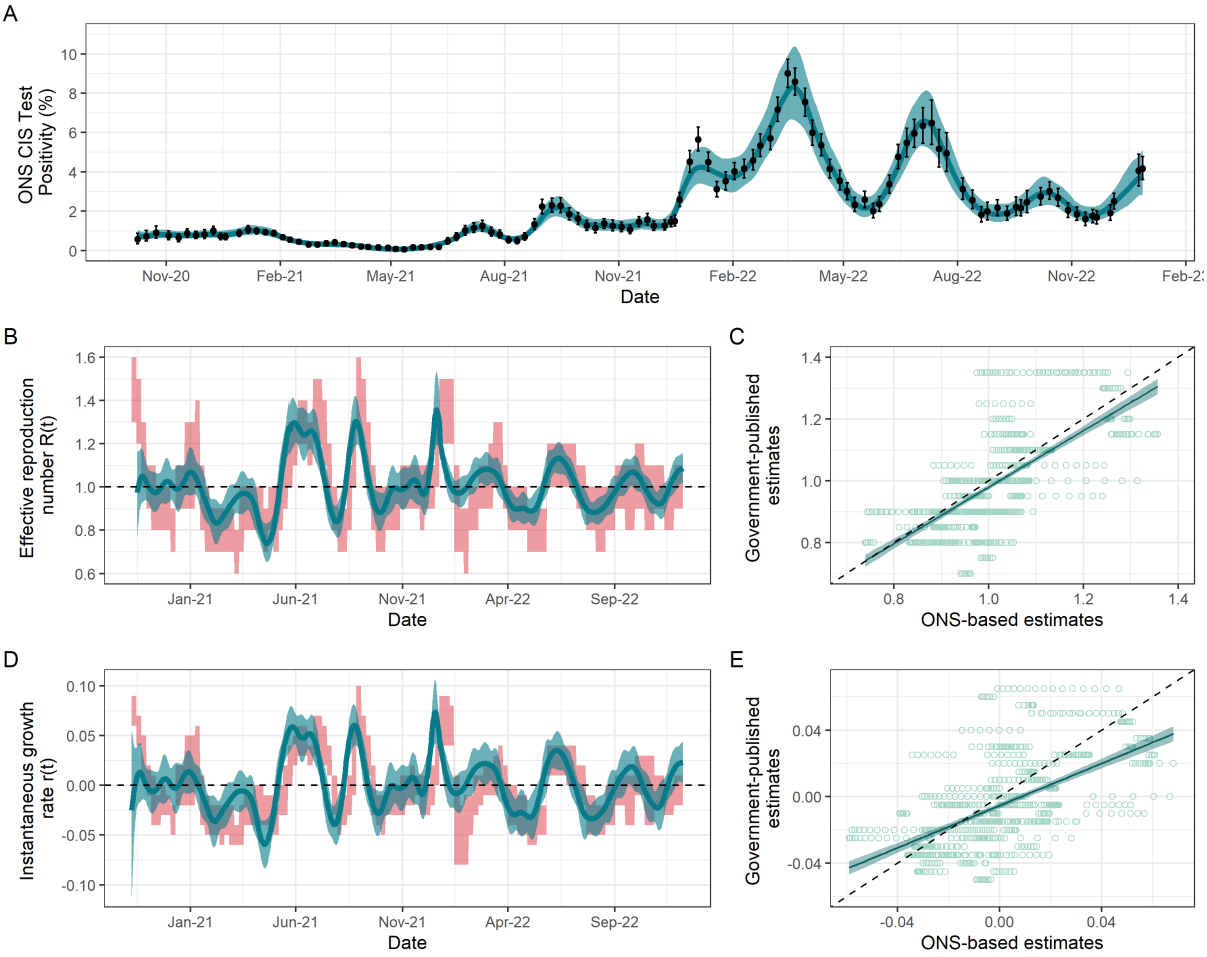

Figure F.12: Fit to ONS CIS data and resulting estimates of $R(t)$ and $r(t)$ for England without any lags. (A) Spline model fit ($\hat{p}(t)$) (blue; median line with 95% credible intervals) to data (black points; ONS point estimate with 95% credible intervals) with the number of knots totalling 40% of the total data points. (B) The ONS-based estimated effective reproduction number ($\hat{R}(t)$) (blue; median line with 95% credible intervals) alongside the government-published estimates ($\tilde{R}(t)$) (red; 90% confidence intervals). The black dashed line highlights the epidemic growth threshold of 1. (C) Scatterplot (light green points) showing the relationship between the ONS-based and government-reported estimates of $R(t)$. The dark green line shows the trend line from the linear model (with dark green shading corresponding to 95% confidence intervals) regressing the government-published estimates on the ONS-based estimates. Rows of points are due to the limited precision available (due to rounding before publication) for government-published estimates. (D) The ONS-based estimated instantaneous growth rate ($\hat{r}(t)$) (blue; median line with 95% credible intervals) alongside the government-published estimates ($\tilde{r}(t)$) (red; 90% confidence intervals). The dashed line highlights the epidemic growth threshold of 0. (E) Scatterplot (light green points) showing the relationship between the ONS-based and government-reported estimates of $r(t)$. The dark green line shows the trend line from the linear model (with dark green shading corresponding to 95% confidence intervals) regressing the lagged government-published estimates on the ONS-based estimates. Rows of points are due to the limited precision available (due to rounding before publication) for government-published estimates.

the ONS-based estimates are unbiased predictors of the government-published estimates. If $\hat{\beta}_1$ or $\hat{\gamma}_1$ tend to 0, this implies that the government-published estimates are not predicted by our ONS-based estimates.

Figure F.15 presents the resulting $\hat{\beta}_1$ or $\hat{\gamma}_1$ for these models under the same lags (plus and minus 20

