An epidemiological investigation of the burden of and facility-level risk factors for SARS-CoV-2 infection and outbreaks in care home staff and residents

Dr Maria Krutikov
University College London
PhD Thesis in Infectious Disease Epidemiology
July 2023

Supervisors:
Prof Laura Shallcross
Prof Andrew Copas
Prof Andrew Hayward

I, Maria Krutikov, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Acknowledgements

The COVID-19 pandemic dramatically changed the course of many people's lives including my own. I am grateful that I was able to contribute to the national pandemic response - some of that work is presented in this thesis. Over the last three years I have had the opportunity to work with and learn from highly experienced people across different sectors who were extremely generous with their time, regardless of how much of it they had to spare. This especially relates to the people living and working in care homes who I was privileged to work with, including those who participated in the study, whose ingenuity and resilience in the face of adversity is inspiring. I would particularly like to pay tribute to those who lost their lives to COVID-19, including my grandfather.

I would like to thank my supervisors, Laura, Andrew, and Andrew whose guidance and expertise have been a constant. In particular I am grateful to Laura, my primary supervisor, for her tireless efforts establishing the VIVALDI study, her mentorship, and for giving me the opportunity to do this work.

I have benefited from support from the wider VIVALDI team. This includes Oliver Stirrup and Tom Palmer who provided me with statistical support and teaching, Borscha Azmi and Chris Fuller who project managed the study, and Hector Altamirano who drafted the building survey and whose insights helped me to interpret the results of my risk factor analysis. The wider team at University of Birmingham including Paul Moss, Rachel Bruton, Gokhan Tut, and Dave Bone undertook antibody testing of blood samples and shared these data with me, some of which I present in this thesis. Support from colleagues at the UK Health Security Agency, including Alasdair Donaldson and Aidan Irwin-Singer, helped to overcome logistical challenges with study delivery at pace and facilitated wide dissemination of results. This work would not have been possible without the data pipelines for which I would particularly like to thank Mark Marshall and Igor Monakhov.

Finally, I would like to acknowledge all the support that I have had over the years from my family and friends. In particular I am grateful to my wife Ali, who spurs me on every day.

UCL Research Paper Declaration Form

referencing the doctoral candidate's own published work(s)

Please use this form to declare if parts of your thesis are already available in another format, e.g. if data, text, or figures:

- *have been uploaded to a preprint server*
- *are in submission to a peer-reviewed publication*
- *have been published in a peer-reviewed publication, e.g. journal, textbook.*

This form should be completed as many times as necessary. For instance, if you have seven thesis chapters, two of which containing material that has already been published, you would complete this form twice.

1. **For a research manuscript that has already been published** (if not yet published, please skip to section 2)

a) **What is the title of the manuscript?**

Study Protocol: Understanding SARS-CoV-2 infection, immunity and its duration in care home residents and staff in England (VIVALDI)

b) **Please include a link to or doi for the work**

https://doi.org/10.12688/wellcomeopenres.16193.2

c) **Where was the work published?**

Wellcome Open Research

d) **Who published the work?** (e.g. OUP)

The Wellcome Trust

e) **When was the work published?**

29 January 2021

f) **List the manuscript's authors in the order they appear on the publication**

Maria Krutikov, Tom Palmer, Alasdair Donaldson, Fabiana Lorencatto, Gill Forbes, Andrew J Copas, James Robson, Susan Hopkins, Paul Moss, Jeremy Farrar, Andrew Hayward, Laura Shallcross

g) **Was the work peer reviewed?**

Yes

h) **Have you retained the copyright?**

Yes

i) **Was an earlier form of the manuscript uploaded to a preprint server?** (e.g. medRxiv). If 'Yes', please give a link or doi)

Yes – Wellcome Open Research (see above)

If 'No', please seek permission from the relevant publisher and check the box next to the below statement:

☐

*I acknowledge permission of the publisher named under **1d** to include in this thesis portions of the publication named as included in **1c**.*

2. **For a research manuscript prepared for publication but that has not yet been published** (if already published, please skip to section 3)

a) **What is the current title of the manuscript?**

Click or tap here to enter text.

b) **Has the manuscript been uploaded to a preprint server?** (e.g. medRxiv; if 'Yes', please give a link or doi)

Click or tap here to enter text.

c) **Where is the work intended to be published?** (e.g. journal names)

Click or tap here to enter text.

d) **List the manuscript's authors in the intended authorship order**

Click or tap here to enter text.

e) **Stage of publication** (e.g. in submission)

Click or tap here to enter text.

3. **For multi-authored work, please give a statement of contribution covering all authors** (if single-author, please skip to section 4)

LS, AH, SH, JF, AS, FL, GF, and MK conceptualised the study. LS, AH, SH, AD, and JF secured funding. First draft was written by MK, all authors revised and edited the manuscript. LS and MK shared the final responsibility for the decision to submit for publication.

4. **In which chapter(s) of your thesis can this material be found?**

Chapters 3 - 7

5. **e-Signatures confirming that the information above is accurate** (this form should be co-signed by the supervisor/ senior author unless this is not appropriate, e.g. if the paper was a single-author work)

Candidate

Date:
8th June 2023

Supervisor/ Senior Author (where appropriate)
L.Shallcross
Date
5/7/23

UCL Research Paper Declaration Form
referencing the doctoral candidate's own published work(s)

Please use this form to declare if parts of your thesis are already available in another format, e.g. if data, text, or figures:

- *have been uploaded to a preprint server*
- *are in submission to a peer-reviewed publication*
- *have been published in a peer-reviewed publication, e.g. journal, textbook.*

This form should be completed as many times as necessary. For instance, if you have seven thesis chapters, two of which containing material that has already been published, you would complete this form twice.

6. For a research manuscript that has already been published (if not yet published, please skip to section 2)

j) **What is the title of the manuscript?**

Spread of a Variant SARS-CoV-2 in Long-Term Care Facilities in England

k) **Please include a link to or doi for the work**

DOI: 10.1056/NEJMc2035906

l) **Where was the work published?**

The New England Journal of Medicine

m) **Who published the work?** (e.g. OUP)

Massachusetts Medical Society

n) **When was the work published?**

29th April 2021

o) **List the manuscript's authors in the order they appear on the publication**

Maria Krutikov, Andrew Hayward, Laura Shallcross

p) **Was the work peer reviewed?**

Yes

q) **Have you retained the copyright?**

No

r) **Was an earlier form of the manuscript uploaded to a preprint server?** (e.g. medRxiv). If 'Yes', please give a link or doi)

No

If 'No', please seek permission from the relevant publisher and check the box next to the below statement:

☒

*I acknowledge permission of the publisher named under **1d** to include in this thesis portions of the publication named as included in **1c**.*

7. For a research manuscript prepared for publication but that has not yet been published (if already published, please skip to section 3)

f) **What is the current title of the manuscript?**

Click or tap here to enter text.

g) **Has the manuscript been uploaded to a preprint server?** (e.g. medRxiv; if 'Yes', please give a link or doi)

Click or tap here to enter text.

h) **Where is the work intended to be published?** (e.g. journal names)

Click or tap here to enter text.

i) **List the manuscript's authors in the intended authorship order**

Click or tap here to enter text.

j) **Stage of publication** (e.g. in submission)

Click or tap here to enter text.

8. **For multi-authored work, please give a statement of contribution covering all authors** (if single-author, please skip to section 4)

LS, AH, and MK conceptualised the study. MK and LS developed the statistical analysis plan. MK did the statistical analysis. LS and MK wrote the first draft of the manuscript. All authors revised and edited the manuscript.

9. **In which chapter(s) of your thesis can this material be found?**

4

10. **e-Signatures confirming that the information above is accurate** (this form should be co-signed by the supervisor/ senior author unless this is not appropriate, e.g. if the paper was a single-author work)

Candidate

Date:
8th June 2023

Supervisor/ Senior Author (where appropriate)
L.Shallcross
Date
5/7/23

UCL Research Paper Declaration Form
referencing the doctoral candidate's own published work(s)

Please use this form to declare if parts of your thesis are already available in another format, e.g. if data, text, or figures:

- *have been uploaded to a preprint server*
- *are in submission to a peer-reviewed publication*
- *have been published in a peer-reviewed publication, e.g. journal, textbook.*

This form should be completed as many times as necessary. For instance, if you have seven thesis chapters, two of which containing material that has already been published, you would complete this form twice.

11. For a research manuscript that has already been published (if not yet published, please skip to section 2)

s) **What is the title of the manuscript?**

Prevalence and duration of detectable SARS-CoV-2 nucleocapsid antibodies in staff and residents of long-term care facilities over the first year of the pandemic (VIVALDI study): prospective cohort study in England

t) **Please include a link to or doi for the work**

DOI: 10.1016/S2666-7568(21)00282-8

u) **Where was the work published?**

Lancet Healthy Longevity

v) **Who published the work?** (e.g. OUP)

Elsevier Ltd

w) **When was the work published?**

16th December 2021

x) **List the manuscript's authors in the order they appear on the publication**

Maria Krutikov, Tom Palmer, Gokhan Tut, Christopher Fuller, Borscha Azmi, Rebecca Giddings, Madhumita Shrotri, Nayandeep Kaur, Panagiota Sylla, Tara Lancaster, Aidan Irwin-Singer, Andrew Hayward, Paul Moss, Andrew Copas, Laura Shallcross

y) **Was the work peer reviewed?**

Yes

z) **Have you retained the copyright?**

No

aa) **Was an earlier form of the manuscript uploaded to a preprint server?** (e.g. medRxiv). If 'Yes', please give a link or doi)

Yes - https://www.medrxiv.org/content/10.1101/2021.09.27.21264166v1

If 'No', please seek permission from the relevant publisher and check the box next to the below statement:

☐

*I acknowledge permission of the publisher named under **1d** to include in this thesis portions of the publication named as included in **1c**.*

12. For a research manuscript prepared for publication but that has not yet been published (if already published, please skip to section 3)

k) **What is the current title of the manuscript?**

Click or tap here to enter text.

l) **Has the manuscript been uploaded to a preprint server?** (e.g. medRxiv; if 'Yes', please give a link or doi)

Click or tap here to enter text.

m) **Where is the work intended to be published?** (e.g. journal names)

Click or tap here to enter text.

n) **List the manuscript's authors in the intended authorship order**

Click or tap here to enter text.

o) **Stage of publication** (e.g. in submission)

Click or tap here to enter text.

13. **For multi-authored work, please give a statement of contribution covering all authors** (if single-author, please skip to section 4)

LS, AH, AC, TP, and MK conceptualised the study. MK, TP, LS, and AC developed the statistical analysis plan. MK did the formal statistical analysis. MK, CF, BA, RG, MS, and AI-S were involved with project administration. MK did the literature review. LS and AH obtained research funding. GT, NK, PS, TL, and PM did laboratory investigations. MK wrote the first draft of the manuscript. All authors revised and edited the manuscript. MK and TP accessed and verified the data. All authors had full access to all the data reported in the study. LS, AC, and MK shared the final responsibility for the decision to submit for publication.

14. **In which chapter(s) of your thesis can this material be found?**

4 and 5

15. **e-Signatures confirming that the information above is accurate** (this form should be co-signed by the supervisor/ senior author unless this is not appropriate, e.g. if the paper was a single-author work)

Candidate

Date:
8th June 2023

Supervisor/ Senior Author (where appropriate)
L.Shallcross
Date 5/7/23

UCL Research Paper Declaration Form

referencing the doctoral candidate's own published work(s)

Please use this form to declare if parts of your thesis are already available in another format, e.g. if data, text, or figures:

- *have been uploaded to a preprint server*
- *are in submission to a peer-reviewed publication*
- *have been published in a peer-reviewed publication, e.g. journal, textbook.*

This form should be completed as many times as necessary. For instance, if you have seven thesis chapters, two of which containing material that has already been published, you would complete this form twice.

16. For a research manuscript that has already been published (if not yet published, please skip to section 2)

bb) **What is the title of the manuscript?**

Severe acute Respiratory Syndrome Coronavirus 2 Anti-Spike Antibody Levels Following Second Dose of ChAdOx1 nCov-19 or BNT162b2 Vaccine in Residents of Long-term Care Facilities in England (VIVALDI)

cc) **Please include a link to or doi for the work**

https://doi.org/10.1093/infdis/jiac146

dd) **Where was the work published?**

The Journal of Infectious Diseases

ee) **Who published the work?** (e.g. OUP)

OUP

ff) **When was the work published?**

16th April 2022

gg) **List the manuscript's authors in the order they appear on the publication**

Oliver Stirrup*, Maria Krutikov*, Gokhan Tut*, Tom Palmer, David Bone, Rachel Bruton, Chris Fuller, Borscha Azmi, Tara Lancaster, Panagiota Sylla, Nayandeep Kaur, Eliska Spalkova, Christopher Bentley, Umayr Amin, Azar Jadir, Samuel Hulme, Rebecca Giddings, Hadjer Nacer-Laidi, Verity Baynton, Aidan Irwin-Singer, Andrew Hayward, Paul Moss, Andrew Copas, Laura Shallcross
*Joint first authors

hh) **Was the work peer reviewed?**

Yes

ii) **Have you retained the copyright?**

No

jj) **Was an earlier form of the manuscript uploaded to a preprint server?** (e.g. medRxiv). If 'Yes', please give a link or doi)

Yes –
https://www.medrxiv.org/content/10.1101/2022.01.26.22269885v1

If 'No', please seek permission from the relevant publisher and check the box next to the below statement:

☐

*I acknowledge permission of the publisher named under **1d** to include in this thesis portions of the publication named as included in **1c**.*

17. For a research manuscript prepared for publication but that has not yet been published (if already published, please skip to section 3)

p) **What is the current title of the manuscript?**

Click or tap here to enter text.

q) **Has the manuscript been uploaded to a preprint server?** (e.g. medRxiv; if 'Yes', please give a link or doi)

Click or tap here to enter text.

r) **Where is the work intended to be published?** (e.g. journal names)

Click or tap here to enter text.

s) **List the manuscript's authors in the intended authorship order**

Click or tap here to enter text.

t) **Stage of publication** (e.g. in submission)

Click or tap here to enter text.

18. **For multi-authored work, please give a statement of contribution covering all authors** (if single-author, please skip to section 4)

LS, PM, AC, OS, GT, and MK conceptualised the study. MK, TP, OS, LS, and AC developed the statistical analysis plan. OS did the formal statistical analysis. GT, NK, PS, ES, CB, UA, AJ, SH, TL, and PM did laboratory investigations. MK, CF, BA, RB, AIS, DB, and VB were involved with project administration. LS and AH obtained research funding. MK and OS wrote the first draft of the manuscript. All authors revised and edited the manuscript. MK, TP, and OS verified the data. All authors had full access to all the data reported. LS, AC, OS, and MK share the final responsibility for the decision to submit for publication.

19. **In which chapter(s) of your thesis can this material be found?**

5

20. **e-Signatures confirming that the information above is accurate** (this form should be co-signed by the supervisor/ senior author unless this is not appropriate, e.g. if the paper was a single-author work)

Candidate

Date:
8th June 2023

Supervisor/ Senior Author (where appropriate)
L.Shallcross
Date
5/7/23

UCL Research Paper Declaration Form

referencing the doctoral candidate's own published work(s)

Please use this form to declare if parts of your thesis are already available in another format, e.g. if data, text, or figures:

- *have been uploaded to a preprint server*
- *are in submission to a peer-reviewed publication*
- *have been published in a peer-reviewed publication, e.g. journal, textbook.*

This form should be completed as many times as necessary. For instance, if you have seven thesis chapters, two of which containing material that has already been published, you would complete this form twice.

21. For a research manuscript that has already been published (if not yet published, please skip to section 2)

kk) **What is the title of the manuscript?**

Click or tap here to enter text.

ll) **Please include a link to or doi for the work**

Click or tap here to enter text.

mm) **Where was the work published?**

Click or tap here to enter text.

nn) **Who published the work?** (e.g. OUP)

Click or tap here to enter text.

oo) **When was the work published?**

Click or tap here to enter text.

pp) **List the manuscript's authors in the order they appear on the publication**

Click or tap here to enter text.

qq) **Was the work peer reviewed?**

Click or tap here to enter text.

rr) **Have you retained the copyright?**

Click or tap here to enter text.

ss) **Was an earlier form of the manuscript uploaded to a preprint server?** (e.g. medRxiv). If 'Yes', please give a link or doi)

Click or tap here to enter text.

If 'No', please seek permission from the relevant publisher and check the box next to the below statement:

☐

*I acknowledge permission of the publisher named under **1d** to include in this thesis portions of the publication named as included in **1c**.*

22. For a research manuscript prepared for publication but that has not yet been published (if already published, please skip to section 3)

u) **What is the current title of the manuscript?**

Built environment risk factors for ingress and transmission of SARS-CoV-2 within long-term care facilities in England (VIVALDI): Prospective cohort study using routine data and cross-sectional survey

v) **Has the manuscript been uploaded to a preprint server?** (e.g. medRxiv; if 'Yes', please give a link or doi)

No

w) **Where is the work intended to be published?** (e.g. journal names)

Journal of the American Medical Directors Association

x) **List the manuscript's authors in the intended authorship order**

Maria Krutikov, Oliver Stirrup, Chris Fuller, Natalie Adams, Borscha Azmi, Aidan Irwin-Singer, Niyathi Sethu, Andrew Hayward, Hector Altamirano, Andrew Copas, Laura Shallcross

y) **Stage of publication** (e.g. in submission)

Submitted

23. For multi-authored work, please give a statement of contribution covering all authors (if single-author, please skip to section 4)

MK, HA, LS, OS, and AC conceptualised the study. MK, HA, and LS designed the survey. MK, BA, CF, and NS collected the data. MK, OS, LS, and AC planned the statistical analysis. MK, BA, CF, AIS, and CF were involved with project administration. LS and AH obtained research funding. CF and MK entered the data. MK conducted the statistical analysis. MK and OS accessed and verified the data. MK, OS, LS, and AC had access to the data. MK, HA, LS, AC, and OS interpreted the results. MK wrote the first draft of the manuscript. All authors revised and edited the manuscript. All authors had full access to all the data reported in the study and accept responsibility for the decision to submit for publication.

24. In which chapter(s) of your thesis can this material be found?

6

25. e-Signatures confirming that the information above is accurate (this form should be co-signed by the supervisor/ senior author unless this is not appropriate, e.g. if the paper was a single-author work)

Candidate

Date:
8th June 2023

Supervisor/ Senior Author (where appropriate)
L. Shallcross
Date
5/7/23

Abstract

Background

The COVID-19 pandemic significantly impacted care homes, highlighting their vulnerability to infection. I described the burden of infection and investigated facility-level risk factors for SARS-CoV-2 infections and outbreaks within care homes.

Methods

I helped to rapidly establish the VIVALDI cohort study in ~330 care homes for older people in England (ISRCTN14447421), which hosted my analyses. I reviewed the literature to investigate risk factors for SARS-CoV-2 in care homes. Using data from asymptomatic SARS-CoV-2 testing and anti-nucleocapsid (from infection) and anti-spike (from vaccination) antibodies in care home staff and residents, I estimated prevalence and spread of SARS-CoV-2 infection across homes and modelled longevity of antibody responses following infection and vaccination. Finally, I designed a built environment survey and evaluated environmental risk factors for ingress and transmission of SARS-CoV-2.

Results

Within VIVALDI, over one-quarter of staff and one-third of residents were infected over 15 months from the pandemic start, increasing to two-thirds after two years. I showed that nucleocapsid-antibodies were negative in half of participants eight months post-infection, suggesting waning immunity, however spike-antibody waning rates following vaccination were comparable between staff and residents. I demonstrated rapid spread of the emergent B.1.1.7 variant in care homes, suggesting introduction of infection from the community. Community incidence of SARS-CoV-2 was also the main risk factor for infection ingress (measured by outbreak incidence) but not transmission (measured by infection incidence, outbreak size, and duration), which was associated with environmental factors like bedroom and storey number, building type, indoor temperature, air quality, and ventilation.

Conclusion

Care homes experienced high SARS-CoV-2 rates despite stringent control measures, with comparable antibody responses between staff and residents that wane following

infection. Although preventing infection entry is challenging, environmental modifications may limit spread. Building on lessons from VIVALDI, controlling infection in care homes should be a research priority.

Impact Statement

The work presented in this thesis has directly benefitted care home staff, residents, their families, social care leaders, public health officials, policymakers, and researchers.

The VIVALDI study which I helped establish (Chapter 3), is one the largest care home cohorts monitoring COVID-19 globally. Over the pandemic, it provided data on infection prevalence, immunity, and vaccine efficacy, that informed national policy decisions to protect care homes from infection. Using asymptomatic SARS-CoV-2 screening, staff and residents could be reliably identified in routine data for the first time. This addressed broader research priorities by permitting linkage across datasets. VIVALDI is registered with the International Standard Randomised Controlled Trial Number (ISRCTN) registry and has achieved over fifteen peer-reviewed publications with significant media attention. The protocol is published in *Wellcome Open Research*.

The three analytical chapters of my thesis (Chapters 4,5,6) outline work that impacted on policy. In Chapter 4, I present the first description of emergent Alpha variant spread into care homes, despite control measures. This was presented to New and Emerging Respiratory Viral Threats Advisory Group (NERVTAG) and the Chief Medical Officer, directly informing decisions to impose a national lockdown in January 2021. Publication in *New England Journal of Medicine* had significant international and national media coverage. My estimate of SARS-CoV-2 seroprevalence in care home residents and staff was amongst the largest and most representative globally, demonstrating higher prevalence in care homes than in the community. This was published in *Lancet Healthy Longevity* and was highlighted with an accompanying commentary. Estimates of cumulative incidence of SARS-CoV-2 in residents and staff over two years were presented at the European Congress of Clinical Microbiology and Infectious Diseases 2023 in Denmark, featuring as a conference highlight.

In Chapter 5, I describe waning of infection-induced antibody responses, which is greater amongst staff than residents and model vaccine-induced antibody levels, showing comparable responses between groups. These results were presented to

academics and Ministers at the Department of Health & Social Care Data Debrief Group (DDG), directly informing re-vaccination strategies. Findings were disseminated through publication in *Lancet Healthy Longevity* and in *the Journal of Infectious Diseases* and distributed amongst participating care homes as leaflets and posters.

In Chapter 6, I demonstrate substantial diversity in care home built environments and highlight the influence of environmental features on infection transmission for the first time, generating data and identifying gaps for future proposals. Modelling infection and ingress using more than one outcome separately is more informative for control measures and future study designs. Results were presented to the DDG and submitted to a peer-reviewed journal. The lay summary will be shared with key stakeholders.

This work has showcased how collaboration between researchers, the care sector, and policymakers can maximise impact by ensuring results are relevant and timely. These relationships have built trust and raised the research profile in care settings. The VIVALDI data platform also has wider applications for infections beyond SARS-CoV-2. This has set a precedent for future research collaborations and will strengthen grant applications for academic programmes in care homes.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

95% CI	95% Confidence Interval
CAG	Confidentiality Advisory Group
CI	Chief Investigator
COG-UK	COVID-19 Genomics UK Consortium
COPI	Control of Patient Information Regulations 2002
COVID-19	Coronavirus Disease 2019
CQC	Care Quality Commission
CRN	Clinical Research Network
Ct	Cycle threshold
DHSC	Department of Health & Social Care
DNA	Deoxyribonucleic Acid
DSH	Data Safe Haven
ECDC	European Centre for Disease Prevention and Control
FOTE	Friends of the Elderly
FSHC	Four Seasons Health Care
GRADE	Grading of Recommendations Assessment Development and Evaluation
HES	Hospital Episode Statistics
HR	Hazards Ratio
HRA	Health Research Authority
IMD	Index of Multiple Deprivations
IPC	Infection Prevention & Control
IRR	Incidence Rate Ratio
ISRCTN	International Standard Randomised Controlled Trial Number
LA	Local Authority
LFD	Lateral Flow Device
MSD	Meso Scale Discovery
NERVTAG	New & Emerging Respiratory Viral Threats Advisory Group
NHS	National Health Service
NHSD	National Health Service Digital
NHSE	National Health Service England
NICE	National Institute for Clinical Excellence

NIHR	National Institute for Health and Care Research
NIMS	National Immunisations Management System
ONS	Office for National Statistics
OR	Odds ratio
OSJCT	The Orders of St John Care Trust
PCR	Polymerase Chain Reaction
PHE	Public Health England
PPE	Personal Protective Equipment
PPIE	Patient & Public Involvement and Engagement
R_0	Basic Reproduction Number
RBD	Receptor-Binding Domain
REC	Research Ethics Committee
RNA	Ribonucleic Acid
SAGE	Scientific Advisory group for Emergencies
SARS-CoV-2	Severe Acute Respiratory Syndrome Coronavirus 2
SE	Standard Error
SGTF	S-Gene Target Failure
SOP	Standard Operating Procedure
TDL	The Doctor's Laboratory
UCL	University College London
UKHSA	United Kingdom Health Security Agency
USA	United States of America
WHO	World Health Organisation

Chapter 1

Introduction

I am an infectious diseases and medical microbiology specialist registrar and, in this thesis, I will describe the original research that I undertook during a 3-year Wellcome Trust funded clinical PhD between November 2020 and November 2023. My PhD has been based at the Institute of Health Informatics at University College London and focusses on infectious disease epidemiology. My research was hosted within the VIVALDI study, a government-funded national surveillance study of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) in care homes, that I played an integral role in establishing and running over the first three years of the global Coronavirus Disease 19 (COVID-19) pandemic.

In this chapter, I describe the aim and objectives of my PhD. To provide context, I first review the characteristics of SARS-CoV-2 and COVID-19, present a short overview of the pandemic with a focus on the UK, and then give a more detailed description of how the pandemic evolved in care homes.

1.1 An overview of the SARS-CoV-2 pandemic

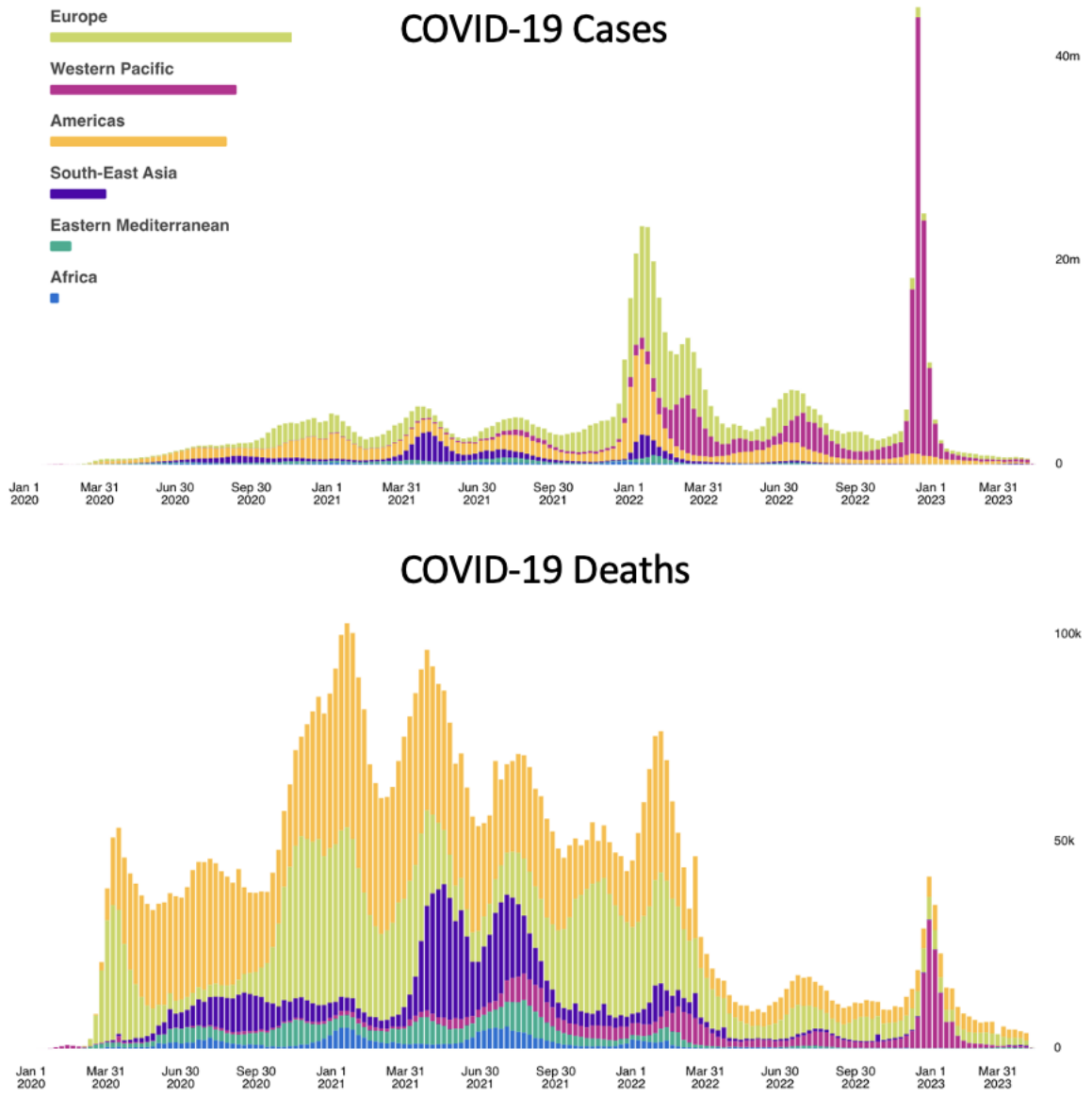
The SARS-CoV-2 pandemic was the largest global public health emergency since the Spanish influenza pandemic of 1918, which infected a third of the world's population of whom one in ten died.[1] The SARS-CoV-2 virus was first identified in China in December 2019 and spread rapidly across the globe.[2] COVID-19 is the clinical syndrome caused by the virus and is characterised by respiratory illness with progression to respiratory and multi-organ failure in some cases.[3] The high transmissibility and severity of infection, especially in vulnerable hosts, caused great concern as the infection count and mortality increased at an accelerated rate, prompting the declaration of a pandemic by the World Health Organisation (WHO) in March 2020.[4]

As of the end of March 2023, 762 million confirmed cases and 6.9 million related deaths have been reported globally,[5] although this is probably an underestimate as

there has been geographic and temporal variation in access to testing, Figure 1.1. It is estimated that by March 2022 at least half of the European population had evidence of prior infection.[6] The economic and health impacts of the pandemic have included school closures and associated impacts on education, loss of income due to closure of businesses, social isolation, abuse within the home, and mental health issues exacerbated by all of these. These have penetrated all layers of society and highlighted stark inequality in the availability of resources across borders and within countries, such as access to testing, vaccines and treatments, financial support, ability to isolate safely, and access to the internet and digital technologies.[7–9] Many of the effects of COVID-19 are likely to be felt in the longer-term.

Figure 1-1: Weekly number reported by region to the WHO of COVID-19 cases and number of deaths attributed to COVID-19, 13 January 2020 – 31 March 2023.

(Reproduced and adapted from WHO[5] under CC BY-NC-SA 3.0 IGO licence)

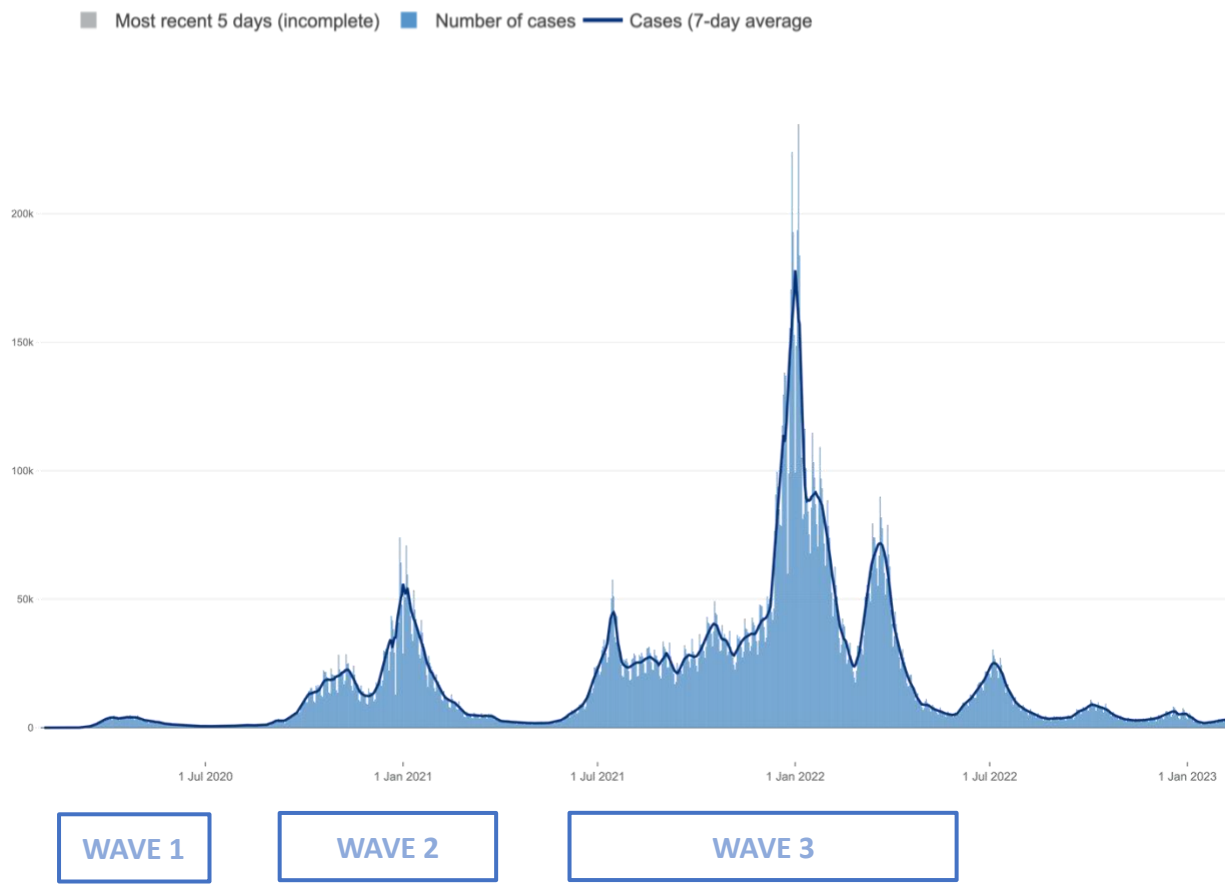


As illustrated in Figure 1.2, epidemiological trends in the UK mirrored the waves seen in Europe and North America, however, patterns of infection varied widely by age-group and region over time which may also reflect temporal differences in testing uptake. Early in the pandemic it became clear that specific groups were at greatest risk of severe outcomes either because of their vulnerability, such as care home residents, homeless people, people with chronic illnesses, or because of their exposure to infection which was usually occupational such as health and social care

workers, public transport workers, or a combination of both.[10] Age is one of the most important predictors for severe infection and the key determinant of outcome.[11,12] There is also clear evidence of increased risk amongst males and some evidence of greater risk among people of non-white ethnicity, although the causality is uncertain.[13–15] In the UK, the ability to track infection rates in different age-groups and regions, variation over successive waves of the pandemic, and the influence of contact patterns, vaccines, variants, and lockdowns, has been made possible by a number of large-scale epidemiological studies which have regularly sampled the general population including the national COVID-19 Infection Survey (CIS), REACT and VirusWatch.[16]

Figure 1-2: 7-day rolling incidence rate of SARS-CoV-2 in England, 1st February 2020 – 1st February 2023

(Reproduced and adapted from [17] which contains public sector information licensed under the Open Government Licence v3.0)



1.2 SARS-CoV-2 virus and syndrome

1.2.1 Viral structure

SARS-CoV-2 virus is a *betacoronavirus* that belongs to the *Coronaviridae* family, a large family of enveloped positive-sense single-stranded RNA viruses that mainly infect mammals.[18] The virus has four viral structural proteins (spike (S), envelope (E), membrane (M), and nucleocapsid (N)) and 15 non-structural proteins, which are sometimes known as viral antigens.[19,20] The spike protein is of particular significance as it is used by the virus to gain entry to host cells and is made up of the S1 and S2 subunits. The S1 subunit consists of the N-terminal domain (NTD) and receptor-binding domain (RBD), the latter binds to angiotensin-converting enzyme 2 (ACE2) receptors on the host cell surface which are mainly found in the lung and upper airway epithelial cells. This RBD-ACE2 binding triggers conformational changes that facilitate the S2 subunit to mediate fusion of the viral and host membranes so that the virus can enter the cell.[21] As the spike protein has been the target of most vaccines and therapeutics, mutations in these domains can have implications for disease control measures.

1.2.2 Clinical syndrome

The SARS-CoV-2 clinical syndrome is characterised by respiratory symptoms which result from viral invasion of lower respiratory tract cells.[22,23] Anosmia was initially considered pathognomonic of infection[23] however, is less commonly seen with newer variants.[24] Severe outcomes are usually related to respiratory failure, thromboembolism, myocardial damage, and encephalitis[3] which can be exacerbated by uncontrolled inflammatory responses to infection and cytokine storms.[25] Conversely, it is also estimated that up to 40% of cases exhibit no symptoms or very mild symptoms.[26]

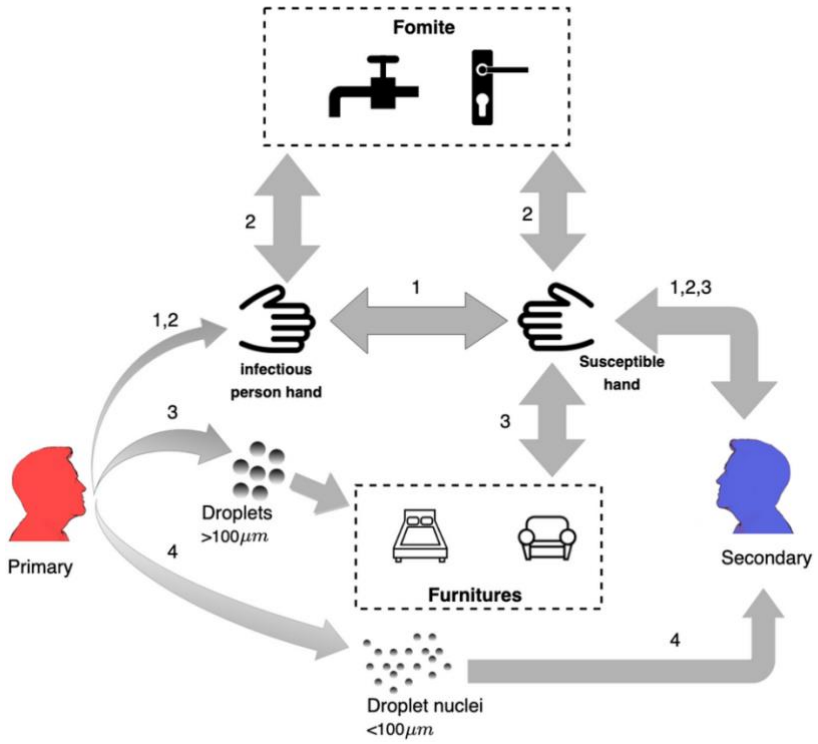
1.2.3 Transmission

Transmission of SARS-CoV-2 occurs mainly through inhalation of airborne droplets or aerosols [27–29] or contact with contaminated surfaces or fomites,[30] although faecal-oral[31] and vertical transmission[32] has been described, Figure 1.3.

Figure 1-3: Modes of SARS-CoV-2 transmission from primary to secondary case

(Reproduced from[33] under CC BY 4.0 licence)

Primary case is shown in red, secondary case in blue, grey arrows marked with number to demonstrate transmission routes: 1) direct contact; 2) indirect contact from fomites; 3) indirect contact through surfaces; and 4) droplet nuclei.

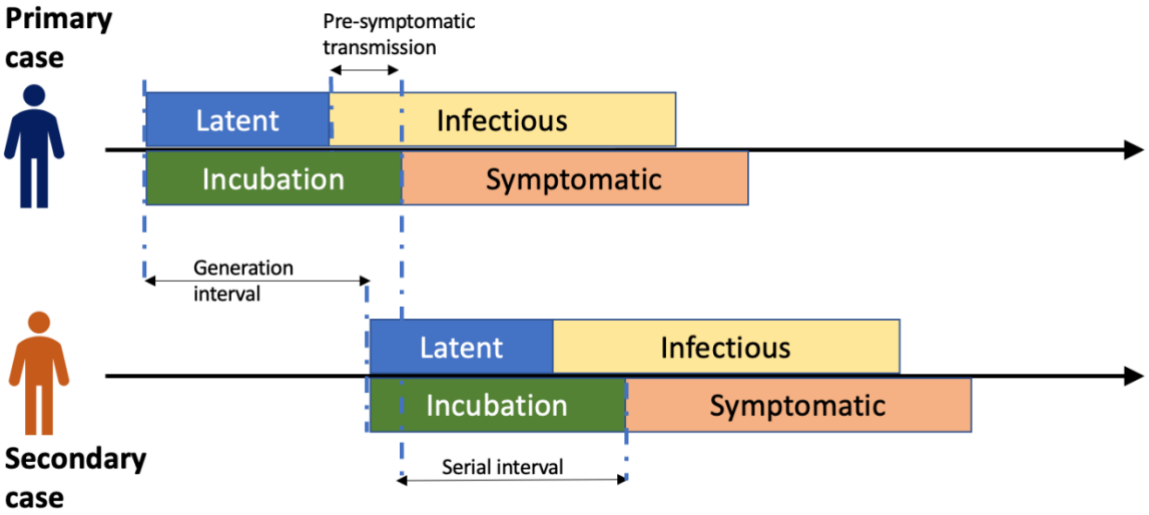


Asymptomatic and pre-symptomatic cases are likely to contribute to transmission in the population, however as these cases are difficult to identify, quantifying their relative contribution has been challenging.[34–36] The SARS-CoV-2 transmission chain between two cases is illustrated in Figure 1.4. The latent period is the time from infection to becoming infectious to others, whereas the incubation period is the interval from infection to developing symptoms.[37] The generation interval is the time between infection in the primary and secondary case, which is very difficult to measure, whereas the serial interval is the time between symptom development in primary and secondary case and is much easier to measure. This is an important consideration for the control of SARS-CoV-2 as there is evidence that the serial interval is shorter than the incubation period (pooled mean 5.2 vs 6.5 days) allowing pre-symptomatic transmission to occur as pre-symptomatic or asymptomatic individuals who do not know they are infected may continue to mix with susceptible individuals and transmit infection.[38,39] The serial interval is also shorter with newer variants which may explain

their increased transmissibility as the time from becoming infected to infectious is shorter.[40] In the first few months of the pandemic it was estimated in 97.5% of symptomatic infections, symptoms developed within 11.5 days of infection, and in 99% within 14 days,[41] which formed the basis of quarantine recommendations for cases, described under Section 1.4.2.

Figure 1-4: SARS-CoV-2 transmission chain

Primary case is shown on top row, secondary case on bottom row, black arrow denotes time. Of note serial interval is shorter than incubation period suggesting that asymptomatic transmission can occur.



1.3 Waves of infection and viral variants in the UK

Viral variants occur because of variations in the viral RNA, such as changes, insertions, or deletions of nucleic acid bases, that occur naturally during viral replication. These alterations can code for mutated viral proteins and persistent replication of these variants can form viral sub-populations. This process can also occur within a single immunocompromised host as the delay in viral clearance by their immune system allows ongoing replication and mutation and as the host continues to shed the virus, sometimes asymptomatically, this facilitates spread to other hosts.[42,43] In some cases, these mutations confer a survival advantage such as increased transmissibility or evasion of host immune responses. This facilitates expansion of sub-populations as transmission persists between hosts and replication continues. Eventually, this variant overcomes the main circulating variant and becomes dominant.

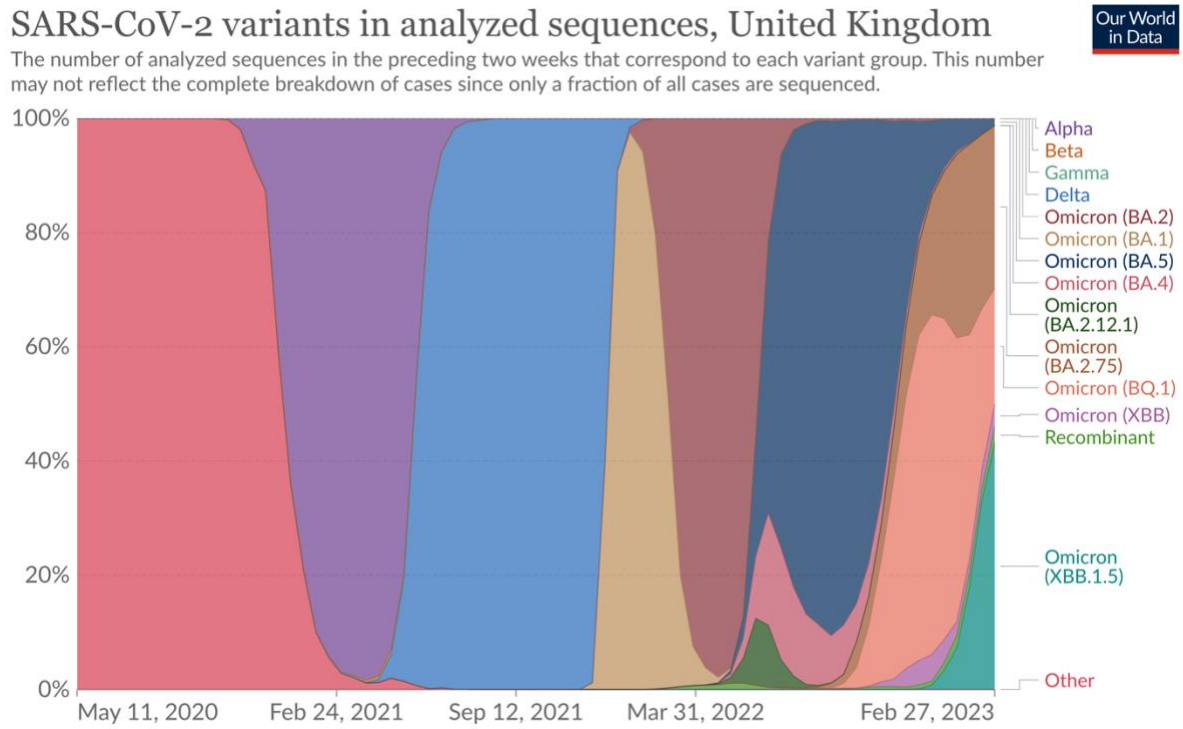
Although RNA viruses are known to mutate faster than DNA viruses, coronaviruses have a proofreading enzyme that is absent from other viruses which slows its mutation rate.[21,44] However due to the large number of SARS-CoV-2 infections that occurred within a short space of time over the pandemic, mutations were observed within the spike protein, the main target for vaccination, within a few months.[21] High rates of transmission within the population applied a selection pressure as viral sub-populations that could survive vaccine-derived immune responses or were more transmissible were able to replicate more easily.[45] It is also considered that persistent viral shedding and replication within immunocompromised hosts, applied further selection pressures and can account for development of viral variants like Alpha.[46]

It is possible to read the sequence of RNA that makes up the SARS-CoV-2 viral genome using whole genome sequencing. Sequencing of viral isolates from clinical samples allows identification of new mutations in the genome that can alter viral proteins and result in emergence of new variants. Variants that are closely related and share similar mutations can be grouped into lineages which are classified according to the Pango lineage naming system.[47,48] These can become variants of concern which may require further public health action if they become sufficiently widespread or exhibit concerning mutations.[49] The COVID-19 Genomics UK Consortium (COG-UK) was the largest genomic surveillance programme in the world and sequenced over 1.5 million isolates with a turnaround time of a few days.[50,51] These genomes were made openly accessible and played a critical role in tracking the emergence of new variants and considering factors such as transmissibility, severity, and ability to evade vaccine-induced immunity. In turn these data informed policy decisions such as booster vaccination programmes in the UK and globally.

Figure 1-5: Timeline of circulating SARS-CoV-2 variants in United Kingdom between 11th May 2020 and 27th February 2023.

(Reproduced under CC-BY licence from Our World in Data[52])

Proportion of variants identified in analysed sequences shown for each date on x-axis. Key variants: wild-type Wuhan shown in dark pink, Alpha in purple, Delta in bright blue, Omicron sub-variants in beige, mauve, dark blue, brown, salmon pink, and teal.



After the first wave of infection, caused by wild-type (Wuhan) strain, subsequent waves of infection have largely been attributed to emergence of SARS-CoV-2 variants; only the most significant ones are described here,[53] Figure 1.5. The wild-type Wuhan variant accounted for most of the infections occurring globally over the first wave with a drop in cases and mortality over the summer of 2020. The Alpha variant of B1.1.7 lineage was first described in Kent in September 2020 and became dominant in the UK and globally over subsequent months.[54,55] This variant initially exhibited several amino acid mutations (notably 14 nucleotide replacements and three deletions),[56] most significantly in the RBD and NTD regions of the spike protein which increased binding to host cells and allowed evasion of host immune responses.[57,58] Transmissibility of this variant was 50% greater than the Wuhan variant with greater infection severity and more cases described in under-20-year-olds.[59,60] This accounted for a rapid rise in cases and hospitalisations in the UK from October 2020

onwards. Incidence and mortality began to drop again from January 2021 once national vaccination had commenced.[61] However, a further surge in infections occurred after the emergence of the B.1.617.2 (Delta) variant in March 2021, first described in India.[62,63] Additional mutations in the spike protein made this variant more resistant than its predecessors to neutralisation from infection-acquired and vaccine-derived antibodies, therefore infections occurred in the vaccinated population, albeit with a lower severity than in the unvaccinated.[64]

In November 2021, the BA.1 Omicron variant with 35 mutations in the spike protein (compared to the Wuhan strain)[65] was first identified in South Africa and was responsible for a rapid and large surge in cases. These mutations allowed the virus to evade antibodies against pre-Omicron variants derived from infection and vaccination, and actions of most anti-viral therapeutic agents.[66–68] Although incidence was high, mortality and hospitalisation rates were relatively low compared with prior variants, largely due to some protection from booster vaccinations.[69–73] Omicron has developed a number of sub-lineages, BA.2-5, that have accounted for ongoing transmission with frequent reinfections, however clinical acuity has remained low.[68]

1.4 Disease control measures

1.4.1 Testing

Significant advances have been made in diagnostic testing for SARS-CoV-2, which is required to identify cases and prevent spread to susceptible individuals. Although access to testing was limited early on, a monumental effort meant that tests were developed rapidly after the virus was first identified in December 2019.[74–76] It is important to assess the diagnostic performance of these tests and consider applications. Tests either assess for evidence of current infection by detecting virus in clinical specimens from the upper respiratory tract or measure immune responses to infection by detecting antigen-specific antibodies or T-cell responses in blood.[77] Where specificity describes the proportion of truly negative samples that have been correctly identified as negative by the test, sensitivity describes the proportion of truly positive samples that are positive with the test.[78]

Reverse Transcriptase Polymerase Chain Reaction (RT-PCR) testing detects viral RNA in clinical samples and is the gold standard. Although this test has high sensitivity and specificity, there are significant disadvantages including cost and need for specialised equipment and highly trained staff.[77] Turnaround times are usually less than 24 hours, although can take up to 48 hours depending on the distance to the laboratory and at times during the pandemic were longer than this due to laboratory capacity.[79] In addition, a large meta-analysis conducted in June 2020 reported that RT-PCR can continue to detect viral RNA for up to 83 days from symptom onset in samples from the upper respiratory tract, however the mean was 17 days.[80] As it is only possible to culture live virus for up 9 days following symptom onset, it is likely that any viral RNA isolated after this time is non-viable.[80] This has implications for the diagnosis of infections occurring within 90 days of each other, as reliably distinguishing discrete episodes may only be possible using whole genome sequencing to identify distinct genomes.

In the UK, mass national scale-up of testing capacity early in the pandemic meant that tests were only available to all symptomatic individuals in the community from the summer of 2020. Although early tests were performed at National Health Service (NHS) hospital and public health surveillance laboratories, a national network of new 'Lighthouse' laboratories were established, and existing private services were commissioned to perform PCR tests as part of the NHS Test and Trace programme (described in more detail under Section 1.4.2 and in Chapter 4).[81] In April 2020, the UK government announced their national testing strategy based on five pillars, three of which focussed on PCR testing: NHS inpatient settings (Pillar 1); NHS staff and social care settings and later at satellite testing sites and using home testing kits for the wider population (Pillar 2); and population surveillance testing (Pillar 4).[82,83] Limited testing early on had implications for incidence estimates that could only rely on hospitalisation and mortality data to model the number of cases.

Point-of-care tests using lateral flow technology to detect viral antigens became available in the second half of 2020 and have since been widely deployed in community and health and social care settings. These tests are easy to use so can be operated without training, are cheap, and results are ready within 15-30 minutes.[77] A systematic review that pooled 24 studies found that sensitivity of Lateral Flow Devices

(LFDs) using PCR as gold standard was 38-99% and specificity was consistently > 93%.[84,85] LFD sensitivity was greater amongst symptomatic individuals and in those infected with the Omicron variant in a large national evaluation funded by the UK government.[85] LFDs changed the testing landscape as they were freely available to the public and facilitated prompt case isolation, therefore limiting further transmission.

Detection of virus-specific IgM – the antibody produced within a few days of infection – has been used to diagnose infection however, due to lag in antibody production recommendations advise against their use in early infection.[86,87] Presence of virus-specific IgG – antibody produced from two weeks after infection – can identify previously-exposed individuals. So far vaccines have targeted the viral spike protein meaning that presence of anti-spike antibodies cannot distinguish prior exposure from vaccination, therefore anti-nucleocapsid antibody testing is recommended for this purpose.[87]

1.4.2 Non-pharmaceutical control measures in the general population

The strain on health services from a rapid, unprecedented increase in the number of patients requiring inpatient care and critical care, triggered urgent public health action with the aim of bringing down the basic reproduction number (R_0). The R_0 is an epidemiological term used to describe the transmissibility of an infectious agent. It is estimated by calculating the number of susceptible contacts that would become infected from a single case without control measures in place.[88] This can be affected by biological, socio-behavioural, and environmental factors so it varies as behavioural patterns change and the viral mutations affect transmissibility - a particular challenge over this pandemic.[89,90] This concept is mainly applied to outbreak modelling where an outbreak is an increase beyond the expected number of cases in a specific location and can trigger additional control measures.[91]

The R_0 can be used to predict the size of an outbreak (by considering the size of the infected and susceptible population at one time), estimate the population proportion that need to be vaccinated to control an outbreak (by reducing the susceptible population), and monitor whether control measures are working (depending on the direction of R_0 growth). R_0 greater than 1 suggests an outbreak is growing and if it falls

below one, it is expected to stop.[90] As asymptomatic and pre-symptomatic SARS-CoV-2 infections were commonly underestimated, this impacted on reliability of R_0 estimates which varied significantly in line with the setting and were subject to reporting lag.[89]

To bring the reproductive number below one, national lockdowns were enforced to break the chain of transmission. During lockdowns in the UK, people could only leave their homes for essential activities, schools and non-essential businesses were closed, and social distancing was enforced in public areas. Because of school closures classrooms moved online, however this has significantly disadvantaged children from more socio-economically deprived homes.[92] In the UK there were three national lockdowns in response to reductions in hospital bed capacity, the first was between 23rd March 2020 and 4th July 2020, the second between 5th November and 2nd December 2020 and the final one was between 6th January and 8th March 2021, with varying limits on socialising in the intervening periods.[93]

In addition to these lockdowns, non-pharmaceutical interventions (NPIs) were introduced on a national level. As described earlier, mass testing was introduced for symptomatic cases and in health and social care settings to allow detection and isolation of cases. In view of known asymptomatic transmission, cases and their contacts were quarantined initially for 14 days from first symptoms (the maximum incubation period)[41] to prevent onward transmission which was supported by evidence from a rapid Cochrane review.[94,95] Initially quarantine was recommended based on symptoms alone, however once testing capacity increased, PCR tests were made available to the public through the NHS Test and Trace programme.[81] This enabled active case finding in the community and contact tracing by a team of trained assessors. Test and Trace also introduced a mobile phone application that used GPS and manual location check-in to monitor proximity to cases and notify contacts.[96]

To limit social contacts, varying restrictions were enforced on indoor and outdoor mixing over the course of the pandemic outside of lockdowns. Recommendations to maintain a 2-metre distance were based on experimental data describing the distance the viral droplets and aerosols can travel.[97,98] Individuals belonging to groups at higher risk of severe outcomes, including pregnant women, were advised to avoid leaving

their house for any reason.[99] Public health campaigns were also launched to inform the public of the importance of hand-washing and avoiding contact with others.[99] Guidance on ventilation in all public spaces and health and social care facilities suggested opening windows if possible and ensuring ventilation systems introduced fresh outdoor air instead of re-circulating indoor air.[100]

As the virus is predominantly transmitted through airborne droplets, with some transmission from fomites (see 1.2.3), recommendations on personal protective equipment (PPE) were introduced. In healthcare settings, aprons or gowns, gloves, face shields or goggles, and masks were recommended for all clinical contact.[101] The types of masks varied from filtering facepieces (FFP) to fluid-resistant surgical masks depending on the level of exposure. However, PPE shortages early on meant that several health and social care facilities were unable to meet these standards.[102] In July 2020, face coverings became mandatory in all indoor public areas including public transport.[103]

To prevent ingress of cases and novel variants from overseas, a “stay in the UK” regulation was brought in between March 2020 and May 2021, which advised against all overseas travel.[104] A traffic light system was implemented for different countries based on genomic surveillance data. This dictated the UK entry requirements which ranged from quarantine in a hotel or at home, screening test only, or no restrictions.[105] Samples from PCR-based screening at the border were prioritised for viral sequencing as part of the ongoing surveillance programme.

Taken together, the total actions across the globe were the largest scale simultaneous implementation of NPIs to date.[94]

1.4.3 Vaccination and therapeutics

An unprecedented effort by researchers across the scientific community led to development of 242 vaccine candidates globally, 50 of which have been approved for clinical use.[106] At the time of writing, eight vaccines had been approved in the UK, three of which were deployed in late 2020 / early 2021, Table 1.1.[106]

The first was the BNT162b2 mRNA vaccine, produced by the Pfizer-BioNTech collaboration, which was approved by the Medicines and Healthcare products Regulatory Agency (MHRA) and rolled out on 8th December 2020.[107] This was followed closely by the ChAdOx1 nCOV-19 Oxford-AstraZeneca vaccine, a non-replicating viral vector vaccine on 30th December 2020[108] and mRNA-1273 Spikevax on 8th January 2021,[109] an mRNA vaccine produced by Moderna. These vaccines all consisted of primary course of two vaccine doses with a recommended interval of four weeks. Initial vaccine rollout and subsequent boosters were prioritised for high-risk groups including care home residents (top priority), health and social care workers, adults older than 65 years, individuals with chronic diseases, immunosuppression, severe mental illness, pregnant people, household contacts of immunosuppressed individuals, and informal carers.[110]

Table 1-1: SARS-CoV-2 vaccine candidates approved for use in UK (as of 4th March 2023)

Manufacturer *(vaccine candidate)*	**Vaccine type**	**Country of development**	**Number of countries approved in**	**Date of approval in UK**
Janssen, Johnson & Johnson *(Jcovden)*	N-R VV	Netherlands	113	28th May 2021
Moderna *(Spikevax)*	RNA	USA	88	8th January 2021
Moderna *(Spikevax Bivalent Original / Omicron BA/1)*	RNA	USA	38	15th August 2022
Novovax *(Nuvaxovid)*	Protein subunit	USA	40	3rd February 2022
Oxford / AstraZeneca *(Vaxzevria)*	N-R VV	UK	149	30th December 2020
Pfizer / BioNTech *(Cominarty)*	RNA	USA	149	8th December 2020
Pfizer / BioNTech *(Cominarty Bivalent Original / Omicron BA.1)*	RNA	USA	35	3rd September 2022
Valneva *(VLA2001)*	Inactivated	France	33	14th April 2022

N-R VV non-replicating viral vector

In view of a rapid increase in cases in the winter of 2020, the dosing interval was extended from the manufacturer-approved four weeks to eight-to-twelve weeks to increase coverage.[111] Due to unparalleled operational coordination nationally, 70% of the UK population had received their first dose by 4th August 2021 and second dose by 16th December 2021.[112] In common with other countries, the UK has adopted a national programme of booster vaccination, primarily driven by evidence of some additional protection during surges of infection with new variants within immune populations.[113] A full description of the rollout of vaccination in the UK is beyond the scope of this thesis, however a timeline illustrating the deployment of vaccinations in the UK is listed in Table 1.2.

Table 1-2: UK COVID-19 vaccination timeline (December 2020 – September 2022)

Date	Vaccination policy
	Primary vaccination course
December 2020	Phase 1: groups according to risk of COVID-19 mortality*±
April 2021	Phase 2: adults 16-50-year-olds not in high-risk groups±
August-September 2021	Young people <16 years old at higher risk#
August-December 2021	Young people 5-16 years old not in a high-risk group
	Booster vaccinations – autumn 2021
September 2021	Residents of care homes for older adults, ≥50-year-olds, health & social care workers, 16–49-year-olds from higher risk group#, carers, and household contacts ≥16 years old
November 2021	18-49-year-olds
December 2021	16-17-year-olds, 12-15-year-olds at higher risk# or household contacts of immunosuppressed individuals
	Booster vaccinations – spring 2022
February 2022	≥75-year-olds, residents of care homes for older adults, ≥12-year-old and immunosuppressed
	Booster vaccinations – autumn 2022
September 2022	Residents of care homes for older adults, ≥50-year-olds, health & social care workers, 5-49-year-olds in high-risk group#, 5-49-year-old household contacts of people with#, 16-49-year-old carers

*High risk groups include (in order of priority): (1) residents of care homes for older adults, (1) staff in care homes for older adults, (2) frontline health and social care workers, (2) ≥80-year-olds, (3) ≥75-year-olds, (4) ≥70-year-olds, (4) 16-69-year-olds in high risk group#, (5) ≥65-year-olds, (6) 16-65-year-olds in at-risk group, (7) ≥60-year-olds, (8) ≥55-year-olds, (9) ≥50-year-olds.

±Rolled out incrementally.

#Higher risk groups include individuals with chronic diseases (respiratory, cardiac, renal, digestive, neurological, endocrine), obesity, immunosuppression, pregnancy.

Although vaccines have had the greatest impact on infection incidence, developments in the therapeutics pipeline have improved outcomes.[114] Treatments aim to reduce disease severity and duration and reduce risk of transmission. These can target the virus itself (antivirals, monoclonal antibodies), the immune response to infection (corticosteroids, monoclonal antibodies) or the sequelae of infection such as coagulopathies (anticoagulants).[115] Differing routes of administration and sites of action can influence the choice of clinical setting that these treatments are administered in. Monoclonal antibodies have received particular attention - these are specific proteins that can treat infections, cancers, and chronic immunological conditions as they target pathogenic proteins from an infectious agent or from the host's own immune system.[116]

Rapidly established multi-site clinical trials such as RECOVERY[117] and PRINCIPLE[118] have reported efficacy and informed guidance on a range of treatments including corticosteroids, anticoagulants, repurposed drugs, antivirals, and monoclonal antibodies.[115] On 20th August 2021, Ronapreve which attaches to the viral spike protein to prevent cell entry by the virus was the first monoclonal antibody approved for use by the MHRA however subsequent evidence of limited efficacy against Omicron variants has limited its use. [115,119] This was closely followed on 4th November of the same year by Molnupiravir, the first approved antiviral in the UK, which prevents viral replication.[120] Community therapies are also recommended to treat non-severe illness in high-risk immunocompromised populations.[121] However, data are still being collected on efficacy of these different treatments on varying populations and the implications on a population-level have not been described.[122,123]

1.5 SARS-CoV-2 in care homes

Early reports from the UK and internationally described severe outbreaks of SARS-CoV-2 in care home residents associated with high mortality. However, the scale of the problem in the UK only became fully apparent at the end of April 2020 when mortality data were first reported.[124] This delay was largely related to lack of

surveillance and research infrastructure in care homes (described in Section 1.7) which contributed to disproportionate mortality in this population. By April 2022, between 15% to 45% of all COVID-19 associated deaths in the UK and Europe had occurred in care home residents.[125] To put this in context, care home residents in the UK account for just 0.7% of the population in England (410,000 residents in 10,000 care homes for older adults).[124]

1.5.1 Overview of care homes in England

In England, care homes for older adults provide a mixture of residential and nursing care alongside specialist dementia care.[124] Residential care homes provide a home for people who may require assistance with personal care or additional support for those who cannot live independently. For those with specific medical needs additional care from qualified nurses and carers is provided in nursing homes. The majority of residents are female and older than 80 years.[126] Average life expectancy of a resident in a UK care home is 12 to 42 months, lower for those requiring nursing care.[127–129] Frailty was found to affect between 19% and 75% of residents in one meta-analysis from seven countries (excluding the UK)[130] and is more prevalent than in age-matched community-dwelling peers.[131] Multi-morbidity is common; a cohort study of 11 UK care homes found that residents had an average of 6.2 medical diagnoses, one-third were malnourished and median Barthel Index of 9, suggesting total physical dependence.[132]

On average care homes in England have 29.5 beds, ranging from 1 to 215.[126] In 2019/2020, approximately 50% of beds for older adults were funded by the Local Authority (LA) predominantly in smaller care homes, whereas the rest were self-funded.[133] Approximately 95% of beds are provided by the independent sector, made up of for-profit companies and not-for-profit voluntary organisations, with the remainder provided by local councils.[134] Staff turnover is high, in 2021/2022 mean turnover rate was 29% overall, 52.6% for junior carers and 44% for registered nurses.[135]

1.5.2 SARS-CoV-2 clinical syndrome in care home residents

The immunopathogenesis of severe COVID-19 disease in older people is not fully understood however in part it is likely to be related to age-associated changes to the

immune response, such as inflammageing and immunosenescence.[136] Inflammageing is a chronic, low-level inflammation which is probably caused by the accumulation of abnormal or senescent immune cells over the life course in response to physiological stressors.[137] These abnormal cells secrete pro-inflammatory cytokines, engaging key signalling pathways which become chronically and inappropriately stimulated. This causes impaired responses to infectious challenges producing hyper-inflammation and tissue damage.[136,137]

Immunosenescence can affect both the adaptive and innate immune responses. The adaptive immune response is impaired as the T and B cell repertoire, cells responsible for antibody production, is limited by poor differentiation of their naïve precursor cells in the bone marrow. Reduced phagocytosis and altered Interferon-gamma responses result in impaired innate immune response.[136,138] Taken together, these changes to immune responses of older people make them susceptible to severe disease from SARS-CoV-2 infection.

In addition to age-related changes to immunity, factors that are known to impair response to infection in this population are nutritional deficiencies, reduced mobility, frailty, and medical co-morbidities.[139–141] Features of the built environment, such as air flow and ventilation, are probably associated with the risk of acquiring or transmitting infection.[142,143] Proximity to staff who are in contact with infected individuals in the community, or to newly admitted residents who were exposed before admission (in the community or in hospital), increase the risk of infection.[144]

1.5.3 Measuring the impact of SARS-CoV-2 in care homes.

Lack of testing in the first wave of the pandemic makes it challenging to estimate the full impact of SARS-CoV-2 in care home residents, however there was a significant peak in all-cause mortality among care home residents when compared with the 5-year average, Figure 1.6. These excess deaths are likely to represent deaths from COVID-19 although the cause of death was often based on clinical diagnosis alone, impairing accuracy of death certificates which mortality estimates are based on.[145] The subsequent drop in mortality may be because deaths in more vulnerable residents occurred earlier in the first wave, leaving a cohort of immunologically robust residents

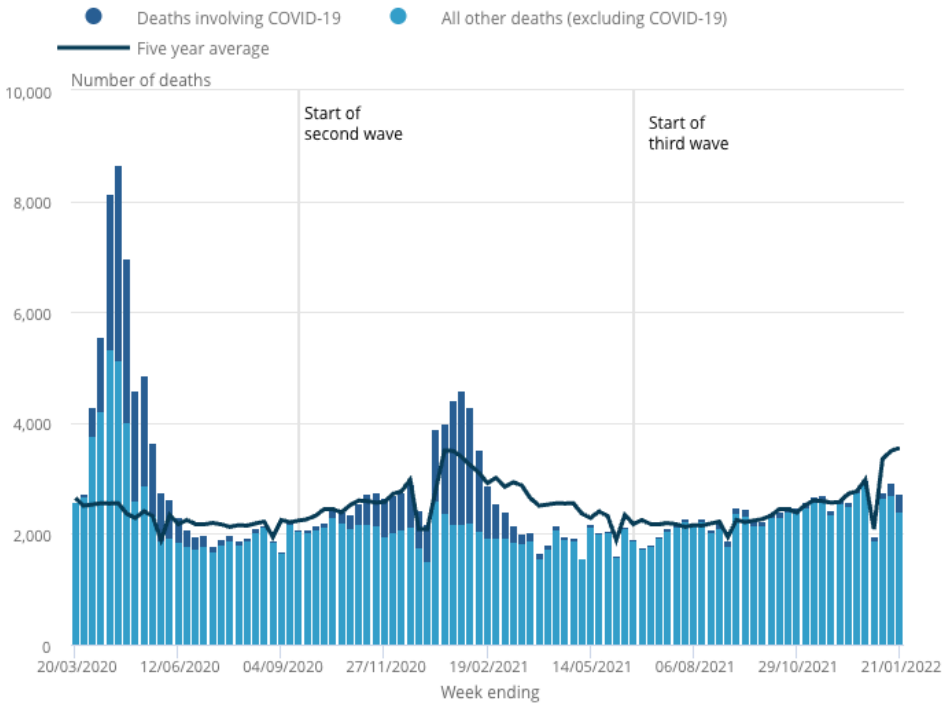
who had survived COVID-19.[145,146] This mortality-displacement has also been known as the 'harvesting' effect,[147] although given the devastating loss of lives over the pandemic, this term is not considered acceptable.

Although data are limited, bed occupancy was also lower than average following the first wave.[148] This is probably because of the large number of deaths at the start of the pandemic, restrictions on bed capacity in efforts to social distance residents, and perceptions of risk which led to avoidance of care home admission.[149] Dementia and Alzheimer's disease remained the leading cause of death for female residents over the pandemic, however COVID-19 was the leading cause of death amongst male residents in the first wave and second leading cause in all residents in subsequent waves.[149]

Figure 1-6: Weekly number of deaths amongst care home residents between 14th March 2020 and 21st January 2022 in England

(Reproduced from[149] which contains public sector information licensed under the Open Government Licence v3.0)

Bars in dark blue show deaths where COVID-19 was indicated as a cause on the death certificate, pale blue bars show all other deaths. The black line indicates the 5-year average number of deaths.



