Longer-term effects of SARS-CoV-2 infection on blood vessels and blood pressure.

Stefanie Zhao Lin Lip

BSc Med Sci, MBChB, MRCP (UK)

Submitted in fulfilment for the requirements for the degree of Doctor of Philosophy

BHF Cardiovascular Research Centre

School of Cardiovascular and Metabolic Health

University of Glasgow

January 2025

Abstract

Introduction

This thesis aims to address critical gaps in our understanding of the interplay between COVID-19 infection, Renin-Angiotensin-Aldosterone System (RAAS) dysregulation, and hypertension. The first objective investigates whether COVID-19 infection increases the risk of developing hypertension post-recovery by conducting a prospective follow-up of non-hypertensive individuals over a 12-month period, assessing blood pressure and endothelial function. The second objective explores the role of RAAS dysregulation in post-COVID-19 hypertension by analysing RAAS pathway components and blood pressure changes in the same cohort. The third objective seeks to determine if whether individuals post COVID-19 have any differences in QoL utilising EQ-5D-3L instrument (EQ-5D-3L Index and Visual Analog Scale (VAS) scores). The fourth objective examines the association of ACE inhibitors and different antihypertensive drug classes and statins on the risk of SARS-CoV-2 infection using longitudinal studies and machine learning techniques to analyse linked electronic health records, adjusting for confounding variables.

Methods

To assess whether COVID-19 increases the risk of hypertension post-recovery, we conducted a 12-month follow-up of non-hypertensive individuals. Blood pressure and endothelial function were monitored using ambulatory blood pressure monitoring and brachial flow-mediated dilation. Quality of life (QoL) was measured with the EQ-5D-3L questionnaire. RAAS components and blood pressure changes were analysed to investigate RAAS dysregulation. Advanced machine learning techniques were applied to estimate individual treatment effects (ITE) for the four major classes of antihypertensive drugs in comparison to statins across two distinct time frames of the COVID-19 pandemic.

Results

There was a significant rise in blood pressure among recovered COVID-19 patients, with systolic pressure increasing by 4.57 mmHg and diastolic by 4.46 mmHg over 12 months compared to controls. A 3.15% reduction in flow-mediated dilation (FMD) suggested endothelial dysfunction. No significant difference in RAAS fingerprinting was observed. While recovered COVID-19 participants reported lower QoL, this was not statistically significant at 12 months follow-up. Our machine learning model found ACE inhibitors and statins were associated with increased SARS-CoV-2 infection risk, while thiazides showed

mixed effects, and beta blockers and calcium channel blockers were associated with decreased risk.

Discussion

The increase in blood pressure seen 12 months after recovery from COVID-19 suggest the need for prioritising cardiovascular monitoring in post-COVID-19 era. Although RAAS fingerprinting showed no significant difference, the blood pressure rise and reduced FMD suggest RAAS dysregulation may contribute to post-infection hypertension. Machine learning-based ITE estimation could potentially revolutionise studies of drug efficacy and adverse reactions, especially when randomised controlled trials are impractical.

Conclusion

This thesis advances our understanding of COVID-19's cardiovascular consequences and provides insights for future mechanistic studies and clinical and public health policies. The observed blood pressure rise, and potential endothelial dysfunction post-recovery indicate the need for vigilant cardiovascular monitoring. The use of machine learning to estimate individual effects of antihypertensive drugs on COVID-19 risk underscores the importance of personalised treatment. Further research should elucidate long-term cardiovascular impacts and develop targeted interventions.

Publications relating to this work

1. Lip S, Mccallum L, Delles C, et al. Rationale and Design for the LOnger-term effects of SARS-CoV-2 INfection on blood Vessels And blood pRessure (LOCHINVAR): an observational phenotyping study. Open Heart 2022;9:e002057. doi: 10.1136/openhrt-2022-002057
2. Lip, S. A tale of two diseases. J Hum Hypertens 37, 248–251 (2023). https://doi.org/10.1038/s41371-022-00798-3

Publications arising during this PhD

Papers Published

3. Tran TQB, **Lip S**, du Toit C et al. Assessing Machine Learning for Diagnostic Classification of Hypertension Types Identified by Ambulatory Blood Pressure Monitoring. CJC Open 2024.
4. McCallum L, **Lip S**, Rostron M, Hanna R, Bin Pg Md Salimin N, Nichol S, Padmanabhan S. OPTIMA-BP: empOwering PaTients in MAnaging Blood Pressure - protocol for a randomised parallel group study comparing use of Kvatchii web-based patient education portal as an addition to home blood pressure monitoring. Open Heart. 2024 Mar 1;11(1):e002535.
5. du Toit C, Tran TQB, Deo N, Aryal S, **Lip S**, Sykes R, Manandhar I et al. Survey and Evaluation of Hypertension Machine Learning Research. Journal of the American Heart Association. 2023;12:e027896
6. Bird G, D'Agostin D, Alsanosi S, **Lip S**, Padmanabhan S, Parekh AB. A Reappraisal of the Effects of L-type Ca2+ Channel Blockers on Store-Operated Ca2+ Entry and Heart Failure, *Function*, Volume 4, Issue 6, 2023, https://doi.org/10.1093/function/zqad047
7. **Lip, S**, Dempster, G, Jain, S. et al. Underrepresentation of ethnic minorities in hypertension research—a survey of enablers and barriers among South Asian and African communities in Glasgow. Trials 23, 609 (2022). https://doi.org/10.1186/s13063-022-06542-z
8. Le NN, Tran TQB, **Lip S**, McCallum L, McClure J, Dominiczak AF, Gill D, Padmanabhan S. Unravelling the Distinct Effects of Systolic and Diastolic Blood Pressure Using Mendelian Randomisation. *Genes*. 2022; 13(7):1226. https://doi.org/10.3390/genes13071226

9. **Lip, S**, McCallum, L, & Padmanabhan, S. (2023). Hypertension Residual Risk and Beyond – Five Decades of Insights from the Glasgow Blood Pressure Clinic. Hypertension Journal, 7(2), 64–70.
10. Toit, C. du, McCallum, L., **Lip, S**., & Padmanabhan, S. (2023). Transforming Hypertension Care through the Lens of Planetary Health. Hypertension Journal, 7(2), 82–87.

Abstracts

1. **Lip. S**, McCallum L, Strachan K, McKinstry B, Webb D, Padmanabhan S. Epidemiological prevalence estimates of hypertension of 856,878 adults in Glasgow. Journal of Hypertension 39():p e84-e85, April 2021. | DOI: 10.1097/01.hjh.0000745232.47837.d4
2. **Lip, S**, McCallum, L., & Padmanabhan, S. Serum Chloride and Adverse Cardiovascular Risk in the West of Scotland General Population. Journal of Hypertension 39():p e14, April 2021. | DOI: 10.1097/01.hjh.0000744508.81161.2e
3. McCallum L, **Lip S**, Rios et al. COVID-19 blood pressure endothelium interaction study (OBELIX)-blood pressure assessment, vascular phenotyping and RAAS-fingerprinting of COVID-19 survivors, Journal of Human Hypertension

Awards and prizes arising during this PhD

1. Inspirational Role Model, Commended, Medical Directorate Awards 2024 (29th April 2024)
2. John Munro Medal Prize, Royal College of Physicians of Edinburgh (10th October 2023)
3. Stanley Peart Essay Prize, British and Irish Hypertension Society (13th September 2022)
4. Rang Prize for Outstanding Achievement in Teaching, British Pharmacological Society (13th of September 2022)
5. Winner, MVLS Public Engagement Awards 2021, Breaking a lifetime of disadvantage by enhancing engagement of ethnic minorities in health research (25th of January 2022)
6. Top 4 Finalists, Climate Change and Health Webinar, Royal College of Physicians and Surgeons, Glasgow, 29th October 2021.

Table of Contents

List of Tables

List of Figures

Acknowledgements

I extend my heartfelt gratitude to my supervisors, Professor Sandosh Padmanabhan for his unwavering support and encouragement throughout this project. Professor Sandosh Padmanabhan's guidance, spanning over a decade, has been instrumental in shaping my academic journey, instilling in me inspiration, and providing invaluable opportunities. I would like to thank Professor Colin Berry for being the co-supervisor for my PhD project. I also want to acknowledge Dr. Linsay McCallum for her invaluable advice, friendship, and dedicated time; her contributions have been immensely appreciated.

Special appreciation is extended to HEART Research UK (RG2690/21/24) for their funding support, and to the School of Cardiovascular and Metabolic Health for appointing me as a Clinical Research Fellow. My gratitude also extends to the staff at the Glasgow Clinical Research Facility, particularly Ms. Susan Yip, Ms. Caitlin Gray for their assistance in the execution of this study.

The selflessness of the trial participants from OBELIX and LOCHINVAR, who generously devoted their time to partake in the clinical phenotyping study, made this research possible. I would also like to acknowledge Tran Quoc Bao Tran for his invaluable contributions to the machine learning methodology outlined in Chapter 6 and Appendices. This thesis is dedicated to all healthcare professionals tirelessly serving within the NHS including those I worked closely with at the frontline at the Queen Elizabeth University Hospital of whom there are too many names to mention, whose unwavering dedication and selfless commitment have been instrumental in combating the COVID-19 pandemic. I would like to thank Dr Erica Peters and Dr Beth White involving me in developing COVID-19 infographics for healthcare professionals in NHS Greater Glasgow and Clyde. This thesis is also dedicated to the patients, frontline workers and the public who have bravely endured the hardships of COVID-19 and to those who have tragically lost loved ones during this unprecedented crisis. Your resilience, courage, and sacrifices inspire us all and serve as a poignant reminder of the profound impact of this global health emergency.

Lastly, I dedicate this thesis to my family. To my parents, Michael Lip and Jean Lip, and my brothers, Sean Lip and Quinton Lip, I am profoundly indebted for the unwavering support and love they have bestowed upon me. My husband, Ronald Gordon Edward Brown, deserves special recognition for his enduring patience and understanding throughout the years of my intense focus on this project. His decision to marry me during my second year of the PhD reflects his boundless kindness and support, without which I would be incomplete. To him, I owe an eternal debt of gratitude.

Author’s declaration

The research detailed in this thesis was conducted during my employment as a Clinical Research Fellow at the BHF Glasgow Cardiovascular Research Centre, University of Glasgow, within the School of Cardiovascular and Metabolic Health. Under the guidance of Professor Sandosh Padmanabhan and Professor Colin Berry, I undertook the study protocol, which was primarily designed by Professor Sandosh Padmanabhan, with my collaborative input. I was actively involved all aspects of this research project which includes preparing the research project for ethical approval, being trained in all study procedures, screening, recruitment, obtaining informed consent from study participants, conducting all study procedures and conducting follow-up visits.

The analysis was a collaborative effort, conducted jointly with Professor Sandosh Padmanabhan, with valuable input from Professor John McClure.

I affirm that this thesis is my original work, completed independently, and has not been previously submitted for any other degree at the University of Glasgow or elsewhere. All sources of information utilised within this thesis are duly acknowledged.

Dr. Stefanie Zhao Lin Lip

January 2025

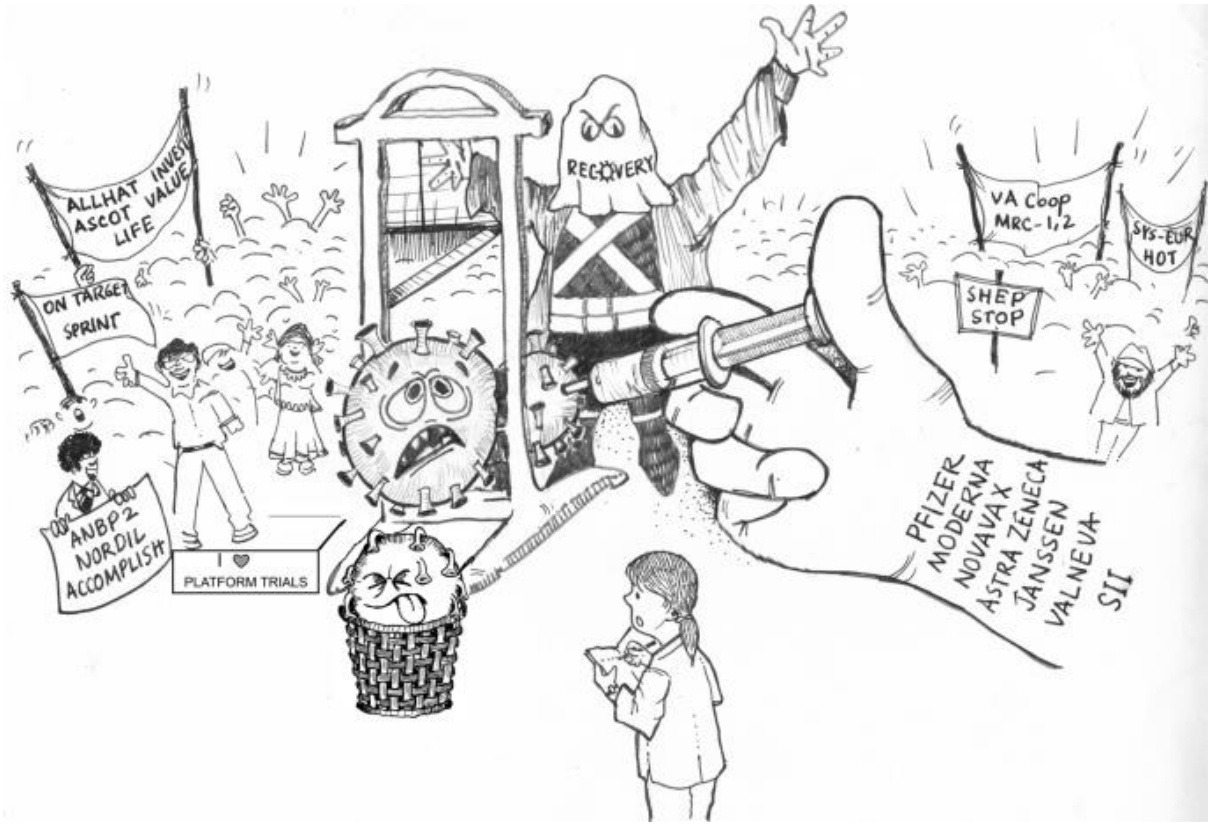
List of Abbreviations

%FMD Assessment of endothelial function
6MWT 6-Minute Walk Test
95% CI 95% Confidence interval
AA2 Ratio of aldosterone/Ang II
ABPM Ambulatory blood pressure monitoring
ACE2 Angiotensin-converting enzyme 2
ACE-S Angiotensin converting enzyme
ACEI Angiotensin converting enzyme inhibitor
AI Artificial Intelligence
AIDS Acquired Immunodeficiency Syndrome
ALT Serum Alanine Transaminase
Ang 1 Angiotensin I
Ang II Angiotensin II
Angiotensin 1-7 (Ang 1-7)
aOR Adjusted Odds Ratio
ARB Aldosterone receptor blockers
ATE Average Treatment Effect
AVRI Acute viral respiratory illness
BB Beta blockers
BIHS British and Irish Hypertension Society
BMI Body Mass Index
BP Blood pressure
Ca(adj) Serum adjusted Calcium
CASTOR EDC (www.castoredc.com)
CATE Conditional Average Treatment Effect
CCB Calcium channel blocker
CDC Centres for Disease Control and Prevention
CFS Chronic Fatigue Syndrome
CHI Community health index
Cl Serum Chloride
COVAX COVID-19 Vaccine Global Access
COVID Coronavirus disease
COVID-19 Coronavirus disease 2019
DBP Diastolic Blood Pressure
DID Difference-in-Differences
EQ-5D EuroQol-5 Dimension
EQ5D-3L Euroqol (EQ5D-3L) Questionnaire
FMD Brachial flow mediated dilatation
GCRF Glasgow Clinical Research Facility
Hb Haemoglobin
HbA1c Glycosylated haemoglobin
HBPM Home blood pressure monitoring
HDL High density lipoprotein
HIV Human immunodeficiency virus
HR Hazard Ratio
IQR Interquartile range
ITE Individual treatment effect
K Serum Potassium
LOCHINVAR Longer-term effects of SARS-CoV-2 infection on blood vessels and blood pressure
LPAC Safe Haven Local Privacy and Advisory Committee
ME Myalgic encephalitis
MERS-CoV Middle East Respiratory Syndrome Coronavirus
Mg Serum Magnesium
ML Machine Learning
Na Serum Sodium
NETs Neutrophil Extracellular Traps
NHS GG&C NHS Greater Glasgow & Clyde
nsps non-structural proteins
OBELIX COVID-19 Blood pressure EndotheLium Interaction
oBP Office Blood Pressure
oDBP Office Diastolic Blood Pressure
oSBP Office Systolic BP
PASC Post-Acute Sequelae of SARS-CoV-2 Infection
POTS Postural orthostatic tachycardia syndrome
PRA-S Angiotensin based markers for renin
QALYs Quality-adjusted life years
QEUH Queen Elizabeth University Hospital
QoL Quality of Life
RAAS Renin-Angiotensin-Aldosterone System
RCTs randomised controlled trials
RDD Regression discontinuity designs
RNA Ribonucleic acid
ROS Reactive oxygen species
RT-PCR Reverse transcription polymerase chain reaction
SARS Severe Acute Respiratory Syndrome
SARS-CoV-1 Severe acute respiratory syndrome coronavirus 1
SARS-CoV-2 Severe acute respiratory syndrome coronavirus 2
SBP Systolic blood pressure
SD Standard Deviation
SIMD Scottish Index of Multiple Deprivation
THZ Thiazide
uClCr Urine Chloride Creatinine ratio
UK United Kingdom
uKCr Urine Potassium Creatinine ratio
uNaCr Urine Sodium Creatinine ratio
VAS Visual Analog Scale
WBC White Blood Cell Count
WHO World Health Organization

Chapter 1 Introduction

As we navigate through the aftermath of the coronavirus disease 2019 (COVID-19) pandemic, the significant human cost of this viral illness becomes increasingly apparent. However, the history of humankind is marked by pandemics that have repeatedly ravaged societies, influenced the outcomes of wars, annihilated entire populations, and yet, paradoxically, paved the path for groundbreaking advancements in various fields, including medicine, public health, the economy, and political systems. In my early years of my PhD, I submitted a prize-winning Stanley Peart Essay which summarised the nearly 2 millennia years of hypertension and 2 years of COVID-19 experiences which evoked hidden parallels between these two conditions. (Figure 1) (1)

Figure 1 COVID-19 pandemic and Hypertension



"It is a far, far better thing that I do, than I have ever done; it is a far, far better rest I go to than I have ever known." Dr. Stefanie Lip draws inspiration from the sacrifices and scientific advances during the COVID-19 pandemic to envision how hypertension research can be transformed. Illustration by Dr. Kushal K Choudhuri (The copyright belongs to the author – Stefanie Lip).

This thesis describes a set of studies investigating the long-term cardiovascular consequences of COVID-19, RAAS dysregulation with a specific emphasis on

hypertension as a complication following infection. It investigates if COVID-19 increases hypertension risk post-infection by studying recovered non-hypertensive patients over a year, explores Renin-Angiotensin-Aldosterone System (RAAS) dysregulation's role in mediating post-COVID-19 hypertension risk, and examines how different antihypertensive drug classes affect COVID-19 risk and severity.

In this introduction, I start with a historical review of pandemics and their profound health, social, economic, and political ramifications followed by describing the COVID-19 pandemic and focusing on its cardiovascular consequences and specifically its relationship with hypertension. By examining the potential long-term health effects on hypertension—a widely prevalent cardiovascular risk factor—we can derive insights into public health prevention strategies. These strategies are pertinent for reducing the burden of future cardiometabolic risks.

1.1 Definitions

The distinctions among the terms "pandemic," "epidemic," "endemic," and "outbreak" are important for understanding the scope and impact of diseases.(2, 3) These terms, each with their specific definitions, help public health officials, researchers, and the general public grasp the extent of health challenges posed by various diseases and conditions.(3)

Outbreak: This marks the beginning of an epidemiological concern and is characterised by an unexpected increase in the number of cases of a disease or the emergence of cases in a new geographical area.(4) It serves as an early warning for potential epidemic or pandemic situations, depending on the disease's subsequent spread and control measures implemented.

Epidemic: Defined by the Centres for Disease Control and Prevention (CDC) as a sudden increase in the number of disease cases within a specific geographical area, an epidemic exceeds the normal expectancy of disease occurrence.(5) Not confined to contagious diseases alone, epidemics can also include non-infectious diseases or health-related behaviours, such as obesity rates, that surge beyond typical levels in a community or region.(4) Diseases like yellow fever, smallpox, measles, and polio are historic examples of epidemics.

Endemic: When a disease consistently exists within a particular region or population but remains relatively stable in its spread and impact, it is considered endemic.(4) These diseases are predictable and manageable to a certain extent due to their localised nature. Malaria in certain tropical regions is a prime example of an endemic disease.(4)

Pandemic: A pandemic occurs when a disease's growth becomes exponential across several countries and continents, surpassing the epidemic stage.(4) The World Health Organisation (WHO) declaration of a pandemic focuses on the geographic spread and exponential growth rate of cases, rather than the disease's severity or the population's immunity levels.(6) Pandemics lead to significant social, economic, and public health disruptions due to their wide-reaching effects.(7)

1.2 Cause of outbreaks of infectious diseases:

The emergence and outbreak of infectious diseases are influenced by a myriad of factors involving interactions between humans, animals, and the environment.(8) These factors create a complex web that can facilitate the spread of diseases under certain conditions. Key contributors to the outbreak of infectious diseases include:

Weather Conditions: The prevalence of certain infectious diseases is closely linked to specific weather patterns.(9) For instance, whooping cough shows increased incidence in the spring, while measles outbreaks are more common during the winter months.(10, 11) This pattern can be attributed to the direct effects of weather on the survival and transmission of pathogens, as well as indirect effects on human behaviour and immunity.

Exposure to Chemicals or Radioactive Materials: Diseases can also result from exposure to toxic substances.(12) A notable example is Minamata disease, caused by consuming seafood contaminated with mercury.(13) Such environmental exposures highlight the intersection between industrial activity, environmental health, and infectious diseases.

Social Aftermath of Disasters: Natural disasters such as storms, earthquakes, and droughts can disrupt societies, leading to conditions that favour the spread of infectious diseases.(14) The displacement of populations, overcrowding in shelters, and the breakdown of sanitation and healthcare systems create fertile grounds for disease transmission.(15)

Environmental Factors: The quality of water supply, food, air, and sanitation facilities plays a crucial role in the spread of infectious diseases.(16) Poor sanitation and contaminated water sources are well-known catalysts for outbreaks of diseases such as cholera and typhoid fever.(17) Similarly, air quality and food safety are vital in preventing respiratory and foodborne illnesses, respectively.(18, 19)

Disease Origins of Unknown Causes: Some infectious diseases emerge without a clear understanding of their cause.(16) These may result from new or modified pathogens, natural toxins, undetected chemical releases, or unknown exposure to ionising radiation. For example, the Nodding disease identified in Southern Sudan, where it is a progressive neurological disease with seizures and causes growth retardation where the aetiology remains a mystery.(20) The emergence of such diseases poses significant challenges to public health systems, requiring vigilant surveillance, research, and response strategies to identify and mitigate their spread.(21)

Understanding the multifactorial nature of infectious disease outbreaks is important for developing effective prevention and control measures.(22) It necessitates a multidisciplinary approach involving epidemiology, environmental science, public health policy, and community engagement to address the diverse factors contributing to disease spread.(22)

1.3 Historical Overview of Pandemics and their impacts

Throughout history, humanity has been periodically ravaged by pandemics, each leaving its indelible mark on society, economy, and long-term health consequences.(23) Unlike natural disasters such as tsunamis, earthquakes, and floods, which are typically localised and have immediate impacts, pandemics have the unique ability to cross borders, affecting global populations and systems over extended periods. The WHO categorises pandemics as disasters that not only cause immediate health crises but also disrupt societies by causing extensive material, economic, and environmental losses, often surpassing local capacities to respond effectively.(24)

A key aspect that distinguishes pandemics from other disasters is their dual impact: direct effects stemming from the infectious agent itself and indirect consequences on the social, economic, and environmental fabric of society demonstrated by the West African Ebola, Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2 virus) or Middle East respiratory syndrome (MERS) pandemic.(25) Moreover, critical periods of exposure, such as during childhood, can have profound implications for later-life disease risk, suggesting that pandemics can initiate a cascade of risk or accumulate risks across an individual's life-course.(25)

I shall now provide a brief overview of the major pandemics that have affected humanity over history, causative agents and their estimated death tolls on humanity (Table 1

adapted from Piret and Boivin et al) (26, 27)and describe their relevant long-term consequences.

The Antonine Plague

The Antonine Plague weakened the Roman Empire significantly, undermining its military and economic strength and paving the way for new religious movements, including the spread of Christianity.(28)

The Justinian Plague

The Justinian Plague is often considered the first recorded pandemic. It led to a drastic population decline in the Byzantine Empire, eroding its physical, economic, and cultural foundations, which facilitated the rapid expansion of Islam.(28)

The Black Death

The Black Death caused a global population to decrease (from 450 million to below 350 million).(29) It struck Europe in the mid-14th century, claiming up to 60% of the European population.(29) The cultural impact of this pandemic was profound, influencing art, religion, and societal structures and leading to significant economic and social changes. (30) The emergence of Danse Macabre, or the Dance of Death, in visual arts and religious scripts, reflects the societal preoccupation with mortality and the transient nature of life during this period.(30)

Quarantine

The concept of quarantine, a critical public health measure still in use today, traces its origins back to this era. In 1377, the chief physician of Ragusa, the city-state now known as Dubrovnik, established a facility outside the city limits to isolate patients based on the theory of contagion.(31) Initially, affected individuals were isolated for 30 days (trentine), a period that was later extended to 40 days (quarantine), giving birth to the term we use today.(31) This measure proved to be one of the few effective strategies for controlling the spread of the plague, and the practice of quarantine quickly spread throughout Europe. (31)

Table 1 Brief overview of major pandemics that have affected humanity over history

This table presents an overview of the pandemics, including their time periods, associated pathogens, transmission vectors, and estimated death tolls.

Period	Pandemics	Pathogens	Vectors	Estimated Death Toll
430 BC	Plague of Athens	?Typhoid		750K – 110K
165-180 AD	The Antonine Plague	Unknown	Believed to be either small pox or measles	5 million
541-543	Plague of Justinian	Yersinia Pestis	Fleas (Wild Rodents)	30 – 50 million
1347-1351	Black Death	Yersinia Pestis	Fleas (Wild Rodents)	200 million
1817-1824	First Cholera pandemic	Vibrio cholerae	Contaminated Water	1 million
1827-1835	Second cholera pandemic	Vibrio cholerae	Contaminated Water	
1839-1856	Third Cholera pandemic	Vibrio cholerae	Contaminated Water	
1863 – 1875	Fourth Cholera pandemic	Vibrio cholerae	Contaminated Water	
1881 – 1886	Fifth cholera pandemic	Vibrio cholerae	Contaminated Water	
1885-ongoing	Third Plague	Yersinia Pestis	Fleas (Wild Rodents)	12 million
1889 – 1893	Russian Flu	Influenza A/H3N8?	?Avian	1 million
1899 – 1923	Sixth cholera pandemic	Vibrio cholerae	Contaminated Water	800K
1918 – 1919	Spanish Flu	Influenza A/H1N1	Avian	40 – 50 million
1957-1959	Asian Flu	Influenza A/H1N2	Avian	1.1 million
1961-ongoing	Seventh Cholera Pandemic	Vibrio cholerae	Contaminated Water	570K
1968-1970	Hong Kong Flu	Influenza A/H3/N2	Avian	1 million
21st Century				
2002-2003	Severe acute respiratory syndrome (SARS-CoV virus)	SARS-CoV	Bats, palm civets	770K
2009-2010	Swine Flu	Influenza A/H1N1	Pigs	200K

2013-2016	Ebola Pandemic	Ebola viruses	Bats	11.3K
2015 - ongoing	Zika Virus Epidemic	Zika virus	Aedes mosquitoes (A. aegypti and A albopictus)	
2015 – ongoing	Middle East respiratory syndrome (MERS)	MERS-CoV	Bats, dromedary camels	850K
2019-ongoing	COVID-19	SARS-CoV-2	Bats, pangolins	18.2 to 33.5 million

The Spanish Flu

The Spanish Flu of 1918-1919 stands as one of the most devastating pandemics in human history, with a mortality rate estimated between 10% and 20%.(32) Given that over a quarter of the world's population contracted the flu at some point, the death toll reached an astonishing figure, possibly up to 100 million people, surpassing the fatalities caused by the Black Death across a century.(32) Unlike typical flu outbreaks that severely impact the very young, the elderly, or those with preexisting health conditions, the Spanish Flu notably targeted young, healthy adults, a phenomenon attributed to the triggering of a cytokine storm, an overreaction of the body's immune system.(32)

The pandemic's impact extended beyond the immediate health crisis, influencing global events and societal structures, including potentially affecting the outcome of World War I.(33) It marked the first instance where the prolonged consequences of a pandemic on subsequent generations could be quantitatively assessed. (33)Studies leveraging United States census data from 1960 to 1980 revealed that individuals born to mothers who were pregnant during the pandemic faced more health issues, earned lower incomes, and achieved lower educational and socioeconomic statuses compared to those born just before or after the pandemic.(34) A 2006 study published in the Journal of Political Economy highlighted the significant long-term impacts on the cohorts exposed in utero, including reduced educational attainment, increased physical disabilities, and a reliance on social welfare programs.(35)

Despite its severe impact, the Spanish Flu receded as swiftly as it had emerged, concluding its devastating run within a mere nine months. Its rapid disappearance from public and scientific discourse led to its characterisation as the "forgotten pandemic" by some historians, like Alfred W. Crosby, emphasising the transient collective memory of

such global health crises.(36) This historical episode underscores the importance of remembering and learning from past pandemics to better prepare for future outbreaks, acknowledging the profound and lasting effects they can have on society.(36)

HIV/AIDS pandemic

The Human Immunodeficiency Virus/Acquired Immunodeficiency Syndrome (HIV/AIDS) pandemic represents a unique and prolonged crisis in global health, distinguishing itself by its slow progression and broad reach across continents and populations, each phase introducing new challenges.(37) Unlike the swift onslaught seen in historical pandemics, HIV/AIDS has unfolded over decades, allowing for an extensive and focused public health response from both national governments and international bodies.(38) This sustained attention has facilitated significant advancements in treatment and understanding of the disease, including its psychological impacts, making HIV/AIDS one of the few infectious diseases that have been closely examined through the lens of mental health.(39)

HIV/AIDS and chronic mental health

The intersection of HIV/AIDS with mental health reveals profound insights into the broader challenges posed by infectious diseases.(39) Notably, the lifetime prevalence of depression among individuals living with HIV is approximately 22%, more than double that of the general population.(39) This elevated risk underscores the complex interplay between physical health and mental well-being, where depression is not only a comorbid condition but also a factor that can influence the effectiveness of treatment for HIV/AIDS through its association with substance abuse and adherence to antiretroviral therapy.(40)

The stigma, guilt, and shame associated with HIV/AIDS further compound the challenges faced by those living with the disease.(41) These psychosocial factors can deter individuals from seeking diagnosis and treatment, exacerbating the public health challenge of controlling the pandemic.(41) Recognising these issues, there has been a concerted effort to address the mental health needs of HIV-positive individuals, with a wealth of research focusing on the medical treatment of depression and the development of psychotherapeutic interventions tailored to this population.(41)

The comprehensive approach to HIV/AIDS, encompassing both the medical and psychological aspects of the disease, exemplifies a holistic model of infectious disease management.(42) It highlights the necessity of integrating mental health care into the response to pandemics, acknowledging the intricate connection between psychological well-being and the success of public health interventions.(43) Through the study of HIV/AIDS and its impact on mental health, valuable lessons can be learned about

addressing the multifaceted challenges posed by infectious diseases, emphasising the importance of comprehensive care strategies that include mental health support for affected individuals.(43)

Severe Acute Respiratory Syndrome (SARS) and Middle East Respiratory Syndrome Coronavirus (MERS-CoV)

The outbreak of Severe Acute Respiratory Syndrome (SARS) in the early 21st century highlighted the mental health impacts of infectious diseases on patients, families, and healthcare workers, underscoring the need to integrate mental health considerations into public health responses.(44) As the first major outbreak of the century to draw global attention, SARS illuminated the psychological burdens of isolation, such as loneliness, anxiety, and distress, and revealed that survivors often faced long-term effects, including post-traumatic stress disorder, depression, and anxiety. Studies across various cultures deepened understanding of these impacts, stressing the importance of mental health support alongside physical health interventions.(45, 46)

Healthcare providers working on the frontlines during the SARS outbreak also faced significant mental health challenges.(45) The high-stress environment, coupled with the fear of contracting the virus and spreading it to loved ones, led to increased reports of burnout, anxiety, and depression among medical staff.(45) These findings underscored the need for mental health support systems and interventions tailored to healthcare workers exposed to high-risk environments during infectious disease outbreaks.(45)

In addition to the mental health aspects, respiratory sequelae emerged as a common physical aftermath of SARS, as well as other respiratory viruses such as influenza and MERS-CoV. A systematic review reported respiratory sequelae as a prevalent finding, with 80% of the literature on influenza, 67% on Severe acute respiratory syndrome coronavirus 1 (SARS-CoV-1), and 50% on MERS-CoV highlighting this issue.(46) This indicates the lasting impact of respiratory viruses on physical health, further complicating the recovery process for affected individuals and necessitating comprehensive care that addresses both physical and mental health needs.

The SARS outbreak served as a critical learning opportunity, demonstrating the importance of integrating mental health considerations into public health responses to infectious diseases. It highlighted the need for preparedness plans that include mental health support for patients, survivors, and healthcare providers, ensuring a holistic approach to managing health crises.

2009 H1N1 Pandemic

The 2009 H1N1 pandemic, often reminiscent of the 1918 Spanish Flu in its causative agent, presented a stark contrast in its outcomes, infecting over 10% of the global population but resulting in a considerably lower mortality rate, with estimates ranging from 20,000 to over 500,000 deaths.(47) Unlike the catastrophic impact of its predecessor, the 2009 pandemic's mortality rate was lower than that typically observed for seasonal influenza. However, the threat perception was heightened due to its unusual impact on healthy young adults, leading to severe respiratory issues more rapidly than expected.(48) This phenomenon, partly attributed to a cytokine storm similar to that of the 1918 pandemic, was also thought to be mitigated in older adults who might have had residual immunity from a similar H1N1 strain in the 1970s.(49)

The 2009 H1N1 pandemic notably advanced the understanding and inclusion of mental health in pandemic preparedness and response. For one of the first times, policy reports recognised mental health as a critical element of public health strategies during pandemics, highlighting the importance of addressing the psychological impacts of pandemics alongside physical health challenges.(50)

However, the pandemic also highlighted the complexities of managing public perception and response to health crises.(51) There was a noticeable dissonance between public sentiments, which initially veered towards panic due to the WHO’s and national health institutions' warnings, and the eventual realisation that the pandemic's impact was less severe than feared.(52, 53) This discrepancy led to widespread discontent and mistrust, with accusations of overreaction ("panicdemic") and skepticism towards the motivations behind the rapid production and distribution of vaccines, suggesting an undue influence of pharmaceutical companies.(53)

The experience of the 2009 H1N1 pandemic underscored the challenges inherent in communicating the risks and recommended precautions of a pandemic to the public.(54) It demonstrated the pitfalls of simplifying complex public health threats into terms like "mild," "moderate," or "severe," which may not adequately convey the nuances of the threat and can lead to confusion or mistrust among the public.(54) This pandemic served as a crucial learning opportunity, highlighting the need for clear, transparent, and nuanced communication strategies in public health efforts and the importance of including mental health considerations in the planning and response to global health emergencies.(54, 55)

F

Zika

The Zika virus outbreak serves as a prime example of the complexities and challenges posed by global disease transmission in the modern era.(56) Initially emerging from Micronesia, the virus swiftly traversed across the Pacific to Brazil, from where it spread further, demonstrating the rapid pace at which infectious diseases can disseminate globally in today's interconnected world.(56) However, what sets the Zika outbreak apart is its distinction as a "media pandemic," with social media platforms playing a pivotal role in shaping public awareness and response to the crisis.(57, 58)

In early 2016, at the height of the outbreak, mentions of Zika surged to an unprecedented 50 times a minute on Twitter, illustrating the intense public and media interest in the virus. (59, 60)Social media served as a useful tool for disseminating information, educating the public, and expressing concerns. For the first time, the widespread use of social media during a public health crisis allowed researchers to study public sentiment and emotional epidemiology in real time. (59) This analysis provided insights into how people perceived and reacted to the outbreak, blending traditional epidemiological studies with an understanding of the social and emotional dimensions of public health emergencies.(59, 60)

While both public health institutions and the general populace utilised social media to voice their concerns about Zika, a notable divide emerged in the content being shared.(61) Scientists and health officials primarily focused on the educational aspect, striving to inform and guide the public with accurate information.(62) In contrast, the general public's posts often reflected a desire to have their emotional and personal concerns acknowledged and addressed.(61) This dichotomy highlights the dual role of social media as both a source of reliable information and a platform for expressing personal anxieties and fears.

Interestingly, despite the predominance of accurate information—with four out of five posts on Zika being factually correct—the posts that gained significant traction and "trended" were often those containing inaccurate content or "fake news."(63-65) This phenomenon underscores a critical challenge for public health communication in the digital age: the propagation of misinformation can easily outpace and overshadow fact-based discourse, potentially undermining public health efforts.(64)

The experience of the Zika outbreak underscores the importance of addressing the phenomenon of misinformation and the role of social media in shaping public perceptions during health crises.(66) As the world prepares for future outbreaks, understanding and

leveraging social media's influence will be crucial not only for effective preparedness but also for the successful execution of public health plans, including measures like quarantine and immunisation.(67) Navigating the intricacies of digital communication and public sentiment will be essential for public health officials to ensure that accurate, science-based information prevails over misinformation, thereby safeguarding public health and well-being.(67)

Each of these pandemics has contributed to our understanding of the multifaceted impact of infectious diseases, not only in terms of immediate health effects but also in their ability to alter the course of history, influence public health policy, and shape societal norms and behaviours. The legacy of past pandemics underscores the importance of preparedness, effective communication, and comprehensive public health strategies to mitigate the impacts of future outbreaks.

1.4 Cardiovascular impact of pandemics

Pandemics can cause long-term health consequences, including a variety of cardiovascular complications.(68) Historical and recent pandemics offer valuable insights into these effects, with varying degrees of documentation and understanding. Here are examples of cardiovascular consequences from pandemics other than COVID-19 which will be described separately:

1.4.1 1918 Influenza Pandemic (Spanish Flu)

Research has shown that survivors of the 1918 influenza pandemic experienced higher rates of cardiovascular disease in the following decades, attributed to both the virus's direct damage to cardiovascular tissues and the strain placed on the cardiovascular system.(69, 70) Studies suggest that individuals exposed to the Spanish flu early in life had a heightened risk of heart diseases later, including heart failure and myocardial infarction.(71, 72)Two key studies offer insight into this lasting cardiovascular effect, particularly for those who have been exposed in utero. Mazumder et al and colleagues conducted a longitudinal analysis comparing infants born during the height of the pandemic to those born shortly after, revealing an increased prevalence of diabetes and heart disease in adulthood, particularly notable in the second quarter of 1919 with a 36.7% excess in diabetes cases.(34) Conversely, Myrskyla et al and colleagues broadened the comparative scope to include infants born between 1917 and 1919 against a control group from 1920 to 1924, focusing on cardiovascular mortality.(71) Their findings did not show a significant uptick in cardiovascular mortality hazard ratios, indicating a nuanced picture of the pandemic's long-term cardiovascular impacts.(71)

1.4.2 H1N1 Influenza Pandemic (2009)

During the H1N1 influenza pandemic, there was an observed increase in hospital admissions for acute myocardial infarction during the H1N1 pandemic.(73) The inflammation triggered by the viral infection is believed to have played a role in plaque destabilisation and subsequent heart attacks.(74) Myocarditis which causes inflammation of the heart muscle that can affect the heart's electrical system and reduce the heart's pumping function were reported following H1N1 infection.(73) There have been multiple larger studies have associated influenza with a temporarily elevated risk of acute heart failure. Acute heart failure may be the most common cardiovascular influenza-associated complication.(74)

1.4.3 HIV/AIDS Pandemic

Individuals with HIV/AIDS are at an increased risk of early atherosclerosis, partly due to the chronic inflammatory state induced by the virus and potential effects of antiretroviral therapy.(75) Due to the increased risk of cardiovascular diseases, HIV-infected individuals have a higher risk of developing cardiovascular diseases including myocardial infarction, stroke, peripheral arterial disease when compared to the general population. (75)

1.4.4 H5N1 Avian Influenza

Although, H5N1 is primarily a respiratory disease, H5N1 has been associated with direct viral invasion of myocardial tissue, leading to myocarditis and other cardiac complications in severe cases.(76)

1.4.5 Zika Virus Epidemic

While the primary concern with Zika virus infection has been its association with microcephaly and other congenital anomalies in baby born to infected mothers, there have also been reports of congenital heart defects among these infants.(56)

1.4.6 SARS-CoV-1 and MERS-CoV

The cardiovascular sequelae of SARS-CoV-1 infection remain underexplored, with the literature largely comprising anecdotal evidence and isolated reports of acute coronary syndrome, myocardial infarction, transient diastolic dysfunction, and other cardiovascular abnormalities.(77) A notable postmortem study highlighted thromboembolic disease in patients. Similarly, detailed information on cardiovascular involvement in MERS is limited, with existing knowledge primarily derived from case reports or investigations into the prevalence of comorbidities among affected individuals.(77)

These examples highlight the importance of monitoring and managing cardiovascular health in the aftermath of pandemics, given the potential for long-lasting health impacts beyond the immediate effects of the viral infections themselves.(77)

1.4.7 Putative Pathways in Pandemic Viruses and Cardiovascular Effects

The mechanisms through which pandemic viruses exert their cardiovascular effects are complex and multifaceted. For SARS-CoV-2 virus, as well as its predecessors, the interaction with the host's immune system, direct viral effects on cardiovascular tissues, and the exacerbation of pre-existing conditions are all contributing factors.(78) Table 2 outlines the putative pathways involved in the cardiovascular impact of pandemic viruses adapted from Savedchuk et al.(78)

1.5 COVID-19: A Modern Pandemic

Emergence and Global Impact

The COVID-19 pandemic, caused by the SARS-CoV-2 virus which was first identified in December 2019 in Wuhan, Hubei Province, China. Reports of a cluster of cases of pneumonia of unknown cause prompted immediate investigation, leading to the identification of a novel coronavirus in early January 2020.(79) The WHO declared the outbreak a Public Health Emergency of International Concern on January 30, 2020, and a pandemic on March 11, 2020, as cases surged globally.(80)

Epidemiology of COVID-19

COVID-19 spread rapidly across the globe, affecting millions of people in every continent. As of late 2021, there have been over 250 million confirmed cases and more than 5 million deaths globally, marking it as one of the deadliest pandemics in history.(6) By 2024 there have been 703 million affected with nearly 7 million deaths globally.(81) The disease manifests with a wide range of symptoms, from asymptomatic or mild to severe respiratory distress and death, particularly in older adults and those with underlying health conditions.(82,83)

Virology of COVID-19

SARS-CoV-2, the virus responsible for the COVID-19 pandemic, belongs to the coronavirus family, a group of viruses known for their crown-like appearance under an electron microscope. (83) This appearance is due to the spike (S) proteins on the virus's surface. (83) The virus is a positive-sense, single-stranded ribonucleic acid (RNA) virus,

with its genome encoding both structural and non-structural proteins critical for its replication and pathogenesis.(84)

Table 2 Putative pathways involved in the cardiovascular impact of pandemic viruses

This table outlines pandemic viruses linked to cardiovascular impacts, detailing their host receptor(s), cellular targets, species tropism, and proposed mechanisms of hypertension pathogenesis.

Virus	Host Receptor(s)	Cell Targets	Species Tropism	Proposed mechanism of hypertension pathogenesis
Influenza	Sialic Acid	Tracheal Epithelium	Human	Inflammatory response leading to systemic complications. - Direct viral effects on cardiovascular tissues. - Increase in vascular intimal cellularity and endothelial infiltrates - Plaque destabilisation - Hypoxaemia - Hypercoagulability - Increased myocardial oxygen demand
			Birds	Ang II upregulation, ACE2 downregulation
MERS-CoV	Dipeptidyl-peptidase 4 (DPP4)	Non-ciliated epithelial cells	Human	Similar to SARS-CoV-1, with a focus on pre-existing comorbidities enhancing susceptibility to cardiovascular complications.
		Bronchiolar epithelial cells alveolar epithelial cells	Camel	
SARS-CoV-1 SARS-CoV-2	Angiotensin-Converting Enzyme 2 (ACE2)	Bronchial epithelia	Human	- Direct interaction with ACE2 receptors, leading to endothelial dysfunction. - Cytokine storm contributing to acute cardiovascular events. - Long-term sequelae including hypertension and myocarditis.
	TMPRSS2			
	CD147	Nasal epithelia	Bats	Downregulation of ACE 2 after viral infection leading to activation of the RAAS and increased Ang II
		Type II Alveolar Cells		

Genome Organisation and Structure

The genome of SARS-CoV-2 is approximately 30,000 nucleotides long, making it one of the largest among RNA viruses.(83) The genome organisation is characteristic of coronaviruses, with two-thirds of the genome at the 5’ end encoding for two large

polyproteins, pp1a and pp1ab, which are cleaved into 16 non-structural proteins (nsps).(83) These nsps form the replication-transcription complex essential for viral RNA synthesis.(83) The remaining one-third of the genome encodes for four main structural proteins as described by Hu et al and colleagues :

- Spike (S) Protein: The spike protein facilitates viral entry into host cells by binding to the angiotensin-converting enzyme 2 (ACE2) receptor. The S protein is a primary target for neutralising antibodies and vaccines.
- Envelope (E) Protein: Involved in viral assembly and release.
- Membrane (M) Protein: The most abundant structural protein, it gives shape to the viral envelope.
- Nucleocapsid (N) Protein: Binds to the viral RNA, forming the nucleocapsid.

Life Cycle of the SARS-CoV-2 virus

The SARS-CoV-2 virus is a zoonotic coronavirus characterised by a complex life cycle that facilitates its entry and replication within host cells.(85) The life cycle of SARS-CoV-2 virus can be delineated into five main steps, pertinent for understanding the virus's pathogenicity and implications for treatment and prevention strategies which is demonstrated below in Table 3.(85) The virus specifically targets cells expressing the ACE2 receptor, such as type 2 pneumocytes, macrophages, perivascular pericytes, cardiomyocytes, and endothelial cells. (86) This indicates the varied routes of infection the SARS-CoV-2 virus can exploit within the human body.(85)

1.5.1 Pathophysiology of the SARS-CoV-2 virus and implications on the cardiovascular system

The interaction between the SARS-CoV-2 virus and the cardiovascular system has highlighted several mechanisms through which COVID-19 can exacerbate existing cardiovascular conditions or potentially give rise to new ones, such as hypertension.(78) This is demonstrated as a broad overview in Figure 2.(87, 88)

Transmission

The SARS-CoV-2 virus is predominantly transmitted through respiratory droplets, with the virus entering the body through the nasal or oral mucosa and then binding to the ACE2 receptors found on the surface of cells, particularly in the lungs, heart, kidneys and blood vessels.(89) (Figure 3)

Table 3 Life cycle of the SARS-CoV-2 virus

This table outlines the life cycle of the SARS-CoV-2 virus.

1. Viral Entry	The initial step involves the binding of the viral spike (S) protein to the ACE2 receptor on the host cell surface. The S protein's interaction with ACE2 is facilitated by the host cell's serine protease TMPRSS2, which primes the S protein for membrane fusion
2. Release of RNA	Following membrane fusion, the viral RNA is released into the host cell cytoplasm.
3. Translation and Replication	The viral RNA serves as a template for the translation of viral replicase polyproteins and the synthesis of negative-sense RNA intermediates, which are then used to produce genomic and sub-genomic RNAs.
4. Assembly	Viral structural proteins and genomic RNA assemble at the endoplasmic reticulum and Golgi apparatus to form new viral particles.
5. Release	Newly formed viral particles are transported in vesicles and released from the host cell via exocytosis.
6. Endocytosis and ACE2 shedding	Following the virus's entry into the host cell via endocytosis, there is an increase in ADAM metallopeptidase domain 17 (ADAM17) activity. This leads to the shedding of ACE2's ectodomain from the cell surface, disrupting the physiological balance between ACE and ACE2 activities.

Figure 2 ACE2 Expression

This figure depicts the expression of ACE2 receptors across various organ systems in the body.

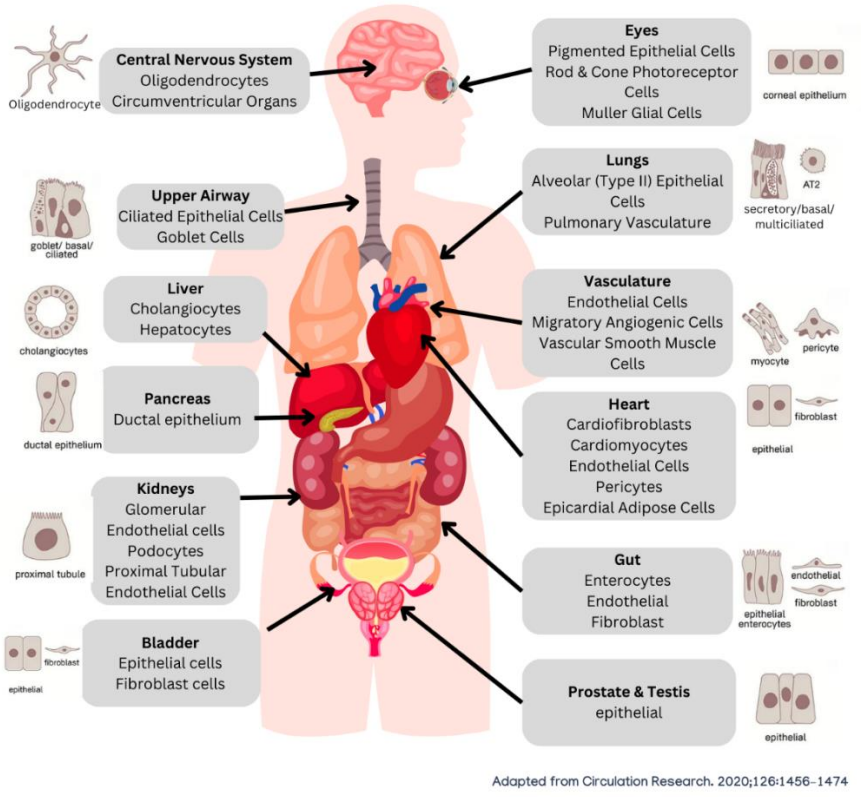
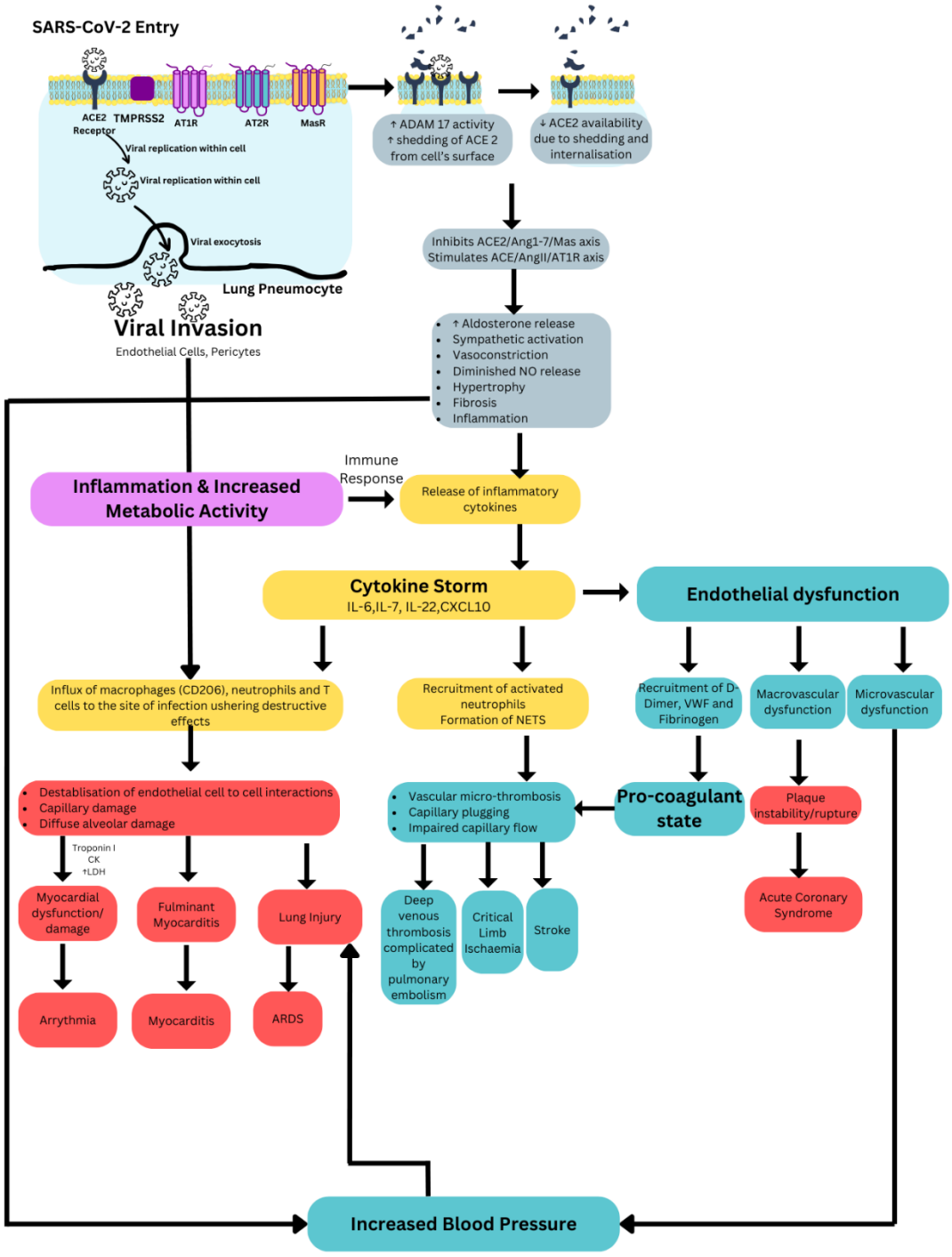


Figure 3 Interactions between SARS-CoV-2 virus and the Cardiovascular System

This figure illustrates the interactions and pathogenesis of the SARS-CoV-2 virus, highlighting its effects on the cardiovascular system and its impact on blood pressure.



