# Social health inequalities in vaccine-preventable viral respiratory emerging infectious diseases

## The case of the COVID-19 pandemic

Lisa Cavillot

Thesis submitted in fulfilment of the requirements for the degree of Doctor in Public Health Sciences

Belgium, September 2025

*“How do we build a fair distribution of benefits during a pandemic in a society so burdened by inequality?”*

– Harvey Kayman –

## *Remerciements*

J’ai décompté les semaines, les jours, et les heures jusqu’à ce moment, symbolisant la fin d’un long parcours vallonné dont chaque doctorante et doctorant peut être fier.ère. Bien que je ressente effectivement un sentiment de fierté, je ressens également une profonde nostalgie et tristesse de quitter toutes ces personnes qui m’ont accompagnée, formée, et inspirée, et je suis reconnaissante d’avoir l’occasion de les remercier aujourd’hui.

Tout d’abord, merci à mon promoteur, Niko, et à mes co-promoteurs Brecht et Koen, qui, bien qu’initialement peu familière avec l’étude des inégalités sociales, les biostatistiques, ou encore l’anglais, ont cru en moi et m’ont donné ma chance. Durant mon parcours doctoral, vous m’avez constamment entourée d’une bulle justement composée de soutien, de guidance, et de confiance, qui m’a permis d’évoluer en tant que chercheuse et en tant que personne. Il est désormais temps pour moi de sortir de cette bulle, mais vous m’y avez bien préparée. Merci pour tout.

Merci à Laura et Robby, officiellement membres internes de mon comité de thèse, mais tellement plus que ça pour moi. Robby, merci pour ta positivité et ton soutien constants, mais également pour ton incroyable pédagogie quand il s’agit d’expliquer des méthodes statistiques complexes, me permettant de beaucoup

progresser. Laura, merci pour la collègue et surtout pour la personne exceptionnelle que tu es, emplie de bienveillance, de soutien, d'encouragements, et de conseils toujours avisés. Vous m'avez tous les deux tellement apporté et j'ai eu beaucoup de chance de réaliser ne serait-ce qu'une petite partie de mon chemin professionnel à vos côtés. Chemins qui, je l'espère, se recroiseront un jour.

Merci aux Prof Buffel, Prof Vansteelandt, et Prof Vlieghe d'avoir accepté d'être membres externes de mon comité de thèse, d'avoir lu attentivement ma thèse, et pour vos retours constructifs qui en ont suivi. Je suis reconnaissante d'avoir pu bénéficier de vos expertises respectives, qui ont rendu ma thèse bien plus aboutie.

Merci également à toi Sandy, pour avoir présidé mon comité de thèse avec brio et pour m'avoir aidée à naviguer dans ce processus grâce à tes notes structurées, tes conseils, et tes encouragements.

Merci à tous mes collègues des projets HELICON et LINKVACC pour ces belles collaborations qui ont mené à des publications mettant en lumière des résultats essentiels. Elin, travailler à tes côtés durant la dernière année du projet a été une source précieuse d'entraide et de soutien, merci pour tout. Jinane et Yasmine, merci pour ces quatre années durant lesquelles nous avons évolué en parallèle sur notre thèse, partagé nos doutes, nos difficultés, et été présentes l'une pour l'autre. Courage pour la finalisation de ce long parcours, je ne manquerai pas d'être présente le jour de votre défense pour célébrer cette belle réussite à vos côtés.

Dans le cadre de ma thèse, j'ai eu la chance d'être affiliée à deux instituts, dupliquant les rencontres, et j'en ai faites de si belles. Tout d'abord, merci à tous mes collègues de l'unité d'informations en santé de Sciensano. Aline, Vanessa, Masja, Eduardo, Louise, Sarah, Rani, Noura et bien d'autres. Merci pour ces quatre années enrichissantes à vos côtés. Merci également à tous mes collègues de l'Institut de Recherche Santé et Société de l'UCLouvain. Margo, Cécilia, Catherine, Céline, Sylvie, Alexia, Madeleine, Ruth, Julien, Pierre, Mégane, Periklis, Camille, Nathan, et Emilie. Merci pour nos teambuilding et drinks plein de rires, nos temps de midi passés ensemble au soleil, et cette ambiance chaleureuse au sein de l'institut, allégeant le sentiment de solitude que l'on peut parfois ressentir en tant que doctorant.e. Camille, Nathan, et Emilie, j'aimerais vous dire spécifiquement merci. Merci d'avoir tant coloré mes journées et d'avoir été là pour me soutenir et me remonter le moral dans certains moments plus compliqués. Votre présence quotidienne va beaucoup me manquer.

En parallèle de ce monde professionnel enrichissant, il y a cet autre monde, rempli d'amour et de partages, porté par mes amis et ma famille, qui me permet de garder un équilibre, même dans les moments plus difficiles.

Merci à mes amis d'enfance, Thomas, Sylvain et Cédric, ainsi qu'à ma cousine Sophie, qui ont toujours été présents dans les moments importants de ma vie, et qui j'espère le seront encore pendant très longtemps. Thomas et Sylvain, on ne se

le dit pas souvent ou pour ainsi dire jamais, mais j'aimerais vous dire aujourd'hui à quel point ce lien aussi fort qu'unique que nous avons créé depuis maintenant presque 15 ans m'est précieux.

Merci à mes meilleures amies, Caroline, Louise, et Shirley, que j'ai la chance d'avoir à mes côtés depuis nos premières années de bachelier, et Louise, je dirai même depuis le tout premier jour. Je suis retournée lire les remerciements que je vous avais adressés dans mon TFE, et j'avais, entre autres, écrit ceci: « *Merci à mes amies d'ergothérapie avec qui une entraide et un soutien solides, ainsi qu'une belle amitié, se sont créé.e.s au fil des ans*». Je suis heureuse de voir qu'aujourd'hui, presque 10 ans plus tard et malgré les chemins différents que nous avons empruntés, rien n'a changé.

Caroline, j'aimerais te dire spécifiquement merci. Merci pour ton soutien quotidien indéfectible depuis tant d'années. Merci pour ton écoute, pour tes conseils, et pour tous ces moments que nous avons partagés, allant de magnifiques voyages autour du monde, à simplement regarder la télé en pyjama avec April. Merci pour notre amitié, et merci d'être toi, tout simplement.

A mes parents, maman, papa, merci de m'avoir toujours soutenue, et de m'avoir toujours donné la force dont j'avais besoin pour avancer. Malgré que je sois sujette à avoir peu confiance en moi, vous ne m'avez jamais laissé douter de la personne

que j’étais ou que j’avais envie d’être. Merci pour tout, et merci d’être les parents exceptionnels que vous êtes.

Pour finir, merci à ma personne préférée sur Terre, mon copain, Joffrey, qui m’a soutenue et supportée pendant une dizaine de blocus, un mémoire, et une thèse. Je te promets que des jours plus tranquilles nous attendent. Tu sais déjà tout mais j’aimerais te dire à quel point être à tes côtés m’est essentiel, et à quel point la vie y est belle et ensoleillée.

Merci à tous

# *Table of content*

# Jury members

| | |
|---|---|
| **President** | **Prof. Dr. Sandy Tubeuf** |
| | Health and Society Research & Louvain Data Analysis and Modeling in Economics and Statistics Institutes |
| | *Catholic University of Louvain, Brussels, Belgium* |
| **Promotor** | **Prof. Dr. Niko Speybroeck** |
| | Health and Society Research Institute |
| | *Catholic University of Louvain, Brussels, Belgium* |
| **Copromotors** | **Prof. Dr. Brecht Devleesschauwer** |
| | Department of Epidemiology and Public Health |
| | (Head of service Health Information) |
| | *Sciensano, Brussels, Belgium* |
| | Department of Translational Physiology, Infectiology and Public Health |
| | *Ghent University, Merelbeke, Belgium* |
| | **Dr. Koen Blot** |
| | Department of Epidemiology and Public Health |
| | (Head of Service Infectious Diseases Epidemiology) |
| | *Sciensano, Brussels, Belgium* |

**Internal members** **Dr. Laura Van den Borre**

Department of Epidemiology and Public Health

*Sciensano, Brussels, Belgium*

**Prof. Dr. Robby De Pauw**

Department of Epidemiology and Public Health

*Sciensano, Brussels, Belgium*

Department of Rehabilitations Sciences

*Ghent University, Ghent, Belgium*

**External members** **Prof. Dr. Erika Vlieghe**

Department of General Internal Medicine, Infectious Diseases and Tropical Medicine

*University Hospital Antwerp, Antwerp, Belgium*

**Prof. Dr. Stijn Vansteelandt**

Department of Applied Mathematics, Computer Sciences and Statistics

*Ghent University, Ghent, Belgium*

**Prof. Dr. Veerle Buffel**

Brussels Institute for Social and Population Studies

*Vrije Universiteit Brussel, Brussels, Belgium*

# Scientific curriculum vitae

## Peer reviewed international journal articles

- **Cavillot L**, Speybroeck S, Van den Borre L, Ghattas J, Boiy E, van Loenhout J AF, Hubin P, Smith P, De Smedt D, Vanthomme K, Gadeyne S, De Pauw R, Devleesschauwer B. One-year post-acute organ complications following COVID-19 hospitalization and related socioeconomic inequalities. Submitted to *Nature Communications* (March 2025).
- **Cavillot L**, Moerkerke B, Devleesschauwer B, Ghattas J, van Loenhout J AF, Van den Borre L, Speybroeck N, Loeys T, De Pauw R. The role of vaccination, underlying health conditions and working in healthcare in the socioeconomic disparities in COVID-19 hospitalization: a mediation analysis using interventional effect models. Submitted to *BMC Infectious Diseases* (September 2024).
- **Cavillot L**, Van den Borre L, Vanthomme K, Scohy A, Deboosere P, Devleesschauwer D, Speybroeck N, Gadeyne S. Unravelling demographic and socioeconomic patterns of COVID-19 death and other causes of death: results of an individual-level analysis of exhaustive cause of death data in Belgium, 2020. *Arch Public Health* 82, 209 (2024).

- **Cavillot L**, van Loenhout J AF, Devleesschauwer B, Wyndham-Thomas C, Van Oyen H, Ghattas J, Blot K, Van den Borre L, Billuart M, Speybroeck N, De Pauw R, Stouten V, Catteau L, Hubin P. Sociodemographic and socioeconomic disparities in COVID-19 vaccine uptake in Belgium: a nationwide record linkage study. *J Epidemiol Community Health* 2024;78:176-183.
- Boiy E, **Cavillot L**, Devleesschauwer B, De Smedt D, Gadeyne S, Speybroeck N, Van den Borre L, Vanthomme K, Verhaege N, De Pauw R. Social inequalities in in-hospital mortality following COVID-19 hospitalization in Belgium: a competing risk survival analysis. Submitted to the *Journal of Epidemiology and Community Health* (August 2025).
- Boiy E, **Cavillot L**, Devleesschauwer B, De Pauw R, De Smedt D, Gadeyne S, Gorasso V, Schmidt M, Vanthomme K, Verhaege N & Van den Borre L. Social inequalities in post-acute healthcare expenditure following COVID-19 hospitalization in Belgium. Submitted to the *International Journal for Equity in Health* (June 2025).
- Ghattas J, Makovski TT, Monnier-Besnard S, **Cavillot L**, Ambrožová M, Vašinová B, Feteira-Santos R, Bezzegh P, Bollmann FP, Cottam J, Haneef R, Devleesschauwer B, Speybroeck N, Nogueira PJ, Forjaz MJ, Coste J, Carcaillon-Bentata L. The role of socio-economic determinants in SARS-CoV-2 health

outcomes: systematic review of population-based studies. Submitted to the *Journal of Epidemiology and Community Health* (January 2025).

- Hubin P, Van den Borre L, Braeye T, **Cavillot L**, Billuart M, Stouten V, et al. Area and individual level analyses of demographic and socio-economic disparities in COVID-19 vaccination uptake in Belgium. Vaccine X. 2024 Jun 1;18:100496.
- Makovski TT, Ghattas J, Monnier-Besnard S, **Cavillot L**, Ambrožová M, Vašinová B, Feteira-Santos R, Bezzegh P, Bollmann FP, Cottam J, Haneef R, Devleesschauwer B, Speybroeck N, Nogueira PJ, Forjaz MJ, Coste J, Carcaillon-Bentata L. Multimorbidity and frailty are associated with poorer SARS-CoV-2-related outcomes: systematic review of population-based studies. Aging Clin Exp Res. 2024 Feb 14;36(1):40.
- De Pauw R, Van den Borre L, Baeyens Y, **Cavillot L**, Gadeyne S, Ghattas J, De Smedt D, Jaminé D, Khan Y, Lusyne P, Speybroeck N, Racape J, Rea A, Van Cauteren D, Vandepitte S, Vanthomme K, Devleesschauwer B. Social inequalities and long-term health impact of COVID-19 in Belgium: protocol of the HELICON population data linkage. BMJ Open. 2023 May 18;13(5):e069355.
- Khan Y, Verhaeghe N, Devleesschauwer B, **Cavillot L**, Gadeyne S, Pauwels N, Van den Borre L, De Smedt D. The impact of the COVID-19 pandemic on delayed care of cardiovascular diseases in Europe: a systematic review. Eur Heart J Qual Care Clin Outcomes. 2023 Nov 2;9(7):647-661.

## Book chapters

- **Cavillot L**, van Loenhout J AF, Devleesschauwer B, Wyndham-Thomas C, Van Oyen H, Ghattas J, Blot K, Van den Borre L, Billuart M, Speybroeck N, De Pauw R, Stouten V, Catteau L, Hubin P. (2024). *La couverture vaccinale contre la COVID-19 en Belgique*. In C. Burton-Jeangros, Y. Jackson, J. Racapé, M.-F. Raynault, & A. Rea (Eds.), *Les inégalités sociales et de santé. Gestion de la crise de la COVID-19 à Bruxelles, Genève et Montréal* (pp. 73–84). Éditions de l'Université de Bruxelles.

## Oral presentations at (inter)national conferences

- **15th European Conference of Public Health**

  Sociodemographic and socioeconomic disparities in COVID-19 vaccine uptake in Belgium. *Berlin, Germany (2022).*

- **Health and Society Research Institute seminar**

  Sociodemographic and socioeconomic disparities in COVID-19 vaccine uptake in Belgium. *Brussels, Belgium (2022).*

- **Health and Society Research Institute seminar**

  Demographic, socioeconomic and clinical risk factors for all-cause mortality up to 12 months following COVID-19 hospitalization: A competing risk survival analysis. *Brussels, Belgium (2023).*

- **BY-COVID seminar**

  Sociodemographic and socioeconomic disparities in COVID-19 vaccine uptake in Belgium. *The Hague, Netherlands (2023).*

- **HELICON webinar**

  Sociodemographic and socioeconomic disparities in COVID-19 vaccine uptake in Belgium. *Online (2023).*

- **International symposium "Social and health inequalities. Managing the COVID-19 crisis in Brussels, Geneva and Montreal"**

  Sociodemographic and socioeconomic disparities in COVID-19 vaccine uptake in Belgium. *Brussels, Belgium (2024).*

- **Health and Society Research Institute seminar**

  The mediation role of vaccination, underlying health conditions and working in healthcare on the socioeconomic disparities in COVID-19 hospitalization. *Brussels, Belgium (2024).*

- **17th European Conference of Public Health**

  Unravelling demographic and socioeconomic patterns of COVID-19 death and other causes of death: results of an individual-level analysis of exhaustive cause of death data in Belgium, 2020*. Lisbon, Portugal (2024).*

- **Interuniversity day for PhD students in public health**

  One-year post-acute organ complications following COVID-19 hospitalization and related social inequalities. *Brussels, Belgium (2024).*

## Certificates obtained / Courses attended

- **University certificate in Junior Data Analyst**, Université Catholique de Louvain, *September – February 2021.*
- **Advanced epidemiology**, Master in Public Health Sciences, Université Catholique de Louvain, *February – June 2022.*
- **Scientific writing**, Université Catholique de Louvain, *March 2022.*
- **University certificate in Statistics and Data Sciences,** Master in Biostatistics, Université Catholique de Louvain, *September 2023 – June 2024.* Courses included:
  - Advanced programming in R
  - Elements of mathematics for statistics
  - Elements of probability and mathematical statistics
  - Analysis of survival and duration data
- **Linear models**, Master in Biostatistics, Université Catholique de Louvain, *September – December 2024*.
- **Data analysis: principal component analysis**, Master in Biostatistics, Université Catholique de Louvain, *September – December 2024*.

- **Concept and processing of random vectors**, Master in Biostatistics, Université Catholique de Louvain, *September – December 2024*.

## Teaching

- **Advanced epidemiology**, Master of Public Health Sciences, Université Catholique de Louvain, *February 2023 – June 2025.*

# Abstract

The spread of Severe Acute Respiratory Syndrome 2 (SARS-CoV-2) and the resulting Coronavirus Disease 2019 (COVID-19) pandemic caused unprecedented health impacts, with over four million cases, 150,000 hospitalizations, and 30,000 deaths recorded over the past four years in Belgium. The clear social gradient observed in SARS-CoV-2 outcomes worldwide led to the pandemic being rapidly described as a *syndemic*, highlighting the interplay between the virus and preexisting health and social conditions. In Belgium, however, major knowledge gaps remained regarding both the extent and drivers of social health inequalities in the COVID-19 pandemic, as well as the long-term health impacts of the disease. This thesis addresses these gaps by leveraging individual-level data from the Belgian adult population, combined with innovative causal inference and machine learning techniques. The findings revealed significant social inequalities in COVID-19-related health behaviors, mortality, and morbidity, the latter being partly driven by unequal vaccine uptake and comorbidity burden, as well as long-term organ complications following severe infection. These results underscore the critical need to systematically integrate social equity into pandemic preparedness and response policies, and to implement long-term care monitoring for post-COVID organ damage.

# List of abbreviations

**A**

| | |
|---|---|
| **ACE** | Angiotensin Converting Enzyme |
| **ACE2** | Angiotensin Converting Enzyme 2 |
| **ARB** | Angiotensin II Receptor Blocker |
| **ARDS** | Acute Respiratory Distress Syndrome |
| **ATC** | Anatomical Therapeutic Chemical |
| **AUC** | Area Under the Curve |

**B**

| | |
|---|---|
| **BELSPO** | Belgian Science Policy Office |

**C**

| | |
|---|---|
| **CBSS** | Crossroads Bank of Social Security |
| **CHS** | Clinical Hospital Survey |
| **CI** | Confidence Interval |
| **CIF** | Cumulative Incidence Function |
| **CIT** | Conditional Inference Tree |
| **CML** | Clinical Microbiology Laboratory |
| **CMR** | Crude Mortality Rate |
| **CoBRHA** | Common Base Registry for HealthCare Actors |

| | |
|---|---|
| **CT** | Classification Tree |
| **CV** | Cardiovascular |
| **D** | |
| **DALY** | Disability-Adjusted Life Years |
| **DAG** | Directed Acyclic Graph |
| **DE** | Direct Effect |
| **E** | |
| **ECDC** | European Center for Disease Prevention and Control |
| **ECMO** | Extra-Corporeal Membrane Oxygenation |
| **EFTA** | European Free Trade Association |
| **EID** | Emerging Infectious Disease |
| **EMA** | European Medicine Agency |
| **EU** | European Union |
| **F** | |
| **FN** | False Negative |
| **FP** | False Positive |
| **FPCI** | Fundamental Problem of Causal Inference |
| **G** | |
| **GLM** | Generalized Linear Model |
| **GP** | General Practitioner |

**H**

| | |
|---|---|
| **HDA** | Health Data Agency |
| **HR** | Hazard Ratio |

**I**

| | |
|---|---|
| **ICD** | International Standard Classification of Disease and Related Health Problems |
| **ICU** | Intensive Care Unit |
| **IDE** | Interventional Direct Effect |
| **IE** | Indirect Effect |
| **IIE** | Interventional Indirect Effect |
| **IMA** | Intermutualistic Agency |
| **IPTW** | Inverse Probability of Treatment Weighting |
| **IQR** | Interquartile Range |
| **ISC** | Information Security Committee |
| **ISCED** | International Standard Classification of Education |

**J**

| | |
|---|---|
| **JIE** | Joint Indirect Effect |

**K**

| | |
|---|---|
| **KM** | Kaplan Meier |

**L**

| | |
|---|---|
| **LSEM** | Linear Structural Equation Modeling |

**M**

| | |
|---|---|
| **MERS-CoV** | Middle East Respiratory Syndrome Coronavirus |

**N**

| | |
|---|---|
| **NCD** | Non-Communicable Diseases |
| **NISS** | National Identification number of the Social Security |
| **NPI** | Non-Pharmaceutical Interventions |
| **NRC** | National Reference Center |

**O**

| | |
|---|---|
| **OCOD** | Other Causes Of Death |
| **OR** | Odds Ratios |
| **OSE** | European Social Observatory |
| **OW** | Overlap Weight |

**P**

| | |
|---|---|
| **PHEIC** | Public Health Emergency of International Concern |
| **PASC** | Post-Acute Sequelae of COVID-19 |
| **PCC** | Post-COVID-19 Conditions |
| **PPE** | Personal Protective Equipment |
| **PCR** | Polymerase Chain Reaction |

| | |
|---|---|
| **PS** | Propensity Score |
| **PI** | Pharmaceutical Intervention |

**R**

| | |
|---|---|
| **RAT** | Rapid Antigen Test |
| **RCT** | Randomized Controlled Trial |
| **RIZIV-INAMI** | National Institute for Health and Disability Insurance |
| **RNA** | Ribonucleic Acid |
| **ROC** | Receiver Operating Characteristic |
| **RT-PCR** | Reverse Transcription Polymerase Chain Reaction |

**S**

| | |
|---|---|
| **SARS-CoV** | Severe Acute Respiratory Syndrome Coronavirus |
| **SARS-CoV-2** | Severe Acute Respiratory Syndrome Coronavirus 2 |
| **SCS** | Surge Capacity Survey |
| **SD** | Sociodemographic |
| **SDOH** | Social Determinants of Health |
| **SE** | Socioeconomic |
| **SES** | Socioeconomic Status |
| **StE** | Standard Error |
| **SHR** | Subdistribution Hazard Ratio |
| **SMD** | Standardized Mean Difference |

| | |
|---|---|
| **Statbel** | Statistic Belgium |
| **T** | |
| **TE** | Total Effect |
| **TMPRSS2** | Transmembrane Protease Serine 2 |
| **TN** | True Negative |
| **TP** | True Positive |
| **U** | |
| **UK** | United Kingdom |
| **US** | United States |
| **V** | |
| **VE** | Vaccine Effectiveness |
| **VOC** | Variant of Concern |
| **VOI** | Variant of Interest |
| **W** | |
| **WHO** | World Health Organization |
| **WP** | Work Package |
| **Y** | |
| **YLD** | Years Lived with Disabilities |
| **YLL** | Years of Life Lost |

# List of figures

# List of tables

# PART I

# Introduction, objectives, and outline

# General introduction

## 1. Characterization of SARS-CoV-2

### 1.1. Emergence and origins

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the 7$^{th}$ member of the coronaviruses family, is one of the three **highly pathogenic zoonotic coronaviruses endemic to humans**, alongside with severe acute respiratory syndrome coronavirus (SARS-CoV) and Middle East respiratory syndrome coronavirus (MERS-CoV) [1]. SARS-CoV was responsible for an epidemic that originated in China and lasted from 2002 to 2004, while MERS-CoV was first identified in 2012 in the Middle East and is since causing intermittent outbreaks [2,3]. SARS-CoV-2 was first detected in Wuhan, China, in December 2019, among a cluster of patients suffering from pneumonia of unknown origin and is the **causative agent of coronavirus disease 2019** (COVID-19) [4].

These three coronaviruses, i.e., SARS-CoV, MERS-CoV, and COVID-19 are classified as **emerging infectious diseases** (EIDs) [5]. According to the World Health Organization (WHO), an EID is defined as "*one that either has appeared and affected a population for the first time, or has existed previously but is rapidly spreading, either in terms of the number of people getting infected, or to new geographical areas*" [6].

## 1.2. Pathogenesis

SARS-CoV-2 is an enveloped ribonucleic acid (RNA) virus characterized by the presence of Spike proteins on its surface. These proteins allow viral entry via two mechanisms, each resulting in viral replication and subsequent infection of the host cell (**Figure 1**) [7–9]:

- **Endocytosis**: the Spike protein first binds to the host cell receptor, angiotensin converting enzyme 2 (ACE2), further activated by the enzyme Cathepsin L, enabling the fusion between the endosomal and viral membranes;
- **Membrane fusion**: after binding to ACE2, the Spike protein triggers the direct fusion of the viral and cellular membranes. This process is mediated by the transmembrane protease serine 2 (TMPRSS2), co-expressed with ACE2.

ACE2 receptors are expressed on the surface of different type of organic cells, including pneumocytes, cardiomyocytes, small intestinal enterocytes, and glomerular and peritubular capillary endothelial cells in the kidney [1,10,11].

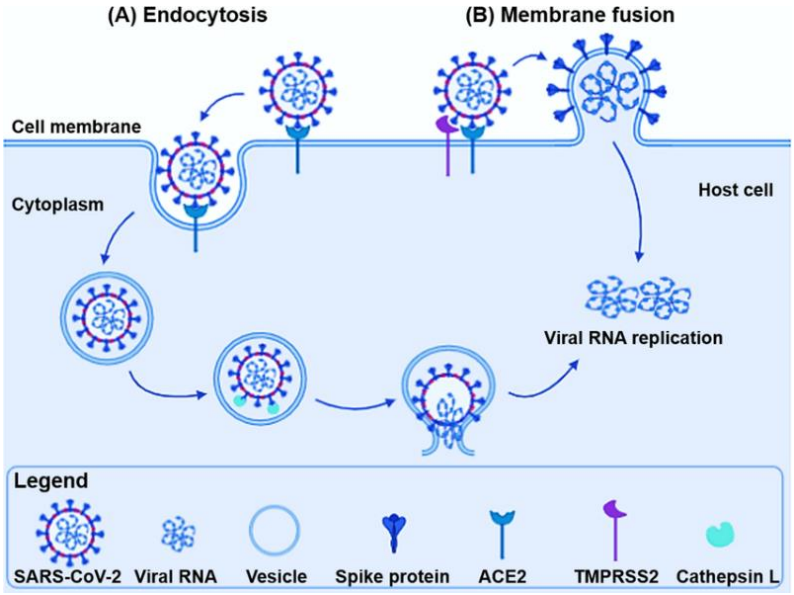



***Figure 1. SARS-CoV-2 viral pathways***. *Source: Pišlar et al. (2020) [12].*

## 1.3. Transmission

Human-to-human transmission of SARS-CoV-2 occurs mainly via the emission of microdroplets when sneezing or coughing. These droplets can be inhaled at close range (i.e., airborne transmission) or come into direct contact with eyes, nose or mouth (i.e., droplet transmission) infecting the exposed individual [13]. The virus enters the body through the nasal or oral cavities, initially infecting cells in the upper respiratory tract [1].

## 1.4. Symptomatology

While some individuals may develop an asymptomatic infection, this early stage typically leads to mild symptoms such as cough, fever or headache, emerging after an incubation period of 2 to 14 days and usually resolving within a few weeks [14].

In some individuals, the virus can further replicate in the lower respiratory tract, infecting the pulmonary alveoli and progressing to moderate symptoms, such as hypoxemia or dyspnea [1,15]. In some cases, SARS-CoV-2 infection can causes critical symptoms, such as pneumonia or respiratory failure, requiring hospital care and which can lead to death [1,15].

Beyond pneumocytes, the presence of ACE2 receptors on other human cells, such as cardiovascular cells or central nervous system neurons, broadens the spectrum of acute SARS-CoV-2 symptoms, with severity ranging from mild to critical [16,17]. As a result, acute cardiovascular complications (e.g., arrythmia, myocardial infarction or cardiovascular death) and neurological complications (e.g., anosmia, stroke, encephalitis) may occur among infected individuals [18–20].

## 2. The COVID-19 pandemic: historical context and epidemiological evidence

The spread of the SARS-CoV-2 resulted in a COVID-19 outbreak, characterized by the WHO as a public health emergency of international concern (PHEIC) in January 2020. On 11 March 2020, the WHO declared COVID-19 a pandemic and officially ended the PHEIC on 5 May 2023 [21,22]. By the end of 2024, more than 777 COVID-19 cases and 7 million COVID-19 deaths were recorded worldwide [23].

## 2.1. Epidemiological situation of COVID-19 in Belgium

In Belgium, the first COVID-19 case was reported on 3 February 2020 in a group of individuals repatriated from Wuhan, China [24]. The virus spread rapidly leading to a first COVID-19 wave, from 1 March to 21 June 2020, in response to which a first lockdown was implemented from 18 March to 5 April 2020. A second lockdown was announced from 2 November to 13 December 2020 in response to the second COVID-19 wave, which occurred from 31 August 2020 to 14 February 2021. This was followed by eight more COVID-19 waves, resulting in a total of ten waves by May 2023 (**Figure 2**).

By May 2023, 5,287,822 COVID-19 cases, 100,613 COVID-19 hospitalizations, including 8,046 intensive care unit (ICU) admissions, and 34,227 COVID-19 deaths were recorded in Belgium. The highest daily number of COVID-19 cases was recorded during the 5$^{th}$ wave, with a total of 76,079 cases recorded on 24 January 2022. The peak in the daily number of COVID-19 hospitalizations occurred during the 2$^{nd}$ wave with a total of 7,461 hospital admissions recorded on 3 November 2020. The first COVID-19 death was reported on 11 March 2020 in a 90-year-old woman. The daily number of COVID-19 deaths peaked during the 1$^{st}$ wave, with a total of 323 deaths occurring on 8 April 2020, before gradually decreasing over time, particularly following the introduction of the vaccine on 28 December 2020 (**Figure 2**).

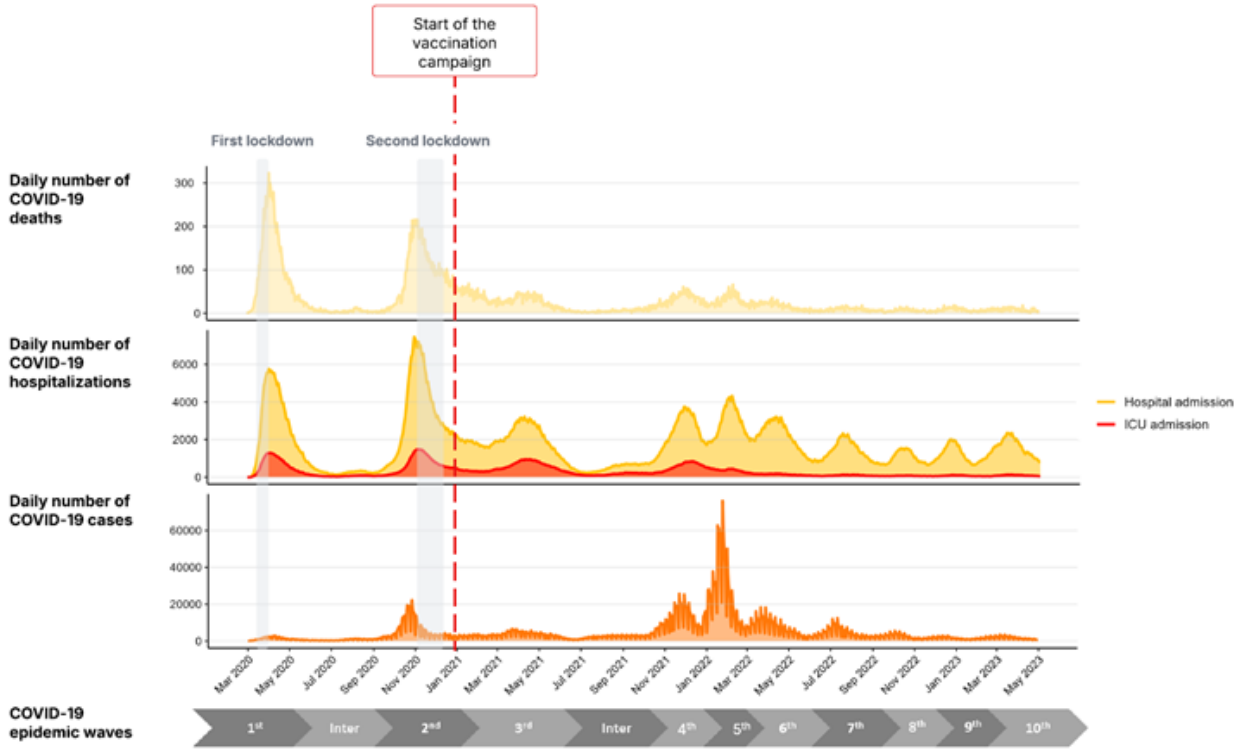



***Figure 2. Overview of the epidemiological situation of COVID-19 in Belgium, March 2020 - May 2023.*** *Source: figure created using Sciensano open data [25].*

## 2.2. The burden of COVID-19 in Belgium

The COVID-19 crisis led to significant excess mortality and morbidity in Belgium, particularly among individuals aged over 65 years [26–28]. COVID-19 was particularly lethal during the first two waves with approximatively 67% of all cumulative COVID-19 deaths occurring during this time period, predominantly in hospitals and nursing homes [28].

### *2.2.1. COVID-19 mortality rate*

In 2020 and 2021, COVID-19 was the first leading cause of death in Belgium for both males and females, with annual mortality rates per 100,000 persons of 191.31 and 86.32, respectively [29]. **Table 1** shows the crude mortality rates (CMRs) of COVID-19 per 100,000 persons from the 1st to 7th waves in Belgium. Across all waves combined, the CMR of COVID-19 was 281. The highest COVID-19 CMRs occurred during the 1st and 2nd epidemic waves, with CMRs of 84 and 104, respectively. The COVID-19 CMR during the 2nd wave was higher due to its length, but the highest peak in deaths occurred during the 1st wave (**Figure 2**).