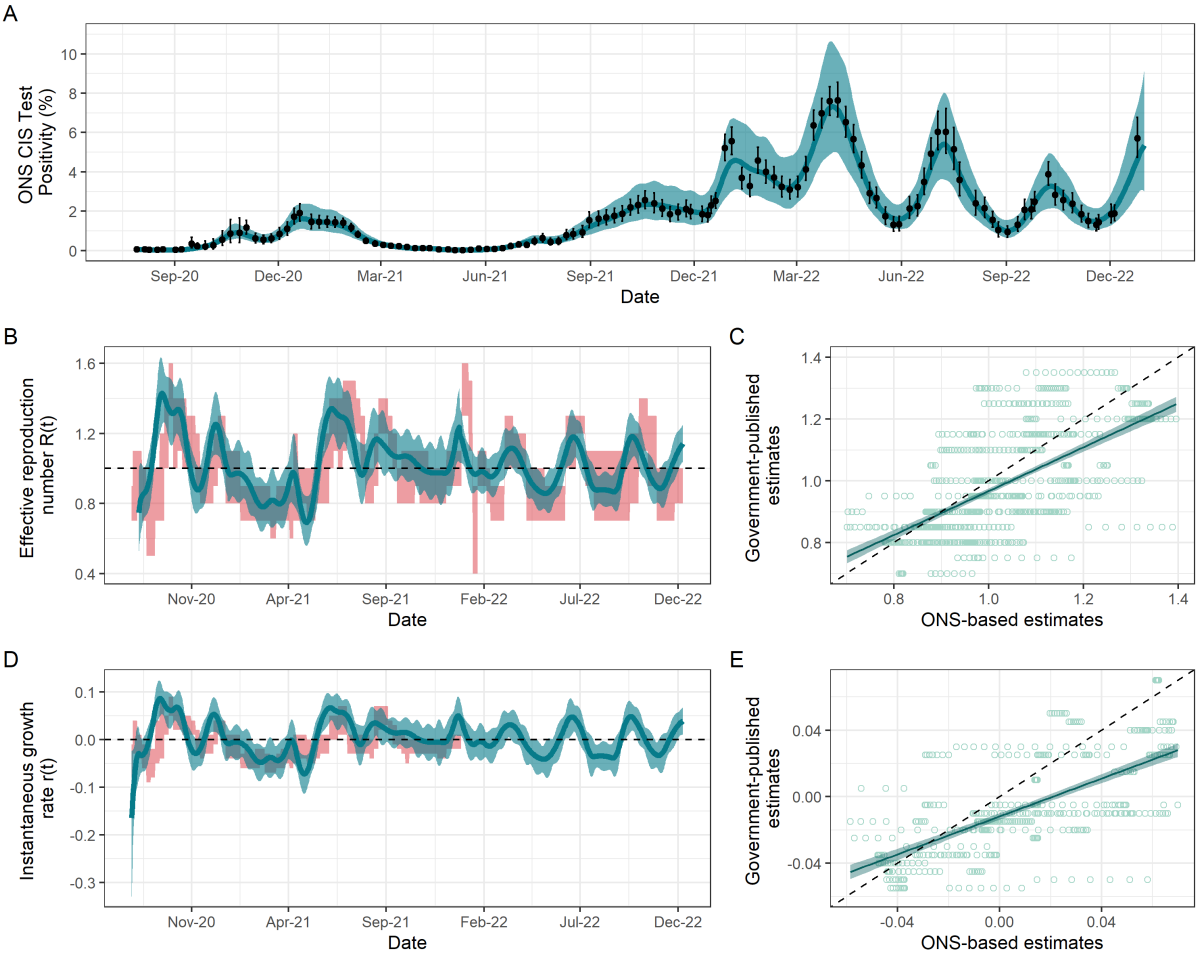

Figure F.13: Fit to ONS CIS data and resulting estimates of $R(t)$ and $r(t)$ for Wales without any lags. (A) Spline model fit ($\hat{p}(t)$) (blue; median line with 95% credible intervals) to data (black points; ONS point estimate with 95% credible intervals) with the number of knots totalling 40% of the total data points. (B) The ONS-based estimated effective reproduction number ($\hat{R}(t)$) (blue; median line with 95% credible intervals) alongside the government-published estimates ($\tilde{R}(t)$) (red; 90% confidence intervals). The black dashed line highlights the epidemic growth threshold of 1. (C) Scatterplot (light green points) showing the relationship between the ONS-based and government-reported estimates of $R(t)$. The dark green line shows the trend line from the linear model (with dark green shading corresponding to 95% confidence intervals) regressing the government-published estimates on the ONS-based estimates. Rows of points are due to the limited precision available (due to rounding before publication) for government-published estimates. (D) The ONS-based estimated instantaneous growth rate ($\hat{r}(t)$) (blue; median line with 95% credible intervals) alongside the government-published estimates ($\tilde{r}(t)$) (red; 90% confidence intervals). The dashed line highlights the epidemic growth threshold of 0. (E) Scatterplot (light green points) showing the relationship between the ONS-based and government-reported estimates of $r(t)$. The dark green line shows the trend line from the linear model (with dark green shading corresponding to 95% confidence intervals) regressing the lagged government-published estimates on the ONS-based estimates. Rows of points are due to the limited precision available (due to rounding before publication) for government-published estimates.

days) applied to the government-published estimates as in the main text. For $R(t)$, we see that the two time series have a strong relationship regardless of lag, with $\hat{\beta}_1$ sitting consistently around 1. This is in contrast to $r(t)$, for which $\hat{\gamma}_1$ approaches 1 at approximately the same lag shown to maximise R^2 in the models with intercepts in the main text.

