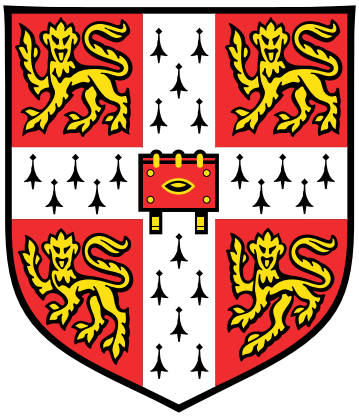
Multi-state modelling to estimate infectious disease burden



Peter Douglas Kirwan

Clare Hall
University of Cambridge

This thesis is submitted for the degree of
Doctor of Philosophy

March 2024

For Mum.

Declaration

I hereby declare that this thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the preface and specified in the text. It is not substantially the same as any work that has already been submitted before for any degree or other qualification except as declared in the preface and specified in the text. It does not exceed the prescribed word limit for the School of Clinical Medicine and Clinical Veterinary Medicine Degree Committee.

Peter Douglas Kirwan
March 2024

Acknowledgements

This PhD thesis is a product of the support of my supervisors, advisors, collaborators, family, and friends, and would not have been possible without all their encouragement and support.

First and foremost, I am incredibly grateful to my supervisors Dr Anne Presanis, Dr Paul Birrell, and Professor Daniela De Angelis who have challenged, supported, and inspired me in equal measure throughout four extraordinary years. I would like to especially thank Anne whose thoughtful guidance and encouragement has shaped me as a researcher.

I am also indebted to my advisors and collaborators, both at the MRC Biostatistics Unit and further afield. Particular thanks to Dr Shaun Seaman, Dr Chris Jackson, Dr Brian Tom, and Dr Andre Charlett for all the advice and useful discussion; to Dr Sara Croxford, Meaghan Kall, Dr Alison Brown, Cuong Chau, and Dr Valerie Delpech for encouraging me to begin this journey and for their support ever since; and to Dr Victoria Hall, Sarah Foulkes, Jameel Khawam, Dr David Wyllie, Ranya Mulchandani, Dr Sema Mandal, Suzanne Elgohari, and Dr Ferdinando Insalata for stimulating conversation and collaboration during the height of the pandemic.

Thank you to the UK Medical Research Council for supporting my research, to Dr Brian Tom, Dr Rob Goudie, and Dr Angela Talbot for their support of the PhD programme, and to my college tutors Dr Lori Passmore and Dr Holly Hedgeland for pastoral support.

Thank you to my dad, brother, and all of my family for their genuine interest and support over the years. Thank you to my oldest friends Jez, Chris, Ollie, Rafal, and Gorkem, who I could always rely on to not take things too seriously. Thank you to all my friends in Cambridge, especially to Debbie, Anna, Hubert, Simone, Esther, Sarah, Lukas, Ruta, Nia, and Ciara. And to Katie, Krathika, Tori, and Guilia thank you for making our house a home. Thank you to all my friends at the MRC Biostatistics Unit, especially Josh, Chiara, and Andrew who shared the ups and downs of PhD-ing during a pandemic. Thank you to all my friends at Clare Hall and CHBC for always keeping life interesting, and to the Newnham Beavers and leaders for reminding me of the value of community.

Last but not least, I would also like to acknowledge all those living with HIV, those who were admitted to hospital during the COVID-19 pandemic, and all the SIREN study participants who contributed their data, as well as the doctors, nurses, data reporters, and research staff without whom this research would not have been possible.

Abstract

Infectious diseases have plagued humanity for millennia, with outbreaks such as cholera and smallpox claiming countless lives. As was demonstrated by COVID-19, our interconnected world is now more susceptible than ever to the global spread of an infectious disease pandemic.

Despite advances in statistical methods and comprehensive public health surveillance, understanding the many facets of infectious disease burden remains a challenge. Questions persist of how fast a disease is spreading, who in society is most at risk, and which interventions might have the capability to save lives.

In this thesis I aim to develop and apply multi-state models to estimate infectious disease burden from observational data. Multi-state models are a flexible approach to estimating risks and lengths of stay, and capable of accounting for data biases and censoring. I demonstrate the potential of these models to address certain limitations of classical survival analysis methods, and their application to answer causal questions using observational data.

Firstly, I develop an established multi-state back-calculation method to better estimate the incidence processes underlying the HIV epidemic in England. The updated method is shown, via simulation, to account for measurement bias in markers of disease progression. This method is also shown to be robust to interruptions in HIV testing activity and transmission which may have occurred as a result of COVID-19 lockdowns.

Next, I apply multi-state models to patient admissions for COVID-19 to estimate hospitalised fatality risks and lengths of stay in hospital using a variety of linked hospital datasets. Changes in patient prognosis are estimated throughout each wave of infection, as well as the beneficial effect of vaccination, and the deleterious effect of hospital pressure on patient outcomes.

Finally, I use multi-state models to estimate the real-world effectiveness of SARS-CoV-2 vaccination among a cohort of healthcare workers in England. These models are applied in a causal inference framework and a 14% effectiveness of fourth dose vaccination against infection is estimated for this cohort, relative to a waned third dose.

Multi-state models are a powerful tool for the estimation of infectious disease burden and have proven invaluable for research concerning the COVID-19 pandemic. The methods and applications in this thesis will inform future areas of research: accounting for population migration in HIV incidence estimates, developing an improved metric of hospital pressure, and assessing the effectiveness of future SARS-CoV-2 vaccine roll-outs.

Preface

Much of the work included in this thesis has been published in scientific journals. These publications were invariably collaborative, involving other researchers and public health professionals. My contributions are outlined below.

- The late HIV diagnosis reclassification methodology described in Chapter 3 and Appendix C was developed in collaboration with epidemiologists at the UK Health Security Agency (UKHSA). I helped to define the research question, extracted the data and did the analysis, defined the final reclassification methodology, and wrote the manuscript (Kirwan, Croxford *et al.* 2022).

- The hospitalised severity analyses described in Chapter 4 were completed in collaboration with epidemiologists and statisticians at UKHSA and the MRC Biostatistics Unit. I contributed to the design of the analysis and the statistical models, cleaned the data and ran the statistical analysis, and wrote the two manuscripts (Kirwan, Charlett *et al.* 2022; Kirwan, Elgohari *et al.* 2021).

- The Cox proportional hazards analyses described in Chapter 5 and Appendix E was a collaboration with epidemiologists and statisticians at UKHSA. I gave significant input into the research: preparing the dataset, implementing the models in R, and generating the estimates, and was a co-author on the two manuscripts (Hall, Foulkes, Insalata *et al.* 2022; Hall, Insalata *et al.* 2024).

- The multi-state model analysis described in Chapter 5 was undertaken in collaboration with epidemiologists and statisticians at UKHSA and the MRC Biostatistics Unit. I designed the multi-state models, cleaned the data and ran the statistical analysis, and wrote the manuscript (Kirwan, Hall *et al.* 2024).

Impact statement

The overarching objective of this thesis has been to develop and apply statistical modelling approaches for public health data monitoring, thereby enhancing our collective understanding of ongoing disease epidemics and providing evidence for decision makers and the wider community. The incidence, burden of disease, and effectiveness of interventions are perhaps the three most important quantities of interest for managing any infectious disease. Improving the precision and accuracy of these metrics whilst accounting for biases in biomedical and public health data is, therefore, a goal for public health data science research. The specific contributions that this thesis has made to infectious disease research are two-fold: increasing knowledge for public health action, and methodological developments in monitoring disease epidemics.

Increasing knowledge for public health action

I generated annual estimates of HIV incidence and undiagnosed HIV prevalence for national surveillance reports and the national action plan on HIV elimination, which have been used to inform HIV prevention campaigns and national policy decisions.

My estimates of hospitalised fatality rates and lengths of stay in hospital for COVID-19 informed expert scientific panels mobilised during the initial months of the pandemic. Meanwhile, my estimates of SARS-CoV-2 vaccine effectiveness among a healthcare worker cohort informed decisions about national vaccine policy, specifically the roll-out of fifth booster vaccines to healthcare workers.

Methodological developments in monitoring disease epidemics

I advanced established statistical methods in HIV estimation; developing and evaluating a revised definition of late HIV diagnosis, enhancing the precision of an existing multi-state model through the inclusion of additional biomarker data, and exploring the sensitivity of model estimates to counterfactual assumptions. The revised late HIV diagnosis definition has since become the standard approach for assessment of this measure, both in the UK and in Europe, and estimates from the dual biomarker method for HIV estimation continue to inform progression towards England's Action Plan on HIV.

The COVID-19 pandemic necessarily required the application of novel methodologies to better understand the newly emerging virus and its sub-variants. I developed multi-state models of hospitalisation and vaccination which were tailored to the specifics of different surveillance systems, and extended vaccine effectiveness models in formal causal inference analyses.

Research publications

The research presented in this thesis has been disseminated via three first-author publications in peer-reviewed journals, alongside several co-authored publications and conference presentations (see Appendix A). My work to increase our understanding of the emerging COVID-19 pandemic was recognised in 2022 with the Scientific Pandemic Influenza Group on Modelling (SPI-M) Award for Modelling and Data Support (SAMDS).

***

Aside from the contributions to research generated by this thesis, working on infectious disease modelling during a pandemic has provided me with the unique opportunity to undertake research whilst understanding of this novel pathogen was still developing. I was fortunate to collaborate with a wide range of statisticians, scientists, doctors, epidemiologists, public health professionals, and government advisors in the national and international response to COVID-19. The various projects I have been involved with have broadened my knowledge of infectious disease statistics and epidemiology, improved my understanding of the role of statistics in public health decision making, and enhanced my ability to communicate statistical concepts.

Nevertheless, there has been an unavoidable impact of the pandemic on my PhD and this thesis is a combination of my original project focused on HIV back-calculation, and my subsequent research into COVID-19 severity and protection against infection. Whereas the HIV research had a clearer set of aims and well-defined methodology, the COVID-19 research was faster-paced, with an evolving research question, and recurrent data updates. Central to this research has been the application of statistical methods to estimate infectious disease burden, and this theme is explored in different ways throughout this thesis.

Contents

List of Figures

List of Tables

Nomenclature

Statistical nomenclature

α, β, γ Greek letters represent parameters

$\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}$ bold Greek letters represent vectors of parameters

$\hat{\cdot}$ a hat over an estimand represents the estimator

$\mathbf{P}, \mathbf{Q}$ bold capital Latin letters represent transition matrices

$\mathbb{1}\{A\}$ indicator function: $\mathbb{1}\{A\} = \begin{cases} 1, & \text{if } A \text{ is true} \\ 0, & \text{if } A \text{ is false} \end{cases}$

$|$ given/conditional on

$\Pr(A)$ probability that event A takes place

$\propto$ is proportional to

$\sim$ distributes according to

$F(x)$ distribution function of the random variable X evaluated at x

$f(x)$ probability density function of the random variable X evaluated at x

X, Y, Z capital Latin letters represent random variables

x, y, z small Latin letters represent a given value of a random variable

Acronyms / Abbreviations

AIC Akaike information criterion

AIDS Acquired Immune Deficiency Syndrome

AMTI Artificial manipulation of transition intensities

ART Antiretroviral therapy

CD4 Cluster of differentiation 4

CDSC Communicable Disease Surveillance Centre

CF Correction factor

CFR Case fatality risk

CI Confidence interval

CIS Coronavirus Infection Survey

COVID-19 Coronavirus disease 2019

CrI Credible interval

DAG Directed acyclic graph

ECDC European Centre for Disease Prevention and Control

EHR Electronic hospital record

EU European Union

EW&NI England, Wales, and Northern Ireland

FRR False recency rate

GBM Gay, bisexual and other men who have sex with men

GDPR General Data Protection Regulation

GP General practitioner

GUMCAD Genitourinary medicine clinic activity dataset

HDU High dependence unit

HFR Hospitalised fatality risk

HIV Human Immunodeficiency Virus

i.i.d. Independent and identically distributed

ICNARC Intensive Care National Audit & Research Centre

ICU Intensive care unit

IFR Infection fatality risk

IMD Index of multiple deprivation

JCVI Joint Committee on Vaccination and Immunisation

LAg-Avidity EIA Limiting Antigen Avidity Enzyme Immunoassay

MDRI Mean duration of recent infection

MLE Maximum likelihood estimate

MPES Multi-parameter evidence synthesis

MTB *Mycobaterium tuberculosis*

NHS National Health Service

PCR Polymerase chain reaction

PEP Post-exposure prophylaxis

PHE Public Health England

PMSE Predictive mean squared error

PrEP Pre-exposure prophylaxis

RCT Randomised controlled trial

RITA Recent infection testing algorithm

SARS-CoV-2 Severe acute respiratory syndrome coronavirus 2

SHS Sexual health services

SIREN SARS-CoV-2 Immunity & Reinfection Evaluation study

SPI-M Scientific Pandemic Influenza Group on Modelling

STI Sexually transmitted infection

SUTVA Stable unit treatment value assumption

U=U Undetectable = untransmittable

UK United Kingdom

UKHSA United Kingdom Health Security Agency

VE Vaccine effectiveness

VL Viral load

WHO World Health Organisation

1. Introduction

1.1 Infectious diseases

1.1.1 The nature of infectious diseases

Infectious diseases are caused by pathogens: microorganisms such as viruses, bacteria, or fungi which spread from one individual to another. Pathogens may transmit directly from person-to-person (in the case of respiratory and blood-borne pathogens), or indirectly (as for vector-borne pathogens) (Leung 2021). Once a pathogen enters the body it invades cells and multiplies, and this process of infection can lead to illnesses ranging in severity from mild to life-threatening (Alberts *et al.* 2002). The classic model of infectious disease considers a combination of three areas: the infectivity and virulence of the pathogen, the susceptibility of the individual, and their vulnerability due to environmental factors (van Seventer and Hochberg 2017).

Infectious diseases play a significant role in the global burden of disease, and, given their disproportionate impact on vulnerable populations and outbreak potential, are a key area for public health research (GBD 2019 Diseases and Injuries Collaborators 2020). The vector-borne infectious disease malaria has had perhaps the greatest impact of any infectious disease in history, causing the deaths of an estimated 1 in 10 affected individuals during the 19th century (Carter and Mendis 2002). Malaria continues to affect millions worldwide and, alongside Human Immunodeficiency Virus (HIV), *Mycobaterium tuberculosis* (MTB), and emerging infectious diseases such as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), is nowadays among the leading causes of mortality globally (van Seventer and Hochberg 2017).

1.1.2 Key concepts in epidemiology

Epidemiology is the study and analysis of patterns of diseases and determinants of health in a population. In the context of infectious diseases, epidemiologists and statisticians employ several standard measures of disease burden to assess the impact of a disease, and to guide the appropriate public health response. These measures include (Coggon *et al.* 1978):

1. the **incidence of infection**, defined as the rate of new infections occurring over a specific period time, e.g. new cases per 1,000 people per year;

2. the **prevalence of infection**, defined as the proportion of the population that have the disease at a specific point in time, regardless of when they became infected;

3. for diseases with adverse outcomes, the **mortality associated with infection**, defined as the incidence of death from the disease.

A key measure of the spread of a disease, related to the incidence of infection, is the basic reproduction number, denoted R_0. The R_0 value measures the expected number of cases directly generated by one case in a population where all individuals are susceptible to infection (Lipsitch, Cohen *et al.* 2003). For instance, if $R_0 = 3$ for a particular disease then one infected individual will, on average, infect three others. When $R_0 > 1$ an infection will spread in a population, whereas when $R_0 < 1$ the infection will eventually die out.

Importantly, not every individual within a population will be equally susceptible to infection. Immunity to a pathogen might be acquired through vaccination or recovery from previous infection, both of which can provide a level of protection against future infections. Through widespread vaccination, immunity can be conferred to a large portion of a population, reducing susceptibility and thereby decreasing R_0 (Mongin *et al.* 2023).

1.1.3 Historic and contemporary epidemics

The terms epidemic, pandemic, and endemic are used to describe the geographical spread and prevalence of infectious diseases. Epidemic refers to an increase in the number of cases of a disease, above what is usually expected; a pandemic is an epidemic which has crossed country borders and affects a large proportion of people; and endemic is the constant or usual prevalence of a disease within a geographic area (van Seventer and Hochberg 2017).

The bubonic plague epidemic: 1346–1352

The bubonic plague, also known as the 'Black Death', is a notable example of a severe and highly infectious epidemic. Caused by the bacterium *Yersinia pestis*, which is spread by fleas living on small mammals, the Black Death epidemic that spread through Europe between 1346 to 1352 had an estimated mortality rate of up to 60% (Benedictow 2004).

The HIV epidemic: 1981–today

A more recent epidemic is that of Human Immunodeficiency Virus (HIV). HIV is a retrovirus that targets the immune system, particularly CD4+ (Cluster of differentiation 4) T cells, and is transmitted through contact with blood or other bodily fluids; most often via sexual contact, contact with blood and blood products, or vertically from mother to child (Delpech 2022). Without effective treatment, HIV infection leads to a progressive decline in CD4+ T cells,

resulting in a weakened immune system and increased susceptibility to opportunistic infections and cancers; a condition known as Acquired Immune Deficiency Syndrome (Lipman 2006).

The first cases of AIDS illness were reported among young men in Los Angeles in 1981 (The Lancet 2021), and the viral cause discovered in 1983 (Barré-Sinoussi *et al.* 1983; Gallo *et al.* 1984). According to statistics from the Joint United Nations Programme on HIV/AIDS (UNAIDS) 2023a, since the start of the HIV epidemic, 85.6 million (95% confidence interval (CI) 64.8 to 113.0 million) people have acquired HIV worldwide, and 40.4 million (95% CI 32.9 to 51.3 million) have died from AIDS-related illnesses. HIV remains an urgent issue in much of the world, with substantial disparity between regions — adult HIV prevalence in sub-Saharan African countries with endemic HIV may exceed 15%, whereas prevalence in many western European countries is below 0.1% (Joint United Nations Programme on HIV/AIDS (UNAIDS) 2023b).

The COVID-19 pandemic: 2020–2023

The most widespread pandemic of modern times has been coronavirus disease 2019 (COVID-19). First detected in November 2019 in Wuhan, China, the SARS-CoV-2 virus which causes COVID-19 would subsequently spread to all areas of China, then around the world (South China Morning Post 2020). SARS-CoV-2 was designated a Public Health Emergency of International Concern by the World Health Organisation (WHO) in January 2020, and officially declared a pandemic in March 2020, a designation which would last for more than 3 years, with the impact of COVID-19 felt in every country around the globe (World Health Organisation 2020; World Health Organisation 2023).

Although unclear at the time, it is now known that SARS-CoV-2 is primarily transmitted through airborne virus particles, with symptoms usually appearing within 1–2 days of infection (UK Health Security Agency 2022b). Most people recover from SARS-CoV-2 infection, but it can be a fatal condition, especially for those with pre-existing comorbidities (Docherty, Mulholland *et al.* 2021).

As of March 2024, there have been 775 million confirmed cases of COVID-19 globally, and more than 7 million deaths attributed to the disease (World Health Organization 2020b). The total number of infections is likely to be substantially higher, however, due to a combination of under-reporting of confirmed cases and insufficient testing capacity for symptomatic cases.

1.2 Public health response to infectious diseases

Whilst the characteristics, infection mechanisms, and scale of the HIV and SARS-CoV-2 epidemics are dissimilar, parallels can be drawn: both viruses arose from an unknown (probable zoonotic) source (Contini *et al.* 2020), both have caused extreme prejudice towards marginalised

communities (Bhanot *et al.* 2020; Hibbert *et al.* 2018), and government intervention and new public health measures have been required to mitigate the impact on healthcare services (Brauner *et al.* 2021; Brown, Nash *et al.* 2018).

1.2.1 Strategies for infectious disease prevention and control

Given the immense global impact of both HIV and SARS-CoV-2, the importance of the public health response and efforts to effectively prevent, diagnose, and manage infectious diseases cannot be overstated. In May 2020, Hargreaves *et al.* 2020 proposed three 'lessons learnt' from the HIV epidemic for the COVID-19 response: anticipating health inequalities; supporting behavioural changes; and multidisciplinary efforts to design and evaluate public health interventions. The public health interventions applied in both epidemics have so far included: testing, physical barriers, social distancing, and pharmaceutical interventions.

Testing, physical barriers, and social distancing

The United Kingdom (UK) has a long history of HIV prevention and control. In 1986, the UK became one of the first countries to introduce a national needle exchange programme, designed to reduce HIV transmission among people who inject drugs (Vlahov and Junge 1998). 'Don't Die of Ignorance', a national HIV prevention campaign, launched the following year in the UK (Figure 1.1) and has been credited with raising awareness of HIV and AIDS among the general public (Fowler 2014).

Figure 1.1: AIDS: Don't die of ignorance leaflets that were sent to all British households in 1987. *Source: www.campaignlive.co.uk.*

Testing is vital to the public health response to HIV, since early diagnosis provides the best opportunity for treatment initiation and prevention interventions to limit onward transmission (Cohen, Gay *et al.* 2010). In recent years, public health guidelines have recommended that all patients in 'high' (>2 per 1,000) and 'very high' (>5 per 1,000) diagnosed prevalence areas

be offered routine HIV testing when accessing healthcare services or registering with a general practitioner (GP) (Gulland 2011; Health Protection Agency 2011; National Institute for Health and Care Excellence 2016b).

Today, HIV prevention policies advocate for a combination of strategies, including regular community point of care testing, self-testing, and home-testing (National Institute for Health and Care Excellence 2016b), as well as condom usage (Pebody 2021a), partner notification, and pharmaceutical interventions. Campaigns to normalise regular HIV testing, such as European Testing week, are run each year with promotion by UK-based third-sector organisations (European Testing Week 2023).

Testing has also formed the cornerstone of the response to the COVID-19 pandemic, with the number of weekly polymerase chain reaction (PCR) tests undertaken in England being scaled up to over a million per week during the initial months of the pandemic (UK Government 2021). Individuals who tested positive were instructed to self-isolate for up to 10 days at home to reduce the risk of passing on the infection (UK Health Security Agency 2022c). Other interventions aimed at containing the spread of the virus included: social distancing guidance (Figure 1.2); advice on the use of face masks and hand-washing; and local and national 'lockdowns' which encompassed restrictions on movement and the closure of schools, workplaces, and hospitality venues (Knock *et al.* 2021).

Figure 1.2: Social distancing notice on a pavement in Cambridge, UK. *Source: www.cambridgenetwork.co.uk.*

Treatment and vaccination programmes

Once developed, pharmaceutical interventions including vaccination and treatment can be highly effective, if costly, tools to reduce the longer-term impact of infectious diseases (Schrappe and Lauterbach 1998). Monitoring their effectiveness in the real-world is, therefore, a major focus of public health research (UK Health Security Agency 2021).

In the early years of the HIV epidemic, being diagnosed with HIV was considered a death sentence (Collaborative Group on AIDS Incubation and HIV Survival 2000; Pezzotti *et al.* 2003; Rothenberg *et al.* 1987), and, while no cure currently exists, the introduction of effective antiretroviral therapy (ART) in the mid-1990s, which suppresses the HIV virus in the body, has transformed HIV into a manageable, if chronic, condition (Deeks *et al.* 2013). Provided they are diagnosed early in infection and adhere to treatment, the life expectancy of people diagnosed with HIV is now comparable to that of the general population (Nakagawa *et al.* 2013). With treatment access becoming more widespread across HIV-endemic regions of sub-Saharan Africa, the number of AIDS-related deaths worldwide has plummeted; from an estimated 1.3 million (95% CI 970,000 to 1.8 million) in 2010 to 630,000 (95% CI 480,000 to 880,000) in 2022 (Joint United Nations Programme on HIV/AIDS (UNAIDS) 2023a).

ART is also an effective way of preventing onwards transmission of HIV, with studies demonstrating that individuals with an undetectable viral load are unable to pass on HIV (known as undetectable = untransmittable, or U=U) (Cohen, Chen *et al.* 2016; World Health Organization 2023). As part of combination prevention, pre-exposure prophylaxis (PrEP), which can help to prevent HIV acquisition for a HIV-negative individual, has been targeted towards key HIV prevention groups in the UK and worldwide (Brown, Mohammed *et al.* 2017; Rivet Amico and Bekker 2019).

The development of pharmaceutical interventions for SARS-CoV-2 has been incredibly rapid, with an unprecedented international effort by private and public institutions (Krammer 2020). As a result, less than a year after the detection of SARS-CoV-2, two COVID-19 vaccine candidates were granted approval in the UK (Medicines and Healthcare products Regulatory Agency 2020; Sharma *et al.* 2020). Within a year of this approval, 40 million people in England had received two vaccine doses, and 30 million had also received a third dose (UK Government 2021). This extensive vaccination campaign dramatically changed the outlook for COVID-19 in England, reducing the risk of infection, lessening symptoms, and reducing morbidity and mortality for those who became infected (Hall, Foulkes, Saei *et al.* 2021; Lopez Bernal *et al.* 2021; Public Health England 2021a).

1.2.2 The role of surveillance data

In the context of public health, surveillance data play a pivotal role in monitoring and responding to infectious diseases. The UK Health Security Agency (UKHSA) is responsible for infectious disease monitoring and surveillance in the UK as part of its mission to prepare for, prevent, and respond to health threats, with the goal of saving lives and protecting livelihoods (UK Health Security Agency 2023c).

Surveillance data can take various forms, including case and aggregate reporting. Case reporting involves the collection of detailed individual-level data about each case of a disease,

such as demographic information, clinical symptoms, and outcomes. Aggregate reporting, meanwhile, involves the collection of summary data about groups of cases (Murray and Cohen 2017). In emerging public health threats, timely case-surveillance data enable rapid response by local public health teams, e.g. helping to pinpoint outbreaks of diseases such as measles, and isolating individuals who may be affected to prevent further spread (Jajosky and Groseclose 2004). Case-surveillance data can also provide a comprehensive picture for routine monitoring of infectious diseases, enabling record linkage over time, and follow-up of individuals through different care pathways (Trostian *et al.* 2022).

National public health agencies such as UKHSA routinely publish surveillance data as national statistics. These provide a snapshot of the current state of infectious diseases within the country, trends in recent years, and are used to inform public health policies (UK Health Security Agency 2020). At an international level, the World Health Organisation (WHO) plays a similar role in the monitoring of infectious disease, collecting data from countries and providing a global perspective on infectious diseases (World Health Organization 2020b).

1.3 Estimating the burden of infectious disease

To understand how quickly a virus is spreading through a population, and the corresponding burden of disease, timely and robust estimates are required (De Angelis *et al.* 2015). The incidence of new cases, prevalence of undiagnosed infection, and extent of disease severity can all inform policy makers seeking to implement effective and appropriate public health measures.

1.3.1 Incidence and prevalence

Incidence of infection and the prevalence of an infectious disease in a population are generally unobserved, with observable information restricted to testing outcomes among subsets of a population. Testing processes may themselves be influenced by unobserved parameters, such as an individual's propensity to test, and the time from infection to symptom presentation (if symptoms present at all) (Lewis and White 2017).

To reconstruct the incidence and prevalence of infection from the available testing data, statistical and mathematical modelling techniques which adjust for sampling biases are employed. These techniques combine information from a range of sources to make valid statistical inference about unobserved incidence and prevalence (Hallett *et al.* 2008; Karon *et al.* 2008).

1.3.2 Severity

Common metrics to understand the severity of infectious diseases and the factors associated with adverse outcomes are the infection fatality risk (IFR) and case fatality risk (CFR). The IFR

is defined as the conditional probability of death given infection, whilst the CFR is defined as the conditional probability of death given a *confirmed case* of infection (Perez-Guzman *et al.* 2023). Since infections resulting in death are more likely to be recorded, and mild infections (particularly asymptomatic infections) less likely to be recorded, the CFR will almost always be higher than the IFR. Statistical modelling can be used to relate these measures of severity to underlying characteristics collected through case-surveillance data (Wong *et al.* 2013).

1.3.3 Effectiveness of interventions

Monitoring outcomes among a target group after the introduction of a public health intervention allows for the real-world effectiveness of the intervention to be assessed, and changes over time to be detected. For example, effectiveness estimates might consider reductions in the transmission or acquisition of an infection among a treated group. Making these estimates requires reliable testing data, recognition of potential biases, and methodologies which can account for these biases (Brauner *et al.* 2021).

1.3.4 Challenges in estimating infectious disease burden

The majority of the data available for infectious disease monitoring relies on observational data, where a population is followed without actively intervening on an exposure (Baker 2023). As a result, estimates of infectious disease burden are often considerably complicated by selection bias, which arises when the selected population observed is different from the general population of interest.

Common causes of selection bias include: testing policies (which may vary nationally and sub-nationally), individual health-seeking behaviour, test availability, clinician decision-making, and detection rates of different tests (Hernán, Hernández-Díaz *et al.* 2004). Inaccuracies in the available case-surveillance data may also arise: at data input, due to incomplete/inaccurate reporting or misdiagnosis; and at data collection, due to insufficient surveillance infrastructure and data processing methods.

Unless the potential sources of bias are well-understood and accounted for they can affect the validity of estimates (Westreich 2012). For example, the IFR measure of severity requires complete knowledge of both the number of infections and number of deaths. Particularly in the early stages of a pandemic, when testing is either unavailable or sub-optimal, these quantities may be impossible to measure. The CFR is more straightforward to assess, since confirmed infections are more likely to be recorded, but this measure is highly sensitive to testing. This was particularly the case during the initial wave of the COVID-19 pandemic, when biases in case-surveillance led to wide variability in national estimates of the case-fatality rate, from 0.1% to over 25% (UK Health Security Agency 2020; World Health Organization 2020a).

Even when knowledge of these multi-faceted biases is available, statistical methodologies must be carefully chosen to properly control for the biases.

1.3.5 The role of multi-state models

Multi-state models are a type of statistical model which may be used to describe health-related processes over time, particularly those relating to disease epidemics (van den Hout 2016; Vynnycky and White 2010). The benefit of multi-state models for these applications is their flexibility, which allows for a wide array of disease processes to be modelled (Hougaard 1999). In a multi-state model, an individual begins in a given state, spends a period of time in that state, before moving to the next state. Figure 1.3 shows a three-state model where individuals move between states 1, 2, and 3 according to transitions $\{q_{r,s} : r, s \in (1,2,3)\}$.

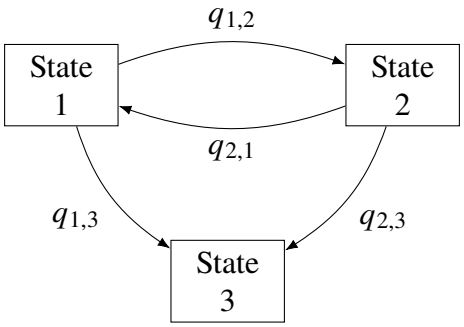


Figure 1.3: Example of a three-state model with transitions between states according to transitions $\{q_{r,s} : r, s \in (1,2,3)\}$.

Multi-state survival models can be used to model linear processes (e.g. disease progression), cyclic processes (e.g. infection and recovery), and may include absorbing states, such as death (Andersen and Keiding 2002). As with any statistical model, a detailed understanding of the data-generating processes for the available data compared to the target population, and careful model specification is required to control for potential sources of bias (Westreich 2012). However, by incorporating information on intermediate events, these models can provide a more nuanced understanding of disease progression and the effect of interventions than classic survival analyses (see Section 2.4 for further details).

Specific recent applications of multi-state models to infectious disease epidemics have included HIV, pandemic influenza, and COVID-19 (Birrell, Blake, van Leeuwen E, Gent *et al.* 2021; Brizzi, Birrell, Plummer *et al.* 2019; Corbella *et al.* 2018; Jackson *et al.* 2022; Presanis, Harris *et al.* 2021).

1.4 Objectives and thesis structure

1.4.1 Research aims and questions

The research objective for my PhD is to develop statistical modelling approaches which can be applied to public health data, thereby enhancing our collective understanding of ongoing disease epidemics and providing evidence for decision makers and the wider community.

Specifically, I aim to apply multi-state survival models to estimate key quantities of infectious disease burden and the effectiveness of public health interventions in three areas:

1. the incidence of HIV among gay and bisexual men in England, and trends in HIV incidence and undiagnosed HIV prevalence over time;

2. the severity of COVID-19 among individuals hospitalised with the disease at various stages of the pandemic;

3. the effectiveness of SARS-CoV-2 vaccination against infection among a cohort of healthcare workers.

1.4.2 Thesis structure

In this introductory chapter I have introduced key concepts in infectious disease research, with specific reference to two of the most urgent public health epidemics of our time: HIV and COVID-19. I have outlined challenges in estimating infectious disease burden, and the role that multi-state models can play. In the final section of this chapter I describe the data governance procedures I have followed when working with infectious disease data.

In Chapter 2 I detail the statistical methods involved in survival analysis and the multi-state survival models I have applied as part of my PhD research. The nature and biases of various epidemiological study designs are discussed, with an introduction to key concepts in causal inference from observational data.

In Chapter 3 I describe the development of a disease-staged, multi-state HIV back-calculation model to provide a more nuanced interpretation of markers of HIV disease progression. I investigate the sensitivity of this model to interruptions in HIV testing and HIV transmission, as may have occurred during the COVID-19 lockdowns in England.

In Chapter 4 I apply competing risks multi-state survival models to estimate trends in hospital pathways and lengths of stays in hospital for patients with COVID-19, before and after widespread vaccination. I estimate how the prognosis for hospitalised individuals has changed based on patient characteristics and hospital pressures, accounting for specific data biases.

In Chapter 5 I apply multi-state survival models for interval-censored data to estimate fourth dose SARS-CoV-2 vaccine effectiveness, and protection from prior infection, against

symptomatic and asymptomatic infection in a cohort of healthcare workers. These models are set in a causal inference framework to estimate the causal effect of vaccination in this cohort.

Lastly, in Chapter 6, I summarise the specific contributions of my research, explaining what was already known, what this thesis adds, and whether my research aims have been met. I also highlight future applications and areas for future methodological research.

1.5 Data governance

1.5.1 Legal basis for data collection

The data used in my PhD were collected and supplied by UKHSA via an honorary contract. Statutory instrument 2002 No. 1483 in The Health Service Regulations provides the legal basis for data handling by UKHSA. UKHSA is also obligated to comply with the rules and regulations set out in the General Data Protection Regulation (GDPR).

1.5.2 Ethical approval

The surveillance data used in Chapters 3 and 4 were collected by NHS England and the UK Health Security Agency, with permissions granted under Regulation 3 of The Health Service (Control of Patient Information) Regulations 2002, and without explicit patient permission under Section 251 of the National Health Service (NHS) Act 2006. My research was classified as surveillance undertaken as part of UKHSA's legal responsibility to monitor infectious disease, and in response to an urgent public health emergency, and exempted from full ethical review by the UKHSA Research Support and Governance Office (Kirwan, Charlett *et al.* 2022; Kirwan, Croxford *et al.* 2022).

For the SIREN study data used in Chapter 5, ethical approval was granted by the Berkshire Research Ethics Committee on May 22, 2020. The study collected consent from all participants and is registered under ISRCTN clinical trial registry number: 11041050. The trial protocol was published by Wallace *et al.* 2022.

1.5.3 Data protection

All of the data used in my PhD were stored and analysed under agreed data governance protocols, as set out below and in accordance with the Data Protection Act 2018 (Spencer and Patel 2019).

Data sharing agreements, data permissions, and an honorary contract with UKHSA were required before receiving or processing any patient data. All patient-level data were stored securely on UKHSA network drives with restricted access on an encrypted UKHSA laptop. All data analyses were carried out on this laptop, or, where agreed under the data sharing

agreements, fully anonymised summary datasets were encrypted and transferred for analysis on a secure University of Cambridge server with limited permissions. I undertook Caldicott training for data protection and completed UKHSA information governance training on an annual basis.

A data governance policy specific to HIV and STI surveillance systems is that no NHS numbers or patient names are collected. Instead, records are pseudoanyonymised, using a Soundex code (an algorithm for indexing names) (Mortimer and Salathiel 1995) and first initial which allow for unique, non-identifying linkage of an individual's records between systems and over time (GOV.UK 2023). This policy did not impact upon my analyses of HIV data.

Over the course of my PhD, the UK national public health organisation transitioned from Public Health England to UKHSA and I sought renewed permissions and an updated honorary contract to continue using the data.

2. Statistical methods for survival data

This chapter provides an introduction to the statistical methods used throughout the thesis, beginning with a description of survival data and of two classical survival methods: the Kaplan-Meier estimator and Cox proportional hazards model. I explain how the presence of competing risks can be a challenge for these classical methods, and describe two analogous survival methods which can account for competing risks. These competing risks methods are generalised to multi-state survival models through the theory of counting processes.

Following this is an introduction to key concepts in causal inference and a description of study designs used for observational data collection, with examples of common biases seen in these data.

2.1 Survival analysis

Survival analysis involves the analysis of time-to-event data, also termed 'survival data', where the event could be any outcome of interest, and the time refers to the duration from a given origin to the occurrence of an event (the endpoint) (Collett 2023).

This section introduces survival data in the context of infectious disease studies and outlines methods to account for data censoring and truncation.

2.1.1 Overview of survival data

In studies of infectious disease, survival data are often used to describe clinical origins and endpoints, for example: from study recruitment to the time of infection; from hospitalisation to discharge; or from infection to death. These data may be combined with information on patient characteristics, such as age, gender, or socioeconomic status, and clinical variables, such as uptake of vaccination, or initiation of treatment (Collett 2023).

2.1.2 Censoring and truncation

Two common features of survival data are censoring and truncation. These arise due to incomplete or non-observation of time-to-event information and may take a number of forms.

Censoring

Censoring in survival data occurs when information about an individual is only known within certain intervals, or 'censoring times', (C_l, C_r) (Klein and Moeschberger 2005). Various categories of censoring exist, which lead to different specifications of the likelihood function (discussed further in Section 2.2.1).

Right-censoring is when we lack information to the 'right' (or future) of the right-censoring time C_r. For example, if study follow-up finishes before the event of interest, X, has taken place for an individual, this individual is right-censored at time $C_r < X$ (Klein and Moeschberger 2005).

Left-censoring, on the other hand, is when the event of interest, X, has occurred at an unknown point before the left-censoring time, C_l, e.g. if the event of interest occurs before the person is first observed in the study, this individual is left-censored at time $C_l > X$ (Collett 2023).

Censoring may also be intentionally imposed on a dataset, termed **administrative censoring**. Administrative censoring may be used to limit the time-to-event data for each individual to a pre-specified cut-off, with any events beyond this period not considered, e.g. mortality within 30 days of hospital admission (Collett 2023). Examples of right-censoring, left-censoring, and administrative censoring in a patient-level time-to-event dataset are shown in Figure 2.1.

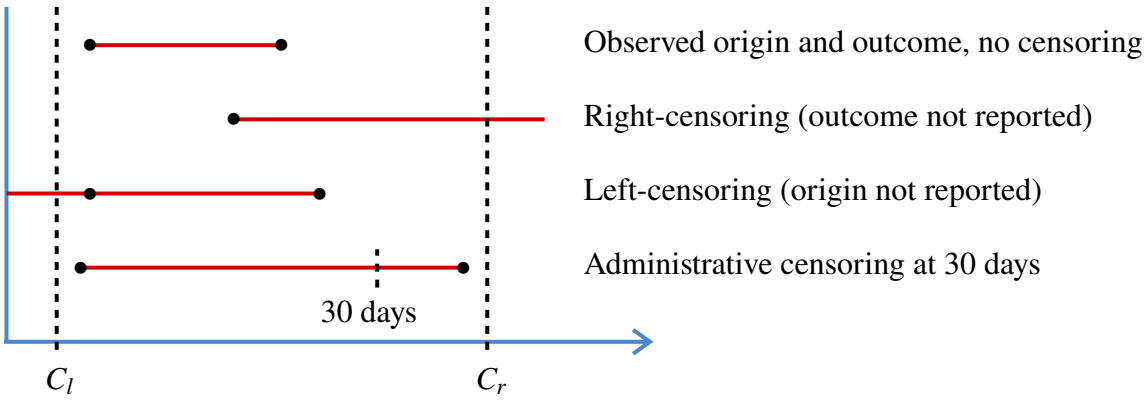


Figure 2.1: Examples of right, left, and administrative censoring for patients in time-to-event data with origin and outcome information reported. C_l: left-censoring time, C_r: right-censoring time, •: reported origin, outcome, or intermediate event.

A more general form of censoring is **interval-censoring**, which occurs when the event of interest, X, takes place in-between measurements, i.e. within a censoring interval $(L, R]$, $L < X < R$. An example is the infection time for an individual — typically not directly observed, but known to occur in-between a negative and positive test. Interval censoring is a feature of 'intermittently-observed' data, where individuals are tested for the presence of infection at several time-points, and can be used to detect changes in an individual's infection status when

testing is sufficiently frequent (van den Hout 2016; Klein and Moeschberger 2005), as shown in Figure 2.2.

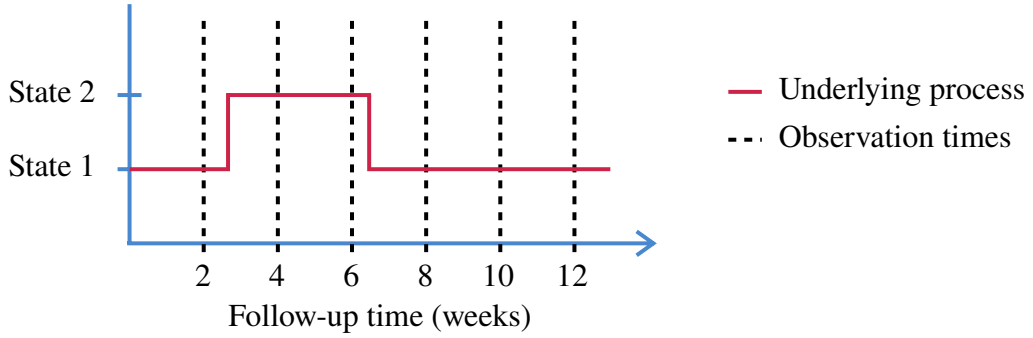


Figure 2.2: Example of interval censoring with intermittently-observed data for an underlying process.

Truncation

A second common feature of survival data is truncation. Truncation is related to censoring, although whereas with censored data there is at least partial information about an individual, with truncated data information outside of the truncation interval (Y_l, Y_r) is not observed at all (Klein and Moeschberger 2005). As with censoring, both left, right, and interval-truncation are possible.

Left-truncation occurs when only event-times which take place after the left-truncation time, Y_l, are available, e.g. if individuals whose event of interest takes place prior to the study, $X < Y_l$, are not recorded, they are left-truncated.

Right-truncation occurs when only event-times which take place before the right-truncation time, Y_r, are available, e.g. if records are only submitted once an outcome has taken place, with no knowledge of individuals at risk but whose event of interest occurs after the study endpoint, $X > Y_r$, these individuals are right-truncated (Klein and Moeschberger 2005). Examples of left and right-truncated patient-level time-to-event data are shown in Figure 2.3.

2.1.3 Cumulative incidence, survivor, and hazard functions

Because censored and truncated survival data are not fully observed, estimating a distribution for the time-to-event presents a challenge for standard statistical methods. For valid statistical estimation of these data, a specific class of survival methods have been developed which can account for different types of censoring and truncation.

These methods are based on the concept of the *survivor function*, which describes the probability of surviving beyond a given time, and the *hazard function*, which describes the instantaneous rate of failure at time t, given survival up until time t (Collett 2023).

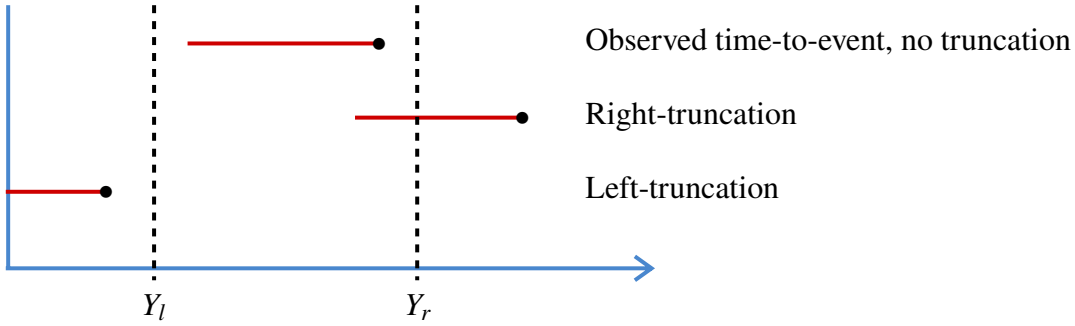


Figure 2.3: Examples of left and right-truncation for patients in time-to-event data, assuming only information within the truncation interval (Y_l, Y_r) is observed. Y_l: left-truncation time, Y_r: right-truncation time, •: outcome time.

Cumulative incidence function

Let T be the independent and identically distributed (i.i.d) random variable representing the survival time, $T = t > 0$, for an individual, and assume this random variable has a probability distribution with probability density function $f(t)$.

The distribution function of T, also known as the cumulative incidence function, is the probability of 'failure' before time t, defined as (Collett 2023):

$$F(t) = \Pr(T < t) = \int_0^t f(u)\,\mathrm{d}u$$

Survivor function

The survivor function is the probability of surviving past time t, defined as:

$$S(t) = \Pr(T \geq t) = 1 - F(t)$$

Hazard function

The hazard function is the instantaneous rate of failure at time t, given survival up until time t, defined as:

$$h(t) = \lim_{\delta t \downarrow 0} \left\{ \frac{\Pr(t \leq T < t + \delta t \mid T \geq t)}{\delta t} \right\} \tag{2.1}$$

By conditional probability, $\Pr(A \mid B) = \Pr(AB)/\Pr(B)$, so the hazard function can also be expressed as:

$$\begin{aligned} h(t) &= \lim_{\delta t \downarrow 0} \left\{ \frac{\Pr(t \leq T < t + \delta t)}{\delta t \Pr(T \geq t)} \right\} \\ &= \lim_{\delta t \downarrow 0} \left\{ \frac{F(t+\delta t) - F(t)}{\delta t S(t)} \right\} \\ &= \lim_{\delta t \downarrow 0} \left\{ \frac{F(t+\delta t) - F(t)}{\delta t} \right\} \frac{1}{S(t)} \end{aligned}$$

This limit is the definition of the derivative of $F(t)$ with respect to t, and therefore equal to $f(t)$ (Collett 2023):

$$\lim_{\delta t \downarrow 0} \left\{ \frac{F(t+\delta t) - F(t)}{\delta t} \right\} = \frac{\mathrm{d}}{\mathrm{d}t} F(t) = f(t)$$

hence the hazard function can be re-written:

$$h(t) = \frac{f(t)}{S(t)} \tag{2.2}$$

2.2 Statistical inference in survival analysis

Statistical inference is the process of using data to infer the properties of an underlying probability distribution. In this section I describe the 'likelihood function', used to connect statistical models to observed data in the frequentist and Bayesian frameworks. Two classical methods of statistical inference for survival analysis data are shown: the Kaplan-Meier estimator and Cox proportional hazards model. These methods are well-suited to studies of simple disease processes, for example when we are only interested in the effect of an intervention on a final outcome, irrespective of any intermediate events (Austin 2017).

2.2.1 Likelihood estimation

Frequentist and Bayesian likelihood

The likelihood function, L, is used to link statistical models to observed data. The likelihood is the probability of observing the data, D, given the parameters, θ, of the model and formally expressed as:

$$L(\theta \mid D) = f(D \mid \theta) = \prod_i f(X_i \mid \theta)$$

where f is the probability density function, and X_i are assumed to be i.i.d observations in the data (Wasserman 2004).

In the 'frequentist' framework the parameters of a statistical model are considered to have a fixed 'true' unknown value to be estimated, and the data are realisations of random variables. The parameters of the model are estimated by maximising the likelihood to obtain the *maximum likelihood estimate* (MLE), denoted $\hat{\theta}$ (Wasserman 2004).

Bayesian methods are an alternative to maximum likelihood estimation, used to estimate the probability distribution of the parameters of the model given the data, known as the 'posterior distribution'. In the Bayesian framework, all quantities (both observable quantities and parameters) are considered as random variables, with observed data being realisations of these random variables. The posterior distribution, $f(\theta \mid D)$, is related to the likelihood function, $f(D \mid \theta)$, via Bayes formula (Wasserman 2004):

$$f(\theta \mid D) = \frac{f(D \mid \theta) f(\theta)}{f(D)}$$

Likelihoods for censored and truncated data

Constructing likelihoods for censored or truncated data requires careful consideration of the information provided by each observation. The likelihood components provided by censored observations are described by Klein and Moeschberger 2005:

- for exactly observed data: information on the probability of the event occurring at this time, i.e. $f(X)$;
- for right-censored observations: information on the survivor function up to a certain time, i.e. $S(C_r)$;
- for left-censored observations: information on the cumulative incidence up to a certain time, i.e. $F(C_l)$;
- for interval-censored observations: information on the probability that the event is within this interval, i.e. $S(L) - S(R)$.

As data may be observed with different levels of censoring or truncation, the likelihood function is constructed from the joint product of each component for each observation i:

$$L \propto \prod_i f(X_i) \prod_i S(C_{r,i}) \prod_i F(C_{l,i}) \prod_i (S(L_i) - S(R_i))$$

The equivalent likelihood components for truncated data provide information on conditional probabilities of events, i.e. $f(X)/F(Y_r)$ for right-truncated observations, where observation is conditional on the event of interest having occurred by truncation time Y_r (Klein and Moeschberger 2005).

2.2.2 Parametric and non-parametric models

The statistical models used for estimating model parameters may be parametric or non-parametric. In general, parametric models rely on assumptions about the distribution of the underlying survival times, so careful selection of this distribution is needed for valid inference.

In contrast, non-parametric models require fewer assumptions about the survival distribution, so can be applied when the functional form of the model is not known, but may require more data and not allow for extrapolation beyond the time period of the observed data. Comparisons to non-parametric estimates can be made to assess the fit of a parametric model (Aalen, Borgan *et al.* 2008; Jackson *et al.* 2022).

2.2.3 Kaplan-Meier estimator

The Kaplan-Meier estimator is a widely-used, non-parametric method to estimate the survivor function from censored data, and defined as:

$$\hat{S}(t) = \prod_{t_k \leq t} \left(\frac{n_k - d_k}{n_k} \right) \tag{2.3}$$

where d_k is the number of events, and n_k is the number of individuals at risk at time $t_k \leq t$ (Collett 2023).

2.2.4 Cox proportional hazards model

The Cox proportional hazards model is used to relate survival times to explanatory variables via regression. The model is 'semi-parametric', as the baseline hazard is estimated non-parametrically while the effects of covariates are estimated parametrically (Cox 1972). The Cox proportional hazards model assumes that the hazard of death at any given time for an individual in one group is *proportional* to the hazard of death at the same time for an individual in a different group. This is expressed by writing the hazard function conditional on a given set of m constant or time-varying covariates $\boldsymbol{z} = \{z_1 \ldots, z_m\}$ as:

$$h(t \mid \boldsymbol{z}) = h^{(0)}(t) \exp\left(\sum_{m=1}^{M} \beta_m z_m \right) \equiv h^{(0)}(t) \exp(\boldsymbol{\beta}^{\mathsf{T}} \boldsymbol{z})$$

where $h^{(0)}(t)$ is the baseline hazard function, and β_m are the regression coefficients. Notably that the 'hazard ratio' for the mth covariate is a constant (Klein and Moeschberger 2005):

$$\frac{h(t \mid z_m = 1)}{h(t \mid z_m = 0)} = \frac{h^{(0)}(t) \exp(\beta_m 1)}{h^{(0)}(t) \exp(\beta_m 0)} = \exp(\beta_m)$$

The regression coefficients β_m are estimated by maximising the partial likelihood $L(\boldsymbol{\beta})$:

$$L(\boldsymbol{\beta}) = \prod_{i \in D} \frac{\exp(\boldsymbol{\beta}^\mathsf{T} \boldsymbol{z}_{d_i})}{\sum_{j \in R_i} \exp(\boldsymbol{\beta}^\mathsf{T} \boldsymbol{z}_j)}$$

where $D = \{T_1, T_2, \ldots, T_n\}$ are the set of distinct failure times, R_i is the set of all individuals who are at risk of failure immediately before time T_i, $\boldsymbol{z}_{d_i}$ is the covariate vector for an individual who failed at time T_i, and $\boldsymbol{z}_j$ is the covariate vector for the jth individual at risk at time T_i (Klein and Moeschberger 2005).

Further details of the Cox proportional hazards model, specifically of how to account for non-proportional hazards and correlation in survival times, are included in Appendix B.

2.3 Competing risks survival models

In medical cohort studies, multiple end-points (e.g. death, intensive care admission, and discharge) may be considered. These multiple end-points are known as *competing risks* and may hinder the event of interest, or modify the chance that this event occurs (Collett 2023). Kaplan-Meier and Cox proportional hazards models treat competing risks as censored observations, and do not account for dependencies between events. Hence the assumption of independent times-to-events in these conventional survival analyses is violated in the presence of competing risks (Andersen, Geskus *et al.* 2012; Putter *et al.* 2007).

To correctly address competing risks, a competing risks survival model may be used instead. In these models an individual is observed over time, with several possible events 'competing' until one takes place and the individual transitions to the corresponding state (Lipsitch, Donnelly *et al.* 2015). Importantly there is no assumption of independence for the distribution of the time to competing events and censoring can still be appropriately accounted for (Prentice *et al.* 1978; Putter *et al.* 2007).

In this section I define the cause-specific hazard and cause-specific cumulative incidence functions, and describe two competing risks methods, analogous to the Kaplan-Meier and Cox proportional hazards model, which are used to estimate these functions from observed data.

2.3.1 Cause-specific hazard and cumulative incidence functions

Cause-specific hazard function

Let T be a random variable for the survival time and C be a random variable for the cause of failure. The *cause-specific* hazard function for the jth cause, $j = \{1, 2, \ldots, m\}$ is defined by:

$$h_j(t) = \lim_{\delta t \downarrow 0} \frac{\Pr(t \leq T < t + \delta t, C = j \mid T \geq t)}{\delta t}$$

Defining $f_j(t)$ as the *cause-specific* density function, and $S(t) = \Pr(T \geq t)$ as the overall survivor function, the relationship given in Equation 2.2 holds in the presence of competing risks (Collett 2023):

$$h_j(t) = \frac{f_j(t)}{S(t)} \tag{2.4}$$

Cause-specific cumulative incidence function

The *cause-specific* cumulative incidence function, i.e. the probability of surviving until time t and failure from cause j, in the presence of all other risks, is given by:

$$F_j(t) = \Pr(T < t, C = j)$$

for $j = \{1, 2, \ldots, m\}$, with $\Pr(C = j)$ often written as π_j. An expression for $F_j(t)$ in terms of the cause-specific hazard function may be derived using Equation 2.4 (Collett 2023):

$$\begin{aligned} h_j(t) &= \frac{f_j(t)}{S(t)} \\ f_j(t) &= S(t) h_j(t) \\ F_j(t) &= \int_0^t S(u) h_j(u) \, \mathrm{d}u \end{aligned} \tag{2.5}$$

2.3.2 Aalen-Johansen estimator

The standard non-parametric estimator of the cause-specific incidence function is the Aalen-Johansen estimator, also described as the 'multi-state version' of the Kaplan-Meier estimator (Aalen and Johansen 1978).

Firstly, by the non-parametric Nelson-Aalen estimator of the cause-specific hazard function for cause j:

$$\hat{h}_j(t) = \frac{d_j(t)}{n(t\text{-})}$$

where $d_j(t)$ is the number of deaths due to cause j at time t, and $n(t\text{-})$ is the number of individuals at risk just prior to t (Borgan 2014).

With the Kaplan-Meier estimate of the survivor function defined in Equation 2.3, $\hat{S}(t)$, the Aalen-Johansen estimator for the cause-specific cumulative incidence function then follows from Equation 2.5:

$$\hat{F}_j(t) = \sum_{t_k<t} \hat{S}(t_{k-1}) \frac{d_j(t_k)}{n_k}$$

for all times $t_k < t$ where transition events are observed to occur (Kalbfleisch and Prentice 2011; Lambert 2017).

2.3.3 Fine-Gray proportional hazards model

Analogous to the Cox proportional hazards model, the Fine-Gray proportional hazards model may be used estimate the hazard of a competing event (termed the sub-distribution hazard) among those yet to experience an event by time t (Fine and Gray 1999). The risk set for the sub-distribution hazard consists of both those who have yet to experience any event, and those who have yet to experience the event of interest, but who have experienced a competing event.

The sub-distribution hazard is defined as the instantaneous risk of experiencing a competing event j given that the individual has not already experienced event j (Austin *et al.* 2016):

$$\lambda_j(t) = \lim_{\delta t \downarrow 0} \left\{ \frac{\Pr\left(t \leq T < t+\delta t, C = j \mid (T > t) \cup (T \leq t \cap C \neq j)\right)}{\delta t} \right\}$$

where, as before, C is the random variable for the event that occurs. Fine-Gray regression links the sub-distribution hazard, λ_j, to the cause-specific cumulative incidence function, F_j, through the relationship (Putter *et al.* 2007):

$$\lambda_j(t) = -\frac{\mathrm{d}}{\mathrm{d}t} \log\left(1 - F_j(t)\right)$$

As with the Cox proportional hazards model, a proportional hazards regression model is assumed, where the hazard of cause j at time t for individual i is:

$$\lambda_{i,j}(t) = \exp(\boldsymbol{\beta}^\mathsf{T} \boldsymbol{z}) \lambda_j^{(0)}(t)$$

with $\lambda_j^{(0)}(t)$ being the baseline sub-distribution hazard for cause j. Covariate coefficients are estimated by maximising the weighted partial likelihood $L(\boldsymbol{\beta})$:

$$L(\boldsymbol{\beta}) = \prod_{i \in D} \frac{\exp(\boldsymbol{\beta}^\mathsf{T} \boldsymbol{z}_{d_i})}{\sum_{j \in R_i} w_{i,j} \exp(\boldsymbol{\beta}^\mathsf{T} \boldsymbol{z}_j)}$$

where $w_{i,j}$ are weights which account for the increasing probability of censoring with increasing follow-up time, $D = \{T_1, T_2, \ldots, T_n\}$ are the set of distinct failure times, R_i is the set

of all individuals who are at risk of failure immediately before time T_i, $\boldsymbol{z}_{d_i}$ is the covariate vector for an individual who failed at time T_i, and $\boldsymbol{z}_j$ is the covariate vector for the jth individual at risk at time T_i (Lambert 2017). Covariate effects on the sub-distribution hazard may be interpreted as covariate effects on the cumulative incidence of a competing event.

2.3.4 Stratification

Whilst proportional hazards models are a common method for incorporating covariates in survival analyses, a more straightforward approach is through stratification. In a stratified model the population is subdivided according to covariate group (or strata), the survival is compared within each stratum, and the differences within stratum are combined to give an overall comparison (Bradburn *et al.* 2003).

As stratification allows the baseline hazard to vary across strata it is sometimes used to accommodate non-proportional hazards in Cox and Fine-Gray proportional hazards models (see Appendix B for further details) (Hosmer and Lemeshow 1999).

2.4 Multi-state survival models

When disease processes are more complex, or when intermediate events may influence the final outcome, the standard survival methods described so far may be insufficient to explore the effects of treatment on these outcomes. In these cases, multi-state models, which are more flexible and hence better able to investigate the different pathways that patients may experience, can be applied (Le-Rademacher *et al.* 2018). These models allow for joint estimation of features of the underlying process and the associated hazards of transition for a set of given covariates, and implicitly account for competing risks (Andersen and Keiding 2002; van den Hout 2016; Jackson 2011).

This section provides an introduction to the theory of multi-state processes and describes how these relate to continuous-time and discrete-time multi-state models.

2.4.1 Stochastic processes underlying multi-state models

Multi-state models combine statistical inference with the theory of stochastic processes (van den Hout 2016).

Discrete and continuous parameter processes

A stochastic process is formally defined as a collection of random variables $\{X(t),\ t \in T\}$, indexed by a parameter t which varies in a mathematical index set T. The variable $X(t)$ is the

state of the process at time t, and the set $\mathcal{S}$ of all possible values of $X(t)$ is termed the *state space*.

Two important cases of stochastic processes are discrete parameter processes, when $T = \{\pm 1, \pm 2, \pm 3, \dots\}$, and continuous parameter processes, when $T = \{t : -\infty < t < \infty\}$. Both may be restricted to the positive domain, in which case the parameter t may be used to represent time, and these processes are termed 'discrete-time' and 'continuous-time' processes (Parzen 1999).

Counting processes

In survival analyses, the focus is typically on observing events as they occur over time, forming a class of stochastic process known as a *point process*. If the number of events that occur over time are counted, then the stochastic process is instead known as a *counting process*. Counting processes form the basis for multi-state models, and linkage to the theory of martingales provides a theoretical framework for these models (Aalen, Borgan *et al.* 2008; Andersen, Borgan *et al.* 1996).

Markov process

A stochastic process is called a 'Markov process' if the future state of a process depends only on the present state, and not on the sequence of events that preceded it. Equivalently, for a Markov process the conditional probability of a future state $X(t_{i+j})$ given $X(t_i)$ is independent of previous states $X(t_0), X(t_1), \dots, X(t_{i-1})$ (Chiang 1980).

Time-homogeneous Markov processes

Markov processes are said to be 'time-homogeneous' if the probability of transition from one state to the next (termed the *transition probability*) is independent of the time parameter t (Chiang 1980). Equivalently, for $\Pr(X(t+u) = s \mid X(t) = r)$ denoting the conditional probability of a process $X(\cdot)$ being in state s at time $t+u$, given the process was in state r at time t, a Markov process is time-homogeneous if:

$$\Pr(X(t+u) = s \mid X(t) = r) = \Pr(X(u) = s \mid X(0) = r)$$

A well-known time-homogeneous Markov counting process is the Poisson process, where events occur randomly and independently of each other. A homogeneous Poisson process is described by its intensity λ, which is the probability of an event occurring in a given time interval divided by the length of the time interval (Aalen, Borgan *et al.* 2008).

Multi-state processes

Multi-state processes are discrete-state stochastic processes with a finite state space $\mathcal{S} = \{1, 2, \ldots, m\}$ (Andersen and Keiding 2002). In these processes, an individual begins in one state and spends a random, continuously-distributed time in that state before moving to a random next state. Multi-state processes are defined by the initial distribution of the states $\Pr(X(0) = r)$, and the transition probabilities $\Pr(X(t+u) = s \mid X(t) = r, \mathcal{X}_t)$ from state r to state s during the interval $(t, t+u]$, with $r, s \in \mathcal{S}$, and where $\mathcal{X}_t$ represents the history of the process $X(\cdot)$ up to time t (Aalen, Borgan *et al.* 2008).

Figure 2.4 shows how a counting process may be built on top of a multi-state process: panel A shows a multi-state process with transitions between two states, and the corresponding counting process for the transition event from State 1 → State 2 is shown in panel B.

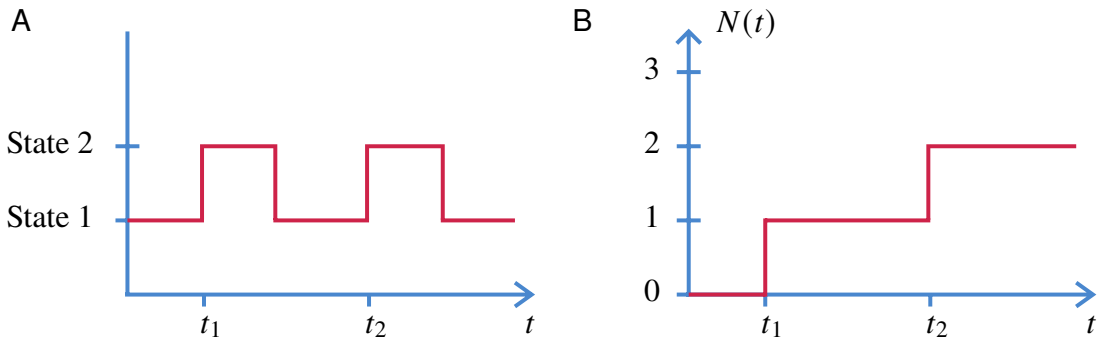


Figure 2.4: Example of transitions between two states in a multi-state process (panel A), and the corresponding counting process for the State 1 → State 2 transition (panel B).

2.4.2 Multi-state models

In a multi-state model, each state represents a different stage or condition that an individual may experience. The transitions between these states are modelled as a multi-state process, with the time spent in each state and the transitions between states being random variables that can be estimated from data (Putter *et al.* 2007).

Multi-state models are particularly useful in the context of disease progression, as each state in the model can be used to represent a different level of disease severity, with transitions between states representing progression. This may allow for a more nuanced understanding of disease processes, capturing the different pathways that patients may experience (Matsena Zingoni *et al.* 2021).

2.4.3 Continuous-time multi-state models

Continuous-time multi-state models are defined by transition intensities, $q_{r,s}(t, \boldsymbol{z}(t))$, which represent the instantaneous hazard of moving from state r to state s at time t, dependent on a set of explanatory, potentially time-varying, covariates $\boldsymbol{z}(t)$ (Putter *et al.* 2007). Defining an individual's state at time t as $X(t)$ then:

$$q_{r,s}(t, \boldsymbol{z}(t)) = \lim_{\delta t \downarrow 0} \frac{\Pr(X(t+\delta t) = s \mid X(t) = r)}{\delta t} \tag{2.6}$$

From Equation 2.1, this transition intensity is equivalent to the hazard of transition from one state to another. These transition intensities form a *transition intensity matrix* $\mathbf{Q}$, whose rows sum to zero, with diagonal entries defined by:

$$q_{r,r}(t, \boldsymbol{z}(t)) = -\sum_{s \neq r} q_{r,s}(t, \boldsymbol{z}(t))$$

For the example three-state model shown in Figure 1.3 this matrix would be:

$$\mathbf{Q} = \begin{bmatrix} -q_{1,2} - q_{1,3} & q_{1,2} & q_{1,3} \\ q_{2,1} & -q_{2,1} - q_{2,3} & q_{2,3} \\ 0 & 0 & 0 \end{bmatrix}$$

Here, state 3 is known as an 'absorbing' state, as individuals do not leave the state after arrival, while states 1 and 2 are 'transient' states. The *transition probability matrix*, which contains the probabilities of moving between states within a time-interval of length t is derived from the transition intensity matrix by the matrix exponential (van den Hout 2016):

$$\mathbf{P}(t) = \exp(t\mathbf{Q}) = \sum_{n=0}^{\infty} \frac{t^n}{n!} \mathbf{Q}^n$$

Covariates

The (potentially time-varying) effect of a given set of m covariates, $\boldsymbol{z}(t) = \{z_1(t), \ldots, z_m(t)\}$, on the transition intensities $q_{r,s}(t, \boldsymbol{z}(t))$ can be expressed via a proportional hazards regression model. Typically these combine a parametric baseline hazard, $q^{(0)}_{r,s}(t)$, with log-linear regression:

$$\begin{aligned} q_{r,s}(t, \boldsymbol{z}(t)) &= q^{(0)}_{r,s}(t) \exp\left(\sum_{m=1}^{M} \beta_{r,s,m} z_m(t)\right) \\ &= q^{(0)}_{r,s}(t) \exp\left(\boldsymbol{\beta}^{\mathsf{T}}_{r,s} \boldsymbol{z}(t)\right) \end{aligned}$$

where $\exp(\beta_{r,s,m})$ is the hazard ratio for the mth covariate on the $r \rightarrow s$ transition (van den Hout 2016).

Sojourn time

The time spent in a state prior to transition is termed the *sojourn time*, with its mean given by the inverse of the transition intensity for remaining in the state (Jackson 2011):

$$E(T_r) = -\frac{1}{q_{r,r}}$$

2.4.4 Discrete-time multi-state models

For population studies which measure duration in discrete time, discrete-time multi-state models may applied (Steele *et al.* 2004). Since the concept of instantaneous risk does not apply in discrete time, transition probabilities are instead defined which represent the probability of transition from state r to state s in the time interval $(t_i, t_j]$, $t_j > t_i$:

$$p_{r,s}(t_i, t_j) = \Pr(X(t_j) = s \mid X(t_i) = r)$$

2.4.5 Semi-Markov and hidden Markov models

Multi-state models are typically specified to be Markov, i.e. $q_{r,s}(t; \mathcal{X}_t) = q_{r,s}(t)$, where $\mathcal{X}_t$ is the observation history of the process up to time t. Semi-Markov models relax the Markov assumption, allowing dependence on the time of entry into the current state, i.e. $q_{r,s}(t; \mathcal{X}_t) = q_{r,s}(t - t_r)$, where t_r is the time of entry into the current state r (Meira-Machado *et al.* 2009).

Hidden Markov models may also be specified to allow for unobserved, or 'latent', states to be included in the model. These can be particularly useful for modelling complex disease processes, where progression of the disease is influenced by unobserved factors, and the time to event depends on the current state of the disease, for instance in models of HIV progression (Aalen, Farewell *et al.* 1997).

2.5 Statistical inference for multi-state models

The non-parametric Aalen-Johansen estimator, parametric multi-state mixture models, and parametric multi-state cause-specific hazards models were demonstrated to provide comparable estimates of competing risks when applied to real-world hospital admissions data (Jackson *et al.* 2022). Mixture models have the advantage that easily interpretable quantities, such as event probabilities and times to events, are directly estimated (Andersen, Abildstrom *et al.* 2002).

Whereas, for cause-specific hazards models, intensive simulation is required to derive these quantities from parameter estimates that are not directly interpretable (Larson and Dinse 1985).

This section describes statistical inference for non-parametric and parametric multi-state survival models. I discuss how these multi-state methods are well-suited to the investigation of intermittently-observed data and explain how these models have been implemented during my PhD.

2.5.1 Aalen-Johansen estimator

The non-parametric Aalen-Johansen estimator can be used to estimate the transition probability matrix, $\mathbf{P}(t)$ for a Markov process with a finite number of states (Andersen, Borgan *et al.* 1996). Firstly, let $\mathbf{A}(t)$ be the matrix of cumulative transition intensities from state r to state s:

$$\mathbf{A}(t) = A_{r,s}(t) = \begin{cases} \int_0^t q_{r,s}(u)\,\mathrm{d}u & \text{if } r \neq s \\ -\sum_{r \neq s} A_{r,s}(t) & \text{if } r = s \end{cases}$$

The transition probability matrix for this process can be represented as:

$$\mathbf{P} = \prod_{k=1}^{K} (\mathbb{I} + \Delta\mathbf{A}(t_k))$$

where $\mathbb{I}$ is the identity matrix, $\Delta\mathbf{A}(t_k) = \mathbf{A}(t_k) - \mathbf{A}(t_k\text{-})$, and t_k- represents the instant before time t_k (Andersen, Borgan *et al.* 1996). Substituting the Nelson-Aalen estimator for $\mathbf{A}(t)$, then the Aalen-Johansen estimate of the transition probability matrix is:

$$\hat{\mathbf{P}}(t) = \hat{P}_{r,s}(t) = \prod_{k=1}^{K} \left(\mathbb{I}_{r,s} + \Delta\hat{A}_{r,s}(t_k) \right)$$

with:

$$\Delta\hat{A}_{r,s}(t_k) = \begin{cases} \dfrac{\Delta N_{r,s}(t_k)}{Y_r(t_k\text{-})} & \text{if } r \neq s \\ \dfrac{-\Delta N_s(t_k)}{Y_s(t_k\text{-})} & \text{if } r = s \end{cases}$$

where $\Delta N_{r,s}(t_k)$ is the number of transitions from state r to state s at time t_k, $\Delta N_s(t_k)$ is the number of transitions away from state s at time t_k, and $Y_r(t_k\text{-})$ is the number of individuals in state r just before time t_k (Borgan 2014).

2.5.2 Multi-state mixture models

For a general multi-state competing risks mixture model, assume that an individual i who begins in state r makes a transition to a (pre-assigned) destination state s. Let $I_{i,r}$ be the indicator variable which determines which transition will occur, the transition intensity at time t is:

$$q_{i,r,s}(t) = \begin{cases} q^*_{i,r,s}(t) & \text{if } I_{i,r} = s \\ 0 & \text{otherwise} \end{cases}$$

where $q^*_{i,r,s}(t)$ is the transition intensity for the transition which occurs. The probability of transition to state s is defined as:

$$\pi_{r,s} = \Pr(I_{i,r} = s)$$

Let $T_{r,s}$ be the the time from entering state r to moving to state s, given that this transition occurs (i.e. the conditional sojourn time). A parametric distribution with parameters $\theta_{r,s}$ and conditional density $f_{r,s}(. \mid \theta_{r,s})$ may be used to model $T_{r,s}$ (Ghani *et al.* 2005).

The probabilities of competing events $\pi_{r,s}$ can be related to a set of m covariates $\boldsymbol{z} = \{z_1, \ldots, z_m\}$ via multinomial logistic regression. Define as S_r the set of all competing states after r, then the logit of transition to a given state $j \in S_r$ vs. transition to a baseline state $0 \in S_r$ is:

$$\ln\left(\frac{\pi_{r,j}(\boldsymbol{z})}{\pi_{r,0}(\boldsymbol{z})}\right) = \alpha_{r,j} + \boldsymbol{\beta}^{\mathsf{T}}_{r,j}\boldsymbol{z}$$

with $\alpha_{r,j}$ and $\boldsymbol{\beta}_{r,j} = \{\beta_{r,j,k};\ k = 1, \ldots m\}$ being regression coefficients for the probability of transition to state j (Fagerland *et al.* 2008).

Meanwhile, parameters $\theta_{r,s}$ of the time to transition distribution can be related to a different (or identical) set of l covariates $\boldsymbol{z} = \{z_1, \ldots, z_l\}$ via a log-linear model with coefficients $\gamma_{r,s}$ and $\boldsymbol{\omega}_{r,s} = \{\omega_{r,s,k};\ k = 1, \ldots, l\}$:

$$\log(\theta_{r,s}(\boldsymbol{z})) = \gamma_{r,s} + \boldsymbol{\omega}^{\mathsf{T}}_{r,s}\boldsymbol{z}$$

For most parametric time to transition distributions 'accelerated failure time' models are considered, where the covariates affect the rate of progression of time, i.e. the survival function $S(t \mid \boldsymbol{z})$ for an individual at time t with covariates $\boldsymbol{z}$ is related to the baseline survival $S^{(0)}$ via:

$$S(t \mid \boldsymbol{z}) = S^{(0)}\left(\exp(\boldsymbol{\theta}^{\mathsf{T}}\boldsymbol{z})t\right)$$

where $\boldsymbol{\theta}$ is a vector of regression coefficients (Klein and Moeschberger 2005).

2.5.3 Accounting for censoring

As described in Section 2.2.1, survival models may be specified to account for censoring and truncation in observed data by careful consideration and construction of the likelihood function. Multi-state survival models, in particular, are well-suited to investigating interval-censored, or *intermittently-observed* data (van den Hout 2016).

In Figure 2.2, changes in an underlying process were detected with an intermittent observation scheme. Depending on the nature of the underlying process, less frequent and/or irregular observation of the underlying process may result in transitions being undetected, as shown in Figure 2.5.

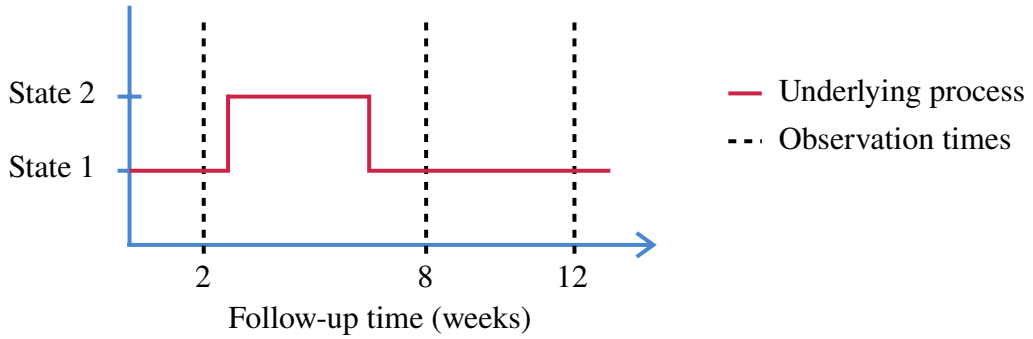


Figure 2.5: Example of intermittently-observed data with missed observations for an underlying process.

An observation scheme is said to be *non-informative* if the likelihood is proportional to a scenario where observation times were fixed in advance and chosen independently of the underlying process (Grüger *et al.* 1991). With a non-informative observation scheme then multi-state models can correctly account for intermittently-observed data (van den Hout 2016).

2.5.4 Implementing multi-state models

A number of sophisticated statistical software packages exist to implement frequentist and Bayesian survival and multi-state models using the R programming language (R Core Team 2020). Frequentist maximum likelihood estimation tends to be less computationally intensive than Bayesian inference, and the latter may require integrating over a vast number of parameters, which is usually impracticable for higher dimensions. Instead, algorithms which simulate from the posterior distribution are often used to evaluate features of these integrals. Markov Chain Monte Carlo (MCMC) algorithms have been widely used for this application (Young and Smith 2005).

The specialised modelling packages used during my PhD have included: `survival`, `msm` and `flexsurv` which fit models using maximum likelihood (Jackson 2011; Jackson 2016;

Therneau 1999), and `JAGS` and `STAN` (Carpenter *et al.* 2017; Plummer *et al.* 2003) which fit Bayesian models using MCMC.

In the spirit of open research, a repository containing code samples for implementing the multi-state methods used during my PhD, and the `R` code used to generate the figures and tables for this thesis is available at: https://github.com/pkirwan/PhD-thesis.

2.6 Key concepts in causal inference

In this final section I introduce the concept of causal inference from observational data and provide an outline of three common methods to compute the average causal effect.

Ideal randomised experiments, where the treated and untreated are exchangeable (or 'exogenous'), provide a rigorous way to assess causal effects in a population, as any association between between the treatment and the outcome can be regarded as causation. Whilst this is generally not true for observational data, if an explicit causal question is defined (e.g. the effect of a well-specified hypothetical intervention), and certain conditions fulfilled, the same assessment of causation can be made from observational data (Hernán and Robins 2006).

2.6.1 Causal diagrams and local independence

Causal diagrams are a graphical tool used to visualise causal relationships (Aalen, Røysland *et al.* 2016). One class of causal diagrams are causal directed acyclic graphs (causal DAGs). In these graphs, nodes represent measurements or interventions at discrete times, and edges between nodes indicate causal effects, an example is shown in Figure 2.6. For causal DAGs, the lack of an edge between nodes can be interpreted as the absence of a direct effect (Hernán and Robins 2023).

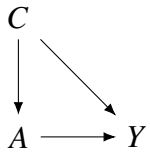


Figure 2.6: Example of a causal DAG representing discrete measurements for treatment, A, causes of treatment, C, and outcome, Y.

Whilst causal DAGs are useful to describe relationships between discrete-time processes, they are poorly suited to describing causality in continuous-time as changes in quantities or the occurrence of events are not represented (Aalen, Røysland *et al.* 2016). Instead, for continuous-time processes, the concept of *local independence* may be applied to consider how well a future value of an intensity process (such as a transition intensity in a multi-state model)

is predicted by the past (Didelez 2008). Local independence can be viewed as 'immediate causation', and can be represented in causal diagrams known as a local independence graph. In these graphs, nodes represent stochastic processes and edges between nodes represent immediate causal influence (Aalen, Røysland *et al.* 2016), as shown in Figure 2.7.

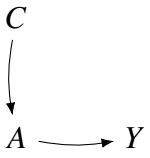


Figure 2.7: Example of a local independence graph representing continuous processes for treatment, A, causes of treatment, C, and outcome, Y. In local independence graphs only immediate causal influences are shown.

2.6.2 Causal inference from observational data

To make causal inference from observational data we rely on being able to analyse the data as if treatment had been randomly assigned, conditional on a set of measured covariates Z. This is termed a 'conditionally randomised' experiment (Hernán and Robins 2023).

Let $Y = (0 : \text{doesn't experience outcome}, 1 : \text{experiences outcome})$ be a random variable for an observed outcome, and $A = (0 : \text{untreated}, 1 : \text{treated})$ be a random variable for an observed treatment, then:

$$\Pr(Y = 1 \mid A = 1)$$

is the probability of experiencing the outcome given treatment. Define the potential outcome under treatment a as Y^a, then the potential outcome under the observed treatment:

$$Y^A = Y$$

The three 'identifiability' conditions required for an observational study to be treated as a conditionally randomised experiment are (Hernán and Robins 2023):

1. **Consistency**: the values of treatment under comparison correspond to well-defined interventions that, in turn, correspond to the versions of treatment in the data, i.e.

$$Y^a = Y \text{ for individuals with } A = a$$

2. **Exchangeability**: the conditional probability of receiving every value of treatment depends only on measured covariates Z. This is alternatively phrased as: conditional on a set of measured covariates Z, the untreated group, had they been treated, would

experience the same average outcome as the treated group, i.e.

$$\Pr(Y^a = 1 \mid A = 1, Z = z) = \Pr(Y^a = 1 \mid A = 0, Z = z)$$

or equivalently, $Y^a \perp\!\!\!\perp A \mid Z = z$, the counterfactual outcome and the observed treatment are independent within the covariate levels $Z = z$;

3. **Positivity**: the probability of receiving every value of treatment a conditional on $Z = z$ is positive.

$$\Pr(A = a \mid Z = z) > 0 \quad \text{for all } z \text{ where } \Pr(Z = z) \neq 0$$

Among these conditions, consistency and positivity are usually straightforward to check in observational studies whilst exchangeability relies on the assumption that all predictors of an outcome have been measured. It is usually impossible to guarantee exchangeability, but understanding the potential for biases such as confounding and collider bias (defined below) can inform which covariates to include such that the assumption will be approximately true (Hernán and Robins 2023).

If the three identifiability conditions are fulfilled, a (hypothetical) randomised experiment, known as a 'target trial', may be emulated using causal inference from observational data, with the average causal effect quantified as $E(Y^{a=1}) - E(Y^{a=0})$ (Hernán and Robins 2016).

2.6.3 Methods to compute the average causal effect

Three commonly-used methods to compute the average causal effect are inverse probability weighting, matching, and standardisation.

The concept of **inverse probability weighting** is to assign weights to each individual in the cohort based on the inverse of the probability of receiving the observed treatment level, and conditional on their covariates, Z. For example, a treated individual, $A = 1$, with a set of covariates $Z = z$ would be assigned a weight of $1/\Pr(A = 1 \mid Z = z)$. Conversely, an untreated individual, $A = 0$, with covariates $Z = z'$ would be assigned a weight of $1/\Pr(A = 0 \mid Z = z')$. These weights are used to adjust for potential baseline confounders when estimating the causal effect of treatment, allowing the creation of a 'pseudo-population' where treatment assignment is independent of the observed covariates (Hernán and Robins 2016).

Matching involves constructing a subset of the population in which the distribution of covariates, Z, is the same for the treated and untreated groups. This is typically achieved by pairing each treated individual with an untreated individual with the same or similar covariate values. The causal effect can be estimated as the average difference in outcomes among the matched pairs (Hernán and Robins 2023).

Standardisation involves estimating the counterfactual risk using a weighted average of the risks (or standardisation of the risks) in each covariate level. For example, for a covariate Z with two levels $(0, 1)$:

$$\Pr(Y^a = 1) = \Pr(Y = 1 \mid Z = 1, A = a)\Pr(Z = 1) + \Pr(Y = 1 \mid Z = 0, A = a)\Pr(Z = 0)$$

which generalises to:

$$\Pr(Y^a = 1) = \sum_z \Pr(Y = 1 \mid Z = z, A = a)\Pr(Z = z)$$

The standardised mean $E(Y^a) = \sum_z E(Y \mid Z = z, A = a)\Pr(Z = z)$ is known as the **g-formula** (Hernán and Robins 2016; Naimi *et al.* 2017).

2.7 Study designs and biases

In this final section of the chapter I describe a number of observational study designs and several potential sources of bias common to these studies.

Observational studies are common in epidemiology and a valuable resource for understanding infectious disease burden. Several different types of observational study design exist, which may be more or less suitable depending on the research question, the need to control for specific biases, and the feasibility of data collection (Baker 2023).

2.7.1 Observational study designs

Table 2.1 shows a 2×2 table which may be used to group participants in observational studies and to inform the calculation of standard statistical measures.

Table 2.1: A 2×2 table of outcomes and exposures.

		Outcome	
		Yes	No
Exposure	Yes	A	B
	No	C	D

Cross-sectional studies

Cross-sectional studies may be used to assess the prevalence of an outcome, defined as: $(A+C)/(A+B+C+D)$, and the prevalence of various different exposures, defined as: $(A+$

$B)/(A+B+C+D)$, within a population. They provide a snapshot of the population at a single point in time and can be relatively straightforward to conduct using surveillance data (Baker 2023).

Case-control studies

Case-control studies are used to investigate the effects of specific exposures on an outcome. In these studies the outcome of interest has already occurred, individuals who have experienced the outcome are termed 'cases', and are compared to individuals without the outcome, known as 'controls'. The odds of being exposed are compared in cases vs. controls via the odds ratio, defined as: AD/BC (Woodward 2013).

Test-negative case-control studies

Test-negative case-control studies are a sub-type of case-control study. These studies aim to minimise bias specifically as a result of test-seeking behaviour by selecting only individuals who test for a disease. As for case-control studies, the cases are individuals who test positive, controls are those who test negative, and again the odds of being exposed are compared between cases and controls.

Cohort studies

In cohort studies participants are selected based on their exposure status and followed up over time. The incidence of the outcome of interest is compared between individuals with and without the exposure using the relative risk (Woodward 2013):

$$\frac{A/(A+B)}{C/(C+D)}$$

For further discussion of test-negative case-control and cohort study designs, see Section 5.4.

2.7.2 Biases in observational studies

Bias can be defined as "any process at any stage of inference which tends to produce results or conclusions that differ systematically from the truth" (Sackett 1979). All of the observational study designs described can include potential sources of bias, which are often grouped into three broad types: information bias, collider bias, and confounding (Hernán and Robins 2023).

Information bias

Information, or measurement, bias arises due to inaccuracies in data and may be induced by errors or missing information in either the outcome or exposure variables, for example, incorrect coding of a diagnosis (Hernán and Robins 2023). Information bias can be mitigated through careful study design and data collection mechanisms, these could include: double entry of data, and linking information across different systems for validation (Page and Henderson 2008).

Collider bias

Collier bias, or selection bias, occurs when individuals are selected into an analysis by conditioning on a common 'collider' variable (a variable which is influenced by two other variables). Figure 2.8 shows an example of collider bias: people who are regular smokers and people at high risk of severe COVID-19 illness are both more likely to be admitted to hospital. By conditioning on hospitalisation a distorted association between smoking and COVID-19 mortality could be induced, and this bias may occur in either a positive or negative direction[1].

Collider bias is most easily avoided by not conditioning on the collider, but this may be impractical depending on the nature of data collection (Holmberg and Andersen 2022). In the context of COVID-19, Griffith *et al.* 2020 have suggested several methods for assessing the sensitivity of model results to collider bias, such as inverse probability weighting. Where information is available, the extent of the bias can also be examined by comparing the profile of selected individuals to the wider population of interest, e.g. whether hospitalised individuals tend to be older or more likely to have comorbidities.

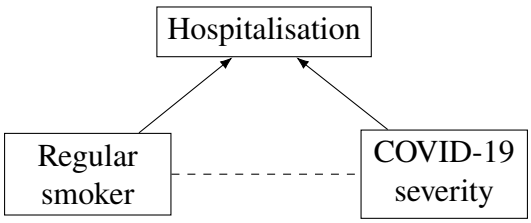


Figure 2.8: Causal diagram showing the effect of collider bias, where conditioning on hospitalisation induces a distorted association between smoking and COVID-19 severity. Conditioned on variables have bounding boxes, distorted associations are shown as dashed lines.

[1]A negative association between smoking and COVID-19 severity was reported by more than one early risk-factor study for COVID-19, with a review concluding this result may have been a result of collider bias (Wenzl 2020).

Confounding

Confounding is related to collider bias, but occurs by not conditioning on an explanatory variable. As a result confounding is sometimes termed 'latent variable' bias. Confounding introduces a 'backdoor path', an additional source of association between the explanatory variable and the outcome, so the association cannot be interpreted as a causal effect (Hernán and Robins 2023).

An example of confounding is shown in the causal diagram in Figure 2.9: during vaccine roll-out older people were more likely to receive a COVID-19 vaccine, but also more likely to experience severe COVID-19. If age group is not adjusted for then a distorted association between vaccination and COVID-19 severity may be induced. Confounding can be avoided by ensuring that factors that might be associated with the outcome are measured and adjusted for (Hernán and Robins 2023).

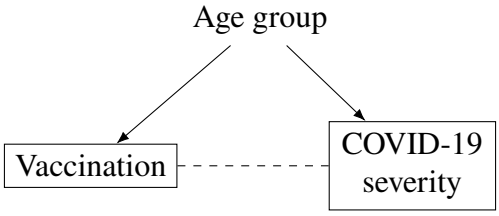


Figure 2.9: Causal diagram showing the effect of confounding, where not conditioning on age group induces a distorted association between vaccination and COVID-19 severity. Conditioned on variables have bounding boxes, distorted associations are shown as dashed lines.