Adapted from Cardiovasc Res. 2020 Aug 1;116(10):1666–1687
Front Physiol. 2021 May 3:12:665064

ACE2 and interaction of SARS-CoV-2 virus with ACE2

ACE2 plays a critical role in the renin-angiotensin-aldosterone system (RAAS) and is widely expressed across major organ systems.(90) Its significance is rooted in the historical discovery of key RAAS components, starting with renin and leading to the identification of angiotensin (hypertensin), Ang I, Ang II, the counter-regulatory axis of RAAS, ACE2, and the Ang-(1-7)/Mas receptor axis, as outlined in Figure 4.(90) In addition to its cardioprotective effects, ACE2 has been identified as the receptor for the SARS-CoV-2 virus, a dual role illustrated in Figure 5.(90)

Under normal conditions, ACE2 converts Ang II—a potent vasoconstrictor and pro-inflammatory molecule—into angiotensin-(1-7), which exerts vasodilatory and anti-inflammatory effects.(91, 92) However, by binding to ACE2, the SARS-CoV-2 virus reduces its surface expression, thereby diminishing these protective effects.(78)

Elevated plasma levels of ACE2, which are associated with cardiovascular disease and mortality, suggest that COVID-19 triggers RAAS modulation.(93) This modulation likely increases blood pressure (BP) through mechanisms linked to the correlation between Ang II levels and the severity of COVID-19.(94, 95)

Shift in Balance in ACE2 activity and RAAS Dysregulation

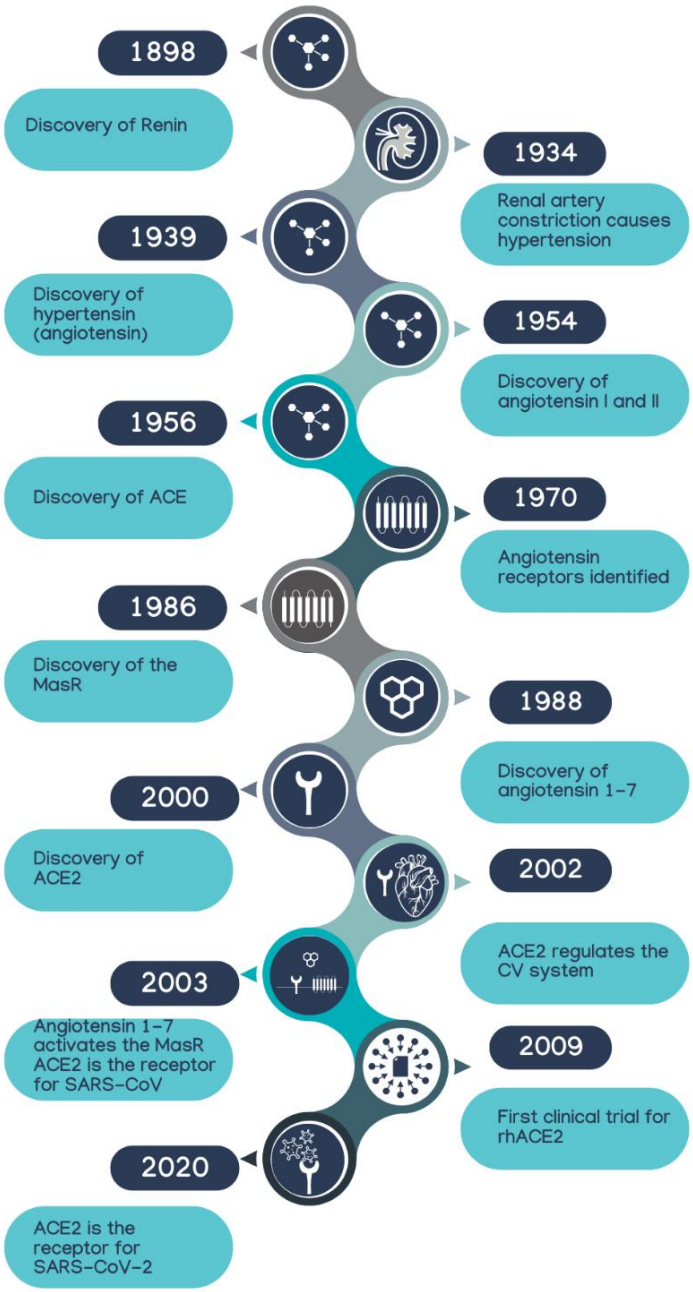
The interaction of the virus with ACE2 receptors leads to a shift in balance of ACE 2 activity and RAAS dysregulation, which is crucial in regulating BP and vascular function. (Figure 5) (78) This observed shift in ACE2 activity is decreased leading to an accumulation of Ang II. The shedding and internalisation of ACE2 enhance Angiotensin II (Ang II) activity due to the reduced availability of ACE2 to convert Ang II into Angiotensin 1-7 (Ang 1-7). This results in a pivotal shift from the protective ACE2/Ang 1-7/Mas axis to the deleterious ACE/Ang II/AT1R axis. This results in increased Ang II effects, including pulmonary vasoconstriction, sodium retention, aldosterone secretion, increased BP, pulmonary oedema, and potentially leading to acute respiratory distress syndrome and death.(78, 91, 96) Additionally, elevated Ang II levels have proinflammatory effects and fibrosis which is further explained below.

Direct Viral Invasion and Pathogenesis

The SARS-CoV-2 virus employs multiple strategies to evade the host immune response, including interfering with interferon production and signalling, which delays the immune response and allows for higher viral replication and spread. The immune evasion mechanisms, combined with the host's immune response, contribute to the pathogenesis

Figure 4 Timeline of discovery of the major renin-angiotensin system (RAAS) components and ACE 2

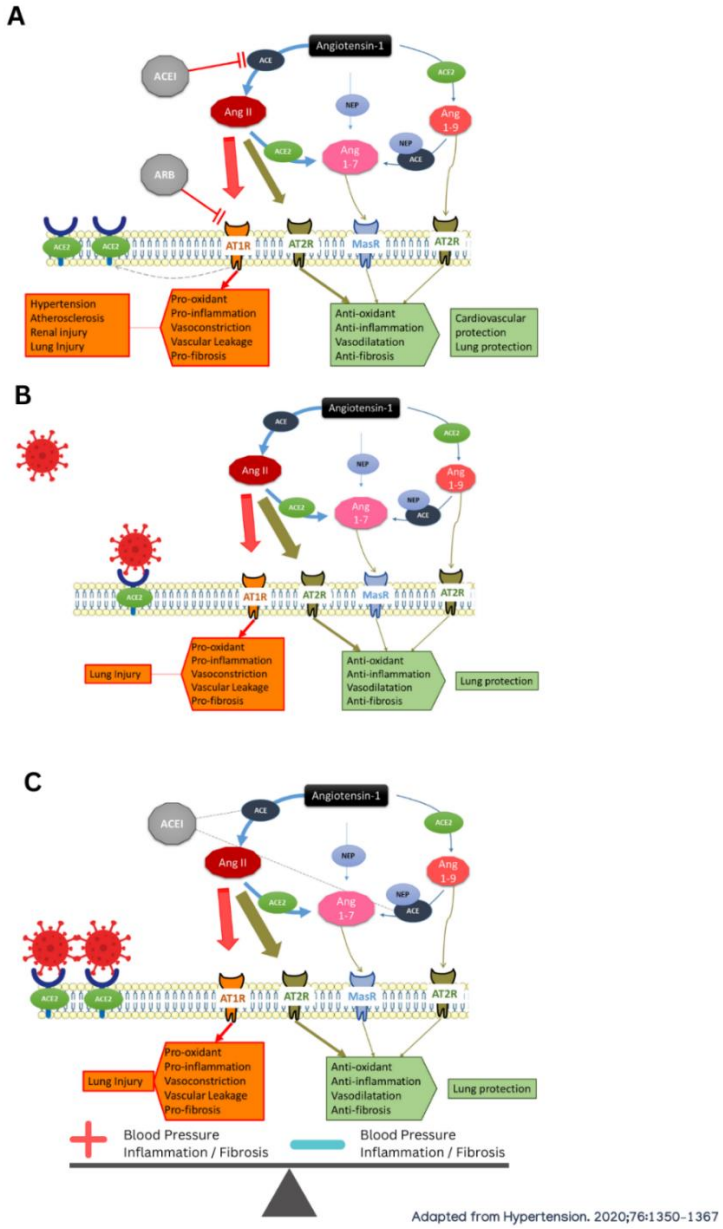
This figure presents a historical timeline showcasing the discovery of key components of the renin-angiotensin-aldosterone system (RAAS), including ACE2, along with significant advancements in understanding their physiological roles.



Adapted from Circulation Research. 2020;126:1456–1474
Nat Rev Immunol. 2020 Jul;20(7):389–391.

Figure 5 The classical renin-angiotensin-aldosterone system (RAAS) enzymatic cascade

(A): In a normal state the RAAS is in equilibrium and blood pressure is normal. The mechanism of action of common antihypertensives ACEIs and ARBs are shown. (B): The SARS-CoV-2 virus causes a change following binding to ACE2 receptor which causes a cascade of changes where there is a shift towards the ACE (angiotensin-converting enzyme)/Ang (angiotensin) II pathway away from the ACE2(angiotensin-converting enzyme 2)/Ang-(1-7) pathway. AT1R indicates type 1 Ang II receptor. This triggers a cascade of pro-oxidant, pro-inflammation, vasoconstriction, vascular leakage and profibrotic cascade leading to lung injury. (C): The effect of SARS-CoV-2 virus pathway is seen in conjunction with the use of ACEI/ARBs and the effects on the RAAS indicative there is a protective/harmful effect leading to debates. There is an observed upregulation or overactivation of the vasoactive part of RAAS contributing to the pathogenesis and underlying mechanisms of cardiovascular disease, hypertension and kidney diseases with Ang II stimulating the AT1 receptor being the key end player.



and clinical manifestations of COVID-19. The SARS-CoV-2 virus can directly infect endothelial cells, leading to inflammation, cell injury, and apoptosis.

Pro-inflammatory Effects and Fibrosis

The dysregulation of RAAS with elevated Ang II levels, can stimulate the production of pro-inflammatory cytokines which contributes to the systematic inflammtory observed in those with severe COVID-19.(78, 97) This inflammation can further damage cardiovascular tissues and exacerbate heart and kidney diseases.(78) This hypothesis is supported by observations of inflammatory infiltrates in vessel walls and endothelial apoptosis in COVID-19 patients.(98) The persistent inflammation accompanied, with elevated Ang II levels, can promote fibrotic changes in the heart and lungs, contributing to long-term complications such as heart failure and pulmonary fibrosis.(99)

Cytokine Storm

A hyperactive immune response to SARS-CoV-2 infection triggers a cytokine storm, characterised by excessive release of pro-inflammatory cytokines such as IL-6, TNF-α, and MCP-1, as well as chemokines.(100) This overwhelming inflammatory response induces widespread endothelial dysfunction, contributing to vascular permeability, inflammation, and a procoagulant state.(100) The elevated cytokines not only exacerbate systemic inflammation but also promote the generation of reactive oxygen species (ROS).(100)

This cytokine storm, in conjunction with the virus-induced downregulation of ACE2, has significant implications for BP regulation. By disrupting the balance between vasodilatory and vasoconstrictive factors, it promotes vascular inflammation, increases vascular resistance, and leads to hypertension.(99, 101) Together, these mechanisms further aggravate endothelial dysfunction, amplifying the cardiovascular complications associated with COVID-19.

Endothelial Dysfunction

The endothelium plays a pivotal role in vascular health by regulating blood flow, thrombosis, and inflammation.(102) Endothelial dysfunction is a hallmark of COVID-19, characterised by reduced nitric oxide availability, an imbalance between vasoconstrictors and vasodilators, and manifestations such as inflammation, oxidative stress, hyperpermeability, and a prothrombotic state, which significantly contribute to the cardiovascular complications associated with the disease.(102, 103) This dysfunction may arise directly from the interaction between the SARS-CoV-2 virus and ACE2, which

disrupts the renin-angiotensin-aldosterone system (RAAS) and leads to elevated levels of angiotensin II (Ang II) which has been previously described.

The increased Ang II promotes endothelial dysfunction through multiple mechanisms, including the generation of reactive oxygen species (ROS).(102) The overwhelming production of ROS during COVID-19 can deplete the body's antioxidant reserves, including glutathione, superoxide dismutase, and catalase.(104) The reduction in antioxidant capacity exacerbates oxidative stress by leaving cells less equipped to neutralise ROS.(104)

The virus-induced downregulation of ACE2 exacerbates these effects, promoting increased vascular permeability, leukocyte adhesion, and vascular inflammation through heightened expression of pro-inflammatory cytokines and adhesion molecules.(99)

Viral proteins, particularly those involved in the replication process, can induce mitochondrial dysfunction, resulting in increased production of reactive oxygen species (ROS) and impaired antioxidant defences.(105) Additionally, activated immune cells generate ROS as part of their antimicrobial defence mechanisms, further amplifying oxidative stress.(100) The infection induces a significant oxidative stress burden on the endothelium, reducing nitric oxide bioavailability and contributing to vascular dysfunction.(106) This oxidative stress is a key factor in the development of hypertension. (87, 107)

Hypercoagulability and Thrombosis

COVID-19 enhances pro-thrombotic mechanisms through endothelial injury, leading to microthrombi formation and increased risk of major thrombotic events, including deep venous thrombosis and pulmonary embolism.(102) The disruption of endothelial integrity by the infection triggers a cascade of events that promote a prothrombotic state, including the exposure of pro-coagulant surfaces and the release of von Willebrand factor and factor VIII, which are pivotal in clot formation.(102) The virus-induced endothelial dysfunction leads to decreased production of anticoagulant factors like nitric oxide and prostacyclin.(103) At the same time, there is an increased expression of tissue factor, the primary initiator of the coagulation cascade, and a reduction in the activation of protein C, an important natural anticoagulant. These alterations favour thrombin generation and fibrin clot formation.(103) Complement activation leads to the formation of the membrane attack complex, which can cause endothelial cell damage and further promote thrombosis.(108) The virus can directly activate platelets or do so indirectly through the action of inflammatory cytokines.(108) Activated platelets release prothrombotic

microparticles and form aggregates, both of which are potent promoters of coagulation. (108)

Hypoxia

Severe COVID-19 often leads to respiratory distress and hypoxia.(109) Hypoxia itself can induce oxidative stress by altering mitochondrial electron transport chain activity, leading to an overproduction of ROS.(109) Moreover, the hypoxia-inducible factor (HIF) pathway, activated in response to low oxygen levels, can also contribute to oxidative stress through various mechanisms, including the upregulation of enzymes involved in ROS production. (109) COVID-19-related lung injury and ARDS can result in hypoxia, which has been shown to enhance pro-coagulant activity and reduce fibrinolysis, creating a predisposition to clot formation.(110)

Neutrophil Extracellular Traps, Cell Death and Tissue Damage.

The formation of neutrophil extracellular traps (NETs), composed of DNA, histones, and neutrophil-derived proteins, is increased in COVID-19.(110) While NETs play a protective role by trapping pathogens, their formation involves the generation of reactive oxygen species (ROS). Excessive NETs formation can exacerbate tissue damage and further amplify oxidative stress.(110) This process, combined with direct viral effects, inflammation, endothelial dysfunction, and hypoxia, leads to cell death through both apoptotic and necrotic pathways, resulting in significant tissue damage.(98) The breakdown of cells releases additional intracellular ROS and diminishes cellular antioxidant defences, further compounding the oxidative stress burden.(111)

1.5.2 Clinical Manifestations and Disease Classification

The incubation period of SARS-CoV-2 virus, from infection to the onset of symptoms, typically spans 2-14 days, with the disease progression varying significantly among individuals.(Figure 2) (112) Certain populations are at a higher risk of developing severe or critical illness, including older adults and individuals with underlying health conditions such as cardiovascular disease, diabetes, chronic respiratory disease, hypertension, and cancer.(113) COVID-19 can be broadly categorised based on the severity of the disease, as outlined in Table 4, which helps in understanding the diverse clinical outcomes and guiding treatment strategies.(114)

Table 4 COVID-19- disease category.

This table outlines disease category for COVID-19 with associated clinical features, clinical manifestations and proportion of patients globally.

Disease category	Clinical Features	Clinical manifestations	Proportion of patients Globally
Mild disease	Non/mild pneumonia	-	75–85%
Moderate disease	Pneumonia	Gastrointestinal symptoms (diarrhoea, cramps)	
Severe disease	Severe pneumonia, Dyspnoea, Respiratory rate ≥ 30/min, Oxygen saturations ≤ 93%, Peak Flow ratio < 300, Lung infiltrates > 50% (within 24–48 hours)	Neurological symptoms such as headache, altered mental status, Guillain–Barre syndrome (GBS) and Stroke	10–15%
Critical disease	Acute respiratory distress syndrome, Respiratory failure, Sepsis (multiple organ dysfunction / failure) Septic shock	Cardiovascular events such as myocarditis, arrhythmias and heart failure. Ocular manifestations such as conjunctival hyperaemia, chemosis Anosmia and dysgeusia	5–10%

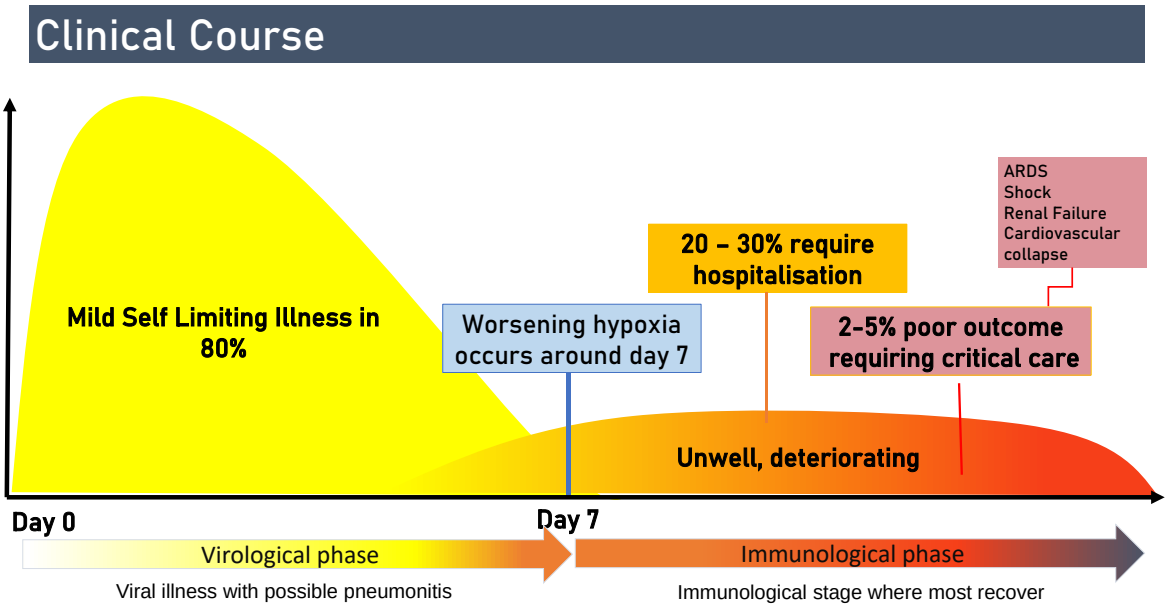
1.5.3 Important COVID-19 Variants

The emergence and evolution of the SARS-CoV-2 virus have been marked by the development of numerous variants, each with mutations in critical structural proteins such as the spike glycoprotein, envelope protein, membrane protein, and various non-structural and accessory proteins.(114) The spike (S) protein, responsible for virus entry into host cells via the ACE2 receptor, has been a focal point of mutation and consequential changes in virus transmissibility, pathogenicity, virulence, or resistance to neutralising antibodies.(114) Significant mutations in the S1 and S2 domains of the spike protein, notably the D614G mutation, have led to the emergence of variants with altered properties, impacting the spread of the virus and the severity of disease they cause.(116) The emergence of the SARS-CoV-2 virus variants poses challenges to existing diagnostic

methods, therapeutic interventions, and vaccine efficacy. (117) Variants with mutations that lead to antigenic drift can result in decreased sensitivity of diagnostic tests, reduced effectiveness of therapeutic antibodies, and diminished vaccine-induced immunity. Consequently, ongoing surveillance of variant evolution and prompt adaptation of diagnostic and therapeutic strategies are critical to managing the pandemic. (118) Below are descriptions of several important COVID-19 variants that have been designated as Variants of Concern (VOCs) or Variants of Interest (VOIs) by health authorities worldwide. (118, 119)

Figure 6 Clinical Course of COVID-19

This figure illustrates the clinical progression of COVID-19, divided into the virological phase (Day 0–7) and the immunological phase (from Day 7 onward). It highlights the acute symptom phase, often requiring hospitalization or critical care. This visualization is based on the author's original work, designed for community assessment of acute COVID-19 and referral to secondary care as outlined in the Acute COVID-19 Scottish Primary Care Hub Triage Guide.(115)



Alpha Variant (B.1.1.7)

First identified in the United Kingdom in September 2020, the Alpha variant quickly became dominant due to its increased transmissibility. (119) It contains a significant mutation (N501Y) in the spike protein, which enhances the virus's ability to bind to human ACE2 receptors.(120) Studies indicated that the Alpha variant was associated with an increased risk of hospitalisation and death, although vaccines have remained effective against it.(121)

Beta Variant (B.1.351)

Detected in South Africa in May 2020, the Beta variant carries mutations that allow it to partially evade the immune response generated by previous infection or vaccination. (122) Key mutations include K417N, E484K, and N501Y in the spike protein.(122) These changes have raised concerns about vaccine efficacy, with some vaccines showing reduced neutralisation capacity against the Beta variant.(123)

Gamma Variant (P.1)

This variant emerged in Brazil in November 2020 and contains several mutations of concern, including K417T, E484K, and N501Y.(124) Like the Beta variant, Gamma has shown the ability to escape from neutralising antibodies, posing challenges for vaccine effectiveness. It has been associated with increased transmissibility and potential increases in disease severity.(125)

Delta Variant (B.1.617.2)

First identified in India in October 2020, the Delta variant has been linked to the devastating second wave of COVID-19 in the country.(125) Characterised by mutations such as L452R and P681R, Delta is significantly more transmissible than earlier variants and has been associated with increased hospitalisation rates.(125) While vaccines remain effective in preventing severe disease and death, breakthrough infections have been reported.(125)

Omicron Variant (B.1.1.529)

Detected in multiple countries in November 2021, the Omicron variant has an extensive mutation profile, comprising over 46 mutations in the spike protein, which enhances its capacity to bind ACE2 receptors and evade neutralising antibodies.(126) Evidence suggests that Omicron is highly transmissible and may be capable of evading immunity from previous infection or vaccination to some degree.(126) However, the severity of disease it causes appears to be less than that of the Delta variant, and vaccines continue to provide significant protection against severe illness and hospitalisation.(127)

1.6 Therapeutic strategies

The management of COVID-19 has significantly evolved, with therapies tailored to disease stage, symptom severity, and patient-specific factors like age and comorbidities. As new variants emerge, strategies continue to adapt through ongoing research and clinical trials. Below are key therapeutic examples:

1.6.1 Antiviral Therapies

1.6.1.1 Oral antivirals

Remdesivir

The first antiviral drug authorised by the FDA for treating COVID-19, Remdesivir inhibits viral replication by targeting the RNA-dependent RNA polymerase.(128) Clinical trials have shown that it can shorten the recovery time in hospitalised patients, particularly those requiring oxygen therapy.(128)

Molnupiravir

An oral antiviral that introduces errors into the viral RNA, impairing its ability to replicate. (129) Molnupiravir has been authorised for use in certain countries for treating mild to moderate COVID-19 in adults at risk of developing severe disease.(129)

Paxlovid (Nirmatrelvir/Ritonavir)

This oral antiviral combination has shown high efficacy in preventing severe COVID-19 when administered early in the disease course.(130) Nirmatrelvir inhibits a protease enzyme that the virus needs to replicate, while ritonavir slows down nirmatrelvir's breakdown to help it remain in the body for a longer period at higher concentrations.(130)

1.6.1.2 Immunotherapies

Corticosteroids (e.g. Dexamethasone)

Dexamethasone has been shown to reduce mortality among severely ill COVID-19 patients requiring supplemental oxygen or mechanical ventilation. (131) It acts by reducing inflammation and the immune system's overreaction to the infection.(131)

Monoclonal Antibodies

Drugs like casirivimab and imdevimab (administered together) have been used to neutralise the virus, preventing it from entering cells.(132) Tocilizumab acts by preventing IL-6 damage to cells.(114) These are particularly effective in early-stage infection and for patients at high risk of progressing to severe disease.(114) Their usage has been evolving with the emergence of new variants with varying levels of susceptibility.

1.6.1.3 Supportive Care

Oxygen Therapy and Mechanical Ventilation

For patients with severe respiratory distress, supplemental oxygen and, if necessary, mechanical ventilation can be lifesaving.(133)

Anticoagulants

Given the increased risk of thrombosis associated with COVID-19, anticoagulants are commonly used in hospitalised patients to prevent clot formation.(134)

Convalescent Plasma

Early in the pandemic, plasma from recovered COVID-19 patients was used in hopes that the antibodies it contained would help fight the virus. Its efficacy remains subject to ongoing research and debate.(135)

1.6.1.4 Preventative Therapies - Vaccination

While not a treatment per se, vaccination is the cornerstone of COVID-19 prevention. Multiple vaccines have been developed and authorised globally.(136) Vaccinations strategies have focused on achieving widespread immunity to reduce transmission, severe disease, and deaths.(136)

Vaccination strategies

Vaccination strategies include three campaigns, mass vaccination, booster doses and achieving vaccine equity.(137) Mass Vaccination Campaigns focuses on achieving herd immunity which can significantly reduce the virus's spread within communities.(137) By prioritising high-risk groups such as healthcare workers, the elderly, and individuals with comorbidities, these campaigns aim to protect the most vulnerable segments of the population first, thereby reducing the overall burden of disease.(138) Booster doses represent another critical strategy, especially in the context of waning immunity and the emergence of new variants. (138) These additional doses are designed to reinforce the body's immune response, and the recommendation and administration of booster doses are based on scientific evidence and are adjusted as new data emerges about vaccine effectiveness over time.(138) Vaccine equity remains a significant challenge and a priority as ensuring that all countries, especially those with fewer resources, have access to vaccines is important for a global recovery.(139) Initiatives like COVID-19 Vaccine Global Access (COVAX) play a pivotal role in this aspect, striving to distribute vaccines equitably worldwide, thus preventing the emergence and spread of vaccine-resistant variants that could arise in unvaccinated populations.(139)

1.6.1.5 Vaccine types

mRNA Vaccines

The Pfizer-BioNTech (BNT162b2) and Moderna (mRNA-1273) vaccines use messenger RNA (mRNA) technology to instruct cells to produce a harmless spike protein of the virus, triggering an immune response.(140) They have shown high efficacy in preventing COVID-19 infection and severe outcomes.(140)

Viral Vector Vaccines

The viral vector vaccines are Oxford-AstraZeneca (ChAdOx1 nCoV-19) and Johnson & Johnson (Janssen, Ad26.COV2.S). These vaccines use a harmless virus (not the coronavirus) to deliver genetic material that codes for the SARS-CoV-2 spike protein into cells, inducing an immune response.(141, 142)

Inactivated Virus Vaccines

There are two main inactivated virus vaccines which are Sinovac-CoronaVac and Sinopharm BBIBP-CorV.(143) These vaccines contain the inactivated SARS-CoV-2 virus, which cannot cause disease but still prompts an immune response.

Protein Subunit Vaccines

The Novavax (NVX-CoV2373) vaccine contains harmless pieces (proteins) of the SARS-CoV-2 virus instead of the whole virus to stimulate an immune response.(144)

1.7 Long COVID

Long COVID, also referred to as Post-Acute Sequelae of SARS-CoV-2 Infection (PASC), is a complex, multifaceted syndrome following the resolution of acute COVID-19.(145) This condition underscores the enduring impact of the SARS-CoV-2 virus, extending beyond the acute infection phase into a prolonged period of convalescence that can last several months to over a year.(146) Long COVID is characterised by persistent symptoms and health complications that continue for more than three months after the initial infection, affecting a significant portion of COVID-19 survivors.(146) Long COVID encompasses a wide range of new-onset conditions, including but not limited to cardiovascular diseases, thrombotic and cerebrovascular disorders, type 2 diabetes, myalgic encephalitis/chronic fatigue syndrome (ME/CFS), and dysautonomia, notably postural orthostatic tachycardia syndrome (POTS).(147) Among these, cardiovascular symptoms such as stroke, chest pain, electrocardiographic abnormalities, postural tachycardia syndrome, atrial fibrillation, acute coronary syndrome, and venous thromboembolism stand out as particularly common manifestations.(146)

The incidence of Long COVID is notably variable, with estimates suggesting that 10-30% of hospitalised patients, 50-70% of those requiring hospitalisation, and 10-12% of vaccinated individuals may experience these prolonged health issues.(147) Interestingly, Long COVID predominantly affects individuals aged 36-50 years and is more common among non-hospitalised patients who had a mild form of the acute illness.(147, 148) The condition seems to have a predilection for women, with 55-75% of Long COVID cases reported in females aged 40-60 years.(147) Factors predictive of developing Long COVID

include higher BMI, lower levels of SARS-CoV-2 antibodies, older age, severe acute COVID-19 requiring hospitalisation, and the presence of symptoms affecting multiple organ systems during the acute phase.(148)

Research suggests a combination of biological, immunological, and physiological factors contributing to its development.(148) The following outlines several hypothesised mechanisms or causes of Long COVID:

1. Persistent viral reservoirs: It is proposed that remnants of the SARS-CoV-2 virus might persist in some individuals, particularly in tissues where the virus is not easily cleared.(149) This persistent viral presence could stimulate a continuous immune response or cause direct tissue damage over time. (149)
2. Immune dysregulation and autoimmunity: Long COVID might result from an abnormal immune response.(147) Some patients exhibit persistent inflammation, immune activation, or dysregulation even after clearing the virus. Elevated levels of pro-inflammatory cytokines, a phenomenon akin to a cytokine storm, and autoantibodies targeting the body's own tissues have been observed, suggesting that autoimmunity plays a role in symptom persistence.(147) Studies have highlighted alterations in immune cell populations, such as exhausted T cells and activated innate immune cells, which persist for months following the initial infection.(150, 151)
3. Endothelial dysfunction: Endothelial dysfunction can result in impaired blood flow, increased clotting, and other cardiovascular complications that contribute to the symptoms of Long COVID.(152-154)
4. Microvascular injury and thrombosis: COVID-19 can cause damage to the microvasculature, leading to microclots and reduced blood flow to various organs. (155) This vascular damage might underlie some of the persistent symptoms seen in Long COVID, such as fatigue and brain fog.
5. Mitochondrial dysfunction: There is evidence to suggest that SARS-CoV-2 infection could affect mitochondrial function, leading to energy metabolism issues. (156) This could explain the profound fatigue and exercise intolerance seen in many Long COVID patients.(156)
6. Neurological impact: SARS-CoV-2 has been shown to affect the nervous system, potentially leading to neuroinflammation and changes in brain structure and function. (157) These neurological impacts could contribute to a wide range of Long COVID symptoms, including cognitive impairment, headache, and sensory disturbances.(158, 159)

7. Reactivation of latent viruses: In some individuals, the immune system's dysregulation during and after COVID-19 infection could lead to the reactivation of latent viruses like Epstein-Barr virus and human herpesvirus 6 has been observed in Long COVID patients, suggesting a potential link to the chronic fatigue and cognitive dysfunction seen in ME/CFS.(160)
8. Tissue damage and organ injury: Direct damage from the virus, as well as the body's immune response to it, can cause lasting tissue and organ injury.(161) This damage, particularly to the lungs, heart, and kidneys, may underpin some of the long-term health issues associated with Long COVID.(161)
9. Hypoxia and oxygen uptake issues: Persistent issues with oxygen uptake and utilisation post-COVID-19 could contribute to symptoms of Long COVID.(162) Damage to the lungs and alterations in blood flow might impede efficient oxygen exchange, leading to ongoing symptoms of dyspnoea and fatigue.(163, 164)
10. Psychosocial Factors: The psychological impact of COVID-19, including stress from illness, isolation, and the overall pandemic situation, can exacerbate or contribute to the persistence of symptoms.(165, 166) Mental health issues such as anxiety and depression might intersect with physical health to create a complex clinical picture.(167)

Understanding Long COVID requires a multidisciplinary approach, integrating insights from virology, immunology, neurology, and psychology.(147) Ongoing research is crucial to unravel these mechanisms fully and to develop targeted therapeutic strategies to support those affected by Long COVID as it becomes increasingly clear that this syndrome represents a significant public health challenge.

1.8 Impact on Global Health

The COVID-19 pandemic has profoundly impacted global health, placing unprecedented pressure on healthcare systems, disrupting public health initiatives, and exacerbating existing disparities.(168) Hospitals and medical facilities worldwide were overwhelmed, leading to shortages of beds, medical supplies, and healthcare personnel.(169, 170) This has disrupted routine and emergency services for non-COVID conditions, as well as vaccination programmes, maternal and child health services, and chronic disease management, causing a broader health crisis.(171-173) Efforts to control other infectious diseases, such as tuberculosis, HIV/AIDS, and malaria, were severely affected, reversing decades of progress in some regions.(174) Beyond physical health, the pandemic triggered a global mental health crisis, with increased rates of anxiety, depression,

substance misuse, and domestic violence fuelled by fear, grief, economic instability, and social isolation.(175, 176) Economically, widespread lockdowns and restrictions led to business closures, job losses, and a contraction of the global economy.(177) These economic disruptions worsened poverty, food insecurity, and access to healthcare, further aggravating health inequalities.(178, 179) Marginalised groups, including the elderly, racial and ethnic minorities, and those with pre-existing conditions, were disproportionately affected.(180, 181)

Whilst the rapid development and deployment of COVID-19 vaccines marked a historic achievement, disparities in vaccine access and misinformation hindered global efforts to control the virus.(182, 183) The pandemic exposed vulnerabilities in public health infrastructure, emphasising the need for stronger systems for disease surveillance, emergency preparedness, and international cooperation.(3) It highlighted the importance of addressing health inequalities, strengthening global health governance, and balancing public health priorities with economic stability.(3)

1.9 COVID-19 and Hypertension

Hypertension is one of the most common preexisting conditions and comorbidities worldwide, affecting 1.13 billion adults in 2015 due to advancing age, higher prevalence in those individuals older than 50 years of age, with a small proportion affecting those who were younger.(184) Hypertension alone accounts for more than 10 million deaths and 218 million disability-adjusted life years worldwide. Hypertension control appears to be suboptimal, ranging from 7% to 65% worldwide with the prevalence increasing yearly.(185)

COVID-19 affects the cardiovascular system through complications like arrhythmias, myocardial injury, heart failure, pulmonary embolism, and coagulation disorders.(186) It also disrupts RAAS and endothelial dysfunction, key factors in hypertension and related complications which have been mentioned previously. Additionally, stress from COVID-19 can activate the sympathetic nervous system, raising heart rate and vascular resistance, potentially causing transient or persistent BP elevation.(187)

The interplay between COVID-19 and hypertension is complex, with the virus influencing BP and hypertension management through both direct effects on the cardiovascular system and indirect impacts resulting from the societal response to the pandemic. This relationship has been a focal point of medical research since the onset of the pandemic. Early research revealed that individuals with pre-existing hypertension were at a greater risk of mortality from COVID-19. Hypertension emerged as one of the most common

comorbidities among COVID-19 patients, with its prevalence being twice as high in severe cases compared to non-severe ones.(188, 189)

Elderly patients, who are more likely to have hypertension and related cardiovascular comorbidities, account for a higher incidence of severe COVID-19. Contributing factors include cardiac damage from long-standing hypertension, the presence of end-organ damage, interactions between COVID-19 and antihypertensive therapies, and the higher prevalence of hypertension in older age groups.(190, 191)

Although the exact reasons why individuals with pre-existing hypertension are at higher risk for severe COVID-19 remain unclear, it is evident that hypertension management typically involves four major drug classes: angiotensin-converting enzyme inhibitors/angiotensin receptor blockers (ACEIs/ARBs), calcium channel blockers (CCBs), diuretics, and beta blockers (BB).(192) These medications and their potential interactions with COVID-19 continue to be areas of active investigation.

1.9.1 Indirect Effects on Blood Pressure Regulation

The indirect effects of COVID-19 and hypertension stem from the widespread societal, behavioural, and healthcare system changes that have occurred in response to the pandemic.(193, 194)

The psychosocial impact of the pandemic—including anxiety, isolation, and economic stressors—has triggered activation of the sympathetic nervous system. This response has been linked to hypertension and poorly controlled BP, highlighting the interconnected effects of stress on physical health.(187)

Lockdown measures further compounded these challenges by limiting physical activity and disrupting daily routines. Many individuals experienced weight gain and a decline in cardiovascular fitness, both of which indirectly affected BP control.(195-197)

Adding to these difficulties, routine healthcare services were significantly disrupted during the pandemic. For chronic conditions like hypertension, delays in diagnosis, treatment adjustments, and reduced monitoring created additional barriers to effective management, leaving many patients vulnerable to worsening health outcomes.(198)

One of the most notable consequences of healthcare disruption was a decline in routine BP monitoring and management. As lockdowns curtailed in-person healthcare visits, many individuals missed opportunities for regular assessments, leading to poorer control of hypertension and increased risks of complications.(199-201)

This decline in healthcare access was not limited to monitoring alone. Research also documented a marked reduction in visits for hypertension management overall. Limited access to care likely exacerbated BP control issues and elevated cardiovascular risks for affected patients, illustrating the ripple effects of strained healthcare systems.(199)

Compounding these challenges, the pandemic also disrupted medication adherence among hypertensive patients. Interruptions in healthcare access and medication supply chains made it more difficult for individuals to maintain consistent treatment, further jeopardising their ability to manage their condition effectively.(202)

In response to these disruptions, healthcare providers rapidly adopted telemedicine and remote BP monitoring as alternative solutions. While these approaches offered new ways to address gaps in care, they also introduced challenges, particularly for individuals with limited digital literacy or access to technology. This highlights the importance of equitable access to digital tools in maintaining effective hypertension management.(203)

Amid these shifts, studies reported a general increase in BP levels across the population. Reduced physical activity, heightened stress, and disruptions in healthcare services collectively contributed to this trend, underscoring the multifaceted impact of the pandemic on cardiovascular health.(204)

Beyond individual health outcomes, the pandemic has had a profound global impact. While 5.4 million COVID-19-related deaths were reported in the first two years, this figure does not fully account for indirect effects, such as deaths resulting from delayed or forgone care for non-COVID conditions. Analyses of excess mortality suggest that these indirect effects significantly influenced cardiovascular disease outcomes, including hypertension.(205, 206)

These findings are particularly concerning given that patients with hypertension face an increased risk of severe COVID-19 outcomes. Higher rates of hospitalisation, intensive care admission, and mortality among hypertensive individuals further highlight the urgent need for improved strategies to manage both hypertension and its associated risks during public health crises.(205)

1.9.2 Impact on New-Onset Hypertension and Relevant Studies

Emerging evidence suggests a potential link between COVID-19 and new-onset hypertension, which may result from direct viral effects, disruption of the renin-angiotensin-aldosterone system (RAAS), or the physiological stress of illness. (207)

However, this relationship is not yet fully understood, and further research is needed to clarify the underlying mechanisms.

Studies have consistently shown that COVID-19 survivors face an increased risk of cardiovascular sequelae, including hypertension, compared to individuals not exposed to SARS-CoV-2. (207, 208) This suggests that the virus may have lasting impacts on cardiovascular health.

A meta-analysis conducted by Zuin et al. investigated the risk of new-onset hypertension in COVID-19 survivors within one year of the initial infection.(209) The pooled analysis revealed that recovered COVID-19 patients had a significantly increased risk of developing new-onset hypertension (HR 1.70, 95% CI 1.46–1.97, $p < 0.0001$, $I^2 = 78.9\%$) within seven months, supporting a direct association between COVID-19 and hypertension.

A retrospective cohort study by Zhang et al. analysed electronic health records from 277 international hospitals. (210) Their findings indicated that even patients diagnosed with COVID-19 in the outpatient setting had a heightened risk of developing hypertension, suggesting a possible connection between mild COVID-19 cases and subsequent hypertension. (210)

Further research has shown a high risk of incident cardiovascular disease, including hypertension, beyond 30 days of infection in non-hospitalised COVID-19 patients.(211) This highlights the potential for COVID-19 to contribute to new-onset hypertension even in cases that did not require hospital admission.

Comparative studies have demonstrated that hospitalised COVID-19 patients are significantly more likely to develop persistent hypertension than those hospitalised for influenza, emphasising the unique cardiovascular impact of COVID-19.(212)

1.9.3 ACE Inhibitors and RAAS Inhibitors

1.9.3.1 ACE Inhibitors and the COVID-19 pandemic

The COVID-19 pandemic unveiled a surprising and unexpected link between an infectious disease and the treatment of a chronic condition like hypertension. Central to this connection is the role of ACE2, which is pivotal in the pathogenesis of the SARS-CoV-2 virus and the regulation of BP and fluid balance through the RAAS, as previously discussed. Consequently, ACEIs, commonly used to manage hypertension, were thrust into the spotlight. (213) These medications, by modulating the RAAS, inadvertently influence ACE2 expression levels, leading to widespread speculation and concern about

their role in COVID-19 susceptibility and severity. There were fears that the use of ACEIs might exacerbate COVID-19 outcomes or increase risk of COVID-19 infection by increasing ACE2 expression, potentially facilitating viral entry into cells.(213)

Additionally, early in the pandemic, several observational reports highlighted a heightened mortality risk among individuals with pre-existing hypertension. This added additional complexities amplifying the uncertainty in establishing whether ACEIs might adversely impact patients taking these medications during the pandemic. An analysis of 44,672 confirmed COVID-19 cases revealed an overall case-fatality rate of approximately 2.3%, which escalated to 6% among those with pre-existing hypertension. Additionally, hypertension emerged as one of the most prevalent comorbidities in COVID-19 patients, with reported rates ranging from 8.0% to 31.2%. A meta-analysis demonstrated that the incidence of hypertension was twofold in severe cases compared to non-severe cases.(188) Another study found that the proportion of hypertension in 406 patients who died from COVID-19 infections was 39.7%, which was higher than the general population (12.6%) however this study was not adjusted for clinical characteristics.(214) There still remain controversies on whether hypertension is a risk factor for severe COVID-19 where in a retrospective study they observed that those with hypertension were more likely to develop severe COVID-19, however pre-existing hypertension was not independently associated with high risk of severe COVID-19.(189) However, many of these early studies had limitations, including confounding factors such as age, comorbidities, and the severity of hypertension (see section 1.9.3.4 Key Studies). Since elderly patients are more likely to have hypertension and additional cardiovascular comorbidities, this could potentially explain the higher incidence of severe COVID-19 illness observed in this population.

1.9.3.2 Protective effect of ACEIs

Alongside the growing concerns about ACEIs, which were primarily based on observational evidence, a strong opposing view emerged suggesting that ACEIs and ARBs might offer protective effects against severe COVID-19.(201, 215, 216) Upregulation of ACE2 might enhance the protective effects of ACE2 which counterbalances the effects of Ang II by generating Ang 1-7 a peptide with anti-inflammatory and vasodilatory properties. By mitigating the harmful effects of Ang II, and additionally by directly reducing the formation of Ang II, ACEIs may reduce inflammation, fibrosis, and thrombosis, contributing to better cardiovascular outcomes in infected patients. The proponents of this viewpoint supported the continued use of ACEIs and ARBs in patients with hypertension during the pandemic.(217) In contrast to this, cardiopulmonary diseases are associated with a reduction in ACE2 activity which limits

the effects of Ang II on the heart and vasculature, thus highlighting that ACE2 could be protective against the severe complications of COVID-19 infection. (218)

1.9.3.3 Mechanism of ARBs

ARBs are another class of antihypertensive drugs that target the RAAS and its mechanism of action of ARBs which is like ACEIs may influence the expression and activity of ACE 2.(218) ARBs antagonises the action of Ang II by blocking blocks Ang II type 1 receptor (AT1R). There is also a potential decrease in the inflammatory response as ARBs attenuate the inflammatory response associated with COVID-19 by modulating pathways downstream of the AT1R receptor. This in turns reduces the cytokine storm and acute respiratory distress syndrome (ARDS). Recent evidence has suggested ARBs as a potential treatment for COVID-19 where ARBs can increase soluble ACE2 in the blood stream which can bind SARS-CoV-2 virus to reduce the injury of organs with expression of ACE2.(219)

All these sparked intense debate within the medical community and led to fears of disrupting established cardiovascular prevention strategies. Patients and healthcare providers faced difficult decisions regarding the continuation or cessation of ACEIs, balancing the need for effective hypertension management against the potential risk of heightened COVID-19 susceptibility.

1.9.3.4 Key Studies

There have been several studies which looked at the association of ACEI/ARBs with risk of hospitalisation or from death of COVID-19. The investigators of the BRACE-CORONA trial found no significant difference in the number of days alive and out of hospital in patients with mild to moderated COVID-19 between patients continuing or discontinuing ACEI/ARBs.(220) However, their study was underpowered. A systematic review and meta-analysis recently published found that there was no evidence that ACEI/ARBs use affected the risk of COVID-19 infection, severity or mortality.(221) A large Swedish study which observed 1.4 million individuals, found that the use of ACEI/ARBs were not associated with increased risk of hospitalisation or from death from COVID-19.(222) This study acknowledged limitations where observational studies preclude establishing any causal relationship between exposures and outcomes. In addition to this, there were no data on clinical variables – BP, renal function, glycaemic control and variables associated with disease severity which may affect treatment decisions and study outcomes were unavailable.

In a population-based case-control study in Lombardy, Italy, Mancia et al. examined the potential association between the use of ARBs and ACEIs and the risk of COVID-19.(201) The study included 6272 confirmed COVID-19 patients and 30,759 matched controls, with a mean age of 68 years and 37% women in both groups. The use of ACEIs and ARBs was more common among COVID-19 patients, who also had a worse clinical profile. However, logistic regression analysis showed no significant association between the use of ARBs (adjusted odds ratio [aOR], 0.95; 95% CI, 0.86-1.05) or ACEIs (aOR, 0.96; 95% CI, 0.87-1.07) and COVID-19 risk. Similarly, there was no association among those with severe or fatal disease outcomes (aOR, 0.83; 95% CI, 0.63-1.10 for ARBs and aOR, 0.91; 95% CI, 0.69-1.21 for ACEIs), and no differences were observed based on sex. The study concluded that ACEIs and ARBs do not affect the risk of COVID-19.

Following this an international cohort study (Morales et al.) analysed electronic health records from Spain and the USA to determine if ACEI/ARBs were associated with increased COVID-19 susceptibility in hypertensive patients. (223)The study included 1,355,349 antihypertensive users, comprising 363,785 ACEI or ARB monotherapy users, 248,915 CCBs or thiazide diuretic (THZ) monotherapy users, 711,799 ACEIs or ARBs combination users, and 473,076 CCB or THZ combination users. The analysis found no association between COVID-19 diagnosis and ACEI or ARBs monotherapy (HR 0.98, 95% CI 0.84–1.14) or combination use (HR 1.01, 95% CI 0.90–1.15). ACEIs alone showed no significant risk difference compared to CCB or THZ monotherapy (HR 0.91, 95% CI 0.68–1.21) or combination use (HR 0.95, 95% CI 0.83–1.07). A direct comparison of ACEIs with ARBs indicated a moderately lower risk with ACEIs in combination use (HR 0.88, 95% CI 0.79–0.99), but not in monotherapy (HR 0.85, 95% CI 0.69–1.05). There were no significant differences between drug classes for risks of hospital admission with COVID-19, pneumonia, or severe complications. The study concluded that ACEI/ARBs use does not increase COVID-19 risk, suggesting these medications should not be discontinued to reduce COVID-19 risk.

In contrast, Jeffery et al. undertook a large retrospective cohort study to evaluate the clinical outcomes of patients with hypertension taking ACEI/ARBs compared to those on other hypertension medications after acute viral respiratory illness (AVRI) during the 2017-2020 influenza and COVID-19 seasons in the USA.(224) The study included 1,059,474 AVRI episodes, with 653,797 involving ACEI/ARB users and 405,677 involving users of other hypertension medications. The cohort was predominantly women (58.6%) and those aged 65 or older (72.9%). During the COVID-19 influenza season, the ACEI/ARB group experienced a larger increase in risk for inpatient stays (additional 1.5 percentage points,

95% CI 1.2 to 1.9), ICU/CCU use (0.3 to 2.7 pp), acute respiratory distress (0.7 pp, 0.1 to 1.2 pp), and ARD syndrome (ARDS) (0.9 pp, 0.4 to 1.3 pp), compared to the pre-COVID influenza seasons. However, there was no statistically significant difference in absolute risk of death (-0.2 pp, 95% CI -0.4 to 0.1 pp), though the relative risk of death was higher in the 2019/2020 season for the ACEI/ARB group (1.40, 95% CI 1.36 to 1.44) compared to the other hypertension medication group (1.24, 95% CI 1.21 to 1.28). Despite these findings, the small absolute differences do not suggest a need for changes in clinical practice.

Using data from the French National Health Insurance databases, Semenzato et al. followed three exclusive cohorts of ACEI, ARB, and CCB users aged 18 to 80 years from February 15 to June 7, 2020, to determine if the risk of COVID-19 varies by antihypertensive drug class.(225) Patients with a history of diabetes, cardiovascular disease, chronic renal failure, or chronic respiratory disease were excluded to focus on those with uncomplicated hypertension. The study population included nearly 2 million hypertensive patients (566,023 ACEI users, 958,227 ARB users, and 358,306 CCB users). Over the 16-week period, 2338 patients were hospitalised and 526 died or were intubated for COVID-19. The findings indicated that ACEIs and ARBs were associated with a lower risk of hospitalisation for COVID-19 compared to CCBs, with hazard ratios of 0.74 (95% CI, 0.65-0.83) for ACEIs and 0.84 (95% CI, 0.76-0.93) for ARBs. Additionally, these drugs were linked to a lower risk of intubation or death, with slightly better outcomes for ACEI users compared to ARB users. This large observational study suggests that long-term use of ACEIs or ARBs may be associated with a lower risk of severe COVID-19 outcomes in hypertensive patients compared to CCBs, potentially challenging previous hypotheses and prompting new ones.

These large studies highlight the complexities of interpreting observational data and underscore the need for cautious consideration of confounding factors. The consistent finding across these studies is the lack of evidence for increased risk associated with ACE inhibitors and ARBs in the context of COVID-19, provided early reassurance to guide clinical decisions during the pandemic.

Observational studies on the association between antihypertensive drugs and COVID-19 have provided initial insights but are inherently susceptible to biases that impact interpretation and evidence generation. Confounding bias, where variables correlate with both the exposure and the outcome, can create false associations or obscure real ones. Techniques like propensity score methods can partially address this by balancing covariates across treatment groups. Selection bias, resulting from unrepresentative

samples such as only including tested or hospitalised individuals, can be mitigated by using population-based methodologies and registry data. Collider bias, introduced by selecting participants based on criteria influenced by both exposure and outcome, can lead to spurious associations and should be avoided by not conditioning on variables within the causal pathway. Many studies also overlook dose variations and combination therapies, necessitating analyses stratified by dose and therapy type. Early studies focused on the pandemic's first wave may not apply to later stages or emerging variants, highlighting the need for ongoing monitoring and updated analyses. While observational studies rely on correlation, they cannot establish causation, which requires randomised controlled trials (RCTs). RCTs, however, are time-consuming, expensive, and may not represent real-world populations, also facing ethical constraints. To overcome these limitations, accurate causal inference from observational data is crucial. Advances in machine learning and deep learning offer new opportunities to enhance causal inference frameworks. These observational studies were valuable during the early pandemic stages, but more rigorous studies are needed to establish true relationships.

1.9.4 Guidelines and Recommendations

Early in the pandemic, there were several health organisations and societies issued guidelines recommending the continuation of ACEI or ARB therapy in patients already receiving these medications unless clinically contraindicated. This was due to the social media amplification of disinformation which were based on speculation where the use of ACEI/ARB could increase the risk of COVID-19 infection and worsen disease led to confusion and panic in particular amongst hypertensive patients.(213) This eventually led to harmful behaviour, where patients stopped their medications abruptly on their own accord.(213) The British and Irish Hypertension Society (BIHS) advised that all patients who were taking ACEI and ARBs should continue to do so as there was currently no evidence on whether there is increase susceptibility to COVID-19 infection or reduce the risk of serious lung disease following infection.(226) They informed that patients could be put at risk by stopping these effective treatments for their current condition and are encouraged to continue taking their medications.(226) This was also echoed by the International Society of Hypertension and the European Society of Hypertension where they affirmed the above and added that it was pertinent to ensure that management of raised BP is important to reduce cardiovascular risk and may improve outcomes among those infected by COVID.(227)

Following the social media amplification that common antihypertensive medications in particular ACEIs and ARBs that patients with hypertension should stop taking these

medications. This led to the Council of Hypertension of the European Society of Cardiology informed that there was lack of evidence which supported the notion of the harmful effect of ACEI and ARB in the context of the pandemic COVID-19.(228) They strongly recommended that treatment should be continued with current anti-hypertensive therapy as there is currently no robust evidence that the treatment should be discontinued.(228) A joint statement given by the American Heart Association, Heart Failure Society of America and the American College of Cardiology stated that the use of ACEIs and ARBs should be continued as prescribed in those with COVID-19. (229)

1.10 General Reflections on COVID-19 pandemic in the context of previous pandemics

The COVID-19 pandemic, much like historical global health crises, has exerted unprecedented pressure on healthcare systems worldwide, echoing challenges encountered in past pandemics. Healthcare systems have faced substantial challenges including surges in patient numbers, critical shortages of medical supplies, and significant attrition among healthcare workers. This situation has been particularly severe in lower- and middle-income countries where the diversion of resources from non-communicable diseases and ongoing infectious disease control efforts has had profound impacts. These challenges have tested the resilience of healthcare infrastructures and emphasised the need for flexible and collaborative crisis management strategies.