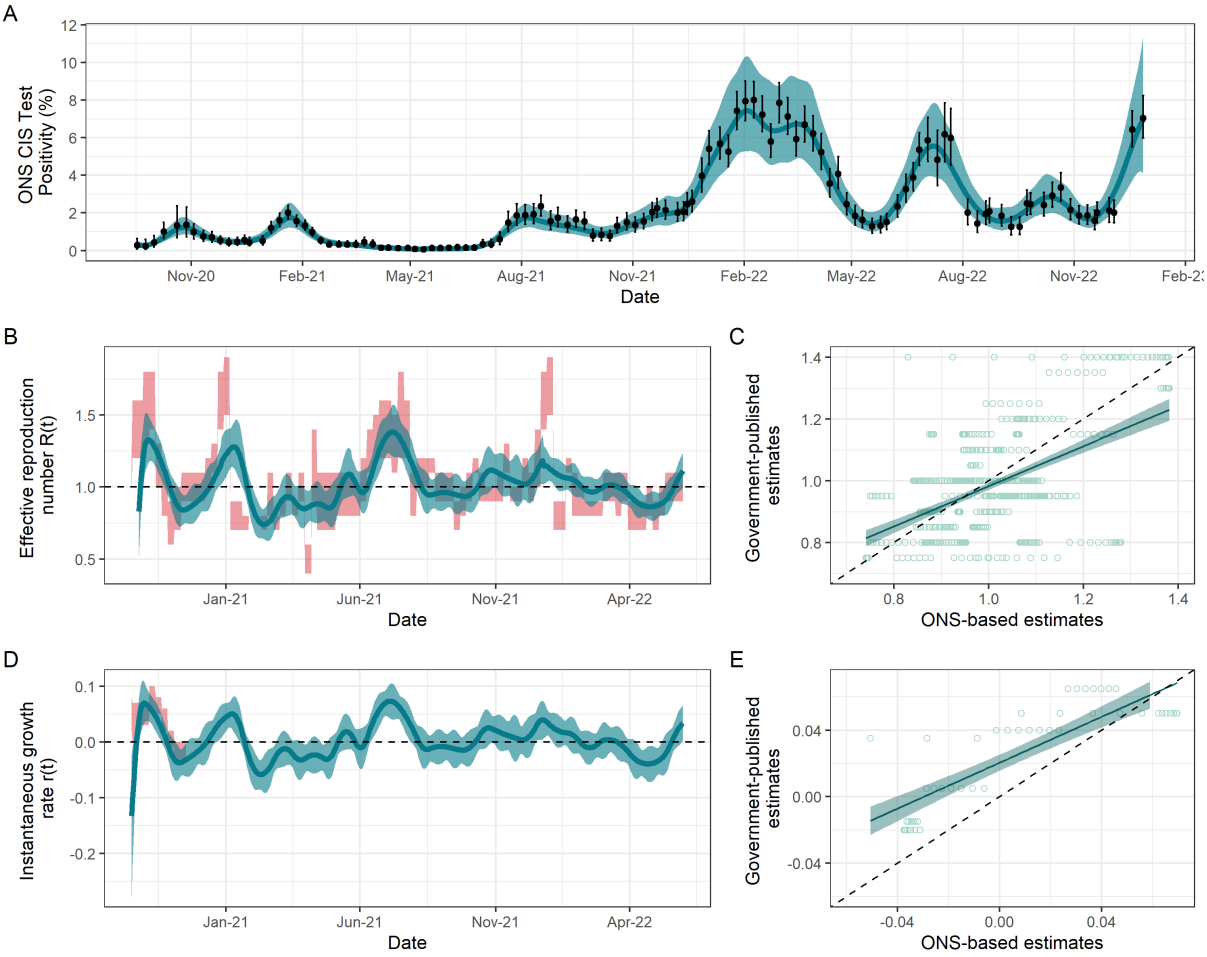

Figure F.14: Fit to ONS CIS data and resulting estimates of $R(t)$ and $r(t)$ for Northern Ireland without any lags. (A) Spline model fit ($\hat{p}(t)$) (blue; median line with 95% credible intervals) to data (black points; ONS point estimate with 95% credible intervals) with the number of knots totalling 40% of the total data points. (B) The ONS-based estimated effective reproduction number ($\hat{R}(t)$) (blue; median line with 95% credible intervals) alongside the government-published estimates ($\tilde{R}(t)$) (red; 90% confidence intervals). The black dashed line highlights the epidemic growth threshold of 1. (C) Scatterplot (light green points) showing the relationship between the ONS-based and government-reported estimates of $R(t)$. The dark green line shows the trend line from the linear model (with dark green shading corresponding to 95% confidence intervals) regressing the government-published estimates on the ONS-based estimates. Rows of points are due to the limited precision available (due to rounding before publication) for government-published estimates. (D) The ONS-based estimated instantaneous growth rate ($\hat{r}(t)$) (blue; median line with 95% credible intervals) alongside the government-published estimates ($\tilde{r}(t)$) (red; 90% confidence intervals). The dashed line highlights the epidemic growth threshold of 0. (E) Scatterplot (light green points) showing the relationship between the ONS-based and government-reported estimates of $r(t)$. The dark green line shows the trend line from the linear model (with dark green shading corresponding to 95% confidence intervals) regressing the lagged government-published estimates on the ONS-based estimates. Rows of points are due to the limited precision available (due to rounding before publication) for government-published estimates.