**Table 1. COVID-19 crude mortality rates (CMRs) per 100,000 persons from the 1st to the 7th waves in Belgium.** *Source: Jurcevic et al (2023) [28].*

| COVID-19 waves | Time period | CMRs per 100,000 persons |
|---|---|---|
| All waves | 01/03/2020 – 11/09/2022 | 281 |
| Wave 1 | 01/03/2020 – 21/06/2020 | 84 |

**Table 1. Continued**

| COVID-19 waves | Time period | CMRs per 100,000 persons |
|---|---|---|
| Inter wave 1-2 | 22/06/2020 – 30/08/2020 | 3 |
| Wave 2 | 31/08/2020 – 14/02/2021 | 104 |
| Wave 3 | 15/02/2021 – 27/06/2021 | 29 |
| Inter wave 3-4 | 28/96/2021 – 03/10/2021 | 4 |
| Wave 4 | 04/10/2021 – 26/12/2021 | 22 |
| Wave 5 | 27/12/2021 – 27/02/2022 | 17 |
| Wave 6 | 28/02/2022 – 29/05/2022 | 13 |
| Wave 7 | 30/05/2022 – 11/09/2022 | 7 |

*The CMRs are calculated as the number of COVID-19 deaths per wave over the Belgian population on January 1st of the corresponding year.*

### *2.2.2. Excess mortality*

Excess mortality is a useful measure to quantify the mortality burden of the pandemic by comparing the number of deaths from all-causes observed with the number of deaths expected based on mortality over the last five years [30].

During the early phases of the pandemic, two episodes of significant excess mortality were recorded. During the 1st epidemic wave (from 20 March to 20 April 2020), Belgium experienced a 64 % excess mortality, with 96% likely attributable to COVID-19. During the 2nd epidemic wave, a significant period of excess mortality was identified from 20 October to 24 December 2020, with a 41.6% excess mortality (**Figure 3**) [31–33].

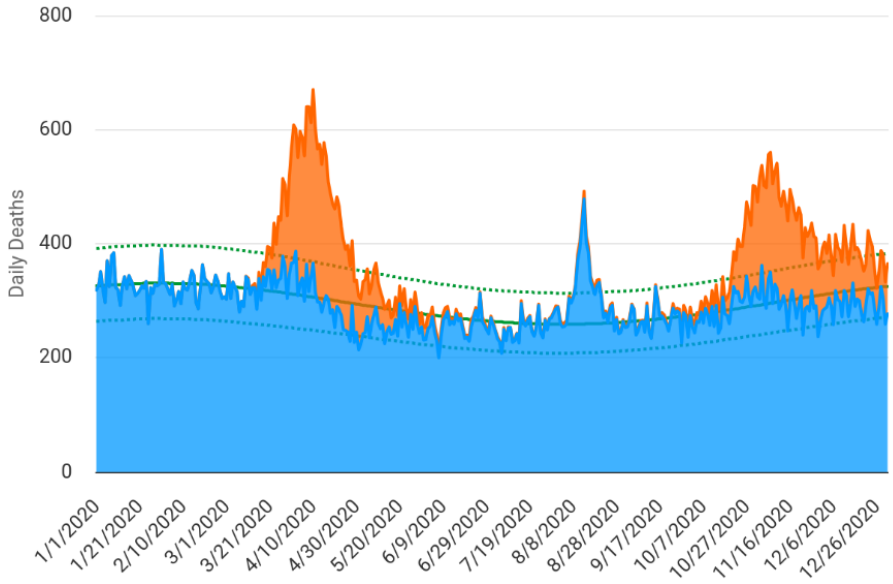



***Figure 3. Expected number of deaths and confidence intervals, observed number of deaths (blue), and COVID-19 deaths (orange), January 2020 – December 2020, Belgium.*** *Significant excess or under-mortality are identified when the total daily number of deaths exceed the lower or upper limits of the prediction interval. Source: Sciensano [30].*

### *2.2.3. Burden of disease of COVID-19*

The burden of disease of COVID-19, measured in disability-adjusted life years (DALYs), which combine years of life lost (YLLs) due to premature mortality and years lived with disabilities (YLDs), was estimated at 253,577 in 2020 and 139,979 in 2021 per 100,000 persons. Premature mortality accounted for 98.5% and 94% of the total burden of disease in 2020 and 2021, respectively [26]. In terms of morbidity, a total of 3,863 and 8,303 YLDs per 100,000 persons due to acute infection and post-acute symptoms following infection were estimated in 2020 and 2021, respectively [26].

## 2.3. COVID-19: a vaccine preventable disease

### *2.3.1. Available vaccines in Belgium*

In response to the COVID-19 pandemic, COVID-19 vaccines were developed in record time. Following robust effectiveness clinical trial data and the meet of the European Union (EU) criteria for efficacy, safety, and quality, five COVID-19 vaccines have been approved to date by the European Medicine Agency (EMA) in Belgium [34–36]. These five vaccines are based on three different technologies:

**(1)** **mRNA vaccines**: the two-dose mRNA vaccines available in Belgium are the mRNA-1273 vaccine by Moderna (approved on 21 December 2020) [37] and the BNT162b2 vaccine by Pfizer and BioNTech (approved on 6 January 2021) [35]. To trigger an immune response, mRNA vaccines deliver synthetic mRNA that provides instructions to the cells on how to produce a harmless portion of the Spike protein. The immune system recognizes this protein as foreign and responds by generating antibodies, leading to an immune response [38].

**(2)** **Viral-vector vaccines**: the viral-vector vaccines available in Belgium are the two-dose AZD1222 vaccine by AstraZeneca (approved on 29 January 2021) [39] and the one-dose Ad26.COV.2.s vaccine by Johnson & Johnson (approved on 11 March 2021) [40]. Viral-vector vaccines use a harmless modified version of the virus, called a vector virus, to introduce genetic

material encoding a piece of the Spike protein into cells. The protein is recognized as foreign by our immune system, leading to the creation of antibodies and triggering an immune response [41].

**(3)** **Protein-based vaccines**: one two-dose protein-based vaccine is available in Belgium, the NVX-CoV2373 vaccine by Novavax (approved on 20 December 2021) [42]. Protein-based vaccines directly deliver a piece of the Spike protein along with an adjuvant into the body, helping the immune system produce antibodies. This triggers the immune system to recognize and respond to the viral protein, building immunity [43].

*2.3.2. The Belgian vaccination campaign*

The first Belgian vaccination campaign started on 28 December 2020, based on a broad automated invitation process, with each Belgian region (i.e., Brussels, Flanders, and Wallonia) responsible for its vaccination rollout. Priority population groups for vaccination included nursing home residents, healthcare staff, individuals aged over 65 years, and individuals with comorbidities. The vaccination campaign was then extended to all individuals aged 18 years and over, in descending order of age [44].

**Figure 4** shows the cumulative COVID-19 vaccination coverage in Belgium as of 28 December 2020. By the end of December 2021, a vaccination plateau for a first vaccine dose and the primary course was reached, with a vaccination coverage of

approximately 80%. The vaccination campaign for the first booster started on 28 April 2021 and reached a vaccination plateau of 60% by March 2022.

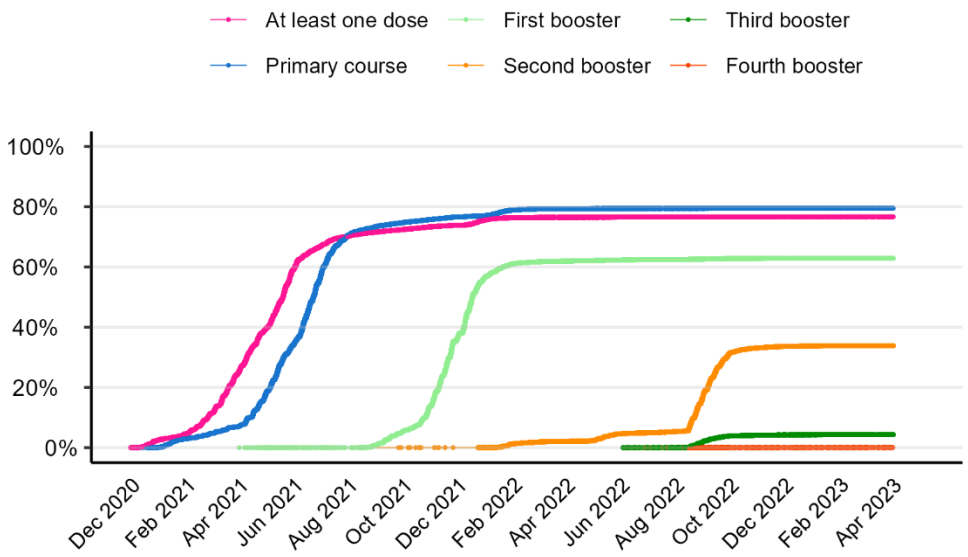



***Figure 4. COVID-19 vaccination coverage (in %) in Belgium, December 2020 – December 2024.*** *Source: figure created using Sciensano Open data [25].*

The rapid development of the vaccine and concerns about its side effects created significant vaccine hesitancy among Belgian residents, with only 34% reporting a definite willingness to be vaccinated before the start of the vaccination campaign [45,46]. However, Belgium achieved the 7$^{th}$ highest vaccination coverage for the primary course in the EU and shows a higher cumulative vaccination coverage for all vaccine doses administered to Belgian residents compared to the EU average (**Figure 5**).

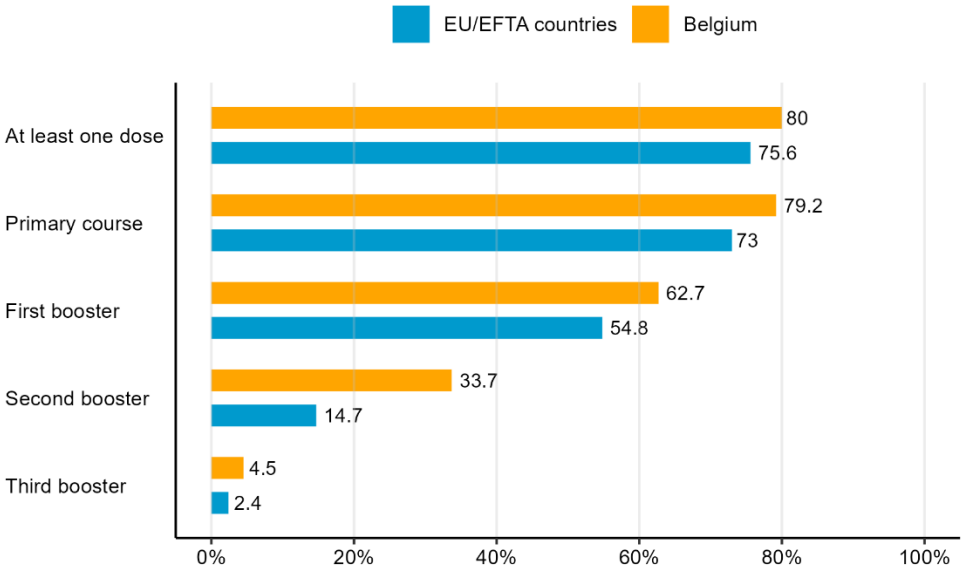



***Figure 5. Cumulative COVID-19 vaccination coverage (%) as of 10 May 2023 for the total population in Belgium compared to the European Union (EU) and European Free Trade Association (EFTA) countries***. *Source: figure created using European Center for Disease Prevention and Control (ECDC) open data [47].*

### *2.3.3. COVID-19 vaccine effectiveness*

The vaccine has proven to be an effective public health tool in reducing symptomatic COVID-19, hospitalizations, and deaths [48–51]. Overall, compared to viral-vector and protein-based vaccines, mRNA vaccines, along with a two-dose vaccination schedule, demonstrated higher vaccine effectiveness (VE) against symptomatic COVID-19 and severe health outcomes [49,50,52,53]. In addition to vaccine-induced immunity, immunity acquired through prior SARS-CoV-2 infection also offers protection against viral transmission and severe health outcomes.

Moreover, hybrid immunity, resulting from both vaccination and prior infection, has been shown to provide the strongest and most durable protection [54].

2.3.3.1. Influence of variants and waning immunity on vaccine effectiveness

When SARS-CoV-2 replicates, it naturally undergoes genetic mutations over time, leading to the emergence of new SARS-CoV-2 strains. While some genetic mutations are neutral, others can impact the virus's transmissibility, severity, or vaccine-induced immunogenicity. Such mutations may result in the classification of a strain as a variant of concern (VOC), as opposed to a variant of interest (VOI) [55,56]. To date, the WHO has classified five variants as VOCs: Alpha, Beta, Gamma, Delta, and Omicron [57]. **Figure 6** shows the proportion of each of the five VOCs over time in Belgium, alongside the daily number of reported COVID-19 cases.

The vaccine was initially developed to target the primary strains of SARS-CoV-2, resulting in an escape from vaccine-induced immunity for Delta and Omicron. The escape of vaccine-induced immunity by Delta and Omicron was the origin of the 4$^{th}$ and 5$^{th}$ COVID-19 waves in Belgium, from 4 October 2021 to 27 February 2022, despite a vaccination coverage of about 80% for the primary course and the start of the vaccination campaign for the first booster (**Figure 6**) [58].

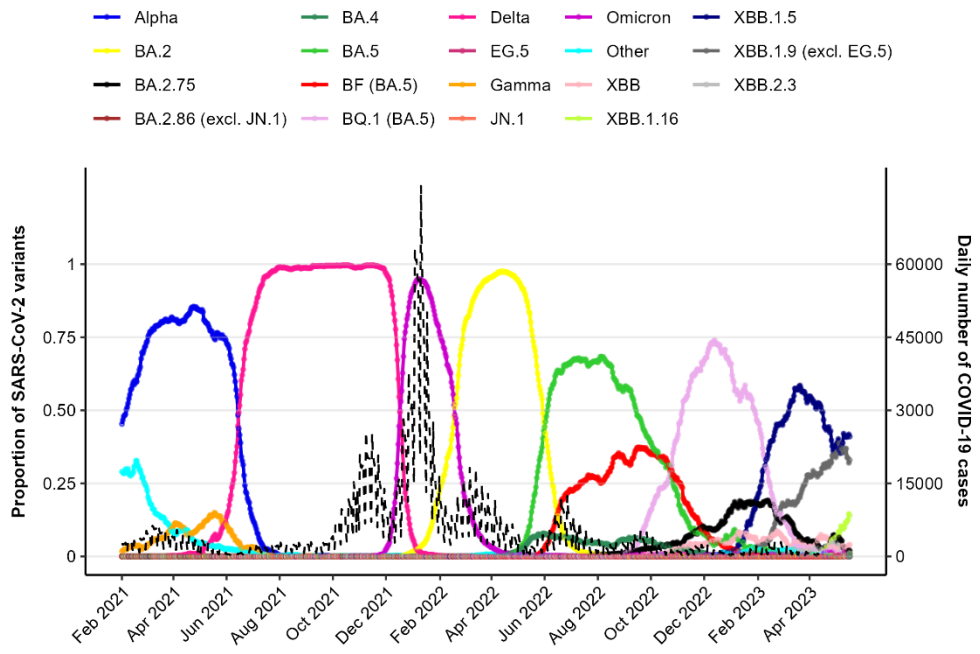



***Figure 6. Proportions of identified variants and daily number of COVID-19 cases over time in Belgium, February 2021 - May 2023.*** *Source: figure created using Sciensano open data [25].*

Delta and Omicron had different impacts on VE and disease severity. The Delta variant resulted in more severe health outcomes than Omicron, with higher hospitalization and death rates [59,60]. However, Omicron, due to numerous mutations in the Spike protein, was more transmissible and resulted in higher case rates (**Figure 6**) [61,62].

These differences in severity and VE between Delta and Omicron are illustrated in **Table 2**, which compares rates of infection, hospitalization, ICU admission, and mortality, as well as VE against symptomatic infection and hospitalization during each VOC period in Belgium.

**Table 2. Comparison of SARS-CoV-2-related infection, hospitalization, intensive care unit (ICU) admission, and mortality rates per 100,000 inhabitants in Belgium, along with vaccine effectiveness (VE), during the Delta and Omicron predominant periods, respectively.** *Sources: Rate estimates: open data available via Sciensano on Epistat [63]; Vaccine effectiveness estimates: Braeye et al. (2022) [64].*

| | Delta | Omicron |
|---|---|---|
| Period of predominance * | 12/07/2021 – 22/12/2021 | 04/01/2022 – 26/05/2022 |
| Infection rate ** | 8,344.4 | 10,130.02 |
| Hospitalization rate ** | 1,198.20 | 965.69 |
| ICU admission rate ** | 476.17 | 121.31 |
| Mortality rate ** | 25.05 | 9.86 |
| VE against symptomatic infection | | |
| *Primary vaccination* | | |
| *0-50 days* | 80% [95% CI 80-81] | 33% [95% CI 30-36] |
| *50-100 days* | 68% [95% CI 67 – 68] | 24% [95% CI 22 - 26] |
| *Booster vaccination* | | |
| *0-50 days* | 85% [95% CI 84-85] | 50% [95% CI 49-50] |
| *50-100 days* | 80% [95% CI 78 – 81] | 34% [95% CI 33 - 35] |
| VE against hospitalization | | |
| *Primary vaccination* | | |
| *0-50 days* | 93% [95% CI 92-95] | 53% [95% CI 29-78] |
| *50-100 days* | 93% [95% CI 92 – 93] | 52% [95% CI 36 – 70] |
| *Booster vaccination* | | |
| *0-50 days* | 96% [95% CI 95-96] | 87% [95% CI 86-89] |
| *50-100 days* | 92% [95% CI 90 – 93] | 81% [95% CI 79 - 83] |

**A variant is considered predominant when it accounted for more than 80% of the infections recorded over the period considered.*

***The total number of Belgian inhabitants as of 1st of January 2021 and 2022 was used as the reference population to calculate the rates for Delta and Omicron, respectively.*

In addition to the decline in vaccine-induced immunity due to viral mutations, VE also decreased over time as immunity waned, as observed in **Table 2** [50,65]. However, the administration of booster doses has been shown to restore immunity and further reduce the risk of severe COVID-19 health outcomes [50,65–67].

### 2.4. Long COVID

The WHO defines long COVID as "*the continuation or development of new symptoms three months after the initial SARS-CoV-2 infection, with these symptoms lasting for at least two months with no other explanation*" [68]. Long COVID may also be referred to as "Post-acute sequelae of COVID-19" (PASC) or "Post-COVID-19 conditions" (PCC).

Several meta-analyses have estimated the worldwide pooled prevalence of long COVID to range between 36% and 43% [69–72]. In Europe, pooled prevalence estimates range between 39% and 46% [69–71]. In Belgium, the COVIMPACT project [73], which follows a cohort of individuals who tested positive for SARS-CoV-2 to unravel its long-term impacts, found that 48% of the participants reported at least one SARS-CoV-2-related symptom three months following the diagnosis [74].

**Figure 7**, based on a meta-analysis by O'Mahoney et al. [72], shows the most frequent PASC among hospitalized and non-hospitalized cases, highlighting the wide range of symptoms characterizing long COVID. Overall, the prevalence of

reporting at least one PASC is largely higher among hospitalized cases (52.63%) compared to non-hospitalized cases (34.46%). Although some symptoms, such as fatigue or dyspnea, are frequently reported in both groups, the symptomatology differs. Non-hospitalized cases more often report classic viral symptoms (e.g., cough, headache, nasal symptoms, gastrointestinal symptoms) (**Figure 7.A**), whereas hospitalized cases frequently report neurocognitive and psychological symptoms, including memory and concentration impairments, post-traumatic stress disorder, anxiety, and depression (**Figure 7.B**).

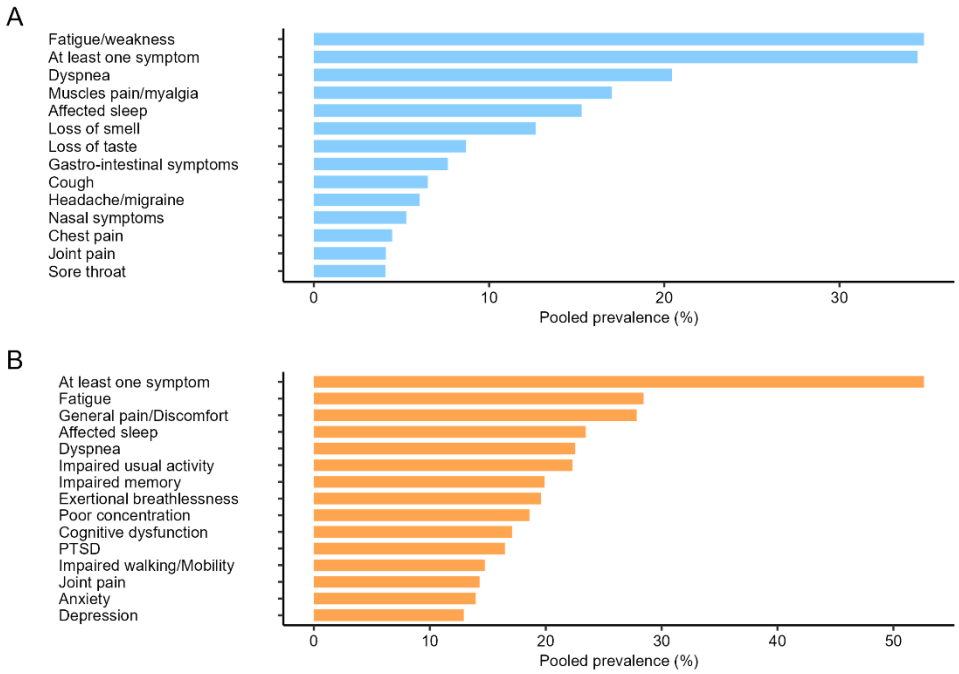



***Figure 7. Global pooled prevalence (in %) of post-acute sequelae of COVID-19 among non-hospitalized infected cases (A) and hospitalized infected cases (B).*** *Source: figure created using data from the meta-analysis by O'Mahoney et al. (2022) [72]. Abbreviations: PTSD: post-traumatic stress disorder.*

### *2.4.1. Pathogenesis of long COVID*

The pathogenesis of PASC is not fully understood yet. However, the following main pathophysiological mechanisms have been suggested to explain post-acute damage across multiple organ systems [75–77]:

- **Direct viral damage**: SARS-CoV-2 directly infects cells expressing high levels of ACE-2 receptors, which enable viral entry into the cell. These receptors have been shown to be particularly overexpressed in cardiomyocytes, pneumocytes, and kidney tubular epithelial cells, making these tissues particularly susceptible to direct viral injury [78];
- **Excessive inflammatory response**: a secondary effect due to excessive inflammation induced by high levels of inflammatory cytokines. Inflammatory responses can lead to viral persistence and tissue damage [77,79];
- **Post-viral autoimmunity**: an autoimmune response triggered by the infection can lead to an increase in autoantibodies, such as antinuclear autoantibodies, which have been identified in COVID-19 patients up to 12 months after acute infection and can target components of the cell nucleus, promoting inflammation and damaging organ systems [80];
- **Hypercoagulability**: endothelial damage, platelet activation, and an exaggerated inflammatory response contribute to a hypercoagulable state,

increasing the risk of thrombotic events. This may be further increased by disseminated intravascular coagulation, immobilization, and venous stasis in ICU patients [77,81];

- **Genetic susceptibility**: emerging evidence suggests that genetic factors may influence variability in long COVID. For instance, a genome-wide study identified a significant association between variations at the FOXP4 locus, a gene critical for lung development and immune regulation, and susceptibility to long COVID. These genetic differences may affect pulmonary function and immune responses, contributing to prolonged symptoms [82].

These pathophysiological mechanisms frequently overlap and may vary across organ systems, with PASC likely resulting from a complex interplay of these factors [76,77,83].

## 3. The social patterns of the COVID-19 pandemic: is COVID-19 a syndemic?

### 3.1. The social determinants of health

The WHO defines the social determinants of health (SDOH) as "*The non-medical factors that influence health outcomes. They are the conditions in which people are born, grow, live, and age, and the wider set of forces and systems shaping the conditions of daily life. These forces and systems include economic policies and*

*systems, development agendas, social norms, social policies, and political systems*" [84]. SDOH can be categorized into five key domains: economic stability (e.g., employment, food security), education access and quality, health care access and quality, neighborhood and built environment (e.g., housing, basic amenities, and the environment), and social and community context (e.g., social inclusion and non-discrimination) [84,85].

### *3.1.1. Social determinants of health inequalities*

Whether in the context of communicable or non-communicable diseases (NCDs), SDOH influence health equity, defined by the Centers for Disease Control and Prevention (CDC) as "*The state in which everyone has a fair and just opportunity to attain their highest level of health*" [86]. Social health inequalities exist between and within countries. Globally, life expectancy at birth was 18.1 years lower in low-income countries compared to high-income ones in 2019 [87]. In Belgium, between 1998 and 2019, men and women living in the most deprived areas, compared to those living in the least deprived, were 1.96 and 1.78 times more likely to die prematurely, respectively, with socioeconomic (SE) inequalities accounting for approximately 28% of all premature deaths [88]. According to Margaret Whitehead, the head of the WHO Collaborating Centre for Policy Research on the SDOH, social inequalities in health are "*socially produced, they are potentially avoidable, and are widely considered unacceptable in a civilized society*" [89].

### 3.2. Social health inequalities in previous viral respiratory emerging infectious diseases: the case of the 1918 Spanish flu and 2009 H1N1 influenza pandemics

The 1918 Spanish and the 2009 H1N1 influenza pandemics resulted in approximately 50 million and 250,000 deaths worldwide, respectively [90,91]. Although both were airborne viral respiratory EIDs, socially vulnerable populations were disproportionally affected.

In the United States (U.S.), during the 1918 Spanish flu pandemic, Grantz et al. [92] found that a higher rate of transmissibility was associated with higher rates of illiteracy, population density, and unemployment, and that a higher mortality rate was associated with a higher rate of illiteracy. In Spain, an income gradient was identified in excess mortality during the 1918 Spanish Influenza pandemic, with average excess mortality rates of 29%, 62%, and 69% among high-, middle-, and low-income groups, respectively [93].

Similar social health inequalities were identified in the 2009 H1N1 influenza pandemic, with higher hospitalization rates among ethnic minorities, individuals with high material deprivation, and those with lower levels of education [94,95]. In England, H1N1 mortality rates were three times higher in the most deprived population than in the least deprived, and 1.7 times higher in urban areas than in rural ones [96].

### 3.3. The COVID-19 pandemic: are we really in this together?

In March 2020, after UK Prime Minister Boris Johnson was infected with SARS-CoV-2, the following declaration was made by politicians: "*The virus does not discriminate. We are all at risk.*" [97]. However, similar to previous viral respiratory EID pandemics, social inequalities in COVID-19-related health outcomes were rapidly identified. Higher risks of SARS-CoV-2 infection and related severe health outcomes (e.g., hospitalizations, deaths) were found among socially vulnerable populations (e.g., lower levels of income and education, unemployment, overcrowded households) and ethnic minorities [98,99]. The COVID-19 pandemic was therefore described as a *syndemic pandemic* [100,101].

#### *3.3.1. Why is COVID-19 a syndemic?*

The term "S*yndemic"* was introduced by Merrill Singer in the 1990s, who defined syndemics as "*The aggregation of two or more diseases or health conditions in a population in which there is some level of deleterious biological or behavior interface that exacerbates the negative health effects of any or all of the diseases involved.*" [102]. In other words, a syndemic arises when there is a synergy between a disease and pre-existing social inequalities in both health conditions and SDOH, exacerbating pre-existing biological, economic, and social conditions and worsening health outcomes [100–102].

**Figure 8**, by Bambra et al. [100], illustrates the syndemic of the SARS-CoV-2 virus and pre-existing health and social conditions, underlying social inequalities in SARS-CoV-2-related health outcomes.

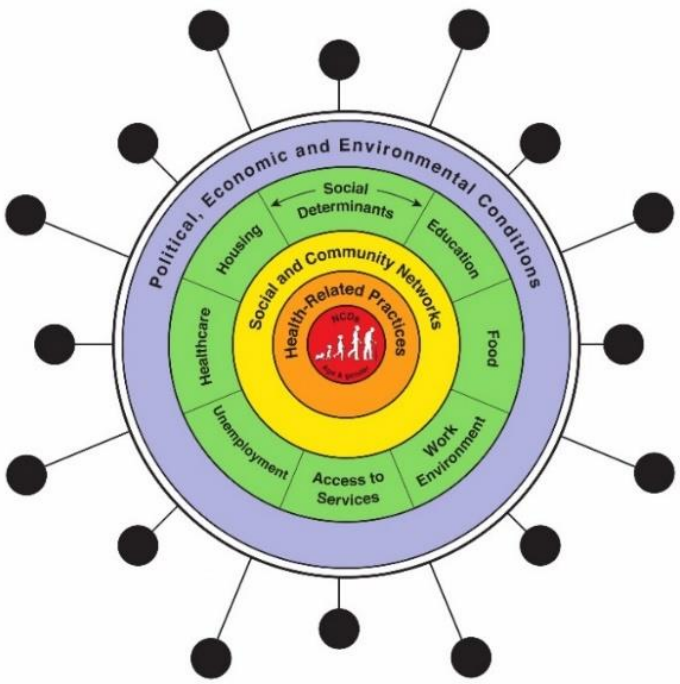



***Figure 8. The syndemic of COVID-19, non-communicable diseases (NCDs), and the social determinants of health.*** *Source: Bambra et al. (2020) [100].*

For example, socially vulnerable populations are more vulnerable to severe COVID-19 outcomes due to the higher burden of comorbidities, such as diabetes or respiratory diseases, as a result of inequalities in SDOH, such as working conditions or access to and quality of healthcare [103–105]. Inequalities in SDOH, such as education or housing, can lead to increased virus transmissibility among socially

vulnerable groups due to lower adoption of preventive measures, overcrowded households, or limited access to outdoor spaces [106–109]. All these factors accumulate and interact, creating social inequalities in SARS-CoV-2 infection and related severe health outcomes, exacerbating those already rooted in our society [105].

# Objectives and outline of the thesis

Sciensano, the Belgian institute for health, was legally responsible for the epidemiological surveillance of the COVID-19 pandemic [110]. In collaboration with regional and federal health authorities, as well as healthcare actors, national electronic health record surveillance systems were established to monitor the evolution of the pandemic in real time [111]. These systems collected key data on laboratory tests (including genome sequencing data), hospitalizations, vaccinations, and deaths [112–116], which were publicly available in aggregated form via the *Epistat* website [63].

The epidemiological data collected by Sciensano allowed real-time monitoring of the COVID-19 crisis and quantification of disease incidence, morbidity, and mortality. However, these surveillance systems focused less on identifying specific vulnerable demographic and SE subgroups during the pandemic.

This thesis therefore **aims to unravel social health inequalities in the COVID-19 crisis and to assess the long-term direct health impacts of the disease in Belgium**. This thesis is structured into three sections, each addressing a specific knowledge gap and contributing to the main research objective. The sections, related knowledge gaps, chapters, and research questions, are outlined below.

## 1. Section I. What are the social patterns of COVID-19-related health behaviors?

### 1.1. Knowledge gap

Internationally, beyond social inequalities in morbidity and mortality, the 2009 H1N1 influenza pandemic demonstrated social disparities in vaccine uptake, with coverage generally increasing with education and income levels [117]. During the COVID-19 pandemic, prior to the rollout of vaccination campaigns, international findings showed that deprived SE groups exhibited lower adoption rates of preventive health behaviors, such as mask-wearing and handwashing, as well as decreased willingness to receive the COVID-19 vaccine [45,107].

In Belgium, despite its high-performing universal health system, social inequalities in health behaviors, including the use of preventive treatments and services, have been documented [118]. Given the syndemic nature of COVID-19, which exacerbates preexisting health and social inequalities, and the proven effectiveness of the COVID-19 vaccine in preventing severe illness [48], it is important to assess whether or not vaccine coverage is equitably distributed across the Belgian population. Yet, to our knowledge, no research has examined social disparities in COVID-19 vaccine uptake in Belgium from the early phases of the pandemic.

Furthermore, beyond assessing equity in vaccination coverage at the national level, it is also important to assess equity at the regional level, as Belgium's federal

structure, in which each region is responsible for the vaccination rollout, may have led to regional differences in outreach and communication. However, each Belgian region has a distinct sociodemographic and socioeconomic profile, making it essential to ensure that vaccination campaigns effectively reached the population in line with its specific characteristics and needs.

This knowledge will provide policymakers with evidence to identify gaps in vaccination coverage, tailor outreach strategies to vulnerable groups, and design future public health interventions that reduce health inequities and improve the overall effectiveness of vaccination campaigns.

### 1.2. Outline of thesis chapters and related research questions

In ***Chapter 1***, we aim to assess the association between SD and SE determinants and COVID-19 vaccination coverage among Belgian adults at the national and regional levels, as each Belgian region was responsible for the vaccination rollout.

## 2. Section II. What are the social patterns of COVID-19-related morbidity and mortality?

### 2.1. Knowledge gap

As described in the introduction, the COVID-19 pandemic was quickly recognized as a syndemic, with emerging international evidence of social inequalities in SARS-

CoV-2 infections, hospitalizations, excess mortality, and COVID-19-related mortality.

In Belgium, however, early research primarily focused on demographic and socioeconomic disparities in COVID-19 incidence at aggregated levels, and on excess mortality during the first wave [119–121]. To our knowledge, no studies at that time had quantified social inequalities across multiple severe COVID-19 health outcomes at the individual level beyond the first wave.

Filling this gap is particularly important, as Belgium recorded one of the highest COVID-19 mortality rates in Europe during the first waves [122,123], and faces a high burden of chronic conditions [29], conditions that are strong risk factors for severe COVID-19 [124], but also unequally distributed across social groups [125]. Leveraging the surveillance systems implemented during the pandemic, collecting data on morbidity and mortality, is therefore crucial to assess the extent to which social inequalities are present in severe COVID-19 outcomes. These insights are critical for identifying the groups at risk of severe COVID-19 in order to design equitable and tailored public health responses for future health crises.