The economic repercussions of the pandemic have been severe, mirroring the socioeconomic disruptions observed in previous pandemics. Lockdowns and quarantine measures have led to business closures, widespread unemployment, and disruptions in global supply chains, exacerbating existing social inequities and vulnerabilities. Amidst these challenges, the use of social media has emerged as a double-edged sword, playing a complex role serving as both an important communication lifeline and a breeding ground for misinformation, complicating efforts to control the pandemic.

Moreover, the urgency of the situation has catalysed clinical research speeding up trials for treatments and the development and distribution of vaccines in a remarkable display of global cooperation. Additionally, emerging evidence suggests a link between COVID-19 and new-onset hypertension, suggesting an increased future cardiovascular burden and complicating the management of all the other health ramifications.

Despite these obstacles, the COVID-19 response has drawn from past pandemic insights, highlighting the critical role of historical knowledge in formulating effective containment

and mitigation strategies. This reflection emphasises the necessity to address several research gaps and future directions. Key areas include exploring the mechanisms of new-onset hypertension following COVID-19 infection, investigating long-term cardiovascular outcomes, assessing the impact of COVID-19 on individuals with pre-existing hypertension, and examining the role of antihypertensive therapy. Additionally, there is a pressing need to evaluate the impact of COVID-19 vaccinations on patients with hypertension, to explore healthcare access and management strategies during and after the pandemic, to address social determinants and disparities in health, and to develop effective prevention and public health strategies to mitigate cardiovascular complications.

Continuing efforts to leverage collective insights and foster innovative solutions that prioritise public health and economic stability is crucial. These efforts will not only enhance the management of the current pandemic but also improve global preparedness for future pandemics. By integrating lessons learned from the COVID-19 crisis and implementing strategies that address both healthcare and economic challenges, the global community can build more resilient systems capable of withstanding future public health emergencies. This approach will ensure a coordinated and efficient response that minimises the impact on human health and the global economy, positioning us better for any similar challenges ahead.

1.11 Research Gaps and Future Directions

Identifying and addressing research gaps in the context of hypertension and COVID-19 is important for advancing our understanding of the disease's pathophysiology, improving patient care, and developing targeted interventions. Multidisciplinary collaboration and innovative research approaches are essential. Below are the prioritised areas for future research:

Mechanisms of New-Onset Hypertension and Long-Term Cardiovascular Outcomes

Research should focus on confirming the link between COVID-19 and new-onset hypertension, exploring the pathophysiological mechanisms such as endothelial dysfunction, RAAS dysregulation, and the impact of viral persistence on cardiovascular health. Additionally, longitudinal studies are needed to assess the long-term cardiovascular outcomes of COVID-19 survivors, particularly those with new-onset hypertension. These studies should examine the progression of hypertension, subsequent cardiovascular diseases, and impacts on mortality and morbidity.

Impact of COVID-19 on Pre-existing Hypertension and Antihypertensive Therapy

Further investigation is required to determine how COVID-19 affects BP control in patients with pre-existing hypertension and its implications for cardiovascular risks and outcomes. There is also a need to evaluate the efficacy and safety of antihypertensive agents, including ACEIs and ARBs, in COVID-19 patients, to identify optimal hypertension management strategies.

Role of Antihypertensive Therapy on COVID-19

While current evidence suggests that ACEIs and ARBs do not worsen COVID-19 outcomes, there is limited data from randomised controlled trials. Furthermore, the impact of evolving viral variants and vaccination is unknown. Future studies should examine the efficacy and safety of different antihypertensive agents in the risk of COVID-19 and its sequelae.

Vaccine Impact on Hypertensive Patients and Healthcare Access

As the global COVID-19 vaccination rollout continues, research should assess its efficacy and safety in hypertensive patients, including potential effects on BP and cardiovascular risk. Concurrently, the pandemic's impact on healthcare access and hypertension management needs to be addressed. Studies should evaluate the effectiveness of telemedicine and remote monitoring, identify barriers to effective care, and develop mitigation strategies.

Social Determinants, Disparities and Public Health Strategies

COVID-19 has underscored significant health disparities, necessitating research into how social determinants of health affect hypertension management and outcomes. This includes the impact of socioeconomic factors, race, and ethnicity on hypertension risk and care accessibility. Research into effective prevention and public health strategies is also critical to mitigate the risk of COVID-19 and its cardiovascular complications, with a focus on lifestyle factors like diet and physical activity to prevent new-onset hypertension and optimise cardiovascular health post-infection.

Chapter 2 Thesis Objectives

The primary aim of this thesis is to address key gaps in our understanding of the interplay between COVID-19 infection, RAAS dysregulation, and hypertension. This work is intended to contribute to the formulation of enhanced clinical management strategies and the development of public health policies aimed at reducing the cardiovascular consequences of the COVID-19 pandemic.

Objective 1: Examining the Association Between COVID-19 and Future Risk of Hypertension. This objective seeks to determine whether COVID-19 infection leads to an increased risk of developing hypertension post-recovery. I carried out a clinical phenotyping study which is a prospective follow-up of non-hypertensive individuals post-COVID-19 recovery, assessing blood pressure and endothelial function periodically over a 12-month period using ambulatory blood pressure monitoring and brachial flow mediated dilatation.

Objective 2: The Association Between RAAS Dysregulation and Hypertension Post-COVID-19. This objective focuses on elucidating the role of RAAS dysregulation in contributing to heightened hypertension risk following COVID-19 infection. This research will analyse RAAS pathway components and blood pressure changes in the cohort established in Objective 1, employing techniques such as RAAS fingerprinting over the same 12-month follow-up period.

Objective 3: Assessment of quality of life in individuals who have recovered from COVID-19. This objective seeks to determine if whether individuals post COVID-19 have any differences in QoL utilising EQ-5D-3L instrument (EQ-5D-3L Index and Visual Analog Scale (VAS) scores)

Objective 4: The evaluation of Transformer-Based Counterfactual Estimation of Individual Treatment Effects - Analysis of Angiotensin Converting Enzyme inhibitors (ACEIs) and risk of SARS-CoV-2 infection. This objective is to examine how the use of various antihypertensive drug classes prior to infection influences the risk of SARS-CoV-2 infection. This will involve longitudinal studies using machine learning techniques to analyse linked electronic health records, focusing on the long-term effects of antihypertensive drugs, while adjusting for confounding variables such as age, existing comorbidities, and other medications.

Chapter 3 General Methods

This chapter is arranged as per the thesis objectives and closely describes the general methods used based on the sequence of the chapters and sections in this thesis. The general approach and methods to address the overall objectives are presented here, additional details, particularly those specific to individual result chapters are provided in detail in the dedicated chapters.

3.1 Methods for Chapter 4: Clinical Phenotyping Study (Addresses Objective 1 and 2)

In this section, I shall describe the methods for recruitment, phenotyping, follow-up and data collection for the Longer-term effects of SARS-CoV-2 infection on blood vessels and blood pressure (LOCHINVAR) clinical phenotyping study (Chapter 4). The study design has been published previously.(230) The details of statistical analyses are presented in the methods section of Chapter 4.

Data sources, preparation, collection, management and retention

All patients will be assigned a unique study identifier which will not contain any identifiable information. Patients in Scotland has a unique community health index (CHI) number. Where linkage is required the key to CHI identification will be kept separate and only on an NHS computer. Data will be recoded in CASTOR EDC (www.castoredc.com) electronic case report forms (eCRF) held centrally on their ISO 27001 certified data management system. CASTOR EDC have been vendor assessed by the study sponsor NHSGG&C in accordance with regulatory guidelines. CASTOR EDC complies with all applicable laws and regulations, including ICH E6 Good Clinical Practice (GCP), 21 CFR Part 11, EU Annex 11, General Data Protection Regulation (GDPR), HIPAA (US), ISO 9001 and ISO 27001. Data entry will be completed by either the study nurse, clinical fellow or an investigator. Data will be validated at the point of entry into the eCRF and can be checked by sponsor audit if required. To enable evaluations and/or audits from regulatory authorities, records will be kept, including the identity of all participating subjects (sufficient information to link records), all original signed informed consent forms, source documents, and detailed records of treatment disposition in accordance with ICH GCP and local regulations. All samples will be retained by the investigators. Data will be retained for a minimum of 20 years.

Study Population and Recruitment

Participants who were admitted to the Queen Elizabeth University Hospital (QEUH) immediate assessment and acute receiving units from 1 September 2020 and 31 December 2020 who presented with COVID-19 or non-COVID illness requiring hospitalisation. The participants were identified via clinical case note review of all patients during this time period. Anonymous data (using study ID) for the case note review and phenotyping study will be held within CASTOR EDC.

Approval was obtained from NHS Information Governance (Caldicott Guardian) to obtain data from case note review and invite letter to be sent to potential participants and Safe Haven LPAC (data deposit and linkage). Identifiable data of participants taking part in the clinical phenotyping study which includes consent forms and patient letters correspondence will be stored in the Glasgow Clinical Research Facility (GCRF) at the QEUH. Identifiable data in electronic form will be stored on a password protected NHS computer. No patient identifiable data will leave the NHS premises or computers. Physical study documents will be stored in a locked filing cabinet within an NHS office with controlled access. The site file will be archived via the NHS R&I office. Each participant will be given a unique identifier which is LVXXX, the similar unique identifier will be used for CASTOR and specimens taken. For the different visits in the study the unique identifier will be LVXXX.Visit_Number to avoid any discrepancies in specimen processing.

COVID-19 Blood Pressure Endothelium Interaction (OBELIX) Study: Participants, Methodology and Key Findings

In Glasgow, the COVID-19 Blood pressure EndotheLium Interaction (OBELIX) study was carried out in non-hypertensive patients ≥12 weeks after admission to the Queen Elizabeth University Hospital between 1 April 2020 to 31 May 2020 with suspected/confirmed COVID-19 or with confirmed non-COVID-19 related condition (clinicaltrials.gov NCT04409847).(231) The study compared 30 non-hypertensive patients exposed to SARS-CoV2 and 18 contemporaneous controls, the study showed higher ABPM in the cases compared to controls (SBP 24-hour (beta [95%CI]: 8.6 [0.9-16.3]; p=0.03), SBP daytime (8.6 mmHg [1.5-17.3]; p0.02), DBP day-time (4.6mmHg [0.1-9.1]; p0.05) and lower plasma renin and Ang-1-10 (-0.4[-0.9-0.1]; p=0.08; -0.7[-1.2 - -0.1]; p=0.02 respectively). Paired analysis of hospital discharge office BP and study visit showed an 11mmHg difference in SBP between groups (11.5[3.12];19.8; p=0.008).

Participants from the OBELIX study were invited to take part in the LOCHINVAR study for a follow up visit.

Data variables collected from clinical case note review

Data variables collected from the clinical case note review will be entered into CASTOR EDC. Data on baseline demographics which include, anthropometric measures (height, weight, BMI), age, sex, ethnicity, smoking status and alcohol intake. From the admissions details, data on temperature, heart rate, hospital BP (systolic BP (SBP) and diastolic BP (DBP)), oxygen saturations, respiratory rate. We also collected data on presenting symptoms and duration of symptoms (fever, cough, dyspnoea, fatigue, myalgia, chest pain, gastrointestinal disturbance, loss of smell, loss of taste and cutaneous symptoms). Details on SARS-CoV-2 test and result was collected and relevant imaging investigations (chest x ray, abdominal x ray, computerised tomography scan, magnetic resonance imaging scan, ultrasound scan) and relevant cardiac investigations (electrocardiogram, echocardiogram) if indicated. Data pertaining to any critical care admission including details on ventilatory support and outcomes (death/discharge) was collected. The hospital admission outcome was also collected on whether the patient was discharged to home or care facility or other. Medication and prescription data was collected.

Eligibility for clinical phenotyping study

From the clinical case note review, participants who are potentially eligible for the clinical phenotyping study will be identified during the case note review. This will be based on the inclusion and exclusion criteria explained in Table 5. The participants if deemed eligible will be sent an invite, advert and participant information leaflet for the study. There will be recruitment of potential participants as controls using advert. Interested participants will contact the clinical research fellow (myself) if they wish to take part and they will be sent a patient information sheet about the study.

3.1.1 Recruitment strategies for the clinical phenotyping study

Due to the dynamic nature of the COVID-19 pandemic, the first year of the PhD was focused on the clinical phenotyping study to ensure there were enough participants recruited at baseline. The study design defined clear inclusion and exclusion criteria, but case control matching was not included. The reason for not including case-control matching was because this would add additional complexity to the recruitment and risk study failure due to not attaining target recruitment within the funded study period. There was a need to adapt analytic strategies which is described in Chapter 4 in more detail.

3.1.2 Clinical phenotyping study visits

Participants who have agreed to take part in the clinical phenotyping study will be screened to determine eligibility and consent will be obtained. (Figure 5) The data

collected from the phenotyping study visits (baseline and 12 months) will be entered into CASTOR EDC.

Table 5 Inclusion and Exclusion Criteria

Inclusion Criteria
SARS-CoV-2 Positive Group • Age 30-60 years. • Admission between 1 September 2020 to 31 December 2021. • Clinically suspected or PCR confirmed COVID-19 on admission. • No history of hypertension or current drug treatment for hypertension. SARS-CoV-2 Negative Group • Age 30-60. • No history of hypertension. • No antihypertensive drugs. • Confirmed RT-PCR negative and admission through QEUH immediate assessment unit and acute receiving units 1 April 2020 to 31 December 2021 or no history of SARS-CoV-2 infection or COVID-19.
Exclusion Criteria
Inability to give informed consent/lack of capacity. Non-English, Arabic, Polish or Urdu speakers. BMI >40. eGFR <60 ml/min. Pregnancy. History of • Cancer within 5 years • Persistent atrial fibrillation • Severe illness, at investigator discretion Prescription of • BP lowering drugs • Oral Corticosteroid (chronic use) • Immunosuppressive agents • Oral NSAIDs (chronic use)

3.1.3 Sampling Procedures

3.1.3.1 Sampling Procedures

Blood Sampling Procedures

Venepuncture will be performed from the antecubital fossa with the limb in a downward position. Blood collection will be conducted using the BD Vacutainer® Push Button Blood Collection Set or BD Vacutainer® Safety-Lok™ Blood Collection Set. Other sites and needle sizes may be used if required. The blood draw will be in the appropriate blood bottles in the following order SST tube, EDTA tube and sodium citrate.

Sampling and Sample Tube Labels

All samples that are ordered through Trakcare for routine blood tests available will have the patient details. Routine clinical biochemistry and haematology samples are sent to the local NHS laboratory using the hospital pod system to the laboratory. The results from

these tests will be entered into CASTOR EDC based on their allocated patient identification number in an anonymised form.

For samples that are not ordered through Trakcare (RAAS fingerprinting, future vascular biomarkers) this would be labelled with the study ID (LOCHINVAR), participant ID (VPLVXXX.Visit Number), date/time, staff initials, and sample identity as above. A laboratory information management system (LIMS) will be employed with samples stored in a linked anonymised form.

Indication for laboratory blood tests

Biochemistry and haematology results were collected to include parameters potentially reflecting the systemic effects of COVID-19. These tests, conducted as part of the COVID-19 study, aimed to assess complications associated with the disease. (Table 6) While not intended for hypothesis testing, these results served as baseline data to assess general background patient health and any changes post COVID-19. This analysis would help decide plans for future follow-on studies.

RAAS Fingerprint

Samples taken for RAAS fingerprinting (Attoquant Diagnostics, Vienna, Austria) and future biomarkers are centrifuged and serum is separated into 1ml aliquots (9 aliquots in total) and then stored in a -80°C freezer. Future biomarkers will be transported to the BHF Glasgow Cardiovascular Research Centre.

The samples for RAAS Fingerprint were shipped to Attoquant Diagnostics, Austria following completion of all baseline visits. The methods for RAAS Fingerprint were provided by Mr Oliver Domenig of Attoquant Diagnostics which is described below.

Equilibrium concentrations of angiotensin peptides (Ang II, Ang 1-7, Ang I, Ang 1-5) and aldosterone were quantified in serum samples by liquid chromatography-mass spectrometry/mass spectroscopy performed at a commercial laboratory (Attoquant Diagnostics, Vienna, Austria), using previously validated and described methods after ex vivo equilibration. (254-256) Briefly, samples were spiked with a stable isotope-labelled internal standard for each angiotensin and a deuterated internal standard for aldosterone (aldosterone D4) after equilibration, and analytes were extracted using C18-based solid-phase extraction. Extracted samples were analysed using mass spectrometry analysis with a reversed-analytical column (Acquity UPLC C18, Waters) operating in line with a XEVO TQ-S triple quadrupole mass spectrometer (Waters Xevo TQ/S, Milford, MA) in multiple reaction monitoring mode. Internal standards were used to correct for analyte

recovery across the sample preparation procedure in each individual sample. Analyte concentrations were calculated from integrated chromatograms considering the corresponding response factors determined in appropriate calibration curves in serum matrix when integrated signals exceeded a signal-to-noise ratio of 10. The lower limits of quantification for the analytes in human serum were 3 pmol/L (Ang I), 2 pmol/L (Ang II), 3 pmol/L (Ang1-7), 2 pmol/L (Ang1-5), and 13.9 pmol/L (aldosterone), respectively. Angiotensin-based markers for renin (PRA-S) and angiotensin converting enzyme (ACE-S), were derived from Ang II and Ang I concentrations by calculating their sum and ratio, respectively, whereas the ratio of aldosterone/Ang II (AA2) was calculated to assess adrenal responsiveness after Ang II signalling resulting in the release of aldosterone.

Urine sampling

For urine samples, the participant will be provided two white top urine collection pots with/without a foil bowl. The participant was asked to wash their hands before and after the urine collection. The participant is advised to provide a mid-stream sample. The urine sample will be dipped for analysis before processing. The urine sample will be divided into 3x 1ml aliquots and stored in the -80^{o}C freezer. A summary of sample collection is shown in Table 6.

Action on blood results

All routine and additional blood results were reviewed within 5 days of being resulted by me. All abnormal results were reviewed by me and will be acted upon if deemed necessary.

3.1.4 Study Procedures

3.1.4.1 Office Blood pressure (oBP)

Office BP (oBP) which is clinic BP will be measured as per the BIHS guidelines.(232) The patient will be seated for at least five minutes, relaxed and not moving or speaking. oBP is measured in both arms; the arm with the highest oBP is selected. oBP is then repeated a further 2 times, at least one minute apart, and the mean of readings 2 and 3 will be used. The appropriate cuff size will be used depending on brachial arm circumference.

3.1.4.2 Ambulatory blood pressure monitoring (ABPM)

Ambulatory blood pressure monitoring (ABPM), using Spacelabs 90217 ABPM will be performed following BIHS guidelines.(232) Patients will undergo ABPM for 24 hours with readings every 30 minutes during the daytime (0800-2159) and every 60 minutes during night-time (2200-0759). ABPM will be valid if 14 measurements are made during the

daytime as per BIHS guidance. The appropriate cuff size will be used depending on brachial arm circumference.

Table 6 Summary of sample collection for every study visit

Sample	Sample Collected	Total Volume	Label
Routine Haematology and Biochemistry			
Blood – 4 ml Serum Separator Tube (Gel) Yellow top x 1	SARS-CoV-2 IgG Antibody test	5 ml	Paper Form
Blood – 4 ml Serum Separator Tube (Gel) Yellow top x 1	Urea and electrolytes, magnesium, bone profile, liver function tests, lipid profile, haematinics	5 ml	Trakcare
Blood – Potassium EDTA Purple Top x 3 (2ml in each)	Full blood count, Renin/Aldosterone, Glycosylated haemoglobin (HbA1c)	6 ml	
Blood – Sodium Citrate x 1	Coagulation Screen	5ml	
Blood – Fluoride/oxalate Grey Top x 1	Glucose	1 ml	
RAAS Fingerprinting			
1 x 4ml Serum Separator Tube (Gel) Yellow (RAAS fingerprinting) https://www.attoquant.com/faq/#sampling	RAAS Fingerprinting serum SST 3x 1ml aliquots	4mls	VPLVXXX.VisitNumber*
Future Biomarkers (Blood and Urine)			
Vascular Biomarkers			
Blood – 1x9ml purple, 1 x yellow top (5ml)	Future Biomarkers	14 ml	VPLVXXX.VisitNumber*
Urine			
Urine (universal container) x 1 x (10ml) containers	Urine Biomarkers	Approx. 10ml Urine 3x 1ml aliquots	VPLVXXX.VisitNumber*
24 hr Urine Collection	Urine Sodium	10 ml of urine will be stored, 2 L for analysis of 24 hr urine sodium	Trakcare VPLVXXX.VisitNumber *

Setting up the 24 hour ABPM before use

The ABPM monitor will be fitted with 4 AA batteries. The ABPM will be connected to a computer with the computer interface cable which has Cardio-Navigator software. A new patient folder will be created by entering the mandatory subject ID, patient number, patient initials, D.O.B and gender fields. The configuration for the ABPM monitor will be set up to determine the start time, time intervals (day and night), display pressure and limits as per study protocol. The configuration settings will be sent to the ABPM monitor. The ABPM monitor will be disconnected from the PC. This will be as per manufacturer guidelines (Spacelabs Medical, Inc Instruction Manual Model number 90207/90217)

Preparation for patient use

A record of the monitor ID number and patient details will be kept with the relevant study documentation and labelled with the participant's LOCHINVAR ID. The ABPM monitor will be fitted, with instructions/documentation provided to the participant.

Completion of ABPM

Following completion of the ABPM, the patient is encouraged to return the ABPM monitor the next day by dropping it off at the clinical research facility or a taxi to collect the device can be arranged.

3.1.4.3 Brachial flow-mediated dilation (FMD)

Brachial flow mediated dilatation (FMD) of the forearm arteries will be used to assess endothelial function. This technique is the most widely used non-invasive method of assessing endothelial function and is considered as the gold standard for non-invasive assessment of peripheral vasoreactivity. This technique relies on brachial artery imaging with high resolution ultrasound during a period of reactive hyperaemia.

Training

I received training to use the UNEXEF38G prior to the commenced on the study by an experienced operator Joanne Flynn and received an educational session from Mr Achim Schwarz (General Manager, UNEX system). I was deemed competent to conduct the FMD procedure independently.

Equipment, setting and test duration

The equipment that is required will be the ultrasound machine (UNEXEF38G). The operation manual UNEXEF38G which will be used in conjunction. The machine is used in conjunction with a BP monitor and ultrasound jelly. The duration of test will last approximately 20 minutes, with a 10-minute set up and baseline recording, 5 minutes for

occlusion and post-occlusion respectively. The FMD procedure will be conducted on the patient couch in the clinical research facility room with curtains to ensure privacy. The light will be slightly dimmed to allow clear visualisation of the UNEXEF38G screen.

Information provided to participants

The participant undergoing the test will be informed that there may be feel some arm discomfort at the site of cuff inflation which may be slight numbness and tingling. When the cuff is deflated, the arm can feel tingly and hot. The participant will be reassured that this is a safe and well tolerated procedure.

Standard operating procedure

The participant is asked to lie supine with their arm outstretched for 10 minutes. A BP cuff is placed on the forearm. ECG clips are applied on each wrist (red clip on the right, green on the left). Brachial BP is measured on the left arm. In the measurement tab, the details for patient ID, initials, birth date and gender will be entered. A new registration is started and the FMD procedure is commenced. The brachial artery is identified in the short axis using the ultrasound probe. It is important to obtain a clear longitudinal image as possible with the intima visible. Once the image of the artery is obtained with good clarity based on operator's discretion, the centre of the artery is tapped in the short axis images which will activate the automatic tracking. The automated tracking system will adjust the probe position. A baseline (rest) image is obtained for 3 minutes, and flow velocity is recorded. When the baseline image has been obtained, ask the participant to lie still and commence the arterial occlusion. The cuff is occluded for 5 minutes. At the end of this 5-minute period, the cuff is deflated. The machine will start to measure the max diameter is measured automatically when the cuff pressure becomes 0 mmHg. When 1 minute is left on countdown, the machine will autotrack to tidy up the image and ensure clear as possible for ongoing measurements. It will measure "base diameter" during this minute. If the machine autotracks off the image, it will be able to refocus by clicking on the centre of the short axes again.

Data collection

The FMD procedure will end and provide automated outputs for %C-FMD, %FMD/L-FMC, Rest Diameter (mm), Base Diameter (mm), Max Diameter (mm), vasodilation amount (seconds), maximum blood flow rate (seconds) and basal intimal wall thickness (bIMT, mm).

Additional manual quality checks and analysis were performed by a single operator (SL) where required. The %FMD is calculated as [(Max Diameter-Rest Diameter)/Rest Diameter]*100.

The result will be saved on a secure password protected USB stick and will be documented on CASTOR EDC based on the participant allocated LOCHINVAR ID. Details on the FMD procedure including cuff inflation pressure, test arm, completion of protocol, documentation of adverse event will be documented on the CASTOR EDC eCRF form.

6-Minute Walk Test (6MWT)

The original purpose of the 6-minute walk test (6MWT) was to assess exercise tolerance in patients with chronic respiratory disease and heart failure however the use has expanded to be used as a performance-based measure of functional exercise capacity in other populations. Level of shortness of breath and fatigue using the Borg scale, BP, oxygen saturations and heart rate will be measured pre and post 6MWT. This procedure has been modified from the ATS and the Australian Lung Foundation/Australian Physiotherapy Association guidelines. (233, 234)

Training

Prior to the study starting, I received training from the respiratory physiotherapists and clinical research nurses and respiratory physiotherapists who are experts in conducting 6MWT at the Golden Jubilee National Hospital. I was deemed competent to conduct the 6MWT independently.

Equipment, setting and test duration

The equipment required for the 6MWT will include a countdown timer, two small cones to mark the turnaround points, clipboard with a 6MWT proforma, Borg scale, BP machine, pulse oximeter and access to oxygen and telephone in case of emergency.

Before the participant starts, the participant will be asked if they have any symptoms of unstable angina or cardiac event in the last month and a routine set of observations are taken to ensure there are no contraindications to the 6MWT (resting heart rate >120 bpm, SBP>180 mmHg, DBP > 100 mmHg, stable angina). I will be the one supervising the test will ensure that it is a safe location where there is appropriate access to the emergency trolley. I am certified in cardiopulmonary resuscitation in Advance Life Support by the Resuscitation Council (UK) – approved cardiopulmonary resuscitation course. The test will be stopped if there are any clinical concerns about the participant’s health. The test will last at least 30 minutes.

Standard operating procedure

The track will be 30 metres in length with turn around points marked with a cone. The starting line will be informed to the participant. The participant should rest for at least 15 minutes before beginning the 6MWT. After the subject has been at rest for 15 minutes, obtain and record measurements of BP, heart rate, oxygen saturation and Borg dyspnoea and fatigue scores. The participant is demonstrated the walking track. The lap counter will be set at zero, and the timer to 6 minutes (or stopwatch to zero).

Standardised phrases will be used as specified below:

"The object of this test is to walk as far as possible for 6 minutes. You will walk back and forth in this corridor. Six minutes is a long time to walk, so you will be exerting yourself. You may get out of breath or feel exhausted. You are permitted to slow down, to stop, and to rest as necessary. You may lean against the wall while resting, but resume walking as soon as you are able.

You will be walking back and forth around the cones. You should pivot briskly around the cones and continue back the other way without hesitation. Now I'm going to show you. Please watch the way I turn without hesitation."

"Are you ready to do that? I am going to use this counter to keep track of the number of laps you complete. I will click it each time you turn around at this starting line. Remember that the objective is to walk as far as possible for 6 minutes but don't run or jog.

Start now, or whenever you are ready."

The timer will start when the participant begins to walk with monitoring for any untoward signs and symptoms during the duration of the test. If the participant stops prior to the test being completed, the participant is asked why they stopped and reason, the time recorded for duration stopped, and if the patient wishes to continue or abandoned the test, the distance, time stopped and reason will be documented.

The following standardised phrases are used below during each minute of the test:

At minute one: “You are doing well. You have five minutes to go.”
At minute two: “Keep up the good work. You have four minutes to go.” At minute three: “You are doing well. You are halfway done.”

At minute four: “Keep up the good work. You have only two minutes left” At minute five: “You are doing well. You have only one minute to go.”

When the timer is 15 seconds from completion, the clinical research fellow (myself) will say:
"In a moment I'm going to tell you to stop. When I do, just stop right where you are and I will come to you".

When the time reaches exactly 6 minutes, the clinical research fellow (myself) will say:
"Stop!". Consider taking a chair over to the subject if they look exhausted. Mark the spot where they stopped by placing a marker on the floor.

Once the test has finished, the participant is invited to sit down. Immediately, oxygen saturation, heart rate, Borg dyspnoea and fatigue rating is recorded. BP is also measured.

The total of the number of lengths/laps walked is counted. The tally of the total distance walked by the subject, rounded to the nearest metre, and record on the proforma and documented on CASTOR EDC.

Borg Dyspnoea and Breathless Scale

The Borg dyspnoea and breathless scale is used at the beginning and at the end of the 6MWT. (Table 7) The scale is shown to the subject and standardised phrases are used to ask their level of breathlessness and fatigue.

Table 7 Borg Scale

SCALE	SEVERITY
0	None
0.5	Very very slight (just noticeable)
1	Very slight
2	Slight
3	Moderate
4	Somewhat severe
5	Severe
6	
7	Very severe
8	
9	Very very severe (almost maximum)
10	Maximum

Standardised phrases that will be used are:

“Please grade your level of shortness of breath using this scale”

Then this is asked: “Please grade your level of fatigue using this scale”

At the end of the exercise, the participant is reminded the number they previously chose before the 6MWT and the participant is asked to grade their shortness of breath level again. The participant is then ask to grade their level of fatigue, after reminding them of their grade before the 6MWT.

Data collection

Data from the 6MWT procedure will be entered into CASTOR eCRF for the screening assessment. Prior to commencing the test, data on clinical observation parameters at baseline will include oxygen saturation, heart rate, the baseline dyspnoea and fatigue BORG scale and BP.

At the end of the test, clinical observations are measured including the completion dyspnoea and fatigue Borg scale and BP.

The total of the number of lengths/laps walked is counted. The tally of the total distance walked by the subject, rounded to the nearest metre, and record on the proforma will be entered into CASTOR EDC.
The optional procedures for the LOCHINVAR study is described below:

3.1.4.4 Home blood pressure monitoring (HBPM)

Home blood pressure monitoring (HBPM) will be performed using OMRON M3 device in accordance with BIHS guidelines.(232) Readings will be recorded, in triplicate, morning (0600-1200) and evening (1800-0000) for seven days. The first day’s readings will be excluded from the mean. A minimum of 5 sets of morning and evening readings will be required for HBPM to be valid within a 10-day period.

3.1.4.5 24-hour Urine Collection

A 24-hour cannister and an instructions sheet will be provided to participants to perform a 24-hour urine collection. The volume of 24-hour urine collected will be measured and 10 mls from the 24-hour urine sample will be aliquoted for storage while the rest will be processed in the local NHS laboratory for measurement of urine electrolytes.

3.1.5 Data Outputs and Cleaning

Following the completion of the last participant visit at 12 months. The data outputs from CASTOR EDC will be downloaded into an excel file format. The data will have the list of all variables created in the study and list of all the option groups with the option group

names and values. The data was cleaned for analysis by ensuring that there were no implausible results (e.g indeterminate SARS-CoV-2 status, biochemistry values which were not within physiological range). Those with missing covariables were excluded; no imputation was performed. The code for analysis will be written in R on the Safe Haven platform. Data will be downloaded from CASTOR EDC will be following completion of last patient, last visit.

3.1.6 Data Linkage with Previous OBELIX Participants.

In the clinical phenotyping study, there will be participants that have taken part in the OBELIX study. The data will be extracted for analysis on Safe Haven, and this will be mapped to pre-existing OBELIX participant IDs. If an OBELIX participant attended the LOCHINVAR baseline visit, this will be re-coded as a 12 month visit.

3.1.7 Study sponsorship, monitoring and audit

The study is sponsored by the NHS GG&C and is coordinated in the GCRF. The study will be subject to audit by the Sponsor to ensure quality of study data and compliance with regulations. Any change in the study protocol will require an amendment which will be initiated by the Chief Investigator in discussion with the sponsor and will be ethically approved.

3.2 Methods for Chapter 5: Quality of Life (Addresses: Objective 3)

In this section, the methods used to assess Quality of Life in participants who have recovered from COVID-19 is explained here.

3.2.1 EQ-5D-3L

The EQ-5D-3L, developed by Williams in 2005, is a widely recognised instrument designed to assess health status, offering a comprehensive measure for evaluating and comparing health outcomes.(235) Its key purpose was to facilitate the calculation of quality-adjusted life years (QALYs), required for economic evaluations of healthcare interventions and policy-making in health. This instrument's uniqueness lies in its simplicity and the ability to assign a single summary value, known as EQ-5D values or values, to each possible health profile, utilising country-specific value sets.

3.2.2 Components of EQ-5D-3L

EQ-5D-3L consists of three primary elements:

1. EQ-5D Descriptive System and EQ-5D-3L Index (236, 237): This encompasses five dimensions of health—mobility, self-care, usual activities, pain/discomfort, and anxiety/depression, measured on a three-level scale (no problems, some problems, and extreme problems). The responses across these dimensions are amalgamated into a 5-digit code, representing the health state of an individual across the five assessed dimensions. To convert this health state code into a single summary index value (EQ-5D-3L Index), a country-specific value set is employed. Each value set contains weights assigned to different levels of severity in each dimension, reflecting the societal preferences for different health states. The index value is calculated by applying these weights to the individual's health state code, often incorporating adjustments for interactions between dimensions (i.e., if the presence of problems in multiple dimensions has a compound effect beyond the sum of individual dimensions' effects).
2. Visual Analog Scale (VAS): The VAS component complements the descriptive system by capturing subjective assessments of overall health. Participants rate their health on a scale from 0 (worst imaginable health) to 100 (best imaginable health), providing an additional perspective on the perceived health-related quality of life (hrQoL).

3.2.3 Uses and Applications

Beyond its use in calculating QALYs for economic health evaluations, EQ-5D-3L's versatility extends to various domains. These include summarising EQ-5D profiles for statistical analysis, describing population health, comparing health across regions or over time, delineating illness severity, and prioritising treatments. Recently, its application has expanded into routine outcome measurements, assessing healthcare performance, and gauging healthcare systems' productivity.

The EQ-5D-3L's major strengths lie in its brevity, generic nature, and the ability to provide a singular summary value for health profiles, enhancing its utility in economic evaluations and policy-making. Additionally, its design allows for broad applications across different health and policy contexts, making it a versatile tool in health assessment.

Despite its strengths, EQ-5D-3L is not without limitations. Its value distributions can exhibit gaps, spikes, or clusters due to ceiling effects and the model's inherent design, potentially complicating data interpretation. Moreover, the three-level scale may not capture the nuances of health states as finely as desired, leading to a potential underestimation of health variations.

For analytical purposes, EQ-5D-3L data often require careful handling due to its specific distribution characteristics, such as skewness and the presence of ceiling effects. Statistical analyses typically involve the use of central tendency measures (mean or median), dispersion measures (SD, IQR), and shape (skewness, kurtosis). However, given the data's unique distribution, median values and non-parametric tests, like the Kruskal-Wallis test, are frequently employed for more accurate analyses. Additionally, graphical illustrations and advanced econometric techniques are utilised to handle heterogeneity and to extrapolate values across the health state spectrum.

The final dataset comprised subjects with EQ-5D-3L assessments at both baseline and 12 months. Prior to analysis, subjects with missing data in any of the dimensions were identified. Incomplete data at either time point were identified and excluded for the paired comparisons. Descriptive analyses were conducted to summarise the EQ-5D-3L Index and VAS scores at baseline and 12 months for both groups. Mean, SD, median, and IQR provided measures of central tendency and dispersion of the data, acknowledging the potential non-normal distribution of EQ-5D-3L values.

3.2.4 Euroqol (EQ5D-3L) Questionnaire

Permission to use the Euroqol (EQ5D-3L) Questionnaire was obtained from the Euroqol Research Foundation on the 21st of April 2021. Terms of use for using the Euroqol (EQ5D-3L has been accepted and approved for use. The questionnaire was submitted for use with the ethics application for the clinical phenotyping study.

At baseline and 12-month visits, the participant will be asked to complete the Euroqol (EQ5D-3L) questionnaire. (260).

3.2.5 Data Collection

Baseline data were collected through structured questionnaires administered in-person or online, depending on participant preference and pandemic-related restrictions. At baseline, demographic and clinical variables, including age, sex, body mass index (BMI), and comorbidities, were recorded. Follow-up data were collected 12 months later using the same structured format. Participants were reminded via email or phone to ensure high follow-up rates.

3.2.6 Data Preparation

Before analysis, data were cleaned and pre-processed to address inconsistencies and missing values. Missing data were managed differently for the full dataset and per protocol analyses. In the full dataset, missing values were assumed to occur at random,

and sensitivity analyses were conducted to assess their potential impact. The per protocol dataset excluded participants with missing data at baseline or follow-up, providing a stricter analysis of those who adhered to the study protocol.

3.2.7 Statistical Analyses

All statistical analyses were conducted using R software (version 4.4.1). Baseline demographic and clinical characteristics were summarised using descriptive statistics. Continuous variables were reported as means with standard deviations, while categorical variables were expressed as frequencies and percentages. Group differences at baseline were assessed using independent t-tests for continuous variables and chi-square tests for categorical variables. Fisher's exact test was applied when expected cell counts were small.

For longitudinal analyses, changes in EQ5D-VAS, EQ5D-Index, and EQ5DL dimensions from baseline to 12 months were analysed using mixed-effects models. These models included a fixed effect for SARS-CoV-2 status (positive or negative) and covariates such as age, sex, and BMI. Random intercepts for participant ID were included to account for within-participant correlations. To further evaluate group differences in change scores, linear regression models were used. Models adjusted for baseline scores to account for initial differences in health status. Separate models were run for the full dataset and per protocol dataset to assess robustness and consistency of the findings.

For EQ5DL dimensions, logistic regression was applied to evaluate the likelihood of reporting problems in each dimension at baseline and 12 months. Change scores for each dimension were also analysed using linear regression, adjusting for covariates and baseline values. Sensitivity analyses were conducted to test the robustness of the models to potential confounders.

Missing data were handled using complete case analysis for descriptive and regression analyses. Sensitivity analyses were performed using the per protocol dataset to examine the robustness of the findings.

Analyses were performed using the lme4, broom. mixed, and dplyr packages in R. Visualization of results was conducted using ggplot2. Model assumptions were assessed using residual plots, and multicollinearity among predictors was evaluated using variance inflation factors. Statistical significance was set at $p < 0.05$.

3.3 Methods for Chapter 6: Evaluation of Transformer-Based Counterfactual Estimation of Individual Treatment Effects - Analysis of Angiotensin Converting Enzyme inhibitors (ACEIs) and risk of SARS-CoV-2 infection (Addresses Objective 4)

In this section I shall describe the methods for data extraction, data curation and machine learning estimation of individual treatment effects (ITE) for the studies presented in Chapter 6. The statistical analyses using ITE are presented in Chapter 6.

3.3.1 Data source

The cohort was identified from routinely collected patient episodes occurring between 1 Jan 2014 and 31 December 2022 in NHS GG&C, the largest health board in Scotland, which provides services to 1.2 million people (approximately a quarter of Scotland's population). The data are held in the West of Scotland Safe Haven (a Trusted Research Environment) and record linkage was used to combine data relating to the same patient.

The Prescribing Information System collects data on all dispensed medications prescribed in the community, coded using the British National Formulary. The Scottish Morbidity Record 01 (SMR01) collects data on all hospitalisations including date of admission and diagnoses coded using the International Classification of Diseases 10 (ICD-10). Death certificates provide date and cause of death, also coded using ICD-10. The Scottish Index of Multiple Deprivation (SIMD) uses postcode of residence to derive a measure of area-based deprivation from aggregated data collected in the census on employment, income, health, education, housing, crime, and access to local services. A lower number indicates a higher socioeconomic deprivation. The digital health records of diabetic patients are obtained from the Scottish Care Information - Diabetes (SCI-Diabetes) dataset.

3.3.2 NHS Greater Glasgow and Clyde Safe Haven

NHS GG&C Safe Haven with the Robertson Centre for Biostatistics provides a secure trusted research environment for data analysis and the opportunity to access various linked datasets for patients across NHS GG&C. NHS GG&C Safe Haven specialise in producing custom datasets linking patient cohorts and their routinely collected electronic health record data for clinical service and academic purposes. The NHS acts as the data custodian. The datasets are available for linkage and access via a tier system. Tier 1 datasets are available to all and are national collected datasets which are used in

everyday health care. Tier 2 datasets are locally/regionally generated by services for patient care and their use in research must be approved by the relevant data controller. Tier 3 datasets are NHS GG&C research databases.

For this thesis, I submitted an application to use reference data sets, as with traditional datasets are reviewed by the Safe Haven Local Privacy and Advisory Committee (LPAC). All access is strictly controlled within the Safe Haven Trusted Research Environment hosted at the University of Glasgow. The data that leaves the environment will be subjected to strict statistical controls, designed to make sure there are no patient identifiable information. I have obtained Safe Haven LPAC approval for the use of the datasets for the project.

3.3.3 Description of Selected Cohort

For the NHS GG&C Safe Haven data set, we obtained data of all patients that were admitted to an NHS GG&C hospital or who were tested positive for SARS-CoV-2. The cohorts that were selected were adults over the age of 17 years, admitted to a West of Scotland Hospital between 1 April 2020 and 31 December 2022. The control group were adults over the age of 17 years, had a SARS-CoV-2 RT-PCR test in the community between 1 Jan 2020 and 31 December 2022. Comparison will be made with a cohort of all patients admitted from the 1 Jan 2019 – 31 December 2019.

3.3.4 Description of Derived Variables

In Chapter 6, the study utilised linked data from four databases: the Prescribing Information System (PIS), the Scottish Morbidity Record 01 (SMR01), death certificates, and patient demographic information. The PIS collects data on all medications dispensed in the community, coded according to the British National Formulary (BNF). The SMR01 records data on hospital admissions, including admission dates and diagnoses, which are coded using the International Classification of Diseases, 10th Revision (ICD-10). Death certificates provide information on the date and cause of death, also coded using ICD-10. Demographic data includes patient age, sex, and the Scottish Index of Multiple Deprivation (SIMD). The SIMD is an area-based measure of deprivation derived from aggregated census data on employment, income, health, education, housing, crime, and access to local services. A lower SIMD value indicates a higher level of socioeconomic deprivation.

Data was also gathered from SCI Store, an information repository interfacing with other local systems that contains laboratory data. In addition, data will be collected in greater detail than is available in the Safe Haven, including broader demographic data (e.g.,

occupation), presenting symptoms, clinical observations, drug history, in-hospital prescribing, hospital progress (including critical care data), enrolment in clinical trials, and discharge information. All data will be anonymised, with participants assigned a study ID. A key linking study IDs to CHI numbers will be stored separately in a password-protected format. This data will then be linked to the Safe Haven dataset and returned to researchers in an anonymised form. Once transferred to the Safe Haven, the CHI link will be destroyed. This process was successfully employed in the OBELIX study, and the number of variables collected for LOCHINVAR has been refined (reduced) following a review of the OBELIX data. All analyses will be conducted on anonymised data.

The dataset will include demographic information, data on prevalent cardiovascular disease and other comorbidities, prescribing data, laboratory results, imaging data, discharge outcomes, and vaccination status. Appendix 1 lists the main data extracted from the NHS Greater Glasgow and Clyde (NHS GG&C) Safe Haven dataset for analysis, along with relevant filters and data ranges. Updates to the dataset will occur at 6, 12, 18, and 24 months.

3.3.5 Data Outputs

After our analysis is complete, the requested outputs (tables/figures/scripts) are placed in a specified “outputs” folder which is then reviewed by the NHS GG&C Safe Haven data management team. This review is to ensure that there is no risk of identification of patients (e.g low numbers in specific group or rare condition). Once approved, the output files have been transferred for the purpose of this thesis.

3.3.6 Missing Data

If there were missing data, this was demonstrated in the results tables in proportions and percentages.

3.3.7 Study population

The study population consisted of patients who had a record in the Patient Information System between 1 October 2019 and 1 October 2020. Inclusion was restricted to patients who were 40 years of age and older on 1 October 2019. Patients were excluded if there was no information on their demographical data, specifically of sex, age, and SIMD.

The study population was separated into two subpopulations: One for the first wave of the pandemic (beginning 1 April 2020) and one for the second wave (beginning 1 October 2020). Patients with COVID-19 were allocated to the first-wave or second-wave population based on the first date of their positive COVID-19 diagnosis. Patients without

recorded COVID-19 diagnosis were stratified into the two study populations using a ratio that mirrors the proportion of COVID-19 cases observed in the first and second waves. This stratified sampling technique aims to achieve representativeness of the original dataset's entire data distribution within the newly created sub-datasets.

3.3.8 Exposure

Exposure to each medication of interest was defined as at least one PIS record of the medication being dispensed during two specified time periods: from 1 October 2019 to 1 April 2020 for the first COVID-19 pandemic wave, and from 1 April 2020 to 1 October 2020 for the second pandemic wave. ACE inhibitors (ACEIs), beta-blockers (BBs), calcium channel blockers (CCBs), thiazides and thiazide-like diuretics (THZs), antiplatelets, biguanides, sulfonylureas, and selective serotonin reuptake inhibitors (SSRIs) were identified using their respective British National Formulary (BNF) codes: 0205051, 0204000, 0206020, 0202010, 0209000, 0601022, 0601021, and 0403030. Statins were identified by the following BNF codes: 0212000B0, 0212000C0, 0212000X0, and 0212000Y0.

Exposure was defined as having at least one prescription of any relevant medication dispensed and recorded in the Prescribing Information System. ACEIs, BBs, CCBs, and THZs were identified using their respective BNF codes: 0205051, 0204000, 0206020, and 0202010. While we initially aimed to include angiotensin receptor blockers (ARBs), their small sample size resulted in the machine learning model being unable to fit ARB data optimally.

3.3.9 Follow-up and outcomes

The primary combined endpoint was the first COVID-19-specific hospitalisation or death (ICD-10 codes: U07.1/U07.2) within 180 days from either 1 April 2020 (the beginning of the first wave) or 1 October 2020 (the beginning of the second wave). Individuals were followed until the earliest occurrence of a COVID-19-specific hospitalisation, 31 December 2020 (end of the first wave analysis), 1 April 2021 (end of the second wave analysis), or death.

The machine learning models developed by Tran Quoc Bao Tran are detailed in Appendix 2, with an overview of the methodology illustrated in Figures 7 and 8.

The first pandemic wave began on 1 April 2020, while the second wave started on 1 October 2020, coinciding with the UK national COVID-19 vaccination rollout, initiated on 8 December 2020. Given the differing circumstances—such as non-vaccinated versus

partially vaccinated populations and evolving public health responses—analysing the data for the first and second waves separately allows for a clearer understanding of the impact of medications on the risk of incident COVID-19.

Patients registered with NHS Greater Glasgow and Clyde (NHS GG&C) and alive as of 1 April 2020 were included in the first pandemic wave analysis. To facilitate separate analyses of the two pandemic waves and address class imbalance, undersampling was performed. This process involved removing samples from the COVID-19-negative group, which constituted the majority of the study population during the first wave. A stratified sampling approach was employed, based on age, sex, and SIMD, using a ratio that mirrored the proportion of COVID-19 cases observed in each wave. Patients excluded during undersampling of the first wave were used to form an independent cohort for the second wave analysis.

The outcome measure was incident COVID-19, defined as a composite endpoint comprising the first positive SARS-CoV-2 test result, the first COVID-19-specific hospitalisation, or death attributed to COVID-19 (ICD-10 codes: U07.1/U07.2) within a 180-day period from the start of either wave. Patients were followed until the earliest of these events or the end of the follow-up period (1 October 2020 for the first wave, and 1 April 2021 for the second wave), whichever occurred first.

3.3.10 Average treatment effect and individual treatment effects

This is further explained in Chapter 6.

3.3.11 Ethics

Delegated research ethics approval was granted for linkage to National Health Service (NHS) patient data by the Local Privacy and Advisory Committee at NHS Greater Glasgow and Clyde. Cohorts and de-identified linked data were prepared by the West of Scotland Safe Haven Research Database at NHS Greater Glasgow and Clyde.

Figure 7 Base Model Inputs
This figure demonstrates the base model inputs. The dates of the first COVID pandemic wave is shown as an example and this is repeated for the second wave. There are two data inputs :(1) The study data set which include medications, admissions and commodities which will be flattened into a time series linking consecutive events. Multi-head attention modelling will be applied to efficiently capture simultaneous interactions among various elements of the dataset simultaneously (2) Demographics which include age, sex and Scottish Index of Multiple Deprivation (SIMD). Feed-forward neural network will be applied which acts as a decision-making pathway. Through the machine learning modelling methods, classification of the COVID-19 hospitalisation/mortality in the next 6 months can be obtained. The base model is then trained in Figure 8 multiple times.

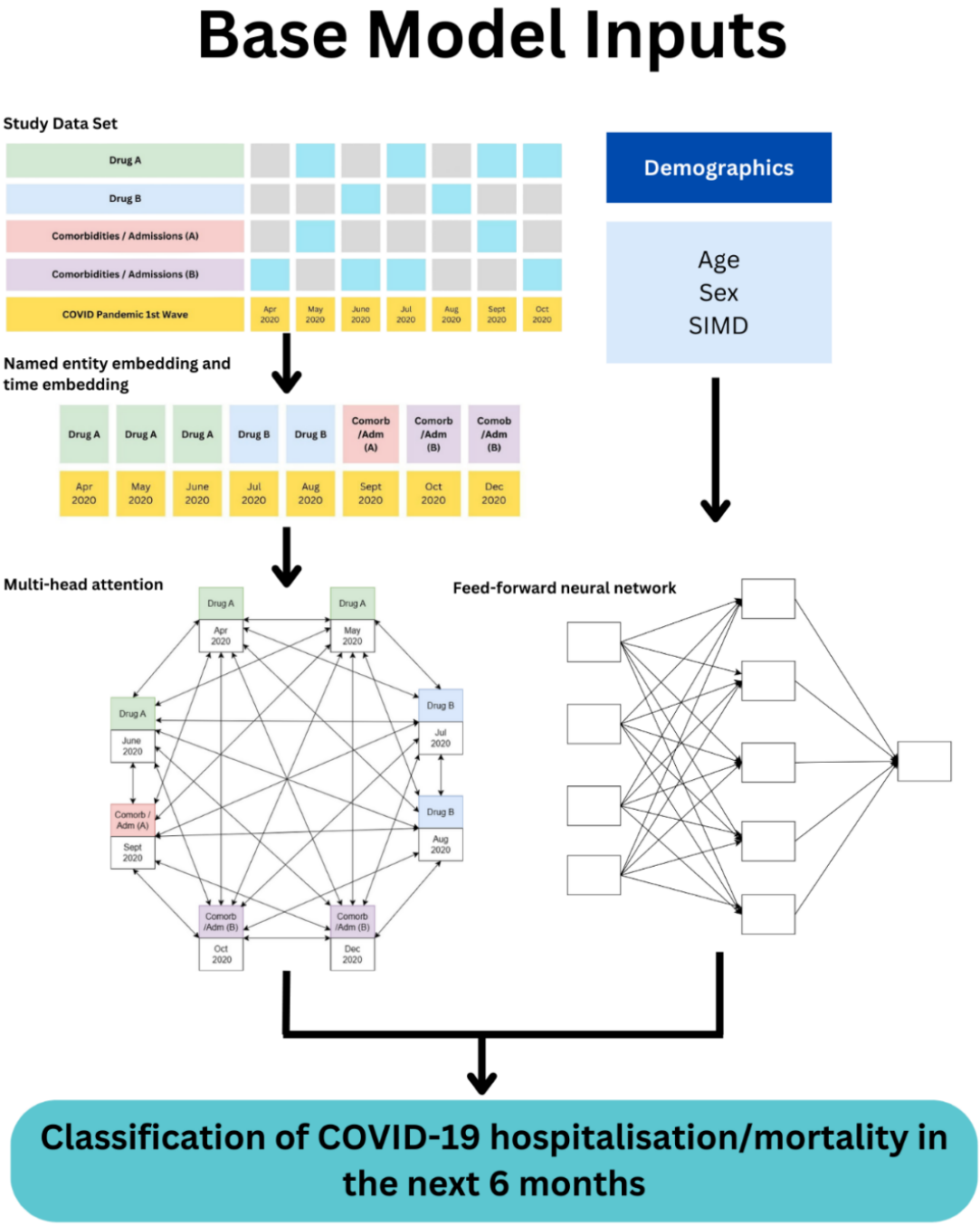
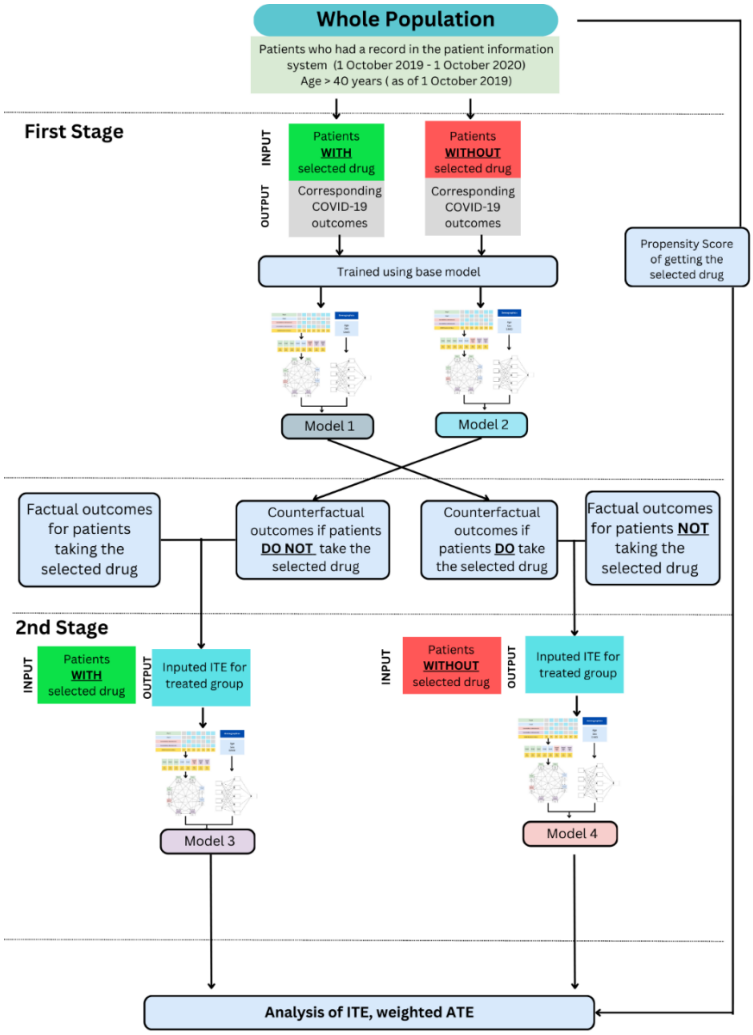


Figure 8 Transformer Model

This figure demonstrates the whole population where patients who had a record in the patient information system from the 1 October 2019 - 1 October 2020, for patients whose age >40 years as of 1 October 2019. The first stage includes study data inputs of selecting patients with the selected drug and patients without the selected drug, to predict their COVID-19 outcomes. The basal model (Figure 7) is then applied to create two distinct models (Model 1 and 2) Model 1 is trained on patients that did not take the selected drug. Model 2 is trained on the patients that did take the selected drug. The two models were subsequently applied to the respective training populations of each other to generate counterfactual outcomes. Specifically, Model 1 was used to predict outcomes on Model 2's training population, and vice versa. The counterfactual outcomes were then compared with the factual outcomes to calculate the imputed ITE for each patient. In the second stage, the ITEs for both counterfactual and factual groups were used to train the base models to develop Model 3 and 4. Combination of Model 3 and Model 4 is then used to provide an analysis of individual treatment effects and weighted ATE. The output of the model undergoes further rigorous statistical analysis. Propensity score matching of the whole population getting the selected drug will be applied to obtain ITE and weighted ATE.