Epidemic phase bias

Epidemic phase bias occurs in studies with time-varying incidence of infection. This bias is induced by a relationship between the time from infection to symptom onset and an individual's eventual outcome, e.g. those who go on to die may experience a more rapid onset of symptoms following infection, an average of c days sooner. Since estimates must be conditioned on the observed symptom onset date, rather than the unobserved infection date, this relationship can introduce bias into results when an epidemic is in a mode of growth or decline (Seaman, Nyberg *et al.* 2022).

Epidemic phase bias is corrected by 'shifting' the symptom onset date among those who experience the more severe outcome to be c days later, so that the time from infection to symptom onset is uncorrelated with the outcome (the time-to-event of interest remains the same for both outcomes.) As the true value of c is typically unknown, sensitivity analysis with differing values of c can be used to assess the susceptibility of results to this bias (Seaman, Nyberg *et al.* 2022).

3. HIV incidence estimation

In common with many infectious disease monitoring programmes, observable data on new HIV diagnoses are not a reliable measure of the underlying incidence of infection in a population. Signals in HIV diagnosis data can be particularly misleading for a number of reasons: the time from infection to diagnosis varies substantially between individuals, diagnoses may be under-reported or delayed by several years, and are subject to external forces such as mass-testing campaigns and migration (Brizzi 2017).

Even with enhanced surveillance systems and widespread availability of testing, the period during which an individual may experience no symptoms (and therefore be unaware of their HIV infection) is on the order of years rather than days or months. Quantifying the incidence and prevalence of the virus in a timely way is therefore far from straightforward.

Statistical techniques which account for these time-varying factors are therefore crucial to reconstruct the underlying prevalence and incidence of HIV from available data.

3.1 Aims of this chapter

This chapter focuses on the development and application of a multi-state modelling method known as 'CD4 back-calculation', which uses: observed numbers of HIV diagnoses over time; the distribution of CD4 counts at, or soon after, diagnosis; and information on disease progression to reconstruct the unobserved HIV incidence and estimate quarterly probabilities of HIV diagnosis (Birrell, Chadborn *et al.* 2012; Brizzi, Birrell, Plummer *et al.* 2019).

In recent years, due to increased frequency of testing and more sensitive diagnostic tests, individuals with recently acquired HIV have been at greater risk of diagnosis during the acute phase of infection. During this acute phase, the CD4 count measurements used in the HIV back-calculation model could present a misleading picture of disease progression, which may impact upon estimates. I aim to quantify the extent to which CD4 count data may incorrectly classify an individual's progression towards AIDS, using additional information on biomarkers of recent infection. I then aim to incorporate these biomarkers into the existing CD4 back-calculation method to ensure the suitability of this model in an era of more frequent testing.

During the COVID-19 pandemic, HIV testing activity was disrupted due to healthcare service reconfiguration, and HIV transmission was likely interrupted as a result of restrictions

on social mixing. The effect was an abrupt change in the number of new HIV diagnoses during 2020. I aim to investigate the sensitivity of the back-calculation model to this change, and estimate if England's ambition to end HIV transmission by 2030 remains on track (HIV Commission 2020).

3.2 Background

3.2.1 HIV in the UK

In the UK, the HIV epidemic is highly concentrated among gay, bisexual and other men who have sex with men (GBM), and black African populations. In 2022, HIV prevalence was estimated to exceed 7% and 1.8% in these groups, respectively, compared to an overall prevalence of 0.2% among 15-to-74 year-olds in England (Martin, Okumu-Camerra *et al.* 2023).

HIV prevention targets

The UK has seen large increases in regular and repeat HIV testing in recent years, particularly among GBM who attend sexual health services (SHS), and PrEP Impact Trial participants who accessed 3-monthly testing during 2017–2020 (Brown, Nash *et al.* 2018; Brown, Mohammed *et al.* 2017; PrEP Impact Trial 2020). Meanwhile, following an update to treatment guidelines in 2015, increased uptake of HIV treatment has led to a reduction in the proportion of people living with detectable levels of HIV (Churchill *et al.* 2016; Martin, Okumu-Camerra *et al.* 2023). As a result, there is now compelling evidence of a trend towards earlier diagnosis and a decline in HIV transmission, detected first among GBM in London, and later confirmed across England (Brizzi, Birrell, Kirwan *et al.* 2021; Brown, Mohammed *et al.* 2017).

In 2020, an independent commission on HIV set ambitious targets to reduce the number of new HIV infections in England to under 600 by 2025 and under 100 by 2030, the respective targets for GBM being under 250 by 2025 and under 50 by 2030 (HIV Commission 2020). This ambition to eliminate HIV as a public health threat by 2030 remains a significant one, requiring continued investment into prevention initiatives and close monitoring of trends in HIV transmission using statistical methodologies (Department of Health and Social Care 2019; HIV Commission 2020).

3.2.2 Natural history of HIV infection

The natural history of HIV infection in the absence of treatment spans several distinct phases: eclipse, acute, chronic, and AIDS (Maartens *et al.* 2014).

The **eclipse phase** lasts around 10 days and begins immediately following HIV transmission. During this period the virus is undetectable in blood serum (also known as the 'window period') (Cohen, Gay *et al.* 2010).

The **acute phase**, or 'seroconversion' phase, lasts around 12 weeks and in this time viruses spread from the site of infection and enter the bloodstream, a process known as viremia. This phase ends with **seroconversion**, when HIV antibodies become detectable in the blood, and some individuals may experience flu-like symptoms around this time, known as seroconversion illness (Cohen, Gay *et al.* 2010; Muema *et al.* 2020).

The **chronic** (generally asymptomatic) **phase** lasts 6-to-7 years, although this can be much shorter or much longer depending on the individual. Throughout this phase, the HIV virus progressively destroys the immune system, depleting CD4+ T cell counts and leading to immune dysfunction (Maartens *et al.* 2014).

Finally, in late-stage HIV infection an individual will progress to **AIDS**, characterised by infection by opportunistic pathogens, including *Pneumocystis jirovecii* and *Candida*, and development of cancers such as Karposi's sarcoma and non-Hodgkin's lymphoma (Dowdle 1983).

3.2.3 Biological markers of disease progression

CD4 counts

CD4+ T cells play a central role in the adaptive and innate response to infection (MacLeod *et al.* 2009). In adults without HIV, CD4 counts typically range between 500–1500 cells/mm^3 (Maini *et al.* 1996). Following HIV infection, as the virus attacks the immune system, an individual's CD4 count follows different trajectories according to the phase of infection: dipping and recovering during the acute phase, followed by a steady decline to zero during the chronic phase. The depletion of CD4 cells during the chronic phase of HIV infection has been well-characterised in clinical studies and follows an approximately quadratic decline, as shown in Figure 3.1 (Maartens *et al.* 2014). An HIV diagnosed adult with a CD4 count of 350 cells/mm^3, and not diagnosed during acute infection, is likely to have been living with HIV for 3-to-5 years, although estimates vary considerably by age, ethnicity, and exposure category (Lodi *et al.* 2011; May *et al.* 2009; Touloumi *et al.* 2013; Yin *et al.* 2021). With adherence to effective ART, the HIV virus can be suppressed and an individual's CD4 count will recover (Ford *et al.* 2015).

Globally, in recognition of the importance of the CD4 marker in determining immune function, most healthcare settings test an individual's CD4 at the point of HIV diagnosis, and at subsequent healthcare visits, with current UK standards of care guidelines recommending annual CD4 measurements as part of routine HIV outpatient follow-up (British HIV Association 2018). Prior to 2015, a low CD4 count (<350 cells/mm^3) was the primary indication for starting

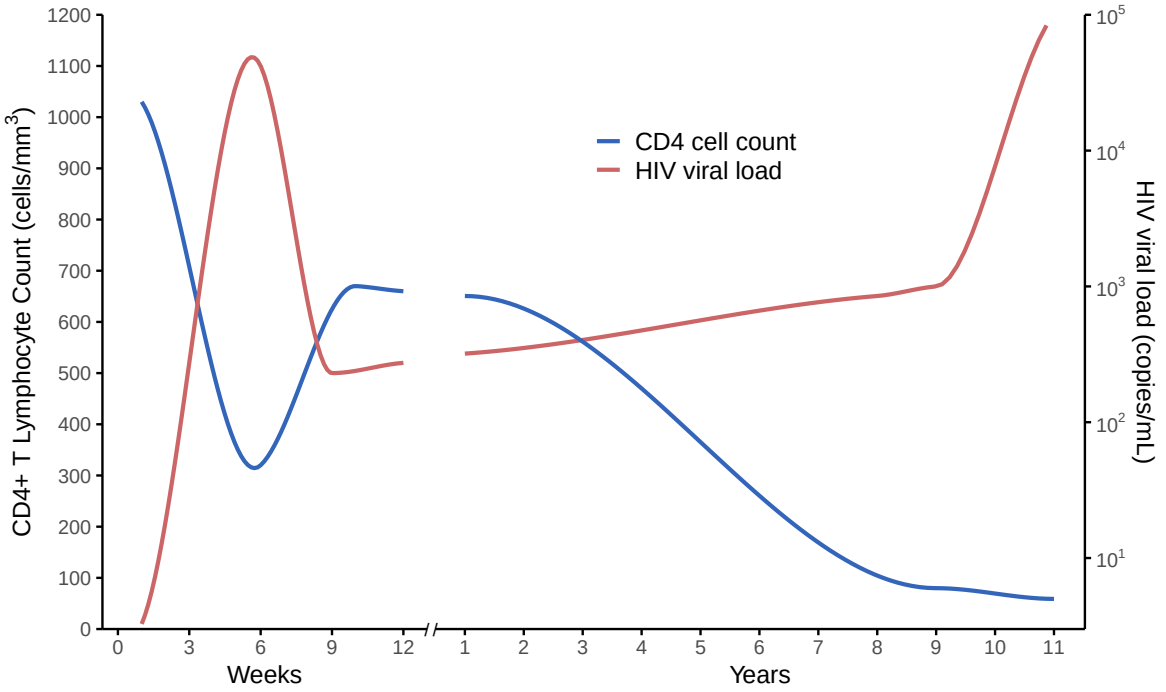


Figure 3.1: CD4 cell count (cells/mm^3) and HIV viral load (copies/mL) trajectories following HIV infection. Figure adapted from Maartens *et al.* 2014.

ART in the UK. In recognition of high adherence to ART, low side-effects of current regimes, and evidence from randomised trials, the CD4-based treatment criteria has since been removed from most European treatment guidelines, meaning that treatment now often starts immediately upon diagnosis (Churchill *et al.* 2016; European AIDS Clinical Society 2023; INSIGHT START Study Group *et al.* 2015).

Despite changes to treatment guidelines, CD4 remains widely accepted as an important public health measure, with the European Late Presenter Consensus working group definition of 'late HIV diagnosis' being a CD4 count <350 cells/mm^3, or an AIDS-defining illness at presentation (Antinori *et al.* 2011), a definition adopted by the WHO and European Centre for Disease Prevention and Control (ECDC) (European Centre for Disease Prevention and Control 2020; World Health Organisation 2010). Late diagnosis rates may be used to evaluate the success of HIV testing strategies, since a decrease in the number and proportion of people diagnosed late alongside sustained or increased testing volumes indicates a trend towards earlier diagnosis (Delpech *et al.* 2013). Reducing late diagnosis is now a priority area for the HIV Action Plan for England 2022-to-2025 (Department of Health and Social Care 2021b).

At an individual level, the measurement of CD4 cells is subject to wide variation; immune cells are known to follow a circadian rhythm, are not uniformly distributed in the blood, and respond to foreign antigens (introduced by diet, pathogens in the air, etc.) and stress (Keshani *et al.* 2023; Moncivaiz 2013). However, when averaged at a population level with appropriate

grouping intervals, progression between subsequent levels of CD4 (post-seroconversion) can be reliably estimated from studies of HIV seroconverters with known dates of HIV acquisition (CASCADE (Concerted Action on SeroConversion to AIDS and Death in Europe) Collaboration 2000).

HIV avidity assays

As the immune response to HIV develops following infection, the body develops more effective antibodies which bind more tightly. The binding strength, or antibody 'avidity', can be measured using avidity assays, and the avidity score used to measure disease progression and distinguish 'recent' from 'non-recent' HIV acquisition (Murphy and Parry 2008; SEDIA Biosciences Corporation 2021). HIV avidity assays are calibrated against a pool of samples with assumed known infection dates to find an appropriate cut-off — an avidity score below which recent HIV acquisition is likely. Samples with avidity scores falling below this cut-off are then used to estimate a mean duration of recent infection (MDRI) for the specific assay. The MDRI is the average time spent recently infected within a time T post-infection, expressed as:

$$\Omega_T = \int_0^T P_R(t)\,\mathrm{d}t$$

where $P_R(t)$ is the probability of being alive and 'recently' infected at time t after infection (Kassanjee, McWalter *et al.* 2012).

Avidity assays have been shown to be sensitive to severe immunosuppression from advanced disease, use of ART, elite controllers with naturally highly suppressed viral load, co-infections, HIV subtype, and pregnancy, with a false recency rate (FRR) of up to 10% (Suligoi, Rodella *et al.* 2011). This misclassification can be minimised if results are considered as part of an algorithm which utilises clinical markers for established infection and treatment status (Aghaizu *et al.* 2014). In England, Wales, and Northern Ireland (EW&NI), avidity test results are routinely incorporated with epidemiological data in a recent infection testing algorithm (RITA) to classify new HIV diagnoses as recent or non-recent (see Appendix C.1 for further details). The FRR after application of the RITA algorithm has been estimated as 1.9% (95% confidence interval 1.0–3.4%) (Aghaizu 2018).

Additional biomarkers collected as part of routine HIV surveillance in the UK include viral load, p24 antigen and HIV testing history, described in Appendix C.1.

3.2.4 UK HIV case surveillance

Surveillance of HIV in the UK began in 1982, with the first case reports of AIDS to the Communicable Disease Surveillance Centre (CDSC), one of the predecessors of UKHSA.

Reporting of new HIV diagnoses was introduced after the first test for HIV became available in the UK in 1984 (British Broadcasting Company 1984).

Reports of new HIV diagnoses have been collected ever since, with data currently submitted electronically to UKHSA on an annual basis by laboratories and clinicians from a variety of diagnosis settings across EW&NI (Rice *et al.* 2017). The covariates collected for new HIV diagnoses include demographic information: sex, age, ethnicity, country of birth, probable route of HIV exposure; and diagnosis data: setting of diagnosis and first CD4 count and date of CD4 test. All reports of HIV diagnoses are validated against a set of rules, for example: no partial dates are accepted and records missing key demographic identifiers are not added to the database. Data cleaning, validation, and de-duplication takes place yearly and records are retained in an annual archive (Public Health England 2013a).

Biomarker data

National surveillance of HIV in EW&NI includes enhanced reporting of biomarker data, with an established record linkage algorithm used to match these biomarker data to HIV diagnosis records for the annual data archive (Rice *et al.* 2017; Winter *et al.* 2016).

The CD4 surveillance scheme, established in 1995, collects longitudinal CD4 count data from 60 laboratories to allow for monitoring of late HIV diagnosis, immunosuppression, and to evaluate the effects of ART (Public Health England 2013b). Samples for HIV avidity testing, alongside pseudonymised demographic information, have been collected centrally by UKHSA since 2011. Avidity testing is used for public health monitoring of the proportion of HIV infections acquired recently, and at the individual level to inform the management of HIV conditions (Aghaizu *et al.* 2014).

Avidity assays have been used for around half of all new HIV diagnoses in EW&NI over this period, with two avidity assays used: the HIV 1/2gO AxSYM assay was used for tests carried out between 2011–2013 (Suligoi, Galli *et al.* 2002), and the Sedia Limiting Antigen Avidity Enzyme Immunoassay (LAg-Avidity EIA) used between 2013–2022 (Sedia Biosciences Corporation 2013). The MDRI on a sample taken within 120 days of HIV diagnosis and prior to ART initiation is around 6 months for both assays[1].

[1]An AxSYM avidity index score <80.0% has an estimated MDRI of 202 days (95% credible interval (CrI) 174–245 days) (Sweeting, De Angelis, Parry *et al.* 2010). A Sedia LAg with normalised optical density score <1.5 has an estimated MDRI of 188 days (95% confidence interval 165–211) (Kassanjee, Pilcher *et al.* 2014).

3.2.5 HIV epidemic models

Back-calculation

As with most infectious diseases, individuals presenting with HIV-related symptoms will typically be tested and receive a diagnosis. Provided an estimate of the incubation period (i.e. the time from exposure to symptoms) is available, this can be combined with observed data on diagnoses to 'back-calculate' the incidence of infection. The general idea of back-calculation is that the time of symptom onset for an infected person is equal to the sum of the time of exposure and the incubation period (Egan and Hall 2015). In continuous time this can be expressed by the following convolution:

$$d(t) = \int_{t_0}^{t} h(s) f(t-s)\,\mathrm{d}s$$

where $d(t)$ is the rate of diagnosis at time t, t_0 is the starting time of the epidemic, $h(s)$ is the rate of new infections at time s, and $f(t-s)$ is the incubation distribution (Brookmeyer, Gail *et al.* 1994).

The occurrence of infections over time is typically modelled as a non-homogeneous Poisson process with the number of new infections in a particular time interval being i.i.d. Poisson random variables with mean $h(s)$ (Becker *et al.* 1991; Rosenberg and Gail 1991). As linear combinations of Poisson random variables are themselves Poisson-distributed, so the number of new diagnoses is Poisson distributed with mean $d(t)$ (Chiang 1980).

The back-calculation method was first developed to estimate HIV prevalence from AIDS cases, with the incubation distribution representing the time from HIV exposure to onset of AIDS (Brookmeyer and Gail 1988). With the introduction of HIV testing, individuals could be diagnosed before AIDS symptoms became apparent and the back-calculation model was therefore extended to consider pre-AIDS diagnoses. This extension introduced a time-dependent incubation distribution, $f(t-s \mid s)$, to capture changes in testing (Aalen, Farewell *et al.* 1997). Further extensions of the model have incorporated HIV progression estimates from CD4 count data to better characterise the time between infection and diagnosis and to estimate trends in the probability of HIV diagnosis by disease stage (Birrell, Chadborn *et al.* 2012; van Sighem, Nakagawa *et al.* 2015; Sweeting, De Angelis and Aalen 2005).

In 2019, Brizzi, Birrell, Plummer *et al.* 2019 extended the model to include age-specific HIV incidence, allowing for the estimation of age groups at higher risk of HIV infection. Estimates from this model have recently provided evidence of a fall in HIV transmission among GBM, with an estimated 80% reduction in the incidence of new HIV infection between 2011 and 2019, concentrated among younger GBM (Brizzi, Birrell, Kirwan *et al.* 2021).

Approaches to HIV incidence estimation

Back-calculation models have been used to estimate HIV incidence from diagnosis data in England, Italy, and Australia (Aalen, Farewell *et al.* 1997; Bellocco and Marschner 2000; Cui and Becker 2000), with a CD4-staged back-calculation model recently adapted for use in European Union (EU) member states (Regine *et al.* 2018; van Sighem, Nakagawa *et al.* 2015; van Sighem, Pharris *et al.* 2017). Aside from back-calculation, several other approaches to HIV incidence estimation exist.

Multi-parameter evidence synthesis (MPES) is a Bayesian approach which combines HIV and sexually transmitted infection (STI) data sources to estimate serial HIV prevalence rates in key populations (Goubar *et al.* 2008; Presanis, Harris *et al.* 2021). Whilst typically used to infer undiagnosed HIV prevalence in the UK, MPES can be applied to estimate HIV incidence, parameterised in terms of prevalence and contact rates (Presanis, De Angelis *et al.* 2011).

Methods to derive HIV incidence from cross-sectional serological assays which detect recent infection, e.g. the p24 antigen or BED capture enzyme immunoassay, were first proposed in 1995 (Brookmeyer and Quinn 1995), and have since been implemented in a number of countries, including Brazil, the United States, and across sub-Saharan Africa (Hall, Song *et al.* 2008; Karita *et al.* 2007; Le Vu *et al.* 2008). Given concerns about the accuracy of these tests, approaches which consider multiple biomarkers have been developed with the aim of minimising misclassification of recent infection. These methods incorporate CD4 counts, HIV viral loads, and avidity assays alongside the p24 antigen assay (Brookmeyer, Konikoff *et al.* 2013; Sun *et al.* 2020).

An individual-based stochastic simulation model, calibrated to multiple sources of data, has been used to estimate HIV incidence in the UK (Phillips *et al.* 2015). Most recently this model was applied to explore the contribution of earlier treatment initiation and expanded testing in reducing HIV incidence (Cambiano *et al.* 2023).

Finally, the Joint United Nations Programme on HIV/AIDS (UNAIDS) Spectrum model is used nationally and internationally to provide annual estimates of HIV mortality and incidence, particularly in countries with high prevalence of paediatric HIV. Spectrum utilises HIV diagnoses in antenatal settings, testing policy and ART coverage information from surveillance and survey data in a joint statistical and mathematical model (Stover *et al.* 2019).

3.3 CD4-staged back-calculation model

The CD4-staged back-calculation model is a Bayesian, discrete-time, multi-state model (Birrell, Chadborn *et al.* 2012). A diagram of states and transitions for the model is shown in Figure 3.2.

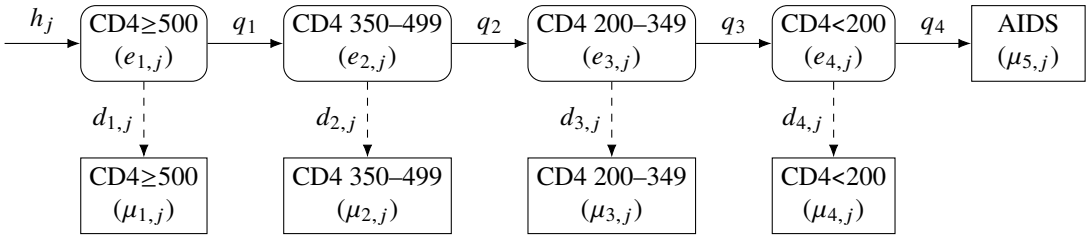


Figure 3.2: CD4-staged back-calculation model. Rounded boxes indicate latent (or undiagnosed) states, square boxes indicate diagnosed states. Solid lines indicate HIV progression. Dashed lines indicate transition from latent to diagnosed state.

In this model HIV infections are assumed to occur according to a non-homogeneous Poisson process with rate $\lambda(t)$. The expected number of infections in the time interval $(t_{j-1}, t_j]$ is given by:

$$h_j = h(t_j) = \int_{t_{j-1}}^{t_j} \lambda(t)\,\mathrm{d}t$$

Following infection, state transitions proceed on a quarterly (i.e. 3-monthly) time interval, with individuals assumed to steadily progress through four latent (or undiagnosed) states $e_{i,j}$ (where i indexes the state), defined by progressively lower CD4 counts. At each CD4 stage, i, individuals are subject to differing, time-dependent, diagnosis probabilities $d_{i,j}$. In the absence of a diagnosis individuals will progress to an eventual AIDS state (Birrell, Chadborn *et al.* 2012).

The transition matrix:

$$\mathbf{Q}_j = \begin{bmatrix} (1-d_{1,j})(1-q_1) & (1-d_{1,j})q_1 & 0 & 0 \\ 0 & (1-d_{2,j})(1-q_2) & (1-d_{2,j})q_2 & 0 \\ 0 & 0 & (1-d_{3,j})(1-q_3) & (1-d_{3,j})q_3 \\ 0 & 0 & 0 & (1-d_{4,j})(1-q_4) \end{bmatrix}$$

describes the probability of progression between the latent states of the model during the time interval $(t_{j-1}, t_j]$. Here q_i are fixed constant progression probabilities — from one latent state to the next, estimated from studies of HIV seroconverters (CASCADE (Concerted Action on SeroConversion to AIDS and Death in Europe) Collaboration 2000); and $d_{i,j}$ are diagnosis probabilities — from a latent state to a diagnosed state. A second transition matrix:

$$\mathbf{D}_j = \begin{bmatrix} d_{1,j} & 0 & 0 & 0 & 0 \\ 0 & d_{2,j} & 0 & 0 & 0 \\ 0 & 0 & d_{3,j} & 0 & 0 \\ 0 & 0 & 0 & d_{4,j} & (1-d_{4,j})q_4 \end{bmatrix}$$

describes the probability of moving between a latent state and a diagnosed state during time interval $(t_{j-1}, t_j]$. The final column of $\mathbf{D}_j$ describes the probability of being diagnosed with AIDS; the implicit assumption is that all people who progress to AIDS will be diagnosed.

Denote as $\boldsymbol{e}_j = (e_{1,j}, \ldots, e_{4,j})$ the expected number of individuals in $(t_{j-1}, t_j]$ in each of the latent states, and as $\boldsymbol{\mu}_j = (\mu_{1,j}, \ldots, \mu_{5,j})$ the expected number of new diagnoses in $(t_{j-1}, t_j]$ in each of the diagnosed states. The model dynamics may be expressed as a set of recursive equations which include the transition matrices defined above:

$$\begin{aligned}\boldsymbol{e}_j &= \mathbf{Q}_j^{\mathsf{T}} \boldsymbol{e}_{j-1} + [h_j, 0, \ldots, 0]^{\mathsf{T}} \\ \boldsymbol{\mu}_j &= \mathbf{D}_j^{\mathsf{T}} \boldsymbol{e}_{j-1}\end{aligned}$$

The expected number of HIV and AIDS diagnoses in $(t_{j-1}, t_j]$ are denoted, respectively, as:

$$\mu_j^{\mathrm{H}} = \sum_{i=1}^{4} \mu_{i,j} \text{ and } \mu_j^{\mathrm{A}} = \mu_{5,j}$$

3.3.1 Likelihood

The number of new infections in each time interval $(t_{j-1}, t_j]$ form a set of independent Poisson random variables with mean h_j, for $j = 1, \ldots, N$. At any time interval during the surveillance period individuals with HIV will be in one of three states:

1. diagnosed with HIV;
2. diagnosed with AIDS;
3. undiagnosed.

By the properties of the non-homogeneous Poisson process, the likelihood contributions of the HIV and AIDS diagnoses are given by independent Poisson random variables Y_j^{H} and Y_j^{A} with expectations μ_j^H and μ_j^A, respectively:

$$\begin{aligned}Y_j^{\mathrm{H}} &\sim \text{Poisson}(\mu_j^{\mathrm{H}}) \\ Y_j^{\mathrm{A}} &\sim \text{Poisson}(\mu_j^{\mathrm{A}})\end{aligned}$$

Denote the observed data on HIV and AIDS diagnoses in time interval $(t_{j-1}, t_j]$ as y_j^{H} and y_j^{A}, with $\boldsymbol{y}^{\mathrm{H}} = (y_1^{\mathrm{H}}, \ldots, y_N^{\mathrm{H}})$ and $\boldsymbol{y}^{\mathrm{A}} = (y_1^{\mathrm{A}}, \ldots, y_N^{\mathrm{A}})$. These are realisations of Poisson random variables, so the likelihood contribution from these data is:

$$L_1(\boldsymbol{y}^{\mathrm{H}}, \boldsymbol{y}^{\mathrm{A}}; \boldsymbol{h}, \boldsymbol{d}) \propto \prod_{j=1}^{T} (\mu_j^{\mathrm{A}})^{y_j^{\mathrm{A}}} \exp(-\mu_j^{\mathrm{A}}) \times (\mu_j^{\mathrm{H}})^{y_j^{\mathrm{H}}} \exp(-\mu_j^{\mathrm{H}})$$

Then, for each of the four CD4 states, let $Y^{\mathrm{C}}_{i,j}$ be the random variable for the number of HIV diagnoses in state i in time interval $(t_{j-1}, t_j]$. These diagnoses are multinomially distributed, with CD4 counts assumed to be missing at random, so:

$$Y^{\mathrm{C}}_{i,j} \sim \text{Multinomial}(n_j, \boldsymbol{c}_j)$$

where $\boldsymbol{c_j}$ is the parameter set of expected probabilities in each CD4 diagnosis state:

$$\boldsymbol{c}_j = \left(\frac{\mu_{1,j}}{\mu^{\mathrm{H}}_j}, \frac{\mu_{2,j}}{\mu^{\mathrm{H}}_j}, \frac{\mu_{3,j}}{\mu^{\mathrm{H}}_j}, \frac{\mu_{4,j}}{\mu^{\mathrm{H}}_j} \right)$$

n_j is the observed total number of HIV diagnoses in $(t_{j-1}, t_j]$ with CD4 count data available:

$$n_j = \sum_{i=1}^{4} y^{\mathrm{C}}_{i,j} \subseteq y^{\mathrm{H}}_j$$

and $y^{\mathrm{C}}_{i,j}$ are the observed CD4 data in each CD4 diagnosis state i, and time interval $(t_{j-1}, t_j]$, with $\boldsymbol{y}^{\mathrm{C}} = \{y^{\mathrm{C}}_{i,j} : i = 1, \ldots, 4, j = 1, \ldots, N\}$.

The likelihood contribution from the observed CD4 data is therefore:

$$L_2(\boldsymbol{y}^{\mathrm{C}} \mid \boldsymbol{y}^{\mathrm{H}}; \boldsymbol{h}, \boldsymbol{d}) \propto \prod_{j=1}^{T} \prod_{i=1}^{4} (c_{i,j})^{y^{\mathrm{C}}_{i,j}}$$

and the full likelihood for this model is the product of the two likelihood terms:

$$L(\boldsymbol{y}^{\mathrm{H}}, \boldsymbol{y}^{\mathrm{A}}, \boldsymbol{y}^{\mathrm{C}}; \boldsymbol{h}, \boldsymbol{d}) = L_1(\boldsymbol{y}^{\mathrm{H}}, \boldsymbol{y}^{\mathrm{A}}; \boldsymbol{h}, \boldsymbol{d}) \times L_2(\boldsymbol{y}^{\mathrm{C}} \mid \boldsymbol{y}^{\mathrm{H}}; \boldsymbol{h}, \boldsymbol{d})$$

3.3.2 Smoothing incidence and diagnosis trends

To address identifiability issues, the time-varying infection ($\boldsymbol{h}$) and diagnosis ($\boldsymbol{d}$) rates are specified according to random-walks on the log-scale. Prior distributions are chosen as an uninformative Gamma$(1, 32)$ for the variance of the random walks, a Normal distribution focused at $\log(2)$ for the starting point of the incidence random walk, and according to results of an earlier study for the mean of the diagnosis random walk at 1985, with a ramp up to this level for previous years (Aalen, Farewell *et al.* 1997; Birrell, Gill *et al.* 2013). This specification serves to smooth these processes over time, ensuring that sudden jumps do not occur between time intervals (Birrell, Chadborn *et al.* 2012). Alternate parametric smoothing methods for the incidence and diagnosis rates may be specified, e.g. Gaussian processes or various spline models (Brizzi 2017).

In the Bayesian framework the priors are updated by the likelihood to obtain the posterior. As the posterior in this case is intractable, samples are obtained via a suitable MCMC algorithm (e.g. as implemented in the JAGS or STAN software).

3.3.3 Latent states

Let $Y^{\mathrm{E}}_{i,j}$ be the random variable for the number of individuals in each of the four latent states after time interval $(t_{j-1}, t_j]$, with expectation $e_{i,j}$. To make statements about these latent states whilst accounting for the full uncertainty about the number of individuals in each state, the *posterior predictive* distribution for the unobserved $\boldsymbol{y}^{\mathrm{E}} = \{y^{\mathrm{E}}_{i,j} : i = 1, \ldots, 4, j = 1, \ldots, N\}$ is needed. This posterior predictive distribution is calculated by marginalising over the posterior distribution of the model parameters, $\theta = (\boldsymbol{h}, \boldsymbol{d})$, given the data:

$$\pi(\boldsymbol{y}^{\mathrm{E}} \mid \boldsymbol{y}^{\mathrm{H}}, \boldsymbol{y}^{\mathrm{A}}, \boldsymbol{y}^{\mathrm{C}}) = \int_{\Theta} p(\boldsymbol{y}^{\mathrm{E}} \mid \boldsymbol{y}^{\mathrm{H}}, \boldsymbol{y}^{\mathrm{A}}, \boldsymbol{y}^{\mathrm{C}}, \theta) \pi(\theta \mid \boldsymbol{y}^{\mathrm{H}}, \boldsymbol{y}^{\mathrm{A}}, \boldsymbol{y}^{\mathrm{C}}) \, \mathrm{d}\theta$$

where $\pi(.)$ represents a posterior probability distribution. The distribution $p(\boldsymbol{y}^{\mathrm{E}} \mid \boldsymbol{y}^{\mathrm{H}}, \boldsymbol{y}^{\mathrm{A}}, \boldsymbol{y}^{\mathrm{C}}, \theta)$ is a non-standard distribution, and computationally expensive to simulate from. Instead, 'post-hoc' methods may be used to derive quantities of interest which involve the posterior predictive distribution for $\boldsymbol{y}^{\mathrm{E}}$. For instance, by expressing undiagnosed HIV prevalence as the total number of infections minus the number of HIV and AIDS diagnoses (Birrell, Chadborn *et al.* 2012).

3.3.4 Limitations

Despite the successful application of the existing CD4 back-calculation model to reveal changes in incidence which precede an observed fall in diagnoses (Brizzi, Birrell, Kirwan *et al.* 2021), the approach has certain limitations which affect both the precision and relevance of estimates for continued HIV incidence estimation.

Misclassification of longstanding infection

The success of combination HIV prevention initiatives during the past decade have led to a higher frequency of testing among those most at risk of infection (Brown, Mohammed *et al.* 2017). Over the same period, more sensitive 'fourth-generation' HIV tests, capable of detecting the virus earlier in infection, were recommended for routine use in the UK (National Institute for Health and Care Excellence 2016a). As a result, the probability of HIV diagnosis occurring during the acute phase of infection (when CD4 counts are prone to dip) is likely to have increased substantially. In a recent Belgian study, for example, a third of people diagnosed with recently-acquired infection had a CD4 count below 350 cells/mm^3 (Sasse *et al.* 2016). Whilst a

trend towards earlier diagnosis is undoubtedly a positive for HIV prevention, and beneficial at an individual-level, the use of CD4 counts at diagnosis to measure progression towards AIDS may be increasingly subject to error.

The CD4 back-calculation model does not account for a transient dip in CD4 cell counts during acute infection. And while similar CD4-staged HIV estimation approaches have included a 'primary infection' stage, individuals are not assumed to be diagnosed during this stage (van Sighem, Nakagawa *et al.* 2015). The issue of CD4 counts dipping early in infection is also a challenge for determining population-level late HIV diagnosis rates, as highlighted by UK and other European reports (Brännström *et al.* 2016; HIV Commission 2020; Sasse *et al.* 2016).

Uncertainty in recent incidence estimates

A second limitation of the back-calculation model stems from the high variability of CD4 cell counts and, as a result, the necessary imprecision of CD4 strata used in the model. In particular, there is a wide time interval following HIV seroconversion during which CD4 cell counts may remain above 500 cells/mm^3 (around 3 years) (Maartens *et al.* 2014). This is particularly evident in recent years which, crucially, are those used to assess current progress towards elimination of HIV transmission.

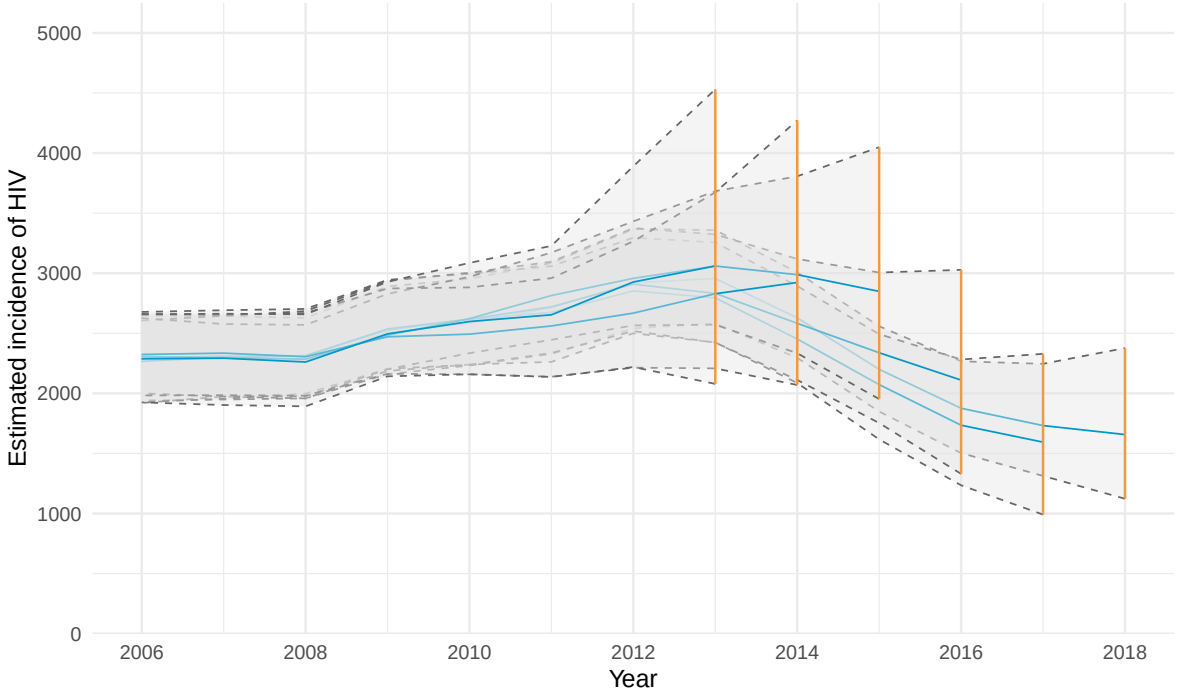


Figure 3.3: Successive annual estimates of HIV incidence (posterior median and 95% CrI) estimated by CD4 back-calculation. The vertical line indicates the range of the credible interval for the the most recent year of available diagnosis data.

As shown in Figure 3.3, for successive annual estimates of HIV incidence estimates the credible intervals are widest for the most recent 2–3 years, narrowing as subsequent diagnosis information becomes available. Being able to better discern the recency of a newly diagnosed infection may help to improve the precision of incidence estimation.

Incorporating other biomarkers of recent infection (such as those described in Section 3.2.3) could help to address both limitations: complementing the CD4 cell count by detecting acute and/or early HIV infection, and improving the precision and identifiability of the model for the most recent years.

The extent of CD4 misclassification has not been previously investigated for UK data. In the next section I investigate the availability and representativeness of the available biomarker data, then use these markers to quantify to what extent an HIV diagnosis occurring during the acute or seroconversion phase may be misclassified as a 'late diagnosis' (due to being diagnosed with a CD4 count <350 cells/mm^3).

3.4 Analysis of biomarker data

An initial investigation of the availability and representativeness of HIV viral loads, HIV avidity assays, and HIV testing history for 10 years of diagnosis data among GBM in EW&NI was undertaken. The key findings are summarised below, with full results included in Appendix C.2.

3.4.1 Key findings

Based on HIV avidity assay results and HIV testing history, an increasing proportion of newly diagnosed GBM with low CD4 counts (<350 cells/mm^3) were found to have evidence of recent HIV acquisition; around a quarter in 2019 compared to <10% in 2011. This finding was corroborated by viral load information — an increasing proportion of those diagnosed with recent infection had elevated baseline viral loads (>100,000 copies/mL), suggestive of diagnosis during acute infection, at a time of peak viremia.

The results of the biomarker data were somewhat age-dependent, with the avidity assay analyses suggesting that younger GBM were more likely than older GBM to be diagnosed promptly. Availability of previous negative test data was similarly age-dependent, with a higher proportion of GBM aged <50 having previous negative tests within 6 months of their HIV diagnosis, compared to older GBM. This reflects the concentration of HIV-prevention campaigns, and high testing volumes at SHS (and therefore greater likelihood of being diagnosed during seroconversion) for younger age groups (Whitlock *et al.* 2020).

Despite excluding individuals known to have initiated treatment, an increasing proportion of newly diagnosed GBM had a baseline viral load <50 copies/mL, highly suggestive of viral

suppression due to ART. Whilst this may be due in part to missing information on treatment start date, widespread use of PrEP among GBM may result in viral load suppression at diagnosis for the small number of breakthrough cases (Ambrosioni *et al.* 2021). Since ART, post-exposure prophylaxis (PEP), and PrEP all affect the measurement of avidity, the application of the RITA algorithm (which accounts for both viral load and treatment start information) will be necessary to maintain a low false recency rate (FRR).

3.4.2 Revised late HIV diagnosis definition

The public health definition of late HIV diagnosis is a CD4 cell count <350 cells/mm^3. The population-level late diagnosis rate is an important public health measure used to target HIV-prevention initiatives, such as testing campaigns, ART initiation, and PrEP (Brown, Mohammed *et al.* 2017; Churchill *et al.* 2016; NHS England 2019), and a key indicator of the English Public Health Outcomes Framework (Office for Health Improvement 2023).

Based on the analysis of HIV biomarker data in this section, the CD4<350 cells/mm^3 definition will increasingly overestimate rates of late HIV presentation in EW&NI. To address this overestimation I developed a reclassification method in collaboration with colleagues at UKHSA. This method incorporates routine tests for recent infection and evidence of a recent negative HIV test into the measurement of population-level late diagnosis rates, with the revised definition of late HIV diagnosis below (Kirwan, Croxford *et al.* 2022).

Revised definition of late HIV diagnosis

Among adults (aged ≥15 years) with a baseline CD4 count (within 14 to 91 days), those with a CD4 count <350 cells/mm^3 and neither:

(a) a recent infection testing algorithm (RITA) result indicating recent HIV acquisition nor

(b) a negative HIV test result within the preceding 24 months.

The impact of the revised late HIV diagnosis definition on late diagnosis rates in EW&NI is described in Appendix C.3. The updated definition is now routinely used for public health monitoring at English local authority-level (Office for Health Improvement 2023), at national level (Martin, Okumu-Camerra *et al.* 2023), and has informed an updated European consensus definition of late HIV diagnosis (Croxford *et al.* 2022).

These findings motivate the next section of this chapter, where avidity test information will be introduced as an additional biomarker for the CD4 back-calculation, with the aim of improving model performance and addressing the misclassification of recent infection.

3.5 Dual biomarker back-calculation model

In this section, I propose an updated 'dual biomarker' back-calculation model for HIV incidence estimation, incorporating both CD4 cell count and HIV avidity test information. The model is described in detail below, followed by a simulation study to explore the performance of the dual biomarker model compared to the original CD4 back-calculation model. The updated version of the back-calculation model is then applied to HIV surveillance data for England.

Of the available biomarkers, the avidity assay biomarker had the least bias in data availability, the RITA algorithm helped to mitigate false recency error, and data were derived from independently-reported laboratory samples. The HIV testing history data, meanwhile, had greater age-dependency, with an unknown degree of error in recall and reporting, and these data were subject to external factors such as HIV testing campaigns and expanded point-of-care testing.

3.5.1 Model description

Addition of recently acquired HIV states

The addition of a recent state to the CD4 back-calculation model should, conceptually, allow for progression through an additional 'recently acquired' latent state, and corresponding diagnosed state, as shown in Figure 3.4, where the (time-varying) probability of diagnosis is $d_{1,j}$.

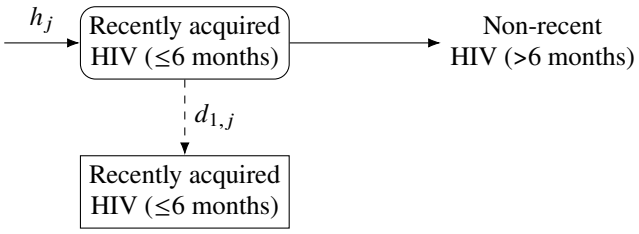


Figure 3.4: Recently acquired state for HIV back-calculation model. Rounded boxes indicate latent states, square boxes indicate diagnosed states. Solid lines indicate HIV progression. Dashed lines indicate transition from latent to diagnosed state.

The MDRI for a recent avidity assay result has previously been estimated as around 6 months, which covers two discrete time steps in the model (Kassanjee, Pilcher *et al.* 2014; Sweeting, De Angelis, Parry *et al.* 2010). For consistency with the CD4 back-calculation and

data reporting mechanisms, it is desirable to retain the quarterly timescale. To accomplish this, the recently acquired HIV latent state is divided in two, reflecting 0–3 and 3–6 months post-HIV acquisition, respectively, as shown in Figure 3.5. The same time-varying probability of diagnosis into a single diagnosis state ($d_{1,j}$) is applied for both latent states.

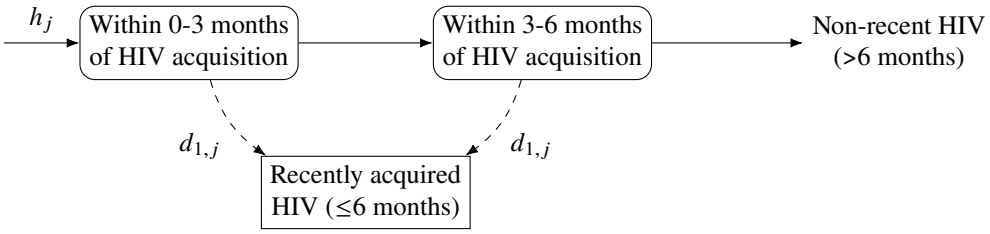


Figure 3.5: Recently acquired states for HIV back-calculation model. Rounded boxes indicate latent states, square boxes indicate diagnosed states. Solid lines indicate HIV progression. Dashed lines indicate transition from latent to diagnosed state.

Finally, to combine these recently acquired states with the existing CD4-staged states, the CD4 progression probabilities $\boldsymbol{q} = (q_1, q_2, q_3, q_4)$ are introduced. In the CD4-staged model progression can occur after 3 months, therefore the latent state 'Within 3–6 months of HIV acquisition' is sub-divided, expanding the number of latent states to three, as shown in Figure 3.6.

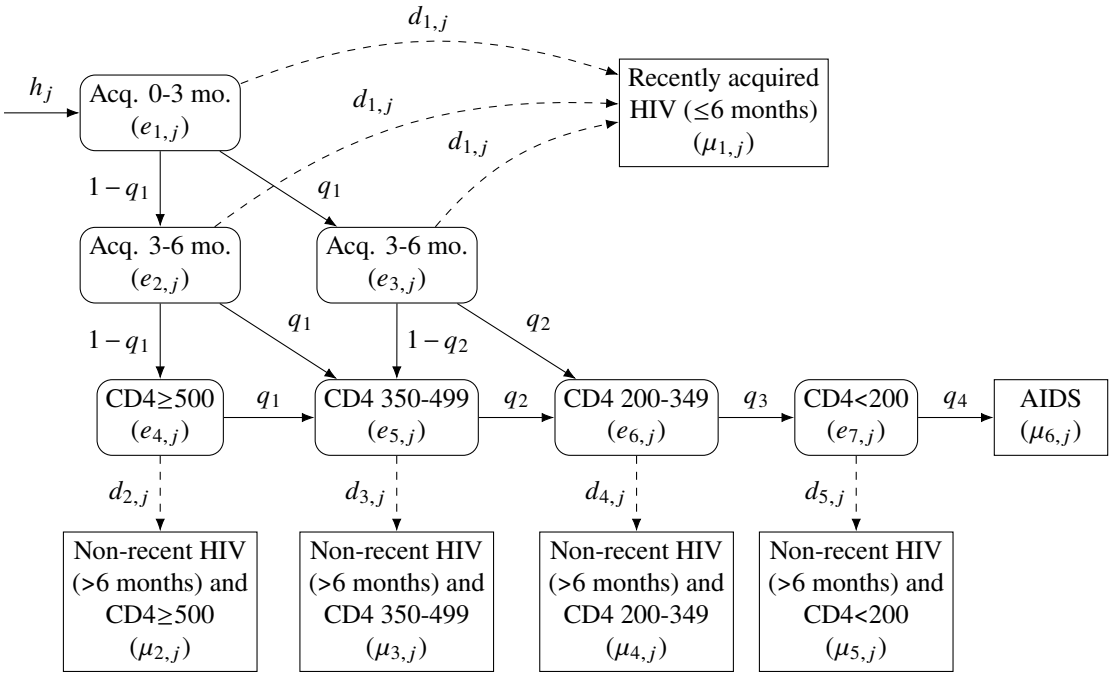


Figure 3.6: Dual biomarker back-calculation model with 'recent incidence assay' and 'non-recent incidence assay' CD4-staged states. Rounded boxes indicate latent states, square boxes indicate diagnosed states. Solid lines indicate HIV progression. Dashed lines indicate transition from latent to diagnosed state. Acq. = Acquired.

The sub-division of states ensures that the same CD4 progression probabilities described in Section 3.3 continue to apply for all individuals in the model. Note the change of orientation in Figure 3.6 so that time proceeds in the vertical direction whilst disease progression proceeds in the horizontal direction. Here, again, the probabilities of diagnosis from all of the three latent states are defined to be equal to $d_{1,j}$.

Individuals who are not recently diagnosed progress to one of the previously described CD4 strata with two time periods having elapsed. For an individual who acquired HIV at time t_j, the probability of entering CD4 stratum S at time t_{j+3}, given they are not recently diagnosed, is:

- $\Pr(S(t_{j+3}) = \text{CD4 500+} \mid \text{not recently diagnosed}) = (1-q_1)(1-q_1)$
- $\Pr(S(t_{j+3}) = \text{CD4 350–499} \mid \text{not recently diagnosed}) = q_1(1-q_1) + q_1(1-q_2)$
- $\Pr(S(t_{j+3}) = \text{CD4 200–349} \mid \text{not recently diagnosed}) = q_1 q_2$
- $\Pr(S(t_{j+3}) = \text{CD4} < 200 \mid \text{not recently diagnosed}) = 0$

The transition matrix $\mathbf{Q}_j$ and diagnosis probability matrix $\mathbf{D}_j$ are updated from the CD4 back-calculation as shown below, with diagonal entries $(1,1)$, $(2,2)$, and $(3,3)$ of $\mathbf{Q}_j$ set equal to zero, indicating no retention in these states at the following time step.

$$\mathbf{Q}_j = \begin{bmatrix} 0 & (1-d_{1,j})(1-q_1) & (1-d_{1,j})q_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & (1-d_{1,j})(1-q_1) & (1-d_{1,j})q_1 & 0 & 0 \\ 0 & 0 & 0 & 0 & (1-d_{1,j})(1-q_2) & (1-d_{1,j})q_2 & 0 \\ 0 & 0 & 0 & (1-d_{2,j})(1-q_1) & (1-d_{2,j})q_1 & 0 & 0 \\ 0 & 0 & 0 & 0 & (1-d_{3,j})(1-q_2) & (1-d_{3,j})q_2 & 0 \\ 0 & 0 & 0 & 0 & 0 & (1-d_{4,j})(1-q_3) & (1-d_{4,j})q_3 \\ 0 & 0 & 0 & 0 & 0 & 0 & (1-d_{5,j})(1-q_4) \end{bmatrix}$$

$$\mathbf{D}_j = \begin{bmatrix} d_{1,j} & 0 & 0 & 0 & 0 & 0 \\ d_{1,j} & 0 & 0 & 0 & 0 & 0 \\ d_{1,j} & 0 & 0 & 0 & 0 & 0 \\ 0 & d_{2,j} & 0 & 0 & 0 & 0 \\ 0 & 0 & d_{3,j} & 0 & 0 & 0 \\ 0 & 0 & 0 & d_{4,j} & 0 & 0 \\ 0 & 0 & 0 & 0 & d_{5,j} & (1-d_{5,j})q_4 \end{bmatrix}$$

As before:

$$\begin{aligned} \boldsymbol{e}_j &= \mathbf{Q}_j^{\mathsf{T}} \boldsymbol{e}_{j-1} + [h_j, 0, \ldots, 0]^{\mathsf{T}} \\ \boldsymbol{\mu}_j &= \mathbf{D}_j^{\mathsf{T}} \boldsymbol{e}_{j-1} \end{aligned}$$

and the expected number of HIV diagnoses in the recent and four non-recent states in $(t_{j-1}, t_j]$ are expressed, respectively, as:

$$\mu_j^{\mathrm{R}} = \mu_{1,j} \text{ and } \mu_j^{\mathrm{L}} = \sum_{i=2}^{5} \mu_{i,j}$$

Likelihood

Following Section 3.3, in each time interval $(t_{j-1}, t_j]$, individuals may be in one of four states:

1. diagnosed with recently acquired HIV, according to RITA (≤6 months);
2. diagnosed with longstanding HIV, according to RITA (>6 months);
3. diagnosed with AIDS;
4. undiagnosed.

Let Y_j^{R} and Y_j^{L} be binomial and multinomial random variables for the number of recent and longstanding HIV diagnoses, respectively:

$$\begin{aligned} Y_j^{\mathrm{R}} &\sim \text{Binomial}(n_j, r_j) \\ Y_{i,j}^{\mathrm{L}} &\sim \text{Multinomial}(n_j^{\mathrm{L}}, \boldsymbol{l}_j) \end{aligned}$$

with parameters:

$$\begin{aligned} n_j &= y_j^{\mathrm{R}} + \sum_{i=1}^{4} y_{i,j}^{\mathrm{L}} & r_j &= \frac{\mu_j^{\mathrm{R}}}{\mu_j^{\mathrm{H}}} \\ n_j^{\mathrm{L}} &= \sum_{i=1}^{4} y_{i,j}^{\mathrm{L}} & \boldsymbol{l}_j &= \left(\frac{\mu_{2,j}}{\mu_j^{\mathrm{L}}}, \frac{\mu_{3,j}}{\mu_j^{\mathrm{L}}}, \frac{\mu_{4,j}}{\mu_j^{\mathrm{L}}}, \frac{\mu_{5,j}}{\mu_j^{\mathrm{L}}} \right) \end{aligned}$$

where y_j^{R} and $y_{i,j}^{\mathrm{L}}$ are the observed data on recent diagnoses and longstanding HIV diagnoses in each of the four CD4 states, respectively, in $(t_{j-1}, t_j]$.

The previously defined likelihood contribution L_1 is retained, the likelihood contribution from the observed CD4 count data (L_2) now relates only to those with longstanding HIV, and a third likelihood contribution (L_3) is now available from the recent diagnosis data, conditional on n_j:

$$\begin{aligned} L_1(\boldsymbol{y}^{\mathrm{H}}, \boldsymbol{y}^{\mathrm{A}}; \boldsymbol{h}, \boldsymbol{d}) &\propto \prod_{j=1}^{T} (\mu_j^{\mathrm{A}})^{y_j^{\mathrm{A}}} \exp(-\mu_j^{\mathrm{A}}) \times (\mu_j^{\mathrm{H}})^{y_j^{\mathrm{H}}} \exp(-\mu_j^{\mathrm{H}}) \\ L_2(\boldsymbol{y}^{\mathrm{L}} \mid \boldsymbol{y}^{\mathrm{R}}, \boldsymbol{y}^{\mathrm{H}}; \boldsymbol{h}, \boldsymbol{d}) &\propto \prod_{j=1}^{T} \prod_{i=1}^{4} (l_{i,j})^{y_{i,j}^{\mathrm{L}}} \\ L_3(\boldsymbol{y}^{\mathrm{R}} \mid \boldsymbol{y}^{\mathrm{L}}, \boldsymbol{y}^{\mathrm{H}}; \boldsymbol{h}, \boldsymbol{d}) &\propto \prod_{j=1}^{T} (r_j)^{n_j} \end{aligned}$$

The full likelihood is the product of all three likelihood terms:

$$L(\boldsymbol{y}^{\mathrm{H}}, \boldsymbol{Y}^{\mathrm{A}}, \boldsymbol{y}^{\mathrm{R}}, \boldsymbol{Y}^{\mathrm{L}}; \boldsymbol{h}, \boldsymbol{d}) = L_1(\boldsymbol{y}^{\mathrm{H}}, \boldsymbol{y}^{\mathrm{A}}; \boldsymbol{h}, \boldsymbol{d}) \times L_2(\boldsymbol{y}^{\mathrm{L}} \mid \boldsymbol{y}^{\mathrm{R}}, \boldsymbol{y}^{\mathrm{H}}; \boldsymbol{h}, \boldsymbol{d}) \times L_3(\boldsymbol{y}^{\mathrm{R}} \mid \boldsymbol{y}^{\mathrm{L}}, \boldsymbol{y}^{\mathrm{H}}; \boldsymbol{h}, \boldsymbol{d})$$

As before, the time-varying $\boldsymbol{h}$ and $\boldsymbol{d}$ are specified according to parametric smoothing method, with appropriate prior distributions. Similar post-hoc methods to those presented by Birrell, Chadborn *et al.* 2012 are used to derive quantities of interest involving the posterior predictive distribution of the latent model states.

Data representativeness

By including only diagnoses with both CD4 and RITA information available to inform HIV progression, the assumption being made is that these data are missing at random. The investigation of the HIV avidity data demonstrated considerable regional differences in data availability, e.g. particularly low availability in Wales. A certain degree of age-dependency in the availability of avidity information and the classification of recent/non-recent using the RITA algorithm was also seen.

These regional and age dependencies contradict the assumption that data are missing at random. To minimise the effect of regional differences only diagnoses in England will be considered when applying the model to national surveillance data. Meanwhile, the difference in data availability by age is relatively small and unlikely to make a meaningful difference to model results.

3.5.2 Simulation study

In this section the results of a simulation study for assessing the performance of the dual biomarker model as compared to the CD4-only model are reported, with a comparison of bias in estimates.

Aim

The aim of the simulation study was to assess the performance of the dual biomarker model and CD4-only model in reconstructing HIV incidence and HIV diagnosis probabilities in the context of HIV diagnosis occurring during acute infection.

Data generating mechanisms

Data were simulated by propagating infections through the convolution equations and transition matrices for the dual biomarker model, as defined in Section 3.5.1:

$$\boldsymbol{e}_j = \mathbf{Q}_j^{\mathsf{T}} \boldsymbol{e}_{j-1} + [h_j, 0, \ldots, 0]^{\mathsf{T}}$$
$$\boldsymbol{\mu}_j = \mathbf{D}_j^{\mathsf{T}} \boldsymbol{e}_{j-1}$$

Data were simulated from these equations 100 times, with independent draws according to the following parametric distributions:

- $h_j \sim \text{Poisson}(\lambda_j)$, with λ_j chosen to approximate a previously estimated pattern of HIV incidence in England, as shown in Figure 3.7 panel A;
- $d_{i,j} \sim N(\delta_{i,j}, 0.001^2)$, with the $\delta_{i,j}$ chosen to represent a scenario in which probability of diagnosis increases, flattens, then decreases, as shown in Figure 3.7 panel B.

The progression probabilities, q_i, were specified to be the same fixed probabilities of disease progression used for the back-calculation models. The initial prevalence in each latent state, $\boldsymbol{e}_0$, was specified as a multiple of the initial incidence simulated at time 0 (h_0), such that $\boldsymbol{e}_0 = \{h_0, 0.8(h_0), 0.1(h_0), 4(h_0), 2(h_0), h_0, 0.25(h_0)\}$. For these simulated data the information on CD4 and RITA were assumed to be complete, both to reduce uncertainty in the model estimates and to aid comparison.

A total of 300 simulated datasets were obtained, using the dual biomarker model as the data generating process. Three model scenarios were constructed, shown below, using 100 datasets each. Each dataset included simulated values for y_j^{H}, y_j^{A}, y_j^{R}, and $y_{i,j}^{\text{L}}$. These values were sufficient for the dual-biomarker model (scenario (i)), but to fit the CD4-only back-calculation model (scenarios (ii) and (iii)) the requisite y_j^{C} needed to be derived from combinations of these values.

(i) No alteration of the simulated data:

- the number of HIV and AIDS diagnoses over time were y_j^{H} and y_j^{A}, respectively;
- the number of diagnoses over time with recently acquired HIV was y_j^{R};
- the distribution of CD4 cell counts over time for those diagnosed with non-recent HIV was $y_{i,j}^{\text{L}}$;

(ii) Specifying a CD4 count distribution for recently acquired HIV diagnoses according to the distribution in Table C.2:

- the number of HIV and AIDS diagnoses over time were y_j^{H} and y_j^{A}, respectively;

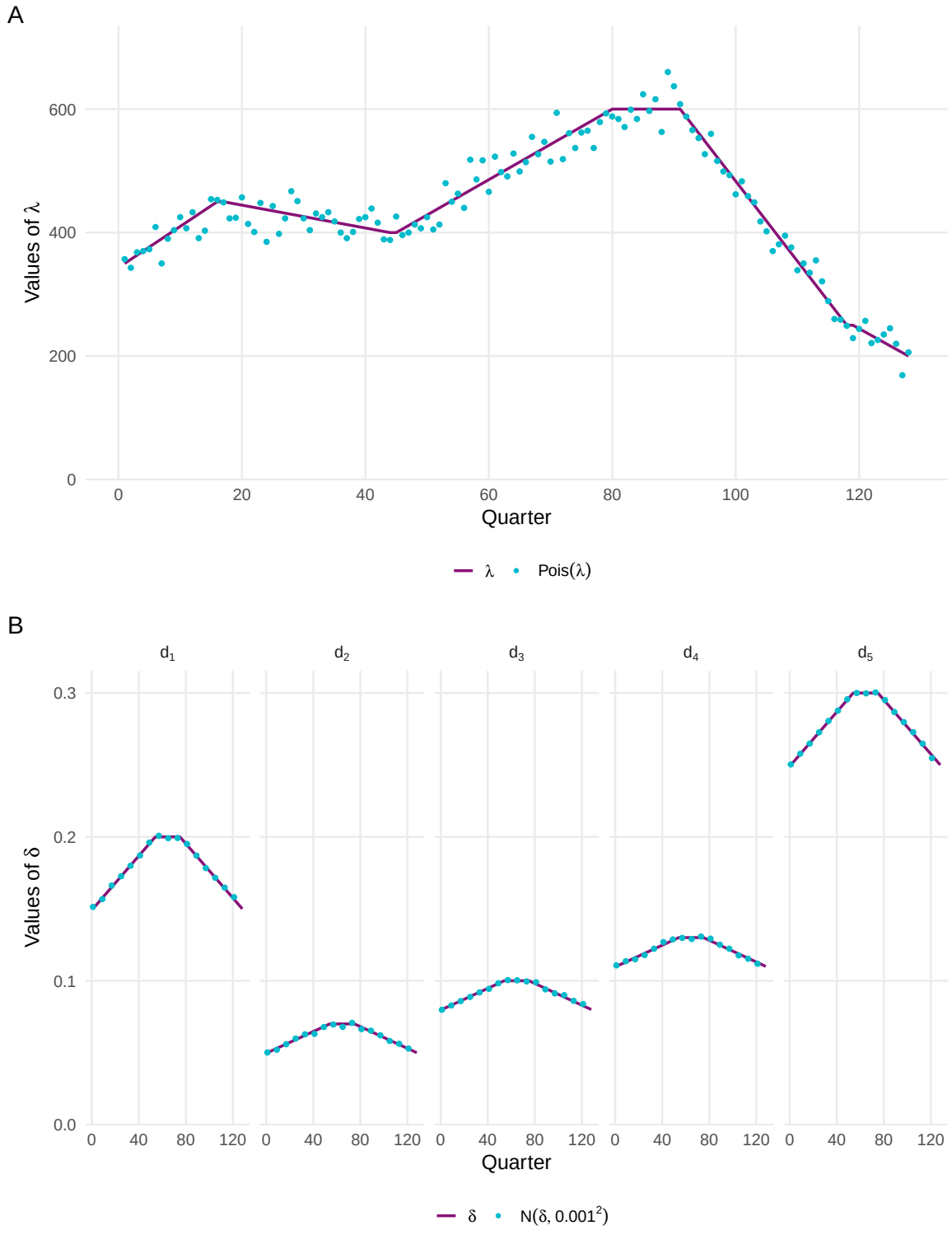


Figure 3.7: Parameter values of λ and sampled values of Pois(λ) (panel A), and parameter values of δ and sampled values of N($\delta, 0.001^2$) (panel B) from one sampling iteration.

- y_j^{R} and $y_{i,j}^{\mathrm{L}}$ were combined to derive an appropriate distribution of CD4 cell counts over time for all HIV diagnoses, $y_{i,j}^{\mathrm{C}}$, according to:

$$y_{i,j}^{\mathrm{C}} = \rho_i y_j^{\mathrm{R}} + y_{i,j}^{\mathrm{L}}$$
$$\boldsymbol{\rho} = \{0.59, 0.27, 0.11, 0.03\}$$

(iii) Reclassifying all recently acquired HIV diagnoses to the 500+ CD4 stratum:

- the number of HIV and AIDS diagnoses over time were y_j^{H} and y_j^{A}, respectively;
- y_j^{R} and $y_{i,j}^{\mathrm{L}}$ were combined to derive an appropriate distribution of CD4 cell counts over time for all HIV diagnoses, $y_{i,j}^{\mathrm{C}}$, according to:

$$y_{i,j}^{\mathrm{C}} = \begin{cases} y_j^{\mathrm{R}} + y_{i,j}^{\mathrm{L}} & \text{for } i = 1 \\ y_{i,j}^{\mathrm{L}} & \text{for } i > 1 \end{cases}$$

Estimands

The estimands compared between the models were the posterior median and 95% credible intervals of HIV incidence, HIV diagnosis probabilities, undiagnosed HIV prevalence, and the posterior predictive median and 95% credible intervals for the number of HIV diagnoses.

Methods

Back-calculation models were fitted to the each of the simulated datasets to reconstruct the underlying HIV incidence and HIV diagnosis probabilities, with prior distributions as specified in earlier sections. The dual biomarker model was fitted to the datasets simulated in scenario (i), and the CD4-only model was fitted to the datasets simulated in scenarios (ii) and (iii). The specified estimands were generated by averaging over the posterior estimates from the 100 model fits in each scenario. Results for scenarios (i) and (ii) are shown below, with results for scenario (iii) included in Appendix C.4

To investigate the effect of missing data, scenarios (i) and (ii) were re-simulated with the CD4 and RITA information missing for 40% of diagnoses. The dual biomarker and CD4-only model were again fitted to these scenarios, results are included in Appendix C.5.

Models were implemented in a Bayesian framework in the `R` and `STAN` software (Carpenter *et al.* 2017; R Core Team 2020). Models were run with 4 chains for 2000 MCMC iterations each and inspection of trace plots and the $\hat{R}$ convergence statistics used to assess convergence (Gelman and Rubin 1992).

The bias and predictive mean squared error (PMSE) of the posterior mean HIV incidence were evaluated for each scenario. The bias is the expected difference between the parameter

value specified for the data generating process, λ_j, and the estimated posterior mean incidence, $\hat{h}_{i,j}$, at each time point, $j = 1, \ldots, N$, over the 100 simulation iterations, $i = 1, \ldots, 100$:

$$\text{Bias}(\hat{h}_j, \lambda_j) = \frac{1}{100} \sum_{i=1}^{100} \hat{h}_{i,j} - \lambda_j$$

The PMSE is the expectation of the squared differences over the 100 simulation iterations:

$$\text{PMSE}(\hat{h}) = \frac{1}{100} \frac{1}{N} \sum_{i=1}^{100} \sum_{j=1}^{N} (\hat{h}_{i,j} - \lambda_{i,j})^2$$

where N is the total number of simulated time points. A lower PMSE indicates that the parameter values are more accurately estimated. The bias and PMSE of the posterior mean HIV incidence are reported for each scenario, and the distribution of the differences and squared differences over the 100 simulations is shown.

Performance measures

Figure 3.8 panel A shows posterior HIV incidence estimates for the dual biomarker model and CD4-only model and fitted to scenarios (i) and (ii). Despite increased precision, the estimates from the dual biomarker model were somewhat more erratic.

Both models were able to reconstruct the trend in undiagnosed HIV prevalence, as shown in Figure 3.8 panel B. However, the posterior estimates from the CD4-only model were a significant overestimate compared to the specified parameter values, which were better reconstructed by the dual biomarker model.

Both models reconstructed the trend in diagnosis probabilities, as shown in Figure 3.8 panel C. Whereas, the CD4-only model over-estimated the diagnosis probabilities in each strata, the dual biomarker model estimates were in line with the specified diagnosis probability parameter values.

Figure 3.9 shows the number of HIV diagnoses derived from the specified parameter values alongside the posterior predictive estimates from the CD4-only and dual biomarker models. For both models the predictive distributions fitted the data with comparable precision.

Figure 3.10 panel A shows the distribution of differences and the bias for each model. On average, over the 100 model fits, there was smaller bias in the dual biomarker model (bias of -3.28) compared to the CD4-only model (bias of -4.48). Figure 3.10 panel B shows the distribution of squared differences and PMSE for each model. The dual-biomarker model had a lower PMSE than the CD4-only model, 340.4 vs. 717.6, indicating a more accurate estimate of the parameter values.

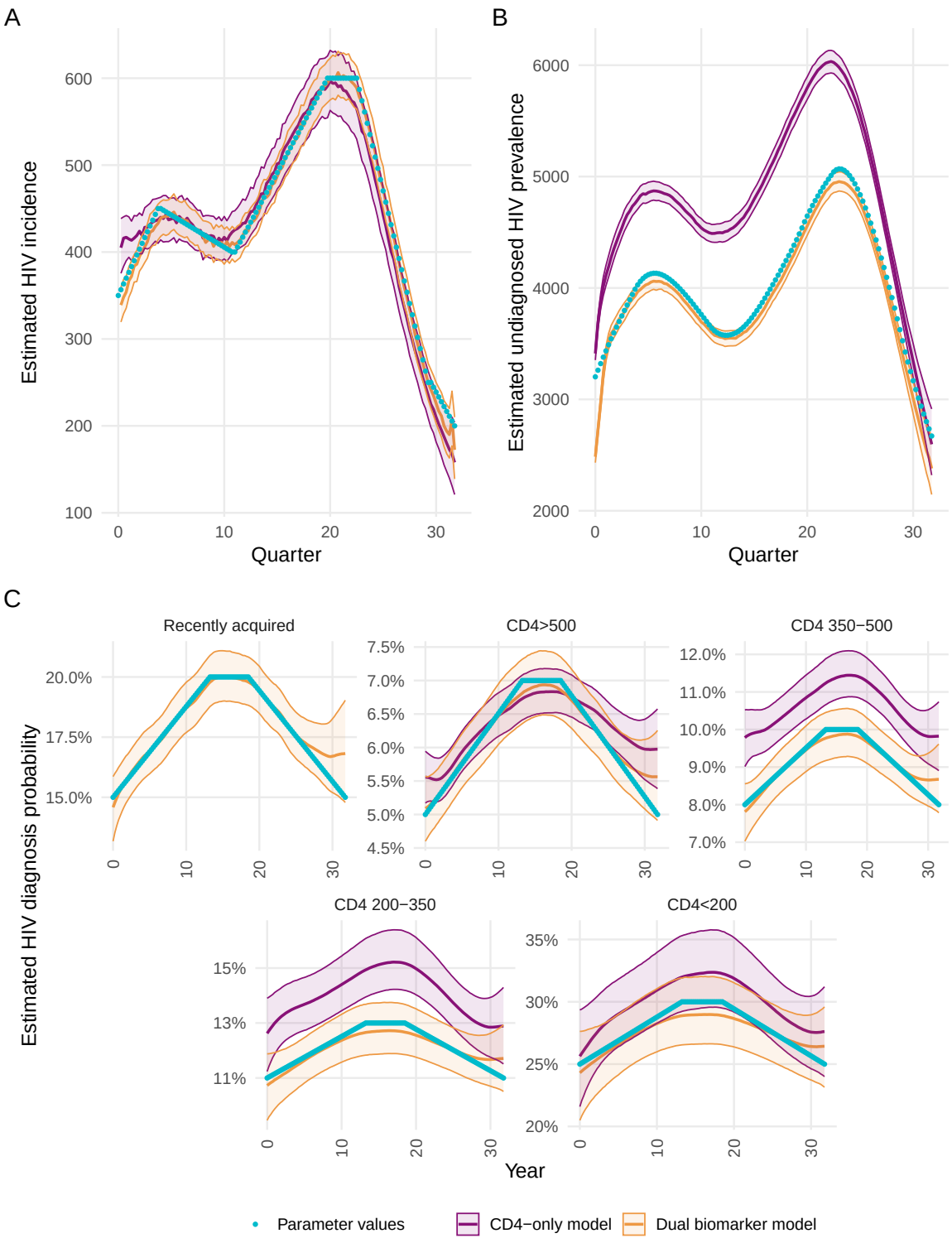


Figure 3.8: Estimated posterior median and 95% CrI of HIV incidence (panel A), undiagnosed prevalence (panel B), and diagnosis probabilities (panel C), averaged over 100 model fits, compared to the specified parameter values, by back-calculation model.

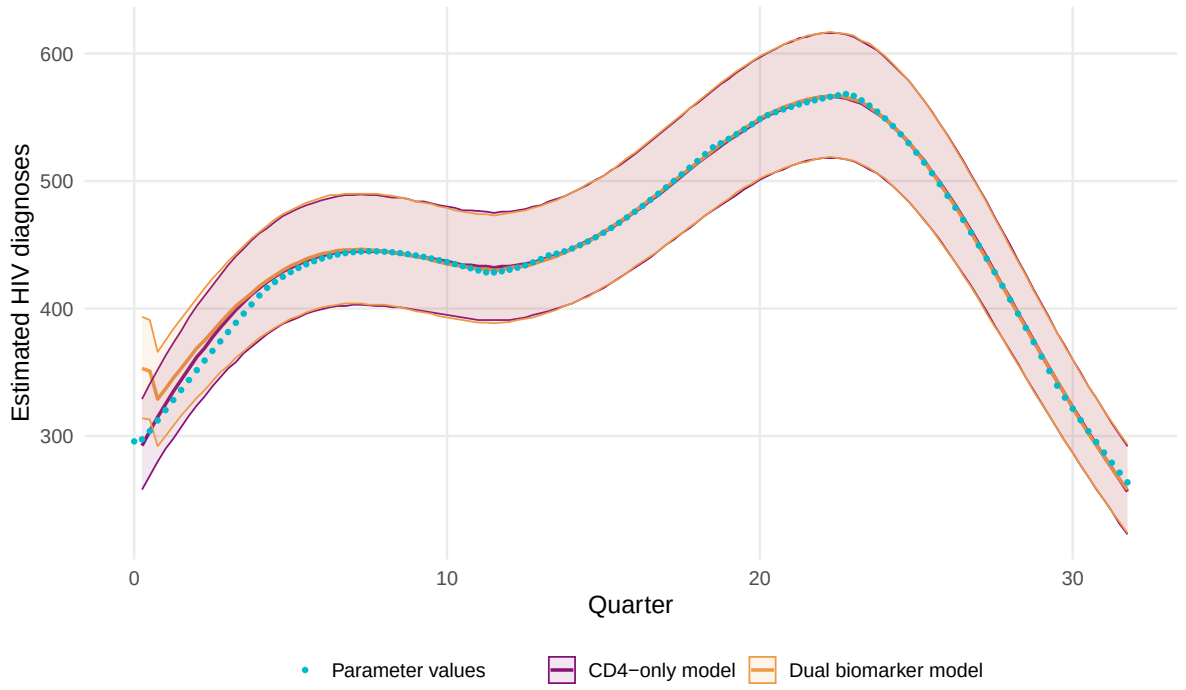


Figure 3.9: Estimated HIV diagnoses (posterior predictive median and 95% prediction interval) averaged over 100 model fits, compared to those derived from the specified parameter values, by back-calculation model.

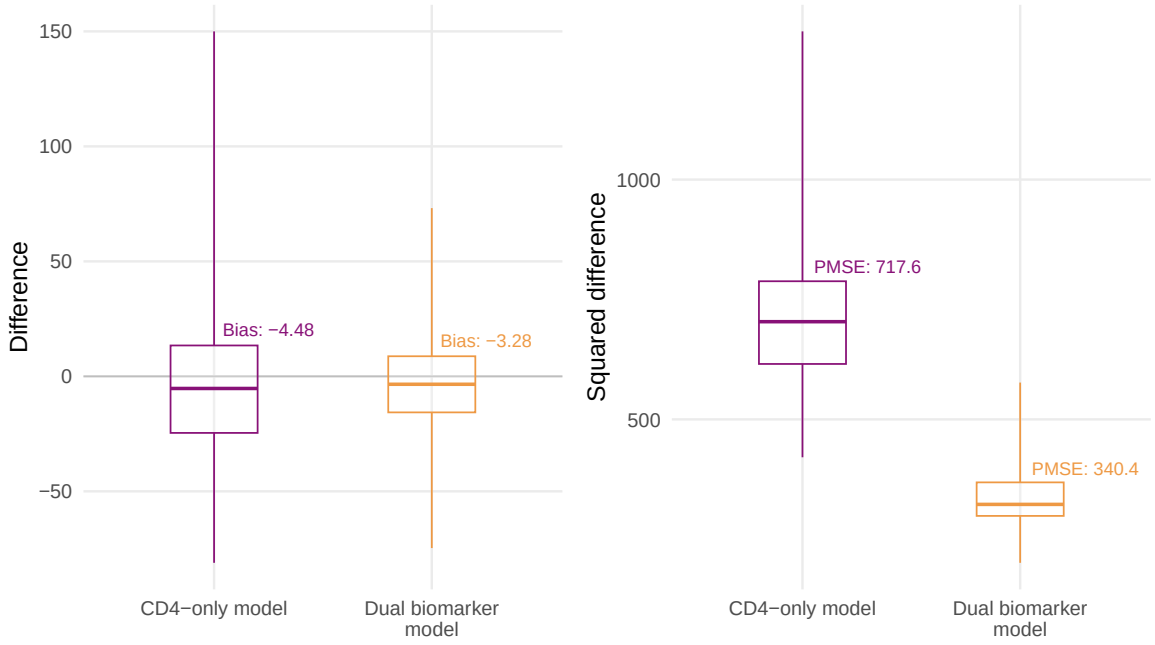


Figure 3.10: Distributions of differences (panel A), and squared differences (panel B) between parameter value and estimated HIV incidence, by back-calculation model. Estimates from 100 model fits. Boxplots show median, inter-quartile range, and range of estimates for each model.

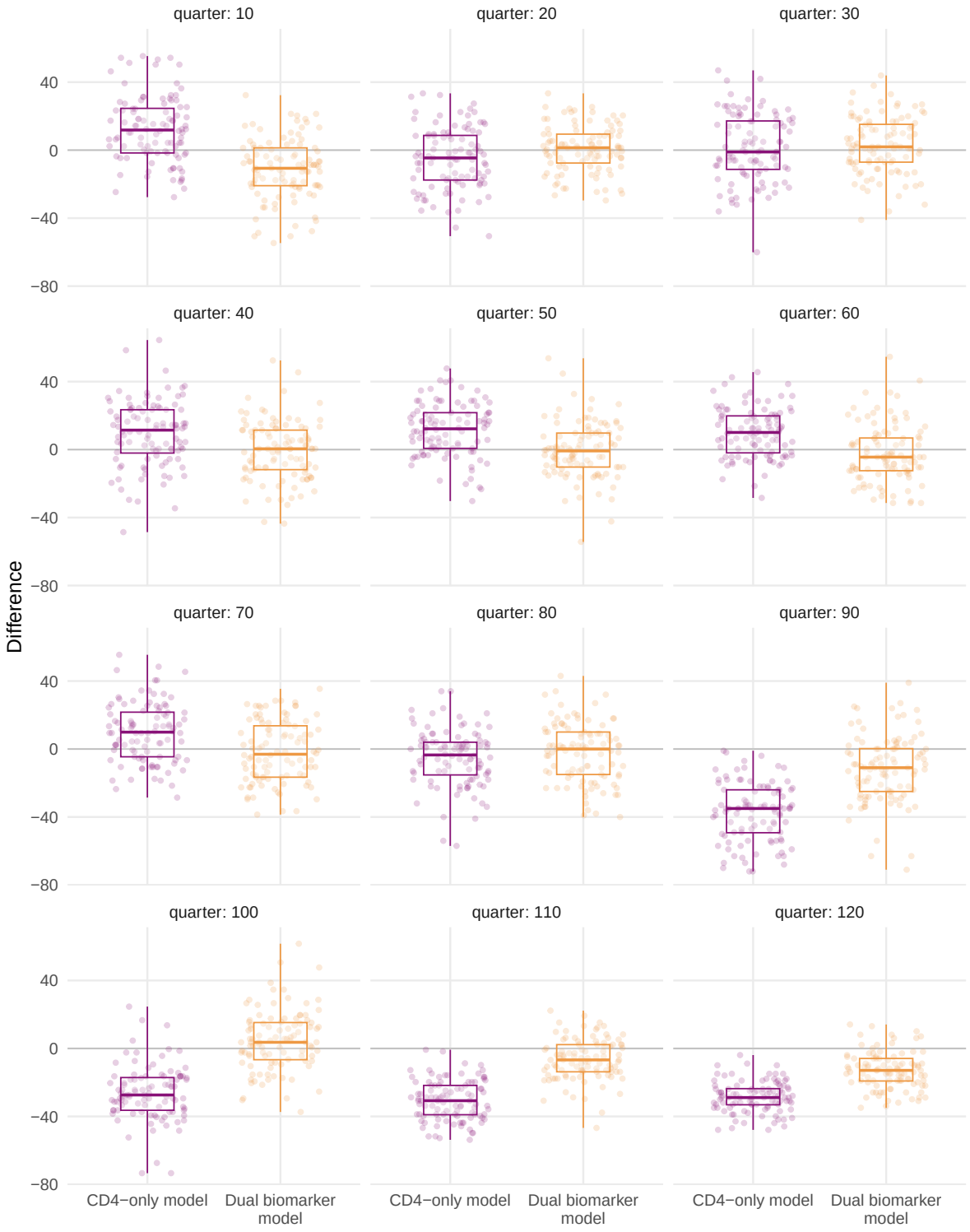


Figure 3.11: Distribution of differences between specified parameter value and estimated HIV incidence for selected quarters, by back-calculation model. Estimates from 100 model fits. Boxplots show median, inter-quartile range, and range of estimates for each model.

Figure 3.11 shows the distribution of differences between the specified parameter value and the estimated HIV incidence for selected quarters, by model. The distributions varied by quarter, although the differences for the dual biomarker estimates tended to have smaller inter-quartile ranges.

3.5.3 Application to national surveillance data

The CD4-only and dual biomarker models were applied to HIV surveillance data for GBM newly diagnosed in England. As is standard practice, GBM born abroad and with a known prior diagnosis abroad were excluded from the analysis since they were unlikely to have been incident cases in England.

Figure 3.12 panel A shows posterior estimates of HIV incidence. Similar trends over time were estimated by both models, with slightly elevated median incidence between 2014–2017 in the CD4-only model. Estimates from the dual biomarker model had narrower 95% credible intervals than the CD4-only model, and this continued up to the most recent quarter of data. For annual HIV incidence estimates, shown in Table 3.1, again the dual biomarker model estimates had narrower 95% credible intervals.

Table 3.1: Estimated annual incidence (posterior median and 95% CrI) of HIV among GBM in EW&NI between 2012–2019, by model.

	Model	
	CD4-only	Dual biomarker
Year		
2012	2540 (2280, 2960)	2400 (2180, 2640)
2013	2740 (2360, 3130)	2570 (2320, 2850)
2014	2270 (1980, 2570)	2410 (2200, 2700)
2015	1670 (1380, 1940)	1900 (1640, 2140)
2016	1300 (1030, 1600)	1560 (1340, 1800)
2017	1150 (900, 1440)	1290 (1070, 1580)
2018	1070 (760, 1440)	1060 (810, 1410)
2019	1020 (540, 1790)	960 (630, 1570)

Undiagnosed HIV prevalence estimates from the dual biomarker model were significantly lower than those estimated by the CD4-only model for the years 2011–2017, with some overlap in more recent years (Figure 3.12 panel B). Meanwhile, HIV diagnosis probability estimates were broadly similar for both models across most CD4 strata, with additional information about recently acquired infections from the dual biomarker model (Figure 3.12 panel C).

Figure 3.13 shows that both models were a good fit to the observed diagnosis data, with the 95% predictive intervals containing almost all of the data points.

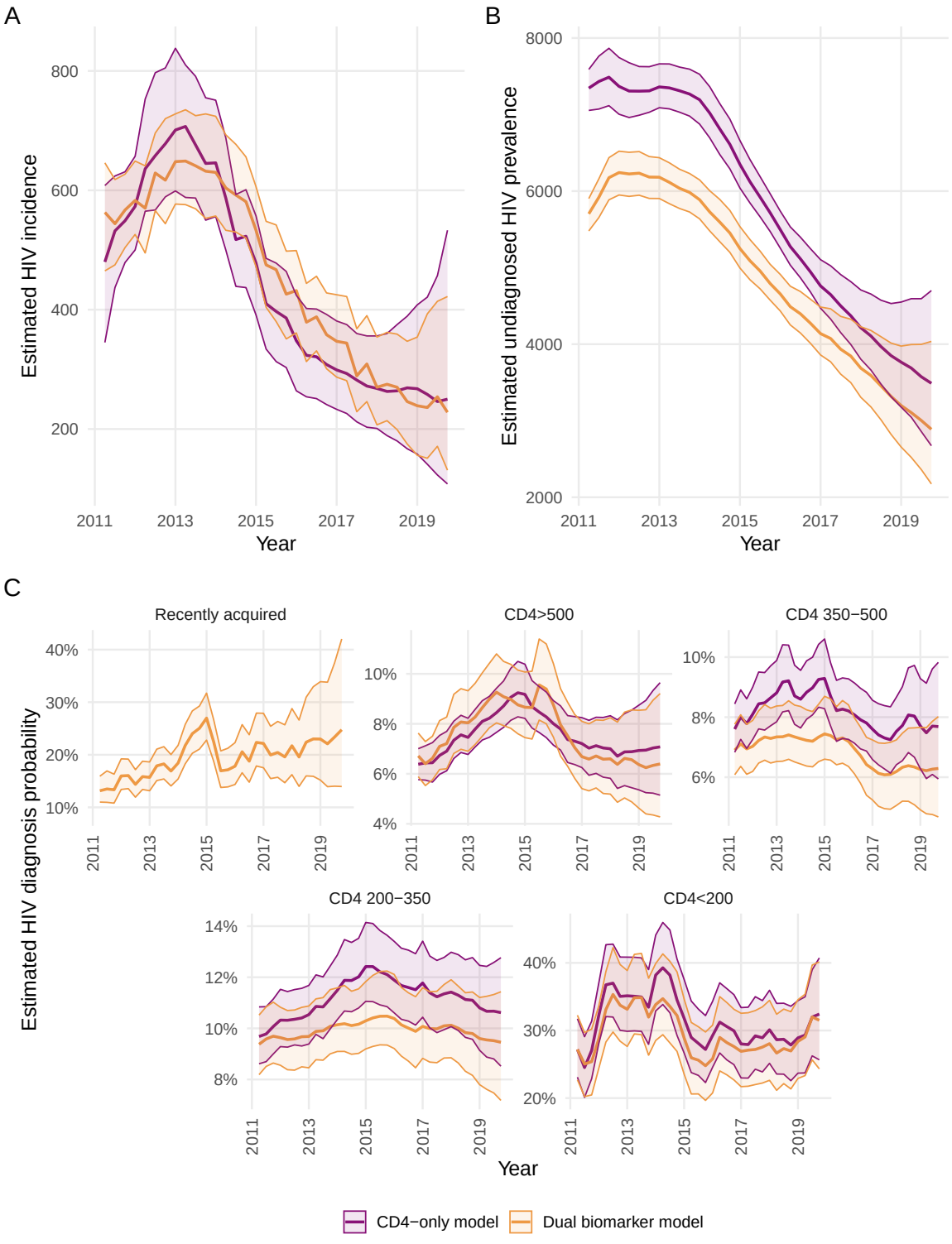


Figure 3.12: Estimated posterior median and 95% CrI of HIV incidence (panel A), undiagnosed prevalence (panel B), and diagnosis probabilities (panel C) fitted to HIV diagnosis data for GBM diagnosed in England, 2011–2019.