## 2.2. Outline of thesis chapters and related research questions

***Chapter 2*** aims to investigate the association between the education level and the likelihood of being hospitalized with a confirmed COVID-19 diagnosis. In addition,

we aim to identify the extent to which this association is mediated by vaccination, working in healthcare, and underlying health conditions.

In ***Chapter 3***, our purpose is to identify the key SD and SE predictors of COVID-19-related death among the Belgian adult population.

## 3. What are the long-term direct health impacts of severe COVID-19 and related social patterns?

### 3.1. Knowledge gap

The previous SARS-CoV pandemic demonstrated persistent impairments in pulmonary function among patients even years after hospitalization [75]. Similarly, evidence on PASC following SARS-CoV-2 infection has been emerging [68]. In Belgium, approximately 4% of the population aged over 15 report PASC three months post-infection [126]. Nevertheless, long-term follow-up studies on PASC among hospitalized COVID-19 patients, who face a higher risk of PASC than non-hospitalized cases, remain limited internationally and are currently lacking in Belgium. Evidence on how SDOH influence PASC among hospitalized patients, a particularly vulnerable group characterized by advanced age and a high burden of chronic conditions, is also scarce. Addressing this gap is essential to quantify the burden of PASC among hospitalized patients and to guide policymakers in planning targeted long-term follow-up, rehabilitation, and care services.

## 3.2. Outline of thesis chapters and related research questions

In ***Chapter 4***, two objectives are considered: 1) assess whether COVID-19 hospitalization leads to the development of one-year post-acute organ complications, compared to non-COVID-19 hospitalization; 2) assess the association between SE determinants and the development of one-year post-acute organ complications among COVID-19 hospitalized patients.

# PART II

# Methodological framework

# Population data linkages

**Table 3** provides an overview of the population data linkages used across the thesis chapters. These three population data linkages, i.e., LINK-VACC, DEMOBEL, and HELICON, are described in detail in the following sections.

**Table 3. Overview of the research outcome, study population, and population data linkage by thesis chapter.**

| Section | Chapter | Main research outcome | Study population | Population data linkage |
|---|---|---|---|---|
| I. What are the social patterns of COVID-19-related health behaviors? | 1 | First COVID-19 vaccine dose uptake | Belgian adult population tested at least once for SARS-CoV-2 | LINK-VACC |
| II. What are the social patterns of COVID-19-related morbidity and mortality? | 2 | COVID-19 hospitalization | Belgian adult population with a positive SARS-CoV-2 test | LINK-VACC |
| | 3 | COVID-19 death | Belgian adult population | DEMOBEL |
| III. What are the long term direct health impacts of severe COVID-19 and related social patterns? | 4 | Post-acute organ complications following hospitalization | Belgian adults hospitalized | HELICON |

## *DEMOBEL*

Demobel is the demographic database of Statistics Belgium (Statbel), the Belgian statistical office. Statbel collects, produces and disseminates reliable data on the national economy, society, and territory via administrative and survey data sources [127]. Demobel consists of a nationwide record linkage of multiple administrative data sources, providing demographic data up to 1991 and covering almost 30 years of Belgian demographic history. It covers the entire Belgian population recorded in the national register on January 1st of each year. The Belgian population includes Belgians and non-Belgians admitted or authorized to settle or reside in the country. However, non-Belgians staying for less than three months, asylum seekers, and non-Belgians in an irregular situation are not included [128].

Demobel is updated annually and has both a structural (e.g., population structure by age, sex, marital status) and dynamic (e.g., births, deaths, migratory movements) component. Demobel's structural component is enriched by the censuses conducted every 10 years, which provide data on education and the labor market [128].

The Demobel data used in this thesis originate from a record linkage of the following administrative data sources: 1) the national register, providing yearly stock files with sociodemographic information for all individuals officially residing in Belgium as of 1st January of each year; 2) the administrative censuses, providing

data on education and employment; 3) the tax register, providing yearly income data; and 4) death certificates, providing yearly data on the date and causes of death using the International Statistical Classification of Diseases and Related Health Problems 10$^{th}$ Revision (ICD-10), a global system for coding causes of death established by the WHO [129].

## *LINK-VACC*

### 1. Project description

The LINK-VACC project was established by Sciensano in January 2021 to monitor post-authorization surveillance of COVID-19 vaccines. Three main objectives are considered in this project: 1) determining the uptake and coverage of COVID-19 vaccines; 2) determining the effectiveness of the vaccine; and 3) contributing to the monitoring of vaccine safety [130]. This thesis contributes to the first objective of the LINK-VACC project through ***Chapter 1*** (see **Table 3**).

### 2. Data linkage process and data sources available

To achieve the objectives of the LINK-VACC project, Sciensano established a retrospective cohort by linking COVID-19 surveillance registries with various existing national health and social sector registries. This linkage was conducted at the individual level using the National Identification Number of Social Security (NISS), within a pseudonymized environment hosted by Healthdata.be. Developed by Sciensano, Healthdata.be is a platform designed to facilitate and standardize the collection of health data from multiple healthcare providers, thereby enhancing research quality and efficiency in a reliable and secure manner through encryption [131,132].

The LINK-VACC project links eight data sources of which five are hosted by Healthdata.be within Sciensano and three come from external data sources (**Figure 9**).

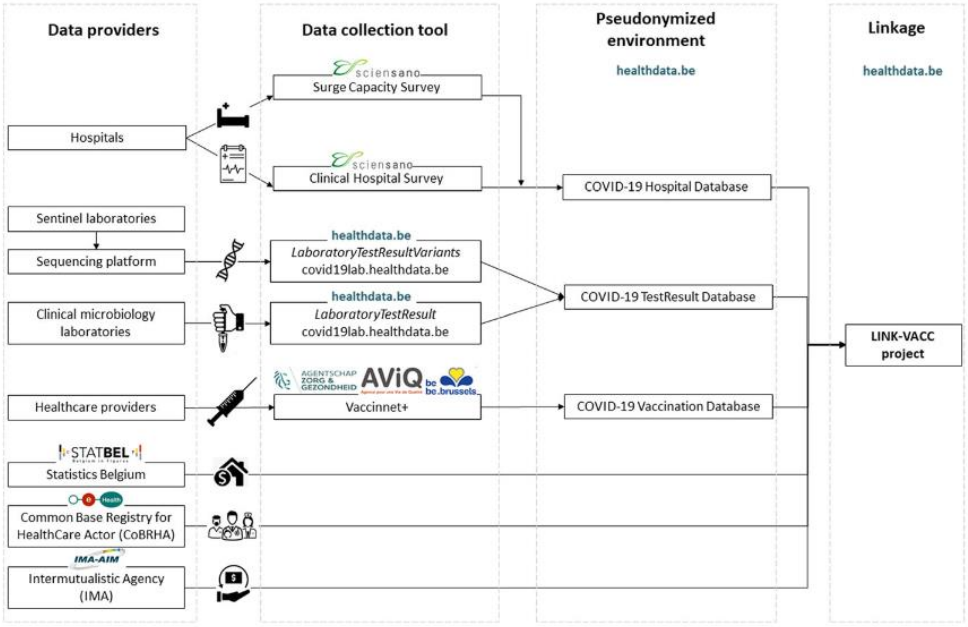



***Figure 9. Data linkage of COVID-19 surveillances registries with existing national social and health sector registries to build the LINK-VACC population data linkage.*** *Source: Van Goethem et al. (2021) [133].*

## 2.1. Data sources hosted by Healthdata.be within Sciensano

### *2.1.1. Surge Capacity and Clinical Hospital surveys*

The Surge Capacity survey (SCS) and the Clinical Hospital survey (CHS) are two complementary hospital-based surveillance systems designed to collect data on COVID-19 hospitalizations. Both surveys started in March 2020. Participation in the SCS was mandatory, whereas participation in the CHS was not mandatory but highly

recommended. To complete both surveys, healthcare staff had to fill out a secure online questionnaire created using LimeSurvey and submit it to Sciensano on a daily basis [112,134].

The objective of the SCS was to monitor bed and medical device occupancy rates in 103 Belgian general hospitals (excluding psychiatric and specialist hospitals) for patients hospitalized with a laboratory confirmed (by Polymerase Chain Reaction (PCR) or antigen test) or suspected COVID-19 diagnosis. Outpatients and day hospitalizations were not included. Prevalence and incidence data obtained through the SCS are described in **Table 4**. The SCS data were collected at an aggregated level, per hospital accreditation number [112].

**Table 4. Aggregated-level data collected through the Surge Capacity Survey.** *Source: Van Goethem et al. [112].*

| Data | Data description |
|---|---|
| Prevalence data | Number of hospitalized COVID-19 patients in all units of the hospital (including ICU) |
| | Number of hospitalized COVID-19 patients in ICU |
| | Number of hospitalized COVID-19 patients under IV |
| | Number of hospitalized COVID-19 patients under ECMO |
| Incidence data | Number of new COVID-19 hospital admissions in the last 24 hours |
| | Number of COVID-19 discharges in the last 24 hours |
| | Number of COVID-19 deaths in the last 24 hours |

*ICU: Intensive Care Unit ; IV: invasive ventilation; ECMO: Extra-Corporeal Membrane Oxygenation.*

The objectives of the CHS were to study the demographic and comorbidity profiles of COVID-19 patients and their acute health outcomes. The CHS collects data at both admission and hospital discharge for hospitalized patients with a COVID-19 diagnosis confirmed by laboratory testing (PCR or antigen test) or computed tomography scan. On 14 September 2020, an ICU discharge form was implemented. Data obtained through the CHS were collected at the individual level and are based on the WHO case report for COVID-19 and recommendations from hospital clinicians [112].

**Table 5** provides a summary of data collected through the admission, hospital discharge, and ICU discharge forms included in the CHS. A detailed version of the three questionnaires, as well as the dates of additions or modifications to the questionnaires as the pandemic evolves can be found in **Appendix A**.

**Table 5. Individual-level data collected through the Clinical Hospital Survey.** *Source: Van Goethem et al. [113].*

| Data | Data description |
|---|---|
| Admission data | Demographic data (data of birth, sex, ethnicity, postal code) |
| | Exposure risk (e.g., travel to risk region) |
| | Clinical presentation (e.g., symptoms at admission) |
| | Pre-existing comorbidities |
| | Uptake of ACE inhibitors and ARBs |
| | Diagnostic information (e.g., method of diagnostic) |

**Table 5. Continued**

| Data | Data description |
|---|---|
| Hospital discharge data | Severity indicators (e.g., ARDS, sepsis, stroke) |
| | Laboratory data |
| | COVID-19 treatments details |
| | Admission in ICU |
| | Use of IV or ECMO |
| | Status at discharge (deceased or alive) |
| ICU discharge data | Use of respiratory support (invasive vs non-invasive) |
| | Use of ECMO |

*ACE: Angiotensin-Converting Enzyme; ARBs: Angiotensin II Receptor Blockers; ICU: Intensive Care Unit; IV: Invasive Ventilation; ECMO: Extra-Corporeal Membrane Oxygenation.*

Since participation in the CHS was not mandatory, and given the limited testing capacity during the early phases of the pandemic, as well as the criteria for case definition, the CHS is not exhaustive. Overall, it covers approximately 50% of all COVID-19 hospitalized patients in Belgium, based on comparisons with the exhaustive SCS. However, this coverage can vary significantly depending on the time period considered.

### *2.1.2. The COVID-19 laboratory test result database*

The COVID-19 laboratory test result database collects PCR and antigen SARS-CoV-2 test results since 5 May 2020. Test result data are gathered from general practitioners (GPs), hospital doctors, sampling-triage centers, and clinical microbiology laboratories (CMLs). Positive, negative, and inconclusive test results

are reported. Reporting diagnostic test results is mandatory for reimbursement. In addition to test results, the database includes information on the sample collection date, result date, type of diagnostic test, and demographic variables (postal code, date of birth, and gender) [114,135].

### *2.1.3. The COVID-19 sequencing result database*

The National Reference Center (NRC) for respiratory diseases established genomic surveillance for SARS-CoV-2 in Belgium, later expanding into a federal sequencing consortium with 17 laboratories [133].

In December 2020, sequencing focused on returning travelers, atypical PCR results, and outbreaks investigations due to the emerging VOCs. Baseline surveillance started in January 2021, with a network of laboratories sending a representative proportion of positive samples for sequencing to ensure geographical and clinical diversity. The goal is to sequence about 10% of positive cases, supplemented by targeted active surveillance of reinfections, vaccine breakthrough cases, and immunocompromised patients. Some hospitals also conduct additional sequencing on ICU patients or all hospitalized cases [133].

Since March 2021, sequencing laboratories have been reported SARS-CoV-2 variant data to Sciensano. After sequencing, variants are classified using the Pangolin lineage system for standardized nomenclature [133,136].

### *2.1.4. The Belgian vaccine registry*

All COVID-19 vaccine doses administered to Belgian residents are recorded in the national vaccine registry (Vaccinnet+) by healthcare providers. Vaccinnet+ includes data on administered vaccine doses, including the brand, lot number, administration date, and registration date. It also contains demographic information about the vaccinated person, including age, gender, and place of residence [133,137].

## 2.2. Data sources outside Sciensano

### *2.2.1. Statbel*

Statbel provides demographic, SE, and mortality data for all individuals officially residing in Belgium as of January 1st of each year via the Demobel database. This database is described in detail in the previous section "DEMOBEL", page 38.

### *2.2.2. Common Base Registry for Healthcare Actors*

The Common Base Registry for HealthCare Actors (CoBRHA) database identifies all individuals licensed to practice a healthcare profession in Belgium. It includes data on the type of profession, authorized practices, and responsibilities associated with each healthcare provider or institution [138].

### *2.2.3. InterMutualistic Agency*

The InterMutualistic Agency (IMA) database collects data from the seven health insurances in Belgium, which are organized by the public health system. As the health insurance system is compulsory in Belgium, IMA covers more than 99% of the total Belgian population [139]. IMA collects and provides data for analysis and research purposes on healthcare expenditures, hospitalization, and prescribed and reimbursed medicines dispensed in public pharmacies [140].

Due to the lack of information on pre-existing chronic conditions in Belgium, IMA data on prescribed and reimbursed medicines have been used to create a proxy, called pseudo pathologies. This proxy was created by a group of experts from the National Institute for Health and Disability Insurance (RIZIV-INAMI). Groups of pseudopathologies are created on the basis of Anatomical Therapeutic Chemical (ATC) codes established by the WHO [141]. An individual is assigned to a pseudopathology group when the total number of defined daily doses (estimated quantity of active ingredient for a 70 kg adult per day of a drug prescribed for its main action) for all ATC codes in this group is greater than or equal to 90 [142]. A summary of all pseudo pathologies included in IMA, along with corresponding ATC codes, can be found in **Appendix B**.

# *HELICON*

## 1. Project description

The HELICON project was set up by Sciensano in January 2021 to unravel the social inequalities and the long-term and indirect health effects of the COVID-19 crisis in Belgium. HELICON involves six national public institutions, fostering collaboration among clinicians, epidemiologists, sociologists, demographers, and health economists [143]. To meet its main objective, HELICON includes three scientific work packages (WPs): (WP1) identifying the social patterns of COVID-19; (WP2) investigating the long-term health impacts of severe COVID-19 infections; and (WP3) assessing the indirect health effects of the COVID-19 crisis on healthcare resource use and associated costs [143].

This thesis contributes to WP1 and WP2 of the HELICON project. Consequently, the following description of data sources and study design focuses on the population data linkage established to support objectives within these two WPs. The objectives within WP3 have been explored through other data linkages, as described elsewhere by Khan et al. [144–146].

## 2. Study design and data sources available

To investigate the social patterns of COVID-19 and the long-term health impact of severe infections, HELICON established an individual-level data linkage between three national health and social sector registries using the pseudonymized NISS number. Those three data sources are as follows: 1) the CHS, described in detail in the section "*Surge Capacity and Clinical Hospital surveys*", page 41; 2) the IMA database, described in detail in the section "*InterMutualistic Agency*", page 47; and 3) Demobel, provided by Statbel, described in detail in the section "*DEMOBEL*", page 38.

The HELICON project relies on a case-cohort study design, as depicted in **Figure 10**. According to this study design, a 10% random subcohort from the Belgian population as of 1st January was sampled. This subcohort includes cases (i.e., any individual hospitalized with a confirmed COVID-19 diagnosis in a Belgian general hospital) and controls (i.e., individuals not hospitalized with a confirmed COVID-19 diagnosis). In addition to the controls and cases from the subcohort, remaining cases from the general Belgian adult population were added. Case identification started from July 2020 to July 2022. All participants were followed over a 4-year period with annual updates of all data sources.

The use of a case-cohort study design offers several advantages by combining the design of a cohort study and a nested case-control study. First, compared to a

cohort study, data on exposure and covariates are only needed for cases from the population, as well as for cases and controls from the subcohort, leading to time and cost savings. In addition, compared to a nested case-control study, controls from the subcohort can be used as the comparative group to study multiple health outcomes, preventing the creation of a new control group for each research question [147].

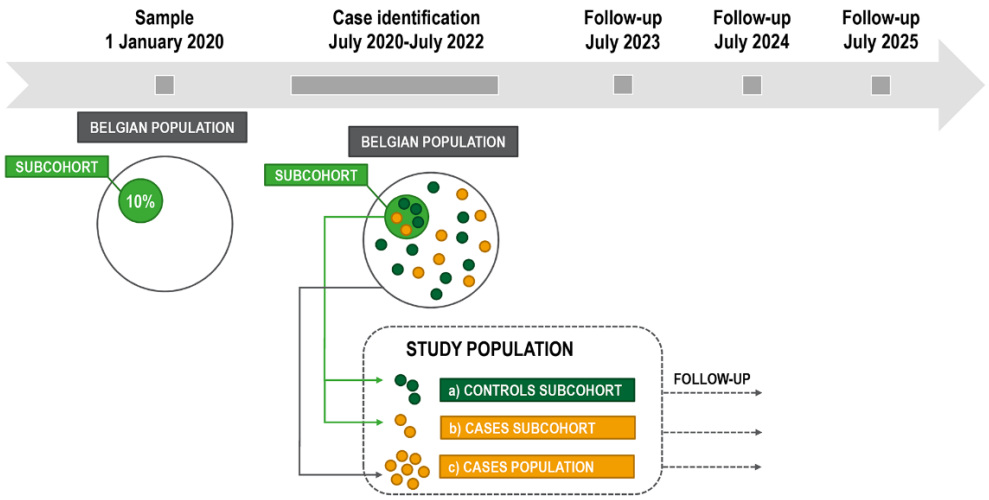
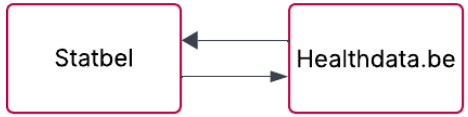
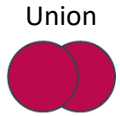
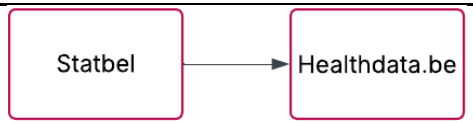
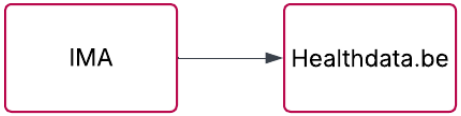



***Figure 10. Schematical overview of the case-cohort study design within the HELICON project.*** *Source: De Pauw et al. (2023) [148].*

## 3. Data linkage process

**Table 6**, adapted from De Pauw et al. [148], describes the three-step process to build the HELICON population data linkage.

***Table 6.*** **Data linkage process to build the HELICON population data linkage.** *Source: adapted from De Pauw et al. [143].*

| STEPS LINKAGE | DATASETS LINKED | SQL JOIN |
|---|---|---|
| 1. A list of identifiers of cases and controls from the subcohort is sent to Healthdata.be by Statbel via eHealth, an electronic platform. eHealth is a federal public health institution that aims to promote and support the mutual electronic exchange of information between healthcare professionals, while guaranteeing the security of information and the protection of patients and healthcare providers and medical confidentiality [149]. A list of identifiers for cases from the Belgian population is sent back from Healthdata.be to Statbel via eHealth. Both are joined to create the full cohort. |  |  |
| 2. Statbel transfers data on the subcohort and cases from the population to Healthdata.be. eHealth is responsible for the linkage through a series of (de)pseudonymization steps. Healthdata.be hosts the linked data. |  |  |
| 3. The Crossroads Bank of Social Security (CBSS) transfers IMA data on the subcohort and cases from the population to Healthdata.be. The CBSS acted as a Trusted Third Party (TTP) of the health insurance agencies. This federal public institution collects, validates, and manages data from over 3000 actors in the social sector [150]. eHealth was responsible for linkage through a series of (de)pseudonymization processes. Healthdata.be hosts the final HELICON population data linkage. |  |  |

## 4. Characterization of the HELICON cohort

The full HELICON cohort includes 1,230,103 individuals, of which 1,149,264 individuals from the subcohort (1,121,448 and 27,816 controls and cases, respectively), and 80,839 cases from the Belgian population (**Figure 11**).

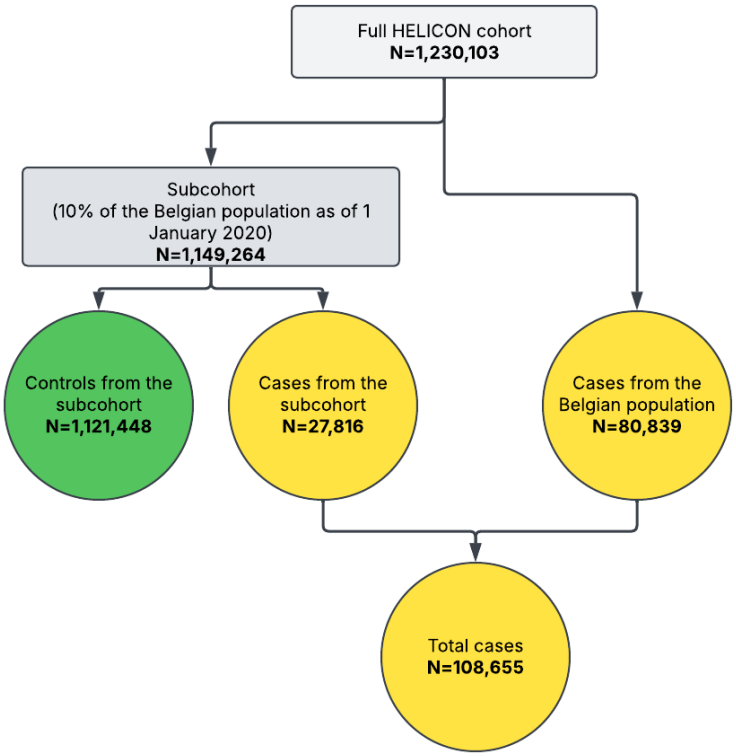



***Figure 11. Number of controls from the subcohort, cases from the subcohort, and cases from the Belgian population constituting the HELICON cohort.*** *Cases are defined as any individuals hospitalized with a confirmed COVID-19 diagnosis as recorded in the Clinical Hospital Survey (CHS) database. Controls are defined as any individual without a history of COVID-19 hospitalization, as recorded in the CHS database.*

**Table 7** presents a comparison between the HELICON cohort and the Belgian population as of January 1st 2020. On that date, the entire Belgian population included 11,492,641 inhabitants. Females accounted for 50.8% of the Belgian population (N=5,832,577) and 50% of the HELICON cohort (N=620,273). In terms of regional distribution of Belgian inhabitants as of 1st January 2020, 10.6% (N=1,218,255) resided in Brussels, 57.7% (N=6,629,143) in Flanders, and 31.7% (N=3,645,243) in Wallonia. Similar regional distributions were observed in the HELICON cohort.

**Table 7. Demographic characteristics of the HELICON entire cohort compared to the Belgian population as of 1st of January 2020.**

| Demographic characteristics | HELICON cohort | Belgian population as of 1st of January 2020 |
|---|---|---|
| All | 1,230,103 (100%) | 11,492,641 (100%) |
| Sex | | |
| Females | 620,273 (50.4%) | 5,832,577 (50.8%) |
| Males | 603,770 (49.1%) | 5,660,064 (49.2%) |
| Missing | 6,060 (0.5%) | 0 (0%) |
| Region | | |
| Brussels | 130,102 (10.6%) | 1,218,255 (10.6%) |
| Flanders | 682,596 (55.5%) | 6,629,143 (57.7%) |
| Wallonie | 389,894 (31.7%) | 3,645,243 (31.7%) |
| Missing | 27,511 (2.2%) | 0 (0%) |

## *Summary of variables included in each thesis chapter by population data linkage and time frame*

**Table 8** provides a summary of the population data linkages from which each variable used in each thesis chapter is derived, along with their respective time frames.

***Table 8.*** **Summary of variables and time frame according to each of the three population data linkages used within each thesis chapter.**

| Section | Chapter | Variables | LINK-VACC | DEMOBEL | HELICON |
|---|---|---|---|---|---|
| I. What are the social patterns of COVID-19-related health behaviors? | 1 | Vaccination status | Vaccinnet+ (31/08/2021) | | |
| | | Age groups | National register (2021) | | |
| | | Gender | National register (2021) | | |
| | | Region | National register (2021) | | |
| | | Migration background | Statbel (2021) | | |
| | | Household type | Statbel (2021) | | |
| | | Income | Statbel, tax register (2018) | | |
| | | Education | Statbel, census (2017) | | |
| | | Employment status | Statbel (2019) | | |
| | | Healthcare degree | CoBRHA (2021) | | |
| II. What are the social patterns of COVID-19-related morbidity and mortality? | 2 | Positive SARS-CoV-2 tests | Test Results database (01/09/2021-01/09/2022) | | |
| | | COVID-19 hospitalization | CHS database (01/09/2021-01/09/2022) | | |

**Table 8. Continued**

| Section | Chapter | Variables | LINK-VACC | DEMOBEL | HELICON |
|---|---|---|---|---|---|
| II. What are the social patterns of COVID-19-related morbidity and mortality? | 2 | Age | National register (01/01/2021) | | |
| | | Gender | National register (01/01/2021 | | |
| | | Province | National register (01/01/2021) | | |
| | | Migration background | Statbel (01/01/2021) | | |
| | | Education | Statbel, census (2017) | | |
| | | Vaccination status | Vaccinnet+ (14 days before the date of hospital admission) | | |
| | | Underlying health conditions | IMA (2020) | | |
| | | Healthcare degree | CoBRHA (2021) | | |
| II. What are the social patterns of COVID-19-related morbidity and mortality? | 3 | Date and causes of death | | Death certificates (01/01/2020) | |
| | | Age | | National register (01/01/2020) | |
| | | Gender | | National register (01/01/2020) | |

**Table 8. Continued**

| Section | Chapter | Variables | LINK-VACC | DEMOBEL | HELICON |
|---|---|---|---|---|---|
| | | Living situation | | National register (01/01/2020) | |
| | | Migration background | | National register (01/01/2020) | |
| | | Income | | Tax register 2018 | |
| | | Education | | Census 2011 | |
| III. What are the long-term direct health impacts of severe COVID-19 and related social patterns? | 4 | COVID-19 hospitalization | | | CHS (01/07/2020 – 31/12/2020) |
| | | ICU admission/ARDS | | | CHS (01/07/2020 – 31/12/2020) |
| | | Age | | | Demobel (01/01/2020) |
| | | Gender | | | Demobel (01/01/2020) |
| | | Region | | | Demobel (01/01/2020) |
| | | Migration background | | | Demobel (01/01/2020) |

**Table 8. Continued**

| Section | Chapter | Variables | LINK-VACC | DEMOBEL | HELICON |
|---|---|---|---|---|---|
| | | Living situation | | | Demobel (01/01/2020) |
| | | Education | | | Demobel, census 2017 |
| | | Income | | | Demobel, tax register 2018 |
| | | Date of death | | | Demobel (2020 & 2021) |
| | | Underlying health conditions | | | IMA (2020 & 2021) |

*Abbreviations : Statbel: Statistics Belgium ; CoBRHA: Common Based Registry for Healthcare Actors; IMA: InterMutualistic Agency; CHS: Clinical Hospital Survey*

# Statistical methodologies

**Table 9** provides an overview of the statistical methods and frameworks used, according to the research questions and thesis chapters.

Within the **associational inference framework**, multivariable logistic regression models were used in ***Chapters 1***, ***3***, and ***4***, as well as competing risk models in ***Chapter 3***. Associational inference refers to the process of estimating **associations** among variables under **static conditions**. These methods typically rely on (semi-)parametric **statistical assumptions** that are, in principle, **testable** [151,152].

In ***Chapters 2*** and ***4***, two statistical methods within the **causal inference framework** were used: mediation analyses using interventional effect models in ***Chapter 2***, and overlap propensity score weighting in ***Chapter 4***. Causal inference aims to estimate the effect of an exposure on an outcome by comparing what would happen under **different exposure conditions**. In observational studies, the **potential outcomes framework** is the most widely used framework for causal inference. Within this framework, the causal effect for a given unit is defined as the difference between its outcome value under the exposure and under the control condition, both measured at the same point in time [151,153,154].

This leads to the **Fundamental Problem of Causal Inference** (FPCI), as for each unit, only one of these potential outcomes can be observed, while the other remains

counterfactual [154]. Nevertheless, valid inference of causal effects from observational data is possible by relying on **three key** and often **untestable assumptions** [151,153–157]:

- *Conditional exchangeability*: conditional on observed covariates, the assignment to treatment is independent of the potential outcomes, mimicking the randomness of a controlled trial;
- *Positivity*: each level of covariates for which exchangeability is required has a non-zero probability of receiving each treatment level;
- *Consistency:* variability in the exposure should not lead to different outcomes, in other words, the exposure should be sufficiently well-defined.

In **Chapter 3**, a **machine learning** method was applied, i.e., *conditional inference tree algorithms*. Machine learning refers to a set of computational methods that enable algorithms to **identify patterns and make predictions** based on data. Machine learning methods can handle high-dimensional data and complex, non-linear relationships between variables [158].

This section of the thesis aims to describe the following statistical methods: mediation analyses using interventional effect models, overlap weighting based on propensity scores, conditional inference tree algorithms, and competing risk models.

To ensure consistency throughout this section, notations have been harmonized: the exposure is denoted by $A$, the outcome by $Y$, covariates by $X$, and mediators by $M$ (**Figure 12**).

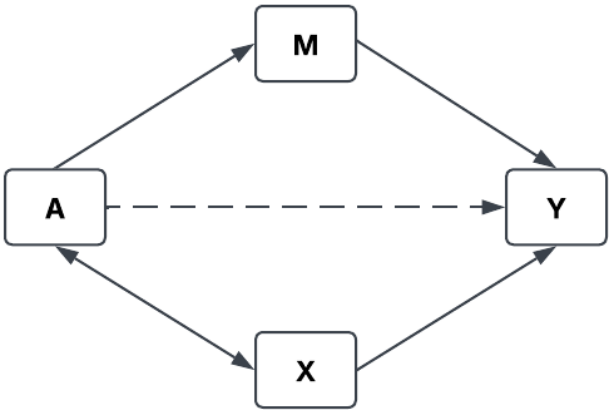



***Figure 12. Schematic overview of notations used in this thesis: exposure (A), outcome (Y), covariate (X), and mediator (M).***

**Table 9. Overview of statistical methods and frameworks used by research question and thesis chapter.**

| Sections | Chapters | Research questions | Statistical methods | Statistical frameworks |
|---|---|---|---|---|
| I. What are the social patterns of COVID-19-related health behaviors? | 1 | What are the SD and SE determinants associated with the uptake of a first COVID-19 vaccine dose? | Multivariable logistic regression models | Associational inference |
| | | Does the association between SD and SE determinants and the uptake a first COVID-19 vaccine dose differ by region? | Multivariable logistic regression models with interaction effect | Associational inference |
| II. What are the social patterns of COVID-19-related morbidity and mortality? | 2 | Does a lower level of education increase the likelihood of being hospitalized for COVID-19, and is this causal effect mediated by vaccination status, underlying health conditions, and working in healthcare? | Mediation analysis using interventional effect models | Causal inference |
| | 3 | What are the SD and SE determinants associated with COVID-19 death and all OCOD, respectively? | Multivariable logistic regression models | Associational inference |
| | | What are the strongest SD and SE predictors of COVID-19 death and all OCOD, respectively? | Conditional inference tree algorithms | Machine learning |

**Table 9. Continued**

| Sections | Chapters | Research questions | Statistical methods | Statistical frameworks |
|---|---|---|---|---|
| | | What are the relative hazards of COVID-19 death, accounting for all OCOD as competing risk, and of all OCOD, accounting for COVID-19 death as competing risk, respectively? | Competing risk models (i.e., Fine and Gray subdistribution hazard multivariable regression models) | Associational inference (survival analysis) |
| III. What are the long-term direct health impacts of severe COVID-19 and related social patterns? | 4 | Are COVID-19 hospitalized patients, compared to non-COVID-19 hospitalized controls, more likely to develop post-acute organ complications after balancing confounders (i.e., SD and SE characteristics, and baseline health status)? | Overlap propensity score-weighted logistic regression model | Causal inference |
| | | What are the SE and SD determinants associated with the onset of post-acute organ complications among COVID-19 hospitalized patients? | Multivariable logistic regression models | Associational inference |
| | | Does the effect of SD and SE determinants on post-acute organ complications differ between COVID-19 and non-COVID-19 hospitalized patients? | Multivariable logistic regression models with interaction effect | Associational inference |

*Abbreviations : SD=sociodemographic ; SE=socioeconomic ; OCOD=other causes of death.*

# *Mediation analyses using interventional effect models*

## 1. Mediation analysis: a valuable method for unraveling causal mechanisms

Mediation analysis is a valuable method for disentangling the mechanisms by which an exposure affects an outcome, through both **direct and indirect pathways**. Indirect pathways involve one or multiple mediators, which are intermediate variables that lie on the causal pathway between the exposure and the outcome. These mediators can partially or fully explain the exposure-outcome association, thereby helping to unravel the underlying causal mechanisms [159,160]. By offering insights into **how** and **why** an effect occurs, mediation analysis moves beyond the **"black-box" view of causality** and enables causal explanations that support evidence-based decision-making across a variety of disciplines [159].