Chapter 4 Clinical Phenotyping Study

4.1 Introduction

This chapter explores the long-term effects of COVID-19 on BP, presenting findings from a clinical phenotyping study aimed at addressing Objective 1 and 2 outlined in Chapter 2.

The LOCHINVAR study, a single-centre clinical investigation, builds upon the findings of the OBELIX pilot study (clinicaltrials.gov NCT04409847).(231) The detailed methods for this chapter is explained in Chapter 3. This study aims to further explore the longitudinal changes in blood pressure post-COVID-19 and the underlying mechanisms thereof. It compares these individuals to a control group without a history of COVID-19 infection, by firstly establishing if COVID-19 results in long-term increased BP and then focus on the possible involvement of the RAAS pathways in any observed BP changes.

4.2 Study Outcomes

4.2.1 Primary outcome

The primary outcome is the average 24-hour ABPM SBP (all day and night) at 12 months in both SARS-CoV-2 positive and SARS-CoV-2 negative groups.

4.2.2 Secondary outcomes

The secondary outcomes are the average of the following measures in both SARS-CoV-2 positive and SARS-CoV-2 negative groups at 12 months: 24-hour ABPM DBP; Day ABPM SBP; Day ABPM DBP; Night ABPM SBP; Night ABPM DBP; 24-hour ABPM heart rate and 24-hour Urine Sodium.

4.2.3 Tertiary outcomes

The tertiary outcomes are the assessment of endothelial function (% FMD), 6MWT, quality of life (QoL), and HBPM metrics at 12 months.

4.3 Ethics approval and study registration

Ethical approval was granted by the West of Scotland Research Ethics Committee 5; 21/WS/0075, Scotland United Kingdom. The study is registered on clinicaltrials.gov; trial identifier NCT05087290 and UK Clinical Research Network; GN20CA501. Current protocol version 1.1 (24/06/2021) Written informed consent will be given by each study participant and participants are allowed to withdraw at any given time.

4.4 Funding

This study is funded by HEART Research UK (Registered Charity, No.1044821, RG2690/21/24)

4.5 Statistical Methods

4.5.1 Sample size calculation

Based on the OBELIX pilot study's findings, we anticipate the ABPM SBP to follow a normal distribution with a standard deviation (SD) of 10mmHg. Considering a clinically significant mean difference of 5mmHg in ABPM SBP between the SARS-CoV-2 positive and negative groups, we estimate that a sample size of 64 participants per group is necessary to achieve 80% power to detect this difference at a 5% significance level. Factoring in an anticipated 15% attrition rate due to missing data and ABPM discrepancies, our recruitment target is set at 150 participants, aiming for 128 completions. This calculation incorporates the recruitment of OBELIX participants into the LOCHINVAR study to fulfil the required sample size.

4.5.2 Statistical analysis

The primary outcome, the average 24-hour ABPM SBP at 12 months, will be analysed using linear regression to compare between SARS-CoV-2 positive and negative groups, adjusting for age, sex, and BMI. We will report the adjusted mean difference with a 95% confidence interval (95% CI) and p-value. Secondary outcomes will employ similar regression models to explore the effect of baseline characteristics and potential interactions with COVID-19 status. Longitudinal changes in ABPM SBP will be assessed using linear mixed models, adjusting for baseline ABPM SBP and other covariates and this will be used for analysis of longitudinal changes for other variables. Group characteristics will be summarised with means ± standard deviation (SD) for continuous variables and frequencies (percentages) for categorical data. Non-normally distributed variables (renin, NT-pro-BNP, Ang II [1-8], Ang [1-7], Ang I [1-10], Ang [1-5]) will undergo log transformation for analysis. Aldosterone will be analysed as a binary variable, split at the local NHS laboratory threshold, to address its non-normal distribution. Group comparisons will utilise independent t-tests or chi-square tests as appropriate. Linear regression will explore the association between COVID-19 status and various outcomes, progressively adjusting for demographics and clinical measurements. Significance is set at $p < 0.05$. We applied the Bonferroni correction to account for multiple testing, setting a threshold for significance to minimise over-interpretation and reduce the risk of missing

true results. Given the sample size, this approach was deemed appropriate to identify potential signals warranting further investigation.

All statistical analyses will be performed using R Software version 4.3.3.

4.5.3 Adapting analytic strategies to the changing landscape of COVID-19

The evolving nature of the COVID-19 pandemic introduced complexities into our study design and participant recruitment, necessitating a revised approach to our analysis to maintain the integrity and relevance of our findings. The OBELIX study enrolled participants from 23 July 2020 to 10 December 2020, all of whom had been hospitalised due to COVID-19 or with a non-COVID-19 diagnosis between April and September 2020. For the LOCHINVAR study, recruitment began with the first participant on 16 November 2021 and concluded with the enrolment of the last participant on 17 November 2022.

4.5.4 Strategy and Justification

The LOCHINVAR study was conceptualised to expand on the OBELIX study's initial observations. However, the dynamic landscape of the COVID-19 pandemic, marked by the emergence of new viral variants, changes in public health policies, the introduction of vaccines, and changes in public risk perception about COVID-19 led to a heterogeneity in the participant population that was not initially anticipated. To address these challenges, we bifurcated our analysis into two distinct datasets: the full dataset analysis and the per-protocol analysis. This is demonstrated graphically in Figure 9.

4.5.5 Full Dataset Analysis

The full dataset incorporates all participants from the OBELIX study who consented to continue their participation in the LOCHINVAR study and new recruits to the LOCHINVAR study (83 participants), totalling 97 participants. This dataset aims to provide a comprehensive overview of our research population, encompassing a broad spectrum of COVID-19 experiences and outcomes.

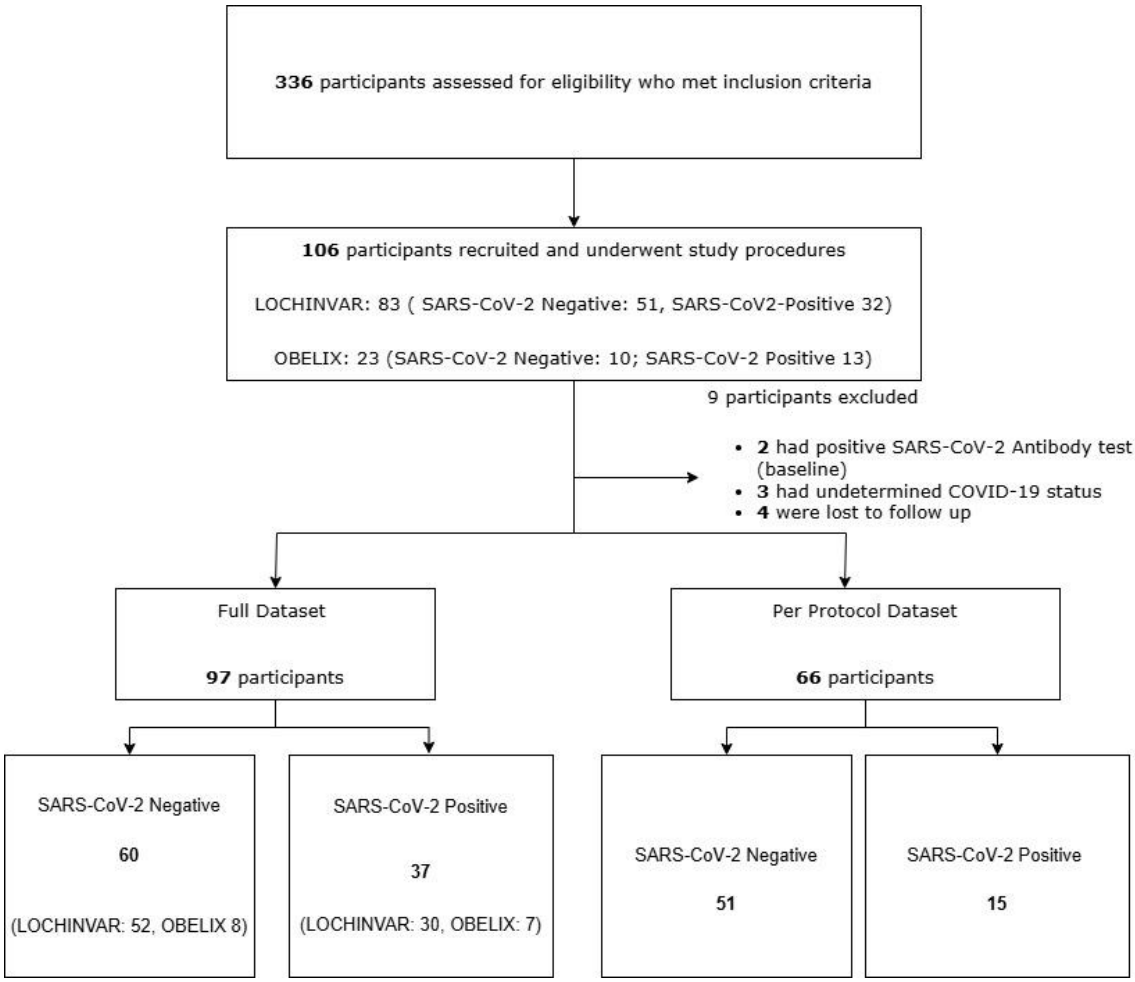
4.5.6 Per-Protocol Analysis

The per-protocol analysis is more stringent, including only those participants who met specific eligibility criteria, such as a baseline visit timeframe of 41-200 days post-COVID-19 infection, with all baseline visits conducted after the year 2020, totalling 66 participants. The baseline visit timeframe was limited to 41-200 days due to significant delays in scheduling participants for their initial visits within a consistent period. This restriction was necessary for standardising the estimation of BP changes. Participants who enrolled too

soon after infection might not have fully recovered, making it premature to assess whether any observed BP variations are indeed long-term effects of COVID-19. Conversely, participants joining the study more than six months post-infection might have BP influenced by factors unrelated to COVID-19, complicating the analysis of the virus's direct impact. This dataset is designed to minimise the variability introduced by the evolving pandemic context, allowing for a more controlled examination of the long-term effects of COVID-19 on BP.

Figure 9 LOCHINVAR study flow diagram

This figure outlines participant flow in the LOCHINVAR study, including eligibility assessment, exclusions, and study procedures. It highlights SARS-CoV-2 status (negative or positive) in the full and per-protocol datasets.



The decision to split the analysis may draw criticism for potentially diluting the statistical power of our findings. Critics may argue that separating the participants into two groups could introduce bias or limit the ability to generalise the results to the broader population affected by COVID-19. Furthermore, the per-protocol analysis, by its nature, excludes a portion of the study population, which may lead to questions about the representativeness

of its findings. Despite these critiques, the bifurcated approach was deemed necessary to account for the substantial heterogeneity introduced by the intervening variables since the onset of the pandemic. The full dataset analysis allows us to capture a broad snapshot of the pandemic's impact, while the per-protocol analysis provides a more controlled environment to discern the specific effects of COVID-19 infection on BP, accounting for variables such as vaccination status, variant pathogenicity, and changing public health guidelines.

The rationale for adopting this dual-analysis strategy lies in its ability to balance comprehensiveness with specificity. By analysing the full dataset, we can ensure that our findings are inclusive of the diverse experiences of our participants. Conversely, the per-protocol analysis allows for a focused examination of the pandemic's impact under more controlled conditions, enhancing the robustness of our conclusions regarding COVID-19's long-term effects on BP. Importantly, this analytical strategy and the statistical analysis plan were established prior to the commencement of any data analysis. This pre-emptive planning was critical to avoid post-hoc adjustments that could introduce bias or undermine the study's integrity. The bifurcated approach to analysing the LOCHINVAR study data was a strategic decision made in response to the unprecedented and evolving challenges posed by the COVID-19 pandemic. It represents a balanced effort to ensure that our findings are both comprehensive and methodologically sound, providing valuable insights into the long-term BP consequences of COVID-19 infection.

4.6 Results

In the LOCHINVAR study, participant demographics were analysed both as a full dataset and a per-protocol dataset, the latter defined by specific criteria due to the dynamic landscape of COVID-19 and vaccination roll out. This has been described previously in the Methods Section 4.5.3: Adapting analytic strategies to the changing landscape of COVID-19. This chapter will explain results as per full data set and per-protocol dataset.

4.6.1 Baseline characteristics (Full Data Set and Per-protocol Data Set)

There were 97 participants included in the full dataset and 66 included in the per-protocol dataset. Table 8 presents the baseline characteristics of the full and per-protocol datasets, demonstrating that both datasets are matched in terms of demographics and laboratory features. This provides reassurance that there are no obvious systematic biases introduced in the selection of the per-protocol dataset.

The participants in SARS-CoV-2 negative and SARS-CoV-2 positive groups in both datasets were of similar age range group 49.5 (IQR: 41.0-54.0) years vs 48.0 (IQR: 44.0-54.0) years. The majority were females in both groups (SARS-CoV-2 positive group: 48 (80%) vs SARS-CoV-2 negative group: 19 (51%)) and this difference was statistically significant (p=0.006) within groups. The major ethnic group was white Caucasian in both groups. The majority were not smokers, had no history of diabetes and admitted to consuming 1-14 units of alcohol a week. There was no significant difference in BMI between the groups (SARS-CoV-2 positive group: 28.2 (SD: 4.7) vs SARS-CoV-2 negative group: 26.6 (SD: 5.0)). Office SBP and oDBP were also similar between the groups. ABPM SBP were 114.9 mmHg (SD: 9.8) and 119.4 mmHg (SD:12.5) for SARS-CoV-2 negative group and positive group respectively, p=0.052. ABPM DBP was similar between the groups as were day and night ABPM SBP and DBP. Haematology laboratory parameters (Hb, WBC, neutrophils, lymphocytes) were not statistically different between the groups while some biochemistry parameters demonstrated significant differences (Serum sodium (Na), urea, creatinine, urine ratios – urine sodium creatinine (uNaCr), urine potassium creatinine (uKCr) and urine chloride creatinine (uClCr). Renin, aldosterone and RAAS Fingerprinting parameters (Ang II [1-8], Ang 1-7, Ang 1(1-10), Ang 1-5, AA2-Ratio, PRA-S, ACE-S and log of each parameter was not significantly different between the groups. The SARS-CoV-2 positive group had a lower FMD median (IQR): 4.5 (2.4-6.5) compared to SARS-CoV-2 negative participants which was 4.7 (2.0-8.5) however this did not attain statistical significance (p=0.461). Both the SARS-CoV-2 negative and positive participants had slightly different waking distances at 6 minutes (Median: 654 (IQR: 570.0-750.0) vs 624.0 (544.5 to 727.5)).

4.6.2 12 months characteristics (Full Data Set and Per-Protocol Data Set)

Table 9 presents the characteristics of participants who attended their 12-month visit including both the full and per-protocol datasets. At 12 months, the participants in SARS-CoV-2 negative and SARS-CoV-2 positive groups were similar in terms of age, gender, ethnicity, smoking status, diabetes status, alcohol intake and BMI. Both oSBP and oDBP at 12 months were similar between the groups for both datasets at 12 months. ABPM SBP 12 months were 114.7 mmHg (SD: 9.5) and 121.6 mmHg (SD: 10.8) for SARS-CoV-2 negative and positive group respectively, p=0.007. ABPM DBP, day ABPM SBP and DBP, and night ABPM DBP and SBP were similar between the groups at 12 months. Haematology laboratory parameters (Hb, WBC, neutrophils, lymphocytes) were not statistically different between the groups. Serum urea, serum creatinine, urine ratios –

urine NaCr and urine ClCr were the only biochemistry parameters which demonstrated significant differences. At 12 months, the SARS-CoV-2 virus positive group demonstrated a lower %FMD 2.4 (1.1-4.0) compared to SARS-CoV-2 negative group 5.0 (2.6-7.4) which attained statistical significance p=0.002. Both the SARS-CoV-2 negative and positive group, had higher walking distances at 6 minutes compared to baseline with the positive group (897.0 metres (816.0 to 1011.0)) having a higher walking distance at 12 months.

4.6.3 Multivariable linear regression at baseline and 12 month visits for full dataset and per-protocol dataset.

Multivariable linear regression analysis was conducted to assess the difference between the groups on a range of health outcomes after adjusting for relevant confounders, at baseline and at 12-month for both datasets (full dataset and per-protocol dataset). The findings from the univariable and multivariable analyses of the full and per-protocol data sets are detailed in Table 10 and Table 11 respectively.

Primary Outcomes

Ambulatory Blood Pressure Monitoring (ABPM)

At baseline and 12 months, ABPM SBP and ABPM DBP did not reach statistical significance in either dataset. In the full dataset, both ABPM SBP (1.41 [-2.88 to 5.70], p=0.515) and ABPM DBP (0.87 [-2.00 to 3.74], p=0.549) showed no significant differences. Similarly, in the per-protocol dataset, ABPM SBP (-1.15 [-6.96 to 4.66], p=0.694) and ABPM DBP (-2.54 [-6.33 to 1.25], p=0.185) also did not achieve statistical significance.

Secondary Outcomes

Brachial Flow Mediated Dilation (FMD)

At baseline in the full dataset, flow mediated dilation was not different between the groups across all models (-5.78 (-14.38 to 2.82, p=0.185)) and this was mirrored in the per-protocol dataset. At 12 months both the full and the per-protocol datasets showed significantly lower %FMD between groups (Full: -1.98 (-3.92 to -0.04, p=0.046); Per-protocol: -2.90 (-5.49 to -0.31, p=0.029)).

6-Minute Walk Test (6MWT)

At baseline, there were no significant differences in the 6MWT distances, suggesting preserved physical function at this point (-16.19 (-88.57 to 56.19, p=0.657)) which was similar in both datasets. At 12 months in the full dataset, the 6MWT distance was similar

between both datasets in the fully adjusted model (126.39 (18.64 to 234.14, p=0.022)) and reached significance.

Office Blood Pressure (Office Systolic Blood Pressure and Office Diastolic Blood Pressure)

Office systolic blood pressure (oSBP) was comparable between the two groups across all models, with no statistically significant differences observed (5.73 [-1.15 to 12.62], p=0.101). Similarly, office diastolic blood pressure (oDBP) showed no significant differences. A marginally higher value in the unadjusted model (3.87 [-0.26 to 7.99], p=0.066) disappeared after adjustment for covariates (2.25 [-1.87 to 6.37], p=0.281). Both oSBP and oDBP were also comparable between groups in the per-protocol dataset.

Laboratory tests

All haematological markers ((Haemoglobin (Hb), white blood cells (WBC), Neutrophils, Lymphocytes) showed no differences between both datasets at baseline and 12 months. For electrolytes, sodium levels in the full dataset were significantly higher at baseline (1.22 (0.44 to 2.00, p=0.002)) and this was similar in the per-protocol dataset. However, at 12 months, sodium, potassium and chloride were similarly observed in both the full and per-protocol dataset. At baseline, urea levels, log-transformed to account for non-normal distribution, was significantly higher in the univariable model (0.17 (0.06 to 0.29, p=0.003)) and the significance persisted in the multivariable model (0.12 (0.01 to 0.24, p=0.037)). However, this lost significance in the per-protocol dataset. Creatinine levels showed a higher level in the univariable model (8.11 (2.03 to 14.19, p=0.009)), but this disappeared in the multivariable model 1.48 (-3.41 to 6.36, p=0.550). This was similar in the per-protocol dataset. Urea levels at 12 months demonstrated similar patterns to baseline as described (univariable model: 0.19 (0.06 to 0.32, p=0.004) multivariable model: 0.09 (-0.05 to 0.22, p=0.215) in the full data set. This was similarly observed in the per-protocol dataset. Creatinine at 12 months demonstrated similar patterns to baseline (univariable model: 8.11 (2.03 to 14.19, p=0.009), multivariable model: 1.48 (-3.41 to 6.36, p=0.550) which was similarly observed in the per-protocol dataset at 12 months. For NT-proBNP (log-transformed), glucose, HbA1c, cholesterol, triglycerides and HDL showed no significant differences between groups across all models for both datasets. This finding was the same for urinary markers uNaCr , uKCr, and uClCr.

Table 8 Baseline demographics for full and per-protocol data set

This table presents the baseline demographics and clinical characteristics of the participants in the full dataset and per-protocol dataset stratified by SARS-CoV-2 status (positive or negative). Continuous variables are reported as mean (standard deviation), median (IQR) if data is skewed and does not follow a normal distribution, and categorical variables are presented as frequency (percentage). P-values indicate group differences assessed using independent t-tests for continuous variables and chi-square tests for categorical variables.

		Full Dataset (n=97)			Per-Protocol Dataset (n=66)		
Label (Baseline)	**levels**	**SARS-CoV-2 Neg (n = 60)**	**SARS-CoV-2 Pos (n = 37)**	**p**	**SARS-CoV-2 Neg (n = 51)**	**SARS-CoV-2 Pos (n = 15)**	**p**
Age (years)	Median (IQR)	49.5 (41.0 to 54.0)	48.0 (44.0 to 54.0)	0.680	50.0 (42.0 to 54.0)	49.0 (43.0 to 53.5)	0.939
Sex (n,%)	Female	48 (80.0)	19 (51.4)	0.006	42 (82.4)	6 (40.0)	0.003
	Male	12 (20.0)	18 (48.6)		9 (17.6)	9 (60.0)	
Ethnicity (n,%)	Asian Indian	1 (1.7)	1 (2.7)	0.196			
	Asian Pakistani	1 (1.7)	2 (5.4)			1 (6.7)	
	Caucasian	58 (96.7)	32 (86.5)				
	Afro-Carribean		2 (5.4)				
Smoking (n,%)	Ever Smoker	6 (10.0)	7 (18.9)	0.233	3 (5.9)	1 (6.7)	1.000
	Never Smoker	54 (90.0)	30 (81.1)		48 (94.1)	14 (93.3)	
Alcohol (n,%)	0	28 (46.7)	17 (45.9)	0.712	24 (47.1)	7 (46.7)	1.000
	1-14	30 (50.0)	20 (54.1)		25 (49.0)	8 (53.3)	
	14	2 (3.3)			2 (3.9)		
Hospital SBP (mmHg)	Mean (SD)	135.1 (18.0)	122.8 (16.8)	0.058	156.0 (NA)	123.7 (12.0)	0.022
Hospital DBP (mmHg)	Mean (SD)	75.9 (14.7)	76.0 (11.4)	0.985	107.0 (NA)	76.0 (8.5)	0.004
Year of Visit (n, %)	2020	7 (11.7)	7 (18.9)	0.112			
	2021	12 (20.0)	2 (5.4)				

	2022	41 (68.3)	28 (75.7)		39 (76.5)	14 (93.3)	
Days to Visit (days)	Median (IQR)	146.0 (120.8 to 167.0)	190.0 (145.0 to 309.0)	<0.001	147.0 (127.0 to 167.0)	147.0 (143.0 to 173.0)	0.270
Visit Interval (days)	Median (IQR)	371.5 (365.2 to 386.5)	363.0 (356.0 to 369.0)	0.002	370.0 (365.0 to 380.0)	358.5 (356.0 to 366.0)	0.002
Vaccination Status (n, %)	No	2 (3.3)	7 (18.9)	0.013	1 (2.0)	3 (20.0)	0.034
	Yes	58 (96.7)	29 (78.4)		50 (98.0)	12 (80.0)	
	(Missing)	0 (0.0)	1 (2.7)		4 (7.8)	0 (0.0)	
	(Missing)	0 (0.0)	1 (2.7)		4 (7.8)	0 (0.0)	
Height (cm)	Mean (SD)	1.7 (0.1)	1.7 (0.1)	0.115	1.7 (0.1)	1.7 (0.1)	0.075
Weight (kg)	Mean (SD)	73.6 (16.5)	80.8 (16.7)	0.040	73.5 (16.4)	82.1 (17.3)	0.080
BMI (kg/m^2)	Mean (SD)	26.6 (5.0)	28.2 (4.7)	0.129	26.6 (4.9)	27.9 (3.8)	0.325
Office SBP (mmHg)	Mean (SD)	122.1 (12.8)	123.6 (12.9)	0.594	122.6 (13.2)	122.8 (13.1)	0.945
Office DBP (mmHg)	Mean (SD)	75.4 (10.1)	79.3 (9.6)	0.066	74.8 (10.4)	78.7 (5.6)	0.169
Office Heart Rate	Mean (SD)	66.9 (11.8)	69.5 (11.4)	0.290	65.4 (10.2)	69.9 (11.3)	0.155
ABPM SBP (mmHg)	Mean (SD)	114.9 (9.8)	119.4 (12.5)	0.052	114.7 (10.0)	117.8 (12.3)	0.327
ABPM DBP (mmHg)	Mean (SD)	73.1 (6.4)	75.1 (7.3)	0.148	73.3 (6.7)	72.9 (4.6)	0.827
ABPM SBP (day) (mmHg)	Mean (SD)	117.9 (10.1)	121.5 (13.0)	0.127	117.6 (10.3)	121.9 (12.0)	0.177
ABPM DBP day) (mmHg)	Mean (SD)	75.6 (6.9)	76.6 (8.4)	0.514	75.5 (7.2)	76.1 (4.5)	0.760
ABPM SBP (night) (mmHg)	Mean (SD)	104.7 (10.1)	108.6 (14.2)	0.121	104.7 (10.1)	105.3 (14.0)	0.862
ABPM DBP (night) (mmHg)	Mean (SD)	64.6 (6.0)	66.5 (9.4)	0.223	64.9 (6.1)	62.2 (5.8)	0.137
Hb (g/L)	Mean (SD)	135.2 (11.5)	139.6 (13.0)	0.082	134.6 (10.7)	141.8 (12.9)	0.033

WBC (x10^9/L)	Mean (SD)	5.8 (1.4)	5.8 (1.1)	0.912	5.7 (1.3)	5.7 (0.9)	0.980
Neutrophils (x10^9/L)	Mean (SD)	3.5 (1.2)	3.4 (0.9)	0.544	3.4 (1.2)	3.4 (0.6)	0.824
Lymphocytes (x10^9/L)	Mean (SD)	1.7 (0.4)	1.8 (0.6)	0.241	1.7 (0.4)	1.7 (0.5)	0.642
Na (mmol/L)	Mean (SD)	138.8 (1.8)	140.2 (1.6)	<0.001	138.6 (1.8)	140.4 (1.7)	0.001
K (mmol/L)	Mean (SD)	4.2 (0.3)	4.2 (0.3)	0.760	4.2 (0.3)	4.2 (0.2)	0.478
Cl (mmol/L)	Mean (SD)	104.8 (2.0)	104.6 (2.0)	0.704	104.9 (2.0)	104.8 (1.7)	0.911
Urea (mmol/L)	Median (IQR)	3.9 (3.4 to 4.6)	4.7 (4.2 to 5.4)	0.003	3.9 (3.5 to 4.7)	4.9 (4.3 to 5.5)	0.011
Creatinine (µmol/L)	Mean (SD)	65.2 (13.0)	73.3 (17.0)	0.009	65.2 (11.7)	75.7 (18.7)	0.010
Mg (mmol/L)	Mean (SD)	0.8 (0.1)	0.8 (0.1)	0.402	0.8 (0.1)	0.9 (0.1)	0.400
Ca(adj) (mmol/L)	Mean (SD)	2.3 (0.1)	2.3 (0.1)	0.352	2.3 (0.1)	2.3 (0.1)	0.962
Albumin (g/L)	Mean (SD)	40.9 (2.4)	39.2 (6.9)	0.089	40.9 (2.3)	41.1 (1.8)	0.819
Bilirubin (µmol/L)	Median (IQR)	11.5 (9.0 to 14.2)	10.0 (8.0 to 14.0)	0.401	12.0 (9.0 to 14.5)	10.0 (9.0 to 15.0)	0.884
ALT (U/L)	Mean (SD)	20.0 (11.7)	23.6 (12.5)	0.151	18.9 (10.7)	21.7 (11.6)	0.377
Glucose (mmol/L)	Mean (SD)	8.7 (15.6)	13.2 (20.8)	0.241	9.3 (16.7)	20.8 (26.8)	0.052
HbA1C (mmol/mol)	Mean (SD)	34.5 (2.6)	37.1 (5.5)	0.003	34.3 (2.6)	37.1 (2.9)	0.001
Cholesterol	Mean (SD)	3.6 (13.0)	5.3 (1.0)	0.441	3.4 (14.1)	5.2 (0.6)	0.616

(mmol/L)							
Triglyceride (mmol/L)	Median (IQR)	0.9 (0.7 to 1.1)	1.1 (0.8 to 1.4)	0.207	0.8 (0.7 to 1.1)	0.8 (0.8 to 1.2)	0.454
HDL (mmol/L)	Mean (SD)	1.6 (0.5)	1.5 (0.6)	0.502	1.6 (0.5)	1.7 (0.7)	0.908
Renin (mIU/L)	Median (IQR)	18.7 (11.9 to 24.8)	20.8 (13.5 to 28.4)	0.380	18.1 (10.9 to 24.1)	21.4 (14.5 to 31.4)	0.227
Aldosterone (pmol/L)	Median (IQR)	163.5 (130.0 to 246.8)	177.0 (130.0 to 312.2)	0.348	156.5 (130.0 to 246.8)	183.0 (130.0 to 276.0)	0.708
NT-pro-BNP (ng/L)	Median (IQR)	48.0 (34.2 to 84.5)	42.5 (30.2 to 78.5)	0.537	49.0 (37.5 to 80.5)	38.0 (32.0 to 63.5)	0.281
Ang II (1-8) (pmol/L)	Median (IQR)	80.1 (49.4 to 109.5)	65.3 (39.5 to 111.6)	0.491	71.7 (48.2 to 106.0)	70.4 (45.2 to 111.1)	0.941
Ang 1-7 (pmol/L)	Median (IQR)	2.5 (2.5 to 2.5)	2.5 (2.5 to 2.5)	0.927			
Ang I (1-10) (pmol/L)	Median (IQR)	21.1 (14.1 to 32.2)	16.3 (12.3 to 26.7)	0.166	21.0 (14.3 to 30.6)	19.8 (14.5 to 26.0)	0.605
Ang 1-5 (pmol/L)	Median (IQR)	2.6 (1.5 to 3.6)	3.0 (1.5 to 5.0)	0.237	2.4 (1.5 to 3.4)	2.9 (1.5 to 4.3)	0.294
Aldosterone (pmol/L)	Median (IQR)	151.3 (109.1 to 223.2)	181.8 (106.4 to 261.8)	0.467	150.4 (105.5 to 223.3)	147.4 (98.7 to 253.7)	0.824
AA2-Ratio (pmol/L)	Mean (SD)	2.7 (2.2)	3.3 (2.6)	0.279	2.9 (2.3)	3.1 (3.1)	0.743
PRA-S (pmol/L)	Median (IQR)	110.1 (63.9 to 138.0)	86.7 (49.9 to 138.3)	0.375	94.6 (63.6 to 135.6)	92.2 (59.8 to 137.1)	0.915
ACE-S (pmol/L)	Mean (SD)	3.9 (2.1)	4.5 (2.3)	0.209	3.6 (1.3)	4.1 (1.6)	0.212
AngII_1_8 (log)	Mean (SD)	4.3 (0.6)	4.2 (0.7)	0.649	4.3 (0.6)	4.3 (0.6)	0.926
Ang1_7 (log)	Mean (SD)	0.8 (0.2)	0.8 (0.2)	0.848			
AngI_1_10 (log)	Mean (SD)	3.0 (0.7)	2.8 (0.8)	0.154	3.0 (0.6)	2.9 (0.5)	0.477
Ang1_5 (log)	Mean (SD)	0.9 (0.6)	1.1 (0.7)	0.218	0.9 (0.5)	1.0 (0.6)	0.317
PRA-S (log)	Mean (SD)	4.6 (0.6)	4.5 (0.7)	0.490	4.5 (0.6)	4.5 (0.6)	0.845
Aldosterone (log)	Mean (SD)	5.1 (0.5)	5.1 (0.7)	0.552	5.1 (0.6)	5.0 (0.6)	0.585

uNaCr (mmol/L)	Median (IQR)	11.3 (8.7 to 15.6)	11.9 (9.0 to 14.2)	0.922	11.3 (8.7 to 15.6)	9.4 (8.1 to 14.4)	0.712
uNaCr(log)	Mean (SD)	2.5 (0.4)	2.4 (0.3)	0.816	2.5 (0.4)	2.4 (0.4)	0.736
uKCr (mmol/L)	Mean (SD)	7.0 (2.2)	5.6 (2.2)	0.022	7.0 (2.2)	5.6 (2.2)	0.067
uClCr (log)	Mean (SD)	2.5 (0.4)	2.4 (0.3)	0.527	2.5 (0.4)	2.4 (0.5)	0.569
%FMD	Median (IQR)	4.7 (2.0 to 8.5)	4.5 (2.4 to 6.5)	0.461	4.7 (2.0 to 8.5)	4.6 (2.6 to 6.8)	0.662
6MWT Distance (metres)	Median (IQR)	654.0 (570.0 to 750.0)	624.0 (544.5 to 727.5)	0.656	654.0 (570.0 to 750.0)	642.0 (534.0 to 747.0)	0.800

Table 9 12 month Characteristics (Full Data set and Per-Protocol Data Set)

This table presents the 12-month demographics and clinical characteristics of the participants in the full dataset and per-protocol dataset stratified by SARS-CoV-2 status (positive or negative). Continuous variables are reported as mean (standard deviation), median (IQR) if data is skewed and does not follow a normal distribution and categorical variables are presented as frequency (percentage). P-values indicate group differences assessed using independent t-tests for continuous variables and chi-square tests for categorical variables.

		Full Dataset (n=97)			Per-protocol Dataset (n=66)		
Label (12 months)	**Levels**	**SARS-CoV-2 Neg n = 60**	**SARS-CoV-2 Pos n = 37**	**p**	**SARS-CoV-2 Neg n = 51**	**SARS-CoV-2 Pos n = 15**	**p**
Age	Median (IQR)	49.5 (41.0 to 54.0)	48.0 (44.0 to 54.0)	0.680	50.0 (42.0 to 54.0)	49.0 (43.0 to 53.5)	0.939
Sex (n, %)	Female	48 (80.0)	19 (51.4)	0.006	42 (82.4)	6 (40.0)	0.003
	Male	12 (20.0)	18 (48.6)		9 (17.6)	9 (60.0)	
Ethnicity (n, %)	Asian Indian	1 (1.7)	1 (2.7)	0.196			
	Asian Pakistani	1 (1.7)	2 (5.4)			1 (6.7)	
	Caucasian	58 (96.7)	32 (86.5)				
	Afro-Carribean		2 (5.4)				
Smoking (n, %)	Ever Smoker	6 (10.0)	7 (18.9)	0.233	3 (5.9)	1 (6.7)	1.000
	Never Smoker	54 (90.0)	30 (81.1)		48 (94.1)	14 (93.3)	
Alcohol (n, %)	0	28 (46.7)	17 (45.9)	0.712	24 (47.1)	7 (46.7)	1.000
	1-14	30 (50.0)	20 (54.1)		25 (49.0)	8 (53.3)	
	14	2 (3.3)			2 (3.9)		
Days to Visit (n, %)	Median (IQR)	146.0 (120.8 to 167.0)	190.0 (145.0 to 309.0)	<0.001	147.0 (127.0 to 167.0)	147.0 (143.0 to 173.0)	0.270

Visit Interval	Median (IQR)	371.5 (365.2 to 386.5)	363.0 (356.0 to 369.0)	0.002	370.0 (365.0 to 380.0)	358.5 (356.0 to 366.0)	0.002
Vaccination Status (n, %)	No	2 (3.3)	7 (18.9)	0.013	1 (2.0)	3 (20.0)	0.034
	Yes	58 (96.7)	29 (78.4)		50 (98.0)	12 (80.0)	
	(Missing)	0 (0.0)	1 (2.7)				
BMI	Mean (SD)	26.8 (4.9)	27.7 (5.1)	0.432	26.6 (4.5)	27.9 (4.5)	0.343
Office SBP (mmHg)	Mean (SD)	119.9 (12.8)	127.8 (14.6)	0.017	119.5 (13.2)	128.0 (15.0)	0.051
Office DBP (mmHg)	Mean (SD)	76.7 (8.9)	80.5 (8.3)	0.074	76.3 (8.9)	81.2 (6.2)	0.069
ABPM SBP (mmHg)	Mean (SD)	114.7 (9.5)	121.6 (10.8)	0.007	114.2 (9.5)	123.9 (9.6)	0.004
ABPM DBP (mmHg)	Mean (SD)	72.4 (6.8)	75.3 (5.6)	0.072	72.4 (6.9)	76.3 (4.7)	0.085
ABPM SBP(day) (mmHg)	Mean (SD)	117.2 (10.4)	124.5 (11.2)	0.008	116.9 (10.3)	127.5 (10.4)	0.004
ABPM DBP(day) (mmHg)	Mean (SD)	74.2 (7.9)	77.6 (6.2)	0.071	74.5 (7.3)	78.7 (5.9)	0.077
ABPM SBP(night) (mmHg)	Mean (SD)	104.0 (10.6)	110.8 (10.7)	0.012	104.0 (10.7)	112.8 (9.3)	0.015
ABPM DBP(night) (mmHg)	Mean (SD)	64.7 (7.6)	66.5 (6.4)	0.332	65.0 (7.7)	67.5 (4.1)	0.304
Hb (g/L)	Mean (SD)	134.1 (12.0)	140.4 (14.0)	0.044	134.0 (11.6)	141.2 (13.5)	0.060
WBC (x10^9/L)	Mean (SD)	6.2 (1.4)	5.6 (1.3)	0.103	6.2 (1.5)	5.5 (1.2)	0.098
Neutrophils (x10^9/L)	Mean (SD)	3.8 (1.2)	3.3 (1.0)	0.125	3.8 (1.2)	3.3 (0.9)	0.150

Lymphocytes (x10^9/L)	Mean (SD)	1.8 (0.4)	1.7 (0.4)	0.205	1.8 (0.4)	1.6 (0.4)	0.203
Na (mmol/L)	Mean (SD)	138.8 (2.1)	139.6 (1.9)	0.107	138.9 (2.1)	139.5 (1.7)	0.272
K (mmol/L)	Mean (SD)	4.2 (0.3)	4.2 (0.3)	0.441	4.2 (0.3)	4.2 (0.2)	0.486
Cl (mmol/L)	Mean (SD)	104.9 (2.2)	105.2 (1.7)	0.530	105.1 (2.2)	105.2 (1.8)	0.969
Urea (mmol/L)	Median (IQR)	4.1 (3.3 to 4.8)	5.2 (4.1 to 5.5)	0.004	4.1 (3.3 to 4.8)	5.2 (4.2 to 5.4)	0.041
Creatinine (µmol/L)	Mean (SD)	67.2 (13.1)	77.8 (15.3)	0.002	67.7 (11.6)	80.0 (16.3)	0.003
Mg (mmol/L)	Mean (SD)	0.9 (0.1)	0.9 (0.1)	0.302	0.9 (0.1)	0.9 (0.0)	0.983
Ca(adj) (mmol/L)	Mean (SD)	2.3 (0.1)	2.4 (0.1)	0.355	2.3 (0.1)	2.4 (0.1)	0.392
Albumin (mmol/L)	Mean (SD)	39.8 (5.5)	41.2 (2.3)	0.220	39.8 (5.8)	41.8 (1.6)	0.234
Bilirubin (µmol/L)	Median (IQR)	10.0 (8.0 to 12.0)	11.0 (9.0 to 14.0)	0.613	10.0 (8.0 to 12.5)	11.0 (9.0 to 12.0)	0.801
ALT (U/L)	Mean (SD)	20.2 (10.5)	25.7 (13.5)	0.053	19.8 (10.6)	26.5 (16.8)	0.087
Glucose (mmol/L)	Mean (SD)	12.3 (22.1)	19.0 (27.3)	0.251	13.2 (23.4)	20.5 (25.0)	0.331
HbA1C (mmol/mol)	Mean (SD)	35.6 (2.8)	37.5 (5.1)	0.043	35.5 (2.9)	37.5 (2.0)	0.023
Cholesterol (mmol/L)	Mean (SD)	5.4 (1.0)	5.3 (1.0)	0.519	5.4 (1.0)	5.4 (0.8)	0.948
Triglyceride (mmol/L)	Median (IQR)	1.1 (0.8 to 1.5)	1.2 (0.8 to 1.8)	0.177	1.0 (0.8 to 1.3)	1.1 (0.7 to 1.5)	0.583
HDL (mmol/L)	Mean (SD)	1.6 (0.4)	1.4 (0.3)	0.055	1.6 (0.4)	1.4 (0.3)	0.116
Renin (mIU/L)	Median (IQR)	20.6 (11.8 to 30.9)	19.1 (16.3 to 24.3)	0.750	20.6 (11.7 to 29.2)	21.6 (16.4 to 30.3)	0.602
Aldosterone (pmol/L)	Median (IQR)	270.0 (191.2 to 365.2)	298.5 (149.8 to 410.8)	0.627	266.0 (189.8 to 355.8)	255.0 (153.0 to 360.0)	0.939

NT-pro-BNP (ng/L)	Median (IQR)	63.0 (42.0 to 96.0)	49.0 (32.0 to 77.0)	0.104	63.0 (41.0 to 91.0)	38.0 (32.0 to 78.0)	0.251
uNaCr (mmol/L)	Median (IQR)	11.8 (8.9 to 14.7)	8.8 (7.5 to 10.7)	0.007	11.9 (9.0 to 14.8)	8.6 (7.5 to 9.2)	0.014
uKCr (mmol/L)	Mean (SD)	7.2 (3.4)	5.9 (2.0)	0.129	7.2 (3.2)	5.1 (1.6)	0.063
uNaCr(log)	Mean (SD)	2.3 (0.7)	2.1 (0.4)	0.167	2.3 (0.7)	2.0 (0.5)	0.253
uClCr(log)	Mean (SD)	2.5 (0.5)	2.1 (0.4)	0.032	2.4 (0.5)	2.1 (0.4)	0.081
%FMD	Median (IQR)	5.0 (2.6 to 7.4)	2.4 (1.1 to 4.0)	0.002	4.6 (2.6 to 7.0)	1.8 (1.4 to 3.5)	0.009
6MWT Distance (metres)	Median (IQR)	750.0 (672.0 to 912.0)	897.0 (816.0 to 1011.0)	0.015	768.0 (708.0 to 930.0)	900.0 (858.0 to 1002.0)	0.022

RAAS Fingerprinting

RAAS fingerprinting parameters ((renin (log-transformed), aldosterone (categorised and log-transformed), Ang II (1-8) (log-transformed), Ang 1-7 (log-transformed), Ang 1 (1-10) (log-transformed), Ang 1-5 (log-transformed), PRA-S (log-transformed), AA2-Ratio, and ACE-S) showed no differences at baseline. RAAS fingerprinting was not carried out at 12 months as explained in our methods section that this was only conducted at baseline. Renin and aldosterone showed no significant difference between the groups across all models. Given that the RAAS fingerprinting parameters were not significant, further analysis at 12 months was not conducted.

4.6.4 Longitudinal Analysis of Blood Pressure and Cardiometabolic Parameters post-COVID

Longitudinal analysis of BP, cardiometabolic variables compared changes between baseline and 12-month measurements in SARS-CoV-2 negative and positive groups. In these analyses, results are presented for the 12-month measurement after adjustment for covariates including the baseline measurement of the dependent variable. Additionally, paired analyses were conducted for each parameter and presented as box plots with Wilcoxon test results presented for each group separately. Table 12 (full dataset) and Table 13 (per-protocol dataset) demonstrates the multivariable regression analysis of the effect of SARS-CoV-2 status and longitudinal cardiometabolic parameters for 12 months measurements adjusted for baseline after accounting for potential confounders such as age, sex, BMI.

Primary Outcomes

Ambulatory Blood Pressure Monitoring (ABPM)

The clustered boxplots (Figure 10 and Figure 11) depict the distribution of ABPM SBP and DBP at baseline and 12 months across SARS-CoV-2 status for both the full dataset and the per-protocol dataset respectively. Though paired analysis showed higher median values for both ABPM SBP and DBP, these did not reach statistical significance. In the multivariable adjusted analyses, both ABPM SBP and DBP showed directionally consistent point estimates but only the per-protocol dataset showed clear statistical significance for ABPM DBP (4.46 (1.01 to 7.90, p=0.012)) while ABPM SBP 4.57 (-0.04 to 9.18, p=0.052) (Table 13).

Figure 10 ABPM SBP Paired

This figure illustrates the paired ambulatory blood pressure monitoring (ABPM) systolic blood pressure (SBP) for the full dataset (left) and the per-protocol dataset (right). Each graph includes two panels: the left panel displays box plots for the SARS-CoV-2 negative group at baseline and 12 months, while the right panel shows box plots for the SARS-CoV-2 positive group at the same time points. The paired t-test p-values are indicated above the square brackets in each graph.

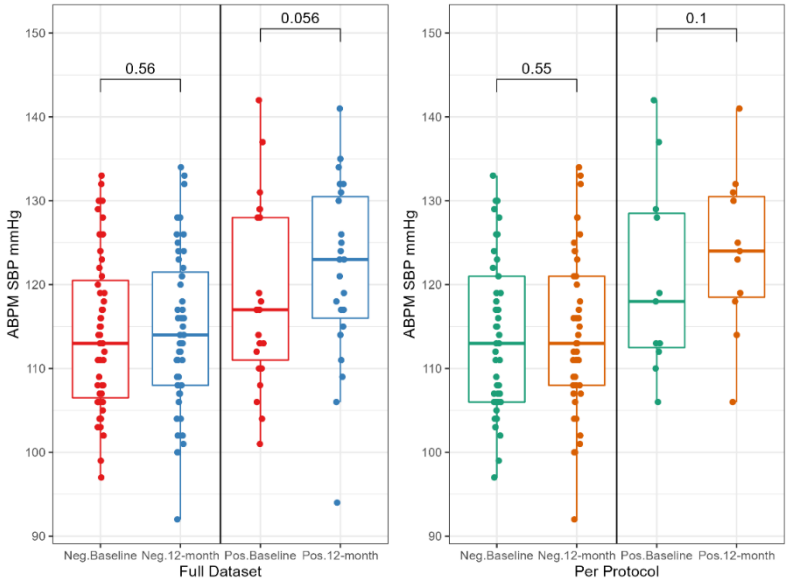


Figure 11 ABPM DBP Paired

This figure illustrates the paired ambulatory blood pressure monitoring (ABPM) diastolic blood pressure (DBP) for the full dataset (left) and the per-protocol dataset (right). Each graph includes two panels: the left panel displays box plots for the SARS-CoV-2 negative group at baseline and 12 months, while the right panel shows box plots for the SARS-CoV-2 positive group at the same time points. The paired t-test p-values are indicated above the square brackets in each graph.

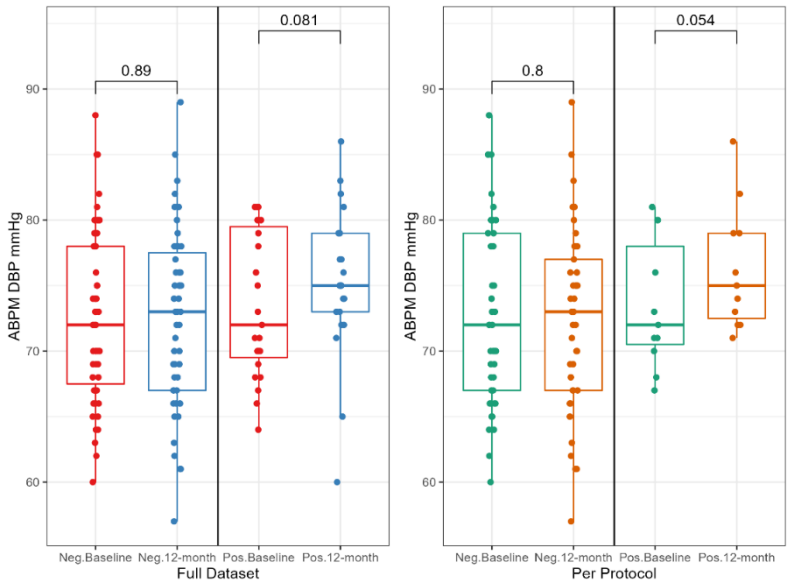


Table 10 Baseline Characteristics Univariable and Multivariable Regression Analyses

This table shows the univariable, and multivariable regression analyses based on dependent variables at baseline for both the full and per-protocol dataset.