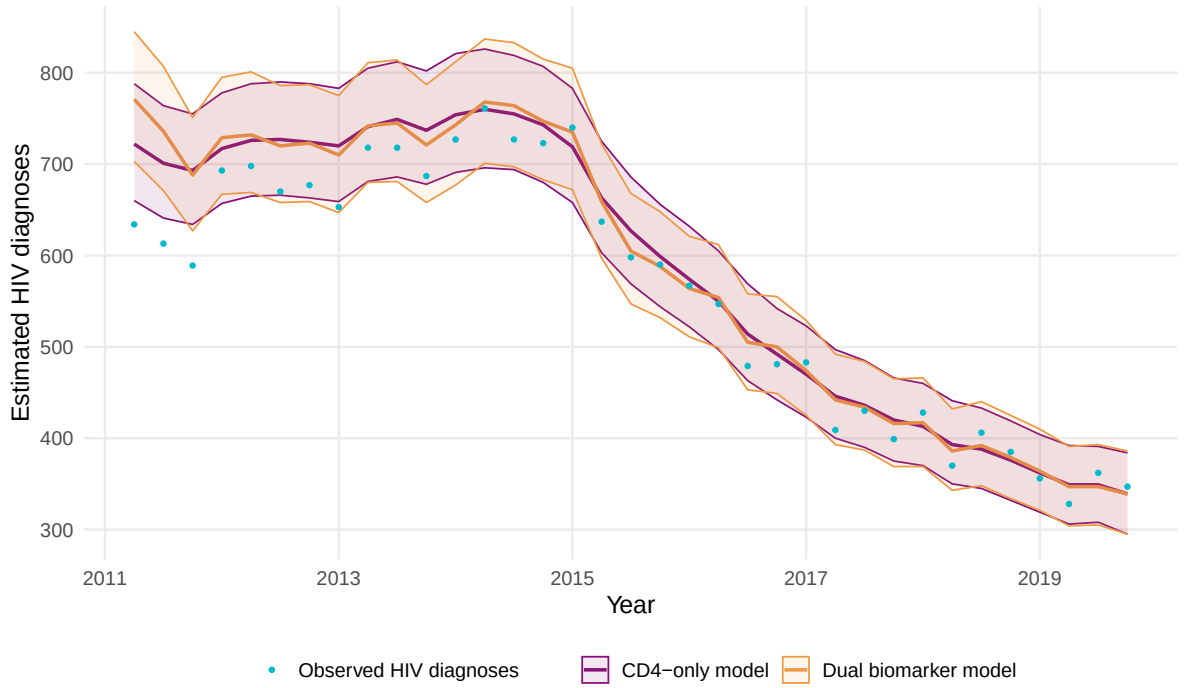


Figure 3.13: Estimated HIV diagnoses (posterior predictive median and 95% prediction interval), compared to observed HIV diagnosis data for GBM diagnosed in England, 2011–2019.

3.5.4 Conclusions

The CD4-only and dual biomarker model were used to reconstruct trends in HIV incidence, undiagnosed prevalence, probability of diagnosis, and quarterly numbers of diagnoses for simulated scenarios and real-world surveillance data.

In the simulation study, all relevant estimates from the dual biomarker model were in-line with the specified parameter values, and accurate trends could be inferred from the estimates. The CD4-only model gave accurate, although somewhat less precise, estimates of HIV incidence, but over-estimated the HIV diagnosis probabilities and undiagnosed HIV prevalence as compared to the specified parameter values. This lack of fit of the CD4-only model is not surprising, since the model does not account for diagnosis during the acute stage.

Estimates of HIV incidence, the key quantity of interest in current back-calculation models, were similar for both models. However, the dual biomarker model estimates were closer to the simulated values on average, and estimated with greater precision.

By applying these models to surveillance data for GBM newly diagnosed in England, trends could be estimated for unobserved quantities. As for the simulation study, whilst trends were similar in both models, the dual biomarker model gave more precise estimates of HIV incidence and a significantly lower estimate of undiagnosed HIV prevalence before 2017. Based on the simulation study, and knowing that diagnosis occurring during the acute and seroconversion

phases can lead to low CD4 counts, the estimates from the dual biomarker model are likely to be more accurate estimates of these unobserved quantities.

Whilst comparison of these estimates with a 'true' unobserved quantity is not possible for real-world data, the back-calculation estimates can be compared with estimates obtained via different methods. The MPES model has provided estimates of undiagnosed HIV prevalence among GBM in England since 2013 (Presanis, Harris *et al.* 2021). As shown in Figure 3.14, undiagnosed prevalence estimates from both the CD4-only and dual biomarker back-calculation models fall within the 95% credible interval from MPES, with the median MPES estimate being closer to the dual biomarker model, especially in more recent years. The wider credible range in MPES compared to the back-calculation reflects uncertainty in several of the MPES data sources which are not considered in the back-calculation model.

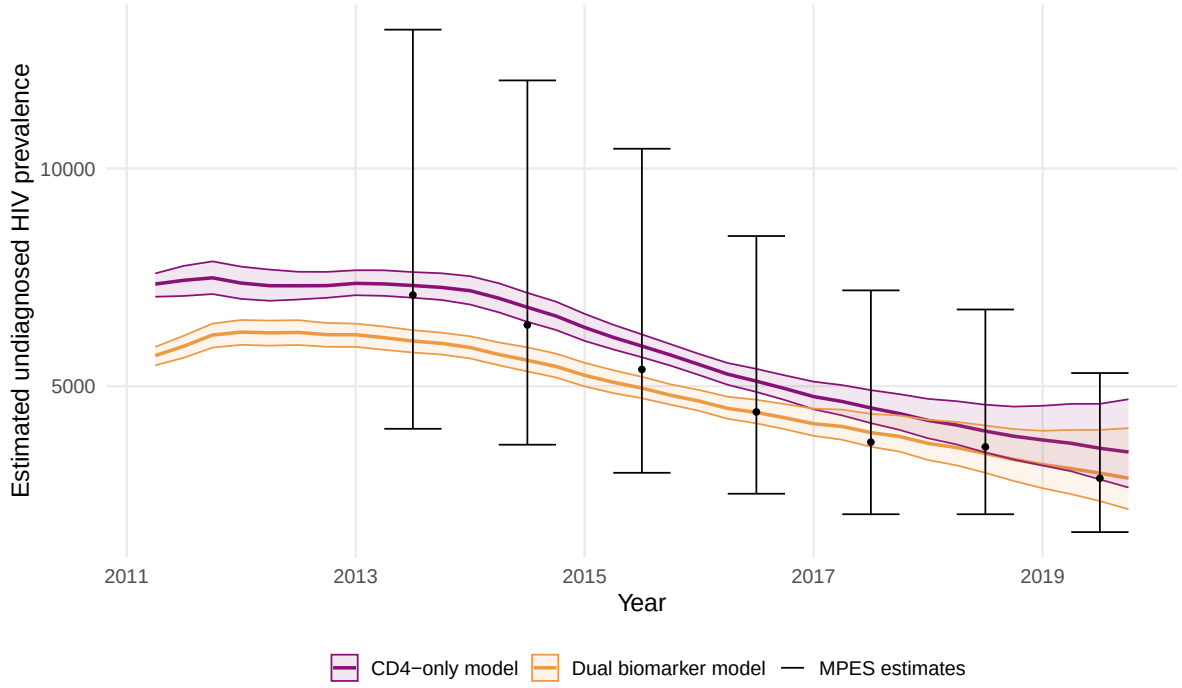


Figure 3.14: Estimated undiagnosed HIV prevalence (posterior median and 95% CrI) from dual biomarker and CD4-only back-calculation, compared to MPES estimates for GBM diagnosed in England, 2011–2019.

In addition to improved estimate precision and accuracy, the dual biomarker model provides additional information on the probability of diagnosis among individuals with recently acquired HIV. These estimates are of interest for epidemiologists as trends can be correlated with HIV prevention programmes. In the English estimates for 2011–2019, there was a spike in the probability of being recently diagnosed during 2015, followed by a slower rise to the end of 2019. The 2015 spike coincides with a rapid increase in testing and in diagnoses of recently

acquired HIV at a major combined sexual health and HIV service in an area of central London known for its high concentration of LGBTQI venues (Girometti *et al.* 2021).

Unsurprisingly, given the small value of the progression probability q_1, a simple reclassification of recent diagnoses shown in Appendix C.4 yielded very similar estimates to the dual biomarker model, with minimal loss of precision. The dual biomarker model remains the more accurate method to incorporate additional biomarker data, reflecting the 6-month MDRI and providing valuable information about HIV diagnosis probability in the early infection period.

Limitations

The methodology presented here makes use of a single biomarker classification (recent/non-recent infection) although does not address false recency error in this classification, which was previously estimated as 1.9% (95% CI 1.0–3.4%) (Aghaizu 2018). Accounting for this error may lessen the change in undiagnosed HIV prevalence, although the difference is likely to be small.

As previously noted, there were regional differences in the availability of HIV avidity information, with particularly low availability in Wales. HIV testing policies also vary between different devolved nations which results in differences in the proportion of recent diagnoses (Kirwan, Croxford *et al.* 2022). To mitigate these regional biases, and for consistency with the HIV action plan for England and HIV estimation methods applied annually (Department of Health and Social Care 2021b; Martin, Okumu-Camerra *et al.* 2023), the models were applied only to surveillance data for England. There was also observed to be some age-dependency in the availability of the recency assay data which was not accounted for. Whilst the effect on the model estimates is likely to be small, development of the age-dependent back-calculation model to include avidity information may help to more formally address this potential bias.

GBM diagnosed with HIV in England are an increasingly diverse group, with 34% of those newly diagnosed in 2022 born abroad and potentially having acquired HIV outside of England (Martin, Okumu-Camerra *et al.* 2023). However, a limitation of all currently existing models is the assumption of a closed system which does not account for external processes such as immigration and emigration. Currently this is addressed by excluding GBM born abroad and first diagnosed outside of the UK from the diagnosis dataset. However, the method remains unsuited for use among heterosexual groups due to the large proportion born abroad, and may become unreliable as the GBM population becomes increasingly diverse.

This chapter has developed the CD4 back-calculation model into a dual biomarker model, taking account of biases in the data and improving the precision and accuracy of HIV incidence estimates for England. Avenues for further development, for example including age-specific back-calculation and the incorporation of migration data are discussed further in Chapter 6.

3.6 Impact of COVID-19 lockdown on HIV

3.6.1 Background

The analyses in this chapter have so far considered HIV diagnosis data prior to 2020. In 2020 the epidemiology of many infectious diseases was severely affected by the COVID-19 pandemic, with changes to healthcare services and so-called 'lockdown' strategies and other measures being implemented worldwide to control the spread of COVID-19 (Institute for Government 2022). The emergency response to COVID-19 likely impacted transmission (due to restrictions on social mixing) and testing activity (due to services being reconfigured and people being less able to access testing) for a multitude of non-COVID-related infectious diseases (Mude *et al.* 2023; Simões *et al.* 2020). Whilst healthcare access and social mixing have now largely returned to pre-pandemic levels, understanding the effect of an abrupt change in testing and diagnosis for other conditions requires further research.

One concern is that existing methodologies which rely on time-series data to uncover trends in infectious disease may not be sufficiently sensitive to account for such sudden changes, or may incorrectly attribute their cause. Other analyses of English HIV diagnosis data, using a synthetically-controlled Bayesian structural time series model to infer the causal impact of the first lockdown in England, estimated 51 missed diagnoses (95% prediction interval -310–393) among GBM during 2020 (Muhammed *et al.* 2024). However, the sensitivity of the back-calculation estimates to this drop and to different assumptions about the effect of COVID-19 lockdown is so far untested.

Estimating the specific impact of COVID-19 mitigation on HIV and other infectious diseases is important for re-evaluation of public health priorities post-pandemic. The public health response to COVID-19 diverted efforts away from England's HIV elimination agenda. Assessing if England remains on-target to eliminate HIV transmission by 2030 (as was previously estimated using the CD4 back-calculation (Brizzi, Birrell, Kirwan *et al.* 2021)), or if further investment is required (as has been recently suggested by a stochastic simulation-based approach (Cambiano *et al.* 2023)), is vital to inform ongoing HIV prevention efforts.

3.6.2 Aims

This section aims to assess the sensitivity of HIV estimates for 2020 and 2021 to different lockdown effect assumptions, to evaluate the impact of COVID-19 on HIV incidence, and to assess if England remains on track to eliminate HIV as a public health threat by 2030. While not a formal causal inference study, similar sensitivity or stochastic scenario analyses have been conducted for influenza as a type of 'counterfactual' analysis (Cooper, Pitman *et al.* 2006; Wu *et al.* 2017).

3.6.3 COVID-19 lockdown effect assumptions

Data

The HIV surveillance data used in these analyses relate to GBM first diagnosed in England, with GBM born abroad and with a known prior diagnosis abroad being excluded.

This analysis considers changes in HIV transmission and probability of diagnosis during the 2-year period post-lockdown. Information on the timing of national and local COVID-19 lockdowns in England were available from the Institute for Government (Institute for Government 2022). Legal COVID-19 lockdown measures instructing people to stay at home were introduced in England on 26 March 2020, i.e. around the start of the second calendar quarter of 2020 (2020Q2).

Figure 3.15 shows quarterly HIV diagnoses among GBM in England before and after lockdown restrictions were introduced. A reduction in the number of new HIV diagnoses was particularly apparent during 2020Q2 (April to June) and 2020Q3 (July to September). Diagnoses recovered somewhat during 2020Q4 (October to December) and 2021Q1 (January to March), but were lower again during 2021Q2 to 2021Q4.

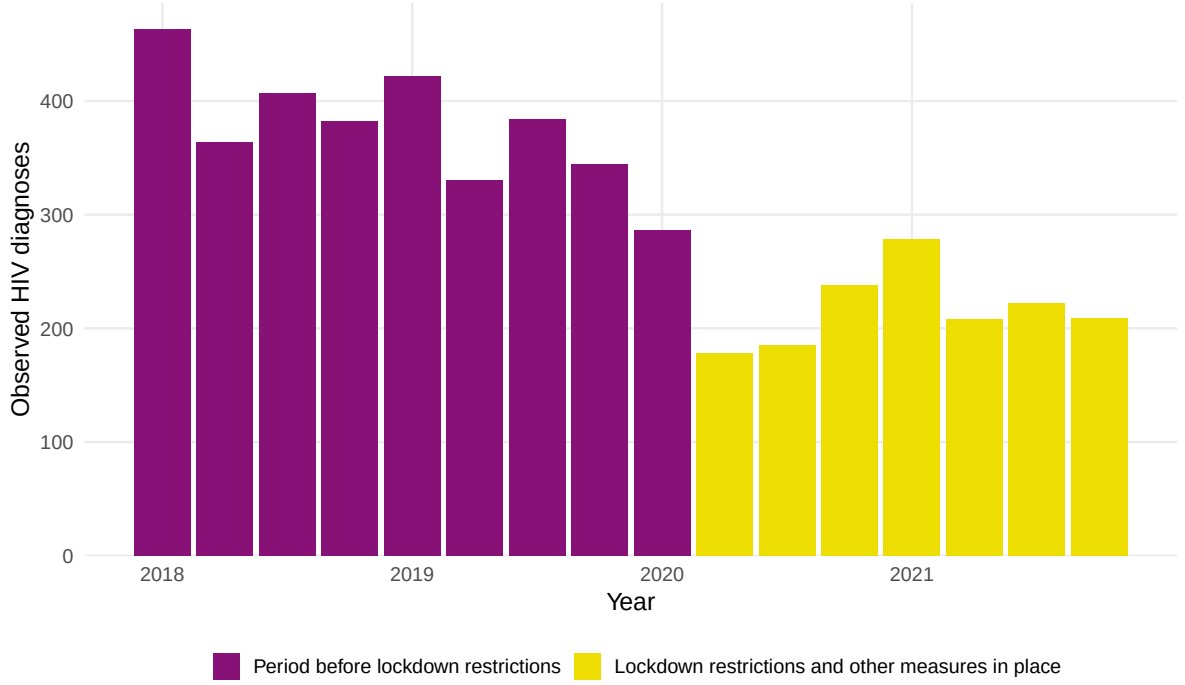


Figure 3.15: Quarterly number of HIV diagnoses among GBM diagnosed in England, before and after introduction of lockdown restrictions, 2018–2021.

Methods

Observed HIV and AIDS diagnoses in the CD4 back-calculation are modelled as realisations of thinned Poisson processes from an unobserved, time-varying, incidence process, h_j. Expected HIV and AIDS diagnoses, $\mu_{i,j}$, are related to this incidence process via prevalent latent infections, $e_{i,j}$, known disease progression probabilities, q_i, and estimated diagnosis probabilities, $d_{i,j}$. Figure 3.16 shows the relationship between underlying HIV infections, HIV diagnosis probabilities, and HIV diagnoses in a schematic diagram.

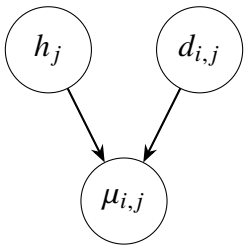


Figure 3.16: Schematic diagram linking HIV infections h_j and HIV diagnosis probabilities $d_{i,j}$ to HIV diagnoses $\mu_{i,j}$.

In the CD4 back-calculation model, estimates of the contributions of the unknown incidence (h_j) and diagnosis probability ($d_{i,j}$) parameters to the expected diagnoses depend on all of: the diagnosis data, the CD4 count data, the prior distributions (including point priors) on the progression rates (q), and the incidence and diagnosis probabilities themselves. To assess the sensitivity of the estimates to different assumptions about the effect of lockdowns on the h_j and $d_{i,j}$ parameters, the model can be modified with additional prior information about h_j and $d_{i,j}$.

Three model scenarios were considered:

- the unmodified model, with any disruption in diagnosis data taken at face-value and apportioned to the h_j and $d_{i,j}$ parameters.

(a) a scenario where the incidence of HIV was unaffected by the lockdowns and remained the same as in 2019, whereas probability of HIV diagnosis was allowed to be affected. In this scenario, the prior for h_j in the calendar quarters $j \in \{2020\text{Q}2, \ldots, 2021\text{Q}4\}$ was specified as $h_j \sim \text{Poisson}(\lambda_{2019\text{Q}k})$, where $\lambda_{2019\text{Q}k}$ is the HIV incidence in the kth quarter of 2019. This approximated HIV incidence in 2020 and 2021 remaining at the same level as 2019.

(b) a scenario where the probability of HIV diagnosis was unaffected by the lockdowns and remained unchanged from 2019, whereas incidence was allowed to be affected. In this scenario, the prior for $d_{i,j}$ in the calendar quarters $j \in \{2020\text{Q}2, \ldots, 2021\text{Q}4\}$ was specified as $d_{i,j} \sim \text{N}(\mu_{i,2019\text{Q}k}, 0.001^2)$, where $\mu_{i,2019\text{Q}k}$ is the diagnosis probability in the

ith CD4 stratum and kth quarter of 2019. This approximated the probability of diagnosis in 2020 and 2021 in each CD4 stratum remaining at the same level as 2019.

The unmodified model and scenarios (a) and (b) were each fitted to the observed HIV diagnosis data for GBM newly diagnosed in England between 2011–2021. Models were run with 4 chains and for 4000 MCMC iterations, with inspection of trace plots and the $\hat{R}$ convergence statistics used to assess convergence (Gelman and Rubin 1992).

Visual comparison and leave-one-out cross-validation (LOO-CV) using Pareto smoothed importance sampling were used to compare models (Vehtari *et al.* 2024). The LOO Information Criterion (LOOIC) is similar to the frequentist Akaike information criterion (AIC), and enables the 'best fitting' model to be identified whilst balancing parsimony and over-fitting.

Results

Figure 3.17 panels A and B show a downwards trend in estimated HIV incidence between 2016–2021 in both the unmodified model and scenario (a) (where HIV incidence was assumed unaffected by lockdown), with estimates being more precise in scenario (a). The incidence estimates in scenario (b) (where diagnosis probabilities were assumed to be unaffected by lockdown) were more uncertain, but a similar downwards trend in incidence was estimated.

During 2020 and 2021 the 95% credible intervals of annual incidence for all scenarios overlapped, with the posterior median being lowest in scenario (b) (Table 3.2).

Table 3.2: Estimated annual incidence (posterior median and 95% CrI) of HIV among GBM in EW&NI between 2012–2021, by model.

	Scenario		
	Unmodified model	(a) Incidence unaffected	(b) Diagnosis probabilities unaffected
Year			
2016	1420 (1290, 1560)	1430 (1300, 1530)	1310 (1140, 1500)
2017	1090 (960, 1210)	1080 (980, 1180)	950 (780, 1160)
2018	860 (720, 1030)	740 (660, 820)	770 (570, 1020)
2019	650 (490, 840)	550 (500, 610)	530 (330, 830)
2020	530 (360, 780)	460 (380, 530)	370 (180, 790)
2021	410 (210, 830)	360 (250, 480)	260 (80, 890)

HIV diagnosis probabilities in the unmodified model and in scenario (a) showed high sensitivity to the impact of COVID-19 lockdowns, with a reduction in the probability of diagnosis being observed across all CD4 strata (Figure 3.17 panels C and D). In scenario (b) the diagnosis probabilities were more stable (by design), but particularly uncertain after 2019.

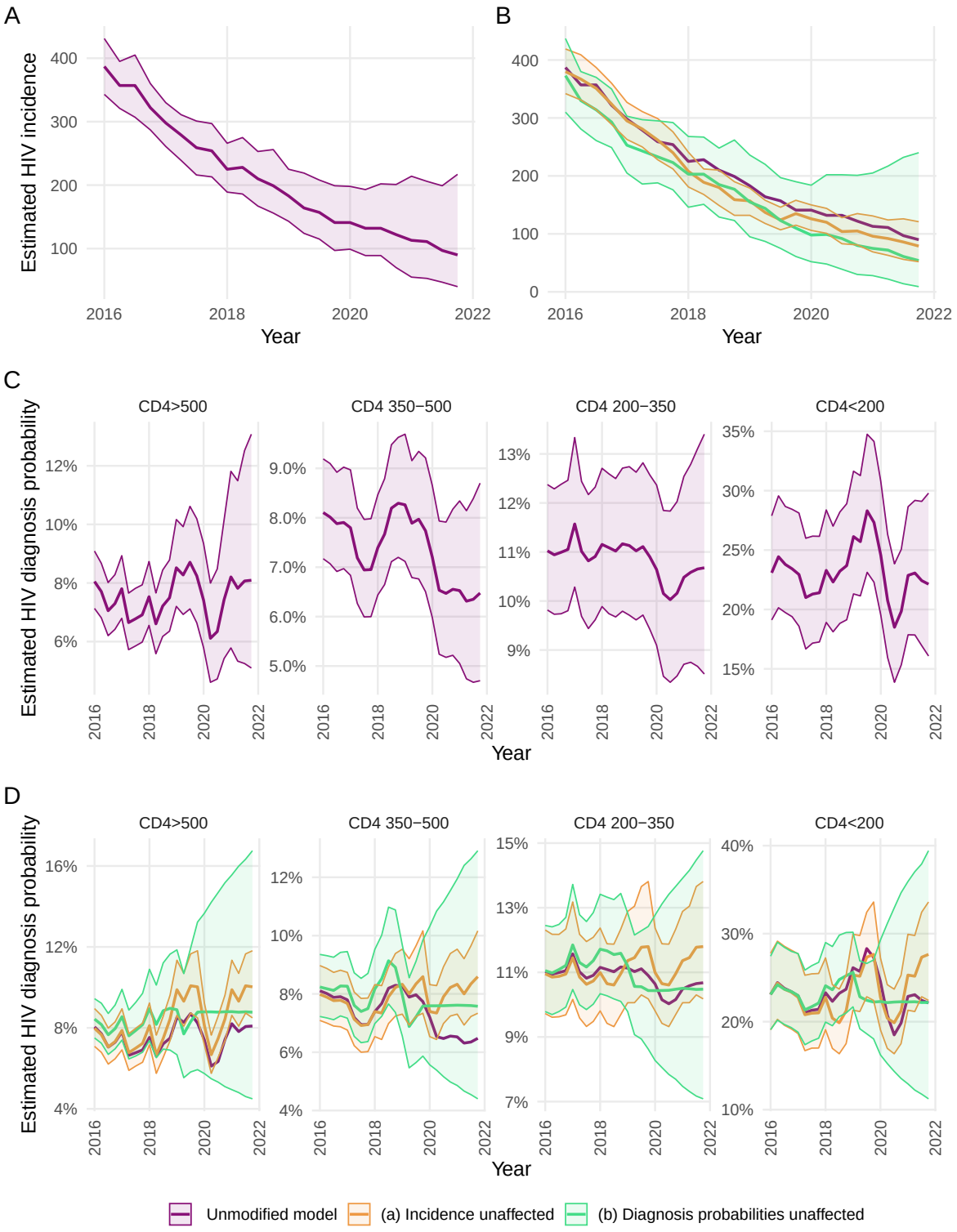


Figure 3.17: Estimated posterior median and 95% CrI of HIV incidence (panels A and B) and diagnosis probabilities (panels C and D) fitted to HIV diagnosis data for GBM diagnosed in England, 2016–2021.

The posterior predictive estimates of HIV diagnoses in the unmodified model are shown in Figure 3.18 panel A, with corresponding estimates for scenarios (a) and (b) shown in panel B. The unmodified model had the closest fit to the data visually, whilst estimates for scenario (b) were less precise.

The LOOIC and the difference in the estimated log-likelihood between models is shown in Table 3.3. Scenario (a) had the lowest LOOIC, although the difference in the log-likelihood between scenario (a) and the unmodified model was not significant. The LOOIC for scenario (b) was significantly higher than the other models indicating a poorer fit to the data.

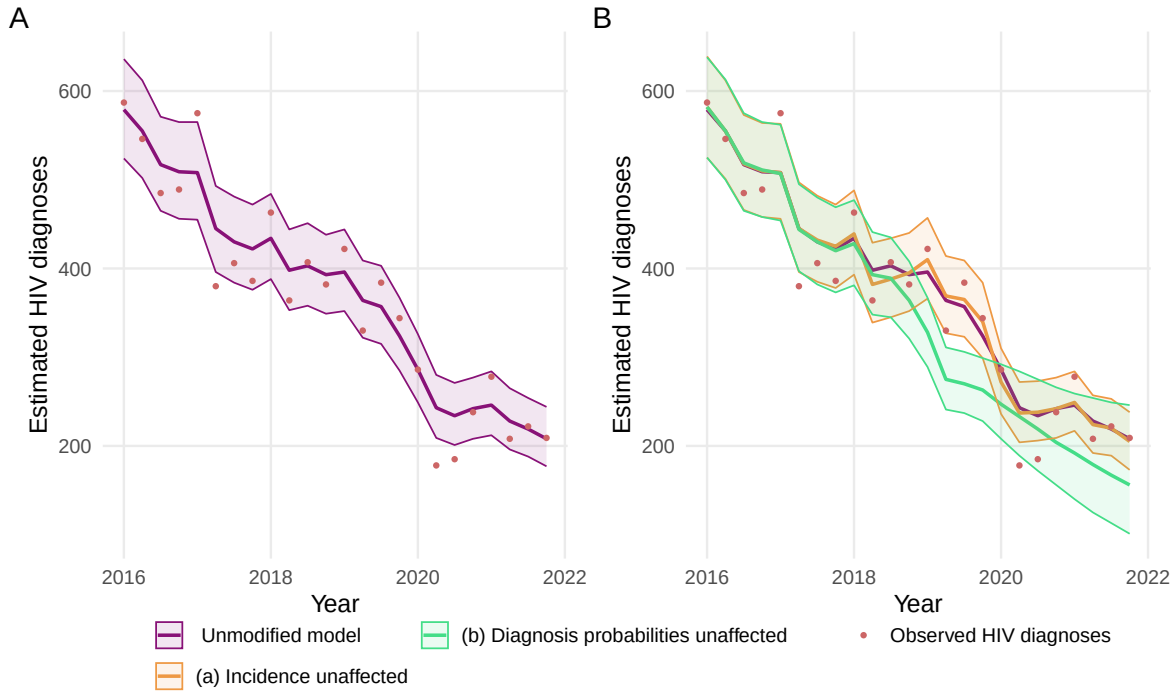


Figure 3.18: Estimated HIV diagnoses (posterior predictive median and 95% prediction interval), compared to observed HIV diagnosis data for GBM diagnosed in England, 2016–2021.

Table 3.3: LOOIC and difference in estimated log-liklihood values for unmodified model and each model scenario.

	LOOIC	Difference in estimated log-likelihood (Standard error)
Model scenario		
(a) Incidence unaffected	4262.1	0 (0)
Unmodified model	4269.8	-3.8 (8.7)
(b) Diagnosis probabilities unaffected	4495.7	-116.8 (42.6)

3.6.4 Projected HIV incidence and progress towards HIV elimination targets

Methods

Projected HIV incidence (to 2030) was previously estimated by Brizzi, Birrell, Kirwan *et al.* 2021 using an age-dependent back-calculation model fitted to data to the end of 2018. These projections were re-run in light of the COVID-19 pandemic, using the CD4-only back-calculation model fitted to data to the end of 2022 (by which time HIV testing in SHS had largely returned to pre-pandemic levels (Martin, Okumu-Camerra *et al.* 2023)).

To generate the projected incidence estimates, the incidence splines estimated by the back-calculation were extrapolated forwards using the `R` package `mgcv` (Wood 2016). Projected HIV incidence was compared to the previously published projections, and to the HIV Commission targets (HIV Commission 2020). The proportion of extrapolated incidence profiles falling below the 250 and 50 thresholds were used to assess the likelihood of reaching the respective targets by 2025 and 2030.

Results

Projections of HIV incidence based on data to the end of 2018 are shown in Figure 3.19 panel A (Brizzi, Birrell, Kirwan *et al.* 2021), with updated projections based on data to the end of 2022 shown in panel B. Extrapolating forwards using the 2022 estimates, HIV incidence in 2025 was projected to be 582 (90% CrI 125 to 2108) in 2025 and 593 (90% CrI 52 to 5198) in 2030. Based on these projections, 17.1% of the extrapolated incidence profiles for 2025 were under 250, and 4.9% of those for 2030 were under 50.

3.6.5 Conclusions

The COVID-19 lockdowns in 2020 and 2021 likely influenced both HIV transmission and testing activity. Two model scenarios with contrasting assumptions demonstrated that the back-calculation estimates were sensitive to the effect of lockdown being fixed entirely one way or the other. The results of these scenarios were included in UKHSA reports to show the potential impact of COVID-19 on HIV transmission in England (Martin, Lester *et al.* 2022).

The best fitting model was scenario (a), in which HIV incidence was unaffected by COVID-19 lockdowns, but diagnosis probabilities were affected. In this scenario HIV diagnosis probabilities were estimated to have recovered by the end of 2020. Although the the reduction in HIV testing at sexual health services persisted beyond 2020, this may have been offset by the rapid expansion in self-sampling and home HIV testing, with around two thirds of HIV

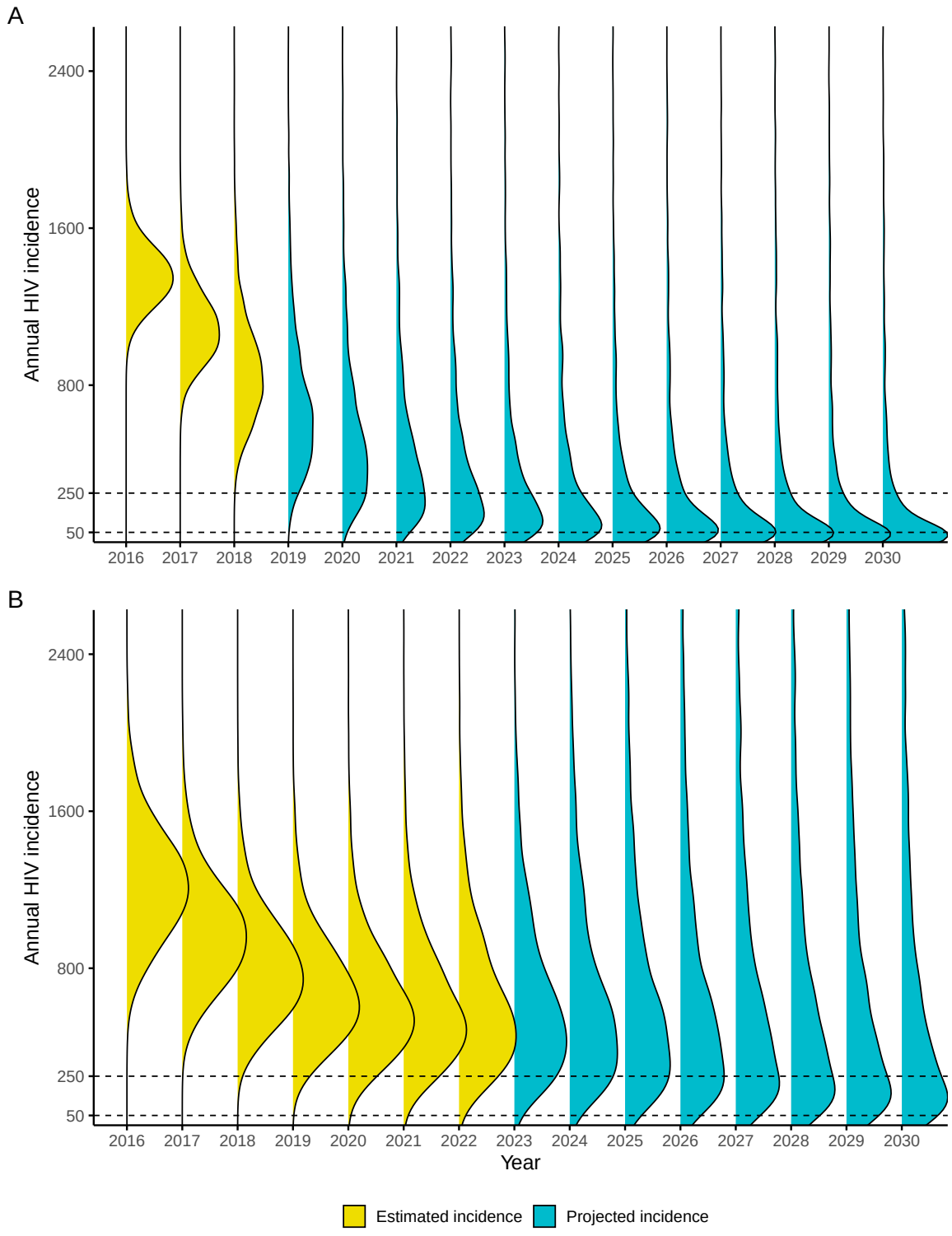


Figure 3.19: Density plot of estimated and projected annual HIV incidence among GBM in England, 2016–2030, projections based on data to the end of 2018 (panel A), and data to the end of 2022 (panel B). Dotted lines indicate the HIV commission targets for HIV incidence among GBM of <250 by 2025, and <50 by 2030 (HIV Commission 2020).

tests being accessed via internet services and conducted at home during the first national lockdown (Public Health England 2020c; Wenlock *et al.* 2022).

Although priors were specified differently for each scenario, all three models estimated a reduction in the posterior median for HIV incidence during 2020 and 2021. A continued reduction in incidence may be due to a combination of: lockdown, more widespread use of PrEP, a reduction in the number of condomless sex partners (Pebody 2021a), and a high proportion of those living with HIV being unable to pass it on due to undetectable levels of virus (World Health Organization 2023).

A formal method of causal inference for a multi-state model is described in Chapter 5, whereby elements of the transition matrix are fixed to pre-specified values to investigate the effectiveness of an intervention. The model scenarios presented in this section are not a formal causal analysis, but there are conceptual similarities: instead of fixing the elements of the transition matrices, strong prior information is placed on them, and the estimates corresponding to these different counterfactual scenarios are compared.

In previous projections, using data to the end of 2018, around 60% of extrapolated incidence profiles reached the <250 target in 2025, and around 40% reached the <50 target in 2030 (Brizzi, Birrell, Kirwan *et al.* 2021). Projecting forwards from 2022 suggests that England is no longer on track to eliminate HIV transmission among GBM by 2030: under 17% and under 5% of extrapolated incidence profiles met the 2025 and 2030 targets, respectively. Estimates from the corresponding age-dependent model were more uncertain.

Whilst credible intervals overlap, the median projections for 2025 and 2030 from the back-calculation model are somewhat above those predicted by a simulation model which incorporates rates of testing, ART, and PrEP usage (Cambiano *et al.* 2023). There is potential for future changes in HIV prevention and testing initiatives to accelerate the decline in HIV transmission, and progress towards the 2030 target should continue to be measured.

Limitations

Some lack of fit to the HIV diagnosis data was observed in all three models. Incorporating additional data sources, such as behavioural surveys and testing data around the time of lockdown, may help to to improve the precision and identifiability of these estimates.

As well as limiting the number of diagnoses occurring in 2020 and 2021, service disruption as a result of COVID-19 lockdowns may have led to under-reporting of new HIV diagnoses. UKHSA HIV surveillance teams cross-validate the HIV diagnosis data with an HIV commissioning database which collects service information among all people accessing care, and follow-up on individuals with missing diagnosis records. In the 2021 data archive, data quality was not considered to have been particularly affected by under-reporting as compared to the pre-COVID-19 pandemic period (C. Chau, personal communication).

As before, the back-calculation is not currently suitable for estimating HIV incidence as a result of heterosexual HIV transmission, and therefore not able to provide estimates of the impact of the COVID-19 pandemic on overall HIV incidence in England. A method utilising synthetically-controlled Bayesian structural time series to estimate the effect of COVID-19 lockdowns on HIV diagnoses has been applied to HIV diagnosis data for England and estimated a greater effect of COVID-19 lockdowns on diagnoses among GBM compared to heterosexual men and women (Muhammed *et al.* 2024). There is substantial uncertainty in these estimates, however, and diagnosis data suggests a slowdown in the progress towards HIV elimination among heterosexual groups in particular (Martin, Okumu-Camerra *et al.* 2023).

3.7 Chapter summary

In this chapter I have developed the CD4 back-calculation model to incorporate additional biomarker information, and applied the resulting dual biomarker model to HIV surveillance data for GBM newly diagnosed in England. The dual biomarker model was shown to provide more precise estimates of HIV incidence, undiagnosed prevalence, and diagnosis probabilities compared to the CD4-only model. The new model also provided information about the probability of diagnosis in individuals with recently acquired HIV, which can be correlated with HIV prevention programmes for programme evaluation.

The sensitivity of the back-calculation to different assumptions about the effect of COVID-19 lockdowns in 2020 and 2021 was assessed, with the most likely scenario being a substantial reduction in the probability of diagnosis in 2020, with a swift recovery thereafter. Forward projections using data from the post-lockdown period suggest that England is no longer on track to eliminate HIV transmission among GBM by 2030, although this may not be only as a result of the pandemic.

At present the back-calculation model is only used for estimating HIV incidence among GBM, and not applied to other key prevention groups due to the assumption of no migration. The model also cannot be used to assess the specific impact of HIV prevention initiatives, such as PrEP, on HIV transmission. Further development to incorporate migration and transmission dynamics would enhance the utility and relevance of the model for all key prevention groups in England.

4. Severity estimation

Compared to the infection fatality rate (IFR) and case fatality rate (CFR) introduced in Chapter 1, the hospitalised fatality risk (HFR) has a better-defined denominator (admission to hospital). In the UK, hospital admissions for COVID-19 were subject to significantly less reporting bias than infections among the general population, due to centralised and high-quality reporting systems. Nonetheless, the case-mix at hospital admission can vary substantially and the presence of competing risks and other biases may lead to distorted estimates unless these factors are addressed.

4.1 Aims of this chapter

A segment of the population who acquire SARS-CoV-2 in the community will require hospitalisation, potential escalation to intensive care facilities, and may die in hospital or soon after discharge. Knowledge of the risks of the different hospital pathways experienced by COVID-19 patients, the lengths of stay in hospital prior to a final outcome, and how these change over time, are crucial to: (i) evaluate the effect of clinical and public health strategies on reducing disease burden (e.g. by reconstructing the evolution of the pandemic); and (ii) inform future policies, through medium and long-term predictions of hospital demand (Birrell, Blake, van Leeuwen E, MRC Biostatistics Unit COVID-19 Working Group *et al.* 2021; Scientific Pandemic Influenza Group on Modelling, Operational sub-group 2021).

In this chapter I aim to apply multi-state models to estimate trends in hospital pathways and outcomes among patients hospitalised with community-onset COVID-19 during the initial waves of the pandemic. In the first analysis I investigate the changing risks of different pathways through hospital over the first two waves (i.e. first 12-months) of the pandemic in England, prior to widespread vaccination. In the second analysis I estimate the longer-term changes in HFR and hospital lengths of stay including the third wave (i.e. first 26-months), according to vaccination status and hospital load. Both analyses account for competing risks and censoring due to missing outcome information.

4.2 Background

4.2.1 Timeline of COVID pandemic measures

In England, the first cases of COVID-19 were detected in March 2020 and the country experienced three distinct 'waves' of infection over the following two years (Figure 4.1). By the end of April 2022, there had been over 18 million confirmed cases of COVID-19 in England, resulting in 730,000 hospital admissions, and 190,000 people had died with COVID-19 on their death certificate (UK Government 2021).

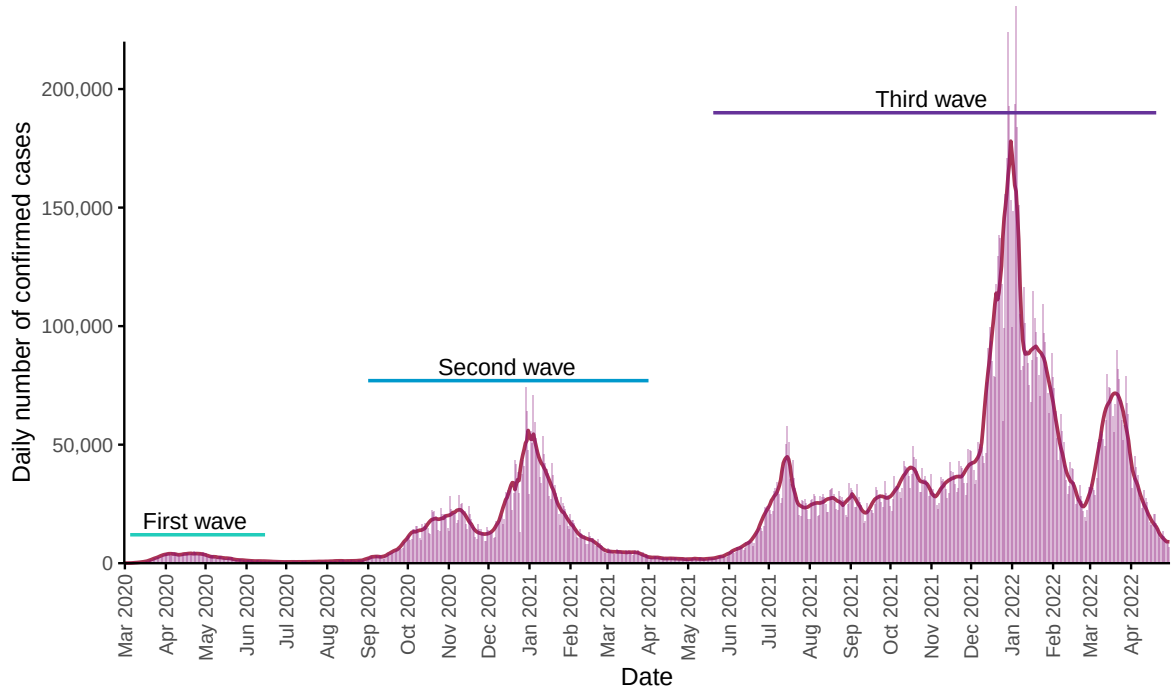


Figure 4.1: Daily number and 7-day average of confirmed SARS-CoV-2 cases over successive pandemic waves, March 2020 to April 2022, data from the UK Health Security Agency (UK Government 2021).

The various interventions introduced during the first wave of COVID-19 pandemic, such as testing and social distancing, are described in Section 1.2.1. Over the same period, hospitals were reconfigured in recognition of the severity of illness among people presenting to emergency care. An example of this was an scale up in the number of intensive care unit (ICU) and high dependency unit (HDU) beds, with more than 3500 mechanical ventilation beds occupied in January 2021 (UK Government 2021).

COVID-19 vaccination began in December 2020, initially prioritised to those most at risk of adverse outcomes and later offered to all adults. Widespread vaccination has been highly effective in reducing the risk of hospitalisation, morbidity, and mortality for those who become

infected (Lopez Bernal *et al.* 2021; Nyberg, Ferguson *et al.* 2022; Public Health England 2021a).

4.2.2 Measures of severity

The HFR considers only those who are ill enough to require hospital admission, so is likely to be much higher than either the IFR or CFR. Other measures of severity related to the HFR may look at endpoints other than death, e.g. the risk of hospital or critical care admission. Severity can also be quantified as a time delay, such as the time from infection to recovery, with a longer delay indicating greater severity. In these analyses, infection time is typically unknown and delays from measurable events, such as the date of symptom onset or date of first positive test, are used (Harrison *et al.* 2020; Nyberg, Twohig *et al.* 2021). Electronic hospital records (EHR) and hospital surveillance data can similarly be used to estimate delays, for instance the delay from admission to critical care admission, or from admission to final outcome.

Regardless of which metric is used, or which timescale considered, severity measurement is often subject to bias as a result of: incomplete case-ascertainment, changes in the frequency of testing and healthcare monitoring, differences in population-level immunity, and variation in case-mix over time (Bhattacharya *et al.* 2021). Studies of HFR are also typically not representative of the general population, since they do not consider the disproportional burden of infection on certain groups. Controlling for and assessing the potential impact of these biases can help to ensure accuracy in severity estimates.

Finally, as the evolution of new variants, immunity acquired through past infection or vaccination, and improved therapeutics may all lead to changes in disease severity over time, regular updates of severity are important for national policy makers, and a crucial input to models reconstructing pandemic evolution and predicting hospital demand (Birrell, Blake, van Leeuwen E, Gent *et al.* 2021; Birrell, Blake, van Leeuwen E, MRC Biostatistics Unit COVID-19 Working Group *et al.* 2021).

4.2.3 Severity estimation

During the first wave of infections in England, COVID-19 was observed to disproportionately impact older people and those with multiple co-morbidities, compared to younger, healthier individuals (Public Health England 2020a; Williamson *et al.* 2020). Preliminary evidence among those admitted to hospital in England and Scotland demonstrated that these factors, alongside sex, ethnicity, and socioeconomic deprivation, also heavily influenced an individual's prognosis following hospital admission for COVID-19 (Agrawal, Azcoaga-Lorenzo *et al.* 2021; Docherty, Mulholland *et al.* 2021; Ferrando-Vivas *et al.* 2021; Gray *et al.* 2021; Mathur *et al.* 2021).

The studies referenced above investigated the effect of different covariates on COVID-19 HFR, with Cox proportional hazards models used to estimate hazard ratios in two studies (Ferrando-Vivas *et al.* 2021; Mathur *et al.* 2021), and logistic regression used to estimate odds ratios in several others (Agrawal, Azcoaga-Lorenzo *et al.* 2021; Docherty, Mulholland *et al.* 2021; Gray *et al.* 2021). Most included age, sex, ethnicity, socioeconomic deprivation, and comorbidity, with a few using cubic splines to model covariates with non-linear effects, particularly age (Ferrando-Vivas *et al.* 2021; Gray *et al.* 2021; Mathur *et al.* 2021).

Limitations

Data collection mechanisms were a limitation in many of these studies, with most datasets only receiving information on completed hospital episodes (i.e. at death or discharge), and reporting of these completed episodes may also have been subject to delay. This issue is known as right-truncation, and, if not correctly accounted for, may bias HFR and length of stay estimates, as discussed in Section 2.1.2. Seaman, Presanis *et al.* 2022 have presented several methods to account for right-truncation when estimating time-to-event distributions in the context of an infectious disease epidemic.

Most of the studies performed administrative right-censoring of patient follow-up times by only considering cases which had experienced an outcome within 28 (or 30) days of hospital admission (Agrawal, Azcoaga-Lorenzo *et al.* 2021; Docherty, Mulholland *et al.* 2021; Ferrando-Vivas *et al.* 2021). Whilst right-censoring of patient follow-up is common in survival studies, this relatively short censoring window may ignore information on patients with more complex needs, who spend longer in hospital prior to a final outcome, and truncated patient stays would be completely excluded.

These studies also mostly focused on the first wave of the pandemic, and tended not to explicitly include calendar time as a covariate. The Intensive Care National Audit & Research Centre (ICNARC), which considered a longer period, reported that the mortality risk for patients admitted to intensive care in England had dropped throughout the first wave, but increased into the start of the second wave (Intensive Care National Audit and Research Centre 2021). While extending to longer time periods is desirable, the case mix of people presenting to hospital has varied throughout the pandemic and improvements have been made in therapies for COVID-19. Studies which extend to longer time periods should, therefore, consider the impact of collider bias (see Section 2.7.2), mentioned as a limitation by Gray *et al.* 2021.

None of these initial studies considered the impact of hospital load on COVID-19 outcomes. A study by the King's Fund concluded that a shortage of overnight and acute bed availability prior to the pandemic had already put hospitals under increased strain (Ewbank *et al.* 2021). A study across 207 hospitals during the pandemic found that ICU strain was associated with higher acute hospital mortality (Wilcox *et al.* 2022), and similar findings have been reported in

Switzerland, where an ICU occupancy of 70% or greater was estimated to be a tipping point at which outcomes become adversely affected (Anderegg *et al.* 2022).

4.3 COVID-19 surveillance data

Surveillance of the emergence and ongoing variation in the SARS-CoV-2 virus has been vital for public health monitoring during the COVID-19 pandemic. In England during March 2020, hospital surveillance data systems were swiftly reconfigured to collect information specifically related to COVID-19 admissions. Data from two of these systems will be described in further detail in this section.

4.3.1 Hospital surveillance data

Data from two hospital surveillance systems were available: a case-reporting dataset from a subset of sentinel clinics with patient details uploaded upon admission and updated in real-time; and a nationally representative dataset of all hospital admissions, with a reporting delay as records were only reported at final hospital outcome (i.e. at death or discharge). These datasets collected a wide array of data from the point of hospital admission to enable assessment of infection severity following hospitalisation.

As the focus for these analyses was understanding the severity and impact of community-acquired COVID-19, hospital-onset (or nosocomial) COVID-19 cases were excluded. Those with hospital-onset infection tended to be older and have longer lengths of stay than the community-onset cases, and transmission dynamics were likely to be significantly different for hospital-acquired as compared to community-acquired cases (Cooper, Evans *et al.* 2023; Evans *et al.* 2022).

SARI-Watch

The Severe Acute Respiratory Infections in England (SARI-Watch), previously known as the COVID-19 Hospitalisations in England Surveillance System (CHESS), is run by UKHSA and collects detailed individual-level data on hospital admissions for COVID-19 and influenza from a subset of sentinel NHS trusts in England, with mandatory reporting of admissions to ICU (and HDU) (Public Health England 2021b).

Data are reported in real-time via a secure online reporting tool, with data reporters able to update records as changes in care pathways and/or hospital outcomes occur. Covariate information collected in SARI-Watch includes dates of hospital and critical care admission, date and type of outcome, baseline characteristics, and NHS number.

As above, data were limited to admissions for non-nosocomial (i.e. community-acquired) admissions to sentinel NHS trusts with known admission and discharge or death date, or right-censored at the date of extraction.

SUS

The NHS England Secondary Uses Service (SUS) dataset, run by NHS England, contains well-completed, individual-level data on hospital admissions for COVID-19 across all NHS trusts in England. SUS data are only reported upon completion of a hospital stay (i.e. at the point of discharge from hospital or death) and are therefore subject to a reporting delay. To partially account for this, linkage to the Emergency Care Dataset for England (ECDS), which promptly records all emergency care attendances and onwards destinations (i.e. to home or admittance to hospital), may be used to obtain information for patients admitted via emergency care and remaining in hospital (NHS Digital 2023).

Covariate information collected in SUS includes: date of hospital admission, age group, Charleson comorbidity index (CCI) (Charlson *et al.* 1987), baseline characteristics, and NHS number. As data on all hospital admissions are collected, a crude measure of demand (or hospital load) can be obtained from these data, e.g. the number of admissions within a 1-day period.

4.3.2 Data linkage

Obtaining complete information on a patient's testing history, vaccination status, pathways through hospital, and eventual outcome requires data linkage to augment the hospital records. Both SARI-Watch and SUS collected NHS number, which was the basis for this linkage (Bhattacharya *et al.* 2021). Data validation on the linked datasets was used to rule-out spurious matches, helped to minimise missing data, and improved data quality.

The data available for linkage included:

- COVID-19 testing data from the UKHSA Second Generation Surveillance System (SGSS) (NHS Digital 2024) collected via two routes: **Pillar 1**, tested on arrival at hospital, and **Pillar 2**, tested within the community;
- vaccination data from the National Immunisation and Vaccine System (NIVS) (NHS Digital 2022);
- data on deaths from UKHSA's COVID-19 mortality dataset (Covid-19 EpiCell 2020). These data comprised deaths from four sources, with death outcomes included if they occurred within 60 days of a positive COVID-19 test, or if COVID-19 was mentioned on the death certificate:

1. Office for National Statistics (ONS) death registrations;
2. NHS England deaths in hospitals;
3. SGSS;
4. local Health Protection Teams.

A third hospital dataset, the NHS Hospital Episodes Statistics (HES), also recorded information on hospital admissions for COVID-19, but due to substantial reporting delay (of >3 months) these data had limited utility for pandemic response (NHS Digital 2020). As HES contains more granular information these data were used for validation, and to minimise missing information.

4.4 Pre-vaccine hospitalised severity

In this section I investigate COVID-19 hospitalised severity during the first two waves (12-months) of the COVID-19 pandemic, pre-vaccine introduction. The aim of this analysis was to assess how competing pathways through hospital, and the corresponding lengths of stay in hospital have changed over time. Changes in the hospital-fatality risk (HFR) and time to final event were estimated using Aalen-Johansen and parametric multi-state mixture models.

4.4.1 Data

The SARI-Watch data include all dates of movement between levels of care (e.g. Hospital/ICU), a patient's current status (e.g. still in hospital/still in ICU/transferred/unknown) or final outcome (e.g. discharged/died/unknown), and detailed information on patient characteristics. Data for March 2020 to February 2021 were extracted at the end of March 2021, with data linkage to the UKHSA mortality dataset and HES used to improve the completeness of outcome information and patient characteristics.

Figure 4.2 panel A shows the final outcomes and information on missing outcomes for patients reported in SARI-Watch between March 2020 and February 2021, and panel B shows the proportion of patients who were admitted to ICU over this period.

Censoring

For this investigation data were updated in real-time and outcomes may have been unknown either (a) because the outcome was not reported despite having occurred (termed 'truly missing'), or (b) because the outcome had not occurred by the end of the period covered by the data ('right-censoring'). For past admissions (several months prior), unknown outcomes are more likely to be truly missing, as the majority of patients will have completed their hospital episode.

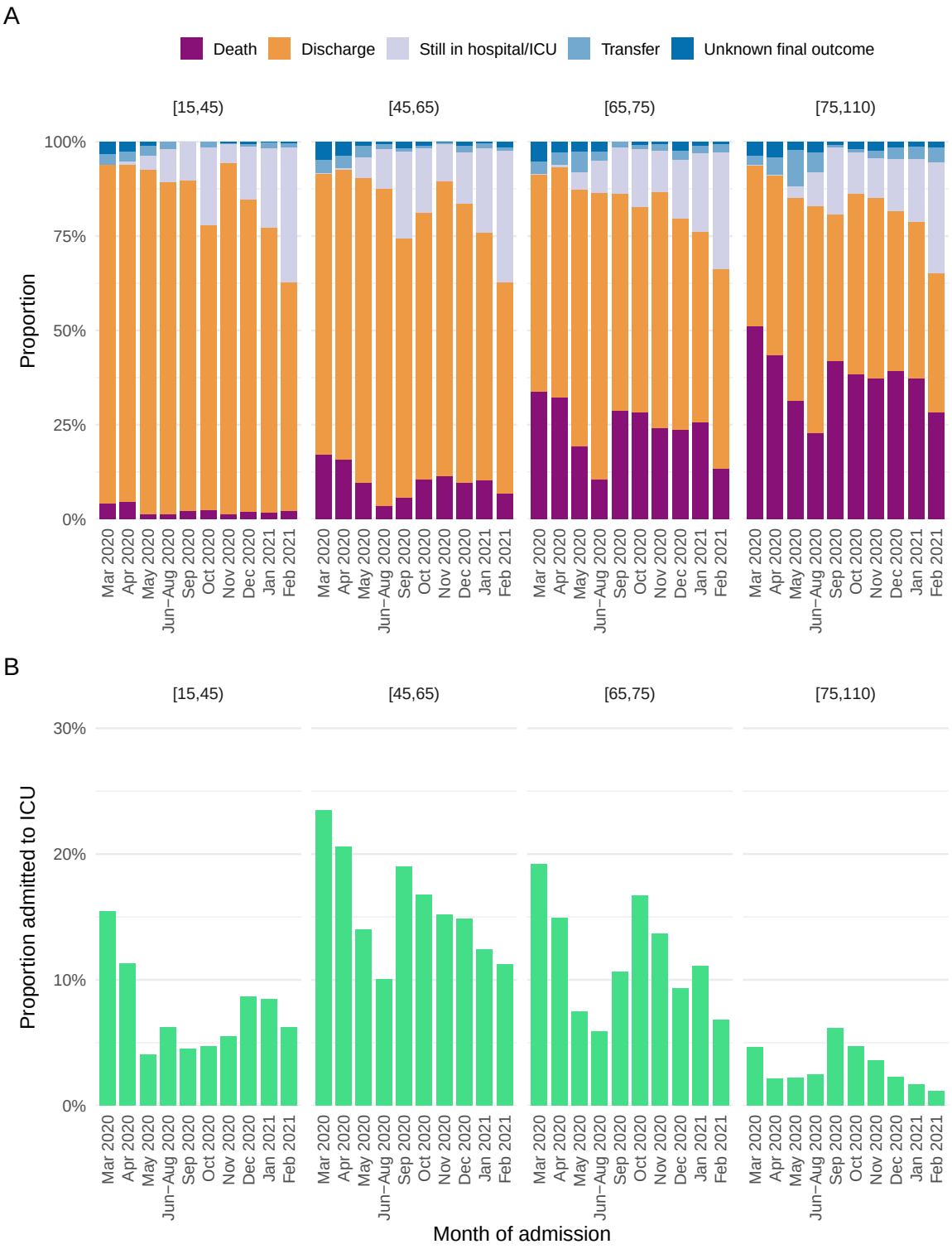


Figure 4.2: Proportion of COVID-19 hospital admissions by age group and final outcome (Panel A), and ICU admission status (panel B), March 2020 to February 2021. Admissions occurring during June, July and August 2020 are grouped due to small numbers.

In more recent months, unknown outcomes are more likely to represent patients who remain in hospital care, i.e. right-censoring.

The last known status for each patient was used to determine an appropriate cut-off for censoring. Patients reported as 'still in hospital/still in ICU' were right-censored at the date of data extraction, or at 60 days if they were reported as still in non-critical care beyond 60 days, or at 90 days if they were reported as still in critical care beyond 90 days. This approach was a compromise between excluding these records from the analysis, which may bias outcome estimation, and censoring at the date of data extraction, which would allow for unrealistically long hospital stays of several months after admission. Transferred patients with unknown outcomes were censored at the date of transfer, patients with unknown status and unknown outcome were censored at the date of their last observed event.

Figure 4.3 shows the different transitions experienced by individuals in the SARI-Watch data, with right-censoring and missing outcomes indicated.

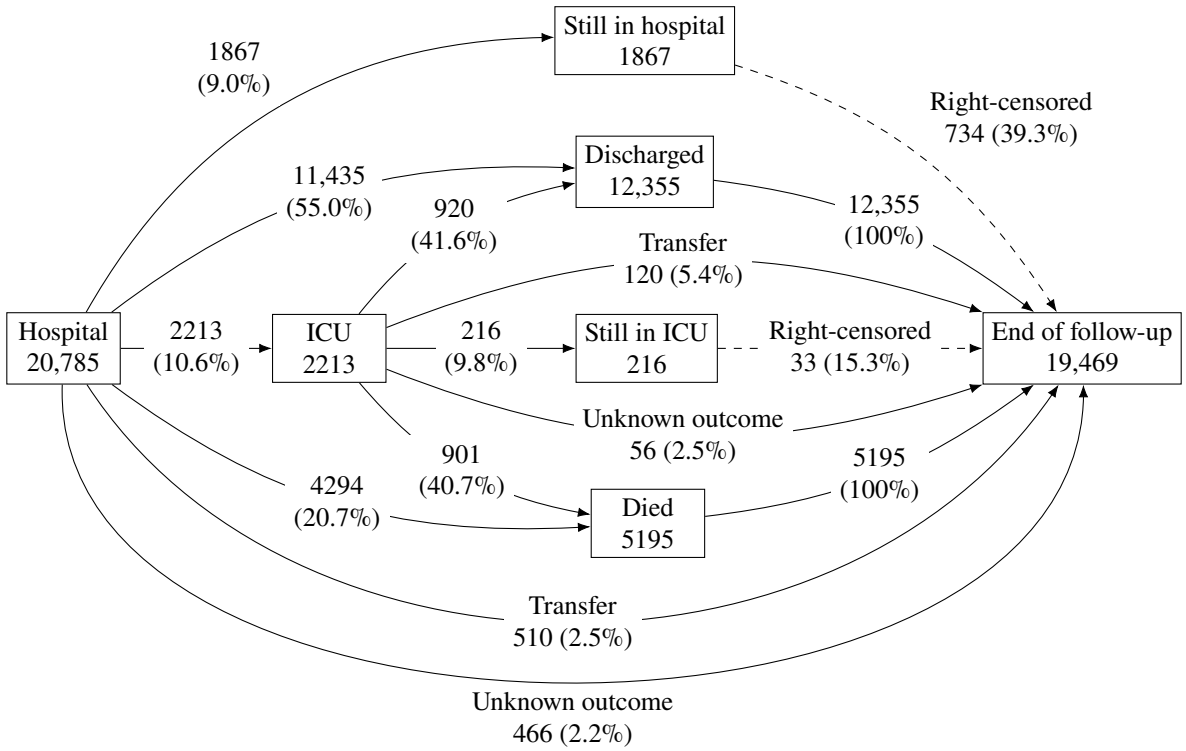


Figure 4.3: Flow diagram showing transitions between hospital states and outcomes in the SARI-Watch data, with observed number and proportion of individuals in and moving between each state. Dashed lines indicate right-censoring for individuals reported as 'still in hospital' after 60 days or 'still in ICU' after 90 days.

4.4.2 Multi-state model

Model description

Non-parametric Aalen-Johansen and parametric multi-state mixture models, specified to account for the competing risks of each pathway and missing outcome information were applied to the SARI-Watch data. The multi-state models had four states, with the transition intensities to subsequent states depending on how long a person had spent in the current state, but not on any events before entering the current state (i.e. a Semi-Markov model). Two sub-models represented competing risks of the next event (i) from Hospital admission, and (ii) from ICU admission, as shown in Figure 4.4.

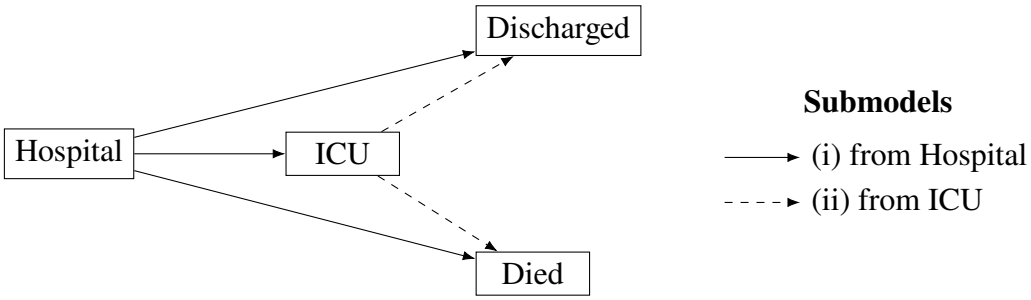


Figure 4.4: Multi-state model for pathways through hospital following admission for COVID-19 in the SARI-Watch data. Arrows indicate transitions between states in each of the submodels: (i) from Hospital, and (ii) from ICU.

Likelihood for multi-state mixture model

The likelihood for a competing risks model may be extended to our multi-state model (Jackson *et al.* 2022). For each patient i, with event observations indexed by j, the transitions between states $r_{i,j}$ and $s_{i,j}$ occurring at time $y_{i,j}$ are observed in the data as one of three types, indexed by $\delta_{i,j}$:

1. exact transition times: $\boldsymbol{y}_{i,j} = \{y_{i,j}, r_{i,j}, s_{i,j}, \delta_{i,j} = 1\}$, where the transition from state $r_{i,j}$ to state $s_{i,j}$ occurs at time $y_{i,j}$;

2. right-censoring: $\boldsymbol{y}_{i,j} = \{y_{i,j}, r_{i,j}, \delta_{i,j} = 2\}$, where follow-up ends while the individual is in state $r_{i,j}$ at time $y_{i,j}$, and the next state and time of transition are unknown;

3. partially-known outcome: $\boldsymbol{y}_{i,j} = \{y_{i,j}, \delta_{i,j} = 3\}$, where it is known whether or not an individual went to ICU, but unknown if they are still in hospital at time $y_{i,j}$.

As in Section 2.3, let $\pi_{r,s}$ be the probability that the next state for an individual i in state r is state s and let $f_{r,s}(. \mid \theta_{r,s})$ be the densities of the parametric distributions for the time to

transition from state r to state s, given that this transition occurs. The likelihood contributions for each $\delta_{i,j}$ are:

$$\begin{aligned}
\delta_{i,j} = 1: &\quad l_{i,j} = \pi_{r_{i,j}s_{i,j}} f_{r_{i,j}s_{i,j}}(y_{i,j} \mid \theta_{r_{i,j}s_{i,j}}) \\
\delta_{i,j} = 2: &\quad l_{i,j} = \textstyle\sum_{s \in S_r} \pi_{r_{i,j}s} \left(1 - F_{r_{i,j}s}(y_{i,j} \mid \theta_{r_{i,j}s})\right) \\
\delta_{i,j} = 3, s \neq \text{Discharge}: &\quad \text{as for } \delta_{i,j} = 2 \\
\delta_{i,j} = 3, s = \text{Discharge}: &\quad l_{i,j} = \pi_{r_{i,j}s}
\end{aligned}$$

The full likelihood is the product of these likelihood contributions:

$$L(\boldsymbol{\pi}, \boldsymbol{\theta} \mid \boldsymbol{y}) = \prod_{i,j} l_{i,j}$$

where $\boldsymbol{\pi}$ includes all of the competing event probabilities $\pi_{r,s}$, and $\boldsymbol{\theta}$ includes the parameters of the time to transition distributions.

Parametric distributions

Figure 4.5 shows the empirical distribution of time to transition (measured in days) from hospital admission to the competing outcomes of ICU admission, death, and discharge.

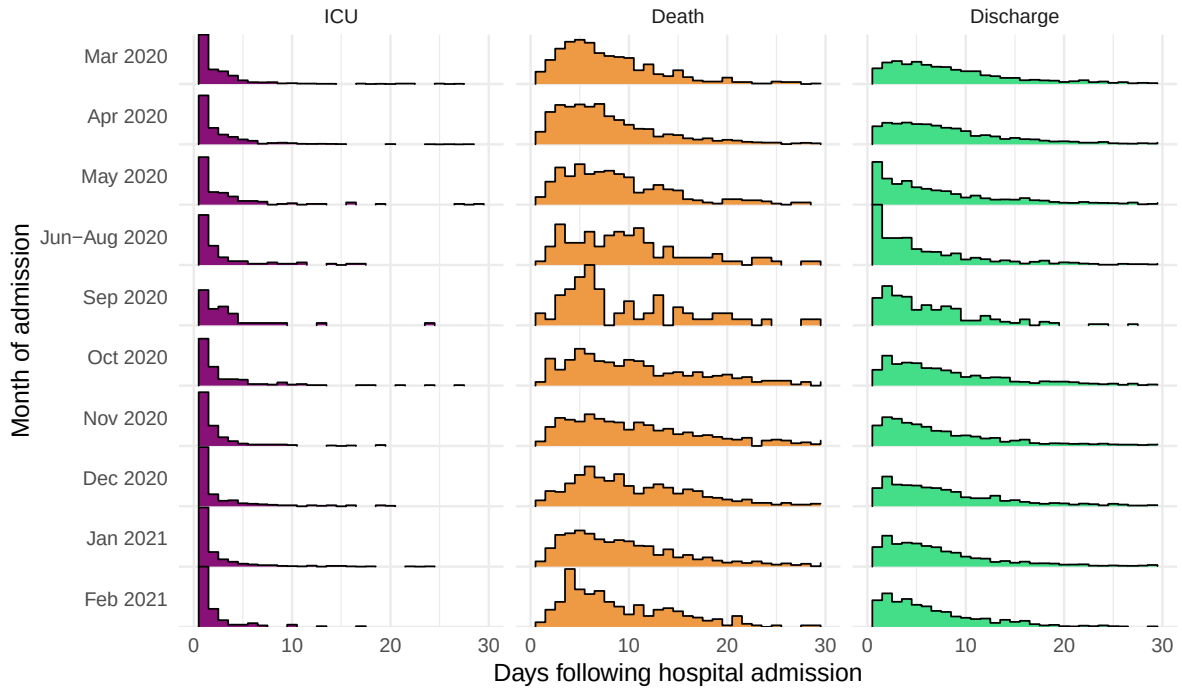


Figure 4.5: Distribution of time to event in SARI-Watch data for the first 30 days following hospital admission, by month of hospital admission, March 2020 to February 2021. Admissions occurring during June, July and August are grouped due to small numbers.

In selecting parametric distributions to model these times to transition a trade-off must be made between over-parameterisation and goodness of fit. The flexibility of an extra parameter can allow for a better fit to the data, but using too many parameters compared to the data may result in parameters being unidentifiable.

Generalized gamma distributions are common choices for right-censored survival data due to their flexibility, which avoids too many assumptions on the shape of the distribution (Pal *et al.* 2020). One parameterisation of this distribution is expressed in terms of the parameters μ, σ, and Q, representing the mean (location), scale, and shape of the distribution, respectively (Prentice 1974):

$$f(x \mid \mu, \sigma, Q) = |Q| \frac{(Q^{-2})^{Q^{-2}}}{\sigma \cdot x \cdot \Gamma(Q^{-2})} \exp(Q^{-2}(Q \cdot w - \exp(Q \cdot w)))$$
$$\text{where } w = \frac{\log(Q^2 \cdot g)}{Q} \quad \text{and} \quad g \sim \text{Gamma}(Q^{-2}, 1)$$

As their name might suggest, generalized gamma distributions are a *generalisation* of other common survival distributions, namely the gamma, log-normal, exponential, and Weibull distributions. These distributions are obtained from the generalized gamma under the following constraints:

$$\begin{aligned} \text{Weibull}: &\quad Q = 1 \\ \text{Log-normal}: &\quad Q = 0 \\ \text{Gamma}: &\quad Q = \sigma \\ \text{Exponential}: &\quad Q = \sigma = 1 \end{aligned}$$

The specification of the mixture model allows for different distributions to be used for each transition. Table 4.1 shows a selection of distributions which were considered for each transition in each of the sub-models, fitted without covariate effects. Goodness of fit and parsimony of each sub-model can be assessed through comparison to non-parametric Aalen-Johansen cumulative incidence curves, likelihood ratio tests, and Akaike information criterion (AIC) values.

The hospital sub-model with the lowest AIC used a log-normal distribution for the 'to ICU' transition and generalized gamma distributions for the 'to death', and 'to discharge' transitions. The difference in AIC between the two best fitting ICU sub-models was not significant (with a difference in AIC <5), so the generalized gamma distribution was chosen for both transitions based on a comparison to the Aalen-Johansen estimates.

Table 4.1: Log-likelihood and AIC values for parametric time to transition distributions, sorted by sub-model and AIC. Models with unidentifiable parameters not shown.

Sub-model	ICU	Death	Discharge	Log-likelihood	AIC
Hospital	log-normal	gengamma	gengamma	-77841.96	155731.93
Hospital	log-normal	log-normal	gengamma	-77843.40	155732.80
Hospital	log-normal	gengamma	log-normal	-77982.90	156011.80
Hospital	Weibull	gengamma	gengamma	-78281.82	156611.63
Hospital	log-normal	gamma	gengamma	-78324.22	156694.44
Hospital	log-normal	Weibull	gengamma	-78329.02	156704.05
Hospital	log-normal	exponential	gengamma	-78340.72	156725.44
Hospital	gamma	gengamma	gengamma	-78367.64	156783.28
Hospital	exponential	gengamma	gengamma	-78397.36	156840.73
Hospital	log-normal	gengamma	Weibull	-79322.13	158690.27
Hospital	log-normal	gengamma	gamma	-79905.44	159856.87
Hospital	log-normal	gengamma	exponential	-81066.20	162176.41
ICU		gengamma	log-normal	-8710.57	17447.14
ICU		gengamma	gengamma	-8710.62	17449.24
ICU		log-normal	gengamma	-8717.79	17461.59
ICU		gamma	gengamma	-8764.12	17554.24
ICU		gengamma	gamma	-8765.89	17557.78
ICU		gengamma	exponential	-8770.58	17565.17
ICU		gengamma	Weibull	-8769.90	17565.80
ICU		exponential	gengamma	-8772.38	17568.76
ICU		Weibull	gengamma	-8772.14	17570.28

Covariate effects

Covariates available in the SARI-Watch dataset included month of admission, age group, sex, region of residence (London, South of England, Midlands and East of England, and North of England), ethnicity, and number of comorbidities (see Table D.1).

Probabilities of competing events $\pi_{r,s}$ could be related to one or more of these covariates via multinomical logistic regression, with parameters of the time to transition distributions related to covariates through accelerated failure time models, as described in Section 2.3.

For the generalized gamma distribution, covariates can be specified on the mean, scale, and shape parameters, μ, σ, and Q, respectively. For the log-normal distribution, covariates can be specified on the log(mean) and log(standard deviation) parameters, μ and σ, respectively. Table 4.2 shows the effect on the log-likelihood and AIC of regressing the probability (π) and the parameters of each time to transition distribution on the month of admission covariate. For example: the best-fitting 'from Hospital' sub-model included regression on month of admission for each of:

- the competing event probability (π);
- the mean and scale parameters of the generalized gamma (μ and σ);
- the log(mean) parameter of the log-normal (μ).

When multiple covariates are considered, covariate interactions can also be included. Table 4.3 shows the effect on the log-likelihood and AIC of including an interaction between month of admission and age group on the probability, and on each parameter of the time to transition distributions. The best-fitting 'from Hospital' sub-model included interaction terms between month and age group on both parameters of the time to discharge transition, whilst the best-fitting 'from ICU' sub-model included interaction terms on the scale parameter of the time to death transition.

Appendix D.1.1 shows a comparison between the parametric and non-parametric estimates for the month, and month and sex covariates. There was a high level of agreement for the 'to death' and 'to ICU' transitions, with some lack of agreement for the 'to discharge' transition occurring due to right-censoring.

Estimation

The models used to estimate hospital pathways and lengths of stay were selected by comparing AIC and goodness of fit. Since the focus of this analysis was understanding trends over time, the covariates included in each model were the month of admission and one additional covariate (models with more than two covariates were generally not identifiable).

Final outcome probabilities were obtained by averaging over the hospital and ICU submodels. For example, to obtain an estimate of the overall HFR, the probability of death (D) given hospitalisation (H) is averaged over ICU admission (I) and non-ICU admission ($\bar{I}$):

$$\begin{aligned}\Pr(D \mid H) &= \Pr(D \mid H, \bar{I}) \cdot \Pr(\bar{I} \mid H) + \Pr(D \mid H, I) \cdot \Pr(I \mid H) \\ &= \Pr(D, \bar{I} \mid H) + \Pr(D \mid H, I) \cdot \Pr(I \mid H)\end{aligned}$$

The parameters of the models were estimated using maximum likelihood estimation with the open-source `flexsurv` and `survival` packages in `R` (Jackson 2016; R Core Team 2020; Therneau 1999). Quantiles of the event-specific parametric time-to-event distributions for specified combinations of covariates were obtained from the fitted mixture models.

Table 4.2: Log-likelihood and AIC values for log-normal and generalized gamma distributions, by location of covariate effects, sorted by sub-model and AIC. Models with unidentifiable parameters not shown.

Sub-model	π	ICU	Death	Discharge	Log-likelihood	AIC
Hospital	✓	μ	μ, σ	μ, σ	-77328.69	154775.39
Hospital	✓	σ	μ, σ	μ, σ	-77330.15	154778.30
Hospital	✓	μ	μ	μ, σ	-77349.89	154803.78
Hospital	✓	σ	μ	μ, σ	-77351.18	154806.35
Hospital	✓	μ	σ	μ, σ	-77412.52	154929.04
Hospital	✓	σ	σ	μ, σ	-77414.11	154932.22
Hospital	✓	μ	μ, σ	σ	-77450.65	155005.30
Hospital	✓	σ	μ, σ	σ	-77452.02	155008.04
Hospital	✓	μ	μ	σ	-77470.37	155030.74
Hospital	✓	σ	μ	σ	-77471.77	155033.54
Hospital	✓	μ	σ	σ	-77534.72	155159.44
Hospital	✓	σ	σ	σ	-77535.97	155161.95
Hospital	✓			σ	-77561.32	155184.63
Hospital	✓	μ	μ, σ	μ	-77595.90	155295.79
Hospital	✓	σ	μ, σ	μ	-77596.80	155297.60
Hospital	✓	μ	μ	μ	-77615.72	155321.45
Hospital	✓	σ	μ	μ	-77616.98	155323.96
Hospital	✓	μ	σ	μ	-77680.16	155450.32
Hospital	✓	σ	σ	μ	-77681.18	155452.35
Hospital	✓			μ	-77706.51	155475.02
Hospital	✓		μ		-77756.14	155574.27
Hospital	✓		σ		-77822.23	155706.45
Hospital	✓	μ			-77833.38	155728.77
Hospital	✓	σ			-77834.38	155730.75
ICU	✓		μ, σ	μ	-8683.10	17436.20
ICU	✓		μ	μ	-8691.38	17438.75
ICU	✓		μ, σ	μ, σ	-8677.96	17439.91
ICU	✓		σ	μ	-8692.44	17440.88
ICU	✓		μ, σ	σ	-8685.68	17441.35
ICU	✓		μ	μ, σ	-8686.14	17442.27
ICU	✓			μ	-8700.25	17442.51
ICU	✓		μ	σ	-8693.51	17443.02
ICU	✓		σ	μ, σ	-8687.15	17444.29
ICU	✓		μ		-8701.24	17444.48
ICU	✓		σ	σ	-8694.51	17445.02
ICU	✓		σ		-8702.09	17446.19
ICU	✓			σ	-8702.24	17446.47

Table 4.3: Log-likelihood and AIC values for for log-normal and generalized gamma distributions, by inclusion of interaction of month and age group covariate, sorted by sub-model and AIC. Models with unidentifiable parameters not shown.

Sub-model	π	ICU	Death	Discharge	Log-likelihood	AIC
Hospital				σ, μ	-74507.16	149258.33
Hospital	✓			σ, μ	-74472.88	149273.76
Hospital				μ	-74539.05	149280.10
Hospital				σ	-74545.95	149293.89
Hospital	✓			μ	-74506.44	149298.89
Hospital					-74569.99	149299.98
Hospital		μ			-74550.31	149302.62
Hospital	✓			σ	-74514.02	149314.05
Hospital			μ		-74558.15	149318.29
Hospital	✓				-74538.34	149320.68
Hospital	✓	μ			-74517.84	149321.68
Hospital			σ		-74562.25	149326.51
Hospital	✓		μ		-74525.54	149337.09
Hospital	✓		σ		-74530.10	149346.20
Hospital			σ, μ		-74551.19	149346.39
Hospital	✓		σ, μ		-74515.30	149358.60
ICU			σ		-8471.84	17099.69
ICU					-8497.76	17109.53
ICU	✓		σ		-8461.73	17121.46
ICU			μ		-8483.82	17123.65
ICU	✓				-8486.04	17128.07
ICU	✓		σ, μ		-8445.41	17130.83
ICU	✓		μ		-8474.92	17147.85

4.4.3 Results

A total of 28,363 patients were admitted to a sentinel NHS trust between March 2020 and February 2021. Baseline characteristics by outcome following hospital admission are shown in Table D.1. Excluding those with missing outcomes, 26.5% (849/3198) of those admitted to a sentinel NHS trust in March 2020 died outside of critical care. This decreased to 11.1% (125/1125) in June-August 2020, before rising to 24.1% (510/2112) in December 2020. The proportion of hospital admissions with missing outcome information increased over time, from 5.6% (189/3387) in March 2020 to 32.8% (689/2089) in February 2021.

Following hospital admission, death was more common in older patients: 43.9% (4290/9767) of those aged 75+, compared to 1.5% (49/3251) of those aged 15–49, whereas ICU admission was less common: 3.1% (305/9767) vs. 9.8% (319/3251), respectively. Outcomes were poorer among those with multiple comorbidities: 31.2% (2219/7123) of those with 3 or more

comorbidities died outside of critical care, and an additional 11.8% (842/7123) were admitted to ICU, this compared to 18.5% (1561/8419) and 8.1% (681/8419), respectively, for those with no comorbidities recorded.

Estimated outcome probabilities following hospital admission by month of admission and sex, age group, and number of comorbidities are shown in Figure 4.6, panels A–D. Probability of ICU admission following hospital admission decreased between March 2020 and May 2020, from 14.2% (95% CI 13.1 to 15.5%) to 6.0% (95% CI 5.0 to 7.1%); rising to 10.4% (95% CI 9.0 to 11.8%) in October 2020 before plateauing at around 8% in November and December 2020. These trends were similar across the different sex, age, and comorbidity subgroups with younger patients being more likely to be admitted to ICU than older patients, e.g. in March 2020: 23.8% (95% CI 21.1 to 26.2%) for those aged 45–65 compared to 4.6% (95% CI 3.7 to 5.9%) for those aged 75+.

Figure 4.7, panels A–D, show estimated median times from hospital admission to next event by month of admission, sex, age group, and number of comorbidities. Time to ICU admission remained low throughout the period, at around 1.1 days after hospital admission. Median length of stay in hospital prior to death in non-critical care increased from 6.5 days (95% CI 6.0 to 7.0 days) in March 2020 to 9.3 days (95% CI 8.6 to 10.1 days) in December 2020, with a similar trend across all subgroups. Meanwhile, median lengths of stay in non-critical care prior to discharge fell from 6.3 days (95% CI 5.9 to 6.7 days) for admissions during March to 3.9 days (95% CI 3.6 to 4.3 days) in June/July/August, but lengthened to 8.0 days (95% CI 7.5 to 8.6 days) by December. Stays in non-critical care prior to discharge were longest for older patients, e.g. 10.8 days (95% CI 10.2 to 11.6 days) in March 2020 among those aged 75+ compared to 3.2 days (95% CI 3.0 to 3.5 days) for those aged 15–45.

Corresponding probabilities and lengths of stay following ICU admission are included in Appendix D.1.2.

HFR estimates which combine the two sub-models (with and without ICU admission) are shown in Figure 4.8. Between March and June-August 2020, HFR reduced from 32.0% (95% CI 30.5 to 33.7%) to 11.3% (95% CI 9.6 to 13.1%); this was followed by an increase to 23.1% (95% CI 21.6 to 24.9%) for patients admitted in December 2020 and a fall to 15.9% (95% CI 14.4 to 17.7%) by February 2021. Whilst the initial fall in HFR was estimated across all subgroups, the increase in December was estimated particularly in the oldest age group (75+) and those with three or more comorbidities, with estimated HFRs of 39.4% (95% CI 36.3 to 42.3%) and 35.9% (95% CI 30.9 to 41.5%) for admissions in the final month of 2020, respectively.

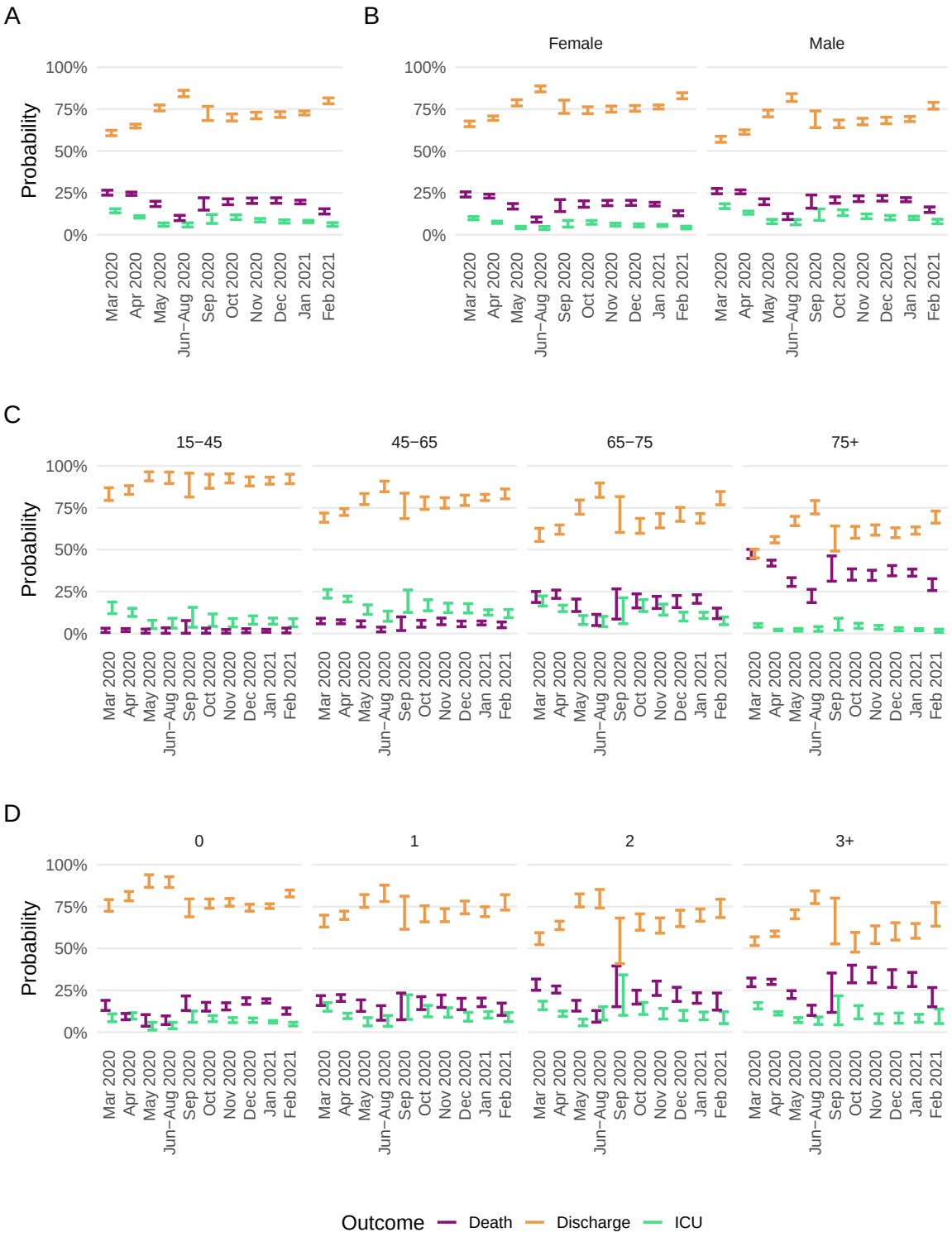


Figure 4.6: Probability of next event and 95% CI following hospital admission in SARI-Watch data estimated using multi-state mixture model, March 2020 to February 2021. By month of admission (panel A), sex (panel B), age group (panel C), and number of comorbidities (panel D).

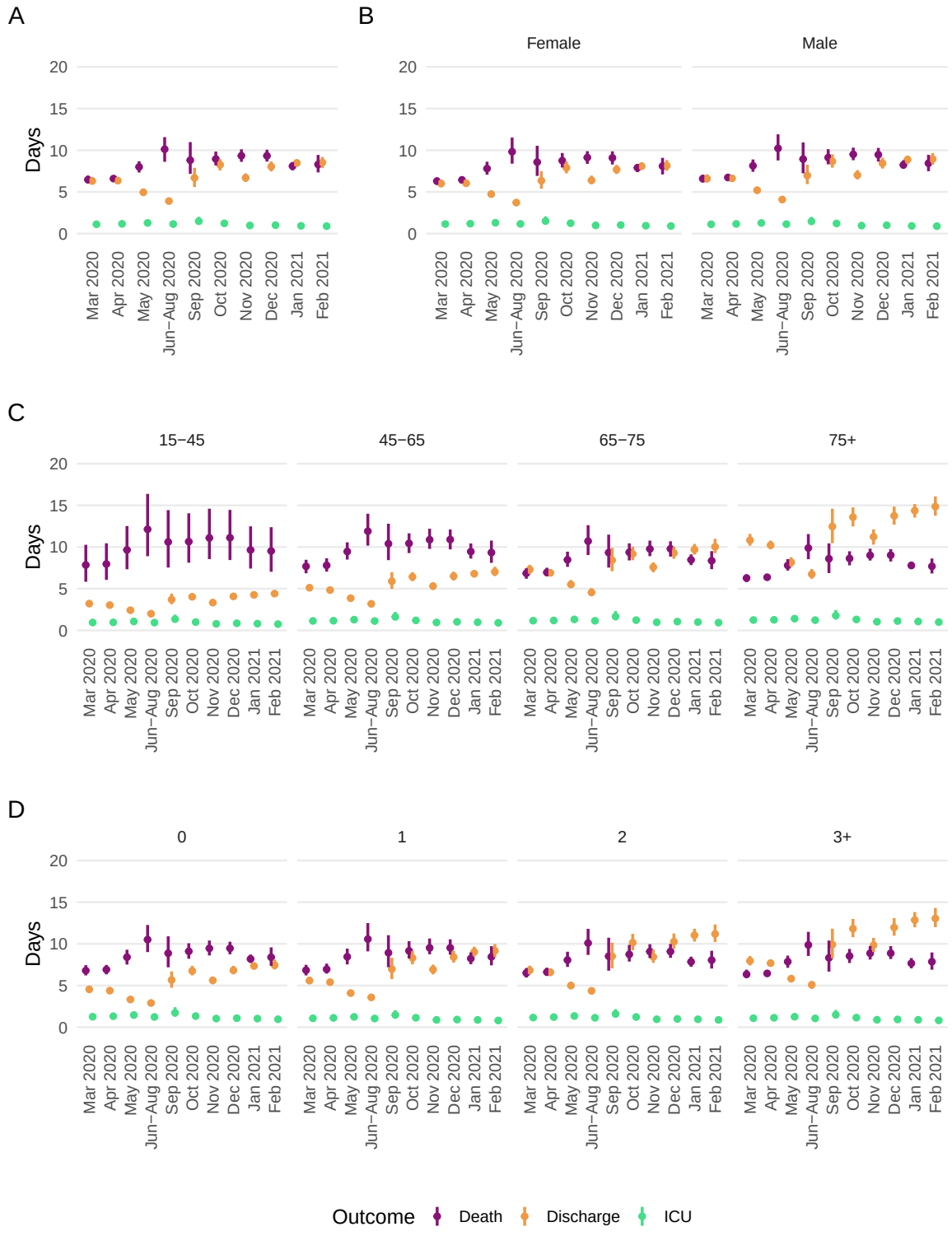


Figure 4.7: Median time to event and 95% CI following hospital admission in SARI-Watch data estimated using multi-state mixture model, March 2020 to February 2021. By month of admission (panel A), sex (panel B), age group (panel C), and number of comorbidities (panel D).