## 2. Traditional parametric mediation analyses

Traditional mediation analysis uses a **model-based** approach within the framework of **linear structural equation modeling** (LSEM), often relying on the *product-of-coefficients* or the *difference-in-coefficients* methods to estimate the mediation effect [159,161].

In the simplest setting, where both the outcome and the mediator are continuous, traditional mediation analysis using the *product-of-coefficients* method is based on three parametric linear models (**Eq. 1 to 3**) [161–163]:

$$E(Y|A,X) = \beta_0 + \beta_A A + \beta_X X \quad \textbf{\textit{(Eq. 1)}}$$

where $Y$ denotes the outcome, $A$ the exposure, and $X$ a set of covariates. Here, $\beta_A$ represents the total effect (TE) of $A$ on $Y$.

$$E(M|A,X) = \alpha_0 + \alpha_A A + \alpha_X X \quad \textbf{\textit{(Eq. 2)}}$$

where $M$ denotes the mediator. The regression coefficient $\alpha_A$ captures the effect of $A$ on $M$.

$$E(Y|A,M,X) = \gamma_0 + \gamma_A A + \gamma_M M + \gamma_X X \quad \textbf{\textit{(Eq. 3)}}$$

where $\gamma_A$ represents the direct effect (DE) of $A$ on $Y$ after adjusting for $M$, while $\gamma_M$ represents the effect of $M$ on $Y$.

The indirect effect (IE) of $A$ on $Y$ through $M$ is estimated as the product of the effect of $A$ on $M$ and the effect of $M$ on $Y$ (i.e., $\alpha_A\gamma_M$). Thus, the TE is the sum of the DE and the IE. Using a Directed Acyclic Graph (DAG) [164], **Figure 13** illustrates the decomposition of the TE into IE and DE through the *product-of-coefficients* method [161,162].

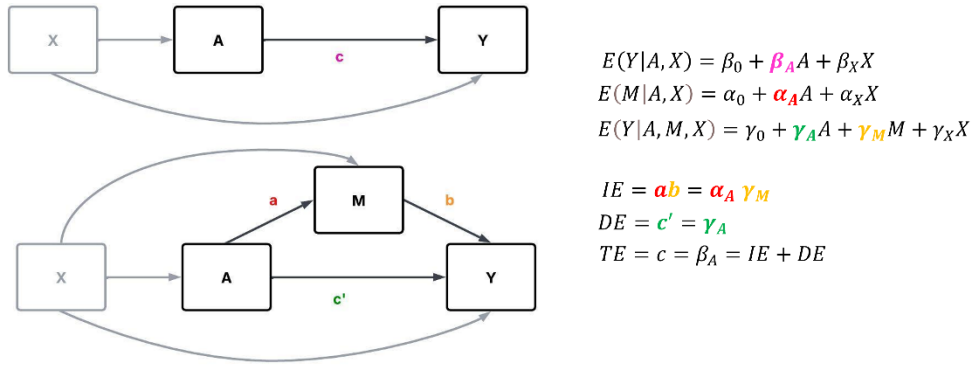



***Figure 13. Traditional approach using the product-of-coefficients method to estimate the direct effect (DE), indirect effect (IE), and total effect (TE) as functions of regression coefficients.*** *The outcome is denoted by $Y$, the mediator by $M$, the exposure by $A$ and the covariate by $X$. Source: adapted from Nguyen et al. [161].*

In a similar setting, traditional mediation analysis using the *difference-in-coefficients* method estimates the IE as the difference between $\beta_A$ from **Eq. 1** and $\gamma_A$ from **Eq. 3** [162,163].

When both the outcome and the mediator are continuous and modeled using linear regression models, the relationship $\alpha_A\gamma_M = \beta_A - \gamma_A$ holds.

## 2.1. Limitations of traditional parametric mediation analysis

Since traditional parametric mediation models are developed within the LSEM framework, the relationship $\alpha_A\gamma_M = \beta_A - \gamma_A$ does not generalize to non-linear regression models (e.g., logistic regression models) or to non- or semi-parametric models [159,162,163,165].

Additionally, traditional approaches rely on the assumption of no interaction between the exposure and the mediator, an assumption that is rarely satisfied in real-world settings [162,163,165,166].

## 3. Causal mediation analyses

To overcome the limitations of traditional mediation analysis, **causal mediation analysis** has been developed, distinguishing the **definition of causal effects** (i.e., whether effects are identifiable) from their **estimation** (i.e., how to infer them from data) [159,161,162]. Since the effect definitions are **non-parametric**, causal mediation analyses are broadly applicable than traditional parametric methods, even in non-linear models or when exposure-mediator interactions are present [165,167].

### 3.1. Causal mediation within the potential outcomes framework

A formal approach to causal mediation analysis has been developed within the potential outcome framework, enabling the decomposition of the TE into the **natural direct effect (NDE)** and the **natural indirect effect (NIE)**. Using a binary outcome and within a single mediator setting, the NDE quantifies the change in the outcome due to a shift in the exposure from $a = 0$ to $a = 1$, while holding the mediator fixed at the value it would naturally take when $a = 0$. Conversely, the NIE quantifies the change in the outcome solely due to a shift in the mediator from $M(a = 0)$ to $M(a = 1)$, when holding the exposure fixed at $a = 0$ [162,165,168].

### *3.1.1. Causal assumptions in a single mediator setting*

The estimation of natural (in)direct effects requires key assumptions for valid causal inference, known as "*sequential ignorability*" assumptions. These include the absence of confounders for the exposure-mediator, exposure-outcome, and mediator-outcome relationships [162,165,167–169].

### *3.1.2. Causal assumptions in a multiple mediator setting*

Defining the natural indirect effects of each mediator in a multiple mediators setting is challenging, as **strong extra assumptions** are required regarding mediators: 1) known causal ordering among the mediators; and 2) no unmeasured confounding between the mediators [170]. Defining the correct causal structure among the mediators and adjusting for all potential confounders is challenging and usually not feasible in practice. When these assumptions do not hold, relying on natural effects may lead to biased causal inference [169,171].

## 3.2. Causal mediation within the interventional effects framework

To address these limitations, causal mediation analysis has been developed within the **interventional effects framework**. This framework extends the potential outcomes framework by estimating indirect effects through **simulated interventions on mediator distributions**. This enables a more flexible approach when the true causal structure among the mediators is unknown [161,169,172].

### *3.2.1. Interventional effects definition*

While natural effects are defined in terms of individual-specific counterfactual mediator values, interventional effects shift the focus to the population level [170,173]. Interventional effects address the question: *"What would happen to $Y$ if we randomly drew the mediator $M$ from the distribution of $M$ conditional on $A$?"*. By contrast, natural effects aim to answer: *"What would happen to $Y$ if we changed $A$, while holding $M$ fixed at the value it would take under the alternative treatment/exposure?"*

Specifically, the mediator is treated as a random draw from its conditional distribution given $A$ and baseline covariates $X$, denoted $\mathcal{M}(a^*|X)$ [161,173]. This yields to the definition of four interventional (in)direct effects [161,171]:

two interventional direct effects (IDE):

$$\rightarrow \; IDE(0) = E\big[Y\big(1, \mathcal{M}(a^* = 0|X)\big)\big] - E\big[Y\big(0, \mathcal{M}(a^* = 0|X)\big)\big]$$

$$\rightarrow \; IDE(1) = E\big[Y\big(1, \mathcal{M}(a^* = 1|X)\big)\big] - E\big[Y\big(0, \mathcal{M}(a^* = 1|X)\big)\big]$$

and two interventional indirect effects (IIE):

$$\rightarrow \; IIE(0) = E\big[Y\big(0, \mathcal{M}(a^* = 1|X)\big)\big] - E\big[Y\big(0, \mathcal{M}(a^* = 0|X)\big)\big]$$

$$\rightarrow \; IIE(1) = E\big[Y\left(1, \mathcal{M}(a^* = 1|X)\right)\big] - E\big[Y\left(1, \mathcal{M}(a^* = 0|X)\right)\big]$$

Because these effects rely on random draws from conditional mediator distributions, they are **interpretable only at the population level** [161,169,173].

### *3.2.2. Causal assumptions*

Interventional (in)direct effects are identified under **three weaker assumptions** than those required for natural effects in a multiple mediator setting: 1) no unobserved exposure-outcome confounders; 2) no unobserved mediators-outcome confounders; and 3) no unobserved exposure-mediators confounders [171,172].

### *3.2.3. Interventional effects estimation*

Interventional (in)direct effects can be estimated using a set of linear regression models. Then, the interventional effect estimates align with those obtained from parallel path models, commonly used in a multiple mediator setting but assuming that they are not causally linked, using the *product-of-coefficient*s method [169,171,172], described previously (see p. 64). ***Chapter 2*** constitutes a practical example of the estimation of the interventional effects using a set of linear regressions for a setting with three mediators.

# *Overlap weights based on propensity scores*

## 1. Propensity score to correct for confounding bias

Confounding bias can be mitigated using propensity scores (PSs), which estimate the conditional probability of an individual being exposed, given a set of measured confounders [174–176]. By balancing the distribution of confounders between exposed and unexposed groups, PS aims to replicate the effect of random assignment achieved in randomized controlled trials (RCTs), thereby reducing baseline differences between groups and enabling more accurate comparisons. However, unlike RCTs, PS methods can only adjust for measured confounders and cannot account for unmeasured baseline characteristics [175,177].

### 1.1. Propensity score model specification for a binary exposure

When considering a binary exposure, the PS model is typically specified for each individual using logistic regression as follows:

$$logit\left\{P\left(A = 1 \mid X_1, X_2, \ldots, X_p\right)\right\} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p \quad \textbf{(Eq. 4)}$$

where $P\left(A = 1 \mid X_1, X_2, \ldots, X_p\right)$ represents the individual probability of being exposed, conditional on a set of measured confounders. Interaction terms can also be added to the PS model if necessary. Once the PS model is specified, the individuals probabilities of exposure, conditional on measured baseline confounders, are obtained [175,177].

## 2. Inverse probability of treatment weighting as the standard propensity score-based balancing weight method

Inverse probability of treatment weighting (IPTW) is the most commonly used PS-based weighting method applied to control for baseline differences between exposed and unexposed groups, by creating a weighted pseudo-population in which confounders are equally distributed [177–179]. Conventional IPTWs are estimated for each individual as $1/PS$ for exposed individuals and $1/(1-PS)$ for unexposed individuals. As a result, larger weights are assigned to exposed individuals with a lower probability of exposure, and vice versa, thereby increasing their relative influence in the weighted pseudo-population [177].

### 2.1. Limitations of inverse probability of treatment weighting

When the exposure is rare in certain groups of individuals, extreme PSs may arise, resulting in large weights when using IPTW, violating the positivity assumption. This can lead to poor covariate balance and biased causal effect estimates, with inflated variances [177,178,180–182].

A common approach to address this issue is to exclude observations with extreme PS values. However, this raises concerns about selecting an appropriate PS-based cutoff criterion and the resulting reduction in sample size [181,183]. Moreover, the impact of existing methods for determining this cutoff on bias and the precision of causal effect estimates remains uncertain [181,183].

## 3. Overlap weights for unbiased causal effect estimates in the presence of extreme propensity scores

Overlap weights (OWs) have been shown to mitigate the aforementioned issues caused by extreme PSs when using IPTW by down-weighting observations at the tails of the PS distribution, thereby reducing their influence in the weighted pseudo-population without excluding any observations. Simultaneously, OWs up-weight observations in regions where the PS distributions overlap, i.e., including individuals whose characteristics are similar regardless the exposure status. Specifically, OWs are calculated as $1 - PS$ for exposed individuals and $PS$ for unexposed individuals [180,181,183,184]. As a result, OWs are bounded, leading to reduced variance and greater precision compared to IPTW [181,183]. Notably, unlike IPTW, OWs focus the causal estimates on the overlap population, meaning the results primarily apply to individuals who are sufficiently similar to plausibly belong to either treatment group, rather than to the entire study population [185].

**Practical example**

This example, adapted from Thomas et al. [180], uses a dataset of 100 individuals simulated based on the distribution of variables included in the study by Mehta et al. [186].

This study aimed to assess whether the use of angiotensin-converting enzyme inhibitors (ACEIs) is associated with the likelihood of testing positive for SARS-CoV-2, as well as other SARS-CoV-2-related outcomes, after balancing for a set of confounders (i.e., age, sex, high blood pressure, diabetes, coronary heart disease, heart failure, and COPD) [186].

The PSs are estimated as the conditional probabilities of receiving ACEI based on age, sex, high blood pressure, diabetes, coronary heart disease, heart failure, and COPD. **Figure 14** shows the density distributions of PSs by ACEI uptake, demonstrating the presence of extreme PS values.

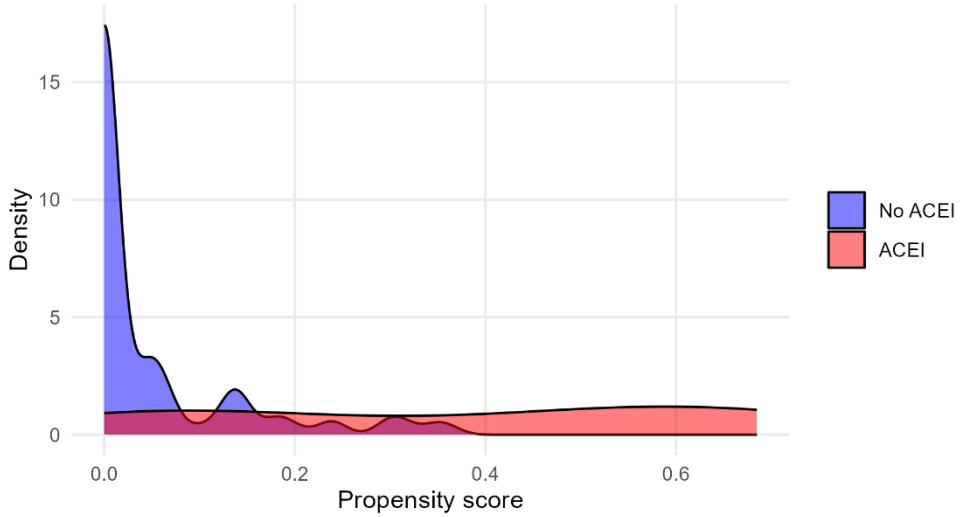



***Figure 14. Density distributions of propensity scores by angiotensin-converting enzyme inhibitors (ACEI) uptake status.*** *Source: adapted from Thomas et al. [180] and based on data simulated using the distribution of the variables used in the study by Mehta et al. [186].*

Based on two variables (i.e., age and diabetes), whose mean and frequency were higher at baseline in patients who received ACEIs, **Figure 15** shows their distributions in the unweighted population according to ACEI uptake, as well as the relative contribution of each patient when using OWs compared to IPTWs. In the overlap weighted pseudo-population (**Figure 15 B**), compared to the unweighted population (**Figure 15 A**), outliers contribute less to the weighted population, while those with overlapping characteristics between treatment groups contribute more. Conversely, IPTWs increase the weight of individuals who are underrepresented in the opposite treatment group (**Figure 15 C**). The R code to reproduce this example is available in **Appendix C**.

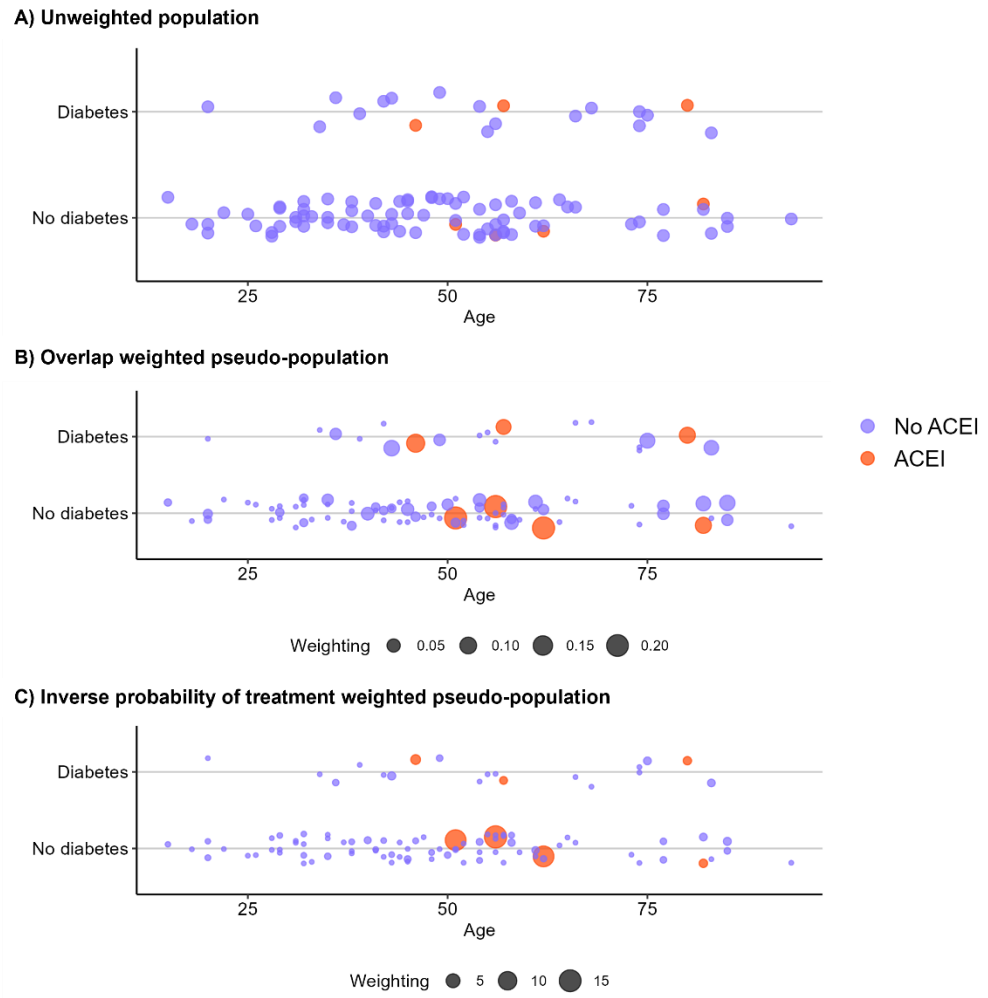



***Figure 15. Effect of overlap weighting and inverse probability of treatment weighting on the relative contributions of 100 patients simulated according to the distribution of the variables included in the study by Mehta et al. (2020) [186] according to diabetes status and age.*** *Source: adapted from Thomas et al. (2020) [180].*

### 3.1. Standardized mean difference to assess balance of baseline characteristics

To assess the effectiveness of balancing weight methods in achieving balance in baseline characteristics between exposed and unexposed groups, the standardized mean difference (SMD) is commonly used [177,179,180]. The SMD quantifies the difference in the means of baseline characteristics between exposed and unexposed groups, for both continuous and categorical covariates, and is expressed in units of standard deviation. A SMD lower than 10% usually indicates a negligeable difference in baseline characteristics between groups after weighting [175,177,179]. Based on our previous working example, **Figure 16** shows a SMD of zero for each covariate when using OWs, whereas IPTWs result in SMDs greater than 0.1 for several covariates, indicating poor balance.

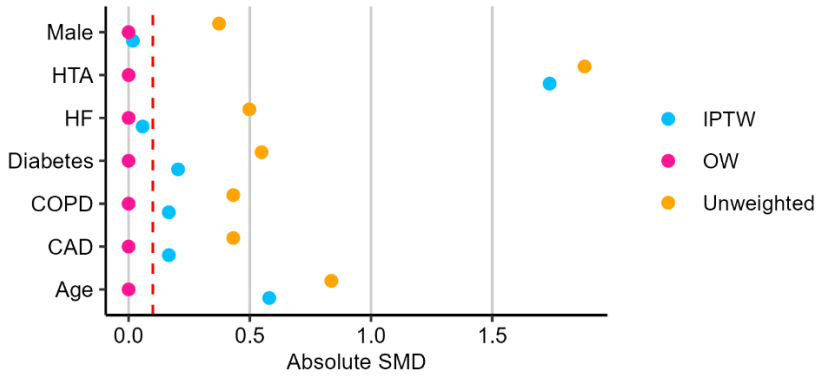



***Figure 16. Absolute standardized mean difference (SMD) assessing the effectiveness of inverse probability of treatment weighting (IPTW) and overlap weighting (OW) based on propensity scores in achieving covariate balance between treated and untreated individuals.*** *Abbreviations: HBP= high blood pressure; HF=heart failure; COPD=chronic obstructive pulmonary disease; CAD=coronary artery disease. Source: adapted from Thomas et al. (2020) [180], using data simulated according to the distribution of the variables included in the study by Mehta et al. (2020) [186].*

### 3.2. Overlap propensity score-weighted outcome model specification

After calculating OWs based on PS estimation, overlap PS-weighted logistic regression models can be applied by incorporating OWs as weights. The outcome model is specified as follows:

$$logit\,\{P(Y = 1 \mid A)\} = \beta_0 + \beta_A A \quad \textbf{\textit{(Eq. 5)}}$$

where each observation is weighted by its corresponding OW. Based on this outcome model, the overlap PS-weighted ORs can be computed as $exp(\beta_A)$, estimating disparities in the outcome between exposed and unexposed groups after balancing for confounders [183].

### 3.3. Propensity score model misspecification

Compared to IPTW, OW estimators are more sensitive to misspecification of the PS model, which can lead to biased causal effect estimates [182]. Such misspecification may result from the omission of important confounders or interactions and should be assessed using the SMD, as explained above [177]. Additionally, combining the outcome regression model with overlap PS weighting can improve the robustness of the estimators beyond weighting alone, provided that at least one of the models is correctly specified. However, this approach does not generally yield doubly robust estimations, which require specific methods, such as fitting a canonical link GLM fitted via maximum likelihood with weighted score equations, followed by regression standardization [187].

# *Conditional inference tree algorithms*

## 1. Conditional inference trees: a class of classification trees

Conditional inference trees (CITs) are a class of classification trees (CTs), a powerful machine learning technique that complements traditional regression models. CTs are particularly useful for evaluating the **predictive relationships** between a categorical outcome variable and a set of categorical or continuous predictors, especially when the **associations are non-linear or involve complex interactions** [188,189].

### 1.1. Limitations of classification trees

While CTs offer a clear and intuitive approach to modeling complex interactional relationships within a non-parametric framework, they face two major issues [188–191]:

- **Overfitting:** CTs do not incorporate the concept of statistical significance when determining splits. As a result, even minimal and potentially non-significant improvements in node purity can lead to additional splits. This can cause the tree to overfit the training data by capturing noise rather than meaningful patterns, thereby reducing the model's generalizability and robustness [189,191];

- **Biased variable selection**: CT algorithms search for the optimal split by simultaneously evaluating all possible splits across all variables, selecting the one that maximizes a chosen splitting criterion (e.g., the Gini index). This method tends to favor variables with more potential split points or a higher number of missing values, regardless of their true predictive value [191,192].

Although pruning procedures have proven effective in reducing overfitting and generating more optimal trees [188], biased variable selection remains an issue in CT interpretation, ultimately limiting the accuracy and reliability of its predictions [191]. To overcome these issues, Horton et al. [191] introduced the CIT algorithm, which incorporates statistical significance testing into the splitting process.

## 2. Conditional inference trees and the concept of statistical significance

CITs integrate **statistical significance testing** into the tree-building process by using p-values derived from conditional permutation tests. This approach, also known as unbiased conditional recursive partitioning, recursively partitions the data based on the strength of the association between the predictor variables and the outcome, modeling regression relationships within a conditional inference framework [190,191,193].

2.1. Conditional inference tree algorithm mechanisms

The CIT algorithm operates through three main steps [190,191,194]:

(1) **Global null hypothesis testing and variable selection:** the algorithm initiates with the entire sample and seeks to identify whether there is any association between the response variable and each predictor. For this purpose, it assesses the correlation ($\rho_{X_j,Y}$) between the response variable (*Y*) and any random variable from the set of predictors ($X_j$). The global null hypothesis ($H_0$*:* $\rho_{X_j,Y} = 0$) between *Y* and $X_j$ is assessed using a statistical test (*i.e.*, conditional permutation test). If $H_0$ is rejected ($\mathrm{H}_1$: $\rho_{\mathrm{X_j,Y}} \neq 0$), the predictor $X_j$ with the strongest association (i.e., the lowest p-value) with *Y* is selected to constitute the root node.

(2) **Binary split on the selected input variable:** once a predictor is selected based on its strongest association with the outcome, permutations tests are used to evaluate all possible binary splits of that predictor.

(3) **Recursive application of steps 1 and 2:** for each resulting subset, the algorithm recursively applies conditional permutation tests to all predictor variables $X_j$ and selects the one that maximizes the pre-specified significance level α for splitting that subset. The process continues until the stopping criteria are met.

**Practical example**

**Figure 17** presents the graphical output from a CIT fitted on a training dataset including 70% of a simulated sample of 1,000 individuals, considering education and income as predictors of disease status.

The CIT algorithm first identifies education as the predictor most strongly associated with disease onset (p-value < 0.001), partitioning the **root node** into one terminal node including individuals with high education, and one **internal node** including those with low education. The algorithm then selects income as the next most significant predictor (p-value<0.001), further splitting the internal node into two **terminal nodes** based on whether individuals have low or high income.

Using a predictive threshold of 0.5, the model predicts disease onset among individuals with both low income and low education, with a 70% disease prevalence in this terminal node, compared to 40% in the other terminal nodes. The R code to reproduce this example is provided in **Appendix D**.

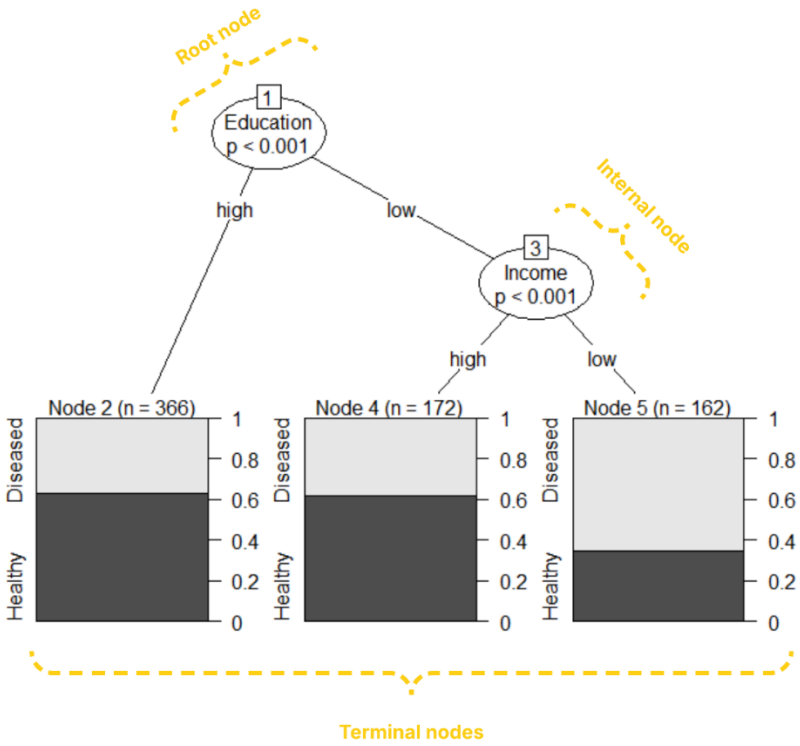



***Figure 17. Conditional inference tree output considering diseased individuals as cases, healthy individuals as controls, and income and education as predictors, fitted on a simulated training dataset (N=700).*** *A significance level of 5% ($\alpha = 0.05$) and a predictive threshold of 0.5 have been used.*

### *2.1.1. Stopping criteria*

At each recursive step, the algorithm evaluates whether specific stopping criteria have been met. Once these criteria are satisfied, further splitting ceases, and the current internal nodes become terminal nodes of the tree. Common splitting criteria include: a) the null hypothesis ($H_0$) can no longer be rejected at a pre-specified level α; b) a minimum node size is reached; or c) a maximum tree depth is set to prevent the model from becoming overly complex [190,191,194].

## 2.2. Evaluating the predictive performance of the conditional inference tree

Once the CIT model is trained on the training dataset, its ability to generalize and make accurate predictions on new data that it has not encountered during training (i.e., the test dataset) is assessed [189].

To this end, two commonly used metrics are **sensitivity** and **specificity** [195]. Sensitivity measures how well the algorithm classified truly diseased individuals (i.e., true positives), while specificity identifies how well the algorithm classified truly healthy individuals (i.e., true negatives) [195]. The closer the values of sensitivity and specificity are to 1, the better the predictive performance of the model. However, in practice, there is often a trade-off governed by the predictive threshold: increasing sensitivity typically reduces specificity and vice versa [196].

One limitation of using sensitivity and specificity is their dependence on a single threshold, requiring the selection of an appropriate trade-off between the two statistical measures [197]. The **Receiver Operating Characteristic** (ROC) curve provides a graphical representation of the trade-off by plotting sensitivity against (1 - specificity) across a range of threshold values. The **Area Under the Curve** (AUC) is then calculated and summarizes the model's overall predictive performance. The closer the AUC is to 1, the better the model's predictive performance. A diagonal line from $(0,0)$ to $(1,1)$ represents the ROC curve for random chance. An AUC value below 0.5 indicates predictive performances worse than random chance [196].

**Practical example**

Based on the previous working example, **Table 10** presents the classification results using a predictive threshold of 0.5. The CIT model has a sensitivity of 44% $\big(63/(63+80)\big)$, meaning that 44% of truly diseased individuals were correctly classified. The specificity was 83% $\big(131/(131+26)\big)$, meaning that 83% of truly healthy individuals were correctly identified. The R code to calculate the sensitivity and specificity and their 95% CIs for this example can be found in **Appendix D**.

**Table 10. Contingency table comparing actual and predicted health status based on the conditional inference tree model using a classification threshold of 0.5.** TN: true negative; FP: false positive, FN: false negative; TP: true positive.

| | | **Predicted health status** | |
|---|---|---|---|
| | | *Healthy* | *Diseased* |
| **Actual health status** | *Healthy* | 131 (TN) | 26 (FP) |
| | *Diseased* | 80 (FN) | 63 (TP) |

**Figure 18** shows the ROC curve for the CIT model, along with the AUC and their 95% CI. The AUC was 0.66 [95% CI 0.61 – 0.73], meaning that the model performs better than random chance, though with moderate discriminative performance. The R code for reproducing this analysis is available in **Appendix D**.

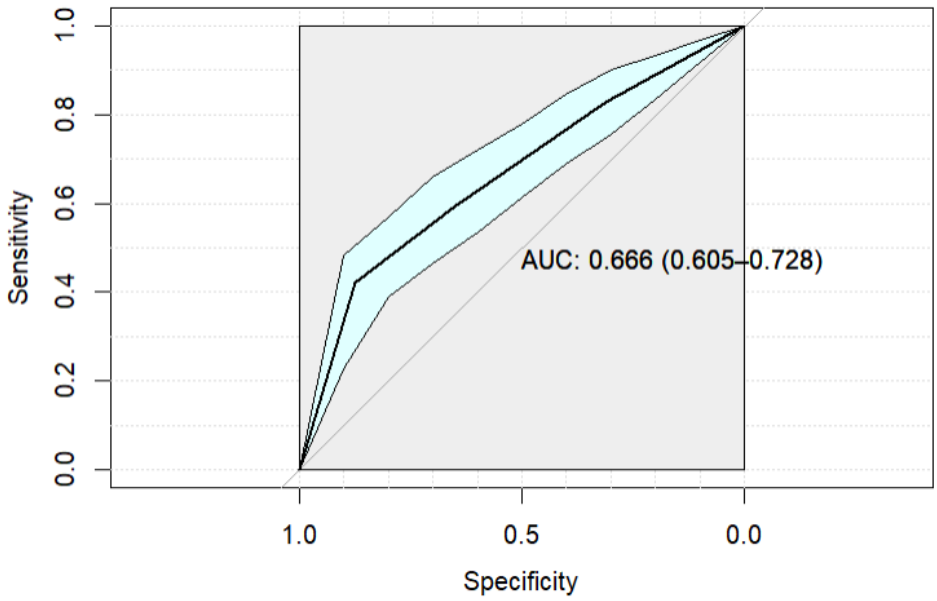



***Figure 18. ROC curve and AUC with 95% CI for the conditional inference tree model applied to the fictive test dataset (N=300), evaluating model's ability to predict disease status based on income and education.***

## *Competing risk models*

### 1. Limitations of traditional survival analysis

Traditional survival analyses, such as the Kaplan-Meier (KM) estimator or the Cox proportional hazard regression model, assume that individuals who do not experience the event of interest remain at risk indefinitely. However, in many real-world scenarios, individuals may experience an alternative event that prevents the occurrence of the primary event, known as a **competing risk**. Once such an event occurs, the individual is no longer at risk for the event of interest.

Traditional survival analyses treat these competing events as censored in the same way as other conventional censoring, even though their true survival time for the primary event is no longer observable [198,199]. However, this might lead to biased estimates of survival probabilities and hazard rates by violating a key assumption of conventional survival analysis: the assumption of **non-informative (or independent) censoring** [198,199]. This assumption implies that the survival time and the censoring time must be independent.

However, in competing risk settings, the assumption of non-informative censoring does not hold because the occurrence of a competing event directly alters the likelihood of the primary event occurring [200]. For example, in traditional survival analyses, individuals who die from causes unrelated to the cause of death of

interest are treated as censored, assuming that their removal from observation is unrelated to their survival time. As a result, traditional survival analyses violate the assumption of non-informative censoring and estimate what the survival probability would have been if the competing event had not occurred, rather than the actual survival probability of experiencing the event of interest [200].