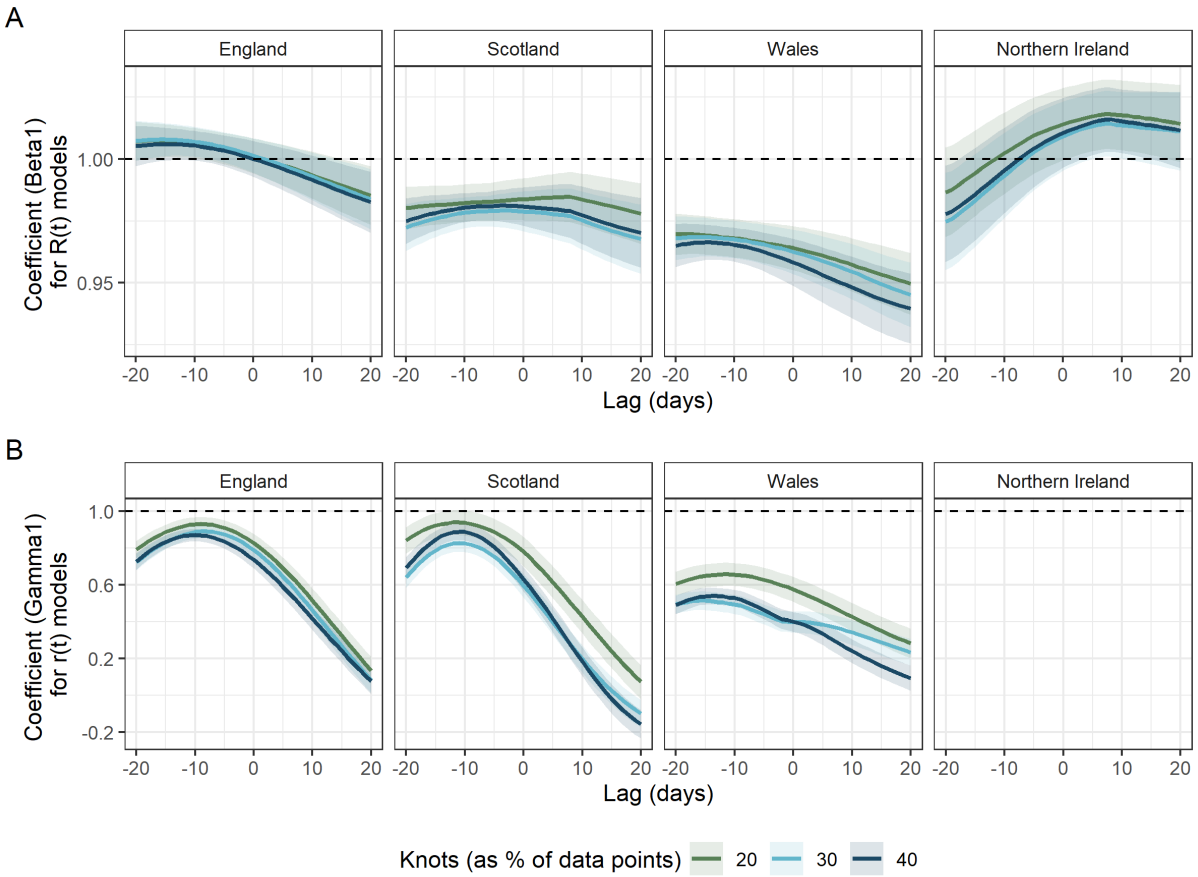

Figure F.15: Coefficients (ordinary least squares (OLS) estimate (solid lines) with 95% confidence intervals (shaded areas) for $\hat{\beta}_1$ and $\hat{\gamma}_1$ from linear regression models regressing ONS-based estimates on lagged government-published estimates without the inclusion of an intercept, for England, Scotland, Wales and Northern Ireland under different levels of knots, given as a percentage of the total data points. Dashed line indicates coefficient value of 1. (A) Output for $R(t)$. (B) Output for $r(t)$. Northern Ireland is omitted due to the lack of government-published estimates for which to compare to.

Appendix G

Additional work published during the DPhil

1. G Cuomo-Dannenburg*, K McCain*, **R McCabe*** et al. Marburg Virus Disease outbreaks, mathematical models, and disease parameters: a Systematic Review. *The Lancet Infectious Diseases* 2023. https://doi.org/10.1016/S1473-3099(23)00515-7.

2. MA Hassan, H Yarow, **R McCabe** et al. Mapping the Middle East Respiratory Syndrome (MERS) related Research - A Scoping Review (2012-2023). *medRxiv* 2023. https://doi.org/10.1101/2023.11.08.23298197.

3. S Hayes*, **R McCabe***, D De Angelis et al. Modelling hepatitis C infection acquired from blood transfusions in the UK between 1970 and 1991 for the Infected Blood Inquiry. *medRxiv* 2023. https://doi.org/10.1101/2023.11.03.23298029.

4. Statistics Expert Group. Expert Report to the Infected Blood Inquiry: Statistics. *The Infected Blood Inquiry* 2022. Available from: https://tinyurl.com/2ddhyfed.