The delay in detecting the degree of severity of the SARS-CoV-2 pandemic in care home residents was largely attributed to the absence of data infrastructure and surveillance in care homes. Coupled with limited testing in the first wave of the pandemic this made it difficult to determine the extent of infection and outbreaks in this setting. Lack of data also undermined efforts to understand which care home factors increase or decrease rates of transmission to inform public health interventions. As there is no national registry of care home residents or care home 'flag' it is not straightforward to link deaths occurring in hospital to care homes.[150] Although all care providers report deaths to the Care Quality Commission (CQC), the independent national regulator of health and social care, this has a reporting lag of a few days and it was only on 10th April 2020 that they started reporting COVID-19 deaths separately, with a subsequent rapid escalation in the number of deaths reported.[151] Additionally, up-to-date bed occupancy figures were unavailable therefore registered beds were used as the mortality denominator.[150] It is likely that mortality was under-estimated as bed occupancy was low and approximately a quarter of deaths occurred in hospital.[145]

During the pandemic a range of data collection methods from care homes were set up to inform public health policy and strategy in England. These included the CQC mortality tracker, as described above, which collected data on deaths.[152] Capacity Tracker, launched in 2019, was mandated over the pandemic and collected data on bed occupancy, staffing, hospitalisations, and PPE access from care homes who completed the tool almost daily.[153] The Department of Health & Social Care (DHSC) introduced a COVID-19 care home dashboard for policymakers that was regularly updated through results of Pillar 1 and 2 tests and hospitalisation records. In May 2020, the VIVALDI survey was conducted across all care homes for older people in England collecting data on infection prevalence, hospitalisations and deaths, and disease control measures.[154] These results were reported nationally and informed the establishment of the VIVALDI observational study in a subset of care homes, a national UCL-led surveillance study funded by DHSC[155] that provided up-to-date data on infection incidence, reinfection, variants of concern, and vaccine effectiveness (described in Chapter 3).

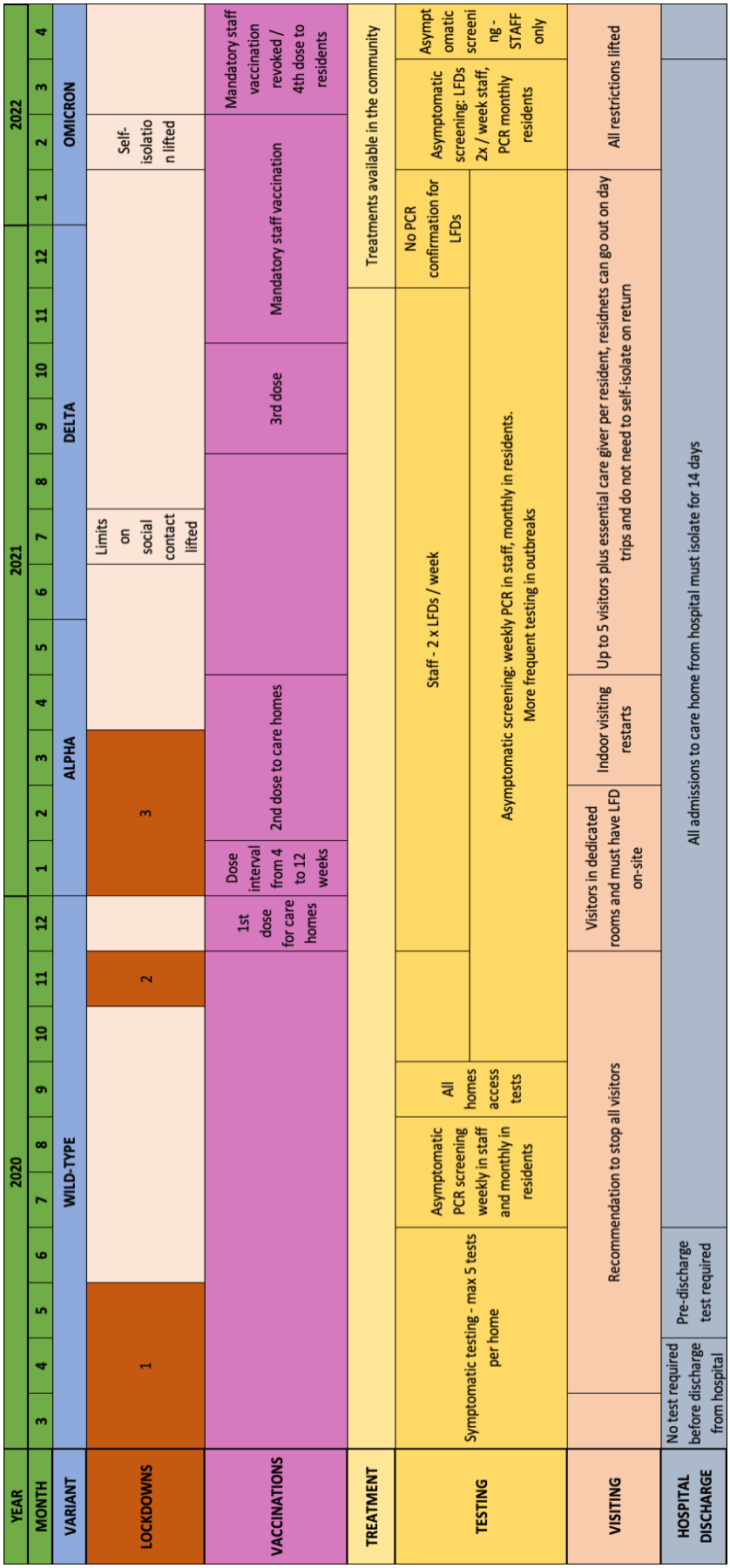
1.6 Strategies to reduce transmission of SARS-CoV-2 in care homes.

Interventions to limit the introduction and spread of SARS-CoV-2 infection within care homes can be divided into five main domains, as summarised in a recent Cochrane review (published in 2021).[156] These comprise measures to regulate infection entry such as screening and isolating new admissions, restricting visitors and new admissions; measures aiming to reduce transmission by regulating contacts including barrier nursing, use of PPE, and enhanced cleaning; surveillance measures like asymptomatic screening or syndrome-based testing; measures aiming to contain outbreaks such as cohorting where different staff members care for infected and uninfected residents in separate areas of the facility; and multi-component measures where these are implemented in combination. Key government initiatives to limit infection spread included introduction of the Infection Control fund which supported care homes in paying sickness payments to staff,[157] and the Capacity Tracker tool which allowed identification of facilities that needed additional support to access resources such as PPE.[153,158] Figure 1.7 outlines key care home policy changes in England over the pandemic.[156,159]

Restrictions on non-essential visitors were first recommended in mid-March 2020 and were shortly followed by policy advising closure of care homes to visits from family and friends, which remained in place for one year.[160]

Despite frequent breakthrough infections with newer viral variants such as Omicron, mounting evidence suggests that vaccination of residents and staff reduced the incidence of infection and severe outcomes.[161–164] As previously described, vaccinations against SARS-CoV-2 were rapidly deployed to care home staff and residents from December 2020 onwards with booster vaccinations to target waning immunity.[165] Evidence that vaccination of staff can protect residents from infection prompted mandatory staff vaccination policies, however these were removed five months later due to ethical implications and concerns of staffing shortages.[166]

Figure 1-7: SARS-CoV-2 control in care homes - key policies March 2020–April 2022



1.6.1 Testing for SARS-CoV-2 in care homes

Systematic SARS-CoV-2 testing of hospital inpatients prior to discharge into care homes was mandated three weeks after national lockdown. It is possible that infections were brought into care homes by people who were discharged from hospital before this policy was introduced.[167–170] Within care homes, testing was initially limited to five tests per home following a symptomatic case.[158] Asymptomatic screening of all residents and staff was announced in July 2020, and widely available from September 2020.[171] This involved weekly testing of staff and monthly testing of residents using PCR. Additional more frequent LFD testing of staff was introduced in December 2020. Asymptomatic testing of residents was retracted from April 2022 onwards and asymptomatic screening of staff was stopped shortly after in December 2022.[172] The Cochrane review found some evidence of a reduction in outbreaks, hospitalisations, and deaths amongst residents in homes performing asymptomatic screening from one observational and four modelling studies.[156] A modelling study published in April 2022 found little impact of asymptomatic testing on outbreak risk when compared with no testing however suggested that daily LFD testing of staff was a more effective strategy.[169] As NPIs were brought in simultaneously, estimating impacts in isolation is challenging.

1.7 Challenges of data collection and research in care homes

Due to logistic barriers such as the absence of research infrastructure to facilitate blood or data collection, performing research on SARS-CoV-2 in care homes is challenging.[173] Strategies are required to ensure those who lack capacity are not excluded from research studies and to consent new staff and residents who have a high turnover however, this can be particularly challenging during a pandemic.[174–176]

Linkage of pseudonymised routine datasets from the general population has facilitated research and modelling on SARS-CoV-2 at scale that is generalisable, timely, and relevant. However this approach was difficult to replicate in care homes as there is no official database of care home residents or staff in the UK and using primary care databases to identify these individuals is unreliable.[150,171,177,178] Address-based matching can be inaccurate: a diagnostic accuracy study found that sensitivity of postcode matching, the most common approach to linking residents to hospital

admissions, was 78.2-90.2% as it often identified private residences as care homes.[179] In addition, residents may be missed as care homes frequently change name and ownership, there can be errors in recording of the postcode, and individuals may only reside in a care home for a short period while awaiting care in the community or as respite. Linkage between datasets also requires a common identifier such as the NHS number, a unique identifier allocated to every person in the UK.[178] Prior to the pandemic most care homes did not record NHS numbers for residents. Matching based on multiple variables like name, address, and date of birth is error prone as small differences in spelling or format may prevent linkage.[150] These problems are partly because of the very limited integration of health and social care, with most care homes relying on paper records and only 40% using fully digital records,[180] although this number is increasing.

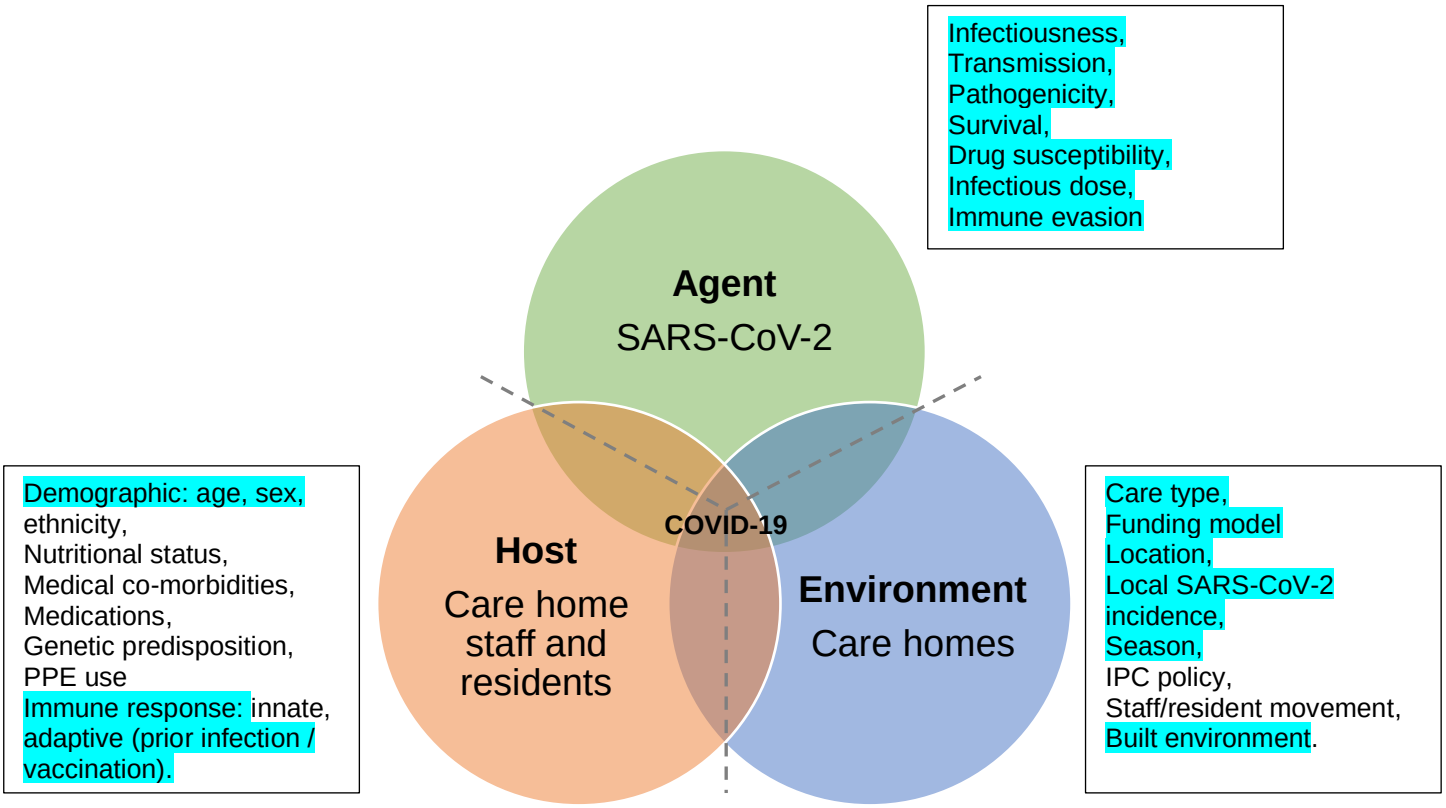
1.8 Modelling risk factors for SARS-CoV-2 infection in care homes

Factors associated with SARS-CoV-2 spread within different settings can be summarised using the epidemiological triad. This demonstrates how a susceptible individual within a conducive environment can become infected when exposed to a sufficiently virulent pathogen.[181] The triad consists of three main components that are all required for disease to occur: agent, host, and environment.[182]

Figure 1.8 outlines the application of the epidemiological triad to COVID-19 in care homes. Key agent factors that can contribute to risk of infection are also associated with the viral variant and include pathogenesis, transmission mode and survival, susceptibility to treatments, immune evasion, and infectious dose. Host factors within the care home population include demographic characteristics, nutritional status, medical co-morbidities, medications, genetic predisposition to infection, use of PPE, and immune response to infection which can also be affected by prior exposure or vaccination. Environmental factors in care homes include geography, local infection incidence, season, funding model and care type, infection prevention policy, staff / resident movement, and features of the built environment.[181,183]

Figure 1-8: Epidemiological triad for COVID-19 infections and outbreaks in care homes

Each circle represents a component of the triad, with overlap showing where factors converge to cause COVID-19 disease. Breaking the connection (grey dotted line) may stop further infections. Factors associated with each component are listed in the boxes. Factors that I have explored in my PhD thesis are highlighted in turquoise.



PPE Personal Protective Equipment IPC Infection Prevention & Control

Using this model, it is possible to consider the interplay between different factors and to assess which situation-specific components can be altered to control disease. It can also help to identify significant gaps in existing knowledge and areas that should be prioritised for research and surveillance.

1.9 Research gaps at start of PhD

In November 2020 when I started my PhD, a disproportionate number of older residents of care homes had died from COVID-19. Although this was likely to be related to the widespread circulation of SARS-CoV-2 within care settings in the first wave and the vulnerability of these residents to infection, specific factors that had contributed to the magnitude of infection in these settings were not known, largely

because of incomplete data. Incidence of infection was difficult to measure because early testing programmes did not link tests to care homes, and access to testing in the first pandemic wave was restricted to outbreaks, therefore asymptomatic cases were missed. This made it challenging to quickly identify why some care homes were experiencing the highest burden of infection which could inform public health control measures. As there was no routine surveillance for infection in the care sector prior to the pandemic and variation in care structures meant that research infrastructure was patchy, it was clear that setting-specific research was crucial.

My PhD was hosted within the VIVALDI study which I played a key role in establishing. This was a government-funded, UCL-led national surveillance study of COVID-19 in care homes. This involved data linkage for participating homes and undertook serial blood sampling in care home residents and staff to measure SARS-CoV-2 infection and immunity. The study was swiftly established in the first lockdown soon after the asymptomatic SARS-CoV-2 screening programme in care homes was announced. This programme made it possible to link tests to care homes at scale, creating a 'registry' of care home staff and residents for the first time ever, a detailed description of methods and my involvement can be found in Chapter 3.

I was able to capitalise on my role in establishing VIVALDI to develop the research aim and objectives of my PhD. I decided to consider the agent, host, and environmental factors on a facility-level that were associated with SARS-CoV-2 infections and outbreaks in care homes (Section 1.8) to identify significant factors that may help to distinguish homes at highest risk. A full analysis of this epidemiological triad is beyond the scope of a PhD thesis, and I was limited in the data that were available to me. I chose to focus on facility-level factors as individual-level factors had been described in the literature and I had limited access to host factors such as medical co-morbidities, and medications and agent factors like sequencing data. However, describing significant associations could inform disease control strategies and identify important directions for future research. In Figure 1.8 I have highlighted factors that I have considered in my thesis, based on those with limited published evidence for which I could access reliable data.

To start, I planned to describe the existing evidence on facility-level risk factors for SARS-CoV-2 infections and outbreaks in care homes from around the world (Objective 1, Chapter 2). I subsequently considered agent factors by applying different approaches to measuring disease burden in care homes and distinguishing between viral variants (Objective 2, Chapter 4). This was based around the hypothesis that there is variation in the burden of SARS-CoV-2 infection between care homes which is also affected by the variant, and it can be estimated from seroprevalence surveys and asymptomatic PCR/LFD screening. Next, I considered host factors that protect staff and residents against infection, by modelling the immune responses to infection and vaccination and considering how to measure them on a facility-level (Objective 3, Chapter 5). Finally, in relation to environmental factors, I tested the hypothesis that variation in environmental factors between the homes affects their susceptibility to the introduction and transmission of infection (Objective 4, Chapter 6). I aimed to test these hypotheses in my PhD, and to this end have designed a series of studies which are hosted within the VIVALDI study.

1.10 Aims and objectives of my thesis.

Aim

To describe the burden of SARS-CoV-2 infection in care home staff and residents and investigate facility-level factors associated with infection and outbreaks in residents over the first two years of the pandemic.

Objective 1

To undertake a scoping review of the existing literature on care home factors that are associated with the SARS-CoV-2 infections, outbreaks, and large outbreaks within care homes (Chapter 2).

Objective 2

To test the hypothesis that it is possible to measure the proportion of staff and residents infected with SARS-CoV-2 and that there is substantial variation between care homes (Chapter 4).

Objective 3

To test the hypothesis that care home residents and staff develop durable SARS-CoV-2 antibody responses following infection and vaccination and these responses can be measured on a facility-level (Chapter 5).

Objective 4

To test the hypothesis that care home characteristics are risk factors for SARS-CoV-2 infection and outbreaks and that factors associated with infection ingress differ from those associated with transmission (Chapter 6).

Chapter 2

Objective 1: Scoping review of care home factors associated with SARS-CoV-2 infection, outbreaks, and large outbreaks in care homes.

To identify knowledge gaps and inform the analytical objectives of my thesis (Objectives 2,3,4) I performed a scoping review of the published literature. I considered facility-level agent, host, and environmental factors associated with SARS-CoV-2 risk in care homes. I split outcomes into those describing infection ingress and transmission as this may provide more granular insights into areas to focus strategies to prevent infection.

I found 31 eligible longitudinal or cross-sectional studies. The majority accessed routine surveillance data and only one-fifth detected asymptomatic infections. None considered how viral variants or vaccine-induced immunity affected infection risk, although one reported the relationship between naturally acquired immunity and outbreak risk. Few studies considered IPC measures or the built environment in detail. The main environmental factors associated with infection ingress were local factors such as infection incidence and socio-demographic composition of the local population, or larger size of the facility. Host and staffing factors were additionally associated with infection transmission, although data were limited.

This review has highlighted that infection burden varies between care homes and is likely to be associated with facility-level factors. Factors associated with ingress were predominantly related to the local population and appeared to differ from those associated with transmission of infection. This suggests that preventing infection ingress into care homes is extremely challenging but limiting transmission may be possible. However, there remain substantial gaps in evidence on factors including the influence of viral variants (agent), immunity (host), and built environment (environment) on infection risk which I will explore in Chapters 4-6 of this thesis.

2.1 Introduction

The high mortality associated with COVID-19 in care homes across many countries, made it apparent that measures were urgently needed to curb the spread of infection into these settings. Although risk factors for infection and severe outcomes had been described in the literature, older people and care home residents were under-represented precluding reliable conclusions.[150–152] Factors consistently associated with worse outcomes from infection are advanced age and presence of multiple co-morbidities;[11,12] features that predominate within the care population.

Care homes are unique settings as they are residential communities that provide a home for people with different care needs. As described in Chapter 1, they vary in their design and layout, the type of funding they receive, and the population they serve: who range from requiring residential care only, assistance with activities of daily living such as washing or dressing, to more specialised nursing care.[126,184] Most have a high resident turnover due to new admissions from hospitals and from the community and low life expectancy in this frail population.[126,185] Staff usually live in the community, but wages are low, which means they are more likely to rely on public transport or car sharing, live in households of multiple occupancy, and co-habit with other frontline workers, which substantially increases their exposure to the virus relative to the general population beyond their occupational exposure.[186] They may also work across multiple sites and for care agencies that cover unfilled rota gaps.[135,154] The building layout and design may also influence spread of infection, as has been well-described in other settings where factors such as crowding, and air flow have been associated.[187–189]

Ingress of infection into the care home can result in a case or an outbreak. SARS-CoV-2 outbreak definitions vary across settings however in the absence of widely available testing, the occurrence of one or two cases in the care home is often used.[190,191] The occurrence of a case or an outbreak shows that infection has entered the facility, however the size and duration of outbreaks vary, and this is likely to be influenced by multiple and complex care home factors. Genomic sequencing data can support inferences about the chain of transmission, as linked cases with genetically similar isolates suggest that transmission has occurred from one source.[192–194]

However as sequencing is time-consuming and not always available, prompt action to contain the outbreak is required without access to these data.[195]

Public health disease control measures deployed once an outbreak has been identified are described in Chapter 1 in more detail. These are mainly designed to prevent further infection ingress such as pre-admission and visitor screening, isolation of new admissions, care staff living on-site, and closure to visitors,[156,184,196] and to limit spread, such as cohorting of infected residents, PPE, and additional testing.[156,197] As such, it may be valuable to consider the factors associated with ingress and spread of infection separately to inform more specific and targeted use of disease control measures against SARS-CoV-2, with the added potential of extending findings to other respiratory infections.

To inform the research studies outlined in my thesis, my first objective was to conduct a scoping review to identify studies that have identified facility-level factors associated with introduction and spread of SARS-CoV-2 within care homes.

2.2 Methods

I performed the initial literature search in July 2020, and it formed the basis of my MSc dissertation which was submitted in September 2020. To disseminate research findings as quickly as possible, most research was pre-printed in the early pandemic therefore most of the studies that I included in my review had not been peer-reviewed. To improve the quality of my review, I repeated the search two years later in June 2022, using similar search terms however only included peer-reviewed studies. Findings from this second, more comprehensive review are presented in this chapter.

To focus the review and inform my planned research, I decided to investigate factors associated with infection ingress (occurrence of a case or outbreak) and those describing infection transmission (outbreak size, secondary attack rate, incidence rate, reproductive number). This is because a single case or an outbreak (where a small number of cases have been identified in the facility), probably suggests infection has entered the facility, whereas the number of cases probably describes the extent of transmission within the facility. Although my thesis aims focusses on risk factors for

infections amongst residents, I included studies that reported infections in residents and/or staff in this review because infections in staff play a key role in infection ingress and transmission within care homes.

The methods for this scoping review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) checklist.[198] I conducted the literature search in Ovid (MEDLINE) on 23rd June 2022, Table 2.1. The COVID-19 search term was taken from the search strategy recommended by the National Institute for Clinical Excellence (NICE) in March 2020.[199] I reviewed my search terms with a medical librarian at UCL. Date was restricted to 2019 onwards as this was when SARS-CoV-2 was first identified, and language was restricted to English. I included studies if they reported facility-level risk factors for SARS-CoV-2 infections and / or outbreaks and if they had been peer-reviewed. I excluded studies if they did not report original data or did not include any care homes for older people, although studies that presented a mixture of care homes serving both older and younger populations were included.

I assessed study quality using the Newcastle-Ottawa Scale case-control or cohort tool[200] in line with study design and the National institute of Health (NIH) tool for cross-sectional studies.[201] These tools can be used to assess sample selection, comparability, and exposure selection and evaluation in studies.

I extracted data on study location, design, dates, population, inclusion and exclusion criteria, case definitions, exposures, statistical methods, and results. Meta-analysis was not attempted in this scoping review due to heterogeneity in study populations, case definitions, and outcome measures. In addition, most studies only presented summary measures and did not report raw data.

Table 2-1: Ovid (MEDLINE) Search strategy

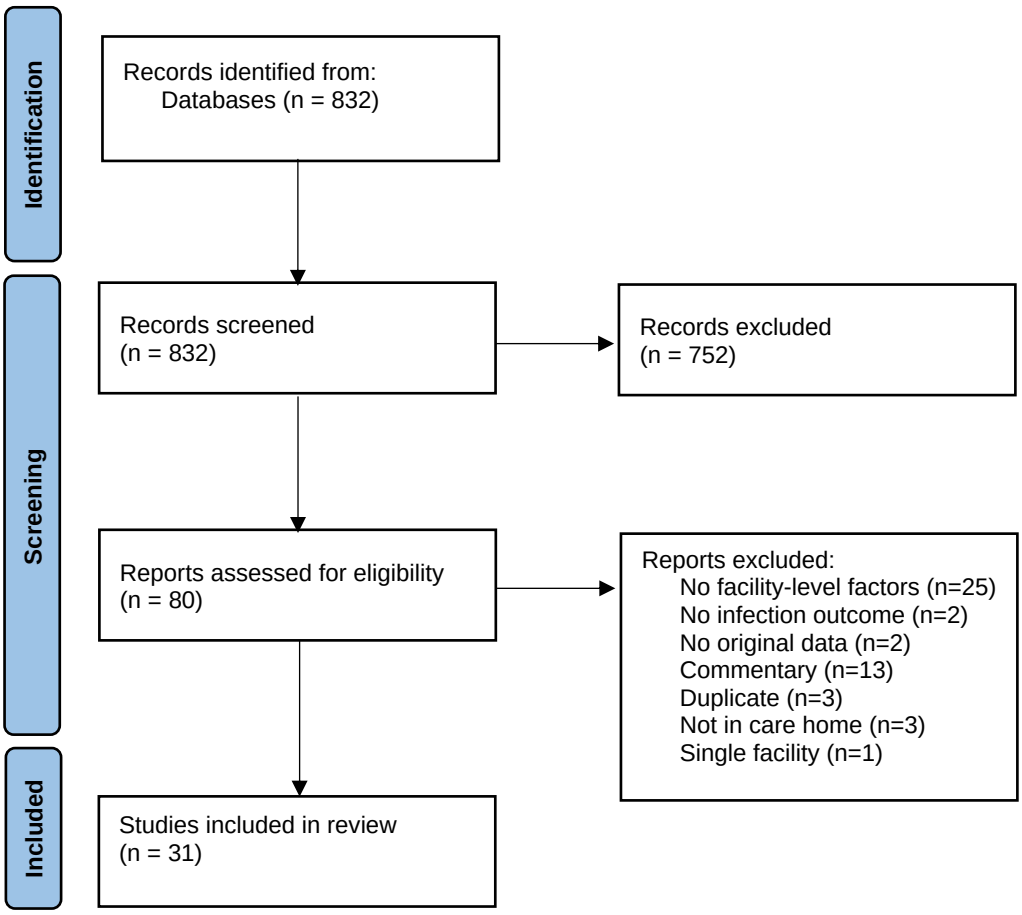
(Also published in MSc dissertation)

Step	Search term	No. of records
1	exp coronavirus/	160636
2	((corona* or corono*) adj1 (virus* or viral* or virinae*)).ti,ab,kw.	5354
3	(coronavirus* or coronovirus* or coronavirinae* or Coronavirus* or Coronovirus* or Wuhan* or Hubei* or Huanan or "2019-nCoV" or 2019nCoV or nCoV2019 or "nCoV-2019" or "COVID-19" or COVID19 or "CORVID-19" or CORVID19 or "WN-CoV" or WNCoV or "HCoV-19" or HCoV19 or CoV or "2019 novel*" or Ncov or "n-cov" or "SARS-CoV-2" or "SARSCoV-2" or "SARSCoV2" or "SARS-CoV2" or SARSCov19 or "SARS-Cov19" or "SARSCov-19" or "SARS-Cov-19" or Ncovor or Ncorona* or Ncorono* or NcovWuhan* or NcovHubei* or NcovChina* or NcovChinese*).ti,ab,kw.	348037
4	(((respiratory* adj2 (symptom* or disease* or illness* or condition*)) or "seafood market*" or "food market*") adj10 (Wuhan* or Hubei* or China* or Chinese* or Huanan*)).ti,ab,kw.	954
5	((outbreak* or wildlife* or pandemic* or epidemic*) adj1 (China* or Chinese* or Huanan*)).ti,ab,kw.	489
6	"severe acute respiratory syndrome*".ti,ab,kw.	40196
7	or/1-6	359908
8	limit 7 to yr="2019 -Current"	339087
9	(care home* or residential home* or nursing home* or long term care facilit*).ti,ab,kw.	46529
10	(risk factor* or risk* or factor*).ti,ab,kw.	5745601
11	and/8-10	832

2.3 Results

The study selection process is illustrated in Figure 2.1. I screened 832 citations, of which 80 were selected for further review based on the title and abstract alone. Following full-text review, 31 were eligible for inclusion.

Figure 2-1: PRISMA diagram of study selection process [198]



All included studies were peer-reviewed and are summarised in Table 2.2. Where stated, the number of facilities ranged from 6[202] to 13079 [203] and the number of resident participants from 409[202] to 160,000.[154] 11/31 studies were conducted in countries in Europe (Spain, Belgium, Denmark, Sweden, Finland, Norway, Ireland, France, Italy),[204–214] 6/31 in UK,[154,170,215–218] 9/31 in USA,[203,219–226] 3/31 in Canada,[227–229] one in Iran,[202] and one in Korea.[230] Studies' designs were either cross-sectional (9/31),[203,206,211,212,215,219,220,225,230] cohort (12/31),[170,209,213,216,217,221,223,224,226–229] case-control (2/31)[202,210] or cross-sectional with cohort (8/31).[154,204,205,207,208,214,218,222] 8/20 cohort studies used postcode matching to identify residents.[170,203,209,216,218,223,226,230] Most studies were conducted within the first year of the pandemic (January 2020 – January 2021) before SARS-CoV-2 vaccination had been rolled out and before widespread circulation of novel variants such as Alpha and Delta. There were two studies that continued beyond the first month of 2021 - Lane et al ended data collection

in February 2021[226] and Aghili et al continued until July 2021.[202] However as the latter was based in Iran, widespread vaccination had not yet occurred.[231,232]

Table 2-2: Summary of included studies

Red shading of author box indicates studies of low quality, orange indicates medium / fair quality, blue indicates good quality.

Author, year	No. participants / CHs	Location	Study dates (Study design)	Case / Outbreak definition	Exposures	Adjusted analysis results
Aghili 2022[202]	409 residents 6 CHs	Tehran, Iran	25 March - 12 July 2021 (CC)	PCR confirmed, possible case if clinical symptoms (cough + fever)	Demographics, IPC measures (social distancing, face mask use, hand washing, education), predisposition to COVID-19 (influenza vaccine, Vitamin D intake), environmental and staff characteristics (beds per room, air conditioning, windows in bedrooms, meal area, shift duration, ratio of nurses / healthcare workers to residents, glass shield in visitors' area), temperature checks amongst residents	*Risk of case:* - **Higher** if masks not used outside the room (aOR 3.37, 95% CI 1.74–6.53, p<0.001), longer staff shifts (3.02, 1.68–5.43, p<0.001), using cloth mask / not wearing a mask (2.47, 1.13–5.42, p=0.024), absence of glass barriers in visitors space (1.95, 1.11–3.50, p=0.025)
Bach-Mortensen 2021[218]	all care homes in 149 upper tier LAs	England	10 April - 19 June 2020 (CS, Ch)	PHE outbreak reports, CQC-reported deaths Outbreak ≥ 2 cases within 14d	Area deprivation (IDAOPI, IMD index), local area demographics (number of people > 65 years, population density, percentage black and minority ethnic population)	No associations found
Brown 2021[229]	>78000 residents 618 CHs	Ontario, Canada	29 March - 20 May 2020 (Ch)	PCR confirmed - symptomatic	Crowding index, ownership type, facility size, ratio staff-to-beds, proportion of 1-bed / 2-bed / 4-bed rooms, design standard, demographics (sex, age, co-morbidities, functional level, education level), local SARS-CoV-2 incidence rate, proportion of local population born outside of Canada	*Incidence rate of infection*: * - **Higher** if higher crowding index (aRR 1.73, 95% CI 1.10-2.72)
Bui 2020[221]	123 CHs	West Virginia, USA	14 March - 11 June 2020 (Ch)	PCR confirmed - symptomatic Outbreak ≥ 2 laboratory-confirmed cases within 14d, ≥ 1 case in a resident	Quality rating (CMS), ownership type, number of residents, staffing hours per resident, cumulative local SARS-coV-2 incidence, number of fines / penalties on inspection	*Odds of outbreak:* * - **Lower** in facilities with higher quality rating (2–3-star vs 1-star (aOR 0.13, 95% CI 0.03-0.54), 4–5-star vs 1-star (0.06, 0.003-0.39))

Author, year	No. participants / CHs	Location	Study dates (Study design)	Case / Outbreak definition	Exposures	Adjusted analysis results
Burgana 2021[213]	842 residents 12 CHs	Sant Cugat del Valles, Spain	15 March - 15 May 2020 (Ch)	PCR confirmed - symptomatic	Facility size, number of symptomatic staff tested, demographics	*Risk of case:* - **Higher** if higher infection rate amongst staff (aOR 1.07, 95% CI 1.03-NS, p<0.001)
Burton 2021[216]	817 CHs	Scotland	1 March - 31 May 2020 (Ch)	PCR confirmed - symptomatic, Outbreak ≥ 1 case in resident	Number of beds, ownership type, duration of care home service, type of care, Risk Assessment Document (RAD) score which determines frequency of inspections, local population density, distance from urban centre, local SARS-CoV-2 incidence rate	*Risk of outbreak: ** - **Higher** if larger facility (>=90 beds vs <20 beds) (aOR 55.4, 95% CI 15.0-251.7), higher local SARS-CoV-2 incidence rate (1.2 per 100 cases/100,000 population ↑, 1.0-1.4), Local Authority or NHS funded vs private (2.0, 1.1-3.7), -**Lower** if rural location vs urban (0.1, 0.03-0.3), longer duration of service – 11-14 years vs 0-2 years (0.4, 0.2-0.9)
Cazzoletti 2021[208]	5145 beds 45 CHs	Trento, Italy	1 March - 1 June 2020 (CS, Ch)	PCR confirmed - symptomatic	Facility size, structure, presence of special care units, urban / rural location, geographic location, ownership type, number of staff, compliance with quality standards, IPC measures	*Cumulative incidence:* - Associated with geographical region only (effect measures not stated)
Corvol 2022[207]	20881 residents 231 CHs	Brittany, France	July 2020 (CS) 1 March 2020 - 31 May 2020 (Ch)	PCR confirmed or chest scan	Facility location (urban vs rural), number of residents, dependence level of residents, proportion double rooms, presence of physician or hygienist, ratio nurses / healthcare assistants / salaried personnel to each resident, lockdown measures, daily access to outside area, in-room meal service, use of PPE (re-use of materials, use of unlicensed materials, systematic mask wearing), delayed closure	*Risk of case:* - **Lower** with in-room meal service (vs meals in communal areas) (aOR 0.10, 95% CI 0.02-0.35, p<0.001), daily access to outdoor space (0.20, 0.04-0.90, p=0.04) -**Higher** if visitors banned when recommended (11th March 2020) compared with earlier (5.27, 1.29-27.63, p=0.03)
Dutey-Magni 2021[217]	9339 residents, 11604 staff 179 CHs	UK	2 March - 14 June 2020 (Ch)	PCR confirmed, Suspected - symptoms only	Demographics (sex, age, type of care), type of care provided, number of beds, occupancy, bed-to-staff ratio, IMD index	*Incidence rate of infection:* - **Higher** if lower staffing (aHR 10.1 per 1↑ in bed:staff, 95% CI 1.64-62.1), higher bed occupancy (60.5 per 1↑ in resident:bedroom, 2-55-1436)
Emmerson 2021[170]	25661 residents 1068 CHs	Wales	22 February - 27 June 2020 (Ch)	National registry Outbreak ≥ 1 case in resident	Recent hospital discharge into facility, number of beds, services available, region	*Risk of outbreak:* - **Higher** if greater number of residents (10-24 vs <10 (aHR 1.99, 95% CI 1.99-5.80), 25-49 bs

Author, year	No. participants / CHs	Location	Study dates (Study design)	Case / Outbreak definition	Exposures	Adjusted analysis results
						<10 (8.25, 4.93-13.81), 50+ vs <10 (17.35, 9.65-31.19))
He 2020[220]	1223 CHs	California, USA	1 January - 2 June 2020 (CS)	National registry	Quality rating, ownership type, bed occupancy, proportion white residents, facility age	*Risk of case:* - **Higher** if greater proportion of residents from white ethnic background - <59.5% vs ≥59.5% (aOR 1.95, 1.49-2.55, p<0.01) -**Lower** if higher quality rating – 5-star vs 3-star (0.41, 0.27-0.62, p<0.01)
Hege 2022[223]	9900 CHs	USA	1 June 2020 - 31 January 2021 (Ch)	Self-reported	Facility-level: quality rating, staffing rating, overall rating, number of fines incurred, total sum of fines, ownership type, recent change in ownership, facility size, number of weeks of nursing / clinical staff / nurse aide / other staff shortages County level: population density, median annual income, age and ethnic composition, SARS-CoV-2 infection rates	*Infection rate:* - **Higher** if nursing (coeff 0.0005, <0.001) / staff (0.002, p<0.001) shortages, greater local SARS-CoV-2 infection incidence (varies over time periods) - **Lower** if locally owned (-0.007, p=0.01)/ state-owned (-0.03, p<0.001) / not-for-profit (-0.011, p<0.001) vs for-profit ownership, higher median annual personal income at county level (varies over time periods)
Lane 2022[226]	2951 CHs	Kentucky, Virginia, Tennessee, North Carolina, South Carolina, Georgia, Mississippi, Alabama, and Florida	Three separate time periods (Ch): (1) Early Pandemic: 24 May - 26 September 2020 (2) Mid-Pandemic: 27 September - 26 December 2020 (3) Late Pandemic: 27 December 2020 - 6 February 2021	Suspected case or PCR confirmed	Resident demographics, quality rating, staff rating, level of care needs, number of beds, ownership type, number of fines incurred by facility, county level demographics (age, ethnicity, poverty), county-level SAS-CoV-2 incidence rate	*Number of cases:* <u>Early pandemic period -</u> - **Higher** if greater county level SARS-CoV-2 incidence rate (aIRR 1.55 95% CI 1.43–1.68, p<0.001), greater number of beds (1.75, 1.64-1.88, p<0.001), greater proportion of county-population are Black (1.19, 1.10–1.29, p<0.001) or Asian (1.25, 1.12–1.38, p<0.001), higher proportion of residents who are Medicaid-funded (1.14, 1.06–1.23, p=0.001), greater nursing aide numbers (1.31, 1.13-1.52, p<0.001), - **Lower** if greater staffing numbers (0.72, 0.61-0.84, p<0.001) <u>Mid-pandemic period -</u> - **Higher** if greater county level SARS-CoV-2 incidence rate (1.55, 1.43–1.68, p<0.001), greater number of beds (1.58, 1.48-1.70, p<0.001), higher proportion of residents who are Medicaid-funded (1.09, 1.01–1.17, p=0.024),

Author, year	No. participants / CHs	Location	Study dates (Study design)	Case / Outbreak definition	Exposures	Adjusted analysis results
						greater proportion of local population below poverty line (1.15, 1.05–1.26, p=0.002) - **Lower** if greater staffing numbers (0.87, 0.80-0.93, p<0.001), counties with higher proportion Black (0.77, 0.70–0.85, p<0.001) and Hispanic population (0.74, 0.68–0.81, p<0.001) <u>Late-pandemic period -</u> - **Higher** if greater county level SARS-CoV-2 incidence rate (1.26, 1.15–1.38, p<0.001), greater number of beds (1.61, 1.48-1.74, p<0.001), higher number of female residents (1.16, 1.07–1.25, p<0.001)
Lee 2022[230]	118315 residents 3396 CHs	Korea	20 January - 20 October 2020 (CS)	PCR confirmed - symptomatic	Facility size, proportion registered nurses, ratio care workers / physicians / physical therapists to residents, quality rating	*Infection rate:* - **Lower** if greater proportion of staff are registered nurses (coeff -0.63, SE 0.31, p=0.049) - Higher if higher ratio care workers to residents (9.27, 4.30, p=0.033)
Lombardo 2021[214]	100806 residents 1356 CHs	Italy	25 March - 5 May 2020 (CS, Ch)	PCR confirmed - symptomatic Outbreak - any cases among staff or residents resulting in hospitalisation / death, ≥ 1 case in a resident	Number of beds, ownership type, occupancy, number of healthcare and social workers, IPC measures	*Risk of outbreak:* - **Higher** if staff shortages (aOR 3.22, 95% CI 2.38-4.36, p<0.001), difficulties with transferring infected residents out (4.66, 2.98-7.31 *p*<0.001), difficulties with isolating infected residents (1.97, 1.42-2.73 *p*<0.001), greater median number of beds (>60 vs <=60) (1.50, 1.09-2.07 p=0.013), geographic region (varies by regions of Italy) - **Lower** if lack of PPE reported at start of study (0.45, 0.29-0.68 *p*<0.001)
Longo 2022[224]	1719 CHs	Illinois, Florida, Massachusetts	1 June 2020 - 17 January 2021 (Ch)	Self-reported	Joint commission accreditation status, state and country level SARS-CoV-2 incidence rate	*Infection rate:* - **Higher** if higher local SARS-CoV-2 incidence (coeff 0.00, SE 0.00, p<0.001)

Author, year	No. participants / CHs	Location	Study dates (Study design)	Case / Outbreak definition	Exposures	Adjusted analysis results
Orlando 2022[210]	100 CHs	Lazio region, Italy	March to December 2020 (CC)	PCR confirmed - symptomatic Outbreak ≥ 2 cases within 14d Case = COVID-19 outbreak in CH Control = 1 or 0 COVID-19 infections in CH	Facility size, number of shared rooms, ownership type, IPC measures (isolation of residents, area for staff changing / donning / doffing, separate entrances for staff in contact with residents, active surveillance for infection of staff and residents i.e., temperature monitoring, asymptomatic screening), local SARS-CoV-2 incidence rate, cases in staff, number of days open to visitors between June and Sept 2020, urban / rural location	*Odds of outbreak:* - **Higher** if greater number of beds (larger facility, >15 beds) compared with smaller (<15 beds) (aOR 5.37, 95% CI 1.58-22.8, p=0.012)
Peckeu-Abboud 2022[206]	66209 residents 62989 staff 695 CHs	Belgium, Flanders region	5 April - 15 May 2020 (CS)	PCR confirmed - asymptomatic	Number of beds, proportion of nursing beds, staff-to-resident ratio, median age of residents and staff, test positivity rate amongst staff, proportion asymptomatic cases amongst staff and residents, home ownership	*Proportion of residents testing positive:* - **Higher** if higher proportion of staff testing positive (IRR 1.89 per % ↑, 95% CI 1.68–2.12, p<0.001), higher proportion of nursing beds ((IRR 1.97 per % ↑, 95% CI 1.00–3.86, p=0.05)
Piet 2021[211]	5189 residents, 4652 staff 74 CHs	French Alps	15 July - 15 November 2020 (CS) covering period 1 March - 31 May 2020	Confirmed - PCR/serology, Probable - clinical signs + thoracic CT Outbreak >3 confirmed/ probable cases among residents over 8-week period	Facility size, ownership type, number of residents, number of dementia beds, average dependency and morbidity scores of residents, human resources and operational management during pandemic, number of staff, agency staff, IPC measures including date of closure to visitors, use of face masks, isolation of residents,	No factors were associated with outcomes in bivariate analysis
Rauhala 2022[205]	1962 CHs	Denmark, Finland, Norway, Sweden	March - June 2020 (Ch) May - June 2020 (CS)	National registry	325 variables from survey relating to preventive measures and resources in facility, local SARS-CoV-2 incidence rate	*Risk of case:* - **Higher** with higher local SARS-CoV-2 incidence (aOR 1.06, 95% CI 1.05–1.08, p≤0.001), greater number of employees manager is responsible for (in 10s) (1.02, 1.01–1.04, p≤0.001), absence of preventive client testing (1.56, 1.07-2.26, p≤0.05) - **Lower** in homes providing residential care vs home care (0.41, 0.29-0.59, p≤0.001)

Author, year	No. participants / CHs	Location	Study dates (Study design)	Case / Outbreak definition	Exposures	Adjusted analysis results
San Roman 2022[204]	23,756 residents 20,795 staff 369 CHs	Spain, Madrid	July to October 2020 (CS) July to December 2020 (Ch)	PCR confirmed - symptomatic Outbreak ≥ 1 case	Facility size, seropositivity (nucleocapsid antibody), local SARS-CoV-2 incidence rate	*Outbreak risk:* - **Lower** if high (aHR 0.22, 95% CI 0.10-0.48, p<0.001) or intermediate (0.45, 0.25-0.80, p=0.007) seroprevalence (vs low) - **Higher** in intermediate (1.91, 1.00-3.65, p=0.05) and larger (4.57, 2.38-8.75, p<0.001) compared with smaller facilities
Shallcross 2021[154]	160 033 residents 248 594 staff 5126 CHs	England	30 April - 13 June 2020 (Ch) 26 May - 19 June 2020 (CS)	Self-reported cases / asymptomatic PCR testing. Outbreak ≥ 1 case in a resident or staff member; Large outbreak (largest of) > 1/3 of all residents + staff positive or >20 residents + staff positive	IMD index, ownership type, chain status, staff-to-bed ratio, CQC rating of leadership quality, presence of barrier nursing, any difficulties in isolating residents, cohorting staff, cleaning frequency, sick pay for staff, use of agency staff, use of PPE, cross-site working	*Odds of case in resident:* - **Higher** if no cohorting of staff with infected / uninfected residents (aOR 1·30, 95% CI 1·23–1·37, p<0·001), greater number of new admissions to facility relative to the baseline (1·01 per unit ↑, 1·01–1·01, p<0·001), for-profit ownership vs not-for-profit (1·19, 1·12–1·26, p<0·001), frequent employment of agency staff (vs none) (1·65, 1·56–1·74, p<0·001), difficulty isolating residents (vs no difficulty) (1·33, 1·28–1·38, p<0·001) - **Lower** if paid statutory sick pay to staff (vs those that did not) (0·80, 0·75–0·86, p<0·001) higher staff-to-bed ratio (0·82 per unit ↑, 0·78–0·87, p<0·001) *Odds of case in staff:* - **Higher** no cohorting of staff with infected / uninfected residents (1·20, 1·13–1·29, p<0·001), greater number of new admissions to facility relative to the baseline (1·00 per unit ↑, 1·00–1·01, p=0·0005), for-profit ownership vs not-for-profit (1·19, 1·10–1·29, p<0·001), frequent employment of agency staff (vs none) (1·85, 1·72–1·98, p<0·001), difficulty isolating residents (vs no difficulty) (1·48, 1·41–1·56, p<0·001) - **Lower** if paid statutory sick pay to staff (vs those that did not) (0·70, 0·65–0·77, p<0·001), higher staff-to-bed ratio (0·63 per unit ↑, 0·59–0·68, p<0·001) *Odds of outbreak:* - **Higher** no cohorting of staff with infected /

Author, year	No. participants / CHs	Location	Study dates (Study design)	Case / Outbreak definition	Exposures	Adjusted analysis results
						uninfected residents (2·56, 1·94–3·49, p<0·001, greater number of new admissions to facility relative to the baseline (1·08 per unit ↑, 1·05–1·10, p<0·001), frequent employment of agency staff (vs none) (2·33, 1·72–3·16, p<0·001), difficulty isolating residents (vs no difficulty) (1·84, 1·48–2·30, p<0·001) *Odds of large outbreak:* - **Higher** if for-profit ownership vs not-for-profit (1·65, 1·07–2·54, p=0·024), frequent employment of agency staff (vs none) (2·42, 1·67–3·51, p<0·001), difficulty isolating residents (vs no difficulty) (1·62, 1·24–2·11, p=0·0004) - **Lower** if paid statutory sick pay to staff (vs those that did not) (0·59, 0·38–0·93, p=0·024)
Soldevila 2022[212]	8021 residents 168 CHs	Catalonia, Spain	1 March - 30 June 2020 (CS)	PCR confirmed - symptomatic / contact. Asymptomatic testing from mid-April	Demographics (age, sex, underlying co-morbidities, functional level), facility size, local SARS-CoV-2 incidence rate	*Risk of case*: - **Higher** if larger facility (number of beds) (aOR 1.73, 95% CI 1.6-1.9, p<0.001), higher local SARS-CoV-2 incidence (1.77, 1.0-3.0, p=0.04), low level of functional dependence (vs higher level of dependence) (1.22, 95% CI NS p=0.03)
Stall 2022[228]	75676 residents 623 CHs	Ontario, Canada	29 March - 20 May 2020 (Ch)	National registry Outbreak ≥ 1 case in resident	Number of beds, occupancy, number of shared rooms and number of residents per room, staff-to-bed ratio, chain ownership and size of chain, age of facility design	*Odds of outbreak:** - **Higher** if greater number of residents (aOR 1.38 per ↑50 residents, 95% CI 1.18-1.61), higher local SARS-CoV-2 incidence rate (1.91, 1.19-3.05), older facility design (1.55, 1.01-2.38) *Larger outbreak:* - **Higher** if for-profit vs not-for-profit ownership (aRR 1.96, 95% CI 1.26-3.05) *(mediated by higher number of facilities with older design and belonging to chain amongst for-profit homes*)
Sugg 2021[203]	13079 CHs	USA	1 January - 30 June 2020 (CS, spatial modelling)	national registry	Quality rating, staffing rating, total staff, number of fines, ownership type. County level: ethnicity, average household size, employment rate, average income, average rent,	*Risk of case:* - **Higher** if greater number of fines (aIRR 1.13, 95% CI 1.13-1.13, p<0.001), higher per capita county-level income (2.48, 2.48-2.49, p<0.001), greater average household size in county (1.26, 1.26-1.26, p<0.001), greater proportion African

Author, year	No. participants / CHs	Location	Study dates (Study design)	Case / Outbreak definition	Exposures	Adjusted analysis results
					population density, local SARS-CoV-2 incidence rate	American in county population (1.27, 1.27-1.27, *p*<0.001), greater population density (1.12, 1.12-1.12, *p*<0.001). higher COVID-19 county rates (1.86, 1.86-1.86, *p*<0.001), higher number registered nurses (1.16, 1.16-1.16, *p*<0.001) -**Lower** if higher total staffing levels (0.78, 0.78-0.78, *p*<0.001)
Torres 2022[209]	232 CHs	Barcelona	1 March - 22 June 2020 (Ch, ecological)	Confirmed or suspected (method not stated)	Local area socio-economic status (based on Available Family Income Index), capacity to isolate residents or to cohort, occupancy, crowding (ratio residents-to-room), ownership type	*Risk of case:* - **Higher** if socio-economic status of area low (aRR 1.44, 95% CI 1.34–1.55, p<0.001) or medium (vs high) (1.28, 1.21–1.34, p<0.001), complete vs partial occupancy (1.07, 1.02-1.12, p<0.001), crowding: medium vs low (1.43, 1.35-1.51, p<0.001), high vs low (1.36, 1.28-1.45, p<0.001), public vs private for-profit ownership (1.15, 1.06-1.24, p<0.001)
Travers 2021[222]	11587 CHs	USA	20 January - 19 July 2020 (CS, Ch)	National registry	Proportion of black residents in facility, number of beds, occupancy, chain ownership, ownership type, proportion Medicaid residents, nurse / staffing ratios, staffing shortages, PPE shortages	*Incidence rate of infection:* - **Higher** if greater proportion black residents - 20-49.9% vs none (coeff 0.028, 95% CI 0.006-0.05, p=0.014) *(NB. Attenuated once county-level fixed effects added to model)*
Tulloch 2021[215]	77 CHs	Liverpool, UK	27 April - 3 May 2020 (CS)	PCR confirmed - asymptomatic, suspected - symptoms only Outbreak ≥ 2 possible/confirmed cases within 14d (staff or resident)	Quality rating (CQC)	No associations found in univariable analysis therefore multivariable analysis not performed
Vijh 2022[227]	74 CHs in descriptive analysis, 48 CHs in regression	British Columbia, Canada	1 March 2020 - 10 January 2021 (Ch)	PCR confirmed - symptomatic / contact. Asymptomatic testing from mid-April	Local SARS-CoV-2 incidence rate, year of build, proportion single rooms, score from outbreak prevention assessment tool	*COVID-19 attack rate:* - **Higher** if care home opened before 1972 (vs after) (aRR 5.89, 95% CI 2.33-14.85, p<0.001), if at least one item in outbreak prevention assessment tool not met (strongest for dining room) (6.37, 2.70-15.04, p<0.001) - **Lower** if index case was a staff member (0.34, 0.12-0.94, p<0.05)

Author, year	No. participants / CHs	Location	Study dates (Study design)	Case / Outbreak definition	Exposures	Adjusted analysis results
White 2020[219]	341 Genesis, CHs 3016 non-genesis CHs. 64 CHs had universal testing	25 states in USA	mid-March - 21 April 2020 (non-genesis) or mid-March - 4 May 2020 (genesis) (CS)	Genesis homes - PCR confirmed (symptomatic & asymptomatic); non-Genesis - National registry Outbreak ≥ 1 case in resident	Demographics (age, ethnicity, dementia), quality rating (NHC), any IPC deficiency citations in prior year, local population density and ethnicity distribution, local SARS-CoV-2 incidence rate	Risk of outbreak: - Higher in facilities with greater number of beds (+0.9 pp per 10 beds ↑, 95% CI 0.6-1.2 pp, p<0.001), local SARS-CoV-2 infection incidence (+33.6 pp per 1000 cases/100,000 population, 9.6-57.7, p=0.008) - Lower in facilities with 4 or 5-star quality rating compared with 3-star (-2.9 pp, -5.1- -0.7, p=0.01)
Zhu 2022[225]	7785 CHs	USA	7 June - 20 Dec 2020 (CS)	National registry	Demographics (age, sex, ethnicity, functional level), number of beds, number of ventilator-dependant beds, shared bedrooms, proportion private beds, floor area, previous infection control inspection results, rural / urban location, county-level socio-economic status, percentages SARS-CoV-2 infection / death in county, dummy variables for state, ownership type, quality ratings, staff nursing hours, staff cases, COVID-19 testing, IPC measures	*Incidence rate of infection:* - **Higher** if greater number of residents previously hospitalised with SARS-CoV-2 (aIRR 1.00 per %↑, 95% CI 1.00-1.00, p<0.001), greater number of staff cases (1.03, 1.02-1.03, p<0.001), - **Lower** if greater number of beds (0.95, 0.95-0.96, p<0.001), higher proportion of non-Hispanic Asians in local population (0.99 per %↓, 0.99-1.00, p<0.01), higher proportion of Hispanics in local population (0.99 per %↓, 0.99-1.00, p<0.001), higher quality ratings (0.98, 0.97-1.00, p<0.05), for-profit vs non-for-profit ownership (0.86, 0.82-0.91, p<0.001) *COVID-19 transmissibility:* - **Higher** if greater number of residents previously hospitalised with SARS-CoV-2 (β 0.00005, 95% CI 0.00001-0.00008, p<0.001), greater number of staff cases (0.02, 0.02-0.02, p<0.001), - **Lower** if greater number of beds (-0.03, -0.03- -0.03, p<0.001), higher proportion of non-Hispanic Asians in local population (-0.004, -0.008- -0.00006, p<0.05), higher proportion of Hispanics in local population (-0.005, -0.006- -0.003, p<0.001), higher quality ratings (-0.02, -0.03- -0.009, p<0.001), not-for-profit vs for-profit ownership (-0.08, 0.12 - -0.05, p<0.001)

CS Cross-sectional CC Case-control Ch Cohort CH care home CQC Care Quality Commission NS Not Stated
CMS Centers for Medicare & Medicaid services PHE Public Health England NHC Nursing Home Compare IPC infection Prevention & Control

IMD Index of Multiple Deprivation IDAOPI Income Deprivation Affecting Older People Index PPE personal & protective equipment CT Computed Tomography
aOR adjusted Odds Ratio aRR adjusted Risk Ratio aHR adjusted Hazards ratio IRR Incidence Rate Ratio pp percentage point
SE Standard Error CI confidence intervals

**P*-values not stated
Effect sizes, confidence intervals, and *P*-values presented to significant figures originally reported.

Table 2-3: Quality assessment of included studies: a) Newcastle Ottawa Scale [200] assessment of cohort & case-control studies, b) National Institute of Health score for cross-sectional studies [201]

In a) good quality was assessed for studies with 3 or 4 points in selection domain AND 1 or 2 points in comparability domain AND 2 or 3 stars in exposure domain. Fair rating was assessed for studies with 2 points in selection domain AND 1 or points in comparability domain AND 2 or 3 points in exposure domain (according to rating recommendations published in Newcastle-Ottawa Scale). No studies fulfilled criteria for poor quality rating (0 or 1 point in selection domain OR 0 points in comparability domain OR 0 or 1 points in exposure domain). For b) quality rating was based on total points (out of 14): 0-4 = poor, 5-10 = fair, 11-14 = good.

a)

		Selection					Comparability		Exposure					
Author, year	Study type	Case def	Non-exposed	Controls def	Control selection	Total - selection	Comparability of cohorts	Total – comparability	Ascertainment	Follow-up	Non-response	Total - exposure	Total – overall	Rating
Aghili 2022	Case-control	c	b	a	a	2	a, b	2	d	a	a	2	6	Fair
Brown 2021	Cohort	a	a	a	a	4	a, b	2	b	a	a	2	9	Good
Bui 2020	Cohort	a	a	a	a	4	a	1	b	a	a	2	8	Good
Burgana 2021	Cohort	b	a	a	a	3	a, b	2	a	a	a	3	8	Good
Burton 2021	Cohort	a	a	a	a	4	a, b	2	b	a	a	2	9	Good
Dutey-Magni 2021	Cohort	a	a	a	a	4	a, b	2	b	a	a	2	9	Good
Emmerson 2021	Cohort	a	a	a	a	4	a	1	b	a	a	2	8	Good
Hege 2022	Cohort	a	a	d	b	2	a, b	2	b	a	a	2	7	Fair

		Selection					Comparability		Exposure					
Author, year	Study type	Case def	Non-exposed	Controls def	Control selection	Total - selection	Comparability of cohorts	Total – comparability	Ascertainment	Follow-up	Non-response	Total - exposure	Total – overall	Rating
Lane 2022	Cohort	a	a	a	b	3	a, b	2	b	a	a	2	8	Good
Longo 2022	Cohort	d	a	a	a	3	a	1	b	a	a	2	7	Good
Orlando 2022	Case-control	a	a	a	a	4	a, b	2	a	a	a	3	9	Good
Stall 2022	Cohort	a	a	a	a	4	a	1	b	a	a	2	8	Good
Torres 2022	Cohort	a	a	d	a	3	a, b	2	b	a	a	2	8	Good
Travers 2021	Cohort	a	a	a	a	4	a, b	2	b	a	a	2	9	Good
Vijh 2022	Cohort	c	a	a	a	3	a, b	2	b	a	a	2	8	Good

b)

Author, year	Objective	Population	Participation rate	Recruitment	Sample size	Exposure	Follow-up	Exposure levels	Exposure definitions	Exposure assessment	Outcome	Blinding	Loss to follow-up	Confounding variables	Total score	Rating
Bach-Mortensen 2021	Yes	No	Yes	Yes	No	Yes	Yes	Yes	Yes	No	No	NS	N/A	No	7	Fair
Cazzoletti 2021	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	yes	No	Yes	NS	N/A	Yes	10	Fair
Corvol 2022	Yes	Yes	No	Yes	NS	Yes	yes	Yes	Yes	No	Yes	NS	N/A	Yes	9	Fair
He 2020	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	No	No	No	N/A	No	8	Fair
Lee 2022	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	No	No	NS	N/A	Yes	9	Fair
Lombardo 2021	Yes	Yes	No	Yes	No	No	Yes	Yes	Yes	No	Yes	No	N/A	Yes	8	Fair
Peckeu-Abboud 2022	Yes	Yes	Yes	Yes	No	Yes	No	Yes	Yes	No	Yes	NS	N/A	Yes	9	Fair
Piet 2021	Yes	Yes	No	Yes	No	Yes	Yes	No	Yes	No	Yes	NS	N/A	No	7	Fair

Author, year	Objective	Population	Participation rate	Recruitment	Sample size	Exposure	Follow-up	Exposure levels	Exposure definitions	Exposure assessment	Outcome	Blinding	Loss to follow-up	Confounding variables	Total score	Rating
Rauhala 2022	Yes	Yes	No	NS	NS	Yes	No	Yes	Yes	No	No	NS	No	Yes	6	Fair
San Roman 2022	Yes	Yes	NS	Yes	NS	Yes	Yes	Yes	Yes	No	Yes	No	NS	No	8	Fair
Shallcross 2021	Yes	Yes	Yes	Yes	No	Yes	No	Yes	Yes	No	Yes	NS	N/A	Yes	9	Fair
Soldevila 2022	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	No	No	NS	N/A	Yes	9	Fair
Sugg 2021	Yes	No	Yes	No	No	Yes	Yes	Yes	Yes	No	Yes	No	N/A	Yes	8	Fair
Tulloch 2021	Yes	Yes	Yes	Yes	No	Yes	No	Yes	Yes	No	Yes	NS	N/A	No	8	Fair
White 2020	Yes	Yes	NS	No	No	Yes	No	Yes	No	No	No	No	N/A	Yes	5	Fair
Zhu 2022	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	No	No	NS	N/A	Yes	9	Fair

N/A not applicable NS Not stated

Quality varied and is presented in Table 2.3. Amongst the case-control and cohort studies, 13/15 studies were rated as good,[170,209,210,213,216,217,221,222,224,226–229] 2/15 as fair,[202,223] and none were low quality. Amongst the cross-sectional studies, all 16 were classed as fair quality.[154,203–208,211,212,214,215,218–220,225,230] Fair quality rating was assigned to cross-sectional studies that had lower than 50% response rate (n= 4).[205,207,211,214] Other quality issues were that 11 studies used clinical case definitions, based them on self-reports, or did not report them;[202,205,212,213,218–220,224,225,227,230] and five did not adjust for or consider key important confounders in their analysis.[204,211,215,218,220] Fair rating was assigned to two case-control and cohort studies that did not define their case or exposure.[202,223] No cross-sectional studies reported sample size or power calculations.

Studies that used an outcome describing ingress of infection into the facility included 12/31 that investigated the occurrence of any new SARS-CoV-2 infection, [154,202,203,205,207,209,210,212,213,215,220,226] one of which only included asymptomatic cases,[215] and 12/31 studies,[154,170,204,210,211,214–216,218,219,221,228] the described the occurrence of an outbreak. Studies describing transmission of infection within the facility included 15/31 that considered the total number of cases, or the proportion infected over the study period,[203,206,208,209,211,217,219,222–225,227–230] although five also included infections in staff[154,214,215,217,227] and four considered the size of the outbreak in their analysis,[154,204,227,228] however outbreak definitions varied. One study calculated the basic reproductive number (R_0, see Chapter 1) as a measure of overall transmission.[225]

Although this was not the focus of this scoping review, COVID-19 associated deaths and all-cause mortality amongst residents were the outcome of interest in 10[209,214,216,218–220,222,225,228,230] and four[207,212,217,218] studies respectively.

20/31 studies confirmed SARS-CoV-2 infection using PCR or NAAT test,[154,170,202,204,206–208,210–219,221,227,229] however only six performed asymptomatic testing [154,206,212,215,219,227] and for the majority (26/31), testing was triggered when an individual had symptoms of infection or was a contact of a confirmed case. Due to limited testing early in the pandemic, outbreak definitions were usually at least one case of SARS-CoV-2 in a resident as simultaneous infections in the facility were

probably undiagnosed. In four studies[210,215,218,221] outbreaks definitions were in line with those used by Public Health England (PHE) - at least two cases within 14 days.[233] Most studies obtained testing results from national surveillance systems, however 5/31 relied on care home reports[154,214,217,223,224] which may have been subject to ascertainment bias.

The exposures evaluated fell into the following domains: facility-level; staffing; infection prevention and control; and local / regional, Table 2.2. The most frequently evaluated facility-level factors were number of beds, number of shared bedrooms, type of care (e.g., residential etc), availability of outdoor space, building age, ownership (e.g., for-profit, not-for-profit, private, public), number of facilities in the chain (size of chain), quality rating, hospital discharges into the care home, resident's ethnicity, age, functional dependence. Staffing factors included total number and mix of staff type, shift duration, and use of agency staff. Infection prevention measures related to PPE, staff training, cohorting, and visiting policy. Local / regional factors included community SARS-CoV-2 incidence, urban / rural location, socio-economic deprivation of the area surrounding the care home. One study evaluated the influence of seroprevalence on risk of outbreaks.[204] Most studies conducted multivariate statistical analysis and adjusted for key confounders, although two reported only bivariate analyses.[211,215] None of the studies were able to account for turnover of the population from deaths, new admissions, or staff movement during the study period. Full results are outlined in Table 2.2 and summarised in Figure 2.2.

Studies that described risk factors associated with the introduction of infection into a facility looked at risk of a single case in a resident or of an outbreak occurring. After adjustment, the factor most frequently associated with ingress of infection was the size of the facility as measured by number of beds. Larger facilities were found to have a greater risk of infection in one study[212] and outbreak in five studies.[204,210,214,216,219] The magnitude of this relationship varied and adjusted odds ratios were 1.50 for >60 vs <=60 beds (95% CI 1.09-2.07, *P*=0.013),[214] 55.4 for >=90 beds vs <20 beds (95% CI 15.0–251.7),[216] and 5.37 for >15 vs <15 beds (95% CI 1.58-22.8, *P*=0.012)[210] in the three studies that reported this. Other facility features associated with increased risk of infection or outbreak were greater resident numbers (two studies),[170,228] for-profit vs not-for-profit ownership (two studies),[154,209] Local Authority vs private funding (one

study),[216] and older building design (two studies).[227,228] A higher objective quality rating of the facility[219] and access to in-room meals [207] and outdoor space,[202] and older business (two studies)[216] were all associated with a lower risk of ingress of infection, Table 2.2.

Staffing was associated with the risk of ingress of infection, specifically longer shifts,[202] use of agency staff,[154] staff shortages,[203] and greater number of infections amongst staff.[213] Higher staff-to-bed ratio,[154] greater staffing,[203] and payment of statutory sickness pay[154] were all associated with lower risk of infection. A limited number of resident characteristics were assessed on a facility-level, however higher risk of infection ingress was only associated with a lower overall level of functional dependence[212] and greater proportion of residents from a minority ethnic group.[220] Among the nine studies that considered the association between IPC measures and infection ingress,[154,202,205,207,209–211,214,219] significant factors included difficulty in isolating infected residents,[154,214] absence of glass barriers in visitors areas,[202] incorrect mask use,[202] increase in new admissions,[154] difficulties with transferring infected residents out of the facility,[214] and later closure to visitors.[207] Facilities that cohorted infected residents,[154] as well as those that reported a lack of PPE at the study start[214] appeared to have a lower risk of infection coming in, although reverse causality is possible as findings are from cross-sectional studies.

In terms of local factors, local SARS-CoV-2 incidence was most frequently assessed and was significantly associated with a resident case in three studies[203,205,212] and of an outbreak in three.[216,219,228] Other local factors with a significant relationship to resident infection in two studies were lower socio-economic status of the local population.[203,209]

Shallcross et al considered risk factors for infection in a staff member separately and found these to be absence of staff cohorting with either infected or uninfected residents (aOR 1.20, 95% CI 1.13-1.29, *P*<0.001), high number of new admissions (1·00 per unit ↑, 1·00–1·01, *P*=0·0005), frequent employment of agency staff (1·85, 1·72–1·98, *P*<0·0001), difficulty isolating residents (1·48, 1·41–1·56, *P*<0·0001), for-profit vs not-for-profit ownership (1·19, 1·10–1·29, *P*<0·001). Factors associated with lower odds

of infection were payment of sick pay to staff (0·70, 0·65–0·77, *P*<0·001), and higher staff-to-bed ratios (0·63 per unit ↑, 0·59–0·68, *P*<0·001).[154]

There were 16 studies that considered the factors associated with transmission of SARS-CoV-2 in the care home: 15 considered cumulative number of cases or incidence rate, one calculated the R_0, three evaluated factors associated with larger outbreaks. The most frequently associated factor with increased transmission was for-profit ownership compared with not-for-profit identified in three studies (based in UK, USA, and Canada), however reasons behind this are not clear.[154,225,228] There was one USA-based study accessing national registry data from almost 8000 care homes that found that for-profit vs not-for-profit ownership was associated with lower incidence rate of infection (aIRR 0.86, 96% CI 0.82-0.91, *P*<0.001) which may reflect structural differences between social care systems internationally.[225] Older care homes,[227] those that were more crowded,[229] those with higher bed occupancy[217] and those with greater proportion of nursing beds[206] were also found to have increased transmission of SARS-CoV-2 within the facility. Interestingly, a greater bed number was protective against transmission in one cross-sectional survey in the USA by Zhu et al[225] which considered incidence rate and R_0. However, another USA-based cohort study over the same period by Lane et al consistently found an increased transmission risk with more beds when they split the eight-month follow-up into distinct periods, which may be more representative of the significant epidemiological changes that occurred at this time.[226]

In contrast to outcomes describing infection introduction, resident and staffing factors seemed to influence outcomes for risk of spread. A greater proportion of female residents,[226] Medicaid-funded residents,[226] and Black & minority ethnic residents[222] were associated with a higher transmission risk, as did the use of agency staff[154] and care homes with a greater ratio of care workers to residents,[230] or nursing aides[226] which may reflect that staff with less clinical training may be more likely to spread infection within the facility. Factors appearing to reduce the risk of transmission were greater staffing numbers, which was found in two studies,[217,226] having a higher proportion of nursing staff,[230] and facilities that paid statutory sick pay to staff.[154] Five studies considered the influence of IPC measures on transmission risk, [154,208,222,225,227]

however only difficulty with isolating residents[154] and deficiencies on a bespoke outbreak prevention assessment tool[227] were significant. In common with risk factors from introduction of infection, the local incidence of SARS-CoV-2 was the most frequently associated factor with risk of transmission, found in three studies.[223–225] A greater proportion of people from Black and minority ethnic groups in the local population appeared to reduce the risk of infection spread in two studies conducted in the USA,[225,226] however this is likely to vary by study setting.