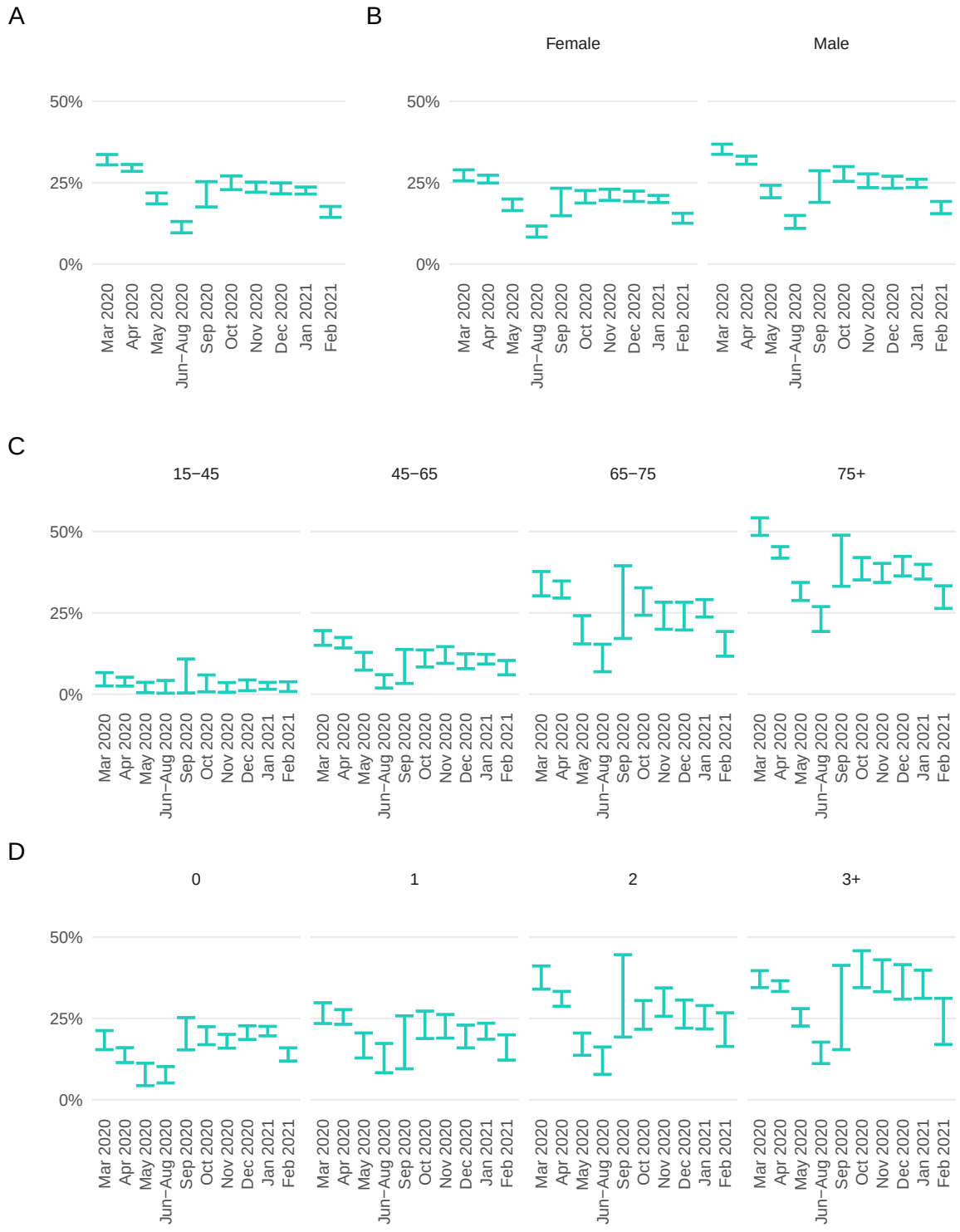


Figure 4.8: Hospitalised case-fatality risk (HFR) and 95% CI averaged over ICU and non-ICU admission in SARI-Watch data, estimated using multi-state mixture model, March 2020 to February 2021. By month of admission (panel A), sex (panel B), age group (panel C), and number of comorbidities (panel D).

Seasonality

A somewhat seasonal trend in HFR was observed and the use of a smoothed functions of time to model seasonality in COVID-19 mortality rates was considered, in addition to the categorical month of admission. The Serfling model has been previously implemented to monitor death registrations in European countries (Nielsen *et al.* 2013; Serfling 1963). For this model, weekly mortality counts (w) were modelled as a function of calendar time, with sine and cosine terms to represent seasonality:

$$\sin(2\pi\frac{w}{52}),\ \ \cos(2\pi\frac{w}{52})$$

Figure 4.9 shows the result of including trigonometric functions of the week of admission as covariates on the probabilities $\pi_{r,s}$, and distribution parameters $\theta_{r,s}$, with comparison to the categorical estimates. HFR estimates based in the trigonometric functions were similar throughout the first half of 2020, but there was a lack of fit to the data in later months.

4.4.4 Discussion

This analysis of COVID-19 hospital surveillance data provides a comprehensive assessment of hospital severity over the first two waves of the pandemic, covering the pre-vaccine period in England. Risks of severe events and lengths of stay in hospital were found to vary over time and according to baseline characteristics.

The hospitalised case-fatality risk fell during the first wave of the pandemic (March to June 2020), with a resurgence estimated across all subgroups in autumn/winter 2020/21, particularly for older patients and those presenting with multiple comorbidities. Lengths of stay in hospital for COVID-19 patients, meanwhile, increased throughout this period, and were longer for older patients and those with multimorbidity.

The estimated decline in hospitalised mortality during the initial months of the COVID-19 pandemic in England supports the findings of Docherty, Harrison *et al.* 2020 and Gray *et al.* 2021, who estimated reductions in HFR from 35% to 15% and from 52% to 17%, respectively, during the first wave. These, and other studies, have reported the same association between COVID-19 severity, age, gender, ethnicity, and comorbidity explored in this analysis (Navaratnam *et al.* 2021; Williamson *et al.* 2020).

By extending the follow-up period through to the end of winter 2021, and accounting for competing risks, longer-term trends in the risks of ICU admission, death, and hospital discharge, and lengths of stay for both critical and non-critical care patients could be estimated. The mortality trends among intensive care patients matched an increase in mortality during autumn 2020 reported by the Intensive Care National Audit and Research Centre 2021.

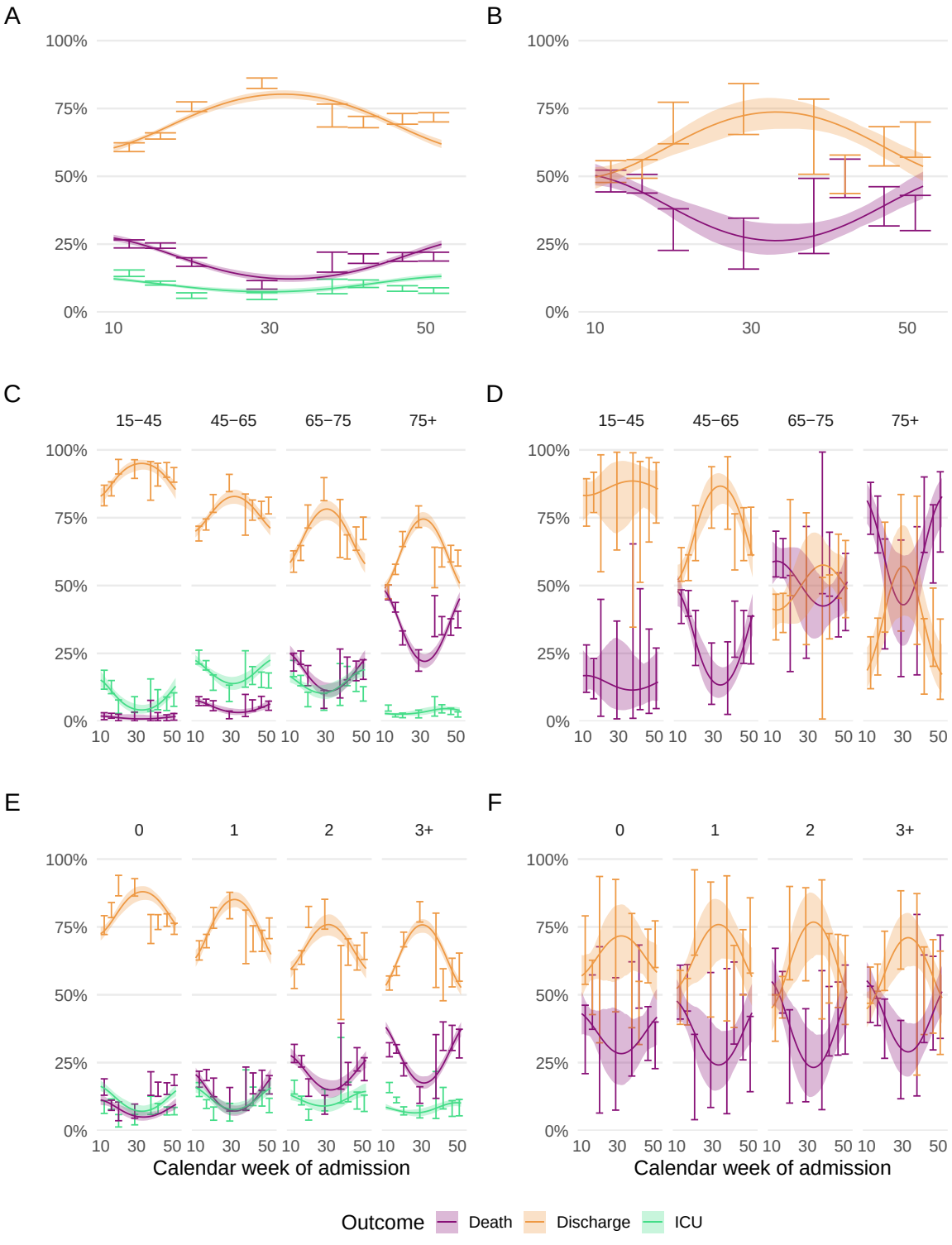


Figure 4.9: Outcome probabilities and 95% CI for hospital and ICU mixture models in SARI-Watch data, March 2020 to February 2021. By month and week (panels A and B), including age (panels C and D), and comorbidity (panels E and F). Lines and shaded regions are trigonometric estimates, line ranges are categorical estimates.

The lack of fit when incorporating seasonality may indicate that assumptions about the 'cyclical' nature of the distributions were too strong. Future work might consider the use of splines to define a smooth function of time with fewer assumptions.

Improved patient management and containment of COVID-19 via lockdown restrictions likely contributed to the reduction in mortality during the first wave of infection in England (Bamford *et al.* 2020; Davies, Barnard *et al.* 2020; Recovery Collaborative Group 2020). Subsequent changes in case-mix, easing of national restrictions, new variants, and changes in critical care admission criteria may have contributed to the resurgence in hospital pressures during autumn and winter 2020/21 reported by Roxby *et al.* 2020, and the estimated increase in mortality (National Health Service 2020). Indeed, the severity of the B.1.1.7 VOC-202012/01 (Alpha) variant, which emerged in in December 2020, was likely higher than the wild type variant (Davies, Jarvis *et al.* 2021; Nyberg, Twohig *et al.* 2021; Scientific Advisory Group for Emergencies 2021).

The substantial drop in HFR observed at the end of the study period coincided with the receipt of second vaccine doses by those most at risk, an intervention which was found to be highly effective in reducing COVID-19 mortality (Lopez Bernal *et al.* 2021). Whilst additional SARI-Watch data covering the post-vaccination period were not available for this analysis, the effect of vaccination on severity is further explored in the next section.

Interim estimates of HFR using these methods have informed several of the expert scientific panels mobilised during the initial months of the COVID-19 pandemic, including SPI-M, SAGE, and the New and Emerging Respiratory Virus Threats Advisory Group (NERVTAG) (GOV.UK 2020a; GOV.UK 2020b; GOV.UK 2020c). The application of competing risks mixture models to COVID-19 data presented here contributed to a joint publication (Jackson *et al.* 2022), and additional results from this analysis have been made available as a pre-print (Kirwan, Elgohari *et al.* 2021).

Biases and limitations

The data used for this analysis were collected from a subset of sentinel clinics, representing only 31 out of >200 NHS trusts in England. Whilst initially chosen to be geographically representative, trusts in SARI-Watch chose whether to continue sentinel reporting during the pandemic, with certain regions being under-represented in the dataset (e.g. London) which could bias the estimates of HFR. To assess the impact of this bias covariate effects were compared with the population-level estimates of death and hospital admission of Williamson *et al.* 2020, with the lower HFR in this study likely due to the lower representation in London. The next section addresses this bias by considering HFR among all NHS trusts, rather than a subset.

Another limitation of the sentinel reporting was the smaller number of patient records than were available for other comparable studies (Docherty, Mulholland *et al.* 2021; Ferrando-Vivas

et al. 2021). This increased the degree of uncertainty in the estimates, particularly among BAME patient groups, and limited analyses to only two covariates for the most part (month of admission plus one demographic factor). The extent to which interactions between covariates could be explored was therefore limited, and these associations must be interpreted carefully due to the potential for confounding.

Whilst covariate information was generally well completed, with validation via linkage to the HES dataset, outcome information was more likely to be missing towards the end of the study. Right censoring and linkage to the UKHSA mortality dataset helped to mitigate this missing information bias, with no obvious systematic bias in reporting.

Finally, the estimated trends in hospital outcomes may have been influenced by epidemic phase. As described in Section 2.7.2, a sensitivity analysis can be used to assess epidemic phase bias. This analysis should condition on the symptom onset date, rather than the hospital admission date, since there are several factors other than epidemic phase which may influence the time from infection to hospital admission. Whilst symptoms data were unavailable in SARI-Watch, this bias will be investigated among all hospitalised individuals in the following section.

4.5 Post-vaccine hospitalised severity

The extensive vaccination campaign in England during 2021 led to a marked reduction in the number of individuals being hospitalised, and in the overall risk of mortality (Lopez Bernal *et al.* 2021; Thygesen *et al.* 2021; UK Health Security Agency 2021). Nonetheless, individuals continue to experience COVID-19 infection severe enough to require hospitalisation. In 2021 there was little information about prognosis following hospital admission in the context of widespread vaccination, or how this might be related to hospital load.

In this section I investigate trends in mortality among all people hospitalised with community-onset COVID-19 in England over the first three waves (two years) of the pandemic, and how these trends have varied according to vaccination status, hospital load, and other factors. Relative and absolute risks of hospitalised fatality and lengths of stay in hospital by outcome are estimated using Aalen-Johansen estimators and Fine-Grey proportional hazards regression.

4.5.1 Data

The SUS dataset contains well-completed, accurate information on all hospitalisations for COVID-19 in England. For this analysis, data on all adults (aged 15+) with community-acquired COVID-19 (defined as a positive test for COVID-19 within -14 to 1 days of hospital admission), admitted to hospital in England for COVID-19 or another non-injury related condition during the first two years of the pandemic (March 2020 to April 2022) were considered. Data were

extracted in August 2022, allowing four months of follow-up to minimise the impact of lags in data reporting.

Since records are only entered into SUS upon completion of a hospital stay (i.e. at the point of discharge from hospital, or death) these were routinely supplemented with information on individuals still in hospital though linkage to ECDS, as well as data on vaccination and PCR testing for COVID-19, as described in Section 4.3. Among those with completed hospital episodes, 77% were admitted via emergency care.

Covariates

Covariates in the linked dataset included vaccination status (as shown in Figure 4.10), date of hospital admission, age group, region of residence, Charlson Comorbidity Index (CCI)(Charlson *et al.* 1987), ethnicity, sex, and index of multiple deprivation (IMD) quintile. A hospital load measure was derived from the number of COVID-19 admissions at an NHS trust within the 7 days around admission (3 before, same day, and 3 after). Hospital load was calculated as a proportion of the busiest 7-day period at that trust, grouped into: 0–20%, 20–40%, 40–60%, 60–80%, 80–90%, and 90–100%.

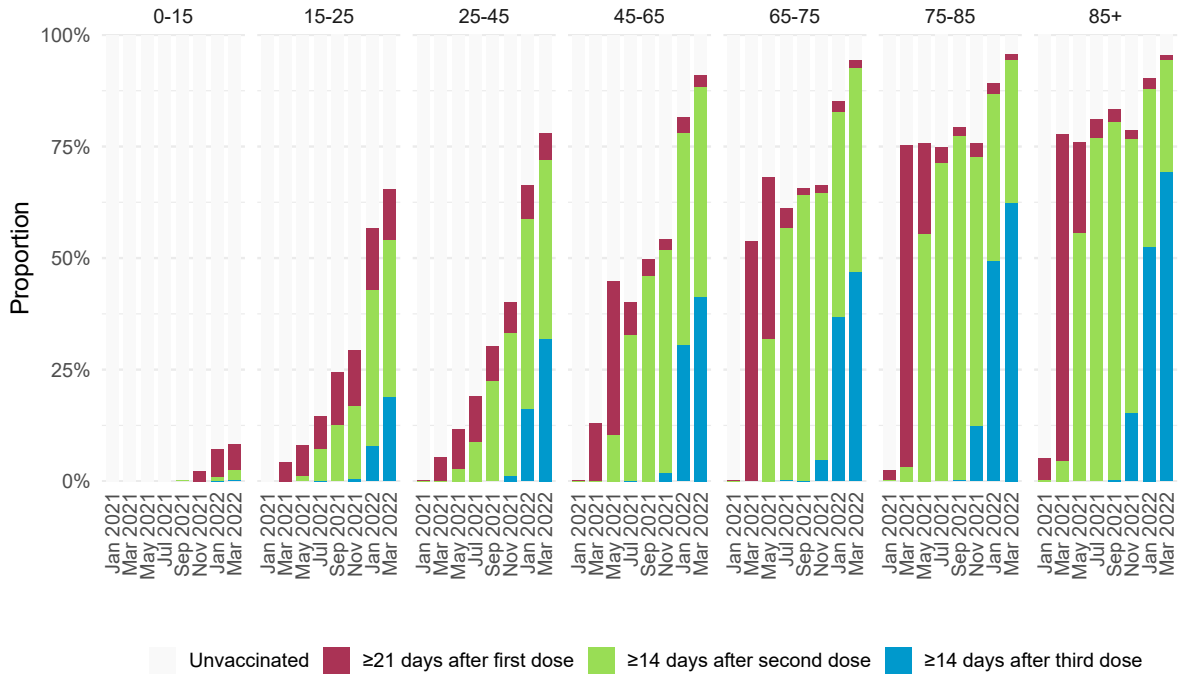


Figure 4.10: Vaccination status of hospitalised individuals in SUS data, by month of admission and age group, March 2020 to April 2022.

The number of reported admissions were compared with the NHS weekly COVID-19 admissions data to ensure data were representative (NHS England 2022). Hospital-onset

COVID-19 (i.e. infection occurring in hospital) cases were excluded: those who acquired COVID-19 infection in hospital during the study period (n = 209,139) tended to be older and have longer lengths of stay than the community-onset cases considered in this study, with a greater proportion remaining in hospital post-90 days (Table D.2). Hospital stays for these individuals may be influenced by conditions other than COVID-19, as described by Bhattacharya *et al.* 2021.

4.5.2 Multi-state model

Two competing risks survival models (Section 2.3) were applied to the SUS data to investigate the competing risks of death and discharge following hospital admission. Given the lack of information about ICU admissions in SUS, these risks are averaged over whether the admission included an ICU stay or not.

Aalen-Johansen cumulative incidence estimation was used to obtain estimates of cumulative HFR and median lengths of stay in hospital for specific sub-sets of the population, unadjusted for other factors. Median lengths of stay were the weighted medians from the Aalen-Johansen estimate with weighted ties (see Section D.2.1 for details).

Stratified Fine-Gray competing risk proportional hazards regression with adjustment for confounders was used to estimate the association of each risk factor with the cumulative incidence of mortality within 90-days of hospitalisation with COVID-19 (i.e. hospitalised fatality). The proportional hazards assumption was checked for each covariate, and stratification used for confounders with non-proportional hazards.

Stratification

For the regression on month of admission, stratification was by age group, region of residence and vaccination status, with regression adjustment on sex, ethnicity, IMD quintile and CCI. For the regression on vaccination status, stratification was by age group, region of residence and month of hospital admission, with regression adjustment on sex, ethnicity, IMD quintile and CCI.

Censoring

To focus these analyses on outcomes following COVID-19 admission, and for consistency with the previous analysis, a pragmatic cut-off of 90 days from first positive specimen date was chosen and only those hospital outcomes (death or discharge) occurring within this cut-off were included. All records with outcomes occurring beyond 90 days (n = 656) were right-censored at 90 days, while individuals who remained in hospital at the date of data extraction (n = 1019) were right-censored at the shorter of this date or 90 days.

Linkage to the UKHSA deaths data enabled post-hospital discharge deaths to be identified. To better account for palliative discharge, deaths occurring within 14 days of discharge from hospital (n = 15,523, 19.1% of all deaths observed) were classified as deaths rather than discharges, and the date of death used as the outcome date.

Consistency in model estimates

Figure 4.11 demonstrates the high degree of agreement between the Aalen-Johansen and Fine-Gray model cumulative fatality estimates for selected months over the study period.

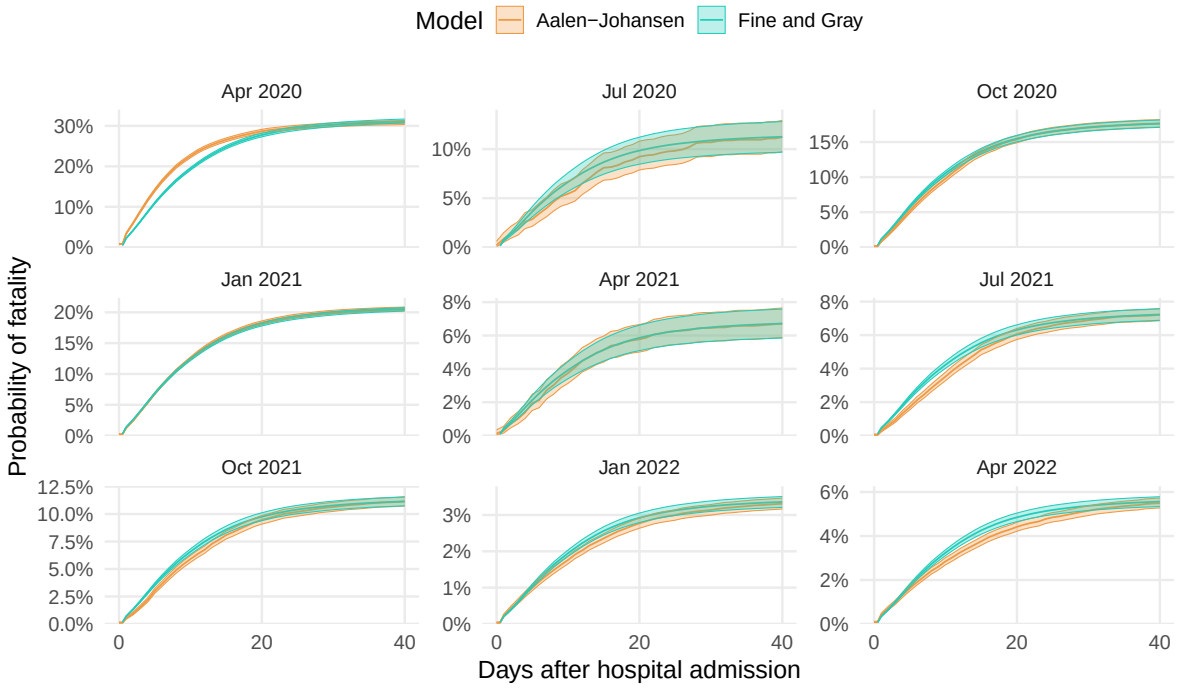


Figure 4.11: Aalen-Johansen and Fine-Gray cumulative fatality estimates for first 60 days following hospital admission in SUS data, with 95% CI, March 2020 to April 2022

Epidemic phase bias

Sensitivity analysis was used to assess the potential effect of epidemic phase bias on the estimated hazard ratios in relative risk analyses. For this analysis the date of symptom onset was conditioned on, rather than the date of hospital admission, to ensure that the sensitivity analysis targeted bias due to epidemic phase, as opposed to any other factors which may influence time from symptom onset to admission.

A time shift of $c = 0, 1, 2, 3, 4$ days was added to records with a mortality outcome, where c should represent the mean difference in time from infection to symptom onset date between

those who died and those who did not. Based on previous analyses of COVID-19 data this time was not expected to be greater than 4 days (Seaman, Nyberg *et al.* 2022).

Estimation

Hospitalised fatality risks and lengths of stay were estimated using the Aalen-Johansen estimator. Sub-distribution hazards were estimated using Fine-Gray proportional hazards regression (see Section 2.3).

Statistical models were implemented using maximum likelihood estimation with the open-source packages `survival` and `matrixStats` in `R` (Bengtsson 2007; R Core Team 2020; Therneau 1999).

4.5.3 Results

Participant characteristics are shown in Appendix D.2. Among 590,313 people hospitalised with COVID-19 between March 2020 and April 2022, a total of 81,343 (13.8%) died, 507,254 (85.9%) were discharged, and the remaining 1716 (0.3%) remained in hospital at the date of data extraction and/or were right-censored at 90 days. Compared to all people with PCR-confirmed community-acquired COVID-19 infection, those hospitalised for COVID-19 were older (44% vs. 10% aged over 65), more likely to live in an area of high deprivation (27% vs. 24%), and to reside in London (18% vs. 16%).

Hospitalised fatality risk over the study period is shown in Figure 4.12 panel A. HFR decreased during the first wave of the pandemic from 34.2% (95% CI 33.5 to 34.9%) in March 2020 to 10.7% (95% CI 9.1 to 12.5%) in August 2020. In the second wave, HFR peaked at 21.0% (95% CI 20.7 to 21.2%) in January 2021, but by March 2021 had halved to 9.5% (95% CI 8.9 to 10.2%). The peak was smaller again in the third wave, at 11.5% (95% CI 11.1 to 11.9%) in October 2021.

Patients with an eventual outcome of death generally had longer stays in hospital compared to those who were discharged (Figure 4.12 panel B). Lengths of stay prior to discharge decreased throughout the first wave, from 6.7 days (95% CI 6.6 to 6.7 days) in March 2020 to 3.4 days (95% CI 3.2 to 3.5 days) in August 2020. Over the same time period, lengths of stay prior to death increased from 6.0 days (95% CI 6.0 to 6.1 days) to 9.8 days (95% CI 9.3 to 10.4 days). Fluctuation in lengths of stay were estimated throughout the second and third waves, in parallel with trends in HFR.

Older individuals and those with multiple co-morbidities were more likely to die or else experienced longer stays prior to discharge (Figures D.6 and D.8). Whilst trends are uncertain in months with fewer admissions, those admitted to hospitals experiencing higher load had slightly greater HFR than to those admitted to hospitals with lower activity, e.g. during March

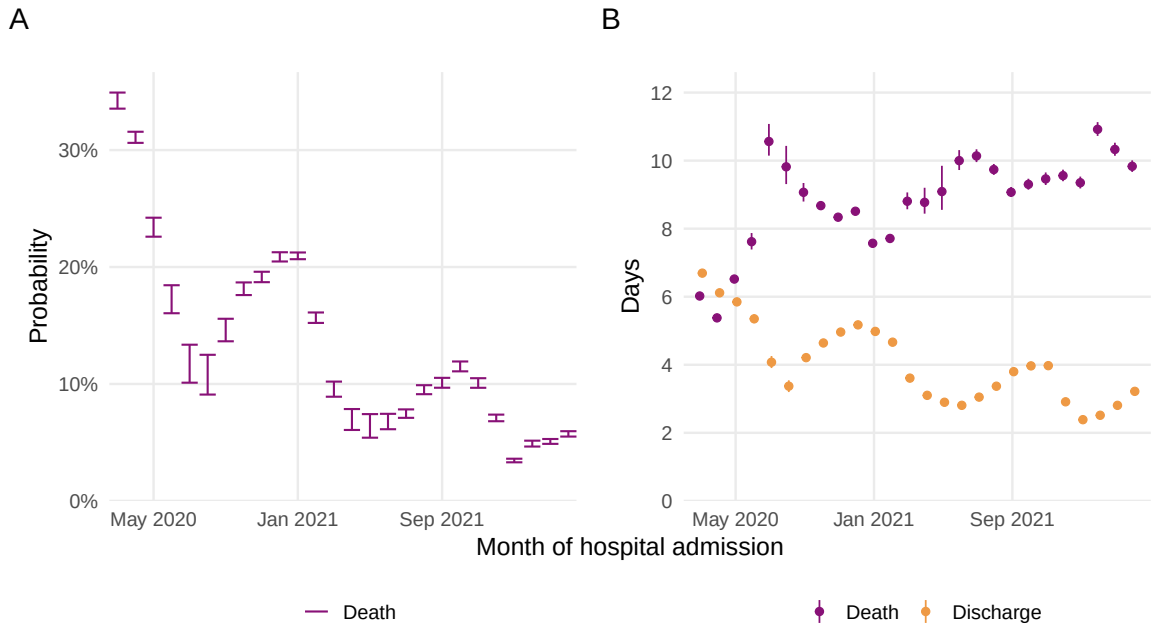


Figure 4.12: Hospitalised fatality risk (panel A) and median length of stay (panel B) by month of hospital admission. SUS data, with 95% CI, March 2020 to April 2022. Unadjusted for other covariates.

2020: 35.7% (95% CI 32.7 to 38.9%) for hospitals at 90–100% of their peak load, compared to 32.8% (95% CI 31.1 to 34.5%) for hospitals at 0–20% of their peak load, and similarly for November 2020: 17.6% (95% CI 15.6 to 19.7%) compared to 15.8% (95% CI 15.0 to 16.7%) (Figure D.9).

Figure 4.13 shows hazard ratios for hospitalised fatality by month of hospital admission. Controlling for age group, region of residence, vaccine status, sex, ethnicity, IMD quintile, CCI, and hospital load, month of hospital admission remained a significant factor for the prognosis of hospitalised individuals. Compared to June 2020, the hazard for hospitalised fatality was significantly increased during March to May 2020, and September 2020 to February 2021, with much lower HFR estimated at the start of 2022.

Compared to unvaccinated people, controlling for month of hospital admission and the factors mentioned above, the hazard ratio of hospitalised mortality was 0.87 (95% CI 0.80 to 0.94) during the first 3–6 weeks post-first dose vaccination, this increased to 0.96 (95% CI 0.80 to 0.94) for individuals hospitalised during the 6–12 weeks post-first vaccine dose. Hazard ratios were lower following a second vaccine dose, 0.71 (95% CI 0.59 to 0.85) for the first 2–6 weeks post-second vaccine dose, although with similar waning of this protection in the following weeks. A third dose reduced the hazard ratio of hospitalised mortality to 0.68 (95%

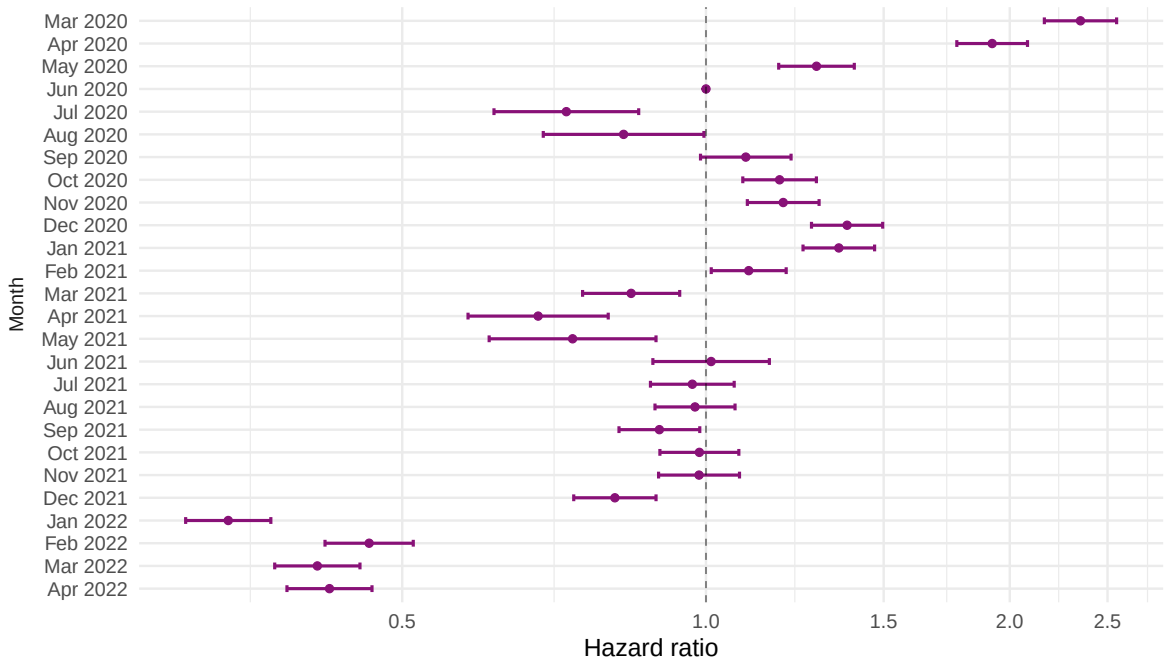


Figure 4.13: Hospitalised fatality sub-distribution hazard ratio and 95% CI by month of hospital admission. SUS data, March 2020 to April 2022. Adjusted for: age group, region, vaccine status, sex, ethnicity, IMD quintile, CCI, and hospital load. Reference group: June 2020.

CI 0.63 to 0.74) in the first 2–6 weeks, and this protection appeared to be sustained for longer (Figure 4.14).

Hazard ratios associated with other patient characteristics are shown Figure D.10. There was an increased hazard of hospitalised fatality among those of Asian ethnicity (1.12 (95% CI 1.08 to 1.17)) but reduced among those of Black ethnicity (0.87 (95% CI 0.83 to 0.92)) compared to reference category White. Females had a lower hazard of fatality compared to males (hazard ratio 0.80 (95% CI 0.78 to 0.81)), and compared to a CCI of 0, those with a higher burden of comorbidity had a greater hazard of hospitalised fatality (hazard ratio 3.91 (95% CI 3.76 to 4.07) for CCI of 5 and above). Those residing in more deprived quintiles had greater hazards for hospitalised fatality (hazard ratio 1.14 (95% CI 1.10 to 1.18) for the most deprived quintile) compared to a reference of least deprived.

Supplementary Figure D.11 shows the shift sensitivity analyses by month of symptom onset, adjusted for the same covariates as above. The greatest effect was observed for the March 2020 estimate, which steadily reduced towards 1 following a shift of $c = 1, 2, 3$ or 4 days. The effect in other months was small, with the previously described monthly trends persisting.

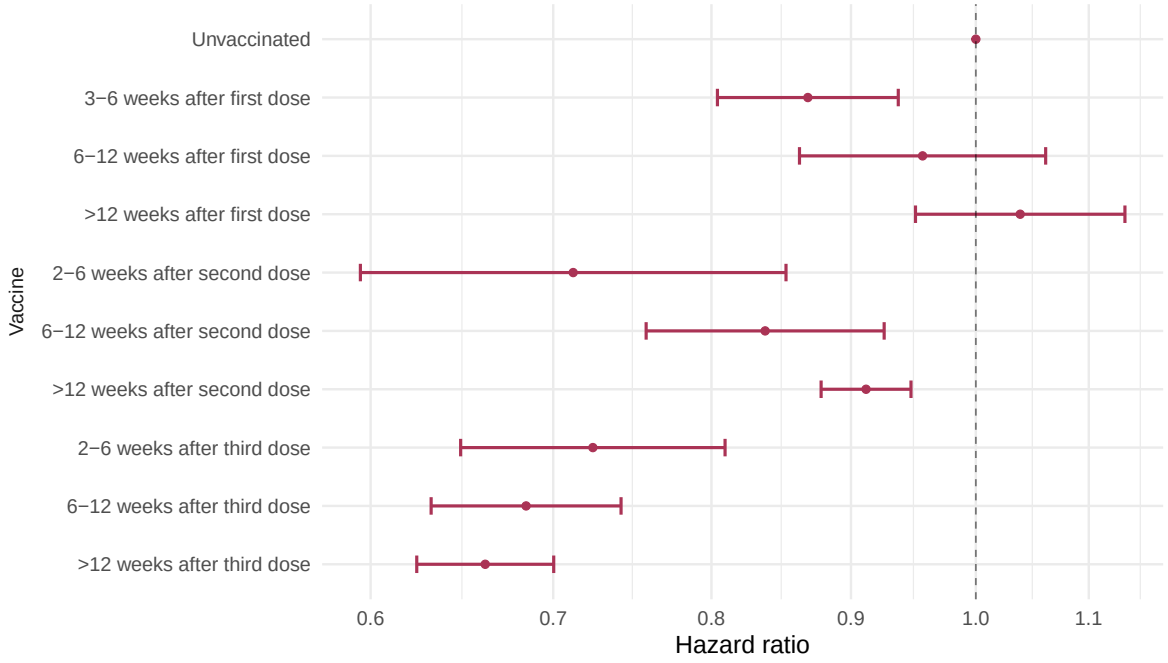


Figure 4.14: Hospitalised fatality sub-distribution hazard ratio and 95% CI by vaccine status. SUS data, March 2020 to April 2022. Adjusted for: age group, region, month of admission, sex, ethnicity, IMD quintile, CCI, and hospital load. Reference group: Unvaccinated.

4.5.4 Discussion

The prognosis for people hospitalised with COVID-19 in England during the first three waves of the COVID-19 pandemic has varied substantially according to month of admission, case-mix, and vaccination.

In this analysis, greater absolute hospitalised fatality risks were estimated for men, older individuals, and those with multimorbidity, and HFR varied according to ethnicity, month of admission, and hospital load. Lengths of stay in hospital were similarly associated with demographic factors, with median lengths of stay prior to death generally longer than those prior to discharge. In relative risk analyses controlling for all measured confounders, baseline comorbidity burden was the strongest predictor of mortality. Being admitted during a period of high hospital load was also associated with poorer outcomes, but patients received a high standard of care overall, with lockdown measures likely protecting against hospitals becoming excessively overburdened (Davies, Barnard *et al.* 2020).

A significant reduction in HFR was seen in early 2022, despite a peak in the number of infections in the community (Figure 4.1). This reduction in HFR may be related to the emergence of the Omicron (B.1.1.529) variant in December 2021, which became the dominant circulating variant in January 2022. In complementary analyses of all PCR-confirmed infections

in England, Nyberg, Ferguson *et al.* 2022 estimated substantially lower risks of hospitalisation and severe events following Omicron infection, as compared to infection with the previous Delta (B.1.617.2) variant.

Estimated risk of hospitalised fatality risk was lower among vaccinated individuals, with the effect most clearly seen in older individuals. For those aged 75 and over, vaccination reduced HFR to approximately the risk of an unvaccinated individual aged 10 years younger. In adjusted estimates, each additional vaccine dose significantly reduced the hazard of hospitalised fatality, with some waning of protection following the first and second doses. A national population cohort study in Scotland reported significantly lower risks of severe events (hospitalisation and mortality) following the first vaccine booster (Agrawal, Katikireddi *et al.* 2021), and waning protection of vaccination against COVID-19-related mortality was reported among all persons in England (Andrews, Tessier *et al.* 2022), and for hospitalised cases in Sweden (Nordström *et al.* 2022). The sustained protection following a third vaccine dose may be related to the lower severity of the Omicron variant.

The estimates in this section have provided an indication for demands on hospital resources during the pandemic, and provide further evidence of the relationship between baseline covariates and hospitalised patient outcomes. Preliminary estimates were presented to expert scientific panels during the vaccine roll-out period, and as new variants of COVID-19 emerged, although real-time analyses of the SUS dataset were somewhat limited as a result of delays in data reporting. The methodology and results based on data for the first two COVID-19 waves were published in 2022 (Kirwan, Charlett *et al.* 2022).

Biases and limitations

Hospital pathway data were unavailable in this analysis, hence risks of critical care admission and need of respiratory support could not be explored. This study also focused only upon individuals hospitalised for COVID-19, excluding nosocomial cases and those not hospitalised. Nosocomial cases tended to be more severe, as shown in Table D.2, whereas the vast majority of those not requiring hospitalisation go on to recover with relatively mild illness.

Data validation was undertaken between the linked datasets to address missing data bias, with no systematic under-reporting or misreporting of patient characteristics. The study population was focused on those hospitalised *for COVID-19*, as opposed to those hospitalised *with COVID-19*, and all NHS trusts in England were considered in order to obtain representative estimates.

All measured confounders were adjusted for, including hospital load, month of study, and vaccination. Data on treatment and variant were unavailable, but as these changes occurred over time they are likely to be included in the temporal variation in hazards. There was very minimal effect of epidemic phase bias on the estimates, grouping by month of admission (rather

than day or week) helps to minimise this bias as only a relatively small proportion of patients are affected each month.

Characterisation of hospital pressures

There are several potential biases which make the measure of hospital load hard to define and interpret: during periods of peak hospital load there is likely a modification in an individual's willingness to attend healthcare services for mild illness, changes occurred in the admission criteria both to wards and ICUs (NHS England 2020a), and individuals with milder disease may have been selected for transfer from overloaded hospitals to those with greater bed availability.

A load-based marker is also a crude measure of hospital pressure, as despite being associated with increased hazard of fatality, it only accounts for the 'demand' on services, and not the 'supply' capacity of the healthcare service. An improved hospital pressure marker would provide actionable information to the NHS about which trusts are feeling the most extreme pressure, in terms of both supply and demand, and may allow for improved targeting of support.

Selection biases

This chapter has included analysis of both a real-time sentinel system (SARI-Watch) and a comprehensive reporting system with delayed reporting (SUS). Combining these two data systems may enable the strengths to be exploited, whilst overcoming selection biases and reporting delay. Evidence synthesis methods are a principled way to combine data from two sources, with the relevance (lack of external bias) and rigour (lack of internal bias) of each dataset considered (Turner *et al.* 2009; Welton and Ades 2005).

Causal inference from observational hospital data

To make causal inference in this study of observational hospital data an explicit causal question should be defined, e.g. the effect of a well-specified hypothetical intervention (Hernán and Robins 2023). Relevant questions of interest might include: the causal effect of vaccination on patient outcomes, or the causal effect of of interventions to mitigate pressure on hospitals (e.g. delaying elective procedures) on patient outcomes.

Further discussion about potential future methodological development is included in Chapter 6.

4.6 Chapter summary

In this chapter I have applied competing risks multi-state models to surveillance data from hospitals in England to estimate hospitalised fatality rates, addressing potential sources of bias

and adjusting for censored outcomes and reporting delay. Whilst only a few hospital pathways were considered due to data limitations, with sufficient data these models could be extended to more detailed hospital analyses, e.g. respiratory support, and step-ups and step-downs from intensive care.

A parametric mixture model was applied to sentinel data from the pre-vaccine era. The application of a parametric model to a prospective dataset enabled projections from the sentinel hospital data to be made. These models were applied in real-time to predict the severity of emerging COVID-19 variants during the pandemic, and could be applied other novel pathogens.

A stratified proportional hazards model was used to condition on several covariates at once across the pre and post-vaccine eras. Hazards were estimated by month and by vaccination status to provide fully adjusted estimates for a comprehensive and well-completed dataset of all hospital admissions in England. Temporal trends were estimated, alongside waning protection of vaccination against hospitalised fatality, and this research has informed vaccination policies for those most at risk of adverse COVID-19 outcomes.

5. Effectiveness of interventions

Whilst the *efficacy* of an intervention can be measured in a controlled setting, the *effectiveness* of the intervention in the real-world is often of greater interest, and vital to inform policy decisions about its roll-out (Ortiz and Neuzil 2022). Assessing the effectiveness of an intervention requires careful consideration of the study design, data collection mechanism, and statistical methodology to be used. Vaccine effectiveness estimation can be particularly challenging, as the reasons individuals choose to receive a vaccine may well be related to their risk of infection. For example, those more likely to be exposed to the virus may be more likely to receive the vaccine, and also more likely to become infected. If not accounted for, this confounding can lead to biased estimates of vaccine effectiveness.

5.1 Aims of this chapter

During the COVID-19 pandemic, healthcare workers have been at increased risk of infection and prioritised for vaccination against SARS-CoV-2. Understanding the effectiveness of COVID-19 vaccines in this working-age group provides important evidence to inform wider potential vaccine roll-outs to the general population.

Randomised controlled trials (RCTs) are considered the gold standard for assessing vaccine efficacy, but can be infeasible in the real-world as they are often time-consuming and costly, and because there are ethical considerations in randomising participants to receive a vaccine once it is known to be effective (Ottoboni and Poulos 2020; Stoehr *et al.* 2021). Observational studies (particularly case-control studies) of vaccine effectiveness are more common, but may be subject to confounding and other types of selection bias (Shi *et al.* 2023; Sullivan *et al.* 2016).

In this chapter I aim to assess COVID-19 vaccine effectiveness and protection from prior infection in a cohort of UK healthcare workers. Vaccine effectiveness among this cohort was initially investigated using Cox proportional hazards models. To account for a reduction in testing frequency later in the study, and to make use of the full information in the data, I aim to develop and apply multi-state survival models to investigate bivalent booster vaccine effectiveness against both symptomatic and asymptomatic infection. I also aim to understand the causal effect of vaccination by modelling counterfactual scenarios.

5.2 Background

5.2.1 UK COVID-19 vaccine roll-out

The SARS-CoV-2 vaccination strategy chosen for the UK aimed to rapidly reduce hospitalisations, severe outcomes, and preventable deaths from COVID-19 by initially targeting those at high risk of severe COVID-19 (including care home residents and people aged 80 years and over), and front-line healthcare workers, recognising this latter group's high occupational exposure, their potential role in nosocomial transmission, and the significant impact of healthcare worker absence on healthcare delivery (Department of Health and Social Care 2020b; Department of Health and Social Care 2021c).

The BNT162b2 mRNA (Pfizer-BioNTech) and ChAdOx1 adenoviral (Oxford-AstraZeneca) COVID-19 vaccines were approved in December 2020 on the basis of phase 3 randomised controlled trials (RCTs) (Polack *et al.* 2020; Voysey *et al.* 2021). That same month, the Joint Committee on Vaccination and Immunisation (JCVI) recommended delaying the second vaccine dose by up to 12 weeks to optimise the public health impact of vaccination (Department of Health and Social Care 2021a). The majority of healthcare workers were vaccinated with a long interval between doses (Hall, Foulkes, Insalata *et al.* 2022).

In autumn 2021, amid mounting evidence of waning vaccine effectiveness several months post-administration of the second vaccine dose (Andrews, Tessier *et al.* 2022; Chemaitelly *et al.* 2022; Goldberg *et al.* 2021; Hall, Foulkes, Insalata *et al.* 2022; Nordström *et al.* 2022; Šmíd *et al.* 2022), third mRNA vaccine doses (first booster doses) were widely deployed. Again, vaccination was initially targeted to front-line healthcare workers and those considered at higher risk of severe disease (Joint Committee on Vaccination and Immunisation 2021).

A year later, during a period of dominance of Omicron sub-variants, which had higher immune escape capability from vaccination and previous infection (Liu *et al.* 2022; Willett *et al.* 2022), fourth 'bivalent' mRNA doses (second booster doses) containing sequences from both the original SARS-CoV-2 strain and Omicron sub-variants were offered to priority population groups (Medicines and Healthcare products Regulatory Agency 2022a; Medicines and Healthcare products Regulatory Agency 2022b).

By the end of 2023, more than 24 million SARS-CoV-2 infections had been reported and more than 150 million doses of COVID-19 vaccine delivered in the UK (Mathieu *et al.* 2020), conferring some degree of immunity (whether from infection, vaccination, or a hybrid of both) to most of the population. Despite the public health emergency being declared over by the WHO in May 2023 (World Health Organisation 2023), and budgets for testing and vaccination being scaled back (HM Government 2022), population-level studies remain relevant to monitor the protection of existing levels of immunity against emerging strains of SARS-CoV-2, and to assess the real-world impact of booster vaccines which target these variants (GOV.UK 2023).

5.2.2 Fourth dose vaccine effectiveness estimation

A variety of study designs and statistical methodologies are used to estimate vaccine effectiveness (VE) against different outcomes. To assess the effectiveness of the bivalent/fourth booster COVID-19 vaccines against infection, recent studies have implemented both Cox proportional hazards models in cohort studies (Auvigne *et al.* 2023; Huiberts *et al.* 2023; Shrestha *et al.* 2023), and multivariable logistic regression in test-negative case-control studies (Link-Gelles *et al.* 2023; Plumb *et al.* 2023).

Each of these these studies reported evidence of protection of bivalent/fourth booster mRNA vaccination against symptomatic COVID-19 infection during the Omicron variant circulating period, with substantial waning of protection from 2-months post-vaccination. The range of VE estimates varies widely, however, from 49% (95% CI 41 to 55%) during December 2022–January 2023 among all adults aged 18-to-49 with COVID-19-like illness (Link-Gelles *et al.* 2023), to 4% (95% CI -12 to 18%) among health clinic employees during September 2022–March 2023 (Shrestha *et al.* 2023).

Hybrid immunity was investigated by several of the cohort studies, with recent prior infection estimated to provide greater protection than booster vaccination, although this protection waned 4–6 months post-infection (Auvigne *et al.* 2023; Huiberts *et al.* 2023).

The studies referenced above relied on self-reporting of symptomatic COVID-19 infection, or on individuals seeking testing (most likely as a result of COVID-19 symptoms). Among these, only one study discussed the limitations of their study in relation to selection bias, test-seeking behaviour, reporting biases, and unmeasured confounding (Plumb *et al.* 2023). Shi *et al.* 2023 have shown how failing to account for these biases in test-negative case-control studies can lead to invalid estimates of VE against symptomatic COVID-19 infection.

In the next section I describe a healthcare worker cohort study, in which participants undergo sampling at pre-specified intervals, enabling both symptomatic and asymptomatic infection to be detected. Further on, I explain how a cohort study design mitigates specific biases of test-negative case-control studies, and the utility of multi-state models for investigating the interval-censored data collected in these studies.

5.3 The SIREN study

The SARS-CoV-2 Immunity & Reinfection Evaluation (SIREN) study, run by UKHSA, is a large-scale prospective cohort study of UK healthcare workers undergoing fortnightly PCR testing and running continuously since June 2020 (Wallace *et al.* 2022). The study aims to understand the extent to which prior infection and vaccination protects against future SARS-CoV-2 infection. Epidemiological and statistical analyses of SIREN data have so far investigated COVID-19 incidence rates, vaccine coverage, and protection from prior infection

and vaccination (Hall, Foulkes, Insalata *et al.* 2022; Hall, Foulkes, Charlett *et al.* 2021; Hall, Foulkes, Saei *et al.* 2021).

5.3.1 Setting and participants

SIREN study participants are staff working in healthcare organisations in the UK. At enrolment, participants completed a demographic questionnaire and provided blood serum and nasal swab samples for antibody and PCR testing. Participants provided fortnightly nasal swab samples for PCR testing and completed a fortnightly symptoms questionnaire, as well as undergoing regular antibody testing.

All NHS hospitals and health boards in the UK were invited to join SIREN, with no random sampling of hospitals, health boards, or participants. Over 40,000 participants were initially recruited (Wallace *et al.* 2022). Participant follow-up was fixed at 12 months, with subsequent extensions to 24 months, and to end-March 2023 for those choosing to remain enrolled. Figure 5.1 shows the number of participants actively enrolled in SIREN between June 2020 and March 2023, alongside their recruitment status.

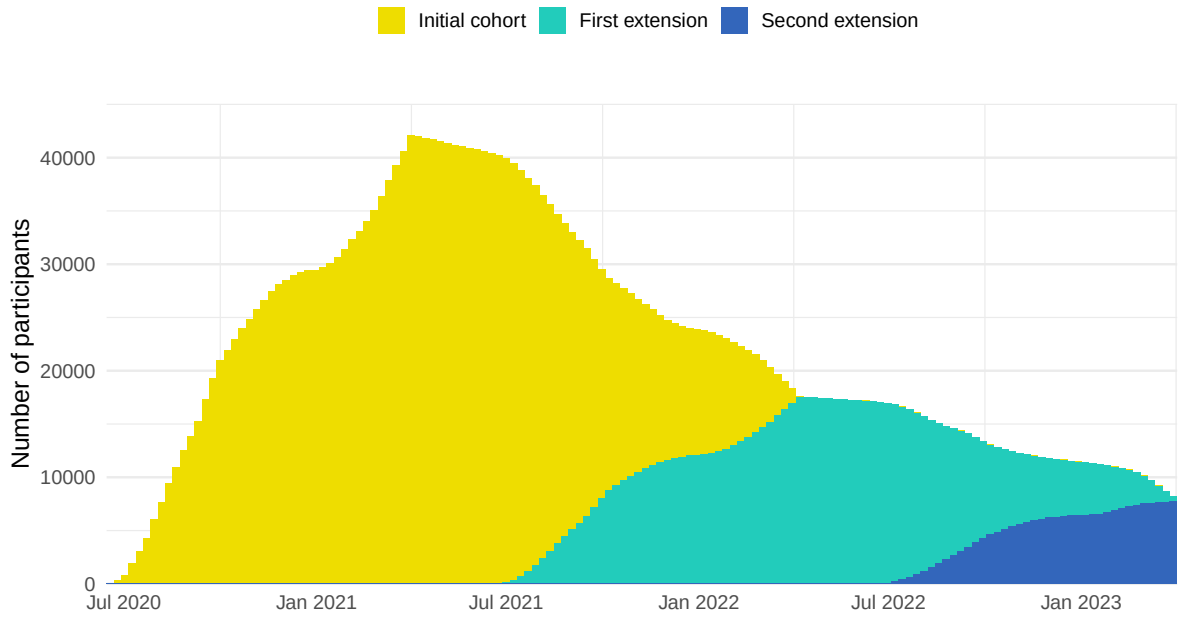


Figure 5.1: Number of participants actively enrolled in SIREN by recruitment status, June 2020 to March 2023.

5.3.2 Information collected

COVID-19 vaccination information, and results from PCR and antibody testing were collected centrally by UKHSA and linked to SIREN participant responses. Results from PCR testing undertaken for other reasons (e.g. symptomatic testing, contact tracing, etc.) or prior to study entry were also included in the linked dataset.

Anti-N (to confirm previous infection) and anti-S (to confirm previous infection and/or response to vaccination) antibody tests were used to identify if participants had an infection prior to initial recruitment into the study. Post-recruitment, information from fortnightly PCR testing was used to identify if an infection has occurred, with the date of a positive PCR test used as a proxy for the onset of infection.

Because of intermittent testing and/or infections clearing within the fortnightly window, participants may have received a positive anti-N result without a positive PCR test. In these cases previous analyses have used the date of a positive antibody result as a proxy for infection, but the gaps in testing (i.e. 'interval-censoring') could be accounted for with an appropriate methodology. Meanwhile, negative anti-N antibody results provide serological evidence of a participant's infection-naïve status (UK Health Security Agency 2022a).

Information on sequenced variant, cycle threshold (Ct) values, and T-cell counts were available for a subset of SIREN study participants. Sequencing data can be used to determine COVID-19 variant and SIREN sequences were cross-referenced with national data to identify variant circulating periods (UK Government 2021). Ct values indicate the concentration of viral genetic material in a sample, with low Ct values (corresponding to a high viral load) indicating acute disease and high infectivity (Public Health England 2020b). However, Ct values are not comparable between different assays and vary considerably according to sampling procedure and laboratory set-up (Dahdouh *et al.* 2021). Alongside antibodies, T-cells provide protection against SARS-CoV-2 infection, and the T-cell immunity induced by vaccination and natural infection was measured and characterised in a sub-study of SIREN participants (Angyal *et al.* 2022).

5.3.3 Characteristics of study population at enrolment

Demographic characteristics of the study population were self-reported at enrolment and included: age, gender identity, ethnicity, region of residence, occupational setting, staff type, medical conditions, and household structure. The linked testing and vaccination data included: PCR and antibody dates and results, and vaccination dose, date, and manufacturer (with most receiving mRNA vaccines).

As shown in Figure 5.2 panel A, the majority of participants who enrolled in the study reported female gender identity, and most were employed as nurses, doctors, or in other

patient-facing roles. Most were aged between aged between 25–65, and among these the majority lived with others (other adults, or both adults and children) (Figure 5.2 panel B).

Most study participants were of white ethnicity (87.8%), and the greatest proportion were from the South of England (South West, South East or London), with smaller numbers of participants in the North East. Three-quarters (75%) had no chronic medical conditions, although 2.2% (n = 969) reported immunosuppression. Compared to the UK population, participants for whom deprivation information was available lived in less socio-economically deprived areas on average: 25.6% in the least deprived quintile and 11.4% in the most deprived quintile (Table E.1).

The cohort broadly reflected the demography of UK healthcare workers and whilst the demographic characteristics of the cohort did not differ greatly between the initial cohort and first extension, younger people (aged <25 and 25–34) were particularly under-represented in the second extension (Table E.1).

5.3.4 Infections

Figure 5.3 shows the weekly number of new COVID-19 infections detected through fortnightly PCR testing among SIREN participants over the course of the study. Peaks in the number of new infections appear soon after a change in the predominant circulating variant (inferred from those with sequencing information available) and coincide with peaks in infections detected among the UK population (UK Government 2021).

Questionnaires completed by participants within a 14-day window of a positive PCR test (to minimise recall-bias) were used to distinguish between infections with and without COVID-19 symptoms. For this investigation, infections were grouped into two categories based on reported symptoms: COVID-19-specific symptoms (any of: cough, fever, sore throat, anosmia (loss of smell), and/or dysgeusia (loss of taste)), and asymptomatic for COVID-19 (absence of symptoms, or only non-specific symptoms such as fatigue and muscle ache).

Symptoms varied according to variant, with a greater proportion of participants reporting COVID-19 symptoms during the Delta, BA.1, and BA.2 variant circulating periods compared to the Wild type or later variants (Figure 5.4 panel A). Prevalence of symptomatic infection did not differ by gender, but younger participants were more likely to report COVID-19-specific symptoms as compared to older participants (Figure 5.4 panel B).

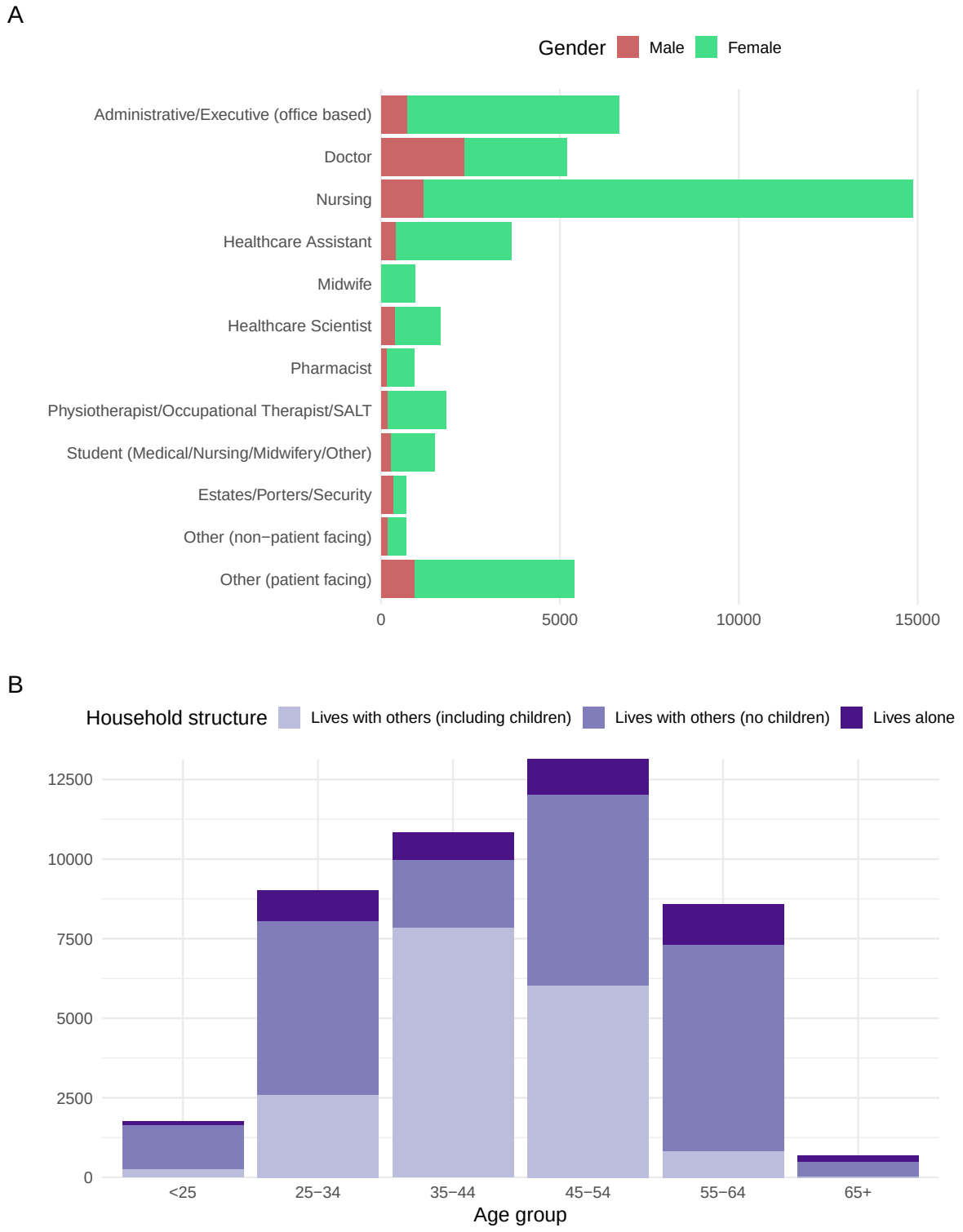


Figure 5.2: SIREN participant demographics, by occupation and gender identity (panel A) and age group and household structure (panel B).

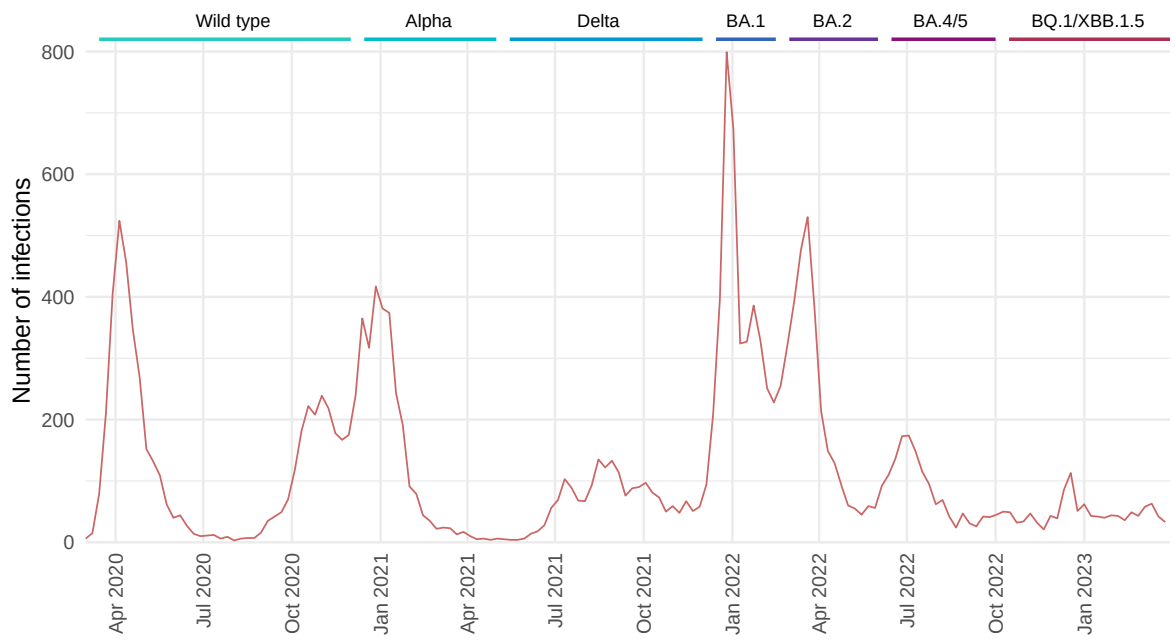


Figure 5.3: Weekly number of new COVID-19 infections detected among SIREN participants, June 2020 to March 2023. Predominant circulating variant(s) are indicated above.

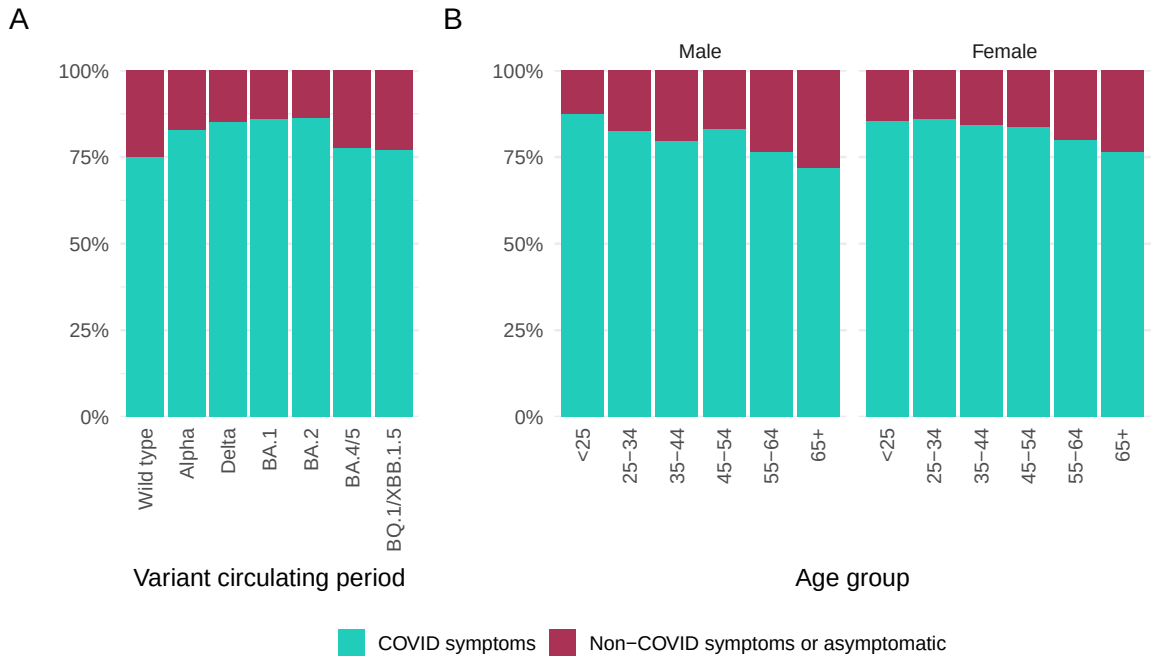


Figure 5.4: Self-reported symptoms among SIREN participants testing positive for COVID-19 between June 2020 to March 2023, by circulating variant period (panel A), and age and gender (panel B).

5.3.5 Vaccination

Primary (first and second dose), and first booster (third dose) COVID-19 vaccine coverage was high and rapid, with the majority of participants accepting a vaccine within 2-months of vaccine availability (Figure 5.5, panels A and B). Coverage of the second booster (fourth dose) was notably slower and less widespread, with a smaller proportion of participants receiving a fourth dose within 60 days of vaccine availability (64.3%), compared to third doses (83.0%). Fourth dose vaccine coverage within 60 days was particularly low in Scotland compared to other UK regions (Figure 5.6).

Younger participants were less likely to receive a fourth dose (59.7% for those aged 25–34 compared to 77.2% for those aged over 65) (Figure 5.7, Panel A), whereas infection-naïve participants were slightly more likely to receive a fourth dose than those with a recent COVID-19 infection (71.8% vs. 65.6%) (Figure 5.7, Panel B). Fourth dose vaccine coverage by other demographic characteristics is shown in Table E.2.

5.3.6 Fortnightly testing

PCR testing in SIREN was performed through several means, which varied by site. Testing pathways included: on-site nurse-led testing, self-testing with samples being dropped off by participants at an on-site drop-box, and self-testing with participants returning the samples by post ($n = 3,500$ individuals).

Figure 5.8 shows a timeline of participation in the SIREN study and vaccination status at each time point for a subset of randomly selected participants. The date of enrolment and PCR test dates are overlaid onto the timeline, demonstrating substantial variation among participants; some tested frequently and regularly, whilst others tested at irregular intervals and/or missed several tests.

Whilst participants may have missed tests for a variety of reasons, including annual leave, shift changes, or illness, there was some evidence of testing fatigue over time: 83.0% of participants underwent PCR testing according to the fortnightly schedule during the Wild type variant circulating period, and 80.7% during the Delta period, but this reduced to 58.7% during the BQ.1/XBB.1.5 period (Figure 5.9).

Lower adherence to the fortnightly testing schedule coincided with the removal of non-pharmaceutical interventions, and return to more normal population mixing (Department of Health and Social Care 2020a; Smith *et al.* 2022).

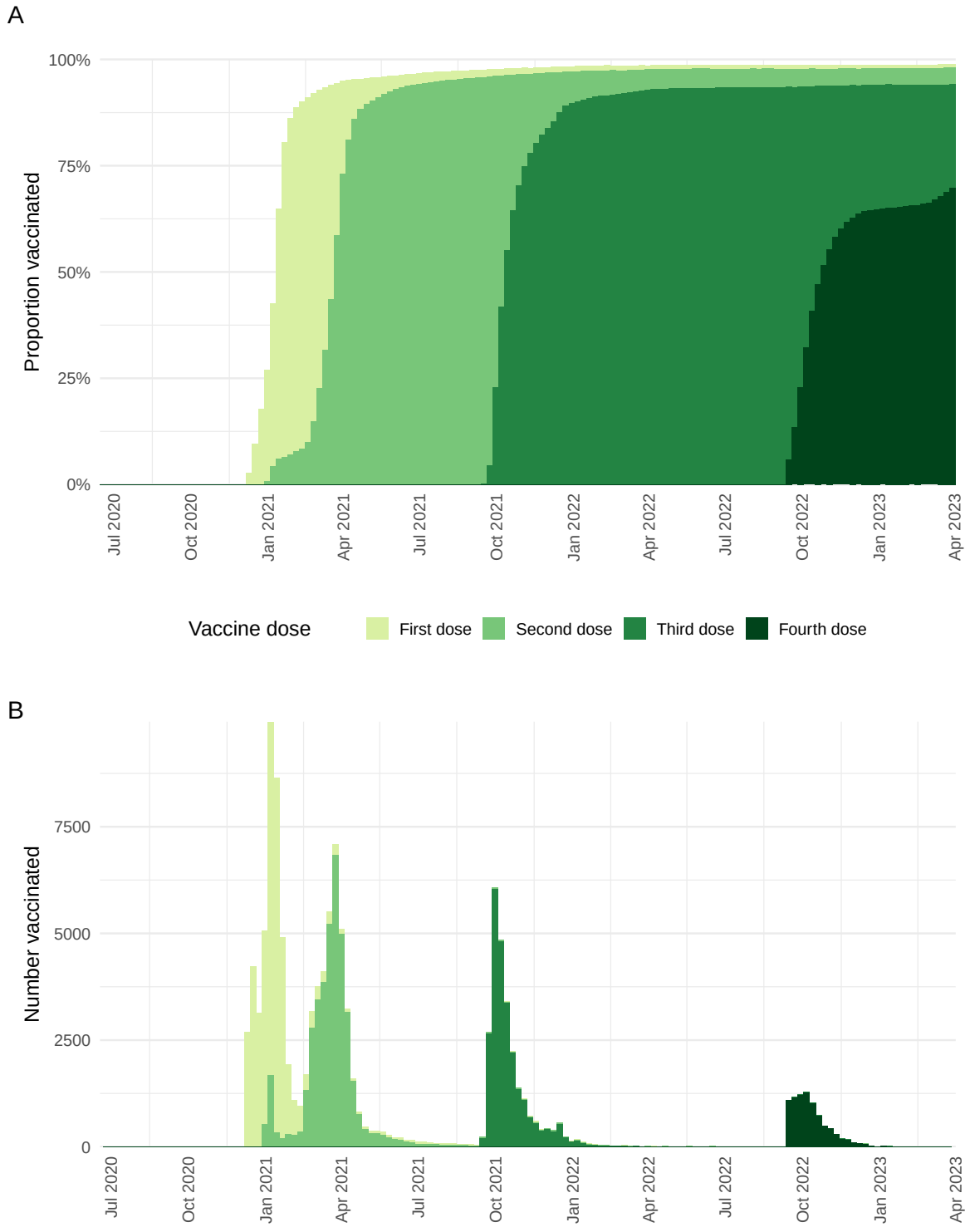


Figure 5.5: Proportion of SIREN participants vaccinated (panel A) and number of participants vaccinated each week (panel B), by vaccine dose, June 2020 to March 2023.

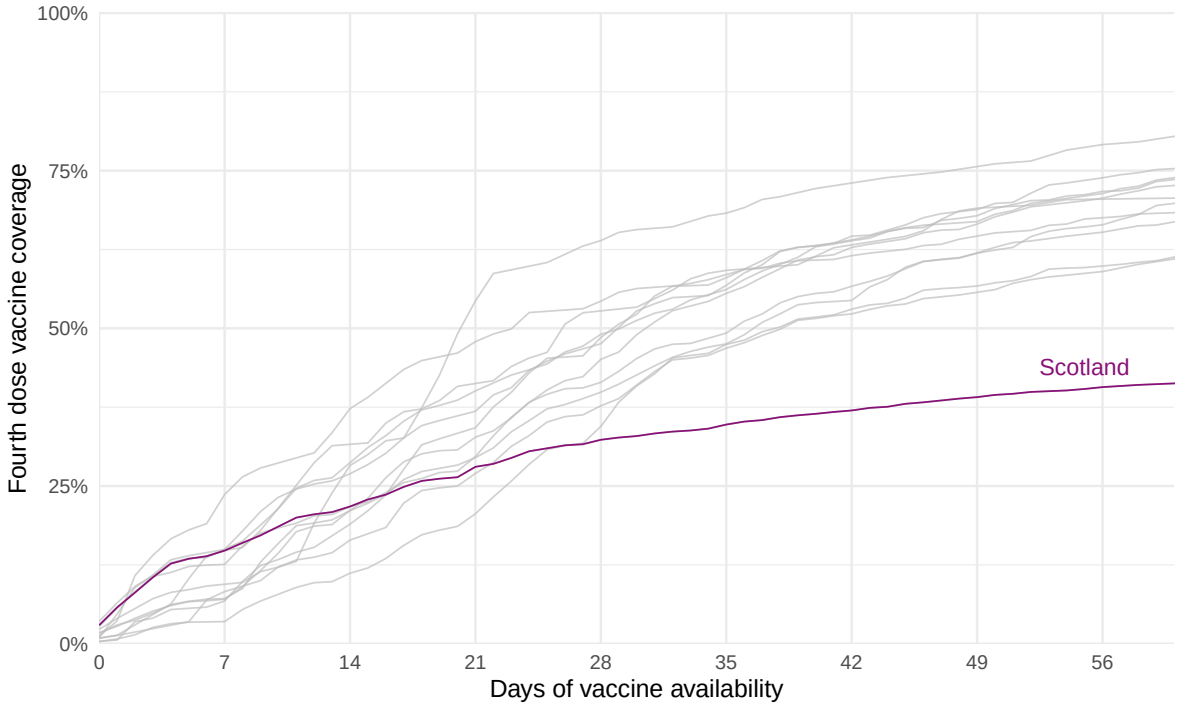


Figure 5.6: Fourth dose vaccine coverage among SIREN participants within 60 days of vaccine availability, by region of residence.

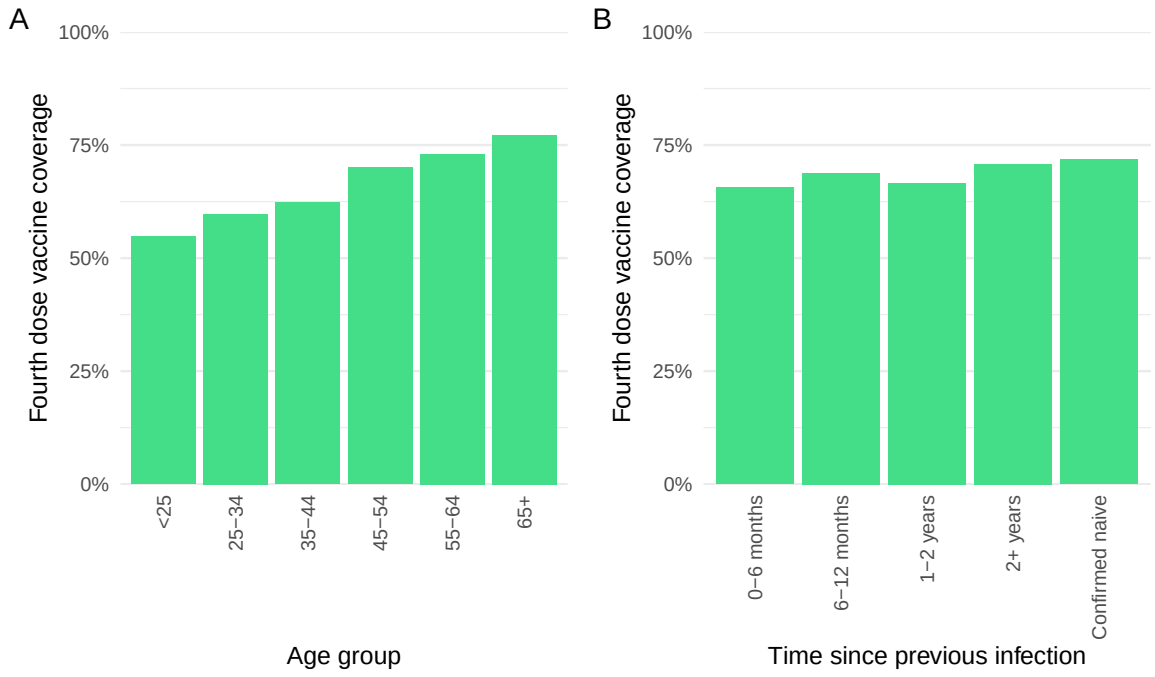


Figure 5.7: Fourth dose vaccine coverage among SIREN participants by age group (panel A) and time since previous infection (panel B).

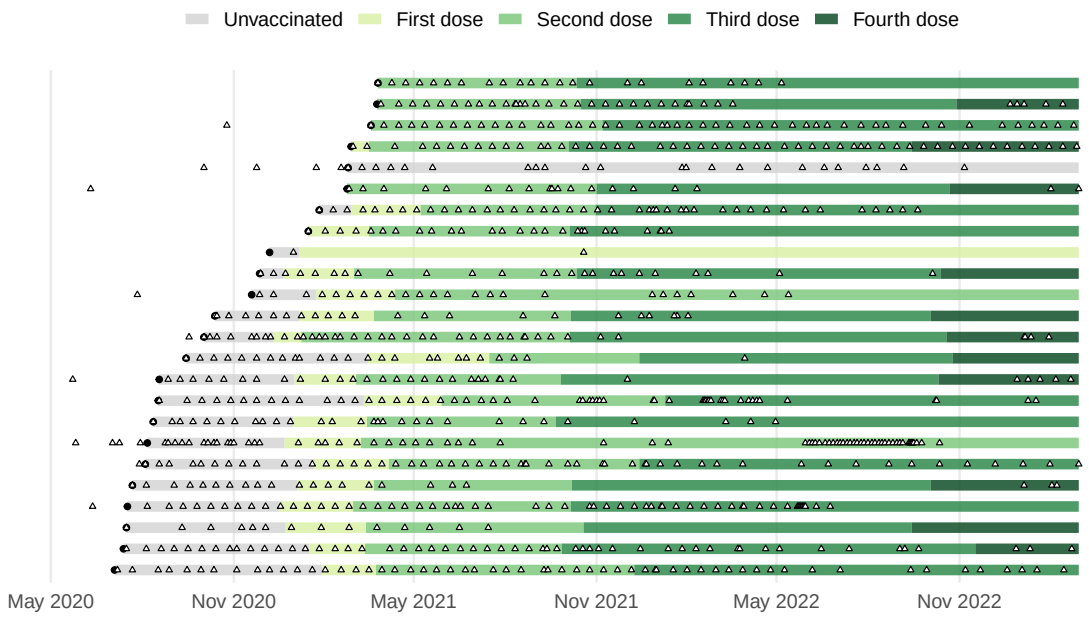


Figure 5.8: Timeline of study participation and vaccination status for randomly selected SIREN participants, June 2020 to March 2023. Key: enrolment (●), PCR testing (△).

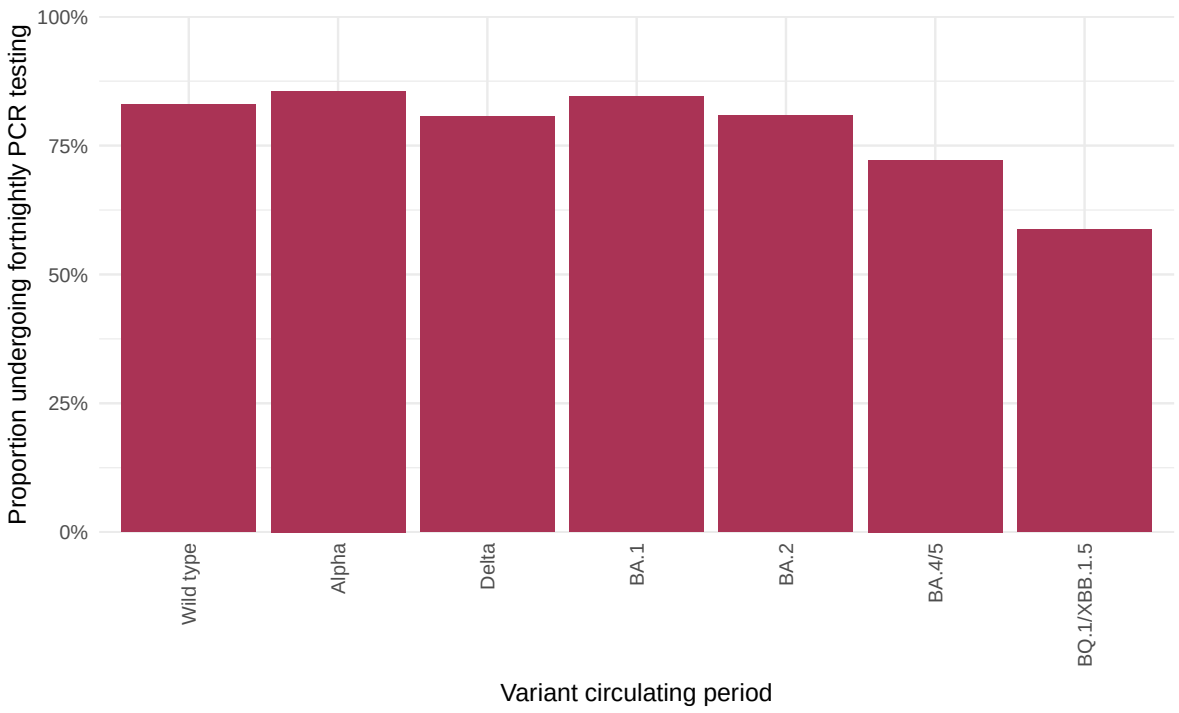


Figure 5.9: Proportion of enrolled SIREN participants undergoing PCR testing on a fortnightly schedule, by variant circulating period.

5.3.7 Cox proportional hazards models

Stratified Cox proportional hazards models, described in Section 2.2.4, were applied to the SIREN study data in two 'interim analyses': firstly to estimate protection from prior infection and vaccination during the Alpha and Delta variant circulating period; and secondly to estimate protection of first booster doses and previous infection against Omicron sub-variants compared to Delta.

The main findings of these analyses are included in Appendix E.1, and further details can be found in the two publications (Hall, Foulkes, Insalata *et al.* 2022; Hall, Insalata *et al.* 2024). These analyses supported the use of booster vaccination to sustain immunity, and highlighted the challenges posed by new variants.

5.4 Cohort study design

5.4.1 Biases in vaccine effectiveness studies

Test-negative case-control study designs are routinely used to assess vaccine effectiveness, both for influenza and COVID-19 (Andrews, Tessier *et al.* 2022; Dean *et al.* 2021; Lopez Bernal *et al.* 2021). Latent healthcare-seeking behaviour may influence vaccine uptake, infection, and testing, as shown in Figure 5.10 panel A. By conditioning on testing, the test-negative case-control study design controls for this potential source of bias, as shown in Figure 5.10 panel B.

Despite controlling for healthcare-seeking behaviour, a test-negative case-control study may be subject to several other uncontrolled biases (Shi *et al.* 2023; Sullivan *et al.* 2016). In particular, 'collider stratification bias' can arise if, say, workplace testing policies require unvaccinated individuals to test on a more regular basis (shown by the blue arrow in Figure 5.10 panel B); or alternatively, if vaccinated individuals are less likely to be symptomatic, and hence less likely to get tested for infection (shown by the red arrow in Figure 5.10 panel B).

Cohort study designs, where people test according to a pre-determined observation scheme, improve on the test-negative case-control study design by addressing the specific 'collider stratification' biases whilst still controlling for health-seeking behaviour. The corresponding causal diagram for a cohort study design is shown in Figure 5.10 panel C. The cohort study design in SIREN also benefits from being able to capture both symptomatic and asymptomatic infection.

Despite accounting for the specific biases described, both study designs remain subject to biases from unmeasured confounding, misclassification, and may suffer from a lack of generalisability.

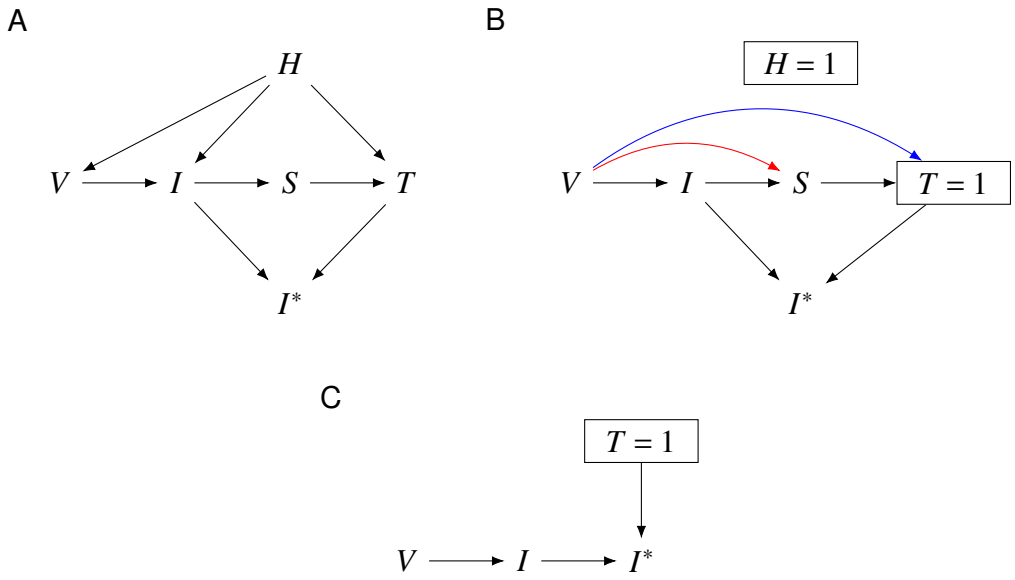


Figure 5.10: Causal DAG describing the relationship between vaccination (V), infection (I), observed infection (I^*), symptoms (S), testing (T), and latent health-seeking behaviour (H) (panel A), the potential biases which may remain after adjusting for test-seeking behaviour in a test-negative case-control study (panel B), and in a cohort study (panel C).