	Full Dataset		Per-protocol Dataset	
Dependent	**Univariable**	**Multivariable**	**Univariable**	**Multivariable**
Office SBP (mmHg)	1.43 (-3.88 to 6.75, p=0.594)	-0.46 (-5.81 to 4.88, p=0.864)	0.27 (-7.48 to 8.02, p=0.945)	-1.33 (-9.45 to 6.79, p=0.744)
Office DBP (mmHg)	3.87 (-0.26 to 7.99, p=0.066)	2.25 (-1.87 to 6.37, p=0.281)	3.90 (-1.70 to 9.51, p=0.169)	2.40 (-3.30 to 8.10, p=0.403)
ABPM SBP (mmHg)	4.48 (-0.03 to 9.00, p=0.052)	1.41 (-2.88 to 5.70, p=0.515)	3.05 (-3.13 to 9.24, p=0.327)	-1.15 (-6.96 to 4.66, p=0.694)
ABPM DBP (mmHg)	2.07 (-0.74 to 4.88, p=0.148)	0.87 (-2.00 to 3.74, p=0.549)	-0.41 (-4.12 to 3.31, p=0.827)	-2.54 (-6.33 to 1.25, p=0.185)
ABPM SBP (day) (mmHg)	3.62 (-1.05 to 8.30, p=0.127)	1.06 (-3.47 to 5.59, p=0.644)	4.28 (-1.98 to 10.54, p=0.177)	0.56 (-5.57 to 6.70, p=0.855)
ABPM DBP (day) (mmHg)	1.03 (-2.08 to 4.14, p=0.514)	0.26 (-2.97 to 3.50, p=0.872)	0.60 (-3.32 to 4.53, p=0.760)	-1.06 (-5.23 to 3.11, p=0.613)
ABPM SBP (night) (mmHg)	3.93 (-1.06 to 8.93, p=0.121)	0.72 (-3.96 to 5.40, p=0.761)	0.57 (-5.99 to 7.14, p=0.862)	-4.09 (-10.03 to 1.85, p=0.174)
ABPM DBP (night) (mmHg)	1.94 (-1.20 to 5.08, p=0.223)	1.02 (-2.21 to 4.25, p=0.532)	-2.70 (-6.28 to 0.88, p=0.137)	-4.30 (-8.02 to -0.58, p=0.024)
Hb (g/L)	4.45 (-0.57 to 9.47, p=0.082)	-1.26 (-5.19 to 2.67, p=0.526)	7.21 (0.62 to 13.80, p=0.033)	-0.95 (-6.10 to 4.20, p=0.713)
WBC (x10^9/L)	0.03 (-0.50 to 0.56, p=0.912)	-0.01 (-0.56 to 0.54, p=0.968)	0.01 (-0.68 to 0.70, p=0.980)	0.17 (-0.59 to 0.93, p=0.657)
Neutrophils (x10^9/L)	-0.14 (-0.61 to 0.32, p=0.544)	-0.13 (-0.61 to 0.36, p=0.602)	-0.07 (-0.70 to 0.56, p=0.824)	0.12 (-0.57 to 0.81, p=0.731)
Lymphocytes (x10^9/L)	0.12 (-0.08 to 0.32, p=0.241)	0.09 (-0.11 to 0.30, p=0.372)	0.06 (-0.18 to 0.29, p=0.642)	0.07 (-0.20 to 0.34, p=0.610)
Na (mmol/L)	1.40 (0.66 to 2.13, p<0.001)	1.22 (0.44 to 2.00, p=0.002)	1.81 (0.76 to 2.86, p=0.001)	1.57 (0.41 to 2.74, p=0.009)

K (mmol/L)	-0.02 (-0.13 to 0.09, p=0.760)	-0.03 (-0.15 to 0.08, p=0.550)	-0.05 (-0.20 to 0.09, p=0.478)	-0.05 (-0.21 to 0.11, p=0.518)
Cl (mmol/L)	-0.16 (-1.00 to 0.68, p=0.704)	-0.03 (-0.89 to 0.83, p=0.947)	-0.06 (-1.19 to 1.06, p=0.911)	0.39 (-0.78 to 1.55, p=0.509)
Urea (log)	0.17 (0.06 to 0.29, p=0.003)	0.12 (0.01 to 0.24, p=0.037)	0.20 (0.05 to 0.34, p=0.009)	0.09 (-0.06 to 0.24, p=0.222)
Creatinine (µmol/L)	8.11 (2.03 to 14.19, p=0.009)	1.48 (-3.41 to 6.36, p=0.550)	10.54 (2.61 to 18.46, p=0.010)	0.20 (-5.77 to 6.16, p=0.948)
Ca(adj) (mmol/L)	0.02 (-0.02 to 0.05, p=0.352)	0.01 (-0.02 to 0.05, p=0.506)	0.00 (-0.05 to 0.05, p=0.962)	0.01 (-0.04 to 0.06, p=0.768)
Glucose (mmol/L)	4.45 (-3.05 to 11.95, p=0.241)	0.99 (-6.81 to 8.79, p=0.802)	11.49 (-0.10 to 23.07, p=0.052)	7.04 (-5.43 to 19.51, p=0.263)
HbA1C (mmol/mol)	2.53 (0.87 to 4.20, p=0.003)	1.97 (0.27 to 3.67, p=0.024)	2.79 (1.22 to 4.35, p=0.001)	2.24 (0.68 to 3.80, p=0.006)
Cholesterol (mmol/L)	1.71 (-2.67 to 6.08, p=0.441)	0.67 (-3.94 to 5.27, p=0.775)	1.84 (-5.46 to 9.15, p=0.616)	0.61 (-7.28 to 8.50, p=0.878)
Triglyceride (mmol/L)	0.15 (-0.07 to 0.36, p=0.170)	0.04 (-0.18 to 0.26, p=0.731)	0.08 (-0.16 to 0.31, p=0.526)	-0.04 (-0.27 to 0.20, p=0.749)
HDL (mmol/L)	-0.08 (-0.30 to 0.15, p=0.502)	-0.01 (-0.24 to 0.22, p=0.937)	0.02 (-0.31 to 0.35, p=0.908)	0.12 (-0.24 to 0.47, p=0.524)
Renin(log)	0.11 (-0.15 to 0.37, p=0.398)	0.14 (-0.13 to 0.41, p=0.306)	0.17 (-0.16 to 0.50, p=0.312)	0.15 (-0.20 to 0.50, p=0.407)
Aldosterone(cat)	0.04 (-0.17 to 0.25, p=0.721)	-0.01 (-0.23 to 0.22, p=0.961)	0.03 (-0.27 to 0.33, p=0.824)	0.06 (-0.26 to 0.39, p=0.710)
BNP(log)	-0.01 (-0.31 to 0.29, p=0.935)	0.09 (-0.22 to 0.39, p=0.563)	-0.17 (-0.55 to 0.21, p=0.367)	-0.07 (-0.48 to 0.34, p=0.739)
Ang II (1-8)(log)	-0.06 (-0.35 to 0.22, p=0.649)	0.03 (-0.25 to 0.31, p=0.844)	-0.02 (-0.39 to 0.36, p=0.926)	-0.02 (-0.40 to 0.37, p=0.929)

Ang 1-7(log)	-0.01 (-0.10 to 0.08, p=0.848)	-0.01 (-0.11 to 0.09, p=0.866)	-0.00 (-0.06 to 0.06, p=0.978)	-0.00 (-0.07 to 0.06, p=0.940)
Ang I (1-10)(log)	-0.23 (-0.54 to 0.09, p=0.154)	-0.14 (-0.46 to 0.18, p=0.395)	-0.13 (-0.47 to 0.22, p=0.477)	-0.08 (-0.44 to 0.28, p=0.647)
Ang 1-5(log)	0.17 (-0.10 to 0.43, p=0.218)	0.21 (-0.07 to 0.48, p=0.134)	0.15 (-0.15 to 0.46, p=0.317)	0.05 (-0.25 to 0.36, p=0.730)
PRA-S(log)	-0.10 (-0.37 to 0.18, p=0.490)	-0.00 (-0.28 to 0.27, p=0.972)	-0.04 (-0.40 to 0.32, p=0.845)	-0.03 (-0.39 to 0.34, p=0.889)
Aldosterone(log)	0.08 (-0.19 to 0.35, p=0.552)	0.11 (-0.17 to 0.40, p=0.432)	-0.09 (-0.43 to 0.25, p=0.585)	-0.09 (-0.47 to 0.29, p=0.634)
AA2-Ratio (pmol/L)	0.55 (-0.45 to 1.55, p=0.279)	0.24 (-0.79 to 1.26, p=0.649)	0.25 (-1.26 to 1.76, p=0.743)	0.01 (-1.61 to 1.62, p=0.992)
ACE-S (pmol/L)	0.60 (-0.34 to 1.53, p=0.209)	0.57 (-0.43 to 1.57, p=0.258)	0.50 (-0.30 to 1.30, p=0.212)	0.35 (-0.53 to 1.23, p=0.429)
uNaCr(log)	-0.02 (-0.21 to 0.17, p=0.816)	0.05 (-0.16 to 0.26, p=0.646)	-0.04 (-0.31 to 0.22, p=0.736)	0.08 (-0.19 to 0.35, p=0.555)
uKCr (mmol/L)	-1.35 (-2.49 to -0.20, p=0.022)	-0.17 (-1.31 to 0.96, p=0.761)	-1.41 (-2.93 to 0.10, p=0.067)	-0.53 (-2.02 to 0.96, p=0.477)
uClCr(log)	-0.06 (-0.24 to 0.12, p=0.527)	0.02 (-0.18 to 0.23, p=0.806)	-0.07 (-0.33 to 0.19, p=0.569)	0.06 (-0.21 to 0.33, p=0.648)
%FMD	-6.76 (-14.75 to 1.24, p=0.096)	-5.78 (-14.42 to 2.86, p=0.186)	-6.30 (-17.37 to 4.76, p=0.259)	-4.18 (-16.39 to 8.04, p=0.496)
6MWT Distance (metres)	-10.74 (-77.22 to 55.73, p=0.748)	-16.19 (-88.57 to 56.19, p=0.657)	-10.59 (-98.04 to 76.86, p=0.810)	-27.91 (-121.88 to 66.06, p=0.555)

Table 11 12-month Characteristics Univariable and Multivariable Regression Analyses

This table shows the univariable, and multivariable regression analyses based on dependent variables at 12-months for both the full and per-protocol dataset.

	Full Dataset		Per-protocol Dataset	
Dependent (12 months)	**Univariable**	**Multivariable**	**Univariable**	**Multivariable**
Office SBP (mmHg)	7.91 (1.43 to 14.39, p=0.017)	5.73 (-1.15 to 12.62, p=0.101)	8.50 (-0.02 to 17.02, p=0.051)	4.94 (-4.09 to 13.96, p=0.278)
Office DBP (mmHg)	3.81 (-0.38 to 8.00, p=0.074)	2.33 (-2.26 to 6.92, p=0.315)	4.88 (-0.40 to 10.16, p=0.069)	2.32 (-3.38 to 8.02, p=0.419)
ABPM SBP (mmHg)	6.94 (1.99 to 11.90, p=0.007)	3.86 (-1.47 to 9.19, p=0.153)	9.71 (3.30 to 16.12, p=0.004)	5.38 (-1.96 to 12.71, p=0.147)
ABPM DBP (mmHg)	2.98 (-0.27 to 6.22, p=0.072)	2.22 (-1.51 to 5.96, p=0.239)	3.87 (-0.55 to 8.30, p=0.085)	2.81 (-2.59 to 8.21, p=0.301)
ABPM SBP (day) (mmHg)	7.35 (2.00 to 12.69, p=0.008)	4.18 (-1.68 to 10.04, p=0.159)	10.54 (3.61 to 17.48, p=0.004)	5.39 (-2.61 to 13.39, p=0.182)
ABPM DBP (day) (mmHg)	3.41 (-0.30 to 7.13, p=0.071)	2.79 (-1.53 to 7.10, p=0.202)	4.26 (-0.48 to 9.00, p=0.077)	2.80 (-2.94 to 8.54, p=0.332)
ABPM SBP (night) (mmHg)	6.82 (1.52 to 12.13, p=0.012)	3.47 (-2.40 to 9.33, p=0.242)	8.86 (1.82 to 15.91, p=0.015)	4.40 (-3.73 to 12.52, p=0.283)
ABPM DBP (night) (mmHg)	1.77 (-1.85 to 5.39, p=0.332)	0.36 (-3.81 to 4.54, p=0.863)	2.50 (-2.33 to 7.33, p=0.304)	1.09 (-4.79 to 6.97, p=0.712)
Hb (g/L)	6.28 (0.17 to 12.40, p=0.044)	-3.98 (-8.78 to 0.81, p=0.102)	7.20 (-0.33 to 14.72, p=0.060)	-5.65 (-11.49 to 0.19, p=0.058)
WBC (x10^9/L)	-0.56 (-1.24 to 0.12, p=0.103)	-0.43 (-1.18 to 0.32, p=0.254)	-0.75 (-1.64 to 0.14, p=0.098)	-0.54 (-1.57 to 0.50, p=0.304)
Neutrophils (x10^9/L)	-0.43 (-0.97 to 0.12, p=0.125)	-0.30 (-0.90 to 0.30, p=0.329)	-0.53 (-1.25 to 0.20, p=0.150)	-0.36 (-1.18 to 0.46, p=0.389)
Lymphocytes (x10^9/L)	-0.13 (-0.34 to 0.07, p=0.205)	-0.12 (-0.35 to 0.11, p=0.310)	-0.17 (-0.44 to 0.10, p=0.203)	-0.14 (-0.46 to 0.18, p=0.387)

Na (mmol/L)	0.81 (-0.18 to 1.79, p=0.107)	0.36 (-0.74 to 1.46, p=0.521)	0.69 (-0.55 to 1.93, p=0.272)	0.17 (-1.26 to 1.60, p=0.813)
K (mmol/L)	0.05 (-0.08 to 0.19, p=0.441)	-0.00 (-0.16 to 0.15, p=0.972)	0.06 (-0.11 to 0.22, p=0.486)	-0.02 (-0.21 to 0.16, p=0.814)
Cl (mmol/L)	0.32 (-0.68 to 1.31, p=0.530)	0.20 (-0.93 to 1.34, p=0.723)	0.03 (-1.31 to 1.37, p=0.969)	-0.03 (-1.60 to 1.53, p=0.965)
Urea (log)	0.19 (0.06 to 0.32, p=0.004)	0.09 (-0.05 to 0.22, p=0.215)	0.14 (-0.02 to 0.30, p=0.094)	-0.01 (-0.18 to 0.16, p=0.913)
Creatinine (µmol/L)	8.11 (2.03 to 14.19, p=0.009)	1.48 (-3.41 to 6.36, p=0.550)	10.54 (2.61 to 18.46, p=0.010)	0.20 (-5.77 to 6.16, p=0.948)
Ca(adj) (mmol/L)	0.02 (-0.02 to 0.05, p=0.352)	0.01 (-0.02 to 0.05, p=0.506)	0.00 (-0.05 to 0.05, p=0.962)	0.01 (-0.04 to 0.06, p=0.768)
Glucose (mmol/L)	4.45 (-3.05 to 11.95, p=0.241)	0.99 (-6.81 to 8.79, p=0.802)	11.49 (-0.10 to 23.07, p=0.052)	7.04 (-5.43 to 19.51, p=0.263)
HbA1C (mmol/mol)	2.53 (0.87 to 4.20, p=0.003)	1.97 (0.27 to 3.67, p=0.024)	2.79 (1.22 to 4.35, p=0.001)	2.24 (0.68 to 3.80, p=0.006)
Cholesterol (mmol/L)	1.71 (-2.67 to 6.08, p=0.441)	0.67 (-3.94 to 5.27, p=0.775)	1.84 (-5.46 to 9.15, p=0.616)	0.61 (-7.28 to 8.50, p=0.878)
Triglyceride(log)	0.18 (-0.05 to 0.42, p=0.124)	0.13 (-0.12 to 0.39, p=0.290)	0.13 (-0.17 to 0.42, p=0.398)	0.10 (-0.21 to 0.42, p=0.515)
HDL (mmol/L)	-0.08 (-0.30 to 0.15, p=0.502)	-0.01 (-0.24 to 0.22, p=0.937)	0.02 (-0.31 to 0.35, p=0.908)	0.12 (-0.24 to 0.47, p=0.524)
Renin (log)	0.11 (-0.15 to 0.37, p=0.398)	0.14 (-0.13 to 0.41, p=0.306)	0.17 (-0.16 to 0.50, p=0.312)	0.15 (-0.20 to 0.50, p=0.407)
Aldosterone (cat)	0.04 (-0.17 to 0.25, p=0.721)	-0.01 (-0.23 to 0.22, p=0.961)	0.03 (-0.27 to 0.33, p=0.824)	0.06 (-0.26 to 0.39, p=0.710)
NT-pro-BNP (log)	-0.01 (-0.31 to 0.29, p=0.935)	0.09 (-0.22 to 0.39, p=0.563)	-0.17 (-0.55 to 0.21, p=0.367)	-0.07 (-0.48 to 0.34, p=0.739)

uNaCr(log)	-0.24 (-0.58 to 0.10, p=0.167)	-0.17 (-0.55 to 0.21, p=0.381)	-0.29 (-0.79 to 0.21, p=0.253)	-0.05 (-0.64 to 0.55, p=0.873)
uKCr (mmol/L)	-1.30 (-3.00 to 0.39, p=0.129)	-0.88 (-2.60 to 0.84, p=0.309)	-2.09 (-4.29 to 0.11, p=0.063)	-1.11 (-3.57 to 1.34, p=0.365)
uClCr (log)	-0.31 (-0.59 to -0.03, p=0.032)	-0.25 (-0.56 to 0.06, p=0.106)	-0.34 (-0.73 to 0.04, p=0.081)	-0.13 (-0.57 to 0.32, p=0.572)
%FMD	-2.51 (-4.07 to -0.95, p=0.002)	-1.98 (-3.92 to -0.04, p=0.046)	-2.52 (-4.45 to -0.59, p=0.012)	-2.90 (-5.49 to -0.31, p=0.029)
6MWT Distance (metres)	109.25 (16.98 to 201.52, p=0.021)	126.39 (18.64 to 234.14, p=0.022)	118.45 (21.01 to 215.90, p=0.018)	153.86 (41.29 to 266.44, p=0.008)

Table 12 Longitudinal Regression Analyses - Full Dataset (Coefficient SARS-COV-2 Positive vs SARS-COV-2 Negative)

This table shows the univariable, and multivariable regression analyses for both SARS-CoV-2 status (positive and negative) based on dependent variables at baseline after adjusting for relevant confounders (age, sex, BMI, baseline measure) for both the full dataset.

Full Dataset	Mean (SD)	Multivariable (age, sex, BMI, baseline measure)
Office SBP(mmHg)	127.8 (14.6)	6.30 (0.81 to 11.79, p=0.025)
Office DBP(mmHg)	80.5 (8.3)	1.86 (-1.46 to 5.19, p=0.267)
ABPM SBP(mmHg)	121.6 (10.8)	2.63 (-1.11 to 6.37, p=0.165)
ABPM DBP(mmHg)	75.3 (5.6)	2.30 (-0.39 to 4.99, p=0.092)
Na (mmol/L)	139.6 (1.9)	-0.67 (-1.55 to 0.20, p=0.128)
HbA1c (mmol/mol)	37.5 (5.1)	0.40 (-0.86 to 1.65, p=0.529)
%FMD	2.5 (2.1)	-2.32 (-4.82 to 0.17, p=0.067)
6MWT Distance (metres)	919.6 (138.1)	132.13 (48.79 to 215.48, p=0.002)
Urea (log)	1.6 (0.3)	0.05 (-0.07 to 0.16, p=0.417)
Hb (g/L)	140.4 (14.0)	-1.06 (-4.67 to 2.54, p=0.559)
uNaCr(mmol/L)	9.2 (3.0)	-1.31 (-4.33 to 1.71, p=0.385)
uKCr(mmol/L)	6.2 (2.1)	-0.08 (-1.93 to 1.76, p=0.926)
Creatinine(μmol/L)	77.8 (15.3)	0.15 (-2.95 to 3.24, p=0.926)

Table 13 Longitudinal Regression Analyses - Per-protocol (Coefficient SARS-COV-2 Positive vs SARS-COV-2 Negative)

This table shows the univariable, and multivariable regression analyses for both SARS-CoV-2 status (positive and negative) based on dependent variables at baseline after adjusting for relevant confounders (age, sex, BMI, baseline measure) for both the per-protocol dataset.

Per-protocol Dataset	Mean (SD)	Multivariable (age, sex, BMI, baseline measure)
Office SBP (mmHg)	128.0 (15.0)	5.13 (-1.10 to 11.37, p=0.105)
Office DBP (mmHg)	81.2 (6.2)	0.61 (-3.22 to 4.44, p=0.750)
ABPM SBP (mmHg)	123.9 (9.6)	4.57 (-0.04 to 9.18, p=0.052)
ABPM DBP (mmHg)	76.3 (4.7)	4.46 (1.01 to 7.90, p=0.012)
Na (mmol/L)	139.5 (1.7)	-1.12 (-2.19 to -0.05, p=0.040)
HbA1C (mmol/mol)	37.5 (2.0)	0.95 (-0.76 to 2.66, p=0.271)
%FMD	2.5 (1.8)	-3.15 (-6.33 to 0.04, p=0.053)
6MWT Distance (metres)	925.4 (108.3)	145.60 (49.14 to 242.06, p=0.004)
Urea (log)	1.5 (0.2)	-0.03 (-0.18 to 0.12, p=0.698)
Hb (g/L)	141.2 (13.5)	-3.73 (-8.61 to 1.16, p=0.132)
uNaCr (mmol/L)	8.4 (3.0)	-1.63 (-5.71 to 2.46, p=0.424)
uKCr (mmol/L)	5.0 (1.3)	-0.62 (-3.11 to 1.87, p=0.615)
Creatinine (µmol/L)	80.0 (16.3)	0.15 (-4.01 to 4.30, p=0.944)

4.6.4.1 Secondary Outcomes

Out of all the measurements carried out, we carried out longitudinal analyses on the set of variables that showed nominally significant differences between groups at baseline. The variables analysed are serum sodium, HbA1c, %FMD, 6MWT Distance, oSBP, oDBP, serum urea, haemoglobin, uNaCr, uKCr, and creatinine and the results are presented in Figures (17 to 27) and Tables (12 and 13). Given the multiple tests carried out, a Bonferroni correction was applied to determine significance threshold, and this was set at $p<0.0055$. Only the 6MWT crossed the Bonferroni threshold for significance and this will be presented first.

6-Minute Walk Test (6MWT)

Figure 12 presents the paired t-test comparison showing the distribution across groups for both datasets. There was a significantly higher 6MWT at 12 months ($p <0.0055$) compared to baseline in both the SARS CoV-2 positive and negative groups for both datasets. Table 12 and Table 13 presents the multivariable regression analysis between COVID-19 status and longitudinal 6MWT between baseline and 12 month after accounting for potential confounders such as age, sex, BMI and baseline 6MWT for both datasets. In the full dataset, the multivariable analysis adjusted for potential confounders including baseline 6MWT, indicated a more pronounced increase of 132.13 metres in 6MWT at 12 months from baseline in the SARS CoV-2 positive subjects compared to negative subjects (132.13 (48.79 to 215.48, p=0.002)). (Table 12) This was observed also in the per-protocol dataset 145.60 (49.14 to 242.06, p=0.004) (Table 13).

Office Systolic Blood Pressure and Office Diastolic Blood Pressure

The clustered boxplots (Figure 13 and Figure 14) depict the distribution of oSBP and oDBP at baseline and 12 months across SARS-CoV-2 status for both the full dataset and the per-protocol dataset respectively. The Wilcoxon paired t-test p-values are presented separately for the SARS-CoV-2 positive and negative participants. Paired analysis for oSBP in the per-protocol dataset showed a significant difference (p=0.048) (Figure 13). In the multivariable adjusted analyses, oSBP showed a significant longitudinal increase of 6.3 mmHg increase in the SARS-CoV-2 positive participants compared to SARS-CoV-2 negative participants (6.30 (0.81 to 11.79, p=0.025), this was not evident in the per-protocol dataset although the point estimate of 5.13 was directionally concordant (Table 12 and Table 13). oDBP did not demonstrate any difference in longitudinal trajectory in both datasets.

Figure 12 6-Minute Walk Test

This figure illustrates the paired 6-MWT for the full dataset (left) and the per-protocol dataset (right). Each graph includes two panels: the left panel displays box plots for the SARS-CoV-2 negative group at baseline and 12 months, while the right panel shows box plots for the SARS-CoV-2 positive group at the same time points. The paired t-test p-values are indicated above the square brackets in each graph.

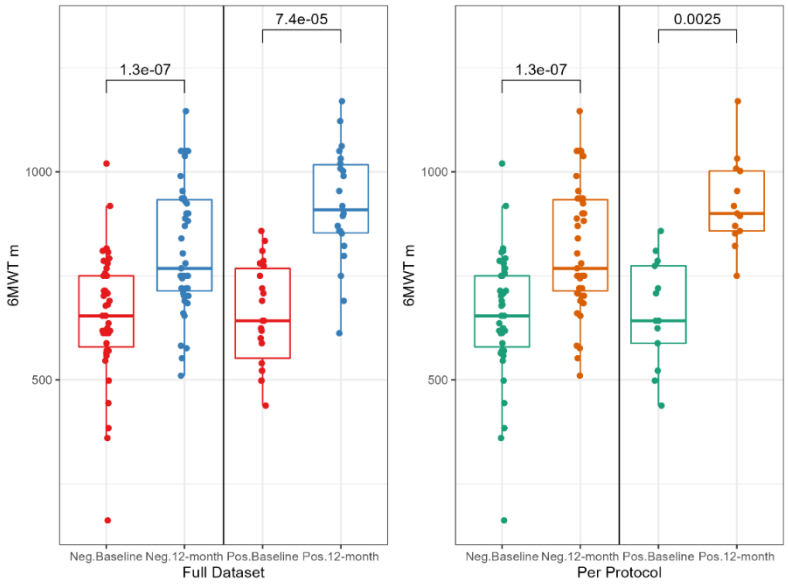


Figure 13 Office SBP Paired

This figure illustrates the paired office SBP for the full dataset (left) and the per-protocol dataset (right). Each graph includes two panels: the left panel displays box plots for the SARS-CoV-2 negative group at baseline and 12 months, while the right panel shows box plots for the SARS-CoV-2 positive group at the same time points. The paired t-test p-values are indicated above the square brackets in each graph.

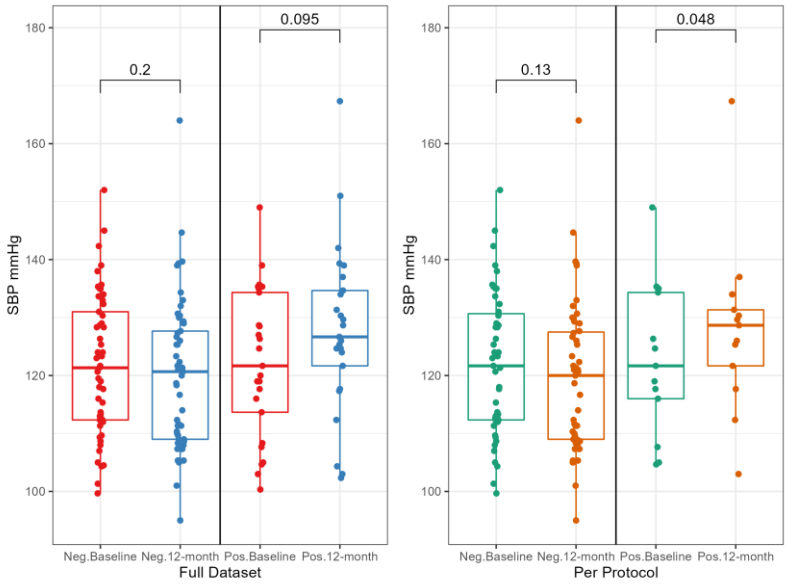
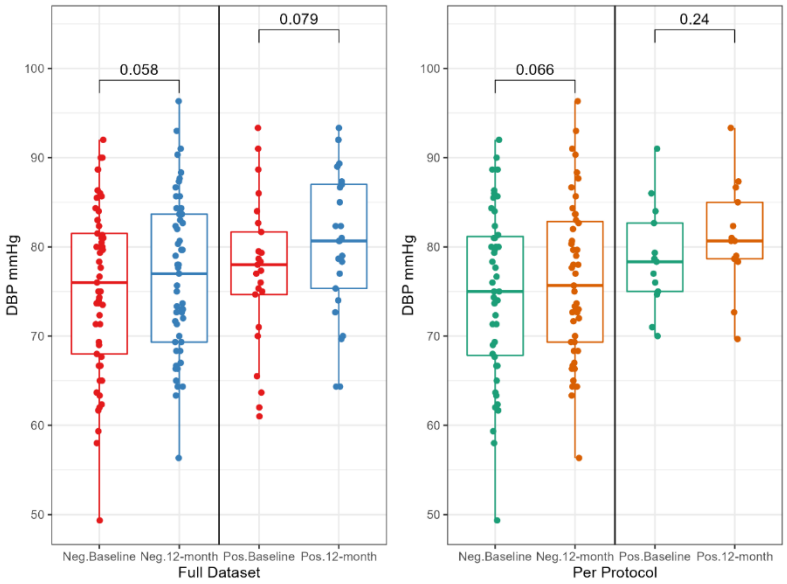


Figure 14 Office DBP Paired

This figure illustrates the paired office DBP for the full dataset (left) and the per-protocol dataset (right). Each graph includes two panels: the left panel displays box plots for the SARS-CoV-2 negative group at baseline and 12 months, while the right panel shows box plots for the SARS-CoV-2 positive group at the same time points. The paired t-test p-values are indicated above the square brackets in each graph.



Brachial Flow Mediated Dilation (FMD)

Though FMD did not cross the Bonferroni threshold for significance, the results are interesting in the context of the significant differences observed in the ABPM data. Figure 15 provide box plots for the full and per-protocol datasets respectively, showing the distribution of FMD across groups. Figure 15 presents the paired t-test comparison, with lower FMD at 12 months compared to baseline in only the SARS CoV-2 positive groups which attain statistical significance in both data sets (Full: p=0.0076; Per-protocol: p=0.039). In the full and per-protocol datasets, multivariable analysis showed a similar point estimate for reduction in FMD at 12 months (Full: -2.32 (-4.82 to 0.17, p=0.067); Per-protocol: -3.15 (-6.33 to 0.04, p=0.053)) (Table 12 and 13).

Serum Sodium

Serum sodium was only nominally significant in the per-protocol multivariable analyses reflecting the results in the paired analysis. Figure 16 provide box plots for the full and per-protocol datasets respectively, showing the distribution of serum sodium across groups. It illustrates the paired t-test comparison, with lower serum sodium at 12 months compared to baseline in the SARS CoV-2 positive group reaching statistical significance (Full: p=0.033, Per-protocol: 0.047) while there was no difference in the SARS-CoV-2 negative group. In the full dataset (Table 12), the multivariable full model adjustment demonstrated

Figure 15 FMD Paired

This figure illustrates the paired FMD for the full dataset (left) and the per-protocol dataset (right). Each graph includes two panels: the left panel displays box plots for the SARS-CoV-2 negative group at baseline and 12 months, while the right panel shows box plots for the SARS-CoV-2 positive group at the same time points. The paired t-test p-values are indicated above the square brackets in each graph.

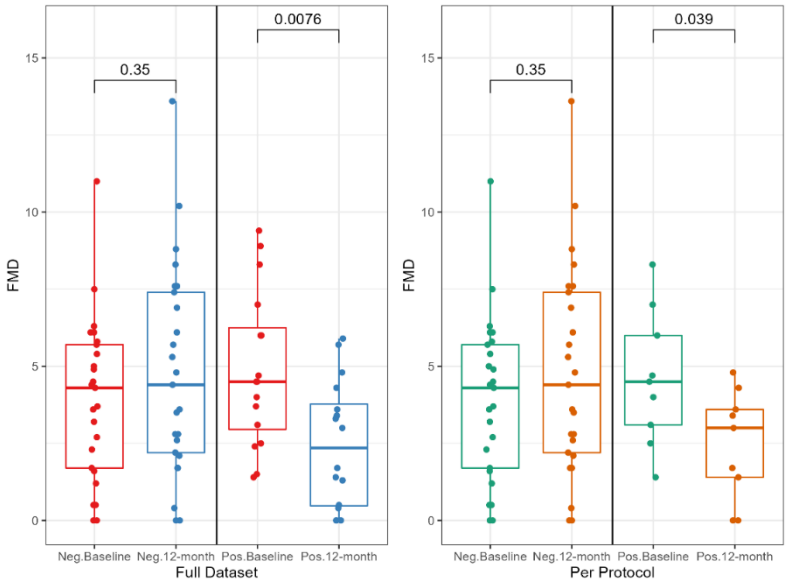
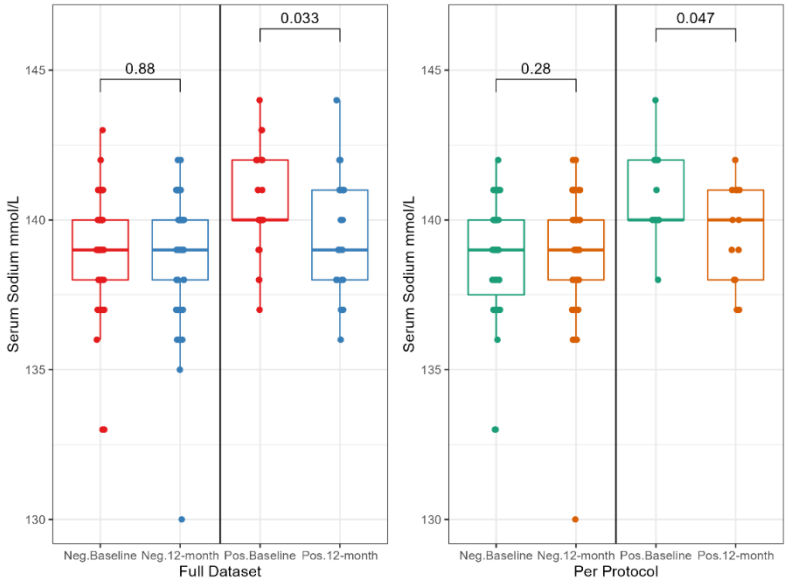


Figure 16 Serum Sodium Paired

This figure illustrates the serum sodium for the full dataset (left) and the per-protocol dataset (right). Each graph includes two panels: the left panel displays box plots for the SARS-CoV-2 negative group at baseline and 12 months, while the right panel shows box plots for the SARS-CoV-2 positive group at the same time points. The paired t-test p-values are indicated above the square brackets in each graph.



a 0.67 mmol/L decrease in serum sodium at 12 months compared to baseline in the SARS CoV-2 positive participants compared to negative participants (-0.67 (-1.55 to 0.20, p=0.128)) (Table 12). In the per-protocol dataset (Table 13), there was a reduction in serum sodium by 1.12 mmol/L at 12 months compared to baseline in the SARS-CoV-2 positive participants (-1.12 (-2.19 to -0.05, p=0.040)).

Serum HbA1c (Glycosylated Haemoglobin)

Figure 17 provide box plots for the full and per-protocol datasets respectively, showing the distribution of serum HbA1c across groups. It illustrates the paired t-test comparison, with higher serum HbA1c at 12 months in both SARS-CoV-2 negative (p=0.00047) and positive (0.011) in the full data set. Similar effects were seen in the per-protocol dataset for both SARS-CoV-2 negative (p=0.00023) and positive (0.041) participants. The multivariable full model adjustment demonstrated a 0.40 mmol/L increase in serum HbA1C at 12 months compared to baseline in the SARS CoV-2 positive subjects compared to negative subjects. (0.40 (-0.86 to 1.65, p=0.529)) in Table 12. This was similarly observed in the per-protocol dataset. 0.95 (-0.76 to 2.66, p=0.271) in Table 13.

Serum Urea (log), creatinine, haemoglobin (Hb), urine NaCr and urine KCr

Paired t-test comparisons for the following variables serum urea (log) (Figure 18), creatinine (Figure 19), haemoglobin (Hb) (Figure 20), urine NaCr (Figure 21) and urine KCr (Figure 22) were not significant in both full and per-protocol datasets. For both univariable and multivariable analysis for both data sets (Table 12 and Table 13), serum urea (log), creatinine, haemoglobin (Hb), urine NaCr and urine KCr were not significant across all models.

4.7 Discussion

4.7.1 Study Findings

In this study, we conducted a comprehensive assessment of longitudinal BP changes after recovery from SARS-CoV-2 infection. This study measured a wide range of clinical (BP, cardiometabolic, and laboratory) parameters at the first visit after recovery from SARS-CoV-2 infection and after 12 months. We compared with a contemporaneous group of subjects who were not exposed to SARS-CoV-2. All subjects did not have hypertension and were at low cardiovascular risk.

Figure 17 Serum HbA1c Paired

This figure illustrates the paired HbA1c for the full dataset (left) and the per-protocol dataset (right). Each graph includes two panels: the left panel displays box plots for the SARS-CoV-2 negative group at baseline and 12 months, while the right panel shows box plots for the SARS-CoV-2 positive group at the same time points. The paired t-test p-values are indicated above the square brackets in each graph.

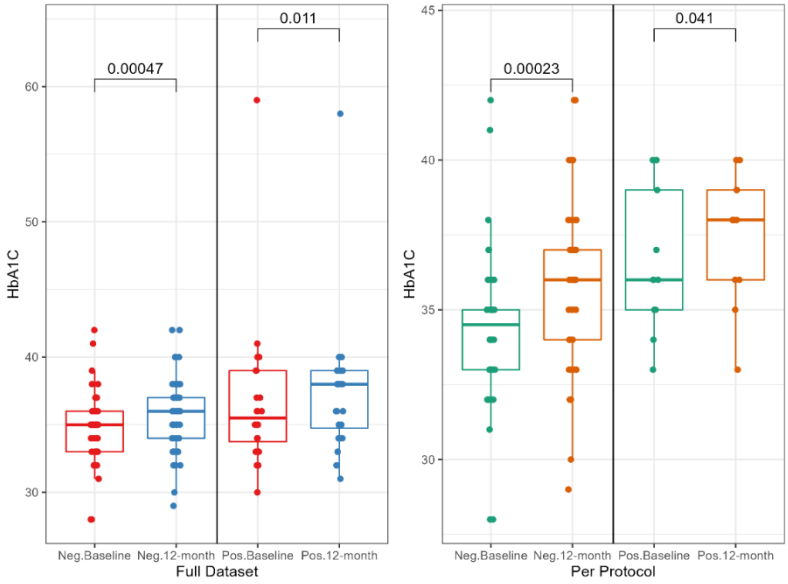


Figure 18 Serum Urea Paired

This figure illustrates the paired serum urea for the full dataset (left) and the per-protocol dataset (right). Each graph includes two panels: the left panel displays box plots for the SARS-CoV-2 negative group at baseline and 12 months, while the right panel shows box plots for the SARS-CoV-2 positive group at the same time points. The paired t-test p-values are indicated above the square brackets in each graph

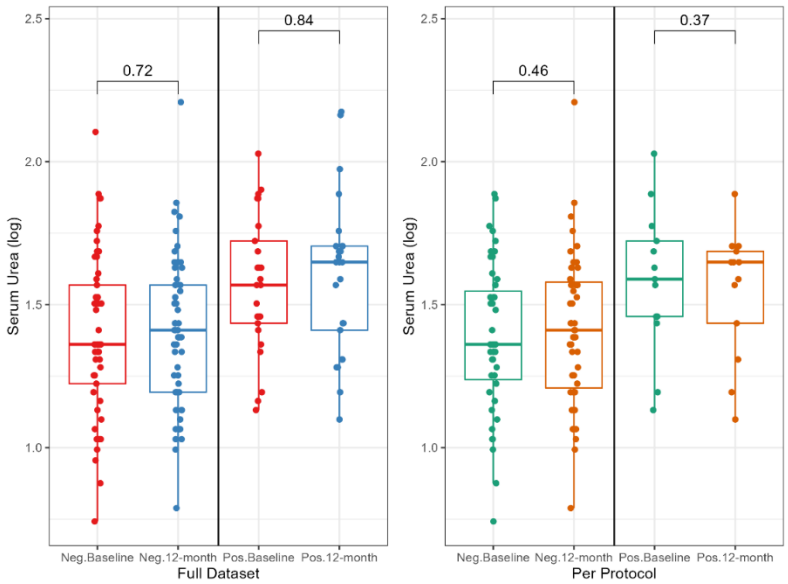


Figure 19 Serum Creatinine Paired

This figure illustrates the paired serum creatinine for the full dataset (left) and the per-protocol dataset (right). Each graph includes two panels: the left panel displays box plots for the SARS-CoV-2 negative group at baseline and 12 months, while the right panel shows box plots for the SARS-CoV-2 positive group at the same time points. The paired t-test p-values are indicated above the square brackets in each graph.

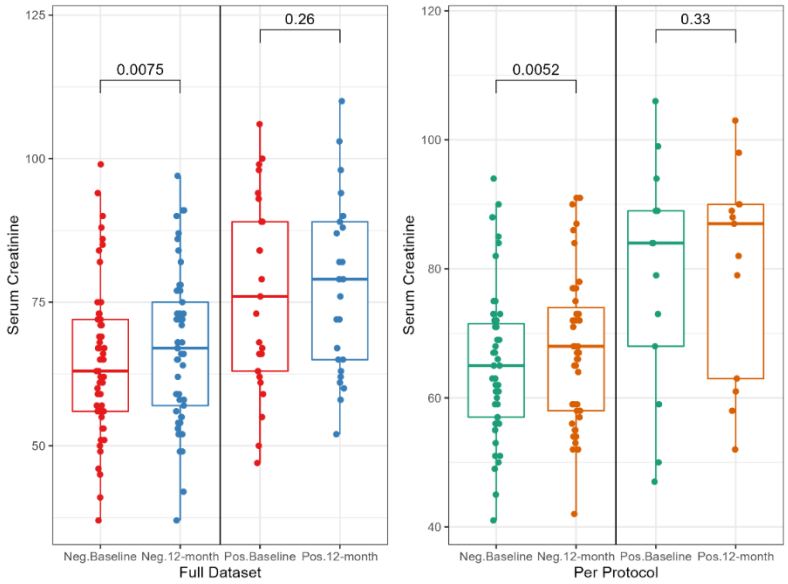


Figure 20 Serum Hb Paired

This figure illustrates the paired serum Hb for the full dataset (left) and the per-protocol dataset (right). Each graph includes two panels: the left panel displays box plots for the SARS-CoV-2 negative group at baseline and 12 months, while the right panel shows box plots for the SARS-CoV-2 positive group at the same time points. The paired t-test p-values are indicated above the square brackets in each graph.

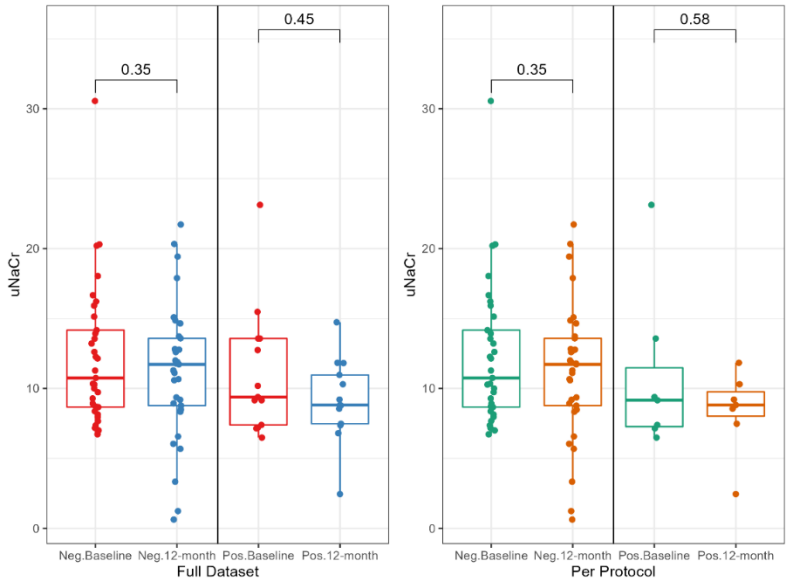


Figure 21 Urine NaCr Paired

This figure illustrates the paired urine NaCr for the full dataset (left) and the per-protocol dataset (right). Each graph includes two panels: the left panel displays box plots for the SARS-CoV-2 negative group at baseline and 12 months, while the right panel shows box plots for the SARS-CoV-2 positive group at the same time points. The paired t-test p-values are indicated above the square brackets in each graph.

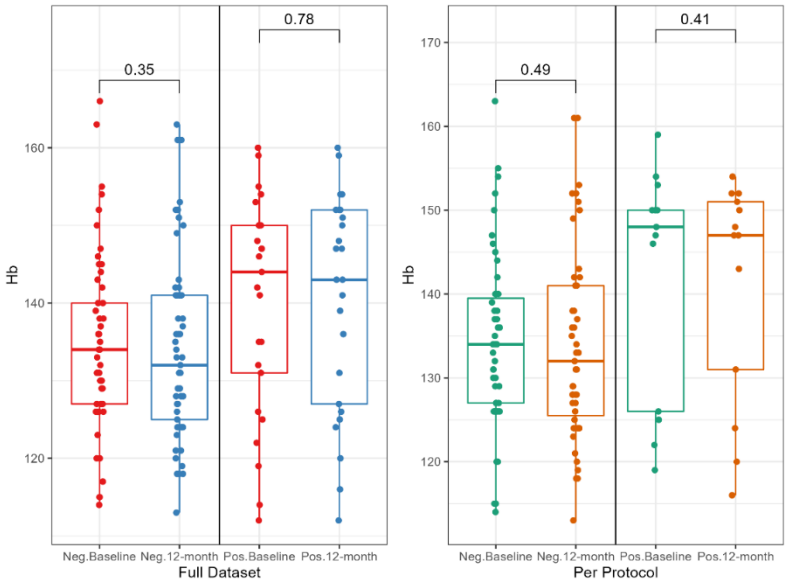
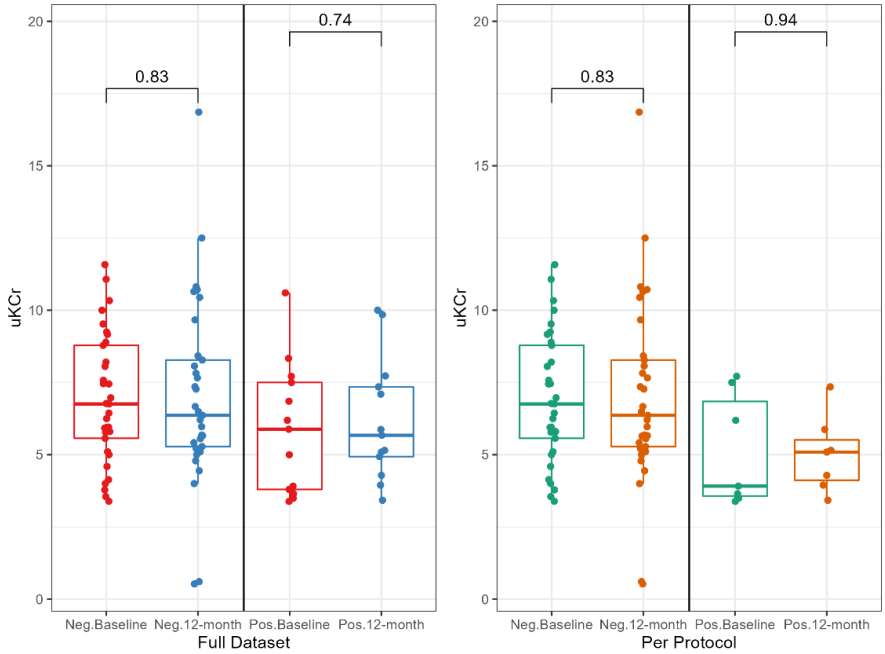


Figure 22 Urine KCr Paired

This figure illustrates the paired urine KCr for the full dataset (left) and the per-protocol dataset (right). Each graph includes two panels: the left panel displays box plots for the SARS-CoV-2 negative group at baseline and 12 months, while the right panel shows box plots for the SARS-CoV-2 positive group at the same time points. The paired t-test p-values are indicated above the square brackets in each graph.



Our study highlights an increase in ambulatory BP, with ABPM SBP rising by 4.57 mmHg and DBP by 4.46 mmHg over a 12-month period post-SARS-CoV-2 recovery compared to SARS-CoV-2 negative counterparts. This rise coincided with a 3.15% reduction in FMD among SARS-CoV-2 positive participants, suggesting a potential vascular basis for the observed BP changes. This study observed a significantly improved 6MWT distance of 145 m, which is an improvement compared to SARS-CoV-2 negative participants, suggesting an overall improvement in overall health in the SARS-CoV-2 positive participants. These results were obtained after adjustment for confounders and were mirrored in the parallel paired analyses.

Cross-sectionally, baseline oBP, ABPM and FMD were matched between the two groups in the full and per-protocol analyses. At 12 months, BP was similar between groups and %FMD was significantly lower in the SARS-CoV-2 positive group in both the full and per-protocol analyses. Our FMD findings are similar to a previous study that compared participants who had COVID-19 infection with normal controls; they found that %FMD was higher in the control group compared to those with COVID-19 at a single time point. (238) A study of 12 female participants (COVID-19 cases) matched with 11 age-matched controls without COVID-19 had FMD and brachial BP measured where they demonstrated higher brachial BP (SBP 126 ± 19 vs.109 ± 8 mmHg; P = 0.010 and lower FMD (cases: 4.69 ± 2.68 vs. control: 5.73 ± 2.69%; P= 0.381) however, this particular study was limited by its sample size. (239) In this study, we did not find significant differences in RAAS fingerprinting results at baseline between the two groups. Thus, we did not repeat the RAAS fingerprinting at 12 months.

Among laboratory measurements as secondary outcomes, we found nominally significant changes in Hb, HbA1c and serum sodium cross-sectionally at baseline and in paired analyses but these parameters did not survive multiple testing correction (Bonferroni corrected for the p-value P<0.0055). There was an observed increase in HbA1c at 12 months post SARS-CoV-2 virus infection in both groups, but this did not pass multiple testing considerations, which is interesting and worth pursuing. A recent meta-analysis by Chen et al and colleagues provides evidence that severe COVID-19 is associated with increased blood glucose (weighted mean difference 2.21, 95% CI: 1.30-3.13, P < 0.001) however this was not demonstrated in those with mild COVID-19 (weighted mean difference 0.29, 95% CI: -0.59 to 1.16, P = 0.52), highlighting the need for monitoring not only BP but also glycaemic control. (240) Another meta-analysis, demonstrated that there is a higher risk of incident diabetes after hospital discharge or at least 28 days after

COVID-19 diagnosis and screening is essential regardless of disease severity or history of steroid treatment use. (241) It is postulated that blood glucose may be high in recovered individuals after COVID-19 infection where overt hyperglycaemia was identified in those without a diagnosis of diabetes, but persistent hyperglycaemia could lead to the development of new onset diabetes or put the individual at high risk of developing diabetes remains debatable, highlighting the need for further longitudinal studies. (242)

In current literature, studies linking COVID-19 with elevated BP or new-onset hypertension exhibit several limitations, including retrospective and observational study designs, reliance on self-reporting, small sample sizes, absence of control groups, and short follow-up durations. Hence, a cautious interpretation of these findings is essential.(243) For instance, Akpek et al. found elevated BP in only 18 out of 153 confirmed COVID-19 patients after a brief follow-up period of 31.6 ± 5.0 days. (94) Similarly a single-centre prospective study, reported no significant increase in BP after three and six months, although their study was limited by a small sample size. (244) A cross-sectional observational study in COVID-19 positive adults, with over a quarter reporting hypertension post-infection; however, caution is warranted due to variability in cardiac-related symptoms and study design. (211) Another study observed a heightened risk of incident cardiovascular disease beyond 30 days post-infection, with hypertension showing a high hazard ratio HR: 15.18 (95% CI: 11.53 to 18.62). (245)Two studies, a longitudinal prospective study found only 29.7% of survivors developing hypertension at one year and another observed new-onset hypertension in 33.2% of hospitalized COVID-19 patients at one year. (246, 247). Additionally, Alfadda et al. (2022) reported an increase in SBP at six months in COVID-19 survivors (124.68 ± 14.9 vs. in follow-up 131.26 ± 15.3, $p < 0.001$). (248) Another retrospective and prospective observational cohort study that looked at 185 participants who had been discharged 23 days following COVID-19 infection found that 40 (21.6%) had uncontrolled BP that required therapeutic change. (249) Gameil et al who conducted a study observed that SBP was elevated in COVID-19 survivors (Control 120.63 ± 8.49 vs. Cases 126.70 ± 10.31,) in the univariate analysis (crude odds ratio 1.07 (1.03–1.109), $p<0.001$) but lost significance in the multivariate analysis. (250)These findings underscore the need for further research with robust study designs and longer follow-up periods to elucidate the true relationship between COVID-19 and hypertension.

This leads to demonstrates that our study presents several notable strengths when compared to existing literature. Firstly, we offer longitudinal data points collected at both baseline and 12 months, complemented by sequential ABPM measurements, widely acknowledged as the gold standard for diagnosing and classifying hypertension. It's worth

noting that our participants had no prior history of hypertension, a crucial factor in minimising confounding variables and accurately tracking longitudinal BP changes potentially attributable to COVID-19 infection.

Additionally, our study assessed FMD as a marker of endothelial dysfunction at two time points, providing valuable insights into vascular health dynamics post-SARS-CoV-2 recovery. However, it is necessary to acknowledge the study's recruitment challenges amidst evolving pandemic dynamics, including vaccine rollout and emerging variants with potentially differing pathogenicity. This resulted in longer delays between the diagnosis of COVID-19 and baseline visit assessment and required the formulation of a per-protocol analysis in addition to analysis of the full dataset. This ensured that the participants in the per-protocol data set were similar in terms of the time of their baseline BP and other assessments and their 12-month assessments.

The 6MWT serves as a key marker of recovery, reflecting improvements in physical fitness and cardiopulmonary function. (234)Over time, individuals recovering from COVID-19 may experience enhanced lung function, muscle strength, and overall endurance as they recover from the acute effects of the illness. (146, 251) Additionally, the resolution of systemic inflammation may contribute to improved vascular and endothelial function, thereby reducing exertional fatigue and enhancing physical performance—despite the observed signal in BP at 12 months. Participants may also develop greater awareness of their baseline physical functioning and strive to improve it, addressing muscle deconditioning through intentional efforts. Familiarity with the test protocol and motivation to perform better are additional factors that could influence performance, albeit representing a common limitation in repeated assessments. The observed improvement in the 6MWT highlights the complex interplay between physical recovery, cardiometabolic health, and vascular changes post-COVID-19. This underscores the need for further research to elucidate the mechanisms driving these improvements and to better understand their implications for long-term rehabilitation strategies.

Limitations

We acknowledge several limitations in our study that may influence the interpretation of our findings. The LOCHINVAR study was conducted at a single centre with a relatively small sample size, which included an overrepresentation of females. Additionally, factors such as lifestyle changes, medication use, or psychological stress during the pandemic may have confounded the observed relationship between COVID-19 and cardiovascular outcomes. Recruitment challenges, influenced by the dynamics of the pandemic—such as vaccination rollouts and emerging variants—led to delays between COVID-19 diagnosis

and baseline assessment. Participants were not screened for repeated infections, long COVID-19 symptoms were not quantified, and data on specific SARS-CoV-2 variants were not collected. Recruitment relied on an opt-in approach to minimise participant burden, potentially limiting the diversity of the sample. Furthermore, urinary albumin excretion was not measured, and most cases in our cohort were mild to moderate in severity, which may affect the generalisability of the findings. A per-protocol analysis was conducted to ensure comparability between baseline and 12-month follow-up data, though this reduced statistical power.

The study cohort primarily included individuals who contracted SARS-CoV-2 but did not develop severe illness requiring intensive care. While this reflects the current epidemiological landscape amidst widespread vaccination efforts, it may not fully capture the spectrum of disease severity. Participants were recruited a minimum of 12 weeks post-hospital admission, without an upper limit on enrolment, possibly skewing the sample towards highly motivated individuals. Recruitment strategies were further influenced by the public's evolving perceptions of COVID-19 and the return to normal NHS services. For instance, during the early pandemic OBELIX pilot study, there was significant interest in participation, but recruitment rates declined as mandatory testing was phased out and public risk perception diminished, particularly among those without comorbidities.

These limitations highlight potential systematic errors and selection biases that may impact the internal validity of our results. Measurement errors in ABPM and FMD, despite rigorous quality control, and the lack of analysis of BP dipping and variability, as well as aldosterone-renin ratio, further constrain the depth of insights into hypertension pathophysiology in COVID-19 patients. To our knowledge, no participant was started on BP medication during the follow up period.

Nevertheless, our findings provide important insights into longitudinal changes in BP among individuals recovering from SARS-CoV-2 infection. A statistically significant increase in both SBP and DBP was observed over 12 months, alongside a reduction in FMD, suggesting a vascular underpinning to the hypertensive changes. Improvements in 6MWT further indicate a complex interplay between recovery and cardiometabolic health. While our results do not establish a causal relationship, they underscore the importance of continued cardiovascular monitoring in post-COVID-19 recovery. Future studies leveraging larger, more diverse cohorts and real-world data will be crucial to validate these findings and better understand the long-term cardiovascular risks associated with COVID-19.

4.8 Clinical Implications

Despite the need for validation, our findings underscore the critical importance of vigilant monitoring of BP and vascular health among individuals' post-recovery from SARS-CoV-2 infection. This is required irrespective of pre-existing hypertension status or COVID-19 disease severity. Healthcare providers should prioritise comprehensive post-COVID cardiovascular assessments to proactively address potential long-term complications. Moreover, the observed improvements in the six-minute walk test outcomes suggest a complex interplay between recovery trajectories and cardiovascular health, lending further credence to our BP findings. While these improvements may offer some reassurance, they also underscore the need for continued monitoring and proactive management of cardiovascular health in the post-COVID period.