As already described, Lane at al conducted a study in eight states across the USA and considered how risk factors for greater cumulative incidence of cases changed over the eight-month follow-up between end of May 2020 and February 2021. This period encompassed introduction of vaccination and the transition of variant dominance from wild type to Alpha. The relationship with the factors that they considered remained fairly stable over the three periods, however the association with the proportion of the local population that were Black or Hispanic changed from positive to negative between the first two periods (May-September 2020 and September-December 2020) which may reflect that people from these communities were more likely to work as frontline staff (i.e., in healthcare) and although PPE access to workers was limited in the early pandemic phases therefore increasing infection risk, PPE supply increased in later stages.[226]

Figure 2-2: Heat map of risk factors for infections in residents, outbreaks, and larger outbreaks identified from included studies.

Red colour indicates factors that are associated with increased risk of outcome, blue boxes indicate those associated with reduced risk of outcome. Number in box indicates the number of studies that found this association. If the direction of the significant association varied between studies, the box is split into two (red and blue).

	Infection in resident	Outbreak	Greater number of cases / Larger outbreak	
Facility (environment)				
More beds	1	5	1	1
Higher bed occupancy	1		1	
More nursing beds			1	
Greater number of residents		2		
In-room meals	1			
Daily access to outdoor space	1			
For-profit vs not-for-profit ownership	2		3	1
Local Authority / NHS funded vs private		1	1	
Higher quality rating	2	2	1	

Greater number of fines	1		
Older design	1	1	1
Longer duration of service		1	
Rural vs urban setting		1	
Higher crowding index	1		1
Resident demographics (host)			
More female residents			1
Greater proportion residents Medicaid funded			1
Greater proportion residents belonging to Black & minority ethnic groups	1		1
Lower level of functional dependence	1		
More residents previously hospitalised with COVID-19			1
Higher seroprevalence		1	
Staffing (environment)			
Higher proportion of staff that are registered nurses	1		1
Greater ratio care workers to residents			1
Greater number nursing aides			1
Longer shifts	1		
Use of agency staff	1	1	1
Staff shortages	1		2
Payment of statutory sick pay	1		1
Higher staff-to-bed ratio	1		
Greater staffing numbers	1		2
More staff cases		1	3
Index case is a staff member			1
IPC measures (environment)			
Later closure to visitors	1		
Incorrect face mask use	1		
Cohorting of infected residents	1	1	
Difficulty isolating infected residents	1	2	1
Absence of glass barriers in visitors' areas	1		
More new admissions to facility relative to baseline	1	1	
Difficulty transferring infected resident out of facility		1	
Lack of PPE		1	
Outbreak prevention assessment tool – any item not met			1
Local / regional (environment)			
Higher local SARS-CoV-2 incidence rate	3	3	3
More deprived areas (socio-economic)	2		2
Geographical region		1	1
Greater proportion Black and Hispanic population in county	1		2
Greater average household size	1		
Greater population density	1		

IPC Infection Prevention & Control PPE Personal Protective Equipment

2.4 Discussion

My review has identified 31 published studies that estimated facility-level risk factors for infection and transmission of SARS-CoV-2 in care homes. There is some evidence

that factors that influence the risk of infection ingress differ from those associated with transmission, however only two studies considered these together.[154,228] Higher local SARS-CoV-2 incidence was most frequently associated with higher risk of both outcomes suggesting that infection usually enters from the community but may be brought into the care environment by several individuals at once. This is supported by the significance of other local factors such as ethnic composition and socio-economic status of the local population, which have all been linked with infection incidence.[234–236] Larger facility size was most frequently associated with ingress of infection, possibly because the likelihood of exposure to infection increases with the number of subjects. Resident and staff demographics appeared to be most associated with risk of infection spread, whereas IPC measures appeared to prevent ingress, although few studies considered the impact of IPC on transmission outcomes. There was, however, substantial heterogeneity in the design and settings of included studies. Further work is required to consider how to expand on this to identify modifiable factors to limit infection in social care settings.

Most studies were conducted early in the pandemic therefore could only access data from routine observational datasets or from cross-sectional self-completed questionnaires, as insufficient time had passed to establish dedicated longitudinal data collection, therefore are subject to substantial bias and confounding. Although the built environment is likely to significantly impact on infection risk, the only features considered were building age, access to outdoor space, in-room meals, and number of bedrooms, as these were data that were available to researchers. As studies considered heterogenous populations and facilities over a period of substantial change, they attempted to account for the confounding from this using multivariable analysis, however there was variation in the factors considered. In addition, case detection was largely based on symptomatic testing as asymptomatic surveillance in care homes had not been introduced. Less than one-fifth accessed results from asymptomatic screening, therefore likely under-reporting infections and causing measurement bias, as it is estimated that between 20% and 40% of infections are asymptomatic.[26,237,238] Although host factors relating to immunity affect infection risk, only one study considered the influence of naturally-acquired immunity, with higher seroprevalence appearing protective against infection ingress.[204] Data collection did not continue beyond vaccination rollout, therefore influence of immunity from

vaccination was not considered, thus limiting the implications for current highly vaccinated care populations. None of the studies were able to access genomic sequencing data from diagnostic clinical samples which may have provided useful additional insights into transmission chains.[194,195,239]

Issues with the quality of evidence collected early in the pandemic are also illustrated in systematic reviews that have been conducted over the pandemic. A rapid review from McMaster University that included 44 studies published before 30th November 2020 classed all the studies as low certainty of evidence according to the Grading of Recommendations, Assessment, Development, and Evaluations (GRADE) criteria, and reported similar findings.[240] Larger facility size, staff movement, for-profit status, and lower staffing levels were all associated with greater risk of infection, outbreaks, and deaths in care homes. However, adjustment for community SARS-CoV-2 prevalence reduced the estimated associations between factors, illustrating the impact of this variable and the complexity of relationships between risk factors. A more recent systematic review that focussed on the association between ownership and outbreaks, transmission, and deaths, from 32 studies found that for-profit status increased the risk of transmission and death but not that of outbreaks. They also stressed the importance of considering confounding from other predictors such as staffing, resident characteristics, and access to PPE.[241] Cochrane published a review of the impact of non-pharmaceutical interventions on risk of infection, transmission, and death from SARS-CoV-2 in care homes which included 22 studies published before January 2021. This found uncertain evidence of efficacy for all the interventions that were evaluated including PPE, asymptomatic screening, outbreak control measures, and regulation of entry of staff and residents.[156]

2.4.1 Strengths & Limitations

Strengths of my review are the systematic nature of the search which was designed with a medical librarian who was experienced in literature search design. As many studies early in the pandemic had been pre-printed, I repeated my search two years after my baseline search to allow time for the publication process and improve the quality of included papers by ensuring they had all been peer-reviewed. In addition, by focussing on facility factors only and splitting outcomes into those describing

ingress and transmission of infection, my findings can help identify specific areas to maximise the effectiveness of preventive measures by focussing on either infection entry or infection spread. This approach has not been taken by other reviews but could facilitate more targeted use of resources and potentially limit the negative consequences of policies, as care homes may have more opportunities to stop infection spread than ingress given the challenges associated with the latter.

Limitations include heterogeneity between studies that prevented direct pooling and comparison of results. Case definitions differed as access to testing varied between sites and over the pandemic. For example, only 65% of studies used virologically confirmed diagnoses, whereas the rest were guided by clinical signs and symptoms. As outbreak definitions varied from any case in the facility, to a case in a resident, to two cases within fourteen days, direct comparison is challenging. In addition, as there is an element of chance in the occurrence of outbreaks, approaches to sampling have a big impact on the number of outbreaks captured. Most studies used an opportunistic sample usually from one geographic area or provider and none applied a random sampling frame, increasing the risk of sampling bias. Most studies were conducted in the first two waves of the pandemic however, as they had been conducted across fourteen countries, there was significant heterogeneity in epidemiology due to rapid changes to policies across borders, circulating variants, and population levels of immunity to infection over this time (described in Chapter 1). As there are large differences in the structure of the adult social care system between countries and regions,[242,243] measurement of predictors was not consistent across studies. For example, ownership type was defined as for-profit or not-for-profit in some studies, whereas others compared publicly to privately funded facilities.

Over half of the studies in this review relied on results from cross-sectional surveys, many of which were self-completed by staff. This increases the risk of reporting bias as these were completed over a strained period therefore staff may have been unable to double-check answers. Most data on control measures were collected from these surveys, therefore practices that fell short of recommendations may not have been reported, increasing the chance of social desirability bias. As surveys are completed at one timepoint, only association and not causation can be inferred from the data and reverse causality is possible.

Few studies presented raw data and there was substantial heterogeneity between studies, therefore meta-analysis was not performed. As such, it is difficult to compare the size of the effect for different predictors as studies used different models and adjusted for different predictors.

2.5 Future Research

This review has highlighted the challenges around data collection and the low availability of quality evidence around the variation in burden of SARS-CoV-2 between care homes and factors associated with infection and outbreaks. Prospective longitudinal collection of data from a care home cohort could help to overcome some of these limitations. It would also facilitate evaluation of the effects of epidemiological and policy changes over successive waves of the pandemic, such as vaccination or emergence of novel variants, which has not been considered in any of the studies included in my search. Although several studies have relied on postcode matching to identify residents, more reliable approaches should be considered. Data on building characteristics, which are absent from currently published studies, could provide deeper insights into infection transmission. Finally, a representative sample should be sought from a range of care home types, providers, and geographic locations, to ensure that conclusions are generalisable to the wider population.

The VIVALDI study was established in England in May 2020 and aimed to address many of the gaps that I have identified in this literature search, specifically in relation to understanding disease burden, how levels of immunity vary, and how the built environment of care homes impacts on the risk of SARS-CoV-2 infections and outbreaks. In Chapter 3, I will describe the study design and the integral role that I played in establishing it. I then use the study to address my remaining PhD objectives.

Chapter 3

Establishing the VIVALDI cohort study in care homes

At the start of the pandemic, there were significant gaps in the research and surveillance infrastructure in care homes, outlined in Chapter 1 and highlighted in my scoping review (Chapter 2). Given the unique challenges associated with research in care homes and the diversity of the sector, a large and setting-specific study was required to consider the burden of SARS-CoV-2 infection in care homes and identify associated agent, host, and environmental factors.

The UCL-led VIVALDI study is a prospective open cohort study that was established as a SARS-CoV-2 seroprevalence study in care homes for older people in England. Consenting participants underwent sequential blood sampling over a period of up to three years and samples were tested for humoral and cellular aspects of the host immune response to infection and vaccination. Additional data linkage was undertaken for all staff and residents in care homes to data on SARS-CoV-2 testing and vaccinations, as well as hospitalisations and deaths. This took advantage of a national asymptomatic screening programme where tests were linked to care homes, allowing identification of staff and residents for the first time. Linkage to facility-level data provided more granular data on the environment in different care homes.

Data from VIVALDI played a significant role in national policy decisions over the pandemic. Having helped to establish the study, I was able to use this as the setting for the analytical objectives of my PhD (Objectives 2-4) where I considered the burden of infection in care homes and facility-level host and environmental factors associated with infection and outbreaks.

3.1 Research landscape

Care home residents' vulnerability to SARS-CoV-2 was highlighted by higher excess mortality, when compared with community-dwelling older adults in the first wave of the pandemic.[244] In England, the separation of social care from the National Health Service (NHS) accentuated the disparity in availability and quality of surveillance and research across each of these sectors. Although clinical research has been embedded

in healthcare, the challenges of conducting large-scale studies in the fragmented social care system have hampered research in these settings.[150] As described in Chapter 1, approaches to identifying staff and residents using routinely available data are unreliable which has reduced the accuracy of surveillance and any associated research.[179] Early in the pandemic it became clear that a large prospective study in care homes could overcome some of these challenges and help to estimate infection prevalence and compare and describe immune responses in staff and residents. In addition, once SARS-CoV-2 vaccinations had been deployed, their efficacy could be evaluated within this cohort.

3.2 Establishing the VIVALDI cohort study.

The VIVALDI study was established in late May 2020 to collect information about seroprevalence, immunity, and outcomes following infection and vaccination amongst the staff and residents of care homes for older people in England.[245] The study is one of the government-funded national core COVID-19 surveillance studies established amongst key populations to directly inform policy decisions.[246] VIVALDI is led by UCL researchers who collaborate with researchers at the University of Birmingham and the Francis Crick Institute.

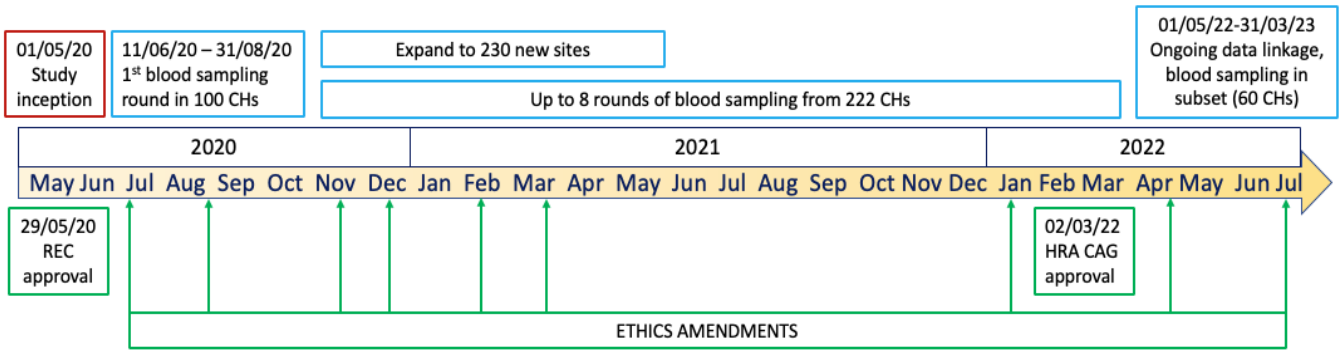
Additional funding was granted to extend follow up until April 2023 in a smaller group of participants after a new round of consenting, however as this cohort is not included in my PhD, I have not described this in detail.

3.2.1 Study design

The study is a prospective, open cohort study that began as a seroprevalence study in 100 care homes but was extended to become an immunological cohort study (see 3.2.1.1). At the request of the DHSC this was expanded in autumn 2020 to a larger set of homes (additional 122 care homes). In addition, the introduction of the Control of Patient Information (COPI) notice (see 3.3.3) made it possible to access routine data for all staff and residents in participating care homes without consent, for the purposes of responding to the pandemic. Consequently, routine data from all staff and residents in 330 participating care homes (including those involved in the seroprevalence study) were included in the study (see 3.2.1.2). These individuals were

identified using SARS-CoV-2 test results derived from the national care home COVID-19 screening programme, introduced in June 2020 but fully operational from the following September.[247] The study commenced on 29th May 2020 and finished on 31st March 2023, with data linkage for existing participants until 31st March 2024. The study timeline is outlined in Figure 3.1.

Figure 3-1: Study timeline



CH care home REC Research Ethics Committee
HRA CAG Health Research Authority Confidentiality Advisory Group

Only care homes in England were included due to the complexity of gaining ethical approvals for research in all four nations and differences in the databases where testing data are stored. All staff and residents in participating care homes were eligible for inclusion in both the immunological cohort and the broader observational study.

3.2.1.1 Immunological Cohort

Informed consent procedures (see 3.2.2) were undertaken in a subset of individuals from 222 care homes participating in the immunological cohort study (100 from June 2020 and additional 122 from Nov 2020). All staff and residents with valid consent were eligible for inclusion. There were no exclusion criteria.

Sequential serum and plasma samples were drawn at eight-to-twelve-week intervals between 11th June 2020 and 31st March 2022, with further samples collected as part of the study extension in a subset of 60 care homes between 1st May 2022 and 31st March 2023. These samples were analysed for B and T cell responses to SARS-CoV-

2 and other respiratory viruses by study collaborators; The Doctor's Laboratory (TDL), University of Birmingham (UoB), and the Francis Crick Institute (see 3.2.3).

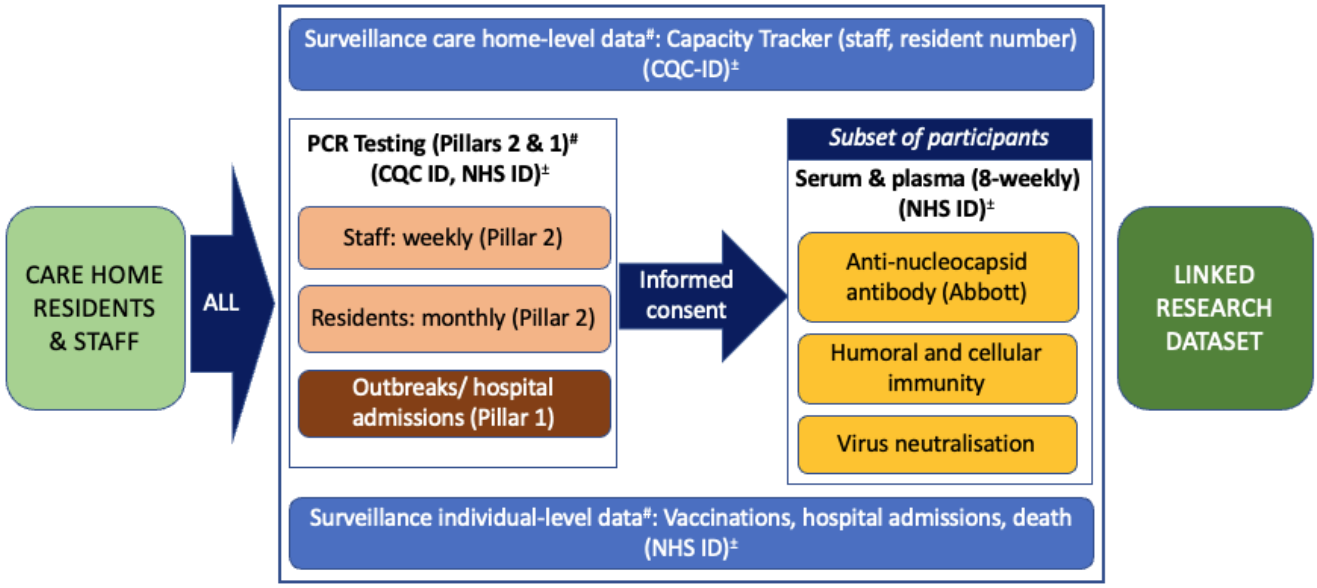
3.2.1.2 Observational Cohort

As part of the national Pillar 2 screening programme, between June 2020 and March 2022, staff were tested weekly and residents monthly using PCR or LFD (Chapter 1).[171,172] Additional testing was undertaken in outbreaks and during hospital admissions under the Pillar 1 programme.[248] As tests were registered with the care home location details, tests and therefore individuals could be linked to care homes based on the CQC-ID. This is a unique identifier that is allocated to every care home in England by the CQC. The frequency and coverage of testing optimised the capture of asymptomatic infections and facilitated accurate identification of the care home population. Due to the frequency of testing (at least monthly in residents, weekly in staff), it was also possible to use testing dates to estimate dates of entry and exit into the care home.

Each test result was linked to the individual's unique COVID-19 pseudo-identifier based on their NHS number. Using this pseudo-identifier, PCR/LFD results could be linked to national, routine datasets on vaccination, hospitalisations, and deaths, as well as care home level datasets.

Subjects were excluded if they had been identified as a visitor or visiting professional or if they had no PCR/LFD tests that could be linked to a participating CQC-ID during the study period. Data linkage was undertaken for all eligible participants between 1st March 2020 (when COPI came into effect) until 31st March 2023 with additional data linkage for existing participants until 31st March 2024. Figure 3.2 outlines the study flow.

Figure 3-2: Study flow diagram



Data available from 1st March 2020
± Enables linkage between datasets.

3.2.2 Consent procedures

VIVALDI was set up extremely rapidly to inform the public health emergency response. When it was initially established, the country was still under national lockdown with recommendations against care home visits.[93,160] These restrictions remained in place in some capacity over most of the period of recruitment to the study. Along with the geographic dispersal of the participating care homes, this prevented members of the small research team from visiting the care homes to assess capacity and to consent participants. Later in the study, the Clinical Research Network (CRN) assisted with consenting in smaller and independent care homes, however this took time to set up. In view of the urgency of study findings, we partnered with care providers who were able to quickly deploy resources.

Consent was received by senior care home staff, such as managers or senior nurses. None had undertaken *Good Clinical Practice* training, which is recommended by the Health Research Authority (HRA),[249] however providing this training within the study timelines was not feasible. All had extensive experience and training in assessing capacity and consent which they applied in their daily practice. Subjects were given a clear description of the study in oral and written form. Capacity to decide on study involvement was assessed by these senior staff and in cases where residents were

assessed as lacking capacity to consent, declaration forms were completed by personal (i.e., next of kin) or nominated (i.e., care professional) consultees. Study information packs were mailed to all next of kin to make them aware of the study and followed up by telephone to address any questions or concerns. The approach of using consultees is consistent with previously published studies, for example during an investigation into a cluster of care home scabies outbreaks,[250] and complies with the five main principles of the Mental Capacity Act 2005.[251] Our approach was assessed as appropriate by the Research Ethics Committee (REC). In Chapter 7, I have reflected on the strengths, challenges, and learning points around our consenting approach.

3.2.3 Laboratory procedures

At each visit, a separate 5ml serum and 5ml plasma sample was drawn from each participant. Serum samples were processed at TDL using the Abbott ARCHITECT i-system (Abbott, Maidenhead UK) semi-quantitative SARS-CoV-2 anti-nucleocapsid IgG antibody immunoassay. Residual samples were sent to the UK Biocentre for aliquoting and storage for future use. A subset was sent to the Francis Crick Institute to investigate neutralisation of SARS-CoV-2 variants.

Plasma samples were sent directly to the Prof Paul Moss's laboratory at the University of Birmingham where they had in-depth assessments of humoral and cellular immune responses to SARS-CoV-2 and other respiratory viruses. These included quantitative assays for antibodies against spike protein and its components using the Meso Scale Discovery system (Meso Scale Diagnostics, Rockville MD USA) based on electrochemiluminescence.

Within the national surveillance programme, nasopharyngeal samples were self-collected by participants (or with assistance where required) and sent to a national network of laboratories, rapidly established for infection surveillance.[252] Samples were analysed for presence of SARS-CoV-2 virus using RT-PCR and cycle thresholds (Ct) were reported for each target. PCR assays and Ct positivity thresholds varied between laboratories in line with local protocols.[253] A subset of viral isolates from nasopharyngeal samples underwent whole genome sequencing by the COVID-19

Genomics UK Consortium (COG-UK) under separate ethical approvals (Public Health England Research Ethics Governance Group (reference: R&D NR0195)).[254] After roll-out of LFD testing, results were self-reported into the national NHS reporting system.

3.2.4 Ethical approvals

Ethical approval for the study was granted by the South Central – Hampshire B REC (20/SC/0238) on 29th May 2020. The study has been registered with the International Standardised Randomised Controlled Trail Number (ISRCTN) registry, number 14447421.

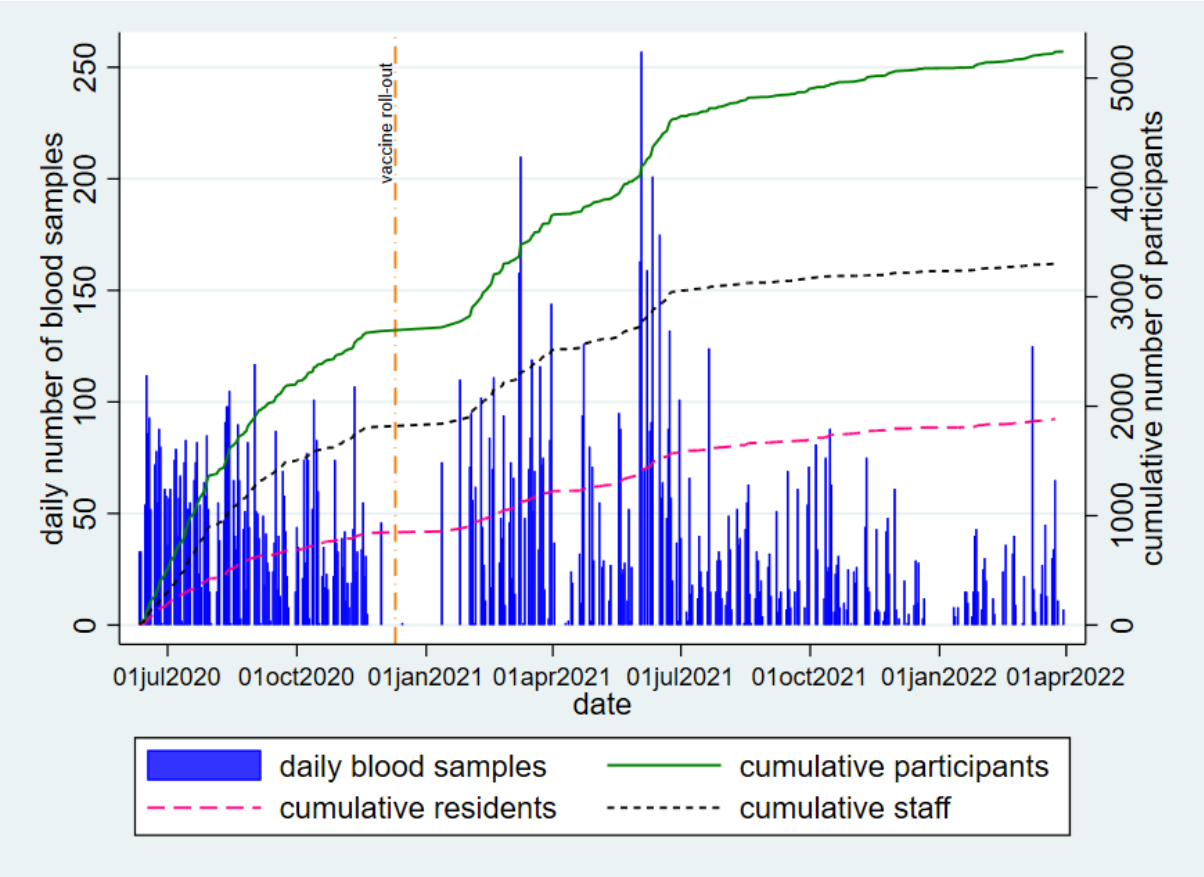
3.2.5 Participant recruitment

Building on an existing research collaboration with the Chief Investigator (CI), in the initial phase of the study care homes were recruited from Four Seasons Health Care (FSHC), a large for-profit care provider. [255] This allowed rapid enrolment of staff and residents in 100 care homes across England. Within the first two months of recruitment, 2334 participants (1561 staff, 725 residents, 49 unknown) had each donated one blood sample.

To improve generalisability, the study was granted additional funding from DHSC to expand to more providers and independent care homes. From November 2020, two additional large chains began recruitment, The Orders of St Johns Care Trust (OSJCT), a medium-sized not-for-profit chain, and HC-One, a large for-profit provider. Through the National Institute for Health and Care Research (NIHR) ENRICH network[256] - a network of research-ready care homes - independent homes and small chains were also recruited. This included Friends of the Elderly (FOTE), a small not-for-profit chain with nine homes, and Evolve specialising in dementia care who enrolled five homes.

By April 2022 when the first part of the study ended, there were approximately 6131 consented participants from 222 care homes, 5298 (86.4%) of these (3554 staff, 1712 residents, 138 unknown) could be reliably linked to a pseudo-identifier. Recruitment of participants and blood sampling over the study period are illustrated in Figures 3.1 & 3.3.

Figure 3-3: Total blood samples per day, cumulative participants overall and by subject type (11th June 2020 to 1st April 2022).



Participants have donated up to eight consecutive blood samples at intervals of six-to-twelve weeks. Due to challenges with data linkage (described in 3.4), exact estimates of the number of individuals who have donated blood samples overall are not possible as approximately 2649/15457 (17.1%) of samples could not be linked to an identifier so not all sequential samples from the same individual could be identified.

The observational immunology cohort was complemented by data linkage to routine datasets for all staff and residents in participating care homes, regardless of consent, including 108 care homes that had not participated in blood sampling, but were willing to take part in the study (see 3.3.3 for legal basis for accessing data). In total, there were approximately 70,000 individuals tested for SARS-CoV-2 at least once in the 330 included care homes.

3.2.6 Outcomes

The primary outcome was the proportion of staff and residents in participating care homes with naturally acquired antibodies against SARS-CoV-2 over sequential blood sampling rounds, suggesting prior infection. Secondary outcomes have changed over the course of the study but most notably included duration of antibody response against SARS-CoV-2; magnitude of SARS-CoV-2 humoral and cellular responses amongst residents; COVID-19 related mortality in residents; effectiveness of SARS-CoV-2 vaccinations against infection, hospitalisation and death in residents and staff.

3.2.7 Recruitment target and sample size

The recruitment target was initially based on sample size calculations for precision of seroprevalence estimates. This was based on an assumed 80% participation rate amongst staff and residents (38 residents, 49 staff per care home) across 105 care homes (the planned number of care homes from FSHC), therefore 9135 participants. The intra-cluster correlation was estimated to be 0.36 (twice the published figures from seasonal influenza[257]). Using these estimates to account for loss of precision due to clustering, the study would have an effective sample size of 279 residents and 285 staff. Estimating an antibody prevalence of 30%, this would give a precision of estimates of +/- 5.4% for both residents and staff.

Sample size calculations were performed by the study statistician. However, given the urgency of the study, consenting was performed very rapidly so the recruitment target was not met in every care home. Given the changing research priorities, the initial sample size calculations were less applicable to later analyses that relied on the data linkage as all staff and residents were included in these under COPI (see 3.3.3).

3.3 Data sources and data linkage

3.3.1 Outline of datasets

The study accessed individual and care home level datasets, Table 3.1. Data were predominantly collected for routine surveillance, although results of immunological assays were available from samples donated for the study. Aggregate data on staff and resident turnover and bed funding were also collected directly from care homes, in addition to the care home building survey described in Chapter 6.

Table 3-1: Data items and sources

Item	**Source; owner**
Individual level	
Demographics - Pseudo-identifier (linkage) - Age - Sex - Ethnicity	All
CQC-ID	Pillar 2; DHSC / UKHSA
Resident or staff	Pillar 2; DHSC / UKHSA
PCR / LFD - Sample date - Sample result (binary) - Sample identifier - Testing laboratory - Cycle threshold by target	Pillar 1; NHSE / PHE Pillar 2; DHSC / UKHSA
COVID-19 vaccination - Date - Dose number - Vaccine type	National Immunisations Management System; NHSD
Hospital admission - Admission date - Discharge date - Primary diagnosis - Underlying conditions	Hospital Episode Statistics; NHSE
Death - Date - Cause of death, primary - Cause of death, secondary / contributing	ONS
Nucleocapsid antibody (Abbott) - Sample date - Sample result (binary) - Sample identifier - Antibody titre (semi-quantitative)	The Doctors Laboratory; UCL / DHSC
Whole genome sequencing - Viral lineage	COG-UK; UKHSA
Immunology - Antibody testing using MSD - ELISpot tests	University of Birmingham; UCL, DHSC
Care home level	
Capacity - Number of beds - Number of staff - Number of occupied beds - Record date	Capacity Tracker; NHS England and the Better Care Fund
Address - Care home name - Provider - Postcode	Care home registry; CQC
Deaths - Monthly COVID-19 associated resident deaths	CQC

Bed funding & turnover - Monthly staff turnover - Monthly resident turnover - Proportion of beds funded for dementia care and by the Local Authority	Directly from care homes; care homes

DHSC Department of Health & Social Care
PHE Public Health England
NHSD National Health Service Digital
ONS Office for National Statistics
UKHSA UK Health Security Agency
NHSE National Health Service England
CQC Care Quality Commission
COG-UK COVID-19 Genomics UK Consortium

3.3.2 Data pipeline

National surveillance datasets are updated and stored within the COVID-19 datastore (https://data.england.nhs.uk/covid-19/),[258] a secure data platform established and maintained by NHS Digital as part of the pandemic response. Within this datastore, staff and residents from care homes participating in VIVALDI were identified using the CQC-IDs and pseudo-identifiers linked to samples tested within the Pillar 2 screening programme. Using pseudo-identifiers, linkage was undertaken to other stored datasets including Pillar 1 and 2 testing data reported through the National Pathology Exchange (NPEx) for Pillar 1 and 2 results;[259] the National Immunisations Management System (NIMS) of all COVID-19 vaccinations in England;[260] Hospital Episode Statistics (HES) which details all attendances to NHS hospitals in England;[261] and mortality data collected by the Office for National Statistics (ONS).

Personal details for subjects undergoing blood sampling (including NHS number, name, date of birth) were sent securely by care home managers to TDL. This was necessary to ensure antibody results could be fed back to participants and linked to the participant's pseudo-identifier (based on their NHS number) to allow linkage across other datasets. Staff who provided a mobile phone number received results by text message, results from residents were sent by mail to the care home manager. Antibody results were uploaded weekly by TDL to the Foundry and pseudonymised by NHSE prior to ingress and subsequent data linkage. Many care homes could only provide name, date of birth, registered GP, and address as they did not hold NHS numbers for residents or staff. Due to discrepancies in spelling or outdated information, it was more challenging to use these variables to link to pseudo-identifiers (despite multiple attempts at correcting the data) therefore not all samples could be linked to a participant in the final dataset. To address this issue, we made NHS number a mandatory data field and asked the TDL project manager to ensure these data were

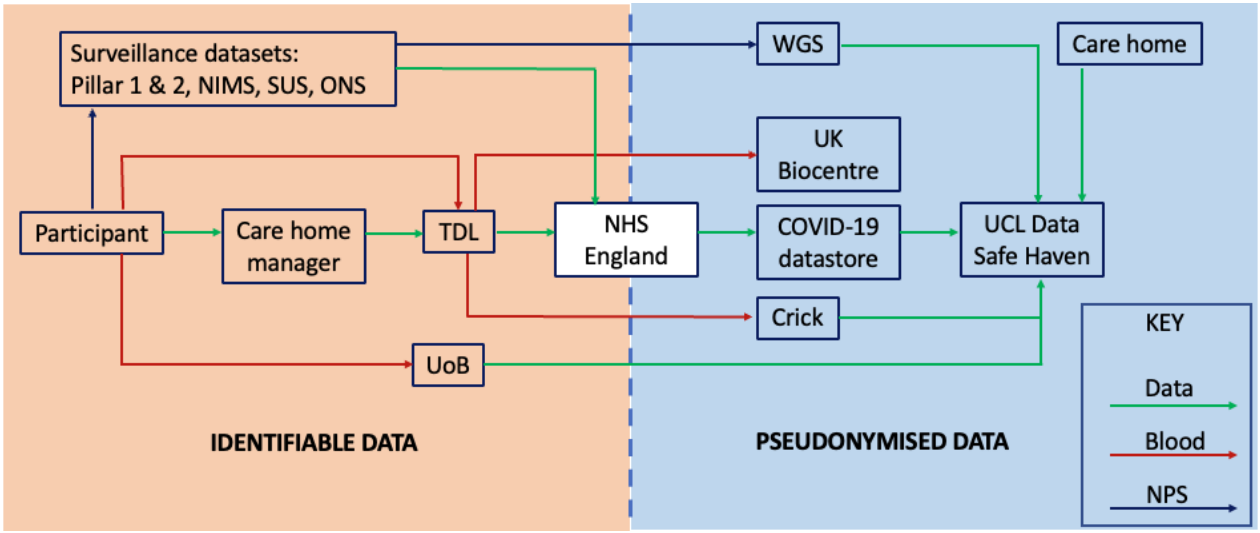
provided, which resulted in improved linkage over time. Nonetheless, we had to make exceptions for care homes who had already enrolled in the study and could not access these data.

The CQC-ID number was also used to link to the Capacity Tracker dataset (https://www.necsu.nhs.uk/capacity-tracker) stored within the COVID-19 datastore, a tool established in the pandemic to allow care homes to report capacity centrally and inform provision of additional support.

Data extracts from the linked datasets were regularly uploaded to the UCL Data Safe Haven (DSH) where they are stored, a secure datastore that has been certified to the ISO27001 information security standard. The datastore uses a "walled garden" approach and has been approved by NHS Digital's Information Governance Toolkit.[262] Additional datasets were imported into the DSH on an ad-hoc basis including results of additional immunological assays, Cycle threshold (Ct) values, genome sequences from viral isolates, and care home turnover and funding. Data flows are outlined in Figure 3.4.

All analyses were performed in the COVID-19 Datastore or in the UCL DSH.

Figure 3-4: VIVALDI study data flows



TDL The Doctors Laboratory
UoB University of Birmingham
NIMS National Immunisations System
ONS Office for National Statistics
WGS Whole Genome sequencing
NPS Nasopharyngeal swab
SUS Secondary Uses Service

3.3.3 Legal basis for accessing data.

When the study was set up, the legal basis for accessing pseudonymised data regardless of whether participants have consented was under Regulation 3(4) of the Health Service (Control of Patient Information) Regulations 2002 (COPI) which was in place between March 2020 and June 2022.[263] This allowed collection of patient data by organisations engaged in the national COVID-19 response. From July 2022 the legal basis transitioned to Regulation 5 of the COPI Regulations 2002, for which we were granted approval from the Health Research Authority Confidential Advisory Group (HRA CAG) ref 21/CAG/0156. I was responsible for drafting the CAG application with support from the CI.

The joint data controllers for the study are DHSC and UCL, data processors are NHS England.

The study privacy notice is available here: https://www.ucl.ac.uk/health-informatics/research/vivaldi/vivaldi-privacy-notice

3.4 Operational challenges

Although Chapter 7 includes detailed reflections on the operational challenges with the study, a brief overview is outlined below.

To inform urgent policy decisions, the timescales for study set-up were very tight, Figure 3.1. Coordinating study delivery whilst ensuring compliance with research and information governance frameworks was particularly challenging as this was conducted during a period of national lockdown with strict restrictions on social contacts and movement. Nonetheless almost 2000 participants were consented within the first month and blood collection commenced less than a fortnight from REC approval. This was possible through close collaboration between researchers at UCL and policymakers at DHSC, strong support from senior decisionmakers in DSHC, and by working in partnership with the care homes themselves. From the outset, an internal project manager from each provider oversaw recruitment within care homes. These project managers applied their experience to inform standard operating procedures (SOPs) that were developed in partnership with the study team for the sites. This

facilitated rapid roll-out by building trust across the organisation and through internal promotion of the study. This was also aided by close integration with policymakers (for example by assisting in securing equipment, outlined below) which ensured that study questions addressed national policy priorities.

3.4.1 Procurement & contracts

The speed with which the study was established meant that operational challenges, a few of which I have outlined here, had to be addressed very quickly. To allow rapid blood sampling during the first lockdown, three private phlebotomy companies were contracted. These were based in different parts of the country to facilitate multiple care homes visits on the same day (up to five per day for the first year). As UCL had not contracted with any phlebotomy providers at the time, a tendering process was completed within three weeks. Although contracts with multiple parties were required at short notice, such as data sharing and data processing agreements and separate contracts with each care provider, outsourcing of legal services to a private firm who allocated one solicitor to the study increased the efficiency of this process.

Due to national shortages, clinical equipment, and personal protective equipment (PPE) was rapidly secured and distributed to phlebotomists, with assistance from DHSC. A system was developed for care homes to send participant details to the laboratory so sample labels could be distributed in advance of phlebotomist visits, recognising that most care homes lacked the time and resources to print labels or participant lists locally. Samples were collected by designated couriers after every visit and delivered on the same day to laboratories in London and Birmingham. To reduce the risk of delivery of samples to the wrong laboratory, phlebotomists used different coloured sample bags according to the testing laboratory (serum to TDL, plasma to University of Birmingham). In view of limited storage for clinical waste on care home sites, we arranged for additional waste collection after each phlebotomy visit.

3.4.2 Ethics amendments

In view of the rapidly changing epidemiological situation, over three years there were nine amendments to the study protocol submitted for REC review - three substantial and six minor, Figure 3.1. This has allowed the study to adapt and address new research questions which can inform policy in near real-time, such as estimating the

effectiveness of newly developed vaccines in the care home population.[161] However, this has impacted on our ability to plan and conduct analyses in the medium to long-term, as amendments were often required at very short notice.

3.5 Patient & Public Involvement & Engagement (PPIE)

In this section I have described the PPIE activities that have contributed to the initial study set up. Since then, VIVALDI has benefited from a broad programme of engagement with participants, sector representatives, and the wider community, described in detail in Chapter 7.

3.5.1 Consultation with stakeholders

The study design was discussed with members of the National Care Forum (NCF) - a forum of not-for-profit providers across the country, senior care staff at FSHC, and senior policymakers within DHSC, PHE and later the UK Health Security Agency (UKHSA). Examples of how this feedback informed practice are that antibody results were communicated to staff by text message instead of post (as originally planned), changes to participant-facing study documents were made to improve clarity, and translation of study materials into Braille to facilitate the needs of a particular care home.

3.5.2 Establishing a PPIE group.

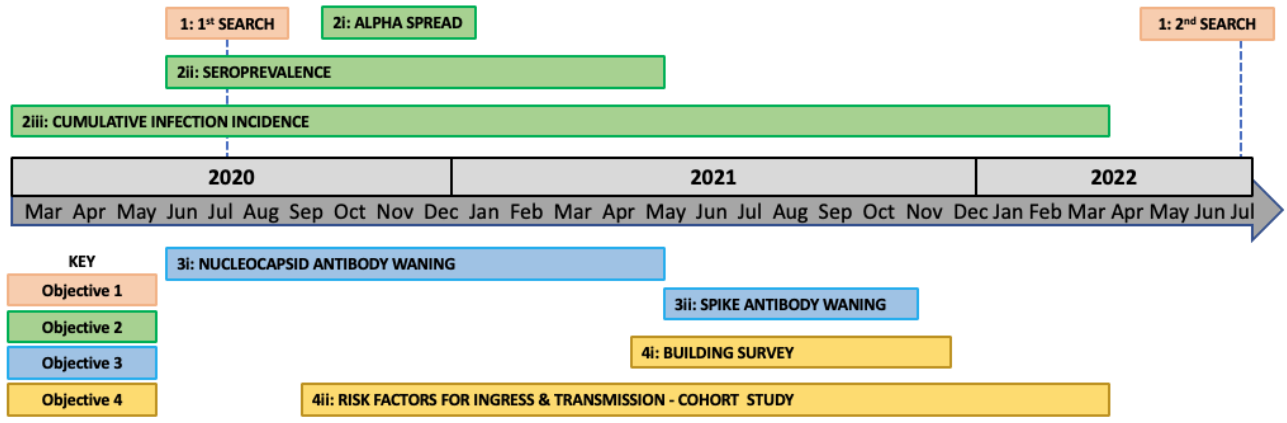
The short timescales for study establishment were insufficient for the recruitment of a PPIE group. However, a PPIE group consisting predominantly of family members of care homes residents was convened a few months into the study. We advertised widely through the providers we were working with, the UCL website, UCL Institute of Health Informatics Twitter account, the NCF, and the CRN and ENRICH networks. Despite this, as recruitment took place in the first half of 2021 during the second wave of infection, we were not successful in recruiting any staff or residents. Two groups of ten relatives of care home residents attended introductory sessions about the study. Subsequently, lay summaries of key research papers were developed with their input and these were distributed amongst study participants and displayed on the study website. Feedback was also obtained from them on plans for study expansion and data privacy concerns (described in Chapter 7).

3.6 Conclusion

Globally, VIVALDI is the largest cohort study of SARS-CoV-2 infection in a geographically representative sample of care homes. The study has capitalised on asymptomatic testing to improve the accuracy of estimates of infection burden and enable linkage to a range of routine data for everyone in the care home. This has created a blueprint for future research in care homes and has set a precedent for inclusive research that gives equal importance to social care when compared with healthcare within the NHS where data are readily accessible.

The VIVALDI data infrastructure facilitated the analyses that I have outlined in Chapter 1 and all studies described in Chapter 4-6, Figure 3.5. Chapter 4 relies on a combination of sero-sampling and linkage to PCR/LFD testing data and facility-level capacity data to consider approaches to measuring the burden of infection in care homes and how this varies by care home and by variant. Chapter 5 uses detailed antibody testing data with linkage to PCR/LFD testing data, vaccination records and facility-level data to consider host responses to infection and vaccination. Chapter 6 combines linkage to individual and facility-level records with results from a care home level survey of the built environment to consider how agent, host, and environmental factors affect the risk of infections and outbreaks.

Figure 3-5: Timeline of cohorts for PhD objectives



3.7 Contribution statement

The study protocol has been published in an open-access peer-reviewed journal, I am the lead author.[245]

Since the inception of VIVALDI, I have worked to establish and manage the study and relationships with study collaborators. I worked with the CI to gain the necessary ethical approvals from the REC and to register the study with the ISCRTN. This involved developing the study protocol, the participant-facing materials, completing a data protection impact assessment (DPIA), and developing standard operating procedures for the study sites and for the phlebotomists attending the sites. I led the application to the HRA-CAG for section 5 support. This required close working with the UCL Information Governance team, review of data security systems in place, and engagement with the key stakeholders to ensure acceptability of the proposed approach. I developed a bespoke opt-out system for the study in response to feedback from the initial CAG review, which has been implemented across the study.

Operationally, I led the procurement and contracting with the phlebotomists and was responsible for day-to-day management of study finances for the first six months when I was employed as the project manager before starting my PhD. I also established and maintained working relationships with couriers, laboratories, equipment suppliers, and the care homes themselves. As VIVALDI works closely with the government-run national testing programme, I liaised closely with the operations team at the DHSC to develop new PCR testing kits and sample bar-codes that can allow laboratories to identify and cherry-pick VIVALDI samples for further sequencing. I have also led the process of identifying, contracting and project managing the long-term storage of serum samples in a biobank for use by other researchers (described in Chapter 7). Although this unusual for a PhD, it reflects the extreme pressure of the pandemic which necessitated intensive input to deliver the study quickly.

The protocol is published in Wellcome Open Research:
Krutikov M, Palmer T, Donaldson A, *et al.* Study Protocol: Understanding SARS-Cov-2 infection, immunity and its duration in care home residents and staff in England

(VIVALDI). *Wellcome Open Research 2021 5:232* 2021; **5**: 232 DOI: 10.12688/wellcomeopenres.16193.2

Chapter 4

Objectives 2: Testing the hypothesis that it is possible to measure the proportion of care home staff and residents infected with SARS-CoV-2 and that there is substantial variation between care homes.

As described in Chapter 3, the VIVALDI study infrastructure established in the early pandemic, allowed accurate identification of care home staff and residents for the first time. Routine SARS-CoV-2 screening meant that asymptomatic infections could be identified alongside symptomatic cases, thereby optimising the accuracy of estimates of infection burden. However, this programme was only introduced towards the end of the first pandemic wave following a period of high infection mortality in the care home population. Limited access to PCR testing over the first wave impaired the accuracy of incidence estimates. However, it is possible that serological surveys of a representative sample of the care home population could identify the proportion of care home staff and residents who have been infected with SARS-CoV-2.

The second objective of my thesis was to demonstrate variation in the burden of SARS-CoV-2 infection between care homes; a measure of the agent. As serological surveys can also describe population-level immunity (host factors) that may be protective against infection, I split my second objective into two sections. The first considered how to measure the burden of infection and the second estimated population-level immunity using serological surveys.

I implemented three approaches to measuring infection that analysed data from serological surveys or PCR tests. I also considered how infection burden varied with variant, geography, and over time. First, I used PCR data to describe spread of the emergent Alpha variant across regions of England over three months. Second, I calculated point estimates of weighted seroprevalence at two-month intervals over the first 15 months of the pandemic. Finally, I combined PCR and antibody data to estimate the cumulative incidence of infection for staff and residents over the first two years of the pandemic.

My analysis demonstrated rapid spread of Alpha variant into care homes in areas of high community incidence, despite control measures, demonstrating how agent factors play a significant role in infection risk and supporting the findings from my scoping review that preventing infection ingress may not be possible. I showed some variation in seroprevalence over time, however after 15 months at least one-third of residents and one-quarter of staff remaining in the care home, had evidence of infection, rising to two-thirds in the care home after two years. These measurements of infection burden will lay the foundations for my remaining objectives (3 and 4) to identify host and environmental factors associated with infection and outbreaks in care homes.

4.1 Background

In Chapter 1, I described the epidemiology of SARS-CoV-2 in the first two years of the pandemic, and how this impacted on care homes. To inform policy throughout this period, it was essential to measure how much infection had entered care homes, whether non-pharmaceutical measures to prevent infection were effective, and how much infection-induced immunity the population had. This could help to identify areas in need of additional support that could be prioritised for preventive measures such as regional lockdowns and later vaccinations. In addition, less effective measures could be scaled back if their risks appeared to outweigh their benefits. As previously described in Chapter 3, the VIVALDI study played a key role in addressing these gaps.

During the second lockdown in the final quarter of 2020, rapid spread was described of a new B.1.1.7 viral variant from the East and South of England to the rest of the country[264] and then globally over the subsequent months.[265,266] This variant, renamed Alpha, first emerged in September 2020 in Kent, England, with multiple mutations in the spike protein (used by the virus to gain cell entry)[267,268] which increased the transmissibility of the virus.[269] There was also early evidence of increased disease severity when compared with wild-type strains.[270,271] As spread occurred very quickly in a largely unvaccinated population, greater morbidity and mortality from infection was described nationally.[272–275] Care home residents are known to be at greatest risk from severe complications of infection, therefore it was urgently important to understand whether Alpha variant had spread to the care home population and identify the worst-

affected areas. These regions could benefit from additional resources and be prioritised for stricter measures. This could also inform future work modelling infection spread and immunity.

Alongside these efforts to monitor the spread of infection based on PCR testing, serial blood sampling of care home staff and residents and linkage to routine datasets within the VIVALDI study, provided the ideal setting for addressing questions on infection exposure, immunity, and spread within this population. Although mortality from infection was substantial in the first wave, estimating the true incidence of infection over this time was challenging because of restricted access to infection diagnostics.[276] However, it is possible that measuring the prevalence of antibodies against the SARS-CoV-2 could help to estimate the number of people who had survived infection. As described in Chapter 1, following SARS-CoV-2 infection antibodies directed against different viral proteins including the nucleocapsid protein are produced.[277] In contrast, vaccination only elicits production of anti-spike antibodies. As such, detection of anti-nucleocapsid antibodies can distinguish between exposure through vaccination or infection[278] and probably correlates with protection against re-infection.[279] However waning of antibody levels over time and turnover of staff and residents in the home may affect these estimates.

The second objective of my PhD thesis was to consider approaches to measuring infection burden and investigate whether there was variation in the proportion of care home staff and residents with SARS-CoV-2. Once I had demonstrated this variation, I could consider facility-level host and environment factors associated with infection (Objectives 3 and 4). In addition, by considering how infection burden varied with the variant, I was able to describe how the agent influenced infection risk (as described in Chapter 1).

Although the focus of this objective was measuring variation in SARS-CoV-2 infection, seroprevalence estimates also provide some measure of population-level immunity of the host, which is an important component of the epidemiological triad and which I will explore further in Chapter 5. I therefore decided to address two main questions in my second PhD objective using three sub-studies (i-iii). The first question was addressed in all three sub-studies and considered what proportion of care home staff and

residents had SARS-CoV-2 infection over the first one to two years of the pandemic and how this varied with variant, geography, and time. The second question on how infection spread affected seroprevalence across care homes, which could have implications for immunity, was only addressed in the second and third sub-studies (ii and iii).

Given the difficulties with measuring the burden of infection in care homes, I applied different approaches to measuring infection in each sub-study, outlined below. These used data from routine PCR testing linked to facility-level data (i and iii) or data from seroprevalence surveys (ii and iii).

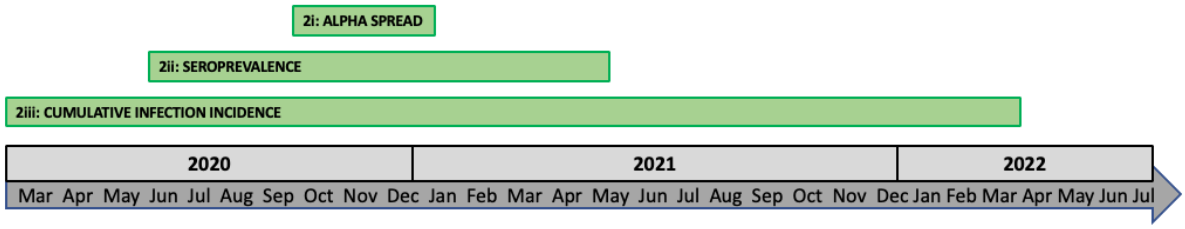
i) I used testing data (PCR) to describe the early spread of the novel B.1.1.7 variant in care homes between October and December 2020 at a time of strict infection control measures, therefore rapidly informing public health policy.

ii) I estimated the proportion of staff and residents infected with SARS-CoV-2 between March 2020 and May 2021 (when all participating care homes had undergone at least one round of blood sampling), by calculating serial point estimates of anti-nucleocapsid seroprevalence in the VIVALDI cohort.

iii) I estimated the cumulative incidence of infection for any given staff member or resident in a participating care home from March 2020 until March 2022 (when asymptomatic SARS-CoV-2 testing in care homes ended)[172] as a measure of the attack rate of the virus and to inform estimates of the proportion with some level of natural immunity.

Timeline of analyses is shown in Figure 4.1.

Figure 4-1: Timeline of cohorts for Objective 2



4.2 Methods

4.2.1 Measuring B.1.1.7 variant spread

In the fourth quarter of 2020, S-gene detection was considered a reliable marker to distinguish Alpha from wild-type variant as 90-100% of samples with S-gene target failure (SGTF) were confirmed as B.1.1.7 on sequencing.[267,280] Although a national programme of genomic surveillance was established early on,[50] the coverage for samples from care homes was insufficient to provide meaningful estimates of Alpha variant prevalence. Therefore, I used SGTF in my analysis to describe the spread of the novel B.1.17 variant.

I included all SARS-CoV-2 PCR results from nasopharyngeal samples collected from care homes in England between 5th October and 17th December 2020. At that time, there were six national 'Lighthouse' laboratories that processed most community PCR tests, however laboratories used different assays.[281] I included samples that were tested at one of these, the National Biocentre in Milton Keynes, which processed approximately 20% of all Pillar 2 samples.[248] This laboratory along with two other Lighthouse laboratories used an Applied Biosystems 7500 fast RT-PCR system and the Applied Biosystems TaqPath™ 1-Step Multiplex Master Mix (No ROX) (Cat. A28523) and TaqPath COVID-19-ASY-KIT 1000 (Cat. A47817) that targeted ORF1ab, nucleocapsid (N) and spike (S) protein. Samples were categorised as "positive" if at least one of the three target genes (ORF1ab, N-gene, S-gene) was detected by RT-PCR and the internal control was valid. If only ORF1ab and N-genes were detected, this sample was classified as SGTF.

Cycle threshold (Ct) values for the three gene targets and the control were collected and linked to data in the Pillar 2 dataset including sex, age, and geographic location. Ct describes the number of cycles required before viral RNA is detectable and positivity thresholds for target detection are set by these values. Ct is considered to correlate with clinical severity as well as the probability of culturing live virus.[282–284]

For this analysis, individuals younger than 65 years were classed as staff, those who were 65 years or older were considered residents.

4.2.2 Estimating seroprevalence

To estimate seroprevalence within the VIVALDI cohort, results of sequential blood samples collected between 11th June 2020 and 7th May 2021 were included, each participant contributed a maximum of four samples. Blood samples were tested with the Abbott ARCHITECT IgG anti-nucleocapsid semi-quantitative assay (Maidenhead UK) using an index value threshold of 0.8 for positivity. This value was lower than the manufacturer's recommended threshold, however, was selected to increase sensitivity whilst preserving specificity based on published evidence.[285,286] I undertook linkage to bed occupancy, bed number, and staffing from Capacity Tracker using the care home CQC-ID (Chapter 1, Chapter 3). Aggregate data on staff and resident turnover were collected directly from participating care homes. Monthly turnover was calculated using the following equation: Number leaving care home / ((total number at start of month + total number at end of the month) /2) X 100.

Residents older than 65 and staff younger than 65 from care homes taking part in the study were included (recruitment described in Chapter 3). In cases where subject type was not stated, participants 65 years and older were classed as residents and those under 65 were staff.

4.2.3 Measuring cumulative incidence of infection.

To investigate the individual risk of SARS-CoV-2 infection within care homes, I estimated the cumulative incidence of infection over two years. This approach accounted for loss to follow-up if the participant declined subsequent blood tests, left the care home, or died, and for waning of the antibody response over time. I included nucleocapsid antibody and PCR/LFD tests performed in participants who had donated at least one blood sample between 22nd March 2020 and 23rd March 2022. I included samples that could be linked to a pseudo-identifier which enabled linkage between serial tests and to results of symptomatic and asymptomatic PCR and LFD screening conducted under Pillar 1 and 2. Classification of staff and residents was based on the approach described in 4.2.2.

4.3 Statistical analysis

4.3.1 Measuring B.1.1.7 variant spread

I calculated the weekly proportion of samples with SGTF to estimate the change in the distribution of infection with wild-type and B.1.1.7 variants over time. I stratified this by age and by region. I chose the Ct value of the gene target with the highest value to calculate the overall median Ct value according to number of gene amplicons detected (one target only, ORF + N, all three targets). I presented the change in median Ct over time. Although this approach has not been published by other groups, at the time of analysis SGTF was a relatively new approach to classifying Alpha variant, therefore the highest Ct value was used to reduce the probability that absence of S-gene was due to low viral concentration within the sample and not Alpha variant.

Ct may correlate with clinical severity[282–284] and transmission had been demonstrated from individuals with Ct $\leq$ 30 in clinical isolates (and not higher).[287] I performed a sensitivity analysis to increase the probability that isolates with SGTF represented genetic variation and were of the B.1.1.7 variant and not a result of incomplete detection of viral targets from low viral volume in the sample (due to low infectiousness). I therefore repeated my main analysis but classified samples with highest Ct value lower than 30 as positive and the rest as negative.

4.3.2 Estimating seroprevalence

I estimated unweighted and weighted seroprevalence and used cluster-robust estimation of standard errors to account for clustering at the care home level. Weights were applied using the inverse of the proportion of the care home that had been sampled, to estimate the seroprevalence across all residents and staff in participating care homes, accounting for the level of missing data, i.e., those not tested. Weights were calculated using monthly self-reported data in Capacity Tracker on number of staff and number of occupied beds. To maximise the number of samples in the analysis, I included all antibody test results regardless of whether they could be linked to their unique pseudo-identifier (which allows linkage between samples from the same individual). I reported unweighted and weighted seroprevalence in two-month intervals to ensure individuals were included no more than once in each interval in

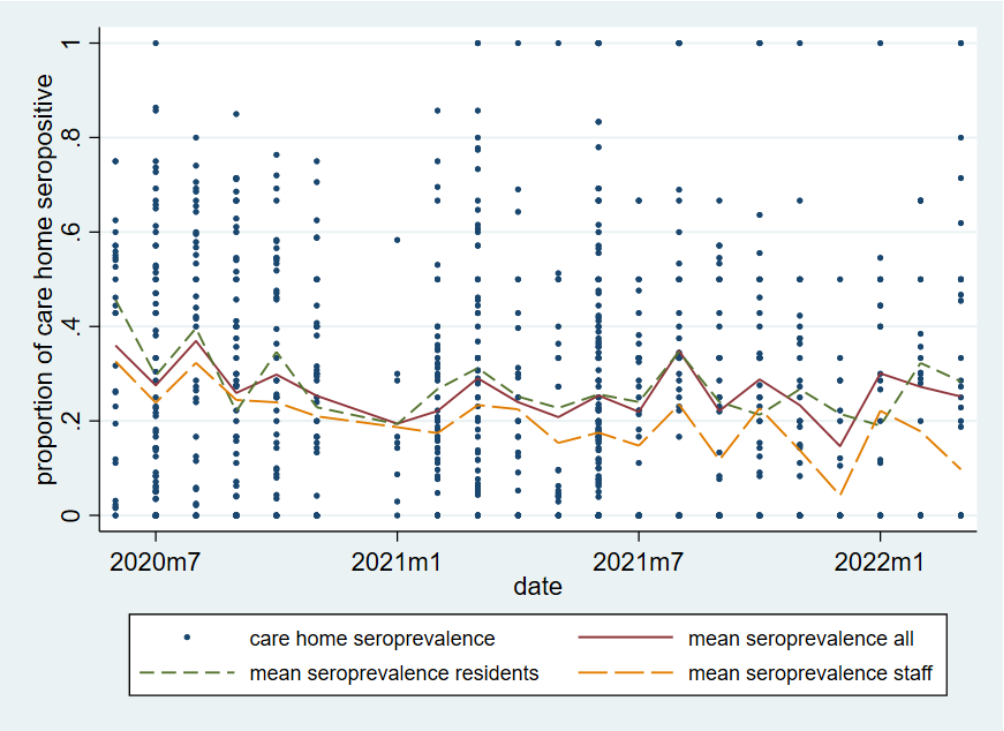
view of the eight-weekly sampling schedule. In cases where one individual had two samples available from the same interval, only the first sample was included.

I estimated an individual-level cumulative incidence of antibody-positivity based on the proportion of individuals who tested positive at least once over the study period. As this was only feasible if serial samples could be linked to the participant, this analysis was limited to the samples where linkage to a pseudo-identifier was possible.

4.3.3 Measuring cumulative incidence of infection.

I extended unweighted seroprevalence estimates performed at the end of the second pandemic wave to include the first two years of the pandemic (22nd March 2020 – 23rd March 2022) up until the end of the first serosurvey and the end of asymptomatic screening.[172] Having plotted the individual and mean monthly care home seroprevalence over the study period, the mean appeared stable, Figure 4.2. Based on published literature and the results from my seroprevalence analysis, I hypothesised that this was likely to be a result of resident turnover from new admissions and death, loss to follow-up, and waning of antibody response over time.[288–290]

Figure 4-2: Mean care home level seroprevalence by month; overall and by staff and resident groups separately (June 2020 – March 2022).



*Seroprevalence determined from detection of anti-nucleocapsid antibody

To estimate the infection risk in care homes at any given point in the pandemic, I estimated cumulative infection incidence from calendar-scale Kaplan-Meier curves and estimated individual-level time at-risk using PCR/LFD and antibody testing dates. This described the hypothetical cumulative risk of infection, providing that participants did not die of non-COVID causes or leave the care home. The infection outcome (fail) was considered to occur on the date of first positive antibody or PCR/LFD test, and I compared cumulative incidence between staff and residents using the log-rank test.

Participants were allowed to enter after the study start date (22nd March 2020). I considered the date of entry to be the earliest of the date of the first blood sample or PCR/LFD test within that care home. Individuals left the at-risk cohort on the date of first positive PCR/LFD test or date of seroconversion, which was estimated to be the mid-point between the date of the first positive antibody test and the preceding negative test. In cases where first positive PCR/LFD test occurred before the seroconversion date, this was taken as the exit date. Individuals were censored on the latest of the date of their final PCR/LFD test within that care home, the date of their final antibody test within that care home (inflated by 60 days to account for two-monthly

blood sampling cycle) and all were censored on the cohort end date (22nd March 2022). Individuals with seropositive baseline test who entered the care home after 1st October 2020 were dropped (n=14) as the national screening programme was not fully operational before this therefore inferring the date of primary infection in these cases was not possible. If date of infection was the same as the entry date, one day was added for inclusion in the survival analysis.

Although the original sample size calculations were based on the precision of seroprevalence estimates (see Chapter 3), the study size has changed to address other research questions, therefore sample size calculations were not performed for these analyses.

The analysis of B.1.1.7 spread was performed in the COVID-19 datastore using Contour plots. The seroprevalence analyses were performed in the UCL Data Safe Haven using STATA 16.0 and the complex survey function for weighted analyses. Cumulative incidence analysis was performed in STATA 17.0. A significance threshold of 0.05 was applied in all analyses.

Ethical approvals are described in Chapter 3.

4.4 Results

4.4.1 Measuring B.1.1.7 variant spread

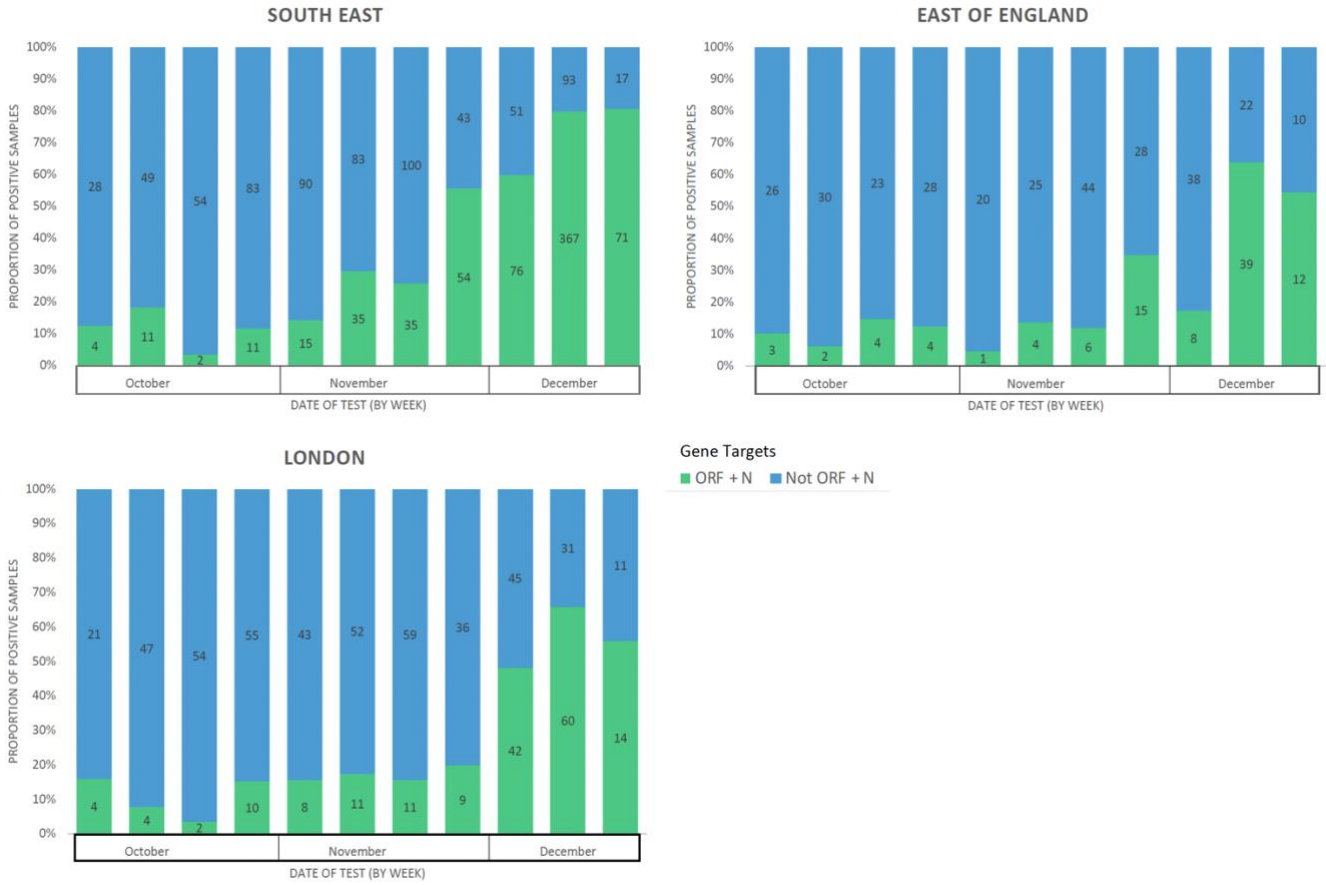
There were 143,994 PCR samples eligible for inclusion by 17th December 2020, of which 4442 (3.1%) were positive (4053 aged < 65 years (91.2%), 389 ≥ 65 years (8.8%)). More than half of the samples were from care homes in the South-East, London, and East of England with the remainder from the North, South-West, East, and West Midlands, likely reflecting the catchment area for the Milton Keynes laboratory. The prevalence of infection rose rapidly over the study period from 132/20240 (0.7%) in the first week of October, to 813/27728 (2.9%) in the second week of December. The proportion of samples with SGTF (ORF + N) remained stable over October and the first half of November, however, began to rise from 12% (13/132) in the last week of November to 60% (491/813) in the week commencing 7th December 2020. When stratified by age, those older than 65 (residents) had a lower prevalence

of SGTF when compared with those younger than 65 (staff), however by the week commencing 7th December 2020 this was 119/157 (75%) and 372/656 (57%) respectively. When stratified by region, SGTF was more prevalent in London, South-East, and East of England, in line with national data from the community, and increased rapidly between mid-November and mid-December, Figure 4.3. This pattern was not seen in the other four regions included.

Figure 4-3: Proportion of SARS-CoV-2 positive samples with SGTF by week in South-East, London, and East of England between 5th October and 17th December 2020.

(Reproduced with permission from [291]*, Copyright Massachusetts Medical Society.)*

'ORF + N' (green bar) denotes samples with SGTF, 'Not ORF + N' (blue bars) denotes samples without SGFT. Number in the bar area shows the number of samples in this category, number on top of bar shows total number of samples reported as positive in that week.



Over the 11-week study period, median highest Ct values for samples with detection of all three genes were static between 21.0 and 23.4, whereas the median highest Ct

values of samples with SGTF dropped from 32.9 (IQR 32.0-34.1) in October to 20.4 (16.9-24.9) in December, Table 4.1.

Table 4-1: Weekly median (IQR) Ct value amongst SARS-CoV-2 positive samples, by number of detectable gene amplicons (5th October to 17th December 2020).*

(Reproduced with permission from [291], Copyright Massachusetts Medical Society.)