5.4.2 Intermittent observation

An additional challenge, specific to observational cohort data, is *intermittent observation* of the underlying process, also termed interval-censoring.

In the Cox proportional hazards models described in Section E.1 only the positive test information in SIREN was used, with the assumption that participants remained uninfected in the absence of a positive test. However, as described in Section 2.5.3, intermittent (and irregular) observation of participants' PCR status could result in infections being undetected.

Adherence to the fortnightly testing schedule during the Wild type, Alpha, Delta, and Omicron BA.1/2 variant circulating periods was high (Figure 5.9), so the regular PCR and antibody testing in SIREN was judged to be sufficient to detect the majority of COVID-19 infections. During the BA.4/5 and BQ.1/XBB.1.5 variant circulating periods, however, a greater proportion of participants missed PCR tests and there was less frequent monitoring of antibodies, which may limit the validity of Cox proportional hazards model estimates during this period. To address the challenge of intermittent observation multi-state models can be considered. Multi-state models will correctly account for interval-censoring when this censoring is non-informative (van den Hout 2016).

At the start of autumn 2022, in anticipation of a new wave of SARS-CoV-2 infections, second vaccine boosters (i.e. fourth doses) were offered to healthcare workers alongside their seasonal influenza vaccinations. In the next section I develop multi-state models which are

applied to estimate fourth dose vaccine effectiveness against SARS-CoV-2 infection during the autumn/winter 2022/23 period.

5.5 Multi-state models

In Section 2.4 I described how parameters related to infection may be estimated by fitting a multi-state model to the set of observed state transitions for a group of individuals. In this section I describe several 'SIR-type' multi-state models which account for the specific challenges of intermittent observation and misclassification in observational cohort data.

I apply these models to the SIREN data to estimate the protection of second booster vaccination and prior SARS-CoV-2 infection against infection (both symptomatic and asymptomatic), and mean duration of PCR positivity among triple-vaccinated healthcare workers between September 2022 and March 2023 (i.e. the period of Omicron BA.4/5 and BQ.1/XBB.1.5 sub-variant circulation).

The public health research question was whether healthcare workers should be routinely offered autumn vaccine boosters, given the short-term nature of protection and uncertainty about future variant emergence.

5.5.1 Multi-state model frameworks

SIREN participants test according to pre-specified dates, with observations representing testing rather than the occurrence of infection. For an individual with a PCR-negative status at time L and PCR-positive status at time R, infection is assumed to occur at some time $T \in [0, \infty)$ during the time interval $(L, R]$. The infection time is therefore interval-censored, $T \in (L, R]$. If the time intervals are fixed in advance then this interval censoring is independent of T, formally (van den Hout 2016):

$$\Pr(L < T \leq R \mid L = l, R = r) = \Pr(l < T \leq r)$$

As independent interval-censoring implies a non-informative observation scheme, multi-state models are a suitable choice for this analysis (van den Hout 2016). These models implicitly account for incomplete testing (i.e. gaps in an individual's testing history), without the assumption that an individual remains uninfected despite not testing. Since observation in SIREN occurs at fixed 2-weekly time intervals, continuous-time multi-state models were used which allow for changes of state (from susceptible to infected, and infected to recovered) to occur at any point during the time interval.

Two-state model

A two-state 'SI' model, where individuals move between a susceptible (S) and an infected (I) state, is one of the simplest forms of multi-state model. Individuals with a negative test result (indicating lack of infection) are observed to be currently in the susceptible state, and those with a positive test result (indicating infection) are observed to be currently in the infected state. As shown in Figure 5.11, the model is defined by transition intensities q_{12} and q_{21} between these two states, and covariate effects may be included on these transitions.

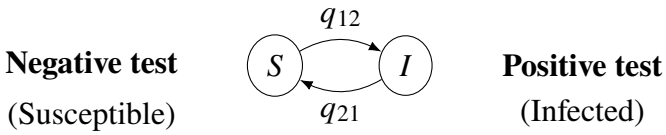


Figure 5.11: Two-state 'SI' model, defined by transition intensities q_{12} and q_{21}.

Three-state model

Immediately following recovery from infection there is typically very little risk of reinfection (Hall, Foulkes, Charlett *et al.* 2021). Whilst this 'convalescence' period can be represented through a covariate in the two-state model, an extension to a three-state 'SIR' model is also possible through the inclusion of a recovery state (R). Within the recovery state participants are not at risk of infection and must first progress to the susceptible state, as shown in Figure 5.12.

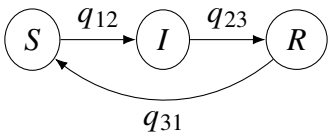


Figure 5.12: Three-state 'SIR' model defined by transition intensities q_{12}, q_{23}, and q_{31}.

Symptom status model

Self-reported symptoms information may be used to classify infections as symptomatic/asymptomatic and these states modelled as separate outcomes, I_s and I_a, as shown in Figure 5.13. Whilst q_{12} and q_{13} represent the incidence of asymptomatic and symptomatic infection, typically of greater interest in these models are the covariate effects on these transition intensities, and the mean sojourn time in each state.

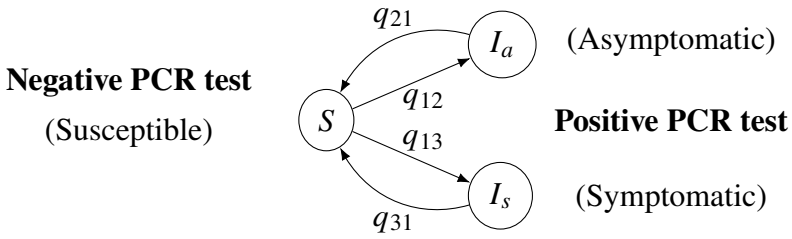


Figure 5.13: Symptom status multi-state model defined by transition intensities q_{12}, q_{13}, q_{21}, and q_{31}.

Semi-Markov model

The models considered so far in this section are Markov, where transition to the next state does not depend on the time spent in a state. Relaxing this assumption may allow the model to better represent recovery from a disease, as transition from the I to the S state typically becomes more likely as time progresses i.e. for the two-state model, allowing q_{21} to depend on the time spent in state I.

In a semi-Markov multi-state model, transition intensities depend on the time spent in the current state, but not on previous states:

$$q_{rs}(t;\mathcal{X}_t) = q_{rs}(t-t_r)$$

where $\mathcal{X}_t$ is the observation history of the process up to time t, and t_r is the time of entry into the current state r (Jackson 2011).

To achieve this in the two-state model the I state is sub-divided into two 'hidden' states, with paths through these states representing short-stayers ($S \to I_1 \to S$) and long-stayers ($S \to I_1 \to I_2 \to S$), as shown in Figure 5.14.

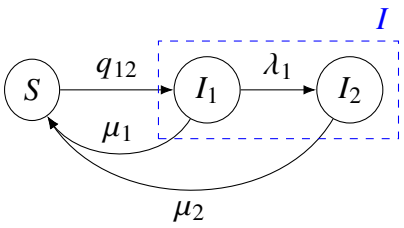


Figure 5.14: Semi-Markov multi-state model with two hidden states I_1 and I_2 and associated transition intensities.

This is equivalent to making the sojourn distribution in the I state more flexible than an exponential distribution, such that the hazard of leaving I depends on how long has been spent there. In this model an individual is a long-stayer with probability $\lambda_1/(\lambda_1+\mu_1)$.

Misclassification model

Multi-state models may also be extended to account for misclassification in the observed states, reflecting the non-zero probability of an infected person being observed as uninfected, and vice-versa. The Coronavirus Infection Survey (CIS) previously estimated the test sensitivity (true-positive rate) of SARS-CoV-2 PCR tests to be 99.9%, and test specificity (true-negative rate) to be between 85–98% (Wei *et al.* 2023).

These probabilities may be included through a hidden-Markov, or misclassification model, with true positive (P), true negative (N), observed PCR-negative (S), and observed PCR-positive (I) states. The corresponding misclassification rates $\Pr(I \mid N)$ and $\Pr(S \mid P)$ may be fixed (as shown in Figure 5.15) or may exist as unknown parameters to be estimated.

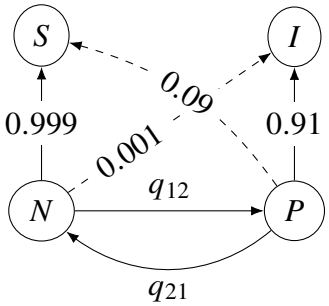


Figure 5.15: Hidden-Markov multi-state model with hidden states, P (true-positive) and N (true-negative), and observed states, S and I. Dashed lines indicate misclassification.

Combining several of these models is possible, for example a hidden semi-Markov model incorporating symptom status. However, in real-world applications there may not be sufficient information to identify each of the parameters in these more complex models.

5.5.2 Application to SIREN study: Methods

The multi-state models described above were used to estimate the hazards associated with infection for selected covariates, and the mean sojourn time in the PCR positive state. The `R` package `msm` provides functions for fitting continuous-time Markov and hidden Markov multi-state models to intermittently observed longitudinal data using maximum-likelihood, with comprehensive online documentation available (Jackson 2011; Jackson 2021). The package has been implemented and validated by numerous disease studies investigating longitudinal cohort data (Steyerberg *et al.* 2019; Tran *et al.* 2022; Vermunt *et al.* 2019; Webb *et al.* 2020).

Participants

All participants enrolled in SIREN between 12 September 2022 and 31 March 2023 were included in the analysis, provided they had at least 24 weeks since receiving a third vaccine

dose (those with fewer than 24 weeks only entered the study once 24 weeks had elapsed). To account for study withdrawal, participants were right-censored at their last recorded PCR test.

Piecewise-constant hazards

The risk of community transmission of SARS-CoV-2, as well as the risk of within-hospital transmission among healthcare workers, varied throughout the study period (UK Government 2021). Piecewise-constant hazards may be used to model changes in hazards over time, with hazards defined to be constant within specified time intervals but allowed to change between the intervals (van den Hout 2016).

To model the changing risk of infection in this study a time-varying covariate $z(t)$ was used for the calendar month, with piecewise-constant hazards within each month. The resulting Markov model is time-inhomogeneous, with the probability of being in state s at time t_n, conditional on being in state r at time t_1, given by the (r, s) entry of the transition probability matrix $\mathbf{P}(t_1, t_n)$:

$$\mathbf{P}(t_1, t_n) = \mathbf{P}(t_1, t_2)\mathbf{P}(t_2, t_3)\ldots\mathbf{P}(t_{n-1}, t_n)$$

where $\mathbf{P}(t_i, t_{i+1}) = \exp((t_{i+1} - t_i)\mathbf{Q})$, and each of the (t_i, t_{i+1}) are intervals over which $\mathbf{Q}$, the transition intensity matrix, is constant (van den Hout 2016).

Covariates

Covariates were selected based on knowledge from previous vaccine effectiveness studies in this cohort (Hall, Foulkes, Insalata *et al.* 2022; Hall, Insalata *et al.* 2024; Hall, Foulkes, Saei *et al.* 2021), data availability, and by comparing Akaike information criterion (AIC) values and likelihood-ratio tests (Tables 5.1 and 5.2 show AIC and log-likelihood for the two-state and three-state models).

The covariates selected were: vaccination status, time since previous infection, age group, gender, region, household status, and occupation/setting. Whist this combination of covariates did not have the smallest AIC, this choice was also based on prior knowledge of which covariates might explain an individual's susceptibility to infection.

The occupation/setting covariate was chosen instead of the binary patient contact covariate as it provided more detailed information. For the time since previous infection covariate '2+ years' was chosen as the baseline category as the 'confirmed naïve' group was much smaller, and as historical testing patterns indicated different behavioural trends e.g. more shielding, making them less suitable as a baseline for the wider cohort.

Hazard ratios (HR) were converted into vaccine effectiveness (VE) and relative protection estimates using the formula: $\text{VE} = 1 - \text{HR}$.

Table 5.1: Covariates, log-likelihood and AIC values used for two-state model covariate selection, sorted by AIC.

Month (piecewise-constant)	Vaccination	Time since infection	Region	Age group	Household status	Gender	Occupation/setting	Clinical risk group	Staff type	Ethnicity	Patient contact	Log-likelihood	AIC
✓	✓	✓	✓	✓								-3882.14	7830.27
✓	✓	✓	✓	✓	✓							-3881.46	7832.91
✓	✓	✓	✓	✓	✓	✓						-3881.17	7834.35
✓	✓	✓	✓	✓	✓						✓	-3881.42	7834.85
✓	✓	✓	✓	✓	✓	✓					✓	-3881.14	7836.28
✓	✓	✓	✓	✓	✓		✓					-3876.82	7837.64
✓	✓	✓	✓	✓	✓			✓				-3881.16	7838.31
✓	✓	✓	✓	✓	✓					✓		-3880.30	7838.60
✓	✓	✓	✓	✓	✓	✓	✓					-3876.56	7839.12
✓	✓	✓	✓	✓	✓				✓			-3873.69	7839.37
✓	✓	✓	✓	✓	✓	✓		✓				-3880.85	7839.71
✓	✓	✓	✓	✓	✓	✓			✓			-3872.87	7839.75
✓	✓	✓	✓	✓	✓	✓				✓		-3879.88	7839.77
✓	✓	✓	✓	✓	✓	✓	✓				✓	-3876.56	7841.12
✓	✓	✓	✓	✓	✓	✓	✓		✓			-3868.21	7844.42
✓	✓	✓	✓	✓	✓	✓	✓	✓				-3876.29	7844.58
✓	✓	✓	✓	✓	✓	✓	✓			✓		-3875.46	7844.91
✓	✓	✓	✓									-3900.46	7856.93
✓	✓	✓										-3919.87	7873.73

Table 5.2: Covariates, log-likelihood and AIC values used for symptom-status model covariate selection, sorted by AIC.

Month (piecewise-constant)	Vaccination	Time since infection	Region	Age group	Household status	Gender	Occupation/setting	Clinical risk group	Staff type	Ethnicity	Patient contact	Log-likelihood	AIC
✓	✓	✓	✓	✓								-3488.13	7064.27
✓	✓	✓	✓	✓	✓							-3487.41	7066.83
✓	✓	✓	✓	✓	✓	✓						-3487.12	7068.24
✓	✓	✓	✓	✓	✓						✓	-3487.39	7068.77
✓	✓	✓	✓	✓	✓		✓					-3482.89	7071.77
✓	✓	✓	✓	✓	✓			✓				-3487.12	7072.24
✓	✓	✓	✓	✓	✓					✓		-3486.15	7072.30
✓	✓	✓	✓	✓	✓				✓			-3479.58	7073.16
✓	✓	✓	✓	✓	✓	✓	✓					-3482.62	7073.24
✓	✓	✓	✓	✓	✓	✓	✓	✓				-3482.35	7078.70
✓	✓	✓	✓									-3506.38	7090.76
✓	✓	✓										-3525.79	7107.58

Likelihood for intermittently observed data

For interval-censored data, the likelihood is calculated from the transition probability matrix, $\mathbf{P}(t)$ (Kalbfleisch and Lawless 1985).

For individual i in the SIREN cohort observed at n_i observation times $(t_{i,1}, \ldots, t_{i,n_i})$, denote the corresponding states as $\left(X(t_{i,1}), \ldots, X(t_{i,n_i})\right)$, which may take the values PCR negative (Susceptible) or PCR positive (Infected)

The likelihood contribution for this individual occupying a given pair of observed states $X(t_j), X(t_{j+1})$ at times t_j, t_{j+1} is given by:

$$L_{i,j} = p_{X(t_j),X(t_{j+1})}(t_{j+1} - t_j)$$

This is the entry for the transition probability matrix $\mathbf{P}(t)$ at the $X(t_j)$th row and $X(t_{j+1})$th column at $t = t_{j+1} - t_j$. The full likelihood, $L(\mathbf{Q})$, for the transition intensity matrix $\mathbf{Q}$ is the

product of the $L_{i,j}$ over all individuals and all transitions. The estimate for the transition intensity matrix is obtained by maximising this likelihood (Jackson 2011).

Model fitting

Several challenges emerged during model fitting. Firstly, since the amount of data available for the study was limited, fitting complex models with many parameters was often infeasible. The hidden Markov model specifications (semi-Markov and misclassification models), in particular, exhibited convergence issues and were sensitive to initial values. With the available (intermittently observed) data and number of observed infections it was unlikely that robust estimates could be obtained from these models (see Appendix E.3).

Secondly, while multi-state models are better-suited to examining intermittently observed data, it was notable that these models took significantly longer to fit than to the Cox proportional hazards models (several hours compared to under a minute), and careful tuning of the optimizer was usually required to converge to the maximum likelihood.

Finally, the inclusion of site-specific random effects (i.e. a shared frailty model) was not possible within the `msm` package (Jackson 2021). These models are able to be specified in a Bayesian framework (van den Hout 2016), although experience has shown that fitting complex Bayesian models within a reasonable time-scale to provide estimates during a pandemic can be challenging (Birrell, Blake, van Leeuwen E, Gent *et al.* 2021).

Goodness of fit

Pearson-type goodness-of-fit hypothesis tests can be used to assess the goodness of fit of multi-state models fitted to intermittently-observed data (Aguirre-Hernández and Farewell 2002). These test the hypothesis that the data were generated by the fitted model, and the test statistic, T, is the sum of the deviances between the observed, O, and expected, E, number of state transitions. For participants i moving between states r and s (Titman 2009):

$$
\begin{aligned}
O_{rs} &= \sum_i \mathbb{1}\{X_i(t_i) = s, X_i(t_{i-1}) = r\} \\
E_{rs} &= \sum_i \Pr\{X_i(t_i) = s \mid X_i(t_{i-1}) = r\}\mathbb{1}\{X_i(t_{i-1}) = r\} \\
T &= \sum_r \sum_s \frac{(O_{rs} - E_{rs})^2}{E_{rs}}
\end{aligned}
$$

The deviance can be stratified by covariate, time from study start, and time interval categories. Here time intervals refer to the times between consecutive observations, i.e. two time interval categories would group frequent and less frequent testers.

Model fit can also be assessed by comparing the number of individuals forecast to be occupying a given state in the fitted model to a more empirical estimate, derived from the observed data and the estimated transition probability matrix. This empirical estimate combines the observed number of individuals under observation at time t with the estimated mean sojourn time from the transition probability matrix to infer the the number of individuals occupying a given state. The number of individuals occupying the infected state as forecast by the fitted model was compared to the more empirical estimate at a series of times t (Jackson 2011).

The expected number of infections can also be simulated from the fitted model and compared to the number that were observed in the data. This summary measure can be simulated under different sampling schemes to explore model fit and the impact of intermittent observation.

Multi-state models

The two-state model (Figure 5.11) and the symptom status model (Figure 5.13) were specified to include proportional hazards for each covariate (with piecewise-constant hazards for the month covariate) and each of the models was run under four scenarios (Table 5.3):

1. Not including time since previous infection as a covariate and using a binary indicator of vaccination, to estimate overall vaccination effectiveness in a cohort with heterogeneous time since previous infection;

2. Not including time since previous infection as a covariate and using a time-varying indicator of vaccination, to estimate vaccine effectiveness by time since vaccination in a cohort with heterogeneous time since previous infection;

3. Including time since previous infection and a binary indicator of vaccination, to estimate protection associated with prior infection;

4. Including time since previous infection and an interaction effect between this variable and vaccination, to estimate the marginal effectiveness of vaccination by time since prior infection.

Participants without indication of prior infection and for whom a recent serological sample (within the last 6 months) was not available for testing ($n = 2100$) were excluded from analyses exploring time since infection as their infection-naïve status could not be confirmed. This is indicated by a smaller number of individuals for these models in Table 5.3.

Estimates from the three-state model (Figure 5.12), semi-Markov model (Figure 5.14), and misclassification model with fixed misclassification probabilities (Figure 5.15) are included in Appendix E.3.

Table 5.3: List of multi-state models used to generate estimates, with number of individuals, inclusion of symptom information, and list of covariates. ✓: covariate included as main effect in regression, p: covariate included with piece-wise constant hazards.

	Multi-state model						
	M1	M2	M3	M4	M5	M6	M7
N (individuals in model)	9560	9560	7549	7549	9560	9560	7549
Symptom data (Y/N)	N	N	N	N	Y	Y	Y
Covariates							
Month	p	p	p	p	p	p	p
Binary vaccination status	✓		✓		✓		✓
Vaccination status		✓				✓	
Time since infection			✓	✓			✓
Vaccination:time since infection (interaction term)			✓	✓			✓
Region	✓	✓	✓	✓	✓	✓	✓
Age group	✓	✓	✓	✓	✓	✓	✓
Household status	✓	✓	✓	✓	✓	✓	✓
Gender	✓	✓	✓	✓	✓	✓	✓
Occupation/setting	✓	✓	✓	✓	✓	✓	✓

5.5.3 Application to SIREN study: Results

Symptom status

Most (87%; 1130/1298) of the SARS-CoV-2 infections detected had symptom information reported. The proportion symptomatic was slightly higher among those with a booster dose (78.1%, 525/672) compared to those with a waned third dose (74.2%, 340/458). Among those with an infection in the past 6 months, 55.9% (38/68) reported symptoms compared to >70% of those confirmed naïve or with an infection 6+ months prior (Figure 5.16).

Vaccine effectiveness

Vaccine effectiveness of the fourth (booster) dose relative to protection at least six-months after a third dose was estimated as 13.1% (95% CI 0.9 to 23.8%) over the entire analysis period (Figure 5.17 panel A). VE was highest in the two months post-vaccination at 24.0% (95% CI 8.5 to 36.8%), reducing to 10.3% (95% CI -11.4 to 27.8%) in the period two-to-four months post-vaccination, and 1.7% (95% CI -17.0 to 17.4%) in the period four-to-six months post-vaccination (Figure 5.17 panel B). Relative to a waned third dose, VE was 8.6% (95% CI -5.8 to 20.9%) against symptomatic infection and 28.0% (95% CI 6.6 to 44.5%) against asymptomatic infection (Figure 5.17 panel D and E).

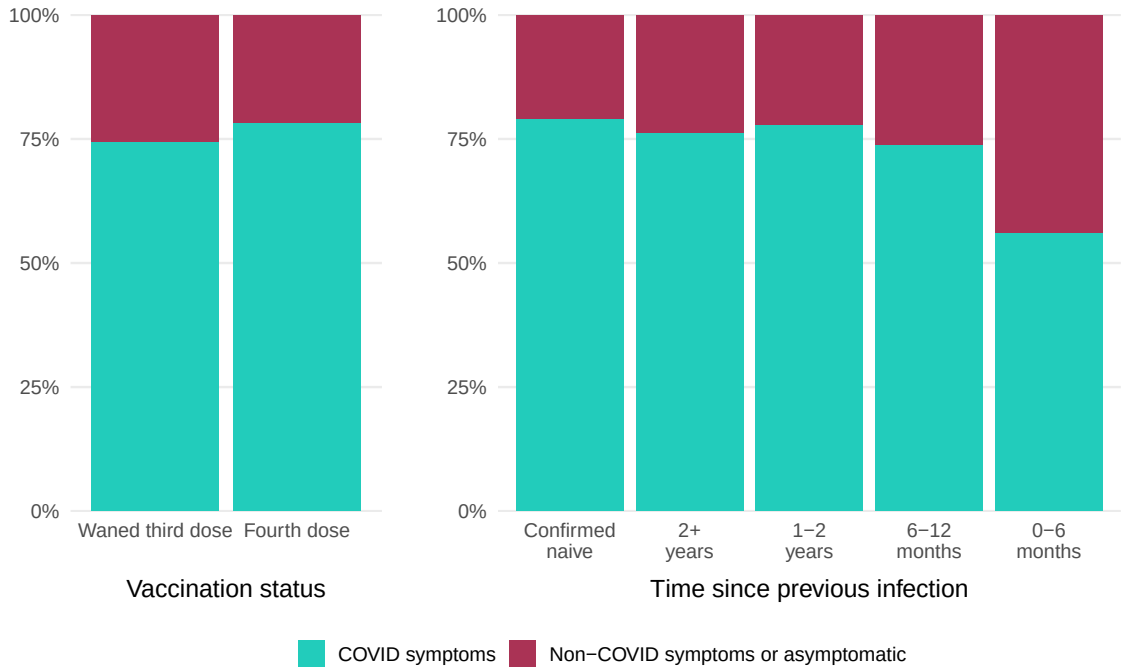


Figure 5.16: Self-reported symptom status among SIREN participants testing positive for COVID-19, by vaccination status (panel A), and time since previous infection (panel B).

SARS-CoV-2 infection within the last 6 months was associated with a 63.6% (95% CI 46.8 to 75.0%) increase in protection compared to individuals who had an infection more than 2 years ago (Figure 5.17 panel C). This protection was 72.0% (95% CI 56.1 to 82.1) against symptomatic infection and 42.2% (95% CI -17.8 to 72.7) against asymptomatic infection (Figure 5.17 panel F). Corresponding estimates with the confirmed naïve group used as the baseline are shown in Figure E.6. Table E.3 includes these estimates alongside crude PCR positivity rates (number of detected PCR positive results per 10,000 person-days of follow-up).

Mean sojourn time in PCR positive state

The mean sojourn time in the PCR positive state was estimated as 7.5 days (95% CI 6.9 to 8.1) overall (Figure 5.18 panel A). When averaged over the study population, the mean sojourn time was 6.9 days (95% CI 5.9 to 8.1) among those with a booster vaccination, compared to 8.5 days (95% CI 6.8 to 10.6) for those with a waned third dose (Figure 5.18 panel B).

The mean sojourn time was 9.5 days (95% CI 7.1 to 12.8) for confirmed naïve participants and 7.3 days (95% CI 6.3 to 8.3) for those with an infection within the past 0–6 months (Figure 5.18 panel C). Infections reported as symptomatic had a mean sojourn time of 8.1 days (95% CI 7.4 to 8.9), compared to 4.7 days (95% CI 4.0 to 5.5) for asymptomatic (Figure 5.18 panel D).

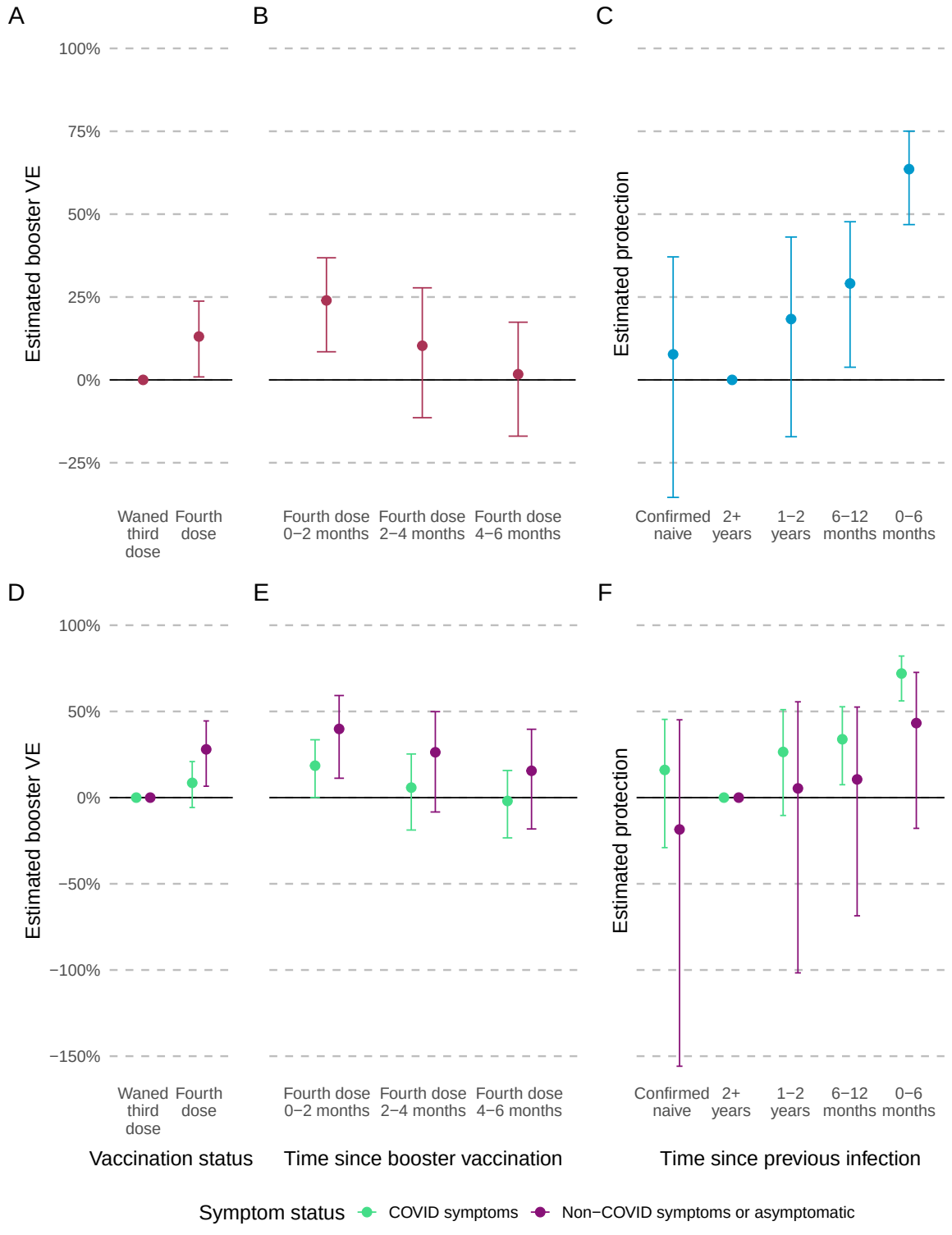


Figure 5.17: Estimated booster VE by booster vaccination status (panel A, model M1), time since booster vaccination (panel B, model M2), and estimated protection from previous infection (panel C, model M3), and symptom status (panels D–F, models M5–M7). Error bars show the 95% CI around the estimates.

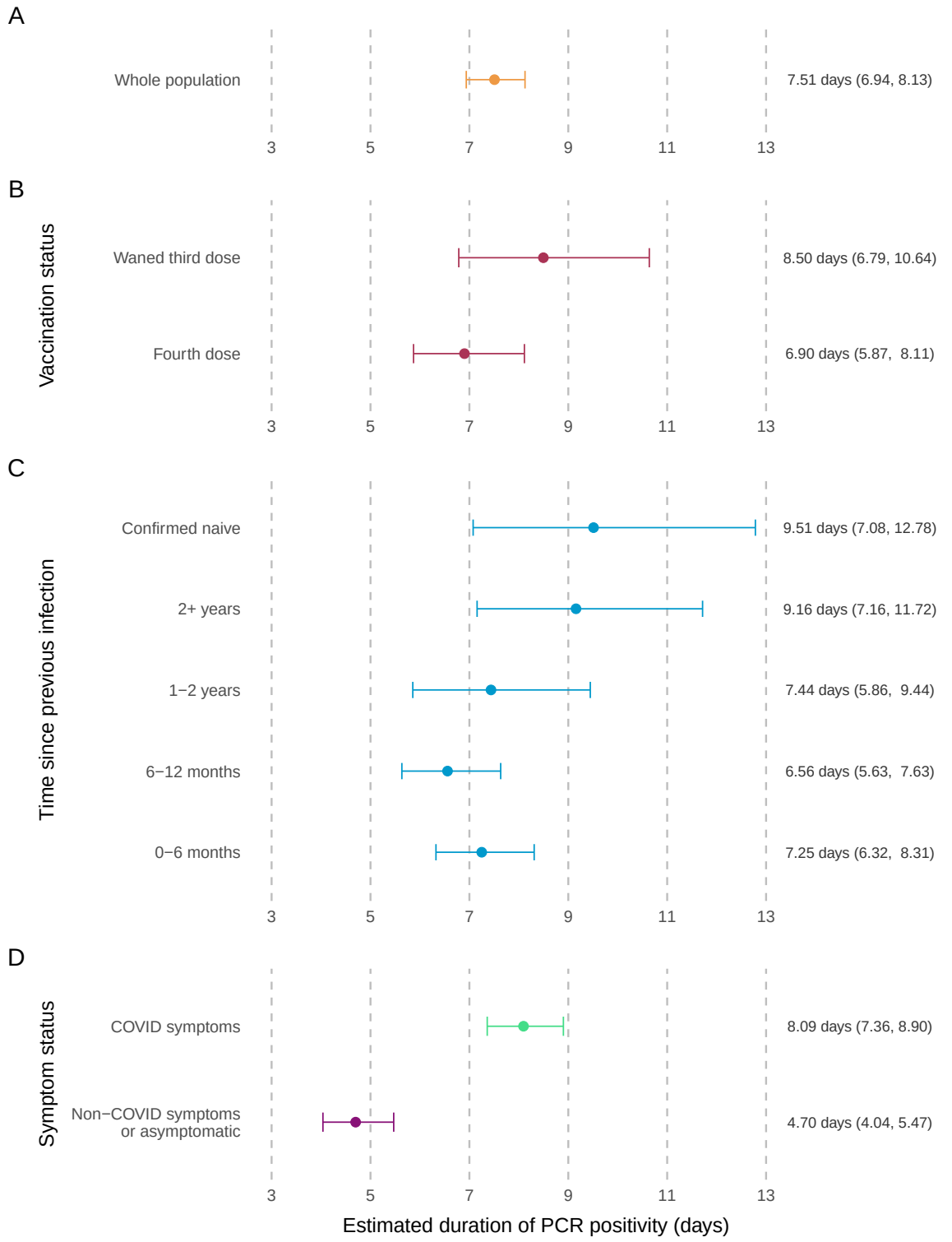


Figure 5.18: Estimated mean sojourn time in PCR positive state, averaged across the study population, overall (panel A, model M1), by vaccination status (panel B, model M1), time since previous infection (panel C, model M2), and symptom status (panel D, model M5). Error bars show the 95% CI around the estimates.

Table E.4 includes the mean sojourn time estimates alongside empirical estimates of the duration of PCR positivity, calculated as the median time from initial PCR positive to subsequent PCR negative.

Goodness of fit

A comparison between the forecast number of individuals occupying the infected state and the more empirical estimate derived from the observed data was used to assess model fit. For the two-state model, the empirical estimate was within the confidence intervals of the fitted model throughout most of the study period, but fell below the lower 95% CI of forecast prevalence in weeks 12–16 (Figure 5.19).

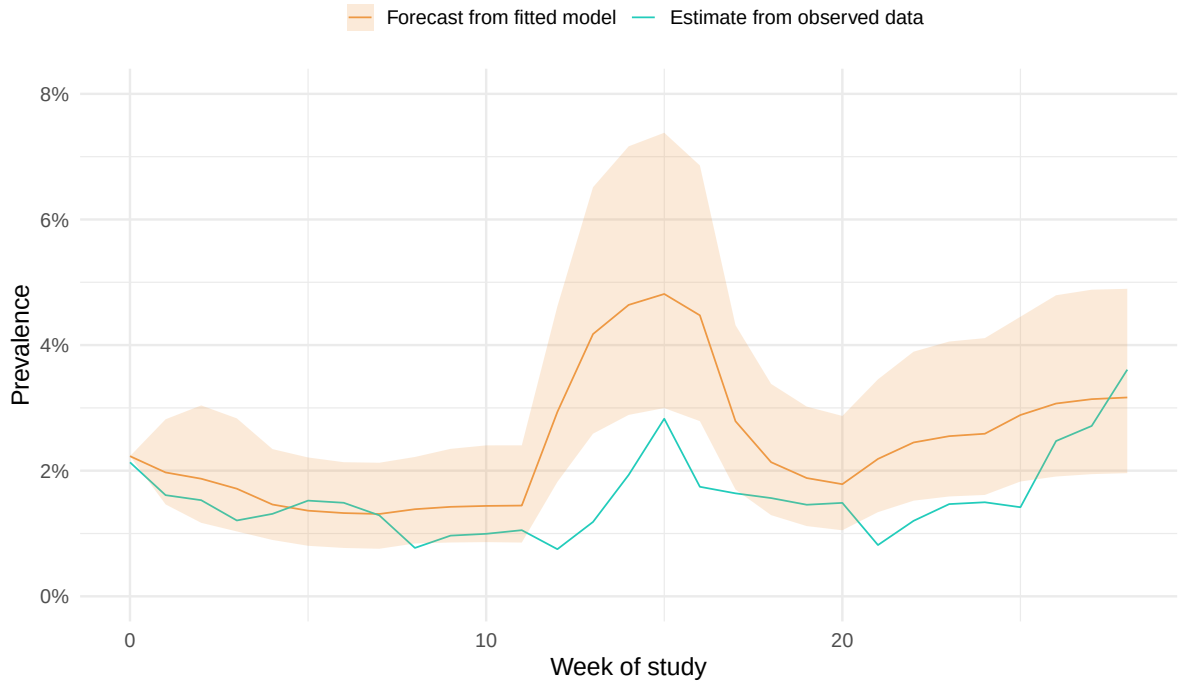


Figure 5.19: Comparison of prevalence in the infected state forecast from fitted model M1 and more empirical estimate over time. Shaded area shows the 95% CI around the expected prevalence.

Table 5.4 shows the deviance test statistic for the two-state model fitted to the SIREN data. Time intervals 1 and 2 represent more and less frequent testers, respectively. Covariate categories A and B represent are partitions of the data based on quantiles of the estimated transition intensities $q_{r,s}$ (Titman 2009). The test statistic varied most according to time interval, with the greatest deviance among less frequent testers (time interval 2) and for the susceptible to infection transition.

Table 5.4: Deviance statistics for fitted multi-state model M1, by time from study start, time interval, and covariate strata.

Time from study start (weeks)	Time interval	Covariate category	Susceptible to infected transition	Infected to susceptible transition
[0.14,16.29)	1	A	5.71	-5.46
[16.29,28.57]	1	A	7.11	-0.37
[0.14,16.29)	2	A	-18.39	-0.01
[16.29,28.57]	2	A	-25.59	0.02
[0.14,16.29)	1	B	0.04	-6.53
[16.29,28.57]	1	B	0.04	-2.30
[0.14,16.29)	2	B	0.49	-0.64
[16.29,28.57]	2	B	0.00	-0.28

Lastly, Appendix E.5 shows the expected number of infections simulated under different sampling schemes, for instance all participants testing on a bi-weekly basis.

Comparison of model estimates

The Cox proportional hazards model is equivalent to a two-state model without recovery or interval censoring. Figure 5.20 compares the hazards estimated in multi-state model M1 to those from the corresponding stratified Cox proportional hazards model, where the same covariates are included in both models. Point estimates were very similar between the two models, with estimates from the multi-state model having slightly wider confidence intervals. Additional comparisons are included in Figures E.8 and E.7.

5.5.4 Discussion

Multi-state models were successfully applied to the SIREN data to jointly estimate the effectiveness of fourth dose mRNA vaccines against infection and the mean sojourn time in the PCR positive state over the period September 2022 to March 2023. Models were applied in real-time throughout the study period, with interim estimates presented to UKHSA and the JCVI to inform booster vaccination policy (Department of Health and Social Care 2023).

Controlling for prior infection, fourth dose mRNA vaccine effectiveness against infection was estimated around 24% in the first 2 months post-vaccination, with significant waning of protection thereafter. These estimates are in-line with bivalent booster vaccine effectiveness estimates among healthcare workers in the US (Shrestha *et al.* 2023), and corroborate the results of other cohort studies (Auvigne *et al.* 2023; Huiberts *et al.* 2023).

Mean duration of PCR positivity estimates were also consistent with previous studies, most of which have estimated a mean duration of positivity for Omicron-era infections of

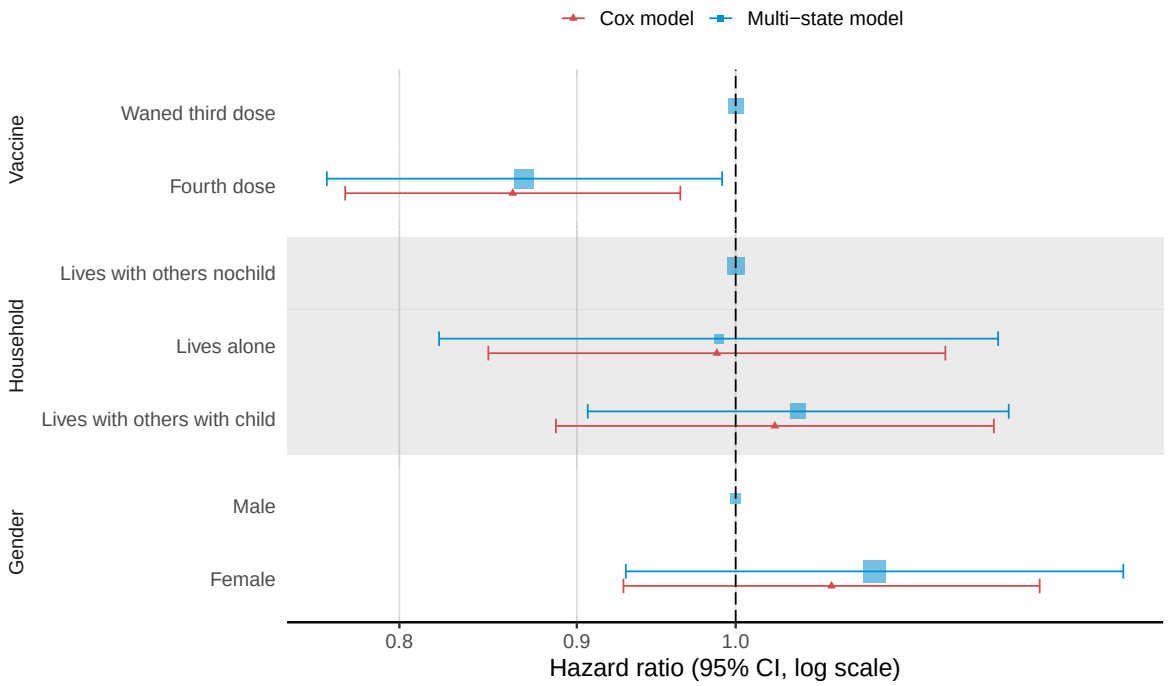


Figure 5.20: Comparison of hazards estimated by multi-state model M1 and the corresponding Cox proportional hazards model, for covariates common between the two models. Error bars show the 95% CI around the estimated hazards.

around 7 days (Boucau *et al.* 2022; Hay *et al.* 2022; Kojima *et al.* 2022; Pei *et al.* 2023; UK Health Security Agency 2023a). The estimates from the multi-state model uncover additional distinctions between sub-groups which are missed by an empirical approach and have not been previously reported, such as mean duration of PCR positivity by symptom status.

The divergence in infection prevalence estimated by the observed data and forecast from the fitted multi-state model coincided with the December and new year period, when lower adherence to the fortnightly testing schedule was seen in the data. This highlights the utility of the multi-state model, which is able to account for potentially unobserved infections in intermittently observed data. With optimal fortnightly testing it is likely that a greater number of infections would have been detected. Future analyses of the this cohort will use multi-state models to account for intermittent observation, thereby reducing biases from sub-optimal sampling.

Similar hazard ratio estimates were obtained from the multi-state and Cox proportional hazards models. The slightly wider confidence intervals in the multi-state model estimates may be a more accurate reflection of their uncertainty, given that they account for intermittent observation and missed testing.

Additional results from this analysis are available in the published manuscript and supplementary materials (Kirwan, Hall *et al.* 2024).

5.5.5 Limitations

Whilst this cohort analysis addressed specific biases of test-negative case-control studies, as described in Section 5.4, certain limitations remain. Firstly, while piecewise-constant hazards were used to account for non-proportionality, the proportional hazards assumption may still be unrealistic for the covariates we have studied. Non-proportional hazards are common in medical studies, and Stensrud and Hernán 2020 have recommended interpreting modelled hazard ratios as a weighted average of the true hazard ratios over the entire follow-up period.

Secondly, the Markov assumption was unlikely to hold true immediately post-infection or immediately post-recovery, although the extent of this issue is likely to be small in comparison to the long follow-up time. As shown in Appendix E.3, the convalescent model did not provide significantly different hazards estimates, perhaps as the time since previous infection covariate already accounts for some of this variability. The semi-Markov model relaxes the Markov assumption but this model exhibited identifiability issues and was sensitive to initial values, likely a result of insufficient data on infections.

Another potential limitation of the chosen models was the assumption of no state misclassification. The Coronavirus Infection Survey (CIS) has previously estimated PCR test sensitivity of between 85–98% (Wei *et al.* 2023). Results from the misclassification model with fixed misclassification probabilities are included in Appendix E.3; estimated booster VE and protection from prior infection were similar to estimates from the two-state model without misclassification. As the SIREN study includes both serial PCR and linked antibody testing data, and allows for manual review in case of inconsistent results, the extent of misclassification in this cohort was expected to be small.

Lastly, non-informative sampling is an assumption of the model. The presence of informative sampling would lead to an overestimate of the probability of detecting infection, e.g. if participants choose to test because of symptoms. For this study the majority of SIREN study participants did test according to the fortnightly schedule (albeit with reduced frequency during December), and only a small degree of out-of-schedule PCR testing was observed.

5.6 Causal effect of vaccination

So far this chapter has investigated the effects associated with an intervention, specifically vaccination, using observational data. Commonly of greater interest, however, are the causal effects of such an intervention. By explicitly stating the causal assumptions and controlling for confounding and selection bias, causal inference methods allow findings from observational studies to be interpreted in a causal way (Hernán and Robins 2006; Ryalen, Stensrud, Fosså *et al.* 2020).

Causal inference methods have been applied in several COVID-19 vaccine effectiveness studies to estimate the causal effect of vaccination at different stages in the pandemic (Dagan *et al.* 2021; Hulme *et al.* 2023; Xie *et al.* 2022). In this section I specify a causal inference method and the necessary assumptions for making causal statements using continuous-time multi-state models. This method is applied to the SIREN data to estimate the causal effect of vaccination in this cohort.

5.6.1 Artificial manipulation of transition intensities

Artificial manipulation of transition intensities (AMTI) is a method of causal inference for continuous-time multi-state models. AMTI has been used to investigate counterfactual scenarios in multi-state models, for example the effect of a hypothetical change in sickness benefits on absence, and the effect of hypothetical treatment regimes for transplant patients (Gran *et al.* 2015; Keiding *et al.* 2001).

AMTI involves intervening on specific entries of an estimated transition intensity matrix, $\mathbf{Q}$. These entries may be fixed to particular values which represent a hypothetical intervention, for instance complete vaccination of a cohort. The intervened-upon transition intensity matrix is $\tilde{\mathbf{Q}}$ and quantities of interest can be estimated from the corresponding transition probability matrix, $\tilde{\mathbf{P}}$.

Before deriving causal estimates using AMTI, I will firstly explicitly state the required causal assumptions. The causal assumptions for a continuous-time parametric proportional hazards model with piecewise-constant baseline hazards are a special case of those formalised by Ryalen, Stensrud, Fosså *et al.* 2020 to infer the causal effect of treatment on survival for cancer patients.

Local independence graphs

A local independence graph which represents short-term causal relationships between continuous processes in the SIREN study is shown in Figure 5.21. Here A represents vaccination, Z are confounders (e.g. age, occupation, prior infection), and Y represents infection. The local independence graph represents the assumptions that vaccination has a causal relationship with infection, and that the confounders (baseline covariates) have a causal relationship with both vaccination and infection.

Participant information

Participants in the study are assumed to be i.i.d. over the study period $[0, \tau]$. For each participant, i, with baseline covariates Z_i, define a set of counting processes (Aalen, Borgan *et al.* 2008):

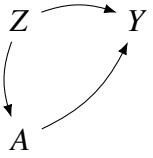


Figure 5.21: Local independence graph with A representing vaccination, Z representing confounders, and Y representing infection.

- $\left(N_i^I(t) : t \geq 0\right)$ the counting process for entries into the infected state by time $t \geq 0$, i.e. $N_i^I(t)$ is the number of times the participant has entered the infection state by time t;
- $\left(N_i^S(t) : t \geq 0\right)$ the counting process for returning to the susceptible state by time $t \geq 0$;
- $(A_i(t) : t \geq 0)$ the counting process for vaccination at time $t \geq 0$, i.e. $A_i(t) = 1$ if the participant has been vaccinated by time t, and $A_i(t) = 0$ otherwise.

Henceforth the i notation will be omitted for brevity, i.e. $N^I(t) \equiv N_i^I(t)$, $Z \equiv Z_i$, and $A(t) \equiv A_i(t)$.

Let

$$\bar{N}^I(t) = \{N^I(s) : s \leq t\}$$
$$\bar{N}^S(t) = \{N^S(s) : s \leq t\}$$
$$\bar{A}(t) = \{A(s) : s \leq t\}$$

represent the history of each counting process up to time t. Define the vaccination status for an individual at the end of the study period as $\bar{A} = \bar{A}(\tau)$, and let $N^{I,\bar{a}}(t)$ and $N^{S,\bar{a}}(t)$ be the counting processes at time t when we intervene to set $\bar{A} = \bar{a}$, i.e. the number of times the participant would have entered the infected state/returned to the susceptible state by time t when we intervene to set $\bar{A} = \bar{a}$.

Let $q^I(t)$ be the transition intensity at time t for counting process $N^I(t)$:

$$\begin{aligned} &q^I(t \mid Z, \bar{A}(t\text{-}), \bar{N}^I(t\text{-}), \bar{N}^S(t\text{-})) \\ &= \lim_{\delta t \downarrow 0} \left\{ \frac{\Pr\left(N^I(t) - N^I(t-\delta t) > 0 \mid Z, \bar{A}(t-\delta t), \bar{N}^I(t-\delta t), \bar{N}^S(t-\delta t)\right)}{\delta t} \right\} \end{aligned}$$

where t- represents the instant before time t. The transition intensities for counting processes $N^S(t)$, $N^{I,\bar{a}}(t)$ and $N^{S,\bar{a}}(t)$ are, respectively:

$$q^S(t \mid Z, \bar{A}(t\text{-}), \bar{N}^I(t\text{-}), \bar{N}^S(t\text{-}))$$
$$q^{I,\bar{a}}(t \mid Z, \bar{N}^{I,\bar{a}}(t\text{-}), \bar{N}^{S,\bar{a}}(t\text{-}))$$
$$q^{S,\bar{a}}(t \mid Z, \bar{N}^{I,\bar{a}}(t\text{-}), \bar{N}^{S,\bar{a}}(t\text{-}))$$

Causal assumptions

Suppose that we make a hypothetical, but meaningful, intervention on a transition intensity q^I to set $\bar{A} = \bar{a}$, then under the assumption of 'local independence' this intervention will not change the 'local' characteristics for other transition intensities and covariate effects (Ryalen, Stensrud and Røysland 2019; Seaman and Keogh 2023):

$$\begin{aligned} &q^I(t \mid Z = z, \bar{A}(t\text{-}) = \bar{a}(t\text{-}), \bar{N}^I(t\text{-}) = \bar{n}^I(t\text{-}), \bar{N}^S(t\text{-}) = \bar{n}^S(t\text{-})) \\ &= q^{I,\bar{a}}(t \mid Z = z, \bar{N}^{I,\bar{a}}(t\text{-}) = \bar{n}^I(t\text{-}), \bar{N}^{S,\bar{a}}(t\text{-}) = \bar{n}^S(t\text{-})) \end{aligned}$$

and, analogously, for q^S:

$$\begin{aligned} &q^S(t \mid Z = z, \bar{A}(t\text{-}) = \bar{a}(t\text{-}), \bar{N}^I(t\text{-}) = \bar{n}^I(t\text{-}), \bar{N}^S(t\text{-}) = \bar{n}^S(t\text{-})) \\ &= q^{S,\bar{a}}(t \mid Z = z, \bar{N}^{I,\bar{a}}(t\text{-}) = \bar{n}^I(t\text{-}), \bar{N}^{S,\bar{a}}(t\text{-}) = \bar{n}^S(t\text{-})) \end{aligned}$$

for all possible values $z \in Z$, $\bar{a}(t\text{-}) \in \bar{A}(t\text{-})$, $\bar{n}^I(t\text{-}) \in \bar{N}^I(t\text{-})$, and $\bar{n}^S(t\text{-}) \in \bar{N}^S(t\text{-})$.

Now make the simplifying assumption that the transition intensities only depend on the binary indicator of vaccination, and not on the history of the vaccination process:

$$q^I(t \mid Z, \bar{A}(t\text{-}), \bar{N}^I(t\text{-}), \bar{N}^S(t\text{-})) = q^I(t \mid Z, A(t\text{-}), \bar{N}^I(t\text{-}), \bar{N}^S(t\text{-}))$$
$$q^S(t \mid Z, \bar{A}(t\text{-}), \bar{N}^I(t\text{-}), \bar{N}^S(t\text{-})) = q^S(t \mid Z, A(t\text{-}), \bar{N}^I(t\text{-}), \bar{N}^S(t\text{-}))$$

'Consistency' and 'no interference' is also assumed (known as the Stable Unit Treatment Value Assumption (SUTVA)). By SUTVA, there are no other forms of treatment, i.e. $N^I(t) = N^{I,\bar{a}}(t)$ for individuals with $\bar{A}(t\text{-}) = \bar{a}(t\text{-})$, and a participant's outcomes do not depend on whether or not other participants were vaccinated (VanderWeele and Hernán 2013).

Lastly, parametric models are assumed for q^I and q^S, with proportional hazards and a piecewise-constant baseline hazard.

5.6.2 Application to SIREN

An intervention to set $(A(t) = a(t) : t \geq 0)$ corresponds to either all participants receiving a fourth vaccine booster dose at time t or none receiving a booster dose. To investigate the causal effect of vaccination in this cohort, two specific interventions were considered:

A. **No vaccination**: all participants begin in the susceptible state and no vaccination is received, corresponding to $a(t) = 0$ for all $t \geq 0$;

B. **Complete vaccination at time zero**: all participants begin in the susceptible state and receive a vaccine booster at the start of the study, corresponding to $a(t) = 1$ for all $t \geq 0$.

Estimation procedure

Let $N_{\text{TOT}}^{I,a}$ be the total number of infections across all individuals and the whole study period, $[0, \tau]$, when we set $A(t) = a(t)$. The expected number of infections under each hypothetical scenario is then defined as:

$$\mathbb{E}\left(N_{\text{TOT}}^{I,a}; q^{I,a}\right) = \sum_i \left\{ \int_{t=0}^{\tau} q^{I,a}(t \mid Z_i, \bar{N}_i^{I,a}(t\text{-}), \bar{N}_i^{S,a}(t\text{-}))\, \mathrm{d}t \right\}$$

Since, in practice, $q^{I,a}$ is unknown, artificial manipulation of transition intensities (AMTI) is applied to the maximum likelihood estimate $\hat{q}^I$ that was obtained from the fitted model to estimate $\mathbb{E}(N_{\text{TOT}}^{I,a}; \hat{q}^{I,a})$, which corresponds to the intervention $A(t) = a(t)$.

Confidence intervals may be calculated by drawing a a random sample from the multivariate normal distribution with mean $\hat{q}$ and covariance matrix as estimated using the maximum likelihood estimator (by the asymptotic normality of this estimator (Young and Smith 2005)) and taking the 2.5 and 97.5 percentiles.

The causal effect of vaccination was calculated by comparing the total number of expected infections under the 'no vaccination' and 'complete vaccination' counterfactual scenarios. For this study, the cohort was assumed to remain under follow-up for the whole study period, so that all infections (regardless of whether they would have been detected) were counted.

Covariates

For a causal analysis, all relevant confounders which could be causal to both an individual in the study choosing to be vaccinated and becoming infected should be included as covariates, whilst not conditioning on common effects of vaccination and infection. Time-varying covariates (i.e. $Z(t)$) may be included provided there is no time-varying confounding (Ryalen, Stensrud, Fosså *et al.* 2020).

Comparable COVID-19 vaccine trial emulation studies have included similar covariates to the association study described in Section 5.5 (calendar month, vaccination history, time since previous infection, age group, gender, and region) (Barda *et al.* 2021; Hulme *et al.* 2023; Xie *et al.* 2022). Two studies set in the UK additionally included clinical risk group and index of multiple deprivation (IMD) in their analyses (Hulme *et al.* 2023; Xie *et al.* 2022). Whilst these covariates were available they were not chosen for the association study as the IMD measure was unavailable for certain regions and as conditioning on clinical risk group did not improve model fit.

To capture potential infection risk among this cohort, both within and outside of the clinical workplace setting, the covariates considered in the causal model were: calendar month, vaccination history, time since infection, age group, gender, region, clinical risk group, IMD, occupation/setting, household status, and patient contact. Calendar month and time since infection were allowed to be time-varying with the time since infection covariate only relating to infections acquired prior to study start.

For the hypothetical scenarios described: the vaccination covariate was fixed to the specified intervention value; the calendar month covariate was included and changed linearly with time; the time since prior infection covariate was specified to relate only to infections prior to the study and increased linearly with time; all other covariates were fixed to the values they were observed as in the data.

5.6.3 Results

The expanded model, including clinical risk group and index of multiple deprivation, was fitted to a subset of the data with complete information for all covariates. This subset included 6553 participants, among whom 1023 infections were detected. A comparison between the forecast number of individuals occupying the infected state and an estimate derived from the observed data was used to assess model fit for the expanded model, included in Appendix E.5.

Figure 5.22 shows the estimated number of infections under each counterfactual scenario. The expected number of infections when no participants received vaccination was 3101 (95% CI 2326 to 3981), and the expected number of infections when all participants received vaccination at time zero was 2660 (95% CI 1992 to 3477). The causal effect of fourth dose vaccination for this cohort was therefore 441 averted infections, albeit with substantial overlap of confidence intervals.

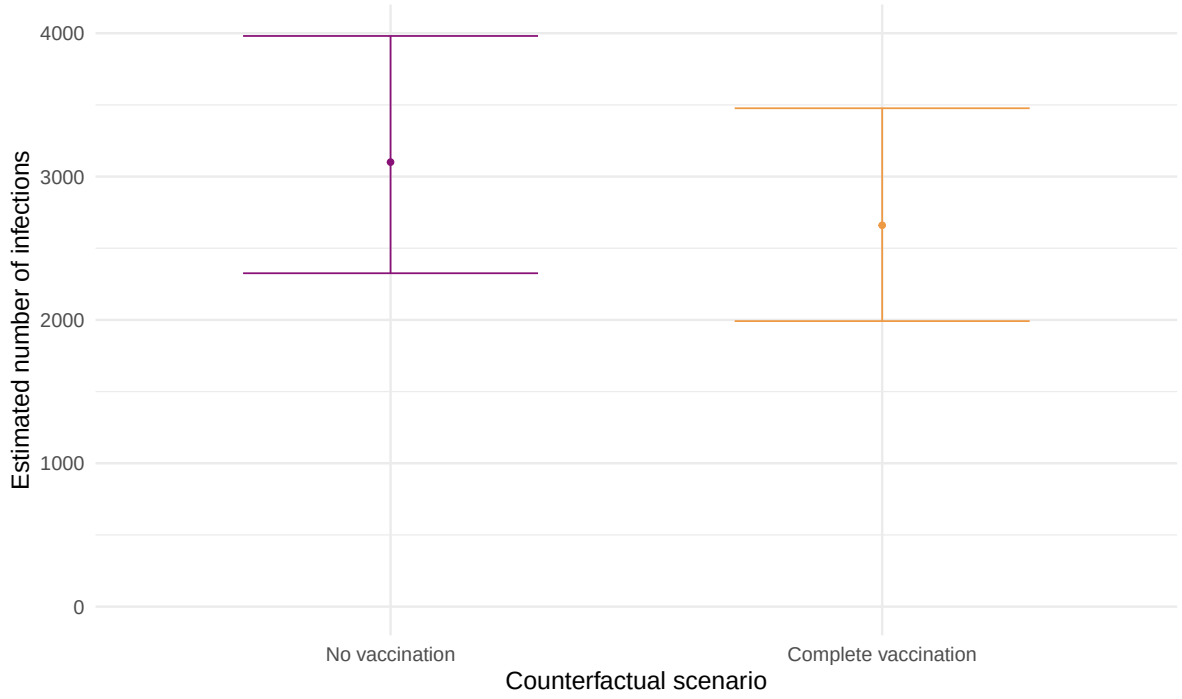


Figure 5.22: Estimated total number of infections under the 'No vaccination' and 'Complete vaccination' counterfactual scenarios.

5.6.4 Discussion

In this analysis I have estimated the causal effect of vaccination upon SARS-CoV-2 infection among SIREN study participants by comparing the total number of expected infections under two counterfactual scenarios.

The proportion of infections estimated to have been averted through vaccination was relatively small (around 14%), with overlapping confidence intervals due to the uncertainty in the parameter estimates. The counterfactual estimates reflect the total number of infections in each hypothetical cohort over the entire time period, regardless of whether or not the infection would be detected. The detected number of infections in the association study was significantly lower than in this causal analysis (n = 1023), both because the data in the association study were intermittently observed, and as participants may have dropped-out of follow-up.

Despite a modest number of averted infections, there are accompanying causal effects of vaccination which may also reduce the burden of infection in this cohort. The mean duration of PCR positivity was estimated to be reduced among those who were vaccinated, which may result in a reduction in sick leave. Other metrics, including cost and impact on patient infections, have previously been used to investigate the utility of SARS-CoV-2 interventions (Heath *et al.* 2023) and vaccination may mitigate adverse events such as hospital admission, as discussed

in Chapter 4. These metrics could provide a wider understanding of the causal effects of vaccination, if sufficient data were available.

Limitations

The same potential biases of informative sampling and state misclassification discussed in Section 5.5.5 are present in this analysis but, as before, the effects from these biases are likely to be small.

Neither the i.i.d. assumption, nor the assumption of no interference are credible for this study (or infectious disease studies in general): healthcare workers may well transmit infection between themselves, and there may well be 'spillover' effects of one individual's vaccination on another individual's infection risk. Clustering data at hospital-level, rather than individual-level, could account for this interference, but this would reduce the sample size considerably and means individual-level confounders cannot be accounted for (Tchetgen Tchetgen and VanderWeele 2012).

For this causal analysis it was assumed that transition intensities did not depend on time since vaccination. A more realistic assumption would be to allow for the waning (i.e. time-dependent) effect of vaccination. However, this would entail significantly greater complexity and is unlikely to increase the precision of the estimates.

Despite controlling for a broad range of covariates, it was not possible to fully account for a number of unobserved confounders. For instance, it was unknown whether participants were 'shielding' (not mixing with others outside of their household) which would likely affect both their decision to be vaccinated and likelihood of being infected. This may be partially controlled by the clinical risk group covariate, and given participants typically needed to attend workplace clinics for PCR testing it is unlikely that many (if any) would have been shielding at the time of this study. Secondly, participants' mixing behaviour and risk tolerance levels could not be measured or controlled for. Participants with a greater risk tolerance may be less likely to receive a fourth booster and/or more likely to mix with others, exposing them to greater risk of infection. It is possible that age and household status may partially account for these factors.

5.7 Chapter summary

In this chapter I have applied Cox proportional hazards and multi-state models to the SIREN study data to estimate the effectiveness of booster vaccination against COVID-19 infection and the mean duration of PCR positivity. The multi-state model was specified to account for intermittent observation, with mean sojourn times jointly estimated. I have also implemented a causal inference method (AMTI) to assess the causal effect of vaccination in this cohort.

In both the Cox proportional hazards and multi-state models, booster vaccination was shown to offer modest but tangible increased protection against SARS-CoV-2 infection in the short-term, with prior infection conferring more robust and sustained protection. This finding was consistent with other studies of fourth dose vaccine effectiveness and highlights the potential role of vaccination in boosting population immunity in advance of expected periods of high prevalence.

To provide causal inference of the effectiveness of vaccination against infection, two counterfactual scenarios were simulated. Vaccination was estimated to provide around a 14% reduction in the overall number of infections, as compared to no vaccination. This average causal effect could be considered alongside other metrics when assessing vaccine effectiveness, including cost, time off work, and prevalence of severe outcomes.

6. Conclusions

The urgent public health response to the COVID-19 pandemic has involved significant biostatistical research, and led to numerous methodological advancements. Major inroads into our understanding of current and future infectious disease epidemics continue to be made through careful selection and application of these statistical tools.

In this final chapter I summarise the main findings of the thesis and suggest potential developments and applications following on from my PhD research.

6.1 Overall summary

In this thesis I have explored the development of survival models and their application to increase knowledge for public health action. The specific focus of my research has been multi-state models, which I have applied in the context of associative and causal inference studies using Bayesian and frequentist methods, with this research adding to the published literature on the burden of infectious diseases.

In each chapter of this thesis I have shown that detailed understanding of the data-generating processes and careful specification of multi-state models are required to address limitations in observational data sources. In each case the validity of model estimates depend on the ability to control for certain biases, such as misclassification, competing risks, and censoring.

Causal inference provides a rigorous framework in which to identify and control for these biases, with the aim of making meaningful causal statements from observational data. Causal questions were considered throughout this thesis: the effect of lockdown, the impact of hospital pressure, and the effectiveness of vaccination. A key take-away from this work is that defining causal questions as clearly as possible in advance can help to inform both the study design and the choice of statistical method.

The next section discusses the specific findings of each chapter, highlighting what was already known and what this thesis adds.

6.2 Main findings

6.2.1 HIV incidence estimation

What was already known

CD4 counts were known to dip during the acute (or seroconversion) phase of HIV infection. Whilst not an issue for earlier generation tests, the introduction of fourth-generation HIV tests and a rapid increase in regular HIV testing among GBM has implications for use of the CD4 marker as a measure of HIV progression. Sasse *et al.* 2016 estimated that misclassification of late HIV diagnosis among gay and bisexual men in Belgium had increased from 5% in 1998 to 41% in 2012. This assessment was based on HIV testing history and clinical presentation among men diagnosed with a CD4 count <350 cells/mm^3. Meanwhile, a modified definition of late HIV diagnosis which reclassified individuals with primary HIV infection had been previously implemented in Sweden (Brännström *et al.* 2016). The extent of this bias among individuals newly diagnosed with HIV in England, and the potential implications for HIV incidence estimation methods utilising CD4 counts at diagnosis were untested.

Secondly, as a result of the COVID-19 lockdowns in England during 2020 and 2021 there was widespread disruption of both routine healthcare service provision and population mixing (Mude *et al.* 2023). These effects likely impacted both the probability of diagnosis for HIV and the incidence of new cases (Martin, Okumu-Camerra *et al.* 2023). Since the HIV back-calculation parameterises diagnoses in terms of incidence of infection, fixed disease progression, and probability of diagnosis, the impact of a large-scale disruption on the unobserved aspects of the model needed to be tested.

What this thesis adds

I estimated the extent of bias in the late HIV diagnosis measurement among those diagnosed in England by comparing individual-level CD4 count data against evidence for recent HIV acquisition. Misclassification was found to have increased over time, from 10% in 2015 to 18% in 2018, and to vary according to route of exposure (25% for GBM compared to 6% and 7% respectively for heterosexual men and women) and age group (up to 46% for GBM aged 15–24 years). I published a revised definition of late HIV diagnosis (Kirwan, Croxford *et al.* 2022), which has since been adopted by UKHSA and ECDC (Croxford *et al.* 2022; Martin, Okumu-Camerra *et al.* 2023).

To address the effect of this bias on the CD4 back-calculation model, I extended the multi-state model to incorporate additional recent infection states (see Figure 3.6). A simulation study was undertaken to verify that the resulting dual biomarker model could reconstruct

underlying patterns of HIV incidence and diagnosis probabilities, and the model was applied to HIV diagnosis information for England.

I investigated the sensitivity of the CD4 back-calculation model to large-scale disruptions in HIV testing and incidence during the COVID-19 lockdowns, demonstrating that the relative contributions of incidence and probability of diagnosis were appropriately weighted. This investigation establishes a methodological basis for examining counterfactual assumptions about future trends in HIV incidence and diagnosis.

6.2.2 Severity estimation

What was already known

Hospitalised severity for individuals admitted to hospital for COVID-19 had been investigated during the first wave of the pandemic in the UK, with studies drawing associations between hospitalised fatality risk and demographic characteristics, such as age and comorbidity burden (Agrawal, Azcoaga-Lorenzo *et al.* 2021; Docherty, Mulholland *et al.* 2021; Ferrando-Vivas *et al.* 2021; Gray *et al.* 2021; Mathur *et al.* 2021).

The data collection mechanisms used in these studies precluded real-time analysis of hospitalised severity, and none investigated trends in hospitalised severity or lengths of stay in hospital after the first wave. In mid-2020, discussions with clinicians and NHS England indicated an urgent need to understand the various pathways through hospital and how hospital pressures were impacting upon patient care.

In 2021, subsequent to the introduction of widespread vaccination in England, estimates of vaccine effectiveness against adverse outcomes were reported in several studies (Lopez Bernal *et al.* 2021; Thygesen *et al.* 2021; UK Health Security Agency 2021). However, information concerning the impact of vaccination for individuals severe enough to require hospitalisation was scarce.

What this thesis adds

I estimated hospitalised severity and lengths of stay among people admitted to hospital for COVID-19 during the first two waves of infection and prior to vaccine introduction using a flexible, competing-risks, multi-state mixture model. This model was applied to real-time sentinel surveillance data to estimate in real-time the severity of emerging COVID-19 variants, and the association of hospitalised mortality with month of admission and other covariates.

The third wave of COVID-19 in the UK occurred in a context of widespread population vaccination. I estimated the prognosis for hospitalised individuals over all three waves using a stratified Fine-Gray proportional hazards model, with adjustment for month of admission, patient characteristics, and vaccination status. Hospitalised severity and lengths of stay were

found to vary substantially according to these covariates. Evidence of waning vaccine protection against hospitalised fatality had not previously been reported, and informed vaccine policy for those most at risk.

6.2.3 Effectiveness of interventions

What was already known

By 2022 the effectiveness of first and second dose vaccination against SARS-CoV-2 infection and severe outcomes was well-established (Andrews, Tessier *et al.* 2022; Goldberg *et al.* 2021; Hall, Foulkes, Insalata *et al.* 2022). Several UK studies had also investigated the effectiveness of booster vaccination, finding modest protection of third mRNA vaccines against symptomatic COVID-19 infection during the Omicron variant circulating period, with significant waning of protection from 2 months post-vaccination (Andrews, Stowe *et al.* 2022; Hall, Insalata *et al.* 2024; Kirsebom *et al.* 2022). Vaccine effectiveness of fourth (bivalent) doses had not yet been investigated in the UK and international estimates varied considerably (Link-Gelles *et al.* 2023; Shrestha *et al.* 2023).

Several studies had also investigated acquired immunity from infection, finding that a recent SARS-CoV-2 infection provided greater protection than booster vaccination, with waning of this protection beginning at 4–6 months post-infection (Auvigne *et al.* 2023; Huiberts *et al.* 2023). Estimates of the mean duration of PCR positivity varied in study-size and statistical methodology, and no studies had considered the duration of symptomatic vs. asymptomatic infection (Boucau *et al.* 2022; Hay *et al.* 2022; Kojima *et al.* 2022).

The majority of these vaccine effectiveness studies relied on symptomatic testing and did not account for the potential impact of selection bias, test-seeking behaviour, and unmeasured confounding. Whilst a few studies had undertaken a formal causal assessment of vaccination, controlling for confounders (Dagan *et al.* 2021; Hulme *et al.* 2023), none had implemented this in a continuous-time multi-state model framework.

What this thesis adds

I estimated booster vaccine effectiveness and mean duration of SARS-CoV-2 PCR positivity by applying multi-state models for interval-censored data to the SIREN cohort. The effectiveness of a fourth booster vaccination was modest, and waned significantly after the first 2 months, whereas protection from a recent prior infection was more sustained. Duration of PCR positivity was estimated around 7 days, and was shorter for asymptomatic compared to symptomatic infections.

In a causal analysis of vaccine protection, I estimated that the hypothetical scenario in which all healthcare worker received a fourth dose vaccination, compared to everyone remaining on a

third dose, would avert around 14% of infections in this cohort, although there was substantial uncertainty in this estimate.

I presented interim vaccine effectiveness estimates from the SIREN study to UKHSA and the JCVI and these estimates have informed UK vaccination policy (Department of Health and Social Care 2023).

6.3 Methodological developments

This section brings together future methodological developments, motivated by different public health challenges and building on the main findings of the thesis. Common themes are highlighted, followed by a more detailed description of specific developments in each area.

6.3.1 Modelling cluster-level random effects

Both the hospital severity analysis in Chapter 4 and vaccine effectiveness analysis in Chapter 5 were likely to be subject to a degree of unobserved correlation at the hospital (or cluster) level. Accounting for this correlation as a fixed effect led to the model parameters being unidentified. Incorporating hospital-specific random effects (i.e. a shared frailty model) instead would be desirable, and these frailty models can also be used to test the Markov assumption (Titman and Putter 2022). At present, frailty models cannot be implemented within the `msm` package, and the implementation in the `survival` package remains somewhat limited (Jackson 2021; Therneau 1999).

A novel method for fitting a flexible Cox proportional hazards model with a shared frailty term was proposed by Gasperoni *et al.* 2020 and may be applicable to the hospital severity analysis. An alternative approach would be specifying a Bayesian multi-state frailty model (van den Hout 2016), but the MCMC sampling which is required for Bayesian models can present significant computational challenges. These challenges are not unique to this setting and other authors have proposed scalable approaches for random effects in the BUGS modelling framework, and for joint longitudinal and multi-state models (Chen *et al.* 2023; Goudie, Turner *et al.* 2020).

6.3.2 Evidence synthesis

Evidence synthesis methods are a principled way to combine multiple sources of information. These methods have previously been used to model HIV prevalence by combining multiple data sources (Presanis, De Angelis *et al.* 2011; Welton and Ades 2005). Future development could draw on the ideas of 'Markov melding' to join the MPES prevalence model with the HIV back-calculation described in Chapter 3 (Goudie, Presanis *et al.* 2019).

For the hospital severity analyses in Chapter 4, evidence synthesis methods might also be used to explicitly model biases in the real-time SARI-Watch sentinel system, by combining these data with the comprehensive SUS hospital surveillance system (De Angelis *et al.* 2015; Turner *et al.* 2009). Future analyses of the sentinel system could use estimates derived from this evidence synthesis as priors for a set of bias parameters in a Bayesian model, thereby obtaining bias-adjusted or 'weighted' estimates in real time.

6.3.3 Observational study design

Cochran 1965 suggested that a well-designed observational study should resemble, as closely as possible, a controlled or randomised experiment.

The observational cohort study design used for the SIREN study was particularly well-suited to a causal investigation of vaccine effectiveness, with a well-defined analysis plan and clear research question. The other studies in this thesis have instead utilised routine epidemiological data, with the primary analysis being less clear. A risk is that routine data may fail to collect important information on risk factors or case definitions, and may be limited in size or generalisability, limiting the scope of any potential research study. Careful study design can help to overcome some of these biases and may enhance epidemiological data collection, for instance stratifying the study population in a cross-sectional survey so that subgroups of interest are adequately represented (Woodward 2013).

6.3.4 HIV incidence estimation

Age-dependent dual biomarker back-calculation

A previous extension to the CD4-only back-calculation model incorporated information on rates of HIV progression by age group to estimate age-specific HIV incidence and diagnosis rates (Brizzi, Birrell, Plummer *et al.* 2019). Incorporating avidity assay information into this age-dependent back-calculation, as was done for the age-independent model in Chapter 3, may lend additional precision to the estimation of those age-groups at higher risk of infection, whilst accounting for age-dependency in the availability of the avidity assay data.

Incorporating migration in back-calculation

The back-calculation models presented in Chapter 3 assume that individuals are 'at risk' of diagnosis at any point following infection. This limits the model from being applied to populations with substantial incidence outside of England, since the estimated quantities would have a less clear interpretation, i.e. incidence estimates would include infections acquired both within and outside of England. Heterosexual populations, in particular, include a large

proportion of migrants from sub-Saharan African countries with high prevalence of HIV (Martin, Okumu-Camerra *et al.* 2023).

Incorporating a migration process as part of the model would be necessary to enable estimation of HIV incidence among groups with a large proportion of migrants, and would contribute to the estimation of pre and post-migration incidence for an increasingly diverse GBM population (Palich *et al.* 2024). This development requires detailed information on the country of acquisition for newly diagnosed individuals, with patterns of migration incorporated to inform recent trends.

Existing methods to assign country of HIV acquisition rely on CD4 decline slopes inferred from databases of HIV seroconverters, alongside information on age, ethnicity, and country of birth, to estimate the probable year of HIV acquisition at an individual level and compare this to the year of UK arrival. These methods apply similar HIV progression assumptions to the back-calculation, and have been implemented in both frequentist and Bayesian frameworks (Pantazis *et al.* 2019; Yin *et al.* 2021).

Experimental statistics on international migration since 2012 were recently published by the Office for National Statistics (ONS), albeit with information on migrants' nationality currently being very limited (Cheatham 2023). Recent trends in these estimates indicate a shift in migration patterns following the COVID-19 pandemic, with a rise in non-European Union migrants. These trends mirror a rise in HIV diagnoses among non-UK born heterosexual men and women since 2022 (Martin, Okumu-Camerra *et al.* 2023).

A sketch diagram describing a potential multi-state model specification which includes migration is shown in Figure 6.1. Here HIV incidence in the UK is separated from incidence outside of the country, with individuals entering a latent state following migration.

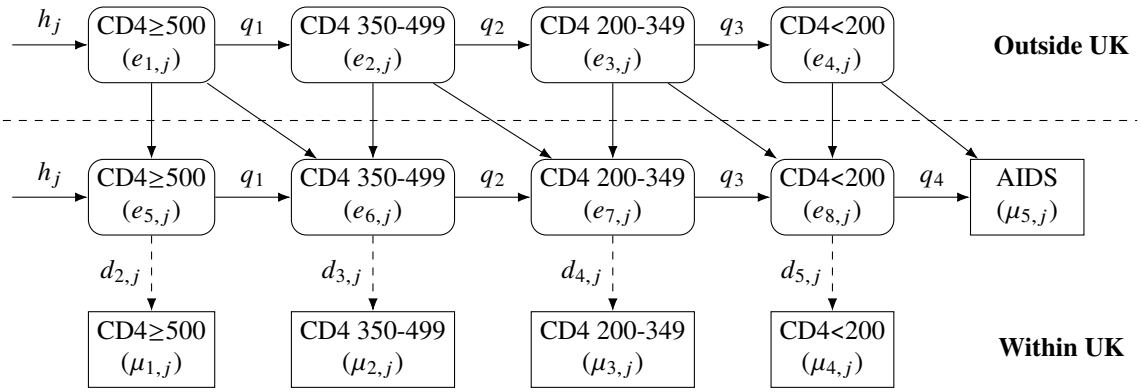


Figure 6.1: Sketch of CD4-staged back-calculation model including migration. Rounded boxes indicate latent states, square boxes indicate diagnosed states. Solid lines indicate HIV progression. Dashed lines indicate transition from latent to diagnosed state.

Evaluation of interventions

Evidence-based assessments of HIV testing campaigns, provision of self-sampling and home HIV testing, PrEP, and condom usage have helped to inform HIV prevention efforts (HIV Commission 2020). Most recently, Cambiano *et al.* 2023 used a mathematical model of HIV incidence to explore the contribution of expanded testing and earlier treatment initiation in reducing HIV incidence, and others have explored the specific impacts of PrEP in reducing HIV transmission (Punyacharoensin *et al.* 2016; Rozhnova *et al.* 2018).

Incorporating transmission dynamics into the back-calculation model might enable similar assessment about the effectiveness of interventions in reducing HIV transmission. Following the example of Presanis, De Angelis *et al.* 2011 and Birrell, Ketsetzis *et al.* 2011, susceptible, infected, and treated components could be included in the model, with interaction between the infected and susceptible components for the generation of new infections.

HIV testing history data

Due to their cost, and the lack of standardisation between laboratories, HIV avidity assays are only undertaken in a few countries (European Centre for Disease Prevention and Control 2023). Data on HIV testing history can provide similar information about recent HIV acquisition, and incorporating these data into a back-calculation model may help to improve the precision of HIV incidence estimates in countries which lack avidity assay testing. As discussed in Chapter 3, a challenge in incorporating last negative HIV test information is that the equivalent 'non-recent' classification cannot be inferred from testing history, and wider trends in testing would likely need to be modelled alongside incidence and diagnosis processes.

6.3.5 Severity estimation

Robust measure of hospital pressure

In analyses of hospitalised fatality during the COVID-19 pandemic, increased hospital load, measured by number of admissions, was associated with poorer outcomes. This metric of load was relatively crude, however, and did not incorporate nuances in staffing ratios, ward capacity, or Red/Green ward demarcation (NHS England 2020b). Four key elements which limit ICU surge capacity during a pandemic have been pointed to: the availability of staff, consumables and essential equipment, bed-spaces, and management systems (so-termed 'staff, stuff, space, and systems') (Fong *et al.* 2024; Gabriel and Farrar 2019).

An improved measure of hospital pressure should consider the demand and supply sides of these four elements, and could be a composite of several metrics to indicate which trusts are experiencing (or likely to experience) extreme pressure. This measure would provide

actionable information to the NHS, allowing targeted interventions to better balance load across the healthcare service.

Whilst the range of data currently available are limited, measures of hospital demand might include: the number of occupied beds, number of patients arriving at emergency care, and transfers from other hospitals. Information on staffing levels (or patient:staff ratios) on any given day, delays in emergency care admission, and the number of continuous days that clinicians have been on duty might be examples of supply (Figure 6.2).

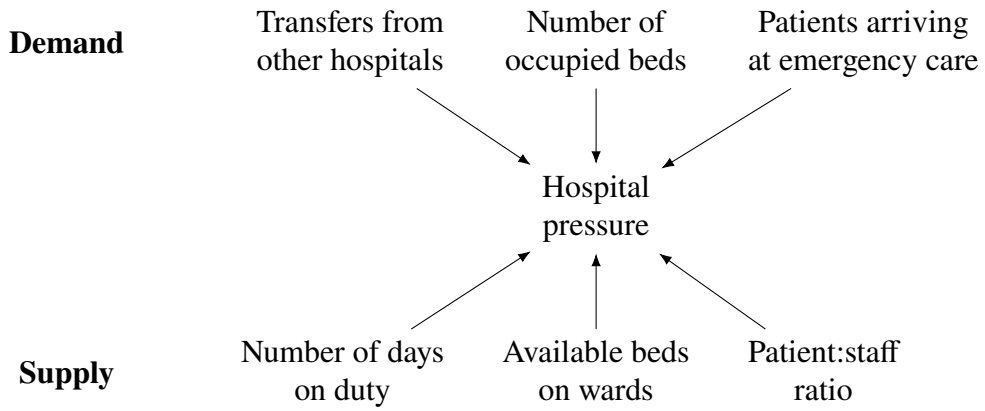


Figure 6.2: Factors of supply and demand which may be indications for hospital pressure.

Of course, the complexities of hospital pressure expand well beyond these factors, and it is infeasible that any single quantitative measure could succinctly represent all aspects of pressure within a healthcare system. Qualitative surveys can help to refine our understanding about the effects of hospital pressure: during the COVID-19 pandemic qualitative surveys among ICU clinicians found substantial rates of poor mental health, with half reporting symptoms consistent with probable post-traumatic stress disorder, leading to operational impairment, and likely to further increase pressure on services (Greenberg *et al.* 2021; Hall, Milward *et al.* 2022). Implementation of qualitative surveys and assessments of the limits of capacity should be a key consideration for future studies (Fong *et al.* 2024).

Causal relationships for hospital severity

A causal diagram, introduced in Section 2.6, can aid in visualising which relationships to consider to make causal inference from observational hospital data. Figure 6.3 shows the causal relationship between a range of covariates collected at hospital admission and hospitalised severity.

In this causal diagram, the 'Hospital' covariate is a confounder which influences both hospital pressure and hospitalied severity, with the path 'Hospital pressure ← Hospital → Severity' known as a 'backdoor path'. For a causal analysis this covariate should be conditioned

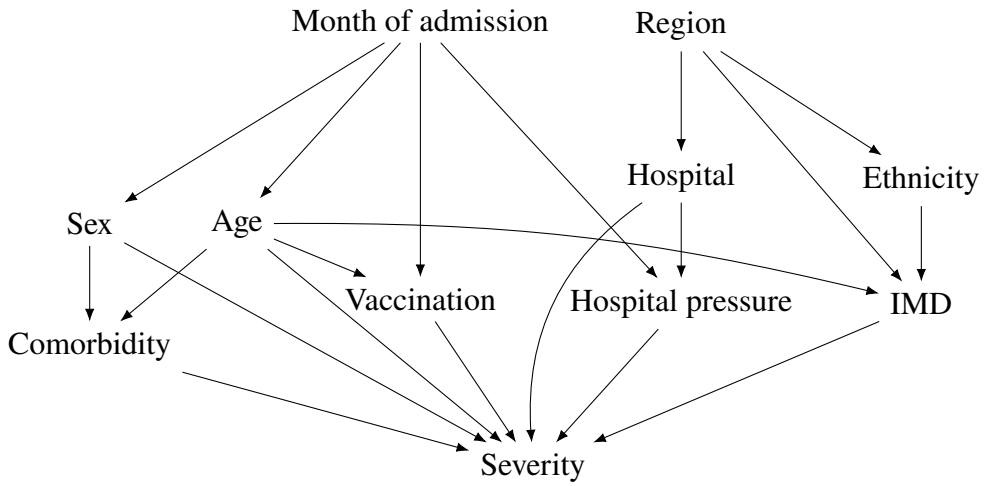


Figure 6.3: Causal DAG of discrete measurements for hospitalised patients.

on (Hernán and Robins 2023). Where this is not possible, e.g. due to identifiability issues, other summaries of the hospital covariate could be considered, or the inclusion of cluster-level random effects as discussed earlier in this section.

6.3.6 Effectiveness of interventions

Retrospective assessment of SARS-CoV-2 vaccine effectiveness

So far, multi-state models have only been applied to investigate the vaccine effectiveness of fourth doses in the SIREN dataset. However, with a large quantity of available testing data, a retrospective analysis could also be undertaken. This analysis might utilise all of the available PCR and antibody testing data to estimate the time-varying protection of vaccination and prior infection throughout the COVID-19 pandemic. Given the large number of parameters that would be included in these models, and limited availability of data, the scalable approaches to include fixed and random effects previously discussed should be investigated.

6.4 Applications of multi-state models

As well as methodological developments, there are several future applications of multi-state models where the findings of this thesis can be directly applied.

6.4.1 Ongoing monitoring of infectious disease burden

Ongoing monitoring of HIV incidence, COVID-19 hospital pressures, and vaccine effectiveness utilising the multi-state models described in this thesis will continue to provide information for public health action.

For HIV, annually updated incidence estimates will help to inform progress towards the UK's stated goal of eliminating HIV as a public health threat by 2030, with projections of future incidence helping to gauge progress towards these goals (HIV Commission 2020).

The well-documented and actively maintained competing risks multi-state models used to estimate risks of hospitalised fatality and lengths of stay address specific challenges in observational data (Jackson *et al.* 2022; Kirwan 2022). These methods will continue to be applied to routinely collected data to assess the changing burden of COVID-19 and provide information about novel variants and diseases.

The multi-state models developed for the SIREN study will be applied to data collected over the autumn/winter 2023/24 period to provide updated fifth dose vaccine effectiveness estimates (UK Health Security Agency 2023b). As well as protection against acquisition of infection, this research will aim to answer key policy questions about the effect of vaccination on the clinical significance of infection.

6.4.2 Collaborations and knowledge transfer

Disseminating information about multi-state models and enhancing the user-friendliness of these methods has led to increased adoption by public health agencies (Kirwan, Hall *et al.* 2024; van Sighem, Nakagawa *et al.* 2015). Collaboration between researchers and public health agencies will continue to promote high quality research and knowledge transfer.

The importance of this collaboration has been recognised by the Health Protection Research Units (HPRUs). These research partnerships between universities and UKHSA encourage high quality research to mitigate the impact of public health emergencies (National Institute for Health and Care Research 2024).

6.4.3 Statistical communication

Finally, communicating the strengths and limitations of multi-state model estimates to the wider community of public health practitioners, healthcare professionals, the media, and the public is important to ensure correct interpretation of research findings, and to understand when and why estimates are more useful than data alone.

For HIV, incidence estimates are regularly used by government bodies and third sector organisations, such as the National AIDS Trust and Terrence Higgins Trust, in their reports and prevention campaigns (Martin, Okumu-Camerra *et al.* 2023; Terrence Higgins Trust 2023).

Biostatisticians have a role in supporting these organisations to understand and communicate the uncertainty around different estimates.

During the COVID-19 pandemic, much of the reporting of statistical estimates by media organisations in the UK was of low scientific quality and may have failed to alert readers to public-health risks (Mach *et al.* 2021). Uncertainty around estimates tended to be particularly poorly communicated, and may influence the perceived trustworthiness of statistics (van der Bles *et al.* 2019; Schneider *et al.* 2022).

References

Aalen, O. O., Farewell, V. T., De Angelis, D., Day, N. E. *et al.* (1997). 'A Markov model for HIV disease progression including the effect of HIV diagnosis and treatment: application to AIDS prediction in England and Wales'. *Stat. Med.* 16 (19), pp. 2191–2210. DOI: 10.1002/(sici)1097-0258(19971015)16:19<2191::aid-sim645>3.0.co;2-5.