4.9 Future Plans and Considerations

Looking ahead, our study opens avenues for further investigation into several aspects of post-COVID cardiovascular health. Firstly, increasing the sample size and extending the follow-up period beyond 12 months could provide valuable insights into the long-term implications of SARS-CoV-2 infection on BP and vascular function. Additionally, incorporating a broader range of cardiovascular markers, such as measures of arterial stiffness, direct vascular function assessments, and evaluations of metabolic parameters like HbA1c, can offer a more comprehensive understanding of the underlying mechanisms driving the observed changes. Furthermore, exploring the impact of different SARS-CoV-2 variants on cardiovascular health and assessing the potential mitigation role of vaccination in future cardiovascular risk warrants attention. Additionally, considering the evolving nature of the pandemic, ongoing research is essential to elucidate the specific effects of emerging variants and evolving public health measures on post-COVID cardiovascular outcomes. Moreover, investigating the potential benefits of lifestyle interventions, pharmacological treatments, or rehabilitation programs tailored to post-COVID cardiovascular health management could offer actionable strategies for improving patient outcomes.

Given the myriad challenges and uncertainties associated with the evolving landscape of COVID-19, future research endeavours must explore innovative approaches to generate robust evidence on the impact of SARS-CoV-2 recovery on BP. One promising avenue for advancing our understanding is the utilisation of real-world evidence derived from big data studies of healthcare records or large cohort studies. These studies can leverage sophisticated statistical methodologies to adjust for confounding variables and explore the

association between SARS-CoV-2 infection and subsequent changes in BP while accounting for factors such as age, sex, comorbidities, medication use, and socioeconomic status. Moreover, by incorporating data from multiple healthcare systems and geographic regions, these studies can enhance the generalisability and external validity of the findings, providing a more comprehensive understanding of post-COVID BP dynamics.

Additionally, prospective cohort studies involving large and diverse participant cohorts can further elucidate the temporal relationship between SARS-CoV-2 recovery and BP changes over time. Longitudinal assessments conducted at multiple time points post-recovery can capture the trajectory of BP alterations and identify potential predictors or modifiers of these outcomes, shedding light on the underlying mechanisms driving cardiovascular sequelae following COVID-19.

Furthermore, integrating machine learning and artificial intelligence, can enable the identification of novel risk factors or biomarkers associated with post-COVID hypertension, facilitating personalised risk stratification and targeted interventions for at-risk individuals.

4.10 Conclusion

In conclusion, while our findings highlight the importance of post-COVID cardiovascular assessments, it is essential to recognise the need for validation and the potential limitations of our study. By incorporating real-world evidence into future research endeavours, we can enhance our understanding of post-COVID cardiovascular dynamics and improve clinical care practices to mitigate long-term cardiovascular risks in individuals recovering from SARS-CoV-2 infection.

Chapter 5 Quality of Life

5.1 Background

Long COVID-19 significantly impacts psychosocial factors, impairing daily functioning and well-being (Chapter 1). This chapter extends Chapter 4 by providing a more detailed exploration of quality of life (QoL). Findings from a 12-month follow-up post-recovery indicate increased BP and reduced flow-mediated dilation (FMD) among COVID-19 survivors. These observations suggest a critical link to hypertension; however, the causal relationship requires clarification, as long COVID may act as a confounder for BP changes, potentially exacerbating or being exacerbated by diminished QoL.

QoL is a multidimensional construct reflecting overall well-being, encompassing physical health, psychological state, independence, social relationships, and personal beliefs. The EuroQol-5 Dimension (EQ-5D) is a validated tool used here to assess self-reported health. Its five dimensions—Mobility, Usual Activities, Self-care, Pain & Discomfort, and Anxiety & Depression—enable comprehensive health evaluations through questionnaires.

A deeper exploration of QoL in the context of long-term BP changes post-COVID has two main objectives. First, it provides insights into the lived experiences of individuals, offering a broader perspective beyond clinical measurements and capturing recovery trajectories. Second, it examines the interplay between cardiovascular health and QoL, highlighting a potential vicious cycle where declining physical health leads to reduced QoL, which in turn further impacts physical health.

This chapter presents QoL measures, assessed using the EQ-5D-3L (including the EQ-5D-Index and EQ-5D-VAS), from participants in the LOCHINVAR study. The analysis compares QoL between SARS-CoV-2-positive and negative participants at baseline and at 12 months, assessing whether changes persist over time. These findings complement BP and FMD results, broadening the understanding of COVID-19 recovery and emphasising the importance of patient-centred approaches to managing long COVID.

5.2 Methods

The detailed methods for patient recruitment and phenotyping are presented in the General Methods Chapter 3. The statistical analysis full and per-protocol datasets are described in the methods section of Chapter 4.

5.3 Results

5.3.1 Demographics and Baseline Characteristics

In the full dataset, demographic and baseline characteristics (Tables 14 and 15) show that the mean age of participants was comparable between SARS-CoV-2 positive and SARS-CoV-2 negative groups (48.6 vs. 47.9 years, $p = 0.221$). A significantly greater proportion of SARS-CoV-2 positive participants were male (48.6%) compared to SARS-CoV-2 negative groups participants (20%, $p < 0.001$), and SARS-CoV-2 positive participants had a higher mean BMI (28.2 vs. 26.6, $p < 0.001$).

5.3.2 EQ5DL Dimensions

At baseline (Table 14 and 15), SARS-CoV-2 positive participants reported significantly more problems across all EQ5DL dimensions, except self-care, than SARS-CoV-2 negative participants. At 12 months, significant differences persisted for activity, mobility, and pain in the full dataset, though these associations were attenuated compared to baseline. In the per protocol dataset, the significance of these associations was reduced, particularly for mobility and pain, reflecting potential heterogeneity in adherence or recovery trajectories. (Table 17)

5.3.3 EQ5D-VAS and EQ5D-Index

Baseline EQ5D-VAS scores were significantly lower in SARS-CoV-2 positive participants than SARS-CoV-2 negative participants (74.0 vs. 85.4, $p < 0.001$). EQ5D-Index scores were also lower in SARS-CoV-2 positive participants (0.95 vs. 0.85, $p < 0.001$). By 12 months, both groups showed slight declines in EQ5D-VAS and EQ5D-Index scores, with SARS-CoV-2 positive participants continuing to lag behind (EQ5D-VAS: 74.9 vs. 82.0, $p < 0.001$; EQ5D-Index: 0.95 vs. 0.86, $p = 0.011$). (Table 14 and 15) The box plots of EQ5D-VAS and EQ5D-Index over time for each dataset are presented in Figure 23 and Figure 24.

Adjusted Analyses

Linear regression models adjusting for age and BMI revealed that SARS-CoV-2 positive status was associated with significant reductions in EQ5D-VAS at 12 months in both the full dataset (-6.39, $p < 0.001$) and per protocol dataset (-7.9, $p < 0.001$). For EQ5D-Index, the association was significant only in the full dataset (-0.02, $p = 0.049$). (Table 16)

Table 14 : Overall Demographics – Full Dataset

This table presents the baseline demographic and clinical characteristics of participants in the full dataset, stratified by SARS-CoV-2 status (positive or negative). Continuous variables are reported as mean (standard deviation), and categorical variables are presented as frequency (percentage). P-values indicate group differences assessed using independent t-tests for continuous variables and chi-square tests for categorical variables.

		Baseline			12 months		
label	levels	SARS-CoV-2 Negative	SARS-CoV-2 Positive	P	SARS-CoV-2 Negative	SARS-CoV-2 Positive	P
Age (years)	Mean SD)	47.9 (7.5)	48.6 (6.7)	0.221			
Sex (n, %)	Female	288 (80.0)	114 (51.4)	<0.001			
	Male	72 (20.0)	108 (48.6)				
BMI (kg/m^2)	Mean SD)	26.6 (5.0)	28.2 (4.7)	<0.001			
Mobility	No problems	339 (94.2)	177 (79.7)	<0.001	328 (91.1)	179 (80.6)	0.008
	Some Problems	20 (5.6)	44 (19.8)		25 (6.9)	30 (13.5)	
	(Missing)	1 (0.3)	1 (0.5)		7 (1.9)	13 (5.9)	
Self-care	No problems	359 (99.7)	215 (96.8)	-	318 (88.3)	144 (64.9)	-
	Some Problems	0 (0.0)	6 (2.7)		0 (0.0)	0 (0.0)	
	(Missing)	1 (0.3)	1 (0.5)		42 (11.7)	78 (35.1)	
Activity	No problems	340 (94.4)	160 (72.1)	<0.001	314 (87.2)	167 (75.2)	0.004
	Some Problems	19 (5.3)	55 (24.8)		39 (10.8)	42 (18.9)	
	A lot of Problems	0 (0.0)	6 (2.7)		0 (0.0)	0 (0.0)	
	(Missing)	1 (0.3)	1 (0.5)		7 (1.9)	13 (5.9)	
Pain	No problems	352 (97.8)	147 (66.2)	<0.001	279 (77.5)	148 (66.7)	0.005
	Some Problems	0 (0.0)	68 (30.6)		68 (18.9)	61 (27.5)	
	A lot of Problems	7 (1.9)	6 (2.7)		6 (1.7)	0 (0.0)	
	(Missing)	1 (0.3)	1 (0.5)		7 (1.9)	13 (5.9)	
Anxiety	No problems	347 (96.4)	148 (66.7)	<0.001	287 (79.7)	162 (73.0)	0.052
	Some Problems	12 (3.3)	67 (30.2)		60 (16.7)	47 (21.2)	
	A lot of Problems	0 (0.0)	6 (2.7)		6 (1.7)	0 (0.0)	
	(Missing)	1 (0.3)	1 (0.5)		7 (1.9)	13 (5.9)	
EQ5D_VAS	Mean (SD)	85.4 (11.3)	74.0 (15.3)	<0.001	82.0 (15.3)	74.9 (15.6)	<0.001
EQ5D_index	Mean (SD)	0.95 (0.1)	0.85 (0.2)	<0.001	0.95 (0.1)	0.86 (0.1)	0.011

Table 15 Overall Demographics – Per Protocol Dataset

This table summarizes the baseline demographic and clinical characteristics of participants included in the per protocol dataset. The table includes age, sex distribution, BMI, and EQ5D scores (VAS, Index, and dimensions). Statistical comparisons between SARS-CoV-2 positive and negative groups were conducted using similar methods to Table 14.

		Baseline			12 months		
label	**levels**	**SARSCoV2-Neg**	**SARSCoV2-Pos**	**P**	**SARSCoV2-Neg**	**SARSCoV2-Pos**	**P**
Age	Mean (SD)	48.2 (7.3)	48.5 (7.1)	0.698			
Sex	Female	252 (82.4)	36 (40.0)	<0.001			
	Male	54 (17.6)	54 (60.0)				
BMI	Mean (SD)	26.6 (4.9)	27.9 (3.7)	0.014			
Mobility	No problems	293 (95.8)	72 (80.0)	<0.001	290 (94.8)	88 (97.8)	0.076
	Some Problems	12 (3.9)	18 (20.0)		12 (3.9)	0 (0.0)	
	(Missing)	1 (0.3)	0 (0.0)		4 (1.3)	2 (2.2)	
Selfcare	No problems	300 (98.0)	90 (100.0)	-	282 (92.2)	78 (86.7)	-
	Some Problems	0 (0.0)	0 (0.0)		0 (0.0)	0 (0.0)	
	(Missing)	6 (2.0)	0 (0.0)		24 (7.8)	12 (13.3)	
Activity	No problems	293 (95.8)	60 (66.7)	<0.001	295 (96.4)	76 (84.4)	<0.001
	Some Problems	12 (3.9)	30 (33.3)		7 (2.3)	12 (13.3)	
	A lot of Problems	0 (0.0)	0 (0.0)		0 (0.0)	0 (0.0)	
	(Missing)	1 (0.3)	0 (0.0)		4 (1.3)	2 (2.2)	
Pain	No problems	299 (97.7)	66 (73.3)	<0.001	256 (83.7)	68 (75.6)	0.059
	Some Problems	0 (0.0)	24 (26.7)		40 (13.1)	20 (22.2)	
	A lot of Problems	6 (2.0)	0 (0.0)		6 (2.0)	0 (0.0)	
	(Missing)	1 (0.3)	0 (0.0)		4 (1.3)	2 (2.2)	
Anxiety	No problems	293 (95.8)	54 (60.0)	<0.001	266 (86.9)	74 (82.2)	0.154
	Some Problems	12 (3.9)	36 (40.0)		30 (9.8)	14 (15.6)	
	A lot of Problems	0 (0.0)	0 (0.0)		6 (2.0)	0 (0.0)	
	(Missing)	1 (0.3)	0 (0.0)		4 (1.3)	2 (2.2)	
EQ5D_VAS	Mean (SD)	86.8 (10.3)	75.0 (14.1)	<0.001	84.7 (12.1)	76.2 (13.0)	<0.001
EQ5D_index	Mean (SD)	0.9 (0.1)	0.9 (0.1)	<0.001	0.9 (0.1)	0.9 (0.1)	0.414

Figure 23 Changes in EQ5D-VAS Over Time

The box plot below shows the changes in EQ5D-VAS scores for SARS-CoV-2 positive (SARSCoV2-Pos) and negative (SARSCoV2-Neg) groups from baseline to 12 months for both the full dataset and per protocol dataset

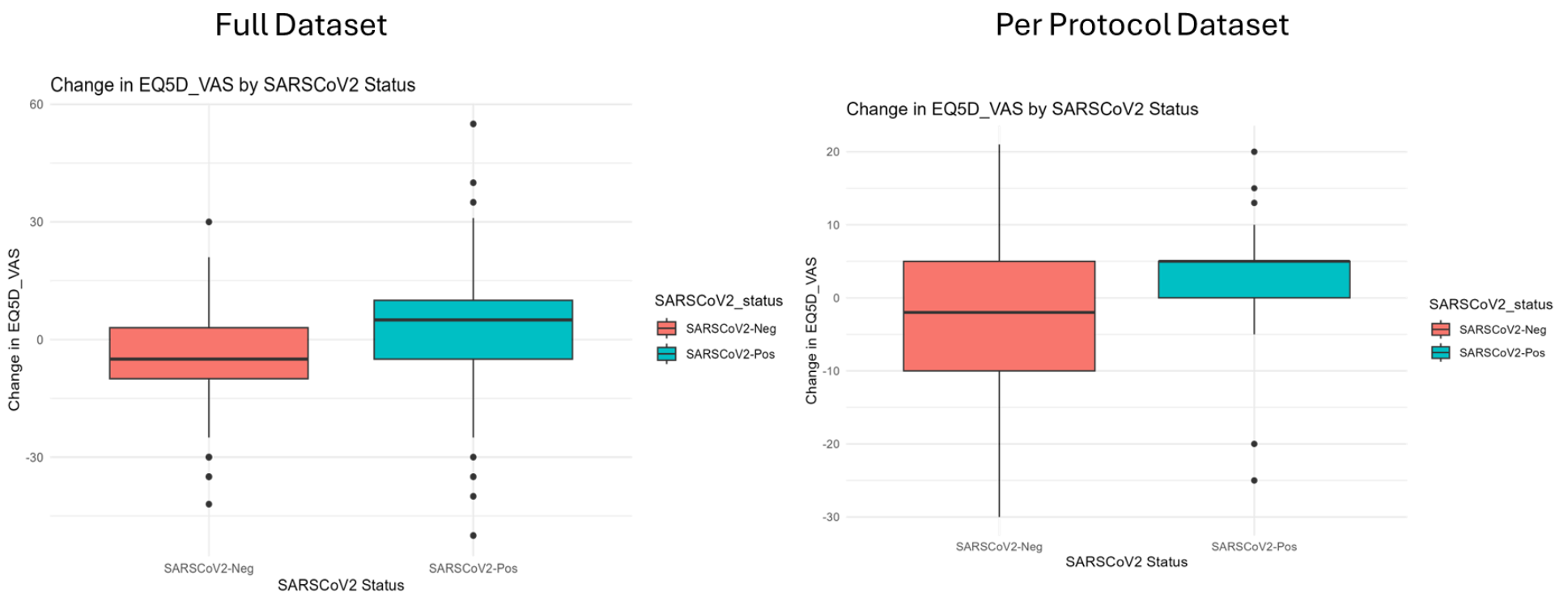


Figure 24 Changes in EQ5D-Index Over Time

The box plot below shows the changes in EQ5D-Index scores for SARS-CoV-2 positive (SARSCoV2-Pos) and negative (SARSCoV2-Neg) groups from baseline to 12 months.

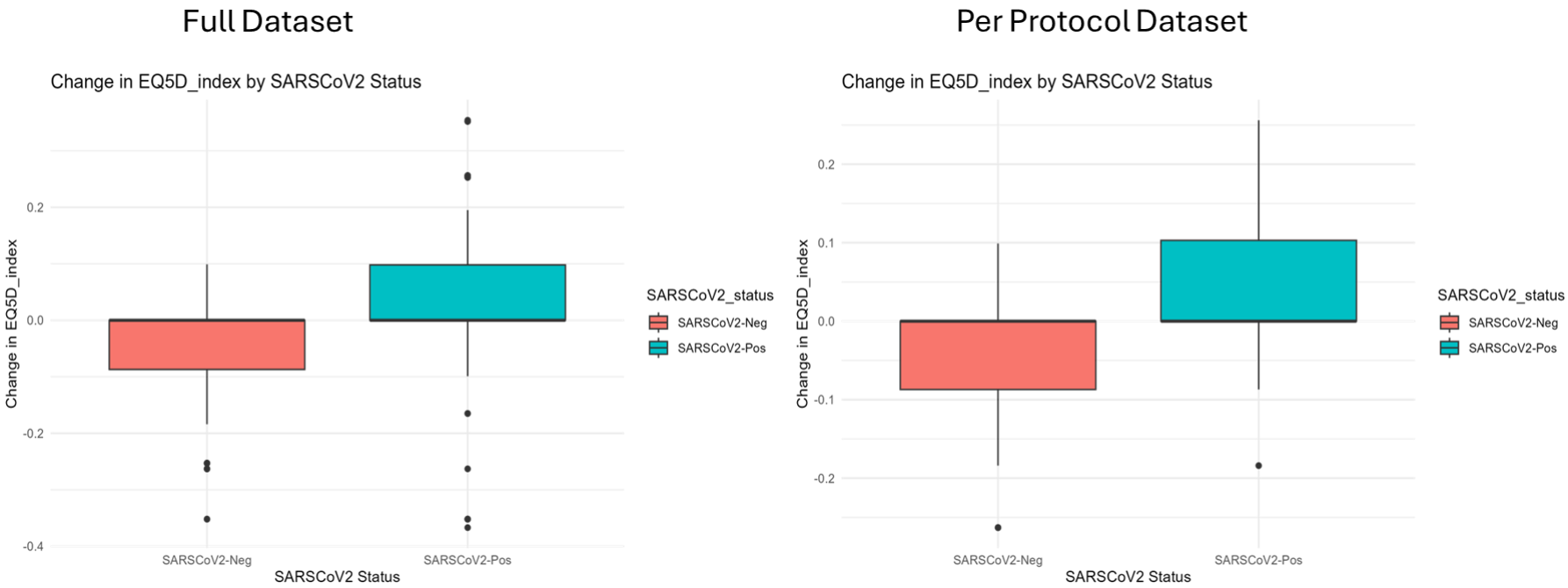


Table 16 Adjusted Analyses for EQ5D VAS and EQ5D Index at 12 Months

This table shows the adjusted regression analyses evaluating the association between SARS-CoV-2 status and EQ5D outcomes (VAS and Index) at 12 months. Models adjust for age, sex, BMI, and baseline scores (where applicable). Results are reported as adjusted estimates with standard errors (SE) and p-values. Separate analyses were conducted for the full dataset and per protocol dataset.

12 months	Full Dataset		Per Protocol	
	Estimate (SE)	P value	Estimate (SE)	P value
EQ5D-VAS adjusted for Age and BMI	-6.39 (1.35)	< 0.001	-7.9 (1.5)	< 0.001
EQ5D-Index adjusted for Age and BMI	-0.02 (0.01)	0.049	0 (0.01)	0.605
EQ5D-VAS adjusted for Age, BMI and Baseline EQ5D- VAS	1.24 (1.18)	0.293	-0.9 (1.37)	0.511
EQ5D-Index adjusted for Age, BMI and Baseline EQ5D-Index	0.03 (0.01)	0.003	0.04 (0.01)	< 0.001

Table 17 Likelihood of reporting problems in EQ5DL Dimensions at baseline and at 12 months.

This table provides results from logistic regression models assessing the likelihood of reporting problems in EQ5DL dimensions at 12 months. Adjusted estimates (standard error) and p-values are shown for each EQ5DL dimension, adjusted for covariates including age, sex, and BMI. Results are presented for baseline and 12 months for both datasets separately.

	Full Dataset				Per Protocol			
EQ5DL Dimension	Baseline		12 Months		Baseline		12 Months	
	Adj. Estimate (SE)	P	Adj. Estimate (SE)	P	Adj. Estimate (SE)	P	Adj. Estimate (SE)	P
Activity	1.94 (0.29)	<0.001	0.69 (0.25)	0.005	2.48 (0.37)	<0.001	2.05 (0.52)	<0.001
Anxiety	2.58 (0.33)	<0.001	0.19 (0.21)	0.376	2.84 (0.38)	<0.001	0.23 (0.35)	0.523
Mobility	1.38 (0.29)	<0.001	0.92 (0.28)	<0.001	1.79 (0.41)	<0.001	-17.45 (1 867.32)	0.993
Pain	3.33 (0.44)	<0.001	0.46 (0.2)	0.025	2.94 (0.49)	<0.001	0.41 (0.31)	0.188

When further adjusted for baseline values, SARS-CoV-2 positive status was not significantly associated with EQ5D-VAS change, but a significant positive association was observed for EQ5D-Index in both datasets. (Table 16) Table 17 demonstrates the results from logistic regression models assessing the likelihood of reporting problems in EQ5DL dimensions at 12 months for both datasets separately with activity being significant in both datasets.

Change in EQ5DL Dimensions

Analysis of changes from baseline to 12 months (Table 4) indicated significant differences between groups for anxiety (full dataset: -0.262, p = 0.005; per protocol: -0.347, p = 0.006) and pain (full dataset: -0.203, p = 0.024). Changes in other dimensions (e.g., mobility, activity) were not statistically significant. (Table 18)

Table 18 Changes in EQ5DL Dimensions from Baseline to 12 Months

This table reports the changes in EQ5DL dimensions (mobility, self-care, activity, pain, and anxiety) between baseline and 12 months for SARS-CoV-2 positive and negative groups. Results are stratified by dataset type (full vs. per protocol). P-values reflect the significance of group differences in change scores, adjusted for age, sex, and BMI.

EQ5DL Dimension	Full Dataset		Per Protocol	
	Adj. Estimate (SE)	P	Adj. Estimate (SE)	P
Activity	-0.128 (0.076)	0.095	-0.18 (0.027)	<0.001
Anxiety	-0.262 (0.092)	0.005	-0.347 (0.121)	0.006
Mobility	-0.034 (0.066)	0.609	-0.2 (0.052)	0.992
Pain	-0.203 (0.088)	0.024	-0.168 (0.103)	0.11

5.4 Discussion

This study highlights the substantial impact of SARS-CoV-2 infection on quality of life (QoL), as measured by EQ5DL dimensions, EQ5D-VAS, and EQ5D-Index, both at baseline and over a 12-month period. SARS-CoV-2 positive participants consistently reported worse baseline QoL across all dimensions, particularly in pain and anxiety, supporting existing evidence that SARS-CoV-2 infection exerts a profound physical and psychological toll.

By 12 months, improvements were observed in SARS-CoV-2 positive participants across dimensions such as activity and mobility, indicating recovery from the acute phase of infection. However, persistent disparities in pain and anxiety suggest potential long-term sequelae of infection, aligning with reports of “long COVID.”

The per protocol dataset revealed slightly larger effects in some dimensions, particularly in anxiety and activity. This may reflect a selection bias in participants who adhered to the full study protocol, potentially those with better health or motivation to recover.

Strengths

A key strength of this study lies in its use of the validated EQ5D framework, which allowed for a comprehensive and multidimensional assessment of QoL. This framework not only measured global health outcomes through summary scores such as EQ5D-VAS and EQ5D-Index but also captured granular details on specific health dimensions, including mobility, self-care, activity, pain, and anxiety. This dual approach provided a holistic view of the QoL impacts of SARS-CoV-2 infection, addressing both overall health perceptions and specific functional impairments. The ability to link summary measures with detailed dimension-level data adds depth to the analysis, enabling nuanced insights into recovery trajectories that would be overlooked by aggregate measures alone.

The longitudinal design is another major strength, as it allowed the study to capture both short-term and medium-term changes in QoL. By assessing participants at baseline and 12 months, this design provided a temporal perspective on recovery, distinguishing between acute effects of SARS-CoV-2 infection and longer-term residual impacts. This approach is particularly valuable given the emerging recognition of long COVID, which necessitates tracking persistent health deficits over time. The inclusion of both SARS-CoV-2 positive and negative participants further enhanced the study's robustness by offering a comparative lens to examine pandemic-related health impacts beyond infection status.

The study employed a rigorous statistical approach, leveraging adjusted regression analyses to account for potential confounders such as age, sex, and BMI. This method reduced the risk of spurious associations and provided a clearer understanding of the independent effects of SARS-CoV-2 infection on QoL outcomes. Moreover, the statistical significance observed in baseline comparisons and longitudinal changes bolsters the reliability of the findings. These results align with existing literature, which has consistently documented poorer QoL among SARS-CoV-2 positive individuals during acute infection and partial recovery over time. Notably, our findings corroborate the persistence of issues such as pain and mobility impairments, particularly in individuals experiencing post-acute sequelae of SARS-CoV-2 infection.

Despite these strengths, it is important to acknowledge that the study's reliance on the EQ5D framework may introduce certain limitations, as discussed below.

Limitations

One of the principal limitations of this study is the potential confounding by unmeasured factors. While the analysis adjusted for key variables such as age, sex, and BMI, it did not incorporate psychosocial determinants of health, including socioeconomic status, mental health history, or access to healthcare. These factors are particularly relevant for dimensions like anxiety and pain, where contextual influences may play a significant role. The absence of these variables limits the ability to fully interpret the drivers of QoL differences, particularly in the SARS-CoV-2 negative group, where worsening anxiety and pain trends were observed.

Another limitation arises from the use of self-reported measures such as EQ5D-VAS and EQ5DL dimensions. While these tools are validated and widely used, self-reported data are inherently subject to recall bias, especially for retrospective baseline assessments. Participants may have underreported or over reported their baseline QoL, leading to potential inaccuracies in the measurement of changes over time. Furthermore, the subjective nature of these measures means that individual perceptions of health can vary widely, introducing variability that may obscure subtle group differences.

Unaccounted heterogeneity within the SARS-CoV-2 positive group is another critical issue. The analysis treated all SARS-CoV-2 positive participants as a homogeneous cohort, ignoring potential variability based on disease severity, vaccination status, or pre-existing comorbidities. These factors likely influence both the immediate impact of SARS-CoV-2 infection and the trajectory of recovery. For instance, individuals with severe initial infections or underlying health conditions may experience prolonged or more severe impairments compared to those with milder cases. The lack of stratification by these key variables may have diluted the observed associations and limits the generalisability of the findings.

Additionally, while the study captured significant QoL impairments compared to the SARS-CoV-2 negative group, the underlying causes of these trends were not explored in depth. Worsening anxiety and pain in this group could reflect broader pandemic-related stressors such as social isolation, economic instability, or restricted access to healthcare. However, without additional data to contextualize these findings, the conclusions remain speculative. This highlights the need for more nuanced investigations into the indirect health impacts of the pandemic.

Finally, while the findings align with much of the existing literature, some discrepancies warrant further exploration. For example, while many studies report full recovery among SARS-CoV-2 survivors, our results indicate persistent impairments in mobility and activity. (252-255) These differences may reflect variations in study populations, methodological approaches, or follow-up durations. (256) Similarly, while our findings on anxiety in SARS-CoV-2 negative participants align with general population studies during the pandemic, the observed decline in physical health measures, such as pain, remains less well-documented, raising questions about potential unmeasured confounders or biases in our dataset.

5.5 Conclusion

This study demonstrates that SARS-CoV-2 infection significantly impairs QoL at baseline, with partial recovery over 12 months. While the study provides valuable evidence on the impact of SARS-CoV-2 infection on QoL, it also highlights the complexity of recovery and the multifaceted nature of post-COVID health challenges. Persistent deficits in pain and anxiety underscore the need for targeted interventions addressing both physical and psychological health post-infection. Overall, the study's strengths lie in its comprehensive scope, robust statistical methodology, and alignment with existing evidence. The use of a validated framework, combined with a longitudinal design, allowed for a detailed and multidimensional analysis of SARS-CoV-2's impact on QoL. However, the study is limited by its inability to fully account for contextual and psychosocial factors, its reliance on self-reported measures, and its lack of stratification within the SARS-CoV-2 positive group. These limitations underscore the need for future research to adopt more granular and context-sensitive approaches.

To address these gaps, future studies should incorporate a wider range of covariates, including socioeconomic and clinical data, and stratify analyses by disease severity and vaccination status. Moreover, qualitative methods could complement quantitative findings by providing deeper insights into the lived experiences of participants. Despite these limitations, the study provides valuable evidence on the long-term impacts of SARS-CoV-2 infection and offers a robust foundation for future research.

The findings from this study underscore the significant and multidimensional impact of SARS-CoV-2 infection on QoL, both during acute illness and in the recovery period. However, they also highlight critical gaps that require further investigation and action.

Future research and clinical strategies should aim to address these gaps through robust, multidimensional approaches that account for the complexities of post-COVID recovery.

5.6 Future Research Directions

A key priority for future work is extending the longitudinal follow-up beyond the 12-month period explored in this study. While the observed improvements in some QoL dimensions among SARS-CoV-2 positive participants suggest partial recovery, the persistence of deficits in areas such as pain and anxiety raise concerns about long-term sequelae. Extending follow-up to 24 months or beyond will provide a clearer picture of recovery trajectories, particularly for individuals experiencing long COVID symptoms. Additionally, this extended timeline will allow for the exploration of delayed or secondary impacts that may not manifest within the first-year post-infection.

Another critical area for future research is the inclusion of a broader range of covariates, particularly those capturing psychosocial and socioeconomic determinants of health. This study demonstrated the importance of adjusting for factors such as age, sex, and BMI, but the exclusion of variables like mental health history, access to healthcare, and economic stability limits the interpretability of findings, especially for anxiety and pain dimensions. By incorporating these variables into future analyses, researchers can better disentangle the direct effects of SARS-CoV-2 infection from the broader contextual impacts of the pandemic.

Stratification of participants based on disease severity, vaccination status, and comorbidities is another essential avenue for exploration. This study treated SARS-CoV-2 positive participants as a homogeneous group, but evidence suggests that recovery trajectories differ significantly between individuals with mild versus severe disease and between vaccinated and unvaccinated individuals. Stratified analyses will provide a more nuanced understanding of which subgroups are most vulnerable to prolonged impairments and inform tailored interventions.

Finally, future studies should integrate qualitative methods to complement the quantitative findings presented here. While validated tools like the EQ5D provide valuable insights into QoL changes, they cannot fully capture the lived experiences of individuals navigating post-COVID recovery. Qualitative interviews or focus groups with participants could shed

light on the specific challenges they face and the factors that facilitate or hinder their recovery.

5.7 Clinical Implications

The findings from this study have potential implications for clinical practice, if validated independently, particularly in the management of post-COVID recovery. The persistence of impairments in pain and anxiety among SARS-CoV-2 positive participants suggests the need for integrated, multidisciplinary rehabilitation programs. These programs should address both physical and mental health, incorporating pain management strategies alongside psychological support. Given the multidimensional nature of QoL deficits, such interventions should be designed to target specific domains of impairment rather than adopting a one-size-fits-all approach.

The worsening trends in pain and anxiety observed among SARS-CoV-2 negative participants also warrant clinical attention. These findings suggest that the broader impacts of the pandemic—such as social isolation, economic instability, and disrupted access to healthcare—have significant consequences for population health, even among those not directly affected by the virus. Clinicians should consider screening for and addressing these indirect effects as part of routine care, particularly for patients presenting with new or worsening symptoms during the pandemic period.

From a policy perspective, the results highlight the importance of investing in long-term post-COVID care infrastructure. This includes expanding access to rehabilitation services and mental health support for individuals recovering from SARS-CoV-2 infection, as well as developing public health initiatives to address the indirect impacts of the pandemic on non-infected populations. Furthermore, these findings underscore the need for ongoing surveillance of QoL outcomes at the population level, both to monitor recovery trajectories and to identify emerging health disparities.

Chapter 6 Evaluation of Transformer-Based Counterfactual Estimation of Individual Treatment Effects - Analysis of Angiotensin Converting Enzyme inhibitors (ACEIs) and risk of SARS-CoV-2 infection.

6.1 Introduction

The mechanism of action of ACEIs/ARBs, their protective effects, and their relevance to the COVID-19 pandemic, along with key studies, have been comprehensively detailed in Chapter 1. This chapter will address Objective 4.

6.2 Role of Machine Learning (ML)

Machine learning (ML) and Artificial Intelligence (AI) algorithms are rapidly replacing traditional logistic regression-based methods as the preferred tool for estimating propensity scores in observational studies. The benefits of ML and AI models include the ability to effectively capture complex relationships between covariates and treatment assignment, handle high-dimensional data, and automatically identify relevant covariates for propensity score estimation.(257-259) Additionally, AI and ML models do not require a priori assumptions about the true underlying form of the propensity model.(260) These models thus offer greater flexibility in estimating propensity scores and exhibit greater robustness against model misspecification compared to logistic regression methods.

Traditional statistical models (such as autoregressive models) often struggle with capturing complex dependencies over extended time intervals. Deep learning algorithms can model long-range dependencies more effectively than traditional statistical methods. Specific neural network architectures allow them to maintain memory over long-time intervals. (261) Furthermore, deep neural networks allow for end-to-end training, eliminating the need for manually crafting spatial contextual features. This streamlined approach allows the model to learn directly from raw data, making it more adaptable to temporal variations.(262) Finally, neural networks are efficient in capturing relationships between distant elements in a sequence, making them suitable for accurately capturing complex temporal dynamics.(263)

Thus, neural networks have emerged as powerful tools that can emulate the rigor and causal inference capabilities of a randomised controlled trial. By integrating various techniques, such as propensity score matching and temporal dynamics analysis, into these models, we can achieve accurate estimations of treatment risks. These approaches address inherent confounding due to the lack of randomisation, allowing us to separate design from analysis and explicitly examine confounder overlap. The result is a robust framework that ensures scientific rigour in assessing drug treatments.

6.3 Average and individual treatment effects.

6.3.1 Average Treatment Effect (ATE)

The concept of Average Treatment Effect (ATE) is central to causal inference and statistical analysis, serving as a key measure to estimate the overall impact of a treatment or intervention across a population.(264, 265) Essentially, ATE represents the difference in average outcomes between those who receive the treatment and those who do not, thereby providing insight into the causal effect of the treatment on the outcome of interest. (264, 265)

6.3.1.1 Considerations in Estimating ATE

A fundamental aspect of estimating ATE is the random assignment of participants into treatment and control groups. Randomisation ensures that any differences observed between the groups can be attributed to the treatment effect rather than pre-existing differences among participants. This method aims to eliminate biases and confounding factors, leading to an unbiased estimation of the ATE.

In scenarios where randomisation is not feasible, such as observational studies, confounding variables that influence both the treatment and the outcome must be carefully accounted for. Various statistical techniques, including regression adjustment, matching, stratification, and instrumental variables, are employed to mitigate the impact of these confounders and achieve an accurate estimation of the ATE.

While ATE provides an average effect across the population, it may not capture the variability in individual responses to the treatment. Different subgroups within the population might respond differently to the treatment, which means that the average effect might not be representative of the effect on any specific individual. Understanding this heterogeneity is necessary for an accurate interpretation of the ATE.

For more complex observational studies, advanced statistical methods such as propensity score matching, difference-in-differences (DiD), and regression discontinuity designs

(RDD) are used. These methods are designed to control for confounding variables and biases, providing a more robust estimation of the ATE in the absence of randomisation.

6.3.1.2 Interpreting the Average Treatment Effect

Interpreting the ATE involves a comprehensive analysis that goes beyond the numerical value. It requires an understanding of the magnitude, direction, and significance of the treatment effect, as well as its practical implications. Additionally, careful consideration of the study design, context, and population is necessary to accurately interpret and generalise the findings.

6.3.1.3 Magnitude of the Effect

The value of the ATE quantifies the average change in the outcome variable attributed to the treatment. For example, an ATE of 5 points implies that, on average, the treatment results in a 5-point change in outcomes compared to no treatment.

6.3.1.4 Direction of the Effect

The sign of the ATE indicates whether the treatment has a positive or negative impact on the outcome. A positive ATE suggests a beneficial effect, while a negative ATE indicates a detrimental effect on the treatment group relative to the control group.

6.3.1.5 Statistical Significance

Assessing the statistical significance of the ATE involves hypothesis testing to determine if the observed effect is unlikely to be due to chance. A statistically significant ATE provides evidence that the treatment has a real effect on the outcome, beyond random variation.

6.3.1.6 Clinical Significance

Beyond statistical significance, the clinical significance of the ATE must be considered. This involves evaluating whether the magnitude of the effect is large enough to be meaningful in real-world contexts. A statistically significant ATE might be practically insignificant if the effect size is too small to have real-world implications.

6.3.1.7 Generalisability and contextual factors

Understanding the contextual factors and mechanisms through which the treatment affects the outcome is essential. This context provides valuable insights into why and how the treatment works, enhancing the interpretation of the ATE. The generalisability of the ATE depends on the representativeness of the study sample. If the sample accurately reflects a larger population, the findings can be generalised more broadly. However, if the sample is highly specific, the generalisability of the ATE may be limited.

Exploring the variability of treatment effects across different subgroups within the population is important. This analysis can reveal who benefits most or least from the treatment, offering a deeper understanding of the ATE's implications.

6.3.2 Individual treatment effect (ITE)

The Individual Treatment Effect (ITE), also referred to as the Conditional Average Treatment Effect (CATE), refers to the effect of a treatment or intervention on an individual or a specific subgroup within the population, as opposed to the entire population. (266) It aims to capture the heterogeneity in treatment effects across individuals, recognizing that different people may respond differently to the same treatment. The ITE for an individual is defined as the difference in outcomes for the same individual if they were to receive the treatment versus if they were not to receive the treatment. Since both potential outcomes cannot be observed for the same individual (a problem known as the fundamental problem of causal inference), statistical methods and models are employed to estimate ITEs. Unlike ATE, which provides an overarching estimate of a treatment's impact on the entire population, ITE focuses on the unique response of each individual or subgroup to the treatment, recognising that the same intervention can have different effects on different people, in other words ITE captures the heterogeneity in treatment effects across individuals.

In practice, for each subject, only one of the potential outcomes is observed (the factual outcome if they have actually received treatment), while the other remains unobserved (the counterfactual outcome), because it is impossible for a subject to simultaneously experience both the treatment and the control conditions.(267) Mathematically, it is the difference between both potential outcomes - under treatment and control conditions for a given individual.

$$\mathrm{ITE}_i = Y_{1i} - Y_{0i}$$

Where:

- Y_{1i} is the potential outcome for individual i if they receive the treatment.
- Y_{0i} is the potential outcome for individual i if they do not receive the treatment.

6.3.2.1 Estimation of ITE

Estimating ITE involves more complexity than estimating ATE because it requires capturing the response of each individual to the treatment. This often necessitates the use of advanced statistical techniques and machine learning models, which can predict individual outcomes based on specific characteristics and covariates. Techniques such as

Bayesian hierarchical models, random forests, and causal machine learning algorithms are commonly used to estimate ITE, leveraging rich datasets that include detailed individual-level information.

6.3.3 Differences Between ITE and ATE

6.3.3.1 Level of Aggregation

The most significant difference between ITE and ATE lies in their level of aggregation. ATE measures the average effect of the treatment across the entire population, yielding a single summary statistic that represents the overall impact. In contrast, ITE measures the effect of the treatment at the individual level or for specific subgroups, reflecting the heterogeneity in treatment responses.

6.3.3.2 Purpose and Application

ATE is particularly useful for understanding the overall impact of a policy or intervention, making it a valuable tool for policymakers and researchers focused on broad-scale outcomes. It informs decisions at the policy level by providing a general picture of the treatment's effectiveness. On the other hand, ITE is essential for personalised decision-making and targeted interventions. It is critical in fields such as precision medicine, where treatments must be tailored to individual patients based on their unique characteristics, and in education and social policy, where interventions may need to be customised for different demographic groups. The strengths and limitations of ATE versus ITE is demonstrated in Table 19.

Table 19 Strengths and Limitations for ATE and ITE

	Strengths	**Weaknesses**
Average Treatment Effect	Provides a clear, concise measure of the overall treatment effect; useful for policy decisions and broad-scale interventions; relatively easier to estimate from experimental and observational data.	Does not capture individual variability in treatment response; may obscure important differences among subgroups; less useful for personalised decision-making.
Individual Treatment Effect	Captures the heterogeneity in treatment effects; essential for personalised medicine and targeted interventions; allows for more nuanced and effective decision-making.	More complex to estimate; requires sophisticated statistical and machine learning models; depends on detailed individual-level data, which may not always be available.

6.3.3.3 Estimation Complexity

The estimation of ATE is generally more straightforward than that of ITE. ATE can often be estimated directly from randomised controlled trials or well-designed observational studies with appropriate statistical adjustments to account for confounding factors. Estimating ITE, however, requires more sophisticated modelling to predict individual-level outcomes accurately. These models must account for the variability in treatment effects and often rely on high-dimensional data and complex algorithms.

6.3.3.4 Use in Precision Medicine and Personalised Policies

ITE estimation is paramount in precision medicine, where the goal is to identify which treatments work best for which patients. This personalised approach can lead to more effective and efficient healthcare interventions. Similarly, in education, social policy, and marketing, understanding the ITE allows for the design of interventions that are tailored to the specific needs and characteristics of different subgroups, thereby optimising outcomes.

The rationale for this chapter stems from the need to understand the actual impact of ACEIs on the risk of COVID-19 infection. Given the widespread use of ACEIs in managing hypertension and the COVID-19 pandemic, it is critical to determine how these medications influence COVID-19. The evolving landscape of the COVID-19 pandemic with the spread of new variants with differing infectivity and pathogenicity, the roll-out of vaccination makes any randomised controlled trial challenging. In this chapter, I propose to leverage the power of machine learning and ITE analysis to provide insights into the specific impact of ACEI on COVID-19.

The aim of this chapter is to investigate the effect of ACEIs on the risk of COVID-19 infection using routinely collected healthcare data from NHS GG&C, the largest health board in the UK. By applying advanced machine learning techniques, this analysis will estimate ITEs for the four major classes of antihypertensive drugs over two distinct timeframes of the COVID-19 pandemic. This approach aims to elucidate the specific effects of ACEIs, thereby contributing to advancing our understanding of RAAS inhibition and SARS-CoV-2 infection and help support effective clinical decision-making in the management of hypertension during this pandemic and future pandemics that involve the RAAS system.

6.4 Methods

The methods for this chapter have been described in detail in Chapter 3 and in the appendices. The Machine Learning models for estimation of ITEs and ATEs were developed by Tran QB Tran, a BHF MBPhD student working in my supervisor's research group. Tran's expertise is ML and he developed and tested the novel models to estimate ITE from the SafeHaven dataset. I was provided the estimated ITE and ATE values and I performed the subsequent analyses to test the relationship between ACEI and other antihypertensive drugs on the risk of SARS-CoV-2 infection. The evaluation and assessment of ML models that was utilised in the ITE estimation is beyond the scope of this PhD project and will not be described. The machine learning (ML) model incorporated time series data using the X Learner Framework for all hospital admissions and prescriptions, excluding cardiovascular drugs. This comprehensive approach enabled the model to automatically adjust for a full range of comorbidities, inferred either through hospitalisation records or prescription history.

6.4.1 Statistical Analysis

Demographic characteristics of the cohort are summarised using descriptive statistics, including medians and IQRs for continuous variables (e.g., age) and frequencies and percentages for categorical variables (e.g., gender, SIMD decile). These statistics were presented for the overall population and stratified by drug exposure. The distribution of ITEs was assessed using box plots and density plots. Box plots summarised the distribution of ITEs for each drug, showing medians, quartiles, and outliers, providing a clear picture of central tendency and variability. Density plots illustrated the overall distribution shape, while conditional density plots visualised how ITE distributions varied across different values of covariates such as age, gender, and SIMD decile. Paired tests of ITEs between ACEIs and each of the other drugs (BBs, CCBs, THZs, statins) were conducted using the Wilcoxon signed-rank test. This non-parametric test compared the median differences in ITEs, accounting for the paired nature of the data. The magnitude of differences in ITEs between ACEIs and the other drugs was assessed using the effect size, estimated by the correlation coefficient (r). Effect sizes were classified as small ($0.1 \leq r < 0.3$), medium ($0.3 \leq r < 0.5$), or large ($r \geq 0.5$), providing an interpretation of the clinical significance independent of sample size.

6.4.2 Software and Tools

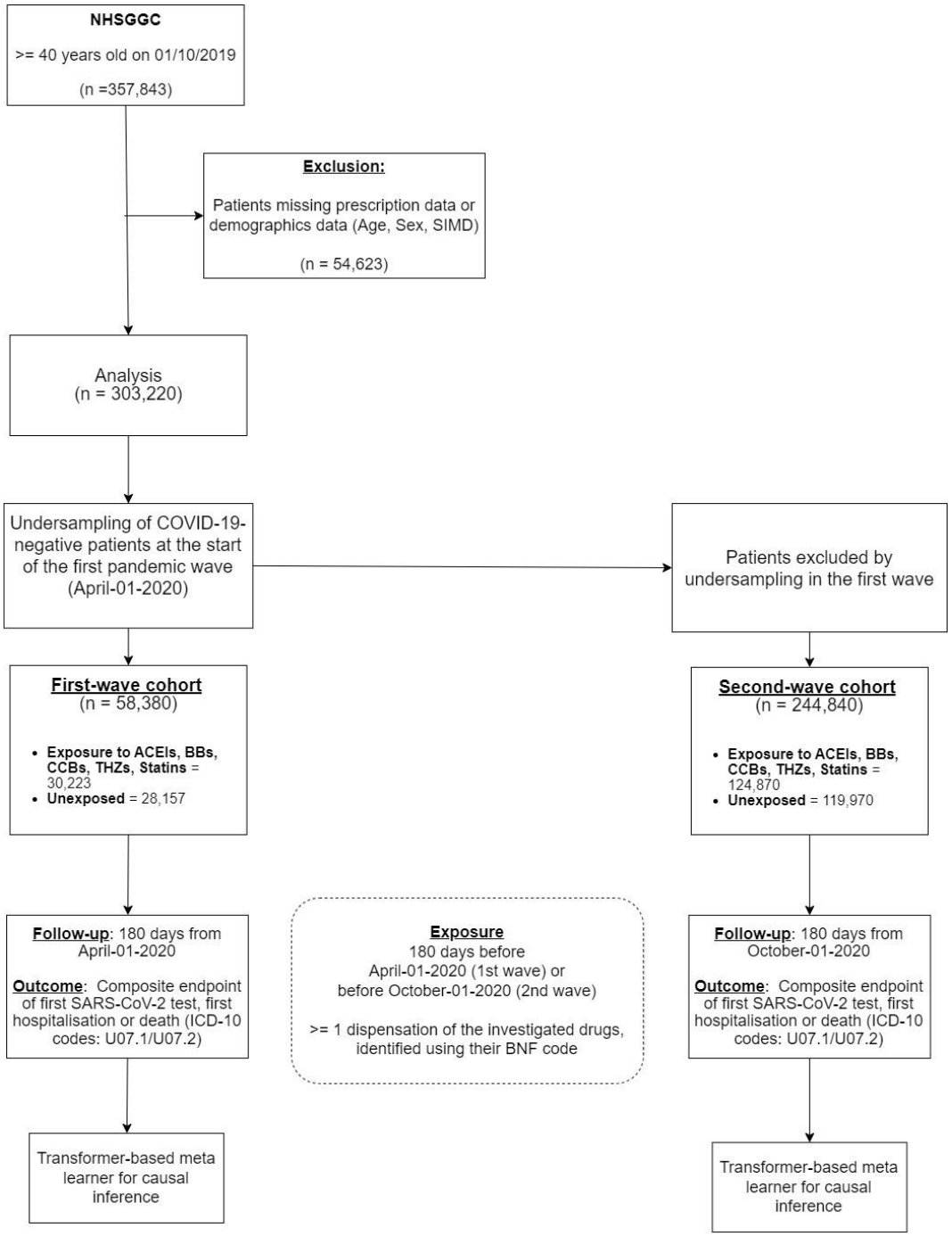
All statistical analyses were conducted using R (version 4.3.3) and Python (version 3.8).

6.5 Results

For this chapter, I will be focus on the comparisons of ACEIs with BBs, CCBs, THZs and statins. The study flow diagram is demonstrated in Figure 25.

Figure 25 Study Flow Diagram

This figure demonstrates the study flow diagram for the number of participants included in the analysis with exposure to antihypertensive drugs and statins for the first-wave and second wave cohort.



6.5.1 Overall study population

The overall study population comprised 303,220 individuals, with 58,380 (19.3%) from the first wave and 244,840 (80.7%) from the second wave and the patient characteristics are presented in Table 20. The top 10 comorbidities incorporated in the ML model is in Table 21. The median age of participants was 60.5 years (IQR: 51.5 to 72.3), with a slightly higher median age in the first wave (61.5 years; IQR: 52.2 to 73.3) compared to the second wave (60.2 years; IQR: 51.4 to 72.0) ($p < 0.001$). The gender distribution was balanced, with 56.2% female (170,263) and 43.8% male (132,957), showing no significant difference between the waves ($p = 0.063$). Socioeconomic status, as measured by the SIMD 2016 decile, was varied across the population. Approximately 22.8% of the participants were in the most deprived decile (decile 1), with similar distributions between the first (22.7%) and second waves (22.8%) ($p = 0.003$). Other deciles ranged from 6.0% to 10.2%, indicating a broad representation of socioeconomic backgrounds in the study.

6.5.2 First wave study population

The first wave study population (Table 22) was stratified by the use of different drugs, including ACEIs, BBs, CCBs, THZs and statins by factual drug exposure. The total number of participants varied across these drug categories, with the highest number of participants using statins (19,270; 33.0%) and the lowest using THZs (4247; 7.3%). The median age of participants ranged from 67 years (IQR: 58.0 to 76.5) for those using ACEIs to 70.5 years (IQR: 61.6 to 78.8) for those using statins indicating that these drugs were predominantly used by older individuals. Gender distribution also varied among the different drug categories. The proportion of females was highest among those using THZs. For ACEIs, BBs, CCBs, and statins, the female proportions were 45.2%, 52.5%, 50.8%, and 47.5%, respectively. Socioeconomic status, as measured by the SIMD 2016 decile, showed a diverse distribution across the drug categories. Participants using THZs and statins had a relatively lower representation in the most deprived decile (21.3% and 25.6%, respectively). The proportions in the least deprived decile for ACEIs, BBs, CCBs, THZs and statins were respectively 6.8%, 6.6%, 7.6% and 8.2%.

Table 20 Demographic and Socioeconomic Characteristics of the Overall Study Population Across Two Waves of the COVID-19 Pandemic

This table summarises the demographic and socioeconomic characteristics of 303,220 participants, stratified by the first and second waves of the COVID-19 pandemic. Variables include median age (with IQR), gender distribution, and socioeconomic status (measured by SIMD 2016 deciles). Statistical significance between waves is indicated.

label	levels	First Wave	Second Wave	Total	p
Total N (%)		58380 (19.3)	244840 (80.7)	303220	
Age	Median (IQR)	61.5 (52.2 to 73.3)	60.2 (51.4 to 72.0)	60.5 (51.5 to 72.3)	<0.001
SEX	F	32982 (56.5)	137281 (56.1)	170263 (56.2)	0.063
	M	25398 (43.5)	107559 (43.9)	132957 (43.8)	
SIMD	1	13078 (22.7)	55024 (22.8)	68102 (22.8)	0.003
	2	8560 (14.9)	34993 (14.5)	43553 (14.6)	
	3	5207 (9.1)	22236 (9.2)	27443 (9.2)	
	4	4559 (7.9)	19578 (8.1)	24137 (8.1)	
	5	4004 (7.0)	16409 (6.8)	20413 (6.8)	
	6	3755 (6.5)	15623 (6.5)	19378 (6.5)	
	7	3553 (6.2)	14358 (5.9)	17911 (6.0)	
	8	4303 (7.5)	17665 (7.3)	21968 (7.4)	
	9	5672 (9.9)	24760 (10.3)	30432 (10.2)	
	10	4816 (8.4)	20679 (8.6)	25495 (8.5)	
Diabetes	No	53660 (91.9)	224920 (91.9)	278580 (91.9)	0.692
	Yes	4720 (8.1)	19920 (8.1)	24640 (8.1)	
Incident COVID-19	No	53036 (90.8)	221319 (90.4)	274355 (90.5)	<0.001
	Yes	5344 (9.2)	23521 (9.6)	28865 (9.5)	

Table 21 Comorbidities in the ML Model

The top 10 comorbidities are shown - all other comorbidities, although not shown in the table, were also incorporated as inputs for the neural network model.