Therefore, to address this limitation, **competing risk models** were developed, providing a more accurate framework for analyzing survival data in the presence of multiple possible outcomes [199].

## 2. Competing risk models to consider multiple outcomes

To take into account the competing nature of multiple events in modeling survival data, competing risk models include [198,200–202]:

- **The cumulative incidence function (CIF)**, which should be used instead of the KM estimator in the presence of competing events. In addition, the **Gray's test** can be used instead of the log-rank test to compare the equality of CIF curves between groups;
- **The Fine & Gray subdistribution hazard regression model**, which can be used instead of the Cox proportional hazard model.

2.1. The cumulative incidence function

The CIF is the cumulative probability of experiencing the event of interest at any point between baseline and time $t$, while accounting for the presence of competing events. For example, a CIF of 10% at time $t$ means that the estimated probability of experiencing the event of interest by time $t$ is 10%, accounting for competing events [198,199,201,203]. Mathematically, the CIF for the $k^{th}$ event is defined as follows [201]:

$$CIF_k(t) = \sum_{t_j \le t} \hat{S}(t_{j-1})\, \hat{h}_k(t_j) = \sum_{t_j \le t} \hat{S}(t_{j-1}) \frac{d_k(j)}{r(j)} \quad \textbf{\textit{(Eq. 6)}}$$

where $\hat{S}(t_{j-1})$ is the estimate of the overall survival probability at time $t_{j-1}$ and $\hat{h}_k(t_j) = \frac{d_k(j)}{r(j)}$ is the estimate of the hazard defined as the ratio between the number of events of type $k$ at time $j$, denoted $d_k(j)$, divided by the number of individuals at risk just before time $j$, denoted $r(j)$. This means that, at each event time, the contribution to the cumulative incidence is determined by the probability of surviving without any event up to that time, i.e., $\hat{S}(t_{j-1})$, multiplied by the instantaneous probability of experiencing the event of type $k$ at that time, i.e., $\hat{h}_k(t_j)$. Summing these contributions across all time points up to $t$ yields the cumulative probability of experiencing the event of interest by time $t$, while appropriately accounting for the presence of competing risks.

To better understand this mathematical definition, and to show how the CIF properly accounts for competing risks, unlike the KM estimator which treats them as censored, we consider the following example adapted from Satagopan et al. [204].

**Practical example**

Let us take a sample of 10 individuals followed over a 10-month period, where the event of interest is death from COVID-19, denoted $k = 1$, and the competing event is death from all other causes, denoted $k = 2$ (**Table 11**).

**Table 11. Sample of 10 individuals followed up over a 10-month period to assess the probability of dying from COVID-19 (k=1) in the presence of a competing event, i.e., other causes of death (k=2).** OCOD=all other causes of death; CENS= censored.

| $\mathbf{t(1)}$ | $\mathbf{t(2)}$ | $\mathbf{t(3)}$ | $\mathbf{t(4)}$ | $\mathbf{t(5)}$ | $\mathbf{t(6)}$ | $\mathbf{t(7)}$ | $\mathbf{t(8)}$ | $\mathbf{t(9)}$ | $\mathbf{t(10)}$ |
|---|---|---|---|---|---|---|---|---|---|
| COVID | CENS | OCOD | COVID | COVID | CENS | OCOD | OCOD | COVID | CENS |
| $k = 2$ | | $k = 2$ | $k = 1$ | $k = 1$ | | $k = 2$ | $k = 2$ | $k = 1$ | |

For example, at time $t(4)$, the risk set decreases from 10 to 7 individuals as one individual died from COVID-19 at $t(1)$, one was censored at $t(2)$, and another died from OCOD at $t(3)$, therefore considered as censored in the KM approach. The survival probability at time $t(4)$ is therefore calculated as $1 - 1/7 = 0.86$. The overall survival probability at four months is then estimated by multiplying the survival probabilities at $t(1)$ and $t(4)$: $0.9 * 0.86 = 77\%$.

As a result, the incidence of COVID-19 death at $t(4)$ is estimated as $1 - 0.77 = 23\%$. This process continues for each event time until the last one $t(9)$, where the incidence of COVID-19 death is estimated at 68% (**Table 12**).

**Table 12. Estimation of the overall probability of surviving COVID-19 death $(\boldsymbol{k = 1})$ using the Kaplan-Meier estimator $\widehat{\boldsymbol{S}}_{\boldsymbol{KM}}(\boldsymbol{t_j})$ considering the competing event, i.e., other causes of death $(\boldsymbol{k = 2})$, as censored.** *Source: adapted from Satagopan et al. [204].*

| Event times | Risk set | Number of events of type $k = 1$ | Survival probability | KM estimator of overall survival probability | Incidence |
|---|---|---|---|---|---|
| $\boldsymbol{t(j)}$ | $\boldsymbol{r(j)}$ | $\boldsymbol{d_{k=1}(j)}$ | $\boldsymbol{1 - \frac{d_{k=1}}{r(j)}}$ | $\boldsymbol{\widehat{S}_{KM}(t_j)}$ | $\boldsymbol{1 - \widehat{S}_{KM}(t_j)}$ |
| $t(0)$ | 10 | 0 | 1 | 1 | 0 |
| $t(1)$ | 10 | 1 | 0.9 | 0.9 | 0.1 |
| $t(4)$ | 7 | 1 | 0.86 | 0.77 | 0.23 |
| $t(5)$ | 6 | 1 | 0.83 | 0.64 | 0.36 |
| $t(9)$ | 2 | 1 | 0.5 | 0.32 | 0.68 |

Let us now consider deaths from OCOD as competing events in the estimation of the cumulative incidence of death from COVID-19 (**Table 13**). The first step is to estimate the overall survival probability using the KM estimator, but this time accounting for all events, meaning that competing events are no longer treated as censored. In this approach, the survival probability is updated at each event time, whether the event is COVID-19 death or OCOD.

As a result, the overall survival probability at 9 months is estimated at 14%, compared to 32% when competing events were treated as censored (see **Tables 12** and **13**, respectively).

The second step involves estimating the cumulative incidence of death from COVID-19 (**Table 13**). In this step, only the event times at which the event of interest occurs are considered. At each event times, the cause-specific failure probability (i.e., the probability of failing from COVID-19 death at time $t(j)$, given that the individual was still at risk just prior to $t(j)$) is calculated by dividing the number of COVID-19 deaths at $t(j)$ by the size of the risk set at $t(j)$. Then, the incidence of COVID-19 death at $t(j)$ is obtained by multiplying this failure probability by the overall survival probability (from Step 1) just prior to that time point, such as: $\hat{S}_{CR}(t_{j-1})\,\hat{h}_{k=1}(j)$. The CIF is then estimated as the sum of these products over all relevant time points: $CIF_{k=1}(t) = \sum_{t_j \leq t} \hat{S}_{CR}(t_{j-1})\,\hat{h}_{k=1}(j)$.

Using this method, the cumulative incidence of death from COVID-19 at 9 months is estimated at 47% (**Table 13**), compared to 68% when using the KM estimator (**Table 12**). This illustrates how the KM estimator can overestimate the incidence of the event of interest when competing risks are present. However, when competing risks are absent, both estimators are equivalent, therefore $CIF_k(t) = \sum_{t_j \leq t} \hat{S}_{CR}(t_{j-1})\,\hat{h}_{k=1}(j) = 1 - \hat{S}_{KM}(t_j)$.

***Table 13.*** **Estimation of the cumulative incidence of COVID-19 death $(k = 1)$ when considering other causes of death $(k = 2)$ as competing risks**. *Source: adapted from Satagopan et al. [204].*

| **Step 1: estimate the overall survival probability using KM estimator considering all events $k = 1, 2$.** | | | | |
|---|---|---|---|---|
| **Event times** | **Risk set** | **Number of events of type $k = 1$ or $k = 2$** | ***Survival probability*** | **KM estimator of overall survival probability** |
| $t(j)$ | $r(j)$ | $d_{k=1,2}(j)$ | $1 - \frac{d_{k=1,2}(j)}{r(j)}$ | $\widehat{S}_{KM}(t_j)$ |
| $t(0)$ | 10 | 0 | 1 | 1 |
| $t(1)$ | 10 | 1 | 0.9 | 0.9 |
| $t(3)$ | 8 | 1 | 0.88 | 0.79 |
| $t(4)$ | 7 | 1 | 0.86 | 0.68 |
| $t(5)$ | 6 | 1 | 0.83 | 0.56 |
| $t(7)$ | 4 | 1 | 0.75 | 0.42 |
| $t(8)$ | 3 | 1 | 0.67 | 0.28 |
| $t(9)$ | 2 | 1 | 0.5 | 0.14 |

**Table 13. Continued**

| Step 2: estimate the cumulative incidence of experiencing the event of interest $k = 1$ | | | | | | |
|---|---|---|---|---|---|---|
| **Event times** | **Risk set** | **Number of events of type $k = 1$** | ***Failure probability for event $k = 1$*** | **KM estimator of overall survival probability *at time $t_{j-1}$*** | ***Incidence*** | ***Cumulative incidence*** |
| $t(j)$ | $r(j)$ | $d_{k=1}(j)$ | $\widehat{h}_{k=1}(j) = \frac{d_{k=1}(j)}{r(j)}$ | $\widehat{S}_{KM}(t_{j-1})$ | $\widehat{S}_{CR}(t_{j-1})\,\widehat{h}_{k=1}(j)$ | $\sum_{t_j \le t} \widehat{S}_{CR}(t_{j-1})\,\widehat{h}_{k=1}(j)$ |
| $t(0)$ | 10 | 0 | 0 | 1 | 0 | 0 |
| $t(1)$ | 10 | 1 | 0.1 | 1 | 0.1 | 0.1 |
| $t(4)$ | 7 | 1 | 0.14 | 0.79 | 0.11 | 0.21 |
| $t(5)$ | 6 | 1 | 0.17 | 0.68 | 0.11 | 0.33 |
| $t(9)$ | 2 | 1 | 0.5 | 0.28 | 0.14 | 0.47 |

**Figure 19** shows the estimated cumulative incidence of COVID-19 over time using the KM estimator compared to the competing risk estimator. The KM estimator clearly overestimates the probability of death from COVID-19, as the non-informative assumption is violated. The R code to reproduce this example can be found in **Appendix E**.

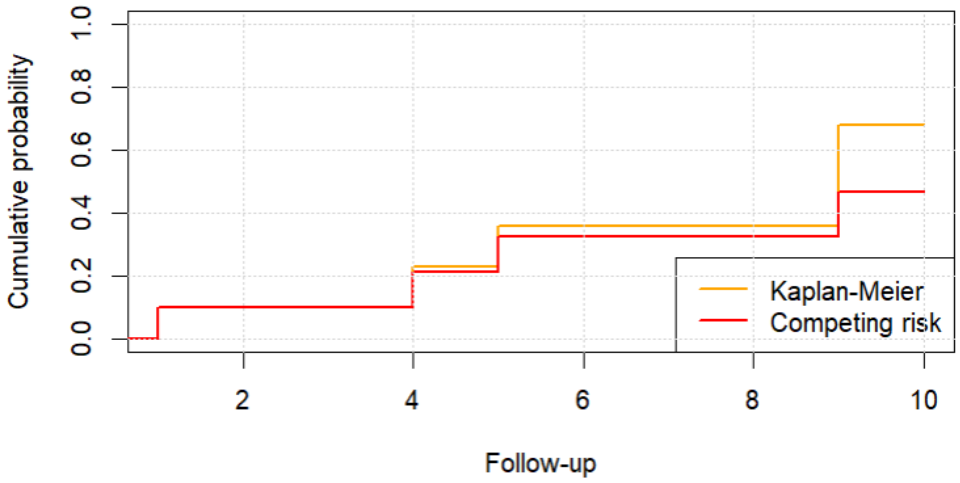



***Figure 19. Cumulative probability of dying from COVID-19 over a 10-month follow-up period estimated via the Kaplan-Meier and competing risk methods, respectively.***

*2.1.1. Gray's test to compare cumulative incidence functions*

The **Gray's test**, consisting in modified Chi-squared test, is a non-parametric method to compare CIFs between groups. It tests the null hypothesis $H_0$ that the CIFs of the groups are equal. This test does not rely on the assumption of noninformative censoring, thus reducing the risk of bias in the comparison of the CIFs [201,205].

**Practical example**

Building on the previous example, where COVID-19 death is the event of interest and OCOD are treated as competing risks, we now increase the sample size to 1,000 individuals and introduce a new variable: the education level (low vs. high). **Figure 18** presents the CIFs for COVID death by education level, accounting for OCOD as competing risks. At 10 months, the estimated CIF for individuals with low education is 45% [39% - 51%] versus 25% [18% - 33%] for those with high education.

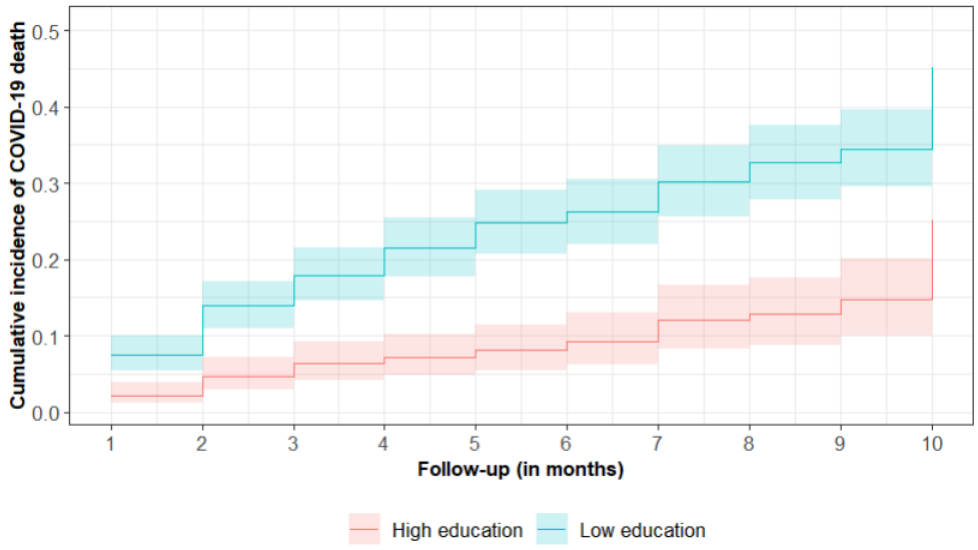



***Figure 20. Cumulative incidence function curves and 95% confidence intervals for COVID-19 death by education level, accounting for all other causes of death as competing risks.***

The Gray's test indicates that the difference between the two CIFs is statistically significant (p-value < 0.001). The R code to reproduce this example can be found in **Appendix F**.

## 2.2. Subdistribution hazard regression models

Traditional survival models, such as the Cox proportional hazard model, estimate the effect of covariates on the cause-specific hazard function, where competing events are considered as censored. These models therefore violate the assumption of non-informative censoring, which may lead to biased estimates, as previously described.

To address this issue, the **Fine & Gray model** was developed to directly **assess the effect of covariates on the CIF** for the event of interest. This model is based on a modified version of the hazard function, known as the **subdistribution hazard function** [206].

The subdistribution hazard function represents the instantaneous rate at which the event of interest occurs among individuals who have not yet experienced that event, including those who have already experienced a competing event. Mathematically, the subdistribution hazard function is expressed as follows [202]:

$$h_{c,\,CIF}(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T_c < t+\Delta t \mid T_c > t \cup T_{c'} \leq t, c' \neq c)}{\Delta t}, \quad \forall\, t > 0 \quad \textbf{\textit{(Eq. 7)}}$$

where $T_c$ is the time to the event of interest of type $c$, and $T_{c'}$ is the time to a competing event of type $c'$.

This formulation ensures that individuals who experience a competing event at time $t$ remain in the risk set for the event of type $c$ at time $t + 1$, in contrast to the

standard hazard function, where such individuals are treated as censored and removed from the risk set [198]. Based on the subdistribution hazard function, the effect of covariates on the CIF can be directly modeled through a CIF-based proportional hazard model, expressed as follows [198,199,201]:

$$h_{c,\,CIF}(t) = h_{0c,\,CIF}(t)\, exp\left(\sum_{i=1}^{P} X_i * \lambda_i\right) \quad \textbf{\textit{(Eq. 8)}}$$

where:

- $h_{c,\,CIF}(t)$ is the subdistribution hazard function for the event of interest $c$ at time $t$;
- $CIF$ is the cumulative incidence function, representing the probability of experiencing the event of interest $c$ while accounting for competing events;
- $h_{0c,\,CIF}(t)$ is the baseline subdistribution hazard function for cause $c$;
- $X_{\mathrm{i}}$ represents a covariate;
- $\lambda_{\mathrm{i}}$ are the regression coefficients estimating the effect of covariates in the presence of competing event.

The results from the Fine & Gray subdistribution hazard model are usually reported using subdistribution hazard ratios (SHRs), calculated as $exp(\lambda_{\mathrm{i}})$. The SHR quantifies the relative change in the subdistribution hazard for a one-unit change in the corresponding covariate [207].

# PART III

# Observational studies on social health inequalities in the COVID-19 pandemic and its long-term direct health impacts in Belgium

# SECTION I

# What are the social patterns of COVID-19-related health behaviors?

# *Chapter 1. Social health inequalities in COVID-19 vaccination coverage*

***Adapted from***

**Lisa Cavillot**, Joris A.F. van Loenhout, Brecht Devleesschauwer, Chloé Wyndham-Thomas, Herman Van Oyen, Jinane Ghattas, Koen Blot, Laura Van den Borre, Matthieu Billuart, Niko Speybroeck, Robby De Pauw, Veerle Stouten, Lucy Catteau & Pierre Hubin. **Sociodemographic and socioeconomic disparities in COVID-19 vaccine uptake in Belgium: a nationwide record linkage study.** *J Epidemiol Community Health*. 2023 Dec 26; 78(3): 176-83.

## Abstract

**Background**: recent studies have identified important social inequalities in SARS-CoV-2 infection and related COVID-19 outcomes in the Belgian population. The aim of our study was to investigate the sociodemographic and socioeconomic characteristics associated with the uptake of COVID-19 vaccine in Belgium.

**Methods**: we conducted a cross-sectional analysis of the uptake of a first COVID-19 vaccine dose among 5,342,110 adults (≥18 years) in Belgium on August 31st 2021. We integrated data from four national data sources: the Belgian vaccine register (vaccination status), COVID-19 Healthdata (laboratory test results), DEMOBEL (sociodemographic/socioeconomic data), and the Common Base Register for HealthCare Actors (individuals licensed to practice a healthcare profession in Belgium). We used multivariable logistic regression analysis for identifying characteristics associated with not having obtained a first COVID-19 vaccine dose in Belgium and for each of its three regions (Flanders, Brussels, and Wallonia).

**Results**: during the study period, 10% (536,716/5,342,110) of the Belgian adult population included in our study sample was not vaccinated with a first COVID-19 vaccine dose. A lower COVID-19 vaccine uptake was found among young individuals, men, migrants, single parents, one-person households, and disadvantaged socioeconomic groups (with lower levels of income and education,

unemployed). Overall, the sociodemographic and socioeconomic disparities were comparable for all regions.

**Conclusions**: the identification of sociodemographic and socioeconomic disparities in COVID-19 vaccination uptake is critical to develop strategies guaranteeing a more equitable vaccination coverage of the Belgian adult population.

## 1. Background

The presence of a social gradient in the Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) infections and subsequent coronavirus disease 2019 (COVID-19) outcomes has been clearly demonstrated through various international studies with a higher risk of getting infected and developing severe outcomes (e.g. hospitalization, death) among certain sociodemographic (SD) groups (e.g. men, elderly, migrants) and socioeconomic (SE) disadvantaged groups [208–210]. In Belgium, the same patterns have been identified, whereby the most disadvantaged SE groups presented a higher incidence of SARS-CoV-2 infection as well as higher levels of excess mortality during the COVID-19 pandemic [121,211,212].

Vaccination is a powerful public health instrument to decrease the transmission of the virus and lower the probability of developing severe COVID-19 health outcomes [213,214]. However, vaccine equity concerns remain as some studies have shown a lower COVID-19 vaccine uptake among SE disadvantaged groups [215–217]. The vaccine uptake is influenced by a range of factors, including vaccine accessibility, awareness (information about the disease and the vaccine), and willingness which can be influenced by the socio-cultural environment, political and religious orientation, or pre-existing health needs [218–220]. Understanding the social patterns in COVID-19 vaccine uptake is crucial for future vaccination campaigns against COVID-19 or other diseases and in the context of pandemic preparedness.

Belgium's nationwide vaccination campaign started on 28 December 2021, targeting first nursing home residents and staff, hospital-based and frontline health professionals, residents and staff of collective care facilities, individuals aged 65 and over, and individuals with comorbidities, before extending the vaccination campaign to all individuals aged 18 and over in order of decreasing age [44,47]. By 31 August 2021, 90% of the Belgian adult population invited for the vaccination had received the first COVID-19 vaccine dose. The objective of the present study was to investigate SD and SE disparities in the uptake of the first COVID-19 vaccine dose in the Belgian adult population.

## 2. Materials & methods

### 2.1. Study design and data sources

In this nationwide record linkage study, we investigated SD and SE disparities in the uptake of a first COVID-19 vaccine dose using data from the LINK-VACC project. LINK-VACC was set-up by Sciensano, the Belgian Institute of Health, and contains, within a pseudonymized environment, selected variables from multiple existing national health and social sector registers linked at an individual level using the Belgian social security number [130,133]. For the present study, four databases were used: 1) The Belgian vaccine register (Vaccinnet+), containing data on COVID-19 vaccine doses administered to Belgian residents as well as demographical data on the vaccinated person, 2) the COVID-19 Healthdata test database, containing

data from COVID-19 laboratory test performed in Belgium as well as demographical data on the tested person, 3) DEMODEL database provided by the Statistics Belgium (Statbel) containing variables related to SD and SE characteristics as well as information on the status in the national register, 4) The Common Base register for HealthCare Actors (CoBRHA) allowing the identification of individuals licensed to practice a healthcare profession in Belgium.

### 2.2. Study population definition

The study population consists of all individuals residing in Belgium aged 18 years and over tested at least once for COVID-19 (PCR and rapid antigen tests) before the 31st of August 2021, as recorded in the COVID-19 Healthdata test database (n=5,661,661). Thanks to an individual linkage with DEMOBEL, SD and SE data were available for 97.6% (n=5,525,634/5,661,661) of them. After this linkage, we excluded individuals who were deregistered, migrated, and deceased based on their status in the national register available in DEMOBEL database as well as individuals with either age, sex, or region unknown. Our final study population was composed of 5,342,110 adults (≥ 18 years) tested at least once in Belgium before August 31st 2021, for whom Vaccinnet+ was consulted in order to determine their vaccination status for a first COVID-19 vaccine dose. A flow chart demonstrating the process for selecting our final study population is available in **Figure 21**.

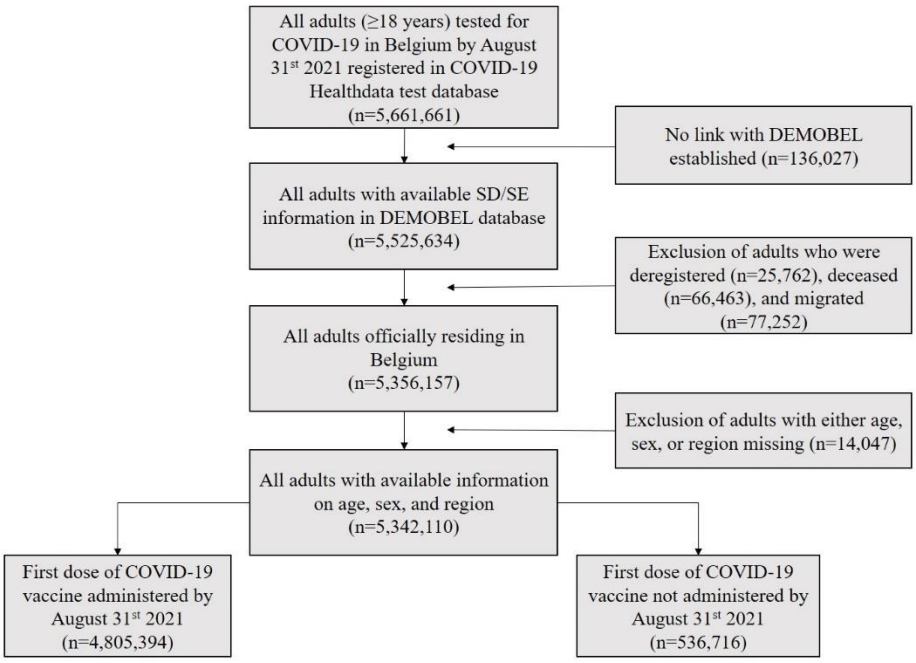



***Figure 21. Flowchart of the study population based on the COVID-19 Healthdata test database, Belgium, December 28th 2020 – August 31st 2021.***

## 2.3. Outcome

The outcome of this study was defined as the uptake of a first dose of any COVID-19 vaccine (approved by the EU) between the start of the vaccination campaign (December 28th 2020) and August 31st 2021. By this date, all individuals of 18 years and over officially residing in Belgium should have received an invitation to be vaccinated with a first dose and the opportunity to receive it. **Figure 22** shows COVID-19 first-dose vaccination coverage over time by region and age group in Belgium. Moreover, administration of a first dose is a good predictor of a full

primary course of vaccination. The proportion of those with a full primary course among those who received at least one dose is close to 100% [25].

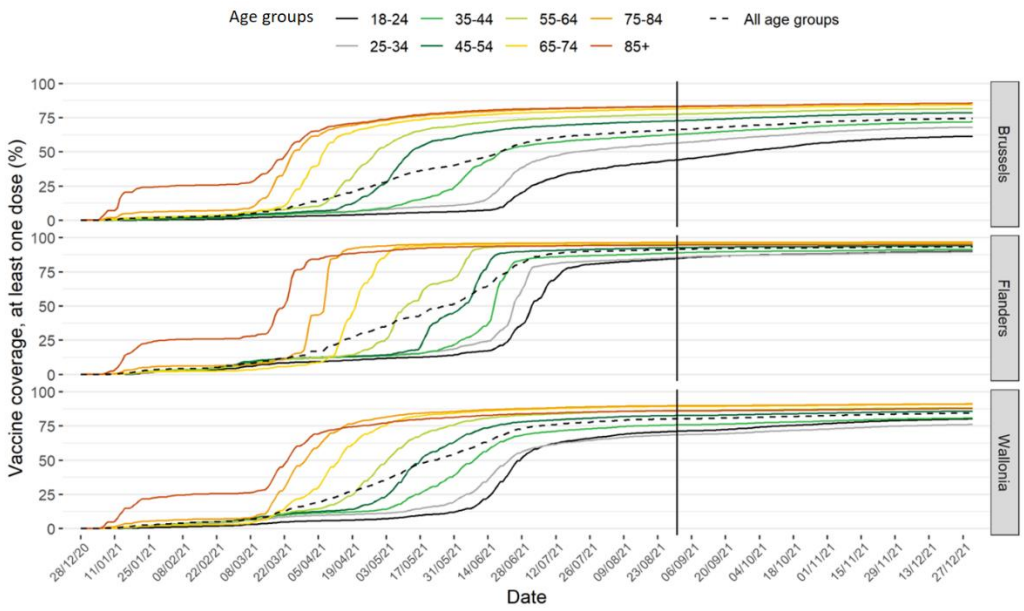



***Figure 22. Vaccination coverage of a first dose of COVID-19 vaccine over time by age groups and regions, Belgium, December 28th 2020 – December 27th 2021.***

## 2.4. Covariates

To identify characteristics associated with the uptake of a first COVID-19 vaccine dose, we focused on several relevant SD and SE characteristics (**Supplementary Table 1**). Age, sex, and region of residence were available in Vaccinnet+ and COVID-19 Healthdata test database. Household type, migration background, income, education, and employment status were available in DEMOBEL. Household type was divided in one-person, couples without children, couples with children, single parents, collectivity (e.g., prison, nursing homes), and other. Migration background

was divided in Belgian natives, second-generation migrants, first generation European migrants, and first-generation non-European migrants. Income was available as deciles of the net household income and was further categorized into low income (deciles 1 to 4), moderate income (deciles 5 to 7), and high income (deciles 8 to 10). Education was classified in eight categories using the International Standard Classification of Education (ISCED). We merged these different categories into three main education levels: low (ISCED0 to ISCED2), moderate (ISCED3 to ISCED4), and high (ISCED5 to ISCED 8). The employment status was defined as employed or unemployed. Finally, we were able to determine whether individuals had a healthcare degree using a variable provided by CoBRHA. Missing SD or SE data were considered as a separate category as they were associated with certain SD and SE characteristics and not randomly distributed in the population (see **Supplementary Table 2**).

### 2.5. Statistical analyses

For each SD and SE level, the number and proportion of first COVID-19 vaccine doses (not) administered were calculated. To compute adjusted odds ratios (aORs) and Wald 95% confidence intervals (CIs), we fitted a first logistic regression model including age, sex, region, migration background, household type, and income as covariates. A second model was fitted only on individuals aged 25-years and older adding education level and healthcare degree as additional covariates. Finally, a

third model was fitted only on individuals aged between 25- and 65-years adding employment status as additional covariate and removing income (highly correlated variables). Following the same procedures, we fitted logistic regression models stratified by regions (Flanders, Brussels, Wallonia) based on the individual postal code of residence, in view of the important differences in vaccination coverage between them (see **Table 14** and **Figure 22**), and the responsibility of each regional health authority for the roll-out and practical implementation of the vaccination campaign [15].

We performed sensitivity analyses to determine whether adding interaction terms between region and each covariate of the model, instead of stratifying by region, had an influence on the results. To do this, we extracted and compared, for both methods, the conditional probabilities of not having obtained a first COVID-19 vaccine dose according to all SD and SE characteristics (**Supplementary Table 3**).

All analyses were performed in R version 4.0.5 [221]. The package "Effects" was used to compute conditional probabilities [222].

## 3. Results

### 3.1. Descriptive statistics

The final analyses included 5,342,110 Belgian residents. By August 31st 2021, 89. 95% of them received a first COVID-19 vaccine dose (**Table 14**).

The vaccination coverage was lower among young age groups, with 84.88% vaccinated among 18-24-years-olds compared to 97.06% vaccinated among 85-years-olds and over. The vaccination coverage was similar between males and females. The lowest vaccination coverage was observed in Brussels (78.02% vaccinated), compared to Flanders (92.83% vaccinated) and Wallonia (88.03% vaccinated). Second-generation migrants, first-generation European migrants, and first-generation non-European migrants had a lower vaccination coverage (80.90%, 81.15%, 77.59% vaccinated, respectively), compared to Belgian native (94.79% vaccinated). A lower vaccination coverage was observed among one-person households (89.52% vaccinated), couples with children (89.08% vaccinated), and single parents (84.50% vaccinated), compared to collectivities (96.59% vaccinated) and couples without children (94.44% vaccinated).

The vaccination coverage was also lower among individuals with low education (89.15% vaccinated), low income (84. 55% vaccinated), and without a healthcare degree (89.59% vaccinated), compared to individuals with high education (94.13% vaccinated), high income (94.94% vaccinated), and a healthcare degree (89.59% vaccinated). The vaccination coverage was similar between employed and unemployed individuals. Overall, the vaccination coverage was relatively low in categories with data indicated as missing: missing household type (76.91%

vaccinated), missing education level (80.93% vaccinated), missing income (82.71% vaccinated), missing employment status (75.60% vaccinated).