Number of gene amplicons detected	5-12 Oct	12-19 Oct	19-26 Oct	26 Oct-2 Nov	2-9 Nov	9-16 Nov	16-23 Nov	23-30 Nov	30 Nov-7 Dec	7-14 Dec	14-17 Dec
One gene	34.2 (33.4-35.0)	33.4 (32.6-35.0)	33.1 (32.6-33.8)	33.4 (32.4-34.2)	33.5 (32.6-34.2)	33.7 (33.2-34.6)	34.5 (33.9-35.4)	32.7 (32.2-34.4)	33.5 (32.3-34.4)	33.1 (30.9-33.3)	NA
2 genes (ORF+N only)	32.9 (32.0-34.1)	32.5 (31.3-33.6)	23.4 (21.8-32.1)	31.3 (20.6-32.8)	26.4 (20.6-31.9)	28.1 (19.4-33.2)	22.5 (18.0-29.5)	22.6 (17.4-29.8)	21.0 (17.3-26.4)	20.4 (16.9-24.9)	20.0 (15.7-23.4)
3 genes (ORF + N + S)	22.4 (18.2-27.9)	23.4 (19.4-27.2)	21.0 (18.6-25.4)	22.2 (19.6-25.8)	22.6 (19.1-26.3)	21.3 (18.5-24.9)	22.7 (19.2-26.9)	22.2 (19.2-27.6)	23.0 (19.6-27.0)	22.5 (19.2-27.4)	22.4 (20.3-25.2)
All samples in analysis	27.3 (20.5-32.9)	26.1 (21.0-32.4)	21.8 (18.9-26.5)	23.2 (19.6-28.8)	23.0 (19.4-27.3)	22.3 (18.5-27.4)	23.7 (19.6-29.9)	22.8 (18.6-28.7)	22.9 (19.0-27.3)	20.9 (17.7-26.3)	20.8 (16.7-24.6)

* Highest Ct value detected out of three targets

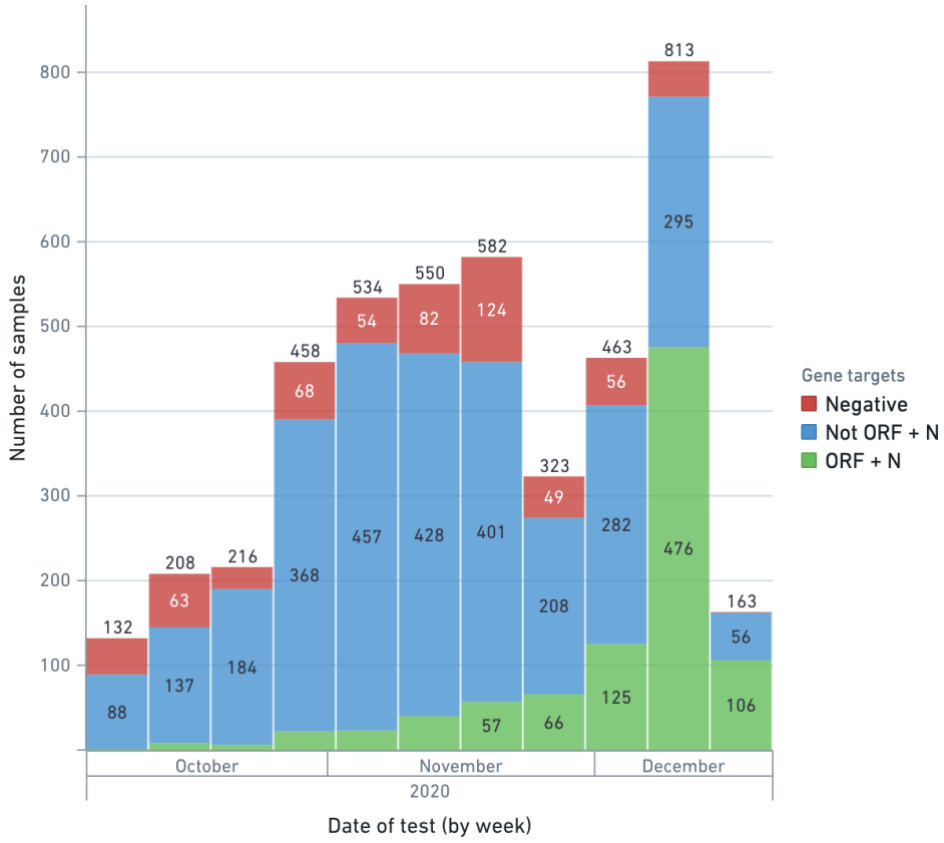
SGTF S-Gene Target Failure ORF Open Reading Frames 1a & 1b gene
N Nucleocapsid gene S Spike gene NA Not Applicable

The sensitivity analysis, where samples with highest Ct ≥ 30 were re-classified as negative, revealed the same pattern as the main analysis, with an increase in the proportion of samples with SGTF over time, Figure 4.4. This supported the hypothesis that samples with SGTF belonged to the B.1.1.7 lineage. Furthermore, sequencing was performed on two isolates from the main cohort, both of which were confirmed as B.1.1.7 lineage.

Figure 4-4: Weekly proportion of samples with SGTF by week amongst all samples reported as SARS-CoV-2 positive. Only samples with highest Ct < 30 are classified as positive, samples with highest Ct value ≥ 30 re-classified as negative (5th October to 17th December 2020).

(Reproduced with permission from [291], Copyright Massachusetts Medical Society.)

'Negative' samples are those reported as positive however highest Ct ≥30. 'ORF + N' denotes S-gene Target Failure (SGTF). 'Not ORF + N' denotes samples with any other combination of detected gene amplicons besides ORF1ab and with highest Ct < 30. Number in the bar area shows the number of samples in this category, number on top of bar shows total number of samples reported as positive in that week.



4.4.2 Estimating seroprevalence

By 7th May 2021, there was a total of 9488 serum samples eligible for inclusion collected from 201 care homes, Figure 4.5. Of these, 8636 (91.0%) could be linked to a pseudo-identifier (2833 from 1434 residents and 5803 from 3288 staff). Median age of staff was 48 years (IQR 35-56) and in residents was 87 years (81-92), Table 4.2. 30.4% of the sampled population were residents (1434/4722) which was lower than the proportion of residents in the overall population of participating care homes

(9059/20299, 44.6%). Care homes took part in up to four rounds of sampling, Table 4.3.

Figure 4-5: Study inclusion flow diagram

(Reproduced from [292], under CC BY-ND licence with permission from Elsevier Ltd)

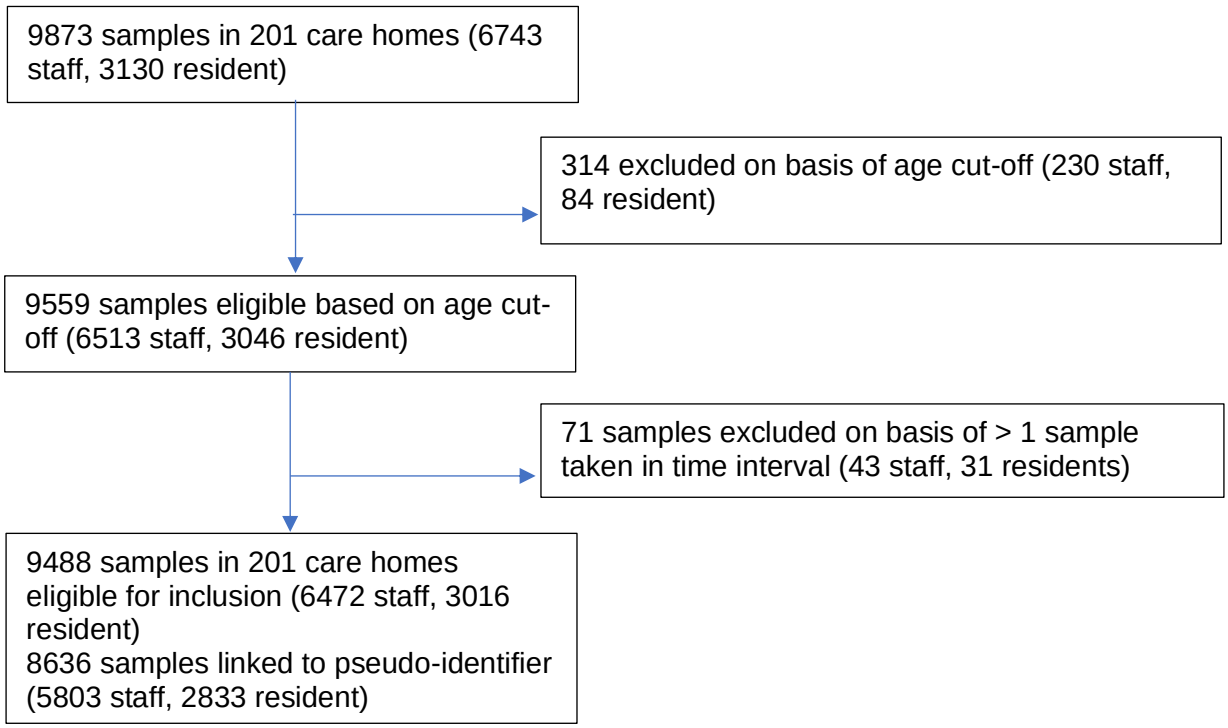


Table 4-2: Demographic details of cohort

(Reproduced from [292], under CC BY-ND licence with permission from Elsevier Ltd)

	Proportion of samples that were antibody-positive* (%)	Proportion of individuals that were antibody-positive at any time* (%)
Total	2516/ 9488 (26·5)	1361 / 4722 (28·5)
Sex		
Female	1895 / 7089 (26·7)	1124 / 3892 (28·9)
Male	424 / 1545 (27·4)	237 / 828 (28·6)
Unknown	197 / 854 (16·7)	0 / 2 (0)
Care home role		
Resident	965 / 3016 (32·0)	484 / 1434 (33·8)
Staff member	1551 / 6472 (16·7)	877 / 3288 (26·7)
Region		
London	291 / 809 (36·0)	155 / 350 (44·3)
South-East	473 / 1683 (28·1)	259 / 861 (30·1)
East of England	95 / 574 (16·6)	47 / 261 (18·0)
East Midlands	164 / 911 (18·0)	96 / 521 (18·4)
West Midlands	124 / 607 (20·4)	73 / 318 (23·0)
South-West	328 / 1616 (20·3)	184 / 873 (21·1)
North-West	297 / 1042 (28·5)	168 / 538 (31·2)
North-East	590 / 1571 (37·6)	290 / 654 (44·3)
Yorkshire & Humber	154 / 675 (22·8)	89 / 346 (25·7)
Care home type		
For-Profit chain	1860 / 6503 (28·6)	966 / 2989 (32·3)
Not-for-Profit chain	563 / 2429 (23·2)	320 / 1326 (24·1)
Independent	93 / 506 (18·4)	75 / 407 (18·4)
Care home size		
Small (<50 beds)	916 / 3843 (23·8)	492 / 1939 (25·4)
Medium (50-100 beds)	1549 / 5491 (28·2)	839 / 2688 (31·2)
Large (≥100 beds)	51 / 154 (33·1)	30 / 95 (31·6)
Interval		
1: June-July 2020	694 / 2225 (31·2)	NA
2: August-September 2020	495 / 1794 (27·6)	NA
3: October-November 2020	360 / 1349 (26·7)	NA
4: December 2020-January 2021	200 / 1136 (17·6)	NA
5: February 2021	221 / 920 (24·0)	NA
6: March-April 2021	546 / 2064 (26·5)	NA

*Based on nucleocapsid antibody detected using Abbott assay

NA Not Applicable

Table 4-3: Characteristics of included care homes

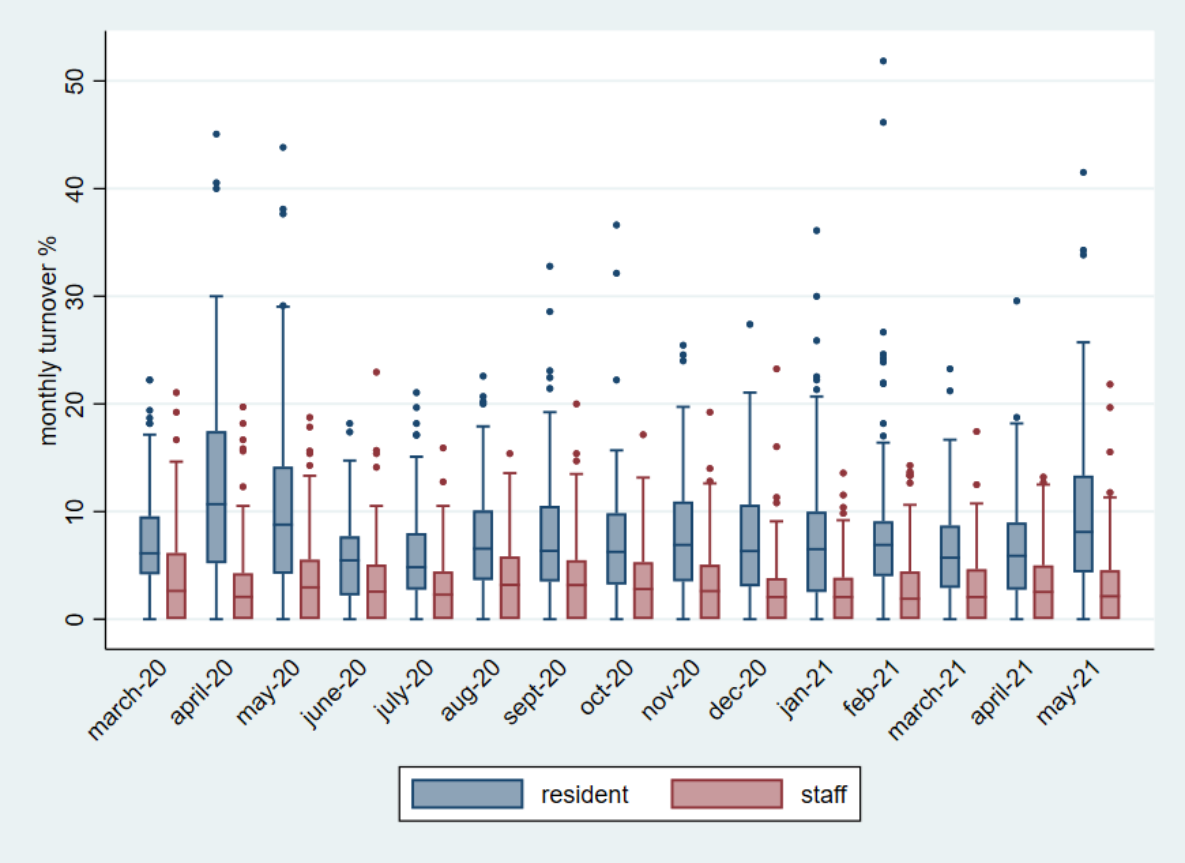
(Reproduced from [292], under CC BY-ND licence with permission from Elsevier Ltd)

	Number of care homes (%)
Total	201
Proportion where residents sampled	176 (87·6)
Proportion where staff sampled	201 (100)
Region	
London	10 (5·0)
South-East	40 (19·9)
East of England	13 (6·5)
East Midlands	26 (12·9)
West Midlands	10 (5·0)
South-West	39 (19·4)
North-West	29 (14·4)
North-East	20 (10·0)
Yorkshire & Humber	14 (7·0)
Care home type	
For-profit chain	118 (58·7)
Not-for-Profit chain	64 (31·8)
Independent	19 (9·5)
Round of testing	
1	201 (100)
2	175 (87·1)
3	84 (41·8)
4	39 (19·4)
Interval	
1: June-July 2020	96 (19·6)
2: August-September 2020	94 (19·2)
3. October-November 2020	87 (17·8)
4. December 2020-January 2021	53 (10·8)
5. February 2021	59 (12·0)
6. March-April 2021	101 (20·6)
Occupied beds per care home, mean (SD)	44·36 (16·5)
Number of staff per care home, mean (SD)	56·86 (21·9)
Number of samples per care home per round, mean (SD):	
Staff	13·18 (8·61)
Residents	7·60 (6·05)
Monthly staff % turnover per care home, median (IQR, range)	2.41 (0-4.80, 0-23.26)
Monthly resident % turnover per care home, median (IQR, range)	6.59 (3.45-10.53, 0-51.86)

Median monthly turnover amongst residents was 6.6% but ranged up to 51.9% (IQR 3.45-10.53%) and was higher than median monthly turnover amongst staff which was 2.41% (IQR 0.00-4.80%), Table 4.3. Resident turnover varied over time and mirrored the pattern of COVID-19 associated mortality in the same care homes, Figures 4.6 & 4.7. For-profit homes had higher median monthly resident and staff turnover than not-for-profit homes (6.72% vs 3.85% in residents and 2.56% vs 1.77% in staff).

Figure 4-6: Box plot showing monthly staff and resident turnover amongst participating care homes (March 2020 to May 2021).

Horizontal line represents median monthly turnover, upper threshold of box shows 75th percentile, lower threshold shows 25th percentile. Lowest and highest whisker show most extreme values within 1.5 interquartile range of the nearer quartile, dots show outlying values.



The proportion of positive tests over the study period was higher in for-profit than not-for-profit care homes (1860/6503, 28.6% vs 563/2429, 23.2%). In addition, smaller care homes (<50 beds) had a lower proportion of positive tests than care homes larger than 100 beds (916/3843, 23.8% vs 51/154, 33.1%), although the latter sample was small. More positive tests and individuals were from London and the North-East and

the fewest were from East of England, East Midlands, and the South-West, which was in line with the national prevalence of infection, Table 4.2.[293]

Cumulative incidence of nucleocapsid antibody-positivity over the study period was 28.2% (95% CI 25.0-31.7) overall; 34.6% (29.6-40.0) in residents and 26.1 % (23.0-29.5, prevalence ratio test *P*<0.0001) in staff. Seroprevalence estimates followed the pattern of pandemic waves, with greater seroprevalence towards the end of the first wave (June – July 2020) with 37.9% (95% CI 30.2-46.2) in residents and 29.8% (24.5-35.7) in staff, troughing during the second wave (December 2020-January 2021) with 21·8% (15·0–30·6) in residents and 16·5% (12·8–20·9) in staff and peaking again after the second wave (March-April 2021) 37·4% (30·5–44·9) in residents and 23·0% (19·1–27·5) in staff, Figure 4.7 & Table 4.4.

Figure 4-7: Weighted seroprevalence with 95% confidence intervals stratified by interval of testing and staff / resident compared against monthly COVID-19 associated deaths in care homes included in study.

(Reproduced from [292], under CC BY-ND licence with permission from Elsevier Ltd)

Seropositivity defined by presence of anti-nucleocapsid antibody. Aggregate COVID-19 associated deaths reported to CQC from care homes included in the study - defined as deaths occurring within 28 days of COVID-19 diagnosis.[294] These are represented as a blue line. Red dashed line represents start of national vaccination programme.

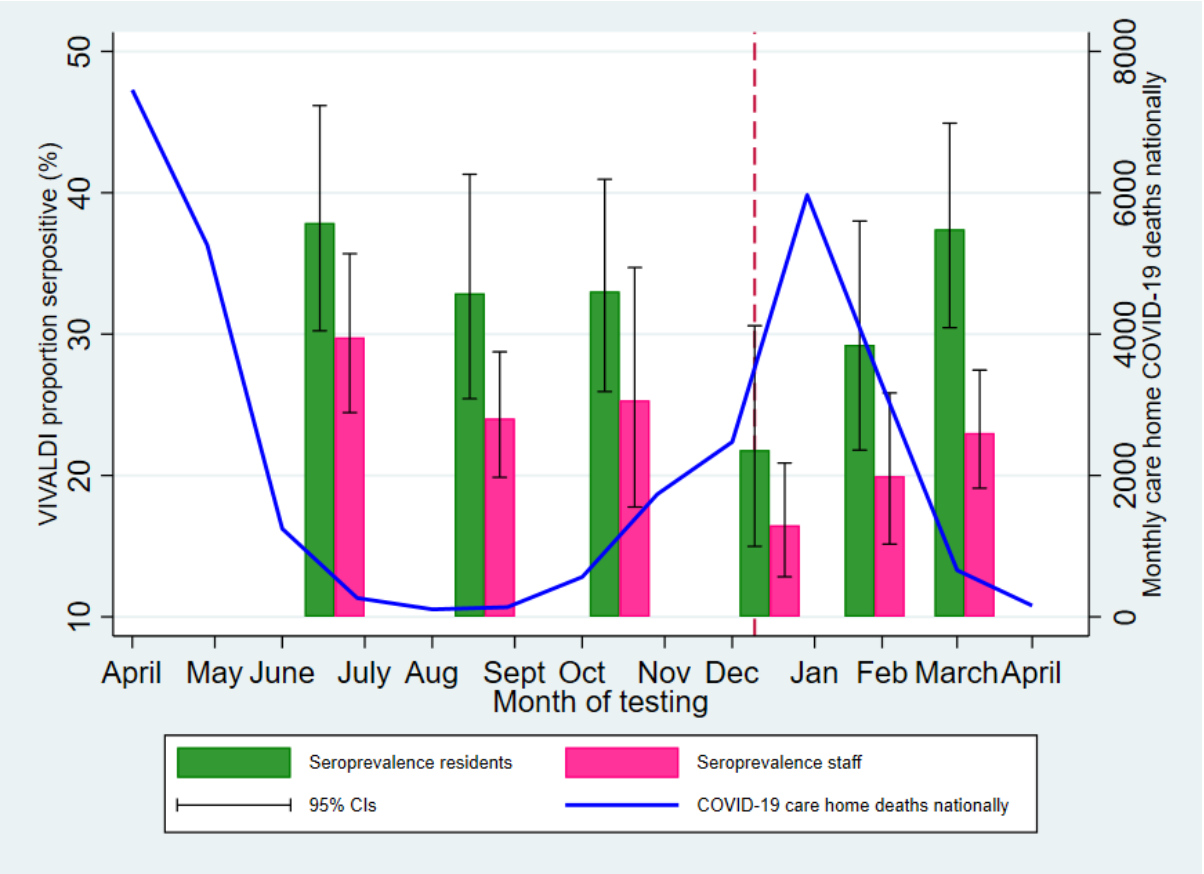


Table 4-4: Seroprevalence estimates with 95% confidence intervals by testing interval, overall and by subject type: a) weighted and b) unweighted.

(Reproduced from [292], under CC BY-ND licence with permission from Elsevier Ltd)
Estimates account for clustering at care home level.

a)

Weighted	**All**		**Residents**		**Staff**	
Number of tests	9488		3016		6472	
Interval	**Proportion positive**	**95% CI**	**Proportion positive**	**95% CI**	**Proportion positive**	**95% CI**
1	30·6	25·4-36·4	37·9	30·2-46·2	29·8	24·5-35·7
2	26·4	22.0-31·5	32·9	25·4-41·3	24·0	19·9-28·8
3	26·5	21·5-32·2	33·0	25·9-41.0	25·3	17·8-34·7
4	17·4	13·6-22·0	21·8	15.0-30·6	16·5	12·8-20·9
5	22.0	17·2-27·6	29·3	21·8-38·0	20.0	15·1-25·8
6	26·1	21·9-30·7	37·4	30·5-44·9	23·0	19·1-27·5

b)

Unweighted	**All**		**Residents**		**Staff**	
Number of tests	9488		3016		6472	
Interval	**Proportion positive**	**95% CI**	**Proportion positive**	**95% CI**	**Proportion positive**	**95% CI**
1	31.2	25·7-37·2	34.0	26.1-43.1	29·9	24·8-35·4
2	27·6	22.8-32.9	32·6	24.8-41·4	25.3	21.1-30.0
3	26·7	22.4-31.5	30.3	22·9-38.9	24.7	20.7-29.3
4	17·6	13·6-22·5	21·8	14.7-31.0	15.8	12·4-20·1
5	24.0	19.0-29.9	30.7	23.6-38·8	21.0	15·8-27.3
6	26·5	22.1-31.3	36.3	30·0-43.3	21.9	17.9-26·5

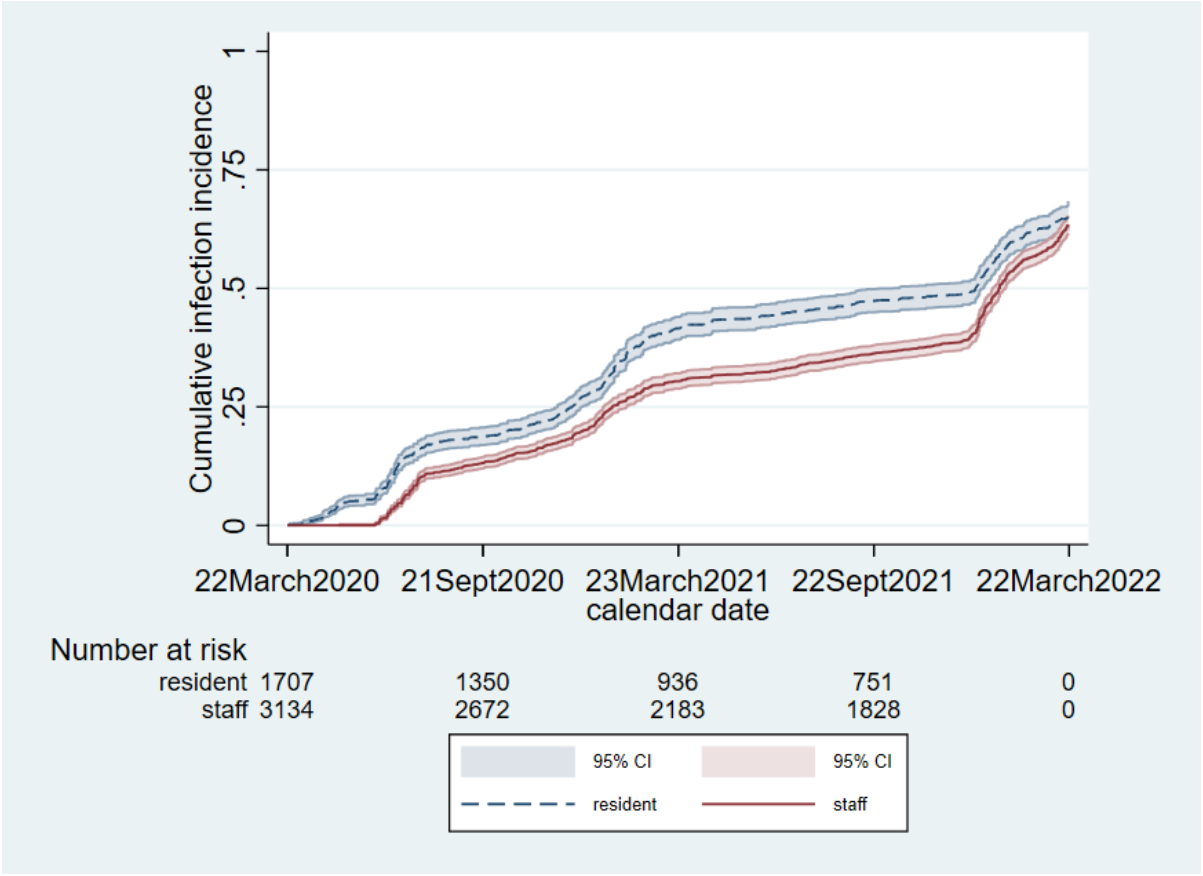
Testing intervals; 1: June-July 2020; 2: August-September 2020; 3: October-November 2020; 4: December 2020-January 2021; 5: February 2021; 6: March-April 2021.

4.4.3 Measuring cumulative incidence of infection.

There were 5179 individuals from 220 care homes who were sampled before 23rd March 2022 who were included in the analysis of cumulative infection incidence analysis, 1794 residents (34.6%) and 3385 staff (65.4%). Participants donated a median of three blood samples each over the period with a minimum of one and maximum of eight samples per participant. The median age at the end of the follow-up period was 87 (IQR 79-92) and 49 (36-57) among residents and staff respectively. Median time-at-risk was 438.5 days (244-691) in residents and 610 days (291-721) in staff. 2780 individuals had ≥1 positive test, 998 (35.9%) residents and 1782 (64.1%)

staff. Overall incidence rate was 0.12 cases per 100 person-days, higher in residents than staff (0.13 vs 0.11, *P*<0.0001). Cumulative incidence over two years amongst all participants was 65%, Figure 4.8.

Figure 4-8: Cumulative incidence of infection with 95% confidence intervals from 22 March 2020 to 22 March 2022, by subject type.



4.5 Discussion

The SARS-CoV-2 B.1.1.7 variant spread rapidly from the community into care homes in England over two to four weeks. This followed the regional patterns of spread[264] as the South-East, London, and the East of England were first to describe widespread transmission of the new variant. This illustrates the rapid ingress of infection from the community into care homes despite strict control measures already in place and supports findings from my scoping review on the critical influence of local infection incidence. My analysis also shows that more than one-third of residents who survived the first wave of infection and one-quarter of staff in care homes had evidence of SARS-CoV-2 infection in the first year of the pandemic based on detection of anti-

nucleocapsid antibodies. Extension of the analysis showed that after two years in the care home, almost two-thirds of the population would have been infected. This estimate suggests that there is currently a large reservoir of individuals in care homes who have survived infection and have some level of natural immunity to SARS-CoV-2. As previously exposed care home residents elicit greater T and B cell responses against SARS-CoV-2 following vaccination[295,296] and are at lower risk of severe outcomes from infection, as demonstrated in this cohort among others,[297,298] this is particularly reassuring. However, If SARS-CoV-2 continues to circulate this finding may have important implications for the future, as new, presumably non-immune residents move into care homes.

This descriptive analysis of B.1.1.7 spread demonstrates the value of near real-time surveillance data to rapidly inform public health policy. At the time, as sequencing did not have universal coverage, SGTF was widely adopted as a marker of B.1.1.7 variant.[272,273,275] The differential drop in Ct values over the study period for samples with SGTF compared with non-SGTF supported this as it coincided with a national rise in B.1.1.7 incidence, suggesting that absence of S-gene was not from low viral volume. Evans et al reported that at a time of increasing transmission or outbreak, on a population-level Ct values are useful markers of higher viral burden as a greater proportion of infected people test earlier in their infection.[299,300] Our analysis was the first globally to describe Alpha variant spread into care homes, suggesting that existing preventive measures were ineffective. As these results were the only source of surveillance from care homes at the time, they were rapidly shared with national decision makers including the Chief Medical Officer (CMO), Chief Scientific Officer (CSO), and the New and Emerging Respiratory Viral Threats Advisory group (NERVTAG). This, in turn, informed the decision to enforce a national lockdown from 6th January 2021.

Estimates of seroprevalence reported in this analysis are greater than those in the general population of England over the same period. Among 70–84-year-old blood donors (who are unlikely to have been shielding from infection at home), 16.6% anti-nucleocapsid antibody-positivity was described between 14th June and 11th July 2021.[301] This is consistent with national data suggesting greater exposure to SARS-CoV-2 within care homes than in the general population.[302] However my results are

lower than those from the largest comparable seroprevalence study, conducted in 9332 residents in 362 care homes in Madrid, Spain between 7th July and 23rd October 2020 and reporting nucleocapsid antibodies in 55.4%.[303] This study also used the Abbott immunoassay and applied the higher manufacturer recommended positivity threshold of 1.4 index. However, as all sites were in one densely populated urban location, this sample may not be representative of the wider care home population.

My analysis demonstrated lower seroprevalence amongst independent care homes when compared with those belonging to for-profit chains and in smaller compared with larger care homes. As I found in my literature review, these factors have been widely associated with increased SARS-CoV-2 transmission risk within care homes.[154,216,228,241] One possible reason for this is that for-profit care homes are usually larger than not-for-profit homes and have lower staff-to-resident ratios.[304] This means they were more reliant on agency staff to cover absence due to sickness. Larger care homes may also have had a greater number of external visitors (even in lockdown) from healthcare professionals, contractors, and staff. It has also been reported that up to a quarter of staff moved into the care homes over the first lockdown to limit the infection ingress risk, largely from smaller independent care homes.[305] Finally there is evidence that staff in for-profit care homes experienced a greater increase in workload and greater staff shortages over the pandemic,[306] which may have impacted on their capacity to prevent infection transmission. However, disentangling the impacts of these multiple and overlapping factors is challenging especially in the context of the changes to staffing levels and bed occupancy over the pandemic. It will be important to learn from these lessons to inform our response to future pandemics.

Nationally, patterns in mortality appear to follow behind infection incidence by a few weeks with a larger gap seen with greater age,[307,308] reflecting the time from infection to development of life-threatening complications. In my analysis, the apparent four-to-six-week lag in seroprevalence peaks following mortality peaks, may reflect the time to IgG antibody development following infection,[277,309] although confidence intervals are wide suggesting some uncertainty in these estimates.

Upon extension of the analysis, I estimated that after two years in the care home, 65% of staff and residents would have been infected (accounting for varying time-at-risk). This is also slightly lower than national period prevalence estimates from the general population.[310,311] This may be because strict infection prevention measures were retained in care homes for longer than in the general population. There is evidence that as the clinical severity of infection declined following vaccination, the public's perceived risk from acquiring COVID-19 and compliance with social distancing measures dropped, therefore risk of infection in the community increased.[312] It is possible that some residents were protected from ever acquiring infection by the preventive measures that were in place or because they took greater steps to self-isolate at home before entering the care homes in light of their clinical vulnerability.

However, it is likely that I have underestimated cumulative infection incidence in the study population as individuals who had died before blood sampling commenced could not be included. In addition, as seronegative individuals were able to enter the cohort later, they may have sero-reverted before joining. In England it is estimated that ~20,000 COVID-19 deaths occurred in care home residents over the first pandemic wave, which was 23.2% of all deaths in this population at that time.[313] The absence of a significant increase in mean seroprevalence estimates over the study period is likely to reflect antibody waning over time (described in more detail in Chapter 5) and high turnover within this population.[135,314] The aggregate data that I collected from the care homes demonstrated significant fluctuations in resident turnover over the first year of the pandemic, likely due to high mortality amongst residents from COVID-19 and large numbers of people who were discharged from acute hospitals for intermediate step-down care to alleviate extreme bed pressures.[315,316] This is further complicated by a sharp decline in new admissions to care homes and bed occupancy over the pandemic.[317] As I could not access individual-level data on care home entry and exit dates or discharge destinations, or pre-admission exposure to infection, it was not possible to account for the potential confounding effect of this in my analysis.

4.5.1 Strengths and limitations

The major strength of this work is the timeliness of results which addressed key questions as these arose informing national and local policy. This was possible due to

research partnerships with a large network of care homes and policymakers, an effective data pipeline, agility in accessing new data streams The seroprevalence analysis sampled a number of diverse and geographically dispersed facilities and was the largest seroprevalence survey from care homes in the UK. It included residents who have been most adversely affected by the pandemic but have traditionally been excluded from research due to logistical challenges around consent and lack of research infrastructure. To my knowledge, the study also benefits from a longer follow-up than any other care home cohort to date, as participants donated up to eight samples over 22 months with linkage to data spanning most of the pandemic. This informed estimates of how the immune reservoir changed over time and will be explored in more detail in the rest of my PhD work.

The seroprevalence study is limited by coverage as an average of 31.8% of residents and 42.1% of residents in the included care homes donated blood. As 9% of samples could not be linked to a pseudo-identifier these were excluded from the estimates of cumulative infection incidence. In addition, 25 (12%) of 201 care homes did not sample any residents. Consenting residents is challenging, particularly in the context of a pandemic, and it is possible that residents without capacity were under-represented as contacting consultees was more labour-intensive for the already overstretched senior care home staff, which may have introduced selection bias. I accounted for unrecruited individuals by weighting estimates using care home size and stratifying estimates to the subject type and time interval. Another possible source of selection bias may have been that the study did not employ a sampling frame to identify and recruit care homes because it was established very rapidly to generate urgently required data. Nonetheless, I found little difference between weighted and unweighted estimates, suggesting this sample was fairly representative.

There is substantial diversity in the care sector, however I attempted to account for any ascertainment bias that this might introduce and increase the generalisability of estimates by including for-profit and not-for-profit homes that were geographically dispersed. The blood sampling schedule was based on availability of care homes, therefore testing rounds may have missed seroprevalence peaks in a particular care home. During outbreaks, visits were often postponed or cancelled as care homes did not allow phlebotomists to visit. This may have reduced seroprevalence estimates over

periods of high community incidence as it was only possible to sample care homes that were not experiencing outbreaks at the time. In addition to the influence of temporal differences in regional infection peaks on risk of infection and outbreaks in care homes, outbreaks also occur by chance. I attempted to account for this sampling bias and confounding by selecting periods for prevalence estimates to include at least one sampling round from as many care homes as possible. To account for attrition bias from loss to follow-up because participants declined subsequent sampling, left the care home, or died, or from antibody waning, I modelled cumulative incidence of seroconversion. This included individuals who donated at least one sample and predicted the rate of seroconversion in seronegative individuals who were lost to follow-up.

Representativeness may also have affected the reliability of my analysis of B.1.1.7 spread. Although I was able to access PCR data in near real-time from one large testing site, this was only one of 42 laboratories that performed testing as part of the Pillar 2 programme at the time. This laboratory processed over 20% of all PCR tests in this programme, however its geographic location meant that 1372/4442 (30.9%) of positive samples came from the South-East. Nevertheless, we responded to the public health emergency and accessed Ct results for genetic PCR targets directly from the largest laboratory. At the time of my analysis, it had not been possible to effectively establish large-scale reliable sequencing surveillance from care homes therefore I used SGTF as a proxy for B.1.1.7 variant. It is very unlikely that these samples were misclassified as lineage was confirmed in two of the included samples with SGTF and this approach had been validated in more than one large cohort.[267,280,318] My findings heralded the emergence of the B.1.1.7 variant in care homes and were critical for policy decisions.

4.5.2 Policy implications

This work demonstrates that a large proportion of the care home population in England have been exposed to SARS-CoV-2 virus over the first two years of the pandemic. This substantial proportion with natural immunity is likely to impact on vaccine effectiveness and protection against re-infection in this vulnerable population. As prior infection and vaccination protects against severe outcomes from re-infections, the

predominance of previously-exposed residents may explain the recent reduction in SARS-CoV-2 mortality in care homes.[319] However as the mean length of stay for a care home resident England can be as low as a year,[127–129] it is possible that new admissions of infection-naïve individuals who had been shielding in their homes, may lead to a rise in mortality with the next pandemic wave. These findings also highlight that despite the strict disease control measures that were in place, it was still not possible to keep SARS-CoV-2 out of care homes as evidenced by the rapid spread of B.1.1.7 and significant seroprevalence that I have described. Although prompt public health action was possible through timely identification of B.1.1.7 spread, this work highlights the importance of continued vigilance and surveillance and the beneficial relationships that can be forged through close collaboration between policymakers and researchers.

In this chapter I used the VIVALDI study to demonstrate substantial variation in the proportion of staff and residents with SARS-CoV-2 infection between care homes. Consistent with findings from my scoping review, this is influenced by region, and funding model, and I additionally described the impact of time and variant. It is clear that infection risk changes substantially with different variants that exhibit properties which may confer survival advantages and higher transmissibility. However, I also identified challenges in using anti-nucleocapsid seroprevalence to estimate care home levels of naturally acquired immunity to SARS-CoV-2. In the context of high community transmission over the study period, I expected to see an increase in seroprevalence over time, which this was not demonstrated in my analyses. Conversely, when I estimated cumulative incidence of infection using both PCR and antibody data (which accounts for waning and loss to follow up), there was a clear temporal increase. Waning of antibody responses may explain this difference in results, and I plan to explore this in more detail in Chapter 5. I will estimate the longevity of naturally acquired and vaccine-induced antibodies in care home staff and residents. Host factors relating to immunity play an important role in preventing infection and severe consequences in care homes. Understanding the longevity of these responses will inform future models on the extent to which immunity protects care homes against infections and outbreaks.

4.6 Contribution, Dissemination & Impact

4.6.1 Measuring B.1.1.7 variant spread

I designed the study in collaboration with the study Chief Investigator, I designed the analysis plan and analysed the data. I co-wrote the manuscript with the Chief Investigator.

The B.1.1.7 spread analysis has been published as a research letter in NEJM:
Krutikov M, Hayward A, Shallcross L. Spread of a Variant SARS-CoV-2 in Long-Term Care Facilities in England. *New England Journal of Medicine* 2021; **384**: 1671–3. DOI: 10.1056/NEJMc2035906

It was also presented as a report to the Scientific Advisory Group for Emergencies (SAGE) who advise the UK Government. This informed decisions to instigate a national lockdown which was implemented on 6th January 2021.

This received national and international media coverage including articles in Mirror, Mail Online, Mail Online (2), Independent, Yahoo! News, Wales Online, Hindustan Times, Business Standard (India), UCL News

4.6.2 Estimating seroprevalence

I designed the study and the statistical analysis plan, project managed the establishment of the cohort and coordinated blood sampling, carried out data analysis, and drafted the manuscript for publication.

The seroprevalence analysis has been published as a full article in Lancet Health Longevity:

Krutikov M, Palmer T, Tut G, *et al.* Prevalence and duration of detectable SARS-CoV-2 nucleocapsid antibodies in staff and residents of long-term care facilities over the first year of the pandemic (VIVALDI study): prospective cohort study in England. *Lancet Healthy Longev* 2022; **3**: e13–21. DOI: 10.1016/S2666-7568(21)00282-8

With commentary piece here:

Verschoor CP, Bowdish DME. Estimating SARS-CoV-2 seroprevalence in long-term care: a window of opportunity. *Lancet Healthy Longev* 2022; **3**: e2-3. DOI: 10.1016/S2666-7568(21)00304-4

This received national media coverage in the Daily Telegraph.

4.6.3 Measuring cumulative incidence of infection.

I designed the study and statistical analysis plan, led project management of sample collection, carried out data analysis and drafted the abstract. This work was presented as a short oral presentation in the e-poster sessions at the international ECCMID conference in April 2023 (see Supplementary Appendix for poster).

Chapter 5

Objective 3: Testing the hypothesis that care home staff and residents develop durable SARS-CoV-2 antibody responses following infection and vaccination and these responses can be measured on a facility-level.

In Chapter 4, I showed that infection prevalence can be estimated from surveillance PCR testing which is particularly valuable when considering variation in infection burden according to agent factors such as variant. Although anti-nucleocapsid seroprevalence surveys can help to identify individuals who were infected over the early pandemic when PCR testing coverage was incomplete, absence of a temporal increase suggests that these estimates are at least partially affected by antibody waning. The dynamics of antibody responses to SARS-CoV-2 amongst care home residents and staff may inform estimates of the levels of immune protection against infection in this population and improve precisions of models considering infection risk and describing prevalence.

In my third PhD objective I explored the longevity and magnitude of antibody responses to infection and vaccination in the care home population, a key host factor. I addressed this objective over two sub-studies, hosted within the VIVALDI study. The first estimated cumulative incidence of semi-quantitative nucleocapsid antibody sero-reversion from primary infection date (using linked data on PCR testing and hospitalisations) and compared this between staff and residents. The second modelled magnitude and rate of loss of quantitative spike antibody responses following full vaccination against SARS-CoV-2 (two vaccine doses) and predicted variation according to key factors such as age, sex, prior infection, and vaccine type.

Nucleocapsid antibodies remained detectable for approximately eight months following infection in around half of participants, but sero-reversion incidence was greater amongst staff than residents. This difference was probably related to greater peak antibody titres amongst residents than staff. In contrast, spike antibody responses appeared comparable between the two groups, although there was some

evidence of greater antibody peaks with faster rate of decline amongst Pfizer-BioNTech vaccine recipients. Amongst Oxford-AstraZeneca recipients, lower antibody peak was demonstrated in older compared with younger participants and more rapid decline amongst males compared with females. Prior infection consistently predicted greater magnitude and slower decline of antibody responses.

In the context of high levels of infection amongst the care population demonstrated in Chapter 4, the boosting effect of prior infection on antibody responses to vaccination suggests a high level of immunity amongst this highly vaccinated population. However rapid sero-reversion that I have demonstrated suggests that seroprevalence surveys must be applied with caution when estimating levels of naturally acquired immunity. This will have important policy implications when considering prioritisation of care homes for re-vaccination strategies and other NPIs. In addition to the agent and host factors that I have considered so far, environmental factors are also likely to play an important role in the risk of infection and outbreaks, which I plan to explore in Chapter 6.

5.1 Background

Over one-third of residents and one-quarter of staff acquired infection-induced antibodies to SARS-CoV-2 over the first fifteen months of the pandemic, rising to two-thirds after two years in the care home, Chapter 4. Seroprevalence may be a useful measure of population-level immunity to infection, however the duration of seropositivity had not been described in this population as most studies in care homes are cross-sectional. The longest longitudinal serological study, published at the time of this analysis, followed up participants for four and seven months.[127320,321] Studies from the general population suggest anti-nucleocapsid antibodies decay rapidly following infection and approximately half remain positive for up to 8-10 months following infection.[322–324] It is not clear how this decline correlates with immunity against infection and severe outcomes against infection, however anti-nucleocapsid antibodies have been widely used as a measure of prior infection and immunity.

Care home residents exhibit important differences in their immune responses when compared with their community-dwelling peers or to younger populations as a result

of immune-senescence, nutritional status, and presence of medical co-morbidities.[132,139] This means that in some cases, immune responses to infection and vaccination are less effective, described in Chapter 1, emphasising the need for studies investigating the immunological response to infection and vaccination in care home residents

As previously described, vaccinations target the SARS-CoV-2 spike protein and therefore elicit production of spike antibodies, whereas infection-induced immunity can be inferred based on presence of non-spike antibodies like anti-nucleocapsid.[278] Evidence from large studies in the general population suggest that vaccination induces durable humoral immune responses for at least six months, with some evidence of waning antibody levels over time, which may be greater with age.[325,326] There is also evidence that larger humoral and cellular responses to vaccination, which may provide greater protection against infection, are generated in those who have experienced prior SARS-CoV-2 infection and have some naturally-acquired immunity.[327] This has also been demonstrated amongst care home residents.[295,328,329]

In addition to monitoring of vaccination responses, seroprevalence surveys are commonly used to measure infection prevalence in a population and are particularly useful for detecting undiagnosed individuals who experienced asymptomatic infections or where diagnostic testing was not universally available (such as in SARS-CoV-2 - described in Chapters 1 & 4).[330] Determining whether nucleocapsid antibodies remain positive over time can inform the application of this approach to SARS-CoV-2.[331] My third objective, which I address in this chapter, is therefore to assess the duration and magnitude of antibody responses following SARS-CoV-2 infection and/or vaccination in care home residents. This analysis could inform population-based estimates of immunity from natural infection, and policy decisions on the need for re-vaccination.

I have addressed Objective 3 through two sub-studies: (Figure 5.1):

i) I modelled the cumulative incidence of loss of antibodies (sero-reversion) in individuals who had been infected with SARS-CoV-2, between March 2020 (data collection start) and May 2021 (when all care homes had been sampled at least once) and compared this between staff and residents, as

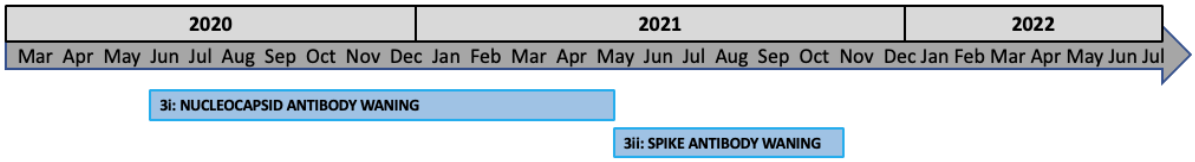
a measure of longevity of naturally-acquired immune responses to infection that may provide some protection from severe infection.

ii) I modelled dynamics of antibody responses over 12 months following two-dose SARS-CoV-2 vaccination amongst staff and residents using data collected between March 2020 and October 2021 (when third booster doses were administered) and predicted how these varied with significant confounders like age, sex, vaccination type, and prior exposure to infection.

Immunity is an important host factor associated with infection risk (described in Chapter 1). Although other host factors include medical co-morbidities, ethnicity, nutritional status, genetic predisposition, and PPE use, I could not access reliable data on these. The impact of demographic factors on infection risk have been extensively described in the literature therefore I chose not to explore these in my PhD thesis, however I have considered their influence on antibody responses.

Although I designed, conducted, and drafted the first analysis (i), I had assistance from Dr Oliver Stirrup, statistician on the VIVALDI study for the second analysis in view of complexity of the modelling approach (ii). For the second analysis (ii), I designed the study objectives, assisted with the analysis plan, and drafted the manuscript, and Dr Stirrup built the models and prepared the tables and figures.

Figure 5-1: Objective 3 Timeline



5.2 Methods

5.2.1 Naturally acquired immunity.

Eligibility for inclusion is outlined under the seroprevalence analysis for Objective 2 in Chapter 4. Consenting individuals from participating care homes had successive rounds of blood sampling approximately every eight weeks as part of the VIVALDI study between 11th June 2020 (when blood sampling began) and 7th May 2021. Each

participant donated a maximum of four samples over the study period, as described in 4.2.2.

Serum samples were tested for anti-nucleocapsid IgG antibodies using the Abbott ARCHITECT immunoassay and I applied a 0.8 index value threshold for positivity to preserve sensitivity and specificity, (see 4.2.2). A subset of samples was tested for antibodies against proteins within the S1 subunit of the virus, the spike, and receptor binding domain (RBD) proteins. These were tested using the Meso Scale Discovery (MSD) V-PLEX COVID-19 Respiratory Panel 2 kit (Rockville MD, USA) by collaborators at the University of Birmingham (UoB). Cut-off for spike-antibody positivity of 1200 arbitrary units (AU) /ml was applied as testing of pre-pandemic samples by the Birmingham laboratory found a 96% specificity using this threshold (48/50). RBD cut-off of 600 AU/ml was applied in line with previously published data from the same group.[332]

To identify dates of primary infection from 1st March 2020 onwards, I linked antibody results to PCR test results from the national care home testing programme (Pillar 2) and outbreak investigations (Pillar 1) as well as hospitalisation records using the individual's unique pseudo-identifier. As LFD tests were only introduced in December 2020 and all the primary infections included in my analysis occurred before then, I did not include LFD results. Age, sex, subject type, and CQC-ID were taken from data recorded on blood samples by the care home managers, however discrepancies in subject type were managed as outlined in 4.2.2. Detailed data linkage and processing methods are described in Chapters 3 and 4.

I included participants if they could be linked to a pseudo-identifier (allowing linkage to successive samples from the same individual), had donated at least one sample with detectable anti-nucleocapsid antibody and at least one subsequent sample, and it was possible to estimate the primary infection date. I applied these following criteria in order of preference to estimate the date of seroconversion:

a) a positive PCR test before the antibody test date (seroconversion was 14 days after this date),

b) any hospital admissions with diagnostic codes of confirmed or suspected COVID-19 (seroconversion was the admission date),
c) if a positive antibody test followed a negative test (the mid-point between tests was considered the seroconversion date),
d) first wave infection if positive antibody test was taken before 1st August 2020 and a-c criteria could not be applied (seroconversion date was 14 days after the first wave peak of cases nationally - 5th May 2020).[301]

I considered sero-reversion (antibody loss) to be the mid-point between a positive and subsequent negative antibody test in a participant. I excluded individuals with an estimated seroconversion date more than 120 days before the date of their first antibody test as I considered it likely that sero-reversion could have occurred over this period therefore participants with a negative first antibody test may have actually already sero-reverted.

To assess whether nucleocapsid antibody-loss was also associated with loss of other infection-associated antibodies, blood samples from a convenience sample of individuals who had sero-reverted were analysed using the MSD assay for antibodies against spike and RBD within the S1 subunit of the virus. All available samples from each individual were tested, however selection of individuals was led by the team at UoB in line with sample availability. In this cohort the latest sample date was 12th November 2020, which preceded the administration of the first SARS-CoV-2 vaccine therefore antibody responses reported in this analysis were only naturally acquired.

5.2.2 Vaccine-induced immunity

For the analysis of vaccine-induced immunity, the analysis period was extended to include samples taken after the second vaccination dose, administered from March 2021 in most care homes. In line with national guidance, two doses of vaccination constituted a full course.[110] Staff and residents taking part in the VIVALDI study who donated blood samples between 11th June 2020 (start of blood sampling) and 21st October 2021 (when most had received third booster vaccine dose) and had received both vaccines in this period were included. Although additional vaccines were subsequently approved, the initial vaccine rollout in the care home population used

the Oxford-AstraZeneca or Pfizer-BioNTech monovalent vaccine (see Chapter 1) therefore all participants had received one of these.

All serum samples in this analysis were tested for anti-nucleocapsid IgG antibodies using the Abbott ARCHITECT immunoassay and quantitative anti-spike antibody titres were measured using the MSD assay (see 5.2.1 for positivity thresholds).[285] All samples from 11th March 2021 onwards were tested using the MSD assay along with a convenience sample of samples received before this date. Only samples that had been tested using the MSD assay were included in this analysis.

As before, detailed data linkage and data processing methods are described in Chapter 3. Using the pseudo-identifier, antibody results were linked to pseudonymised records from the national PCR/LFD testing datasets dating back to 1st March 2020 as outlined in section 5.2.1 and Chapters 3 and 4, as well as vaccination records which include date, dose number and manufacturer of vaccine.[245] Individuals were included from 21 days following their second vaccine dose (as this correlates with peak antibody response) [325] up until the third vaccine date, if applicable. Samples taken following third vaccine were excluded. As LFDs were introduced in December 2020 for staff, positive LFD tests were included in the analysis as evidence of prior infection. Where linkage to national datasets was not possible, I obtained individual vaccination records directly from the care homes.

Participants who had evidence of new infection after their second vaccine dose (based on a positive PCR or LFD test or a newly positive anti-nucleocapsid antibody) were excluded from the analysis as their immune responses may have been boosted by these breakthrough infections.

5.3 Statistical analysis

5.3.1 Naturally acquired immunity.

To estimate the cumulative incidence of nucleocapsid-waning over the study period, I fitted Kaplan-Meier curves to the whole study cohort and separately to staff and resident groups, which I compared using the log-rank test. Sero-reversion was the outcome of interest and time-at-risk was estimated in days from seroconversion

(based on the a-d criteria described in methods). Individuals who did not reach the outcome (sero-reversion) were censored at the date of their final serum sample.

As participants in the group d category had a less reliable estimate of seroconversion date which was based on the peak of first wave cases, I carried out a sensitivity analysis to assess reliability of my results. I repeated the survival analysis and compared subject type however, I only included participants in whom seroconversion dates had been estimated according to the a-c criteria.

In view of evidence that stronger immune responses are generated in response to more severe infections,[333,334] I assessed whether this impacted on sero-reversion rates. I performed a sub-group analysis that compared sero-reversion rate between severe and non-severe infection (using hospitalisation with COVID-19 as a proxy for severity). As before, I fitted Kaplan-Meier curves to the group that had been hospitalised with COVID-19 over the study period and those that had not and compared the groups.

To investigate whether differences in sero-reversion rates were related to the magnitude of antibody response, I compared antibody titres between staff and residents. Following the initial positive sample (allocated to round 1) successive samples from an individual participant were allocated to a round of testing according to the order that they were taken (1 for the first round, 2 to second etc). I calculated the distribution of the nucleocapsid antibody titres in each round of testing according to staff / resident group and plotted these. I estimated the inter-round difference in antibody titres according to subject type group and compared means for each round between groups using t-tests.

To compare nucleocapsid antibody loss to other spike antibodies, quantitative antibody titres against spike and RBD in a subset of sero-reverted participants were compared over time. These were grouped by time from sero-reversion (using Abbott nucleocapsid antibody test) into baseline (date of sero-reversion), 0-30 days, and 60-90 days and were plotted in these groups according to staff and resident status.

5.3.2 Vaccine-induced immunity

Participants were classed into those with and without evidence of past SARS-CoV-2 infection. Presence of anti-nucleocapsid antibodies (using either Abbott or MSD assay) was classed as evidence of past infection. Any individuals with a positive PCR or LFD test or hospitalisation with COVID-19 before their second vaccine dose, were also defined as having had past infection.

As MSD spike antibody levels reached 10^6 AU/ml, logarithmic values were used for analysis. None of the post-vaccine samples were classed as negative, therefore we modelled the peak antibody levels and the rate of decline in these levels over time. Log10-transformed anti-spike antibody levels based on MSD testing were modelled using linear mixed effects models. These had random intercept and slope terms for each participant and time was centred at 21 days from second vaccination (based on data showing this is when vaccine-induced antibody responses peak).[325] Independent effects were assumed for the predictors that were considered to influence the magnitude and rate of decline in antibody level; these included vaccine type, sex, past SARS-CoV-2 infection, staff / resident status. Another model with interaction terms between vaccine type and the other predictors was fitted and a further model by vaccine type with age centred at 70 years (selected to represent the resident population) as a linear variable. Model fit was compared using Likelihood Ratio Test (LRT).

A *P*-value of 0.05 was taken as the cut-off for significance in all analyses.

Formal sample size for these analyses were not calculated. All available samples were included in this analysis, recognising the need to generate findings quickly.

All analyses were performed in STATA 16.0. Graphs in the vaccine-induced immunity analysis were prepared using *ggplot* package in R.

Ethical approvals and legal basis of accessing data for the VIVALDI study are described in Chapter 3.

5.4 Results

5.4.2 Naturally acquired immunity.

As described in Chapter 4, 8636 individuals had at least one blood sample that could be linked to a pseudo-identifier. Of these, 619 participants were eligible for inclusion in the time to sero-reversion analysis (239 residents, 380 staff). There were 503 females (81.3%) and 116 males (18.7%) with a median age of 59 years (IQR: 45-82) overall, 87 (79-91) in residents and 48 (38-57) in staff. Median follow-up in the main analysis was up of 149 days (IQR 107-169) overall, similar between residents and staff (154 days vs 147 days respectively). Individuals belonged to 93 care homes (70 for-profit, 4 independent, 19 not-for-profit), across all nine regions of England – most from South-East (19, 20.4%) and least from East of England (4, 4.3%).

Table 5-1: Number of samples included in the analysis by care home role and mean time between samples in days.

(Reproduced from [292], under CC BY-ND licence with permission from Elsevier Ltd)

Testing round	**Mean time to next sample in days (SD)±**	**No. Residents with N-antibodies (%)**	**No. Staff with N-antibodies (%)**
1	62·5 (26·9)	239/239 (100)	377/380 (99·2)
2	61·8 (24·0)	211/239 (88·3)	303/380 (79·7)
3	157·2 (24·9)	119/154 (77·3)	140/218 (64·2)
4	NA	23/37 (62·2)	8/18 (44·4)

NA = Not Applicable

± If individuals had donated a further blood sample.

Almost two thirds of participants had at least three rounds of blood sampling (Table 5.1) and the mean number of samples per person was 2.5 (SD 1.1). The mean time between samples was similar between the first three rounds (62.5 and 61.8 days) but was longer between rounds three and four (157.2 days) reflecting the period taken to consent participants for a further 3-5 rounds of blood samples.

377 staff tested positive at round one and 137 (63.7%) of these individuals remained positive in their third testing round approximately four months later. In residents, 239 tested positive at their first testing round and 119 (77.3%) remained positive by round three of testing. There were three cases of sero-conversion during follow-up, and all occurred in staff, these only entered the survival analysis from the estimated date of

seroconversion however all rounds of testing from these individuals are described, Table 5.2.

Table 5-2: Antibody results by round of testing and round 1 antibody result in a) staff b) residents.

(Reproduced from [292], under CC BY-ND licence with permission from Elsevier Ltd)

a)

		Round 2			Round 3			Round 4		
		+ve (%)	**-ve (%)**	**Total**	**+ve (%)**	**-ve (%)**	**Total**	**+ve (%)**	**-ve (%)**	**Total**
Round 1	**+ve (%)**	300 (79·6)	77 (20·4)	377	137 (63·7)	78 (36·3)	215	7 (41·2)	10 (58·8)	17
	-ve (%)	3 (100)	0 (0)	3	3 (100)	0 (0)	3	1 (100)	0 (0)	1
Total		303	77	380	140	78	218	8	10	18

b)

		Round 2			Round 3			Round 4		
		+ve (%)	**-ve (%)**	**Total**	**+ve (%)**	**-ve (%)**	**Total**	**+ve (%)**	**-ve (%)**	**Total**
Round 1	**+ve (%)**	211 (88·3)	28 (11·7)	239	119 (77·3)	35 (22·7)	154	23 (62·2)	14 (37·8)	37
	-ve (%)	0	0	0	0	0	0	0	0	0
Total		211	28	239	119	35	154	23	14	37

Sero-reversion occurred in 55 (23%) of residents and 133 (35%) of staff; with most sero-reversions occurring between 90 and 180 days from estimated seroconversion (129/188, 69%), Table 5.3.

Table 5-3: Distribution of time to sero-reversion for residents and staff.

(Reproduced from [292], under CC BY-ND licence with permission from Elsevier Ltd)

Time to sero-reversion	**Residents (%)**	**Staff (%)**	**Total (%)**
< 90 days	13 (23·6)	23 (17·3)	36 (19·2)
90-180 days	30 (54·6)	99 (74·4)	129 (68·6)
180-270 days	11 (20·0)	8 (6·0)	19 (10·1)
≥ 270 days	1 (1·8)	3 (2·3)	4 (2·1)
Total	55	133	188

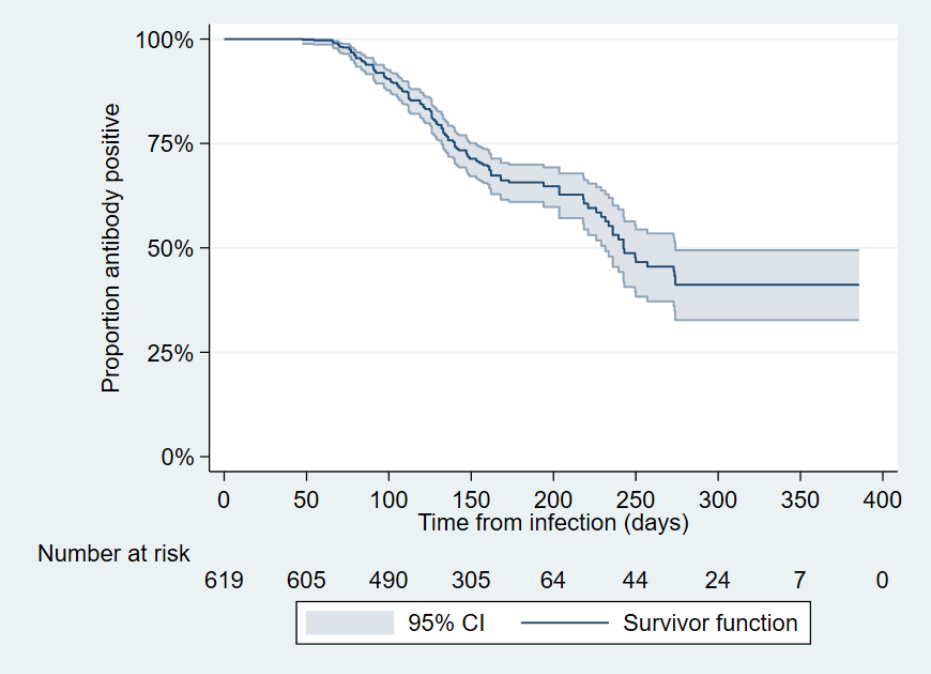
Among staff and residents that sero-reverted during study period.

Seroconversion date was defined using PCR in 108 participants (a), hospitalisation with confirmed COVID-19 in 5 participants (b), conversion from negative to positive antibody test over subsequent testing rounds in 3 (c) and assumed to have occurred in first pandemic wave in 503 (d). The median time between date of assumed seroconversion and first positive antibody test was 57.5 days (IQR: 47.5-70).

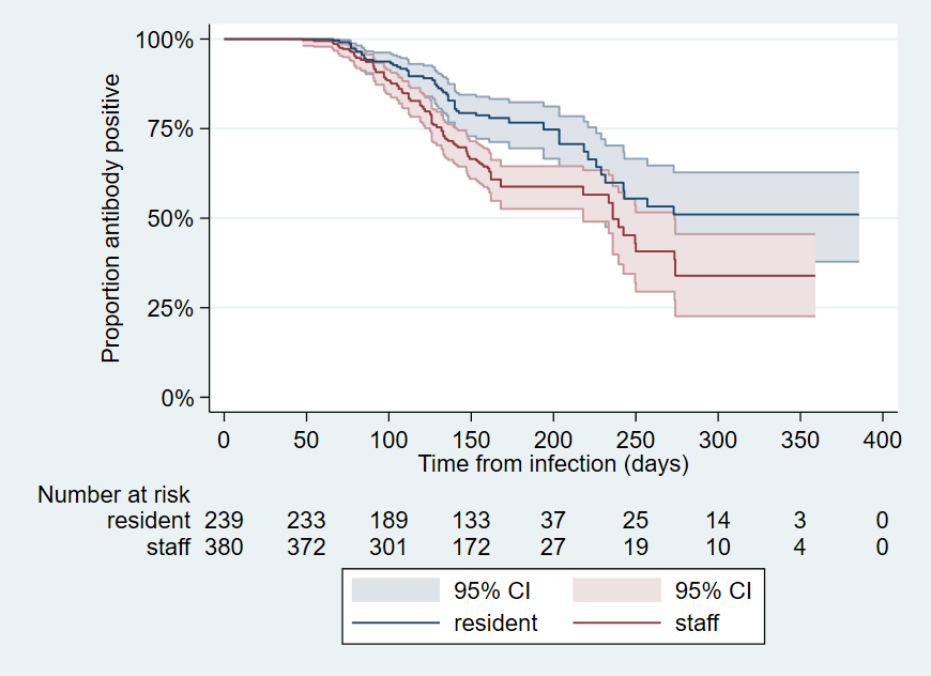
The median time to antibody loss (Kaplan-Meier estimate) was 242.5 days. Time at-risk was 91684 person-days, 54543 in staff and 37141 in residents. The overall incidence rate of sero-reversion was 2.1 / 1000 person-days at-risk, greater in staff than residents (2.4 vs 1.5 / 1000 person-days at-risk, log-rank *P*=0.00034), Figure 5.2.

Figure 5-2: Kaplan-Meier plot with risk table showing time to sero-reversion with 95% confidence intervals a) in whole cohort; b) in staff compared with residents.

(Reproduced from [292], under CC BY-ND licence with permission from Elsevier Ltd)
a)



b)

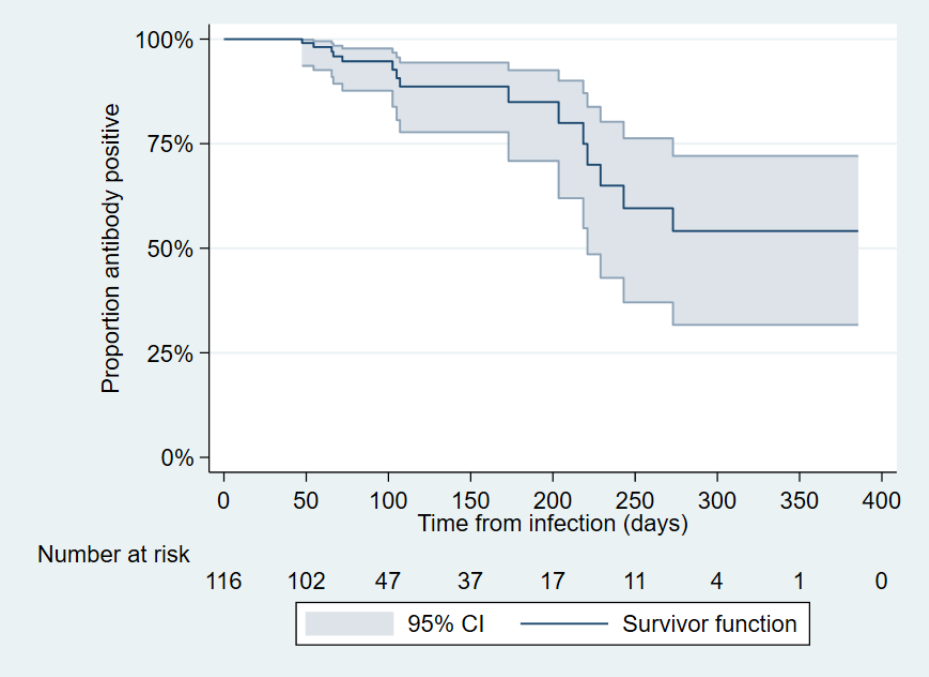


In the sensitivity analysis that only considered participants from groups a-c (n=116), sero-reversion occurred in 10/73 residents and 5/43 staff. However median follow-up was shorter than in the main cohort (89.5 days, IQR 66.3-159.5) so follow-up may have been insufficient for antibody-loss to occur. The significant difference between cumulative incidence of sero-reversion in staff and residents in the main analysis was retained in the sensitivity analysis (1.5 vs 0.9 / 1000 person-days at-risk, *P*=0.0096), Figure 5.3.

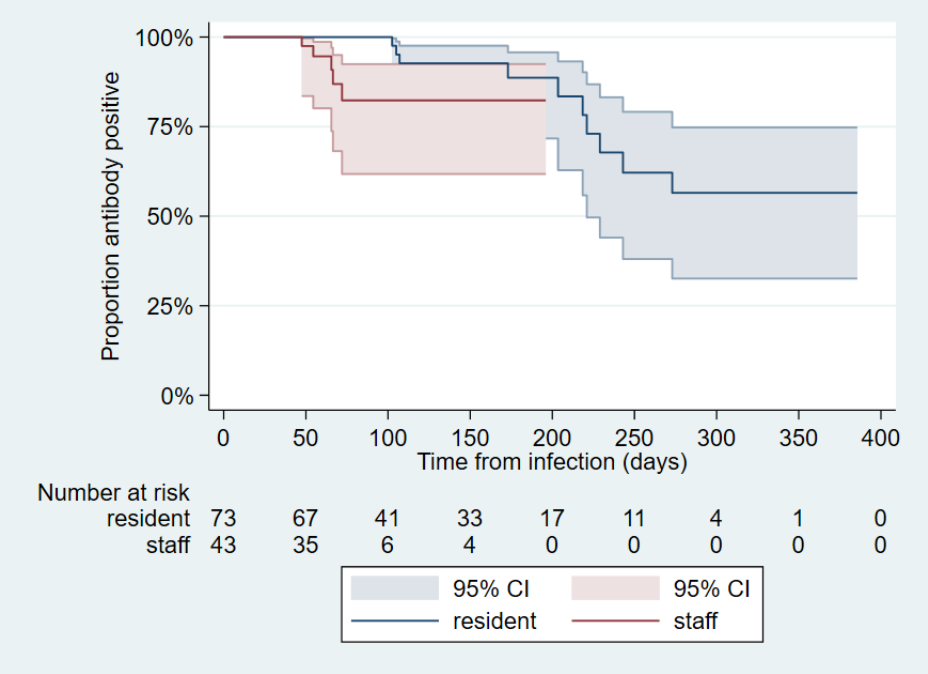
Figure 5-3: Kaplan-Meier plot with risk table showing time to sero-reversion with 95% confidence intervals in groups A-C only a) overall and b) in staff compared with residents.

(Reproduced from [292], under CC BY-ND licence with permission from Elsevier Ltd)

a)



b)



In the sub-group analysis comparing hospitalised (n=20) and non-hospitalised (n=599) participants as a proxy for infection severity, cumulative incidence of sero-reversion in

hospitalised was 1.1 / 1000 person-days at-risk compared with 2.1 / 1000 person-days at-risk in the non-hospitalised cohort. This difference did not reach statistical significance (*P*=0.15) although the sample was probably too small to detect a difference.

To investigate the difference between cumulative incidence of sero-reversion among staff and residents further, I analysed magnitude of antibody responses. Nucleocapsid antibody titres were greater in residents than staff over all four rounds. The fourth-round median antibody titre amongst staff but not residents was below the positivity threshold, Figure 5.4, Table 5.4. In rounds 1-3, differences in the mean titres were statistically significant however significance was lost for round four. However, inter-round differences are very similar between staff and residents suggesting that the rate of antibody decay was comparable, Table 5.4.