Aalen, O. O., Røysland, K., Gran, J. M., Kouyos, R. *et al.* (2016). 'Can we believe the DAGs? A comment on the relationship between causal DAGs and mechanisms'. *Stat. Methods Med. Res.* 25 (5), pp. 2294–2314. DOI: 10.1177/0962280213520436.

Aalen, O., Borgan, O. and Gjessing, H. (2008). *Survival and Event History Analysis: A Process Point of View*. Springer. 540 pp.

Aalen, O. O. and Johansen, S. (1978). 'An Empirical Transition Matrix for Non-Homogeneous Markov Chains Based on Censored Observations'. *Scand. Stat. Theory Appl.* 5 (3), pp. 141–150.

Aghaizu, A., Murphy, G., Tosswill, J., De Angelis, D. *et al.* (2014). 'Recent infection testing algorithm (RITA) applied to new HIV diagnoses in England, Wales and Northern Ireland, 2009 to 2011'. *Euro Surveill.* 19 (2). DOI: 10.2807/1560-7917.es2014.19.2.20673.

Aghaizu, A. R.-L. (2018). 'The public health utility of assays to test for recent HIV infection: an evaluation on UK case-based surveillance data'. Imperial College London. DOI: 10.25560/96930.

Agrawal, U., Azcoaga-Lorenzo, A., Fagbamigbe, A. F., Vasileiou, E. *et al.* (2021). 'Association between multimorbidity and mortality in a cohort of patients admitted to hospital with COVID-19 in Scotland'. *J. R. Soc. Med.*, p. 1410768211051715. DOI: 10.1177/01410768211051715.

Agrawal, U., Katikireddi, S. V., McCowan, C., Mulholland, R. H. *et al.* (2021). 'COVID-19 hospital admissions and deaths after BNT162b2 and ChAdOx1 nCoV-19 vaccinations in 2·57 million people in Scotland (EAVE II): a prospective cohort study'. *Lancet Respir. Med.* 9 (12), pp. 1439–1449. DOI: 10.1016/S2213-2600(21)00380-5.

Aguirre-Hernández, R. and Farewell, V. T. (2002). 'A Pearson-type goodness-of-fit test for stationary and time-continuous Markov regression models'. *Stat. Med.* 21 (13), pp. 1899–1911. DOI: 10.1002/sim.1152.

Alberts, B., Johnson, A., Lewis, J., Raff, M. *et al.* (2002). 'Introduction to Pathogens'. *Molecular Biology of the Cell. 4th edition*. Garland Science.

Ambrosioni, J., Petit, E., Liegeon, G., Laguno, M. *et al.* (2021). 'Primary HIV-1 infection in users of pre-exposure prophylaxis'. *Lancet HIV*. 8 (3), e166–e174. DOI: 10.1016/S2352-3018(20)30271-X.

Anderegg, N., Panczak, R., Egger, M., Low, N. *et al.* (2022). 'Survival among people hospitalized with COVID-19 in Switzerland: a nationwide population-based analysis'. *BMC Med.* 20 (1), p. 164. DOI: 10.1186/s12916-022-02364-7.

Andersen, P. K., Abildstrom, S. Z. and Rosthøj, S. (2002). 'Competing risks as a multi-state model'. *Stat. Methods Med. Res.* 11 (2), pp. 203–215. DOI: 10.1191/0962280202sm281ra.

Andersen, P. K., Borgan, O., Gill, R. D. and Keiding, N. (1996). *Statistical models based on counting processes*. 1st ed. Springer series in statistics. New York, NY: Springer. 784 pp. DOI: 10.1007/978-1-4612-4348-9.

Andersen, P. K., Geskus, R. B., de Witte, T. and Putter, H. (2012). 'Competing risks in epidemiology: possibilities and pitfalls'. *Int. J. Epidemiol.* 41 (3), pp. 861–870. DOI: 10.1093/ije/dyr213.

Andersen, P. K. and Keiding, N. (2002). 'Multi-state models for event history analysis'. *Stat. Methods Med. Res.* 11 (2), pp. 91–115. DOI: 10.1191/0962280202SM276ra.

Andrews, N., Stowe, J., Kirsebom, F., Toffa, S. *et al.* (2022). 'Covid-19 Vaccine Effectiveness against the Omicron (B.1.1.529) Variant'. *N. Engl. J. Med.* 386 (16), pp. 1532–1546. DOI: 10.1056/NEJMoa2119451.

Andrews, N., Tessier, E., Stowe, J., Gower, C. *et al.* (2022). 'Duration of Protection against Mild and Severe Disease by Covid-19 Vaccines'. *N. Engl. J. Med.* 386 (4), pp. 340–350. DOI: 10.1056/NEJMoa2115481.

Angyal, A., Longet, S., Moore, S. C., Payne, R. P. *et al.* (2022). 'T-cell and antibody responses to first BNT162b2 vaccine dose in previously infected and SARS-CoV-2-naive UK health-care workers: a multicentre prospective cohort study'. *Lancet Microbe*. 3 (1), e21–e31. DOI: 10.1016/S2666-5247(21)00275-5.

Antinori, A., Coenen, T., Costagiola, D., Dedes, N. *et al.* (2011). 'Late presentation of HIV infection: a consensus definition'. *HIV Med.* 12 (1), pp. 61–64. DOI: 10.1111/j.1468-1293.2010.00857.x.

Austin, P. C. (2017). 'A Tutorial on Multilevel Survival Analysis: Methods, Models and Applications'. *Int. Stat. Rev.* 85 (2), pp. 185–203. DOI: 10.1111/insr.12214.

Austin, P. C., Lee, D. S. and Fine, J. P. (2016). 'Introduction to the analysis of survival data in the presence of competing risks'. *Circulation*. 133 (6), pp. 601–609. DOI: 10.1161/CIRCULATIONAHA.115.017719.

Auvigne, V., Tamandjou Tchuem, C. R., Schaeffer, J., Vaux, S. *et al.* (2023). 'Protection against symptomatic SARS-CoV-2 infection conferred by the Pfizer-BioNTech Original/BA.4-5 bivalent vaccine compared to the mRNA Original monovalent vaccines - A matched cohort study in France'. *Vaccine*. 41 (38), pp. 5490–5493. DOI: 10.1016/j.vaccine.2023.07.071.

Baker, C. (2023). 'Study Designs'. *Epidemiology*. Open Education Initiative of the University Libraries at Virginia Tech.

Balan, T. A. and Putter, H. (2020). 'A tutorial on frailty models'. *Stat. Methods Med. Res.* 29 (11), pp. 3424–3454. DOI: 10.1177/0962280220921889.

Bamford, P., Bentley, A., Dean, J., Whitmore, D. *et al.* (2020). 'ICS guidance for prone positioning of the conscious COVID patient'. Intensive Care Society. URL: https://emcrit.org/wp-content/uploads/2020/04/2020-04-12-Guidance-for-conscious-proning.pdf.

Barda, N., Dagan, N., Cohen, C., Hernán, M. A. *et al.* (2021). 'Effectiveness of a third dose of the BNT162b2 mRNA COVID-19 vaccine for preventing severe outcomes in Israel: an observational study'. *Lancet*. 398 (10316), pp. 2093–2100. DOI: 10.1016/S0140-6736(21)02249-2.

Barré-Sinoussi, F., Chermann, J. C., Rey, F., Nugeyre, M. T. *et al.* (1983). 'Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS)'. *Science*. 220 (4599), pp. 868–871. DOI: 10.1126/science.6189183.

Becker, N. G., Watson, L. F. and Carlin, J. B. (1991). 'A method of non-parametric back-projection and its application to AIDS data'. *Stat. Med.* 10 (10), pp. 1527–1542. DOI: 10.1002/sim.4780101005.

Bellocco, R. and Marschner, I. C. (2000). 'Joint analysis of HIV and AIDS surveillance data in back-calculation'. *Statistics in Medicine*. 19 (3), pp. 297–311. DOI: 10.1002/(SICI)1097-0258(20000215)19:3<297::AID-SIM340>3.0.CO;2-6.

Benedictow, O. J. (2004). *The black death 1346-1353: The complete history*. Woodbridge, England: Boydell Press. 454 pp.

Bengtsson, H. (2007). *matrixStats: R package: Methods that Apply to Rows and Columns of Matrices (and to Vectors)*. URL: https://github.com/HenrikBengtsson/matrixStats (visited on 21/02/2024).

Bhanot, D., Singh, T., Verma, S. K. and Sharad, S. (2020). 'Stigma and discrimination during COVID-19 pandemic'. *Front. Public Health*. 8, p. 577018. DOI: 10.3389/fpubh.2020.577018.

Bhattacharya, A., Collin, S. M., Stimson, J., Thelwall, S. *et al.* (2021). 'Healthcare-associated COVID-19 in England: A national data linkage study'. *J. Infect.* 83 (5), pp. 565–572. DOI: 10.1016/j.jinf.2021.08.039.

Birrell, P., Blake, J., van Leeuwen E, Gent, N. *et al.* (2021). 'Real-time nowcasting and forecasting of COVID-19 dynamics in England: the first wave'. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 376 (1829), p. 20200279. DOI: 10.1098/rstb.2020.0279.

Birrell, P., Blake, J., van Leeuwen E, MRC Biostatistics Unit COVID-19 Working Group *et al.* (2021). 'Nowcasting and Forecasting of the COVID-19 Pandemic'. MRC Biostatistics Unit, University of Cambridge. URL: https://www.mrc-bsu.cam.ac.uk/tackling-covid-19/nowcasting-and-forecasting-of-covid-19/.

Birrell, P. J., Chadborn, T. R., Gill, O. N., Delpech, V. C. *et al.* (2012). 'Estimating trends in incidence, time-to-diagnosis and undiagnosed prevalence using a CD4-based Bayesian back-calculation'. *Stat. Commun. Infect. Dis.* 4 (1). DOI: 10.1515/1948-4690.1055.

Birrell, P. J., Gill, O. N., Delpech, V. C., Brown, A. E. *et al.* (2013). 'HIV incidence in men who have sex with men in England and Wales 2001-10: a nationwide population study'. *Lancet Infect. Dis.* 13 (4), pp. 313–318. DOI: 10.1016/S1473-3099(12)70341-9.

Birrell, P. J., Ketsetzis, G., Gay, N. J., Cooper, B. S. *et al.* (2011). 'Bayesian modeling to unmask and predict influenza A/H1N1pdm dynamics in London'. *Proc. Natl. Acad. Sci. U. S. A.* 108 (45), pp. 18238–18243. DOI: 10.1073/pnas.1103002108.

van der Bles, A. M., van der Linden, S., Freeman, A. L. J., Mitchell, J. *et al.* (2019). 'Communicating uncertainty about facts, numbers and science'. *R. Soc. Open Sci.* 6 (5), p. 181870. DOI: 10.1098/rsos.181870.

Borgan, Ø. (2014). *Nelson-Aalen Estimator*. *Wiley StatsRef: Statistics Reference Online*. Chichester, UK: John Wiley & Sons, Ltd. DOI: 10.1002/9781118445112.stat06045. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118445112.stat06045.

Boucau, J., Marino, C., Regan, J., Uddin, R. *et al.* (2022). 'Duration of Shedding of Culturable Virus in SARS-CoV-2 Omicron (BA.1) Infection'. *N. Engl. J. Med.* 387 (3), pp. 275–277. DOI: 10.1056/NEJMc2202092.

Bradburn, M. J., Clark, T. G., Love, S. B. and Altman, D. G. (2003). 'Survival analysis part II: multivariate data analysis–an introduction to concepts and methods'. *Br. J. Cancer*. 89 (3), pp. 431–436. DOI: 10.1038/sj.bjc.6601119.

Brännström, J., Svedhem Johansson, V., Marrone, G., Wendahl, S. *et al.* (2016). 'Deficiencies in the health care system contribute to a high rate of late HIV diagnosis in Sweden'. *HIV Med.* 17 (6), pp. 425–435. DOI: 10.1111/hiv.12321.

Brauner, J. M., Mindermann, S., Sharma, M., Johnston, D. *et al.* (2021). 'Inferring the effectiveness of government interventions against COVID-19'. *Science*. 371 (6531), eabd9338. DOI: 10.1126/science.abd9338.

British Broadcasting Company (1984). '1984: Scientist finds Aids virus'. URL: http://news.bbc.co.uk/onthisday/hi/dates/stories/april/23/newsid_2524000/2524039.stm (visited on 14/11/2023).

British HIV Association (2018). 'British HIV Association Standards of care for people living with HIV 2018'. British HIV Association. URL: https://www.bhiva.org/standards-of-care-2018.

Brizzi, F. (2017). 'Estimating HIV incidence from multiple sources of data'. University of Cambridge. DOI: 10.17863/CAM.37782.

Brizzi, F., Birrell, P. J., Kirwan, P., Ogaz, D. *et al.* (2021). 'Tracking elimination of HIV transmission in men who have sex with men in England: a modelling study'. *Lancet HIV*. 8 (7), e440–e448. DOI: 10.1016/S2352-3018(21)00044-8.

Brizzi, F., Birrell, P. J., Plummer, M. T., Kirwan, P. *et al.* (2019). 'Extending Bayesian back-calculation to estimate age and time specific HIV incidence'. *Lifetime Data Anal.* 25 (4), pp. 757–780. DOI: 10.1007/s10985-019-09465-1.

Brookmeyer, R. and Quinn, T. C. (1995). 'Estimation of current human immunodeficiency virus incidence rates from a cross-sectional survey using early diagnostic tests'. *Am. J. Epidemiol.* 141 (2), pp. 166–172. DOI: 10.1093/oxfordjournals.aje.a117404.

Brookmeyer, R. and Gail, M. H. (1988). 'A Method for Obtaining Short-Term Projections and Lower Bounds on the Size of the AIDS Epidemic'. *J. Am. Stat. Assoc.* 83 (402), pp. 301–308. DOI: 10.1080/01621459.1988.10478599.

Brookmeyer, R., Gail, M. H., Medical Statistician Epidemiology and Biostatistics Program Mitchell H Gail and . Gail, M. H. (1994). *AIDS Epidemiology: A Quantitative Approach*. Oxford University Press. 354 pp.

Brookmeyer, R., Konikoff, J., Laeyendecker, O. and Eshleman, S. H. (2013). 'Estimation of HIV incidence using multiple biomarkers'. *Am. J. Epidemiol.* 177 (3), pp. 264–272. DOI: 10.1093/aje/kws436.

Brown, A. E., Nash, S., Connor, N., Kirwan, P. D. *et al.* (2018). 'Towards elimination of HIV transmission, AIDS and HIV-related deaths in the UK'. *HIV Med.* 19 (8), pp. 505–512. DOI: 10.1111/hiv.12617.

Brown, A. E., Mohammed, H., Ogaz, D., Kirwan, P. D. *et al.* (2017). 'Fall in new HIV diagnoses among men who have sex with men (MSM) at selected London sexual health clinics since early 2015: testing or treatment or pre-exposure prophylaxis (PrEP)?' *Euro Surveill.* 22 (25). DOI: 10.2807/1560-7917.ES.2017.22.25.30553.

Cambiano, V., Miners, A., Lampe, F. C., McCormack, S. *et al.* (2023). 'The effect of combination prevention strategies on HIV incidence among gay and bisexual men who have sex with men in the UK: a model-based analysis'. *Lancet HIV*. 10 (11), e713–e722. DOI: 10.1016/S2352-3018(23)00204-7.

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D. *et al.* (2017). 'Stan: A probabilistic programming language'. *J. Stat. Softw.* 76 (1), pp. 1–32. DOI: 10.18637/jss.v076.i01.

Carter, R. and Mendis, K. N. (2002). 'Evolutionary and historical aspects of the burden of malaria'. *Clin. Microbiol. Rev.* 15 (4), pp. 564–594. DOI: 10.1128/cmr.15.4.564-594.2002.

CASCADE (Concerted Action on SeroConversion to AIDS and Death in Europe) Collaboration (2000). 'Changes in the uptake of antiretroviral therapy and survival in people with known duration of HIV infection in Europe: results from CASCADE'. *HIV Med.* 1 (4), pp. 224–231. DOI: 10.1046/j.1468-1293.2000.00033.x.

Charlson, M. E., Pompei, P., Ales, K. L. and MacKenzie, C. R. (1987). 'A new method of classifying prognostic comorbidity in longitudinal studies: development and validation'. *J. Chronic Dis.* 40 (5), pp. 373–383. DOI: 10.1016/0021-9681(87)90171-8.

Cheatham, L. (2023). 'Long-term international migration, provisional - Office for National Statistics'. URL: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/internationalmigration/bulletins/longterminternationalmigrationprovisional/yearendingjune2023 (visited on 05/03/2024).

Chemaitelly, H., Ayoub, H. H., AlMukdad, S., Coyle, P. *et al.* (2022). 'Duration of mRNA vaccine protection against SARS-CoV-2 Omicron BA.1 and BA.2 subvariants in Qatar'. *Nat. Commun.* 13 (1), p. 3082. DOI: 10.1038/s41467-022-30895-3.

Chen, S., Marshall, T., Jackson, C., Cooper, J. *et al.* (2023). 'Sociodemographic characteristics and longitudinal progression of multimorbidity: A multistate modelling analysis of a large primary care records dataset in England'. *PLoS Med.* 20 (11), e1004310. DOI: 10.1371/journal.pmed.1004310.

Chiang, C. L. (1980). *An introduction to stochastic processes and their applications*. RE Krieger Publishing Company New York.

Churchill, D., Waters, L., Ahmed, N., Angus, B. *et al.* (2016). 'British HIV Association guidelines for the treatment of HIV-1-positive adults with antiretroviral therapy 2015'. *HIV Med.* 17 Suppl 4, s2–s104. DOI: 10.1111/hiv.12426.

Cochran, W. G. (1965). 'The planning of observational studies of human populations'. *J. R. Stat. Soc. Ser. A*. 128 (2), p. 234. DOI: 10.2307/2344179.

Coggon, D., Rose, G. and Barker, D. J. (1978). 'Epidemiology for the uninitiated, fourth edition'. *BMJ*. 2 (6142), pp. 941–942. DOI: 10.1136/bmj.2.6142.941.

Cohen, M. S., Chen, Y. Q., McCauley, M., Gamble, T. *et al.* (2016). 'Antiretroviral Therapy for the Prevention of HIV-1 Transmission'. *N. Engl. J. Med.* 375 (9), pp. 830–839. DOI: 10.1056/NEJMoa1600693.

Cohen, M. S., Gay, C. L., Busch, M. P. and Hecht, F. M. (2010). 'The detection of acute HIV infection'. *J. Infect. Dis.* 202 Suppl 2, S270–7. DOI: 10.1086/655651.

Collaborative Group on AIDS Incubation and HIV Survival (2000). 'Time from HIV-1 seroconversion to AIDS and death before widespread use of highly-active antiretroviral therapy: a collaborative re-analysis'. *Lancet*. 355 (9210), pp. 1131–1137. DOI: 10.1016/S0140-6736(00)02061-4.

Collett, D. (2023). *Modelling survival data in medical research*. 4th ed. Chapman & Hall/CRC Texts in Statistical Science. Philadelphia, PA: Chapman & Hall/CRC. 540 pp.

Contini, C., Di Nuzzo, M., Barp, N., Bonazza, A. *et al.* (2020). 'The novel zoonotic COVID-19 pandemic: An expected global health concern'. *J. Infect. Dev. Ctries.* 14 (3), pp. 254–264. DOI: 10.3855/jidc.12671.

Cooper, B. S., Evans, S., Jafari, Y., Pham, T. M. *et al.* (2023). 'The burden and dynamics of hospital-acquired SARS-CoV-2 in England'. *Nature*. 623 (7985), pp. 132–138. DOI: 10.1038/s41586-023-06634-z.

Cooper, B. S., Pitman, R. J., Edmunds, W. J. and Gay, N. J. (2006). 'Delaying the international spread of pandemic influenza'. *PLoS Med.* 3 (6), e212. DOI: 10.1371/journal.pmed.0030212.

Corbella, A., Zhang, X.-S., Birrell, P. J., Boddington, N. *et al.* (2018). 'Exploiting routinely collected severe case data to monitor and predict influenza outbreaks'. *BMC Public Health*. 18 (1), p. 790. DOI: 10.1186/s12889-018-5671-7.

Cormen, T. H. and Leiserson, C. E. (2022). *Introduction to Algorithms, fourth edition*. London, England: MIT Press. 1332 pp.

Covid-19 EpiCell (2020). 'PHE data series on deaths in people with COVID-19: technical summary'. Public Health England. URL: https://www.gov.uk/government/publications/phe-data-series-on-deaths-in-people-with-covid-19-technical-summary.

Cox, D. R. (1972). 'Regression models and life-tables'. *J. R. Stat. Soc.* 34 (2), pp. 187–202. DOI: 10.1111/j.2517-6161.1972.tb00899.x.

Croxford, S., Stengaard, A. R., Brännström, J., Combs, L. *et al.* (2022). 'Late diagnosis of HIV: An updated consensus definition'. *HIV Med.* 23 (11), pp. 1202–1208. DOI: 10.1111/hiv.13425.

Cui, J. and Becker, N. G. (2000). 'Estimating HIV incidence using dates of both HIV and AIDS diagnoses'. *Statistics in Medicine*. 19 (9), pp. 1165–1177. DOI: 10.1002/(SICI)1097-0258(20000515)19:9<1165::AID-SIM419>3.0.CO;2-7.

Dagan, N., Barda, N., Kepten, E., Miron, O. *et al.* (2021). 'BNT162b2 mRNA Covid-19 Vaccine in a Nationwide Mass Vaccination Setting'. *N. Engl. J. Med.* 384 (15), pp. 1412–1423. DOI: 10.1056/NEJMoa2101765.

Dahdouh, E., Lázaro-Perona, F., Romero-Gómez, M. P., Mingorance, J. *et al.* (2021). 'Ct values from SARS-CoV-2 diagnostic PCR assays should not be used as direct estimates of viral load'. *J. Infect.* 82 (3), pp. 414–451. DOI: 10.1016/j.jinf.2020.10.017.

Davies, N. G., Barnard, R. C., Jarvis, C. I., Russell, T. W. *et al.* (2020). 'Association of tiered restrictions and a second lockdown with COVID-19 deaths and hospital admissions in England: a modelling study'. *Lancet Infect. Dis.* DOI: 10.1016/S1473-3099(20)30984-1.

Davies, N. G., Jarvis, C. I., CMMID COVID-19 Working Group, Edmunds, W. J. *et al.* (2021). 'Increased mortality in community-tested cases of SARS-CoV-2 lineage B.1.1.7'. *Nature*. 593 (7858), pp. 270–274. DOI: 10.1038/s41586-021-03426-1.

De Angelis, D., Presanis, A. M., Birrell, P. J., Tomba, G. S. *et al.* (2015). 'Four key challenges in infectious disease modelling using data from multiple sources'. *Epidemics*. 10, pp. 83–87. DOI: 10.1016/j.epidem.2014.09.004.

Dean, N. E., Hogan, J. W. and Schnitzer, M. E. (2021). 'Covid-19 vaccine effectiveness and the test-negative design'. *N. Engl. J. Med.* 385 (15), pp. 1431–1433. DOI: 10.1056/NEJMe2113151.

Deeks, S. G., Lewin, S. R. and Havlir, D. V. (2013). 'The end of AIDS: HIV infection as a chronic disease'. *Lancet*. 382 (9903), pp. 1525–1533. DOI: 10.1016/S0140-6736(13)61809-7.

Delaney, K. P., Hanson, D. L., Masciotra, S., Ethridge, S. F. *et al.* (2017). 'Time Until Emergence of HIV Test Reactivity Following Infection With HIV-1: Implications for Interpreting Test Results and Retesting After Exposure'. *Clin. Infect. Dis.* 64 (1), pp. 53–59. DOI: 10.1093/cid/ciw666.

Delpech, V., Brown, A. E., Croxford, S., Chau, C. *et al.* (2013). 'Quality of HIV care in the United Kingdom: key indicators for the first 12 months from HIV diagnosis'. *HIV Med.* 14 Suppl 3, pp. 19–24. DOI: 10.1111/hiv.12070.

Delpech, V. (2022). 'The HIV epidemic: global and United Kingdom trends'. *Medicine*. 50 (4), pp. 202–204. DOI: 10.1016/j.mpmed.2022.01.002.

Department of Health and Social Care (2019). 'Health Secretary announces goal to end HIV transmissions by 2030'. URL: https://www.gov.uk/government/news/health-secretary-announces-goal-to-end-hiv-transmissions-by-2030 (visited on 24/02/2021).

Department of Health and Social Care (2020a). 'Face coverings: when to wear one, exemptions and what makes a good one'. URL: https://www.gov.uk/government/publications/face-coverings-when-to-wear-one-and-how-to-make-your-own (visited on 14/07/2023).

Department of Health and Social Care (2020b). 'Priority groups for coronavirus (COVID-19) vaccination: advice from the JCVI, 30 December 2020'. URL: https://www.gov.uk/government/publications/priority-groups-for-coronavirus-covid-19-vaccination-advice-from-the-jcvi-30-december-2020 (visited on 14/07/2023).

Department of Health and Social Care (2021a). 'Optimising the COVID-19 vaccination programme for maximum short-term impact'. URL: https://www.gov.uk/government/publications/prioritising-the-first-covid-19-vaccine-dose-jcvi-statement/optimising-the-covid-19-vaccination-programme-for-maximum-short-term-impact (visited on 15/01/2024).

Department of Health and Social Care (2021b). 'Towards Zero - An action plan towards ending HIV transmission, AIDS and HIV-related deaths in England - 2022 to 2025'. URL: https://www.gov.uk/government/publications/towards-zero-the-hiv-action-plan-for-england-2022-to-2025/towards-zero-an-action-plan-towards-ending-hiv-transmission-aids-and-hiv-related-deaths-in-england-2022-to-2025 (visited on 07/01/2022).

Department of Health and Social Care (2021c). 'UK COVID-19 vaccine uptake plan'. URL: https://www.gov.uk/government/publications/covid-19-vaccination-uptake-plan/uk-covid-19-vaccine-uptake-plan (visited on 15/01/2024).

Department of Health and Social Care (2023). 'JCVI statement on the COVID-19 vaccination programme for 2023: 8 November 2022'. URL: https://www.gov.uk/government/publications/covid-19-vaccination-programme-for-2023-jcvi-interim-advice-8-november-2022/jcvi-

statement-on-the-covid-19-vaccination-programme-for-2023-8-november-2022 (visited on 05/03/2024).

Didelez, V. (2008). 'Graphical Models for Marked Point Processes Based on Local Independence'. *J. R. Stat. Soc. Series B Stat. Methodol.* 70 (1), pp. 245–264. DOI: 10.1111/j.1467-9868.2007.00634.x.

Docherty, A. B., Harrison, E. M., Green, C. A., Hardwick, H. E. *et al.* (2020). 'Features of 20 133 UK patients in hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: prospective observational cohort study'. *BMJ*. 369, p. m1985. DOI: 10.1136/bmj.m1985.

Docherty, A. B., Mulholland, R. H., Lone, N. I., Cheyne, C. P. *et al.* (2021). 'Changes in in-hospital mortality in the first wave of COVID-19: a multicentre prospective observational cohort study using the WHO Clinical Characterisation Protocol UK'. *Lancet Respir. Med.* 9 (7), pp. 773–785. DOI: 10.1016/S2213-2600(21)00175-2.

Dowdle, W. R. (1983). 'The epidemiology of AIDS'. *Public Health Rep.* 98 (4), pp. 308–312.

Egan, J. R. and Hall, I. M. (2015). 'A review of back-calculation techniques and their potential to inform mitigation strategies with application to non-transmissible acute infectious diseases'. *J. R. Soc. Interface*. 12 (106). DOI: 10.1098/rsif.2015.0096.

European AIDS Clinical Society (2023). 'EACS Guidelines 12.0'. European AIDS Clinical Society. URL: https://www.eacsociety.org/guidelines/eacs-guidelines/.

European Centre for Disease Prevention and Control (2020). 'HIV/AIDS surveillance in Europe'. European Centre for Disease Prevention and Control. URL: https://www.ecdc.europa.eu/en/publications-data/hivaids-surveillance-europe-2020-2019-data.

European Centre for Disease Prevention and Control (2023). 'HIV/AIDS surveillance in Europe 2023 (2022 data)'. European Centre for Disease Prevention and Control. URL: https://www.ecdc.europa.eu/en/publications-data/hivaids-surveillance-europe-2023-2022-data.

European Testing Week (2023). 'European Testing Week'. URL: http://www.testingweek.eu (visited on 27/11/2023).

Evans, S., Stimson, J., Pople, D., Bhattacharya, A. *et al.* (2022). 'Quantifying the contribution of pathways of nosocomial acquisition of COVID-19 in English hospitals'. *Int. J. Epidemiol.* 51 (2), pp. 393–403. DOI: 10.1093/ije/dyab241.

Ewbank, L., Thompson, J., McKenna, H., Anandaciva, S. *et al.* (2021). 'NHS hospital bed numbers past, present, future'. URL: https://www.kingsfund.org.uk/publications/nhs-hospital-bed-numbers (visited on 16/12/2021).

Fagerland, M. W., Hosmer, D. W. and Bofin, A. M. (2008). 'Multinomial goodness-of-fit tests for logistic regression models'. *Stat. Med.* 27 (21), pp. 4238–4253. DOI: 10.1002/sim.3202.

Ferrando-Vivas, P., Doidge, J., Thomas, K., Gould, D. W. *et al.* (2021). 'Prognostic Factors for 30-Day Mortality in Critically Ill Patients With Coronavirus Disease 2019: An Observational Cohort Study'. *Crit. Care Med.* 49 (1), pp. 102–111. DOI: 10.1097/CCM.0000000000004740.

Fine, J. P. and Gray, R. J. (1999). 'A proportional hazards model for the subdistribution of a competing risk'. *J. Am. Stat. Assoc.* 94 (446), pp. 496–509. DOI: 10.1080/01621459.1999.10474144.

Fong, K. J., Summers, C. and Cook, T. M. (2024). 'National Health Service hospital capacity during COVID-19'. *British Medical Journal*. In press.

Ford, N., Meintjes, G., Pozniak, A., Bygrave, H. *et al.* (2015). 'The future role of CD4 cell count for monitoring antiretroviral therapy'. *Lancet Infect. Dis.* 15 (2), pp. 241–247. DOI: 10.1016/S1473-3099(14)70896-5.

Fowler, N. (2014). *AIDS: Don't Die of Prejudice*. Biteback Publishing. 304 pp.

Gabriel, L. and Farrar, J. (2019). 'Pandemic planning and critical care'. *Challenging Concepts in Critical Care*. Oxford University PressOxford, pp. 259–276. DOI: 10.1093/med/9780198814924.003.0017.

Gallo, R. C., Salahuddin, S. Z., Popovic, M., Shearer, G. M. *et al.* (1984). 'Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS'. *Science*. 224 (4648), pp. 500–503. DOI: 10.1126/science.6200936.

Gasperoni, F., Ieva, F., Paganoni, A. M., Jackson, C. H. *et al.* (2020). 'Non-parametric frailty Cox models for hierarchical time-to-event data'. *Biostatistics*. 21 (3), pp. 531–544. DOI: 10.1093/biostatistics/kxy071.

GBD 2019 Diseases and Injuries Collaborators (2020). 'Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019'. *Lancet*. 396 (10258), pp. 1204–1222. DOI: 10.1016/S0140-6736(20)30925-9.

Gelman, A. and Rubin, D. (1992). 'Inference from iterative simulation using multiple sequences'. *Statistical Science*. 7 (4), pp. 457–472. DOI: 10.1214/SS/1177011136.

Ghani, A. C., Donnelly, C. A., Cox, D. R., Griffin, J. T. *et al.* (2005). 'Methods for estimating the case fatality ratio for a novel, emerging infectious disease'. *Am. J. Epidemiol.* 162 (5), pp. 479–486. DOI: 10.1093/aje/kwi230.

Girometti, N., Delpech, V., McCormack, S., Khawam, J. *et al.* (2021). 'The success of HIV combination prevention: The Dean Street model'. *HIV Med.* 22 (10), pp. 892–897. DOI: 10.1111/hiv.13149.

Goldberg, Y., Mandel, M., Bar-On, Y. M., Bodenheimer, O. *et al.* (2021). 'Waning Immunity after the BNT162b2 Vaccine in Israel'. *N. Engl. J. Med.* 385 (24), e85. DOI: 10.1056/NEJMoa2114228.

Goubar, A., Ades, A. E., De Angelis, D., McGarrigle, C. A. *et al.* (2008). 'Estimates of human immunodeficiency virus prevalence and proportion diagnosed based on Bayesian multiparameter synthesis of surveillance data'. *J. R. Stat. Soc. Ser. A Stat. Soc.* 171 (3), pp. 541–580. DOI: 10.1111/j.1467-985x.2007.00537.x.

Goudie, R. J. B., Presanis, A. M., Lunn, D., De Angelis, D. *et al.* (2019). 'Joining and splitting models with Markov melding'. *Bayesian Anal.* 14 (1), pp. 81–109. DOI: 10.1214/18-BA1104.

Goudie, R. J. B., Turner, R. M., De Angelis, D. and Thomas, A. (2020). 'MultiBUGS: A parallel implementation of the BUGS modelling framework for faster Bayesian inference'. *J. Stat. Softw.* 95 (7), pp. 1–20. DOI: 10.18637/jss.v095.i07.

GOV.UK (2020a). 'New and Emerging Respiratory Virus Threats Advisory Group'. URL: https://www.gov.uk/government/groups/new-and-emerging-respiratory-virus-threats-advisory-group (visited on 24/02/2021).

GOV.UK (2020b). 'Scientific Advisory Group for emergencies'. URL: https://www.gov.uk/government/groups/scientific-advisory-group-for-emergencies-sage-coronavirus-covid-19-response (visited on 24/02/2021).

GOV.UK (2020c). 'Scientific pandemic influenza group on modelling (SPI-M)'. URL: https://www.gov.uk/government/groups/scientific-pandemic-influenza-subgroup-on-modelling (visited on 24/02/2021).

GOV.UK (2023). 'STI and HIV surveillance systems: legal basis and confidentiality controls'. URL: https://www.gov.uk/government/publications/sti-and-hiv-surveillance-legal-basis-and-confidentiality-controls/sti-and-hiv-surveillance-systems-legal-basis-and-confidentiality-controls (visited on 18/12/2023).

Gran, J. M., Lie, S. A., Øyeflaten, I., Borgan, Ø. *et al.* (2015). 'Causal inference in multi-state models–sickness absence and work for 1145 participants after work rehabilitation'. *BMC Public Health*. 15 (1), pp. 1–16. DOI: 10.1186/s12889-015-2408-8.

Gray, W. K., Navaratnam, A. V., Day, J., Wendon, J. *et al.* (2021). 'Changes in COVID-19 in-hospital mortality in hospitalised adults in England over the first seven months of the pandemic: An observational study using administrative data'. *Lancet Reg Health Eur*. 5, p. 100104. DOI: 10.1016/j.lanepe.2021.100104.

Greenberg, N., Weston, D., Hall, C., Caulfield, T. *et al.* (2021). 'Mental health of staff working in intensive care during Covid-19'. *Occup. Med. (Lond.)* 71 (2), pp. 62–67. DOI: 10.1093/occmed/kqaa220.

Griffith, G. J., Morris, T. T., Tudball, M. J., Herbert, A. *et al.* (2020). 'Collider bias undermines our understanding of COVID-19 disease risk and severity'. *Nat. Commun.* 11 (1), p. 5749. DOI: 10.1038/s41467-020-19478-2.

Grüger, J., Kay, R. and Schumacher, M. (1991). 'The validity of inferences based on incomplete observations in disease state models'. *Biometrics*. 47 (2), pp. 595–605.

Gulland, A. (2011). 'Test all patients in high prevalence areas for HIV, says NICE'. *BMJ*. 342, p. d1900. DOI: 10.1136/bmj.d1900.

Hall, C. E., Milward, J., Spoiala, C., Bhogal, J. K. *et al.* (2022). 'The mental health of staff working on intensive care units over the COVID-19 winter surge of 2020 in England: a cross sectional survey'. *Br. J. Anaesth.* 128 (6), pp. 971–979. DOI: 10.1016/j.bja.2022.03.016.

Hall, H. I., Song, R., Rhodes, P., Prejean, J. *et al.* (2008). 'Estimation of HIV incidence in the United States'. *JAMA*. 300 (5), pp. 520–529. DOI: 10.1001/jama.300.5.520.

Hall, V., Foulkes, S., Insalata, F., Kirwan, P. *et al.* (2022). 'Protection against SARS-CoV-2 after covid-19 vaccination and previous infection'. *N. Engl. J. Med.* 386 (13), pp. 1207–1220. DOI: 10.1056/NEJMoa2118691.

Hall, V. J., Insalata, F., Foulkes, S., Kirwan, P. *et al.* (2024). 'Effectiveness of BNT162b2 mRNA vaccine third doses and previous infection in protecting against SARS-CoV-2 infections during the Delta and Omicron variant waves; the UK SIREN cohort study September 2021 to February 2022'. *J. Infect.* 88 (1), pp. 30–40. DOI: 10.1016/j.jinf.2023.10.022.

Hall, V. J., Foulkes, S., Charlett, A., Atti, A. *et al.* (2021). 'SARS-CoV-2 infection rates of antibody-positive compared with antibody-negative health-care workers in England: a large, multicentre, prospective cohort study (SIREN)'. *Lancet*. 397 (10283), pp. 1459–1469. DOI: 10.1016/S0140-6736(21)00675-9.

Hall, V. J., Foulkes, S., Saei, A., Andrews, N. *et al.* (2021). 'COVID-19 vaccine coverage in health-care workers in England and effectiveness of BNT162b2 mRNA vaccine against infection (SIREN): a prospective, multicentre, cohort study'. *Lancet*. 397 (10286), pp. 1725–1735. DOI: 10.1016/S0140-6736(21)00790-X.

Hallett, T. B., Zaba, B., Todd, J., Lopman, B. *et al.* (2008). 'Estimating incidence from prevalence in generalised HIV epidemics: methods and validation'. *PLoS Med.* 5 (4), e80. DOI: 10.1371/journal.pmed.0050080.

Hargreaves, J., Davey, C. and Group for lessons from pandemic HIV prevention for the COVID-19 response (2020). 'Three lessons for the COVID-19 response from pandemic HIV'. *Lancet HIV*. 7 (5), e309–e311. DOI: 10.1016/S2352-3018(20)30110-7.

Harrison, E. M., Docherty, A. and Semple, C. (2020). 'COVID-19: time from symptom onset until death in UK hospitalised patients'. Co-CIN. URL: https://assets.publishing.service.gov.uk/media/5f9172cf8fa8f543f5b80e17/S0803_CO-CIN_-_Time_from_symptom_onset_until_death.pdf.

Hay, J. A., Kissler, S. M., Fauver, J. R., Mack, C. *et al.* (2022). 'Quantifying the impact of immune history and variant on SARS-CoV-2 viral kinetics and infection rebound: A retrospective cohort study'. *Elife*. 11. DOI: 10.7554/eLife.81849.

Health Protection Agency (2011). 'Time to test for HIV: expanding HIV testing in healthcare and community services in England'. Health Protection Agency. URL: https://www.bhiva.org/file/gMSwfxmXnFQeb/Time_to_test_final_report__Sept_2011.pdf.

Heath, B., Evans, S., Robertson, D. S., Robotham, J. V. *et al.* (2023). 'Evaluating pooled testing for asymptomatic screening of healthcare workers in hospitals'. *BMC Infect. Dis.* 23 (1), p. 900. DOI: 10.1186/s12879-023-08881-x.

Hernán, M. A., Hernández-Díaz, S. and Robins, J. M. (2004). 'A structural approach to selection bias'. *Epidemiology*. 15 (5), pp. 615–625. DOI: 10.1097/01.ede.0000135174.63482.43.

Hernán, M. A. and Robins, J. M. (2006). 'Estimating causal effects from epidemiological data'. *J. Epidemiol. Community Health*. 60 (7), pp. 578–586. DOI: 10.1136/jech.2004.029496.

Hernán, M. A. and Robins, J. M. (2016). 'Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available'. *Am. J. Epidemiol.* 183 (8), pp. 758–764. DOI: 10.1093/aje/kwv254.

Hernán, M. A. and Robins, J. M. (2023). *Causal Inference: What If*. CRC Press. 312 pp.

Hibbert, M., Wolton, A., Crenna-Jennings, W., Benton, L. *et al.* (2018). 'Experiences of stigma and discrimination in social and healthcare settings among trans people living with HIV in the UK'. *AIDS Care*. 30 (7), pp. 836–843. DOI: 10.1080/09540121.2018.1436687.

HIV Commission (2020). 'The HIV Commission Final Report and Recommendations: How England Will End New Cases of HIV'. HIV Commission. URL: https://www.hivcommission.org.uk/2020/11/30/final-report-and-recommendations-out-now/.

HM Government (2022). 'COVID-19 response: Living with COVID-19'. HM Government. URL: https://www.gov.uk/government/publications/covid-19-response-living-with-covid-19.

Holmberg, M. J. and Andersen, L. W. (2022). 'Collider bias'. *JAMA*. 327 (13), pp. 1282–1283. DOI: 10.1001/jama.2022.1820.

Hosmer, D. W. and Lemeshow, S. (1999). *Applied survival analysis: Regression modeling of time to event data*. Wiley.

Hougaard, P. (1999). 'Multi-state Models: A Review'. *Lifetime Data Anal.* 5 (3), pp. 239–264. DOI: 10.1023/a:1009672031531.

van den Hout, A. (2016). *Multi-State Survival Models for Interval-Censored Data*. CRC Press. 238 pp.

Huiberts, A. J., de Gier B, Hoeve, C. E., de Melker HE *et al.* (2023). 'Effectiveness of bivalent mRNA booster vaccination against SARS-CoV-2 Omicron infection, the Netherlands, September to December 2022'. *Euro Surveill.* 28 (7). DOI: 10.2807/1560-7917.ES.2023.28.7.2300087.

Hulme, W. J., Horne, E. M. F., Parker, E. P. K., Keogh, R. H. *et al.* (2023). 'Comparative effectiveness of BNT162b2 versus mRNA-1273 covid-19 vaccine boosting in England: matched cohort study in OpenSAFELY-TPP'. *BMJ*. 380, e072808. DOI: 10.1136/bmj-2022-072808.

INSIGHT START Study Group, Lundgren, J. D., Babiker, A. G., Gordin, F. *et al.* (2015). 'Initiation of Antiretroviral Therapy in Early Asymptomatic HIV Infection'. *N. Engl. J. Med.* 373 (9), pp. 795–807. DOI: 10.1056/NEJMoa1506816.

Institute for Government (2022). 'Timeline of UK government coronavirus lockdowns and restrictions'. Institute for Government. URL: https://www.instituteforgovernment.org.uk/data-visualisation/timeline-coronavirus-lockdowns.

Intensive Care National Audit and Research Centre (2021). 'ICNARC report on COVID-19 in critical care: England, Wales and Northern Ireland'. Intensive Care National Audit and Research Centre. URL: https://www.icnarc.org/Our-Audit/Audits/Cmp/Reports.

Jackson, C. H. (2011). 'Multi-State Models for Panel Data: The msm Package for R'. *J. Stat. Softw.* 38 (8), pp. 1–29. DOI: 10.18637/jss.v038.i08.

Jackson, C. H. (2016). 'flexsurv: A Platform for Parametric Survival Modeling in R'. *J. Stat. Softw.* 70 (8), pp. 1–33. DOI: 10.18637/jss.v070.i08.

Jackson, C. H. (2021). 'Multi-state modelling with msm: a practical course'. URL: https://chjackson.github.io/msm/msmcourse/ (visited on 16/01/2024).

Jackson, C. H., Tom, B. D., Kirwan, P. D., Mandal, S. *et al.* (2022). 'A comparison of two frameworks for multi-state modelling, applied to outcomes after hospital admissions with COVID-19'. *Stat. Methods Med. Res.* 31 (9), pp. 1656–1674. DOI: 10.1177/09622802221106720.

Jajosky, R. A. and Groseclose, S. L. (2004). 'Evaluation of reporting timeliness of public health surveillance systems for infectious diseases'. *BMC Public Health*. 4 (1), p. 29. DOI: 10.1186/1471-2458-4-29.

Joint Committee on Vaccination and Immunisation (2021). 'JCVI statement regarding a COVID-19 booster vaccine programme for winter 2021 to 2022'. URL: https://www.gov.uk/government/publications/jcvi-statement-september-2021-covid-19-booster-vaccine-programme-for-winter-2021-to-2022/jcvi-statement-regarding-a-covid-19-booster-vaccine-programme-for-winter-2021-to-2022 (visited on 15/01/2024).

Joint United Nations Programme on HIV/AIDS (UNAIDS) (2023a). 'The Path That Ends AIDS - 2023 Global AIDS Update'. Joint United Nations Programme on HIV/AIDS (UNAIDS). URL: https://thepath.unaids.org/.

Joint United Nations Programme on HIV/AIDS (UNAIDS) (2023b). 'UNAIDS data 2023'. Joint United Nations Programme on HIV/AIDS (UNAIDS). URL: https://www.unaids.org/en/resources/documents/2023/2023_unaids_data.

Kalbfleisch, J. D. and Lawless, J. F. (1985). 'The Analysis of Panel Data under a Markov Assumption'. *J. Am. Stat. Assoc.* 80 (392), pp. 863–871. DOI: 10.1080/01621459.1985.10478195.

Kalbfleisch, J. D. and Prentice, R. L. (2011). *The statistical analysis of failure time data*. 2nd ed. Wiley Series in Probability and Statistics. Newy York: Wiley-Interscience. 462 pp.

Karita, E., Price, M., Hunter, E., Chomba, E. *et al.* (2007). 'Investigating the utility of the HIV-1 BED capture enzyme immunoassay using cross-sectional and longitudinal seroconverter specimens from Africa'. *AIDS*. 21 (4), pp. 403–408. DOI: 10.1097/QAD.0b013e32801481b7.

Karon, J. M., Song, R., Brookmeyer, R., Kaplan, E. H. *et al.* (2008). 'Estimating HIV incidence in the United States from HIV/AIDS surveillance data and biomarker HIV test results'. *Stat. Med.* 27 (23), pp. 4617–4633. DOI: 10.1002/sim.3144.

Kassanjee, R., McWalter, T. A., Bärnighausen, T. and Welte, A. (2012). 'A new general biomarker-based incidence estimator'. *Epidemiology*. 23 (5), pp. 721–728. DOI: 10.1097/EDE.0b013e3182576c07.

Kassanjee, R., Pilcher, C. D., Keating, S. M., Facente, S. N. *et al.* (2014). 'Independent assessment of candidate HIV incidence assays on specimens in the CEPHIA repository'. *AIDS*. 28 (16), pp. 2439–2449. DOI: 10.1097/QAD.0000000000000429.

Keiding, N., Klein, J. P. and Horowitz, M. M. (2001). 'Multi-state models and outcome prediction in bone marrow transplantation'. *Stat. Med.* 20 (12), pp. 1871–1885. DOI: 10.1002/sim.810.

Keshani, P., Sarihi, S., Parsaie, N. and Joulaei, H. (2023). 'Dietary pattern association with CD4 cells count in patients living with human immunodeficiency virus: A cross-sectional study'. *J. Public Health Res.* 12 (2), p. 22799036231181200. DOI: 10.1177/22799036231181200.

Kirsebom, F. C. M., Andrews, N., Stowe, J., Toffa, S. *et al.* (2022). 'COVID-19 vaccine effectiveness against the omicron (BA.2) variant in England'. *Lancet Infect. Dis.* 22 (7), pp. 931–933. DOI: 10.1016/S1473-3099(22)00309-7.

Kirwan, P. (2022). *pkirwan/COVID-hospital-outcomes: R code for manuscript*. DOI: 10.5281/zenodo.6856530. URL: https://zenodo.org/record/6856530.

Kirwan, P. D., Charlett, A., Birrell, P., Elgohari, S. *et al.* (2022). 'Trends in COVID-19 hospital outcomes in England before and after vaccine introduction, a cohort study'. *Nat. Commun.* 13 (1), p. 4834. DOI: 10.1038/s41467-022-32458-y.

Kirwan, P. D., Croxford, S., Aghaizu, A., Murphy, G. *et al.* (2022). 'Re-assessing the late HIV diagnosis surveillance definition in the era of increased and frequent testing'. *HIV Med.* 23 (11), pp. 1127–1142. DOI: 10.1111/hiv.13394.

Kirwan, P. D., Elgohari, S., Jackson, C. H., Tom, B. D. M. *et al.* (2021). 'Trends in risks of severe events and lengths of stay for COVID-19 hospitalisations in England over the pre-vaccination era: results from the Public Health England SARI-Watch surveillance scheme'. *arXiv [stat.AP]*. EPRINT: 2103.04867.

Kirwan, P. D., Hall, V. J., Foulkes, S., Otter, A. D. *et al.* (2024). 'Effect of second booster vaccinations and prior infection against SARS-CoV-2 in the UK SIREN healthcare worker cohort'. *Lancet Reg. Health Eur.* 36, p. 100809. DOI: 10.1016/j.lanepe.2023.100809.

Klein, J. P. and Moeschberger, M. L. (2005). *Survival Analysis: Techniques for Censored and Truncated Data*. Springer Science & Business Media. 538 pp.

Knock, E. S., Whittles, L. K., Lees, J. A., Perez-Guzman, P. N. *et al.* (2021). 'Key epidemiological drivers and impact of interventions in the 2020 SARS-CoV-2 epidemic in England'. *Sci. Transl. Med.* DOI: 10.1126/scitranslmed.abg4262.

Kojima, N., Roshani, A. and Klausner, J. D. (2022). 'Duration of COVID-19 PCR positivity for Omicron vs earlier variants'. *J Clin Virol Plus*. 2 (3), p. 100085. DOI: 10.1016/j.jcvp.2022.100085.

Krammer, F. (2020). 'SARS-CoV-2 vaccines in development'. *Nature*. 586 (7830), pp. 516–527. DOI: 10.1038/s41586-020-2798-3.

Lambert, P. C. (2017). 'The estimation and modelling of cause-specific cumulative incidence functions using time-dependent weights'. *Stata J.* 17 (1), pp. 181–207.

Larson, M. G. and Dinse, G. E. (1985). 'A mixture model for the regression analysis of competing risks data'. *J. R. Stat. Soc. Ser. C. Appl. Stat.* 34 (3), p. 201. DOI: 10.2307/2347464.

Le Vu, S., Pillonel, J., Semaille, C., Bernillon, P. *et al.* (2008). 'Principles and uses of HIV incidence estimation from recent infection testing–a review'. *Euro Surveill.* 13 (36), p. 18969. DOI: 10.2807/ese.13.36.18969-en.

Leung, N. H. L. (2021). 'Transmissibility and transmission of respiratory viruses'. *Nat. Rev. Microbiol.* 19 (8), pp. 528–545. DOI: 10.1038/s41579-021-00535-6.

Lewis, J. and White, P. J. (2017). 'Estimating local chlamydia incidence and prevalence using surveillance data'. *Epidemiology*. 28 (4), pp. 492–502. DOI: 10.1097/EDE.0000000000000655.

Lin, D. Y. and Wei, L. J. (1989). 'The Robust Inference for the Cox Proportional Hazards Model'. *J. Am. Stat. Assoc.* 84 (408), pp. 1074–1078. DOI: 10.1080/01621459.1989.10478874.

Link-Gelles, R., Ciesla, A. A., Roper, L. E., Scobie, H. M. *et al.* (2023). 'Early Estimates of Bivalent mRNA Booster Dose Vaccine Effectiveness in Preventing Symptomatic SARS-CoV-2 Infection Attributable to Omicron BA.5- and XBB/XBB.1.5-Related Sublineages Among Immunocompetent Adults - Increasing Community Access to Testing Program, United States, December 2022-January 2023'. *MMWR Morb. Mortal. Wkly. Rep.* 72 (5), pp. 119–124. DOI: 10.15585/mmwr.mm7205e1.

Lipman, M. (2006). 'HUMAN IMMUNODEFICIENCY VIRUS'. *Encyclopedia of Respiratory Medicine*. Ed. by G. J. Laurent and S. D. Shapiro. Oxford: Academic Press, pp. 284–292. DOI: 10.1016/B0-12-370879-6/00179-4.

Lipsitch, M., Cohen, T., Cooper, B., Robins, J. M. *et al.* (2003). 'Transmission dynamics and control of severe acute respiratory syndrome'. *Science*. 300 (5627), pp. 1966–1970. DOI: 10.1126/science.1086616.

Lipsitch, M., Donnelly, C. A., Fraser, C., Blake, I. M. *et al.* (2015). 'Potential Biases in Estimating Absolute and Relative Case-Fatality Risks during Outbreaks'. *PLoS Negl. Trop. Dis.* 9 (7), e0003846. DOI: 10.1371/journal.pntd.0003846.

Liu, L., Iketani, S., Guo, Y., Chan, J. F.-W. *et al.* (2022). 'Striking antibody evasion manifested by the Omicron variant of SARS-CoV-2'. *Nature*. 602 (7898), pp. 676–681. DOI: 10.1038/s41586-021-04388-0.

Lodi, S., Phillips, A., Touloumi, G., Geskus, R. *et al.* (2011). 'Time from human immunodeficiency virus seroconversion to reaching CD4+ cell count thresholds <200, <350, and <500 Cells/mm^3: assessment of need following changes in treatment guidelines'. *Clin. Infect. Dis.* 53 (8), pp. 817–825. DOI: 10.1093/cid/cir494.

Lopez Bernal, J., Andrews, N., Gower, C., Robertson, C. *et al.* (2021). 'Effectiveness of the Pfizer-BioNTech and Oxford-AstraZeneca vaccines on covid-19 related symptoms, hospital admissions, and mortality in older adults in England: test negative case-control study'. *BMJ*. 373, n1088. DOI: 10.1136/bmj.n1088.

Maartens, G., Celum, C. and Lewin, S. R. (2014). 'HIV infection: epidemiology, pathogenesis, treatment, and prevention'. *Lancet*. 384 (9939), pp. 258–271. DOI: 10.1016/S0140-6736(14)60164-1.

Mach, K. J., Salas Reyes, R., Pentz, B., Taylor, J. *et al.* (2021). 'News media coverage of COVID-19 public health and policy information'. *Humanit. Soc. Sci. Commun.* 8 (1), pp. 1–11. DOI: 10.1057/s41599-021-00900-z.

MacLeod, M. K. L., Clambey, E. T., Kappler, J. W. and Marrack, P. (2009). 'CD4 memory T cells: what are they and what can they do?' *Semin. Immunol.* 21 (2), pp. 53–61. DOI: 10.1016/j.smim.2009.02.006.

Maini, M. K., Gilson, R. J., Chavda, N., Gill, S. *et al.* (1996). 'Reference ranges and sources of variability of CD4 counts in HIV-seronegative women and men'. *Genitourin. Med.* 72 (1), pp. 27–31. DOI: 10.1136/sti.72.1.27.

Martin, V., Lester, J., Adamson, L., Shah, A. *et al.* (2022). 'HIV Action Plan Monitoring and Evaluation Framework: Report summarising progress from 2019 to 2021'. UK Health Security Agency, London. URL: https://www.gov.uk/government/publications/hiv-monitoring-and-evaluation-framework/hiv-action-plan-monitoring-and-evaluation-framework.

Martin, V., Okumu-Camerra, K., Bera, S., Shah, A. *et al.* (2023). 'HIV Action Plan monitoring and evaluation framework 2023 report'. UK Health Security Agency, London. URL: https://www.gov.uk/government/publications/hiv-monitoring-and-evaluation-framework/hiv-action-plan-monitoring-and-evaluation-framework-2023-report.

Mathieu, E., Ritchie, H., Rodés-Guirao, L., Appel, C. *et al.* (2020). 'Coronavirus Pandemic (COVID-19)'. URL: https://ourworldindata.org/coronavirus.

Mathur, R., Rentsch, C. T., Morton, C. E., Hulme, W. J. *et al.* (2021). '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'. *Lancet*. 397 (10286), pp. 1711–1724. DOI: 10.1016/S0140-6736(21)00634-6.

Matsena Zingoni, Z., Chirwa, T. F., Todd, J. and Musenge, E. (2021). 'A review of multistate modelling approaches in monitoring disease progression: Bayesian estimation using the Kolmogorov-Chapman forward equations'. *Stat. Methods Med. Res.* 30 (5), pp. 1373–1392. DOI: 10.1177/0962280221997507.

May, M., Wood, R., Myer, L., Taffé, P. *et al.* (2009). 'CD4(+) T cell count decreases by ethnicity among untreated patients with HIV infection in South Africa and Switzerland'. *J. Infect. Dis.* 200 (11), pp. 1729–1735. DOI: 10.1086/648096.

Medicines and Healthcare products Regulatory Agency (2020). 'MHRA guidance on coronavirus (COVID-19)'. URL: https://www.gov.uk/government/collections/mhra-guidance-on-coronavirus-covid-19 (visited on 15/01/2024).

Medicines and Healthcare products Regulatory Agency (2022a). 'Public Assessment Report Comirnaty Original/Omicron BA.1 (15/15 micrograms)/dose dispersion for injection'. PLGB 53632/0010. Medicines and Healthcare products Regulatory Agency.

Medicines and Healthcare products Regulatory Agency (2022b). 'Public Assessment Report Spikevax bivalent Original/Omicron'. PLGB 53720/0004. Medicines and Healthcare products Regulatory Agency.

Meira-Machado, L., de Uña-Alvarez, J., Cadarso-Suárez, C. and Andersen, P. K. (2009). 'Multi-state models for the analysis of time-to-event data'. *Stat. Methods Med. Res.* 18 (2), pp. 195–222. DOI: 10.1177/0962280208092301.

Moncivaiz, A. (2013). 'CD4 count: Normal range, viral load, and what it means for people'. URL: https://www.healthline.com/health/hiv-aids/cd4-viral-count (visited on 02/01/2024).

Mongin, D., Bürgisser, N., Laurie, G., Schimmel, G. *et al.* (2023). 'Effect of SARS-CoV-2 prior infection and mRNA vaccination on contagiousness and susceptibility to infection'. *Nat. Commun.* 14 (1), p. 5452. DOI: 10.1038/s41467-023-41109-9.

Mortimer, J. Y. and Salathiel, J. A. (1995). ''Soundex' codes of surnames provide confidentiality and accuracy in a national HIV database'. *Commun. Dis. Rep. CDR Rev.* 5 (12), R183–6.

Mude, W., Mwenyango, H., Preston, R., O'Mullan, C. *et al.* (2023). 'HIV Testing Disruptions and Service Adaptations During the COVID-19 Pandemic: A Systematic Literature Review'. *AIDS Behav.* DOI: 10.1007/s10461-023-04139-4.

Muema, D. M., Akilimali, N. A., Ndumnego, O. C., Rasehlo, S. S. *et al.* (2020). 'Association between the cytokine storm, immune cell dynamics, and viral replicative capacity in hyperacute HIV infection'. *BMC Med.* 18 (1), p. 81. DOI: 10.1186/s12916-020-01529-6.

Muhammed, F. K., Kirwan, P., Tabassum, T., Lester, J. *et al.* (2024). 'Impact of the first COVID-19 lockdown on HIV diagnosis in England'. *Unpublished*.

Murphy, G. and Parry, J. V. (2008). 'Assays for the detection of recent infections with human immunodeficiency virus type 1'. *Euro Surveill.* 13 (36). DOI: 10.2807/ese.13.36.18966-en.

Murray, J. and Cohen, A. L. (2017). 'Infectious Disease Surveillance'. *International Encyclopedia of Public Health*. Elsevier, pp. 222–229. DOI: 10.1016/b978-0-12-803678-5.00517-8.

Naimi, A. I., Cole, S. R. and Kennedy, E. H. (2017). 'An introduction to g methods'. *Int. J. Epidemiol.* 46 (2), pp. 756–762. DOI: 10.1093/ije/dyw323.

Nakagawa, F., May, M. and Phillips, A. (2013). 'Life expectancy living with HIV: recent estimates and future implications'. *Curr. Opin. Infect. Dis.* 26 (1), pp. 17–25. DOI: 10.1097/QCO.0b013e32835ba6b1.

National Health Service (2020). 'Clinical guide for the management of surge during the Coronavirus pandemic: critical care rapid learning'. URL: https://www.nice.org.uk/Media/Default/About/COVID-19/Specialty-guides/management-of-surge.pdf.

National Institute for Health and Care Excellence (2016a). 'HIV and AIDS testing guidelines'. NICE. URL: https://www.nice.org.uk/guidance/conditions-and-diseases/infections/hiv-and-aids.

National Institute for Health and Care Excellence (2016b). 'HIV testing: increasing uptake among people who may have undiagnosed HIV'. National Institute for Health and Care Excellence.

National Institute for Health and Care Research (2024). 'Research Units'. URL: https://www.nihr.ac.uk/explore-nihr/support/research-units.htm (visited on 28/03/2024).

Navaratnam, A. V., Gray, W. K., Day, J., Wendon, J. *et al.* (2021). 'Patient factors and temporal trends associated with COVID-19 in-hospital mortality in England: an observational study using administrative data'. *Lancet Respir Med*. 9 (4), pp. 397–406. DOI: 10.1016/S2213-2600(20)30579-8.

NHS Digital (2020). 'Hospital Episode Statistics (HES)'. URL: https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/hospital-episode-statistics (visited on 24/02/2021).

NHS Digital (2022). 'National Immunisation and Vaccination System (NIVS)'. URL: https://digital.nhs.uk/services/vaccinations-point-of-care/national-immunisation-and-vaccination-system-nivs (visited on 14/02/2024).

NHS Digital (2023). 'COVID-19 second generation surveillance system (SGSS)'. URL: https://digital.nhs.uk/services/data-access-request-service-dars/dars-products-and-services/data-set-catalogue/covid-19-second-generation-surveillance-system-sgss (visited on 14/02/2024).

NHS Digital (2024). 'Emergency Care Data Set (ECDS)'. URL: https://digital.nhs.uk/data-and-information/data-collections-and-data-sets/data-sets/emergency-care-data-set-ecds (visited on 14/02/2024).

NHS England (2019). 'NHS England » update on the PrEP impact trial'. URL: https://www.england.nhs.uk/2019/01/update-on-the-prep-impact-trial/ (visited on 07/01/2022).

NHS England (2020a). 'Advice on acute sector workforce models during COVID-19'. NHS England. URL: https://www.england.nhs.uk/coronavirus/publication/advice-on-acute-sector-workforce-models-during-covid-19/.

NHS England (2020b). 'Coronavirus: principles for increasing the nursing workforce in response to exceptional increased demand in adult critical care'. NHS England. URL: http://www.cmccn.nhs.uk/files/2115/8521/1349/Specialty_guide__critical_care_workforce_V1_25_March.pdf.

NHS England (2022). 'COVID-19 Hospital Activity'. URL: https://www.england.nhs.uk/statistics/statistical-work-areas/covid-19-hospital-activity/ (visited on 21/06/2022).

Nielsen, J., Mazick, A., Andrews, N., Detsis, M. *et al.* (2013). 'Pooling European all-cause mortality: methodology and findings for the seasons 2008/2009 to 2010/2011'. *Epidemiol. Infect.* 141 (9), pp. 1996–2010. DOI: 10.1017/s0950268812002580.

Nordström, P., Ballin, M. and Nordström, A. (2022). 'Risk of infection, hospitalisation, and death up to 9 months after a second dose of COVID-19 vaccine: a retrospective, total population cohort study in Sweden'. *Lancet*. 399 (10327), pp. 814–823. DOI: 10.1016/S0140-6736(22)00089-7.

Nyberg, T., Ferguson, N. M., Nash, S. G., Webster, H. H. *et al.* (2022). 'Comparative analysis of the risks of hospitalisation and death associated with SARS-CoV-2 omicron (B.1.1.529) and delta (B.1.617.2) variants in England: a cohort study'. *Lancet*. 399 (10332), pp. 1303–1312. DOI: 10.1016/S0140-6736(22)00462-7.

Nyberg, T., Twohig, K. A., Harris, R. J., Seaman, S. R. *et al.* (2021). 'Risk of hospital admission for patients with SARS-CoV-2 variant B.1.1.7: cohort analysis'. *BMJ*. 373, n1412. DOI: 10.1136/bmj.n1412.

Office for Health Improvement (2023). 'Sexual and reproductive health profiles - PHE'. URL: https://fingertips.phe.org.uk/profile/sexualhealth (visited on 07/01/2022).

Ortiz, J. R. and Neuzil, K. M. (2022). 'The value of vaccine programme impact monitoring during the COVID-19 pandemic'. *Lancet*. 399 (10320), pp. 119–121. DOI: 10.1016/S0140-6736(21)02322-9.

Ottoboni, K. N. and Poulos, J. V. (2020). 'Estimating population average treatment effects from experiments with noncompliance'. *Journal of Causal Inference*. 8 (1), pp. 108–130. DOI: 10.1515/jci-2018-0035.

Page, L. A. and Henderson, M. (2008). 'Appraising the evidence: what is measurement bias?' *Evid. Based. Ment. Health*. 11 (2), pp. 36–37. DOI: 10.1136/ebmh.11.2.36.

Pal, S., Yu, H., Loucks, Z. D. and Harris, I. M. (2020). 'Illustration of the flexibility of generalized gamma distribution in modeling right censored survival data: Analysis of two cancer datasets'. *Ann. Data Sci.* 7 (1), pp. 77–90. DOI: 10.1007/s40745-019-00224-5.

Palich, R., Arias-Rodríguez, A., Duracinsky, M., Le Talec, J.-Y. *et al.* (2024). 'High proportion of post-migration HIV acquisition in migrant men who have sex with men receiving HIV care in the Paris region, and associations with social disadvantage and sexual behaviours: results of the ANRS-MIE GANYMEDE study, France, 2021 to 2022'. *Euro Surveill.* 29 (11), p. 2300445. DOI: 10.2807/1560-7917.es.2024.29.11.2300445.

Pantazis, N., Thomadakis, C., Del Amo, J., Alvarez-Del Arco, D. *et al.* (2019). 'Determining the likely place of HIV acquisition for migrants in Europe combining subject-specific information and biomarkers data'. *Stat. Methods Med. Res.* 28 (7), pp. 1979–1997. DOI: 10.1177/0962280217746437.

Parzen, E. (1999). *Stochastic processes*. SIAM.

Pebody, R. (2021a). 'Better access to PrEP in the UK, especially through NHS services'. URL: https://www.aidsmap.com/news/feb-2021/better-access-prep-uk-especially-through-nhs-services (visited on 07/04/2021).

Pebody, R. (2021b). 'What is p24 antigen?' URL: https://www.aidsmap.com/about-hiv/faq/what-p24-antigen (visited on 16/12/2023).

Pei, L., Chen, Y., Zheng, X., Gong, F. *et al.* (2023). 'Comorbidities prolonged viral shedding of patients infected with SARS-CoV-2 omicron variant in Shanghai: A multi-center, retrospective, observational study'. *J. Infect. Public Health*. 16 (2), pp. 182–189. DOI: 10.1016/j.jiph.2022.12.003.

Perez-Guzman, P. N., Knock, E., Imai, N., Rawson, T. *et al.* (2023). 'Epidemiological drivers of transmissibility and severity of SARS-CoV-2 in England'. *Nat. Commun.* 14 (1), p. 4279. DOI: 10.1038/s41467-023-39661-5.

Pezzotti, P., Dorrucci, M., Donisi, A., Cusini, M. *et al.* (2003). 'Survival, progression to AIDS and immunosuppression in HIV-positive individuals before and after the introduction of the highly active antiretroviral therapy (HAART)'. *Epidemiol. Prev.* 27 (6), pp. 348–355.

Phillips, A. N., Cambiano, V., Miners, A., Lampe, F. C. *et al.* (2015). 'Potential impact on HIV incidence of higher HIV testing rates and earlier antiretroviral therapy initiation in MSM'. *AIDS*. 29 (14), pp. 1855–1862. DOI: 10.1097/QAD.0000000000000767.

Plumb, I. D., Briggs Hagen, M., Wiegand, R., Dumyati, G. *et al.* (2023). 'Effectiveness of a bivalent mRNA vaccine dose against symptomatic SARS-CoV-2 infection among U.S. Healthcare personnel, September 2022-May 2023'. *Vaccine*. DOI: 10.1016/j.vaccine.2023.10.072.

Plummer, M. *et al.* (2003). 'JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling'. *Proceedings of the 3rd international workshop on distributed statistical computing*. Vol. 124, pp. 1–10. URL: http://www.ci.tuwien.ac.at/Conferences/DSC-2003/Drafts/Plummer.pdf.

Polack, F. P., Thomas, S. J., Kitchin, N., Absalon, J. *et al.* (2020). 'Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine'. *N. Engl. J. Med.* 383 (27), pp. 2603–2615. DOI: 10.1056/NEJMoa2034577.

Prentice, R. L. (1974). 'A log gamma model and its maximum likelihood estimation'. *Biometrika*. 61 (3), p. 539. DOI: 10.2307/2334737.

Prentice, R. L., Kalbfleisch, J. D., Peterson Jr, A. V., Flournoy, N. *et al.* (1978). 'The analysis of failure times in the presence of competing risks'. *Biometrics*. 34 (4), pp. 541–554. DOI: 10.2307/2530374.

PrEP Impact Trial (2020). 'PrEP Impact Trial Protocol'. PrEP Impact Trial. URL: https://www.prepimpacttrial.org.uk/protocol.

Presanis, A. M., De Angelis, D., Goubar, A., Gill, O. N. *et al.* (2011). 'Bayesian evidence synthesis for a transmission dynamic model for HIV among men who have sex with men'. *Biostatistics*. 12 (4), pp. 666–681. DOI: 10.1093/biostatistics/kxr006.

Presanis, A. M., Harris, R. J., Kirwan, P. D., Miltz, A. *et al.* (2021). 'Trends in undiagnosed HIV prevalence in England and implications for eliminating HIV transmission by 2030: an evidence synthesis model'. *Lancet Public Health*. 6 (10), e739–e751. DOI: 10.1016/S2468-2667(21)00142-0.

Public Health England (2013a). 'HIV & AIDS New Diagnoses & Deaths (HANDD)'. Public Health England. URL: https://webarchive.nationalarchives.gov.uk/ukgwa/20140714084352/http://www.hpa.org.uk/webc/HPAwebFile/HPAweb_C/1317139917763.

Public Health England (2013b). 'The CD4 Surveillance Scheme'. Public Health England. URL: https://webarchive.nationalarchives.gov.uk/ukgwa/20140714084352/http://www.hpa.org.uk/webc/HPAwebFile/HPAweb_C/1317139917578.

Public Health England (2020a). 'COVID-19: review of disparities in risks and outcomes'. URL: https://www.gov.uk/government/publications/covid-19-review-of-disparities-in-risks-and-outcomes (visited on 28/01/2022).

Public Health England (2020b). 'Cycle threshold (Ct) in SARS-CoV-2 RT-PCR'. URL: https://www.gov.uk/government/publications/cycle-threshold-ct-in-sars-cov-2-rt-pcr (visited on 06/05/2021).

Public Health England (2020c). 'The impact of the COVID-19 pandemic on prevention, testing, diagnosis and care for sexually transmitted infections, HIV and viral hepatitis in England'. Public Health England. URL: https://assets.publishing.service.gov.uk/media/5fd39d6b8fa8f54d5c52de43/Impact_of_COVID-19_Report_2020.pdf.

Public Health England (2021a). 'Direct and Indirect Impact of the Vaccination Programme on COVID-19 Infections and Mortality'. Public Health England. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/997495/Impact_of_COVID-19_vaccine_on_infection_and_mortality.pdf.

Public Health England (2021b). 'Weekly national Influenza and COVID-19 surveillance report'. week 6 report (up to week 5 data). Public Health England. URL: https://www.gov.uk/government/statistics/national-flu-and-covid-19-surveillance-reports.

Punyacharoensin, N., Edmunds, W. J., De Angelis, D., Delpech, V. *et al.* (2016). 'Effect of pre-exposure prophylaxis and combination HIV prevention for men who have sex with men in the UK: a mathematical modelling study'. *Lancet HIV*. 3 (2), e94–e104. DOI: 10.1016/S2352-3018(15)00056-9.

Putter, H., Fiocco, M. and Geskus, R. B. (2007). 'Tutorial in biostatistics: competing risks and multi-state models'. *Stat. Med.* 26 (11), pp. 2389–2430. DOI: 10.1002/sim.2712.

R Core Team (2020). *R: A Language and Environment for Statistical Computing*. URL: https://www.R-project.org/.

Le-Rademacher, J. G., Peterson, R. A., Therneau, T. M., Sanford, B. L. *et al.* (2018). 'Application of multi-state models in cancer clinical trials'. *Clin. Trials*. 15 (5), pp. 489–498. DOI: 10.1177/1740774518789098.

Recovery Collaborative Group (2020). 'Dexamethasone in hospitalized patients with Covid-19—preliminary report'. *New England Journal of Medicine*. DOI: 10.1056/NEJMoa2021436.

Regine, V., Dorrucci, M., Pezzotti, P., Mammone, A. *et al.* (2018). 'People living with undiagnosed HIV infection and a low CD4 count: estimates from surveillance data, Italy, 2012 to 2014'. *Euro Surveill.* 23 (15). DOI: 10.2807/1560-7917.ES.2018.23.15.17-00240.

Rice, B. D., Yin, Z., Brown, A. E., Croxford, S. *et al.* (2017). 'Monitoring of the HIV Epidemic Using Routinely Collected Data: The Case of the United Kingdom'. *AIDS Behav.* 21 (Suppl 1), pp. 83–90. DOI: 10.1007/s10461-016-1604-6.

Rivet Amico, K. and Bekker, L.-G. (2019). 'Global PrEP roll-out: recommendations for programmatic success'. *Lancet HIV*. 6 (2), e137–e140. DOI: 10.1016/S2352-3018(19)30002-5.

Rosenberg, P. S. and Gail, M. H. (1991). 'Backcalculation of flexible linear models of the human immunodeficiency virus infection curve'. *J. R. Stat. Soc. Ser. C. Appl. Stat.* 40 (2), pp. 269–282. DOI: 10.2307/2347592.

Rothenberg, R., Woelfel, M., Stoneburner, R., Milberg, J. *et al.* (1987). 'Survival with the Acquired Immunodeficiency Syndrome'. *N. Engl. J. Med.* 317 (21), pp. 1297–1302. DOI: 10.1056/NEJM198711193172101.

Roxby, B. P., Butcher, B. and England, R. (). *Pressure on hospitals 'at a really dangerous point'*. URL: https://www.bbc.co.uk/news/health-55362681 (visited on 18/02/2021).

Rozhnova, G., Heijne, J., Bezemer, D., van Sighem, A. *et al.* (2018). 'Elimination prospects of the Dutch HIV epidemic among men who have sex with men in the era of preexposure prophylaxis'. *AIDS*. 32 (17), pp. 2615–2623. DOI: 10.1097/QAD.0000000000002050.

Ryalen, P. C., Stensrud, M. J. and Røysland, K. (2019). 'The additive hazard estimator is consistent for continuous-time marginal structural models'. *Lifetime Data Anal.* 25 (4), pp. 611–638. DOI: 10.1007/s10985-019-09468-y.

Ryalen, P. C., Stensrud, M. J., Fosså, S. and Røysland, K. (2020). 'Causal inference in continuous time: an example on prostate cancer therapy'. *Biostatistics*. 21 (1), pp. 172–185. DOI: 10.1093/biostatistics/kxy036.

Sackett, D. L. (1979). 'BIAS IN ANALYTIC RESEARCH'. *The Case-Control Study Consensus and Controversy*. Elsevier, pp. 51–63. DOI: 10.1016/b978-0-08-024907-0.50013-4.

Sasse, A., Florence, E., Pharris, A., De Wit, S. *et al.* (2016). 'Late presentation to HIV testing is overestimated when based on the consensus definition'. *HIV Med.* 17 (3), pp. 231–234. DOI: 10.1111/hiv.12292.

Schneider, C. R., Freeman, A. L. J., Spiegelhalter, D. and van der Linden, S. (2022). 'The effects of communicating scientific uncertainty on trust and decision making in a public health context'. *Judgm. Decis. Mak.* 17 (4), pp. 849–882. DOI: 10.1017/s1930297500008962.

Schrappe, M. and Lauterbach, K. (1998). 'Systematic review on the cost-effectiveness of public health interventions for HIV prevention in industrialized countries'. *AIDS*. 12 Suppl A, S231–8.

Scientific Advisory Group for Emergencies (2021). 'NERVTAG: Update note on B.1.1.7 severity, 11 February 2021'. URL: https://www.gov.uk/government/publications/nervtag-update-note-on-b117-severity-11-february-2021 (visited on 24/02/2021).

Scientific Pandemic Influenza Group on Modelling, Operational sub-group (2021). 'SPI-M-O Medium-Term Projections'. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/966763/S1120_SPI-M-O_MediumTermProjections.pdf.

Seaman, S. R. and Keogh, R. H. (2023). 'Simulating data from marginal structural models for a survival time outcome'. *arXiv [stat.ME]*. EPRINT: 2309.05025.

Seaman, S. R., Nyberg, T., Overton, C. E., Pascall, D. J. *et al.* (2022). 'Adjusting for time of infection or positive test when estimating the risk of a post-infection outcome in an epidemic'. *Stat. Methods Med. Res.* 31 (10), pp. 1942–1958. DOI: 10.1177/09622802221107105.

Seaman, S. R., Presanis, A. and Jackson, C. (2022). 'Estimating a time-to-event distribution from right-truncated data in an epidemic: A review of methods'. *Stat. Methods Med. Res.* 31 (9), pp. 1641–1655. DOI: 10.1177/09622802211023955.

SEDIA Biosciences Corporation (2021). 'HIV Incidence Products'. URL: http://www.sediabio.com/products-technologies/hiv-incidence/home (visited on 24/02/2021).

Sedia Biosciences Corporation (2013). 'Sedia HIV-1 LAg-Avidity EIA: single well avidity enzyme immunoassay for detection of recent HIV-1 infection using liquid serum or plasma, Cat. No. 1002'. (Visited on 24/02/2021).

Serfling, R. E. (1963). 'Methods for current statistical analysis of excess pneumonia-influenza deaths'. *Public Health Rep.* 78 (6), pp. 494–506. DOI: 10.2307/4591848.

van Seventer, J. M. and Hochberg, N. S. (2017). 'Principles of infectious diseases: Transmission, diagnosis, prevention, and control'. *International Encyclopedia of Public Health*. Elsevier, pp. 22–39. DOI: 10.1016/b978-0-12-803678-5.00516-6.

Sharma, O., Sultan, A. A., Ding, H. and Triggle, C. R. (2020). 'A Review of the Progress and Challenges of Developing a Vaccine for COVID-19'. *Front. Immunol.* 11, p. 585354. DOI: 10.3389/fimmu.2020.585354.

Shi, X., Li, K. Q. and Mukherjee, B. (2023). 'Current Challenges With the Use of Test-Negative Designs for Modeling COVID-19 Vaccination and Outcomes'. *Am. J. Epidemiol.* 192 (3), pp. 328–333. DOI: 10.1093/aje/kwac203.

Shrestha, N. K., Burke, P. C., Nowacki, A. S., Simon, J. F. *et al.* (2023). 'Effectiveness of the Coronavirus disease 2019 bivalent vaccine'. *Open Forum Infect. Dis.* 10 (6), ofad209. DOI: 10.1093/ofid/ofad209.

van Sighem, A., Nakagawa, F., De Angelis, D., Quinten, C. *et al.* (2015). 'Estimating HIV Incidence, Time to Diagnosis, and the Undiagnosed HIV Epidemic Using Routine Surveillance Data'. *Epidemiology*. 26 (5), pp. 653–660. DOI: 10.1097/EDE.0000000000000324.

van Sighem, A., Pharris, A., Quinten, C., Noori, T. *et al.* (2017). 'Reduction in undiagnosed HIV infection in the European Union/European Economic Area, 2012 to 2016'. *Euro Surveill.* 22 (48). DOI: 10.2807/1560-7917.ES.2017.22.48.17-00771.

Simões, D., Stengaard, A. R., Combs, L., Raben, D. *et al.* (2020). 'Impact of the COVID-19 pandemic on testing services for HIV, viral hepatitis and sexually transmitted infections in the WHO European Region, March to August 2020'. *Euro Surveill.* 25 (47). DOI: 10.2807/1560-7917.ES.2020.25.47.2001943.

Šmíd, M., Berec, L., Přibylová, L., Májek, O. *et al.* (2022). 'Protection by Vaccines and Previous Infection Against the Omicron Variant of Severe Acute Respiratory Syndrome Coronavirus 2'. *J. Infect. Dis.* 226 (8), pp. 1385–1390. DOI: 10.1093/infdis/jiac161.

Smith, L. E., Potts, H. W. W., Amlôt, R., Fear, N. T. *et al.* (2022). 'Patterns of social mixing in England changed in line with restrictions during the COVID-19 pandemic (September 2020 to April 2022)'. *Sci. Rep.* 12 (1), p. 10436. DOI: 10.1038/s41598-022-14431-3.

South China Morning Post (2020). 'China's first confirmed Covid-19 case traced back to November 17'. URL: https://www.scmp.com/news/china/society/article/3074991/coronavirus-chinas-first-confirmed-covid-19-case-traced-back (visited on 24/02/2021).

Spencer, A. and Patel, S. (2019). 'Applying the Data Protection Act 2018 and General Data Protection Regulation principles in healthcare settings'. *Nurs. Manag.* DOI: 10.7748/nm.2019.e1806.

Steele, F., Goldstein, H. and Browne, W. (2004). 'A general multilevel multistate competing risks model for event history data, with an application to a study of contraceptive use dynamics'. *Stat. Modelling*. 4 (2), pp. 145–159. DOI: 10.1191/1471082x04st069oa.

Stensrud, M. J. and Hernán, M. A. (2020). 'Why Test for Proportional Hazards?' *JAMA*. 323 (14), pp. 1401–1402. DOI: 10.1001/jama.2020.1267.

Steyerberg, E. W., Wiegers, E., Sewalt, C., Buki, A. *et al.* (2019). 'Case-mix, care pathways, and outcomes in patients with traumatic brain injury in CENTER-TBI: a European prospective, multicentre, longitudinal, cohort study'. *Lancet Neurol.* 18 (10), pp. 923–934. DOI: 10.1016/S1474-4422(19)30232-7.

Stoehr, J. R., Hamidian Jahromi, A. and Thomason, C. (2021). 'Ethical Considerations for Unblinding and Vaccinating COVID-19 Vaccine Trial Placebo Group Participants'. *Front Public Health*. 9, p. 702960. DOI: 10.3389/fpubh.2021.702960.

Stover, J., Glaubius, R., Mofenson, L., Dugdale, C. M. *et al.* (2019). 'Updates to the Spectrum/AIM model for estimating key HIV indicators at national and subnational levels'. *AIDS*. 33 Suppl 3, S227–S234. DOI: 10.1097/QAD.0000000000002357.

Suligoi, B., Galli, C., Massi, M., Di Sora, F. *et al.* (2002). 'Precision and accuracy of a procedure for detecting recent human immunodeficiency virus infections by calculating the antibody avidity index by an automated immunoassay-based method'. *J. Clin. Microbiol.* 40 (11), pp. 4015–4020. DOI: 10.1128/jcm.40.11.4015-4020.2002.

Suligoi, B., Rodella, A., Raimondo, M., Regine, V. *et al.* (2011). 'Avidity Index for anti-HIV antibodies: comparison between third- and fourth-generation automated immunoassays'. *J. Clin. Microbiol.* 49 (7), pp. 2610–2613. DOI: 10.1128/JCM.02115-10.

Sullivan, S. G., Tchetgen Tchetgen, E. J. and Cowling, B. J. (2016). 'Theoretical Basis of the Test-Negative Study Design for Assessment of Influenza Vaccine Effectiveness'. *Am. J. Epidemiol.* 184 (5), pp. 345–353. DOI: 10.1093/aje/kww064.

Sun, X., Nishiura, H. and Xiao, Y. (2020). 'Modeling methods for estimating HIV incidence: a mathematical review'. *Theor. Biol. Med. Model.* 17 (1), p. 1. DOI: 10.1186/s12976-019-0118-0.

Sweeting, M. J., De Angelis, D. and Aalen, O. O. (2005). 'Bayesian back-calculation using a multi-state model with application to HIV'. *Stat. Med.* 24 (24), pp. 3991–4007. DOI: 10.1002/sim.2432.

Sweeting, M. J., De Angelis, D., Parry, J. and Suligoi, B. (2010). 'Estimating the distribution of the window period for recent HIV infections: a comparison of statistical methods'. *Stat. Med.* 29 (30), pp. 3194–3202. DOI: 10.1002/sim.3941.

Tchetgen Tchetgen, E. J. and VanderWeele, T. J. (2012). 'On causal inference in the presence of interference'. *Stat. Methods Med. Res.* 21 (1), pp. 55–75. DOI: 10.1177/0962280210386779.

Terrence Higgins Trust (2023). 'HIV statistics'. URL: https://www.tht.org.uk/hiv-and-sexual-health/about-hiv/hiv-statistics (visited on 05/03/2024).

The Lancet (2021). '40 years of HIV/AIDS: a painful anniversary'. *Lancet*. 397 (10290), p. 2125. DOI: 10.1016/S0140-6736(21)01213-7.

Therneau, T. (1999). *survival: Survival package for R*. URL: https://github.com/therneau/survival.

Thygesen, J. H., Tomlinson, C., Hollings, S., Mizani, M. *et al.* (2021). 'Understanding COVID-19 trajectories from a nationwide linked electronic health record cohort of 56 million people: phenotypes, severity, waves & vaccination'. *bioRxiv*. DOI: 10.1101/2021.11.08.21265312.

Titman, A. C. (2009). 'Computation of the asymptotic null distribution of goodness-of-fit tests for multi-state models'. *Lifetime Data Anal.* 15 (4), pp. 519–533. DOI: 10.1007/s10985-009-9133-5.

Titman, A. C. and Putter, H. (2022). 'General tests of the Markov property in multi-state models'. *Biostatistics*. 23 (2), pp. 380–396. DOI: 10.1093/biostatistics/kxaa030.

Touloumi, G., Pantazis, N., Pillay, D., Paraskevis, D. *et al.* (2013). 'Impact of HIV-1 subtype on CD4 count at HIV seroconversion, rate of decline, and viral load set point in European seroconverter cohorts'. *Clin. Infect. Dis.* 56 (6), pp. 888–897. DOI: 10.1093/cid/cis1000.

Tran, V.-T., Porcher, R., Pane, I. and Ravaud, P. (2022). 'Course of post COVID-19 disease symptoms over time in the ComPaRe long COVID prospective e-cohort'. *Nat. Commun.* 13 (1), p. 1812. DOI: 10.1038/s41467-022-29513-z.

Trostian, B., McCloughen, A., Lago, L., McAlister, B. *et al.* (2022). 'Challenges and adaptions in longitudinal data linkage to track patient health service use and care pathways after Emergency Department presentation: an exemplar'. *Research Square*. DOI: 10.21203/rs.3.rs-2173257/v1.

Turner, R. M., Spiegelhalter, D. J., Smith, G. C. S. and Thompson, S. G. (2009). 'Bias modelling in evidence synthesis'. *J. R. Stat. Soc. Ser. A Stat. Soc.* 172 (1), pp. 21–47. DOI: 10.1111/j.1467-985X.2008.00547.x.

UK Government (2021). 'Official UK Coronavirus Dashboard'. URL: https://coronavirus.data.gov.uk/ (visited on 24/02/2021).

UK Health Security Agency (2020). 'National COVID-19 surveillance reports'. URL: https://www.gov.uk/government/publications/national-covid-19-surveillance-reports (visited on 22/02/2024).

UK Health Security Agency (2021). 'Monitoring reports of the effectiveness of COVID-19 vaccination'. URL: https://www.gov.uk/guidance/monitoring-reports-of-the-effectiveness-of-covid-19-vaccination (visited on 23/01/2024).

UK Health Security Agency (2022a). 'Antibody testing for SARS-CoV-2: key information'. URL: https://www.gov.uk/government/publications/antibody-testing-for-sars-cov-2-key-information/antibody-testing-for-sars-cov-2-information-for-general-practitioners (visited on 16/01/2024).

UK Health Security Agency (2022b). 'COVID-19: epidemiology, virology and clinical features'. URL: https://www.gov.uk/government/publications/wuhan-novel-coronavirus-background-information/wuhan-novel-coronavirus-epidemiology-virology-and-clinical-features (visited on 25/02/2024).

UK Health Security Agency (2022c). 'Guidance for people with symptoms of a respiratory infection including COVID-19, or a positive test result for COVID-19'. UK Health Security Agency. URL: https://www.gov.uk/guidance/people-with-symptoms-of-a-respiratory-infection-including-covid-19.

UK Health Security Agency (2023a). 'COVID-19 Omicron variant: infectious period and transmission from people with asymptomatic compared with symptomatic infection: a rapid review'. UK Health Security Agency.

UK Health Security Agency (2023b). 'SIREN study expands surveillance of respiratory pathogens ahead of winter'. URL: https://www.gov.uk/government/news/siren-study-expands-surveillance-of-respiratory-pathogens-ahead-of-winter (visited on 16/01/2024).