**Table 14. Sociodemographic and socioeconomic characteristics of the population according to the uptake of a first dose of COVID-19 vaccine in Belgium, December 28th 2020 – August 31st 2021.**

| Variables | All (n=5,342,110 ) | First dose administered by August 31st 2021 (n=4,805,394) | First dose not administered by August 31st 2021 (n=536,716) |
|---|---|---|---|
| | *n (%column)* | *n (%row)* | *n (%row)* |
| Age groups (years) | | | |
| 18-24 | 645,416 (12·08) | 547,826 (84·88) | 97,590 (15·12) |
| 25-34 | 1,051,576 (19·68) | 886,846 (84·33) | 164,730 (15·67) |
| 35-44 | 978,478 (18·32) | 857,111 (87·60) | 121,367 (12·40) |
| 45-54 | 903,514 (16·91) | 827,999 (91·64) | 75,515 (8·36) |
| 55-64 | 797,786 (14·93) | 752,601 (94·34) | 45,185 (5·66) |
| 65-74 | 511,784 (9·58) | 492,092 (96·15) | 19,692 (3·85) |
| 75-84 | 296,974 (5·56) | 288,939 (97·29) | 8,035 (2·71) |
| 85+ | 156,582 (2·93) | 151,980 (97·06) | 4,602 (2·94) |
| Regions | | | |
| Brussels | 579,687 (10·85) | 452,251 (78·02) | 127,436 (21·98) |
| Flemish | 3,348,547 (62·68) | 3,108,510 (92·83) | 240,037 (7·17) |
| Walloon | 1,413,876 (26·47) | 1,244,633 (88·03) | 169,243 (11·97) |
| Sex | | | |
| Women | 2,849,932 (53·35) | 2,565,422 (90·02) | 284,510 (9·98) |
| Men | 2,492,178 (46·65) | 2,239,972 (89·88) | 252,206 (10·12) |
| Migration background | | | |
| Belgian natives | 3,652,357 (68·37) | 3,461,899 (94·79) | 190,458 (5·21) |
| Second-generation migrants | 324,420 (6·07) | 262,465 (80·90) | 61,955 (19·10) |
| First-generation European migrants | 607,586 (11·37) | 493,062 (81·15) | 114,524 (18·85) |
| First-generation non-European migrants | 757,747 (14·18) | 587,968 (77·59) | 169,779 (22·41) |

**Table 14. Continued**

| Variables | All (n=5,342,110 ) | First dose administered by August 31st 2021 (n=4,805,394) | First dose not administered by August 31st 2021 (n=536,716) |
|---|---|---|---|
| | ***n (%column)*** | ***n (%row)*** | ***n (%row)*** |
| Household type | | | |
| One-person households | 878,506 (16·44) | 786,400 (89·52) | 92,106 (10·48) |
| Collectivity | 96,213 (1·80) | 92,929 (96·59) | 3,284 (3·41) |
| Couples without children | 1,302,576 (24·38) | 1,230,134 (94·44) | 72,442 (5·56) |
| Couples with children | 2,363,467 (44·24) | 2,105,396 (89·08) | 258,071 (10·92) |
| Single parents | 548,369 (10·27) | 463,386 (84·50) | 84,983 (15·50) |
| Missing | 26,752 (0·50) | 20,574 (76·91) | 6,178 (23·09) |
| Other | 126,227 (2·36) | 106,575 (84·43) | 19,652 (15·57) |
| Education level | | | |
| Low | 1,567,744 (29·35) | 1,397,598 (89·15) | 170,146 (10·85) |
| Moderate | 1,569,361 (29·38) | 1,423,852 (90·73) | 145,509 (9·27) |
| High | 1,511,494 (28·29) | 1,422,694 (94·13) | 88,800 (5·87) |
| Missing | 693,511 (12·98) | 561,250 (80·93) | 132,261 (19·07) |
| Income | | | |
| Low | 1,815,339 (33·98) | 1,534,829 (84·55) | 280,510 (15·45) |
| Moderate | 1,564,464 (29·29) | 1,431,971 (91·53) | 132,493 (8·47) |
| High | 1,762,607 (32·99) | 1,673,416 (94·94) | 89,191 (5·06) |
| Missing | 199,700 (3·74) | 165,178 (82·71) | 34,522 (17·29) |
| Healthcare degree | | | |
| Yes | 423,852 (7·93) | 398,902 (89·59) | 24,950 (10·41) |
| No | 4,918,258 (92·07) | 4,406,492 (94·11) | 511,766 (5·89) |
| Employment status | | | |
| Unemployed | 1,874,544 (35·09) | 1,695,122 (90·43) | 179,422 (9·57) |
| Employed | 3,364,795 (62·99) | 3,032,575 (90·13) | 332,220 (9·87) |
| Missing | 102,771 (1·92) | 77,697 (75·60) | 25,074 (24·40) |

*n(%column): absolute number with percentage of the total population; n(%row) : absolute number with percentage of the subgroup per sociodemographic and socioeconomic characteristics.*

### 3.2. Predictors of first dose of COVID-19 vaccine uptake

Multivariable Model 1 (**Table 15**), applied on all age groups, showed an age gradient in COVID-19 vaccination coverage from 25- to 84-years-olds: the younger the individuals were, the more likely they were to be unvaccinated (aOR 5.49 [5.37-5.62] for 25-34 age group, compared to 75-84 age group). Men had a slightly lower COVID-19 vaccine uptake, compared to women (aOR 1.05 [1.04-1.06]). Individuals with a migration background had higher odds of being unvaccinated, compared to Belgian natives (aOR 2.25 [2.23-2.28] for second-generation migrants; aOR 2.90 [2.88-2.93] for first-generation European migrants; aOR 2.98 [2.95-3.00] for first-generation non-European migrants). Compared to couples with children, one-person households (aOR 1.18 [1.17-1.19) and single parents (aOR 1.27 [1.26-1.29]) had higher odds of being unvaccinated. Individuals with missing information on the household type had also higher odds of being unvaccinated (aOR 1.04 [1.01-1.08]). A lower COVID-19 vaccine uptake was identified among individuals with low (aOR 2.36 [2.34-2.38]), moderate (aOR 1.54 [1.52-1.55]), or missing income (aOR 1.93 [1.89-1.96]), compared to individuals with high income.

In Model 2 (**Table 15**), a lower COVID-19 vaccine uptake was identified among individuals with a low (aOR 1.37 [1.36-1.39]), moderate (aOR 1.31 [1.30-1.32]), or missing (aOR 1.19 [1.18-1.21]) education level, compared to individuals with a high

education level. Not having a healthcare degree was associated with a lower COVID-19 vaccine uptake (aOR 1.41 [1.39-1.43]).

In Model 3 (**Table 15**), being unemployed was associated with a higher odd of being unvaccinated (aOR 1.46 [1.45-1.47]) as well as having a missing employment status (aOR 1.16 [1.13-1.18]), compared to being employed. Overall, similar patterns could be identified in all regions (**Supplementary Tables 4** to **6**). The crude ORs for the Belgian models and those stratified by regions can be found in **Supplementary Tables 7** and **Supplementary Tables 8** to **10**, respectively.

**Table 15. Adjusted odds ratios (aORs) and their 95% confidence intervals (CIs) for the association between sociodemographic and socioeconomic characteristics and the odds of not having received a first dose of COVID-19 vaccine, Belgium, December 28th 2020 - August 31st 2021.**

| Variables | Model 1 [a] | Model 2 [b] | Model 3 [c] |
|---|---|---|---|
| | *aOR (95%CI)* | *aOR (95%CI)* | *aOR (95%CI)* |
| Age groups | | | |
| 18-24 | 4.78 (4.66-4.89) | - | - |
| 25-34 | 5.49 (5.37-5.62) | 5.98 (5.84-6.12) | 2.75 (2.71-2.78) |
| 35-44 | 3.89 (3.80-3.99) | 4.22 (4.12-4.32) | 1.97 (1.94-1.99) |
| 45-54 | 2.67 (2.60-2.73) | 2.83 (2.76-2.90) | 1.34 (1.32-1.36) |
| 55-64 | 2.15 (2.10-2.21) | 2.23 (2.18-2.29) | 1.00 |
| 65-74 | 1.43 (1.40-1.47) | 1.47 (1.43-1.51) | - |
| 75-84 | 1.00 | 1.00 | - |
| 85+ | 1.19 (1.14-1.23) | 1.15 (1.11-1.20) | - |
| Sex | | | |
| Women | 1.00 | 1.00 | 1.00 |
| Men | 1.05 (1.04-1.06) | 0.99 (0.98-1.00) | 1.07 (1.06-1.07) |

**Table 15. Continued**

| Variables | Model 1 [a] | Model 2 [b] | Model 3 [c] |
|---|---|---|---|
| | ***aOR (95%CI)*** | ***aOR (95%CI)*** | ***aOR (95%CI)*** |
| Regions | | | |
| Flemish | 1.00 | 1.00 | 1.00 |
| Brussels | 1.80 (1.78-1.81) | 1.76 (1.74-1.78) | 1.77 (1.75-1.78) |
| Walloon | 1.63 (1.62-1.64) | 1.67 (1.66-1.68) | 1.67 (1.66-1.68) |
| Migration background | | | |
| Belgian natives | 1.00 | 1.00 | 1.00 |
| Second-generation migrants | 2.25 (2.23-2.28) | 1.94 (1.91-1.96) | 2.12 (2.09-2.15) |
| First-generation European migrants | 2.90 (2.88-2.93) | 2.64 (2.62-2.67) | 3.10 (3.07-3.13) |
| First-generation non-European migrants | 2.98 (2.95-3.00) | 2.76 (2.73-2.78) | 3.45 (3.42-3.48) |
| Household type | | | |
| Couples with children | 1.00 | 1.00 | 1.00 |
| Couples without children | 0.88 (0.87-0.89) | 0.84 (0.83-0.84) | 0.83 (0.82-0.84) |
| Single parents | 1.27 (1.26-1.29) | 1.32 (1.30-1.33) | 1.49 (1.47-1.51) |
| One-person households | 1.18 (1.17-1.19) | 1.17 (1.16-1.18) | 1.26 (1.25-1.27) |
| Collectivity | 0.63 (0.60-0.65) | 0.61 (0.58-0.63) | 0.53 (0.50-0.56) |
| Other | 1.23 (1.21-1.25) | 1.24 (1.22-1.26) | 1.21 (1.19-1.23) |
| Missing | 1.04 (1.01-1.08) | 1.04 (1.00-1.08) | 1.16 (1.13-1.18) |
| Income | | | |
| High | 1.00 | 1.00 | - |
| Moderate | 1.54 (1.52-1.55) | 1.44 (1.43-1.46) | - |
| Low | 2.36 (2.34-2.38) | 2.08 (2.06-2.10) | - |
| Missing | 1.93 (1.89-1.96) | 1.68 (1.65-1.71) | - |
| Education level | | | |
| High | - | 1.00 | - |
| Moderate | - | 1.31 (1.30-1.32) | |
| Low | - | 1.37 (1.36-1.39) | |
| Missing | - | 1.19 (1.18-1.21) | |
| Healthcare degree | | | |
| Yes | - | 1.00 | - |
| No | - | 1.41 (1.39-1.43) | |

**Table 15. Continued**

| Variables | Model 1 [a] | Model 2 [b] | Model 3 [c] |
|---|---|---|---|
| | ***aOR (95%CI)*** | ***aOR (95%CI)*** | ***aOR (95%CI)*** |
| Employment status | | | |
| Employed | - | - | 1.00 |
| Unemployed | - | | 1.46 (1.45-1.47) |
| Missing | - | | 1.16 (1.13-1.18) |

*[a] Logistic regression model applied on all age groups (n=5,342,110). OR are adjusted for age, sex, region, migration background, household type, and income.*

*[b] Logistic regression model applied only to individuals aged over 25 years (n=4,696,694). OR are adjusted for age, sex, region, migration background, household type, income, education level, and healthcare degree.*

*[c] Logistic regression model applied only to individuals aged 25 to 65 years (n=3,792,100). OR are adjusted for age, sex, region, migration background, household type, and employment status. Income being strongly correlated with employment status, OR are not adjusted for income when employment status is included in the model.*

## 4. Discussion

Although only 34% of the Belgian population stated that they were definitively ready to be vaccinated against COVID-19 before the start of the vaccination campaign, Belgium achieved the 7$^{th}$ highest vaccination coverage for the primary vaccination course in the European Union (with 89% of individuals over 18 years old having completed their primary course on 8 April 2022) [44,45,47]. Nevertheless, despite this high rate and free vaccination, important SD and SE disparities in COVID-19 vaccination coverage remained. Indeed, our findings showed a lower COVID-19 vaccination coverage among young individuals, men, migrants, single parents, one-person households, and disadvantaged SE groups

(low income and education, unemployment). Similar patterns were observed for each Belgian region.

A Swedish study also found that COVID-19 vaccination coverage was lower among young individuals, with low income, living alone, and born outside Sweden [215]. A Danish study identified a lower COVID-19 vaccination coverage among men, individuals living in more deprived areas or urban areas, and ethnic groups other than white [223]. In the United States, Williams et al. found a lower COVID-19 vaccine uptake among young people, SE disadvantaged groups, and all racial and ethnic minority groups (except for Asians). They highlighted that age and SE factors (such as health insurance, income, education, employment) accounted for a large proportion of the social disparities and are therefore important key factors [216]. Our findings also partially confirm those of Barry et al. who identified a lower COVID-19 vaccination coverage among single parents and individuals living in counties with lower SE status and with a high percentage of households with children [224]. In the United States, Farah et al. found a higher initial vaccination coverage among healthcare workers compared to the general population national average [225], similar to our results. A cohort study conducted by Azamgarhi et al. during a time of high community COVID-19 prevalence in the United Kingdom also reported high early vaccination rates among healthcare workers [226].

Prior to the implementation of COVID-19 vaccination, numerous studies examined SD and SE disparities in COVID-19 vaccine hesitancy and identified a higher vaccine hesitancy among female, parents, ethnic minority, and disadvantaged SE groups (i.e., with lower levels of education and income, unemployment, and poor knowledge of COVID-19) [46,227–229]. Our study partially supports these findings by showing lower vaccination coverage among migrants, single parents, and SE disadvantaged groups. Although there may have been concerns about vaccination among healthcare workers, Wang et al. [230] shown that they were less hesitant to be vaccinated, compared to non-healthcare workers, which is consistent with our findings.

Vaccine hesitancy in migrants and SE disadvantaged groups may be underlined by several factors: the more severe direct and indirect impact that the crisis has had on them (e.g., higher rate of COVID-19 infection, subsequent negative health outcomes, and unemployment) may have increased distrust of governments, healthcare systems, and immunization [208,209,231,232]; decreased willingness to participate in public health measures as a result of decreased access to healthcare and resources [232]; and raised concerns and negative assumptions about vaccination due to a lack of health literacy and recognition of misinformation, potentially accentuated by the rapid development of the COVID-19 vaccine which has led to a higher level of mistrust of its benefits and concerns about its side effects

(the strongest predictors of COVID-19 vaccine uptake) [220,227,232–235]. Other factors that may explain the disparities identified in COVID-19 vaccine uptake include administrative hurdles which can partly explain the lower coverage in some SE groups. Indeed, our results, demonstrating lower vaccine coverage among individuals with missing SD or SE information, may be indicative of a hard-to-reach population in the context of a broad automated invitation process. Language barriers may also be an obstacle to accessing COVID-19 vaccination information, primarily available in the Belgian official languages.

To our knowledge, this paper is the first large representative study investigating, thanks to an individual data linkage established within the LINK-VACC project, SD and SE disparities in COVID-19 vaccination coverage in Belgium and all regions with up-to-date SD and SE information.

Our study has several limitations. First, individuals never tested for COVID-19 before August 31st 2021 are not included in our study population. Indeed, because of the General Data Protection Regulation, individual's SD/SE information from the DEMOBEL database could not be obtained from the total Belgian population, the master database used for the linkage consists of the COVID-19 Healthdata test database obtained on the 31st of August 2021. This resulted in a slight over-representation of vaccinated individuals in our study population compared to the actual vaccination coverage in Belgium (90.0% vs 85.4%, respectively). However,

despite this slight over-representation of vaccinated individuals in our study population, it should be pointed out that the trends in vaccination coverage per region, age group, and sex are very similar to those of the overall Belgian adult population (see **Supplementary Figure 1**).

A second limitation is that our study sample does not include the unregistered population, a group of about 100,000-150,000 individuals, independent of vaccination status, as no information on this population was available in DEMOBEL. Third, Belgian residents vaccinated abroad (e.g., frontier workers) are not automatically registered in Vaccinnet+.

Future research should be focused on SD and SE disparities in COVID-19 vaccination coverage among adolescents and children and on factors influencing the uptake of the boosters. Another perspective would be to investigate the factors underlying the refusal to be vaccinated to determine whether these inequities are based on personal convictions or inequity in the distribution of the vaccine.

## 5. Conclusion

Despite the success of the vaccination campaign in Belgium (89% of adults vaccinated with primary course), free vaccination, and the efforts made by the regional health authorities to reach all citizens, important SD and SE inequalities in COVID-19 vaccine uptake were identified. Our study contributes to a better

identification of disparities in COVID-19 vaccination and helps to better target vaccination strategies to more vulnerable groups with the goal of acquiring the broadest possible vaccine coverage to limit the circulation of the virus and the development of severe negative health outcomes, whether in the context of the COVID-19 pandemic or other potential health threats.

## 6. Declarations

### 6.1. Acknowledgments

We are very grateful to Statistics Belgium for their essential advice on the recent socioeconomic data. We would also like to thank Healthdata and the Belgian regions.

### 6.2. Contributors

LiC reviewed the literature; LiC, JvL, BDV, LVdB, NS, RDB, LuC, and PH conceived the study; LiC, JvL, LVdB, RDB, LuC, and PH selected the population; LiC and PH reviewed all available data; LiC, JvL, BDV, LVdB, NS, RDB, and PH designed the statistical methodology; LiC conducted the statistical analyses; LiC, JvL, LVdB, RDB, and PH interpreted the findings; LiC and PH designed the figures; LiC wrote the first draft of the paper. All authors revised the text. LiC, JvL, MB, VS, LuC, and PH have directly access to the data of the present study. LiC, JvL, MB, VS, LuC, and PH have full access to all the data in the study. LiC and PH has verified the underlying data of the study.

All authors approved the final version of this manuscript and accepted responsibility for its submission for publication.

### 6.3. Funding

This study was supported by the Belgian Federal Authorities through funding for the LINK- VACC project on vaccine surveillance. In addition, Belgian Science Policy Office (BELSPO) funded this research within the BRAIN-be 2.0 framework supporting pillar 3 Federal societal challenges (grant number B2/202/P3/HELICON). The sponsors did not have a role in the conduction of the study, the writing of the manuscript, or in the decision to submit it for publication.

### 6.4. Competing interest

The authors have no competing interests to declare that are relevant to the content of this article.

### 6.5. Ethics approval

The protocol of the LINK-VACC project was approved by the Medical Ethics Committee from the Vrije Universiteit of Brussels on 3 February 2021 (B.U.N 1432020000371) and obtained authorization from the Information Security Committee (ISC) Social Security and Health (reference number: IVC/KSZG/21/034). The study protocol has been preregistered on ClinicalTrials.gov (ClinicalTrials.gov ID: NCT05373420).

As confirmed by sections 23 and 24 of the Guidelines 03/2020 on the processing of data concerning health for the purpose of scientific research in the context of the COVID-19 outbreak of the European Data Protection Board (V1.0 of 21 April 2020), this survey falls under Article 6 §1(e) and Article 9 §2(i) of the General Data Protection Regulation (GDPR). In appliance with these GDPR legal grounds of data processing, no informed consent had to be signed by the patients.

### 6.6. Patient consent for publication

Not applicable

### 6.7. Data availability statement

The data used in this study are pseudonymized individual data obtained from several national registers hosted by Healthdata.be, an independent institution within Sciensano aiming to bring together all the data currently stored in multiple health registers on a single web-based platform. Due to the General Data Protection Regulation (GDPR) legislations in Belgium, these data are not publicly available. Data request must be addressed to the Information Security Committee (ISC). The code used in these analyses can be provided upon request.

## 7. Supplementary materials

### 7.1. Table of content

#### *7.1.1. Supplementary Figures*

- **Supplementary Figure 1.** Vaccination coverage of a first dose of COVID-19 vaccine according to region, age group, and sex in the 2021 Belgian adult population and in our study population including only adults tested at least once in Belgium by August 31st 2021, Belgium, December 28th 2020 - August 31st 2021.

#### *7.1.2. Supplementary Tables*

- **Supplementary Table 1.** Data holder and time frame for all sociodemographic and socioeconomic variables included as covariates in the logistic regression models, Belgium, December 28th 2020 - August 31st 2021.
- **Supplementary Table 2.** Characterization of the population with at least one missing information regarding household type, education level, income, or employment status relative to the total population by each sociodemographic and socioeconomic characteristic.
- **Supplementary Table 3.** Conditional probabilities of not having receive a first dose of COVID-19 vaccine obtained by logistic regression models stratified by region and logistic regression models including region as an interaction term

between each sociodemographic and socioeconomic characteristic, Belgium, December 28th 2020 - August 31st 2021.

- **Supplementary Table 4.** Adjusted odds ratios (aORs) and their 95% confidence intervals (CIs) for the association between sociodemographic and socioeconomic characteristics and the odds of not having received a first dose of COVID-19 vaccine in Flanders, Belgium, December 28th 2020 - August 31st 2021.

- **Supplementary Table 5.** Adjusted odds ratios (aORs) and their 95% confidence intervals (CIs) for the association between sociodemographic and socioeconomic characteristics and the odds of not having received a first dose of COVID-19 vaccine in Brussels, Belgium, December 28th 2020 - August 31st 2021.

- **Supplementary Table 6.** Adjusted odds ratios (aORs) and their 95% confidence intervals (CIs) for the association between sociodemographic and socioeconomic characteristics and the odds of not having received a first dose of COVID-19 vaccine in Wallonia, Belgium, December 28th 2020 - August 31st 2021.

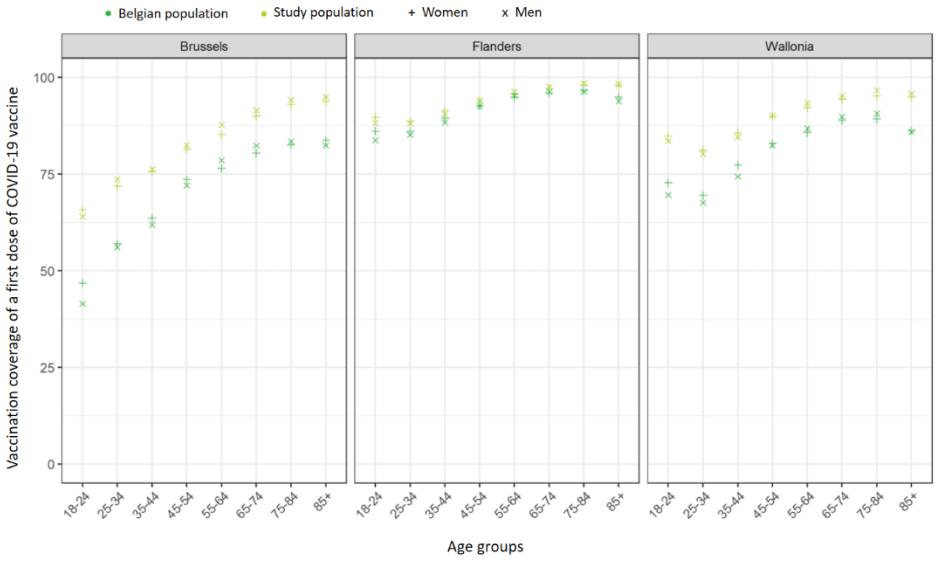
- **Supplementary Table 7.** Crude odds ratios (cORs) and their 95% confidence intervals (CIs) for the association between sociodemographic and

socioeconomic characteristics and the odds of not having received a first dose of COVID-19 vaccine, Belgium, December 28th 2020 - August 31st 2021.

- **Supplementary Table 8.** Crude odds ratios (cORs) and their 95% confidence intervals (CIs) for the association between sociodemographic and socioeconomic characteristics and the odds of not having received a first dose of COVID-19 vaccine stratified in Flanders, Belgium, December 28th 2020 - August 31st 2021.

- **Supplementary Table 9.** Crude odds ratios (cORs) and their 95% confidence intervals (CIs) for the association between sociodemographic and socioeconomic characteristics and the odds of not having received a first dose of COVID-19 vaccine stratified in Brussels, Belgium, December 28th 2020 - August 31st 2021.

- **Supplementary Table 10.** Crude odds ratios (cORs) and their 95% confidence intervals (CIs) for the association between sociodemographic and socioeconomic characteristics and the odds of not having received a first dose of COVID-19 vaccine stratified in Wallonia, Belgium, December 28th 2020 - August 31st 2021.




***Supplementary Figure 1. Vaccination coverage of a first dose of COVID-19 vaccine according to region, age group, and sex in the 2021 Belgian adult population and in our study population including only adults tested at least once in Belgium by August 31st 2021, Belgium, December 28th 2020 - August 31st 2021.***

**Supplementary Table 1. Data holder and time frame for all sociodemographic and socioeconomic variables included as covariates in the logistic regression models, Belgium, December 28th 2020 - August 31st 2021.**

| Variables | Data holder | Time frame |
|---|---|---|
| Age | National register | March 2022 |
| Sex | National register | March 2022 |
| Region | National register | March 2022 |
| Migration background | Statistics Belgium | August 2021 |
| Household type | Statistics Belgium | August 2021 |
| Income | Statistics Belgium | 2018 (fiscal year 2019) |
| Education level | Statistics Belgium | Census 2017 |
| Employment status | Statistics Belgium | Year 2019 |
| Healthcare degree | Common Base Registry for HealthCare Actor | January 2021 |

**Supplementary Table 2. Characterization of the population with at least one missing information regarding household type, education level, income, or employment status relative to the total population by each sociodemographic and socioeconomic characteristic.**

| Variables | Total (n=5,342,110) | At least one missing information (n=764,857) | At least one missing information to the total population |
|---|---|---|---|
| | n | n | $\%_{row}$ |
| Age groups (years) | | | |
| 18-24 | 645,416 | 55,991 | 8·66 |
| 25-34 | 1,051,576 | 182,302 | 17·34 |
| 35-44 | 978,478 | 192,936 | 19·72 |
| 45-54 | 903,514 | 133,289 | 14·75 |
| 55-64 | 797,786 | 76,903 | 9·64 |
| 65-74 | 511,784 | 46,103 | 9·01 |
| 75-84 | 296,974 | 37,353 | 12·58 |
| 85+ | 156,582 | 39,980 | 25·53 |
| Regions | | | |
| Brussels | 579,687 | 205,993 | 35·53 |
| Flemish | 3,348,547 | 392,471 | 11·72 |
| Walloon | 1,413,876 | 166,393 | 11·77 |
| Sex | | | |
| Women | 2,849,932 | 401,457 | 14·09 |
| Men | 2,492,178 | 363,400 | 14·58 |
| Migration background | | | |
| Belgian natives | 3,652,357 | 181,359 | 4·97 |
| Second-generation migrants | 324,420 | 13,868 | 4·27 |
| First-generation European migrants | 607,586 | 265,626 | 43·72 |
| First-generation non-European migrants | 757,747 | 304,004 | 40·12 |

**Supplementary Table 3. Conditional probabilities of not having receive a first dose of COVID-19 vaccine obtained by logistic regression models stratified by region and logistic regression models including region as an interaction term between each sociodemographic and socioeconomic characteristic, Belgium, December 28th 2020 - August 31st 2021.**

| | Flanders | | Brussels | | Wallonia | |
|---|---|---|---|---|---|---|
| **Variables** | **Probabilities (IC95%)** | | **Probabilities (IC95%)** | | **Probabilities (IC95%)** | |
| | *Stratification* | *Interaction* | *Stratification* | *Interaction* | *Stratification* | *Interaction* |
| Age (years) [a] | | | | | | |
| 18-24 | 0·076 (0·075-0·077) | 0·086 (0·085-0·087) | 0·301 (0·298-0·305) | 0·197 (0·194-0·201) | 0·133 (0·132-0·135) | 0·133 (0·131-0·135) |
| 25-34 | 0·089 (0·088-0·090) | 0·010 (0·099-0·101) | 0·262 (0·260-0·264) | 0·169 (0·167-0·171) | 0·176 (0·175-0·178) | 0·176 (0·175-0·178) |
| 35-44 | 0·065 (0·064-0·065) | 0·073 (0·072-0·074) | 0·203 (0·201-0·206) | 0·129 (0·127-0·131) | 0·129 (0·128-0·131) | 0·129 (0·128-0·130) |
| 45-54 | 0·046 (0·045-0·047) | 0·052 (0·051-0·052) | 0·150 (0·148-0·152) | 0·093 (0·091-0·095) | 0·089 (0·088-0·090) | 0·089 (0·087-0·090) |
| 55-64 | 0·037 (0·036-0·038) | 0·041 (0·041-0·042) | 0·130 (0·127-0·132) | 0·079 (0·077-0·081) | 0·071 (0·070-0·072) | 0·071 (0·070-0·072) |
| 65-74 | 0·023 (0·022-0·023) | 0·026 (0·025-0·027) | 0·101 (0·098-0·104) | 0·061 (0·059-0·062) | 0·049 (0·048-0·051) | 0·050 (0·048-0·051) |
| 75-84 | 0·015 (0·014-0·015) | 0·017 (0·016-0·018) | 0·078 (0·075-0·083) | 0·047 (0·044-0·049) | 0·038 (0·036-0·039) | 0·038 (0·036-0·039) |
| 85+ | 0·017 (0·016-0·018) | 0·020 (0·019-0·020) | 0·087 (0·081-0·093) | 0·052 (0·048-0·056) | 0·046 (0·044-0·049) | 0·046 (0·044-0·049) |

**Supplementary Table 3. Continued**

| | Flanders | | Brussels | | Wallonia | |
|---|---|---|---|---|---|---|
| **Variables** | **Probabilities (IC95%)** | | **Probabilities (IC95%)** | | **Probabilities (IC95%)** | |
| | *Stratification* | *Interaction* | *Stratification* | *Interaction* | *Stratification* | *Interaction* |
| Sex [a] | | | | | | |
| Women | 0·047 | 0·053 | 0·190 | 0·110 | 0·096 | 0·097 |
| | (0·046-0·047) | (0·054-0·057) | (0·188-0·191) | (0·108-0·111) | (0·096-0·097) | (0·096-0·102) |
| Men | 0·050 | 0·057 | 0·188 | 0·109 | 0·101 | 0·101 |
| | (0·049-0·050) | (0·057-0·059) | (0·186-0·189) | (0·107-0·110) | (0·100-0·102) | (0·101-0·102) |
| Migration background [a] | | | | | | |
| Belgian natives | 0·036 | 0·038 | 0·106 | 0·084 | 0·080 | 0·080 |
| | (0·036-0·037) | (0·038-0·038) | (0·104-0·107) | (0·082-0·085) | (0·079-0·080) | (0·079-0·080) |
| Second generation | 0·092 | 0·096 | 0·218 | 0·178 | 0·127 | 0·128 |
| migrants | (0·090-0·093) | (0·095-0·097) | (0·215-0·222) | (0·175-0·182) | (0·126-0·129) | (0·126-0·129) |
| First generation European | 0·141 | 0·147 | 0·208 | 0·169 | 0·150 | 0·150 |
| migrants | (0·140-0·143) | (0·146-0·148) | (0·206-0·210) | (0·167-0·171) | (0·149-0·152) | (0·149-0·152) |
| First generation non- | 0·109 | 0·114 | 0·253 | 0·208 | 0·171 | 0·171 |
| European migrants | (0·107-0·110) | (0·113-0·115) | (0·251-0·255) | (0·206-0·210) | (0·169-0·173) | (0·169-0·173) |
| Household type [a] | | | | | | |
| Couples with children | 0·046 | 0·053 | 0·202 | 0·119 | 0·094 | 0·094 |
| | (0·046-0·047) | (0·052-0·053) | (0·200-0·204) | (0·118-0·121) | (0·093-0·095) | (0·094-0·095) |
| Couples | 0·043 | 0·049 | 0·153 | 0·088 | 0·089 | 0·090 |
| | (0·043-0·044) | (0·049-0·050) | (0·151-0·156) | (0·087-0·090) | (0·088-0·091) | (0·089-0·091) |

**Supplementary Table 3. Continued**

| | Flanders | | Brussels | | Wallonia | |
|---|---|---|---|---|---|---|
| **Variables** | **Probabilities (IC95%)** | | **Probabilities (IC95%)** | | **Probabilities (IC95%)** | |
| | *Stratification* | *Interaction* | *Stratification* | *Interaction* | *Stratification* | *Interaction* |
| Single parents | 0·061 | 0·070 | 0·216 | 0·129 | 0·116 | 0·117 |
| | (0·061-0·062) | (0·070-0·071) | (0·213-0·219) | (0·126-0·131) | (0·115-0·118) | (0·116-0·119) |
| One person households | 0·059 | 0·067 | 0·186 | 0·109 | 0·119 | 0·119 |
| | (0·058-0·060) | (0·067-0·068) | (0·184-0·188) | (0·107-0·112) | (0·117-0·120) | (0·118-0·121) |
| Collectivity | 0·031 | 0·035 | 0·142 | 0·083 | 0·055 | 0·055 |
| | (0·030-0·033) | (0·033-0·037) | (0·131-0·153) | (0·076-0·090) | (0·052-0·058) | (0·052-0·058) |
| Other | 0·064 | 0·073 ( | 0·178 | 0·104 | 0·125 | 0·125 |
| | (0·063-0·066) | 0·071-0·074) | (0·173-0·182) | (0·102-0·108) | (0·122-0·129) | (0·121-0·129) |
| Missing | 0·046 | 0·057 | 0·181 | 0·113 | 0·076 | 0·099 |
| | (0·044-0·048) | (0·054-0·059) | (0·173-0·190) | (0·107-0·119) | (0·071-0·081) | (0·093-0·105) |
| Income [a] | | | | | | |
| High | 0·032 | 0·036 | 0·109 | 0·068 | 0·068 | 0·069 |
| | (0·031-0·032) | (0·036-0·036) | (0·107-0·111) | (0·067-0·070) | (0·068-0·069) | (0·069-0·070) |
| Moderate | 0·048 | 0·054 | 0·182 | 0·119 | 0·095 | 0·096 |
| | (0·048-0·049) | (0·054-0·055) | (0·180-0·185) | (0·117-0·121) | (0·094-0·096) | (0·095-0·097) |
| Low | 0·075 | 0·083 | 0·234 | 0·155 | 0·137 | 0·138 |
| | (0·074-0·075) | (0·083-0·084) | (0·232-0·235) | (0·153-0·157) | (0·136-0·138) | (0·137-0·139) |
| Missing | 0·068 | 0·075 | 0·183 | 0·119 | 0·122 | 0·122 |
| | (0·066-0·070) | (0·073-0·076) | (0·179-0·186) | (0·116-0·122) | (0·118-0·126) | (0·119-0·126) |