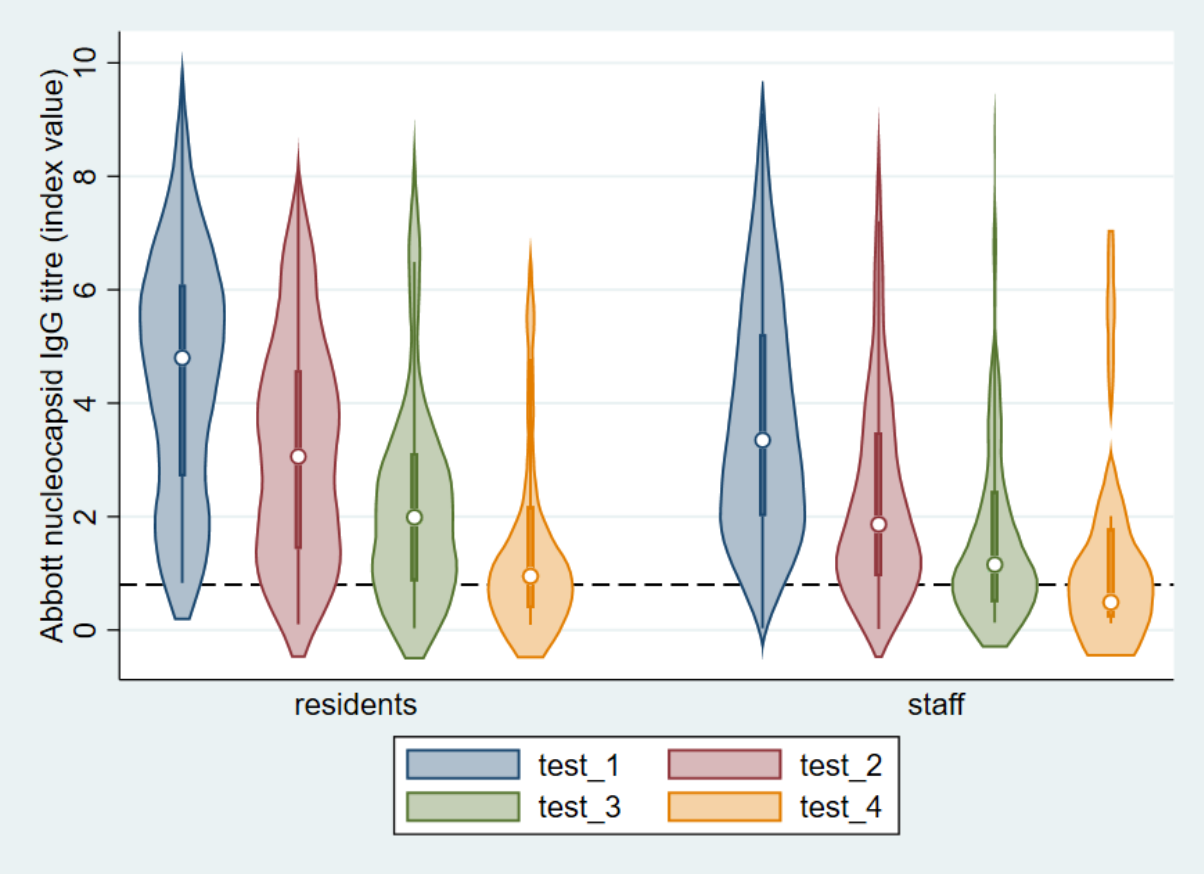
Table 5-4: Mean difference in antibody titre between rounds, by staff and residents with P-value.

Mean antibody titre (index value) (SD)	**Staff**	**Residents**	***P****
Round 1	3.71 (0.11)	4.52 (0.14)	<0.0001
Round 2	2.43 (0.10)	3.21 (0.13)	<0.0001
Round 3	1.87 (0.13)	2.41 (0.16)	0.0042
Round 4	1.34 (0.43)	1.72 (0.32)	0.2443
Mean inter-round difference (index value) (SD)	**Staff**	**Residents**	***P****
Round 1 to 2	-1.28 (0.06)	-1.32 (0.06)	0.6631
Round 2 to 3	-0.88 (0.05)	-0.81 (0.08)	0.2199
Round 3 to 4	-1.16 (0.18)	-1.18 (0.15)	0.5235

*t-test

Figure 5-4: Violin plot showing distribution of Abbott nucleocapsid antibody (IgG) titres in staff and residents over rounds 1-4 of sampling.

Hollow dot in centre of plot is the median, thick vertical lines are the IQR and thin vertical lines are 1.5 x IQR. Dotted horizontal line represents the threshold for positivity (0.8 index value).



Anti-spike and anti-RBD antibodies were measured using the MSD assay in samples from 41 individuals who had sero-reverted. The majority of staff were female, 21/23 (91.3%), and median age was 57 (IQR 48-61). There were 14 residents of whom 4 (28.6%) were female and overall median age was 80.5 (71-88). Of the 16 samples (9 residents, 7 staff) that were tested 60-90 days from loss of anti-nucleocapsid antibody, spike antibodies remained above the positivity threshold in 12 participants (7 residents, 5 staff) and RBD antibodies in 8 participants (5 residents, 3 staff), Figure 5.5.

Figure 5-5: Quantitative antibody titres over 90 days following nucleocapsid antibody sero-reversion for a) spike antibody b) RBD antibody.

(Reproduced from [292]*, under CC BY-ND licence with permission from Elsevier Ltd).* Titres are presented at date of first positive antibody (baseline), T1: 0-30 days and T2: 60-90 days after estimated date of sero-reversion, (n=41 at baseline, and T1 and 16 at T2). Titres are reported from MSD assay according to a logarithmic scale. Red interrupted line denotes cut-off for test positivity (spike = 1200 AU/ml, RBD = 600 AU/ml). Shaded boxes denote IQR, horizontal line in centre of each box is the median, whiskers span all points within 1.5 x IQR of the nearer quartile, dots are outlier data points beyond whiskers.

a)

b)

5.4.2 Vaccine-induced immunity

Quantitative anti-spike antibody titres were measured in 1317 samples from 1034 participants that were eligible for inclusion based on availability of vaccination history, 558 (42.4%) from 402 residents (282 female, 120 male) and 759 (57.6%) from 632 staff (550 female, 82 male), between 15th March 2021 and 22nd October 2021. Median age in staff was 50 (IQR 37-58) years and 86 (78-91) years in residents. Participants contributed a maximum of three samples (23/1034), with three quarters (774/1034) donating one sample to the analysis. Participants from 82 care homes were included, 70 for-profit, 6 independent, and one not-for-profit across all nine regions of England ranging from 3 (3.7%) in East Midlands to 21 (25.6%) in North-West.

Linkage to national datasets was possible for 595 participants (57.5%) and vaccination data were obtained directly from the care homes for 439 participants (42.5%). More than half of staff (61.7%) were administered the Pfizer-BioNTech vaccine (390/632) whereas the opposite pattern was seen amongst residents in whom more than half (63.2%) received the Oxford-AstraZeneca vaccine (254/402). The median time between receipt of second vaccination and antibody measurement date was 136 days (IQR 104-170). There were 24 samples dropped from the analysis because of evidence of breakthrough infection following vaccination; 12 from staff and 12 from residents.

A better fit was demonstrated in the regression model with interactions between predictors and vaccine-type than the model with independent effects for each predictor (LRT P=0.01). This fit was improved when age was added as a linear variable (LRT P=0.03), therefore this model is presented.

Table 5-5: Regression coefficients from final model for anti-spike antibody levels from 21 days following second vaccine dose, fitted to log10-transformed data.

(Reproduced from [335], under CC BY license, with permission from OUP).

	n, n (%*) or median (IQR)	***Intercept*† *(95% CI); P***	***Slope (95% CI); P [annual change]***
Reference coefficients‡		4.12 (3.86 to 4.38)	-0.67 (-1.48 to 0.14)
Oxford-AZ recipients	493	***Difference in intercept (95% CI) #; P***	***Difference in slope (95% CI) **; P***
Prior infection (yes *vs* no)	246 (49.9)	0.68 (0.5 to 0.85); <0.01	0.50 (-0.01 to 1.01); 0.06
Care home resident (*vs* staff)	251 (50.9)	0.22 (-0.14 to 0.59); 0.23	-0.45 (-1.58 to 0.67); 0.43
Male (*vs* female)	105 (21.3)	0.17 (-0.05 to 0.39); 0.13	-0.69 (-1.32 to -0.05); 0.03
Age (per 10y greater than 70)	67 (48–87)	-0.10 (-0.18 to -0.02); 0.01	0.16 (-0.09 to 0.42); 0.20
Pfizer-B. recipients	534	***Difference in intercept (95% CI) #; P***	***Difference in slope (95% CI) **; P***
Difference *vs* Oxford-AZ¶		0.90 (0.56 to 1.23); <0.01	-1.09 (-2.04 to -0.14); 0.02
Prior infection (yes *vs* no)	306 (57.3)	0.44 (0.27 to 0.61); <0.01	0.43 (0.01 to 0.85); 0.04
Care home resident (*vs* staff)	147 (27.5)	-0.05 (-0.36 to 0.26); 0.74	0.06 (-0.7 to 0.82); 0.87
Male (*vs* female)	94 (17.6)	0.11 (-0.1 to 0.31); 0.31	-0.23 (-0.72 to 0.26); 0.36
Age (per 10y greater than 70)	56 (44–71)	-0.01 (-0.08 to 0.06); 0.76	-0.06 (-0.23 to 0.11); 0.49

**% calculated using number with same vaccine type as denominator.*

†Representing average peak value at 21 days after second vaccine dose.

‡Values for Oxford-AZ recipient female staff member at 70 years of age without prior infection.

¶Taken alone, represents the difference for female staff member at 70 years of age without prior infection.

#10^x gives multiplicative difference in intercept associated with each factor.

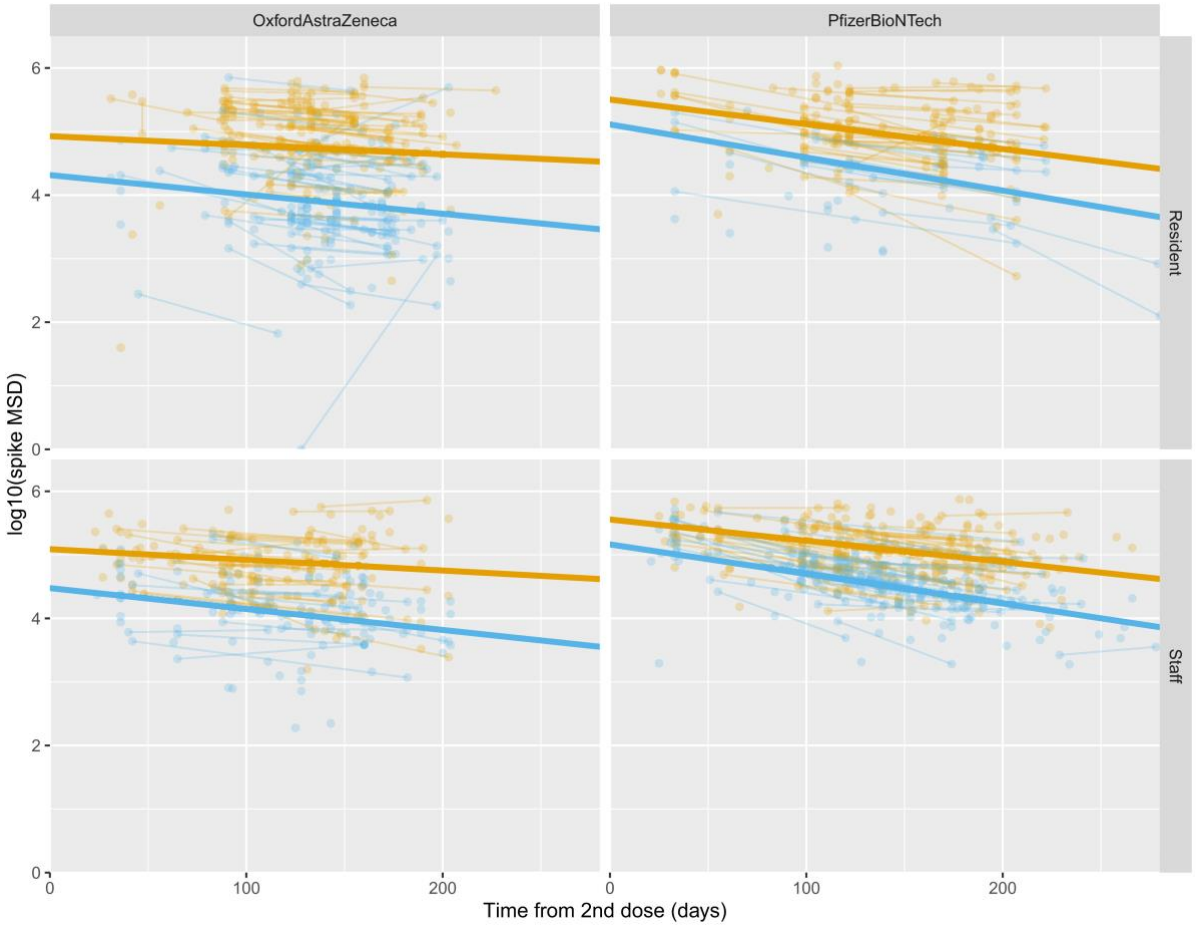
***10^x gives multiplicative difference in value at 12 months from peak level.*

Pfizer-BioNtech recipients (specifically for a female staff member at age 70 years without prior infection) had 7.9 times higher peak antibody titres than Oxford-AstraZeneca recipients (95% CI 3.6-17.0, *P*<0.01), Table 5.5 and Figure 5.6. Prior infection was associated with greater peak antibody titres in both groups, although more strongly in Oxford-AstraZeneca (×4.8, 3.2-7.1; *P*<0.01) than Pfizer-BioNTech recipients (peak ×2.8, 1.9-4.1; *P*<0.01). Although staff / resident grouping was not associated in any difference in peak antibody titres, when other predictors were accounted for, recipients of Oxford-AstraZeneca had lower antibody peak as age increased (peak x0.79 per 10 years above 70-years, 0.66-0.95, *P*=0.01).

Figure 5-6: Log-transformed anti-spike antibody levels based on MSD testing, according to time from second vaccination.

(Reproduced from [335], under CC BY license, with permission from OUP)

The data are divided by vaccine type and staff/resident status. Points or lines in orange represent those with evidence of prior infection and blue are those without. Dots represent individual observations and lines connect those from the same person. Regression fits from the model are shown with bold straight lines to estimate trends from each group (omitting age and sex).



The rate of antibody decline was steeper amongst Pfizer-BioNTech recipients (specifically for a female staff member at age 70 years without prior infection) when compared with Oxford-AstraZeneca vaccinees (×0.08 at 12 months vs equivalent decline from peak, 95% CI 0.01-0.72; *P*=0.02). Prior infection was also associated with a slower rate of decline following both vaccination types when compared with infection-naïve participants (x3.16 at 12 months for Oxford-AstraZeneca, 0.98-10.23, *P*=0.06 and x2.69 for Pfizer-BioNTech, 1.02-7.08, *P*=0.04). Staff / resident status and age did not significantly affect the slope for either vaccine type. Males who had received the Oxford-AstraZeneca vaccine had a slightly faster decline in antibody levels when compared with females (x0.20, 0.05-0.89, *P*=0.03) however this association was not seen in Pfizer-BioNTech recipients, Table 5.5, Figure 5.6.

Table 5-6: Estimated marginal intercept, slope terms and corresponding half-life from the final statistical model for anti-spike antibody levels for participant sub-groups.

(Reproduced from [335], under CC BY license, with permission from OUP).

Subject	**Vaccine**	**Sex and inf. status**	**Average peak (intercept) (95% CI)**	**Average value at 6 months*(95% CI)**	**Slope (95% CI) (annual)**	**Half-life (days)**
Resident	*Ox.-AZ*	*Female: no inf.*	4.07 (3.84 to 4.29)	3.75 (3.56 to 3.95)	-0.63 (-1.33 to 0.06)	174 (83 to inf)
		Female: prior inf.	4.75 (4.55 to 4.95)	4.68 (4.51 to 4.85)	-0.14 (-0.73 to 0.46)	806 (151 to inf)
		Male: no inf	4.30 (4.07 to 4.53)	3.58 (3.39 to 3.77)	-1.45 (-2.12 to -0.78)	76 (52 to 140)
		Male: prior inf.	4.94 (4.7 to 5.19)	4.50 (4.29 to 4.72)	-0.88 (-1.6 to -0.15)	125 (68 to 741)
	Pfizer-B	*Female: no inf.*	4.99 (4.73 to 5.26)	4.07 (3.88 to 4.25)	-1.85 (-2.48 to -1.22)	59 (44 to 90)
		Female: prior inf.	5.42 (5.22 to 5.62)	4.72 (4.57 to 4.87)	-1.40 (-1.9 to -0.91)	78 (58 to 121)
		Male: no inf	5.08 (4.82 to 5.35)	4.05 (3.87 to 4.24)	-2.06 (-2.68 to -1.44)	53 (41 to 76)
		Male: prior inf.	5.51 (5.29 to 5.74)	4.71 (4.55 to 4.86)	-1.62 (-2.16 to -1.07)	68 (51 to 102)

Staff	*Ox.-AZ*	*Female: no inf.*	4.33 (4.16 to 4.49)	3.83 (3.69 to 3.97)	-1.00 (-1.5 to -0.5)	110 (73 to 221)
		Female: prior inf.	5.00 (4.86 to 5.15)	4.75 (4.62 to 4.89)	-0.50 (-0.93 to -0.07)	219 (118 to 1562)
		Male: no inf	4.50 (4.25 to 4.74)	3.65 (3.44 to 3.86)	-1.69 (-2.41 to -0.96)	65 (46 to 115)
		Male: prior inf.	5.17 (4.93 to 5.42)	4.58 (4.37 to 4.79)	-1.19 (-1.91 to -0.46)	92 (57 to 237)
	Pfizer-B	*Female: no inf.*	5.04 (4.9 to 5.17)	4.22 (4.12 to 4.31)	-1.64 (-1.98 to -1.3)	67 (55 to 84)
		Female: prior inf.	5.48 (5.34 to 5.61)	4.87 (4.78 to 4.97)	-1.21 (-1.56 to -0.87)	91 (71 to 127)
		Male: no inf	5.14 (4.93 to 5.36)	4.21 (4.05 to 4.36)	-1.87 (-2.37 to -1.37)	59 (46 to 80)
		Male: prior inf.	5.58 (5.35 to 5.82)	4.86 (4.7 to 5.03)	-1.44 (-2.02 to -0.86)	76 (54 to 127)

Staff age set at 50 and resident age at 86 in line with median age from these groups.

*From peak level 21 days after second vaccine dose.

Anti-spike antibody half-lives were estimated from the model and did not differ significantly between staff and residents, Table 5.6. These mainly ranged between 59 and 125 days with lowest half-life in female Pfizer-BioNTech recipients without prior infection (staff 59 days, 95% CI 46-80, vs residents 59 days, 95% CI 44-90) and longer half-life in Oxford-AstraZeneca recipients, particularly those who had been previously infected and were men (residents 125 days, 68-741, vs staff 92 days, 71-127). Antibodies in female Oxford-AstraZeneca recipients with prior infection had the longest half-life although confidence intervals were wide due to small sample size, so estimates are likely to be imprecise.

5.5 Discussion

Anti-nucleocapsid antibodies remained positive for eight months following infection in around half of study participants, although the rate of sero-reversion was faster in staff than in residents. As the SARS-CoV-2 virus enters its fourth year in circulation, this has implications for the use of seroprevalence studies to inform population-based

estimates of immunity, as these are likely to underestimate the proportion who have been exposed to infection.[331] In contrast, post-vaccine anti-spike antibody responses are comparable in magnitude and longevity for both Pfizer-BioNTech and Oxford-AstraZeneca recipients between care home staff and residents. Although Pfizer-BioNTech recipients exhibit a greater peak antibody titre, they are estimated to have a steeper rate of decline over the subsequent 12 months when compared with Oxford-AstraZeneca. Prior infection results in a greater antibody peak with slower decline across vaccine types which is particularly important in the care home population who have high levels of past infection as demonstrated by the seroprevalence estimates in Chapter 4. Amongst Oxford-AstraZeneca recipients, age was associated with a lower antibody peak, and antibody titres declined faster in males. These findings are similar to those reported from the general population[336–338] and suggest that care home residents mount a durable humoral immune response to SARS-CoV-2 vaccination.

Rates of nucleocapsid sero-reversion were greater in staff than residents in the main survival analysis and this finding was replicated in the sensitivity analysis that used more reliable seroconversion dates. It is possible that residents experienced more severe infections which has been shown to elicit stronger immune responses.[333,334] However I found no difference in the sub-group analysis that compared individuals requiring hospitalisation for their primary infection with those who did not, although numbers were small. As I did not have data on symptoms, I used hospitalisation as a proxy for severity which may have underestimated severe cases and misclassified them into non-severe, especially in the early pandemic when the threshold for admission to hospital was higher due to bed pressures.[339,340]

Residents had greater nucleocapsid antibody titres than staff over the first three rounds of testing. As the decline in antibody titre between rounds was comparable between groups, this may explain why titres in staff dropped below the positivity threshold and sero-reverted sooner than residents. Larger antibody responses and greater time to nucleocapsid sero-reversion in older people has been demonstrated in other cohorts, including a study from China in >500 hospitalised patients followed up for one year,[341] and a community cohort study from England with a sample size of almost 14,000.[342] The difference in size of antibody response may be related to prolonged viral clearance that has been reported in older people[343–345] or because of

disordered antibody production from immunosenescence and inflammaging where antibodies produced are polyreactive and may have lower binding affinity to antigens.[136,346] As I could only include infection survivors in my analysis, it is also possible that participants were immunologically more robust than those that died from their infection which may account for the differences in magnitude in antibody responses.

Spike and RBD antibodies persisted beyond nucleocapsid antibody sero-reversion suggesting more rapid waning of nucleocapsid compared with responses against the S1 subunit (spike and RBD). This was also demonstrated in a longitudinal cohort of 331 hospitalised patients that used the same nucleocapsid-antibody assay, but higher manufacturer recommended threshold of 1.4 index value applied. Nucleocapsid sero-reversion was estimated to occur after one year in half of the population and spike sero-reversion after two years, although severity of disease in this population was probably greater than in my study cohort.[333] This was also demonstrated in the only longitudinal study that I found from care home residents, which followed up 106 over 54-year-old participants for seven months following an outbreak in one facility and found significant waning of nucleocapsid antibodies on both Abbott and MSD assay when compared with spike and RBD.[321]

The absence of a statistically significant difference between the staff and resident group in the analysis of spike antibody responses to vaccination is reassuring. As more than half of both staff and residents in this analysis had evidence of prior SARS-CoV-2 infection, it is likely that their vaccine-induced immune responses were boosted. This has been shown across a number of studies to be the strongest predictor of immune responses to vaccination, regardless of age,[327,347] with previously vaccinated care home residents exhibiting anti-spike antibody titres that are eight times greater than their infection-naïve resident peers.[348]

A greater post-vaccine antibody peak with more rapid rate of decline in titres in Pfizer-BioNTech recipients when compared with Oxford-AstraZeneca has been described in larger community-based random sampling studies from the general population including one with sample size > 50,000[337] , and the other >8,500.[349] These have also shown that male Oxford-AstraZeneca recipients and older individuals had lower peak

antibody titres. These studies described comparable mean half-life following either vaccine types with extended half-life in those with prior infection,[337,349] which is similar to findings from this analysis. Greater immunological responses to vaccination in females, as seen in my study for female Oxford-AstraZeneca recipients, have been described for other vaccinations however the mechanisms behind this difference are not clearly understood.[338,350] A further cohort of 1750 participants found a larger drop in spike antibody levels over six months following vaccination in older participants. However, this may not be comparable to the population in my analysis which consisted of 40% residents, as over 60-year-olds only made up 18% of their cohort and <10% had evidence of prior infection.[351]

5.6 Strengths and Limitations

This study is one of the largest studies to monitor antibody levels in care home residents with the longest follow up. Data linkage to routinely collected datasets on asymptomatic PCR testing and hospitalisation enabled estimation of primary infection dates and prevented recall bias as data collection was in real-time. In addition, the multivariable modelling approach for spike-antibody waning accounted for the effect of confounders, like sex, on the outcome. For both analyses, the sample included care homes from a range of providers across England which limited selection bias and improved generalisability of results. Findings of these analyses were also presented in a timely manner to policymakers and informed decisions at the time, such as decisions about re-vaccination. The nucleocapsid waning analysis was submitted as a report to NERVTAG in May 2021 when they were examining this issue in different populations. It was also one of the first papers to be published that examined antibody waning in the care home population. The spike waning data was presented to Ministers via the DHSC Data Debrief Group (DDG) which was chaired by the DHSC Director of the COVID-19 response and informed national re-vaccination policy.

Limitations of the analyses include incomplete data linkage and data capture on primary infection. Full linkage was not possible for 42.5% of samples included in the analysis of anti-spike antibody levels. This was because many care homes that joined the study early on could not access NHS numbers for staff and residents in the study, therefore they could only provide name, date of birth, and address as identifiers.

Despite extensive efforts to correct errors in spelling or dates, linkage to pseudo-identifiers using these data was incomplete. Although we were able to obtain vaccination records for these individuals directly from providers, care homes did not collect infection dates. This means that prior infections may have been missed for these participants, which may have led to overestimates in the peak antibody titres and rate of decline in the infection-naïve group.

In addition, it was not possible to accurately estimate the date of seroconversion in 503 out of the 619 (81.3%) participants included in the anti-nucleocapsid sero-reversion cohort. Findings from the sensitivity analysis in those with known date of primary infection showed a similar pattern albeit a slightly lower rate of sero-reversion. This is likely to be related to the shorter follow-up in this group which may have missed sero-reversion events that occurred later. Both analyses had a relatively modest sample size which, although larger than any other studies in this population to date, may affect the accuracy of the effects observed.

The nucleocapsid-antibody waning that was demonstrated may also have affected the accuracy with which prior exposure status was classified in the spike antibody model as some participants who were infected in the first wave of the pandemic may have been incorrectly considered infection-naïve. As there is evidence of greater waning with the Abbott assay when compared with other assays,[352] I used a lower Abbott threshold for positivity in both analyses (as described in Chapter 4). The MSD assay has been shown to have better sensitivity than the Abbott assay,[353] therefore in the spike-antibody waning analysis we classified prior infection exposure using the MSD anti-nucleocapsid results, where available. However, as we used post-vaccination samples, we could not be certain that primary infection had occurred before vaccination.

Finally, this analysis does not account for cellular immune responses or neutralising antibodies that play an important role in protection against infection.[354] Evidence to date suggests robust vaccine-induced cellular responses amongst older people.[327,355] In view of persistent cellular responses, there is growing evidence that titres of spike and RBD titres can correlate with magnitude of cellular response[351] and level of protection.[349,356,357]

Although I found differences in the magnitude of nucleocapsid antibody responses between staff and residents, it is possible that these were related to survivor bias as residents that had died before blood sampling may have exhibited less effective immune responses. Exploring these differences by examining underlying medical co-morbidities and ethnicity may have been possible, however the only data that I could access were hospitalisation records with ICD-10 diagnostic codes and ethnicity data (collected from Pillar 2 or hospital records) were incomplete (most participants were coded as “British”). These are usually completed by busy ward staff so complex and often overlooked diagnoses like frailty and dementia may be absent, thereby introducing additional bias into analyses. This is an important direction for future work as early research suggests that frailty is associated with a larger drop in post-vaccine spike-antibody titres in a small study of care home residents.[358] A competing risk analysis may have helped explore the survivor bias, however as mentioned previously participants who died in the first wave could not be identified.

I decided against modelling the rate of nucleocapsid antibody waning using antibody titres as the Abbott ARCHITECT immune-assay is a semi-quantitative assay.[359] This means that although the assay has been validated to report binary results (positive/negative), antibody titres do not directly correlate to antibody level. In addition, the focus of the analysis was the real-world application of nucleocapsid antibody to determine population-level immunity and antibody titres are not routinely available to clinicians and public health officials. However, as the MSD assay is quantitative and spike antibody levels in all the post-vaccination samples were above the positivity cut-off, we chose to model these as a continuous variable. Since these analyses were performed, evidence has been published showing that sex and primary infection variant may impact on waning,[360] however evaluating this relationship in more detail was outside the scope of this thesis.

5.7 Conclusion

Over the nine-month follow-up, spike antibody levels remained positive for all participants, with mean antibody half-life between two and six months. Although the antibody level that is protective against infection has not been well described and could

differ over periods of different viral variant predominance (i.e., Delta, Omicron), it is reassuring that, amongst residents, antibodies remain detectable after vaccination and may have some protective effect. Nucleocapsid antibodies wane in half of the care home population within eight months of infection, suggesting a better test to identify infection-naïve individuals is needed to inform estimates of immunity, particularly on a population-level. Although spike antibodies appear to remain detectable for longer than nucleocapsid antibodies, the widespread administration of booster vaccines has precluded further follow-up and meaningful comparison. Alongside quantification of antibody responses, vaccine efficacy studies are required to describe the functional protection against re-infection and severe outcomes in care home residents, particularly considering how the influence of prior infection varies with variants. As turnover of staff and residents in care homes is high,[127–129] vaccine efficacy may decline as SARS-CoV-2 incidence drops and residents are replaced with infection-naïve individuals who may have been shielding at home, thereby depleting the reservoir of previously infected hosts.

In addition to immunity to infection, environmental characteristics of the care home are likely to be associated with risk of SARS-CoV-2 infection and transmission. Identifying care homes with lower levels of immunity (infection or vaccine-induced) and with facility features that put them at greater risk of ingress and transmission of infection, may inform more targeted use of re-vaccination and other NPIs to protect residents in higher-risk care homes. This will be important given the likely cost implications of further vaccination rounds and concerns about inequity in vaccine access globally. I plan to explore this in more detail in Chapter 6 by describing the built environment of care homes and modelling facility-level risk factors for infection.

5.8 Contribution & Dissemination

5.8.1 Naturally acquired immunity.

I designed the study and the statistical analysis plan, led organisation of blood sample collection, carried out data analysis and drafted the manuscript for publication.

I also summarised the initial analysis of sero-reversion rates in a report that was submitted to NERVTAG, a sub-group of SAGE, that supports the UK government to make decisions around respiratory viral threats.

This analysis has been published as a full article in Lancet Health Longevity:

Krutikov M, Palmer T, Tut G, Fuller C, Azmi B, Giddings R, Shrotri M, Kaur N, Sylla P, Lancaster T, Irwin-Singer A, Hayward A, Moss P, Copas A, Shallcross L. (2022). **Prevalence and duration of detectable SARS-CoV-2 nucleocapsid antibodies in staff and residents of long-term care facilities over the first year of the pandemic (VIVALDI study): prospective cohort study in England**. *Lancet Healthy Longev* 2022; 3:e13-21 DOI: 10.1016/S2666-7568(21)00282-8

With commentary piece here:
Verschoor CP, Bowdish DME. (2022). **Estimating SARS-CoV-2 seroprevalence in long-term care: a window of opportunity**. *Lancet Healthy Longev* 2022; 3:e2-3 DOI: 10.1016/S2666-7568(21)00304-4

5.8.2 Vaccine-induced immunity

I collected the data and designed the study. I worked with Dr Oliver Stirrup, a statistician on the VIVALDI team, to develop the analysis plan and he developed and ran the models. We interpreted the results and drafted the published manuscript together.

The results were presented to the DDG within the DHSC, a group of policymakers and scientists undertaking COVID-19 surveillance work, who meet weekly to share and discuss recent epidemiological data relating to COVID-19.

The manuscript has been published in the Journal of Infectious Diseases and I am joint first author with Oliver Stirrup.

Stirrup O., Krutikov M., Tut G., Palmer T., Bone D., Bruton R., Fuller C., Azmi B., Lancaster T., Sylla P., Kaur N., Spalkova E., Bentley C., Amin U., Jadir A., Hulme S.,

Giddings R., Nacer-Laidi H., Baynton V., Irwin-Singer A., Hayward A., Moss P., Copas A., Shallcross, L. (2022). **SARS-CoV-2 anti-spike antibody levels following second dose of ChAdOx1 nCov-19 or BNT162b2 in residents of long-term care facilities in England (VIVALDI)**. *The Journal of Infectious Diseases* 2022; 226:1877-1881. DOI: 10.1093/infdis/jiac146

Chapter 6

Objective 4: Testing the hypothesis that care home characteristics are risk factors for SARS-CoV-2 infections and outbreaks and that factors associated with infection ingress differ from those associated with transmission.

In Chapter 4, I explored the influence of the agent on infection risk by describing rapid entry of the emergent SARS-CoV-2 Alpha variant into care homes in regions of high community transmission. I also demonstrated that seroprevalence can be a useful tool for measuring infection prevalence when diagnostic testing is limited, however the accuracy of this as a measure of population-level immunity is affected by rapid decline in anti-nucleocapsid antibody titres (Chapter 5). Additional linkage to PCR and hospitalisation data may enhance estimates of prior infection, however under-ascertainment is still an issue. Vaccination elicits detectable spike antibody responses however the threshold that may correlate with protection is unknown. Rapid changes in the population and in antibody levels impact on the ability to estimate how population-level immunity affects infection risk. As such, although I had initially planned to model the relationship between population-level immunity and infections and outbreaks, I elected to adjust for these factors only.

My final objective was to investigate the influence of care home specific environmental factors on the risk of infections and outbreaks. I designed surveys about the built environment which were distributed amongst care homes participating in VIVALDI and I linked responses to routine individual-level data on SARS-CoV-2 infections and vaccinations and facility-level data. I modelled associations between building factors and four outcomes describing infection ingress and transmission, adjusting models for known confounders. Outcomes were incidence of infection amongst residents, outbreaks, outbreak size and duration. As the Omicron variant is significantly more transmissible than preceding variants, I stratified analyses by Omicron-dominance.

I found substantial variability in built environments. The only factor associated with ingress, was local community incidence. Environmental factors consistently associated with transmission included size, number of storeys, building type, ventilation, indoor temperature, and subjective air quality. Relationships were affected by Omicron variant, supporting evidence of the impact of agent factors on infection risk.

This analysis demonstrates that the built environment affects risk of infections but not outbreaks. This supports findings from my scoping review that care homes may not be able to stop infection ingress but can potentially limit spread. It is challenging to model risk factors for SARS-CoV-2 due to confounders and changing epidemiology and policy over the pandemic. Further work beyond my thesis could investigate the built environment in more depth including modifications that may limit infection transmission to prepare for future respiratory viral threats.

6.1 Background

The scoping review in Chapter 2, identified that the main facility-level factors associated with risk of SARS-CoV-2 introduction and transmission were infection incidence in the local community, care home size, ownership type, and crowding. No studies considered the built environment in detail as most relied on routine administrative data. All studies were conducted early in the pandemic so could not account for the impact of vaccination and circulating variants. In addition, most were subject to ascertainment bias as they only considered symptomatic infections and in some, these were based on clinical diagnosis alone. The review also showed differences between factors associated with infection ingress and transmission, despite some overlap. This is relevant to public health teams and policymakers as measures that prevent each of these outcomes could be adapted in line with evidence.

Transmission of SARS-CoV-2 virus is mainly respiratory, through inhaled droplets and aerosols, [27–29] described in Chapter 1. Models of other respiratory pathogens such as *Mycobacterium tuberculosis* have shown that the features of the built environment influence transmission risk.[361–363] The Wells-Riley equation estimates the probability of becoming infected with a respiratory pathogen in a space[364,365] and has been applied with

some adaptations to model SARS-CoV-2 transmission in nail salons, schools, and other public buildings,[187–189] however I am not aware of any such studies from care homes. The equation bases calculations on the air flow, infectious inoculum size, number of infected and susceptible individuals within the space, estimated pulmonary ventilation rate of occupants, and exposure time. It provides a useful framework to identify factors that play a role in indoor transmission. However, some factors featured in this equation probably differ in care homes when compared with other spaces, suggesting variation in transmission dynamics. This includes the physiology and immune responses of older care home residents (affecting ventilation rate and number of susceptible / infected individuals), the fact that residents live in these spaces (affecting number of occupants and exposure time), and diversity in ventilation systems, insulation, and building design (affecting air flow).

In common with other healthcare settings, care homes are semi-closed clusters with a predominantly non-resident staff body, that includes domiciliary and catering staff as well as front-facing clinical staff who are exposed to SARS-CoV-2 in the community (Chapter 2).[186] In addition, as described in Chapters 1 and 4, staff turnover is high – estimated to be 29% in 2021/2022.[135] Care homes are visited by healthcare professionals such as general practitioners and family or friends of residents. Although the resident population is fairly stable, many care homes accept short-term admissions from the community or from hospital for intermediate care while awaiting further placement or respite care and this number fluctuated substantially over the pandemic. Homes also vary in design and layout with some based in residential buildings that have been converted whereas others are purpose-built. Recommendations and preferences for design have varied over the years however over the last three decades these have emphasised adaptations for people living with dementia, reducing risk of falls, and increasing quality of life for residents.[366] To date, this variation in care home design has not been comprehensively documented.

In the first three objectives of my PhD, I have illustrated the significant exposure to infection in care homes over the pandemic, how this varies with the circulating variant and across homes, and the extent and duration of humoral immunity in residents and staff following infection and vaccination. The fourth objective of my PhD aims to investigate associations between the built environment of care homes and the introduction and transmission of SARS-CoV-2. This could identify care homes that would benefit from additional support from local public health teams and inform

recommendations for environmental modifications that may be better tolerated than some of the infection control measures imposed over the pandemic. It may also be possible to apply these results to management of outbreaks with other respiratory pathogens such as influenza, for which there is currently limited evidence.[365,367] As I found many analyses that had considered characteristics such as staffing, resident demographics, and local factors, I focussed my analysis on the features of the built environment, which were not well-described. I also found limited published evidence on how infection risk is affected by facility-level immunity to SARS-CoV-2 and had planned to consider this relationship in more detail. However, I identified substantial limitations with the approaches to estimating immunity that I had access to (as described in Chapters 4 & 5). As such, I decided that it would not be possible to model these associations reliably but, using the data I had, I could adjust for host immunity.

To address my fourth PhD objective, I conducted two sub-studies within the VIVALDI cohort, Figure 6.1.

i) First, I designed and administered a cross-sectional survey to participating care homes, which I analysed to describe the diversity in built environments.

ii) Second, I linked the survey data to the longitudinal VIVALDI dataset and estimated incidence rates of infection and outbreaks. I modelled associations between environmental risk factors and infection ingress and transmission in care homes, accounting for key confounders identified from the epidemiological triad (Chapter 1) and my scoping review (Chapter 2). I modelled ingress and transmission separately by applying four outcomes: infection incidence, outbreak incidence, and outbreak size and duration. An outbreak was defined as at least two cases occurring in one care home over a fourteen-day period, at least one of which a resident. Depending on the outcome, models were built on an individual or facility-level:

<u>Individual level:</u>

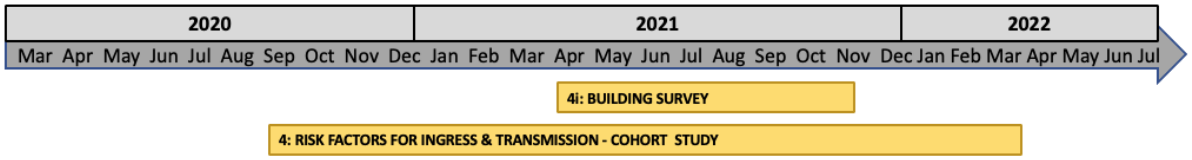
1) Incidence of SARS-CoV-2 infection in a resident (ingress & transmission)

<u>Facility-level:</u>

2) Incidence of SARS-CoV-2 outbreak in a care home (ingress)
3) Size of the outbreak (transmission)

4) Duration of the outbreak (transmission)

Figure 6-1: Objective 4 Timeline



6.2 Methods

Between 4th April and 2nd November 2021, I conducted a cross-sectional survey on the built environment in care homes for older adults in England, participating in VIVALDI. I linked survey responses to routinely collected data from staff and residents on asymptomatic and symptomatic SARS-CoV-2 testing and their vaccination status for the period 1st September 2020 - 31st March 2022. The analysis period was aligned to the dates that the national asymptomatic screening programme was operational (1st September 2020[368] - 31st March 2022), as tests allowed identification of the study cohort by linkage to the participating care homes.

6.2.1 Survey design and administration

In collaboration with Dr Hector Altamirano, building scientist at the UCL School of the Built Environment, I developed surveys to collect information about care home environmental characteristics. These were based around themes from the Wells-Riley equation, Figure 6.2, and included a mix of multiple-choice and short answer questions around size, type of building, layout, air flow, and the use of the space (appendix 6.1).

Although most questions were self-explanatory, these were the options presented to describe ventilation. Central air conditioning units that disperse cooled air from a central unit through a system of ducts around the property (central air conditioning unit); ceiling mounted units that cool or heat recirculated air and can be more efficient at air cooling and distribution (ceiling cassette unit); portable air conditioning units with either a single hose that outputs hot air out of the room usually through the window or two hoses – one for intake and one for output (portable unit exhaust pipe); MVHR (Mechanical ventilation with heat recovery) which is a centralised system that extracts

and filters air and recovers heat to recirculate through the building;[369] units that act similarly to MVHR but adjusts humidity, and recirculate cooled instead of heated air (mechanical extraction unit); or a freestanding fan (freestanding).[370,371]

As the survey was designed early in the pandemic when care homes were experiencing a high workload, I could only pilot it with two managers. These were followed up with telephone interviews to review questions individually. Based on this feedback, wording was clarified, and options added to multiple-choice questions. Survey design and reporting conformed to the CROSS checklist[372] (appendix 6.2).

Figure 6-2: Survey themes, criteria, and indicators based on key components of the Wells-Riley equation.

(Adapted from MSc thesis by Niyathi Sethu with permission)

Theme	Criteria	Survey indicator
Environment	Air Quality	Ventilation
		Perceived air quality
		Presence of windows
		Condensation
		Access to outdoor space
		Heating
	Environmental conditions	Temperature
		Sunlight
Building characteristics	Layout / shared facilities	Shared rooms or bathrooms
		Number of beds
Operation & Maintenance	Timetable	Maximum number of people in common rooms at one time
		Exposure time
	Staffing	Number of staff
	Interventions	Cleaning
		Social distancing
		Use of PPE

PPE Personal Protective Equipment

Although the intention was to distribute the questionnaires in August 2020, as close to the study start as possible, the strain on care homes from rapidly changing policies and successive waves of infection meant that they were only completed the following year, between 4th April and 2nd November 2021. Questionnaires were distributed to a convenience sample of care homes that were identified by project managers from two large for-profit (chain sizes 250 and 300 homes), one medium not-for-profit chain (60 homes) and one small not-for-profit provider (chain size nine homes). In this period, one questionnaire was completed per care home by the manager or member of maintenance staff. No personal data were collected, and providers consented to collection of aggregate data when they enrolled to VIVALDI. There were no incentives for completion, however reminders were sent until 1st November 2021. Completed questionnaires were collated in a database by one of the Vivaldi project managers (CF) and stored in a secure data repository in the UCL Data Safe Haven.

6.2.2 Cohort study

I accessed the VIVALDI dataset, described in detail in Chapter 2. This contains all PCR/LFD results conducted under asymptomatic, symptomatic, and outbreak testing that are linked to care home CQC-IDs taking part in the study. As outlined in Chapter 1, the asymptomatic screening programme was introduced in England in July 2020 but fully established in the following September[368] and continued until March 2022. All staff were tested weekly and residents monthly and from December 2020 when LFD tests were introduced, more frequent testing of staff was recommended. These tests were also linked to unique pseudo-identifiers for each individual which enabled linkage to SARS-CoV-2 vaccination records and SARS-CoV-2 anti-nucleocapsid antibody tests (Abbott ARCHITECT i-system, Maidenhead UK) in those who consented to blood samples.

On a facility-level, I used the CQC-ID to link to daily bed occupancy, staffing, and bed numbers from the Capacity Tracker dataset and linked to the national SARS-CoV-2 incidence data using lower layer super output areas (LSOAs). The UK is split into smaller geographic areas (output areas), that can be used as units for census reporting - LSOAs comprise 4-5 of these output areas consisting of 400-1200 households.[373] I linked each CQC-ID to its corresponding LSOA and used this to link

to the local SARS-CoV-2 incidence from the national dataset. I also linked each care home by its postcode to deciles of socio-economic deprivation described by the Index of Multiple Deprivation (IMD), published by the ONS.[374] I collected data on the proportion of beds funded for dementia care and by the Local Authority (LA) directly from care homes.

All care homes completing the survey that could be linked to the CQC-ID of a facility taking part in VIVALDI were included and individuals with valid pseudo-identifiers linked to these CQC-IDs by at least one PCR or LFD test within the analysis period were included. I defined subject type in line with the methods described in Chapter 4.

6.2.3 Outcomes and covariates

The primary outcomes were the incidence of SARS-CoV-2 infection in residents (who probably acquired the infection in the care home, given the restrictions on excursions) and incidence of SARS-CoV-2 outbreaks. I included two primary outcomes as each describes different elements of infection dynamics: infection incidence among residents probably describes both ingress and transmission of SARS-CoV-2, whereas outbreak incidence describes infection ingress alone (Chapter 2). I defined cases based on a positive PCR or LFD test and, to account for persistent viral shedding, I excluded episodes from the same individual that were less than 90 days apart.[80,375] Where an individual had both a positive PCR and LFD test, the earliest was prioritised for the analysis. National guidance from PHE defines a SARS-CoV-2 outbreak in a care home as two PCR/LFD confirmed cases occurring within 14 days of each other, which triggers additional testing in all staff and residents.[376] The outbreak is declared complete when new cases have not been detected over a 28-day period. However, as my analysis aimed to describe transmission of infection within the care home, I modified this outbreak definition so that at least one of the cases had to be in a resident as these are likely to have been acquired within the care home. I included outbreaks if they started before the study end date.

The secondary outcomes were outbreak size and outbreak duration. Outbreak size included the number of staff and residents infected in the outbreak and the duration was considered as the number of days between the first and last positive test. The

outcomes aimed to describe transmission and to identify more susceptible care homes.

I built a ‘baseline’ time-varying model for each outcome and used this to adjust estimates of association between the building factors from the survey and the primary and secondary outcomes. The baseline model included covariates described in the literature as risk factors for SARS-CoV-2 infection. Full model results are not reported as these relationships have already been widely described and the number of covariates was large.

Table 6.1 outlines the covariates included in the models. If covariates varied over time, such as bed occupancy or vaccination coverage, an average for each month was used. Full vaccination was classed as receipt of two vaccine doses, irrespective of vaccine manufacturer, in line with national guidance.[110] Prior infection was defined as having a positive SARS-CoV-2 PCR/LFD test or anti-nucleocapsid antibody test at any point before the analysis month. In line with previous analyses (Chapters 4 & 5), I used a positivity threshold of 0.8 IU/ml for the Abbott anti-nucleocapsid antibody assay. The bed-to-resident and staff-to-resident ratios were calculated based on monthly averages of bed occupancy, total beds, and staffing numbers recorded in the Capacity Tracker dataset.

Table 6-1: Covariates included in baseline model.

Covariate	Infection incidence (resident)	Outbreak incidence	Outbreak size	Outbreak duration
Individual level				
Sex	X			
Age	X			
2nd vaccine dose*	X			
Prior infection*	X			
Facility-level				
Proportion >80 years*		X	X	X
Proportion residents female*		X	X	X
Proportion with prior infection (res)*	X	X	X	X
Proportion with prior infection (staff)*	X	X	X	X
Proportion fully vaccinated (res)*	X	X	X	X
Proportion fully vaccinated (staff)*	X	X	X	X
Number of residents*	X	X	X	X
Number of staff*	X	X	X	X
Number of beds	X	X	X	X
Staff-to-resident ratio*	X	X	X	X
Bed-to-resident ratio*	X	X	X	X
Local infection incidence*	X	X	X	X
IMD decile	X	X	X	X
Analysis month	X	X	X	X

*Time-varying

IMD Index of Multiple Deprivation

I included building factors if they were at least 80% complete and there was variability in responses (i.e., no more than 90% of responses were the same). As I could not verify the accuracy of responses, I treated temperatures greater than 30°C as missing. I performed listwise deletion for missing data in regression analyses because of the large number of explanatory variables and it was likely that the missingness of data was random (missing at random, MAR) therefore distribution of responses was retained as it was associated with ease of data access for the person completing the survey (i.e., air temperature and ventilation had most missing data and these data were most difficult to access). Although all missing data can introduce bias, regression analyses and listwise deletion for MAR data can be an effective way of dealing with missingness within the explanatory variable.[377,378]

6.2.3 Statistical analysis

I modelled the 7-day rolling incidence rate of SARS-CoV-2 infections among staff and residents and outbreaks in care homes. At-risk time was estimated using the date of the first PCR/LFD test in the care home (entry date) and the last test (exit date). If final tests occurred between 1st January and 31st March 2022 these were inflated by 100 days to account for missed tests (as residents were tested monthly and testing became less regular towards the end of the national testing programme), and individuals were removed for 90 days following a positive test (to account for persistent positivity).

To model primary and secondary outcomes I modelled each survey building factor separately and adjusted for confounders using the baseline model. Because the proportion of completed answers varied for each factor from 24% (32/134) to 97% (130/134), to preserve sample size I decided against building an overall multivariable model that included all or only purposefully selected building factors. I modelled infection incidence among residents using multivariate Poisson regression with frailty terms at individual and facility-levels to account for clustering. The analysis dataset consisted of individual-level monthly observations for residents only. The exposure term was the monthly number of days at-risk for each participant. To model outbreak incidence, I used a facility-level dataset with monthly observations for each care home and modelled risk factors using multivariate Poisson regression. The models had frailty terms at the care home level and an exposure term of monthly number of at-risk days for the care home. Poisson regression was chosen as infection and outbreak incidence are count data and there was good fit when comparing observed and predicted counts.

The secondary outcomes of outbreak size and duration were modelled from a care home level dataset consisting of an observation per outbreak. Risk factors were modelled using multivariable negative binomial regression with a facility-level frailty term and no exposure term. As by definition the minimum outbreak size is two and the minimum outbreak duration is one, these values were subtracted from the outcomes for analysis, as negative binomial regression should be applied for outcomes taking integer values from zero upwards. For time-varying facility-level covariates (see Table 6.1), observations from the month of the outbreak start were assigned to the outbreak. Negative binomial regression was selected as the most appropriate modelling

approach as outbreak size and duration data are non-negative integer data with a negative binomial distribution and variance that is larger than the mean. Despite zero-inflation of data, zero-inflated negative binomial models were not possible in view of the large number of covariates.

Analyses are presented at the individual-level for infection incidence and at facility-level for outbreak-related outcomes. All models were adjusted for calendar month by introducing a coefficient for each month and removing the baseline default intercept term. I median-centred all continuous variables and assessed the linearity of association between these covariates and the outcome using likelihood-ratio tests that compared model fit between linear and polynomial factors. In the baseline models if covariates were non-linearly associated then I retained the polynomials in the model, as these results were not presented for interpretation by the reader. In contrast, to facilitate the interpretation of the building factor models, I did not retain polynomials for non-linearly associated building factors but categorised them into terciles.

As previously mentioned, over the analysis period there were four different dominant SARS-CoV-2 variants in circulation at different periods: Wuhan, Alpha, Delta, and Omicron. Omicron has been documented to be more transmissible than prior variants[379,380] and this is reflected in the substantial increase in the incidence of infections and outbreaks in my data, as well as the national data, from December 2021 onwards when this variant predominated. This period was also distinct from the preceding period as most of the study population had received a third SARS-CoV-2 vaccine dose. I decided to explore the impact of the Omicron variant on the associations by creating a variable to represent the Omicron-dominant period (1st December 2021-31st March 2022). Using likelihood-ratio tests, I assessed for an interaction between Omicron-dominance and variables describing host immunity (prior infection, vaccination) in the baseline model and retained significant interaction terms. I chose these variables as I had shown that host immunity, a key factor from the epidemiological triad, varies over time. I assessed for interactions with each of the building factors and planned to stratify the analyses if I found evidence of multiple interactions with Omicron-dominance.

P-values <0.05 were considered statistically significant. Analyses were conducted using Stata v17.0 in the UCL Data Safe Haven.

Sample size calculation was not possible in advance of the analysis given the rapidly changing epidemiology of the pandemic.

Full details of ethical approvals and legal basis for accessing data are outlined in Chapter 3.

6.3 Results

6.3.1 Cross-sectional survey

Of 151 questionnaires sent out, 137 were completed and 134 could be linked to a CQC-ID. 119 stated the date of completion: 56 in April, 16 in May, 7 in June, 11 in July, 6 in August, 17 in September, 5 in October, and 1 in November, 2021. 105 (78.4%) were completed by a manager, 19 (14.2%) by the maintenance officer, the rest unknown.

The mean number of bedrooms was 54.7 (SD 21.4) and mean number of storeys was 2.2 (SD 0.6). 10/126 (8.0%) reported having shared bedrooms and in 22.5% (27/120) bathrooms were shared between staff and residents (Table 6.2). 104/128 (81.2%) of homes were purpose-built and the remainder had been converted.

Table 6-2: Building survey responses and proportion completed, overall and comparing homes with no or small outbreaks only to those with larger outbreaks.

Building factor	**No. Completed (%)**	**Overall (n=134)**	**No/small outbreaks* (n=11)**	**Larger outbreaks (n=123)**	***P* (no/small vs large outbreaks)**
Number of rooms (mean, SD)					
Bedrooms	123 (91.8%)	54.65 (21.40)	38.00 (12.64)	55.66 (21.44)	0.017
Common rooms	128 (95.6%)	3.97 (2.60)	3.00 (0.93)	4.03 (2.67)	0.14
Dining rooms	130 (97.0%)	2.33 (1.20)	1.63 (0.74)	2.38 (1.21)	0.043
Kitchens	129 (96.3%)	1.53 (1.10)	1.50 (0.93)	1.53 (1.12)	0.47
Toilets	125 (93.3%)	9.34 (6.77)	5.38 (2.88)	9.61 (6.88)	0.044
Staircases	130 (97.0%)	3.37 (2.05)	2.88 (0.83)	3.40 (2.11)	0.24
Corridors	129 (96.3%)	6.39 (4.02)	6.38 (3.66)	6.39 (4.06)	0.50

Building factor	No. Completed (%)	Overall (n=134)	No/small outbreaks* (n=11)	Larger outbreaks (n=123)	*P* (no/small vs large outbreaks)
Storeys	111 (82.8%)	2.21 (0.56)	2.50 (0.58)	2.20 (0.56)	0.86
Building type	128 (95.5%)				<0.001
Purpose-built		104 (81.2%)	2 (28.6%)	102 (84.3%)	
Converted		24 (18.8%)	5 (71.4%)	19 (15.7%)	
Presence of shared bedrooms (% responses)	126 (94.0%)	10 (8.0%)	0 (0%)	10 (8.5%)	0.39
Number of shared bathrooms (between residents) (mean, SD)	102 (76.1%)	10 (6-13, 1-73)	5.60 (1.34)	10.89 (8.28)	0.08
Presence of shared toilets (staff and residents) (% responses)	120 (89.6%)	27 (22.5%)	2 (28.6%)	25 (22.1%)	0.69
Air temp (°C) (mean, SD) ±					
Dining room	38 (28.4%)	22.77 (2.66)	23.25 (2.06)	22.71 (2.75)	0.65
Common room	52 (38.8%)	22.87 (2.45)	23.25 (2.06)	22.84 (2.50)	0.62
Bedroom	32 (23.9%)	22.59 (2.86)	22.00 (1.41)	22.63 (2.95)	0.38
Perceived air quality (common room) (% responses)	115 (85.8%)				0.38
Too humid		5 (4.3%)	0 (0%)	5 (4.6%)	
Humid		9 (7.8%)	1 (16.7%)	8 (7.3%)	
Slightly humid		16 (13.9%)	1 (16.7%)	15 (13.8%)	
Just right		70 (60.9%)	2 (33.3%)	68 (62.4%)	
Slightly dry		10 (8.7%)	2 (33.3%)	8 (7.3%)	
Dry		3 (2.6%)	0 (0%)	3 (2.8%)	
Too dry		2 (1.7%)	0 (0%)	2 (1.8%)	
Perceived air quality (dining room) (% responses)	115 (85.8%)				0.77
Too humid		6 (5.2%)	0 (0%)	6 (5.5%)	
Humid		7 (6.1%)	1 (16.7%)	6 (5.5%)	
Slightly humid		18 (15.7%)	0 (0%)	18 (16.5%)	
Just right		75 (65.2%)	5 (83.3%)	70 (64.2%)	
Slightly dry		6 (5.2%)	0 (0%)	6 (5.5%)	
Dry		2 (1.7%)	0 (0%)	2 (1.8%)	
Too dry		1 (0.9%)	0 (0%)	1 (0.9%)	

Building factor	No. Completed (%)	Overall (n=134)	No/small outbreaks* (n=11)	Larger outbreaks (n=123)	*P* (no/small vs large outbreaks)
Perceived air quality (bedroom) (% responses)	113 (84.3%)				0.86
Too humid		4 (3.5%)	0 (0%)	4 (3.7%)	
Humid		7 (6.2%)	1 (16.7%)	6 (5.6%)	
Slightly humid		10 (8.7%)	0 (0%)	10 (9.3%)	
Just right		82 (72.6%)	5 (83.3%)	77 (72.0%)	
Slightly dry		6 (5.3%)	0 (0%)	6 (5.6%)	
Dry		3 (2.7%)	0 (0%)	3 (2.8%)	
Too dry		1 (0.9%)	0 (0%)	1 (0.9%)	
Cleaning frequency - vacuuming (% responses)	111 (82.8%)				0.93
Daily		108 (97.3%)	5 (100%)	103 (97.2%)	
Several times a week		2 (1.8%)	0 (0%)	2 (1.9%)	
Weekly		1 (0.9%)	0 (0%)	1 (0.9%)	
Several times a month		0 (0)	0 (0%)	0 (0%)	
Monthly		0 (0)	0 (0%)	0 (0%)	
Cleaning frequency - washing floor (% responses)	108 (80.6%)				0.81
Daily		91 (84.3%)	4 (80.0%)	87(84.5%)	
Several times a week		8 (7.4%)	1 (20.0%)	7 (6.8%)	
Weekly		7 (6.5%)	0 (0%)	7 (6.8%)	
Several times a month		1 (0.9%)	0 (0%)	1 (1.0%)	
Monthly		1 (0.9%)	0 (0%)	1 (1.0%)	
Cleaning frequency – sweeping (% responses)	105 (78.4%)				0.96
Daily		103 (98.1%)	4 (100%)	99 (98.0%)	
Several times a week		1 (1.0%)	0 (0%)	1 (1.0%)	
Weekly		1 (1.0%)	0 (0%)	1 (1.0%)	
Several times a month		0 (0)	0 (0%)	0 (0%)	
Monthly		0 (0)	0 (0%)	0 (0%)	
Ventilation type - dining room (% responses) [#]	54 (40.3%)				0.91
Central air conditioning		29 (53.7%)	0 (0%)	9 (17.6%)	
Cassette ceiling unit		2 (3.7%)	0 (0%)	2 (3.9%)	
Portable unit exhaust pipe		1 (1.9%)	0 (0%)	1 (2.0%)	

Building factor	**No. Completed (%)**	**Overall (n=134)**	**No/small outbreaks* (n=11)**	**Larger outbreaks (n=123)**	***P* (no/small vs large outbreaks)**
Mechanical extraction unit		9 (16.7%)	1 (33.3%)	8 (15.7%)	
Freestanding		9 (16.4%)	2 (66.7%)	27 (52.9%)	
Unknown		4 (7.4%)	0 (0%)	4 (7.8%)	
Ventilation type - common room (% responses) [#]	67 (50.0%)				0.21
Central air conditioning		8 (11.9%)	0	8	
Cassette ceiling unit		9 (13.4%)	1	8	
Portable unit exhaust pipe		2 (3.0%)	1	1	
Mechanical extraction unit		8 (11.9%)	0	8	
Freestanding		35 (52.2%)	3	32	
Unknown		5 (7.5%)	0	5	
Ventilation type – bedroom (% responses) [#]	52 (38.8%)				0.78
Central air conditioning		32 (61.5%)	0 (0%)	4 (8.0%)	
Cassette ceiling unit		3 (5.8%)	0 (0%)	3 (6.0%)	
Portable unit exhaust pipe		0 (0)	0 (0%)	0 (0%)	
Mechanical extraction unit		9 (17.3%)	1 (50%)	8 (16%)	
Freestanding		4 (7**.7**%)	1 (50%)	31 (62%)	
Unknown		4 (7.7%)	0 (0%)	4 (8%)	
Heating - dining room (% responses)	128 (95.6%)				0.80
central heating		127 (99.2%)	8 (100%)	119 (99.2%)	
other		1 (0.8%)	0 (0%)	1 (0.8%)	
Heating - common room (% responses)	124 (92.6%)				0.81
central heating		123 (99.2%)	7 (100%)	116 (99.1%)	
other		1 (0.8%)	0 (0%)	1 (0.9%)	
Heating – bedroom (% total responses)	109 (81.3%)				0.81
central heating		108 (99.1%)	6 (100%)	102 (99.0%)	
other		1 (0.9%)	0 (0%)	1 (1.0%)	
Presence of humidifiers / air	20 (14.9%)	2 (10.0%)	0 (0%)	2 (11.1%)	0.62

Building factor	No. Completed (%)	Overall (n=134)	No/small outbreaks* (n=11)	Larger outbreaks (n=123)	*P* (no/small vs large outbreaks)
purifiers - dining room (% responses)					
Presence of humidifiers / air purifiers – bedroom (% responses)	15 (11.2%)	3 (20.0%)	0 (0%)	3 (21.4%)	0.61
Presence of condensation (% responses)	124 (92.6%)	12 (9.7%)	1/8 (12.5%)	11 (9.5%)	0.78
Presence of outdoor space (% responses)	124 (92.6%)	121 (97.6%)	7 (87.5%)	114 (98.3%)	0.06
Maximum people in dining room at one time (mean, SD)	94 (70.2%)	13.71 (7.49)	11.86 (4.14)	13.86 (7.69)	0.25
Maximum people in common room at one time (mean, SD)	101 (75.4%)	12.02 (8.16)	10.67 (4.89)	12.11 (8.33)	0.34

SD Standard Deviation

T-test to compare means, chi-squared to compare proportions.

*No/small outbreaks = care homes that either had no outbreaks over the study period or all outbreaks had <= 2 resident cases or had a duration of <=2 days (suggesting low transmission in home)

± In line with main analysis, values > 30°C were treated as nulls.

MVHR not listed as no responses indicated presence of this ventilation type.

The majority cleaned every day and perceived air quality was ‘just right’ instead of dry’ or ‘humid’. Seasonal changes at the time of survey completion did not appear to affect the recorded indoor temperatures as mean temperatures were not greater in summer months (June, July) than in spring (April, May) or autumn months (September-November) although August measurements did appear higher, Table 6.3. Likewise, air quality did not appear to follow a seasonal pattern of variation, Figure 6.3. 12/124 (9.7%) reported condensation but only 12-18% of care homes with condensation reported humid conditions. Although <15% completed the questions around humidifiers or air purifiers, very few reported using them. Most care homes had outdoor space (121/124, 97.6%) and almost all used central heating (108/109, 99.1%). Over half reported ventilation type, central air conditioning was most common in dining rooms (29/54, 53.7%) and bedrooms (32/52, 61.5%), whereas freestanding fans predominated in common rooms (35/67, 52.2%) (Table 6.2). The mean maximum

number of people in common spaces was similar between the common and dining rooms (12.02 vs 13.71).

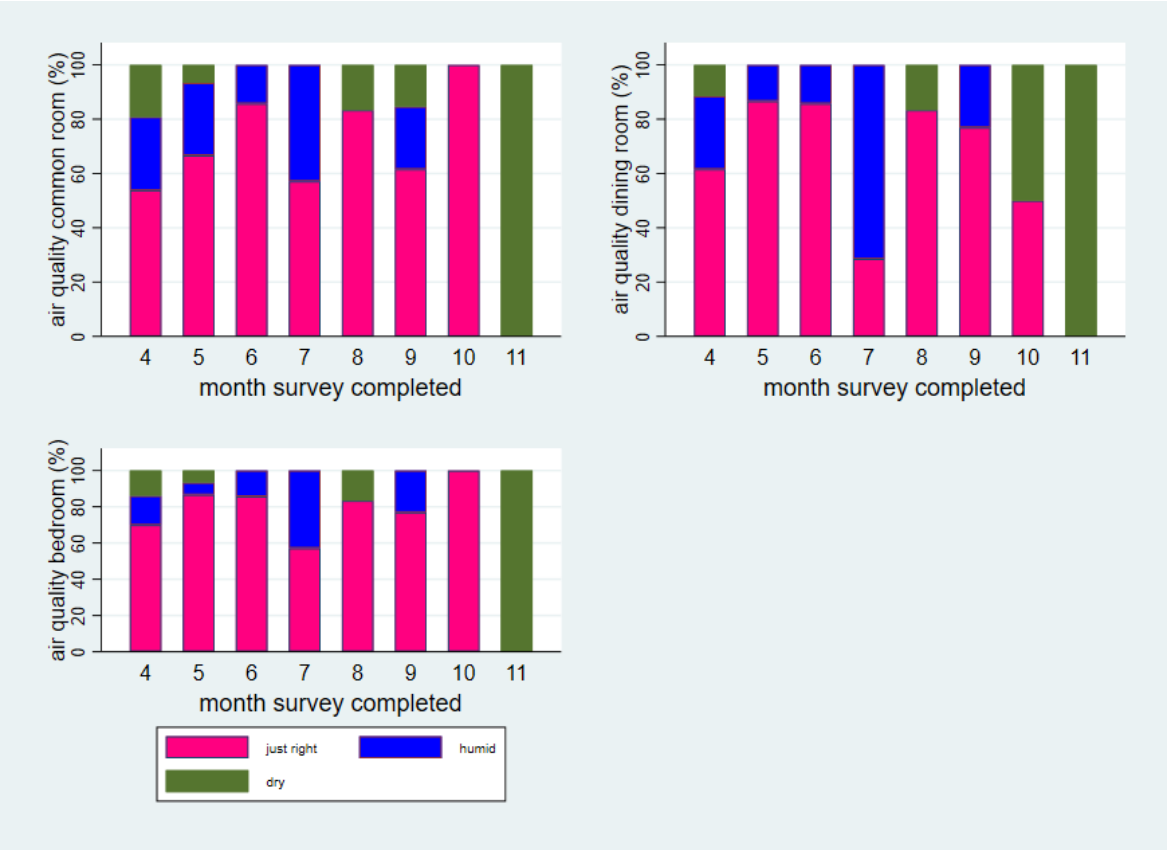
The survey responses were compared between those from care homes that had experienced no outbreaks or only outbreaks that lasted a maximum of two days or had less than or equal to two cases in the outbreak (n=11) and those with larger outbreaks (n=123). Responses appeared similar between groups however small outbreak homes were predominantly converted (7/9, 71.4%) whereas homes with larger outbreaks were more commonly purpose-built (102/121, 84.3%, $P<0.001$). In addition, mean bedroom number was lower in no/small outbreak homes (38.0 vs 55.7, $P=0.017$) as was the mean number of toilets (5.4 vs 9.6, $P=0.044$) which is likely to reflect the overall care home size, Table 6.2. It is possible that this difference was seen because smaller homes were less likely to use agency staff (a known risk factor for outbreaks),[154] as rota gaps were covered internally and may have been less likely to accept new admissions from hospital.

Table 6-3: Mean indoor temperatures and number of responses reported by month of survey completion in dining room, common room, and bedroom.

Month of completion	Dining room (°C)	No. responses	Common room (°C)	No. responses	Bedroom (°C)	No. responses
April	22.6 (2.0)	14	23 (2.0)	18	22 (2.0)	12
May	20.3 (1.5)	3	20.8 (1.5)	4	19.5 (0.7)	2
June	21 (0)	1	23 (2.8)	2	-	0
July	22 (1)	3	21.5 (1.0)	6	22 (1)	3
August	26 (2.8)	2	25 (2.6)	3	26 (2.8)	2
September	22.9 (3.5)	7	23.3 (3.4)	7	23.5 (4.7)	4
October	23 (1)	3	22 (0.8)	4	22 (0.8)	4
November	-	0	25 (0)	1	-	0

*In line with main analysis, values > 30°C were treated as nulls

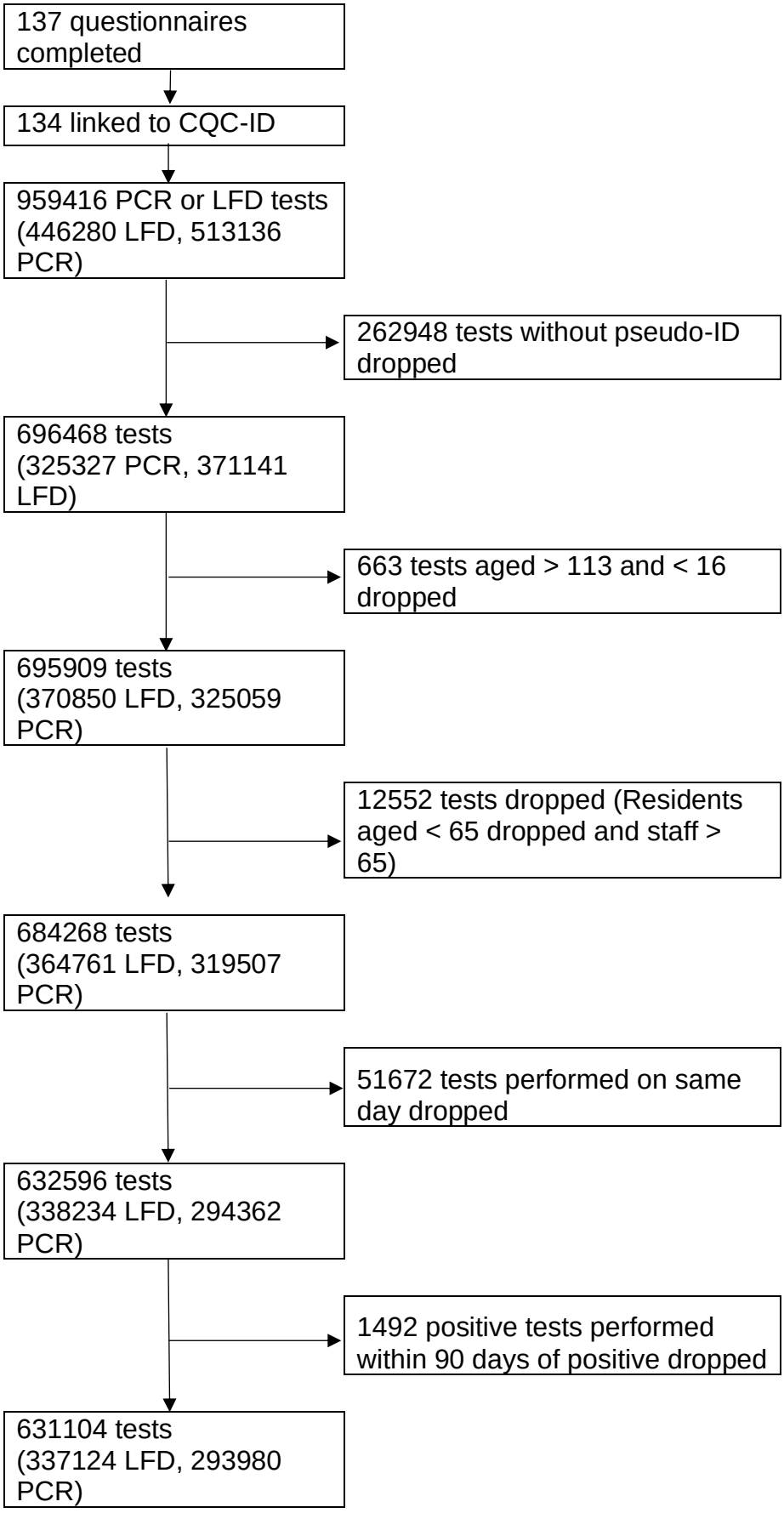
Figure 6-3: Perceived air quality reported by month of survey completion in common room, dining room, and bedroom.