UK Health Security Agency (2023c). 'UKHSA Strategic Plan: 2023 to 2026'. UK Health Security Agency. URL: https://www.gov.uk/government/publications/ukhsa-strategic-plan-2023-to-2026.

VanderWeele, T. J. and Hernán, M. A. (2013). 'Causal Inference Under Multiple Versions of Treatment'. *J Causal Inference*. 1 (1), pp. 1–20. DOI: 10.1515/jci-2012-0002.

Vehtari, A., Simpson, D., Gelman, A., Yao, Y. *et al.* (2024). 'Pareto smoothed importance sampling'. *J. Mach. Learn. Res.* 25, pp. 1–57. EPRINT: 1507.02646.

Vermunt, L., Sikkes, S. A. M., van den Hout, A., Handels, R. *et al.* (2019). 'Duration of preclinical, prodromal, and dementia stages of Alzheimer's disease in relation to age, sex, and APOE genotype'. *Alzheimers. Dement.* 15 (7), pp. 888–898. DOI: 10.1016/j.jalz.2019.04.001.

Vlahov, D. and Junge, B. (1998). 'The role of needle exchange programs in HIV prevention'. *Public Health Rep.* 113 Suppl 1 (Suppl 1), pp. 75–80.

Voysey, M., Clemens, S. A. C., Madhi, S. A., Weckx, L. Y. *et al.* (2021). 'Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK'. *Lancet*. 397 (10269), pp. 99–111. DOI: 10.1016/S0140-6736(20)32661-1.

Vynnycky, E. and White, R. (2010). *An Introduction to Infectious Disease Modelling*. OUP Oxford. 400 pp.

Wallace, S., Hall, V., Charlett, A., Kirwan, P. D. *et al.* (2022). 'Impact of prior SARS-CoV-2 infection and COVID-19 vaccination on the subsequent incidence of COVID-19: a multicentre prospective cohort study among UK healthcare workers - the SIREN (Sarscov2 Immunity & REinfection EvaluatioN) study protocol'. *BMJ Open*. 12 (6), e054336. DOI: 10.1136/bmjopen-2021-054336.

Wasserman, L. A. (2004). *All of statistics*. New York, NY: Springer. 458 pp.

Webb, B. J., Peltan, I. D., Jensen, P., Hoda, D. *et al.* (2020). 'Clinical criteria for COVID-19-associated hyperinflammatory syndrome: a cohort study'. *Lancet Rheumatol*. 2 (12), e754–e763. DOI: 10.1016/S2665-9913(20)30343-X.

Wei, J., Matthews, P. C., Stoesser, N., Newton, J. N. *et al.* (2023). 'Protection against SARS-CoV-2 Omicron BA.4/5 variant following booster vaccination or breakthrough infection in the UK'. *Nat. Commun.* 14 (1), p. 2799. DOI: 10.1038/s41467-023-38275-1.

Welton, N. J. and Ades, A. E. (2005). 'Estimation of markov chain transition probabilities and rates from fully and partially observed data: uncertainty propagation, evidence synthesis, and model calibration'. *Med. Decis. Making*. 25 (6), pp. 633–645. DOI: 10.1177/0272989X05282637.

Wenlock, R. D., Shillingford, C., Mear, J., Churchill, D. *et al.* (2022). 'The impact of COVID-19 on HIV testing in the UK's first Fast-Track HIV city'. *HIV Med.* 23 (7), pp. 790–796. DOI: 10.1111/hiv.13235.

Wenzl, T. (2020). 'Smoking and COVID-19 - A review of studies suggesting a protective effect of smoking against COVID-19'. DOI: 10.2760/564217.

Westreich, D. (2012). 'Berkson's bias, selection bias, and missing data'. *Epidemiology*. 23 (1), pp. 159–164. DOI: 10.1097/EDE.0b013e31823b6296.

Whitlock, G., Scarfield, P. and Dean Street Collaborative Group (2020). 'HIV diagnoses continue to fall at 56 Dean Street'. *EClinicalMedicine*. 19, p. 100263. DOI: 10.1016/j.eclinm.2020.100263.

Wilcox, M. E., Rowan, K. M., Harrison, D. A. and Doidge, J. C. (2022). 'Does Unprecedented ICU Capacity Strain, As Experienced During the COVID-19 Pandemic, Impact Patient Outcome?' *Crit. Care Med.* 50 (6), e548–e556. DOI: 10.1097/CCM.0000000000005464.

Willett, B. J., Grove, J., MacLean, O. A., Wilkie, C. *et al.* (2022). 'SARS-CoV-2 Omicron is an immune escape variant with an altered cell entry pathway'. *Nat Microbiol*. 7 (8), pp. 1161–1179. DOI: 10.1038/s41564-022-01143-7.

Williamson, E. J., Walker, A. J., Bhaskaran, K., Bacon, S. *et al.* (2020). 'Factors associated with COVID-19-related death using OpenSAFELY'. *Nature*. 584 (7821), pp. 430–436. DOI: 10.1038/s41586-020-2521-4.

Winter, J. R., Delpech, V., Kirwan, P., Stagg, H. R. *et al.* (2016). 'Linkage of UK HIV and tuberculosis data using probabilistic and deterministic methods'. *Conference on Retroviruses and Opportunistic Infections. Boston*. Conference on Retroviruses and Opportunistic Infections (Boston). Boston.

Wong, J. Y., Kelly, H., Ip, D. K. M., Wu, J. T. *et al.* (2013). 'Case fatality risk of influenza A (H1N1pdm09): a systematic review: A systematic review'. *Epidemiology*. 24 (6), pp. 830–841. DOI: 10.1097/EDE.0b013e3182a67448.

Wood, S. N. (2016). 'Just Another Gibbs Additive Modeler: Interfacing JAGS and mgcv'. *J. Stat. Softw.* 75, pp. 1–15. DOI: 10.18637/jss.v075.i07.

Woodward, M. (2013). *Epidemiology: Study design and data analysis, third edition*. 3rd ed. Chapman & Hall/CRC Texts in Statistical Science. Philadelphia, PA: Chapman & Hall/CRC. 854 pp. DOI: 10.1201/b16343.

World Health Organisation (2010). 'Antiretroviral therapy for HIV infection in adults and adolescents'. World Health Organisation. URL: https://www.who.int/hiv/pub/arv/adult2010/en/.

World Health Organisation (2020). 'WHO Director-General's opening remarks at the media briefing on COVID-19'. URL: https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020 (visited on 24/02/2021).

World Health Organisation (2023). 'WHO Director-General's opening remarks at the media briefing – 5 May 2023'. URL: https://www.who.int/news-room/speeches/item/who-director-general-s-opening-remarks-at-the-media-briefing---5-may-2023 (visited on 06/07/2023).

World Health Organization (2020a). 'Estimating mortality from COVID-19 (Scientific Brief)'. World Health Organization. URL: https://www.who.int/news-room/commentaries/detail/estimating-mortality-from-covid-19.

World Health Organization (2020b). 'WHO Coronavirus disease (COVID-19) dashboard'. URL: https://covid19.who.int/ (visited on 24/02/2021).

World Health Organization (2023). 'The role of HIV viral suppression in improving individual health and reducing transmission'. World Health Organization. URL: https://apps.who.int/iris/bitstream/handle/10665/360860/9789240055179-eng.pdf.

Wu, P., Presanis, A. M., Bond, H. S., Lau, E. H. Y. *et al.* (2017). 'A joint analysis of influenza-associated hospitalizations and mortality in Hong Kong, 1998-2013'. *Sci. Rep.* 7 (1), p. 929. DOI: 10.1038/s41598-017-01021-x.

Xie, J., Feng, S., Li, X., Gea-Mallorquí, E. *et al.* (2022). 'Comparative effectiveness of the BNT162b2 and ChAdOx1 vaccines against Covid-19 in people over 50'. *Nat. Commun.* 13 (1), p. 1519. DOI: 10.1038/s41467-022-29159-x.

Yin, Z., Brown, A. E., Rice, B. D., Marrone, G. *et al.* (2021). 'Post-migration acquisition of HIV: Estimates from four European countries, 2007 to 2016'. *Euro Surveill.* 26 (33). DOI: 10.2807/1560-7917.ES.2021.26.33.2000161.

Young, G. A. and Smith, R. L. (2005). *Essentials of Statistical Inference*. Cambridge University Press. DOI: 10.1017/cbo9780511755392.

Appendices

A. PhD publications, pre-prints and conference presentations

First-author peer-reviewed publications

Kirwan, P. D., Hibbert, M., Kall, M., Nambiar, K. *et al.* (2021). 'HIV prevalence and HIV clinical outcomes of transgender and gender-diverse people in England'. *HIV Med.* 22 (2), pp. 131–139. DOI: 10.1111/hiv.12987.

Kirwan, P. D., Charlett, A., Birrell, P., Elgohari, S. *et al.* (2022). 'Trends in COVID-19 hospital outcomes in England before and after vaccine introduction, a cohort study'. *Nat. Commun.* 13 (1), p. 4834. DOI: 10.1038/s41467-022-32458-y.

Kirwan, P. D., Croxford, S., Aghaizu, A., Murphy, G. *et al.* (2022). 'Re-assessing the late HIV diagnosis surveillance definition in the era of increased and frequent testing'. *HIV Med.* 23 (11), pp. 1127–1142. DOI: 10.1111/hiv.13394.

Kirwan, P. D., Hall, V. J., Foulkes, S., Otter, A. D. *et al.* (2024). 'Effect of second booster vaccinations and prior infection against SARS-CoV-2 in the UK SIREN healthcare worker cohort'. *Lancet Reg. Health Eur.* 36, p. 100809. DOI: 10.1016/j.lanepe.2023.100809.

Contributing author peer-reviewed publications

Allen, H., Kirwan, P., Brown, A. E., Mohammed, H. *et al.* (2021). 'Does being on HIV antiretroviral therapy increase the risk of syphilis? An analysis of a large national cohort of MSM living with HIV in England 2009-2016'. *Sex. Transm. Infect.* 97 (3), pp. 221–225. DOI: 10.1136/sextrans-2020-054603.

Brizzi, F., Birrell, P. J., Kirwan, P., Ogaz, D. *et al.* (2021). 'Tracking elimination of HIV transmission in men who have sex with men in England: a modelling study'. *Lancet HIV*. 8 (7), e440–e448. DOI: 10.1016/S2352-3018(21)00044-8.

Brizzi, F., Birrell, P. J., Plummer, M. T., Kirwan, P. *et al.* (2019). 'Extending Bayesian back-calculation to estimate age and time specific HIV incidence'. *Lifetime Data Anal.* 25 (4), pp. 757–780. DOI: 10.1007/s10985-019-09465-1.

Brown, A. E., Croxford, S. E., Nash, S., Khawam, J. *et al.* (2022). 'COVID-19 mortality among people with diagnosed HIV compared to those without during the first wave of the COVID-19 pandemic in England'. *HIV Med.* 23 (1), pp. 90–102. DOI: 10.1111/hiv.13167.

Docherty, A. B., Mulholland, R. H., Lone, N. I., Cheyne, C. P. *et al.* (2021). 'Changes in in-hospital mortality in the first wave of COVID-19: a multicentre prospective observational cohort study using the WHO Clinical Characterisation Protocol UK'. *Lancet Respir. Med.* 9 (7), pp. 773–785. DOI: 10.1016/S2213-2600(21)00175-2.

Hall, V., Foulkes, S., Insalata, F., Kirwan, P. *et al.* (2022). 'Protection against SARS-CoV-2 after covid-19 vaccination and previous infection'. *N. Engl. J. Med.* 386 (13), pp. 1207–1220. DOI: 10.1056/NEJMoa2118691.

Hall, V. J., Insalata, F., Foulkes, S., Kirwan, P. *et al.* (2024). 'Effectiveness of BNT162b2 mRNA vaccine third doses and previous infection in protecting against SARS-CoV-2 infections during the Delta and Omicron variant waves; the UK SIREN cohort study September 2021 to February 2022'. *J. Infect.* 88 (1), pp. 30–40. DOI: 10.1016/j.jinf.2023.10.022.

Hall, V. J., Foulkes, S., Charlett, A., Atti, A. *et al.* (2021). 'SARS-CoV-2 infection rates of antibody-positive compared with antibody-negative health-care workers in England: a large, multicentre, prospective cohort study (SIREN)'. *Lancet*. 397 (10283), pp. 1459–1469. DOI: 10.1016/S0140-6736(21)00675-9.

Jackson, C. H., Tom, B. D., Kirwan, P. D., Mandal, S. *et al.* (2022). 'A comparison of two frameworks for multi-state modelling, applied to outcomes after hospital admissions with COVID-19'. *Stat. Methods Med. Res.* 31 (9), pp. 1656–1674. DOI: 10.1177/09622802221106720.

Mulchandani, R., Taylor-Philips, S., Jones, H. E., Ades, A. E. *et al.* (2020). 'Self assessment overestimates historical COVID-19 disease relative to sensitive serological assays: cross sectional study in UK key workers'. *bioRxiv*. DOI: 10.1101/2020.08.19.20178186.

Pantazis, N., Rosinska, M., van Sighem, A., Quinten, C. *et al.* (2021). 'Discriminating between premigration and postmigration HIV acquisition using surveillance data'. *J. Acquir. Immune Defic. Syndr.* 88 (2), pp. 117–124. DOI: 10.1097/qai.0000000000002745.

Presanis, A. M., Harris, R. J., Kirwan, P. D., Miltz, A. *et al.* (2021). 'Trends in undiagnosed HIV prevalence in England and implications for eliminating HIV transmission by 2030: an evidence synthesis model'. *Lancet Public Health*. 6 (10), e739–e751. DOI: 10.1016/S2468-2667(21)00142-0.

Wallace, S., Hall, V., Charlett, A., Kirwan, P. D. *et al.* (2022). 'Impact of prior SARS-CoV-2 infection and COVID-19 vaccination on the subsequent incidence of COVID-19: a multicentre prospective cohort study among UK healthcare workers - the SIREN (Sarscov2 Immunity & REinfection EvaluatioN) study protocol'. *BMJ Open*. 12 (6), e054336. DOI: 10.1136/bmjopen-2021-054336.

First-author pre-prints

Kirwan, P. D., Elgohari, S., Jackson, C. H., Tom, B. D. M. *et al.* (2021). 'Trends in risks of severe events and lengths of stay for COVID-19 hospitalisations in England over the pre-vaccination era: results from the Public Health England SARI-Watch surveillance scheme'. *arXiv [stat.AP]*. EPRINT: 2103.04867.

Conference presentations and posters

Sensitivity of an HIV back-calculation model to counterfactual COVID-19 lockdown assumptions, **Greek Stochastics 2022**, Aug 2022 [Oral presentation], **UKHSA 2022**, Oct 2022 [Poster presentation], and **MRC Biostatistics Unit Armitage week 2023** [Poster presentation].

Protection of SARS-CoV-2 vaccination and prior infection against future infection. **MRC Biostatistics Unit Armitage week 2022**, Sep 2022 [Oral presentation], and **RSS 2022**, Sep 2022 [Poster presentation].

Trends in COVID-19 hospital outcomes in England before and after vaccine introduction, 2020–2021: a cohort study. **UKHSA 2022**, Oct 2022 [Oral presentation].

Protection of COVID-19 booster vaccination and prior infection against symptomatic and asymptomatic SARS-CoV-2. **UKHSA 2023**, Nov 2023 [Poster presentation] and **ESCAIDE 2023**, Nov 2023 [Oral presentation by co-author].

Monitoring protection from SARS-CoV-2 infection and COVID-19 vaccination in the UK SIREN healthcare worker cohort 2020–2024: Assessing the impact of vaccine boosters. **ECCMID 2024**, Apr 2024 [Poster presentation].

Other academic activities

MRC Biostatistics Unit PhD student representative and organised the annual PhD student Symposium, 2021–2022.

Invited speaker at Scientific Pandemic Influenza Group on Modelling (SPI-M) sub-group meeting, Feb 2022.

Interviewer for Armitage week: In conversation with... Prof. Nicholas Jewell, Sep 2022. https://www.mrc-bsu.cam.ac.uk/event/19th-armitage-workshop-and-lecture/

Course demonstrator on Bayesian Modelling short course delivered to PhD students in Addis Ababa, Ethiopia, Oct 2022. https://www.mrc-bsu.cam.ac.uk/blog/bsu-researchers-teach-bayesian-statistics-course-in-ethiopia/

Recipient of SPI-M Award for Modelling and Data Support (SAMDS), Oct 2022. https://www.mrc-bsu.cam.ac.uk/blog/mrc-biostatistics-unit-covid-19-working-group-receive-spi-m-o-award/

Invited speaker and panellist on European Testing Week/EuroTEST 'Updated definition of HIV late diagnosis' webinar, Nov 2022. https://www.testingweek.eu/resources/webinars-videos/etw-2022-updated-definition-of-hiv-late-diagnosis/

Course demonstrator on University of Cambridge Population Health Science MPhil, Bayesian Modelling and Infectious Disease Modelling modules, Easter 2022 and 2023.

Invited speaker at Joint Committee on Vaccination and Immunisation (JCVI) COVID-19 meeting, May 2023.

B. Appendix to Chapter 2

B.1 Cox proportional hazards model

In the Cox proportional hazards model introduced in Section 2.2.4, the partial likelihood for regression coefficients β_m was defined as:

$$L(\boldsymbol{\beta}) = \prod_{i \in D} \frac{\exp(\boldsymbol{\beta}^\mathsf{T} \boldsymbol{z}_{d_i})}{\sum_{j \in R_i} \exp(\boldsymbol{\beta}^\mathsf{T} \boldsymbol{z}_j)}$$

The score function $U(\beta)$ and the observed information matrix $I(\beta)$ may be derived from this partial likelihood, with the standard estimate of the covariance matrix being the inverse of the observed information matrix:

$$\mathrm{Cov}(\beta) = I(\beta)^{-1}$$

B.1.1 Stratified Cox model

A key assumption of the Cox proportional hazards model is that hazards are 'proportional'. For the example of two covariates, vaccination and age, the model would state:

$$h(t \mid z) = h^{(0)}(t) \exp(\beta_{\text{age}} z_{\text{age}} + \beta_{\text{vax}} z_{\text{vax}})$$

where, again, $h^{(0)}(t)$ is the baseline hazard function. This implies that the trend in survival over time is the same for all age groups, and that e.g. (hazard for age group 2)/(hazard for age group 1) is a constant.

As described in Section 2.3.4, stratification may be used for covariates which exhibit non-proportionality. Stratification avoids needing to make the proportional hazards assumption by allowing the baseline hazard to differ according to covariate level. This is achieved by fitting two (or more) separate equations with a common regression coefficient β_{vax} (Collett 2023):

$$h_{\text{age=1}}(t \mid z) = h^{(0)}_{\text{age=1}}(t) \exp(\beta_{\text{vax}} z_{\text{vax}})$$
$$h_{\text{age=2}}(t \mid z) = h^{(0)}_{\text{age=2}}(t) \exp(\beta_{\text{vax}} z_{\text{vax}})$$

B.1.2 Accounting for correlation

The Cox proportional hazards model assumes that the failure times of individuals are independent. However, this assumption is unlikely to hold if, e.g. individuals work at the same location, and may be influenced by common effects. Two standard approaches to account for this so-called 'within-cluster' correlation are the robust variance estimator and shared frailty (Balan and Putter 2020).

Robust variance estimator

The robust variance estimator approach is to adjust the standard errors of the estimated hazard ratios to account for the extra variability due to within-cluster correlation. The standard model-based estimate of the variance-covariance matrix of the parameter estimates in a Cox regression model is the inverse of the observed information matrix: $\text{Cov}(\beta) = I(\beta)^{-1}$ (Collett 2023). A more robust variance estimator is the so-called 'sandwich' estimator, given by $\text{Cov}(\beta) = I(\beta)^{-1} S(\beta) I(\beta)^{-1}$, where $S(\beta)$ is a correction term that captures the actual variance of the data (Lin and Wei 1989).

Shared frailty

The shared frailty approach is to directly incorporate cluster-specific random effects into the model via a shared parameter, or 'frailty' term, which applies to all individuals within the cluster (Austin 2017). In a Cox proportional hazards model with shared frailty, the hazard function for individual i in cluster j may be formulated as:

$$\begin{aligned} h_i(t) &= h^{(0)}(t) \exp\left(\sum_{m=1}^{M} \beta_m z_m + \alpha_j\right) \\ &= h^{(0)}(t) \exp(\alpha_j) \exp\left(\sum_{m=1}^{M} \beta_m z_m\right) \end{aligned}$$

where α_j denotes the random effect associated with the jth cluster. The frailty term $\exp(\alpha_j)$ is "shared" among the cluster, not specific to an individual, with a multiplicative effect on the baseline hazard function. The distribution of this shared frailty term is commonly specified as a gamma distribution with mean 1 and variance θ (Balan and Putter 2020).

C. Appendix to Chapter 3

C.1 Biological markers of disease progression

C.1.1 RITA algorithm

The RITA algorithm defines a 'recent' diagnosis as an AxSYM avidity index score less than 80.0% or a LAg normalised optical density score less than 1.5 within 120 days of HIV diagnosis and neither of: a CD4 count <50 cells/mm^3, an AIDS diagnosis, nor an HIV viral load <400 copies/mL within 90 days of HIV diagnosis (Aghaizu *et al.* 2014). The algorithm is shown in the flowchart in Figure C.1.

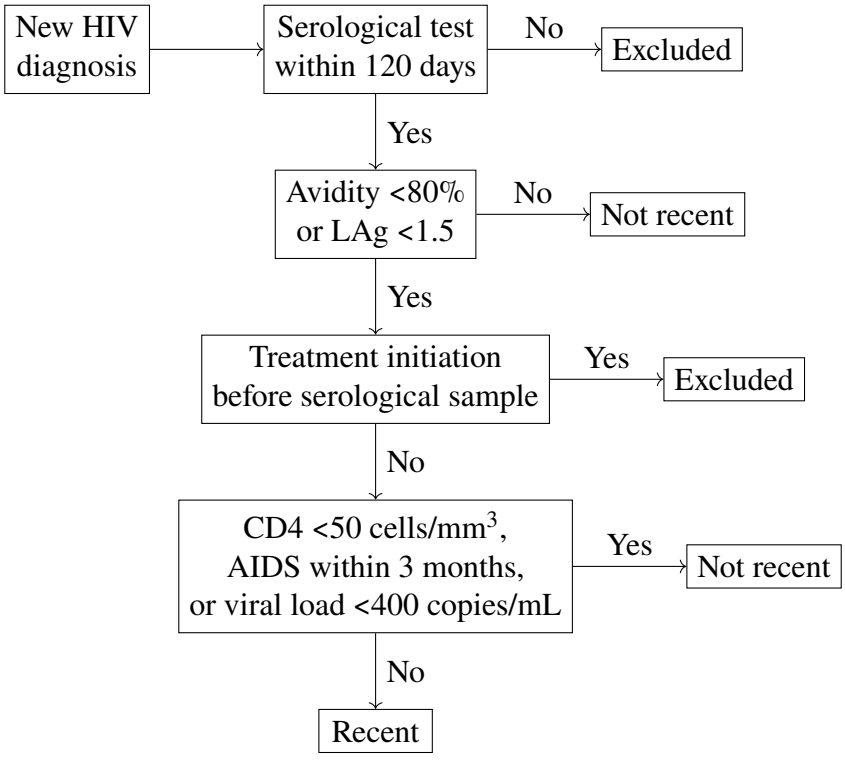


Figure C.1: RITA algorithm applied to EW&NI HIV surveillance data.

C.1.2 Viral load

The HIV viral load measures the number of HIV virus copies in a millilitre of blood, expressed as copies/mL. As shown in Figure 3.1, in the absence of intervention, viral load measurements follow an approximately inverse pattern to the CD4 trajectory; peaking at around 100,000 copies/mL as a result of viremia during acute HIV infection, falling to a lower level post-seroconversion, then rising in late infection (Maartens *et al.* 2014). With initiation of ART, HIV viral load will diminish, preventing onwards transmission of the virus (Delpech 2022).

Viral loads are collected routinely at healthcare visits and, at the population level, inform measurement of the proportion of people with untransmittable levels of virus.

C.1.3 p24 antigen

The p24 HIV antigen is a viral protein which makes up the HIV viral core (Pebody 2021b). p24 antigen levels in the blood spike during acute infection, but then subside as antibodies to p24 start to be produced the body during seroconversion. As a result, the p24 antigen is typically only detectable during early infection and p24 antigen assays alone are not a reliable way of diagnosing HIV after seroconversion.

The latest 'fourth-generation' HIV diagnostic tests make use of the fact that p24 antigens are detectable in blood before HIV antibodies by combining p24 antigen and HIV antibody tests (whereas 'third-generation' tests only detect HIV antibodies). Fourth-generation antibody/antigen laboratory tests are now the standard screening assay in the UK, with an estimated median window period between infection and detection of 17.8 days (IQR 13.0 to 23.6 days) (Delaney *et al.* 2017).

C.1.4 HIV testing history

Whilst not a strictly 'biological' marker of disease progression, HIV testing history can provide useful information about an individual's likely time of HIV acquisition: since individuals with a previous negative HIV test must have acquired HIV in the period between the last negative and first positive test. A more recent negative test indicates a shorter window during which the infection was acquired, for example, a previous negative HIV test within 12 months could be classed as evidence of 'recent' infection.

C.2 Availability of biomarker data

C.2.1 HIV avidity assay data

Samples for HIV avidity assay testing are collected through sentinel surveillance from a subset of clinics. Aghaizu 2018 demonstrated that the population tested for avidity assays prior to 2013 were broadly representative of all EW&NI diagnoses. To assess whether this remains true in more recent years, and in particular among GBM, all available avidity assay scores from national testing and linked to a diagnosis record were considered.

Table C.1 shows the demographic characteristics of GBM diagnosed in EW&NI between 2011–2019 by avidity assay availability. Age and ethnicity characteristics were similar for GBM with and without assay results available, although avidity assay results were slightly more common among younger GBM compared to older GBM. Over the 14-year analysis period, diagnoses in London and Northern Ireland were over-represented for assay result availability, whilst diagnoses in Wales and the Midlands and East of England were under-represented.

Figure C.2 panel A shows the availability of avidity assay data over time among newly diagnosed GBM, which ranged between 51.4% and 63.8% annually between 2011 and 2019.

As shown in Figure C.2 panel B, the CD4 profile among newly diagnosed GBM with and without an assay result available was similar for each calendar year. Diagnoses with an avidity assay result available had fewer missing CD4 counts (as in Table C.1), but, where CD4 data were available, annual trends were similar regardless of the availability of avidity assay.

The RITA algorithm was applied to obtain the classification of recent or non-recent HIV acquisition (see Section C.1 for details). Figure C.2 panel C shows trends in the proportion of diagnoses classified as recently acquired HIV each year, stratified by CD4 count at diagnosis and avidity assay. Whilst a change of avidity assay took place in 2013 (as described in Section 3.2.4), the MDRI was approximately equal between the two assays (around 6 months) and similar proportions were obtained regardless of which assay was used. Trends were relatively consistent over time in each CD4 strata; for the most part the strata did not overlap, and those with a missing CD4 count appear somewhere in the middle of the range. Those with a CD4 count above 500 cells/mm^3 were most likely to be diagnosed with recently acquired HIV. However, a sizeable proportion of diagnoses with low CD4 counts (CD4 <200 or 200–350 cells/mm^3) were also classified as recent.

Figure C.2 panel D shows the same proportion of diagnoses recently acquired broken down by age group. Focussing on those with low CD4 counts (CD4 <200 or 200–350 cells/mm^3), younger individuals were more likely to be classified with recently acquired HIV as compared to older individuals.

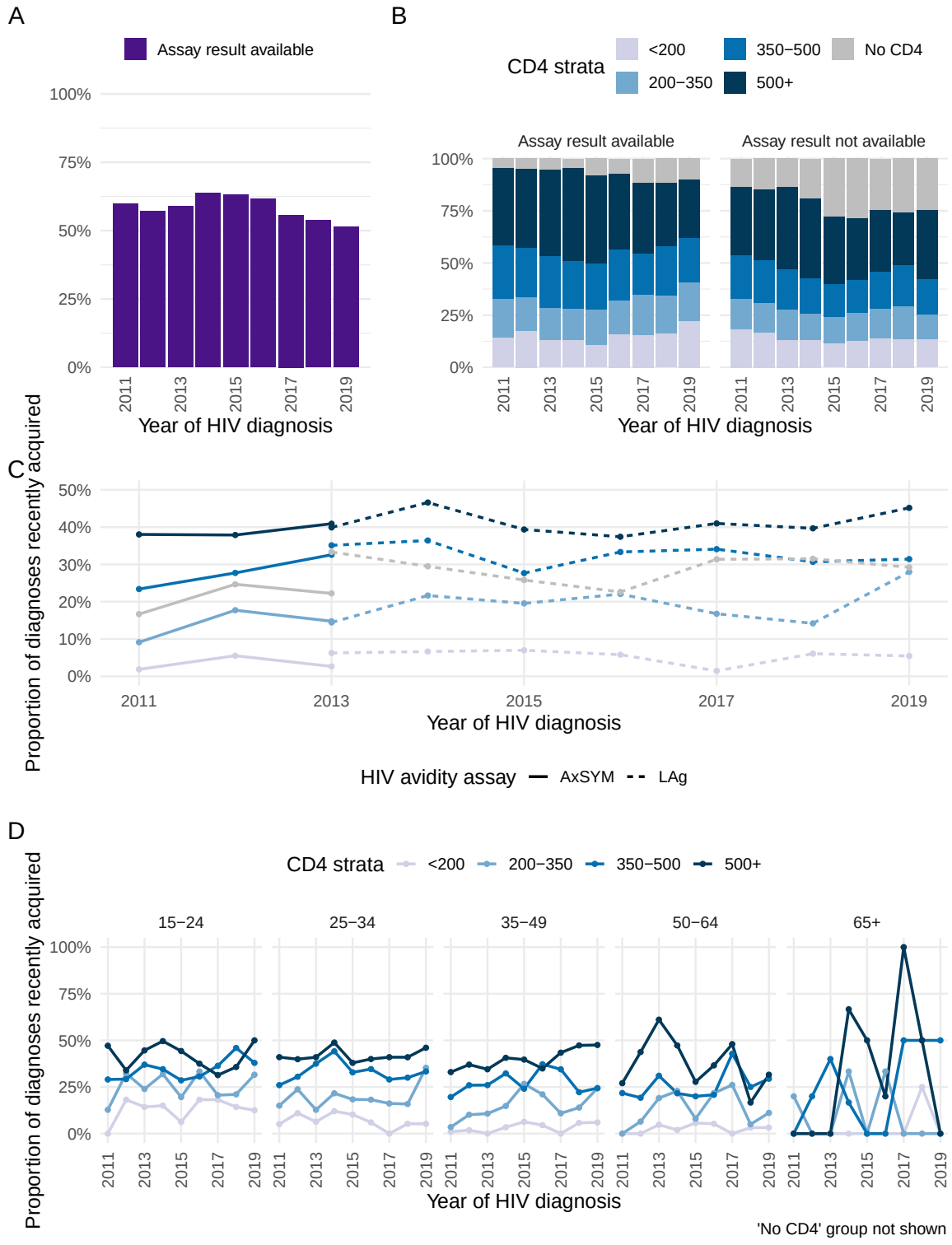


Figure C.2: Availability of HIV avidity assay among GBM, by year of HIV diagnosis (Panel A) and CD4 count at HIV diagnosis (Panel B). Proportion of diagnoses among GBM recently acquired, by year of HIV diagnosis, CD4 count at diagnosis and avidity assay used (Panel C) and age group at HIV diagnosis (both assays combined) (Panel D).

Table C.1: Demographic characteristics of GBM newly diagnosed in EW&NI between 2011–2019 by availability of incidence assay.

Characteristic	**Assay result available** N = 11,653	**Assay result not available** N = 8,045
Age group at diagnosis		
15-24	1,850 (15.9%)	1,205 (15.0%)
25-34	4,533 (38.9%)	3,015 (37.5%)
35-49	3,951 (33.9%)	2,772 (34.5%)
50-64	1,185 (10.2%)	903 (11.2%)
65+	134 (1.1%)	150 (1.9%)
Ethnicity		
White	9,068 (77.8%)	6,262 (77.8%)
Black African	316 (2.7%)	187 (2.3%)
Black Caribbean	236 (2.0%)	150 (1.9%)
Black other	133 (1.1%)	84 (1.0%)
Asian	703 (6.0%)	478 (5.9%)
Other/Mixed	906 (7.8%)	639 (7.9%)
Not Stated	291 (2.5%)	245 (3.0%)
Region of residence		
London	5,728 (49.2%)	3,235 (40.2%)
Midlands and East of England	1,659 (14.2%)	1,432 (17.8%)
North of England	2,217 (19.0%)	1,439 (17.9%)
South of England	1,723 (14.8%)	1,323 (16.4%)
Northern Ireland	282 (2.4%)	69 (0.9%)
Wales	44 (0.4%)	547 (6.8%)
CD4 count at diagnosis		
<200	1,724 (14.8%)	1,139 (14.2%)
200-350	1,965 (16.9%)	1,124 (14.0%)
350-500	2,716 (23.3%)	1,465 (18.2%)
500+	4,469 (38.4%)	2,688 (33.4%)
No CD4	779 (6.7%)	1,629 (20.2%)

Table C.2 shows the overall number and proportion in each CD4 stratum, by age group, for GBM diagnosed with recently acquired HIV between 2011–2019. Whilst numbers are small for certain groups, 15.7% of those aged 15–24 had a CD4 count in the <200 or 200–350 cells/mm^3 range compared to 12.3% among those aged 35–49. Table C.3 shows the same data, broken down by year of HIV diagnosis, demonstrating that CD4 distributions among those recently diagnosed were relatively stable over time, with smaller numbers of tests available in recent years.

Table C.2: Distribution of CD4 cell counts among those diagnosed with recently acquired HIV between 2011–2019, by age group at diagnosis.

	Age group at diagnosis					
	15-24	25-34	35-49	50-64	65+	Total
CD4 strata						
<200	16 (2.7%)	37 (2.7%)	21 (2.3%)	9 (4.2%)	1 (5.6%)	84 (2.7%)
200-350	77 (13.0%)	146 (10.8%)	90 (9.9%)	30 (14.0%)	3 (16.7%)	346 (11.2%)
350-500	167 (28.3%)	367 (27.2%)	236 (26.0%)	56 (26.2%)	7 (38.9%)	833 (27.1%)
500+	331 (56.0%)	800 (59.3%)	559 (61.7%)	119 (55.6%)	7 (38.9%)	1,816 (59.0%)

Table C.3: Distribution of CD4 cell counts among those diagnosed with recently acquired HIV between 2011–2019, by year of diagnosis.

	CD4 strata			
	<200	200-350	350-500	500+
Year of diagnosis				
2011	4 (1.2%)	26 (7.8%)	91 (27.2%)	213 (63.8%)
2012	15 (3.9%)	45 (11.7%)	102 (26.6%)	222 (57.8%)
2013	8 (1.8%)	36 (8.1%)	133 (29.8%)	270 (60.4%)
2014	16 (2.6%)	60 (9.9%)	150 (24.6%)	383 (62.9%)
2015	12 (2.9%)	51 (12.1%)	98 (23.3%)	259 (61.7%)
2016	11 (3.5%)	43 (13.7%)	97 (31.0%)	162 (51.8%)
2017	2 (0.9%)	29 (13.4%)	60 (27.8%)	125 (57.9%)
2018	8 (4.3%)	21 (11.4%)	58 (31.4%)	98 (53.0%)
2019	8 (4.7%)	35 (20.5%)	44 (25.7%)	84 (49.1%)
Total	84 (2.7%)	346 (11.2%)	833 (27.1%)	1,816 (59.0%)

C.2.2 HIV testing history data

HIV testing history information, in particular the date of a previous negative HIV test, are directly collected through HARS whilst records of STI clinic activity (visits and tests) are monitored through the genitourinary medicine clinic activity dataset (GUMCAD). These data represent upper bounds on the time since HIV acquisition for an individual, so whilst a last negative test within 6 months may indicate recent acquisition, the lack of such a result is not an indication of longstanding infection.

HIV testing history was available for a subset of diagnosed individuals and for this analysis, all available last negative HIV test results from HARS and GUMCAD which had been linked to a subsequent HIV diagnosis record for GBM in EW&NI were considered.

Table C.4 shows the demographic characteristics of newly diagnosed GBM by previous negative test availability. Previous negative tests were more commonly available for younger

GBM compared to older GBM, and those with White, Black Caribbean, and Black other ethnicity were more likely to have a previous test reported compared to other ethnicities. Previous test availability was higher among those resident in the South of England compared to other regions in England, and substantially lower in Wales and Northern Ireland. Previous negative tests were also much less common among those with a CD4 count at diagnosis <200 cells/mm^3.

Table C.4: Demographic characteristics of GBM newly diagnosed in EW&NI between 2011–2019 by availability of previous negative test.

Characteristic	**Previous negative test** N = 10,524	**No previous negative test** N = 9,174
Age group at diagnosis		
15-24	1,675 (15.9%)	1,380 (15.0%)
25-34	4,224 (40.1%)	3,324 (36.2%)
35-49	3,609 (34.3%)	3,114 (33.9%)
50-64	914 (8.7%)	1,174 (12.8%)
65+	102 (1.0%)	182 (2.0%)
Ethnicity		
White	8,300 (78.9%)	7,030 (76.6%)
Black African	263 (2.5%)	240 (2.6%)
Black Caribbean	232 (2.2%)	154 (1.7%)
Black other	123 (1.2%)	94 (1.0%)
Asian	560 (5.3%)	621 (6.8%)
Other/Mixed	812 (7.7%)	733 (8.0%)
Not Stated	234 (2.2%)	302 (3.3%)
Region of residence		
London	4,774 (45.4%)	4,189 (45.7%)
Midlands and East of England	1,613 (15.3%)	1,478 (16.1%)
North of England	1,992 (18.9%)	1,664 (18.1%)
South of England	1,826 (17.4%)	1,220 (13.3%)
Northern Ireland	87 (0.8%)	264 (2.9%)
Wales	232 (2.2%)	359 (3.9%)
CD4 count at diagnosis		
<200	1,058 (10.1%)	1,805 (19.7%)
200-350	1,693 (16.1%)	1,396 (15.2%)
350-500	2,505 (23.8%)	1,676 (18.3%)
500+	4,268 (40.6%)	2,889 (31.5%)
No CD4	1,000 (9.5%)	1,408 (15.3%)

Figure C.3 panel A shows the availability of HIV testing history over time for GBM newly diagnosed with HIV between 1995 and 2019. The proportion with information on HIV testing history available rose from 24% in 1995 to a peak of 60% in 2015, falling somewhat in

subsequent years. A corresponding peak in the proportion with a negative test within the past 6 months was seen in 2015 (22% of newly diagnosed GBM). Temporal trends were similar across all age groups, with a greater proportion of younger GBM having HIV testing history available compared to older GBM (Figure C.3 panel B).

Figure C.3 panel C presents the proportion with a previous negative test within 6 months by age group and CD4 count at diagnosis, between 2011–2019. A recent negative test within the past 6 months was less common among HIV diagnoses with a CD4 count <200 cells/mm^3, consistent across all age groups, with some overlap for the other CD4 strata.

C.2.3 Viral load data

A high viral load and low CD4 count are present in both early and late presentation of HIV, as in Figure 3.1, therefore baseline HIV viral load information cannot be used to identify individuals with evidence of recent HIV acquisition. When taken in conjunction with other biomarkers, however, a high viral load may provide supporting evidence of diagnosis occurring during early infection.

Baseline HIV viral load data are reported for the majority of new HIV diagnoses, with additional information available via linkage to HARS. For this analysis, only viral loads taken within 91 days of an initial diagnosis were considered and, where treatment start date information was available, any viral loads taken post-treatment initiation were excluded.

Table C.5 shows the demographic characteristics of GBM newly diagnosed with and without baseline HIV viral load (VL) information available. Viral load availability was similar by age and ethnicity, with some regional differences in data availability (higher in the South of England and Northern Ireland, lower in London). Where CD4 count data was missing, a baseline viral load was also much more likely to be unavailable.

Figure C.4 panel A shows baseline viral load information among newly diagnosed GBM according to RITA classification. Across all CD4 strata, the distribution of baseline viral load was markedly different among those classified as recent compared to non-recent diagnoses. This difference was partly due to viral loads <400 copies/mL being explicitly excluded from the RITA algorithm, but also as a much higher proportion of recently acquired HIV were diagnosed with a viral load above 100,000 copies/mL.

Figure C.4 panel B shows baseline viral load by time since previous negative HIV test (panel B). Again those with a more recent negative test were more likely to have a baseline viral load above 100,000 copies/mL. However, the differences in baseline viral load by testing history were less marked than according to RITA classification.

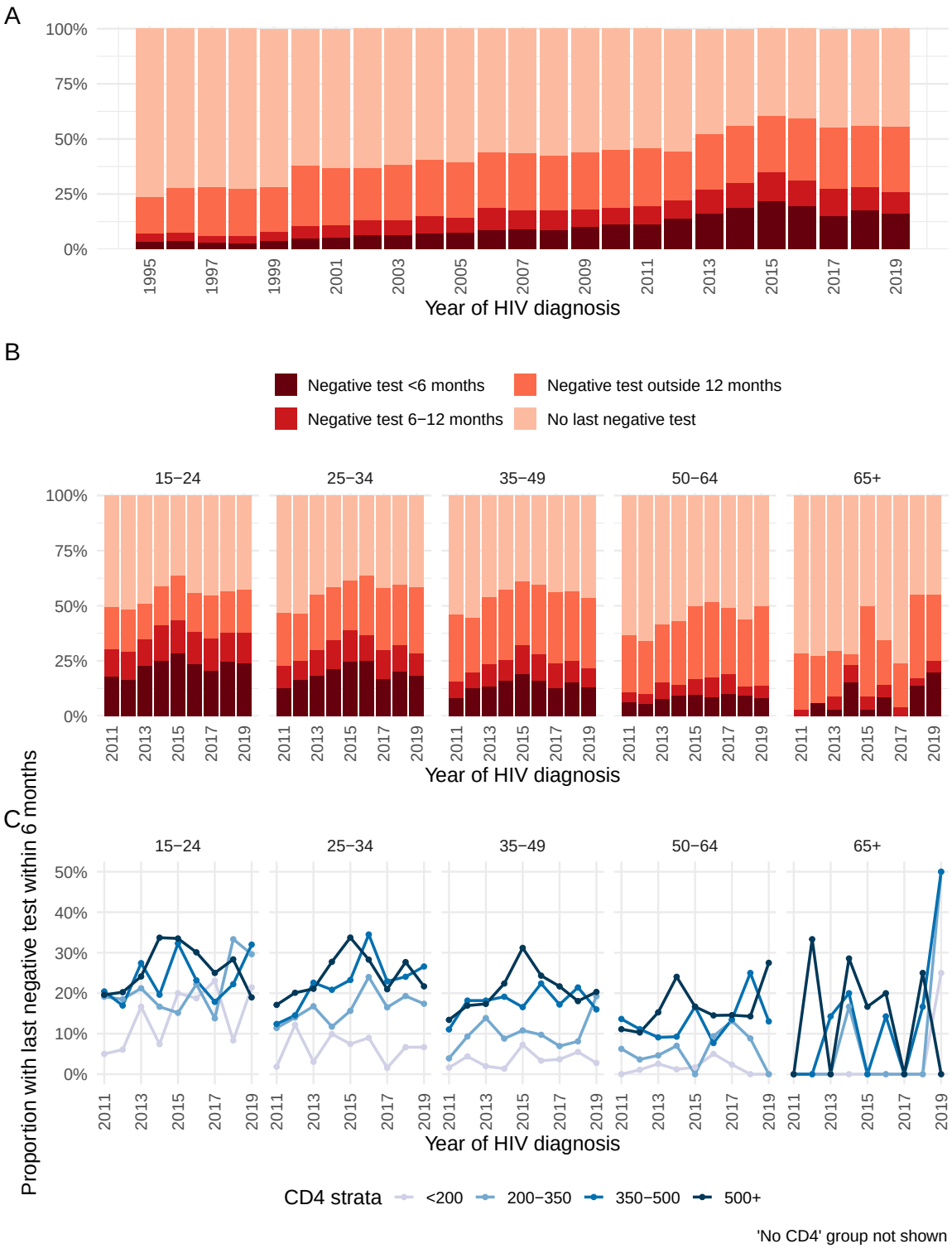


Figure C.3: Time since last negative HIV test among GBM, by year of HIV diagnosis (Panel A), and stratified by age group at HIV diagnosis (Panel B). Proportion of HIV diagnoses with a recent test within past 6 months, stratified by CD4 count and age group at diagnosis (Panel C).

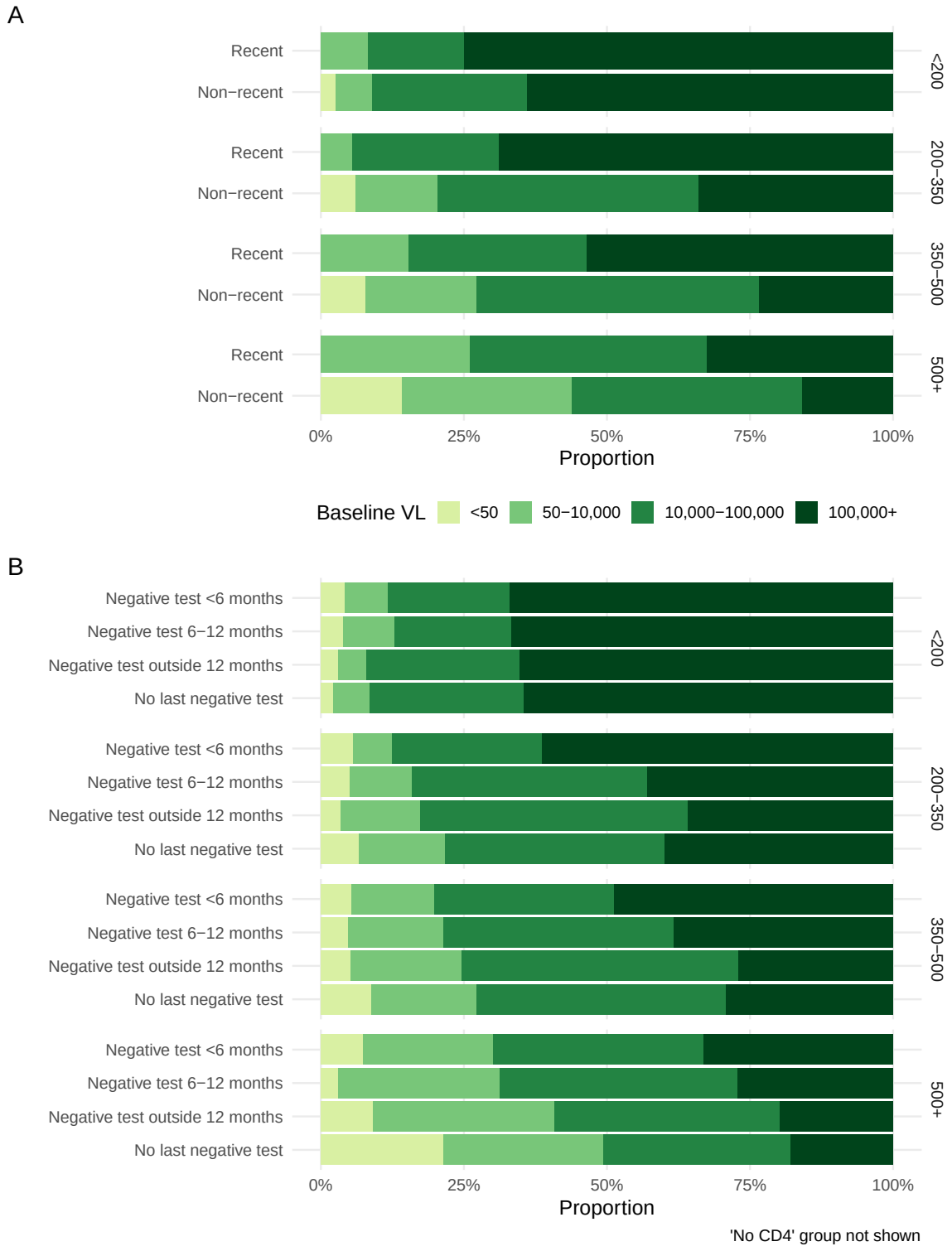


Figure C.4: Baseline HIV viral load and CD4 strata among GBM, by avidity assay and time since previous negative test.

Table C.5: Demographic characteristics of GBM newly diagnosed in EW&NI between 2011–2019 by availability of baseline VL.

Characteristic	**Baseline VL available** N = 12,313	**Baseline VL not available** N = 7,385
Age group at diagnosis		
15-24	1,957 (15.9%)	1,098 (14.9%)
25-34	4,693 (38.1%)	2,855 (38.7%)
35-49	4,138 (33.6%)	2,585 (35.0%)
50-64	1,350 (11.0%)	738 (10.0%)
65+	175 (1.4%)	109 (1.5%)
Ethnicity		
White	9,756 (79.2%)	5,574 (75.5%)
Black African	298 (2.4%)	205 (2.8%)
Black Caribbean	232 (1.9%)	154 (2.1%)
Black other	135 (1.1%)	82 (1.1%)
Asian	744 (6.0%)	437 (5.9%)
Other/Mixed	901 (7.3%)	644 (8.7%)
Not Stated	247 (2.0%)	289 (3.9%)
Region of residence		
London	4,912 (39.9%)	4,051 (54.9%)
Midlands and East of England	2,106 (17.1%)	985 (13.3%)
North of England	2,518 (20.4%)	1,138 (15.4%)
South of England	2,176 (17.7%)	870 (11.8%)
Northern Ireland	245 (2.0%)	106 (1.4%)
Wales	356 (2.9%)	235 (3.2%)
CD4 count at diagnosis		
<200	1,990 (16.2%)	873 (11.8%)
200-350	2,193 (17.8%)	896 (12.1%)
350-500	2,946 (23.9%)	1,235 (16.7%)
500+	4,980 (40.4%)	2,177 (29.5%)
No CD4	204 (1.7%)	2,204 (29.8%)

Figure C.5 panel A shows baseline viral load availability between 2011–2019, which has remained around 60% over this period. Figure C.5 panel B demonstrates the lower completion of CD4 data among diagnoses with missing baseline viral load information, particularly after 2015. Figure C.5 panel C shows baseline viral load by age group and CD4 strata. Higher baseline viral load was seen for lower CD4 strata, e.g. a greater proportion with a viral load above 100,000 copies/mL in the CD4 <200 cells/mm^3 stratum. Examined by age group, older individuals tended to have a higher baseline viral load at HIV diagnosis, regardless of CD4 strata. For more recent diagnosis, an increasing proportion had a baseline viral load <50 copies/mL, particularly in the CD4 >500 cells/mm^3 stratum (Figure C.5 panel D).

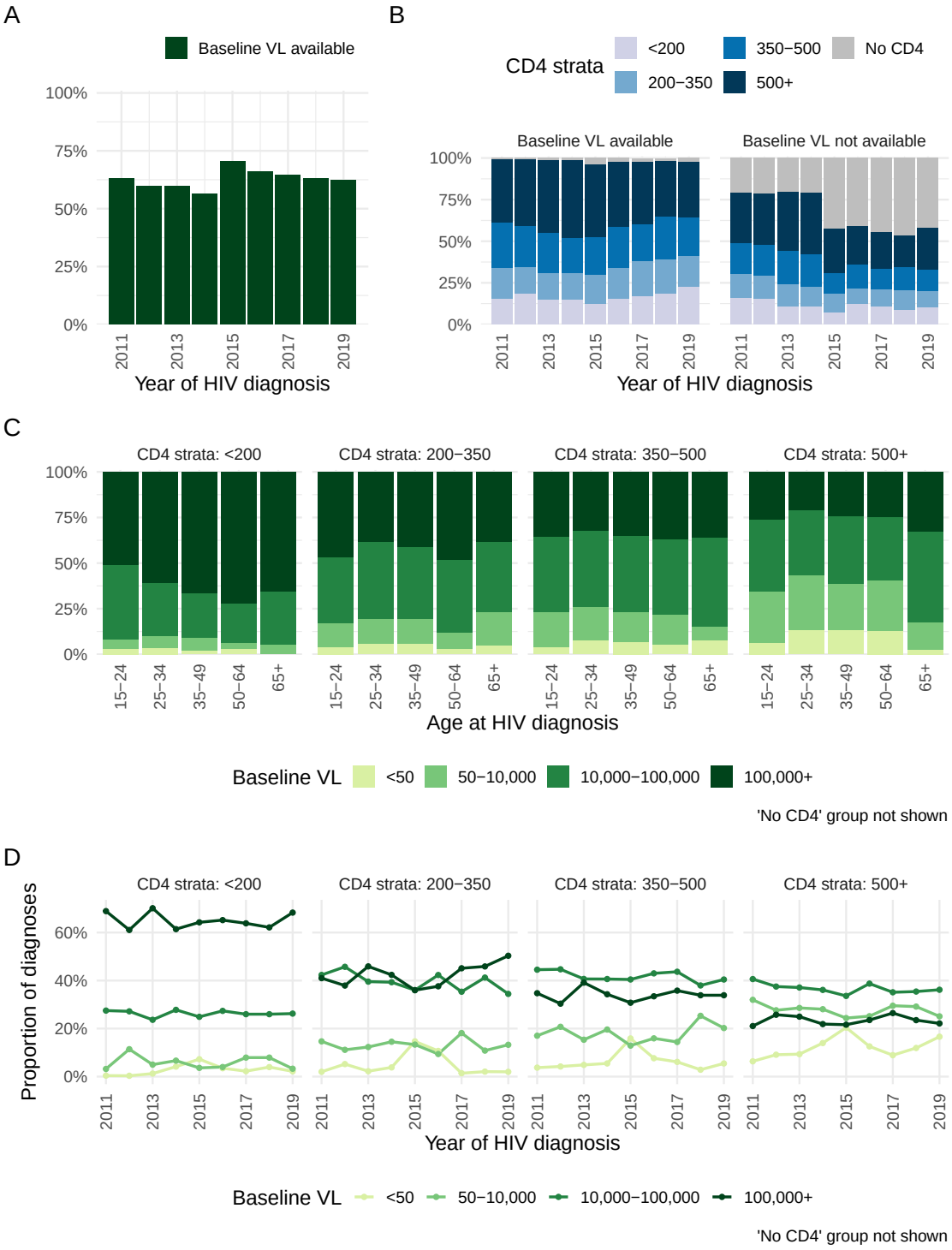


Figure C.5: Availability of baseline HIV viral load (VL) among GBM, by year of HIV diagnosis (Panel A) and CD4 count at HIV diagnosis (Panel B). Baseline VL among GBM by year of HIV diagnosis and age group (Panel C), and by CD4 cell count at diagnosis (Panel D).

C.3 Revised late HIV diagnosis definition

For the reclassification method for late HIV diagnosis, a correction factor (CF) was defined as the percentage change in the late HIV diagnosis rate after reclassification of individuals with evidence of recent HIV acquisition(Kirwan, Croxford *et al.* 2022):

$$\text{correction factor} = \frac{\text{number with CD4 <350 cells/mm}^3\text{ reclassified as 'not late'}}{\text{number with CD4 count <350 cells/mm}^3}$$

In 2019, across all newly diagnosed individuals in EW&NI, reclassification for recent infection resulted in a downwards adjustment of the late diagnosis rate by 14%. This was highest among GBM (26% in 2019), particularly younger GBM (CF of 46% for those aged 15–24 years) and those living in London (CF 30%). Figure C.6 panel A shows the proportion of individuals diagnosed with a CD4 ≥350 cells/mm^3 and CD4 <350 cells/mm^3, by reclassification and probable route of HIV exposure. The evidence used to reclassify is shown in Figure C.6 panel B.

Risk factors for reclassification included exposure through sex between men and a younger age at diagnosis. The refined definition of late HIV diagnosis was shown to focus the late diagnosis measure on those at greatest risk of death.

C.4 CD4 reclassification model

This section presents the results of the CD4-only back-calculation model fitted to the datasets simulated in scenario (iii), in which recent diagnoses were reclassified to the CD4 500+ stratum (as described in Section 3.5.2).

The estimates obtained from this model were very similar to the dual biomarker model, with overlapping trends in estimated incidence and undiagnosed prevalence in Figure C.7, and the PMSE was comparable between the two models (Figure C.8).

C.5 Simulation performance with missing information

This section presents the results of the dual biomarker and CD4-only back-calculation models fitted to simulated data scenarios (i) and (ii) with 40% missing CD4 and RITA information (as described in Section 3.5.2).

The missing information made little difference to the median estimates, as shown in Figure C.9. Again the PMSE was lower for the dual biomarker model compared to the CD4-only model (Figure C.10).

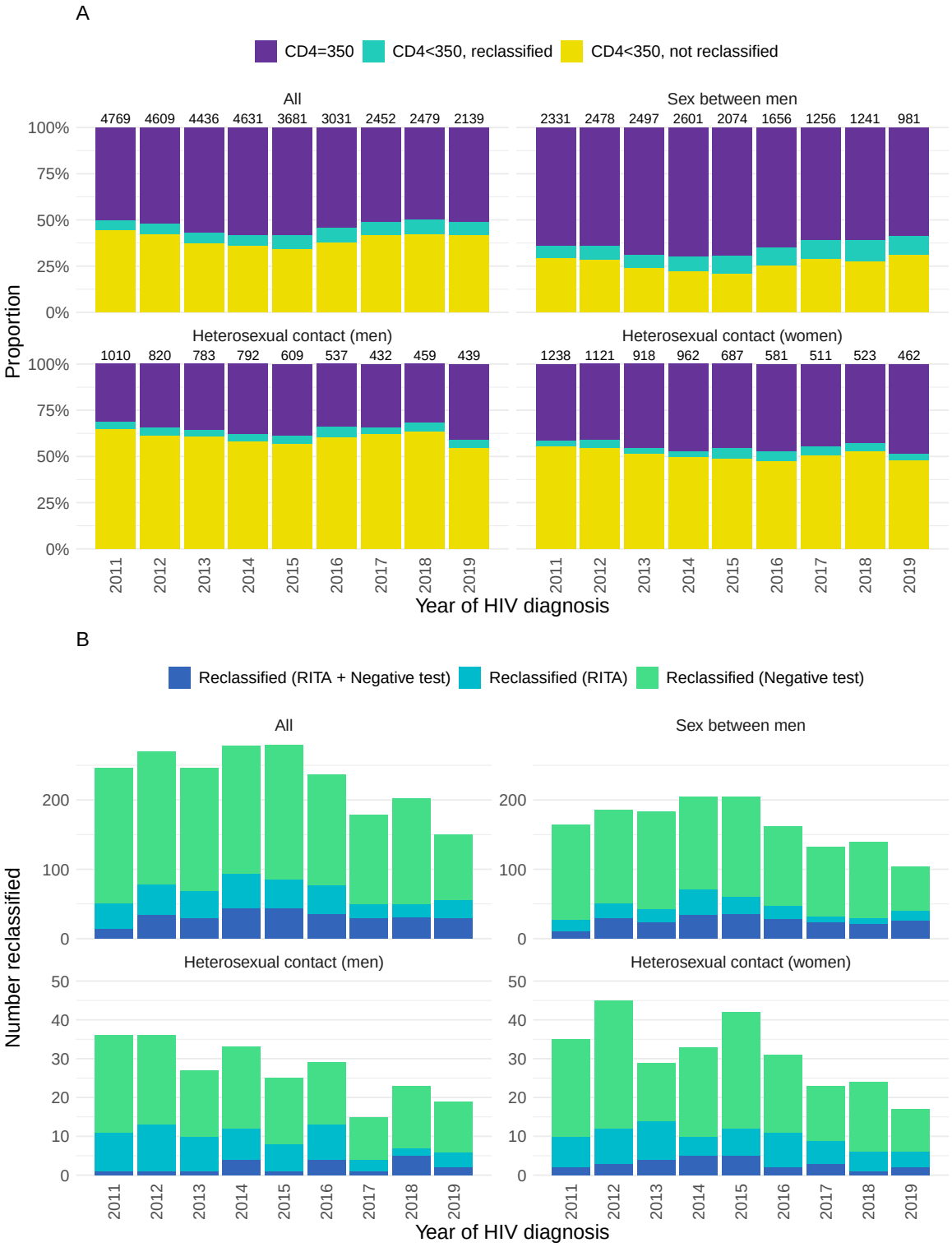


Figure C.6: HIV diagnoses by year of HIV diagnosis, route of HIV exposure, CD4 count, re-classification (Panel A) and reason for re-classification (Panel B). Total number of diagnoses by year and exposure group shown at top of bar in Panel A. RITA: recent infection testing algorithm.

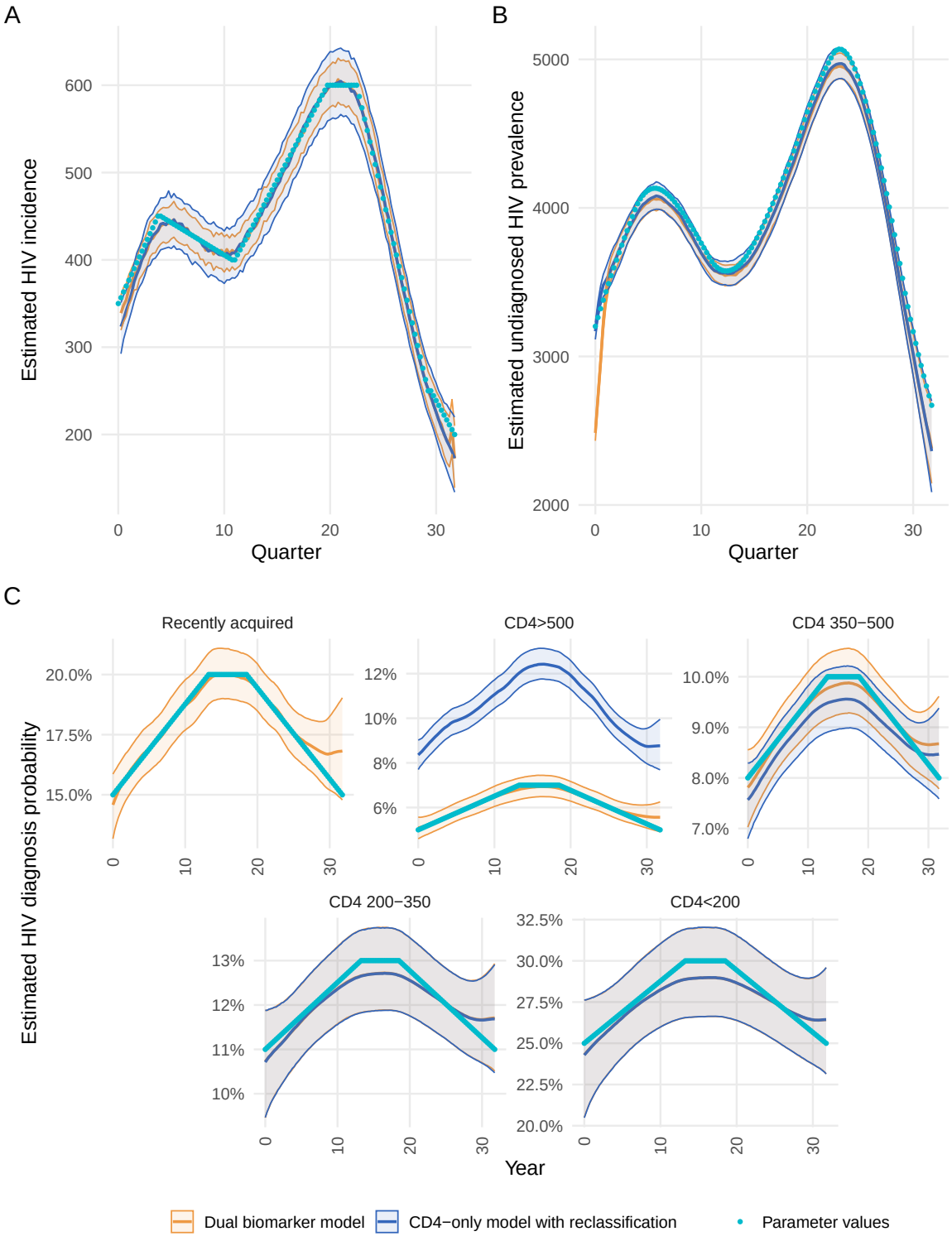


Figure C.7: Estimated posterior median and 95% CrI of HIV incidence, undiagnosed prevalence, and diagnosis probabilities, averaged over 100 model fits, compared to simulated data.

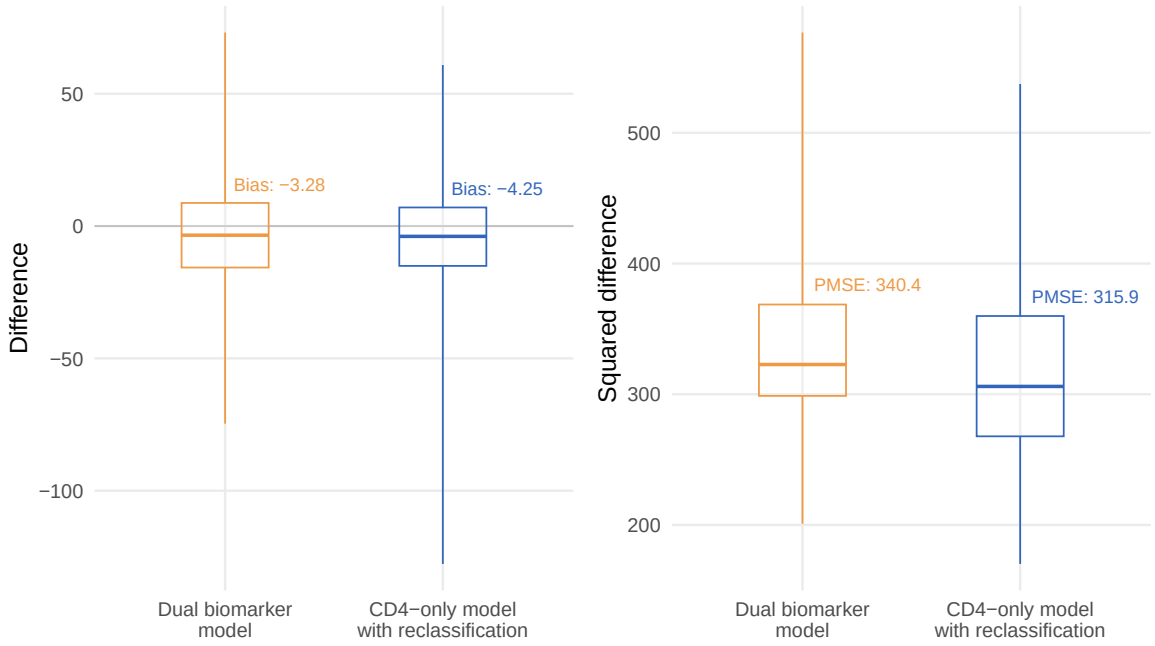


Figure C.8: Distributions of differences (panel A), and squared differences (panel B) between parameter value and estimated HIV incidence, by back-calculation model. Estimates from 100 model fits. Boxplots show median, inter-quartile range, and range of estimates for each model.

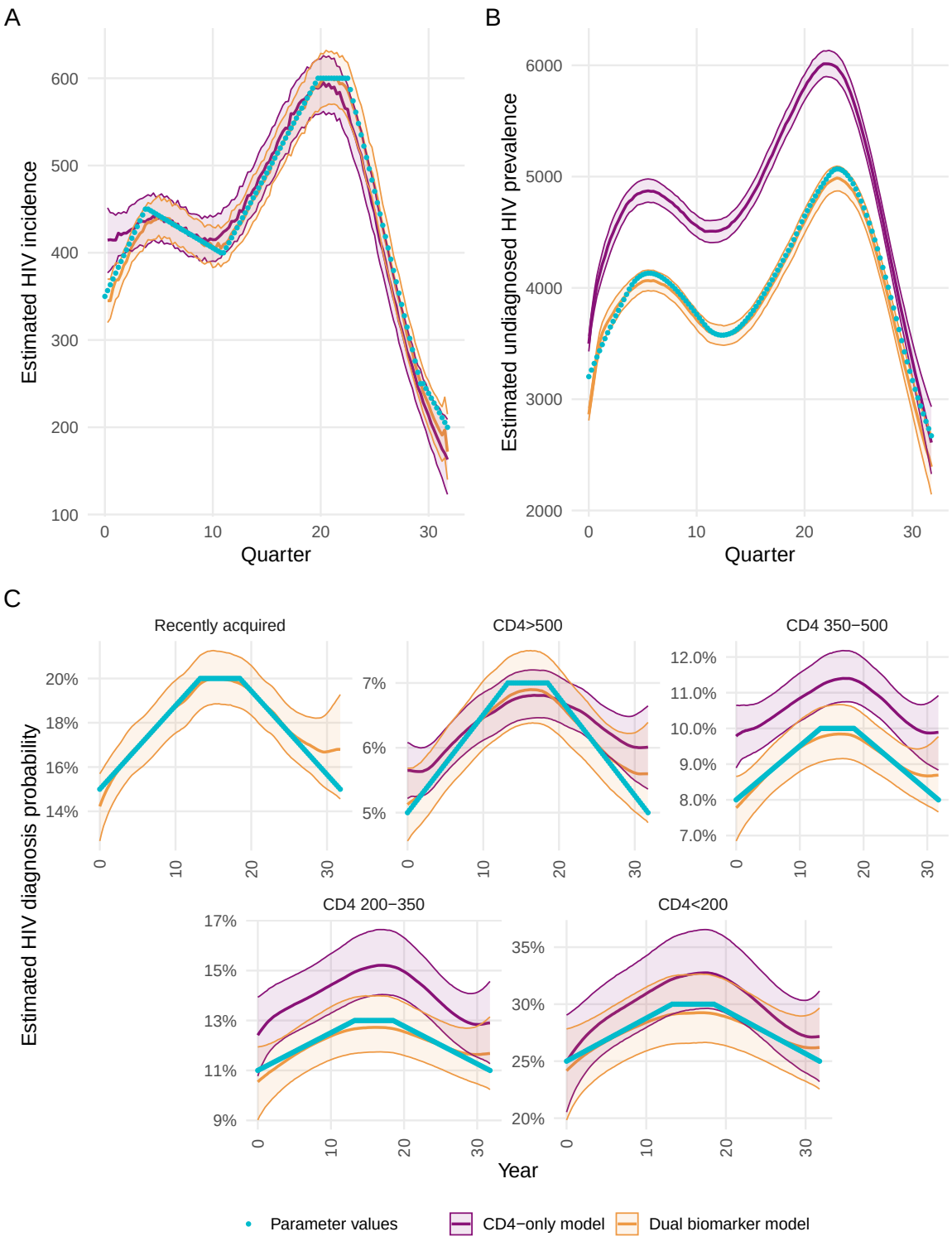


Figure C.9: Estimated posterior median and 95% CrI of HIV incidence, undiagnosed prevalence, and diagnosis probabilities, averaged over 100 model fits, compared to simulated data. 40% of records in simulated data missing CD4 and RITA information.

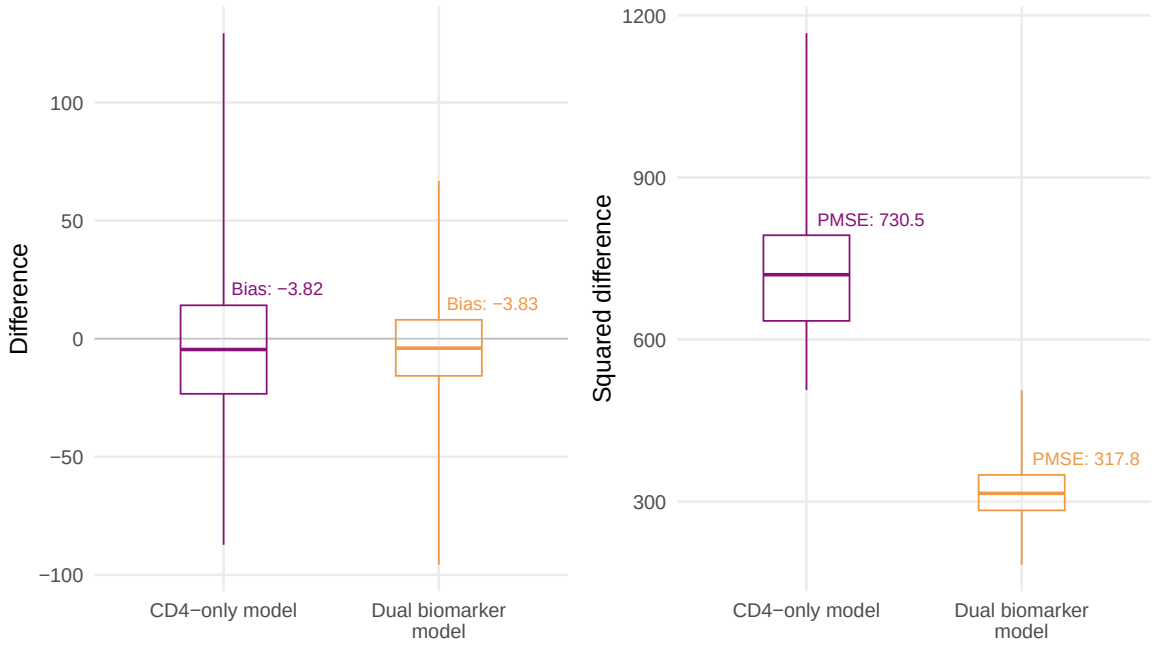


Figure C.10: Distributions of differences (panel A), and squared differences (panel B) between parameter value and estimated HIV incidence, by back-calculation model. Estimates from 100 model fits. 40% of records in simulated data missing CD4 and RITA information. Boxplots show median, inter-quartile range, and range of estimates for each model.

D. Appendix to Chapter 4

D.1 Pre-vaccine hospitalised severity

Table D.1: Baseline characteristics of patients with COVID-19 admitted to sentinel NHS trusts between 15 March 2020 to 28 February 2021, by outcome following hospital admission.

Characteristic	**ICU** N = 2,594	**Death** N = 5,800	**Discharge** N = 16,057	**Unknown** N = 3,912	**Overall** N = 28,363
Sex					
Female	810 (6.3%)	2,446 (18.9%)	7,777 (60.2%)	1,886 (14.6%)	12,919
Male	1,784 (11.6%)	3,354 (21.7%)	8,280 (53.6%)	2,026 (13.1%)	15,444
Month of admission					
Mar 2020	483 (14.3%)	849 (25.1%)	1,866 (55.1%)	189 (5.6%)	3,387
Apr 2020	735 (10.6%)	1,697 (24.5%)	4,020 (58.1%)	470 (6.8%)	6,922
May 2020	134 (6.1%)	399 (18.2%)	1,409 (64.4%)	247 (11.3%)	2,189
Jun-Aug 2020	77 (6.0%)	125 (9.7%)	923 (71.5%)	166 (12.9%)	1,291
Sep 2020	43 (10.2%)	75 (17.8%)	234 (55.5%)	70 (16.6%)	422
Oct 2020	184 (10.4%)	349 (19.7%)	960 (54.1%)	282 (15.9%)	1,775
Nov 2020	204 (8.6%)	481 (20.2%)	1,428 (60.0%)	266 (11.2%)	2,379
Dec 2020	197 (7.8%)	510 (20.2%)	1,405 (55.7%)	410 (16.3%)	2,522
Jan 2021	406 (7.5%)	1,053 (19.6%)	2,796 (52.0%)	1,123 (20.9%)	5,378
Feb 2021	131 (6.2%)	262 (12.5%)	1,016 (48.4%)	689 (32.8%)	2,098
Age group					
15–45	319 (8.5%)	49 (1.3%)	2,883 (77.1%)	486 (13.0%)	3,737
45–65	1,319 (16.5%)	477 (6.0%)	5,201 (65.0%)	1,001 (12.5%)	7,998
65–75	651 (12.6%)	984 (19.0%)	2,801 (54.2%)	731 (14.1%)	5,167
75+	305 (2.7%)	4,290 (37.4%)	5,172 (45.1%)	1,694 (14.8%)	11,461
Ethnicity					
White	1,676 (7.7%)	4,923 (22.7%)	11,890 (54.7%)	3,243 (14.9%)	21,732
Asian	395 (14.2%)	406 (14.6%)	1,741 (62.4%)	247 (8.9%)	2,789
Black	210 (12.8%)	238 (14.5%)	1,049 (64.1%)	140 (8.6%)	1,637
Mixed/Other	162 (15.5%)	104 (9.9%)	667 (63.7%)	114 (10.9%)	1,047
Unreported	151 (13.0%)	129 (11.1%)	710 (61.3%)	168 (14.5%)	1,158
Region					
London/South of England	1,079 (10.2%)	1,841 (17.3%)	6,702 (63.1%)	997 (9.4%)	10,619

Midlands and East of England	913 (8.3%)	2,463 (22.5%)	6,050 (55.3%)	1,517 (13.9%)	10,943
North of England	602 (8.9%)	1,496 (22.0%)	3,305 (48.6%)	1,398 (20.6%)	6,801
Number of comorbidities					
0	681 (6.6%)	1,561 (15.2%)	6,177 (60.1%)	1,860 (18.1%)	10,279
1	541 (10.1%)	931 (17.5%)	3,186 (59.8%)	673 (12.6%)	5,331
2	530 (10.9%)	1,089 (22.5%)	2,632 (54.3%)	599 (12.4%)	4,850
3+	842 (10.7%)	2,219 (28.1%)	4,062 (51.4%)	780 (9.9%)	7,903

D.1.1 Goodness of fit

For the SARI-Watch data, goodness of fit of the parametric mixture models was assessed by comparing non-parametric Aalen-Johansen cumulative incidence curves to estimates derived from the parametric model. A log-normal distribution was used for the hospital to ICU transition, and generalized gamma distributions used for other transitions. Figures D.1 and D.2 show goodness of fit for each of the sub-models, by month of admission (panel A) and month of admission and sex (panel B).

D.1.2 Outcomes following ICU admission

The estimated probability of death among patients admitted to critical care by month of admission, sex, age group, and number of comorbidities is shown in Figure D.3, panels A–D. The probability of death following ICU admission fell from 48.6% (95% CI 44.2 to 52.3%) in March 2020 to 24.1% (95% CI 15.8 to 34.6%) in June-August 2020, increased to 49.6% (95% CI 42.2 to 56.3%) in October and plateaued at around 37% between November and December 2021. The probability of death following ICU admission was estimated to be greater among older patients compared to younger patients, although confidence intervals are wide for these smaller sub-groups.

Figure D.4, panels A–D, shows lengths of stay in ICU by month of admission, sex, age group, and number of comorbidities. Compared to non-critical care patients, those admitted to critical care spent longer in hospital. Median length of stay in ICU prior to discharge initially reduced from 24.0 days (95% CI 20.5 to 28.1 days) in March 2020 to 13.4 days (95% CI 9.8 to 18.0 days) in June-August 2020, then increased to 19.1 days (95% CI 15.3 to 23.2 days) in December 2020. Median length of stay in ICU prior to death showed less variability over time, ranging between 9 and 14 days.

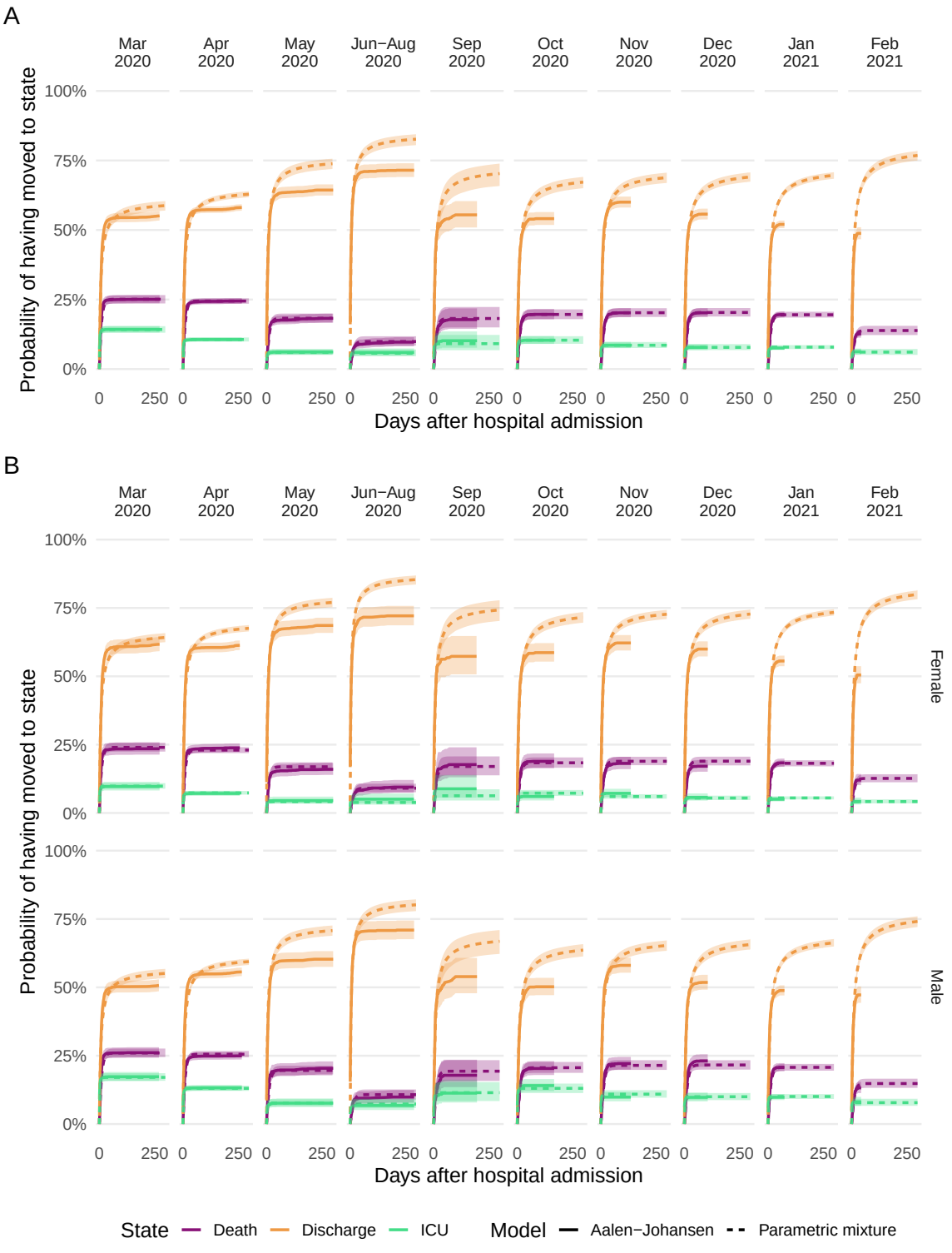


Figure D.1: Goodness of fit for sub-model (i) from Hospital, by month of admission (panel A), and sex (panel B) in SARI-Watch data, March 2020 to February 2021. Solid lines are Aalen-Johansen cumulative incidence curves, dashed lines are parametric mixture model estimates. Error bands are 95% CIs, derived from 1000 simulation replicates.

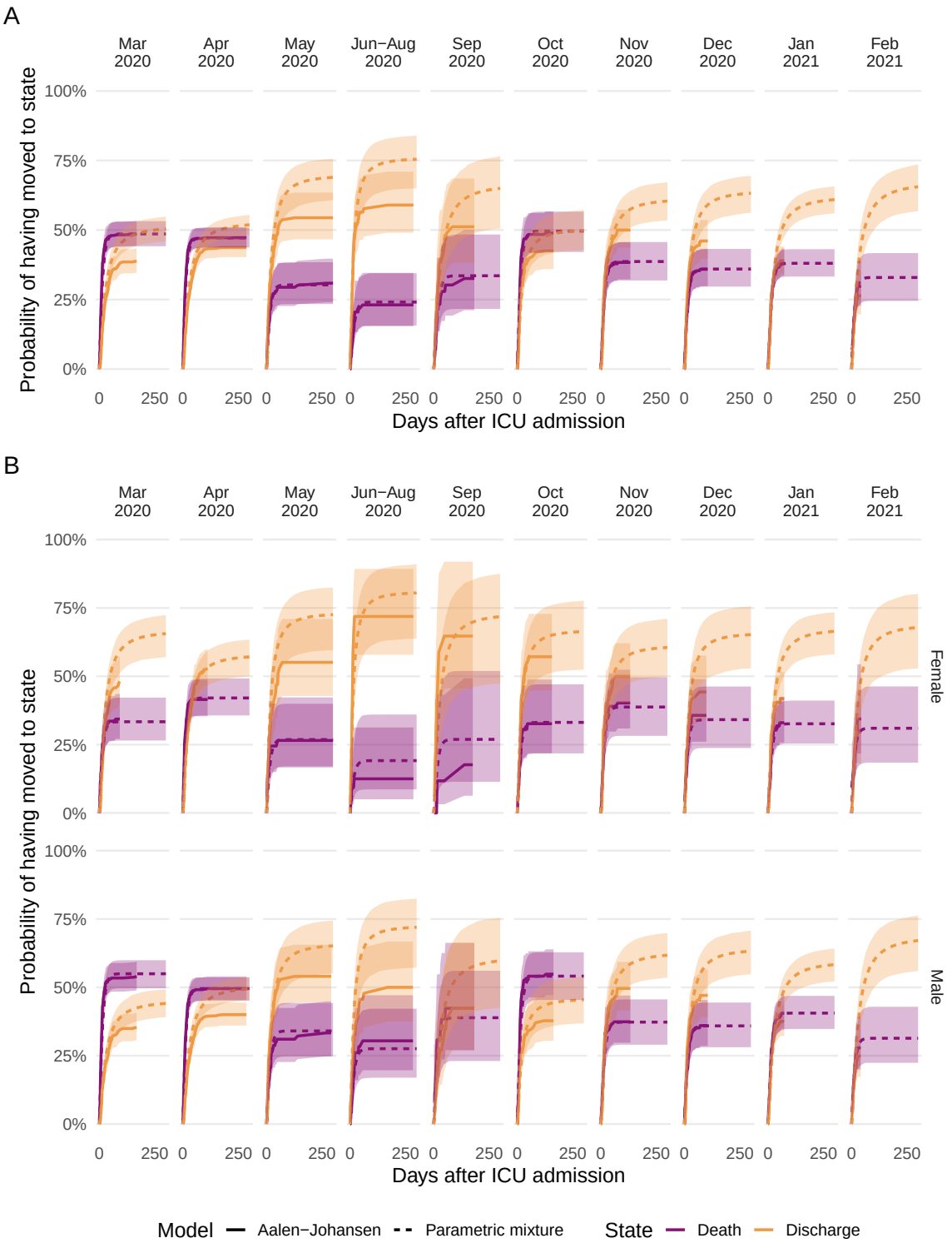


Figure D.2: Goodness of fit for sub-model (ii) from ICU, by month of admission (panel A), and sex (panel B) in SARI-Watch data, March 2020 to February 2021. Solid lines are Aalen-Johansen cumulative incidence curves, dashed lines are parametric mixture model estimates. Error bands are 95% CIs, derived from 1000 simulation replicates.

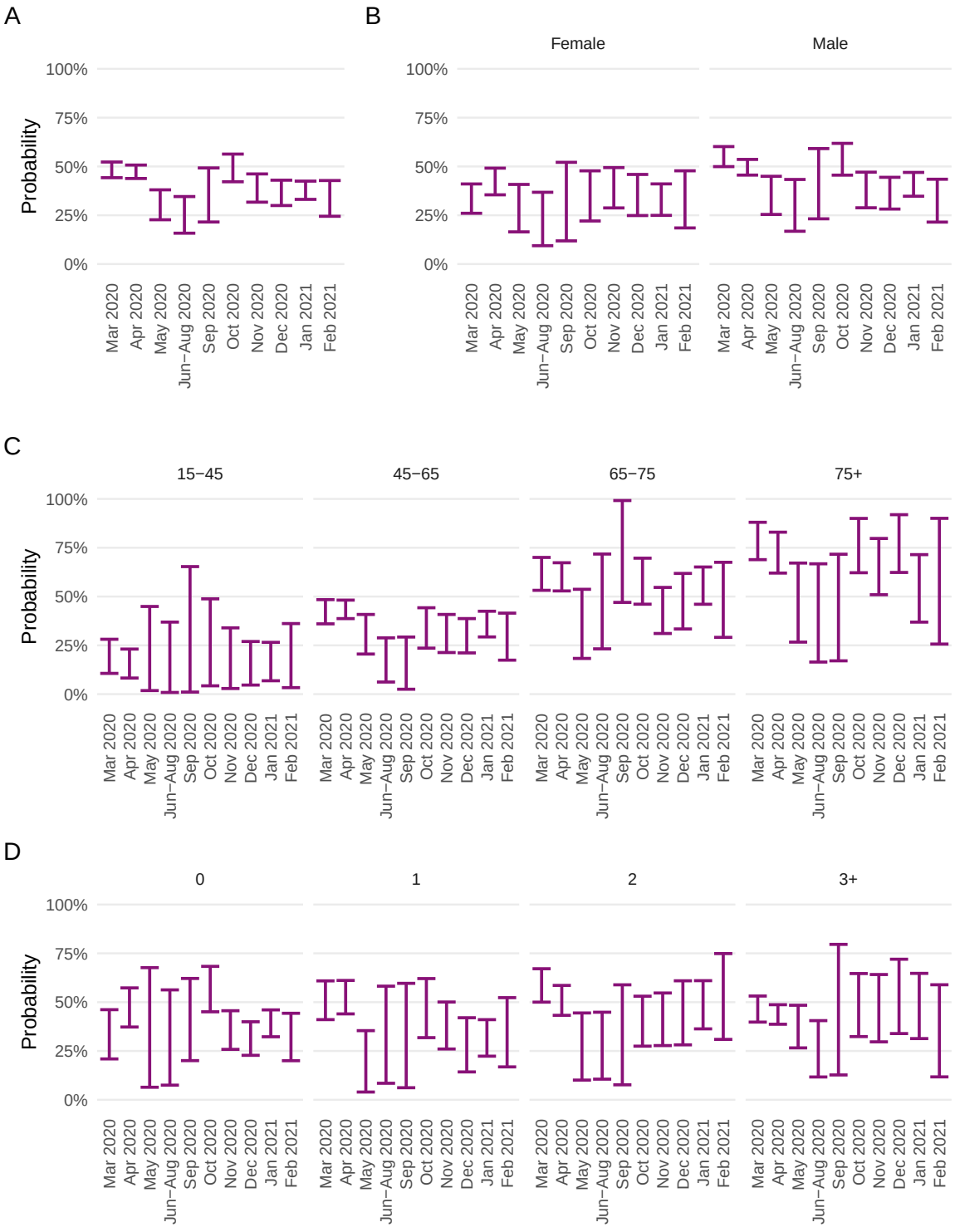


Figure D.3: Probability of death following ICU admission, by month of admission (panel A), sex (panel B), age group (panel C), and number of comorbidities (panel D) in SARI-Watch data, March 2020 to February 2021. Error bands are 95% CIs to represent uncertainty in the estimated probability.