**Supplementary Table 3. Continued**

| | Flanders | | Brussels | | Wallonia | |
|---|---|---|---|---|---|---|
| **Variables** | **Probabilities (IC95%)** | | **Probabilities (IC95%)** | | **Probabilities (IC95%)** | |
| | *Stratification* | *Interaction* | *Stratification* | *Interaction* | *Stratification* | *Interaction* |
| **Population aged 25 years and over (n=4,696,694)** | | | | | | |
| Education level [b] | | | | | | |
| High | 0·038 | 0·043 | 0·144 | 0·082 | 0·081 | 0·081 |
| | (0·037-0·038) | (0·042-0·043) | (0·141-0·146) | (0·081-0·084) | (0·080-0·082) | (0·081-0·083) |
| Moderate | 0·049 | 0·055 | 0·213 | 0·127 | 0·096 | 0·098 |
| | (0·048-0·049) | (0·055-0·056) | (0·211-0·216) | (0·124-0·129) | (0·095-0·097) | (0·096-0·098) |
| Low | 0·050 | 0·057 | 0·210 | 0·125 | 0·103 | 0·104 |
| | (0·049-0·051) | (0·056-0·057) | (0·208-0·213) | (0·123-0·128) | (0·102-0·104) | (0·102-0·105) |
| Missing | 0·049 | 0·055 | 0·153 | 0·088 | 0·095 | 0·096 |
| | (0·048-0·050) | (0·055-0·056) | (0·151-0·155) | (0·087-0·090) | (0·094-0·097) | (0·094-0·097) |
| Healthcare degree [b] | | | | | | |
| Yes | 0·034 | 0·039 | 0·124 | 0·076 | 0·071 | 0·071 |
| | (0·033-0·035) | (0·038-0·039) | (0·120-0·128) | (0·073-0·078) | (0·069-0·072) | (0·069-0·072) |
| No | 0·046 | 0·053 | 0·174 | 0·108 | 0·095 | 0·095 |
| | (0·046-0·046) | (0·052-0·053) | (0·172-0·175) | (0·107-0·109) | (0·095-0·096) | (0·095-0·096) |
| **Population aged between 25 and 65 years (n=3,792,100)** | | | | | | |
| Employment status [c] | | | | | | |
| Employed | 0·055 | 0·062 | 0·193 | 0·122 | 0·109 | 0·111 |
| | (0·055-0·056) | (0·062-0·063) | (0·191-0·194) | (0·120-0·123) | (0·108-0·109) | (0·110-0·111) |

**Supplementary Table 3. Continued**

| | Flanders | | Brussels | | Wallonia | |
|---|---|---|---|---|---|---|
| **Variables** | **Probabilities (IC95%)** | | **Probabilities (IC95%)** | | **Probabilities (IC95%)** | |
| | *Stratification* | *Interaction* | *Stratification* | *Interaction* | *Stratification* | *Interaction* |
| Missing | 0·074 | 0·082 | 0·184 | 0·116 | 0·134 | 0·136 |
| | (0·072-0·077) | (0·080-0·084) | (0·179-0·189) | (0·112-0·120) | (0·129-0·140) | (0·130-0·141) |
| **Population aged between 25 and 65 years (n=3,792,100)** | | | | | | |
| Employment status [c] | | | | | | |
| Employed | 0·055 | 0·062 | 0·193 | 0·122 | 0·109 | 0·111 |
| | (0·055-0·056) | (0·062-0·063) | (0·191-0·194) | (0·120-0·123) | (0·108-0·109) | (0·110-0·111) |
| Unemployed | 0·084 | 0·094 | 0·242 | 0·155 | 0·146 | 0·149 |
| | (0·083-0·085) | (0·093-0·095) | (0·239-0·244) | (0·153-0·158) | (0·144-0·147) | (0·147-0·150) |
| Missing | 0·074 | 0·082 | 0·184 | 0·116 | 0·134 | 0·136 |
| | (0·072-0·077) | (0·080-0·084) | (0·179-0·189) | (0·112-0·120) | (0·129-0·140) | (0·130-0·141) |

*[a] Logistic regression model applied on all age groups (Flanders, n=3,348,547; Brussel, n=579,687; Wallonia, n=1,413,876). Probabilities are adjusted for age, sex, region, migration background, household type, and income;*

*[b] Logistic regression model applied only to individuals aged over 25 years (Flanders, n=2,955,807; Brussel, n=505,082; Wallonia, n=1,235,805). Probabilities are adjusted for age, sex, region, migration background, household type, income, education level, and healthcare degree.;*

*[c] Logistic regression model applied only to individuals aged 25 to 65 years (Flanders, n=2,357,375; Brussel, n=436,356; Wallonia, n=998,369). Probabilities are adjusted for age, sex, region, migration background, household type, and employment status. Income being strongly correlated with employment status, probabilities are not adjusted for income when employment status is included in the model.*

**Supplementary Table 4. Adjusted odds ratios (aORs) and their 95% confidence intervals (CIs) for the association between sociodemographic and socioeconomic characteristics and the odds of not having received a first dose of COVID-19 vaccine in Flanders, Belgium, December 28th 2020 - August 31st 2021.**

| Variables | Model 1* | Model 2** | Model 3*** |
|---|---|---|---|
| | aOR (95%CI) | aOR (95%CI) | aOR (95%CI) |
| Age (years) | | | |
| 18-24 | 5·46 (5·26-5·66) | - | - |
| 25-34 | 6·46 (6·24-6·69) | 7·02 (6·77-7·27) | 2·79 (2·74-2·83) |
| 35-44 | 4·57 (4·41-4·74) | 4·91 (4·73-5·09) | 2·02 (1·98-2·05) |
| 45-54 | 3·19 (3·07-3·30) | 3·35 (3·23-3·47) | 1·39 (1·36-1·41) |
| 55-64 | 2·51 (2·43-2·61) | 2·60 (2·51-2·70) | 1·00 |
| 65-74 | 1·55 (1·49-1·61) | 1·59 (1·53-1·66) | - |
| 75-84 | 1·00 | 1·00 | - |
| 85+ | 1·16 (1·10-1·23) | 1·13 (1·07-1·19) | - |
| Sex | | | |
| Women | 1·00 | 1·00 | 1·00 |
| Men | 1·07 (1·06-1·08) | 1·00 (0·99-1·01) | 1·08 (1·07-1·10) |
| Migration background | | | |
| Belgian natives | 1·00 | 1·00 | 1·00 |
| Second generation migrants | 2·68 (2·63-2·72) | 2·33 (2·28-2·38) | 2·55 (2·49-2·60) |
| First-generation European migrants | 4·35 (4·29-4·40) | 3·77 (3·71-3·82) | 4·76 (4·70-4·83) |
| First-generation non-European migrants | 3·23 (3·20-3·27) | 2·83 (2·79-2·87) | 3·75 (3·70-3·79) |
| Household type | | | |
| Couples with children | 1·00 | 1·00 | 1·00 |
| Couples without children | 0·93 (0·92-0·94) | 0·87 (0·86-0·89) | 0·87 (0·86-0·88) |
| Single parents | 1·35 (1·33-1·37) | 1·41 (1·39-1·44) | 1·60 (1·57-1·62) |
| One-person households | 1·29 (1·27-1·31) | 1·27 (1·25-1·29) | 1·39 (1·37-1·41) |
| Collectivity | 0·65 (0·62-0·69) | 0·64 (0·60-0·67) | 0·58 (0·54-0·63) |
| Other | 1·41 (1·37-1·44) | 1·42 (1·38-1·46) | 1·41 (1·37-1·45) |
| Missing | 1.08 (1.03-1.13) | 1·09 (1·03-1·15) | 1.03 (0.97-1.09) |

**Supplementary Table 4. Continued**

| Variables | Model 1* | Model 2** | Model 3*** |
|---|---|---|---|
| | **aOR (95%CI)** | **aOR (95%CI)** | **aOR (95%CI)** |
| Income level | | | |
| High | 1·00 | 1·00 | - |
| Moderate | 1·54 (1·52-1·56) | 1·45 (1·43-1·47) | - |
| Low | 2·44 (2·41-2·47) | 2·17 (2·14-2·20) | - |
| Missing | 2·17 (2·12-2·22) | 1·84 (1·79-1·90) | - |
| Education level | | | |
| High | - | 1·00 | - |
| Moderate | - | 1·31 (1·29-1·32) | - |
| Low | - | 1·35 (1·33-1·37) | - |
| Missing | - | 1·31 (1·29-1·33) | - |
| Healthcare degree | | | |
| Yes | - | 1·00 | - |
| No | - | 1·38 (1·35-1·41) | - |
| Employment status | | | |
| Employed | - | - | 1·00 |
| Unemployed | - | - | 1·56 (1·54-1·58) |
| Missing | | | 1·35 (1·31-1·39) |

* Logistic regression model applied on all age groups (n=3,348,547). OR are adjusted for age, sex, region, migration background, household type, and income.

** Logistic regression model applied only to individuals aged over 25 years (n=2,955,807) OR are adjusted for age, sex, region, migration background, household type, income, education level, and healthcare degree.

*** Logistic regression model applied only to individuals aged 25 to 65 years (n=2,357,375). OR are adjusted for age, sex, region, migration background, household type, and employment status. Income being strongly correlated with employment status, OR are not adjusted for income when employment status is included in the model.

**Supplementary Table 5. Adjusted odds ratios (aORs) and their 95% confidence intervals (CIs) for the association between sociodemographic and socioeconomic characteristics and the odds of not having received a first dose of COVID-19 vaccine in Brussels, Belgium, December 28th 2020 - August 31st 2021.**

| Variables | Model 1* | Model 2** | Model 3*** |
|---|---|---|---|
| | aOR (95%CI) | aOR (95%CI) | aOR (95%CI) |
| Age (years) | | | |
| 18-24 | 5·01 (4·73-5·32) | - | - |
| 25-34 | 4·15 (3·92-4·39) | 4·64 (4·38-4·91) | 2·44 (2·38-2·50) |
| 35-44 | 3·01 (2·85-3·19) | 3·48 (3·39-3·69) | 1·76 (1·72-1·81) |
| 45-54 | 2·08 (1·97-2·21) | 2·34 (2·21-2·48) | 1·24 (1·20-1·27) |
| 55-64 | 1·75 (1·65-1·86) | 1·88 (1·77-1·99) | 1·00 |
| 65-74 | 1·32 (1·24-1·41) | 1·36 (1·28-1·45) | - |
| 75-84 | 1·00 | 1·00 | - |
| 85+ | 1·12 (1·02-1·24) | 1·07 (0·97-1·17) | - |
| Sex | | | |
| Women | 1·00 | 1·00 | 1·00 |
| Men | 0·99 (0·98-1·00) | 0·94 (0·93-0·96) | 1·01 (1·00-1·03) |
| Migration background | | | |
| Belgian natives | 1·00 | 1·00 | 1·00 |
| Second generation migrants | 2·37 (2·31-2·44) | 1·83 (1·77-1·90) | 2·20 (2·12-2·28) |
| First-generation European migrants | 2·22 (2·17-2·27) | 2·14 (2·08-2·19) | 2·45 (2·39-2·51) |
| First-generation non-European migrants | 2·87 (2·81-2·93) | 2·63 (2·57-2·70) | 3·51 (3·43-3·60) |
| Household type | | | |
| Couples with children | 1·00 | 1·00 | 1·00 |
| Couples without children | 0·72 (0·70-0·73) | 0·71 (0·69-0·73) | 0·67 (0·65-0·69) |
| Single parents | 1·09 (1·07-1·11) | 1·13 (1·10-1·15) | 1·25 (1·22-1·28) |
| One-person households | 0·90 (0·89-0·92) | 0·92 (0·90-0·93) | 0·93 (0·92-0·95) |
| Collectivity | 0·67 (0·61-0·73) | 0·63 (0·57-0·69) | 0·56 (0·49-0·63) |
| Other | 0·86 (0·84-0·89) | 0·89 (0·86-0·92) | 0·83 (0·80-0·86) |
| Missing | 0·94 (0·89-0·99) | 0·93 (0·87-0·99) | 0·94 (0·91-0·98) |

**Supplementary Table 5. Continued**

| Variables | Model 1* | Model 2** | Model 3*** |
|---|---|---|---|
| | aOR (95%CI) | aOR (95%CI) | aOR (95%CI) |
| Income level | | | |
| High | 1·00 | 1·00 | - |
| Moderate | 1·84 (1·79-1·88) | 1·66 (1·62-1·71) | - |
| Low | 2·51 (2·45-2·57) | 2·17 (2·12-2·23) | - |
| Missing | 1·83 (1·78-1·89) | 1·69 (1·63-1·75) | - |
| Education level | | | |
| High | - | 1·00 | - |
| Moderate | - | 1·61 (1·57-1·65) | - |
| Low | - | 1·59 (1·55-1·63) | - |
| Missing | - | 1·08 (1·05-1·11) | - |
| Healthcare degree | | | |
| Yes | - | 1·00 | - |
| No | - | 1·48 (1·43-1·54) | - |
| Employment status | | | |
| Employed | - | - | 1·00 |
| Unemployed | - | - | 1·32 (1·30-1·35) |
| Missing | - | - | 0·94 (0·91-0·98) |

* Logistic regression model applied on all age groups (n=579,687). OR are adjusted for age, sex, region, migration background, household type, and income.

** Logistic regression model applied only to individuals aged over 25 years (n=505,082) OR are adjusted for age, sex, region, migration background, household type, income, education level, and healthcare degree.

*** Logistic regression model applied only to individuals aged 25 to 65 years (n=436,356). OR are adjusted for age, sex, region, migration background, household type, and employment status. Income being strongly correlated with employment status, OR are not adjusted for income when employment status is included in the model.

**Supplementary Table 6. Adjusted odds ratios (aORs) and their 95% confidence intervals (CIs) for the association between sociodemographic and socioeconomic characteristics and the odds of not having received a first dose of COVID-19 vaccine in Wallonia, Belgium, December 28th 2020 - August 31st 2021.**

| Variables | Model 1* | Model 2** | Model 3*** |
|---|---|---|---|
| | aOR (95%CI) | aOR (95%CI) | aOR (95%CI) |
| Age (years) | | | |
| 18-24 | 3·91 (3·75-4·07) | - | - |
| 25-34 | 5·45 (5·24-5·67) | 5·70 (5·48-5·93) | 3·03 (2·97-3·09) |
| 35-44 | 3·77 (3·63-3·93) | 3·97 (3·81-4·13) | 2·11 (2·06-2·15) |
| 45-54 | 2·47 (2·38-2·57) | 2·58 (2·47-2·68) | 1·37 (1·34-1·40) |
| 55-64 | 1·94 (1·87-2·02) | 1·99 (1·91-2·08) | 1·00 |
| 65-74 | 1·32 (1·27-1·38) | 1·35 (1·29-1·41) | - |
| 75-84 | 1·00 | 1·00 | - |
| 85+ | 1·24 (1·17-1·32) | 1·21 (1·14-1·29) | - |
| Sex | | | |
| Women | 1·00 | 1·00 | 1·00 |
| Men | 1·05 (1·04-1·06) | 0·99 (0·98-1·01) | 1·06 (1·05-1·08) |
| Migration background | | | |
| Belgian natives | 1·00 | 1·00 | 1·00 |
| Second generation migrants | 1·69 (1·66-1·72) | 1·54 (1·51-1·57) | 1·62 (1·59-1·66) |
| First-generation European migrants | 2·04 (2·01-2·07) | 1·90 (1·87-1·93) | 2·11 (2·07-2·14) |
| First-generation non-European migrants | 2·38 (2·34-2·41) | 2·24 (2·21-2·28) | 2·62 (2·58-2·66) |
| Household type | | | |
| Couples with children | 1·00 | 1·00 | 1·00 |
| Couples without children | 0·95 (0·93-0·96) | 0·89 (0·88-0·91) | 0·90 (0·88-0·92) |
| Single parents | 1·27 (1·25-1·29) | 1·30 (1·27-1·32) | 1·47 (1·44-1·49) |
| One-person households | 1·30 (1·28-1·32) | 1·26 (1·24-1·28) | 1·39 (1·37-1·42) |
| Collectivity | 0·56 (0·52-0·60) | 0·55 (0·51-0·59) | 0·46 (0·42-0·51) |
| Other | 1·37 (1·32-1·42) | 1·33 (1·28-1·38) | 1·30 (1·25-1·36) |
| Missing | 1·06 (0·99-1·13) | 1·07 (0·99-1·16) | 1·00 (0·92-1·09) |

**Supplementary Table 6. Continued**

| Variables | Model 1* | Model 2** | Model 3*** |
|---|---|---|---|
| | aOR (95%CI) | aOR (95%CI) | aOR (95%CI) |
| Income level | | | |
| High | 1·00 | 1·00 | - |
| Moderate | 1·43 (1·41-1·45) | 1·35 (1·32-1·37) | - |
| Low | 2·15 (2·12-2·18) | 1·91 (1·88-1·94) | - |
| Missing | 1·87 (1·81-1·94) | 1·61 (1·55-1·68) | - |
| Education level | | | |
| High | - | 1·00 | - |
| Moderate | - | 1·20 (1·18-1·22) | - |
| Low | - | 1·30 (1·28-1·32) | - |
| Missing | - | 1·18 (1·16-1·21) | - |
| Healthcare degree | | | |
| Yes | - | 1·00 | - |
| No | - | 1·39 (1·35-1·42) | - |
| Employment status | | | |
| Employed | - | - | 1·00 |
| Unemployed | - | - | 1·40 (1·38-1·42) |
| Missing | - | - | 1·26 (1·20-1·32) |

* Logistic regression model applied on all age groups (n=1,413,876). OR are adjusted for age, sex, region, migration background, household type, and income.

** Logistic regression model applied only to individuals aged over 25 years (n=1,235,805) OR are adjusted for age, sex, region, migration background, household type, income, education level, and healthcare degree.

*** Logistic regression model applied only to individuals aged 25 to 65 years (n=998,369). OR are adjusted for age, sex, region, migration background, household type, and employment status. Income being strongly correlated with employment status, OR are not adjusted for income when employment status is included in the model.

**Supplementary Table 7. Crude odds ratios (cORs) and their 95% confidence intervals (CIs) for the association between sociodemographic and socioeconomic characteristics and the odds of not having received a first dose of COVID-19 vaccine, Belgium, December 28th 2020 - August 31st 2021.**

| Variables | Model 1* | Model 2** | Model 3*** |
|---|---|---|---|
| | cOR (95%CI) | cOR (95%CI) | cOR (95%CI) |
| Age groups | | | |
| 18-24 | 6·41 (6·26-6·56) | | - |
| 25-34 | 6·68 (6·53-6·83) | 6·68 (6·53-6·83) | 3·13 (3·10-3·16) |
| 35-44 | 5·09 (4·98-5·21) | 5·09 (4·98-5·21) | 2·39 (2·36-2·41) |
| 45-54 | 3·28 (3·20-3·36) | 3·28 (3·20-3·36) | 1·54 (1·52-1·56) |
| 55-64 | 2·16 (2·12-2·21) | 2·16 (2·11-2·21) | 1·00 |
| 65-74 | 1·44 (1·40-1·48) | 1·14 (1·40-1·48) | - |
| 75-84 | 1·00 | 1·00 | - |
| 85+ | 1·09 (1·05-1·13) | 1·09 (1·05-1·13) | - |
| Sex | | | |
| Women | 1·00 | 1·00 | 1·00 |
| Men | 1·02 (1·01-1·02) | 0·99 (0·99-1·00) | 0·99 (0·98-0·99) |
| Region | | | |
| Flemish | 1·00 | 1·00 | 1·00 |
| Brussels | 3·65 (3·62-3·68) | 3·52 (3·49-3·55) | 3·34 (3·31-3·37) |
| Walloon | 1·76 (1·75-1·77) | 1·81 (1·80-1·82) | 1·77 (1·75-1·78) |
| Migration background | | | |
| Belgian natives | 1·00 | 1·00 | 1·00 |
| Second-generation migrants | 4·29 (4·25-4·33) | 3·70 (3·65-3·75) | 3·29 (3·25-3·34) |
| First-generation European migrants | 4·22 (4·19-4·26) | 3·98 (3·95-4·01) | 3·73 (3·70-3·76) |
| First-generation non-European migrants | 5·249 (5·211-5·286) | 5·12 (5·08-5·16) | 4·52 (4·48-4·55) |
| Household type | | | |
| Couples with children | 1·00 | 1·00 | 1·00 |
| Couples | 0·48 (0·48-0·49) | 0·47 (0·46-0·47) | 0·63 (0·62-0·63) |
| Single parents | 1·50 (1·48-1·51) | 1·48 (1·46-1·49) | 1·53 (1·51-1·54) |
| One-person households | 0·96 (0·95-0·96) | 0·94 (0·93-0·95) | 1·22 (1·21-1·23) |
| Collectivity | 0·29 (0·28-0·30) | 0·27 (0·26-0·28) | 0·62 (0·58-0·65) |

**Supplementary Table 7. Continued**

| Variables | Model 1* | Model 2** | Model 3*** |
|---|---|---|---|
| | **cOR (95%CI)** | **cOR (95%CI)** | **cOR (95%CI)** |
| Other | 1·50 (1·48-1·53) | 1·46 (1·44-1·49) | 1·59 (1·56-1·62) |
| Missing | 2·45 (2·38-2·52) | 2·31 (2·23-2·39) | 2·33 (2·25-2·41) |
| Income | | | |
| High | 1·00 | 1·00 | - |
| Moderate | 1·74 (1·72-1·75) | 1·70 (1·68-1·72) | - |
| Low | 3·43 (3·40-3·46) | 3·12 (3·10-3·15) | - |
| Missing | 3·92 (3·87-3·97) | 3·35 (3·30-3·40) | - |
| Education | | | |
| High | - | 1.00 | - |
| Moderate | - | 1.65 (1.64-1.67) | - |
| Low | - | 1.56 (1.54-1.57) | - |
| Missing | - | 3.54 (3.51-3.57) | - |
| Healthcare degree | | | |
| Yes | - | 1·00 | - |
| No | - | 1·71 (1·69-1·73) | - |
| Employment status | | | |
| Employed | - | - | 1·00 |
| Unemployed | - | - | 1·67 (1·65-1·68) |
| Missing | - | - | 2·72 (2·68-2·77) |

** Logistic regression model applied on all age groups (n=5,342,110).*

*** Logistic regression model applied only to individuals aged over 25 years (n=4,696,694).*

**** Logistic regression model applied only to individuals aged 25 to 65 years (n=3,792,100).*

**Supplementary Table 8. Crude odds ratios (cORs) and 95% confidence intervals (CIs) for the association between sociodemographic and socioeconomic characteristics and the odds of not having received a first dose of COVID-19 vaccine in Flanders, Belgium, December 28th 2020 - August 31st 2021.**

| Variables | Model 1* | Model 2** | Model 3*** |
|---|---|---|---|
| | cOR (95%CI) | cOR (95%CI) | cOR (95%CI) |
| Age groups | | | |
| 18-24 | 6·90 (6·67-7·15) | - | - |
| 25-34 | 7·29 (7·04-7·54) | 7·29 (7·04-7·54) | 3·22 (3·17-3·27) |
| 35-44 | 5·52 (5·34-5·71) | 5·52 (5·33-5·71) | 2·44 (2·40-2·48) |
| 45-54 | 3·53 (3·41-3·66) | 3·53 (3·41-3·66) | 1·56 (1·54-1·59) |
| 55-64 | 2·29 (2·21-2·37) | 2·29 (2·21-2·37) | 1·00 |
| 65-74 | 1·49 (1·43-1·55) | 1·49 (1·43-1·55) | - |
| 75-84 | 1·00 | 1·00 | - |
| 85+ | 1·08 (1·02-1·14) | 1·08 (1·02-1·14) | - |
| Sex | | | |
| Women | 1·00 | 1·00 | 1·00 |
| Men | 1·05 (1·04-1·06) | 1·02 (1·01-1·03) | 1·02 (1·01-1·03) |
| Migration background | | | |
| Belgian natives | 1·00 | 1·00 | 1·00 |
| Second-generation migrants | 4·55 (4·48-4·62) | 3·96 (3·88-4·04) | 3·51 (3·44-3·59) |
| First-generation European migrants | 5·63 (5·56-5·69) | 5·30 (5·23-5·37) | 5·02 (4·95-5·08) |
| First-generation non-European migrants | 5.54 (5.49-5.60) | 5.36 (5.30-5.42) | 4.64 (4.58-4.69) |
| Household type | | | |
| Couples with children | 1·00 | 1·00 | 1·00 |
| Couples | 0·49 (0·48-0·49) | 0·47 (0·46-0·48) | 0·64 (0·63-0·65) |
| Single parents | 1·52 (1·51-1·54) | 1.51 (1.49-1.53) | 1·56 (1·54-1·59) |
| One-person households | 0.95 (0.94-0.96) | 0.93 (0.92-0.94) | 1.24 (1.22-1.26) |
| Collectivity | 0.30 (0.29-0.32) | 0.28 (0.26-0.29) | 0.72 (0.67-0.78) |
| Other | 1·70 (1·66-1·74) | 1·66 (1·62-1·71) | 1·84 (1·79-1·89) |
| Missing | 2·98 (2·85-3·11) | 2·86 (2·72-3·00) | 2·89 (2·75-3·04) |

**Supplementary Table 8. Continued**

| Variables | Model 1* | Model 2** | Model 3*** |
|---|---|---|---|
| | cOR (95%CI) | cOR (95%CI) | cOR (95%CI) |
| Income level | | | |
| High | 1·00 | 1·00 | - |
| Moderate | 1·73 (1·71-1·76) | 1·69 (1·67-1·72) | - |
| Low | 3·33 (3·29-3·36) | 3·01 (2·98-3·05) | - |
| Missing | 4·54 (4·45-4·63) | 3·91 (3·82-3·99) | - |
| Education level | | | |
| High | | 1·00 | - |
| Moderate | | 1·58 (1·56-1·60) | - |
| Low | - | 1·44 (1·42-1·46) | - |
| Missing | - | 4·31 (4·25-4·36) | - |
| Healthcare degree | | | |
| Yes | - | 1·00 | - |
| No | - | 1·73 (1·70-1·77) | - |
| Employment status | | | |
| Employed | - | - | 1·00 |
| Unemployed | - | - | 1·70 (1·68-1·72) |
| Missing | | | 3·56 (3·47-3·65) |

* Logistic regression model applied on all age groups (n=3,348,547). OR are adjusted for age, sex, region, migration background, household type, and income.

** Logistic regression model applied only to individuals aged over 25 years (n=2,955,807) OR are adjusted for age, sex, region, migration background, household type, income, education level, and healthcare degree.

*** Logistic regression model applied only to individuals aged 25 to 65 years (n=2,357,375). OR are adjusted for age, sex, region, migration background, household type, and employment status. Income being strongly correlated with employment status, OR are not adjusted for income when employment status is included in the model.

**Supplementary Table 9. Crude odds ratios (cORs) and 95% confidence intervals (CIs) for the association between sociodemographic and socioeconomic characteristics and the odds of not having received a first dose of COVID-19 vaccine in Brussels, Belgium, December 28th 2020 - August 31st 2021.**

| Variables | Model 1* | Model 2** | Model 3*** |
|---|---|---|---|
| | cOR (95%CI) | cOR (95%CI) | cOR (95%CI) |
| Age groups | | | |
| 18-24 | 7·74 (7·32-8·19) | - | - |
| 25-34 | 5·38 (5·09-5·68) | 5·38 (5·09-5·68) | 2·42 (2·36-2·48) |
| 35-44 | 4·55 (4·31-4·81) | 4·55 (4·31-4·81) | 2·05 (2·00-2·10) |
| 45-54 | 3·16 (2·99-3·34) | 3·16 (2·99-3·34) | 1·42 (1·38-1·46) |
| 55-64 | 2·25 (2·13-2·39) | 2·25 (2·13-2·39) | 1·00 |
| 65-74 | 1·47 (1·38-1·56) | 1·47 (1·38-1·56) | - |
| 75-84 | 1·00 | 1·00 | - |
| 85+ | 0·89 (0·81-0·98) | 0·89 (0·81-0·98) | - |
| Sex | | | |
| Women | 1·00 | 1·00 | 1·00 |
| Men | 0·96 (0·94-0·97) | 0·93 (0·92-0·94) | 0·91 (0·90-0·92) |
| Migration background | | | |
| Belgian natives | 1·00 | 1·00 | 1·00 |
| Second-generation migrants | 4·80 (4·68-4·93) | 3·26 (3·15-3·38) | 2·88 (2·78-2·99) |
| First-generation European migrants | 3·07 (3·01-3·14) | 2·83 (2·76-2·89) | 2·51 (2·45-2·58) |
| First-generation non-European migrants | 4.38 (4.29-4.47) | 4.25 (4.16-4.34) | 3.85 (3.77-3.94) |
| Household type | | | |
| Couples with children | 1·00 | 1·00 | 1·00 |
| Couples | 0·47 (0·47-0·48) | 0·49 (0·48-0·50) | 0·59 (0·58-0·61) |
| Single parents | 1·19 (1·17-1·21) | 1·18 (1·16-1·21) | 1·20 (1·18-1·23) |
| One-person households | 0.70 (0.69-0.71) | 0.72 (0.71-0.74) | 0.83 (0.81-0.85) |
| Collectivity | 0.23 (0.21-0.25) | 0.22 (0.20-0.24) | 0.45 (0.39-0.51) |
| Other | 0·83 (0·81-0·86) | 0·84 (0·81-0·87) | 0·84 (0·81-0·87) |
| Missing | 1·08 (1·03-1·14) | 1·02 (0·96-1·08) | 1·01 (0·95-1·07) |

**Supplementary Table 9. Continued**

| Variables | Model 1* | Model 2** | Model 3*** |
|---|---|---|---|
| | ***cORs*** | ***cORs*** | ***cORs*** |
| Income level | | | |
| High | 1·00 | 1·00 | - |
| Moderate | 2·23 (2·18-2·29) | 2·15 (2·10-2·21) | - |
| Low | 3·62 (3·54-3·69) | 3·24 (3·17-3·32) | - |
| Missing | 2·94 (2·86-3·03) | 2·53 (2·46-2·61) | - |
| Education level | | | |
| High | - | 1·00 | - |
| Moderate | - | 2·48 (2·42-2·53) | - |
| Low | - | 2·25 (2·20-2·30) | - |
| Missing | - | 2·19 (2·15-2·24) | - |
| Healthcare degree | | | |
| Yes | - | 1·00 | - |
| No | - | 1·74 (1·68-1·80) | - |
| Employment status | | | |
| Employed | - | - | 1·00 |
| Unemployed | - | - | 1·49 (1·47-1·51) |
| Missing | - | - | 1·26 (1·22-1·29) |

* Logistic regression model applied on all age groups (n=579,687). OR are adjusted for age, sex, region, migration background, household type, and income.

** Logistic regression model applied only to individuals aged over 25 years (n=505,082) OR are adjusted for age, sex, region, migration background, household type, income, education level, and healthcare degree.

*** Logistic regression model applied only to individuals aged 25 to 65 years (n=436,356). OR are adjusted for age, sex, region, migration background, household type, and employment status. Income being strongly correlated with employment status, OR are not adjusted for income when employment status is included in the model.

**Supplementary Table 10. Crude odds ratios (cORs) and 95% confidence intervals (CIs) for the association between sociodemographic and socioeconomic characteristics and the odds of not having received a first dose of COVID-19 vaccine in Wallonia, Belgium, December 28th 2020 - August 31st 2021.**

| Variables | Model 1* | Model 2** | Model 3*** |
|---|---|---|---|
| | cOR (95%CI) | cOR (95%CI) | cOR (95%CI) |
| *Age groups* | | | |
| 18-24 | 4·31 (4·14-4·47) | - | - |
| 25-34 | 5·48 (5·28-5·69) | 5·48 (5·28-5·69) | 3·10 (3·04-3·15) |
| 35-44 | 4·01 (3·86-4·17) | 4·01 (3·86-4·16) | 2·27 (2·22-2·31) |
| 45-54 | 2·54 (2·45-2·64) | 2·54 (2·45-2·64) | 1·44 (1·41-1·47) |
| 55-64 | 1·79 (1·72-1·86) | 1·79 (1·72-1·86) | 1·00 |
| 65-74 | 1·26 (1·21-1·32) | 1·26 (1·21-1·32) | - |
| 75-84 | 1·00 | 1·00 | - |
| 85+ | 1·16 (1·10-1·23) | 1·16 (1·10-1·23) | - |
| Sex | | | |
| Women | 1·00 | 1·00 | 1·00 |
| Men | 1·02 (1·01-1·03) | 1·01 (0·99-1·02) | 1·01 (0·99-1·02) |
| Migration background | | | |
| Belgian natives | 1·00 | 1·00 | 1·00 |
| Second-generation migrants | 2·59 (2·55-2·63) | 2·49 (2·44-2·54) | 2·22 (2·17-2·26) |
| First-generation European migrants | 2·18 (2·15-2·22) | 2·13 (2·10-2·16) | 2·03 (2·00-2·06) |
| First-generation non-European migrants | 3.40 (3.35-3.45) | 3.34 (3.29-3.39) | 2.95 (2.91-3.00) |
| Household type | | | |
| Couples with children | 1·00 | 1·00 | 1·00 |
| Couples | 0·56 (0·55-0·57) | 0·53 (0·52-0·54) | 0·71 (0·70-0·72) |
| Single parents | 1·38 (1·36-1·40) | 1·36 (1·33-1·38) | 1·40 (1·37-1·42) |
| One-person households | 0.97 (0.95-0.98) | 0.94 (0.93-0.96) | 1.21 (1.19-1.23) |
| Collectivity | 0.31 (0.29-0.33) | 0.29 (0.27-0.31) | 0.54 (0.49-0.59) |
| Other | 1·37 (1·32-1·41) | 1·30 (1·26-1·35) | 1·45 (1·39-1·50) |
| Missing | 2·20 (2·07-2·34) | 2·11 (1·97-2·27) | 2·15 (2·00-2·31) |

**Supplementary Table 10. Continued**

| Variables | Model 1* | Model 2** | Model 3*** |
|---|---|---|---|
| | **cOR (95%CI)** | **cOR (95%CI)** | **cOR (95%CI)** |
| Income level | | | |
| High | 1·00 | 1·00 | - |
| Moderate | 1·53 (1·51-1·55) | 1·50 (1·48-1·53) | - |
| Low | 2·47 (2·43-2·50) | 2·31 (2·28-2·35) | - |
| Missing | 2·30 (2·24-2·37) | 2·00 (1·93-2·06) | - |
| Education level | | | |
| High | - | 1·00 | - |
| Moderate | - | 1·59 (1·57-1·62) | - |
| Low | - | 1·44 (1·42-1·47) | - |
| Missing | - | 2·42 (2·38-2·46) | - |
| Healthcare degree | | | |
| Yes | - | 1·00 | - |
| No | - | 1·45 (1·42-1·48) | - |
| Employment status | | | |
| Employed | - | - | 1·00 |
| Unemployed | - | - | 1·33 (1·31-1·35) |
| Missing | - | - | 2·26 (2·17-2·35) |

* Logistic regression model applied on all age groups (n=1,413,876).