	Levels	First Wave	Second Wave	Total
Comorbidities included in the ML model – showing Top 10 comorbidities by ICD10 codes based on prior hospital admissions. (n, % of patient population)	No admission	44653 (76.5)	212745 (86.9)	257398 (84.9)
	Cataract	828 (1.42)	654 (0.27)	1482 (0.49)
	Unspecified acute lower respiratory infection	389 (0.67)	619 (0.25)	1008 (0.33)
	Urinary tract infection	374 (0.64)	999 (0.41)	1373 (0.45)
	Chronic obstructive pulmonary disease	280 (0.48)	534 (0.22)	814 (0.27)
	Atherosclerotic heart disease	272 (0.47)	774 (0.32)	1046 (0.34)
	Unspecified sepsis	229 (0.39)	617 (0.25)	846 (0.28)
	Malignant breast neoplasm	206 (0.35)	633 (0.26)	839 (0.28)
	Unspecified chest pain	187 (0.32)	717 (0.29)	904 (0.30)
	Syncope and collapse	140 (0.24)	429 (0.19)	569 (0.19)

6.5.3 Second wave study population

The second wave study population (Table 23) included 244,840 individuals stratified by their use of different drugs by factual drug exposure. Among them, 47,833 (19.5%) were using ACEIs, 51,873 (21.2%) were using BBs, 44,993 (18.4%) were using CCBs, 17,867 (7.3%) were using THZs, and 79,672 (32.5%) were using statins. The median age varied

across drug groups, with ACEIs users having a median age of 66.0 years (IQR: 57.1 to 75.8), BBs users at 67.3 years (IQR: 56.7 to 77.5), CCBs users at 68.7 years (IQR: 59.9 to 77.5), THZs users at 70.0 years (IQR: 61.1 to 77.7), and statins users at 69.7 years (IQR: 61.0 to 77.9). Gender distribution showed that females constituted 45.0% of ACEIs users, 52.8% of BBs users, 50.4% of CCBs users, 61.5% of THZs users, and 47.5% of statins users. Consequently, males made up 55.0% of ACEIs users, 47.2% of BBs users, 49.6% of CCBs users, 38.5% of THZs users, and 52.5% of statins users. Socioeconomic status, as measured by the SIMD 2016 decile, indicated a varied distribution among drug users. For those on ACEIs, 24.1% were in the most deprived decile, while 7.4% were in the least deprived decile. BBs users had 24.3% in the most deprived decile and 7.1% in the least deprived. Among CCBs users, 23.0% were in the most deprived decile, and 8.0% were in the least deprived. THZs users showed 21.6% in the most deprived decile and 8.5% in the least deprived. Statins users had 25.5% in the most deprived decile and 7.0% in the least deprived.

6.5.4 Distribution of ITE

The density plots of ITE by wave are presented in Figure 26. Both plots show BBs and CCBs show ITE effects in the negative range while ACEIs and Statins show a positive effect.

6.5.5 Conditional Density Plots of ITEs

The conditional density plots of ITEs were generated to visualise the distribution of ITEs across different categories of age, gender, and SIMD (Figures 28-30 respectively). These plots allow for an examination of how the treatment effect might vary as a function of these covariates. Conditional density plots of ITEs across age, gender, and SIMD categories revealed consistent treatment effects across these demographic and socioeconomic groups. This uniformity suggests that the impact of the studied drug classes on COVID-19 infection risk does not vary significantly based on age, gender, or socioeconomic status

Table 22 Patient Characteristics Stratified by Drug Exposure During the First Wave of the COVID-19 Pandemic

This table details the demographic and socioeconomic distribution of patients stratified by their use of ACEIs, BBs, CCBs, THZs, and statins during the first wave. Medication counts represent the number of patients exposed to each drug, either as monotherapy or in combination with others. Percentages for each medication were calculated relative to the total study population of 58,380 patients in the first pandemic wave. All 58,380 patients were used in the transformer model to estimate the individual treatment effect of each drug.

label	levels	ACEis	BBs	CCBs	THZs	Statins
Total N (%)		11369 (19.5)	12092 (14.1)	10680 (18.3)	4247 (7.3)	19270 (33.0)
Age	Median (IQR)	67.0 (58.0 to 76.5)	68.9 (58.2 to 79.2)	69.5 (60.7 to 78.0)	70.1 (61.5 to 77.9)	70.5 (61.6 to 78.8)
Sex	Female	5143 (45.2)	6351 (52.5)	5422 (50.8)	2592 (61.0)	9160 (47.5)
	Male	6226 (54.8)	5741 (47.5)	5258 (49.2)	1655 (39.0)	10110 (52.5)
SIMD	1	2726 (24.3)	2856 (23.9)	2396 (22.7)	895 (21.3)	4863 (25.6)
	2	1763 (15.7)	1893 (15.9)	1605 (15.2)	653 (15.5)	3005 (15.8)
	3	1095 (9.7)	1166 (9.8)	1005 (9.5)	381 (9.1)	1818 (9.6)
	4	884 (7.9)	970 (8.1)	886 (8.4)	345 (8.2)	1467 (7.7)
	5	787 (7.0)	820 (6.9)	742 (7.0)	295 (7.0)	1282 (6.7)
	6	735 (6.5)	808 (6.8)	668 (6.3)	261 (6.2)	1233 (6.5)
	7	675 (6.0)	681 (5.7)	640 (6.1)	256 (6.1)	1122 (5.9)
	8	772 (6.9)	815 (6.8)	758 (7.2)	328 (7.8)	1229 (6.5)
	9	1032 (9.2)	1135 (9.5)	1031 (9.8)	450 (10.7)	1698 (8.9)
	10	766 (6.8)	787 (6.6)	802 (7.6)	345 (8.2)	1297 (6.8)

Table 23 Patient Characteristics Stratified by Drug Exposure During the Second Wave of the COVID-19 Pandemic

This table details the demographic and socioeconomic characteristics of patients using ACEIs, BBs, CCBs, THZs, and statins during the second wave. Medication counts represent the number of patients exposed to each drug, either as monotherapy or in combination with others. Percentages for each medication were calculated relative to the total study population of 244,840 patients in the second pandemic wave. All 244,840 patients were used in the transformer model to estimate the individual treatment effect of each drug.

label	levels	ACEIs	BBs	CCBs	THZs	Statins
Total N (%)		47833 (19.5)	51873 (21.2)	44993 (18.4)	17867 (7.3)	79672 (32.5)
Age	Median (IQR)	66.0 (57.1 to 75.8)	67.3 (56.7 to 77.5)	68.7 (59.9 to 77.5)	70.0 (61.1 to 77.7)	69.7 (61.0 to 77.9)
Sex	Female	21526 (45.0)	27386 (52.8)	22692 (50.4)	10994 (61.5)	37818 (47.5)
	Male	26307 (55.0)	24487 (47.2)	22301 (49.6)	6873 (38.5)	41854 (52.5)
SIMD	1	11379 (24.1)	12456 (24.3)	10208 (23.0)	3815 (21.6)	20110 (25.5)
	2	7230 (15.3)	8006 (15.6)	6555 (14.7)	2532 (14.3)	12413 (15.8)
	3	4661 (9.9)	4969 (9.7)	4331 (9.7)	1775 (10.0)	7678 (9.8)
	4	3903 (8.3)	4260 (8.3)	3642 (8.2)	1424 (8.0)	6278 (8.0)
	5	3265 (6.9)	3589 (7.0)	3122 (7.0)	1258 (7.1)	5360 (6.8)
	6	3001 (6.3)	3256 (6.4)	2860 (6.4)	1098 (6.2)	4815 (6.1)
	7	2731 (5.8)	2941 (5.7)	2650 (6.0)	1108 (6.3)	4448 (5.6)
	8	3206 (6.8)	3373 (6.6)	3156 (7.1)	1284 (7.3)	5108 (6.5)
	9	4398 (9.3)	4748 (9.3)	4374 (9.8)	1902 (10.8)	6996 (8.9)
	10	3509 (7.4)	3618 (7.1)	3572 (8.0)	1497 (8.5)	5528 (7.0)

Figure 26 Density Plots of Individual Treatment Effects (ITE) by Drug Class and Pandemic Wave

These density plots illustrate the distribution of ITEs for each drug class during the first and second waves. Drugs with protective effects (negative ITEs) and increased risk (positive ITEs) are visually distinguished.

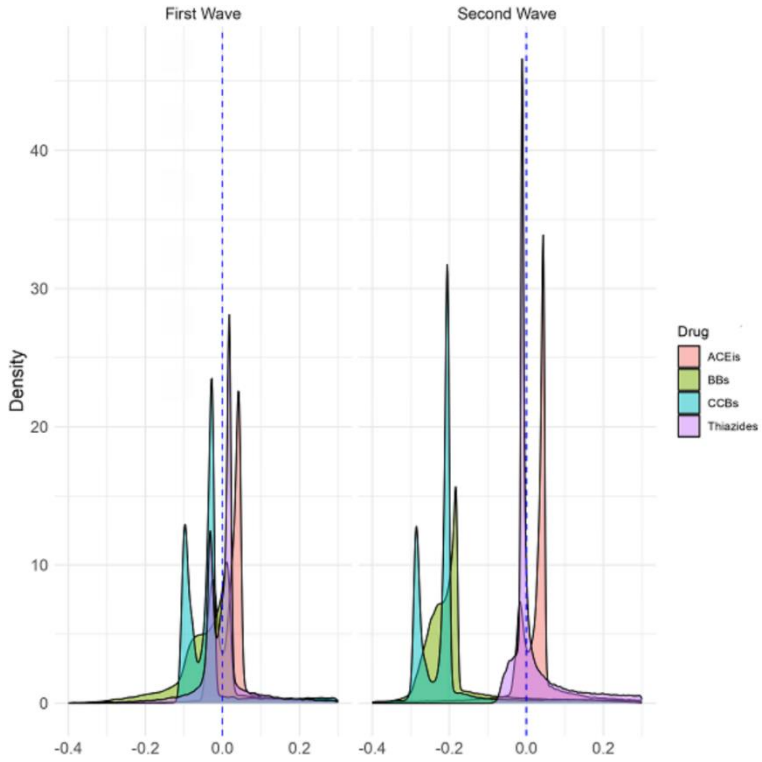


Figure 27 Conditional Density Plot of ITEs by Age Group

This plot demonstrates the distribution of ITEs stratified by patient age assessing differences in treatment effects vary across age groups for each drug class.

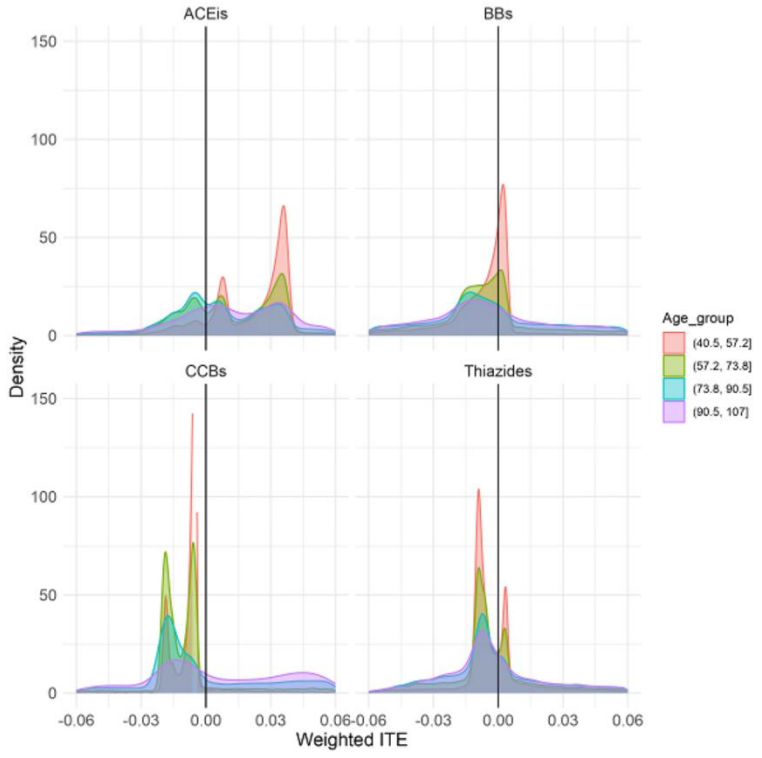


Figure 28 Conditional Density Plot of ITEs by Gender

This figure demonstrates the distribution of ITEs stratified by gender, assessing differences in treatment effects between male and female participants for each drug class.

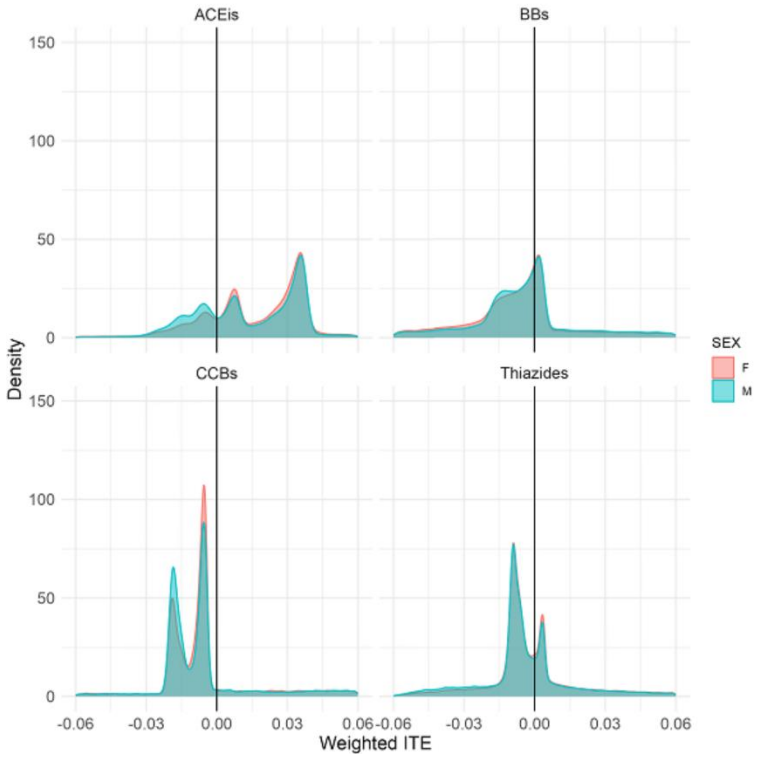
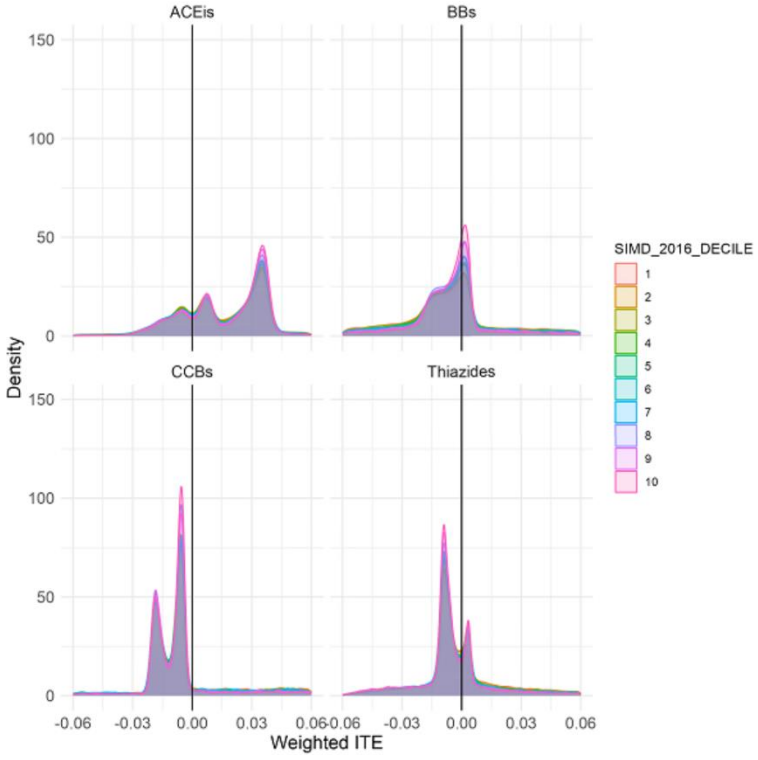


Figure 29 Conditional Density Plot of ITEs by Socioeconomic Status (SIMD Deciles)

The plot demonstrates the distribution of ITEs across SIMD deciles, assessing differences in socioeconomic status on treatment effects.



6.5.6 ATE analyses

Table 24 presents the results of the ATE analyses for four major classes of antihypertensive drugs and statins, a non-antihypertensive cardiovascular drug used as a comparator, during the first and second waves of the COVID-19 pandemic.

Table 25 presents the overall weighted ATE combining both waves. THZs showed the highest risk increase (ATE: 4.3% ± 10.8%), while ACEIs occupied a neutral position (ATE: 0.97% ± 5.5%). BBs (-8.3% ± 7.3%) and CCBs (-9.7% ± 8.1%) continued to exhibit protective effects, whereas statins (3.5% ± 6.1%) demonstrated an increased risk.

Table 24 Average Treatment Effect (ATE) by Drug Class Across the First and Second Waves of the COVID-19 Pandemic

This table shows the mean (SD) ATEs for ACEIs, BBs, CCBs, THZs, and statins during each wave. Differences between the waves are statistically evaluated, and the direction of effect is highlighted.

Drug	First Wave	Second Wave	p
ACEIs	0.0245 (0.0782)	0.0231 (0.1065)	0.004
BBs	-0.0169 (0.1244)	-0.1784 (0.1188)	<0.001
CCBs	-0.0209 (0.1058)	-0.2119 (0.1105)	<0.001
THZs	0.0033 (0.0877)	0.0820 (0.1717)	<0.001
Statins	0.0565 (0.1590)	0.1209 (0.1936)	<0.001

Table 25 Overall Weighted Average Treatment Effect (ATE) for Drug Classes Combining Both Pandemic Waves

The table shows the overall weighted ATEs for each drug class, combining data from both waves. This provides an integrated view of the treatment effect on COVID-19 risk.

Drug	Overall Weighted ATE
ACEIs	0.0097 (0.0552)
BBs	-0.0829 (0.0730)
CCBs	-0.0969 (0.0813)
THZs	0.0427 (0.1083)
Statins	0.0353 (0.0614)

6.5.7 ITE Analyses

Table 26 presents the results of the ITE analyses for four major classes of antihypertensive drugs and statins during the first and second waves of the COVID-19 pandemic. The weighted ITE of the overall population combining both waves and the results are maintained with THZs showing a negative median ITE with the lowest absolute magnitude of ITE of all the drugs.

Table 26 Median Individual Treatment Effect (ITE) by Drug Class During the First and Second Waves of the COVID-19 Pandemic

This table reports the median ITE values (with IQR) for each drug class in both waves. Statistical significance between waves is also presented.

Drug	label	levels	First Wave	Second Wave	p
ACEIs	ITE	Median (IQR)	0.0282 (-0.0156 to 0.0428)	0.0364 (-0.0041 to 0.0452)	<0.001
BBs	ITE	Median (IQR)	-0.0226 (-0.0738 to 0.0116)	-0.2035 (-0.2382 to -0.1817)	<0.001
CCBs	ITE	Median (IQR)	-0.0367 (-0.0861 to -0.0254)	-0.2102 (-0.2759 to -0.2034)	<0.001
THZs	ITE	Median (IQR)	0.0083 (-0.0292 to 0.0189)	-0.0042 (-0.0117 to 0.0993)	<0.001
Statins	ITE	Median (IQR)	0.0238 (-0.0114 to 0.0315)	0.0288 (0.0260 to 0.1386)	<0.001

During the first wave, ITE analyses revealed that patients taking ACEIs experienced a slight increase in COVID-19 risk, with a median risk difference of 2.8% (interquartile range [IQR]: -1.6% to 4.3%) compared to those not on ACEIs. Similarly, statins and THZs were associated with small increases in COVID-19 risk with median risk differences as follows: statins, 2.4% (IQR: -1.1% to 3.2%); THZs, 0.8% (IQR: -2.9% to 1.9. In contrast, BBs and CCBs showed protective effects against COVID-19, with reductions in risk. The median risk reductions were 2.3% (IQR: -7.4% to 1.2%) for BBs and 3.7% (IQR: -8.6% to -2.5%) for CCBs.

During the second wave of the COVID-19 pandemic, the ITEs of most drug classes mirrored those observed in the first wave. ACEIs showed a slight increase in infection risk, with a median increase of 3.6% (IQR: -0.4% to 4.5%). Similarly, statins was associated with a small increase in risk, with median increase of 2.9% (IQR: 2.6% to 13.9%). In contrast, BBs and CCBs continued to demonstrate protective effects, with significant reductions in infection risk. The median reductions were 20.4% (IQR: 18.2% to 23.8%) for BBs and 21.0% (IQR: 20.3% to 27.6%) for CCBs. Interestingly, the second wave revealed a negligible effect for THZs, with a median change of -0.4% (IQR: -1.2% to 9.9%).

6.5.8 Paired ITE Analyses – ACEI versus other drug classes

Table 27 presents the results of paired analyses during the first wave, comparing ACEIs with BBs, CCBs, THZs, and Statins. The table includes the median ITE, Wilcoxon two-sided p-values, Cohen's D, effect sizes, and the magnitude of the effect.

Paired analyses across both waves revealed significant differences between ACEIs and other drug classes. BBs and CCBs consistently exhibited large protective effects compared to ACEIs. For THZs, the effect size was moderate in the first wave but diminished to small in the second wave. Statins consistently demonstrated small effect sizes.(Table 28)

6.6 Discussion

The study evaluated the advanced transformer-based model with counterfactual X-learner framework estimated individual and average treatment effects of antihypertensive medications on the risk of COVID-19 infection. By leveraging real-world data from over 300,000 patients, our findings provide robust insights into the safety profile of ACEIs and highlight differential effects across other drug classes. The model's accuracy indicates its ability to effectively relate infection risk to the dynamic chronological interactions between background medications, comorbidities, and demographics.

Our findings reaffirm the safety of ACEIs during the COVID-19 pandemic, with minimal overall impact on infection risk. These results align with prior observational studies and meta-analyses, which reported no substantial increase in COVID-19 risk among ACEI users.(201, 268, 269) Our findings also align with biological hypothesis, suggesting that ACEIs may enhance SARS-CoV-2 cellular entry by upregulating ACE2 expression, while also being consistent with the predominantly null effect on clinical outcomes reported in observational studies, reinforcing the safety of ACEIs.

BBs and CCBs demonstrated protective effects against COVID-19 infection across both pandemic waves. These findings may reflect their underlying mechanisms, including modulation of immune responses, and attenuation of ACE2 expression by BBs and inhibition of calcium-dependent viral entry pathways by CCBs.(270-272) Previous observational studies showed beneficial effects of BB and CCBs.(273-276) CCBs were also shown to inhibit the post-entry replication events of SARS-CoV-2 in vitro, while no similar effect was observed for ACEIs.(276) The consistency of these effects across demographic and socioeconomic groups highlights their potential as therapeutic options deserving further investigation.

THZs were associated with a slight increase in COVID-19 risk, possibly due to their impact on sodium and RAAS regulation - their diuretic effect decreases sodium concentrations, thereby increasing angiotensin II and ACE2 expression.(277-279) These findings warrant further exploration, particularly in light of the widespread use of these medications in cardiovascular disease management.

Traditional methods for estimating individual treatment effects (ITEs), such as propensity score matching, inverse probability weighting, and linear regression-based approaches, often struggle with high-dimensional, time-series data and nonlinear confounding relationships. While methods like causal forests or generalized additive models address some of these challenges, they often fall short in capturing complex temporal dependencies and interactions between covariates. The transformer-based model integrated with the X-learner framework represents a significant advance over these traditional approaches. Transformers excel at modelling sequential data by leveraging self-attention mechanisms, which enable the identification of intricate temporal patterns and relationships. In this study, these capabilities were utilized to analyse the interplay between medication use, comorbidities, and COVID-19 infection risk, providing a granular understanding of treatment effects.

The Transformer model consistently outperformed other models in terms of F1 score and AUPRC across both waves, making it the most suitable choice for this study's objective of assessing increased COVID-19 risk associated with specific drug classes. Its strong performance in precision-recall balance indicates a higher reliability in identifying true risk cases while minimising false positives in the presence of class imbalances. While Logistic Regression offered the highest accuracy, its lower AUPRC highlights its potential limitations in this study context. Overall, these results suggest that the Transformer model

Table 27 Paired Analysis of ITEs Between ACEIs and Other Drug Classes During the First Wave of the COVID-19 Pandemic.

The paired analysis compares the ITEs of ACEIs with BBs, CCBs, THZs, and statins. Statistical measures, including Wilcoxon p-values and effect sizes, highlight significant differences in treatment effects.

First Treatment	Second Treatment	N	First ITE, median (IQR)	Second ITE, median (IQR)	Wilcox 2-sided	Wilcox First greater	Wilcox First lower	Cohens D (C.I.)	Effect Size	Magnitude
ACEIs	BBs	43,545	0.034 (0.044)	-0.018 (0.083)	0.0E+00	0.0E+00	1.0E+00	-0.43 (-0.44 - -0.42)	0.602	large
ACEIs	CCBs	44,635	0.034 (0.039)	-0.032 (0.039)	0.0E+00	0.0E+00	1.0E+00	-0.48 (-0.49 - -0.43)	0.611	large
ACEIs	THZs	46,536	0.033 (0.042)	0.012 (0.042)	0.0E+00	0.0E+00	1.0E+00	-0.26 (-0.27 - -0.25)	0.497	moderate
ACEIs	Statins	42,301	0.034 (0.039)	0.026 (0.032)	0.0E+00	0.0E+00	1.0E+00	0.13 (0.12 - 0.14)	0.202	small

Table 28 Paired Analysis of ITEs Between ACEIs and Other Drug Classes During the Second Wave of the COVID-19 Pandemic.

This table provides a paired comparison of ACEIs against other drug classes during the second wave. Statistical measures, including Wilcoxon p-values and effect sizes, highlight significant differences in treatment effects.

First Treatment	Second Treatment	N	First ITE, median (IQR)	Second ITE, median (IQR)	Wilcox 2-sided	Wilcox First greater	Wilcox First lower	Cohens D (C.I.)	Effect Size	Magnitude
ACEIs	BBs	182,808	0.039 (0.035)	-0.202 (0.052)	0.0E+00	0.0E+00	1.0E+00	-1.62 (-1.75 - -1.61)	0.794	large
ACEIs	CCBs	188,708	0.039 (0.033)	-0.208 (0.037)	0.0E+00	0.0E+00	1.0E+00	-2.04 (-2.23 - -2.03)	0.829	large
ACEIs	THZs	194,984	0.038 (0.035)	-0.006 (0.069)	0.0E+00	0.0E+00	1.0E+00	0.22 (0.24 - 0.24)	0.128	small
ACEIs	Statins	179,725	0.039 (0.029)	0.028 (0.030)	0.0E+00	1.0E+00	0.0E+00	0.40 (0.44 - 0.44)	0.096	small

provides the best combination of sensitivity, specificity, and predictive power for testing the hypothesis of drug related COVID-19 risk.

Strengths and Limitations

This study represents a significant advance in the application of AI for causal inference in healthcare. By integrating transformers and X-learners, the methodology accounts for both temporal dynamics and confounding factors, enabling more precise estimations of drug effects. Key strengths of this study include a large sample size, enhancing the generalisability of findings, and the use of deep learning and detailed visualisation techniques to strengthen the robustness and interpretability of results. Our counterfactual frameworks create virtual counterparts for each individual, enabling estimation of ITEs and accounting for heterogeneity, thereby providing more robust information on treatment impacts. While studies from the pandemic's first wave may have limited relevance to later stages due to changes driven by new viral variants and global vaccination rollouts, our study showed generally consistent findings across different drugs in both waves, highlighting the model's ability to produce generalisable results.

Despite its strengths, several limitations should be acknowledged. Transformers are inherently complex, and the black-box nature of the model may hinder clinical adoption - clinicians and policymakers often require interpretable outputs to trust and act upon model predictions. The model's reliance on large, high-quality datasets with rich temporal and static variables can limit its applicability in settings with incomplete or inconsistent data. Training transformer-based models is resource-intensive, potentially restricting their use in real-time or resource-constrained environments. Although advanced, the model cannot completely eliminate residual confounding, particularly from unmeasured variables. Additionally, the findings may not fully capture the evolving context of the pandemic, including changes in viral variants, vaccination status, and treatment protocols. Future work should focus on addressing some of the key limitations. This includes enhancing interpretability by integrating explainability techniques, such as SHAP values, to highlight the features driving predictions; reducing computational burden by implementing lightweight transformer variants; expanding the framework to diverse populations and datasets can enhance its generalisability and reliability; incorporating additional causal inference methods, such as targeted maximum likelihood estimation or doubly robust estimators, could enhance the accuracy of ITE estimation; adding dynamic covariates, such as real-time healthcare interventions or vaccination status, would improve the model's ability to adapt to evolving clinical contexts. Finally, our study did not specifically examine ARBs, as clinical practice often initiates ARBs after ACEI failure, introducing potential bias.

Clinical Implications

The findings of this study have several important clinical implications. The identification of differential effects among antihypertensive drugs can inform more personalised treatment strategies for patients with COVID-19, particularly those with hypertension. BBs and CCBs consistently showed positive effects on COVID-19 outcomes across both waves, suggesting their potential benefit in managing hypertensive patients and other cardiovascular conditions during the pandemic. Clinicians can consider these results when selecting antihypertensive medications for patients at risk of COVID-19, potentially opting for drugs associated with lower infection risk. The uniform treatment effects across age, gender, and socioeconomic status underscore the broad applicability of these findings. Insights into the socioeconomic distribution of treatment effects can help tailor public health interventions to ensure equitable healthcare access and outcomes.

Future Directions

The protective effects of BBs and CCBs against COVID-19 infection merit further investigation in randomised controlled trials or pragmatic studies. Mechanistic studies could elucidate the pathways underlying these effects, informing potential repurposing strategies. Furthermore, the integration of advanced machine learning models with real-world data presents opportunities to address other pressing clinical questions in dynamic healthcare settings.

Conclusions

This study not only provides critical evidence on the differential impacts of antihypertensive drugs on COVID-19 infection risk but also highlights the transformative potential of machine learning in medicine. By combining advanced modelling techniques with real-world data, we can move closer to personalised, data-driven healthcare solutions that improve patient outcomes and resilience during global health crises. This study contributes valuable evidence on the differential impacts of antihypertensive drugs on COVID-19 outcomes. While the study supports the continued use of ACEIs in hypertensive patients, it also suggests a re-evaluation of their risk profile during pandemics. The protective effects of BBs and CCBs highlight potential therapeutic benefits, while findings on THZs emphasize the need for further investigation.

Chapter 7 Overall Discussion

This thesis encompasses three interconnected studies that collectively investigate the long-term cardiovascular and quality-of-life consequences of SARS-CoV-2 infection, leveraging advanced machine learning techniques and real-world data. These findings collectively emphasize the intersection of cardiovascular health, post-viral sequelae, and advanced analytical methods in understanding and managing health outcomes during and after pandemics.

Clinical Phenotyping Post-COVID-19: A longitudinal analysis revealed sustained increases in BP and reduced endothelial function (measured by flow-mediated dilation, FMD) among post-SARS-CoV-2 individuals over 12 months. These changes, despite baseline similarity in cardiovascular metrics, underscore potential vascular implications of the virus. However, the study found no significant impact on baseline RAAS markers, necessitating further exploration of mechanistic pathways. The increase in BP and reduction in FMD provide compelling evidence of the vascular impact of COVID-19. These findings resonate with broader literature on endothelial dysfunction in viral infections and emphasize the need for routine cardiovascular monitoring post-COVID.

Quality of Life Impacts: The study documented substantial impairments in quality-of-life dimensions such as pain, mobility, and anxiety among SARS-CoV-2-positive participants. QoL assessments revealed the broader psychosocial and physical toll of SARS-CoV-2 infection. While improvements occurred over a year, persistent issues in pain and anxiety highlighted the recovery trajectories, with potential implications for long COVID management aligning with the emerging narrative of long COVID and highlight the importance of addressing mental health and chronic pain in recovery strategies.

Transformer-Based Counterfactual Modelling for Antihypertensive Drugs: By applying cutting-edge AI methodologies, the analysis demonstrated differential effects of antihypertensive medications on COVID-19 risk. Angiotensin-converting enzyme inhibitors showed neutral effects, while BBs and CCBs exhibited protective properties. THZs were associated with increased risk, suggesting the need for tailored pharmacological strategies in managing cardiovascular conditions during pandemics. The integration of transformer-based models demonstrates the transformative potential of AI in deriving granular insights from complex, high-dimensional datasets. This approach not only advances causal inference methodologies but also provides actionable insights into drug safety and efficacy during pandemics.

Building on the findings, several avenues for research and clinical translation are proposed:

Longitudinal and Multidimensional Follow-Up Studies: Extending follow-up beyond 12 months is critical to capturing delayed or evolving cardiovascular and QoL changes. Incorporating a broader range of biomarkers (e.g., arterial stiffness, inflammatory markers) and stratifying by disease severity, comorbidities, and vaccination status will enhance the understanding of post-COVID sequelae.

Mechanistic Studies and Interventional Trials: The vascular and pharmacological findings necessitate mechanistic investigations into endothelial dysfunction and the protective effects of specific antihypertensive drugs. Randomized controlled trials should evaluate the repurposing potential of BBs and CCBs for mitigating COVID-19 complications.

Integration of Psychosocial and Contextual Data: Future studies should incorporate socioeconomic, mental health, and healthcare access variables to contextualise QoL findings. Qualitative methodologies can complement quantitative analyses, providing deeper insights into lived experiences and recovery barriers.

ML estimation of ITE: The wider potential of using ML based estimation of ITE extends far beyond our first use in this study, offering transformative opportunities in establishing causality and conducting natural experiments. By leveraging advanced ML algorithms, researchers can uncover causal relationships that traditional methods might overlook, enabling a more precise understanding of how specific treatments affect different subpopulations. This approach is particularly valuable in the context of natural experiments, where randomised controlled trials are impractical or unethical. For instance, during a pandemic, ML can be employed to analyse observational data and simulate randomised conditions, thus deriving insights into the effectiveness of various interventions.

Scalable and Explainable AI Models: Enhancing the interpretability of AI models through techniques like SHAP values will address barriers to clinical adoption. Additionally, adapting lightweight transformer variants for resource-constrained settings can democratise the use of advanced analytics in healthcare.

Real-World Evidence and Policy Implications: Leveraging large-scale healthcare datasets can provide robust evidence on long-term outcomes and inform public health strategies. Policymakers should prioritise investments in post-COVID care infrastructure, encompassing multidisciplinary rehabilitation programs and mental health services.

Chapter 8 Overall Conclusions

In conclusion, this thesis advances our understanding of the cardiovascular and QoL impacts of COVID-19, while also showcasing the transformative potential of machine learning in deriving actionable insights The observed increase in BP and endothelial dysfunction post-COVID-19 recovery underscores the need for vigilant cardiovascular monitoring. The differential effects of antihypertensive drugs on COVID-19 risk highlight the importance of personalised treatment approaches. The findings underscore the importance of integrating advanced analytics with robust clinical methodologies to address pressing healthcare challenges. By bridging the gaps between data science, clinical care, and public health, this work lays a strong foundation for improving patient outcomes and system resilience during global health crises. These findings provide a foundation for future research to further elucidate the long-term cardiovascular impacts of COVID-19 and develop targeted interventions to mitigate these risks.

Chapter 9 List of References

1. Lip S. A tale of two diseases. J Hum Hypertens. 2023;37(3):248-51.
2. Morens DM, Folkers GK, Fauci AS. What Is a Pandemic? The Journal of Infectious Diseases. 2009;200(7):1018-21.
3. Sachs JD, Karim SSA, Aknin L, Allen J, Brosbøl K, Colombo F, et al. The Lancet Commission on lessons for the future from the COVID-19 pandemic. Lancet. 2022;400(10359):1224-80.
4. Grennan D. What Is a Pandemic? JAMA. 2019;321(9):910.
5. Prevention CfDCa. 2024 [cited 2024 26 April 2024]. Available from: https://www.cdc.gov/.
6. Organisation WH. WHO Coronavirus (COVID-19) Dashboard. 2021 [cited 2024 28/2/24]. Available from: https://covid19.who.int.
7. Mofijur M, Fattah IMR, Alam MA, Islam A, Ong HC, Rahman SMA, et al. Impact of COVID-19 on the social, economic, environmental and energy domains: Lessons learnt from a global pandemic. Sustain Prod Consum. 2021;26:343-59.
8. Baker RE, Mahmud AS, Miller IF, Rajeev M, Rasambainarivo F, Rice BL, et al. Infectious disease in an era of global change. Nat Rev Microbiol. 2022;20(4):193-205.
9. McMichael AJ. Extreme weather events and infectious disease outbreaks. Virulence. 2015;6(6):543-7.
10. Agency UHS. Pertussis: guidelines for public health management; Guidance for healthcare professionals on the public health management of pertussis (whooping cough). 2012.
11. Agency UHS. National measles guidelines. 2010.
12. Xu H, Jia Y, Sun Z, Su J, Liu QS, Zhou Q, et al. Environmental pollution, a hidden culprit for health issues. Eco Environ Health. 2022;1(1):31-45.
13. Eto K. Pathology of Minamata disease. Toxicol Pathol. 1997;25(6):614-23.
14. Charnley GEC, Kelman I, Gaythorpe KAM, Murray KA. Traits and risk factors of post-disaster infectious disease outbreaks: a systematic review. Sci Rep. 2021;11(1):5616.
15. Walika M, Moitinho De Almeida M, Castro Delgado R, Arcos Gonzalez P. Outbreaks Following Natural Disasters: A Review of the Literature. Disaster Med Public Health Prep. 2023;17:e444.
16. Organisation WH. Environment, Climate Change and Health. 2024.
17. Heymann DL. Social, behavioural and environmental factors and their impact on infectious disease outbreaks. J Public Health Policy. 2005;26(1):133-9.
18. Todd E. Food-Borne Disease Prevention and Risk Assessment. Int J Environ Res Public Health. 2020;17(14).
19. Tran HM, Tsai FJ, Lee YL, Chang JH, Chang LT, Chang TY, et al. The impact of air pollution on respiratory diseases in an era of climate change: A review of the current evidence. Sci Total Environ. 2023;898:166340.
20. Korevaar DA, Visser BJ. Reviewing the evidence on nodding syndrome, a mysterious tropical disorder. Int J Infect Dis. 2013;17(3):e149-52.
21. Gully PR. Pandemics, regional outbreaks, and sudden-onset disasters. Healthc Manage Forum. 2020;33(4):164-9.
22. McCloskey B, Dar O, Zumla A, Heymann DL. Emerging infectious diseases and pandemic potential: status quo and reducing risk of global spread. Lancet Infect Dis. 2014;14(10):1001-10.
23. Dobson AP, Carper ER. Infectious Diseases and Human Population History: Throughout history the establishment of disease has been a side effect of the growth of civilization. BioScience. 1996;46(2):115-26.
24. Ziegler T, Moen A, Zhang W, Cox NJ. Global Influenza Surveillance and Response System: 70 years of responding to the expected and preparing for the unexpected. Lancet. 2022;400(10357):981-2.
25. Ghsa Preparation Task Force T. Global Health Security: The Lessons from the West African Ebola Virus Disease Epidemic and MERS Outbreak in the Republic of Korea. Osong Public Health Res Perspect. 2015;6(6):S25-7.
26. Pan NL. Visualizing the History of Pandemics Visual Capitalist2020 [cited 2024 15 May 2024]. Available from: https://www.visualcapitalist.com/history-of-pandemics-deadliest/#google_vignette.
27. Piret J, Boivin G. Pandemics Throughout History. Front Microbiol. 2020;11:631736.
28. Cunha CB, Cunha BA. Great plagues of the past and remaining questions. Paleomicrobiology: past human infections. 2008:1-20.

29. Zietz BP, Dunkelberg H. The history of the plague and the research on the causative agent Yersinia pestis. International journal of hygiene and environmental health. 2004;207(2):165-78.
30. Getz FM. Black Death and the silver lining: Meaning, continuity, and revolutionary change in histories of medieval plague. Journal of the History of Biology. 1991;24(2):265-89.
31. Conti AA. Quarantine Through History. International Encyclopedia of Public Health. 2008:454-62.
32. Trilla A, Trilla G, Daer C. The 1918 "Spanish flu" in Spain. Clin Infect Dis. 2008;47(5):668-73.
33. Aassve A, Alfani G, Gandolfi F, Le Moglie M. Epidemics and trust: The case of the Spanish Flu. Health Econ. 2021;30(4):840-57.
34. Mazumder B, Almond D, Park K, Crimmins EM, Finch CE. Lingering prenatal effects of the 1918 influenza pandemic on cardiovascular disease. J Dev Orig Health Dis. 2010;1(1):26-34.
35. Almond D. Is the 1918 Influenza Pandemic Over? Long-Term Effects of In Utero Influenza Exposure in the Post-1940 U.S. Population. Journal of Political Economy. 2006;114(4):672-712.
36. Crosby AW. America's Forgotten Pandemic: The Influenza of 1918. 2 ed. Cambridge: Cambridge University Press; 2003.
37. Simon V, Ho DD, Abdool Karim Q. HIV/AIDS epidemiology, pathogenesis, prevention, and treatment. Lancet. 2006;368(9534):489-504.
38. Eisinger RW, Fauci AS. Ending the HIV/AIDS Pandemic(1). Emerg Infect Dis. 2018;24(3):413-6.
39. Remien RH, Stirratt MJ, Nguyen N, Robbins RN, Pala AN, Mellins CA. Mental health and HIV/AIDS: the need for an integrated response. Aids. 2019;33(9):1411-20.
40. Tran BX, Ho RCM, Ho CSH, Latkin CA, Phan HT, Ha GH, et al. Depression among Patients with HIV/AIDS: Research Development and Effective Interventions (GAP(RESEARCH)). Int J Environ Res Public Health. 2019;16(10).
41. Tran BX, Phan HT, Latkin CA, Nguyen HLT, Hoang CL, Ho CSH, et al. Understanding Global HIV Stigma and Discrimination: Are Contextual Factors Sufficiently Studied? (GAP(RESEARCH)). Int J Environ Res Public Health. 2019;16(11).
42. Sherr L, Clucas C, Harding R, Sibley E, Catalan J. HIV and depression--a systematic review of interventions. Psychol Health Med. 2011;16(5):493-527.
43. Pappas G, Kiriaze IJ, Giannakis P, Falagas ME. Psychosocial consequences of infectious diseases. Clin Microbiol Infect. 2009;15(8):743-7.
44. Tzeng N-S, Chung C-H, Chang C-C, Chang H-A, Kao Y-C, Chang S-Y, et al. What could we learn from SARS when facing the mental health issues related to the COVID-19 outbreak? A nationwide cohort study in Taiwan. Translational Psychiatry. 2020;10(1):339.
45. Maunder RG, Lancee WJ, Balderson KE, Bennett JP, Borgundvaag B, Evans S, et al. Long-term psychological and occupational effects of providing hospital healthcare during SARS outbreak. Emerg Infect Dis. 2006;12(12):1924-32.
46. Mak IW, Chu CM, Pan PC, Yiu MG, Chan VL. Long-term psychiatric morbidities among SARS survivors. Gen Hosp Psychiatry. 2009;31(4):318-26.
47. Dawood FS, Iuliano AD, Reed C, Meltzer MI, Shay DK, Cheng PY, et al. Estimated global mortality associated with the first 12 months of 2009 pandemic influenza A H1N1 virus circulation: a modelling study. Lancet Infect Dis. 2012;12(9):687-95.
48. Al Hajjar S, McIntosh K. The first influenza pandemic of the 21st century. Ann Saudi Med. 2010;30(1):1-10.
49. Chowell G, Bertozzi SM, Colchero MA, Lopez-Gatell H, Alpuche-Aranda C, Hernandez M, et al. Severe respiratory disease concurrent with the circulation of H1N1 influenza. N Engl J Med. 2009;361(7):674-9.
50. Bourrier MS, Deml MJ. The Legacy of the Pandemic Preparedness Regime: An Integrative Review. Int J Public Health. 2022;67:1604961.
51. Ransom J. Pandemic influenza preparedness, community engagement, and local public health practice. J Public Health Manag Pract. 2007;13(3):318-20.
52. Gerwin LE. The challenge of providing the public with actionable information during a pandemic. J Law Med Ethics. 2012;40(3):630-54.
53. Abeysinghe S. Vaccine Narratives and Public Health: Investigating Criticisms of H1N1 Pandemic Vaccination. PLoS Curr. 2015;7.
54. Leung GM, Nicoll A. Reflections on pandemic (H1N1) 2009 and the international response. PLoS Med. 2010;7(10).
55. Van Kerkhove MD, Asikainen T, Becker NG, Bjorge S, Desenclos JC, dos Santos T, et al. Studies needed to address public health challenges of the 2009 H1N1 influenza pandemic: insights from modeling. PLoS Med. 2010;7(6):e1000275.

56. Petersen LR, Jamieson DJ, Powers AM, Honein MA. Zika Virus. N Engl J Med. 2016;374(16):1552-63.
57. Wilson AE, Lehmann CU, Saleh SN, Hanna J, Medford RJ. Social media: A new tool for outbreak surveillance. Antimicrob Steward Healthc Epidemiol. 2021;1(1):e50.
58. McGough SF, Brownstein JS, Hawkins JB, Santillana M. Forecasting Zika Incidence in the 2016 Latin America Outbreak Combining Traditional Disease Surveillance with Search, Social Media, and News Report Data. PLoS Negl Trop Dis. 2017;11(1):e0005295.
59. Abouzahra M, Tan J. Twitter vs. Zika—The role of social media in epidemic outbreaks surveillance. Health Policy and Technology. 2021;10(1):174-81.
60. Odlum M, Yoon S. What can we learn about the Ebola outbreak from tweets? Am J Infect Control. 2015;43(6):563-71.
61. Fung IC-H, Tse ZTH, Cheung C-N, Miu AS, Fu K-W. Ebola and the social media. The Lancet. 2014;384(9961):2207.
62. Glowacki EM, Lazard AJ, Wilcox GB, Mackert M, Bernhardt JM. Identifying the public's concerns and the Centers for Disease Control and Prevention's reactions during a health crisis: An analysis of a Zika live Twitter chat. American Journal of Infection Control. 2016;44(12):1709-11.
63. Wood MJ. Propagating and Debunking Conspiracy Theories on Twitter During the 2015–2016 Zika Virus Outbreak. Cyberpsychology, Behavior, and Social Networking. 2018;21(8):485-90.
64. Sharma M, Yadav K, Yadav N, Ferdinand KC. Zika virus pandemic—analysis of Facebook as a social media health information platform. American Journal of Infection Control. 2017;45(3):301-2.
65. Seltzer EK, Horst-Martz E, Lu M, Merchant RM. Public sentiment and discourse about Zika virus on Instagram. Public Health. 2017;150:170-5.
66. Klofstad CA, Uscinski JE, Connolly JM, West JP. What drives people to believe in Zika conspiracy theories? Palgrave Communications. 2019;5(1):36.
67. Wang Y, McKee M, Torbica A, Stuckler D. Systematic Literature Review on the Spread of Health-related Misinformation on Social Media. Soc Sci Med. 2019;240:112552.
68. Bekkering S, Burgner D. Viruses and cardiovascular disease: from bad to worse. Nature Cardiovascular Research. 2022;1(7):601-2.
69. Shanks GD, Brundage JF. Pathogenic responses among young adults during the 1918 influenza pandemic. Emerg Infect Dis. 2012;18(2):201-7.
70. Morens DM, Fauci AS. The 1918 influenza pandemic: insights for the 21st century. J Infect Dis. 2007;195(7):1018-28.
71. Myrskylä M, Mehta NK, Chang VW. Early life exposure to the 1918 influenza pandemic and old-age mortality by cause of death. Am J Public Health. 2013;103(7):e83-90.
72. Helgertz J, Bengtsson T. The Long-Lasting Influenza: The Impact of Fetal Stress During the 1918 Influenza Pandemic on Socioeconomic Attainment and Health in Sweden, 1968-2012. Demography. 2019;56(4):1389-425.
73. Chow EJ, Rolfes MA, O'Halloran A, Anderson EJ, Bennett NM, Billing L, et al. Acute Cardiovascular Events Associated With Influenza in Hospitalized Adults. Annals of Internal Medicine. 2020;173(8):605-13.
74. Skaarup KG, Modin D, Nielsen L, Jensen JUS, Biering-Sørensen T. Influenza and cardiovascular disease pathophysiology: strings attached. European Heart Journal Supplements. 2023;25(Supplement_A):A5-A11.
75. Shah ASV, Stelzle D, Lee KK, Beck EJ, Alam S, Clifford S, et al. Global Burden of Atherosclerotic Cardiovascular Disease in People Living With HIV: Systematic Review and Meta-Analysis. Circulation. 2018;138(11):1100-12.
76. Baral N, Adhikari P, Adhikari G, Karki S. Influenza Myocarditis: A Literature Review. Cureus. 2020;12(12):e12007.
77. Løkke FB, Hansen KS, Dalgaard LS, Öbrink-Hansen K, Schiøttz-Christensen B, Leth S. Long-term complications after infection with SARS-CoV-1, influenza and MERS-CoV – Lessons to learn in long COVID? Infectious Diseases Now. 2023;53(8):104779.
78. Savedchuk S, Raslan R, Nystrom S, Sparks MA. Emerging Viral Infections and the Potential Impact on Hypertension, Cardiovascular Disease, and Kidney Disease. Circ Res. 2022;130(10):1618-41.
79. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. N Engl J Med. 2020;382(8):727-33.
80. Organisation WH. WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020. 2020.
81. Worldometers.info. [WORLDOMETER COVID-19 CORONAVIRUS PANDEMIC]. 2024 [updated 13 April 2024. Available from: https://www.worldometers.info/coronavirus/.

82. Wang D, Hu B, Hu C, Zhu F, Liu X, Zhang J, et al. Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China. Jama. 2020;323(11):1061-9.
83. Hu B, Guo H, Zhou P, Shi Z-L. Characteristics of SARS-CoV-2 and COVID-19. Nature Reviews Microbiology. 2021;19(3):141-54.
84. Hoffmann M, Kleine-Weber H, Schroeder S, Kruger N, Herrler T, Erichsen S, et al. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. Cell. 2020;181(2):271-80.e8.
85. Sparks MA, South AM, Badley AD, Baker-Smith CM, Batlle D, Bozkurt B, et al. Severe Acute Respiratory Syndrome Coronavirus 2, COVID-19, and the Renin-Angiotensin System: Pressing Needs and Best Research Practices. Hypertension. 2020;76(5):1350-67.
86. Fehr AR, Perlman S. Coronaviruses: An Overview of Their Replication and Pathogenesis. In: Maier HJ, Bickerton E, Britton P, editors. Coronaviruses: Methods and Protocols. New York, NY: Springer New York; 2015. p. 1-23.
87. Muhamad SA, Ugusman A, Kumar J, Skiba D, Hamid AA, Aminuddin A. COVID-19 and Hypertension: The What, the Why, and the How. Front Physiol. 2021;12:665064.
88. Guzik TJ, Mohiddin SA, Dimarco A, Patel V, Savvatis K, Marelli-Berg FM, et al. COVID-19 and the cardiovascular system: implications for risk assessment, diagnosis, and treatment options. Cardiovasc Res. 2020;116(10):1666-87.
89. Muralidar S, Ambi SV, Sekaran S, Krishnan UM. The emergence of COVID-19 as a global pandemic: Understanding the epidemiology, immune response and potential therapeutic targets of SARS-CoV-2. Biochimie. 2020;179:85-100.
90. Gheblawi M, Wang K, Viveiros A, Nguyen Q, Zhong JC, Turner AJ, et al. Angiotensin-Converting Enzyme 2: SARS-CoV-2 Receptor and Regulator of the Renin-Angiotensin System: Celebrating the 20th Anniversary of the Discovery of ACE2. Circ Res. 2020;126(10):1456-74.
91. Ferrario CM, Chappell MC, Tallant EA, Brosnihan KB, Diz DI. Counterregulatory Actions of Angiotensin-(1-7). Hypertension. 1997;30(3):535-41.
92. Masi S, Uliana M, Virdis A. Angiotensin II and vascular damage in hypertension: Role of oxidative stress and sympathetic activation. Vascular pharmacology. 2019;115:13-7.
93. Narula S, Yusuf S, Chong M, Ramasundarahettige C, Rangarajan S, Bangdiwala SI, et al. Plasma ACE2 and risk of death or cardiometabolic diseases: a case-cohort analysis. Lancet. 2020;396(10256):968-76.
94. Akpek M. Does COVID-19 Cause Hypertension? Angiology. 2022;73(7):682-7.
95. Chen G, Li X, Gong Z, Xia H, Wang Y, Wang X, et al. Hypertension as a sequela in patients of SARS-CoV-2 infection. PLoS One. 2021;16(4):e0250815.
96. Amraei R, Rahimi N. COVID-19, Renin-Angiotensin System and Endothelial Dysfunction. Cells. 2020;9(7).
97. Kuriakose J, Montezano Augusto C, Touyz Rhian M. ACE2/Ang-(1-7)/Mas1 axis and the vascular system: vasoprotection to COVID-19-associated vascular disease. Clinical Science. 2021;135(2):387-407.
98. Varga Z, Flammer AJ, Steiger P, Haberecker M, Andermatt R, Zinkernagel AS, et al. Endothelial cell infection and endotheliitis in COVID-19. Lancet. 2020;395(10234):1417-8.
99. Xu SW, Ilyas I, Weng JP. Endothelial dysfunction in COVID-19: an overview of evidence, biomarkers, mechanisms and potential therapies. Acta Pharmacol Sin. 2023;44(4):695-709.
100. Hu B, Huang S, Yin L. The cytokine storm and COVID-19. J Med Virol. 2021;93(1):250-6.
101. Gkaliagkousi E, Gavriilaki E, Triantafyllou A, Douma S. Clinical Significance of Endothelial Dysfunction in Essential Hypertension. Current Hypertension Reports. 2015;17(11):85.
102. Libby P, Lüscher T. COVID-19 is, in the end, an endothelial disease. European Heart Journal. 2020;41(32):3038-44.
103. Fodor A, Tiperciuc B, Login C, Orasan OH, Lazar AL, Buchman C, et al. Endothelial Dysfunction, Inflammation, and Oxidative Stress in COVID-19-Mechanisms and Therapeutic Targets. Oxid Med Cell Longev. 2021;2021:8671713.
104. Labarrere CA, Kassab GS. Glutathione deficiency in the pathogenesis of SARS-CoV-2 infection and its effects upon the host immune response in severe COVID-19 disease. Front Microbiol. 2022;13:979719.
105. Montezano AC, Camargo LL, Mary S, Neves KB, Rios FJ, Stein R, et al. SARS-CoV-2 spike protein induces endothelial inflammation via ACE2 independently of viral replication. Scientific Reports. 2023;13(1):14086.
106. Chang R, Mamun A, Dominic A, Le NT. SARS-CoV-2 Mediated Endothelial Dysfunction: The Potential Role of Chronic Oxidative Stress. Front Physiol. 2020;11:605908.