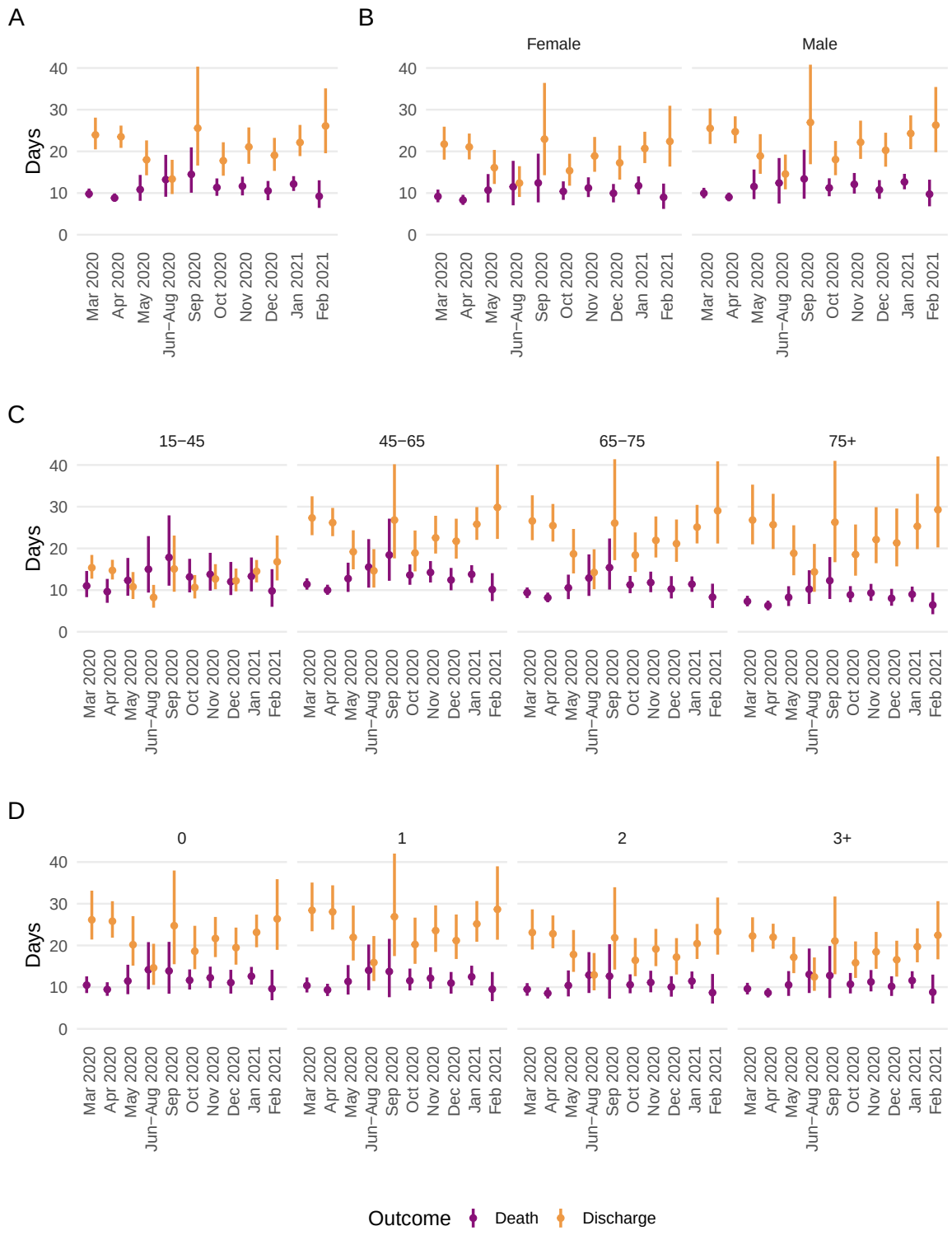


Figure D.4: Estimated median time to event following ICU admission in SARI-Watch data, by month of admission (panel A), sex (panel B), age group (panel C), and number of comorbidities (panel D), March 2020 to February 2021. Line ranges are 95% CIs around the median.

D.2 Post-vaccine hospitalised severity

D.2.1 Weighted median with weighted ties

For the n elements in $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ with positive weights $\boldsymbol{w} = (w_1, w_2, \ldots, w_n)$ and $S = \sum_n w_n$, the weighted median is the element x_k where (Cormen and Leiserson 2022):

$$\sum_{i<k} w_i \leq \frac{S}{2}$$
$$\sum_{i>k} w_i \leq \frac{S}{2}$$

When ties occur between two values x_k which satisfy the weighted median criteria, the median is a weighted average of the two values, weighted by the sum of weights for all values $x_i \leq x_k$, and $x_i \geq x_k$, respectively.

Table D.2: Characteristics of the study population compared with all people hospital-onset COVID-19, and all people with PCR-confirmed community-acquired COVID-19 in England.

Characteristic	**Study population (hospitalised for COVID-19 in England)** N = 590,313	**All people with hospital-onset COVID-19 in England** N = 209,139	**All people with PCR-confirmed, community-acquired COVID-19 in England** N = 6,617,362
Age group			
0–14	19,186 (3.3%)	2,635 (1.3%)	892,380 (13%)
15–24	28,936 (4.9%)	4,437 (2.1%)	1,264,528 (19%)
25–44	122,390 (21%)	13,764 (6.6%)	2,244,393 (34%)
45–64	154,594 (26%)	33,209 (16%)	1,560,517 (24%)
65–74	90,037 (15%)	36,537 (17%)	301,440 (4.6%)
75–84	101,578 (17%)	59,892 (29%)	198,209 (3.0%)
85+	73,592 (12%)	58,665 (28%)	155,895 (2.4%)
Sex			
Male	289,085 (49%)	108,709 (52%)	3,152,281 (48%)
Female	301,228 (51%)	100,430 (48%)	3,465,081 (52%)
Ethnicity			
White	444,615 (75%)	185,426 (89%)	5,100,779 (77%)
Asian	58,411 (9.9%)	9,877 (4.7%)	740,301 (11%)
Black	33,961 (5.8%)	6,362 (3.0%)	278,844 (4.2%)
Mixed/Other/Unknown	41,386 (7.0%)	7,474 (3.6%)	497,438 (7.5%)
Prefer not to say	11,940 (2.0%)	0 (0%)	0 (0%)
Region of residence			
London	104,850 (18%)	28,779 (14%)	1,078,524 (16%)
East Midlands	51,147 (8.7%)	18,774 (9.0%)	586,490 (8.9%)

East of England	59,735 (10%)	22,809 (11%)	679,080 (10%)
North East	33,809 (5.7%)	9,416 (4.5%)	376,501 (5.7%)
North West	94,658 (16%)	42,276 (20%)	1,056,101 (16%)
South East	76,913 (13%)	28,819 (14%)	887,347 (13%)
South West	48,448 (8.2%)	12,880 (6.2%)	496,337 (7.5%)
West Midlands	66,495 (11%)	24,544 (12%)	733,204 (11%)
Yorkshire and Humber	54,258 (9.2%)	20,842 (10.0%)	723,778 (11%)
Index of Multiple Deprivation			
Most deprived	159,904 (27%)	51,314 (25%)	1,564,566 (24%)
2nd quintile	133,835 (23%)	44,996 (22%)	1,440,738 (22%)
3rd quintile	111,840 (19%)	40,584 (19%)	1,288,334 (19%)
4th quintile	99,313 (17%)	38,418 (18%)	1,210,037 (18%)
Least deprived	85,421 (14%)	33,827 (16%)	1,113,687 (17%)
Hospital outcome			
Death	81,343 (14%)	69,469 (33%)	-
Discharge	507,254 (86%)	107,878 (52%)	-
Right-censored in hospital	1,716 (0.3%)	31,792 (15%)	-
Month of hospital admission			
Mar 2020	18,285 (3.1%)	-	-
Apr 2020	36,524 (6.2%)	-	-
May 2020	10,471 (1.8%)	-	-
Jun 2020	3,826 (0.6%)	-	-
Jul 2020	1,506 (0.3%)	-	-
Aug 2020	1,276 (0.2%)	-	-
Sep 2020	5,175 (0.9%)	-	-
Oct 2020	19,525 (3.3%)	-	-
Nov 2020	29,932 (5.1%)	-	-
Dec 2020	40,657 (6.9%)	-	-
Jan 2021	77,186 (13%)	-	-
Feb 2021	25,488 (4.3%)	-	-
Mar 2021	7,941 (1.3%)	-	-
Apr 2021	3,100 (0.5%)	-	-
May 2021	2,260 (0.4%)	-	-
Jun 2021	5,597 (0.9%)	-	-
Jul 2021	20,340 (3.4%)	-	-
Aug 2021	22,388 (3.8%)	-	-
Sep 2021	19,510 (3.3%)	-	-
Oct 2021	21,974 (3.7%)	-	-
Nov 2021	20,707 (3.5%)	-	-
Dec 2021	31,716 (5.4%)	-	-
Jan 2022	54,310 (9.2%)	-	-
Feb 2022	28,412 (4.8%)	-	-
Mar 2022	42,885 (7.3%)	-	-
Apr 2022	39,322 (6.7%)	-	-
Vaccination status at date of admission (January 2021 onwards)			

Unvaccinated	367,082 (62%)	-	-
<21 days after first dose	14,385 (2.4%)	-	-
≤21 days after first dose	20,711 (3.5%)	-	-
≥14 days after second dose	119,559 (20%)	-	-
≥14 days after third dose	68,576 (12%)	-	-
Charlson comorbidity index			
0	232,646 (39%)	-	-
1-2	227,174 (38%)	-	-
3-4	85,131 (14%)	-	-
5+	45,362 (7.7%)	-	-
Hospital load at time of admission (as proportion of busiest week)			
0-20%	120,259 (20%)	-	-
20-40%	192,060 (33%)	-	-
40-60%	123,943 (21%)	-	-
60-80%	80,567 (14%)	-	-
80-90%	34,308 (5.8%)	-	-
90-100%	39,176 (6.6%)	-	-

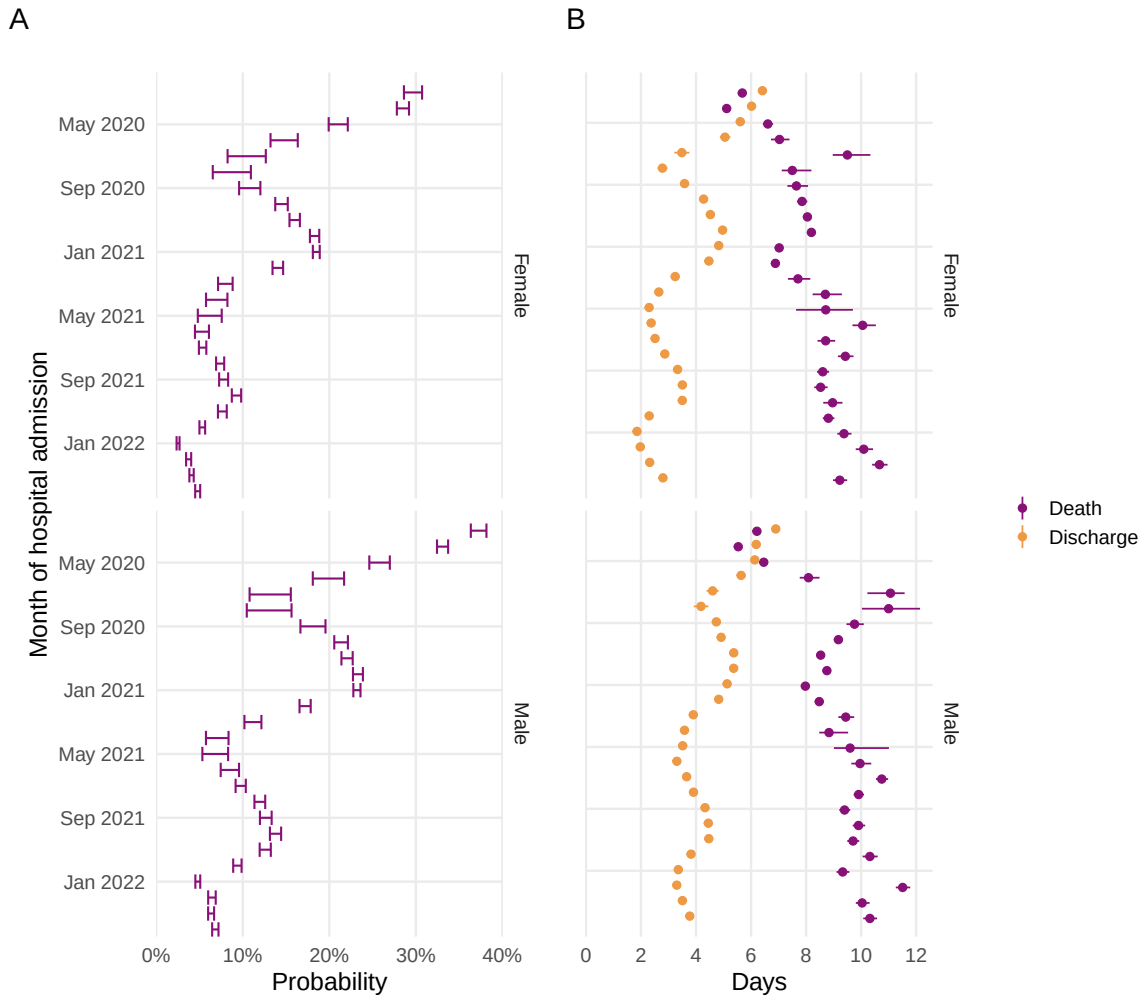


Figure D.5: Hospitalised fatality risk (panel A) and median length of stay (panel B) by month of admission and sex in SUS data, March 2020 to April 2022. Unadjusted for other covariates. Error bars and line ranges are 95% CIs.

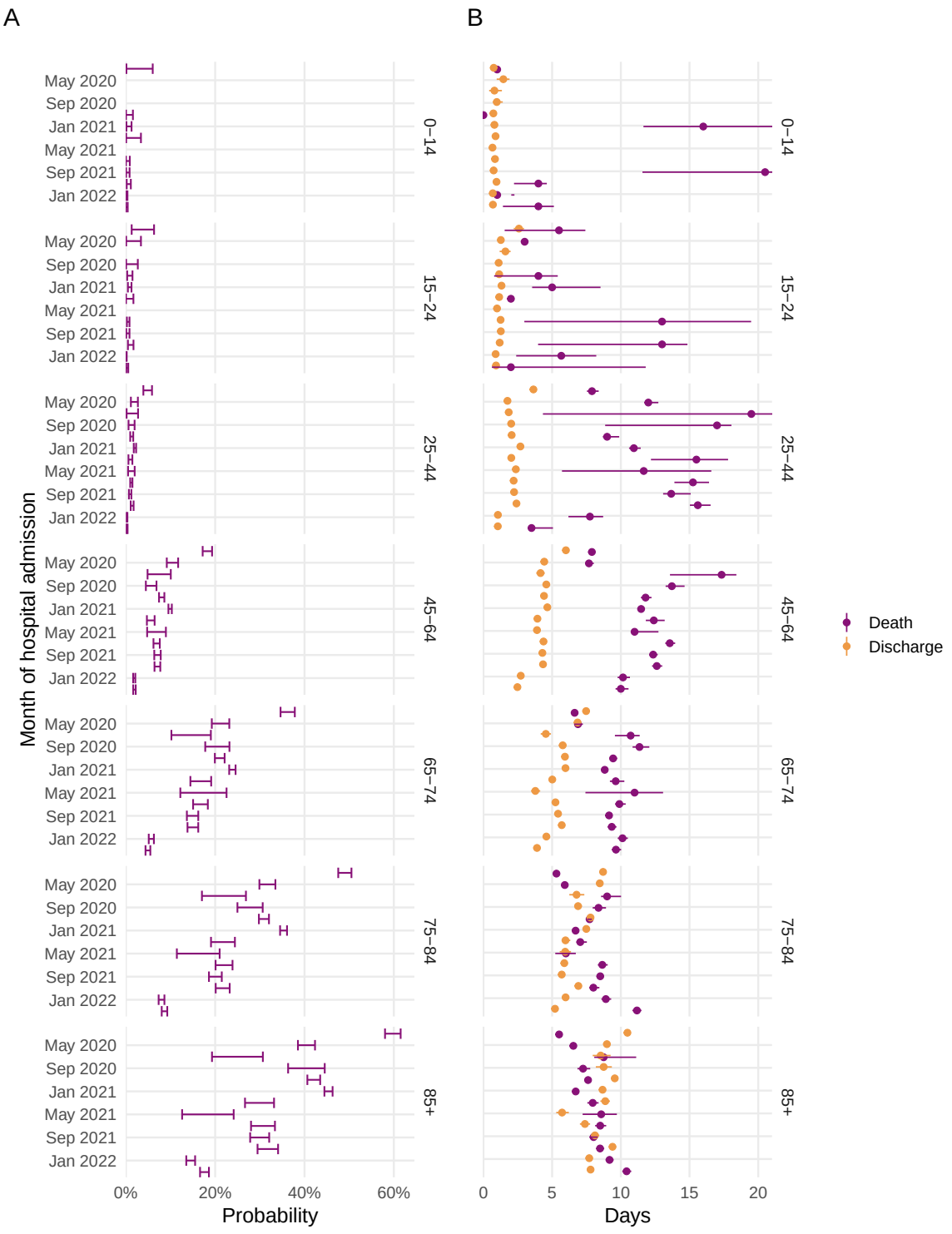


Figure D.6: Hospitalised fatality risk (panel A) and median length of stay (panel B) by month of admission and age group in SUS data, March 2020 to April 2022. Unadjusted for other covariates. Error bars and line ranges are 95% CIs.

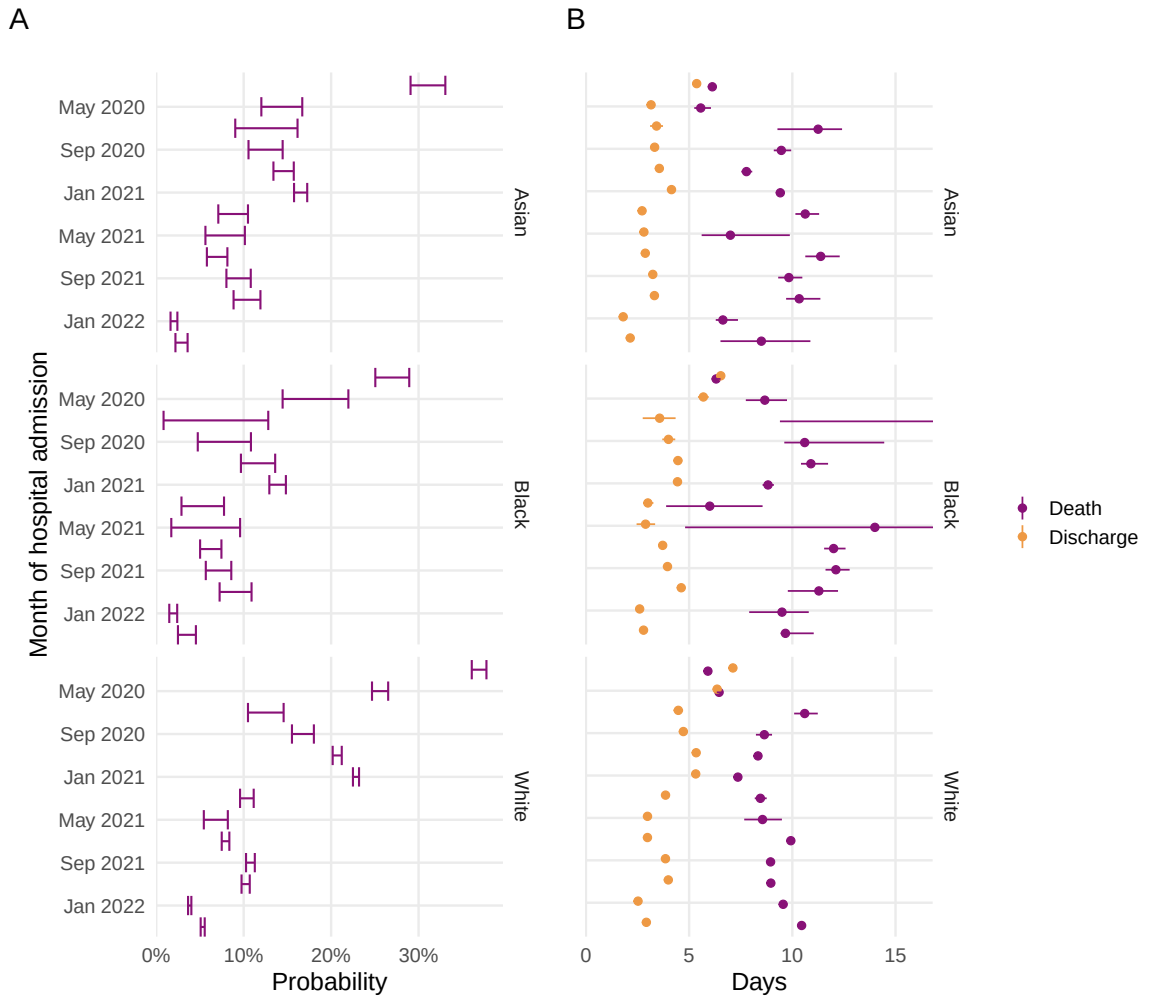


Figure D.7: Hospitalised fatality risk (panel A) and median length of stay (panel B) by month of admission and ethnicity in SUS data, March 2020 to April 2022. Unadjusted for other covariates. Error bars and line ranges are 95% CIs.

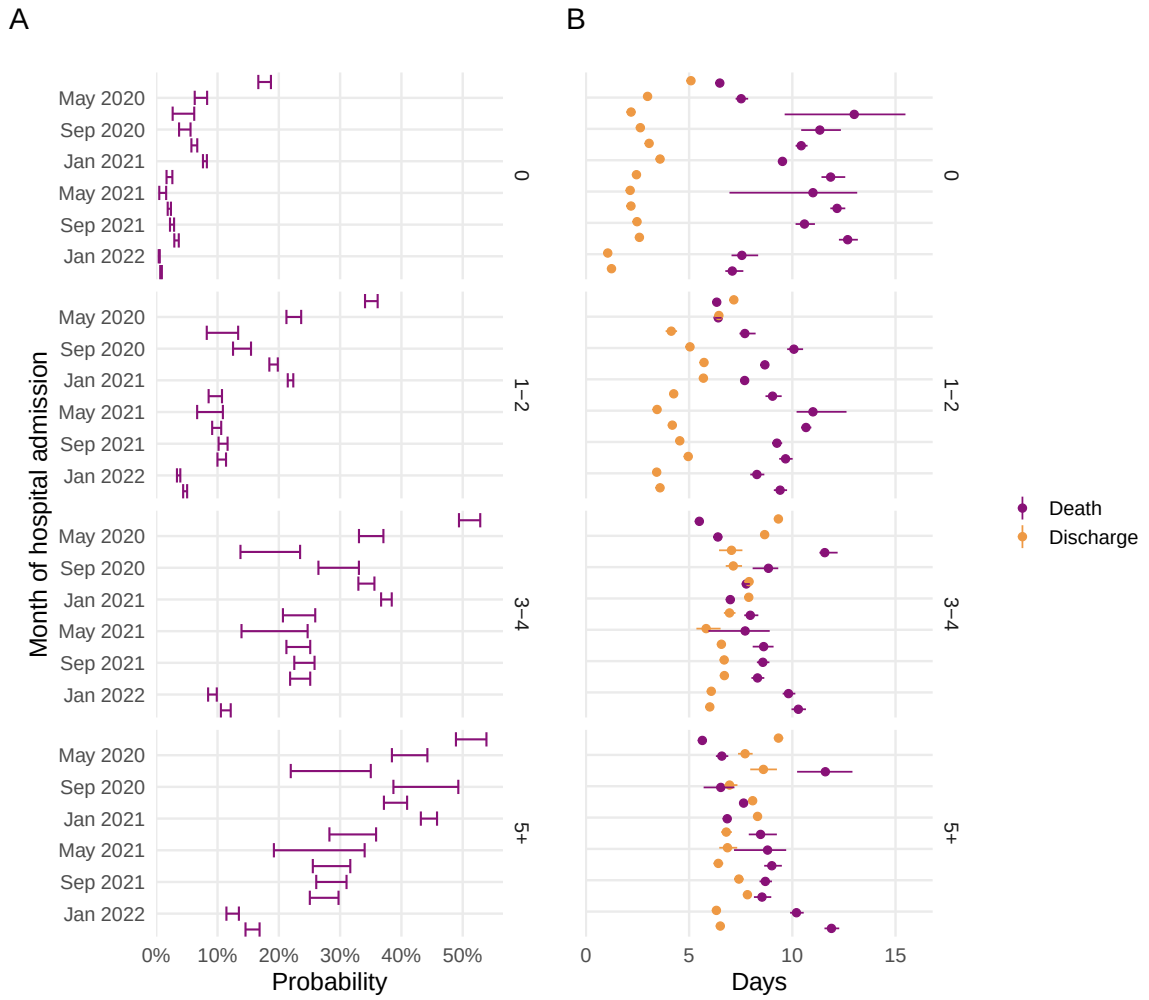


Figure D.8: Hospitalised fatality risk (panel A) and median length of stay (panel B) by month of admission and CCI in SUS data, March 2020 to April 2022. Unadjusted for other covariates. Error bars and line ranges are 95% CIs.

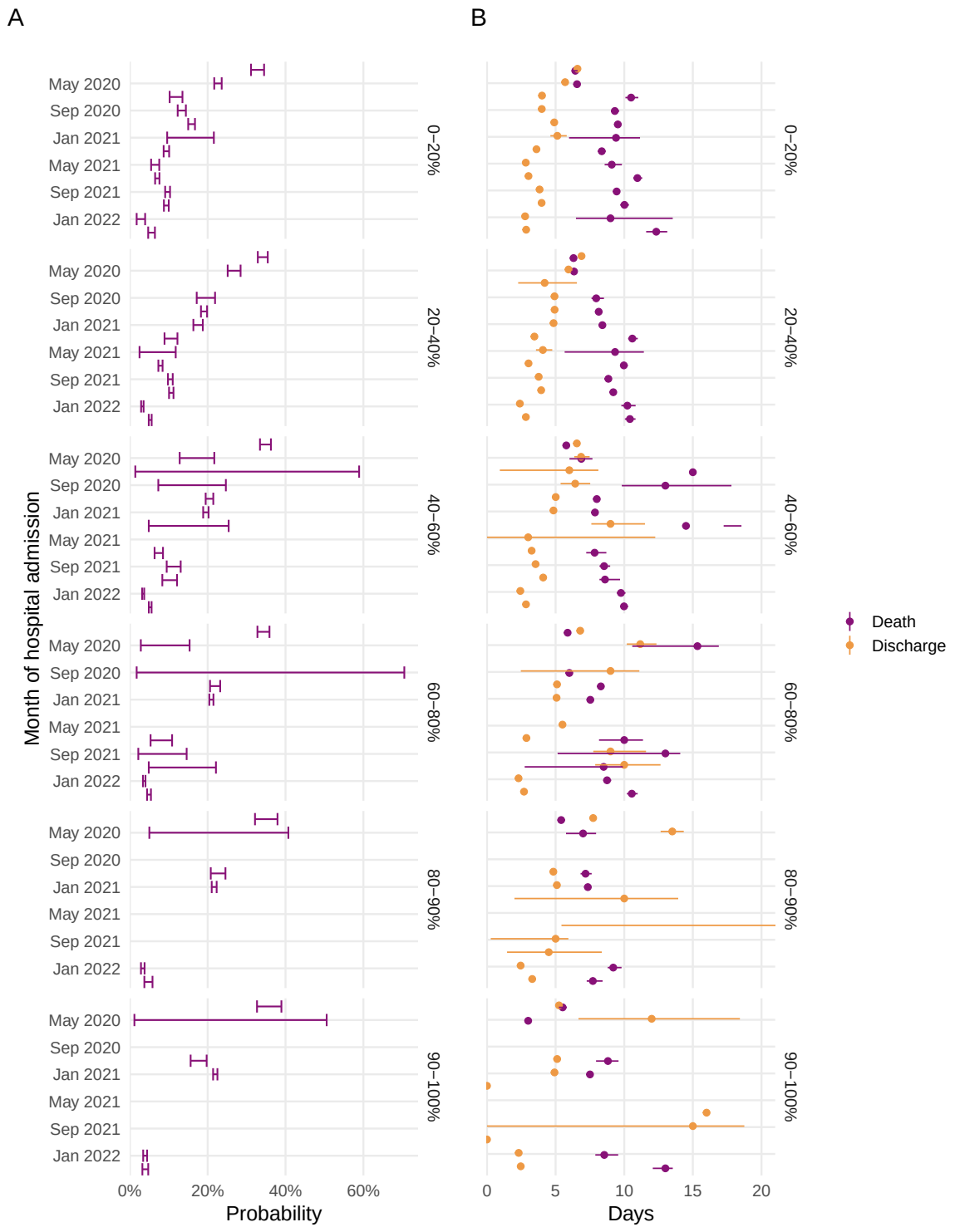


Figure D.9: Hospitalised fatality risk (panel A) and median length of stay (panel B) by month of admission and hospital load in SUS data, March 2020 to April 2022. Unadjusted for other covariates. Error bars and line ranges are 95% CIs.

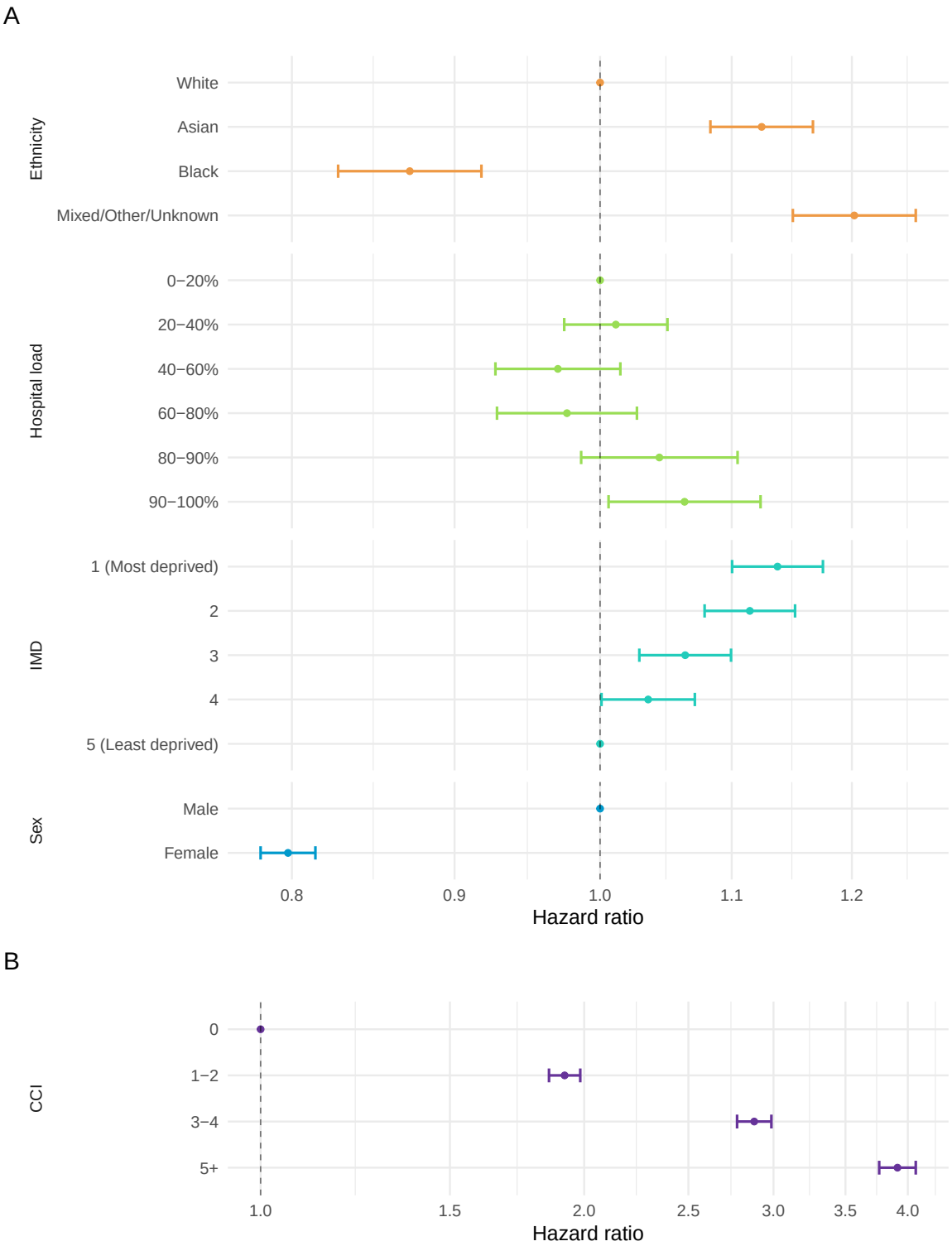


Figure D.10: Hospitalised fatality sub-distribution hazard ratios for sex, ethnicity, IMD quintile, hospital load and CCI in SUS data, March 2020 to April 2022. Figure shows point estimate of hazard ratio with 95% CIs.

D.2.2 Epidemic phase bias

Figure D.11 shows the results of sensitivity analyses where the date of symptom onset was shifted later in time by $c = 0, 1, 2, 3, 4$ days for those who died, but not for those who were discharged. This model included stratification on age group, region of residence, and vaccination status, and regression adjustment on month of hospital admission, sex, ethnicity, IMD quintile, hospital load, and CCI.

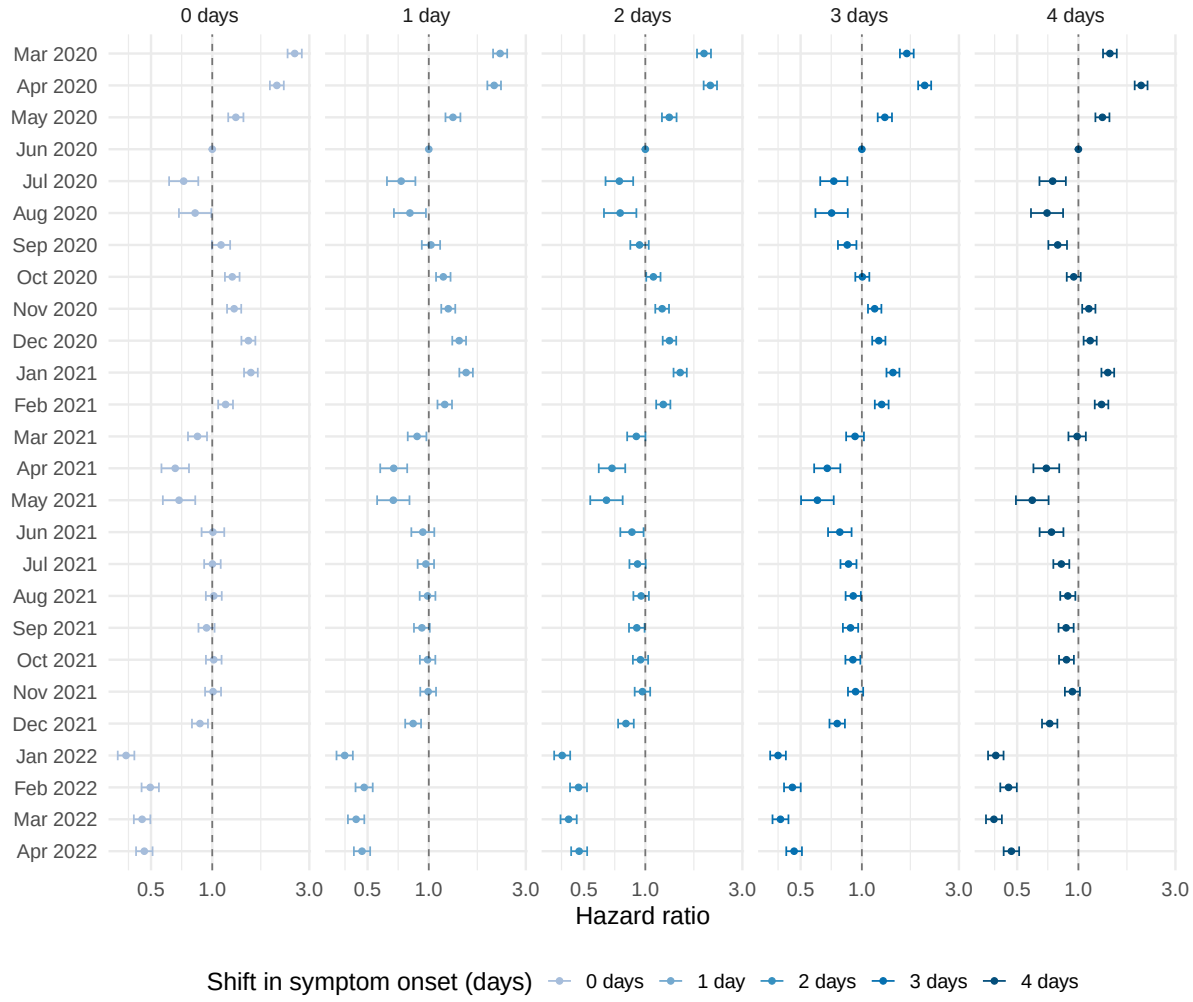


Figure D.11: Hospitalised fatality sub-distribution hazard ratio by month of symptom onset and sensitivity in SUS data, March 2020 to April 2022. Reference group: June 2020. Figure shows point estimate of hazard ratio with 95% CIs.

E. Appendix to Chapter 5

Table E.1: Demographic characteristics of SIREN study participants, by recruitment status.

Characteristic	**Initial cohort** N = 25,717	**First extension** N = 10,356	**Second extension** N = 7,923
Gender			
Male	4,358 (16.9%)	1,613 (15.6%)	1,208 (15.2%)
Female	21,359 (83.1%)	8,743 (84.4%)	6,715 (84.8%)
Age group			
<25	1,473 (5.7%)	216 (2.1%)	70 (0.9%)
25-34	6,666 (25.9%)	1,622 (15.7%)	717 (9.0%)
35-44	6,500 (25.3%)	2,599 (25.1%)	1,737 (21.9%)
45-54	6,661 (25.9%)	3,477 (33.6%)	2,996 (37.8%)
55-64	4,088 (15.9%)	2,283 (22.0%)	2,205 (27.8%)
65+	329 (1.3%)	159 (1.5%)	198 (2.5%)
Ethnicity			
White	22,343 (86.9%)	9,222 (89.0%)	7,079 (89.3%)
Asian	2,108 (8.2%)	671 (6.5%)	457 (5.8%)
Black	512 (2.0%)	189 (1.8%)	182 (2.3%)
Mixed	418 (1.6%)	142 (1.4%)	118 (1.5%)
Other	336 (1.3%)	132 (1.3%)	87 (1.1%)
Medical condition			
No medical conditions	19,400 (75.4%)	7,762 (75.0%)	5,761 (72.7%)
Immunosuppression	552 (2.1%)	231 (2.2%)	186 (2.3%)
Chronic respiratory condition	3,164 (12.3%)	1,250 (12.1%)	1,032 (13.0%)
Chronic non-respiratory condition	2,601 (10.1%)	1,113 (10.7%)	944 (11.9%)
Staff type			
Administrative/Executive (office based)	3,678 (14.3%)	1,608 (15.5%)	1,367 (17.3%)
Doctor	3,042 (11.8%)	1,234 (11.9%)	908 (11.5%)
Nursing	8,714 (33.9%)	3,450 (33.3%)	2,694 (34.0%)
Healthcare Assistant	2,400 (9.3%)	698 (6.7%)	541 (6.8%)
Midwife	562 (2.2%)	233 (2.2%)	157 (2.0%)

Healthcare Scientist	891 (3.5%)	468 (4.5%)	307 (3.9%)
Pharmacist	515 (2.0%)	261 (2.5%)	151 (1.9%)
Physiotherapist/Occupational Therapist/SALT	1,074 (4.2%)	466 (4.5%)	270 (3.4%)
Student (Medical/ Nursing/Midwifery/Other)	927 (3.6%)	262 (2.5%)	312 (3.9%)
Estates/Porters/Security	400 (1.6%)	192 (1.9%)	109 (1.4%)
Other (non-patient facing)	390 (1.5%)	181 (1.7%)	138 (1.7%)
Other (patient facing)	3,124 (12.1%)	1,303 (12.6%)	969 (12.2%)
Occupation setting			
Ambulance/Emergency department/Inpatient wards	6 (0.0%)	635 (6.1%)	221 (2.8%)
Intensive care	1,267 (4.9%)	2,056 (19.9%)	904 (11.4%)
Maternity/Labour ward	326 (1.3%)	353 (3.4%)	154 (1.9%)
Office	4,644 (18.1%)	414 (4.0%)	951 (12.0%)
Outpatient	5,062 (19.7%)	4,657 (45.0%)	2,774 (35.0%)
Patient-facing (non-clinical)	996 (3.9%)	557 (5.4%)	388 (4.9%)
Theatres	646 (2.5%)	65 (0.6%)	80 (1.0%)
Other	12,770 (49.7%)	1,619 (15.6%)	2,451 (30.9%)
Patient contact			
Yes	22,259 (86.6%)	8,797 (84.9%)	6,677 (84.3%)
No	3,458 (13.4%)	1,559 (15.1%)	1,246 (15.7%)
Index of multiple deprivation			
Most deprived (1)	2,982 (12.0%)	1,117 (11.4%)	728 (9.5%)
Deprivation 2	4,430 (17.8%)	1,699 (17.4%)	1,369 (17.8%)
Deprivation 3	5,509 (22.1%)	2,057 (21.1%)	1,719 (22.4%)
Deprivation 4	5,849 (23.5%)	2,302 (23.6%)	1,767 (23.0%)
Least deprived (5)	6,155 (24.7%)	2,585 (26.5%)	2,104 (27.4%)
Unknown	792	596	236
Region			
East Midlands	2,124 (8.3%)	637 (6.2%)	666 (8.4%)
East of England	2,079 (8.1%)	989 (9.6%)	1,010 (12.7%)
London	2,522 (9.8%)	1,307 (12.6%)	1,183 (14.9%)
North East	280 (1.1%)	369 (3.6%)	128 (1.6%)
North West	2,332 (9.1%)	1,331 (12.9%)	772 (9.7%)
Northern Ireland	499 (1.9%)	557 (5.4%)	198 (2.5%)
Scotland	3,641 (14.2%)	2,153 (20.8%)	417 (5.3%)
South East	2,746 (10.7%)	539 (5.2%)	1,071 (13.5%)
South West	4,792 (18.6%)	797 (7.7%)	1,157 (14.6%)
Wales	632 (2.5%)	282 (2.7%)	154 (1.9%)

West Midlands	2,028 (7.9%)	684 (6.6%)	602 (7.6%)
Yorkshire and the Humber	2,042 (7.9%)	711 (6.9%)	565 (7.1%)
Household structure			
Lives with others (including children)	10,529 (40.9%)	4,100 (39.6%)	2,960 (37.4%)
Lives with others (no children)	12,737 (49.5%)	5,100 (49.2%)	4,045 (51.1%)
Lives alone	2,451 (9.5%)	1,156 (11.2%)	918 (11.6%)
Vaccination status at study exit			
Unvaccinated	1,643 (6.4%)	143 (1.4%)	94 (1.2%)
First and second dose	15,503 (60.3%)	614 (5.9%)	359 (4.5%)
Third dose	8,571 (33.3%)	6,846 (66.1%)	1,847 (23.3%)
Fourth dose	0 (0.0%)	2,753 (26.6%)	5,623 (71.0%)

E.1 Cox proportional hazards models

Stratified Cox proportional hazards models (described in Section 2.2.4) were applied to the SIREN study data in two 'interim analyses'. These models included stratification on the age group, household status, and occupation/setting covariates, with the robust variance estimator used to account for correlation between individuals within sites.

The `R` package `survival` was used to implement the Cox proportional hazards models (Therneau 1999). Hazard ratios (HR) were converted into vaccine effectiveness (VE) and relative protection estimates using the formula: VE = 1 – HR.

E.1.1 Vaccine effectiveness by vaccine type

Cox proportional hazards models were first applied to estimate the duration and effectiveness of immunity from prior infection and vaccination with two vaccine doses during the Alpha and Delta variant circulating period (Hall, Foulkes, Insalata *et al.* 2022). A total of 35,768 participants were included in this analysis, of whom 27% had a previous detected SARS-CoV-2 infection. Vaccine coverage was high: 78% received two BNT162b2 mRNA vaccines with a long interval between doses, 9% two BNT162b2 mRNA vaccines with a short interval between doses, and 8% two ChAdOX1 adenoviral vaccines. Between December 2020 and September 2021 there were 2747 primary infections and 210 reinfections detected among this cohort.

For the BNT162b2 vaccine with a long dosing interval, estimated vaccine effectiveness dropped from 85% (95% CI 72 to 92%) within 14-to-73 days post-second dose to 51% (95% CI 22 to 69%) at around 6 months post-second dose. There was no significant difference in vaccine effectiveness between long and short dosing intervals for the BNT162b2 vaccine. However, vaccine effectiveness for the ChAdOX1 vaccine was estimated as 58% (95% CI 23 to 77) at

14-to-73 days post-second dose (Figure E.1). Immunity from previous infection waned after one year in unvaccinated individuals but persisted at above 90% among those who had been vaccinated after infection, even if the infection had occurred >18 months prior.

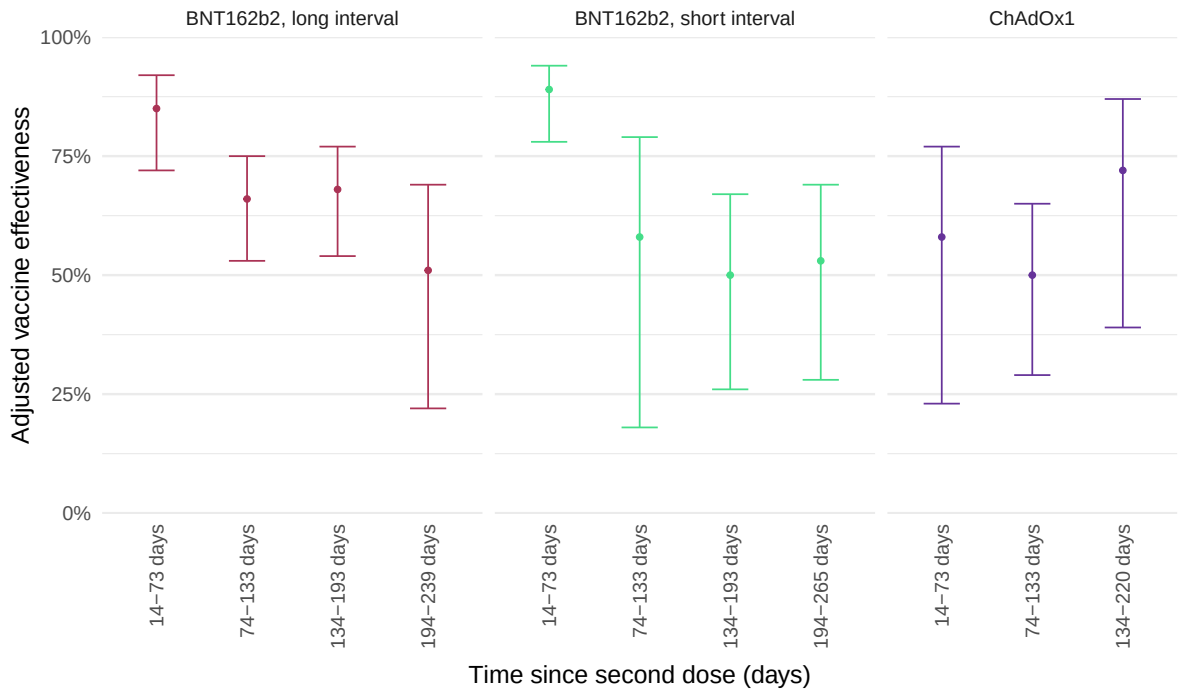


Figure E.1: Adjusted vaccine effectiveness among SIREN participants with no prior infection during the Alpha and Delta variant circulating period, by vaccine type and dosing interval.

E.1.2 Vaccine effectiveness against Omicron

Cox proportional hazards models were next applied to estimate the effectiveness of third BNT162b2 vaccine doses and previous infection against the Omicron BA.1 and BA.2 sub-variants compared to Delta (Hall, Insalata *et al.* 2024). This analysis included 19,614 participants, with 29% having been previously infected. Between September 2021 and February 2022, 95% received a third vaccine dose and there were 2745 primary infections and 966 reinfections detected during the 6-month follow-up period.

Third dose vaccine effectiveness against the Delta variant for participants who had initially received the BNT162b2 vaccine was 63% (95% CI 40 to 77%) within 0-to-2 months post-third dose, compared to 35% (95% CI 21 to 47%) against the Omicron BA.1 and BA.2 sub-variants. Participants who had initially received the ChAdOX1 vaccine had a median vaccine effectiveness ≥68% against both variants (Figure E.2 panel A). Third dose protection waned rapidly against the Omicron variant, however, with only around 4 months of additional protection conferred by vaccination. Previous infection provided longer-lasting protection against Omicron, 67% (95%

CI 56 to 75%) at 3–6 months post-infection, this reduced to 27% (95% CI 4 to 44%) after 9 months, approximately three times lower than against Delta (Figure E.2 panel B).

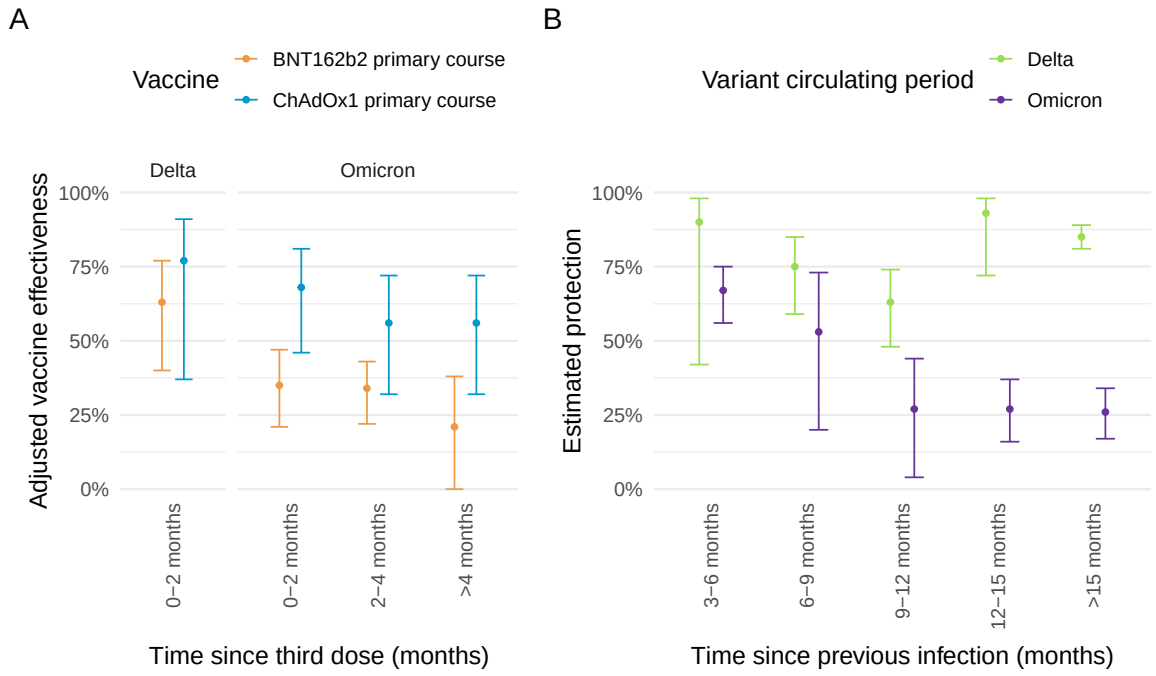


Figure E.2: Third dose vaccine effectiveness, by primary course, variant circulating period, and time since vaccination (panel A) and protection against infection, by time since previous infection and variant circulating period (Panel B).

E.1.3 Discussion

Stratified Cox proportional hazards models were successfully applied to estimate vaccine effectiveness and protection from prior infection in a densely-sampled cohort of healthcare workers with high vaccine coverage.

Two doses of the BNT162b2 vaccine were shown to provide high short-term protection which diminished after six months, whereas protection from two doses of the ChAdOX1 vaccine was substantially lower. Third vaccine doses were found to offer additional short-term protection which again swiftly diminished, particularly against Omicron sub-variants. Meanwhile, infection-acquired immunity (especially when boosted by vaccination) provided longer-lasting protection, although protection was substantially reduced against Omicron as compared to previous variants.

A limitation of the Cox proportional hazards approach under less frequent testing conditions is discussed in Section 5.4, which motivated the development of a multi-state modelling approach.

E.2 Fourth dose vaccine uptake and model results

Table E.2: Demographic characteristics of SIREN study participants enrolled between 12 September 2022 and 31 March 2023, by fourth dose vaccine coverage.

Characteristic	Overall N = 9,560	Fourth dose vaccine coverage	
		Waned third dose N = 2,887	Fourth dose N = 6,673
Gender			
Male	1,524 (15.9%)	394 (13.6%)	1,130 (16.9%)
Female	8,036 (84.1%)	2,493 (86.4%)	5,543 (83.1%)
Age group			
<25	71 (0.7%)	33 (1.1%)	38 (0.6%)
25-34	828 (8.7%)	319 (11.0%)	509 (7.6%)
35-44	2,096 (21.9%)	748 (25.9%)	1,348 (20.2%)
45-54	3,702 (38.7%)	1,067 (37.0%)	2,635 (39.5%)
55-64	2,637 (27.6%)	662 (22.9%)	1,975 (29.6%)
65+	226 (2.4%)	58 (2.0%)	168 (2.5%)
Ethnicity			
White	8,611 (90.1%)	2,522 (87.4%)	6,089 (91.2%)
Asian	543 (5.7%)	189 (6.5%)	354 (5.3%)
Black	179 (1.9%)	91 (3.2%)	88 (1.3%)
Mixed	126 (1.3%)	43 (1.5%)	83 (1.2%)
Other	101 (1.1%)	42 (1.5%)	59 (0.9%)
Medical condition			
No medical conditions	7,038 (73.6%)	2,101 (72.8%)	4,937 (74.0%)
Immunosuppression	247 (2.6%)	135 (4.7%)	112 (1.7%)
Chronic respiratory condition	1,193 (12.5%)	324 (11.2%)	869 (13.0%)
Chronic non-respiratory condition	1,082 (11.3%)	327 (11.3%)	755 (11.3%)
Staff type			
Administrative/Executive (office based)	1,613 (16.9%)	442 (15.3%)	1,171 (17.5%)
Doctor	1,155 (12.1%)	273 (9.5%)	882 (13.2%)
Nursing	3,178 (33.2%)	1,030 (35.7%)	2,148 (32.2%)
Healthcare Assistant	584 (6.1%)	210 (7.3%)	374 (5.6%)
Midwife	197 (2.1%)	74 (2.6%)	123 (1.8%)
Healthcare Scientist	434 (4.5%)	136 (4.7%)	298 (4.5%)
Pharmacist	268 (2.8%)	78 (2.7%)	190 (2.8%)

Physiotherapist/Occupational Therapist/SALT	404 (4.2%)	117 (4.1%)	287 (4.3%)
Student (Medical/ Nursing/Midwifery/Other)	232 (2.4%)	56 (1.9%)	176 (2.6%)
Estates/Porters/Security	185 (1.9%)	66 (2.3%)	119 (1.8%)
Other (non-patient facing)	169 (1.8%)	40 (1.4%)	129 (1.9%)
Other (patient facing)	1,141 (11.9%)	365 (12.6%)	776 (11.6%)
Occupation setting			
Ambulance/Emergency department/Inpatient wards	396 (4.1%)	147 (5.1%)	249 (3.7%)
Intensive care	1,358 (14.2%)	477 (16.5%)	881 (13.2%)
Maternity/Labour ward	303 (3.2%)	96 (3.3%)	207 (3.1%)
Office	829 (8.7%)	188 (6.5%)	641 (9.6%)
Outpatient	3,841 (40.2%)	1,227 (42.5%)	2,614 (39.2%)
Patient-facing (non-clinical)	502 (5.3%)	148 (5.1%)	354 (5.3%)
Theatres	87 (0.9%)	27 (0.9%)	60 (0.9%)
Other	2,244 (23.5%)	577 (20.0%)	1,667 (25.0%)
Patient contact			
Yes	7,992 (83.6%)	2,466 (85.4%)	5,526 (82.8%)
No	1,568 (16.4%)	421 (14.6%)	1,147 (17.2%)
Index of multiple deprivation			
Most deprived (1)	823 (9.1%)	292 (10.7%)	531 (8.4%)
Deprivation 2	1,528 (16.9%)	507 (18.6%)	1,021 (16.2%)
Deprivation 3	1,962 (21.7%)	574 (21.0%)	1,388 (22.0%)
Deprivation 4	2,121 (23.5%)	604 (22.1%)	1,517 (24.1%)
Least deprived (5)	2,593 (28.7%)	754 (27.6%)	1,839 (29.2%)
Unknown	533	156	377
Region			
East Midlands	517 (5.4%)	110 (3.8%)	407 (6.1%)
East of England	1,176 (12.3%)	270 (9.4%)	906 (13.6%)
London	1,199 (12.5%)	394 (13.6%)	805 (12.1%)
North East	223 (2.3%)	60 (2.1%)	163 (2.4%)
North West	794 (8.3%)	212 (7.3%)	582 (8.7%)
Northern Ireland	499 (5.2%)	133 (4.6%)	366 (5.5%)
Scotland	1,701 (17.8%)	919 (31.8%)	782 (11.7%)
South East	947 (9.9%)	210 (7.3%)	737 (11.0%)
South West	958 (10.0%)	193 (6.7%)	765 (11.5%)
Wales	230 (2.4%)	38 (1.3%)	192 (2.9%)
West Midlands	700 (7.3%)	239 (8.3%)	461 (6.9%)
Yorkshire and the Humber	616 (6.4%)	109 (3.8%)	507 (7.6%)

Household structure			
Lives with others (including children)	3,570 (37.3%)	1,192 (41.3%)	2,378 (35.6%)
Lives with others (no children)	4,834 (50.6%)	1,383 (47.9%)	3,451 (51.7%)
Lives alone	1,156 (12.1%)	312 (10.8%)	844 (12.6%)

Table E.3: Crude PCR positivity rates per 10,000 person-days and estimated vaccine effectiveness and protection from prior infection by vaccination status, time since previous infection and reported COVID symptoms.

						Symptomatic		Asymptomatic	
	Number of participants	Positive PCR tests	Exposure (person-days at risk)	Crude PCR positivity rate per 10,000 person-days (95% CI)	Protection relative to baseline (95% CI)	Positive PCR tests	Protection relative to baseline (95% CI)	Positive PCR tests	Protection relative to baseline (95% CI)
Whole population	9560	1298	1521928	8.53 (8.07, 9.01)	N/A	865	N/A	265	N/A
Vaccination status									
Waned third dose	9389	541	603023	8.97 (8.23, 9.76)	Baseline	340	Baseline	118	Baseline
Fourth dose	6345	757	918905	8.24 (7.66, 8.85)	13.08% (0.89, 23.76)	525	8.55% (-5.75, 20.91)	147	27.99% (6.61, 44.48)
Time since fourth dose									
0-2 months	6345	179	338530	5.29 (4.54, 6.12)	23.97% (8.48, 36.83)	131	18.52% (0.08, 33.56)	34	39.85% (11.24, 59.23)
2-4 months	5759	300	305192	9.83 (8.75, 11.01)	10.3% (-11.4, 27.78)	212	5.8% (-18.81, 25.31)	52	26.31% (-8.34, 49.88)
4-6 months	5005	278	275183	10.1 (8.95, 11.36)	1.71% (-16.97, 17.4)	182	-1.95% (-23.36, 15.74)	61	15.56% (-18.18, 39.66)
Time since previous infection									
Confirmed naive	773	173	115165	15.02 (12.87, 17.43)	7.71% (-35.45, 37.11)	124	16.06% (-29.03, 45.39)	33	-18.47% (-155.83, 45.13)
2+ years	1752	196	222783	8.8 (7.61, 10.12)	Baseline	127	Baseline	40	Baseline
1-2 years	2850	209	195715	10.68 (9.28, 12.23)	18.35% (-17.15, 43.09)	136	26.49% (-10.35, 51.04)	39	5.35% (-101.67, 55.58)
6-12 months	4123	347	426815	8.13 (7.3, 9.03)	29.08% (3.82, 47.71)	226	33.85% (7.48, 52.71)	81	10.56% (-68.56, 52.55)
0-6 months	3433	83	254866	3.26 (2.59, 4.04)	63.58% (46.85, 75.04)	38	71.99% (56.11, 82.13)	30	43.23% (-17.83, 72.65)

Note:
Participants joined the analysis at the date of their first PCR test after 12th September 2022. 171 participants had received their fourth dose booster on or before the date of their first PCR test and contributed no follow-up time in the waned third dose vaccine status.

Table E.4: Estimated mean sojourn time in PCR positive state, averaged across the study population, and empirical median duration of PCR positivity by vaccination status and time since previous infection.

	Estimated mean sojourn time in PCR positive state (95% CI)	Empirical median duration of PCR positivity (initial PCR positive to subsequent PCR negative) [interquartile range]
Whole population	7.51 days (6.94, 8.13)	15 days [12, 27]
Vaccination status		
Waned third dose	8.5 days (6.79, 10.64)	15 days [13, 28]
Fourth dose	6.9 days (5.87, 8.11)	14 days [12, 26]
Time since previous infection		
Confirmed naive	9.51 days (7.08, 12.78)	16 days [13, 28]
2+ years	9.16 days (7.16, 11.72)	14 days [12, 27]
1-2 years	7.44 days (5.86, 9.44)	14 days [12, 21]
6-12 months	6.56 days (5.63, 7.63)	14 days [12, 23]
0-6 months	7.25 days (6.32, 8.31)	16 days [14, 28]
COVID symptoms		
COVID symptoms	8.09 days (7.36, 8.9)	15 days [13, 27]
Non-COVID symptoms or asymptomatic	4.7 days (4.04, 5.47)	14 days [12, 23]

E.3 Alternative model specifications

E.3.1 Convalescent model

The studies of the SIREN cohort using Cox proportional hazards methodology described in Section E.1 applied a 90-day episode length, after which point participants were considered to be at risk for reinfection (Hall, Foulkes, Insalata *et al.* 2022; Hall, Insalata *et al.* 2024). To emulate this 90-day episode, participants with a negative PCR can be assigned to a recovery (or convalescence) state for a defined period following a positive PCR result (e.g. 90 days post-infection). Figure E.3 panels A and B show the estimates from the two-state (main) model and three-state convalescent model.

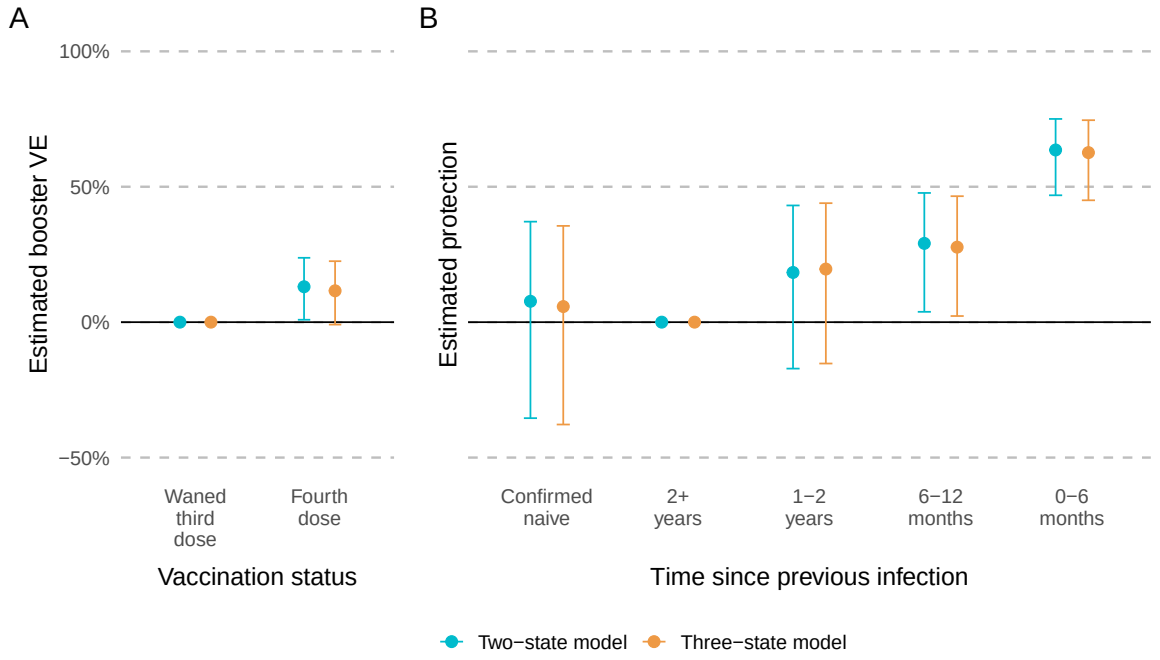


Figure E.3: Estimated booster VE (panel A, model M1) and estimated protection from previous infection (panel B, model M3) in two-state and three-state model. Error bars show the 95% CI around the estimates.

E.3.2 Semi-Markov model

Figure E.4 compares mean sojourn times estimated by the Markov model (panel A) to mean sojourn times for long and short stayers estimated by the semi-Markov model (panel B). As expected, the mean sojourn times for the Markov model fall somewhere in between the mean times for the long and short stayers.

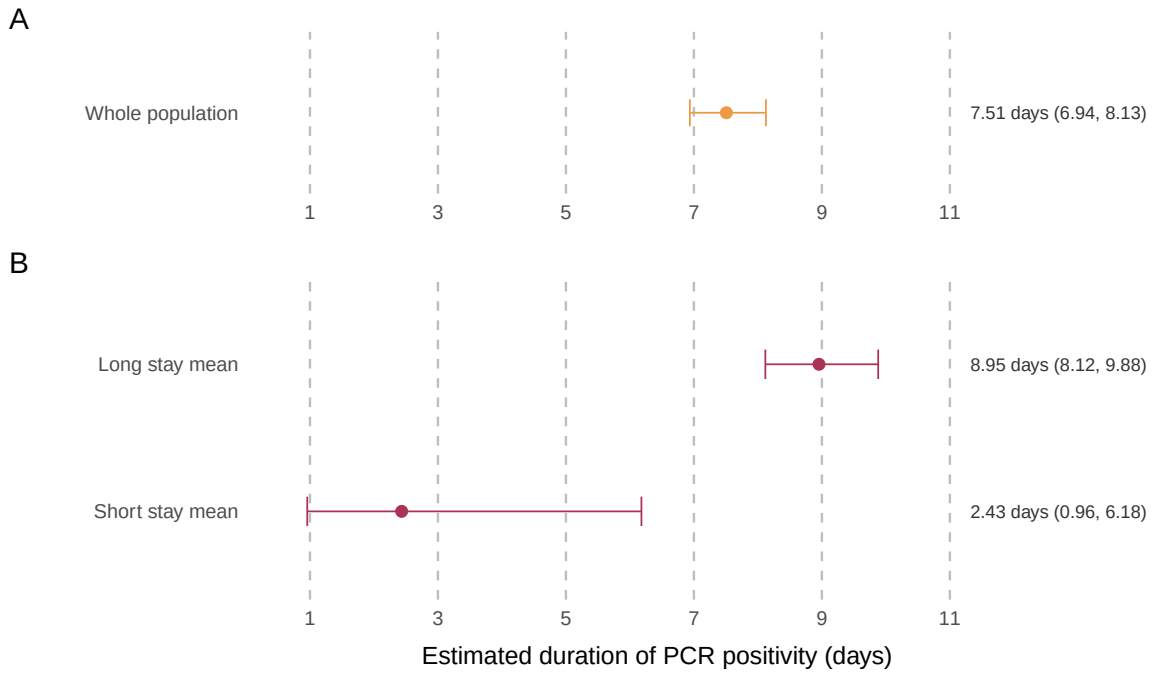


Figure E.4: Estimated mean sojourn time in PCR positive state, averaged across the study population, estimated by Markov model (panel A) and semi-Markov model (panel B). Error bars show the 95% CI around the estimates.

E.3.3 Misclassification model

Hazards estimated via a misclassification model with fixed misclassification probabilities $\Pr(S \mid P) = 0.09$ and $\Pr(I \mid N) = 0.001$ are shown in Figure E.5. Hazards were similar to those estimated in the two-state model without misclassification, with slightly wider confidence intervals.

E.3.4 Naïve baseline model

'2+ years' was chosen as the baseline category for the time since previous infection covariate in the main analysis. Estimates using the 'confirmed naïve' group as the baseline category instead are shown in Figure E.6.

E.4 Cox proportional hazards model comparison

Figures E.8 and E.7 compare hazards estimated by multi-state models M2 and M3 to those from the corresponding stratified Cox proportional hazards model, where the same covariates were included in both models.

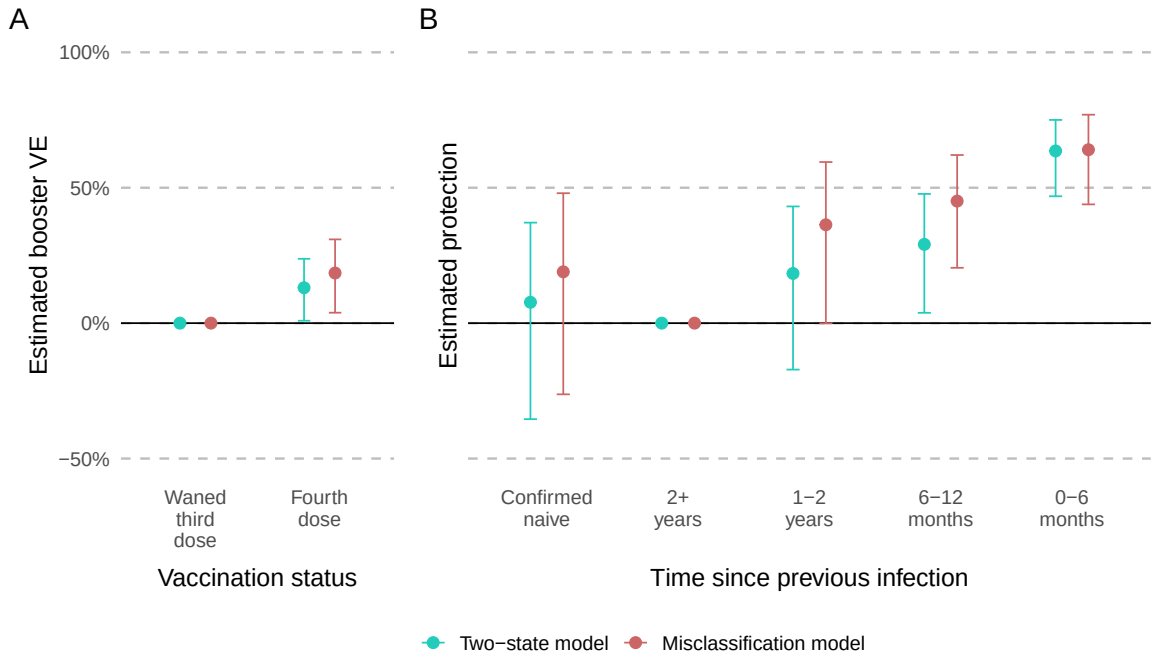


Figure E.5: Estimated booster VE (panel A, model M1) and estimated protection from previous infection (panel B, model M3) in two-state and misclassification model. Error bars show the 95% CI around the estimates.

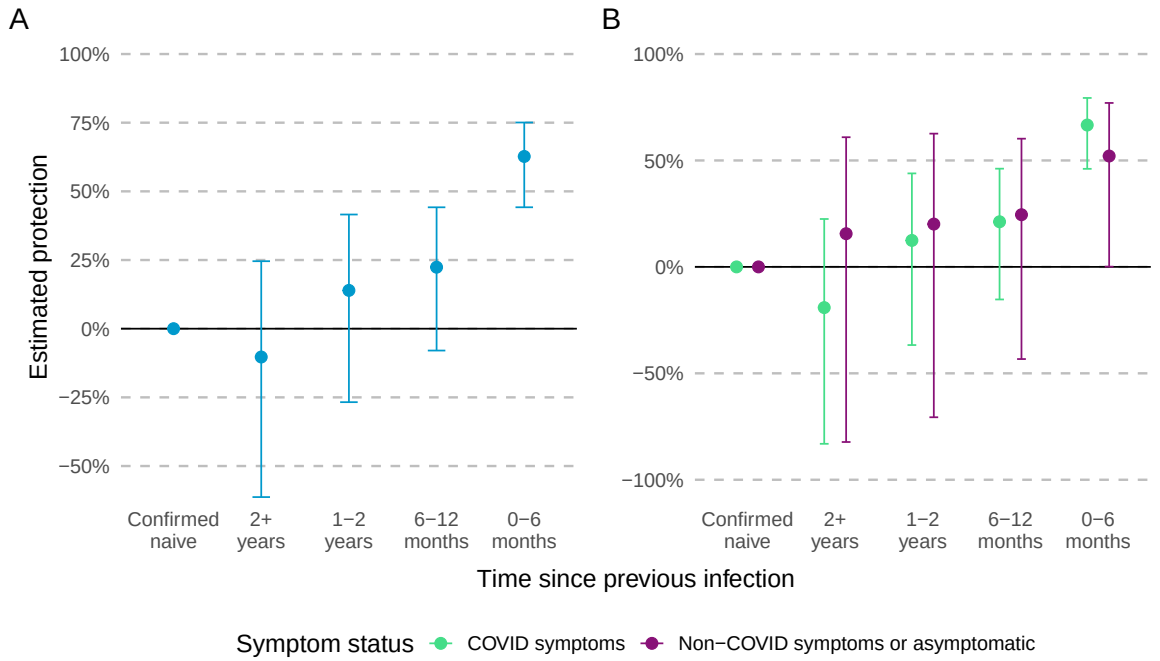


Figure E.6: Estimated protection from previous infection, by time since previous infection (panel A, model M3), and symptom status (panel B, model M7), relative to confirmed naïve baseline category. Error bars show the 95% CI around the estimates.

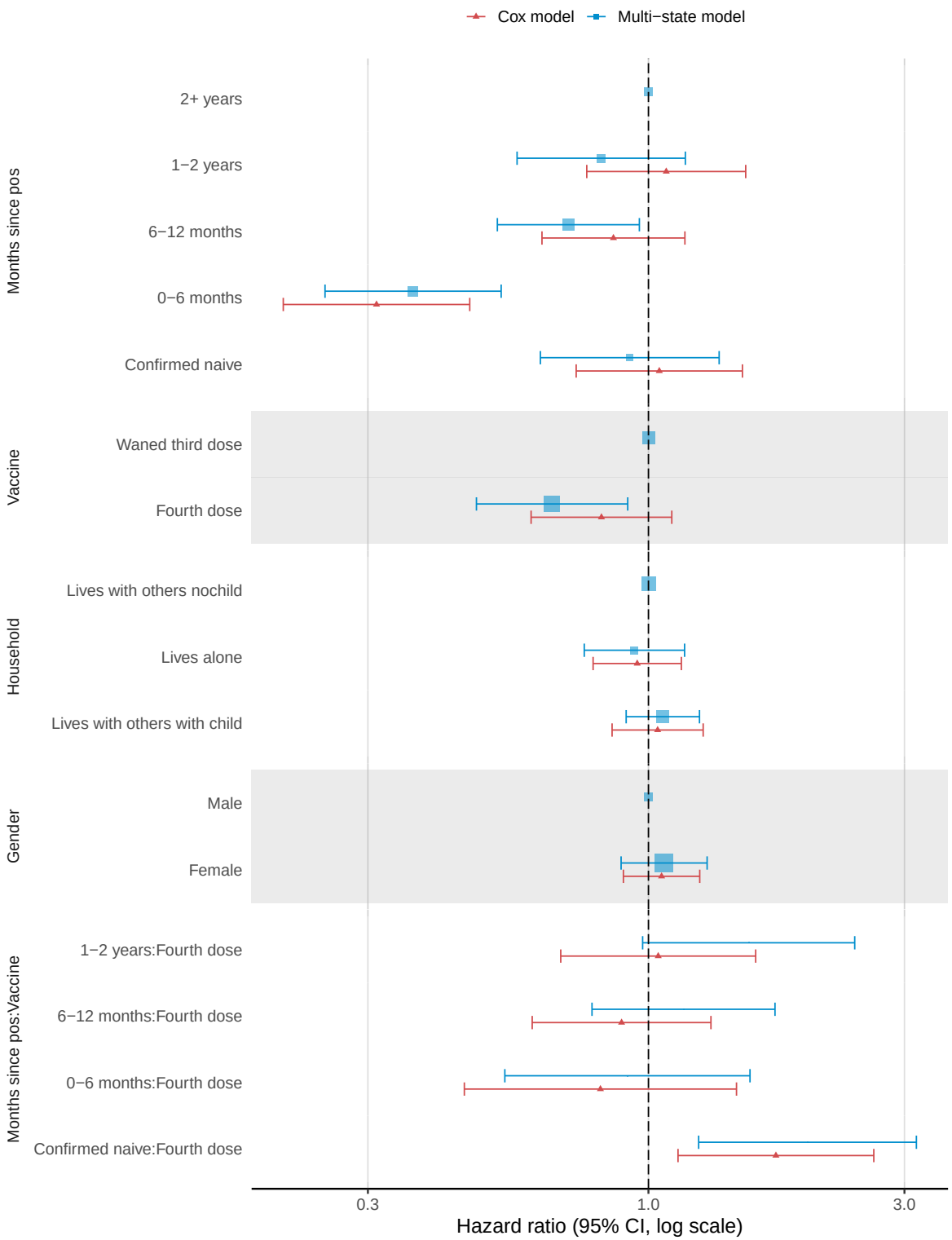


Figure E.7: Comparison of hazards estimated by multi-state model M3 and the corresponding Cox proportional hazards model, for covariates common between the two models. Error bars show the 95% CI around the estimated hazards.

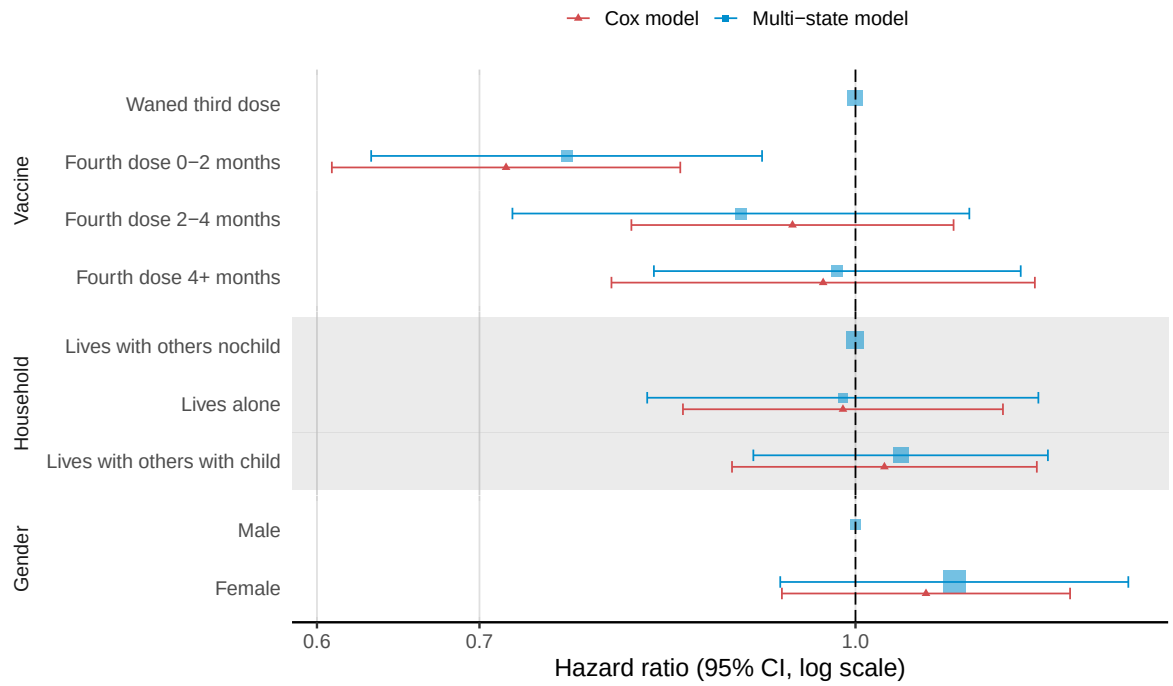


Figure E.8: Comparison of hazards estimated by multi-state model M2 and the corresponding Cox proportional hazards model, for covariates common between the two models. Error bars show the 95% CI around the estimated hazards.

E.5 Goodness of fit

E.5.1 Simulated number of infections

Model M3 was used to simulate the expected number of infections from the fitted model, i.e. $\mathbb{E}(N^I_{\text{TOT}};\hat{q}^I)$, where N^I is the counting process for entries into the infected state and $\hat{q}^I$ is the maximum likelihood estimate for the transition intensity to the infected state as described in Section 5.6. 500 simulations were used to infer the variability in this estimate, $\text{Var}(\mathbb{E}(N^I_{\text{TOT}};\hat{q}^I))$.

Compared to the observed $n = 1023$ infections, the expected number of infections under the fitted model was 1069 infections (95% variance 1013 to 1125). With weekly sampling and 100% adherence to testing, a total of 1310 infections (95% variance 1242 to 1373) would be expected to have been been detected (Figure E.9).

E.5.2 Causal model goodness of fit

To assess model fit of the causal model, the number of individuals forecast to be occupying the infected state in the fitted model at a series of times t was compared to an estimate derived from the observed data (Figure E.10).

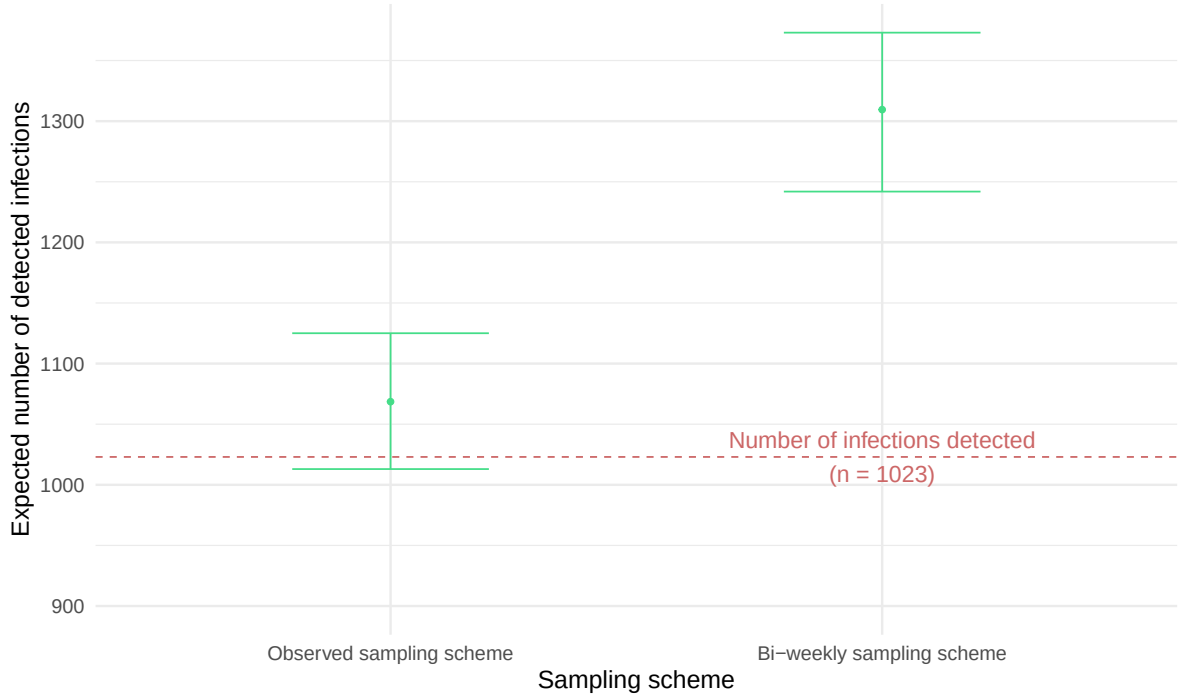


Figure E.9: Expected number of detected infections under observed and bi-weekly sampling schemes with error bars showing variability in the mean estimate, simulated from model M3.

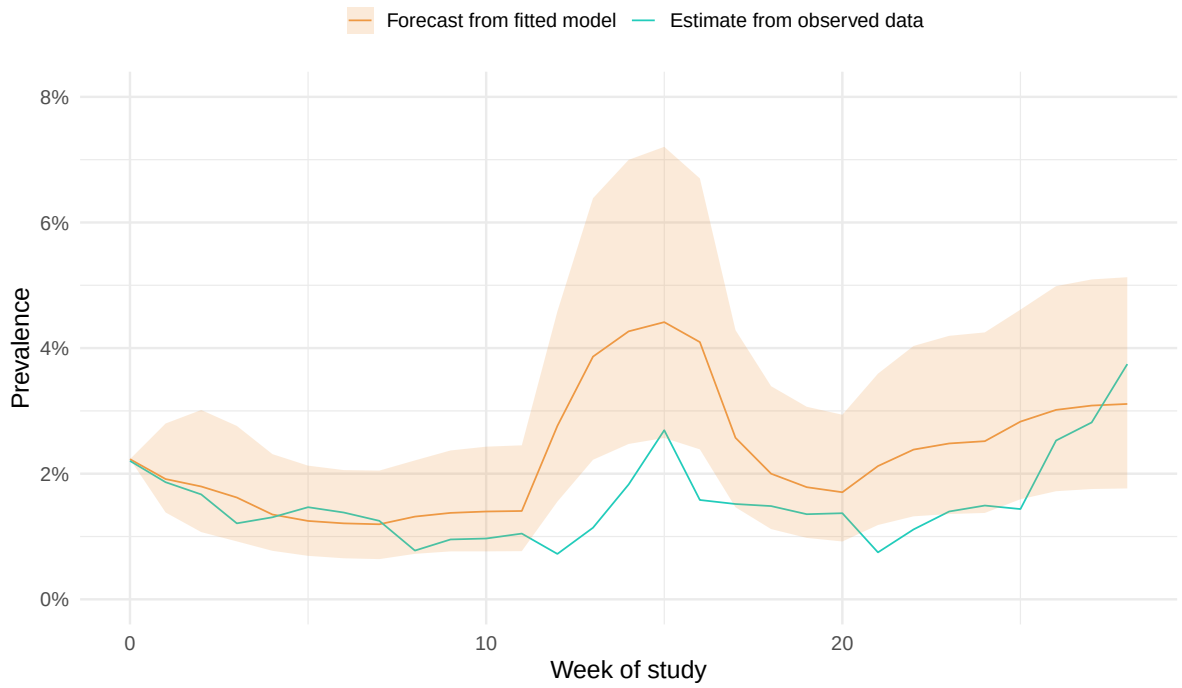


Figure E.10: Comparison of prevalence in the infected state forecast from fitted model and more empirical estimate over time. Shaded area shows the 95% CI around the expected prevalence.