6.3.2 Cohort study

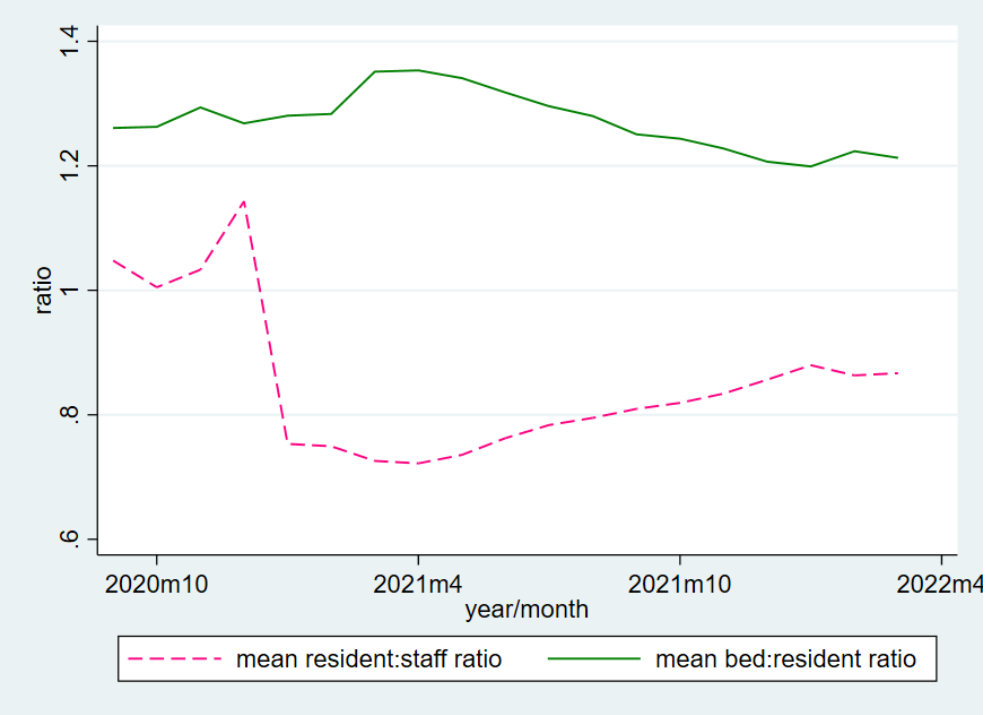
Data on infection and related outcomes were available for 13010 residents and 17766 staff of 134 care homes that had completed the surveys, based on 631104 PCR or LFD tests that were eligible for inclusion, Figure 6.4.

Figure 6-4: Inclusion flow diagram



Overall, 21140/30774 (68.7%) were female, the median age was 47 (IQR 33.6-56.9) in staff and 83.5 (74.6-90) in residents, Table 6.4. Median follow-up was 104 days (9-334) per participant, comparable between staff and residents (102 vs 106 days). Care homes were situated in all regions of England and most (116/134, 86.6%) were for-profit (belonging to two large chains). In each care home, the median number of staff was 48 (32-68) and median number of beds was 50.5 (42-66).

Figure 6-5: Mean resident-to-staff and bed-to-resident ratio over study period.



The median bed-to-resident ratio was 1.2 (1.1-1.4) with a peak seen in April 2021 that coincided with the decrease in mortality nationally. Median resident-to-staff ratio was above one until December 2020 and then fell below one from January 2021 where it remained for the remainder of the study period, probably reflecting an increase in number of staff at this point or high rates of staff sickness in the final months of 2020 and increasing immunity from 2021 onwards from vaccination and infection, Figure 6.5. A median of 73.8% (52.7-85.7%) of resident beds were funded by the Local Authority and 22.9% (0.0-50.0%) were funded for dementia care, Table 6.4.

Table 6-4: Baseline demographics a) individual-level b) care home level

a)

Baseline demographics	**Number (%)**
Number participants	30,774
Staff	17,766 (57.7)
Residents	13,008 (42.3)
Sex	
Male	9,567 (31.1)
Female	21,140 (68.7)
Unknown	68 (0.2)
Age (median, IQR, range)	60 (43-80.6, 16-110.8)
Staff	47 (33.6-56.9, 16-65)
Residents	83.5 (74.6-90, 64-110.8)

b)

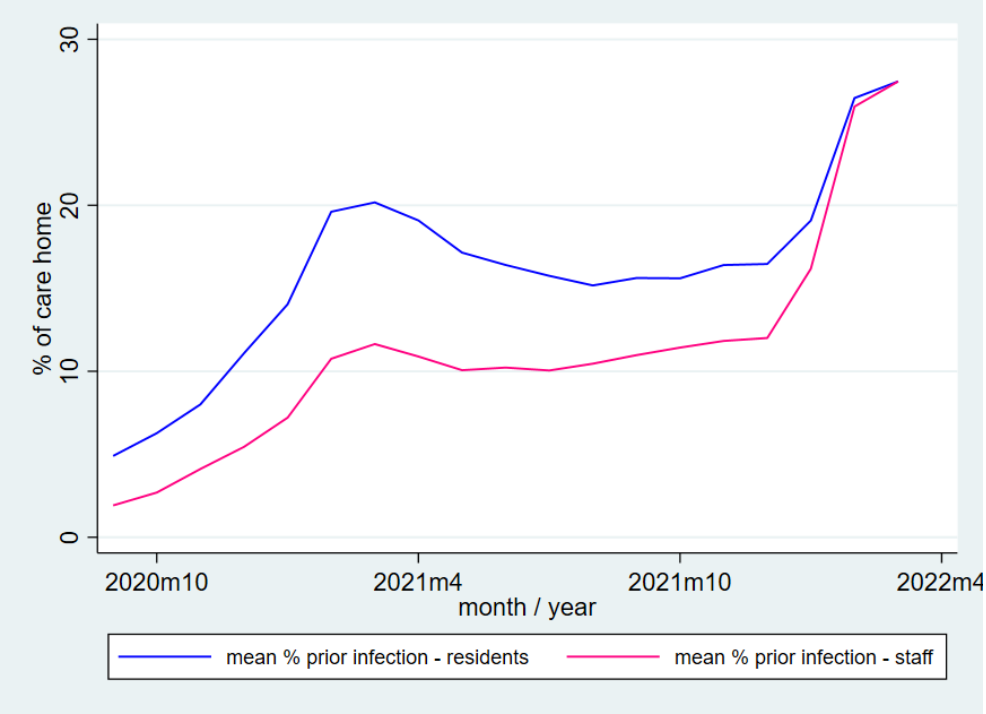
	Number (%) Median (IQR, range)
Number of care homes	134
Region	
London	11 (8.2)
South-East	17 (12.6)
East of England	11 (8.2)
South-West	14 (10.4)
North-West	20 (14.8)
North-East	17 (12.6)
East Midlands	23 (17.0)
West Midlands	11 (8.2)
Yorkshire & Humber	11 (8.2)
IMD index (median, range)	5 (3-8, 1-10)
Care home type	
For-profit	116 (86.6)
Not-for-profit	18 (13.4)
Total staff±	48 (32-68, 0-189)
Total beds±	50.5 (42-66, 7.3-123)
Staff: resident ratio±	0.8 (0.7-1.0, 0.3-2.6)
Bed: resident ratio±	1.2 (1.1-1.4, 1-4.9)
Proportion LA funded beds±	73.8 (52.7-85.7, 0-100)
Proportion dementia beds±	22.9 (0-50, 0-100)
Staff vaccination coverage±# (%)	75.6 (0-92.9, 0-100)
Resident vaccination coverage±# (%)	88.4 (0-96.4, 0-100)
Proportion staff with prior infection± (%)	7.9 (0-17.4, 0-100)

Proportion residents with prior infection± (%)	11.1 (3.3-24.4, 0-100)

± Per care home # receipt of two vaccine doses
IMD index LA Local Authority
Index of Multiple Deprivation[374] – ranges from 1 to 10, 1 is most deprived and 10 is least.

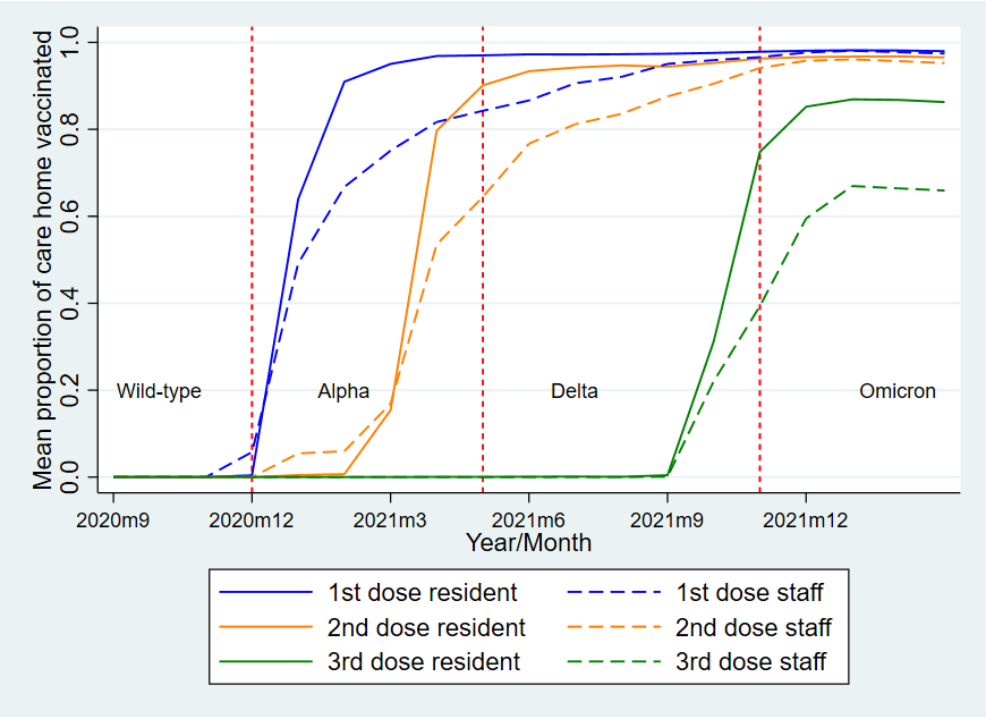
The median proportion of staff with prior infection (based on prior PCR/LFD or antibody positivity) was lower than for residents (7.9% vs 11.1%). This proportion rose over time, with a substantial increase over the Omicron period, Figure 6.6. Figure 6.7 illustrates how vaccination coverage of successive vaccine doses increased over the study period. Average second vaccine dose coverage within the care home exceeded 95% among both staff and residents from December 2021, when full vaccination was mandated in staff.[381]

Figure 6-6: Mean proportion of care home with evidence of prior infection, by staff and resident (1st September 2020 – 31st March 2022).



*Based on prior positive PCR/LFD test or anti-nucleocapsid antibody.

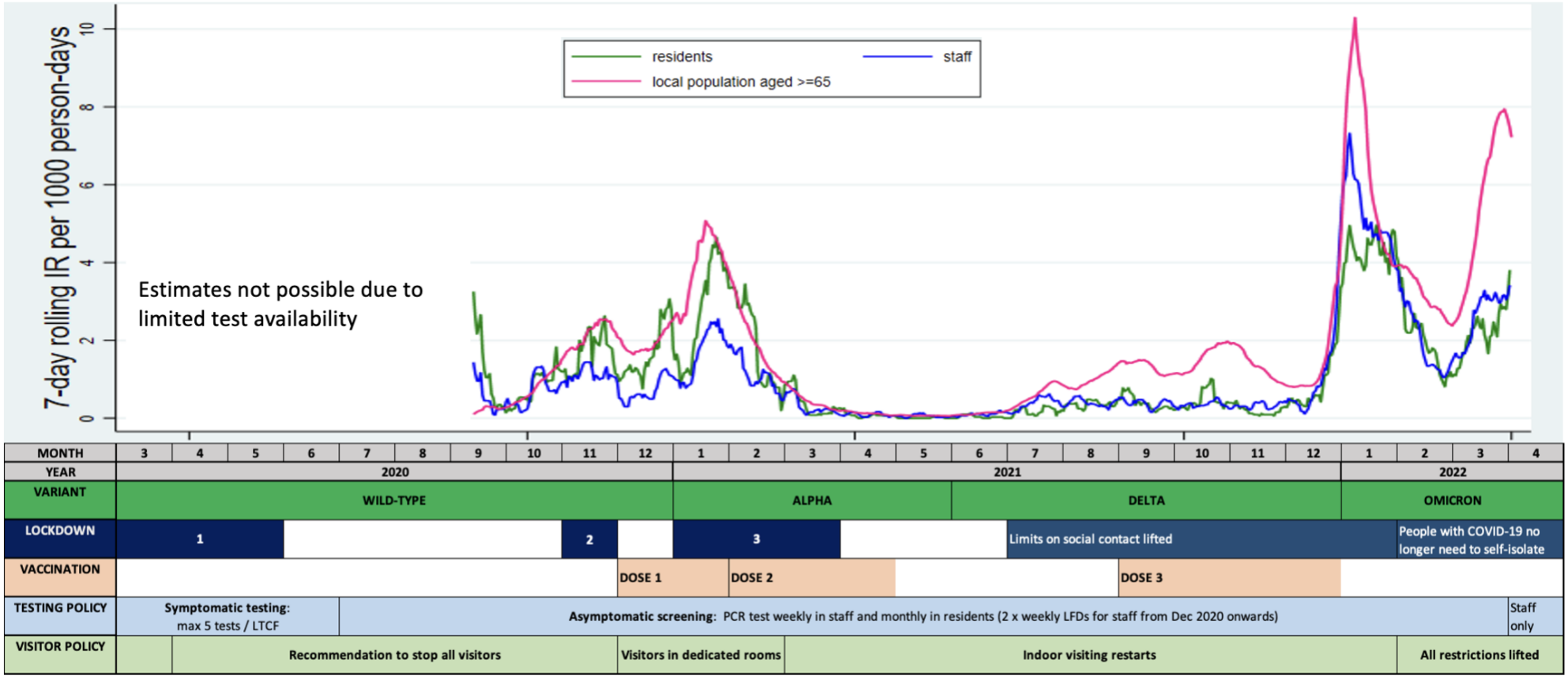
Figure 6-7: Mean proportion of the care homes that received 1, 2, and 3 vaccination doses, by staff and resident (1st September 2020 – 31st March 2022).



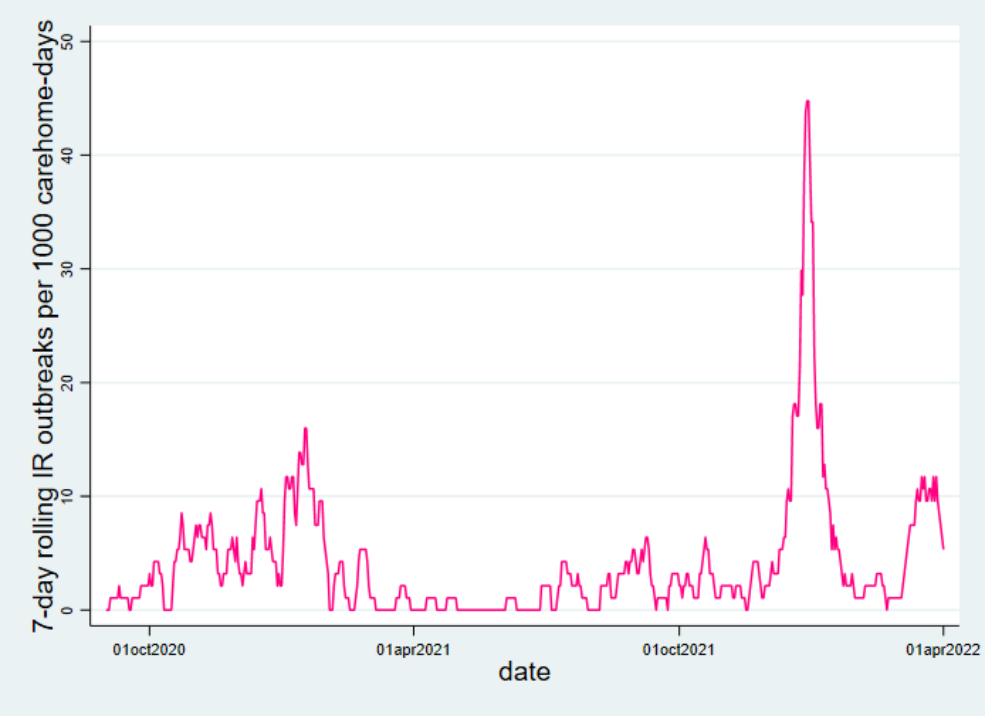
The seven-day rolling SARS-CoV-2 incidence rates among residents and staff followed similar trajectories and reflected national trends amongst over 65-year-olds. Although infection rates in residents were lower than in staff in the Alpha-associated peak between October 2020 and February 2021, this relationship was reversed in the Omicron-associated peak between December 2021 and March 2022, Figure 6.8. The peak seen over the Omicron period was greater than preceding peaks, likely reflecting the increased transmissibility of the variant. Incidence rate in community-dwelling over 65-year-olds that lived in the same areas as the participating care homes appeared greater than those seen in residents from July 2021 onwards. This may be related to higher levels of natural immunity and vaccination coverage among care home residents (where it was very difficult to shield residents from infection), as well as preventive measures that remained in place in care homes as national lockdown restrictions eased, although it is difficult to draw definitive conclusions, Figure 6.8a. Outbreak incidence followed similar trends to infection incidence, with a much larger peak over the Omicron-dominant period when compared with the period of Alpha-predominance, Figure 6.8b.

Figure 6-8: 7-day rolling incidence rate of a) SARS-CoV-2 infection amongst staff and residents compared with local SARS-CoV-2 incidence amongst adults > 65 years in the local community[382], with timeline of SARS-CoV-2 policy changes, b) outbreaks in participating care homes (1st March 2020 – 31st March 2022).

a)



b)



IR incidence rate

Policy timeline is a summary of key policies, outlined in Figure 1.7

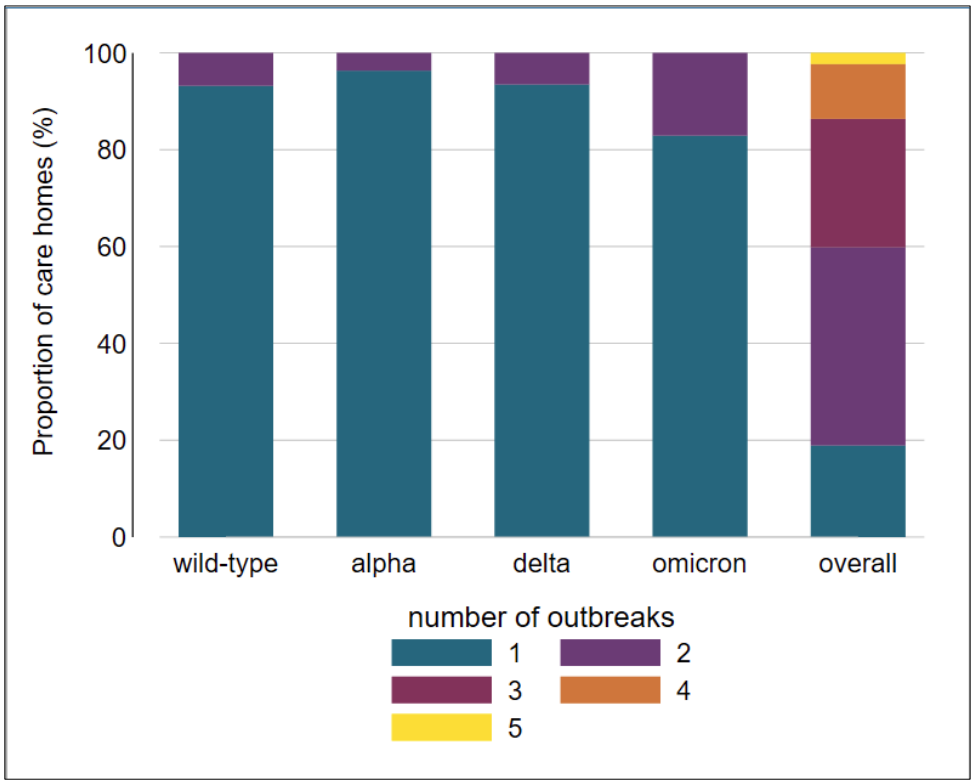
313 outbreaks occurred over the analysis period and the median number per care home was two (IQR 2-3) with a maximum of five, although outbreak characteristics varied over time, Table 6.5 & Figure 6.9. The number of outbreaks, outbreak duration and outbreak size were all greater in the Omicron period compared with preceding periods.

Table 6-5: Characteristics of outbreaks according to dominant variant

Outbreak characteristics (mean, SD)	**Wild type**	**Alpha**	**Delta**	**Omicron**
Number of outbreaks	63	57	49	144
Outbreak size	23.4 (27.5)	15.4 (14.5)	15.6 (25.2)	31.2 (24.1)
- **Residents**	12.5 (14.2)	7.8 (7.6)	6.9 (10.2)	12.4 (10.0)
- **Staff**	9.9 (14.1)	6.6 (7.9)	7.7 (15.5)	17.8 (15.1)
Outbreak duration, days	29.2 (25.4)	18.7 (14.7)	31.7 (36.3)	57.4 (42.1)

Wild type interval 01/09/2020-31/12/2020; Alpha interval 01/01/2021 – 31/05/2021; Delta interval 01/06/2021-31/11/2021; Omicron interval 01/12/2021-31/03/2022

Figure 6-9: Number of outbreaks per care home by dominant variant and overall.



Wild type interval 01/09/2020-31/12/2020; Alpha interval 01/01/2021 – 31/05/2021; Delta interval 01/06/2021-31/11/2021; Omicron interval 01/12/2021-31/03/2022

6.3.3 Risk factors for infection ingress and transmission

Baseline models for all four outcomes are not presented here as these are not the focus of the analysis, however, can be found in the appendix 6.3. Statistically significant interactions with Omicron-dominance were retained in the baseline model of resident infection incidence with individual-level prior infection status, facility-level proportion of staff and residents with prior infection, and facility-level staff vaccination coverage. In the model of outbreak size, interactions between Omicron and facility-level proportion of previously infected residents were retained. There were no interactions retained in the outbreak incidence or the outbreak duration models (Appendix 6.3, Tables S1a-d).

Building characteristics excluded from the model due to low response rate (<20% complete) or lack of variability in responses (>90% the same, defined under Section 6.2.2) were: presence of shared bedrooms, vacuuming and sweeping frequency, heating type, presence of humidifiers, and presence of outdoor space. Significant associations with building characteristics from the survey are summarised below. Only

statistically significant differences from the reference category are described for categorical variables, however as *P*-values are calculated for the trend across the variable, these have not been presented in the text.

Built environment factors associated with rate of infection in residents.

For the primary outcome of incidence of resident infection (a measure of both infection ingress and transmission), the overall baseline model had statistically significant associations with age, local SARS-CoV-2 incidence, individual-level vaccination status, individual-level infection status, bed-to-resident ratio, resident-to-staff ratio, total number of staff, and proportion of residents with prior infection (Table S1a). 14/22 building factors had a significant interaction with Omicron period; therefore, I have presented the unstratified analysis and have also stratified it according to pre-Omicron and Omicron-dominant period, Figure 6.10, Table 6.6a.

In the overall (unstratified) analysis, a greater number of storeys were associated with lower infection rate (aIRR 0.64 per storey, 95% CI 0.43-0.97, *P*=0.036). Factors associated with an increased infection rate were buildings that had been purpose-built compared with converted (1.99, 1.08-3.69 *P*=0.028) and those with a greater number of bedrooms (1.04 per extra bedroom, 1.08-3.69, *P*<0.001) (Table 6.6a, Figure 6.10).

In the stratified analysis, in the pre-Omicron period, care homes with a greater number of storeys (aIRR 0.51 per storey, 95% CI 0.28-0.94, *P*=0.030), and bedrooms (1.04, 1.01-1.07, *P*=0.006) remained significantly associated with infection rate (as found in the overall analysis). Care homes that ventilated dining rooms using portable units with exhaust pipes vs central air conditioning were at increased risk of infection in residents (9.35, 1.06-82.67) however confidence intervals were wide meaning low precision. Ventilating the dining room using a cassette ceiling unit compared with central air conditioning (0.05, 0.00-0.57) and having fewer people in the dining room at one time (0.94, 0.89-0.99, *P*=0.032) were significantly associated with lower infection rate. In the Omicron-dominant period, factors that retained significant association with infection rate were purpose-built buildings compared with converted ones (2.92, 1.36-6.25, *P* =0.006). Washing the floor on a daily compared with less than daily basis was associated with an increased infection risk (2.38, 1.03-5.52, *P*=0.043),

and a 1% increase in LA-funding reduced this risk by 2% (1.02, 1.00-1.03, *P*=0.024), Table 6.6a, Figure 6.10.

Built environment factors associated with incidence of outbreaks.

I also considered the relationship between the other primary outcome, incidence of outbreaks which I considered to be a measure of infection ingress, and the building factors. There was only one factor significantly associated with this outcome in the baseline model, the local incidence of SARS-CoV-2 (*P*<0.001) which was non-linearly associated (Table S1b). The adjusted IRR for outbreak events comparing a high (75th centile: 0.48 cases/100 population) vs low (25th centile: 0.09 cases/100 population) local incidence was 2.84 (95% CI 1.85-4.36, *P*<0.001). None of the building factors had significant associations with this outcome and there were no interactions with Omicron period in the baseline or building factor models, Table 6.6b, Figure 6.10.

Built environment factors associated with outbreak size and duration.

The secondary outcomes that I chose described transmission of infection within the facility. In the baseline model, factors significantly associated with outbreak size were IMD decile (10 vs 1), bed-to-resident ratio, total residents and staff, staff vaccination coverage, and the proportion of staff with prior infection in the pre-Omicron period (Table S1c). In the overall outbreak size risk factor analysis, the only building factors significantly associated with this outcome was perceived air quality in the common room – 'dry' compared with 'just right' air was associated with a larger outbreak (aIRR 1.46, 1.00-2.13).

As there was only one building factor that had significant interactions with Omicron-dominance, the outbreak size analysis was not stratified, Table 6.6c. In the pre-Omicron period, ventilation in the dining room that used portable unit exhaust pipes or if the ventilation type was unknown when compared with central air conditioning were associated with larger outbreak size (aIRR 7.29, 95% CI 2.23-23.83 and aIRR 3.36, 95% CI 1.81-6.26 respectively) although confidence intervals are wide and the "unknown" ventilation type suggests some care homes may have experienced challenges completing surveys, Tables 6.6b&c, Figure 6.10.

For the outbreak duration outcome, which was also considered as a measure of infection transmission within the facility, the only factors in the baseline model that were significantly associated were number of staff and the vaccination coverage amongst residents (supplementary appendix, Table S1d).

In the overall analysis of outbreak duration, there were three significant associations with building factors. The outbreak duration was reduced by 1% with every 1% increase in the number of LA-funded beds and by 2% with every one-person increase in the maximum number of people in the dining room (aIRR 0.99, 95% CI 0.99-1.00, *P* =0.016 and aIRR 0.98, 95% CI 0.96-0.99, *P* =0.009). An increase of 1°C in bedroom temperature increased the outbreak duration by 15% (1.15, 1.01-1.32, *P* =0.033). There were no significant interactions between any of the building factors and Omicron period (Tables 6.6b&c, Figure 6.10).

Table 6-6:

a) *Mixed effects adjusted individual-level Poisson regression models of incidence of infection in a resident[±], overall and stratified by pre-Omicron and Omicron periods.*

b) *Mixed effects adjusted[†] facility-level models of incidence of outbreak (Poisson model), size of outbreak (negative binomial model), and duration of outbreak (negative binomial model), overall.*

c) *Mixed effects adjusted facility-level model of size of outbreak (negative binomial model)[†], reporting associations that differ by pre-Omicron and Omicron periods for factors (significant interaction).*

	Unstratified				Stratified – Pre-Omicron				Stratified – Omicron			
Building Factors	**aIRR**	***P***	**95% CI**		**aIRR**	***P***	**95% CI**		**aIRR**	***P***	**95% CI**	
LA beds (%)	1.01	0.59	0.99	1.01	1.00	0.72	0.99	1.02	1.02	0.024	1.00	1.03
Dementia beds (%)	1.00	0.21	1.00	1.02	1.00	0.37	0.99	1.01	1.01	0.22	0.97	1.02
No. storeys	0.64	0.036	0.43	0.97	0.51	0.030	0.28	0.94	0.85	0.56	0.50	1.45
Purpose built vs converted	1.99	0.028	1.08	3.69	1.06	0.90	0.46	2.42	2.92	0.006	1.36	6.25
No. bedrooms*	1.04	<0.001	1.02	1.06	1.04	0.006	1.01	1.07	1.02	0.08	1.00	1.05
No. common rooms*	1.01	0.83	0.91	1.12	1.04	0.59	0.90	1.20	0.98	0.79	0.86	1.12
No. dining rooms*	1.09	0.48	0.87	1.36	1.04	0.79	0.77	1.40	0.98	0.87	0.73	1.31
Presence of shared bathrooms (staff with residents)	0.75	0.30	0.43	1.30	0.55	0.12	0.25	1.17	0.96	0.92	0.48	1.95
Presence of shared bathrooms (between residents)	1.23	0.73	0.38	3.97	1.51	0.61	0.32	7.17	1.28	0.74	0.30	5.39
Dining room temperature*[#]	0.97	0.73	0.81	1.16	1.00	0.91	0.85	1.21	1.13	0.36	0.87	1.48
Common room temperature*[#]	0.96	0.58	0.81	1.12	1.02	0.83	0.84	1.24	0.96	0.76	0.74	1.25

	Unstratified				Stratified – Pre-Omicron				Stratified – Omicron			
Building Factors	**aIRR**	***P***	**95% CI**		**aIRR**	***P***	**95% CI**		**aIRR**	***P***	**95% CI**	
Bedroom temperature*#	1.14	0.25	0.91	1.43	1.19	0.19	0.92	1.56	1.19	0.09	0.98	1.45
Max people in common room*"	0.97	0.15	0.94	1.01	-	-	-	-	0.99	0.49	0.95	1.03
Low	-	-	-	-	Ref	0.62			-	-	-	-
Medium	-	-	-	-	0.67		0.26	1.76	-	-	-	-
High	-	-	-	-	0.65		0.26	1.66	-	-	-	-
Max people in dining room*	0.99	0.63	0.95	1.03	0.94	0.032	0.89	0.99	1.00	0.86	0.96	1.05
Washing floor frequency												
Less than daily	Ref				Ref				Ref			
Daily	1.63	0.16	0.83	3.22	1.25	0.64	0.49	3.17	2.38	0.043	1.03	5.52
Air quality - common room												
Just right	Ref	0.26			Ref	0.69			Ref	0.32		
Humid	0.61	-	0.34	1.10	0.74	-	0.33	1.64	0.59	-	0.29	1.18
Dry	0.76	-	0.36	1.62	1.08	-	0.39	2.94	0.79	-	0.32	1.93
Air quality - dining room												
Just right	Ref	0.41			Ref	0.98			Ref	0.22		
Humid	0.77	-	0.44	1.34	0.92	-	0.44	1.94	0.73	-	0.39	1.38
Dry	1.40	-	0.56	3.49	0.95	-	0.28	3.22	1.84	-	0.66	5.14
Air quality - bedroom												
Just right	Ref	0.86			Ref	0.58			Ref	0.61		
Humid	0.93	-	0.49	1.76	1.56	-	0.67	3.67	0.79	-	0.38	1.64
Dry	1.22	-	0.52	2.87	0.99	-	0.31	3.14	1.37	-	0.52	3.62

	Unstratified				Stratified – Pre-Omicron				Stratified – Omicron			
Building Factors	**aIRR**	***P***	**95% CI**		**aIRR**	***P***	**95% CI**		**aIRR**	***P***	**95% CI**	
Ventilation - common room												
Freestanding fan	Ref	0.56			Ref	0.97			Ref	0.42		
Cassette ceiling unit	0.84	-	0.33	2.11	0.79	-	0.24	2.57	1.01	-	0.23	4.46
Portable unit exhaust pipe	1.00	-	0.17	5.97	1.04	-	0.12	9.32	1.20	-	0.07	21.34
Mechanical extract units	1.99	-	0.78	5.07	1.15	-	0.37	3.62	6.37	-	1.14	35.52
Central air conditioning	1.67	-	0.58	4.79	0.95	-	0.25	3.57	0.80	-	0.14	4.60
Unknown	0.69	-	0.21	2.20	0.55	-	0.13	2.39	1.16	-	0.18	7.57
Ventilation - dining room												
Central air conditioning	Ref	0.08			Ref	0.037			Ref	0.68		
Cassette ceiling unit	0.37	-	0.11	1.19	0.05	-	0.00	0.57	0.65	-	0.09	4.70
Portable unit exhaust pipe	4.98	-	0.87	28.62	9.35	-	1.06	82.67	1.28	-	0.05	34.88
Mechanical extract units	1.26	-	0.67	2.35	0.64	-	0.28	1.47	2.22	-	0.71	6.90
Freestanding fan	1.74	-	0.86	3.50	0.86	-	0.35	2.13	1.65	-	0.45	6.03
Unknown	1.15	-	0.46	2.89	1.91	-	0.58	6.33	0.72	-	0.14	3.76
Ventilation - bedroom												
Central air conditioning	Ref	0.41			Ref	0.10			Ref	0.83		
Cassette ceiling unit	1.10	-	0.39	3.09	1.37	-	0.32	5.76	0.73	-	0.19	2.84
Mechanical extract units	1.86	-	0.98	3.56	2.28	-	1.03	5.05	1.27	-	0.49	3.26
Freestanding fan	1.29	-	0.43	3.88	0.38	-	0.09	1.64	1.34	-	0.28	6.50
Unknown	0.84	-	0.32	2.21	2.29	-	0.69	7.56	0.61	-	0.17	2.20

b)

	Incidence of outbreaks				Outbreak size				Outbreak duration			
Building Factors	**aIRR**	***P***	**95% CI**		**aIRR**	***P***	**95% CI**		**aIRR**	***P***	**95% CI**	
LA beds (%)	1.00	0.31	1.00	1.01	1.00	0.36	0.99	1.00	0.99	0.016	0.99	1.00
Dementia beds (%)	1.00	0.74	1.00	1.01	1.00	0.56	1.00	1.00	1.00	0.68	1.00	1.00
No. storeys	0.90	0.37	0.71	1.14	0.91	0.39	0.73	1.13	0.98	0.83	0.78	1.22
Purpose built vs converted	1.10	0.59	0.78	1.56	1.16	0.36	0.84	1.57	0.90	0.50	0.66	1.22
No. bedrooms*	1.00	0.94	0.99	1.01	1.00	0.31	1.00	1.02	1.00	0.99	0.99	1.01
No. common rooms*	0.97	0.34	0.91	1.03	1.00	0.93	0.95	1.05	1.02	0.47	0.97	1.08
No. dining rooms*	1.01	0.85	0.89	1.15	1.00	0.99	0.90	1.12	1.06	0.29	0.95	1.19
Presence of shared bathrooms (staff with residents)	0.84	0.31	0.60	1.18	0.93	0.61	0.69	1.24	0.89	0.47	0.66	1.21
Presence of shared bathrooms (between residents)	1.12	0.74	0.56	2.24	0.76	0.37	0.43	1.37	0.66	0.18	0.35	1.22
Dining room temperature*#	1.06	0.41	0.93	1.20	1.10	0.11	0.98	1.23	1.00	0.79	0.90	1.12
Common room temperature*#	1.04	0.46	0.93	1.17	1.05	0.36	0.95	1.16	1.05	0.34	0.95	1.17
Bedroom temperature*#	1.11	0.19	0.95	1.30	1.03	0.61	0.91	1.17	1.15	0.033	1.01	1.32
Max people in common room*	1.00	0.69	0.98	1.01	0.99	0.20	0.97	1.01	0.99	0.16	0.97	1.00
Max people in dining room*	1.00	0.75	0.98	1.02	1.00	0.96	0.98	1.02	0.98	0.009	0.96	0.99
Washing floor frequency												
Less than daily	Ref				Ref				Ref			
Daily	1.20	0.34	0.82	1.76	1.20	0.31	0.84	1.71	1.24	0.25	0.86	1.77

	Incidence of outbreaks				Outbreak size				Outbreak duration			
Building Factors	**aIRR**	***P***	**95% CI**		**aIRR**	***P***	**95% CI**		**aIRR**	***P***	**95% CI**	
Air quality - common room												
Just right	Ref	0.75			Ref	0.036			Ref	0.94		
Humid	0.96	-	0.69	1.32	0.89	-	0.67	1.17	0.95	-	0.71	1.28
Dry	0.85	-	0.56	1.29	1.46	-	1.00	2.13	0.98	-	0.65	1.48
Air quality - dining room												
Just right	Ref	0.94			Ref	0.22			Ref	0.96		
Humid	0.99	-	0.73	1.33	0.88	-	0.68	1.15	1.01	-	0.76	1.33
Dry	0.92	-	0.56	1.49	1.28	-	0.83	1.99	1.07	-	0.66	1.74
Air quality - bedroom												
Just right	Ref	0.27			Ref	0.53			Ref	0.30		
Humid	1.31	-	0.93	1.84	0.88	-	0.66	1.18	1.08	-	0.80	1.47
Dry	0.97	-	0.61	1.54	1.12	-	0.73	1.71	0.74	-	0.47	1.17
Ventilation - common room												
Freestanding fan	Ref	0.69			Ref	0.13			Ref	0.51		
Cassette ceiling unit	0.82	-	0.43	1.57	1.57	-	0.95	2.59	1.36	-	0.81	2.30
Portable unit exhaust pipe	1.58	-	0.54	4.69	1.53	-	0.66	3.54	1.54	-	0.67	3.50
Mechanical extract units	1.32	-	0.75	2.32	1.14	-	0.72	1.79	1.17	-	0.76	1.79
Central air conditioning	1.36	-	0.69	2.70	1.91	-	1.10	3.31	1.17	-	0.72	1.89
Unknown	1.06	-	0.52	2.18	1.41	-	0.81	2.46	1.62	-	0.92	2.84
Ventilation - dining room												
Central air conditioning	Ref	0.72			Ref	0.08			Ref	0.05		
Cassette ceiling unit	0.72	-	0.16	3.32	1.25	-	0.55	2.83	0.70	-	0.26	1.87

	Incidence of outbreaks				Outbreak size				Outbreak duration			
Building Factors	**aIRR**	***P***	**95% CI**		**aIRR**	***P***	**95% CI**		**aIRR**	***P***	**95% CI**	
Portable unit exhaust pipe	3.27	-	0.65	16.35	2.74	-	0.98	7.62	1.84	-	0.63	5.41
Mechanical extract units	0.83	-	0.45	1.50	1.00	-	0.67	1.48	1.17	-	0.77	1.79
Freestanding fan	0.97	-	0.50	1.86	1.30	-	0.84	2.02	1.27	-	0.79	2.06
Unknown	1.26	-	0.53	2.97	1.88	-	1.13	3.11	2.35	-	1.34	4.10
Ventilation - bedroom												
Central air conditioning	Ref	0.72			Ref	0.16			Ref	0.79		
Cassette ceiling unit	0.82	-	0.27	2.49	2.04	-	0.85	4.89	1.40	-	0.57	3.44
Mechanical extract units	1.10	-	0.63	1.90	1.50	-	0.95	2.38	1.18	-	0.72	1.92
Freestanding fan	0.50	-	0.18	1.44	0.93	-	0.40	2.13	0.93	-	0.37	2.32
Unknown	1.00	-	0.45	2.21	1.52	-	0.84	2.75	1.35	-	0.72	2.52

c)

	Outbreak size – Pre-Omicron				Outbreak size - Omicron			
Building Factors	**aIRR**	***P***	**95% CI**		**aIRR**	***P***	**95% CI**	
Ventilation - dining room								
Central air conditioning	Ref	0.0002	-	-	Ref	0.82	-	-
Cassette ceiling unit[§]	-	-	-	-	1.39	-	0.64	3.03
Portable unit exhaust pipe	7.29	-	2.23	23.83	0.87	-	0.21	3.60
Mechanical extract units	1.16	-	0.70	1.93	0.83	-	0.53	1.29

Freestanding fan	1.25	-	0.76	2.07	1.23	-	0.73	2.07
Unknown	3.36	-	1.81	6.26	0.99	-	0.52	1.87

Models adjusted for variables in baseline models shown in Supplementary Appendix tables S1a-S1d, interaction terms between Omicron period and prior immunity / vaccination variables retained in baseline models where statistically significant.

Models presented in table 6.6a include frailty terms at individual and care home level. Models in tables 6.6b and 6.6c include frailty term at care home level only.

*Median-centred

Per °C increase. Temperatures >30°C dropped from analysis.

± adjusted for individual-level: age, prior infection, receipt of 2nd vaccine, sex; facility-level: IMD, local SARS-CoV-2 incidence rate, for-profit status, number of beds, number of staff, number of residents, bed-to-resident ratio, resident-to-staff ratio, proportion residents with prior infection, proportion staff with prior infection, proportion staff vaccinated, proportion residents vaccinated.

† Adjusted for facility-level: median age in residents, proportion females amongst residents, IMD, local SARS-CoV-2 incidence rate, for-profit status, number of beds, number of staff, number of residents, bed-to-resident ratio, resident-to-staff ratio, proportion residents with prior infection, proportion staff with prior infection, proportion staff vaccinated, proportion residents vaccinated.

“Non-linearly associated continuous variables presented as categorical variables in terciles.

§ No outbreaks occurred in the pre-Omicron period in care homes with cassette ceiling unit

Figure 6-10: Heat map of building factors associated with outcomes overall and stratified into Pre-Omicron period and Omicron-dominant period.

Risk factors for outcomes describing ingress risk only are presented in the first column (outbreak), risk factors for transmission only are presented in the final two columns (outbreak size and outbreak duration), risk factors describing both ingress and transmission are presented in the second column (infection). Factors associated with increased risk of the outcome are shaded in orange and factors that are associated with a reduced risk are shaded in blue. Results of overall analysis are shown in the top box, analyses stratified by Omicron period are presented in lower two boxes.

Ingress only | Ingress & Transmission | Transmission only

OVERALL

THEME	BUILDING FACTORS	OUTBREAK	INFECTION	OUTBREAK SIZE	OUTBREAK DURATION
Facilities and size	**More storeys**		Decreases rate		
	Purpose built vs converted building		Increases rate		
	More bedrooms		Increases rate		
	Greater proportion LA beds				Decreases rate
Temperature	**Greater bedroom temperature**				Increases rate
Air quality – comm. room	**Dry vs just right**			Increases rate	
Crowding	**More people in dining room**				Decreases rate

PRE-OMICRON

Facilities and size	More storeys		Decreases rate		
	More bedrooms		Increases rate		
Ventilation – dining room	Cassette ceiling unit vs central air conditioning		Decreases rate		
	Portable unit exhaust pipe vs central air conditioning		Increases rate	Increases rate	
Crowding	More people in dining room		Decreases rate		

OMICRON

Facilities and size	Purpose-built vs converted building		Increases rate		
	Greater proportion LA beds		Increases rate		
Cleaning	Daily floor washing vs less frequent		Increases rate		

$P \geq 0.05$ | Increases rate ($P<0.05$) | Decreases rate ($P<0.05$)

LA Local Authority

6.4 Discussion

My analysis has demonstrated that, after adjusting for confounders, the only clear driver of infection ingress into care homes was the local incidence of SARS-CoV-2, supporting findings from my scoping review in Chapter 2. In contrast, it is likely that environmental factors can influence transmission given the important associations described between these and infection incidence and outbreak size and duration. Environmental features appearing to increase transmission in the overall analysis included facilities that were purpose-built, had more bedrooms, warmer temperatures, and drier perceived air quality. Transmission appeared to be lower in care homes with more storeys, more LA-funded beds, and more people in common spaces. The strength of these associations differed depending on whether Omicron variant was in circulation. A new association with cleaning frequency was identified in the Omicron-dominant period and with ventilation in common spaces in the pre-Omicron period. These factors are mainly indicators of air flow and how well infected residents can be isolated by the care home[383,384], for example by caring for them on different floors; suggesting that limiting spread may be more achievable for care homes than stopping infection ingress. Some associations are difficult to interpret and may reflect underlying confounding or reverse causality, for example the negative association with number of people in the dining rooms may be because care homes managed larger outbreaks by confining residents to their bedrooms. However, I found substantial diversity in built environments, and there were major challenges associated with capturing this information, particularly in the context of a pandemic. Overall, this highlights that consultation between care homes and public health teams on a local level and tailoring infection control strategies is likely to maximise their impact.

Associations between infection outcomes and variables available in administrative datasets used for financial purposes, such as staffing, bed occupancy, for-profit status, rurality, and publicly available data on local infection incidence have been described previously.[154,216–218,229] To my knowledge my study is the first to describe the heterogeneity in care home environment from an infection control perspective in a large sample. I found two systematic reviews investigating the association between the built environment and SARS-CoV-2 risk in care homes, but methods varied. One systematic review conducted mid-way through the pandemic and presented as a

conference poster found 17 studies and identified that key factors associated with SARS-CoV-2 risk were crowding, small cluster dwellings, urban location, ventilation, and outdoor space, however insufficient details on search and data collection methods were available.[385] A further narrative review which included single care home and larger multisite studies, as well as national recommendations, was conducted in the first half of 2022 and identified six important elements based on a hierarchy developed by the author from their professional experience: ventilation, spatial separation, physical barriers, hand hygiene stations, resident room zones, and private rooms. However, these conclusions were only supported by data from predominantly small-scale heterogenous studies.[386]

Prior to the pandemic, most studies of care home design were qualitative, focussing on how the built environment impacts on quality of life and ways that spaces can be optimised for those with dementia and at higher risk of falls.[366] Key features identified have been ease of access, warmth, brightness, natural light, openness, views and or/access to greenery, and home-like design with attractive buildings and furniture.[366,387,388] A recent synthesis recommended that care homes should be made up of separate apartments that allow residents to continue to socialise in a more familiar and smaller unit, which could also optimise the ability to isolate and cohort infected residents.[366,389] This is reflected in the Green House model of care home design which was first developed in 2001, and uses small non-traditional clusters of housing with a maximum of 10-12 residents per unit. Each unit consists of a central entry point, smaller overall space, private bedrooms and bathrooms, and consistent staff. Reported benefits for residents have been improved quality of life, reduced hospital admissions, reduced Medicare (health insurance) spending and lower staff turnover.[390] This model of care has expanded rapidly and by 2020 there were over 300 Green House care homes across the USA. A large study comparing these with more traditional care homes found lower incidence of COVID-19 infections and deaths in the Green House homes over the first wave of the pandemic, suggesting that this may be an effective model for infection prevention.[391]

Most care homes in my study reported having older centralised air conditioning or freestanding fans and, although confidence intervals were wide, portable units compared with centralised ventilation systems were associated with increased

transmission risk. However, as only half of care homes answered the questions on ventilation, it is difficult to draw definitive conclusions from these data. A recent scoping review identified two studies linking central air conditioning with increased risk of nosocomial transmission of infectious pathogens,[392,393] although there appeared to be a protective effect when air was filtered.[394] To my knowledge, no large studies have evaluated the role of ventilation in transmission of infection in care homes specifically.[395]

Characteristics of the care home population (i.e., reduced mobility of residents) suggest that commonly used strategies like monitoring of CO_2 (produced by respiration and used as a proxy for overcrowding), which have been recommended by SAGE to monitor and adjust ventilation (by drawing in outdoor air in response to a rise in CO_2)[396] may need to be recalibrated in view of differences in how older people move around and use spaces when compared with younger, more active populations. This was implicated as a contributory factor in a large outbreak reported from a care home in the Netherlands.[397] To understand the association in ventilation in more detail, future research should include air sampling and sequential data collection on factors that influence air flow, such as: number of air changes, presence and types of filters, frequency of filter cleaning, vent placement, placement of windows and doors, and movement of people within the space.[398–400] A specific question around whether air is recirculated or brought in from outside may benefit the analysis if completed accurately.

Care homes in my sample were predominantly purpose-built (81%), but the adjusted rate ratio of infection in purpose-built buildings was at least double that of converted ones, which is surprising. This may be because although new care home building standards were introduced in 2003[401], 24/34 care homes that completed this question were built before this. It is also possible that air leakage from the less efficient insulation of the external envelope of older converted homes reduced infection transmission by enabling air flow through the facility. Consistent with published literature, [402,403] drier air was associated with lower transmission, although as air quality was assessed subjectively, objective measurements could improve reliability. Although greater indoor temperature was a risk factor for transmission in my analysis, the complex relationship between temperature, humidity, and air flow has precluded

meaningful conclusions in the literature about the association with temperature alone.[383,402,403] As prior to the pandemic, care home regulations only stipulated the use of building ventilation for comfort (maintaining air temperature and humidity),[401,404] future regulations will need to shift priorities and consider how to reduce circulation of airborne infectious droplets and aerosols.

In addition to the physical features of the build environment it is important to consider how the facilities are used. A root cause analysis of four outbreaks in Scottish care homes in the first wave of the pandemic[405] identified a few key issues around isolating residents with dementia who often "wander with purpose" and are at high risk of falls, and the ability of care home staff to rapidly apply infection control principles to facilitate safe isolation of residents. In some, this inadvertently resulted in infection spread, such as in the case of kitchen staff who delivered meals to residents' rooms instead of serving them in the communal space and therefore transmitted infection to these residents. I did not collect data on infection control practices such as handwashing, isolation of residents, and PPE use, which can influence infection risk, although studies that have examined this are described in Chapter 2.

As demonstrated in the stratified analysis, it is likely that properties of different viral variants, vaccination, and infection history influence the association between building characteristics and the risk of transmission. Although there were four variants that dominated at different points in the study period, I chose to only consider the specific impact of the Omicron variant which is substantially more transmissible than prior variants [379,380] and which was associated with the largest peak in cases and outbreaks in my study. To date, this is the only study to consider how different variants affect transmission within specific care home environments. My analysis was limited by samples size however it will be important to consider variant properties in future analyses.

6.5 Strengths & Limitations

The strengths of my study are that I was able to consider how the built environment affects both ingress and transmission of infection by considering more than one outcome and adjusting for known risk factors. As my dataset comprised data from

regular and universal testing during a period of high SARS-CoV-2 incidence nationally, I was uniquely placed to optimise the number of infections in my analysis, thereby increasing power, capturing both asymptomatic and symptomatic infections. In this study, I have also presented the first comprehensive description of the built environment of care homes from a large and geographically dispersed sample across different providers. In contrast to previous studies that reported from facility-level aggregate datasets, I linked to individual-level data on participant characteristics, immune status, and test results from surveillance datasets. This increased the reliability of my results as this limited reporting and recall bias and was able to account for variations in these factors over the analysis period, which I allowed to vary with time. I estimated individual facility entry and exit dates using dates of routine PCR/LFD tests that had been linked to CQC-IDs, which is especially important given the high staff and resident turnover in these settings. However, given the monthly testing schedule of residents, it is possible that short-stay residents were missed. The 19-month follow-up period is longer than any other reported study in this setting therefore I was able to consider the impact of changing epidemiology and emerging variants.

However, my study was limited by missing survey data, more commonly affecting the technical questions as these may have been more challenging. Lowest response rates were seen for questions around air temperature, presence of humidifiers, and ventilation type. Reliability of responses was also an issue, as demonstrated by 5/52 completed surveys that reported temperatures greater than 30°C, with one reporting 60°C in the common room. Social desirability bias is possible as answers may have reflected best practice. Although questionnaires were initially distributed in August 2020, due to significant strain in the care sector they were only completed partway through the analysis period, one year later. This limits the inferences that can be made about factors such as cleaning and use of common spaces as reverse causality is possible, for example care homes with more outbreaks may have cleaned more frequently.

In addition, due to rapid timescales for survey development, I was unable to fully pilot the surveys. This meant free-text and multiple-choice questions were difficult to convert into data that could be easily analysed. Despite input from building scientists into survey design, it is challenging to cross-sectionally capture detailed information

on air flow, humidity, temperature, and person-to-person contacts, particularly from a busy care home. The CONTACT study in Leeds attempted to evaluate the effectiveness of using wearable devices to record number of contacts within a care home, however faced significant challenges.[406] In future, I would consider piloting surveys more extensively, ensuring questions are easy to answer and can be applied to my analysis.

The majority of care homes were for-profit (87%) which may have introduced bias as increased SARS-CoV-2 risk has been well described in these homes compared with not-for-profit homes.[228,241] This was also reflected in my analysis in Chapter 4, where I described greater seroprevalence in for-profit when compared with not-for-profit homes. However, in this current analysis, I found no significant associations between ownership type and any of the study outcomes and I adjusted for the effect of this factor in my final models. LA-funding of beds was both positively and negatively associated with transmission (overall vs stratified analysis), however this may reflect changing numbers of step-down admissions from acute inpatient settings during periods of COVID-19 associated strain on hospital bed capacity.

I included multiple variables in each model, which makes it more likely that a statistically significant result was generated by chance. I considered a Bonferroni correction to account for this where the original alpha level (set to 0.05 for my study) is divided by the number of tests and this new threshold is used to determine significance.[407,408] Given the large number of variables in my study, this would lead to a very conservative cut-off increasing the chance of false-negative results from my analysis. As the focus of this work was primarily descriptive, I chose to retain the original alpha value at 0.05. However, to build on this work, I would limit the number of variables in future models by focussing on aspects of the built environment that appear to influence infection transmission based on findings from this from study.

Finally, although I adjusted for time-varying baseline covariates, it is very likely that the rapidly changing epidemiology and simultaneous introduction of multiple control measures impacted the accuracy of my analyses. This is particularly relevant given the cross-sectional survey design as it was difficult to consider how temporal changes in factors such as indoor temperature and air quality affected study outcomes over

time. As I was inherently unable to adjust for all confounders, this has limited my interpretation of some associations such as between LA-funding and outbreak duration, as short-term admissions to care homes during periods of strained hospital bed capacity may not have been accurately recorded in the monthly bed allocations. Future work should consider how to optimise response accuracy by asking for objective measurements and repeating surveys to track temporal changes.

6.6 Conclusion

This work has highlighted that environmental factors are associated with infection transmission and these relationships are influenced by the variant. Limiting the spread of infection is probably more achievable for homes than preventing ingress, and considering characteristics such as outbreak size and duration may be valuable when identifying care homes that would benefit from targeted public health support.

Based on my findings, care home design should focus on the ability to isolate infected residents, for example on different floors; have fewer bedrooms; ensure good ventilation; humidify indoor air; and reduce indoor temperatures - although maintaining comfort in these conditions for less mobile residents may be challenging. Facilities that house residents in small units like the Green House model, have been shown to improve quality of life for residents whilst reducing risk of infection transmission. Features of these facilities align with characteristics that reduce risk of transmission from my study and could be considered when designing future care homes. Standards were last updated twenty years ago, and new standards should build on the momentum gained in the pandemic to optimise infection prevention and balance this with comfort and dignity for the individual residents for whom these settings are homes.

Over my PhD thesis, I have investigated the role of agent, host, and environmental factors in infection transmission. Delineating these factors may help to identify highest risk care homes that may benefit from additional support. Although agent factors can influence transmission, sequencing data allowing identification of viral lineage are usually not available within the required timeframes to enable direct action. Reliably measuring host immunity at scale is challenging in view of incomplete data capture and antibody waning. However, it is possible that modifying the built environment may

be more effective at preventing transmission and more tolerable. This should be explored in further detail through focussed surveys, environmental sampling studies, and trials of interventions to prevent infection spread.

6.7 Contribution & dissemination

The study was conceived by Laura Shallcross (LS), Hector Altamirano (HA), and I. In view of the urgent need for data during the early stages of the pandemic, the first draft of the survey was designed by HA as he had prior experience of designing environmental surveys, with input from LS and me. I piloted the survey with care home managers and compiled feedback, which HA and I incorporated in the survey. I distributed the survey to project managers from participating care homes and was responsible for sending out reminders. Survey data was entered into a central database by Vivaldi project manager, Chris Fuller, and I reviewed the accuracy of data entry and resolved any queries. I designed the statistical analysis with input from my supervisors and from Oliver Stirrup, post-doctoral statistician for the Vivaldi study. I conducted data analysis, created data visualisations, and have drafted the manuscript. Niyathi Sethu, an MSc student, used a sample of the survey data for her MSc dissertation, however these data have not been published.

I presented this work to the DHSC DDG on 9th February 2023.

I have also submitted the study manuscript for peer-review to the Journal of the American Medical Directors Association (JAMDA).

Chapter 7

Overview, Conclusions & Future Work

7.1 Summary of key findings in context of wider literature

In this thesis, I used one of the largest care home cohort studies internationally to address policy-relevant questions about the epidemiology of the SARS-CoV-2 pandemic. I used the available data to consider agent, host, and environment factors associated with infections and outbreaks in care homes (Chapter 1). I described the seroprevalence and spread of SARS-CoV-2 infection and how this varied, estimated the magnitude and longevity of antibody responses in the care home population, described the variation in built environments of care homes, and investigated risk factors for infection ingress and spread in these settings. I have also described the key role I played in rapidly establishing the cohort in response to the unprecedented need for data, early in the global SARS-CoV-2 pandemic. In this chapter, I discuss the main findings from my PhD, reflect on what I have learned, consider key evidence gaps, and make recommendations for future research.

7.1.1 Scoping literature review on care home factors associated with the risk of SARS-CoV-2 infections, outbreaks, and large outbreaks.

In Chapter 2, I performed a scoping review to evaluate the published literature describing facility-level risk factors for SARS-CoV-2 infections and outbreaks in care homes. I identified 31 peer-reviewed studies mainly from Europe and North America that were published early in the pandemic.

Overall, the existing literature was limited by the following key themes. Insufficient follow-up meant that the impact of emergent variants, population immunity dynamics, and changes to preventive measures was not captured. Cohorts did not extend beyond the introduction of the SARS-CoV-2 vaccine therefore could not evaluate the impact of this intervention. Predominantly symptomatic cases were included therefore underrepresenting asymptomatic or atypical cases. Case ascertainment varied and was often based on clinical diagnosis alone. Finally limited consideration was given to the contribution of the built environment to infection and transmission risk.

The review process highlighted that factors can be split into those that describe risk of the introduction of infection into a facility and those describing transmission once infection has entered. A single case of SARS-CoV-2, particularly in a staff member, is unlikely to have been acquired within the care home however infection may spread from this primary case through the facility and infect others. The number of secondary cases (or secondary attack rate) or the size of the ensuing outbreak can describe the magnitude of transmission. Although not all studies considered these outcomes separately, my review identified that although there was some overlap in risk factors for ingress and for transmission, there were also important differences.

Infection ingress (described by infection in a resident or occurrence of an outbreak) was associated with incidence of SARS-CoV-2 in the local community, the number of beds, ownership type, and the quality rating. Transmission (measured by case number or size of outbreak) also appeared to be associated with staffing levels, suggesting that care home staff play an important role.

Considering risk factors for infection ingress and transmission separately could help to focus preventive measures, conserving already stretched resources and limiting unnecessary disruption. This may also have implications for how SARS-CoV-2 outbreaks are defined in care homes in the future. These definitions are designed to identify linked cases within a setting so that measures can be implemented to limit spread. However, in the context of the unusually high community transmission rates of SARS-CoV-2, two cases within fourteen days (the current PHE definition)[376] may reflect two unrelated introductions of infection into the care home, usually via staff. As preventive measures such as asymptomatic screening and visiting restrictions are lifted, it may be valuable for policymakers to reconsider these outbreak definitions, to differentiate transmission within care homes (which may be mitigated by control measures) from sporadic introduction of infection from the community, which is largely outside of care providers' control.

7.1.2 Establishing a national cohort study of SARS-CoV-2 in care homes.

In Chapter 3, I described how the VIVALDI study, the largest prospective cohort study of SARS-CoV-2 infection and immunity in care homes in England, was rapidly established early in the pandemic.

VIVALDI capitalised on a national programme of asymptomatic SARS-CoV-2 screening in care homes from July 2020 onwards[171] and changes to data regulations over the pandemic that facilitated identification of staff and residents through linkage of tests to participating care homes at scale.[263] The study included 330 care homes for older people from a range of large and small care providers across England, linking routine data on SARS-CoV-2 testing, vaccinations, hospital admissions, and deaths from approximately 70,000 participants. Sequential blood samples were also donated by a subset of consenting individuals and analysed for humoral and cellular components of the immunological response to SARS-CoV-2. This data infrastructure created a unique opportunity for epidemiological research in care homes which we hope to build on in future. However, there were significant challenges associated with establishing and running the study and working at high pressure to such short timelines, which I have described in detail in Section 7.3.

7.1.3 Measuring the proportion of care home staff and residents infected with SARS-CoV-2 and describing variation between care homes.

In Chapter 4, I showed that SARS-CoV-2 spread rapidly from the community into care homes following the emergence of the Alpha variant, despite stringent control measures, suggesting that staff played a role in SARS-CoV-2 ingress from the community and that characteristics of emerging viral variants survival are likely to significantly influence infection risk.[291] I also estimated the prevalence of SARS-CoV-2 in residents and staff at different points in the pandemic, showing that after two years in the care home, two-thirds of the population had been infected. This large reservoir of previously exposed individuals in care homes is comparable to the proportion of exposed individuals in the community.[311] Given the strict control measures implemented early in the pandemic to prevent SARS-CoV-2 infection in care homes, these results are also surprising and raise questions about their effectiveness. Simultaneous introduction of control measures has limited the conclusions that could

be drawn about their benefits in isolation.[156] However, considering the negative psychological and social consequences of policies like restrictions on non-essential visitors that were in place for over a year,[160,196] evaluating the benefits and harms of these measures should be a future research priority.

My results have also highlighted the challenges with measuring infection in this population due to issues with data quality. It was not possible to include residents who died before national screening was established and in the absence of accurate data on care home entry and exit dates, accounting for turnover was unreliable. The large proportion of previously-exposed individuals is reassuring from one perspective as, although vaccination coverage in this population is very high, there is evidence that previously exposed individuals mount stronger immune responses to SARS-CoV-2 and have greater vaccine-elicited protection against severe outcomes from reinfection.[162,295] Given substantial resident turnover in care homes,[314] it will be important to consider the impact of new admissions who may have been avoiding infection (shielding) in the community, on this immune reservoir.

This work has showcased the impact of timely and policy-relevant research on policy decisions, and the importance of establishing data infrastructure early on. For example, our agile response to the emergence of the Alpha variant allowed us to rapidly demonstrate the inadequacy of disease control measures in preventing ingress of infection in the community, directly informing policy decisions to impose a national lockdown, and helping to identify regions in need of additional support. Overall, this demonstrates how ongoing surveillance and collaboration is key to preparedness against future infectious disease threats.

7.1.4 Investigating the durability of SARS-CoV-2 antibody responses in staff and residents of care homes following infection and vaccination and measurement on a facility-level.

In Chapter 5, I demonstrated that naturally acquired (anti-nucleocapsid) antibody responses to SARS-CoV-2 in the care home population remained detectable for around eight months after primary infection, although the rate of sero-reversion was greater in staff than residents.[292] This may be due to greater anti-nucleocapsid

antibody titres following infection in residents compared with staff. Estimating the timing of infection was challenging due to variability in testing access which meant that many early infections were not recorded. This is a problem across studies that utilise routine surveillance data however the high frequency and coverage of testing over the second and third waves of the pandemic provided a unique data source for my study. In contrast, vaccine-induced (anti-spike) antibody responses were more durable, particularly in previously infected hosts and rates of decline did not differ between staff and residents.[335] However, the clinical implications of antibody waning and peripheral antibody titres on protection against infection and severe outcomes are still unclear.

Given rapid antibody loss following infection, seroprevalence surveys may be of limited utility for estimating infection exposure and identifying high risk care homes. The choice of assay and positivity threshold probably impacts on the accuracy of results however, a more reliable test describing population-level immunity is required to provide better evidence for infection models and policy.

Investigating and monitoring real-world effectiveness of immune responses is most important for informing preventive strategies, however this is challenging. Correlates of protection can be estimated by measuring the level of protection afforded by components of the immune response against infection and severe outcomes.[409] For SARS-CoV-2, antibodies against the virus have been the easiest to measure,[410,411] although neutralising antibodies and T-cell responses are also recognised as significant correlates.[412,413]

In VIVALDI, to examine magnitude and longevity of antibody and T-cell responses against different SARS-CoV-2 proteins and variants in more detail, we have collaborated with Professor Paul Moss's laboratory at the University of Birmingham.[295,348,414,415] We are also collaborating with researchers in Professor Rupert Beale's laboratory at the Francis Crick Institute who are monitoring the protective function of antibodies against emerging variants using assays that detect neutralisation of antigens in a subset of VIVALDI samples.[415] Future work could also aim to determine an antibody threshold that protects against clinical infection (correlate of immune protection), as this may be a useful measure of population-level immunity that can be applied for screening, risk-stratification, and modelling purposes.

7.1.5 Investigating how care home characteristics related to the built environment are associated with risk of SARS-CoV-2 infections and outbreaks and whether factors associated with ingress differ from those associated with transmission.

In Chapter 6, I demonstrated substantial diversity in the built environment of care homes in England, which has not previously been described. Addressing the gaps identified in my scoping review, I found that building factors associated with infection ingress differed from those associated with transmission within the facility and that these relationships appeared to vary according to the transmissibility of the viral variant. Whilst introduction of infection was only related to incidence of SARS-CoV-2 in the local community, environmental factors were significantly associated with spread of infection. These associations varied depending on the dominant variant. This supports the notion that whilst care homes may have limited opportunities to prevent infection ingress, reducing transmission of infection within the facility by modifying the environment may be feasible. This emphasises why public health measures should focus on preventing transmission rather than ingress of infection.

Missing data affected quality of surveys, which were also subject to bias from self-completion and reverse causality as pandemic pressures in the care homes meant they could not be completed at the start of follow-up, contrary to initial study plans. It was also challenging to account for the significant variation in epidemiology of infection and policy changes over the study period. The introduction of multiple variables into the final analysis also increased the likelihood of identifying a significant association by chance. Although these factors limit the inferences that can be drawn, the study provides useful preliminary data that can be used for future research in this area and highlights the significant impact of the built environment on SARS-CoV-2 infection in care homes. It also demonstrates how considering multiple study outcomes to describe infection dynamics can be informative for targeting infection prevention policy.

Building on this work, I hypothesise that specific environmental features are associated with infection spread within facilities. Studies like the PROTECT COVID-19 National Core Study have been established to describe how transmission varies in

different environments.[416] Whilst this study includes care homes, their work has focussed on qualitatively understanding current approaches to prevention of transmission. Pandemic-related constraints impacted on the quality of data that could be collected from my survey, therefore I plan to develop a series of studies to collect data on factors associated with air flow in line with published literature.[398–400] This could include number of air changes, presence and type of filters, cleaning of filters, movement within the space, fluctuations in CO_2 concentrations, air temperature and humidity, location of ducts, windows and doors. This would involve in-depth surveys from a representative sample of care homes which include objective measures to allow comparison between sites. The surveys would be piloted extensively to ensure answers are accurate and reproducible. Equipment for measurements would be provided by the study (such as thermometers, CO_2 meters) with some reimbursement to participating sites. Environmental sampling and genomic sequencing of environmental and clinical samples in a subset of care homes would augment surveys and inform analyses of transmission within the facility.

To expand on this work, I hypothesise that specific environmental modifications can be made to limit infection spread within facilities. There are a small number of clinical trials focusing on specific environmental features in care homes, such as the AFRI-c study which is evaluating the effectiveness of air filters in prevention of SARS-CoV-2 transmission.[417] However due to the barriers to recruiting care homes to research studies and the associated workload, their sample sizes are often modest. By capitalising on the care home network that I will help to establish following my PhD (described in Section 7.6), I hope to recruit a representative sample of care homes to a trial of an environmental intervention to limit infection spread. The intervention design will be informed by prior environmental surveys and co-produced with social care stakeholders. This will showcase the potential for collaboration to develop impactful research within adult social care.

7.2 Participant & Public Involvement & Engagement (PPIE)

In Chapter 3, I described initial consultation with key stakeholders to inform study set-up and the establishment of two study specific PPIE groups of care home residents' families. Throughout the study, engagement with participants and the public has

provided useful feedback on study progress and informed plans for future research priorities. Although opportunities for consulting these groups were limited early in the pandemic due to short timelines, below I have described subsequent activities.

7.2.1 Sharing study findings with a wide audience

Study results were summarised in short reports and rapidly presented to government Ministers, advisory groups e.g., NERVTAG, SAGE Social care working group and policymakers to maximise the impact on national decision-making. To rapidly disseminate findings within the scientific community, manuscripts were pre-printed and submitted to peer-reviewed journals simultaneously - the former to avoid delays from the latter. We also created lay summaries that informed study participants and the wider care sector about research findings and potential implications. To increase the visibility of the study, I created a study website,[418] which I initially maintained however this task was subsequently taken over by the study administrator and project manager. This was designed so information about study aims, processes, publications, engagement, and points of contact was publicly accessible. I also pre-printed the study protocol[245] and registered the study to the International Standard Randomised Controlled Trial Number (ISRCTN) registry (ISRCTN14447421),[419] maintaining the profile throughout the study.

The programme for development of lay summaries was primarily devised and delivered by me, the study project manager (BA), a clinical research fellow in the team (RG), and a visiting researcher (GB), with input from the wider study team. They consisted of two sessions per PPIE group with activities between sessions which participants were reimbursed for. The first session described the study setting and aims and outlined the goal of the exercise. Following this introduction, participants were asked to review draft lay summaries that the team had created and to complete a questionnaire. Main areas of focus were around how relevant and engaging the summaries were, the appropriateness of language, and quality of diagrams. The second session enabled further discussion of these key features where alternative graphics, in line with feedback from questionnaires, were also presented. Some of the key suggestions from the group that we have incorporated into all future engagement materials were including a glossary, incorporating both an overall summary and one

with additional detail to cater to different levels of interest and knowledge, taking care with use of bright colours, being aware of variation in colour perception between people, and including a simple schematic to illustrate key findings. Final summaries were created by BA with input from myself and the rest of the team, Figure 7.1. These have been made available on the study website and printed copies were sent out to all study participants. To date, feedback has been overwhelmingly positive from residents, their families, staff, and study collaborators.

Figure 7-1: Lay summaries developed with PPIE groups describing results from a) analysis of risk of reinfection b) first dose vaccine efficacy.

(Also available on VIVALDI website[418]*)*

a)

CAN CARE HOME STAFF AND RESIDENTS GET INFECTED WITH COVID-19 MORE THAN ONCE?

vivaldi UCL

People living and working in care homes who have been infected with COVID-19 before are unlikely to get infected a second time.

Between June 2020 and February 2021, we investigated whether people in care homes can get infected with COVID-19 more than once. We looked at blood and nasal swab test results from 2,000 staff and residents in 100 Four Seasons Health Care homes across England. The results showed that people in care homes who have previously been infected with COVID-19 are unlikely to get infected a second time, but it is important to get the vaccine to get the best level of protection.