5. BM Soumé, DFY Da, **R McCabe** et al. Adapting field-mosquito protocols for near-infrared spectroscopy implementation. *Parasites and vectors* 2022; 15(338). https://doi.org/10.1186/s13071-022-05458-6.

6. L Gray, BC Asay, B Hephaestus, **R McCabe** et al. Back to the future: Quantifying wing wear as a method to measure mosquito age. *American Journal of Tropical Medicine and Hygiene* 2022; 107(3):689-700. https://doi.org/10.4269/ajtmh.21-1173.

7. DF Da, **R McCabe**, MB Soume et al. Detection of Plasmodium falciparum in laboratory and naturally infected mosquitoes using near-infrared spectroscopy. *Scientific Reports* 2021; 11:10289.

https://doi.org/10.1038/s41598-021-89715-1.

8. **R McCabe***, MD Kont*, N Schmit* et al. Modelling ICU capacity under different epidemiological scenarios of the COVID-19 pandemic in three western European countries. *International Journal of Epidemiology* 2021; 50(3):753-767. https://doi.org/10.1093/ije/dyab034.

9. P Christen*, JC D'Aeth*, A Løchen*, **R McCabe***, D Rizmie*, N Schmit* et al. The J-IDEA Pandemic Planner: A Framework for Implementing Hospital Provision Interventions During the COVID-19 Pandemic. *Medical Care* 2021; 59(5):371-378. https://doi.org/10.1097/MLR.0000000000001502.

10. H Fu, X Xi, H Wang et al. A database for the epidemic trends and control measures during the firstwave of COVID-19 in mainland China. *International Journal of Infectious Diseases* 2021 102:463-471. https://doi.org/10.1016/j.ijid.2020.10.075.