107. Zhang Z, Zhao L, Zhou X, Meng X, Zhou X. Role of inflammation, immunity, and oxidative stress in hypertension: New insights and potential therapeutic targets. Front Immunol. 2022;13:1098725.
108. Fletcher-Sandersjöö A, Bellander BM. Is COVID-19 associated thrombosis caused by overactivation of the complement cascade? A literature review. Thromb Res. 2020;194:36-41.
109. Rahman A, Tabassum T, Araf Y, Al Nahid A, Ullah MA, Hosen MJ. Silent hypoxia in COVID-19: pathomechanism and possible management strategy. Mol Biol Rep. 2021;48(4):3863-9.
110. Serebrovska ZO, Chong EY, Serebrovska TV, Tumanovska LV, Xi L. Hypoxia, HIF-1α, and COVID-19: from pathogenic factors to potential therapeutic targets. Acta Pharmacol Sin. 2020;41(12):1539-46.
111. Vecchione C, Carnevale D, Di Pardo A, Gentile MT, Damato A, Cocozza G, et al. Pressure-induced vascular oxidative stress is mediated through activation of integrin-linked kinase 1/betaPIX/Rac-1 pathway. Hypertension. 2009;54(5):1028-34.
112. Guan WJ, Ni ZY, Hu Y, Liang WH, Ou CQ, He JX, et al. Clinical Characteristics of Coronavirus Disease 2019 in China. N Engl J Med. 2020;382(18):1708-20.
113. Clark A, Jit M, Warren-Gash C, Guthrie B, Wang HHX, Mercer SW, et al. Global, regional, and national estimates of the population at increased risk of severe COVID-19 due to underlying health conditions in 2020: a modelling study. Lancet Glob Health. 2020;8(8):e1003-e17.
114. Hebbani AV, Pulakuntla S, Pannuru P, Aramgam S, Badri KR, Reddy VD. COVID-19: comprehensive review on mutations and current vaccines. Arch Microbiol. 2021;204(1):8.
115. Watts M. Community assessment of acute COVID-19 and referral to secondary car 2020 [cited 2924 15 May 2024]. Available from: https://www.sign.ac.uk/media/1792/community_assessment.pdf.
116. Maitra A, Sarkar MC, Raheja H, Biswas NK, Chakraborti S, Singh AK, et al. Mutations in SARS-CoV-2 viral RNA identified in Eastern India: Possible implications for the ongoing outbreak in India and impact on viral structure and host susceptibility. J Biosci. 2020;45.
117. Tao K, Tzou PL, Nouhin J, Gupta RK, de Oliveira T, Kosakovsky Pond SL, et al. The biological and clinical significance of emerging SARS-CoV-2 variants. Nat Rev Genet. 2021;22(12):757-73.
118. Islam MA. A review of SARS-CoV-2 variants and vaccines: Viral properties, mutations, vaccine efficacy, and safety. Infect Med (Beijing). 2023;2(4):247-61.
119. Prevention CfDCa. CDC, 2023; SARS-CoV-2 Variant Classifications and Definitions 2023 [updated 1 September 2023; cited 2024 28 April 2024]. 1 September 2023:[Available from: https://www.cdc.gov/coronavirus/2019-ncov/variants/variant-classifications.html.
120. Zhao S, Lou J, Chong MKC, Cao L, Zheng H, Chen Z, et al. Inferring the Association between the Risk of COVID-19 Case Fatality and N501Y Substitution in SARS-CoV-2. Viruses. 2021;13(4).
121. 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:n579.
122. Tegally H, Wilkinson E, Giovanetti M, Iranzadeh A, Fonseca V, Giandhari J, et al. Detection of a SARS-CoV-2 variant of concern in South Africa. Nature. 2021;592(7854):438-43.
123. Moyo-Gwete T, Madzivhandila M, Makhado Z, Ayres F, Mhlanga D, Oosthuysen B, et al. Cross-Reactive Neutralizing Antibody Responses Elicited by SARS-CoV-2 501Y.V2 (B.1.351). New England Journal of Medicine. 2021;384(22):2161-3.
124. Faria NR, Mellan TA, Whittaker C, Claro IM, Candido DDS, Mishra S, et al. Genomics and epidemiology of the P.1 SARS-CoV-2 lineage in Manaus, Brazil. Science. 2021;372(6544):815-21.
125. Planas D, Veyer D, Baidaliuk A, Staropoli I, Guivel-Benhassine F, Rajah MM, et al. Reduced sensitivity of SARS-CoV-2 variant Delta to antibody neutralization. Nature. 2021;596(7871):276-80.
126. Chakraborty C, Bhattacharya M, Chopra H, Bhattacharya P, Islam MA, Dhama K. Recently emerged omicron subvariant BF.7 and its R346T mutation in the RBD region reveal increased transmissibility and higher resistance to neutralization antibodies: need to understand more under the current scenario of rising cases in China and fears of driving a new wave of the COVID-19 pandemic. Int J Surg. 2023;109(4):1037-40.
127. Soheili M, Khateri S, Moradpour F, Mohammadzedeh P, Zareie M, Mortazavi SMM, et al. The efficacy and effectiveness of COVID-19 vaccines around the world: a mini-review and meta-analysis. Ann Clin Microbiol Antimicrob. 2023;22(1):42.
128. Wang Y, Zhang D, Du G, Du R, Zhao J, Jin Y, et al. Remdesivir in adults with severe COVID-19: a randomised, double-blind, placebo-controlled, multicentre trial. Lancet. 2020;395(10236):1569-78.

129. Jayk Bernal A, Gomes da Silva MM, Musungaie DB, Kovalchuk E, Gonzalez A, Delos Reyes V, et al. Molnupiravir for Oral Treatment of Covid-19 in Nonhospitalized Patients. N Engl J Med. 2022;386(6):509-20.
130. Hammond J, Leister-Tebbe H, Gardner A, Abreu P, Bao W, Wisemandle W, et al. Oral Nirmatrelvir for High-Risk, Nonhospitalized Adults with Covid-19. N Engl J Med. 2022;386(15):1397-408.
131. Horby P, Lim WS, Emberson JR, Mafham M, Bell JL, Linsell L, et al. Dexamethasone in Hospitalized Patients with Covid-19. N Engl J Med. 2021;384(8):693-704.
132. Markowicz S, Trioux T, Rulquin C, Le Guillou C, Ouissa R, Loraux C, et al. Real-world effectiveness of casirivimab plus indevimab in a dedicated ambulatory unit created for patients with early COVID-19 during a massive delta variant wave. Eur J Clin Microbiol Infect Dis. 2022;41(8):1145-9.
133. Ospina-Tascón GA, Calderón-Tapia LE, García AF, Zarama V, Gómez-Álvarez F, Álvarez-Saa T, et al. Effect of High-Flow Oxygen Therapy vs Conventional Oxygen Therapy on Invasive Mechanical Ventilation and Clinical Recovery in Patients With Severe COVID-19: A Randomized Clinical Trial. Jama. 2021;326(21):2161-71.
134. Cuker A, Tseng EK, Nieuwlaat R, Angchaisuksiri P, Blair C, Dane K, et al. American Society of Hematology 2021 guidelines on the use of anticoagulation for thromboprophylaxis in patients with COVID-19. Blood Adv. 2021;5(3):872-88.
135. Tobian AAR, Cohn CS, Shaz BH. COVID-19 convalescent plasma. Blood. 2022;140(3):196-207.
136. Mohamed K, Rzymski P, Islam MS, Makuku R, Mushtaq A, Khan A, et al. COVID-19 vaccinations: The unknowns, challenges, and hopes. J Med Virol. 2022;94(4):1336-49.
137. Organization WH. Vaccine introduction guidelines. Adding a vaccine to a national immunization programme: decision and implementation 2014 [updated 1 April 2014; cited 2024 29 April 2024]. Available from: https://www.who.int/publications/i/item/9789241506892.
138. Bonanni P, Sacco C, Donato R, Capei R. Lifelong vaccination as a key disease-prevention strategy. Clinical Microbiology and Infection. 2014;20:32-6.
139. Manriquez Roa T, Holzer F, Luna F, Biller-Andorno N. Expert Views on COVAX and Equitable Global Access to COVID-19 Vaccines. Int J Public Health. 2021;66:1604236.
140. Wu K, Werner AP, Koch M, Choi A, Narayanan E, Stewart-Jones GBE, et al. Serum Neutralizing Activity Elicited by mRNA-1273 Vaccine. N Engl J Med. 2021;384(15):1468-70.
141. Voysey M, Clemens SAC, Madhi SA, Weckx LY, Folegatti PM, Aley PK, et al. Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK. Lancet. 2021;397(10269):99-111.
142. Logunov DY, Dolzhikova IV, Shcheblyakov DV, Tukhvatulin AI, Zubkova OV, Dzharullaeva AS, et al. Safety and efficacy of an rAd26 and rAd5 vector-based heterologous prime-boost COVID-19 vaccine: an interim analysis of a randomised controlled phase 3 trial in Russia. Lancet. 2021;397(10275):671-81.
143. Tanriover MD, Doğanay HL, Akova M, Güner HR, Azap A, Akhan S, et al. Efficacy and safety of an inactivated whole-virion SARS-CoV-2 vaccine (CoronaVac): interim results of a double-blind, randomised, placebo-controlled, phase 3 trial in Turkey. Lancet. 2021;398(10296):213-22.
144. Keech C, Albert G, Cho I, Robertson A, Reed P, Neal S, et al. Phase 1-2 Trial of a SARS-CoV-2 Recombinant Spike Protein Nanoparticle Vaccine. N Engl J Med. 2020;383(24):2320-32.
145. Singh TK, Zidar DA, McCrae K, Highland KB, Englund K, Cameron SJ, et al. A Post-Pandemic Enigma: The Cardiovascular Impact of Post-Acute Sequelae of SARS-CoV-2. Circ Res. 2023;132(10):1358-73.
146. National Institute for Health and Care Excellence: Clinical Guidelines. COVID-19 rapid guideline: managing the long-term effects of COVID-19. London: National Institute for Health and Care Excellence (NICE)

Copyright © NICE 2020.; 2020.
147. Davis HE, McCorkell L, Vogel JM, Topol EJ. Long COVID: major findings, mechanisms and recommendations. Nat Rev Microbiol. 2023;21(3):133-46.
148. Sudre CH, Murray B, Varsavsky T, Graham MS, Penfold RS, Bowyer RC, et al. Attributes and predictors of long COVID. Nature Medicine. 2021;27(4):626-31.
149. Brodin P, Casari G, Townsend L, O'Farrelly C, Tancevski I, Löffler-Ragg J, et al. Studying severe long COVID to understand post-infectious disorders beyond COVID-19. Nat Med. 2022;28(5):879-82.

150. Guo L, Wang G, Wang Y, Zhang Q, Ren L, Gu X, et al. SARS-CoV-2-specific antibody and T-cell responses 1 year after infection in people recovered from COVID-19: a longitudinal cohort study. The Lancet Microbe. 2022;3(5):e348-e56.
151. Wiech M, Chroscicki P, Swatler J, Stepnik D, De Biasi S, Hampel M, et al. Remodeling of T cell dynamics during long COVID is dependent on severity of SARS-CoV-2 infection. Frontiers in Immunology. 2022;13:886431.
152. Gupta A, Satapathy AK, Bahinipati P. Delayed catastrophic thrombotic events in post-acute COVID-19. Thrombosis Research. 2022;220:60-4.
153. Borrelli MP, Buora A, Frigatti P. The long haul COVID-19 arterial thrombosis. Thrombosis Research. 2022;217:73-5.
154. Osiaevi I, Schulze A, Evers G, Harmening K, Vink H, Kümpers P, et al. Persistent capillary rarefication in long COVID syndrome. Angiogenesis. 2023;26(1):53-61.
155. Martins-Gonçalves R, Campos MM, Palhinha L, Azevedo-Quintanilha IG, Abud Mendes M, Ramos Temerozo J, et al. Persisting platelet activation and hyperactivity in COVID-19 survivors. Circulation Research. 2022;131(11):944-7.
156. Hempel S, Chambers D, Bagnall A-M, Forbes C. Risk factors for chronic fatigue syndrome/myalgic encephalomyelitis: a systematic scoping review of multiple predictor studies. Psychological medicine. 2008;38(7):915-26.
157. Korompoki E, Gavriatopoulou M, Hicklen RS, Ntanasis-Stathopoulos I, Kastritis E, Fotiou D, et al. Epidemiology and organ specific sequelae of post-acute COVID19: a narrative review. Journal of Infection. 2021;83(1):1-16.
158. Yong SJ. Persistent brainstem dysfunction in long-COVID: a hypothesis. ACS chemical neuroscience. 2021;12(4):573-80.
159. Crivelli L, Palmer K, Calandri I, Guekht A, Beghi E, Carroll W, et al. Changes in cognitive functioning after COVID-19: A systematic review and meta-analysis. Alzheimers Dement. 2022;18(5):1047-66.
160. Peluso MJ, Deveau T-M, Munter SE, Ryder D, Buck A, Beck-Engeser G, et al. Impact of pre-existing chronic viral infection and reactivation on the development of long COVID. MedRxiv. 2022.
161. Liu Y, Gu X, Li H, Zhang H, Xu J. Mechanisms of long COVID: An updated review. Chinese Medical Journal Pulmonary and Critical Care Medicine. 2023;1(4):231-40.
162. Zhang H, Li X, Huang L, Gu X, Wang Y, Liu M, et al. Lung-function trajectories in COVID-19 survivors after discharge: a two-year longitudinal cohort study. EClinicalMedicine. 2022;54.
163. Faverio P, Luppi F, Rebora P, D'Andrea G, Stainer A, Busnelli S, et al. One-year pulmonary impairment after severe COVID-19: a prospective, multicenter follow-up study. Respiratory Research. 2022;23(1):65.
164. Aranda J, Oriol I, Feria L, Abelenda G, Rombauts A, Simonetti AF, et al. Persistent COVID-19 symptoms 1 year after hospital discharge: A prospective multicenter study. Plos one. 2022;17(10):e0275615.
165. Taquet M, Sillett R, Zhu L, Mendel J, Camplisson I, Dercon Q, et al. Neurological and psychiatric risk trajectories after SARS-CoV-2 infection: an analysis of 2-year retrospective cohort studies including 1 284 437 patients. Lancet Psychiatry. 2022;9(10):815-27.
166. Taquet M, Geddes JR, Husain M, Luciano S, Harrison PJ. 6-month neurological and psychiatric outcomes in 236 379 survivors of COVID-19: a retrospective cohort study using electronic health records. Lancet Psychiatry. 2021;8(5):416-27.
167. Ley H, Skorniewska Z, Harrison PJ, Taquet M. Risks of neurological and psychiatric sequelae 2 years after hospitalisation or intensive care admission with COVID-19 compared to admissions for other causes. Brain Behav Immun. 2023;112:85-95.
168. Hiscott J, Alexandridi M, Muscolini M, Tassone E, Palermo E, Soultsioti M, et al. The global impact of the coronavirus pandemic. Cytokine Growth Factor Rev. 2020;53:1-9.
169. Wang X, Wu W, Song P, He J. An international comparison analysis of reserve and supply system for emergency medical supplies between China, the United States, Australia, and Canada. Biosci Trends. 2020;14(4):231-40.
170. Gualano MR, Sinigaglia T, Lo Moro G, Rousset S, Cremona A, Bert F, et al. The Burden of Burnout among Healthcare Professionals of Intensive Care Units and Emergency Departments during the COVID-19 Pandemic: A Systematic Review. Int J Environ Res Public Health. 2021;18(15).
171. Lambert H, Gupte J, Fletcher H, Hammond L, Lowe N, Pelling M, et al. COVID-19 as a global challenge: towards an inclusive and sustainable future. Lancet Planet Health. 2020;4(8):e312-e4.

172. Mallah SI, Ghorab OK, Al-Salmi S, Abdellatif OS, Tharmaratnam T, Iskandar MA, et al. COVID-19: breaking down a global health crisis. Ann Clin Microbiol Antimicrob. 2021;20(1):35.
173. Propper C, Stoye G, Zaranko B. The Wider Impacts of the Coronavirus Pandemic on the NHS. Fisc Stud. 2020;41(2):345-56.
174. Formenti B, Gregori N, Crosato V, Marchese V, Tomasoni LR, Castelli F. The impact of COVID-19 on communicable and non-communicable diseases in Africa: a narrative review. Infez Med. 2022;30(1):30-40.
175. Kamble S, Joshi A, Kamble R, Kumari S. Influence of COVID-19 Pandemic on Psychological Status: An Elaborate Review. Cureus. 2022;14(10):e29820.
176. Clemente-Suárez VJ, Martínez-González MB, Benitez-Agudelo JC, Navarro-Jiménez E, Beltran-Velasco AI, Ruisoto P, et al. The Impact of the COVID-19 Pandemic on Mental Disorders. A Critical Review. Int J Environ Res Public Health. 2021;18(19).
177. International Monetary Fund (IMF). (2021). World Economic Outlook Update, January 2021: Policy Support and Vaccines Expected to Lift Activity. 2021 [updated January 2021; cited 2024 29/2/24]. Available from: https://www.imf.org/en/Publications/WEO/Issues/2021/01/26/2021-world-economic-outlook-update.
178. Library UPHoC. Coronavirus: Economic impact 2021 [cited 2024 30 April 2024]. Available from: https://commonslibrary.parliament.uk/research-briefings/cbp-8866/.
179. Bonotti M, Zech ST. The Human, Economic, Social, and Political Costs of COVID-19. Recovering Civility during COVID-19. 2021:1-36.
180. Tai DBG, Sia IG, Doubeni CA, Wieland ML. Disproportionate Impact of COVID-19 on Racial and Ethnic Minority Groups in the United States: a 2021 Update. J Racial Ethn Health Disparities. 2022;9(6):2334-9.
181. Barron GC, Laryea-Adjei G, Vike-Freiberga V, Abubakar I, Dakkak H, Devakumar D, et al. Safeguarding people living in vulnerable conditions in the COVID-19 era through universal health coverage and social protection. Lancet Public Health. 2022;7(1):e86-e92.
182. Yarlagadda H, Patel MA, Gupta V, Bansal T, Upadhyay S, Shaheen N, et al. COVID-19 Vaccine Challenges in Developing and Developed Countries. Cureus. 2022;14(4):e23951.
183. Ngai CSB, Singh RG, Yao L. Impact of COVID-19 Vaccine Misinformation on Social Media Virality: Content Analysis of Message Themes and Writing Strategies. J Med Internet Res. 2022;24(7):e37806.
184. (NCD-RisC) NRFC. Worldwide trends in hypertension prevalence and progress in treatment and control from 1990 to 2019: a pooled analysis of 1201 population-representative studies with 104 million participants. Lancet. 2021;398(10304):957-80.
185. Forouzanfar MH, Liu P, Roth GA, Ng M, Biryukov S, Marczak L, et al. Global Burden of Hypertension and Systolic Blood Pressure of at Least 110 to 115 mm Hg, 1990-2015. Jama. 2017;317(2):165-82.
186. Kole C, Stefanou E, Karvelas N, Schizas D, Toutouzas KP. Acute and Post-Acute COVID-19 Cardiovascular Complications: A Comprehensive Review. Cardiovascular Drugs and Therapy. 2023.
187. Petterson JL, McPhee BN, Wu Y, O'Brien MW. Does COVID-19 influence the sympathetic regulation of blood pressure? J Physiol. 2021;599(22):4951-3.
188. Li B, Yang J, Zhao F, Zhi L, Wang X, Liu L, et al. Prevalence and impact of cardiovascular metabolic diseases on COVID-19 in China. Clin Res Cardiol. 2020;109(5):531-8.
189. Xu XW, Wu XX, Jiang XG, Xu KJ, Ying LJ, Ma CL, et al. Clinical findings in a group of patients infected with the 2019 novel coronavirus (SARS-Cov-2) outside of Wuhan, China: retrospective case series. Bmj. 2020;368:m606.
190. Wei ZY, Qiao R, Chen J, Huang J, Wu H, Wang WJ, et al. The influence of pre-existing hypertension on coronavirus disease 2019 patients. Epidemiol Infect. 2021;149:e4.
191. Liu J, Huang L, Wei W, Bai Y, Chang E, Leng Y. Effects of Antihypertensive Agents on the Clinical Outcome of Hospitalized COVID-19 Patients Concomitant with Hypertension: A Systematic Review and Meta-Analysis. Heart Lung. 2024;63:78-85.
192. Mancia G, Kreutz R, Brunström M, Burnier M, Grassi G, Januszewicz A, et al. 2023 ESH Guidelines for the management of arterial hypertension The Task Force for the management of arterial hypertension of the European Society of Hypertension: Endorsed by the International Society of Hypertension (ISH) and the European Renal Association (ERA). J Hypertens. 2023;41(12):1874-2071.
193. Wang C, Horby PW, Hayden FG, Gao GF. A novel coronavirus outbreak of global health concern. Lancet. 2020;395(10223):470-3.

194. Shibata S, Arima H, Asayama K, Hoshide S, Ichihara A, Ishimitsu T, et al. Hypertension and related diseases in the era of COVID-19: a report from the Japanese Society of Hypertension Task Force on COVID-19. Hypertens Res. 2020;43(10):1028-46.
195. Zeigler Z. COVID-19 Self-quarantine and Weight Gain Risk Factors in Adults. Curr Obes Rep. 2021;10(3):423-33.
196. Park AH, Zhong S, Yang H, Jeong J, Lee C. Impact of COVID-19 on physical activity: A rapid review. J Glob Health. 2022;12:05003.
197. Laffin LJ, Kaufman HW, Chen Z, Niles JK, Arellano AR, Bare LA, et al. Rise in Blood Pressure Observed Among US Adults During the COVID-19 Pandemic. Circulation. 2022;145(3):235-7.
198. Castanon A, Grasic K, Chen S, Ma F, Oboli G, Bray BD, et al. The impact of the COVID-19 pandemic on cardiovascular disease prevention and corresponding geographical inequalities in England: interrupted time series analysis. BMC Public Health. 2023;23(1):2443.
199. Gotanda H, Liyanage-Don N, Moran AE, Krousel-Wood M, Green JB, Zhang Y, et al. Changes in Blood Pressure Outcomes Among Hypertensive Individuals During the COVID-19 Pandemic: A Time Series Analysis in Three US Healthcare Organizations. Hypertension. 2022;79(12):2733-42.
200. Milan W, Victoria S, Christine C, Rose H, Helen JC, Colm A, et al. Impact of COVID-19 on recorded blood pressure screening and hypertension management in England: An analysis of monthly changes in Quality and Outcomes Framework indicators in OpenSAFELY. medRxiv. 2023:2023.07.20.23292883.
201. Mancia G, Rea F, Ludergnani M, Apolone G, Corrao G. Renin-Angiotensin-Aldosterone System Blockers and the Risk of Covid-19. N Engl J Med. 2020;382(25):2431-40.
202. Ruksakulpiwat S, Zhou W, Niyomyart A, Wang T, Kudlowitz A. How does the COVID-19 pandemic impact medication adherence of patients with chronic disease?: A systematic review. Chronic Illn. 2023;19(3):495-513.
203. Nozato Y, Yamamoto K, Rakugi H. Hypertension management before and under the COVID-19 pandemic: lessons and future directions. Hypertens Res. 2023;46(6):1471-7.
204. Savoia C, Volpe M, Kreutz R. Hypertension, a Moving Target in COVID-19: Current Views and Perspectives. Circ Res. 2021;128(7):1062-79.
205. Msemburi W, Karlinsky A, Knutson V, Aleshin-Guendel S, Chatterji S, Wakefield J. The WHO estimates of excess mortality associated with the COVID-19 pandemic. Nature. 2023;613(7942):130-7.
206. Islam N. "Excess deaths" is the best metric for tracking the pandemic. Bmj. 2022;376:o285.
207. Wrona M, Skrypnik D. New-Onset Diabetes Mellitus, Hypertension, Dyslipidaemia as Sequelae of COVID-19 Infection-Systematic Review. Int J Environ Res Public Health. 2022;19(20).
208. Trimarco V, Izzo R, Pacella D, Trama U, Manzi MV, Lombardi A, et al. Incidence of new-onset hypertension before, during, and after the COVID-19 pandemic: a 7-year longitudinal cohort study in a large population. BMC Med. 2024;22(1):127.
209. Zuin M, Rigatelli G, Bilato C, Pasquetto G, Mazza A. Risk of Incident New-Onset Arterial Hypertension After COVID-19 Recovery: A Systematic Review and Meta-analysis. High Blood Press Cardiovasc Prev. 2023;30(3):227-33.
210. Zhang HG, Dagliati A, Shakeri Hossein Abad Z, Xiong X, Bonzel CL, Xia Z, et al. International electronic health record-derived post-acute sequelae profiles of COVID-19 patients. NPJ Digit Med. 2022;5(1):81.
211. Ogungbe O, Gilotra NA, Davidson PM, Farley JE, Dennison Himmelfarb CR, Post WS, et al. Cardiac postacute sequelae symptoms of SARS-CoV-2 in community-dwelling adults: cross-sectional study. Open Heart. 2022;9(2).
212. Zhang V, Fisher M, Hou W, Zhang L, Duong TQ. Incidence of New-Onset Hypertension Post-COVID-19: Comparison With Influenza. Hypertension. 2023;80(10):2135-48.
213. Wang C, Fang B, Regmi A, Yamaguchi Y, Yang L, Cai Y. Text mining online disinformation about antihypertensive agents ACEI/ARB and COVID-19 on Sina Weibo. J Glob Health. 2023;13:06028.
214. [The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (COVID-19) in China]. Zhonghua Liu Xing Bing Xue Za Zhi. 2020;41(2):145-51.
215. Reynolds HR, Adhikari S, Pulgarin C, Troxel AB, Iturrate E, Johnson SB, et al. Renin-Angiotensin-Aldosterone System Inhibitors and Risk of Covid-19. N Engl J Med. 2020;382(25):2441-8.
216. Fosbol EL, Butt JH, Ostergaard L, Andersson C, Selmer C, Kragholm K, et al. Association of Angiotensin-Converting Enzyme Inhibitor or Angiotensin Receptor Blocker Use With COVID-19 Diagnosis and Mortality. Jama. 2020;324(2):168-77.

217. Imai Y, Kuba K, Rao S, Huan Y, Guo F, Guan B, et al. Angiotensin-converting enzyme 2 protects from severe acute lung failure. Nature. 2005;436(7047):112-6.
218. Kjeldsen SE, Narkiewicz K, Burnier M, Oparil S. Potential protective effects of antihypertensive treatments during the Covid-19 pandemic: from inhibitors of the renin-angiotensin system to beta-adrenergic receptor blockers. Blood Press. 2021;30(1):1-3.
219. Jardine MJ, Kotwal SS, Bassi A, Hockham C, Jones M, Wilcox A, et al. Angiotensin receptor blockers for the treatment of covid-19: pragmatic, adaptive, multicentre, phase 3, randomised controlled trial. BMJ. 2022;379:e072175.
220. Macedo AVS, de Barros ESPGM, de Paula TC, Moll-Bernardes RJ, Mendonca Dos Santos T, Mazza L, et al. Discontinuing vs continuing ACEIs and ARBs in hospitalized patients with COVID-19 according to disease severity: Insights from the BRACE CORONA trial. Am Heart J. 2022;249:86-97.
221. Liu Q, Fu W, Zhu CJ, Ding ZH, Dong BB, Sun BQ, et al. Effect of continuing the use of renin-angiotensin system inhibitors on mortality in patients hospitalized for coronavirus disease 2019: a systematic review, meta-analysis, and meta-regression analysis. BMC Infect Dis. 2023;23(1):53.
222. Savarese G, Benson L, Sundstrom J, Lund LH. Association between renin-angiotensin-aldosterone system inhibitor use and COVID-19 hospitalization and death: a 1.4 million patient nationwide registry analysis. Eur J Heart Fail. 2021;23(3):476-85.
223. Morales DR, Conover MM, You SC, Pratt N, Kostka K, Duarte-Salles T, et al. Renin-angiotensin system blockers and susceptibility to COVID-19: a multinational open science cohort study. medRxiv. 2020.
224. Jeffery MM, Cummins NW, Dempsey TM, Limper AH, Shah ND, Bellolio F. Association of outpatient ACE inhibitors and angiotensin receptor blockers and outcomes of acute respiratory illness: a retrospective cohort study. BMJ Open. 2021;11(3):e044010.
225. Semenzato L, Botton J, Drouin J, Baricault B, Vabre C, Cuenot F, et al. Antihypertensive Drugs and COVID-19 Risk: A Cohort Study of 2 Million Hypertensive Patients. Hypertension. 2021;77(3):833-42.
226. Society BaIH. BIHS Statement on ACEi and ARB and COVID-19 (16 March 2020) 2020 [cited 2024 15 May 2024]. Available from: https://bihsoc.org/bihs-statement-on-acei-arb-and-covid19/.
227. (ISH) ISoH. A statement from the International Society of Hypertension on COVID-19 2020 [cited 2024 15 May 2024]. Available from: https://ish-world.com/a-statement-from-the-international-society-of-hypertension-on-covid-19/.
228. Cardiology ESo. Position Statement of the ESC Council on Hypertension on ACE-Inhibitors and Angiotensin Receptor Blockers 2020 [cited 2024 28/2/24]. Available from: https://www.escardio.org/Councils/Council-on-Hypertension-(CHT)/News/position-statement-of-the-esc-council-on-hypertension-on-ace-inhibitors-and-ang.
229. Bozkurt B, Kovacs R, Harrington B. Joint HFSA/ACC/AHA Statement Addresses Concerns Re: Using RAAS Antagonists in COVID-19. J Card Fail. 2020;26(5):370.
230. Lip S, McCallum L, Delles C, McClure JD, Guzik T, Berry C, et al. Rationale and Design for the LOnger-term effects of SARS-CoV-2 INfection on blood Vessels And blood pRessure (LOCHINVAR): an observational phenotyping study. Open Heart. 2022;9(1).
231. Touyz R. COVID-19 Blood Pressure Endothelium Interaction Study (OBELIX) 1 June 2020 [cited 2024 11 March 2024]. Available from: https://classic.clinicaltrials.gov/ct2/show/NCT04409847.
232. (BIHS) BaIHS. [updated 24 May 2024; cited 2024 24 May 2024]. Available from: https://bihsoc.org/guidelines/.
233. Australia LF. [cited 2024 24 May 2024]. Available from: https://pulmonaryrehab.com.au/patient-assessment/assessing-exercise-capacity/the-six-minute-walk-test-6mwt/.
234. Alison JA, McKeough ZJ, Johnston K, McNamara RJ, Spencer LM, Jenkins SC, et al. Australian and New Zealand Pulmonary Rehabilitation Guidelines. Respirology. 2017;22(4):800-19.
235. Devlin N PD, Janssen B. Methods for Analysing and Reporting EQ-5D Data [Internet] Chapter 4, Analysis of EQ-5D Values 2020 [cited 2024 27 May 2024]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK565677/ doi: 10.1007/978-3-030-47622-9_4.
236. McCaffrey N, Kaambwa B, Currow DC, Ratcliffe J. Health-related quality of life measured using the EQ-5D-5L: South Australian population norms. Health Qual Life Outcomes. 2016;14(1):133.

237. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res. 2011;20(10):1727-36.
238. Mansiroglu AK, Seymen H, Sincer I, Gunes Y. Evaluation of Endothelial Dysfunction in COVID-19 With Flow-Mediated Dilatation. Arq Bras Cardiol. 2022;119(2):319-25.
239. Nandadeva D, Skow RJ, Stephens BY, Grotle AK, Georgoudiou S, Barshikar S, et al. Cardiovascular and cerebral vascular health in females with postacute sequelae of COVID-19. Am J Physiol Heart Circ Physiol. 2023;324(6):H713-H20.
240. Chen J, Wu C, Wang X, Yu J, Sun Z. The Impact of COVID-19 on Blood Glucose: A Systematic Review and Meta-Analysis. Front Endocrinol (Lausanne). 2020;11:574541.
241. Banerjee M, Pal R, Dutta S. Risk of incident diabetes post-COVID-19: A systematic review and meta-analysis. Prim Care Diabetes. 2022;16(4):591-3.
242. Gerganova A, Assyov Y, Kamenov Z. Stress Hyperglycemia, Diabetes Mellitus and COVID-19 Infection: Risk Factors, Clinical Outcomes and Post-Discharge Implications. Front Clin Diabetes Healthc. 2022;3:826006.
243. Tadic M, Cuspidi C, Grassi G, Mancia G. COVID-19 and arterial hypertension: Hypothesis or evidence? J Clin Hypertens (Greenwich). 2020;22(7):1120-6.
244. Ozcan S, Ince O, Guner A, Katkat F, Donmez E, Tugrul S, et al. Long-Term Clinical Consequences of Patients Hospitalized for COVID-19 Infection. Anatol J Cardiol. 2022;26(4):305-15.
245. Al-Aly Z, Xie Y, Bowe B. High-dimensional characterization of post-acute sequelae of COVID-19. Nature. 2021;594(7862):259-64.
246. Fernandez-Ortega MA, Ponce-Rosas ER, Muniz-Salinas DA, Rodriguez-Mendoza O, Najera Chavez P, Sanchez-Pozos V, et al. Cognitive dysfunction, diabetes mellitus 2 and arterial hypertension: Sequelae up to one year of COVID-19. Travel Med Infect Dis. 2023;52:102553.
247. Vyas P, Joshi D, Sharma V, Parmar M, Vadodariya J, Patel K, et al. Incidence and predictors of development of new onset hypertension post COVID-19 disease. Indian Heart J. 2023;75(5):347-51.
248. Alfadda AA, Rafiullah M, Alkhowaiter M, Alotaibi N, Alzahrani M, Binkhamis K, et al. Clinical and biochemical characteristics of people experiencing post-coronavirus disease 2019-related symptoms: A prospective follow-up investigation. Front Med (Lausanne). 2022;9:1067082.
249. De Lorenzo R, Conte C, Lanzani C, Benedetti F, Roveri L, Mazza MG, et al. Residual clinical damage after COVID-19: A retrospective and prospective observational cohort study. PLoS One. 2020;15(10):e0239570.
250. Gameil MA, Marzouk RE, Elsebaie AH, Rozaik SE. Long-term clinical and biochemical residue after COVID-19 recovery. Egypt Liver J. 2021;11(1):74.
251. Harenwall S, Heywood-Everett S, Henderson R, Godsell S, Jordan S, Moore A, et al. Post-Covid-19 Syndrome: Improvements in Health-Related Quality of Life Following Psychology-Led Interdisciplinary Virtual Rehabilitation. J Prim Care Community Health. 2021;12:21501319211067674.
252. Rosa RG, Cavalcanti AB, Azevedo LCP, Veiga VC, de Souza D, Dos Santos R, et al. Association between acute disease severity and one-year quality of life among post-hospitalisation COVID-19 patients: Coalition VII prospective cohort study. Intensive Care Med. 2023;49(2):166-77.
253. Valent A, Dudoignon E, Ressaire Q, Depret F, Plaud B. Three-month quality of life in survivors of ARDS due to COVID-19: A preliminary report from a French academic centre. Anaesth Crit Care Pain Med. 2020;39(6):740-1.
254. van Gassel RJJ, Bels J, Remij L, van Bussel BCT, Posthuma R, Gietema HA, et al. Functional Outcomes and Their Association With Physical Performance in Mechanically Ventilated Coronavirus Disease 2019 Survivors at 3 Months Following Hospital Discharge: A Cohort Study. Crit Care Med. 2021;49(10):1726-38.
255. Garrigues E, Janvier P, Kherabi Y, Le Bot A, Hamon A, Gouze H, et al. Post-discharge persistent symptoms and health-related quality of life after hospitalization for COVID-19. J Infect. 2020;81(6):e4-e6.
256. de Oliveira Almeida K, Nogueira Alves IG, de Queiroz RS, de Castro MR, Gomes VA, Santos Fontoura FC, et al. A systematic review on physical function, activities of daily living and health-related quality of life in COVID-19 survivors. Chronic Illn. 2023;19(2):279-303.
257. McCaffrey DF, Ridgeway G, Morral AR. Propensity score estimation with boosted regression for evaluating causal effects in observational studies. Psychol Methods. 2004;9(4):403-25.
258. Lee BK, Lessler J, Stuart EA. Improving propensity score weighting using machine learning. Stat Med. 2010;29(3):337-46.

259. McCaffrey DF, Griffin BA, Almirall D, Slaughter ME, Ramchand R, Burgette LF. A tutorial on propensity score estimation for multiple treatments using generalized boosted models. Stat Med. 2013;32(19):3388-414.
260. Griffin BA, McCaffrey D, Almirall D, Setodji C, Burgette L. Chasing balance and other recommendations for improving nonparametric propensity score models. J Causal Inference. 2017;5(2).
261. Wu Y, Meng X, Zhang J, He Y, Romo JA, Dong Y, et al. Effective LSTMs with seasonal-trend decomposition and adaptive learning and niching-based backtracking search algorithm for time series forecasting. Expert Systems with Applications. 2024;236:121202.
262. Jiang Z. Deep Learning for Spatiotemporal Big Data: A Vision on Opportunities and Challenges. ArXiv. 2023;abs/2310.19957.
263. Islam S, Elmekki H, Elsebai A, Bentahar J, Drawel N, Rjoub G, et al. A comprehensive survey on applications of transformers for deep learning tasks. Expert Systems with Applications. 2024;241:122666.
264. Rubin DB. Estimating causal effects of treatments in randomized and nonrandomized studies. Journal of Educational Psychology. 1974;66:688-701.
265. Hernan MA, Robins JM. Causal Inference: What If: CRC Press; 2023.
266. Künzel SR, Sekhon JS, Bickel PJ, Yu B. Metalearners for estimating heterogeneous treatment effects using machine learning. Proceedings of the National Academy of Sciences. 2019;116(10):4156-65.
267. Morgan SL, Winship C, editors. Counterfactuals and Causal Inference: Methods and Principles for Social Research2007.
268. Macedo AVS, de Barros ESPGM, de Paula TC, Moll-Bernardes RJ, Mendonça Dos Santos T, Mazza L, et al. Discontinuing vs continuing ACEIs and ARBs in hospitalized patients with COVID-19 according to disease severity: Insights from the BRACE CORONA trial. Am Heart J. 2022;249:86-97.
269. Savarese G, Benson L, Sundström J, Lund LH. Association between renin-angiotensin-aldosterone system inhibitor use and COVID-19 hospitalization and death: a 1.4 million patient nationwide registry analysis. Eur J Heart Fail. 2021;23(3):476-85.
270. Al-Kuraishy HM, Al-Gareeb AI, Mostafa-Hedeab G, Kasozi KI, Zirintunda G, Aslam A, et al. Effects of β-Blockers on the Sympathetic and Cytokines Storms in Covid-19. Front Immunol. 2021;12:749291.
271. Vasanthakumar N. Can beta-adrenergic blockers be used in the treatment of COVID-19? Med Hypotheses. 2020;142:109809.
272. Straus MR, Bidon MK, Tang T, Jaimes JA, Whittaker GR, Daniel S. Inhibitors of L-Type Calcium Channels Show Therapeutic Potential for Treating SARS-CoV-2 Infections by Preventing Virus Entry and Spread. ACS Infect Dis. 2021;7(10):2807-15.
273. Yan F, Huang F, Xu J, Yang P, Qin Y, Lv J, et al. Antihypertensive drugs are associated with reduced fatal outcomes and improved clinical characteristics in elderly COVID-19 patients. Cell Discov. 2020;6(1):77.
274. Clemente-Moragón A, Martínez-Milla J, Oliver E, Santos A, Flandes J, Fernández I, et al. Metoprolol in Critically Ill Patients With COVID-19. J Am Coll Cardiol. 2021;78(10):1001-11.
275. Solaimanzadeh I. Nifedipine and Amlodipine Are Associated With Improved Mortality and Decreased Risk for Intubation and Mechanical Ventilation in Elderly Patients Hospitalized for COVID-19. Cureus. 2020;12(5):e8069.
276. Zhang LK, Sun Y, Zeng H, Wang Q, Jiang X, Shang WJ, et al. Calcium channel blocker amlodipine besylate therapy is associated with reduced case fatality rate of COVID-19 patients with hypertension. Cell Discov. 2020;6(1):96.
277. Guragai N, Vasudev R, Hosein K, Habib H, Patel B, Kaur P, et al. Does Baseline Diuretics Use Affect Prognosis in Patients With COVID-19? Cureus. 2021;13(6):e15573.
278. Oliveira FMS, Caetano MMM, de Godoy ARV, de Oliveira LL, de Melo Mambrini JV, Rezende MS, et al. Retrospective cohort study to evaluate the continuous use of anticholesterolemics and diuretics in patients with COVID-19. Front Med (Lausanne). 2023;10:1252556.
279. Esler M, Esler D. Can angiotensin receptor-blocking drugs perhaps be harmful in the COVID-19 pandemic? J Hypertens. 2020;38(5):781-2.

Chapter 10 Appendices

Appendix 1: Data extracted from NHS GG&C Safe Haven

Main Data and Descriptors	Filters	Date Ranges
Demographics	DOB, age, sex, height, weight, smoking status, SIMD	01/01/2019-31/12/2021
Deaths – Combined death data from General Registry Office	Sex, age, ethnic group, Date of death, institution, primary cause of death, secondary cause of death, place of death.	01/01/2014-31/12/2023
Prescriptions	Paid date, BNF item code, BNF item description, BNF chapter code, BNF chapter description, BNF section code, BNF section description, BNF subsection code, BNF subsection description, BNF paragraph description, approved name, drug formulation, item strength, dispensed quantity, number of dispensed items	01/01/2014-31/12/2023
Lab Results	**Serum/blood** Haematology Full blood count (Hb, WCC, neutrophils, lymphocytes, eosinophils, monocytes, basophils, platelets, haematocrit, mean cell volume, red cell distribution width, coagulation screen (PT, APPT,TT, PTR), fibrinogen, INR, d-dimer Biochemistry Sodium, potassium, chloride, urea, creatinine, eGFR,HCO3, Mg, Ca, PO4, Alk phos, AST, ALT, bilirubin, LDH, CRP, glucose, HBA1c, Total cholesterol, HDL, cholesterol, triglycerides, cortisol, hsTNI, NTproBNP, renin, aldosterone, ferritin, vitamin D, CK, myoglobin Virology SARS-Cov-2, Influenza, HIV, Hepatitis B, C, CMV Lighthouse data Urine ACR, PCR, urine electrolytes Imaging (if available in anonymised form) Chest xray, CT chest, chest/abdomen, chest/abdomen/pelvis, CTPA, HRCT chest, CT head, CT venogram, MRI brain, MR venogram, Ultrasound renal	01/01/2014-31/12/2023
SMR00 – General acute outpatient attendance	GGC and WoS. Date, sex, ethnic group, specialty/discipline, significant facility, referral type, attendance status, referral source, attendance follow up, mode of contact	01/01/2014-31/12/2023
SMR01 – General acute inpatient & day case discharges	GGC and WoS Date, sex, ethnic group, specialty, significant facility, management of patient, admission type, admission/transfer from, discharge type, discharge/transfer to, ICD diagnosis	01/01/2014-31/12/2023
SMR04 – Psychiatric and mental handicap hospitals & units	GGC and WoS Sex, ethnic group, specialty/discipline, significant facility, management of patient, admission type, admission/transfer from, admission, referral from, previous psychiatric care, discharge type,	01/01/2014-31/12/2023

	discharge/transfer to, type of psychiatric care provided, arrangements for after care.	
Other Data	**Filters**	**Date ranges**
SMR02 – Maternity inpatient and day case discharges.	GGC and WoS Ethnic group, specialty/discipline, significant facility, management of patient, admission reason, admission type, previous pregnancies, diabetes, booking smoking history, drug misuse during pregnancy, drug used, ever injected illicit drugs, number of births this pregnancy, presentation at delivery, mode of delivery, outcome of pregnancy, resuscitation, Apgar score, sex (baby), neonatal indicator, baby discharged, first feed given, deed on discharge, discharge type,	01/01/2014-31/12/2023
SMR06 – Cancer Registry	GGC and WoS Date of incidence, age, date of death, sex, site, type, clinical stage	01/01/2014-31/12/2023
MINAP – Myocardial Ischaemia National Audit Project	Age, gender, date of admission, initial diagnosis, ECG; determining treatment, cardiac markers raised, systolic BP, pulse rate, Killip class on admission, reperfusion treatment and procedure, date of discharge, discharge diagnosis, death in hospital	01/01/2014-31/12/2023
SERPR – NHS GGC Renal service database	Blood pressure, haemodialysis, haemodialysis date, peritoneal dialysis, peritoneal dialysis date, transplant status, date of transplant, transplant rejection date, vascular procedures, vascular procedure date, immunosuppressant therapy	01/01/2014-31/12/2023
SCI diabetes – National diabetes register (For GGC only)	Sex, age, ethnic group, pulse rate, heart rate, systolic blood pressure, diastolic blood pressure, blood pressure, height, weight, body mass index,waist circumference, diagnosis, family history, type of diabetes, diabetes complications, date of diagnosis diabetes	01/01/2014-31/12/2023
SBR – Scottish Birth Record	GGC and WoS Mother – smoking history, smoking during pregnancy, alcohol intake pre-pregnancy, current alcohol intake, problem with alcohol, ever injected illicit drugs, drugs misuse during this pregnancy, drug used. Birth – onset, induction augmentation,antenatal steroid, pain relief, number of births, membrane rupture, sex, Apgar, resuscitation, place of birth,	01/01/2014-31/12/2023
MUSE – ECGs and Echos	ECG/Echo	
Charlson Matrix - Charlson Comorbidity Index matrix; calculating 1-year mortality risk for all in CHI		01/01/2014-31/12/2023
OBELIX and serology contributed data	All data from previous OBELIX workspace	All – 2023

Appendix 2 Machine Learning Model provided by Tran Quoc Bao Tran

Our work is inspired by the self-attention mechanism of the Transformers neural networks, which is the architecture underlying the success of the large language model chatbot systems such as Chat Generative Pre-trained Transformer (ChatGPT).
By applying the logic of self-attention to longitudinal EHR data, we can capture interactions between different drug prescriptions in space and time. Our proposed model consists of four primary components:

1. A preprocessing pipeline that transforms and encodes the raw time-series data into a suitable format for the neural network.
2. A transformer-based architecture with self-attention mechanism designed to capture the spatiotemporal patterns present in the EHR data.
3. A feed-forward multilayer perceptron to process the static and numerical data.
4. A causal inference system aimed at estimating the causal effect of the selected drug to the risk of hospitalisation or death from COVID-19.

Our model incorporated all diagnoses and medication records prior to the start of two COVID-19 waves for each patient, with static attributes of age, sex, SIMD, and diabetes status at baseline.

Preprocessing pipeline

We create spatiotemporal sequences by flattening the drugs and comorbidities time series into individual “sentences” with drug prescriptions and admission diagnoses being the “words”. Each word was subsequently tokenised and embedded as high-dimensional trainable vectors that allow the representation of concepts in numerical form, for which similar concepts have similar numerical representations. Instead of the positions in the sequence, the positional embedding index for each token was the month and year a drug or a diagnosis appeared in the 6-month study period.

Transformer architecture

When reading a sentence, each word is not processed in isolation, but rather consider its relationship to those before and after. Self-attention in a transformer neural network captures this essence, allowing them to "attend" to relevant parts of an input sequence when making predictions.
In the first step, the embedding for each word in a transformer neural network is generated by adding a word embedding vector and a positional embedding vector. The word embedding vector represents the meaning and context of the “words”, which are drug prescriptions and admission diagnoses in our prosed model. The positional embedding vector represents the location or index of the word in the sequence, which are the drugs dispensing dates or the dates of admissions. The addition of the two vectors thus preserves both the semantic and chronological information of an item in the EHR. The resulted vector was subsequently used as the input for the self-attention layers. Self-attention computes the relevance of each word to every other word in a given sequence, using scaled dot products between query, key, and value vectors. The query and key vectors are derived from the word embeddings. The query vector represents the word that the model wants to focus on or retrieve information from the input. The key vector represents the word that the model wants to compare with the query vector. The dot product between a query and a key vector measures how similar they are. The dot products are then passed through a softmax function, which normalises the dot product values to obtain a probability distribution that sums up to one. The outputs from the softmax function are subsequently multiplied by the value vectors, which are usually the same as the key vector. The results are added together to produce the attention scores. The self-attention output for each word is thus a new vector that combines the information from all the other words in the sentence, weighted by their relevance to the index word. The same process is repeated for the other words to generate the self-attention outputs for every word in the sentence. The result is a matrix of self-attention outputs, where each

row corresponds to a word in the sentence, and each column corresponds to a dimension of the embedding. The matrix has the same shape as the original input matrix, but it has different values that reflect the attention mechanism. In this way, self-attention allows Transformers to capture dependencies between words, regardless of their distance from each other in the sentence. This is particularly useful for understanding the context and semantics of a sentence, making Transformers powerful tools for tasks like machine translation, text summarisation, and sentiment analysis.

In this study, the attention layer was configured to apply the input to itself (i.e., the same input). This self-attention operation allowed each element in the input sequence to attend to all other elements in the same sequence. By computing the attention scores between each pair of elements, our model learned to weigh and combine the relevant information from different drugs or admissions at different time within the study period when producing the output representations.

Finally, the transformer architecture was concatenated to a multi-layer perceptron (MLP), which is more efficient in modelling the static variables, which include sex (male/female), ethnicity (White/non-White), SIMD decile (ordinal), and Diabetes status (0/1).

Causal inference: X-learner

The transformer neural network was then used as the base model for the construction of the X-learner, which is a meta-algorithm specifically designed for studying causal inference and estimating individual treatment effects. While most classification or regression models focus on finding patterns from the input features that are associated with the target variable, meta-learners aim to estimate the effect of an intervention on the outcome, also known as conditional average treatment effect (CATE). Classification or regression models thus often struggle to distinguish between correlations and causal relationships in the data, whereas meta-learners are designed to make causal inferences. Traditionally, X-learner utilises base learners such as logistic regressions, random forests, or boosting algorithms to predict individual outcomes and treatment effects. However, these models often struggle to capture complex temporal and spatial interactions within the data.

In our study, we propose an innovative approach to incorporate the transformer model as the base learners into the X-learner framework. By utilising the ability of transformer models to capture complex spatiotemporal information, we aim to significantly improve the accuracy and interpretability of X-learner's estimated CATEs.

The training of our transformers-based X-learner algorithm consists of three stages.

First stage:

Two models were trained on time-series medication and admission data and demographic data to predict the occurrence of COVID-19-related admission or mortality in the next 6 months. One model was exclusively trained on patients from the treated group, while the other model focused solely on the control group. To estimate counterfactual outcomes, we applied the control model to input features from the treated group and vice versa for the control group. Essentially, this method estimated counterfactuals by predicting what the treated group would have obtained, had they received the control and what the control group would have obtained, had they received the treatment.

Based on the counterfactual predictions, we computed the Individual Treatment Effect (ITE) for each patient. Specifically:

- For the treated group, the ITE represents the difference between the actual outcomes and the counterfactual outcomes predicted by the control model.
- Conversely, for the control group, the ITE corresponds to the difference between the counterfactual outcome predicted by the treated model (from stage one) and the actual outcome.

Second stage: We constructed two other models to predict the ITE, using the same input as the stage-one models. We applied both models to input features from the entire population. By imputing ITEs for both treated and untreated individuals, we essentially create a counterfactual scenario for each person, allowing us to compare their outcomes under both treatment and no treatment.

Third stage:
We developed a separate transformer model to estimate the propensity score of receiving treatment. This model was trained on the same input features as the models from stage-one and stage-two, excluding the time series data for the investigated drug.
The final ITEs will be estimated by getting the propensity-score-weighted average of the stage two model predictions for each patient. The ATE for each medication is the average individual treatment effect across the entire population.

Metrics
The performances of the models were evaluated using stratified five-fold cross-validation. This approach provides a more reliable estimate of a model's performance compared to using a single train-test split. By dividing the data into five non-overlapping subsets (folds), and iteratively using four folds for training and one fold for testing, the evaluation process captures a broader range of data variations. Furthermore, through multiple rounds of training and testing on different subsets, the evaluation becomes less sensitive to any irregularities or biases that may have been coincidentally introduced in a single train-test split.
Performance of models was reported using the area under the receiver operating characteristic curve (AUROC), accuracy, and F1-score (the harmonic mean of precision and recall). The performance of a model was obtained by averaging the performance across all five folds. Results were reported as mean ± standard deviation (SD).
As mentioned earlier, the treatment effects from our causal inference analysis were measured using the Average Treatment Effect (ATE) and the Individual Treatment Effects (ITE). The ATE represents the average difference in outcomes between the treated and control groups, providing an estimate of the overall causal impact of the treatment. This average effect, however, may mask important heterogeneity in how individuals respond to the treatment. The ITE captures the causal effect for each individual, allowing researchers to understand how the treatment impact varies across the population By considering both the ATE and ITEs, causal inference analyses can offer a more nuanced understanding of the treatment effect and identify subgroups that may benefit most from the intervention.

Attention score
To better understand how the neural-network models arrived at the predictions, we extracted the attention weights computed during the forward pass of the model. These tensors were computed as part of the multi-head attention mechanism within our model's encoder layer. The attention weights represent the extent to which each element within the input data attends to other elements during the modelling process for predicting COVID-19 hospitalisation or mortality outcomes. In other words, the attention scores allow The attention weights were retrieved as a 4-dimensional tensor with shape (batch_size, num_heads, sequence_length, sequence_length). Each 2D slice (sequence_length, sequence_length) within this tensor corresponds to the pairwise attention scores between all input positions for a particular patient and attention head. To calculate the final overall attention score, we took the average of all the individual attention scores across the samples (e.g. patients) in the study population. This provided a single aggregated attention score that summarised the attention patterns in the model for the entire dataset.