WHAT IS THE VIVALDI STUDY?

Researchers on the VIVALDI Study are investigating the impact of COVID-19 on care homes and what can be done to prevent infection from spreading among staff and residents. The study was set up in June 2020 and is collecting information from over 50,000 care home staff and residents across more than 300 care homes in England.

WHY DID WE DO THIS STUDY?

The COVID-19 pandemic has hit care homes very hard. Care home staff and residents have higher rates of infection, hospital admissions and deaths compared to other people. Despite this, at the beginning of the pandemic, there was very little information about what was happening in care homes.

WHAT DID WE WANT TO FIND OUT?

Most people who are infected with COVID-19 develop antibodies against the disease 1-2 weeks after infection. Working-age adults with antibodies are unlikely to get COVID-19 a second time. As we get older, however, our immune system slows down and the antibodies that we produce might not protect as well against infection. We wanted to know if older residents in care homes who have COVID-19 antibodies are also protected from new infection.

WHAT DID WE DO?

We invited staff and residents from a sub-set of 100 Four Seasons Health Care homes taking part in the VIVALDI study. At this time most care home staff and residents were not vaccinated, and the new variants were not identified. We collected information on the age and sex of participants but did not have good quality information about their ethnicity.

Between June and October 2020, participants gave us up to three blood samples. The samples were tested for COVID-19 antibodies.

GLOSSARY

Antibody
Produced by the body to fight specific infections.

Care Home
A residential facility for people who need extra help with looking after themselves.

Care Home Resident
A person who lives in a care home.

COVID-19
Coronavirus disease (COVID-19) is a highly contagious respiratory infection caused by the SARS-CoV-2 virus.

COVID-19 Related Death
Death within 28 days of a positive swab test for COVID-19.

Immunity
The ability to resist a particular infection by the action of the body's immune system.

Infection
The invasion of the body by an infectious agent like bacteria or a virus.

Nasal Swab Test
A device inserted in the nose, used to look for active COVID-19 infection.

Older Adult
A person aged 65 years and over.

Working-age Adult
A person aged under 65 years.

Using these samples, we could tell who **had** been infected with COVID-19 before October 2020 (people **with** COVID-19 antibodies in their blood) and who **had NOT** been infected (people **without** COVID-19 antibodies in their blood).

Since July 2020, all care home staff and residents have been tested regularly for COVID-19 infection. This was done with nasal swab tests that were organised through the national testing programme. We used the results from these swab tests to pick out individuals who had a new COVID-19 infection after October 2020. We combined the information from the blood and nasal swab tests and split the participants into two groups. We then compared the number of infections in these groups.

Group 1: Those who did NOT have COVID-19 antibodies in their blood by October 2020, and who later had a COVID-19 infection as shown by a positive swab test.

Group 2: Those who had COVID-19 antibodies in their blood by October 2020, but who later had a second COVID-19 infection as shown by a positive swab test.

WHAT DID WE FIND?

People **with** COVID-19 antibodies at the start of the study (indicating they had been infected before) were much less likely to get a new COVID-19 infection compared to people **without** antibodies.

In an average month between October 2020 and February 2021, **7** out of 100 residents **without** antibodies had a COVID-19 infection. Only **1** out of 100 residents **with** antibodies at the start of the study had a second COVID-19 infection. In the same period, **3** out of 100 staff members **without** antibodies were infected with COVID-19. Only **1** out of 100 staff **with** antibodies at the start of the study had a second COVID-19 infection.

WHAT DOES THIS MEAN FOR CARE HOMES?

The findings suggest that people living and working in care homes who have already been infected with COVID-19 have a reduced risk of further infection. This is good news, but we do not, however, know how long this protective effect lasts. It is therefore important for everyone to get vaccinated (even if you have been infected with COVID-19) to get the best long-term protection against COVID-19.

WHAT HAPPENS NEXT?

We are now working with staff and residents across 300 care homes to investigate how much protection vaccination provides against infection and for how long. This will help us to work out how often people need to be re-vaccinated, and when it might be safe to relax social distancing measures in care homes.

This summary was produced in partnership with patient and public representatives.

Read the full report:
Incidence of SARS-CoV-2 infection according to baseline antibody status in staff and residents of 100 long term care facilities (VIVALDI): a prospective cohort study
www.thelancet.com/journals/lanhil/article/PIIS2666-7568(21)00093-3/fulltext

CONTACT US
VIVALDI Study
UCL Institute of Health Informatics
222 Euston Rd
London, NW1 2DA

Email
vivaldi.covid19@ucl.ac.uk

Website
www.ucl.ac.uk/health-informatics/research/vivaldi-study

Icons made by Freepik from www.flaticon.com

Funded by Department of Health & Social Care

UNIVERSITY OF BIRMINGHAM

vivaldi

CAN CARE HOME STAFF AND RESIDENTS GET INFECTED WITH COVID-19 MORE THAN ONCE?

Funded by Department of Health & Social Care | UNIVERSITY OF BIRMINGHAM | vivaldi UCL

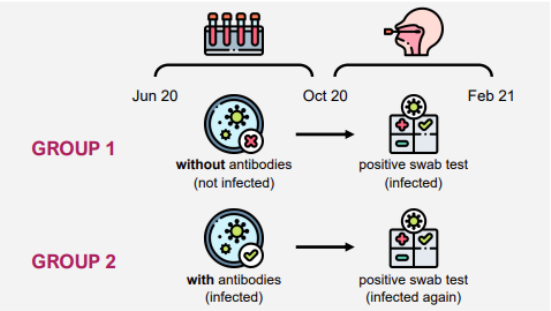
WHAT DID WE WANT TO FIND OUT?

We wanted to know if care home staff and residents who had COVID-19 before can get a second infection.

WHAT DID WE DO?

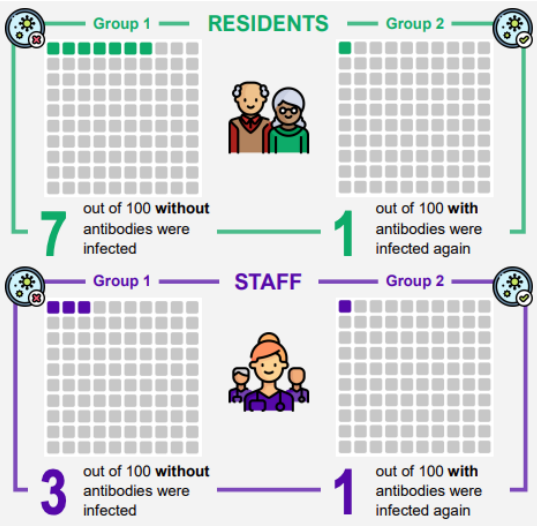
We looked at blood samples from care home staff and residents in 100 care homes across England. We wanted to see who had COVID-19 antibodies in their blood (indicating COVID-19 infection) **before** October 2020. We later looked at nasal swab results to see who had COVID-19 **after** October 2020.

We then compared the blood and nasal swab tests and split the participants into two groups.



WHAT DID WE FIND?

In an average month between October 2020 and February 2021, care home staff and resident who had COVID-19 before were less likely to get a second infection.



This summary was produced in partnership with patient and public representatives

Icons made by Freepik from www.flaticon.com

b)

DOES ONE DOSE OF VACCINE STOP COVID-19 INFECTION IN CARE HOME RESIDENTS?

vivaldi UCL

A single dose of COVID-19 vaccine gives care home residents a good level of protection against COVID-19 infection after one month.

Between December 2020 and March 2021, we investigated whether one dose of vaccine protects care home residents against COVID-19 infection. We looked at vaccination information, blood tests and nasal swab tests from 10,000 residents in 310 care homes across England. This was a period before most people in the UK had been offered their second COVID-19 vaccine. The results showed that one dose of vaccine provides some protection against COVID-19 in care home residents, but it is important to get a second dose to get the best level of protection.

WHAT IS THE VIVALDI STUDY?

Researchers on the VIVALDI Study are investigating the impact of COVID-19 on care homes and what can be done to prevent infection from spreading among staff and residents. The study was set up in June 2020 and is collecting information from over 50,000 care home staff and residents across more than 300 care homes in England.

WHY DID WE DO THIS STUDY?

The COVID-19 pandemic has hit care homes very hard. Care home staff and residents have higher rates of infection, hospital admissions and deaths compared to other people. Vaccines protect people against COVID-19 infection and are especially important for care home residents. However, we know from other studies that the level of protection provided by vaccines can be lower for older adults compared to working-age adults.

WHAT DID WE WANT TO FIND OUT?

Research studies in the general population show that COVID-19 vaccines are safe and give a good level of protection against infection even after just one dose. Care home residents were not included in these studies. We wanted to know whether the first dose of vaccine protects care home residents against COVID-19 infection.

WHAT DID WE DO?

We looked at vaccination records of over 10,000 care home residents across England. We linked these records with results of nasal swab tests which are taken monthly from every care home resident as part of the national testing programme. This allowed us to identify anyone with a current infection.

We also collected information about past infection based on blood tests that looked for antibodies against COVID-19 in some of the residents.

GLOSSARY

Antibody
Produced by the body to fight specific infections.

Care Home
A residential facility for people who need extra help with looking after themselves.

Care Home Resident
A person who lives in a care home.

COVID-19
Coronavirus disease (COVID-19) is a highly contagious respiratory infection caused by the SARS-CoV-2 virus.

Immunity
The ability to resist a particular infection by the action of specific antibodies.

Nasal Swab Test
A device inserted in the nose, used to look for active COVID-19 infection.

Older Adult
A person aged 65 years and over.

Working-age Adult
A person aged under 65 years.

Vaccine
A product which stimulates a person's immune system to protect them from a specific disease.

Each person in the study had at least one nasal swab taken before vaccination and at least one taken after vaccination. This allowed us to compare the risk of infection before vaccination, with the risk of infection after vaccination. Unfortunately, we were not able to collect reliable information on the ethnicity of people taking part in the study.

WHAT DID WE FIND?

By January 2021, most care home residents in our study had received at least one dose of vaccine (either the Astra Zeneca/Oxford or Pfizer vaccines).

We found that approximately one month after vaccination, a single dose of either vaccine provides a good level of protection against COVID-19 infection. We estimated in an average month between December 2020 and March 2021, **6** out of 100 residents who were NOT vaccinated would get infected with COVID-19, whereas only **2** out of 100 residents who WERE vaccinated would get infected with COVID-19.

The Oxford/ Astra Zeneca and Pfizer vaccines gave similar levels of protection against COVID-19 infection. Vaccination does not seem to provide any extra protection from infection in residents who had been infected with COVID-19 before. We also found that vaccinated residents who get infected may be less likely to pass on COVID-19 infection to others, compared to unvaccinated residents who get infected.

WHAT DOES THIS MEAN FOR CARE HOMES?

The findings suggest that vaccination reduces the total number of people who get infected with COVID-19. These results support evidence that vaccinations work well in preventing COVID-19 infection in care homes. It is recommended that people have two doses of these vaccines to give the best level of protection against infection.

WHAT HAPPENS NEXT?

We are now working with staff and residents across 300 care homes to find out how much protection vaccination with two doses provides against infection and for how long. This will help us to work out how often people need to be re-vaccinated, and when it might be safe to relax social distancing measures in care homes.

This summary was produced in partnership with patient and public representatives.

Read the full report:
Vaccine effectiveness of the first dose of ChAdOx1 nCoV-19 and BNT162b2 against SARS-CoV-2 infection in residents of Long-Term Care Facilities
www.thelancet.com/journals/laninf/article/PIIS1473-3099(21)00289-9/fulltext

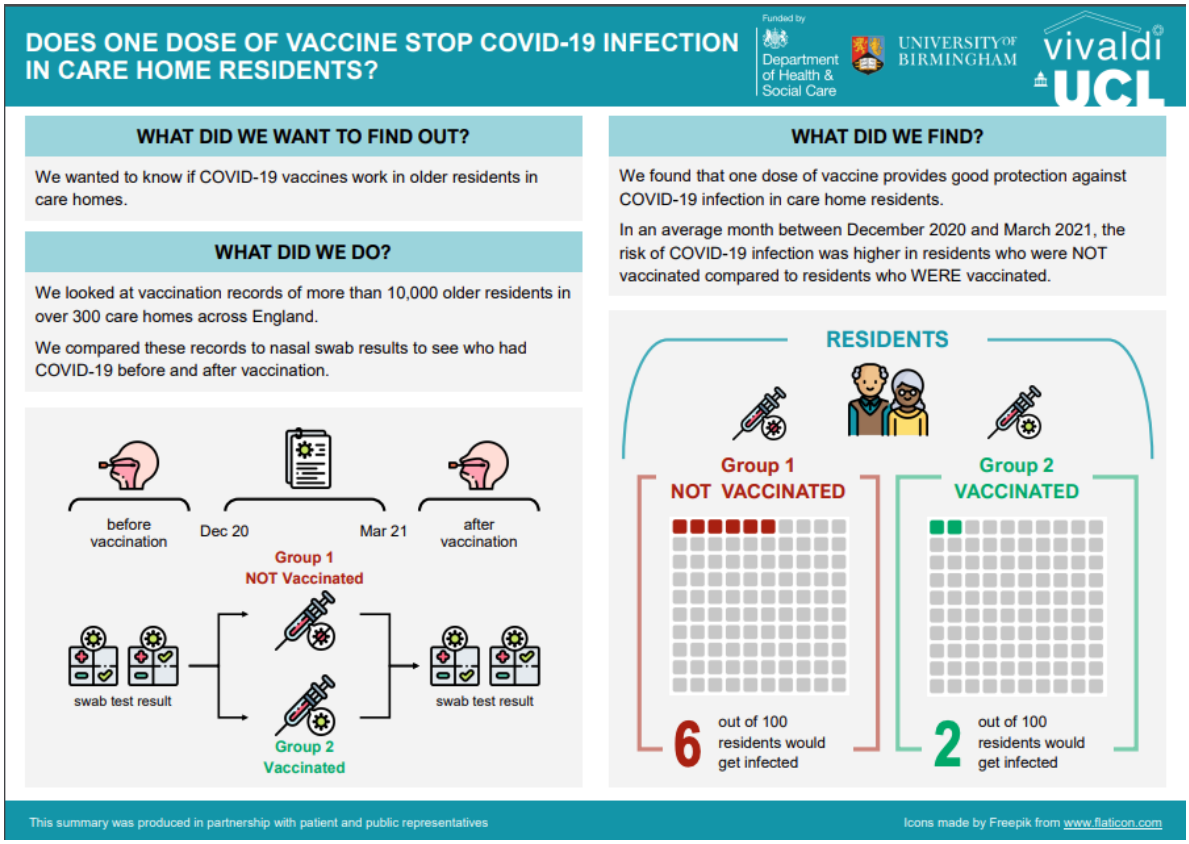
CONTACT US
VIVALDI Study
UCL Institute of Health Informatics
222 Euston Rd
London, NW1 2DA

Email
vivaldi.covid19@ucl.ac.uk

Website
www.ucl.ac.uk/health-informatics/research/vivaldi-study

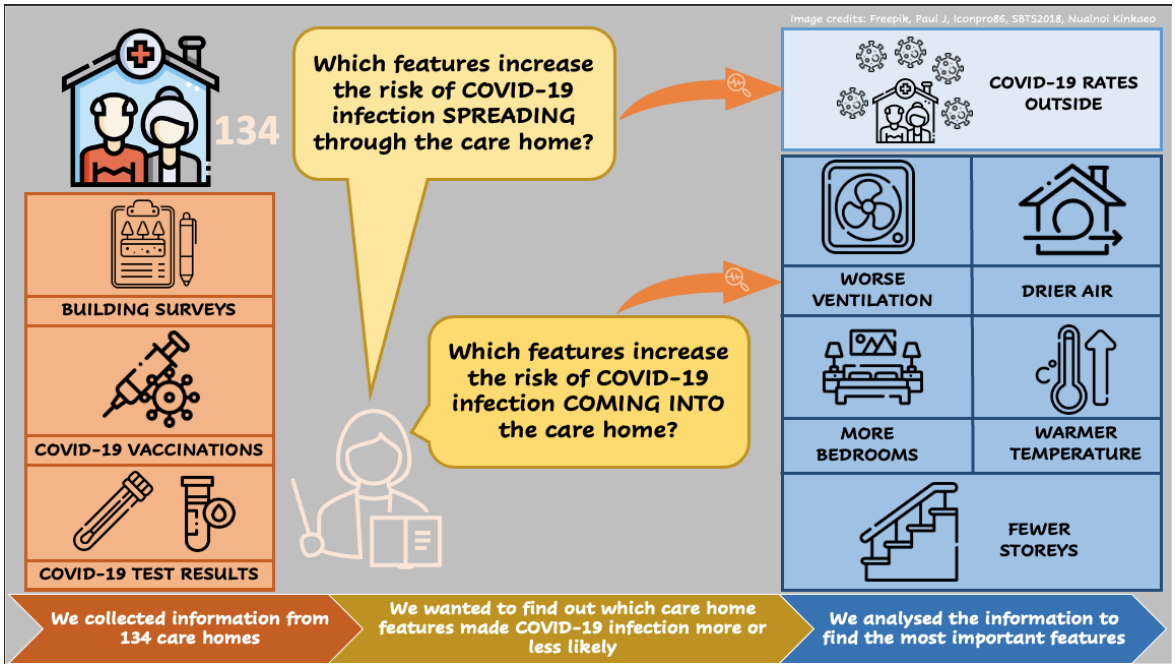
Icons made by Freepik from www.flaticon.com

Funded by Department of Health & Social Care | UNIVERSITY OF BIRMINGHAM | vivaldi



Although these are lay summaries of analyses that fall outside of my thesis (described in Section 7.4), I have used the feedback from these focus groups to develop an infographic summarising my results from Chapter 6 which I will use in engagement work following on from my PhD, Figure 7.2.

Figure 7-2 :Infographic summarising environmental risk factor analysis.



7.2.2 Input into study design and materials

Although rapid initial set-up limited the extent of PPIE and stakeholder consultation regarding study design and materials, their input helped shape revisions and subsequent applications over the rest of the study. Care home staff contributed to the development of the application to the CAG, to allow ongoing access to individual-level anonymised data without participant consent, that I led from November 2021 to March 2022 (Chapter 3). I have listed some of the key engagement activities that I led or participated in as part of this application.

Preparing the first draft of the CAG application, we regularly spoke to five care home managers about the feasibility of data sharing. They provided feedback on the appropriateness of the proposed process to enable individuals to opt-out of sharing data with the study team and on the participant-facing study materials (leaflets and posters).

In May 2022, I visited a large inner-city for-profit care home taking part in the study with the study CI. We discussed the issues around consent for data sharing with six senior manager and regional directors. They strongly agreed with the importance of research in the care setting and that studies like VIVALDI had been crucial to informing

preventive measures throughout the pandemic, particularly when conducted on a large-scale. Whilst they thought that work like this should continue beyond the pandemic, they expressed concerns about the feasibility of consenting all staff and residents on an individual basis, as it would place unmanageable strain on the care home, given the high staff turnover and issues with fluctuating capacity amongst residents.

In November 2022, we visited a small family-run independent care home specialising in dementia care in the East Midlands. We discussed the clarity and impact of leaflets and posters describing and the acceptability of data sharing with approximately ten care staff. They agreed with data sharing in principle however they stressed the importance of ensuring that data are stored anonymously and securely. Overall, they were supportive of the use of their data to answer important research questions.

These discussions, both in person and online, have provided insights into some of the real-life issues and concerns experienced in care homes which can continue to inform my research.

7.2.3 Co-production

As hospitalisations and mortality from SARS-CoV-2 began to decline from May 2021, restrictions on non-essential visits to care homes were eased.[160] This enabled myself and the study team to start visiting care homes and engage with the staff, residents, and their families about the challenges they had faced over the pandemic and identify what research could be most beneficial to them. The first of these visits that I attended was in May 2022 and is described under Section 7.2.1. The attendance of senior policymakers and public health officials demonstrated the impact of the VIVALDI study to care home staff and residents and the interest that senior policymakers were taking in social care research, which was highly valued by all who attended.

To date, we have visited five rural and urban care homes across the country in London, Oxfordshire, East Sussex, Lincolnshire, and Northumbria. They were owned by large for-profit and independent providers, two of which specialised in dementia care and one in complex needs. As a study team, we have been working with Niccola

Hutchinson-Pascal, head of the Co-Production Collective[420] hosted within UCL, to gain a broader understanding of how best to engage with different types of stakeholders. Using a shared reflective log to document our experiences from each visit, we have reflected on successes and areas for improvement between visits. Although these visits have not been directly related to the work presented in the thesis, they have expanded my knowledge of the care sector and therefore strongly informed my interpretation of findings and their policy relevance.

Some of the major challenges that we have identified on these visits were that engaging with residents can be difficult due to significant variation in cognition and interest between residents. Identifying residents to approach, with the help of a staff member who knows them well, can help overcome this. As many care homes specialise in dementia care, we felt it was important to engage residents more widely. Therefore, we have been considering activities and conversational topics that are interesting for the resident. In addition, we have been offering to take residents out into the garden or restaurant to make the interaction more equitable as we are able to assist the resident with an activity that they enjoy and may not have been able to perform alone.

In this way, we are moving away from the traditional paternalistic approach to research and trying to gain insights into what is important to the residents. We plan to continue to develop and co-produce activity packs that we can take with us on care home visits, based on activities that have worked well. It is also clear that having the support of senior staff like the care home manager, is very helpful as they can encourage staff and residents to speak to researchers. We have found that the visits work best when we can independently walk around the care home and speak to residents and staff. Although this was not possible when there were greater concerns about SARS-CoV-2 transmission, this approach has given us better insights into the different settings for our research. As there is no standard approach to co-production with care homes, using our reflective log to improve our approach, we are developing our own ‘toolkit’. These experiences are adding depth to our research, and I plan to apply these principles to my future work by ensuring that the stakeholders are partners over the whole research process.

7.3 Key strengths and limitations of study and lessons for the future

Although I have described the main strengths and limitations associated with each study in the relevant chapters, I have summarised significant general points below.

As described in Chapter 3, the VIVALDI cohort was made possible through the opportunity to reliably identify both staff and residents of care homes by linking test results to the care homes where they were performed. The introduction of the COPI notice[263] provided a legal basis to process data from all staff and residents therefore, we could include the majority of staff and residents and make our findings more generalisable. Once the COPI notice expired, we obtained approval from HRA CAG to continue data collection and linkage. However, now testing has stopped there is currently no reliable way to accurately identify care home residents in routine data, as outlined in Chapter 1. In addition, the ability to identify asymptomatic infections has been very valuable.

VIVALDI was established very rapidly (four weeks) in response to the public health crisis that emerged in care homes. This was possible because of accelerated ethical approval processes in place for COVID-19 associated research[421], assistance from professional services such as procurement, rapid funding decisions, and new processes to expedite contracting (e.g., outsourcing to an external legal firm). The workload associated with rapidly putting in place the necessary contractual arrangements to enable data sharing, including a separate contract with each provider (over twenty separate contracts), was substantial. Rapid data linkage was also facilitated by storage of data in the NHS COVID-19 Foundry.[258] To overcome national shortages in clinical equipment like PPE, colleagues at DHSC supported the study by helping to secure and distribute equipment.

The research team and I worked very closely with policymakers and government officials in DHSC and PHE/UKHSA and regularly presented findings in national meetings. Research questions were relevant and timely (the study CI regularly attended national data debrief meetings and was a core SAGE care home working group member), which meant our results could inform national and international policy. Results also received significant publicity in the national and international media,

providing important lessons in working with media teams to communicate results to the public. However, this also meant that there was significant pressure to complete analyses which often had to be delivered to a high standard at extremely short notice and very tight deadlines. This impacted on capacity to plan and carry out analyses in the medium to long-term as research priorities changed very rapidly and policy-relevant work took precedence.

The study benefited from excellent levels of engagement and recruitment from a range of different types of providers and care homes across the country. COVID-19 related research was prioritised in care homes and nationally which encouraged study participation. Key to the study's success were the project managers from each of the four main care providers in the study (FSHC, HC-One, OSJCT, FOTE), for whom specific funding had been ringfenced in the study budget. This ensured that the study team had a single internal point of contact who was familiar with the organisation's processes, in turn building trust. Senior government Ministers were influential in gaining support of large care providers as they approached CEOs regarding participation in VIVALDI, following a request from the study CI. In addition, research bodies like the Clinical Research Network (CRN) supported recruitment by reaching out to their ENRICH network of research-ready care homes.[256]

National restrictions on movement restricted the ability of the study team to visit sites, therefore we relied heavily on senior care staff, project managers, and later the CRN, to consent participants. Some reflections about this approach are described in Section 7.9, however I am currently leading an audit of consent form completion to evaluate the effectiveness of the study's consenting approach and inform future practice.

Key limitations of the study design were related to data availability and its quality, as described in Chapter 3. Accessing NHS numbers was a key advantage of the study as it allowed data linkage for all staff and residents, therefore limiting bias from under-representative sampling given the high proportion of the population who may have been excluded as they do not have capacity to consent. Data on ethnicity and co-morbidities were unreliable (as described in Chapter 5) and future work should focus on approaches to improving the quality of these. As VIVALDI only considered residents who were older adults, study findings cannot be extended to younger social

care populations such as people with learning difficulties. Building on learning from the pandemic, many care homes are now ensuring that they hold NHS IDs for all their residents which will facilitate integration of health with social care.

The Abbott immunoassay was widely used for anti-nucleocapsid antibody detection at the start of the study. Subsequent evidence has highlighted issues with assay sensitivity and susceptibility to antibody waning.[352] This evidence emerged after the study had already commenced testing therefore switching assay at this stage would have impacted on analyses that compared or pooled results. Although I accounted for this by using a lower positivity threshold and incorporating available MSD results, this highlights that it is difficult to choose a test before performance characteristics have been fully identified. For the final phase of the study (April 2022 - March 2023), we switched to the Elecsys® Anti-SARS-CoV-2 immunoassay (Roche Diagnostics International Ltd, Rotkreuz, Switzerland) for nucleocapsid and spike antibody detection, reported to have higher sensitivity and lower susceptibility to waning than the Abbott assay.[285,422,423]

The low availability of sequencing data from nasopharyngeal viral isolates precluded our ability, and that of PHE/UKHSA, to answer important questions about the chain of transmission into and within care homes. Although we collaborated with the COG-UK, despite substantial efforts to ensure samples could be cherry-picked at testing laboratories, sequencing coverage remained low. Sample selection was challenging given the volume of sample processed by the laboratories and the speed with which the network was established, however lessons about operational management may inform future national testing programmes.

Finally, it was difficult to fully account for confounding and bias that impacted on the study outcomes, as many epidemiological and policy changes occurred simultaneously. SARS-CoV-2 infection rates changed rapidly and were affected by national policies including lockdowns and vaccination rollout, as well as other temporal factors relating to meteorological changes, population movement, and behaviour (such as adherence to IPC recommendations like PPE use). Although I limited selection bias by including all individuals who were tested in participating care homes, there were differences in testing rates between staff and resident groups may have

introduced case-counting or ascertainment bias. Although testing was mandatory during most of the study, this was monitored less carefully later in the pandemic as evidenced by a drop in tests that I identified, which may reflect changes in attitudes towards infection or testing fatigue. Residents having end-of-life care were less likely to be tested in view of the discomfort caused and may not have appeared in my dataset, introducing selection bias. Fluctuations in population and care home level immunity (described in Chapters 4 and 5) impacted on susceptibility to infection and are very difficult to measure. I have considered time-varying factors where possible however, this highlights the challenges of accurately accounting for all significant factors in observational routine data. This is particularly important during a period of substantial instability such as a pandemic and has been recognised, particularly in relation to vaccine efficacy studies.[424,425]

7.4 Wider application of VIVALDI study cohort

Over the pandemic, VIVALDI has made valuable contributions to national policy decisions. Below I have listed VIVALDI research outputs that I have contributed to, manuscripts are included in the Supplementary Appendix.

7.4.1 Reinfection in care home staff and residents after 10 months of the pandemic (February 2021)

In collaboration with the VIVALDI team, I led an analysis estimating the risk of SARS-CoV-2 reinfection in care home staff and residents in the second wave of infection. Using PCR and nucleocapsid-antibody results from the first two waves, a statistician in the VIVALDI team (TP) modelled the hazards of infection in 1417 infection-naïve (nucleocapsid antibody-negative) and 694 previously infected (antibody-positive) individuals. Previously infected residents had an 85% lower risk of infection between October 2020 and February 2021 than infection-naïve residents, whereas previously infected staff had a 60% lower infection risk than infection-naïve staff. This was the largest analysis of reinfection risk in care homes and was performed before vaccine efficacy in this population had been described. Our estimate is comparable to results from a smaller greater London cohort of 13 care homes conducted over the same period that found 85% protection from reinfection amongst 1625 participants.[426]

I was joint first author of the manuscript, published in the *Lancet Healthy Longevity*. It received extensive coverage in the national media, including a live interview with me on the Times radio breakfast show.

7.4.2 Severity of infection with Omicron variant compared with Delta variant in care home residents.

(January 2022)

At the time of Omicron variant emergence, population-level infection control measures had largely been relaxed. Data on Omicron infection severity in the most vulnerable groups was urgently required to inform decisions on whether restrictions should be re-introduced.

I led an analysis, in collaboration with the study team, that compared the clinical severity of Omicron infections to the previously dominant Delta variant in care home residents. I compared and modelled the risk of hospitalisation within 14 days and death within 28 days of a SARS-CoV-2 diagnosis between the Delta-dominant (1st September 2021-12th December 2021) and the Omicron-dominant period (13th December 2021-26th January 2022), which was possible because of the rapid shift in variant dominance. As absence of detectable S-gene (S-gene target failure, SGTF) on PCR testing was found to be a reliable marker of Omicron variant in the Delta-dominant period, I performed a sub-analysis on samples with confirmed variant based on SGTF and/or sequencing, where available, to demonstrate fidelity. In a cohort of 1864 residents infected with Omicron and 400 with Delta, I found risk of hospitalisation was 50% lower and risk of death was 66% lower in Omicron compared with Delta infections.

These data were presented to governmental advisory groups including NERVTAG, COVID Immunology Consortium, SAGE care home working group. The analysis informed decisions against reimposing national restrictions, given the relative drop in risk of severe outcomes that we demonstrated. The study has been published in the *Lancet Healthy Longevity* where I am listed as first author.

7.4.3 Effectiveness of first dose vaccine against infection in care home residents

(March 2021)

Although care home residents were prioritised for vaccination, older people were under-represented and care homes residents were excluded from clinical trials of vaccine efficacy. As such, data on the real-world efficacy of vaccines was urgently needed to inform policy decisions. I led a substantial ethics amendment to enable vaccine efficacy studies within the VIVALDI cohort.

We estimated first dose vaccine effectiveness from 10,412 care home residents who were followed up for seven weeks, of whom 9160 (88%) had received one vaccine dose. We found first-dose vaccine efficacy was 56% after 28 days with similar protection regardless of vaccine type.

I contributed to data collection, study design, and manuscript writing and was second author on the final paper. This study was the first large-scale report on vaccine efficacy from older residents and reinforced decisions around care home prioritisation for vaccination. This was published in *Lancet Infectious Diseases* and had extensive media coverage including a live interview on BBC lunchtime news with the study CI and selection by the Lancet for research highlights at ECCMID in 2021.

7.4.4 Immunological responses to infection and vaccination in care home staff and residents

(March 2021 – present)

I have co-authored four papers with colleagues at the University of Birmingham who have investigated immune responses to SARS-CoV-2 infection and vaccination in the VIVALDI cohort. Findings have demonstrated robust responses to vaccination amongst care home residents that are comparable to those from staff, however there is evidence of waning, and the magnitude and longevity of responses is significantly greater amongst previously exposed residents when compared with the infection-naïve group. These data made a significant contribution to the literature on humoral and cellular responses to SARS-CoV-2 in older people and care home residents and have informed UK vaccination policy. In these studies, I have played a significant role

in study design, project management, and manuscript drafting and have been responsible for linkage to demographic and clinical data.

7.4.5 Creating a Bioresource of samples from VIVALDI participants.

Serum samples donated by VIVALDI participants are valuable to researchers as i) they are from frail, older individuals who are under-represented in research due to challenges with consent and access; ii) longitudinal collection of samples from early in the pandemic allows monitoring of responses to infection over time; iii) linkage to metadata including infection and outcomes data provides important context for meaningful interpretation of results. Existing biobanks of biological samples linked to longitudinal outcomes data from older adults are sparse:[427–429] For example the largest biobank in this country, UK Biobank, excludes participants older than 69 years and does not link sequential samples to data on clinical infections.[430,431]

To optimise the research potential of the VIVALDI samples, I am working with the VIVALDI team to establish a biobank of residual serum samples linked to infection and outcomes data which will be available to other research groups. To date, I have led on the selection and contracting with UK Biocentre, who will store samples and overseen the sample transfer from TDL. To ensure participants have consented to longer-term use of their samples, I am auditing consent forms with colleagues at University of Birmingham. Following submission of my thesis, I will finalise this cohort and metadata and consider longer term approaches to management of the biobank. We plan to publish the cohort profile in a peer-reviewed journal and apply for small grants to support ongoing sample storage costs.

7.5 Impact

The impact of individual analyses has been described in the relevant sections of this thesis. Overall, the VIVALDI cohort has set a precedent for future research in care homes which capitalises on routine data, is timely, and relevant. We have been able to overcome key barriers to research in these settings, outlined in Section 7.3. Through establishment of collaborations with Providers and policymakers in DHSC and UKHSA, the study has consistently informed key policy decisions and raised awareness of the benefits of research in care homes which will affect future practice.

Results from my analyses have received national and international media coverage, been published in high-impact peer-reviewed journals and have directly informed policy decisions. Stored residual samples from this unique population will provide important opportunities for future research focussing on the aging population. Despite the devastating context that the study was borne out of, it has showcased how data-driven research can benefit stakeholders and effect change.

7.6 Future work

The outputs from this thesis and the wider VIVALDI study, lay the foundations for addressing further research questions in care home residents and staff.

In 2018, the WHO identified top priorities for the ageing population in the WHO *Global strategy and action plan on ageing and health.*[432] One of these included *"improving measurement, monitoring, and research on Healthy Ageing"* and has prompted calls for a national care home minimum dataset akin to some of the systems used internationally. An example of this is the detailed and multi-domain Minimum Data Set (MDS) developed by a consortium of clinicians and scientists (InterRAI), used in Europe, North America, and New Zealand to record individual-level resident data for research.[433–436] Research collaborations in the UK such as Developing resources And minimum data set for Care Homes' Adoption (DACHA), that preceded the pandemic, have been working to overcome fragmentation within health and social care to develop a similar dataset [433] however this model of individual-level informed consent places significant strain on the busy care home and may exclude key populations.

The asymptomatic screening programme that enabled the creation of a VIVALDI care home 'registry' was terminated in 2022. Building on this experience, we are exploring the feasibility of collecting NHS numbers directly from digital record suppliers to deliver a future programme of infection surveillance and research in care homes that is underpinned by linked data. Data from this network could be used to research efficacy of SARS-CoV-2 vaccines, incidence and outcomes from SARS-CoV-2 and other infections, as well as analysing patterns of antimicrobial resistance amongst bacteria from clinical isolates. This data infrastructure could also facilitate trials of public health interventions to reduce infections.

Over the last year of my PhD, I have participated in a programme of engagement and co-production activities with social care stakeholders to understand the acceptability, risks, and benefits of this approach. Reponses have been overwhelmingly positive as experiences from the pandemic have showcased the significant impact that data can have on infection prevention. To date, over 800 providers have provisionally volunteered to share data. Over the summer months of 2023, I will be co-leading an application to the HRA CAG with the VIVALDI CI to gain legal support for accessing confidential information about residents under Regulation 5 of the Health Service (Control of Patient Information) Regulations 2002.[437] Given the scale of this project, it is important that the appropriate approvals are in place from the start to ensure that it can be sustainable in the long-term. The hope is that this will forge lasting partnerships between care providers, researchers, and policymakers and drive a new research agenda within social care.

7.7 Conclusions

Over my PhD, I have described variations in the burden of SARS-CoV-2 infection in care homes and considered facility-level factors associated with ingress and transmission of infection according to agent, host, and environment components of the epidemiological triad. This work was performed during a global pandemic when knowledge about the characteristics and of this novel virus was sparse. I contributed to a project that demonstrated the feasibility of generating research in care homes at pace and scale. I have collaborated with national policymakers and providers to increase relevance and impact of my findings. This has laid the foundations for future research on how the built environment of care homes impacts on infection transmission.

As the pandemic eases, care homes must adapt to living with COVID-19, as well as existing widely circulating viral threats such as Influenza. Whilst the impacts of the COVID-19 pandemic in care homes will continue to affect the community for years to come, the volume of research generated from the care sector over this period has been unprecedented. As such, it is important to build on the positive legacy from this devastating time and prioritise data-informed and co-produced research in care

settings. This will provide a key evidence base for preventive strategies to update policy, identify research priorities, and prepare for future pandemics. In this thesis, I have demonstrated that rapidly establishing this type of infrastructure to support national policy is possible. This work showcases the opportunities for data sharing in the longer term and the potential benefits to the residents and staff in care homes who are often excluded from research.

7.8 Outputs from this chapter

Krutikov, M., Palmer, T., Tut, G., et al. (2021). **Incidence of SARS-CoV-2 infection according to baseline antibody status in staff and residents of 100 long-term care facilities (VIVALDI study): a prospective cohort study.** *Lancet Healthy Longev* 2021; 2(6): e362–70. DOI: 10.1016/s2666-7568(21)00093-3.

Krutikov, M., Stirrup, O., Nacer-Laidi, H., et al. (2022). **Outcomes of SARS-CoV-2 Omicron infection in residents of long-term care facilities in England (VIVALDI): a prospective, cohort study**. *Lancet Healthy Longev* 2022; 3(5): e347-e355. DOI: 10.1016/S2666-7568(22)00093-9

Shrotri M, Krutikov M, Palmer T et al. (2021) **Vaccine effectiveness of the first dose of ChAdOx1 nCoV-19 and BNT162b2 against SARS-CoV-2 infection in residents of long-term care facilities in England (VIVALDI): a prospective cohort study.** *Lancet ID* 2021; 21(11): 1529-1538. DOI: 10.1016/S1473-3099(21)00289-9

7.9 Personal reflections on delivering research in a pandemic.

Although I gained invaluable skills and insights from my experience of setting up and running VIVALDI, there have been significant challenges associated with undertaking a PhD in a core national surveillance study during a pandemic.

As the only project manager for the first six months of the study (some of this time preceded my PhD), I was responsible for all study finances, oversaw remote site set-up, day-to-day logistics, and the development of operational paperwork (such as SOPs for phlebotomists and care home sites). As there were up to five visits daily, coordinating the parties involved (sites, phlebotomists, laboratories, and couriers)

required good organisation and time-management. I gained experience in procurement by leading the tendering process for phlebotomy services. I also worked with the CI and institute manager to advertise and interview for project manager and administrator positions for the study going forwards, providing experience in recruitment. I was able to rapidly gain skills that will be hugely valuable for my future academic career, however focussing on my own analyses, particularly early, on was challenging.

Having previously worked in acute healthcare, I had limited experience of social care settings. A key difference was our process of consenting compared with research practice within the NHS. As described in Chapter 3, our approach was borne out of the conditions at the time and was fairly novel. I have obtained feedback from sites around issues that they experienced with this, including that some family members were unable to attend the home to complete consultee forms during lockdown, relatives shielding at home could not return signed forms by post, difficulties with opening and editing consultee forms within e-mails, manual dexterity or visual impairment limiting ability to sign. I will take these lessons forward into future social care research, chiefly the importance of developing consent processes in partnership with stakeholders. This will include exploring options for electronic consent forms, considering alternatives to handwriting for those with physical conditions that impair their ability to read and write (such as using impartial witnesses), and introducing short courses on *Good Clinical Practice* for care staff that can count towards their continuing professional development.

I have also learned about data pipelines and management over the course of my PhD. Although the initial data pipeline was designed by the study CI and built by data engineers within the Foundry, the contract with these data engineers ended shortly after. We had support from data engineers at DHSC/UKHSA, however staff changed rapidly and there were delays while new engineers became familiarised with the platform. I very quickly learned about our data pipeline within Foundry and become accustomed with Contour, which allowed me to develop solutions for data quality issues that were affecting urgent analyses. Alongside the CI, I was also responsible for obtaining the necessary approvals for bringing in new datasets from NHSE. As I have been Information Asset Administrator for the Vivaldi share in the UCL Data Safe

Haven throughout the study, I have worked closely with the UCL Information Governance department to ensure the study is compliant and develop SOPs for secure data transfer with colleagues at University of Birmingham. The knowledge that I have gained has allowed me to take a leading role in the successful CAG application and future data platform development.

Overall, I have had a unique opportunity to gain rapid experience in conducting policy-relevant research that has immediate impacts on the population. As we adapt to living with COVID-19 and return to our previous pace of research, I hope to take forward some of the key lessons that I have learned into my future research career. This includes the importance of working in partnership with key stakeholders to ensure that my research is timely, relevant, and accessible to the people and the communities who will be affected by it.

References

1 Liang ST, Liang LT, Rosen JM. COVID-19: a comparison to the 1918 influenza and how we can defeat it. *Postgrad Med J* 2021; **97**: 273–4.

2 WHO. Archived: WHO Timeline - COVID-19. 2020. https://www.who.int/news/item/27-04-2020-who-timeline---covid-19 (accessed Feb 10, 2023).

3 ECDC. Clinical characteristics of COVID-19. 2022. https://www.ecdc.europa.eu/en/covid-19/latest-evidence/clinical (accessed Feb 10, 2023).

4 WHO. WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020. 2020. https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020 (accessed Feb 18, 2022).

5 WHO. WHO Coronavirus (COVID-19) Dashboard With Vaccination Data. 2020. https://covid19.who.int/ (accessed May 16, 2023).

6 Bergeri I, Whelan MG, Ware H, *et al.* Global SARS-CoV-2 seroprevalence from January 2020 to April 2022: A systematic review and meta-analysis of standardized population-based studies. *PLoS Med* 2022; **19**: e1004107.

7 The British Academy. The COVID Decade: understanding the long-term societal impacts of COVID-19. 2021 https://www.thebritishacademy.ac.uk/publications/covid-decade-understanding-the-long-term-societal-impacts-of-covid-19/ (accessed March 21, 2023).

8 Finch D, Tinson A. The continuing impact of COVID-19 on health and inequalities. 2022 https://www.health.org.uk/publications/long-reads/the-continuing-impact-of-covid-19-on-health-and-inequalities (accessed March 21, 2023).

9 WHO. The impact of COVID-19 on global health goals. 2021. https://www.who.int/news-room/spotlight/the-impact-of-covid-19-on-global-health-goals (accessed March 21, 2023).

10 Mutambudzi M, Niedwiedz C, Macdonald EB, *et al.* Occupation and risk of severe COVID-19: prospective cohort study of 120 075 UK Biobank participants. *Occup Environ Med* 2021; **78**: 307–14.

11 Starke KR, Reissig D, Petereit-Haack G, Schmauder S, Nienhaus A, Seidler A. The isolated effect of age on the risk of COVID-19 severe outcomes: a systematic review with meta-analysis. *BMJ Glob Health* 2021; **6**: e006434.

12 Bhaskaran K, Bacon S, Evans SJ, *et al.* Factors associated with deaths due to COVID-19 versus other causes: population-based cohort analysis of UK primary care data and linked national death registrations within the OpenSAFELY platform. *The Lancet Regional Health - Europe* 2021; **6**: 100109.

13 Holt H, Talaei M, Greenig M, *et al.* Risk factors for developing COVID-19: a population-based longitudinal study (COVIDENCE UK). *Thorax* 2022; **77**: 900–12.

14 Williamson EJ, Walker AJ, Bhaskaran K, *et al.* Factors associated with COVID-19-related death using OpenSAFELY. *Nature 2020 584:7821* 2020; **584**: 430–6.

15 Kim HJ, Hwang H, Hong H, Yim JJ, Lee J. A systematic review and meta-analysis of regional risk factors for critical outcomes of COVID-19 during early phase of the pandemic. *Scientific Reports 2021 11:1* 2021; **11**: 1–13.
16 Office for National Statistics. Coronavirus (COVID-19) Infection Survey. 2023. https://www.ons.gov.uk/surveys/informationforhouseholdsandindividuals/householdandindividualsurveys/covid19infectionsurvey (accessed March 21, 2023).
17 GOV.UK. Cases in England | Coronavirus in the UK. 2023. https://coronavirus.data.gov.uk/details/cases?areaType=nation&areaName=England (accessed Jan 26, 2023).
18 V'kovski P, Kratzel A, Steiner S, Stalder H, Thiel V. Coronavirus biology and replication: implications for SARS-CoV-2. *Nature Reviews Microbiology 2020 19:3* 2020; **19**: 155–70.
19 Wu A, Peng Y, Huang B, *et al.* Genome Composition and Divergence of the Novel Coronavirus (2019-nCoV) Originating in China. *Cell Host Microbe* 2020; **27**: 325–8.
20 da Silva Torres MK, Bichara CDA, de Almeida M de N do S, *et al.* The Complexity of SARS-CoV-2 Infection and the COVID-19 Pandemic. *Front Microbiol* 2022; **13**: 30.
21 Thakur S, Sasi S, Pillai SG, *et al.* SARS-CoV-2 Mutations and Their Impact on Diagnostics, Therapeutics and Vaccines. *Front Med (Lausanne)* 2022; **9**: 35.
22 Vetter P, Vu DL, L'Huillier AG, Schibler M, Kaiser L, Jacquerioz F. Clinical features of covid-19. *BMJ* 2020; **369**. DOI:10.1136/BMJ.M1470.
23 Subramanian A, Nirantharakumar K, Hughes S, *et al.* Symptoms and risk factors for long COVID in non-hospitalized adults. *Nature Medicine 2022 28:8* 2022; **28**: 1706–14.
24 Bartheld CS von, Wang L. Prevalence of Olfactory Dysfunction with the Omicron Variant of SARS-CoV-2: A Systematic Review and Meta-Analysis. *Cells 2023, Vol 12, Page 430* 2023; **12**: 430.
25 Mehta P, McAuley DF, Brown M, Sanchez E, Tattersall RS, Manson JJ. COVID-19: consider cytokine storm syndromes and immunosuppression. *The Lancet* 2020; **395**: 1033–4.
26 Ma Q, Liu J, Liu Q, *et al.* Global Percentage of Asymptomatic SARS-CoV-2 Infections Among the Tested Population and Individuals With Confirmed COVID-19 Diagnosis: A Systematic Review and Meta-analysis. *JAMA Netw Open* 2021; **4**: e2137257–e2137257.
27 Chen W, Zhang N, Wei J, Yen HL, Li Y. Short-range airborne route dominates exposure of respiratory infection during close contact. *Build Environ* 2020; **176**: 106859.
28 Ong SWX, Tan YK, Chia PY, *et al.* Air, Surface Environmental, and Personal Protective Equipment Contamination by Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) From a Symptomatic Patient. *JAMA* 2020; **323**: 1610.
29 Greenhalgh T, Jimenez JL, Prather KA, Tufekci Z, Fisman D, Schooley R. Ten scientific reasons in support of airborne transmission of SARS-CoV-2. *The Lancet* 2021; **397**: 1603–5.
30 GOV.UK. COVID-19: epidemiology, virology and clinical features. 2022. https://www.gov.uk/government/publications/wuhan-novel-coronavirus-background-information/wuhan-novel-coronavirus-epidemiology-virology-and-clinical-features#transmission (accessed Feb 10, 2023).

31 Spencer EA, Heneghan CJ, Brassey J, *et al.* SARS-CoV-2 and the role of orofecal transmission: a systematic review. *F1000Research 2021 10:231* 2021; **10**: 231.
32 Plüddemann A, Spencer EA, Heneghan CJ, *et al.* SARS-CoV-2 and the role of vertical transmission from infected pregnant women to their fetuses: systematic review. www.cebm.net/evidence-synthesis (accessed Feb 8, 2023).
33 Arav Y, Klausner Z, Fattal E. Theoretical investigation of pre-symptomatic SARS-CoV-2 person-to-person transmission in households. *Scientific Reports 2021 11:1* 2021; **11**: 1–12.
34 Jefferson T, Spencer EA, Brassey J, *et al.* Transmission of Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2) from pre and asymptomatic infected individuals: a systematic review. *Clinical Microbiology and Infection* 2022; **28**: 178–89.
35 Mugglestone MA, Ratnaraja N v., Bak A, *et al.* Presymptomatic, asymptomatic and post-symptomatic transmission of SARS-CoV-2: joint British Infection Association (BIA), Healthcare Infection Society (HIS), Infection Prevention Society (IPS) and Royal College of Pathologists (RCPath) guidance. *BMC Infect Dis* 2022; **22**: 1–15.
36 Xiang Y, Jia Y, Chen L, Guo L, Shu B, Long E. COVID-19 epidemic prediction and the impact of public health interventions: A review of COVID-19 epidemic models. *Infect Dis Model* 2021; **6**: 324–42.
37 Zhao S, Tang B, Musa SS, *et al.* Estimating the generation interval and inferring the latent period of COVID-19 from the contact tracing data. *Epidemics* 2021; **36**: 100482.
38 Rai B, Shukla A, Dwivedi LK. Estimates of serial interval for COVID-19: A systematic review and meta-analysis. *Clin Epidemiol Glob Health* 2021; **9**: 157.
39 Alene M, Yismaw L, Assemie MA, Ketema DB, Gietaneh W, Birhan TY. Serial interval and incubation period of COVID-19: a systematic review and meta-analysis. *BMC Infect Dis* 2021; **21**: 1–9.
40 Backer JA, Eggink D, Andeweg SP, *et al.* Shorter serial intervals in SARS-CoV-2 cases with Omicron BA.1 variant compared with Delta variant, the Netherlands, 13 to 26 December 2021. *Eurosurveillance* 2022; **27**: 2200042.
41 Lauer SA, Grantz KH, Bi Q, *et al.* The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. *Ann Intern Med* 2020; **172**: 577–82.
42 Hill V, Du Plessis L, Peacock TP, *et al.* The origins and molecular evolution of SARS-CoV-2 lineage B.1.1.7 in the UK. *Virus Evol* 2022; **8**. DOI:10.1093/VE/VEAC080.
43 Kemp SA, Collier DA, Datir RP, *et al.* SARS-CoV-2 evolution during treatment of chronic infection. *Nature 2021 592:7853* 2021; **592**: 277–82.
44 Peck KM, Lauring AS. Complexities of Viral Mutation Rates. 2018. DOI:10.1128/JVI.
45 Moore JP, Offit PA. SARS-CoV-2 Vaccines and the Growing Threat of Viral Variants. *JAMA* 2021; **325**: 821–2.
46 McLean G, Kamil J, Lee B, *et al.* The Impact of Evolving SARS-CoV-2 Mutations and Variants on COVID-19 Vaccines. *mBio* 2022; **13**. DOI:10.1128/MBIO.02979-21/SUPPL_FILE/MBIO.02979-21-S0003.DOCX.
47 Pango Network. Helping track the transmission and spread of SARS-CoV-2. 2022. https://www.pango.network/ (accessed May 10, 2023).

48 Rambaut A, Holmes EC, O'Toole Á, *et al.* A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nat Microbiol* 2020; **5**: 1403–7.
49 CDC. SARS-CoV-2 Variant Classifications and Definitions. 2023. https://www.cdc.gov/coronavirus/2019-ncov/variants/variant-classifications.html (accessed March 21, 2023).
50 COVID-19 Genomics UK Consortium. Impact. 2023. https://www.cogconsortium.uk/priority-areas/research/impact/ (accessed March 21, 2023).
51 Furuse Y. Genomic sequencing effort for SARS-CoV-2 by country during the pandemic. *International Journal of Infectious Diseases* 2021; **103**: 305–7.
52 Our World in Data. SARS-CoV-2 variants in analyzed sequences, United Kingdom. 2023. https://ourworldindata.org/grapher/covid-variants-area?country=~GBR (accessed March 13, 2023).
53 Young M, Crook H, Scott J, Edison P. Covid-19: virology, variants, and vaccines. *BMJ Medicine* 2022; **1**: e000040.
54 GISAID. hCov19 Variants. 2023. https://gisaid.org/hcov19-variants/ (accessed March 13, 2023).
55 Public Health England. Investigation of novel SARS-CoV-2 variant Variant of Concern 202012/01 Technical briefing 5. 2020 www.gov.uk/government/publications/nervtag-paper-on-covid-19-variant-of-concern-b117 (accessed March 13, 2023).
56 nCoV-2019 Genomic Epidemiology. Preliminary genomic characterisation of an emergent SARS-CoV-2 lineage in the UK defined by a novel set of spike mutations. 2020. https://virological.org/t/preliminary-genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-the-uk-defined-by-a-novel-set-of-spike-mutations/563 (accessed May 16, 2023).
57 Yang TJ, Yu PY, Chang YC, *et al.* Effect of SARS-CoV-2 B.1.1.7 mutations on spike protein structure and function. *Nature Structural & Molecular Biology 2021 28:9* 2021; **28**: 731–9.
58 Socher E, Conrad M, Heger L, *et al.* Mutations in the b.1.1.7 sars-cov-2 spike protein reduce receptor-binding affinity and induce a flexible link to the fusion peptide. *Biomedicines* 2021; **9**: 525.
59 Volz E, Mishra S, Chand M, *et al.* Assessing transmissibility of SARS-CoV-2 lineage B.1.1.7 in England. *Nature 2021 593:7858* 2021; **593**: 266–9.
60 Davies NG, Jarvis CI, van Zandvoort K, *et al.* Increased mortality in community-tested cases of SARS-CoV-2 lineage B.1.1.7. *Nature 2021 593:7858* 2021; **593**: 270–4.
61 Public Health England. Investigation of SARS-CoV-2 variants of concern in England | Technical Briefing 6. 2021 https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/961299/Variants_of_Concern_VOC_Technical_Briefing_6_England-1.pdf (accessed March 21, 2023).
62 Torjesen I. Covid-19: Delta variant is now UK's most dominant strain and spreading through schools. *BMJ* 2021; **373**: n1445.
63 Public Health England. SARS-CoV-2 variants of concern and variants under investigation in England. 2021 https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/984274/Variants_of_Concern_VOC_Technical_Briefing_10_England.pdf (accessed March 13, 2023).

64 Mlcochova P, Kemp SA, Dhar MS, *et al.* SARS-CoV-2 B.1.617.2 Delta variant replication and immune evasion. *Nature 2021 599:7883* 2021; **599**: 114–9.
65 Shrestha LB, Foster C, Rawlinson W, Tedla N, Bull RA. Evolution of the SARS-CoV-2 omicron variants BA.1 to BA.5: Implications for immune escape and transmission. *Rev Med Virol* 2022; **32**: 32.
66 Burki TK. Lifting of COVID-19 restrictions in the UK and the Delta variant. *Lancet Respir Med* 2021; **9**: e85.
67 Karim SSA, Karim QA. Omicron SARS-CoV-2 variant: a new chapter in the COVID-19 pandemic. *Lancet* 2021; **398**: 2126.
68 Carabelli AM, Peacock TP, Thorne LG, *et al.* SARS-CoV-2 variant biology: immune escape, transmission and fitness. *Nature Reviews Microbiology 2023 21:3* 2023; **21**: 162–77.
69 Hyams C, Challen R, Marlow R, *et al.* Severity of Omicron (B.1.1.529) and Delta (B.1.617.2) SARS-CoV-2 infection among hospitalised adults: A prospective cohort study in Bristol, United Kingdom. *The Lancet Regional Health - Europe* 2023; **25**. DOI:10.1016/j.lanepe.2022.100556.
70 Leiner J, Pellissier V, Hohenstein S, *et al.* Characteristics and outcomes of COVID-19 patients during B.1.1.529 (Omicron) dominance compared to B.1.617.2 (Delta) in 89 German hospitals. *BMC Infect Dis* 2022; **22**: 1–8.
71 Wolter N, Jassat W, Walaza S, *et al.* Early assessment of the clinical severity of the SARS-CoV-2 omicron variant in South Africa: a data linkage study. *Lancet* 2022; **399**: 437–46.
72 Jassat W, Abdool Karim SS, Mudara C, *et al.* Clinical severity of COVID-19 in patients admitted to hospital during the omicron wave in South Africa: a retrospective observational study. *Lancet Glob Health* 2022; **10**: e961–9.
73 Skarbinski J, Wood MS, Chervo TC, *et al.* Risk of severe clinical outcomes among persons with SARS-CoV-2 infection with differing levels of vaccination during widespread Omicron (B.1.1.529) and Delta (B.1.617.2) variant circulation in Northern California: A retrospective cohort study. *The Lancet Regional Health - Americas* 2022; **12**. DOI:10.1016/j.lana.2022.100297.
74 WHO. Laboratory testing of 2019 novel coronavirus (2019-nCoV) in suspected human cases: interim guidance, 17 January 2020. 2020. https://www.who.int/publications/i/item/laboratory-testing-of-2019-novel-coronavirus-(-2019-ncov)-in-suspected-human-cases-interim-guidance-17-january-2020 (accessed May 10, 2023).
75 Li Q, Guan X, Wu P, *et al.* Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia. *New England Journal of Medicine* 2020; **382**: 1199–207.
76 Kevadiya BD, Machhi J, Herskovitz J, *et al.* Diagnostics for SARS-CoV-2 infections. *Nature Materials 2021 20:5* 2021; **20**: 593–605.
77 ECDC. Diagnostic testing and screening for SARS-CoV-2. 2022. https://www.ecdc.europa.eu/en/covid-19/latest-evidence/diagnostic-testing (accessed Feb 8, 2023).
78 Mallett S, Halligan S, Matthew Thompson GP, Collins GS, Altman DG. Interpreting diagnostic accuracy studies for patient care. *BMJ* 2012; **345**. DOI:10.1136/BMJ.E3999.
79 UKHSA. Weekly statistics for NHS Test and Trace (England) 2 to 15 June 2022. 2022.
80 Cevik M, Tate M, Lloyd O, Maraolo AE, Schafers J, Ho A. SARS-CoV-2, SARS-CoV, and MERS-CoV viral load dynamics, duration of viral shedding,

and infectiousness: a systematic review and meta-analysis. *Lancet Microbe* 2021; **2**: e13–22.
81 GOV.UK. New Lighthouse Labs to boost NHS Test and Trace capacity. 2020. https://www.gov.uk/government/news/new-lighthouse-labs-to-boost-nhs-test-and-trace-capacity (accessed March 21, 2023).
82 DHSC. Coronavirus (COVID-19) Scaling up our testing programmes. 2020 https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/878121/coronavirus-covid-19-testing-strategy.pdf (accessed March 21, 2023).
83 GOV.UK. [Withdrawn] COVID-19 testing data: methodology note. 2022. https://www.gov.uk/government/publications/coronavirus-covid-19-testing-data-methodology/covid-19-testing-data-methodology-note (accessed March 21, 2023).
84 Mistry DA, Wang JY, Moeser ME, Starkey T, Lee LYW. A systematic review of the sensitivity and specificity of lateral flow devices in the detection of SARS-CoV-2. *BMC Infect Dis* 2021; **21**. DOI:10.1186/S12879-021-06528-3.
85 Eyre D, Futschik M, Tunkel S, *et al.* Performance of antigen lateral flow devices in the United Kingdom during the Alpha, Delta, and Omicron waves of the SARS-CoV-2 pandemic. *medRxiv* 2022; : 2022.11.29.22282899.
86 Fox T, Geppert J, Dinnes J, *et al.* Antibody tests for identification of current and past infection with SARS-CoV-2. *Cochrane Database of Systematic Reviews* 2022; **2022**. DOI:10.1002/14651858.CD013652.PUB2/EPDF/ABSTRACT.
87 UKHSA. Antibody testing for SARS-CoV-2. 2022.
88 Ridenhour B, Kowalik JM, Shay DK. Unraveling R0: Considerations for Public Health Applications. *Am J Public Health* 2018; **108**: S445.
89 Curran-Sebastian J, Das R, Fyles M, *et al.* Challenges in control of COVID-19: short doubling time and long delay to effect of interventions. *Philosophical Transactions of the Royal Society B* 2021; **376**. DOI:10.1098/RSTB.2020.0264.
90 Delamater PL, Street EJ, Leslie TF, Yang YT, Jacobsen KH. Complexity of the Basic Reproduction Number (R0) - Volume 25, Number 1—January 2019 - Emerging Infectious Diseases journal - CDC. *Emerg Infect Dis* 2019; **25**: 1–4.
91 WHO. Environment, Climate Change and Health | Disease Outbreaks. 2005. https://www.who.int/teams/environment-climate-change-and-health/emergencies/disease-outbreaks (accessed May 12, 2023).
92 Patrinos H, Vegas E, Carter-Rau R. COVID-19 school closures fueled big learning losses, especially for the disadvantaged. 2022. https://blogs.worldbank.org/developmenttalk/covid-19-school-closures-fueled-big-learning-losses-especially-disadvantaged (accessed March 13, 2023).
93 House of Commons Library. Coronavirus: A history of English lockdown laws. 2021. https://commonslibrary.parliament.uk/research-briefings/cbp-9068/ (accessed March 7, 2023).
94 GOV.UK. Chapter 8: non-pharmaceutical interventions. 2023. https://www.gov.uk/government/publications/technical-report-on-the-covid-19-pandemic-in-the-uk/chapter-8-non-pharmaceutical-interventions#introduction (accessed March 13, 2023).
95 Nussbaumer-Streit B, Mayr V, Dobrescu AI, *et al.* Quarantine alone or in combination with other public health measures to control COVID-19: a rapid

review. *Cochrane Database Syst Rev* 2020; **2020**. DOI:10.1002/14651858.CD013574.
96 GOV.UK. NHS COVID-19 app. 2020. https://www.gov.uk/government/collections/nhs-covid-19-app (accessed March 21, 2023).
97 Qureshi Z, Jones J, Temple R, Larwood JP, Greenhalgh T, Bourouiba L. What is the evidence to support the 2-metre social distancing rule to reduce COVID-19 transmission? Oxford, 2020 https://www.cebm.net/covid-19/what-is-the-evidence-to-support-the-2-metre-social-distancing-rule-to-reduce-covid-19-transmission/ (accessed March 21, 2023).
98 GOV.UK. Review of two metre social distancing guidance. 2020. https://www.gov.uk/government/publications/review-of-two-metre-social-distancing-guidance/review-of-two-metre-social-distancing-guidance (accessed March 21, 2023).
99 GOV.UK. [Withdrawn] [Withdrawn] Guidance on social distancing for everyone in the UK. 2020. https://www.gov.uk/government/publications/covid-19-guidance-on-social-distancing-and-for-vulnerable-people/guidance-on-social-distancing-for-everyone-in-the-uk-and-protecting-older-people-and-vulnerable-adults (accessed March 21, 2023).
100 GOV.UK. Ventilation to reduce the spread of respiratory infections, including COVID-19. 2021. https://www.gov.uk/guidance/ventilation-to-reduce-the-spread-of-respiratory-infections-including-covid-19 (accessed March 21, 2023).
101 WHO. Technical specifications of personal protective equipment for COVID-19. 2020. https://www.who.int/publications/i/item/WHO-2019-nCoV-PPE_specifications-2020.1 (accessed March 21, 2023).
102 National Audit Office. The supply of personal protective equipment (PPE) during the COVID-19 pandemic. 2020 https://www.nao.org.uk/wp-content/uploads/2020/11/The-supply-of-personal-protective-equipment-PPE-during-the-COVID-19-pandemic-Summary.pdf (accessed March 21, 2023).
103 legislation.gov.uk. The Health Protection (Coronavirus, Wearing of Face Coverings in a Relevant Place) (England) Regulations 2020. 2020. https://www.legislation.gov.uk/uksi/2020/791/made (accessed March 21, 2023).
104 GOV.UK. Traffic light system: safe return to international travel. 2021. https://www.gov.uk/government/speeches/traffic-light-system-safe-return-to-international-travel (accessed March 21, 2023).
105 Think Global Health. UPDATED: Timeline of the Coronavirus. 2023. https://www.thinkglobalhealth.org/article/updated-timeline-coronavirus (accessed Feb 10, 2023).
106 COVID19 Vaccine Tracker. Vaccines. 2023. https://covid19.trackvaccines.org/vaccines/ (accessed March 21, 2023).
107 GOV.UK. UK authorises Pfizer/BioNTech COVID-19 vaccine. 2020; published online Dec 2. https://www.gov.uk/government/news/uk-authorises-pfizer-biontech-covid-19-vaccine.
108 GOV.UK. Oxford University/AstraZeneca COVID-19 vaccine approved. https://www.gov.uk/government/news/oxford-universityastrazeneca-covid-19-vaccine-approved (accessed March 13, 2023).
109 GOV.UK. Regulatory approval of Spikevax (formerly COVID-19 Vaccine Moderna). 2020. https://www.gov.uk/government/publications/regulatory-approval-of-covid-19-vaccine-moderna (accessed March 13, 2023).

110 UKHSA. Greenbook COVID-19 chapter 14a. 2023. https://www.gov.uk/ (accessed March 21, 2023).
111 GOV.UK. Optimising the COVID-19 vaccination programme for maximum short-term impact. 2020 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 (accessed Feb 12, 2021).
112 Financial Times. Covid-19 vaccine tracker. 2022. https://ig.ft.com/coronavirus-vaccine-tracker/?areas=gbr&cumulative=1&doses=first&populationAdjusted=1 (accessed March 13, 2023).
113 GOV.UK. JCVI statement on the adult COVID-19 booster vaccination programme and the Omicron variant: 7 January 2022. 2022. https://www.gov.uk/government/publications/jcvi-statement-on-the-adult-covid-19-booster-vaccination-programme-and-the-omicron-variant/jcvi-statement-on-the-adult-covid-19-booster-vaccination-programme-and-the-omicron-variant-7-january-2022 (accessed March 13, 2023).
114 Investigators WC for the R-C, Florescu S, Stanciu D, *et al.* Long-term (180-Day) Outcomes in Critically Ill Patients With COVID-19 in the REMAP-CAP Randomized Clinical Trial. *JAMA* 2023; **329**: 39–51.
115 NICE. COVID-19 rapid guideline: Managing COVID-19. 2023. https://www.nice.org.uk/guidance/ng191/resources/covid19-rapid-guideline-managing-covid19-pdf-51035553326 (accessed March 13, 2023).
116 Shepard HM, Phillips GL, Thanos CD, Feldmann M. Developments in therapy with monoclonal antibodies and related proteins. *Clinical Medicine* 2017; **17**: 220.
117 ClinicalTrials.gov. Randomised Evaluation of COVID-19 Therapy (RECOVERY). 2020. https://clinicaltrials.gov/ct2/show/NCT04381936 (accessed March 13, 2023).
118 ISRCTN. ISRCTN86534580: PRINCIPLE: A clinical trial evaluating treatments for suspected and confirmed COVID-19 for recovery at home. 2020. https://www.isrctn.com/ISRCTN86534580 (accessed March 13, 2023).
119 GOV.UK. First monoclonal antibody treatment for COVID-19 approved for use in the UK. 2021. https://www.gov.uk/government/news/first-monoclonal-antibody-treatment-for-covid-19-approved-for-use-in-the-uk (accessed March 13, 2023).
120 GOV.UK. First oral antiviral for COVID-19, Lagevrio (molnupiravir), approved by MHRA. 2021. https://www.gov.uk/government/news/first-oral-antiviral-for-covid-19-lagevrio-molnupiravir-approved-by-mhra (accessed March 13, 2023).
121 NHS England. Coronavirus » Interim clinical commissioning policy: Treatments for non-hospitalised patients with COVID-19. 2022. https://www.england.nhs.uk/coronavirus/publication/interim-clinical-commissioning-policy-neutralising-monoclonal-antibodies-or-antivirals-for-non-hospitalised-patients-with-covid-19/ (accessed May 10, 2023).
122 Vegivinti CTR, Evanson KW, Lyons H, *et al.* Efficacy of antiviral therapies for COVID-19: a systematic review of randomized controlled trials. *BMC Infect Dis* 2022; **22**: 1–45.
123 Makhoul M, Abu-Hijleh F, Ayoub HH, Seedat S, Chemaitelly H, Abu-Raddad LJ. Modeling the population-level impact of treatment on COVID-19 disease and SARS-CoV-2 transmission. *Epidemics* 2022; **39**: 100567.

124 Bell D, Comas-Herrera A, Henderson D, *et al.* COVID-19 mortality and long-term care: a UK comparison. 2020 https://ltccovid.org/wp-content/uploads/2020/08/COVID-19-mortality-in-long-term-care-final-Sat-29-v1.pdf (accessed May 10, 2023).
125 LTCCovid. International data on deaths attributed to COVID-19 among people living in care homes – Resources to support community and institutional Long-Term Care responses to COVID-19. 2022. https://ltccovid.org/2022/02/22/international-data-on-deaths-attributed-to-covid-19-among-people-living-in-care-homes/ (accessed March 1, 2022).
126 SCWG. Care homes analysis, 12 May 2020. 2020 https://www.gov.uk/government/publications/care-homes-analysis-12-may-2020 (accessed March 21, 2023).
127 Forder J, Fernandez J-L. Length of stay in care homes. 2011 www.pssru.ac.uk (accessed July 12, 2021).
128 Institute and Faculty of Actuaries. Will the cap fit? What the government should consider before introducing a cap on social care costs. 2017 https://actuaries.org.uk/media/otfnsz1m/will-the-cap-fit-v1-0-nov-2017.pdf (accessed March 8, 2023).
129 Steventon A, Roberts A. Estimating length of stay in publicly-funded residential and nursing care homes: a retrospective analysis using linked administrative data sets. *BMC Health Serv Res* 2012; **12**: 377.
130 Kojima G. Prevalence of Frailty in Nursing Homes: A Systematic Review and Meta-Analysis. *J Am Med Dir Assoc* 2015; **16**: 940–5.
131 Veronese N, Custodero C, Cella A, *et al.* Prevalence of multidimensional frailty and pre-frailty in older people in different settings: A systematic review and meta-analysis. *Ageing Res Rev* 2021; **72**: 101498.
132 Gordon AL, Franklin M, Bradshaw L, Logan P, Elliott R, Gladman JRF. Health status of UK care home residents: a cohort study. *Age Ageing* 2014; **43**: 97–103.
133 Office for National Statistics. Care homes and estimating the self-funding population, England. 2021 https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/socialcare/articles/carehomesandestimatingtheselffundingpopulationengland/2019to2020 (accessed Feb 18, 2022).
134 GOV.UK. Care homes market study: summary of final report. 2017 https://www.gov.uk/government/publications/care-homes-market-study-summary-of-final-report/care-homes-market-study-summary-of-final-report (accessed Feb 18, 2022).
135 Skills for Care. The state of the adult social care sector and workforce in England. 2022. https://www.skillsforcare.org.uk/adult-social-care-workforce-data/Workforce-intelligence/publications/national-information/The-state-of-the-adult-social-care-sector-and-workforce-in-England.aspx (accessed March 21, 2023).
136 Bartleson JM, Radenkovic D, Covarrubias AJ, Furman D, Winer DA, Verdin E. SARS-CoV-2, COVID-19 and the aging immune system. *Nature Aging 2021 1:9* 2021; **1**: 769–82.
137 Teissier T, Boulanger E, Cox LS. Interconnections between Inflammageing and Immunosenescence during Ageing. *Cells 2022, Vol 11, Page 359* 2022; **11**: 359.

138 Witkowski JM, Fulop T, Bryl E. Immunosenescence and COVID-19. *Mech Ageing Dev* 2022; **204**: 111672.
139 Castro-Herrera VM, Lown M, Fisk HL, *et al.* Relationships Between Age, Frailty, Length of Care Home Residence and Biomarkers of Immunity and Inflammation in Older Care Home Residents in the United Kingdom. *Frontiers in Aging* 2021; **0**: 2.
140 Rodríguez-Rejón AI, Ruiz-López MD, Artacho R. Dietary Intake and Associated Factors in Long-Term Care Homes in Southeast Spain. *Nutrients* 2019; **11**. DOI:10.3390/NU11020266.
141 Keller HH, Carrier N, Slaughter SE, *et al.* Prevalence and Determinants of Poor Food Intake of Residents Living in Long-Term Care. *J Am Med Dir Assoc* 2017; **18**: 941–7.
142 Anderson DC, Grey T, Kennelly S, O'Neill D. Nursing Home Design and COVID-19: Balancing Infection Control, Quality of Life, and Resilience. *J Am Med Dir Assoc* 2020; **21**: 1519.
143 Morawska L, Tang JW, Bahnfleth W, *et al.* How can airborne transmission of COVID-19 indoors be minimised? *Environ Int* 2020; **142**: 105832.
144 Pineles L, Perencevich EN, Roghmann MC, *et al.* Frequency of nursing home resident contact with staff, other residents, and the environment outside resident rooms. *Infect Control Hosp Epidemiol* 2019; **40**: 815.
145 Scobie S. Covid-19 and the deaths of care home residents. 2021 https://www.nuffieldtrust.org.uk/news-item/covid-19-and-the-deaths-of-care-home-residents?gclid=Cj0KCQjwk7ugBhDIARIsAGuvgPasb1rsi7heW-Gg5mgkF0n4OpE-pTMUUk7r51wOXnVSYA2y8YITc6saAgwXEALw_wcB (accessed March 13, 2023).
146 Office for National Statistics. Analysis of death registrations not involving coronavirus (COVID-19), England and Wales. 2020. https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/articles/analysisofdeathregistrationsnotinvolvingcoronaviruscovid19englandandwales28december2019to1may2020/28december2019to10july2020 (accessed March 13, 2023).
147 Schultze A, Nightingale E, Evans D, *et al.* Mortality among Care Home Residents in England during the first and second waves of the COVID-19 pandemic: an observational study of 4.3 million adults over the age of 65. *The Lancet Regional Health – Europe* 2022; **14**: 100295.
148 Care Quality Commission. COVID Insight 6: Designated settings and care home capacity. 2022. https://www.cqc.org.uk/publications/major-reports/covid-insight-6-designated-settings-care-home-capacity (accessed March 13, 2023).
149 Office for National Statistics. Deaths involving COVID-19 in the care sector, England and Wales: deaths registered between week ending 20 March 2020 and week ending 21 January 2022. 2022. https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/articles/deathsinvolvingcovid19inthecaresectorenglandandwales/deathsregisteredbetweenweekending20march2020andweekending21january2022 (accessed Feb 8, 2023).
150 Burton JK, Goodman C, Guthrie B, Gordon AL, Hanratty B, Quinn TJ. Closing the UK care home data gap – methodological challenges and solutions. *Int J Popul Data Sci* 2020; **5**. DOI:10.23889/ijpds.v5i4.1391.
151 O' A, London D. Covid-19: Care home deaths in England and Wales rise sharply. *BMJ* 2020; **369**: m1727.

152 Care Quality Commission. Publication of statistics on deaths involving COVID-19 in care homes in England: transparency statement. 2020. https://www.cqc.org.uk/publications/major-reports/publication-statistics-deaths-involving-covid-19-care-homes-england (accessed March 21, 2023).
153 NECS. Capacity Tracker. 2022. https://www.necsu.nhs.uk/capacity-tracker/ (accessed Jan 21, 2022).
154 Shallcross L, Burke D, Abbott O, *et al.* Factors associated with SARS-CoV-2 infection and outbreaks in long-term care facilities in England: a national cross-sectional survey. *Lancet Healthy Longev* 2021; **2**: e129–42.
155 UKHSA. COVID-19: the future of UKHSA surveillance. 2022. https://ukhsa.blog.gov.uk/2022/07/01/covid-19-the-future-of-ukhsa-surveillance/ (accessed March 21, 2023).
156 Stratil JM, Biallas RL, Burns J, *et al.* Non-pharmacological measures implemented in the setting of long-term care facilities to prevent SARS-CoV-2 infections and their consequences: a rapid review. *Cochrane Database of Systematic Reviews* 2021; **2021**. DOI:10.1002/14651858.CD015085.PUB2/EPDF/ABSTRACT.
157 GOV.UK. About the Adult Social Care Infection Control Fund. 2021. https://www.gov.uk/government/publications/adult-social-care-infection-control-fund/about-the-adult-social-care-infection-control-fund (accessed March 13, 2023).
158 GOV.UK. COVID-19: our action plan for adult social care. 2020. https://www.gov.uk/government/publications/coronavirus-covid-19-adult-social-care-action-plan/covid-19-our-action-plan-for-adult-social-care (accessed Feb 8, 2023).
159 Frazer K, Mitchell L, Stokes D, Lacey E, Crowley E, Kelleher CC. A rapid systematic review of measures to protect older people in long-term care facilities from COVID-19. *BMJ Open* 2021; **11**: 47012.
160 Committee on Human Rights J. Care homes: Visiting restrictions during the covid-19 pandemic Fifteenth Report of Session 2019-21 Report, together with formal minutes relating to the report by authority of the House of Commons and House of Lords Joint Committee on Human Rights. 2021. www.parliament.uk/site-information/copyright-parliament/. (accessed Feb 8, 2023).
161 Shrotri M, Krutikov M, Palmer T, *et al.* Vaccine effectiveness of the first dose of ChAdOx1 nCoV-19 and BNT162b2 against SARS-CoV-2 infection in residents of long-term care facilities in England (VIVALDI): a prospective cohort study. *Lancet Infect Dis* 2021; **21**: 1529–38.
162 Shrotri M, Krutikov M, Nacer-Laidi H, *et al.* Duration of vaccine effectiveness against SARS-CoV-2 infection, hospitalisation, and death in residents and staff of long-term care facilities in England (VIVALDI): a prospective cohort study. *Lancet Healthy Longev* 2022; **3**: e470–80.
163 Hollinghurst J, North L, Perry M, *et al.* COVID-19 infection risk amongst 14,104 vaccinated care home residents: a national observational longitudinal cohort study in Wales, UK, December 2020–March 2021. *Age Ageing* 2022; **51**: 1–7.
164 McConeghy KW, Bardenheier B, Huang AW, *et al.* Infections, Hospitalizations, and Deaths Among US Nursing Home Residents With vs Without a SARS-CoV-2 Vaccine Booster. *JAMA Netw Open* 2022; **5**: e2245417–e2245417.
165 Department of Health & Social Care. Priority groups for coronavirus (COVID-19) vaccination: advice from the JCVI, 2 December 2020 - GOV.UK.

https://www.gov.uk/government/publications/priority-groups-for-coronavirus-covid-19-vaccination-advice-from-the-jcvi-2-december-2020/priority-groups-for-coronavirus-covid-19-vaccination-advice-from-the-jcvi-2-december-2020 (accessed Dec 9, 2020).
166 McGarry BE, Barnett ML, Grabowski DC, Gandhi AD. Nursing Home Staff Vaccination and Covid-19 Outcomes. *New England Journal of Medicine* 2022; **386**: 397–8.
167 Public Health England. A data linkage approach to assessing the contribution of hospital-associated SARS-CoV-2 infection to care home outbreaks in England, 30 January to 12 October 2020. 2020 https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1005583/Nosocomial_Seeding_OF_care_home_outbreaks_report_Official_Sensitive-3.pdf (accessed Feb 8, 2023).
168 Baister M, McTaggart E, McMenemy P, Megiddo I, Kleczkowski A. COVID-19 in Scottish care homes: A metapopulation model of spread among residents and staff. *medRxiv* 2022; : 2021.08.24.21262524.
169 Rosello A, Barnard RC, Smith DRM, *et al.* Impact of non-pharmaceutical interventions on SARS-CoV-2 outbreaks in English care homes: a modelling study. *BMC Infect Dis* 2022; **22**: 1–13.
170 Emmerson C, Adamson JP, Turner D, *et al.* Risk factors for outbreaks of COVID-19 in care homes following hospital discharge: A national cohort analysis. *Influenza Other Respir Viruses* 2021; **15**: 371–80.
171 GOV.UK. Coronavirus (COVID-19) testing in care homes: statistics to 8 July 2020. 2020. https://www.gov.uk/government/publications/coronavirus-covid-19-testing-in-care-homes-statistics-to-8-july-2020/coronavirus-covid-19-testing-in-care-homes-statistics-to-8-july-2020#measuring-the-data (accessed Feb 8, 2023).
172 GOV.UK. COVID-19 testing in adult social care. 2022. https://www.gov.uk/government/publications/coronavirus-covid-19-testing-for-adult-social-care-settings/covid-19-testing-in-adult-social-care (accessed Feb 14, 2023).
173 Lam HR, Chow S, Taylor K, *et al.* Challenges of conducting research in long-term care facilities: A systematic review. *BMC Geriatr* 2018; **18**: 1–11.
174 Law E, Ashworth R. Facilitators and Barriers to Research Participation in Care Homes: Thematic Analysis of Interviews with Researchers, Staff, Residents and Residents' Families. *Journal of Long-Term Care* 2022; **2022**: 49–60.
175 Moore DC, Payne S, Van Den Block L, Ten Koppel M, Szczerbińska K, Froggatt K. Research, recruitment and observational data collection in care homes: Lessons from the PACE study. *BMC Res Notes* 2019; **12**: 1–6.
176 Shepherd V, Nuttall J, Hood K, Butler CC. Setting up a clinical trial in care homes: Challenges encountered and recommendations for future research practice. *BMC Res Notes* 2015; **8**: 1–4.
177 Moore DC, Hanratty B. Out of sight, out of mind? A review of data available on the health of care home residents in longitudinal and nationally representative cross-sectional studies in the UK and Ireland. *Age Ageing* 2013; **42**: 798–803.
178 Burton JK, Guthrie B. Identifying who lives in a care home—a challenge to be conquered. *Age Ageing* 2018; **47**: 322–3.
179 Burton JK, Marwick CA, Galloway J, *et al.* Identifying care-home residents in routine healthcare datasets: a diagnostic test accuracy study of five methods. *Age Ageing* 2019; **48**: 114.

180 Department of Health and Social Care. People at the Heart of Care: adult social care reform. *Adult Social Care Reform White Paper* 2021; : 1–108.
181 Khan RF, Meyer JD. How does the hierarchy of controls integrate with the epidemiologic triangle to help address and understand transmission of SARS-COV-2? *J Occup Environ Med* 2020; **62**: E665–8.
182 CDC. Principles of Epidemiology | Lesson 1 - Section 8. 2012. https://www.cdc.gov/csels/dsepd/ss1978/Lesson1/Section8.html#ALT116 (accessed June 27, 2023).
183 Tsui BCH, Deng A, Pan S. Coronavirus Disease 2019: Epidemiological Factors During Aerosol-Generating Medical Procedures. *Anesth Analg* 2020; **131**: e175–8.
184 GOV.UK. Technical report on the COVID-19 pandemic in the UK - Chapter 8.2: care homes. 2023. https://www.gov.uk/government/publications/technical-report-on-the-covid-19-pandemic-in-the-uk/chapter-82-care-homes#fnref:4 (accessed May 12, 2023).
185 British Geriatrics Society. End of Life Care in Frailty: Care homes. 2020. https://www.bgs.org.uk/resources/end-of-life-care-in-frailty-care-homes (accessed March 21, 2023).
186 Office for National Statistics. Coronavirus and key workers in the UK. 2020. https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/earningsandworkinghours/articles/coronavirusandkeyworkersintheuk/2020-05-15 (accessed April 19, 2023).
187 Harrichandra A, Ierardi AM, Pavilonis B. An estimation of airborne SARS-CoV-2 infection transmission risk in New York City nail salons. *Toxicol Ind Health* 2020; **36**: 634–43.
188 Foster A, Kinzel M. Estimating COVID-19 exposure in a classroom setting: A comparison between mathematical and numerical models. *Physics of Fluids* 2021; **33**: 021904.
189 Guo Y, Qian H, Sun Z, *et al.* Assessing and controlling infection risk with Wells-Riley model and spatial flow impact factor (SFIF). *Sustain Cities Soc* 2021; **67**: 102719.
190 GOV.UK. [Withdrawn] Overview of adult social care guidance on coronavirus (COVID-19). 2021. https://www.gov.uk/guidance/overview-of-adult-social-care-guidance-on-coronavirus-covid-19#what-to-do-when-you-suspect-an-outbreak (accessed May 12, 2023).
191 ECDC. Surveillance of COVID-19 in long-term care facilities in the EU/EEA, November 2021. 2021 DOI:10.2900/936267.
192 Grubaugh ND, Ladner JT, Lemey P, *et al.* Tracking virus outbreaks in the twenty-first century. *Nat Microbiol* 2019; **4**: 10.
193 Meredith LW, Hamilton WL, Warne B, *et al.* Rapid implementation of SARS-CoV-2 sequencing to investigate cases of health-care associated COVID-19: a prospective genomic surveillance study. *Lancet Infect Dis* 2020; **20**: 1263–72.
194 Hamilton WL, Tonkin-Hill G, Smith ER, *et al.* Genomic epidemiology of COVID-19 in care homes in the east of England. *Elife* 2021; **10**. DOI:10.7554/ELIFE.64618.
195 Stirrup O, Blackstone J, Mapp F, *et al.* Effectiveness of rapid SARS-CoV-2 genome sequencing in supporting infection control for hospital-onset COVID-19 infection: multicenter, prospective study. *Elife* 2022; **11**. DOI:10.7554/ELIFE.78427.

196 LTCCovid. Rapid review of the evidence on impacts of visiting policies in care homes during the COVID-19 pandemic – Resources to support community and institutional Long-Term Care responses to COVID-19. 2020. https://ltccovid.org/2020/11/01/pre-print-rapid-review-of-the-evidence-on-impacts-of-visiting-policies-in-care-homes-during-the-covid-19-pandemic/ (accessed Feb 8, 2023).
197 SCWG. What are the appropriate mitigations to deploy in care homes in the context of the post vaccination risk landscape?, 26 May 2021. 2021. https://www.gov.uk/government/publications/scwg-what-are-the-appropriate-mitigations-to-deploy-in-care-homes-in-the-context-of-the-post-vaccination-risk-landscape-26-may-2021 (accessed May 12, 2023).
198 Tricco AC, Lillie E, Zarin W, *et al.* PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Ann Intern Med* 2018; **169**: 467–73.
199 NICE. 7 Appendix: search strategy for Medline (Ovid Platform) | Interim process and methods for developing rapid guidelines on COVID-19 | Guidance. 2020. https://www.nice.org.uk/process/pmg35/chapter/appendix-search-strategy-for-medline-ovid-platform (accessed June 1, 2022).
200 Ottawa Hospital Research Institute. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. 2012. https://www.ohri.ca/programs/clinical_epidemiology/oxford.asp (accessed Feb 8, 2023).
201 NIH. Study Quality Assessment Tools | NHLBI, NIH. 2021. https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools (accessed Feb 8, 2023).
202 Aghili MS, Darvishpoor Kakhki A, Gachkar L, Davidson PM. Predictors of contracting COVID-19 in nursing homes: Implications for clinical practice. *J Adv Nurs* 2022; **78**: 2799–806.
203 Sugg MM, Spaulding TJ, Lane SJ, *et al.* Mapping community-level determinants of COVID-19 transmission in nursing homes: A multi-scale approach. *Sci Total Environ* 2021; **752**. DOI:10.1016/J.SCITOTENV.2020.141946.
204 San Román J, Candel FJ, del Mar Carretero M, *et al.* Cross-Sectional Analysis of Risk Factors for Outbreak of COVID-19 in Nursing Homes for Older Adults in the Community of Madrid. *Gerontology* 2023; **69**: 163–71.
205 Rauhala AS, Fagerström LM, Lindholst AC, *et al.* Which factors are associated with COVID-19 infection incidence in care services for older people in Nordic countries? A cross-sectional survey. *Scand J Public Health* 2022; **50**: 787–94.
206 Peckeu-Abboud L, Van Kleef E, Smekens T, *et al.* Factors influencing SARS-CoV-2 infection rate in Belgian nursing home residents during the first wave of COVID-19 pandemic. *Epidemiol Infect* 2022; **150**: e72.
207 Corvol A, Charras K, Prud'homm J, *et al.* Structural and Managerial Risk Factors for COVID-19 Occurrence in French Nursing Homes. *Int J Health Policy Manag* 2022; **11**: 2630–7.
208 Cazzoletti L, Zanolin ME, Tocco Tussardi I, *et al.* Risk Factors Associated with Nursing Home COVID-19 Outbreaks: A Retrospective Cohort Study. *International Journal of Environmental Research and Public Health 2021, Vol 18, Page 8434* 2021; **18**: 8434.

209 Torres ML, Díaz DP, Oliver-Parra A, *et al.* Inequities in the incidence and mortality due to COVID-19 in nursing homes in Barcelona by characteristics of the nursing homes. *PLoS One* 2022; **17**: e0269639.
210 Orlando S, Mazhari T, Abbondanzieri A, *et al.* Characteristics of nursing homes and early preventive measures associated with risk of infection from COVID-19 in Lazio region, Italy: a retrospective case–control study. *BMJ Open* 2022; **12**: e061784.
211 Piet E, Maillard A, Mallaval FO, *et al.* Outbreaks of covid-19 in nursing homes: A cross-sectional survey of 74 nursing homes in a french area. *J Clin Med* 2021; **10**: 4280.
212 Soldevila L, Prat N, Mas M, *et al.* The interplay between infection risk factors of SARS-CoV-2 and mortality: a cross-sectional study from a cohort of long-term care nursing home residents. *BMC Geriatr* 2022; **22**: 1–9.
213 Burgaña Agoües A, Serra Gallego M, Hernández Resa R, *et al.* Risk Factors for COVID-19 Morbidity and Mortality in Institutionalised Elderly People. *International Journal of Environmental Research and Public Health 2021, Vol 18, Page 10221* 2021; **18**: 10221.
214 Lombardo FL, Bacigalupo I, Salvi E, *et al.* The Italian national survey on Coronavirus disease 2019 epidemic spread in nursing homes. *Int J Geriatr Psychiatry* 2021; **36**: 873–82.
215 Tulloch JSP, Green R, Tunnah C, Coffey E, Ashton M, Ghebrehewet S. COVID-19 outbreaks in care homes during the first wave: are Care Quality Commission ratings a good predictor of at-risk homes? *Journal of Hospital Infection* 2021; **111**: 96–101.
216 Burton JK, McMinn M, Vaughan JE, Fleuriot J, Guthrie B. Care-home outbreaks of COVID-19 in Scotland March to May 2020: National linked data cohort analysis. *Age Ageing* 2021; **50**: 1482–92.
217 Dutey-Magni PF, Williams H, Jhass A, *et al.* COVID-19 infection and attributable mortality in UK care homes: cohort study using active surveillance and electronic records (March–June 2020). *Age Ageing* 2021; **50**: 1019–28.
218 Bach-Mortensen AM, Degli Esposti M. Is area deprivation associated with greater impacts of COVID-19 in care homes across England? A preliminary analysis of COVID-19 outbreaks and deaths. *J Epidemiol Community Health (1978)* 2021; **75**: 624–7.
219 White EM, Kosar CM, Feifer RA, *et al.* Variation in SARS-CoV-2 Prevalence in U.S. Skilled Nursing Facilities. *J Am Geriatr Soc* 2020; **68**: 2167–73.
220 He M, Li Y, Fang F. Is There a Link between Nursing Home Reported Quality and COVID-19 Cases? Evidence from California Skilled Nursing Facilities. *J Am Med Dir Assoc* 2020; **21**: 905–8.
221 Bui DP, See I, Hesse EM, *et al.* Association Between CMS Quality Ratings and COVID-19 Outbreaks in Nursing Homes — West Virginia, March 17–June 11, 2020. *MMWR Morb Mortal Wkly Rep* 2020; **69**: 1300–4.
222 Travers JL, Agarwal M, Estrada L V., *et al.* Assessment of Coronavirus Disease 2019 Infection and Mortality Rates Among Nursing Homes With Different Proportions of Black Residents. *J Am Med Dir Assoc* 2021; **22**: 893-898.e2.
223 Hege A, Lane S, Spaulding T, Sugg M, Iyer LS. County-Level Social Determinants of Health and COVID-19 in Nursing Homes, United States, June 1, 2020–January 31, 2021. *https://doi.org/101177/00333549211053666* 2021; **137**: 137–48.

224 Longo BA, Barrett SC, Schmaltz SP, Williams SC. A Multistate Comparison Study of COVID-19 Cases Among Accredited and Nonaccredited Nursing Homes. *Policy Polit Nurs Pract* 2022; **23**: 26–31.
225 Zhu X, Lee H, Sang H, *et al.* Nursing Home Design and COVID-19: Implications for Guidelines and Regulation. *J Am Med Dir Assoc* 2022; **23**: 272-279.e1.
226 Lane SJ, Sugg M, Spaulding TJ, Hege A, Iyer L. Southeastern United States Predictors of COVID-19 in Nursing Homes. *Journal of Applied Gerontology* 2022; **41**: 1641–50.
227 Vijh R, Ng CH, Shirmaleki M, Bharmal A. Factors associated with transmission of COVID-19 in long-term care facility outbreaks. *Journal of Hospital Infection* 2022; **119**: 118–25.
228 Stall NM, Jones A, Brown KA, Rochon PA, Costa AP. For-profit long-term care homes and the risk of COVID-19 outbreaks and resident deaths. *CMAJ* 2020; **192**: E946–55.
229 Brown KA, Jones A, Daneman N, *et al.* Association Between Nursing Home Crowding and COVID-19 Infection and Mortality in Ontario, Canada. *JAMA Intern Med* 2021; **181**: 229–36.
230 Lee J, Shin JH, Lee KH, Harrington CA, Jung SO. Staffing Levels and COVID-19 Infections and Deaths in Korean Nursing Homes. *Policy Polit Nurs Pract* 2022; **23**: 15–25.
231 Wall Street Journal. Iran Hit by Worst Covid-19 Wave Yet as Vaccine Rollout Flounders. 2021. https://www.wsj.com/articles/iran-hit-by-worst-covid-19-wave-yet-as-vaccine-rollout-flounders-11628859038 (accessed Feb 8, 2023).
232 Reuters. Iran: the latest coronavirus counts, charts and maps. 2022. https://www.reuters.com/graphics/world-coronavirus-tracker-and-maps/countries-and-territories/iran/ (accessed Feb 8, 2023).
233 GOV.UK. [Withdrawn] [Withdrawn] COVID-19: epidemiological definitions of outbreaks and clusters in particular settings. 2020. https://www.gov.uk/government/publications/covid-19-epidemiological-definitions-of-outbreaks-and-clusters/covid-19-epidemiological-definitions-of-outbreaks-and-clusters-in-particular-settings (accessed Feb 10, 2023).
234 Lo CH, Nguyen LH, Drew DA, *et al.* Race, ethnicity, community-level socioeconomic factors, and risk of COVID-19 in the United States and the United Kingdom. *EClinicalMedicine* 2021; **38**. DOI:10.1016/j.eclinm.2021.101029.
235 van Ingen T, Brown KA, Buchan SA, *et al.* Neighbourhood-level socio-demographic characteristics and risk of COVID-19 incidence and mortality in Ontario, Canada: A population-based study. *PLoS One* 2022; **17**: e0276507.
236 Sze S, Pan D, Nevill CR, *et al.* Ethnicity and clinical outcomes in COVID-19: A systematic review and meta-analysis. *EClinicalMedicine* 2020; **29–30**: 100630.
237 Antonelli M, Penfold RS, Merino J, *et al.* Risk factors and disease profile of post-vaccination SARS-CoV-2 infection in UK users of the COVID Symptom Study app: a prospective, community-based, nested, case-control study. *Lancet Infect Dis* 2022; **22**: 43–55.
238 Shang W, Kang L, Cao G, *et al.* Percentage of Asymptomatic Infections among SARS-CoV-2 Omicron Variant-Positive Individuals: A Systematic Review and Meta-Analysis. *Vaccines (Basel)* 2022; **10**. DOI:10.3390/VACCINES10071049/S1.

239 Lucey M, Macori G, Mullane N, *et al.* Whole-genome Sequencing to Track Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Transmission in Nosocomial Outbreaks. *Clinical Infectious Diseases* 2021; **72**: e727–35.
240 National Collaborating Centre for Methods and Tools. What risk factors are associated with COVID-19 outbreaks and mortality in long-term care facilities and what strategies mitigate risk? Update 1. 2020 https://www.nccmt.ca/knowledge-repositories/covid-19- (accessed Feb 18, 2022).
241 Bach-Mortensen AM, Verboom B, Movsisyan A, Degli Esposti M. A systematic review of the associations between care home ownership and COVID-19 outbreaks, infections and mortality. *Nature Aging 2021 1:10* 2021; **1**: 948–61.
242 BASW. An international comparison of long-term care funding and outcomes: insights for the social care green paper. 2018 https://www.basw.co.uk/resources/international-comparison-long-term-care-funding-and-outcomes-insights-social-care-green (accessed Feb 18, 2022).
243 Robertson R, Gregory S, Jabbal J. The social care and health systems of nine countries Executive summary. 2014.
244 Schultze A, Nightingale E, Evans D, *et al.* Mortality among Care Home Residents in England during the first and second waves of the COVID-19 pandemic: an observational study of 4.3 million adults over the age of 65. *The Lancet Regional Health - Europe* 2022; **14**. DOI:10.1016/j.lanepe.2021.100295.
245 Krutikov M, Palmer T, Donaldson A, *et al.* Study Protocol: Understanding SARS-Cov-2 infection, immunity and its duration in care home residents and staff in England (VIVALDI). *Wellcome Open Research 2021 5:232* 2021; **5**: 232.
246 GOV.UK. COVID-19 surveillance and immunity studies. 2022. https://www.gov.uk/government/publications/covid-19-surveillance/covid-19-surveillance (accessed March 9, 2023).
247 GOV.UK. Whole home testing rolled out to all care homes in England. 2020. https://www.gov.uk/government/news/whole-home-testing-rolled-out-to-all-care-homes-in-england (accessed Feb 18, 2022).
248 Department of Health and Social Care. Coronavirus (COVID-19) Scaling up our testing programmes. 2020 https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/878121/coronavirus-covid-19-testing-strategy.pdf (accessed Feb 24, 2023).
249 Health Research Authority. UK Policy Framework for Health and Social Care Research. 2023. https://www.hra.nhs.uk/planning-and-improving-research/policies-standards-legislation/uk-policy-framework-health-social-care-research/uk-policy-framework-health-and-social-care-research/ (accessed Feb 24, 2023).
250 Head MG, Walker SL, Nalabanda A, Bostock J, Cassell JA. Researching Scabies Outbreaks among People in Residential Care and Lacking Capacity to Consent: A Case Study. *Public Health Ethics* 2017; **10**: 90–5.
251 legislation.gov.uk. Mental Capacity Act 2005. 2005 https://www.legislation.gov.uk/ukpga/2005/9/contents (accessed Feb 24, 2023).

252 GOV.UK. NHS Test and Trace: how we test your samples. 2023. https://www.gov.uk/government/publications/nhs-test-and-trace-how-we-test-your-samples/nhs-test-and-trace-how-we-test-your-samples (accessed Feb 24, 2023).
253 UKHSA. PCR testing for SARS-CoV-2 in the UK during the COVID-19 pandemic Use of PCR in the National Testing Programme. 2023 https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1152352/Testing_in_the_UK_during_the_SARS-1.pdf (accessed May 23, 2023).
254 COG-UK. COVID-19 Genomics UK Consortium. 2022. https://www.cogconsortium.uk/ (accessed Jan 21, 2022).
255 Smith CM, Williams H, Jhass A, *et al.* Antibiotic prescribing in UK care homes 2016-2017: retrospective cohort study of linked data. *BMC Health Serv Res* 2020; **20**. DOI:10.1186/S12913-020-05422-Z.
256 NIHR. ENRICH. 2023. https://enrich.nihr.ac.uk/ (accessed May 23, 2023).
257 Hayward AC, Fragaszy EB, Bermingham A, *et al.* Comparative community burden and severity of seasonal and pandemic influenza: Results of the Flu Watch cohort study. *Lancet Respir Med* 2014; **2**: 445–54.
258 NHS England. NHS England » NHS COVID-19 Data Store. 2020. https://www.england.nhs.uk/contact-us/privacy-notice/how-we-use-your-information/covid-19-response/nhs-covid-19-data-store/ (accessed Dec 22, 2020).
259 NPEx. NPEx. 2022. https://www.npex.nhs.uk/ (accessed Jan 21, 2022).
260 NHS England. Statistics » COVID-19 Vaccinations. 2022. https://www.england.nhs.uk/statistics/statistical-work-areas/covid-19-vaccinations/ (accessed Jan 21, 2022).
261 NHS Digital. Hospital Episode Statistics (HES). 2022. https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/hospital-episode-statistics (accessed Jan 21, 2022).
262 UCL. Data Safe Haven (DSH) | Information Services Division. 2023. https://www.ucl.ac.uk/isd/services/file-storage-sharing/data-safe-haven-dsh (accessed May 23, 2023).
263 NHS Digital. Control of patient information (COPI) notice. 2022. https://digital.nhs.uk/coronavirus/coronavirus-covid-19-response-information-governance-hub/control-of-patient-information-copi-notice (accessed Feb 18, 2022).
264 Walker AS, Vihta K-D, Gethings O, *et al.* Tracking the Emergence of SARS-CoV-2 Alpha Variant in the United Kingdom. *New England Journal of Medicine* 2021; **385**: 2582–5.
265 O'Toole Á, Hill V, Pybus OG, *et al.* Tracking the international spread of SARS-CoV-2 lineages B.1.1.7 and B.1.351/501Y-V2 with grinch. *Wellcome Open Res* 2021; **6**: 121.
266 Public Health England. Investigation of SARS-CoV-2 variants of concern in England. 2021. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/961299/Variants_of_Concern_VOC_Technical_Briefing_6_England-1.pdf (accessed Feb 18, 2022).
267 Public Health England. Investigation of novel SARS-COV-2 variant Variant of Concern 202012/01. 2020 https://assets.publishing.service.gov.uk/government/uploads/system/uploads/a

ttachment_data/file/947048/Technical_Briefing_VOC_SH_NJL2_SH2.pdf (accessed Dec 23, 2020).
268 COG-UK. COG-UK update on SARS-CoV-2 Spike mutations of special interest: Report 1. 2020 https://www.cogconsortium.uk/wp-content/uploads/2020/12/Report-1_COG-UK_19-December-2020_SARS-CoV-2-Mutations.pdf (accessed Dec 22, 2020).
269 New and Emerging Respiratory Virus Threats Advisory Group. New and Emerging Respiratory Virus Threats Advisory Group NERVTAG/SPI-M Extraordinary meeting on SARS-CoV-2 variant of concern 202012/01 (variant B.1.1.7) Note of meeting. MW, 2020 https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/954945/s0993-nervtag-spi-m-voc-meeting.pdf (accessed Jan 22, 2021).
270 Lin L, Liu Y, Tang X, He D. The Disease Severity and Clinical Outcomes of the SARS-CoV-2 Variants of Concern. *Front Public Health* 2021; **9**: 1929.
271 Grint DJ, Wing K, Houlihan C, *et al.* Severity of SARS-CoV-2 alpha variant (B.1.1.7) in England. *Clin Infect Dis* 2021; published online Sept 6. DOI:10.1093/CID/CIAB754.
272 Davies NG, Jarvis CI, van Zandvoort K, *et al.* Increased mortality in community-tested cases of SARS-CoV-2 lineage B.1.1.7. *Nature 2021 593:7858* 2021; **593**: 270–4.
273 Patone M, Thomas K, Hatch R, *et al.* Mortality and critical care unit admission associated with the SARS-CoV-2 lineage B.1.1.7 in England: an observational cohort study. *Lancet Infect Dis* 2021; **21**: 1518–28.
274 Stirrup O, Boshier F, Venturini C, *et al.* SARS-CoV-2 lineage B.1.1.7 is associated with greater disease severity among hospitalised women but not men: multicentre cohort study. *BMJ Open Respir Res* 2021; **8**: e001029.
275 Challen R, Brooks-Pollock E, Read JM, Dyson L, Tsaneva-Atanasova K, Danon L. Risk of mortality in patients infected with SARS-CoV-2 variant of concern 202012/1: matched cohort study. *BMJ* 2021; **372**. DOI:10.1136/BMJ.N579.
276 GOV.UK. Government launches new portal for care homes to arrange coronavirus testing. 2020. https://www.gov.uk/government/news/government-launches-new-portal-for-care-homes-to-arrange-coronavirus-testing (accessed Feb 18, 2022).
277 Long QX, Liu BZ, Deng HJ, *et al.* Antibody responses to SARS-CoV-2 in patients with COVID-19. *Nature Medicine 2020 26:6* 2020; **26**: 845–8.
278 Houlihan CF, Beale R. The complexities of SARS-CoV-2 serology. Lancet Infect Dis. 2020; **20**: 1350–1.
279 Huang AT, Garcia-Carreras B, Hitchings MDT, *et al.* A systematic review of antibody mediated immunity to coronaviruses: kinetics, correlates of protection, and association with severity. *Nat Commun* 2020; **11**. DOI:10.1038/S41467-020-18450-4.
280 Public Health England. Investigation of novel SARS-CoV-2 variant Variant of Concern 202012/01 Technical briefing 4. 2020. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/959359/Variant_of_Concern_VOC_202012_01_Technical_Briefing_4.pdf (accessed March 7, 2023).
281 PHE. Investigation of novel SARS-COV-2 variant Variant of Concern 202012/01. 2020

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/959438/Technical_Briefing_VOC_SH_NJL2_SH2.pdf (accessed May 16, 2023).
282 Magleby R, Westblade LF, Trzebucki A, *et al.* Impact of Severe Acute Respiratory Syndrome Coronavirus 2 Viral Load on Risk of Intubation and Mortality Among Hospitalized Patients With Coronavirus Disease 2019. *Clin Infect Dis* 2021; **73**: e4197.
283 Faíco-Filho KS, Passarelli VC, Bellei N. Is Higher Viral Load in SARS-CoV-2 Associated with Death? *Am J Trop Med Hyg* 2020; **103**: 2019–21.
284 Jaafar R, Aherfi S, Wurtz N, *et al.* Correlation Between 3790 Quantitative Polymerase Chain Reaction-Positives Samples and Positive Cell Cultures, Including 1941 Severe Acute Respiratory Syndrome Coronavirus 2 Isolates. *Clin Infect Dis* 2021; **72**: E932–3.
285 Ainsworth M, Andersson M, Auckland K, *et al.* Performance characteristics of five immunoassays for SARS-CoV-2: a head-to-head benchmark comparison. *Lancet Infect Dis* 2020; **20**: 1390–400.
286 Bryan A, Pepper G, Wener MH, *et al.* Performance characteristics of the abbott architect sars-cov-2 igg assay and seroprevalence in Boise, Idaho. *J Clin Microbiol* 2020; **58**. DOI:10.1128/JCM.00941-20/ASSET/A74506C5-922E-42D7-8C98-CC4DCC484511/ASSETS/GRAPHIC/JCM.00941-20-F0003.JPEG.
287 Jaafar R, Aherfi S, Wurtz N, *et al.* Correlation Between 3790 Quantitative Polymerase Chain Reaction–Positives Samples and Positive Cell Cultures, Including 1941 Severe Acute Respiratory Syndrome Coronavirus 2 Isolates. *Clinical Infectious Diseases* 2020; published online Sept 28. DOI:10.1093/cid/ciaa1491.
288 Ibarrondo FJ, Fulcher JA, Goodman-Meza D, *et al.* Rapid Decay of Anti–SARS-CoV-2 Antibodies in Persons with Mild Covid-19. *New England Journal of Medicine* 2020; **383**: 1085–7.
289 Seow J, Graham C, Merrick B, *et al.* Longitudinal evaluation and decline of antibody responses in SARS-CoV-2 infection. *Nat Microbiol* 2020; **5**: 1598.
290 Choudhary HR, Parai D, Dash GC, *et al.* IgG antibody response against nucleocapsid and spike protein post-SARS-CoV-2 infection. *Infection* 2021; **49**: 1045–8.
291 Krutikov M, Hayward A, Shallcross L. Spread of a Variant SARS-CoV-2 in Long-Term Care Facilities in England. *New England Journal of Medicine* 2021; **384**: 1671–3.
292 Krutikov M, Palmer T, Tut G, *et al.* Prevalence and duration of detectable SARS-CoV-2 nucleocapsid antibodies in staff and residents of long-term care facilities over the first year of the pandemic (VIVALDI study): prospective cohort study in England. *Lancet Healthy Longev* 2022; **3**: e13–21.
293 Office for National Statistics. Coronavirus (COVID-19) Infection Survey, UK. 2021. https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/coronaviruscovid19infectionsurveypilot/21may2021 (accessed March 8, 2023).
294 Office for National Statistics. Number of deaths in care homes notified to the Care Quality Commission, England. 2021. https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarria

ges/deaths/datasets/numberofdeathsincarehomesnotifiedtothecarequalitycommissionengland (accessed July 12, 2021).
295 Tut G, Lancaster T, Sylla P, *et al.* Antibody and cellular immune responses following dual COVID-19 vaccination within infection-naive residents of long-term care facilities: an observational cohort study. *Lancet Healthy Longev* 2022; **3**: e461–9.
296 Jesús SR, Javier CF, Carlos SJ, *et al.* Humoral and Cellular Response after mRNA Vaccination in Nursing Homes: Influence of Age and of History of COVID-19. *Vaccines (Basel)* 2022; **10**: 383.
297 Bobrovitz N, Ware H, Ma X, *et al.* Protective effectiveness of previous SARS-CoV-2 infection and hybrid immunity against the omicron variant and severe disease: a systematic review and meta-regression. *Lancet Infect Dis* 2023; **0**. DOI:10.1016/s1473-3099(22)00801-5.
298 Shrotri M, Krutikov M, Nacer-Laidi H, *et al.* Duration of vaccine effectiveness against SARS-CoV-2 infection, hospitalisation, and death in residents and staff of long-term care facilities in England (VIVALDI): a prospective cohort study. *Lancet Healthy Longev* 2022; **3**: e470–80.
299 Hay JA, Kennedy-Shaffer L, Kanjilal S, *et al.* Estimating epidemiologic dynamics from cross-sectional viral load distributions. *Science (1979)* 2021; **373**. DOI:10.1126/SCIENCE.ABH0635.
300 Evans D, Cowen S, Kammel M, *et al.* The Dangers of Using Cq to Quantify Nucleic Acid in Biological Samples: A Lesson From COVID-19. *Clin Chem* 2021; **68**: 153–62.
301 Public Health England. Weekly national Influenza and COVID-19 surveillance report Executive summary - Week 29 2021. 2021 https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1005056/Weekly_Flu_and_COVID-19_report_w29.pdf (accessed Aug 17, 2021).
302 Bell D, Comas-Herrera A, Henderson D, *et al.* COVID-19 mortality and long-term care: a UK comparison COVID-19 mortality and long-term care: a UK comparison. 2020 https://ltccovid.org/wp-content/uploads/2020/08/COVID-19-mortality-in-long-term-care-final-Sat-29-1.pdf (accessed March 3, 2021).
303 Candel FJ, Barreiro P, San Román J, *et al.* The demography and characteristics of SARS-CoV-2 seropositive residents and staff of nursing homes for older adults in the Community of Madrid: the SeroSOS study. *Age Ageing* 2021; **50**: 1038–47.
304 Patwardhan S, Sutton M, Morciano M. Effects of chain ownership and private equity financing on quality in the English care home sector: retrospective observational study. *Age Ageing* 2022; **51**: 1–8.
305 carehome.co.uk. One in four say care home staff have left their families to move in with residents during lockdown. 2020. https://www.carehome.co.uk/news/article.cfm/id/1624983/one-in-four-care-home-staff-left-families-move-in-lockdown (accessed May 16, 2023).
306 Fotaki M, Horton A, Rowland D, Ozdemir Kaya D, Gain A. Bailed out and burned out? The financial impact of COVID-19 on UK care homes for older people and their workforce. Coventry, 2023 https://ficch.org.uk/resources/reports/ (accessed May 16, 2023).
307 Jin R. The lag between daily reported Covid-19 cases and deaths and its relationship to age. *J Public Health Res* 2021; **10**: 2049.

308 Office for National Statistics. Coronavirus (COVID-19) latest insights. 2022. https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/articles/coronaviruscovid19latestinsights/overview (accessed March 7, 2023).
309 To KKW, Tsang OTY, Leung WS, *et al.* Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by SARS-CoV-2: an observational cohort study. *Lancet Infect Dis* 2020; **20**: 565–74.
310 UKHSA. COVID-19 vaccine surveillance report: week 15. 2022. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1069256/Vaccine_surveillance_report_-_week_15.pdf (accessed March 7, 2023).
311 Office for National Statistics. Coronavirus (COVID-19) Infection Survey technical article: Cumulative incidence of the number of people who have tested positive for COVID-19, UK. 2022. https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/articles/coronaviruscovid19infectionsurveytechnicalarticlecumulativeincidenceofthenumberofpeoplewhohavetestedpositiveforcovid19uk/22april2022 (accessed March 7, 2023).
312 Savadori L, Lauriola M. Risk perceptions and COVID-19 protective behaviors: A two-wave longitudinal study of epidemic and post-epidemic periods. *Soc Sci Med* 2022; **301**: 114949.
313 Office for National Statistics. Deaths involving COVID-19 in the care sector, England and Wales. 2022. https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/articles/deathsinvolvingcovid19inthecaresectorenglandandwales/deathsregisteredbetweenweekending20march2020andweekending21january2022 (accessed March 7, 2023).
314 Netten A, Bebbington A, Darton R, Forder J. Care Homes for Elderly People: Volume 1. Facilities, Residents and Costs. 2001 https://www.pssru.ac.uk/pub/chop1.pdf (accessed March 8, 2023).
315 Oliver D. David Oliver: The story of covid in care homes—neither mystery nor melodrama. *BMJ* 2021; **373**. DOI:10.1136/BMJ.N1432.
316 Mateen BA, Wilde H, Dennis JM, *et al.* Hospital bed capacity and usage across secondary healthcare providers in England during the first wave of the COVID-19 pandemic: a descriptive analysis. *BMJ Open* 2021; **11**: e042945.
317 Care Quality Commission. The state of health care and adult social care in England 2020/21. 2021 https://www.cqc.org.uk/sites/default/files/20211021_stateofcare2021_print.pdf (accessed May 16, 2023).
318 McMillen T, Jani K, Robilotti E V., Kamboj M, Babady NE. The spike gene target failure (SGTF) genomic signature is highly accurate for the identification of Alpha and Omicron SARS-CoV-2 variants. *Sci Rep* 2022; **12**: 18968.
319 LTC Covid. COVID-19 and the Long-Term Care system in England (UK) – Resources to support community and institutional Long-Term Care responses to COVID-19. 2022. https://ltccovid.org/ltccovid-country-profile-england-uk/ (accessed March 8, 2023).
320 Jeffery-Smith A, Iyanger N, Williams S v., *et al.* Antibodies to SARS-CoV-2 protect against re-infection during outbreaks in care homes, September and October 2020. *Eurosurveillance* 2021; **26**: 2100092.

321 Tanunliong G, Liu A, Vijh R, *et al.* Persistence of Anti-SARS-CoV-2 Antibodies in Long Term Care Residents Over Seven Months After Two COVID-19 Outbreaks. *Front Immunol* 2022; **12**: 5646.
322 Alfego D, Sullivan A, Poirier B, Williams J, Adcock D, Letovsky S. A population-based analysis of the longevity of SARS-CoV-2 antibody seropositivity in the United States. *EClinicalMedicine* 2021; **36**: 100902.
323 van Elslande J, Oyaert M, Ailliet S, *et al.* Longitudinal follow-up of IgG anti-nucleocapsid antibodies in SARS-CoV-2 infected patients up to eight months after infection. *Journal of Clinical Virology* 2021; **136**: 104765.
324 Alasmari F, Mukahal M, Alqurashi AA, *et al.* Seroprevalence and longevity of SARS-CoV-2 nucleocapsid antigen-IgG among health care workers in a large COVID-19 public hospital in Saudi Arabia: A prospective cohort study. *PLoS One* 2022; **17**: e0272818.
325 Eyre DW, Lumley SF, Wei J, *et al.* Quantitative SARS-CoV-2 anti-spike responses to Pfizer–BioNTech and Oxford–AstraZeneca vaccines by previous infection status. *Clinical Microbiology and Infection* 2021; **27**: 1516.e7-1516.e14.
326 Andrews N, Tessier E, Stowe J, *et al.* Vaccine effectiveness and duration of protection of Comirnaty, Vaxzevria and Spikevax against mild and severe COVID-19 in the UK. *medRxiv* 2021; : 2021.09.15.21263583.
327 Parry H, Bruton R, Tut G, *et al.* Immunogenicity of single vaccination with BNT162b2 or ChAdOx1 nCoV-19 at 5–6 weeks post vaccine in participants aged 80 years or older: an exploratory analysis. *Lancet Healthy Longev* 2021; **2**: e554–60.
328 Meyers E, De Rop L, Deschepper E, *et al.* Prevalence of SARS-CoV-2 antibodies among Belgian nursing home residents and staff during the primary COVID-19 vaccination campaign. *European Journal of General Practice* 2022. DOI:10.1080/13814788.2022.2149732/SUPPL_FILE/IGEN_A_2149732_SM4488.DOCX.
329 Vicentini C, Zotti CM, Cornio AR, *et al.* Antibody responses to BNT162b2 SARS-CoV-2 mRNA vaccine among healthcare workers and residents of long-term care facilities: A cohort study in Northern Italy. *Health Sci Rep* 2023; **6**: e1087.
330 Lai CC, Wang JH, Hsueh PR. Population-based seroprevalence surveys of anti-SARS-CoV-2 antibody: An up-to-date review. *International Journal of Infectious Diseases* 2020; **101**: 314.
331 Chen S, Flegg JA, White LJ, Aguas R. Levels of SARS-CoV-2 population exposure are considerably higher than suggested by seroprevalence surveys. *PLoS Comput Biol* 2021; **17**. DOI:10.1371/JOURNAL.PCBI.1009436.
332 Dowell AC, Butler MS, Jinks E, *et al.* Children develop robust and sustained cross-reactive spike-specific immune responses to SARS-CoV-2 infection. *Nature Immunology 2021 23:1* 2021; **23**: 40–9.
333 Van Elslande J, Oyaert M, Lorent N, *et al.* Lower persistence of anti-nucleocapsid compared to anti-spike antibodies up to one year after SARS-CoV-2 infection. *Diagn Microbiol Infect Dis* 2022; **103**: 115659.
334 Peluso MJ, Takahashi S, Hakim J, *et al.* SARS-CoV-2 antibody magnitude and detectability are driven by disease severity, timing, and assay. *Sci Adv* 2021; **7**. DOI:10.1126/SCIADV.ABH3409/SUPPL_FILE/SCIADV.ABH3409_TABLES_S1_TO_S10.ZIP.

335 Stirrup O, Krutikov M, Tut G, *et al.* Severe Acute Respiratory Syndrome Coronavirus 2 Anti-Spike Antibody Levels Following Second Dose of ChAdOx1 nCov-19 or BNT162b2 Vaccine in Residents of Long-term Care Facilities in England (VIVALDI). *J Infect Dis* 2022; **226**: 1877–81.
336 Levin EG, Lustig Y, Cohen C, *et al.* Waning Immune Humoral Response to BNT162b2 Covid-19 Vaccine over 6 Months. *New England Journal of Medicine* 2021; **385**: e84.
337 Wei J, Pouwels KB, Stoesser N, *et al.* SARS-CoV-2 anti-spike IgG antibody responses after second dose of ChAdOx1 or BNT162b2 and correlates of protection in the UK general population. *medRxiv* 2022; : 2021.09.13.21263487.
338 Ward H, Whitaker M, Flower B, *et al.* Population antibody responses following COVID-19 vaccination in 212,102 individuals. DOI:10.1038/s41467-022-28527-x.
339 Daly M. COVID-19 and care homes in England: What happened and why? *Soc Policy Adm* 2020; **54**: 985.
340 Grimm F, Hodgson K, Brine R, Deeny SR. Hospital admissions from care homes in England during the COVID-19 pandemic: a retrospective, cross-sectional analysis using linked administrative data. *Int J Popul Data Sci* 2020; **5**. DOI:10.23889/IJPDS.V5I4.1663.
341 Zeng F, Wu M, Wang J, Li J, Hu G, Wang L. Over 1-year duration and age difference of SARS-CoV-2 antibodies in convalescent COVID-19 patients. *J Med Virol* 2021; **93**: 6506.
342 Navaratnam AMD, Shrotri M, Nguyen V, *et al.* Nucleocapsid and spike antibody responses following virologically confirmed SARS-CoV-2 infection: an observational analysis in the Virus Watch community cohort. *International Journal of Infectious Diseases* 2022; **123**: 104–11.
343 Ueno A, Kawasuji H, Miyajima Y, *et al.* Prolonged viral clearance of severe acute respiratory syndrome coronavirus 2 in the older aged population. *J Infect Chemother* 2021; **27**: 1119–21.
344 Zhong W, Yang X, Jiang X, *et al.* Factors associated with prolonged viral shedding in older patients infected with Omicron BA.2.2. *Front Public Health* 2023; **10**: 5349.
345 Owusu D, Pomeroy MA, Lewis NM, *et al.* Persistent SARS-CoV-2 RNA Shedding Without Evidence of Infectiousness: A Cohort Study of Individuals With COVID-19. *J Infect Dis* 2021; **224**: 1362–71.
346 Selva KJ, van de Sandt CE, Lemke MM, *et al.* Systems serology detects functionally distinct coronavirus antibody features in children and elderly. *Nat Commun* 2021; **12**. DOI:10.1038/S41467-021-22236-7.
347 Feng S, Phillips DJ, White T, *et al.* Correlates of protection against symptomatic and asymptomatic SARS-CoV-2 infection. *Nature Medicine 2021 27:11* 2021; **27**: 2032–40.
348 Tut G, Lancaster T, Krutikov M, *et al.* Profile of humoral and cellular immune responses to single doses of BNT162b2 or ChAdOx1 nCoV-19 vaccines in residents and staff within residential care homes (VIVALDI): an observational study. *Lancet Healthy Longev* 2021; **2**: e544–53.
349 Aldridge RW, Yavlinsky A, Nguyen V, *et al.* SARS-CoV-2 antibodies and breakthrough infections in the Virus Watch cohort. *Nature Communications 2022 13:1* 2022; **13**: 1–8.

350 Fischinger S, Boudreau CM, Butler AL, Streeck H, Alter G. Sex differences in vaccine-induced humoral immunity. *Semin Immunopathol* 2019; **41**: 239.
351 Pérez-Alós L, Armenteros JJA, Madsen JR, *et al.* Modeling of waning immunity after SARS-CoV-2 vaccination and influencing factors. *Nature Communications 2022 13:1* 2022; **13**: 1–11.
352 Nakagama Y, Komase Y, Kaku N, *et al.* Detecting Waning Serological Response with Commercial Immunoassays: 18-Month Longitudinal Follow-up of Anti-SARS-CoV-2 Nucleocapsid Antibodies. *Microbiol Spectr* 2022; **10**. DOI:10.1128/SPECTRUM.00986-22/SUPPL_FILE/REVIEWER-COMMENTS.PDF.
353 Torres Ortiz A, Fenn Torrente F, Twigg A, *et al.* The influence of time on the sensitivity of SARS-CoV-2 serological testing. *Sci Rep* 2022; **12**: 10517.
354 Moss P. The T cell immune response against SARS-CoV-2. *Nature Immunology 2022 23:2* 2022; **23**: 186–93.
355 Tut G, Lancaster T, Butler MS, *et al.* Reduced Antibody and Cellular Immune Responses Following Dual COVID-19 Vaccination Within Infection-Naïve Residents of Long-Term Care Facilities. *SSRN Electronic Journal* 2021. DOI:10.2139/SSRN.3979590.
356 Earle KA, Ambrosino DM, Fiore-Gartland A, *et al.* Evidence for antibody as a protective correlate for COVID-19 vaccines. *Vaccine* 2021; **39**: 4423–8.
357 Feng S, Phillips DJ, White T, *et al.* Correlates of protection against symptomatic and asymptomatic SARS-CoV-2 infection. *Nature Medicine 2021 27:11* 2021; **27**: 2032–40.
358 Dyer AH, Noonan C, McElheron M, *et al.* Previous SARS-CoV-2 Infection, Age, and Frailty Are Associated With 6-Month Vaccine-Induced Anti-Spike Antibody Titer in Nursing Home Residents. *J Am Med Dir Assoc* 2022; **23**: 434.
359 Public Health England. Evaluation of the Abbott SARS-CoV-2 IgG for the detection of anti-SARS-CoV-2 antibodies. 2020 https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/890566/Evaluation_of_Abbott_SARS_CoV_2_IgG_PHE.pdf (accessed March 24, 2023).
360 Matsunaga H, Takeuchi H, Oba Y, Fujimi S, Honda T, Tomonaga K. Waning of Anti-SARS-CoV-2 Spike Antibody Levels 100 to 200 Days after the Second Dose of the BNT162b2 Vaccine. *Vaccines 2022, Vol 10, Page 177* 2022; **10**: 177.
361 Yates TA, Tanser F, Abubakar I. Plan Beta for tuberculosis: it's time to think seriously about poorly ventilated congregate settings. *The International Journal of Tuberculosis and Lung Disease* 2016; **20**: 5.
362 Escombe AR, Oeser CC, Gilman RH, *et al.* Natural Ventilation for the Prevention of Airborne Contagion. *PLoS Med* 2007; **4**: e68.
363 Menzies D, Fanning A, Yuan L, *et al.* Hospital ventilation and risk for tuberculous infection in canadian health care workers. Canadian Collaborative Group in Nosocomial Transmission of TB. *Ann Intern Med* 2000; **133**. DOI:10.7326/0003-4819-133-10-200011210-00010.
364 Zemouri C, Awad SF, Volgenant CMC, Crielaard W, Laheij AMGA, de Soet JJ. Modeling of the Transmission of Coronaviruses, Measles Virus, Influenza Virus, Mycobacterium tuberculosis, and Legionella pneumophila in Dental Clinics. *J Dent Res* 2020; **99**: 1192–8.

365 Andrews JR, Morrow C, Walensky RP, Wood R. Integrating Social Contact and Environmental Data in Evaluating Tuberculosis Transmission in a South African Township. *J Infect Dis* 2014; **210**: 597.
366 Eijkelenboom A, Verbeek H, Felix E, van Hoof J. Architectural factors influencing the sense of home in nursing homes: An operationalization for practice. *Frontiers of Architectural Research* 2017; **6**: 111–22.
367 Hayward AC, Harling R, Wetten S, *et al.* Effectiveness of an influenza vaccine programme for care home staff to prevent death, morbidity, and health service use among residents: cluster randomised controlled trial. *BMJ : British Medical Journal* 2006; **333**: 1241.
368 The Health Foundation. NHS Test and Trace: the journey so far. 2020. https://www.health.org.uk/publications/long-reads/nhs-test-and-trace-the-journey-so-far (accessed April 19, 2023).
369 Green Building Store. What is MVHR Heat Recovery Ventilation? 2023. https://www.greenbuildingstore.co.uk/information-hub/heat-recovery-ventilation-mvhr/ (accessed May 19, 2023).
370 LearnMetrics. 10 Popular Air Conditioner Types (With Pictures, Prices). 2023. https://learnmetrics.com/air-conditioner-types/ (accessed May 19, 2023).
371 Seyam S. Types of HVAC Systems. In: Kandelousi MS, ed. HVAC System. London: IntechOpen, 2016. DOI:10.5772/intechopen.78942.
372 Sharma A, Minh Duc NT, Luu Lam Thang T, *et al.* A Consensus-Based Checklist for Reporting of Survey Studies (CROSS). *J Gen Intern Med* 2021; **36**: 3179.
373 Office for National Statistics. Census 2021 geographies - Office for National Statistics. 2021. https://www.ons.gov.uk/methodology/geography/ukgeographies/censusgeographies/census2021geographies#lower-layer-super-output-areas-lsoas (accessed April 19, 2023).
374 GOV.UK. English indices of deprivation 2019. 2019. https://www.gov.uk/government/statistics/english-indices-of-deprivation-2019 (accessed April 19, 2023).
375 GOV.UK. COVID-19: management of staff and exposed patients or residents in health and social care settings. 2022. https://www.gov.uk/government/publications/covid-19-management-of-exposed-healthcare-workers-and-patients-in-hospital-settings/covid-19-management-of-exposed-healthcare-workers-and-patients-in-hospital-settings#section4-1 (accessed Feb 18, 2022).
376 GOV.UK. COVID-19 care home testing guidance for regular and outbreak testing of staff and residents. 2022. https://www.gov.uk/government/publications/coronavirus-covid-19-testing-in-adult-care-homes/covid-19-care-home-testing-guidance-for-regular-and-outbreak-testing-of-staff-and-residents#outbreak-testing (accessed Feb 18, 2022).
377 Schafer JL, Graham JW. Missing data: Our view of the state of the art. *Psychol Methods* 2002; **7**: 147–77.
378 Hughes RA, Heron J, Sterne JAC, Tilling K. Accounting for missing data in statistical analyses: multiple imputation is not always the answer. *Int J Epidemiol* 2019; **48**: 1294–304.

379 Viana R, Moyo S, Amoako DG, *et al.* Rapid epidemic expansion of the SARS-CoV-2 Omicron variant in southern Africa. *Nature 2022 603:7902* 2022; **603**: 679–86.
380 Liu Y, Rocklöv J. The effective reproductive number of the Omicron variant of SARS-CoV-2 is several times relative to Delta. *J Travel Med* 2022; **29**: 1–4.
381 GOV.UK. [Withdrawn] Coronavirus (COVID-19) vaccination as a condition of deployment for the delivery of CQC-regulated activities in wider adult social care settings. 2022. https://www.gov.uk/government/publications/vaccination-of-workers-in-social-care-settings-other-than-care-homes-operational-guidance/coronavirus-covid-19-vaccination-as-a-condition-of-deployment-for-the-delivery-of-cqc-regulated-activities-in-wider-adult-social-care-settings (accessed May 16, 2023).
382 GOV.UK. England Summary | Coronavirus (COVID-19) in the UK. 2022. https://coronavirus.data.gov.uk/ (accessed Aug 26, 2022).
383 Dietz L, Horve PF, Coil DA, Fretz M, Eisen JA, van den Wymelenberg K. 2019 Novel Coronavirus (COVID-19) Pandemic: Built Environment Considerations To Reduce Transmission. *mSystems* 2020; **5**. DOI:10.1128/MSYSTEMS.00245-20/SUPPL_FILE/REVIEWER-COMMENTS.PDF.
384 Stratil JM, Biallas RL, Burns J, *et al.* Non-pharmacological measures implemented in the setting of long-term care facilities to prevent SARS-CoV-2 infections and their consequences: a rapid review. *Cochrane Database of Systematic Reviews* 2021; **2021**. DOI:10.1002/14651858.CD015085.PUB2/MEDIA/CDSR/CD015085/IMAGE_N/NCD015085-FIG-09.PNG.
385 Murphy E, Burke E, O'Donoghue J, *et al.* 212 BUILT ENVIRONMENT OF NURSING HOMES AND INFECTION CONTROL AND PANDEMIC PREPAREDNESS. *Age Ageing* 2021; **50**: ii9–41.
386 Benbow W. COVID-19 in Long-Term Care: The Built Environment Impact on Infection Control. *Health Environments Research and Design Journal* 2022; **15**: 287–98.
387 Bowes A, Dawson A. Designing Environments for People with Dementia. Emerald Publishing Limited, 2019 DOI:10.1108/978-1-78769-971-720191004.
388 Burton E, Sheehan B. Care-home environments and well-being: Identifying the design features that most affect older residents. *J Archit Plann Res* 2010; **27**: 237–56.
389 Grey X, O'Donoghue B, Puntambekar O. Residential Long-Term care and Covid-19: Role of the built environment in balancing infection control and quality of life - Key Findings and Recommendations 2022. 2022 www.tcd.ie/trinityhaus (accessed April 19, 2023).
390 Zimmerman S, Bowers BJ, Cohen LW, *et al.* New Evidence on the Green House Model of Nursing Home Care: Synthesis of Findings and Implications for Policy, Practice, and Research. *Health Serv Res* 2016; **51**: 475.
391 Zimmerman S, Dumond-Stryker C, Tandan M, *et al.* Nontraditional Small House Nursing Homes Have Fewer COVID-19 Cases and Deaths. *J Am Med Dir Assoc* 2021; **22**: 489–93.
392 Silva DR, Viana VP, Coelho AC, *et al.* Viral Respiratory Infections In Adult Patients Treated In The Emergency Room. *American Thoracic Society International Conference Meetings Abstracts American Thoracic Society International Conference Meetings Abstracts* 2010; : A6167–A6167.

393 Nair V, Sahni AK, Sharma D, *et al.* Point prevalence & risk factor assessment for hospital-acquired infections in a tertiary care hospital in Pune, India. *Indian J Med Res* 2017; **145**: 824.
394 Wu HT, Li QS, Dai RC, *et al.* Effects of air-conditioning systems in the public areas of hospitals: A scoping review. *Epidemiol Infect* 2021; **149**. DOI:10.1017/S0950268821001990.
395 Konetzka RT, White EM, Pralea A, Grabowski DC, Mor V. A systematic review of long-term care facility characteristics associated with COVID-19 outcomes. *J Am Geriatr Soc* 2021; **69**: 2766–77.
396 SAGE-EMG. Role of Ventilation in Controlling SARS-CoV-2 Transmission. 2020 https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/928720/S0789_EMG_Role_of_Ventilation_in_Controlling_SARS-CoV-2_Transmission.pdf (accessed April 19, 2023).
397 de Man P, Paltansing S, Ong DSY, Vaessen N, van Nielen G, Koeleman JGM. Outbreak of Coronavirus Disease 2019 (COVID-19) in a Nursing Home Associated With Aerosol Transmission as a Result of Inadequate Ventilation. *Clinical Infectious Diseases* 2021; **73**: 170–1.
398 Allen JG, Ibrahim AM. Indoor Air Changes and Potential Implications for SARS-CoV-2 Transmission. *JAMA* 2021; **325**: 2112–3.
399 Ghoroghi A, Rezgui Y, Wallace R. Impact of ventilation and avoidance measures on SARS-CoV-2 risk of infection in public indoor environments. *Science of The Total Environment* 2022; **838**: 156518.
400 Bhagat RK, Davies Wykes MS, Dalziel SB, Linden PF. Effects of ventilation on the indoor spread of COVID-19. *J Fluid Mech* 2020; **903**: F1.
401 Department of Health. Care Homes for Older People, National Minimum Standards, Care Home Regulations, 3rd Edition. London, 2003 https://webarchive.nationalarchives.gov.uk/ukgwa/20130107105354/http:/dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_4135403.pdf (accessed Dec 1, 2022).
402 Verheyen CA, Bourouiba L. Associations between indoor relative humidity and global COVID-19 outcomes. *J R Soc Interface* 2022; **19**. DOI:10.1098/RSIF.2021.0865.
403 Park HJ, Lee SG, Oh JS, *et al.* The effects of indoor temperature and humidity on local transmission of COVID-19 and how it relates to global trends. *PLoS One* 2022; **17**. DOI:10.1371/JOURNAL.PONE.0271760.
404 Morawska L, Allen J, Bahnfleth W, *et al.* A paradigm shift to combat indoor respiratory infection. *Science (1979)* 2021; **372**: 689–91.
405 Scottish Government. Coronavirus (COVID-19): care home outbreaks - root cause analysis. Scottish Government, 2021 http://www.gov.scot/publications/root-cause-analysis-care-home-outbreaks/pages/18/ (accessed April 19, 2023).
406 ISRCTN registry. Contact tracing in care homes using digital technology for improving infection control including COVID-19. 2022. https://www.isrctn.com/ISRCTN11204126.
407 Van der Weele TJ, Mathur MB. SOME DESIRABLE PROPERTIES OF THE BONFERRONI CORRECTION: IS THE BONFERRONI CORRECTION REALLY SO BAD? *Am J Epidemiol* 2019; **188**: 617–8.
408 Dunn OJ. Multiple Comparisons among Means. *J Am Stat Assoc* 1961; **56**: 52–64.

409 Goldblatt D, Alter G, Crotty S, Plotkin SA. Correlates of protection against SARS-CoV-2 infection and COVID-19 disease. *Immunol Rev* 2022; **310**: 6–26.
410 Vivaldi G, Jolliffe DA, Faustini S, *et al.* Correlation Between Postvaccination Anti-Spike Antibody Titers and Protection Against Breakthrough Severe Acute Respiratory Syndrome Coronavirus 2 Infection: A Population-Based Longitudinal Study. *J Infect Dis* 2022; **226**: 1903–8.
411 Gilbert PB, Montefiori DC, McDermott AB, *et al.* Immune correlates analysis of the mRNA-1273 COVID-19 vaccine efficacy clinical trial. *Science* 2022; **375**: 43.
412 Peng Y, Felce SL, Dong D, *et al.* An immunodominant NP105–113-B*07:02 cytotoxic T cell response controls viral replication and is associated with less severe COVID-19 disease. *Nature Immunology 2021 23:1* 2021; **23**: 50–61.
413 Khoury DS, Cromer D, Reynaldi A, *et al.* Neutralizing antibody levels are highly predictive of immune protection from symptomatic SARS-CoV-2 infection. *Nat Med* 2021; **27**: 1205–11.
414 Tut G, Lancaster T, Butler MS, *et al.* Robust SARS-CoV-2-specific and heterologous immune responses in vaccine-naïve residents of long-term care facilities who survive natural infection. *Nature Aging 2022 2:6* 2022; **2**: 536–47.
415 Tut G, Lancaster T, Krutikov M, *et al.* Strong peak immunogenicity but rapid antibody waning following third vaccine dose in older residents of care homes. *Nature Aging 2023 3:1* 2023; **3**: 93–104.
416 PROTECT COVID-19 National Core Study. Cross sector study. 2023. https://sites.manchester.ac.uk/covid19-national-project/research-themes/sector-specific-studies/cross-sector-study/ (accessed May 2, 2023).
417 The AFRI-c Study. About the Study. 2023. https://afric.blogs.bristol.ac.uk/about_the_study/ (accessed May 2, 2023).
418 UCL. VIVALDI Study | UCL Institute of Health Informatics. 2023. https://www.ucl.ac.uk/health-informatics/research/vivaldi-study (accessed May 2, 2023).
419 ISRCTN. ISRCTN14447421: COVID-19 in care homes (VIVALDI). 2020. https://www.isrctn.com/ISRCTN14447421 (accessed May 18, 2023).
420 Co-Production Collective. Welcome to the Co-Production Collective! 2021. https://www.coproductioncollective.co.uk/ (accessed May 23, 2023).
421 Health Research Authority. Fast track review guidance for COVID-19 studies. 2022. https://www.hra.nhs.uk/covid-19-research/fast-track-review-guidance-covid-19-studies/ (accessed May 23, 2023).
422 Mohanraj D, Bicknell K, Bhole M, Webber C, Taylor L, Whitelegg A. Antibody responses to SARS-CoV-2 infection—comparative determination of seroprevalence in two high-throughput assays versus a sensitive spike protein ELISA. *Vaccines (Basel)* 2021; **9**: 1310.
423 Robinson A V., Weaving G, Philips BJ, Eziefula AC, Shipman KE, Chevassut T. Real-world experience of SARS-CoV-2 antibody assays in UK healthcare workers. *Clinical Medicine* 2021; **21**: e300–5.
424 Barda N. The ups and downs of observational vaccine research. *Lancet Infect Dis* 2023; **23**: 767–8.
425 Hulme WJ, Williamson E, Horne EMF, *et al.* Challenges in Estimating the Effectiveness of COVID-19 Vaccination Using Observational Data. *Ann Intern Med* 2023; **176**: 685–93.

426 Jeffery-Smith A, Rowland TAJ, Patel M, *et al.* Reinfection with new variants of SARS-CoV-2 after natural infection: a prospective observational cohort in 13 care homes in England. *Lancet Healthy Longev* 2021; **2**: e811.
427 Bijani A, Ghadimi R, Mikaniki E, *et al.* Cohort Profile Update: The Amirkola Health and Ageing Project (AHAP). *Caspian J Intern Med* 2017; **8**: 205.
428 Schooling CM, Chan WM, Leung SL, *et al.* Cohort Profile: Hong Kong Department of Health Elderly Health Service Cohort. *Int J Epidemiol* 2016; **45**: 64.
429 Sijtsma A, Rienks J, van der Harst P, Navis G, Rosmalen JGM, Dotinga A. Cohort Profile Update: Lifelines, a three-generation cohort study and biobank. *Int J Epidemiol* 2022; **51**: e295–302.
430 Hewitt J, Walters M, Padmanabhan S, Dawson J. Cohort profile of the UK Biobank: diagnosis and characteristics of cerebrovascular disease. *BMJ Open* 2016; **6**: e009161.
431 Biobank. Enable your research. 2022. https://www.ukbiobank.ac.uk/enable-your-research (accessed May 2, 2023).
432 WHO. Global strategy and action plan on ageing and health. 2017. https://www.who.int/publications/i/item/9789241513500 (accessed Feb 8, 2023).
433 Burton JK, Wolters AT, Towers AM, *et al.* Developing a minimum data set for older adult care homes in the UK: exploring the concept and defining early core principles. *Lancet Healthy Longev* 2022; **3**: e186–93.
434 Bernabei R, Landi F, Onder G, Liperoti R, Gambassi G. Second and third generation assessment instruments: the birth of standardization in geriatric care. *J Gerontol A Biol Sci Med Sci* 2008; **63**: 308–13.
435 Rahman AN, Applebaum RA. The nursing home Minimum Data Set assessment instrument: manifest functions and unintended consequences--past, present, and future. *Gerontologist* 2009; **49**: 727–35.
436 Dash D, Heckman GA, Boscart VM, Costa AP, Killingbeck J, d'Avernas JR. Using powerful data from the interRAI MDS to support care and a learning health system: A case study from long-term care. *Healthc Manage Forum* 2018; **31**: 153–9.
437 Health Research Authority. Confidential patient information. 2023. https://www.hra.nhs.uk/about-us/committees-and-services/confidentiality-advisory-group/confidential-patient-information-and-regulations/ (accessed May 2, 2023).