** Logistic regression model applied only to individuals aged over 25 years (n=1,235,805).

*** Logistic regression model applied only to individuals aged 25 to 65 years (n=998,369).

# SECTION II

# What are the social patterns of COVID-19-related morbidity and mortality?

# *Chapter 2. Social health inequalities in COVID-19 hospitalization and mediating factors*

***Adapted from***

**Lisa Cavillot,** Beatrijs Moerkerke, Brecht Devleesschauwer, Jinane Ghattas, Joris A.F. van Loenhout, Laura Van den Borre, Niko Speybroeck, Tom Loeys & Robby De Pauw. **The role of vaccination, underlying health conditions, and working in healthcare in the socioeconomic disparities in COVID-19 hospitalization: a mediation analysis using interventional effect models.** Submitted to *BMC Infectious Diseases* (September 2024).

## Abstract

**Background**: prior research has established an association between socioeconomic determinants and COVID-19 hospitalization rates. This study explores the potential mediation of this association by vaccination status, underlying health conditions, and working in healthcare.

**Methods**: we conducted a retrospective cohort study including 148,590 and 430,693 adults living in the Walloon or Brussels regions in Belgium with a confirmed SARS-CoV-2 infection during the Delta-predominant period (1 September 2021 to 22 December 2021) and the Omicron-predominant period (4 January 2022 to 1 September 2022), respectively. Data on COVID-19-related hospital admissions, education, vaccination, underlying health conditions, and individuals working in healthcare were retrieved from multiple national registers, linked at an individual level using the Belgian social security number. Mediation analyses using interventional effect models were used to assess the mediation role of vaccination, underlying health conditions, and working in healthcare in the relationship between education and COVID-19 hospitalization.

**Results**: the interventional direct effects observed were significant and positive across both periods of variant predominance indicating a 0.554% (95% CI 0.449% - 0.662%) and 0.484% (95% CI 0.437% - 0.529%) higher likelihood of COVID-19

related hospitalization among individuals with low educations levels, compared to those with moderate or high levels, during Delta and Omicron periods, respectively. The mediating effects of vaccination and underlying health conditions on the association between education and COVID-19 hospitalization were statistically significant (vaccination: 0.062%, 95% CI 0.054% - 0.071% for the Delta period, and 0.006, 95% CI 0.005% - 0.008% for the Omicron period; underlying health conditions: 0.038%, 95% CI 0.029% - 0.046% for the Delta period, and 0.021%, 95% CI 0.018% - 0.025% for the Omicron period). Working in healthcare did not show a significant mediating effect.

**Conclusions**: this study demonstrates that socioeconomic disparities in COVID-19 hospitalization rates are partially mediated by vaccination and underlying health conditions. This suggests that public health interventions aimed at increasing vaccination rates and preventing underlying health conditions could mitigate some of the effects of socioeconomic disparities in COVID-19 hospitalization outcomes.

## 1. Introduction

The coronavirus disease 2019 (COVID-19), caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has spread rapidly throughout the world, resulting in a pandemic, as announced by the World Health Organization (WHO) in March 2020 [22]. The COVID-19 pandemic was very quickly described as a *syndemic*, emphasizing the interaction between both the purely biological nature of the virus and the social and cultural determinants, exacerbating the social inequalities already entrenched in our society [101,236,237]. This has been confirmed by a large body of literature reporting a social distribution pattern in SARS-CoV-2 infection and in the development of subsequent severe COVID-19 outcomes (e.g., hospitalization, ICU admission, death) to the detriment of individuals belonging to lower socioeconomic (SE) groups (e.g., low levels of income or education, living in overcrowded households, poor housing conditions) [120,210–212,236,238].

SE disadvantaged groups present poorer baseline health conditions with a higher rate of non-communicable diseases (e.g., obesity, cardiovascular diseases, diabetes) [239], which has been shown to increase the risk of developing severe COVID-19 outcomes [124,240]. While the vaccine has proven to be effective in protecting against severe COVID-19 disease [64,241,242], the spread of such preventive measures is accompanied by an increase in social health inequalities,

particularly in its early stages [243]. Indeed, several studies have confirmed lower COVID-19 vaccine uptake among SE disadvantaged groups [213,215,244]. To adopt preventive health behaviors such as vaccination, health knowledge (which is particularly present among healthcare workers), plays an important role [245,246]. Individuals from lower SE groups tend to have lower levels of health literacy, making them less inclined to adopt preventive measures such as vaccination and hence increasing their risk of severe COVID-19 outcomes [234]. In addition to health knowledge, other factors may contribute to the lower vaccination coverage among SE disadvantaged groups, such as lack of trust in government or scientific, transport difficulties, language barriers, or inflexible working hours [247,248]. To our knowledge, no research has yet examined how the association between SE determinants and severe COVID-19 outcomes may be mediated by vaccination, underlying health conditions, and working in healthcare.

The objective of the present study was to identify, among the adult population residing in the Walloon or Brussels regions in Belgium with a confirmed SARS-CoV-2 infection, the mediating effect of vaccination, underlying health conditions, and working in healthcare on the association between education and COVID-19 hospitalization. We have therefore hypothesized that the latter three variables were on the causal pathway between education and COVID-19 hospitalization and could mediate their association.

## 2. Materials & methods

### 2.1. Causal framework and study design

We conducted a retrospective cohort study using data from the LINK-VACC project, a causal conceptual framework set up by Sciensano, the Belgian Institute of Health, containing, within a pseudonymized environment, selected variables from multiple existing national health and social sector registries linked at an individual level using the Belgian social security number [133].

For the present study, six registers were used : 1) the COVID-19 Healthdata test database containing exhaustive data from COVID-19 laboratory tests (polymerase chain reaction or antigenic tests) performed in Belgium (date of sampling, test result) as well as demographical data on the tested person [114], 2) the Clinical Hospital Surveillance (CHS) database containing non-exhaustive patient-level data from admission and discharge forms on approximately 50% of the patients admitted with a confirmed COVID-19 diagnosis across all 103 Belgian hospitals [112], 3) the Belgian Vaccine register for COVID-19 (Vaccinnet+) containing exhaustive data on COVID-19 vaccine doses administered to Belgian residents (brand of the vaccine, lot number, date of administration, date of registration) as well as demographical data on the vaccinated person (sex, age, postal code of residence), 4) the DEMOBEL database provided by Statistics Belgium (Statbel) containing data on sociodemographic (SD) and SE characteristics on everyone

tested and/or vaccinated for COVID-19, 5) the InterMutualistic Agency (IMA) database containing data on reimbursed healthcare and medicines of Belgian residents insured in the country tested and/or vaccinated for COVID-19, 6) the Common Base Registry for HealthCare Actors (CoBRHA) allowing the identification of individuals who have been licensed to practice a healthcare profession in Belgium. Belgium is composed of three regions: Brussels, Flanders, and Wallonia. This research focused only on data from the Walloon and Brussels regions from the six data sources mentioned above, as permission to use the data was not granted for the Flemish region. As of 1 January 2021, the Walloon and Brussels regions covered 42% of the total Belgian population [249].

### 2.2. Study population and study period

All laboratory confirmed SARS-CoV-2 infections recorded in the COVID-19 Healthdata test database between 1 September 2021 and 1 September 2022 among adults (≥ 18 years) living in the Walloon or Brussels regions in Belgium were selected. As of 1 September 2021, all adults had the opportunity to be vaccinated for the primary vaccination course, and the majority had received it, with 71% of the total Belgian adult population having completed their primary vaccination course at that date [25,44]. In addition, 1 September 2021 marks the start of the vaccination campaign for the first booster in Belgium [250].

Through a link with the CHS database, we were able to determine whether a COVID-19-related hospital admission was registered following the date of a positive SARS-CoV-2 test. Therefore, cases were defined as any individuals whose positive test date was followed by a COVID-19 hospitalization within 30 days, or as any individuals whose positive test date was preceded by a COVID-19 hospitalization within one day (in order to not consider individuals with a nosocomial infection as cases).

Only one positive test per individual was considered for the analysis. Within the same individual, we excluded any positive tests not leading to a COVID-19 hospitalization, recorded before or after a positive test leading to a COVID-19 hospitalization. Any positive tests recorded more than one day after the date of hospital admission were dropped. When two or more positive tests led to the same hospital admission within the same individual, only the first was kept as the others probably reflected the same infection. For individuals who have more than one positive test without any registered COVID-19-related hospital admission, infection periods were defined. An infection period included all positive tests counted over a 21-day period. For each infection period, only the first test was kept as the other tests probably reflected the same infection. When several infection periods were recorded for the same individual, only one was chosen at random to avoid

duplication and keep only one record per individual in the control group, as in the case group.

The vaccination rollout for the second booster had begun at the end of the study period with only 5.34% of the Belgian population vaccinated with a second booster by 1 September 2022 [25]. Therefore, individuals who received a second booster, as well as individuals with incomplete vaccination course (i.e., a single dose of a two-dose primary schedule), were excluded. Finally, we excluded individuals who were deregistered, migrated, and deceased based on their status in the national register as well as individuals with incomplete demographic, SE, and clinical data. Based on the baseline surveillance of variant circulation in Belgium [250], the study period was divided according to the variant of concern (VOC) of predominance given the impact of SARS-CoV-2 variants on COVID-19 disease severity and vaccine effectiveness [59,251]. A VOC was considered dominant when it accounted for at least 80% of the infections recorded over a period. The Delta-predominant period included all positive tests sampled between 1 September 2021 (start of the study period) until 22 December 2021 and the Omicron-predominant period included all positive tests sampled between 4 January 2022 until 1 September 2022 (end of the study period). All positive tests sampled between 22 December 2021 and 4 January 2022 were dropped.

## 2.3. Outcome

The study outcome consisted of a COVID-19-related hospital admission defined by a hospital admission form recoded in the CHS database with a confirmed COVID-19 diagnosis by means of a positive RT-PCR or RAT and symptoms at admission. Based on the classification of SARS-CoV-2 infections established by the European Center for prevention and Disease Control (ECDC) [252], only community-acquired, nursing-home acquired, and infections of unknown origin in healthcare workers were kept while indeterminate COVID-19 cases, and healthcare-associated or probable healthcare-associated nosocomial infections were excluded.

## 2.4. Exposure

As SE exposure, we used the educational level provided by the DEMOBEL database and classified in eight categories using the International Standard Classification of Education (ISCED): 'Early childhood education' (less than primary for educational attainment) (ISCED 0), 'Primary education' (ISCED 1), 'Lower secondary education' (ISCED 2), 'Upper secondary education' (ISCED 3), 'Post-secondary non-tertiary education' (ISCED 4), 'Short-cycle tertiary education' (ISCED 5), 'Bachelor's or equivalent level' (ISCED 6), 'Master's or equivalent level' (ISCED 7), 'Doctoral or equivalent level' (ISCED 8). We merged these different categories into two main educational levels: 'Low' (ISCED 0 to ISCED 2) and 'Moderate or high' (ISCED 3 to ISCED 8). 'Moderate or high education' was considered as the reference category.

## 2.5. Mediators

In the causal inference framework for observational studies, an exposure can be associated with an outcome through both direct and indirect pathways [159]. The indirect pathway involves a mediator, an intermediate variable that lies in the causal pathway between the exposure and the outcome, partially or fully explaining their association and making the unraveling of causal mechanisms feasible [159,160]. A mediator can be distinguished from a confounder which is a third variable with a direct causal effect on the outcome and the exposure of interest [253].

As mediator variables, we selected vaccination status, underlying health conditions, and working in healthcare.

The vaccination status was taken 14 days before the date of infection with SARS-CoV-2 and was distinguished between 'No vaccination' and 'Primary or first-booster vaccination'. A primary-vaccination was defined as receiving two doses of BNT162b2, mRNA-1273, or ChAdOx1 (mRNA vaccines), or one dose of Ad26.COV2.S (viral vector vaccine). 'Primary or first-booster vaccination' was considered as the reference category.

Underlying health conditions (e.g., cardiovascular diseases, thrombosis, type 2 diabetes, chronic obstructive pulmonary disease) were deduced from data on reimbursed healthcare and medicine of Belgian residents ensured in the country

available in IMA. Within the LINKVACC project, only underlying health conditions associated with an increased risk of severe COVID-19 were selected from IMA, this is therefore not a complete overview of all underlying health conditions. The overview of underlying health conditions included in IMA can be found in a report by Stouten et al. [254]. Two categories based on the number of underlying health conditions were distinguished: 'No underlying health conditions' and 'At least one underlying health condition'. 'No underlying health conditions' was considered as the reference category.

To identify individuals working in healthcare, we used a proxy variable available in CoBRHA allowing to identify all individuals who have been licensed to practice a healthcare profession in Belgium. 'Working in healthcare' was considered as the reference category.

These three variables were chosen as mediators in view of their presumed association with both SE status and COVID-19 hospitalization. Indeed, despite the known effectiveness of the vaccine against COVID-19 hospitalization [255], a lower vaccine uptake was identified in SE disadvantaged groups [215,244]. Some underlying health conditions (e.g., obesity, cardiovascular diseases, chronic kidney diseases) have been shown to be preponderant in SE disadvantaged groups [239] and a risk factor for COVID-19 hospitalization [256,257]. Healthcare workers are expected to have a higher level of health knowledge, and a higher level of education

has been associated with higher health knowledge [258]. A lower level of health knowledge leads to a lower adoption of preventive behaviors (e.g., social distancing, vaccination), increasing exposure to the virus and potentially leading to more severe disease outcomes [245,246].

### 2.6. Confounders

Age (at the beginning of the study period), gender, province of residence, and migration background were included as potential confounders. Age and gender were available in Vaccinnet+ and COVID-19 Healthdata test database, registers re-using demographic information from the Belgian national register. The province of residence was based on the postal code of residence available in the national register. Concerning the Walloon region, the five provinces were included: Walloon-Brabant, Hainaut, Liege, Luxemburg, and Namur. The Brussels region is not branched into provinces, and has therefore been directly included alongside the five provinces of the Walloon region. CHS hospital coverage during the study period varied by province, with the highest coverage for the province of Hainaut (70%) and the lowest for the province of Liege (15%). Assuming that the province of residence was equivalent to the province of the hospital where the individual was admitted, we used the province of residence as confounder to counter the bias induced by the difference in hospitalization coverage by province. Indeed, province has been considered as confounder in the statistical analyses given the known

association between SE status and province in Belgium [259–261]. Migration background was available in STATBEL and based on a combination of the first nationality and the parent's country of origin. Four groups were distinguished: 'Belgian natives', 'Second-generation migrants', 'First-generation migrants', and 'First-generation non-European migrants'.

### 2.7. Statistical analyses

The study population was first described with the frequency rates and percentages for categorical variables and the median and the interquartile range (IQR) for continuous variables.

To assess the effect of multiple mediators, i.e., vaccination, underlying health conditions, and working in healthcare, on the association between education and COVID-19 hospitalization, we applied causal mediation analysis within the interventional effects framework. Causal mediation analysis are usually applied within the potential outcomes framework, defining the causal effect for a given unit as the difference in its outcome under exposure versus control conditions, measured at the same time point [151,153,154]. However, this approach requires strong assumptions when multiple mediators are involved. Specifically, it assumes a known causal ordering of mediators and no unmeasured confounding, conditions that are rarely satisfied in practice [170]. We therefore applied causal mediation analysis within the interventional effects framework, an extension of the potential

outcomes framework, which offers more flexibility when the true causal structure among the mediators is unknown by estimating indirect effects through simulated intervention on mediator distributions [161,169,172].

Our proposed causal diagram for mediation analyses within the interventional effects framework is presented in **Figure 23**. As explained above, interventional (in)direct effect models allow to consider multiple mediators (M1, M2, M3) on the causal pathway from exposure (A) to outcome (Y) without having to specify any causal structure among the mediators (**Figure 23**) [170,171]. Through this method, the total effect is decomposed into direct and indirect effects and each indirect effect passing through each distinct mediator can be assessed separately [171]. Interventional effects are perfectly suited to assess effects of non-manipulable exposures such as SE determinants and have to be interpreted on the population level rather than the individual level [171].

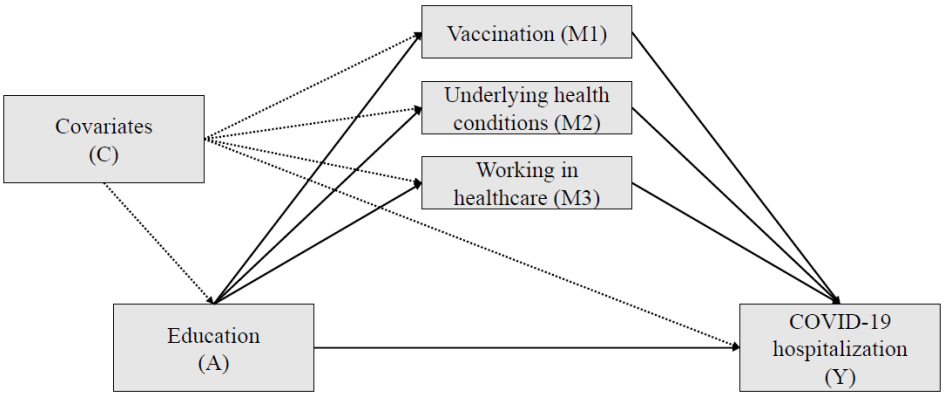



***Figure 23. Proposed causal diagram for the mediating effect of vaccination (M1), underlying health conditions (M2), and working in healthcare (M3) on the association between education (A) and COVID-19 hospitalization (Y) conditional on covariates (C). M1, M2, and M3 are assumed to be independent and not affect each other causally.***

The total effect was estimated by regressing COVID-19 hospitalization (*Y*) on education (*A*) conditional on covariates (*C*) (**Eq. 9**).

$$E(Y|A,C) = \beta_0 + \beta_A A + \beta_{C1} Age + \beta_{C2} Gender + \beta_{C3} Province + \beta_{C4} Migration\ background \quad \textbf{(Eq. 9)}$$

Although the exposure and mediators are binary and generalized linear models can be used to estimate risk differences within the interventional effects framework, we chose to use linear models for their analysis, hence assuming a linear probability model. This approach offered a number of advantages. Firstly, the effects estimated were directly interpretable as probabilities and the decomposition of the total effect into direct and indirect effects was more straightforward when computing

linear regression coefficients [262,263]. Additionally, our proposed method for mediation analysis was computationally more efficient with linear models.

To estimate the interventional (in)direct effects, two sets of linear regression models were used. The first model (**Eq. 10**) was estimated by regressing COVID-19 hospitalization (*Y*) on education (*A*) and mediators (*Ms*) conditional on covariates (*C*).

$$E(Y|A, Ms, C) = \beta_0 + \beta_A A + \sum_{s=1}^{3} \beta_s M_s + \beta_{C1} Age + \beta_{C2} Gender + \beta_{C3} Province + \beta_{C4} Migration\ background \quad \textbf{(Eq. 10)}$$

The second set of models (**Eq. 11**) was estimated by regressing each mediator (*Ms*) separately on education (*A*) conditional on covariates (*C*). The effect of education (*A*) transmitted through each distinct mediator (*Ms*) along the causal pathway between *A* and *Ms* was captured in the regression coefficient $\delta_s$ (**Eq. 11**).

$$E(Ms|A, C) = \delta_{0s} + \delta_s A + \delta_{C1s} Age + \delta_{C2s} Gender + \delta_{C3s} Province + \delta_{C4s} Migration\ background \quad \textbf{(Eq. 11)}$$

The statistical methodology used in this paper was based on Loh. et al. [171] and the code used in this study can be found on GitHub [264]. All statistical analyses were performed in R version 4.0.5. [221]. The package “lavaan” [265] was used to estimate interventional (in)direct effect models.

## 2.8. Sensitivity analyses

Education, vaccination, underlying health conditions, and working in healthcare are interrelated factors that may interact in complex ways to influence COVID-19 hospitalization. Indeed, interactions between mediators (i.e., vaccination, underlying health conditions, and working in healthcare) can be expected. The effect of vaccination may vary for different levels of underlying health conditions due to the prioritization of individuals with comorbidities during Belgium's vaccination campaign [44]. The effect of vaccination may also vary depending on whether an individual work in healthcare as healthcare workers were prioritized during the vaccination campaign in Belgium and were reported to have higher vaccination rate [25,44]. The effect of education on COVID-19 hospitalization can also depend on the value of the mediators as variation in vaccination coverage, number of underlying health conditions, and working in healthcare can be expected according to the level of education [103,234,244].

To take these potential interactions into account, sensitivity analyses were performed including three types of interactions in the outcome model to test the robustness of our results: exposure-mediator, mediator-mediator, and exposure-mediator-mediator interactions.

The outcome model was therefore defined as described in **Eq. 12** where $(A{:}\,M_k)$ represents the exposure-mediator interactions, $\left(M_i{:}\,M_j\right)$ represents the mediator-

mediator interactions, and $(A{:}\,M_i{:}\,M_j)$ represents the exposure-mediator-mediator interactions. The results can be found in **Supplementary Table 11.**

$$E(Y|A, Ms, C) = \beta_0 + \beta_A A + \sum_{s=1}^{3} \beta_s M_s +$$

$$\sum_{k=1}^{3} \left(\beta_{A:M_k}(A{:}\,M_k)\right) + \sum_{i,j=1}^{3} \left(\beta_{M_i:M_j}(M_i{:}\,M_j)\right) +$$

$$\sum_{m,i,j=1}^{3} \left(\beta_{A:M_i:M_j}(A{:}\,M_i{:}\,M_j)\right) + \beta_{C1} Age + \beta_{C2} Gender + \beta_{C3} Province +$$

$$\beta_{C4} Migration\ background \quad \textbf{\textit{(Eq. 12)}}$$

The IIE for a single mediator $Ms$ represents the change in the expected outcome when the distribution of $Ms$ is shifted from what it would be under one treatment level (e.g., $A = 0$) to what it would be under another (e.g., $A = 1$), while holding $A$ fixed. This effect is estimated by simulating the outcome under counterfactual interventions on the mediators, and incorporates not only the main effect of $Ms$ on the outcome but also its interactions with $A$ and with other mediators, as captured by the outcome model coefficients. When mediators are correlated, and their covariance structure differs across treatment levels, an additional component, referred to as the **mutual indirect effect**, arises. This component reflects the portion of the indirect effect that is transmitted jointly through the interdependence among mediators. It is captured by the interaction terms involving the three mediators and the exposure, weighted by the differences in the covariance structure of the mediators between treatment groups [171].

When including mediator-mediator interactions in the outcome model, the indirect effect through each mediator is influenced by the labelling assigned to the mediators, which is presumed arbitrary. This leads to a variation of the effect depending on the specific order in which the mediators are labeled, without any causal ordering imposed on them [171]. To account for this, an additional sensitivity analysis was conducted by examining all possible combinations of labeling the mediators, which involves considering different permutations (**Supplementary Table 12**).

Additionally, as the primary analysis defined cases based on the positive test that led to a COVID-19 hospitalization, there is a potential bias due to outcome-dependent sampling. To assess the robustness of our results, we conducted sensitivity analysis in which cases and controls were redefined by randomly selecting one positive test per individual during each VOC period, or the first positive test of a randomly selected infection period for individuals with multiple infection periods, regardless of hospitalization status. The results of this analysis, both without interaction and with exposure-mediator, mediator-mediator, and exposure-mediator-mediator interactions are available in **Supplementary Tables 13** to **16**.

## 3. Results

### 3.1. Characterization of the study population

The characterization of the study population stratified by VOC-period is available in **Table 16**. The final study population included 148,590 and 430,693 individuals with a confirmed SARS-CoV-2 infection during Delta and Omicron periods, respectively.

The median age of the study population was 41 years (IQR [31-54]) during the Delta period and 41 years (IQR [30-56]) during the Omicron-period. During both VOC periods, the majority of the study population was composed of females (53% and 58% during Delta and Omicron periods, respectively) and Belgian natives (65% and 64% during Delta and Omicron periods, respectively). During both VOC periods, most individuals had moderate or high education (68% and 66% during Delta and Omicron periods, respectively), a primary or first-booster vaccination (73% and 81% during Delta and Omicron periods, respectively), no underlying health conditions (69% and 68% during Delta and Omicron periods, respectively), and did not work in healthcare (92% and 92% during Delta and Omicron periods, respectively).

**Table 16** also provides the characterization of individuals with recorded hospitalizations for COVID-19 in the CHS database by VOC period. During Delta and Omicron periods, 1,059 (0.71%) and 1,828 (0.43%) individuals were recorded as hospitalized for COVID-19, respectively. The median age of COVID-19 patients was 66 years (IQR [53-76.5]) during the Delta period and 75 years (IQR [64-84]) during

the Omicron period. During both VOC periods, the majority of recorded COVID-19 hospitalized patients was males (55% and 51% during Delta and Omicron periods, respectively), Belgian natives (59% and 72% during Delta and Omicron periods, respectively), low educated (56% and 66% during Delta and Omicron periods, respectively), had a primary or first-booster vaccination (58% and 86% during Delta and Omicron periods, respectively), at least one underlying health conditions (69% and 84% during Delta and Omicron periods, respectively), and did not work in healthcare (96% and 98% during Delta and Omicron periods, respectively).

**Table 16. Characterization of the study population, including all Belgian adults living in the Walloon or Brussel regions in Belgium with a SARS-CoV-2 positive test, stratified between the Delta-predominant period (1 September 2021 to 22 December 2021) and the Omicron-predominant period (4 January 2022 to 1 September 2022).**

| | Delta-predominant period | | Omicron-predominant period | |
|---|---|---|---|---|
| | **Total study population (N=148 590)** | **Hospitalized COVID-19 patients (N=1 059)** | **Total study population (N=430 693)** | **Hospitalized COVID-19 patients (N=1 828)** |
| **Covariates** | | | | |
| *Age (in years), median[IQR]* | 41 [31-54] | 66 [53-76.5] | 41 [30-56] | 75 [64-84] |
| *Gender, n(%)* | | | | |
| Females | 82 433 (55.48) | 476 (44.95) | 249 079 (57.83) | 902 (49.34) |
| Males | 66 157 (44.52) | 583 (55.05) | 181 614 (42.17) | 926 (50.66) |
| *Migration background, n(%)* | | | | |
| Belgian natives | 97 104 (65.35) | 618 (58.36) | 275 656 (64.00) | 1 322 (72.32) |
| Second-generation migrants | 15 906 (10.70) | 48 (4.53) | 50 473 (11.72) | 73 (4.00) |
| First-generation European migrants | 17 888 (12.04) | 205 (19.36) | 51 631 (11.99) | 299 (16.36) |
| First-generation non-European migrants | 17 692 (11.91) | 188 (17.75) | 52 933 (12.29) | 134 (7.33) |
| *Province, n(%)* | | | | |
| Walloon-Brabant | 14 793 (9.96) | 102 (9.63) | 43 560 (10.11) | 179 (9.79) |
| Hainaut | 38 600 (25.98) | 442 (41.74) | 130 582 (30.32) | 965 (52.79) |
| Liege | 39 917 (26.86) | 127 (12.00) | 93 751 (21.77) | 219 (11.98) |
| Luxemburg | 8 359 (5.63) | 51 (4.82) | 23 612 (5.48) | 56 (3.06) |
| Namur | 18 153 (12.22) | 92 (8.69) | 50 693 (11.77) | 145 (7.93) |
| Brussels | 28 768 (19.36) | 245 (23.14) | 88 495 (20.55) | 264 (14.44) |

**Table 16. Continued**

| | Delta-predominant period | | Omicron-predominant period | |
|---|---|---|---|---|
| | **Total study population (N=148 590)** | **Hospitalized COVID-19 patients (N=1 059)** | **Total study population (N=430 693)** | **Hospitalized COVID-19 patients (N=1 828)** |
| **Exposure** | | | | |
| Education, n(%) | | | | |
| Moderate or high | 100 873 (67.89) | 462 (43.63) | 282 797 (65.66) | 616 (33.70) |
| Low | 47 717 (32.11) | 597 (56.37) | 147 896 (34.34) | 1 212 (66.30) |
| **Mediators** | | | | |
| *Vaccination status, n(%)* | | | | |
| No vaccination | 40 727 (27.41) | 445 (42.02) | 80 456 (18.68) | 264 (14.44) |
| Primary or first-booster vaccination | 107 863 (72.59) | 614 (57.98) | 350 237 (81.32) | 1 564 (85.56) |
| *Underlying health conditions, n(%)* | | | | |
| No underlying health conditions | 103 036 (69.34) | 325 (30.69) | 292 170 (67.84) | 294 (16.08) |
| At least one underlying health condition | 45 554 (30.66) | 734 (69.31) | 138 523 (32.16) | 1 534 (83.92) |
| *Working in healthcare , n(%)* | | | | |
| No | 136 066 (91.57) | 1 017 (96.03) | 394 466 (91.59) | 1 784 (97.59) |
| Yes | 12 524 (8.43) | 42 (3.97) | 36 227 (8.41) | 44 (2.41) |

3.2. Mediation analyses using interventional (in)direct effect models

**Tables 17** and **18** shows the interventional (in)direct effect estimates, bootstrap standard errors (SEs), and 95% confidence intervals (CIs) for the mediating effect of vaccination (M1), underlying health conditions (M2), and working in healthcare (M3) on the association between education and COVID-19 hospitalization for Delta- and Omicron-predominant period, respectively, along with mediation path diagrams and interventional effects definition.

The estimated total effect (TE), corresponding to the sum of the direct effect (DE) of education on COVID-19 hospitalization and the indirect effects (IE) through each mediator, of shifting from moderate or high education to low education increased the probability of being hospitalized for COVID-19 by 0.653% (95% CI 0.546% – 7.619%) and 0.517% (95% CI 0.469% - 0.565%) during Delta and Omicron periods, respectively.

The estimated interventional direct effect (DE) of shifting from moderate or high education to low education, when holding all mediators under the control conditions, increased the probability of being hospitalized for COVID-19 by 0.554% (95% CI 0.449% - 0.662%) and 0.484% (95% CI 0.437% - 0.529%) during Delta and Omicron periods, respectively. Proportionally, this means that the DE of education on COVID-19 hospitalization accounted for 84.84% ($DE_{Delta}$ / $TE_{Delta}$) and 93.62% ($DE_{Omicron}$ / $TE_{Omicron}$) of the total effect during Delta and Omicron periods, respectively.

The joint indirect effect (JIE) accounted for the remaining proportions and therefore implied that 15.16% (JIE $_{\text{Delta}}$ / TE $_{\text{Delta}}$) and 6.38% (JIE $_{\text{Omicron}}$ / TE $_{\text{Omicron}}$) of the association between education and COVID-19 hospitalization was explained by the total mediating effect through all mediators, during Delta and Omicron periods, respectively.

The estimated indirect effect of shifting the distribution of primary or first-booster vaccination under moderate or high education to low education ($\text{IE}_{\text{M1}}$), increased the probability of being hospitalized by 0.062% (95% CI 0.054% - 0.071%) and 0.006% (95% CI 0.005% - 0.008%) during Delta and Omicron periods, respectively.

The estimated indirect effect of shifting the distribution of underlying health conditions under moderate or high education to low education ($\text{IE}_{\text{M2}}$), increased the probability of being hospitalized by 0.038% (95% CI 0.029% - 0.046%) and 0.021% (95% CI 0.018% - 0.025%) during Delta and Omicron periods, respectively.

The estimated indirect effect via working in healthcare ($\text{IE}_{\text{M3}}$) was not significant during the Delta period (-0.001%, 95% CI -0.012% - 0.008%). However, during the Omicron period, the estimated indirect effect of shifting the distribution of working in healthcare under moderate or high education to low education, increased the probability of being hospitalized by 0.006% (95% CI 0.002% - 0.009%).

Proportionally, this means that, during the Delta period, vaccination and underlying health conditions significantly mediated the association between education and COVID-19 hospitalization by 9.49% ($IE_{M1, Delta}$ / $TE_{Delta}$) and 5.82% ($IE_{M2, Delta}$ / $TE_{Delta}$), respectively. During the Omicron period, underlying health conditions, vaccination, and working in healthcare significantly mediated the association between education and COVID-19 hospitalization by 4.07% ($IE_{M2, Omicron}$ / $TE_{Omicron}$), 1.16% ($IE_{M1, Omicron}$ / $TE_{Omicron}$), and 0.97% ($IE_{M3, Omicron}$ / $TE_{Omicron}$), respectively.

**Table 17. Interventional (in)direct effect estimates, bootstrap standard errors (SEs), and 95% confidence intervals (CIs) for the mediating effect of vaccination, underlying health conditions, and working in healthcare on the association between education and COVID-19 hospitalization for Delta-predominant period, along with mediation path diagrams and interventional effects definition.**

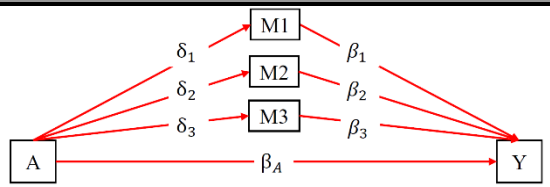
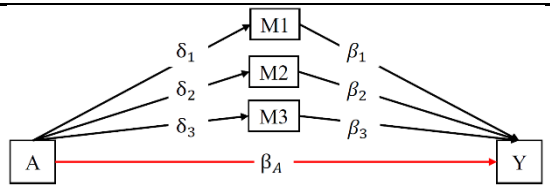
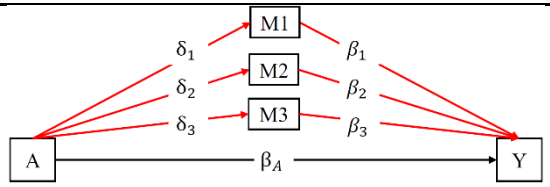
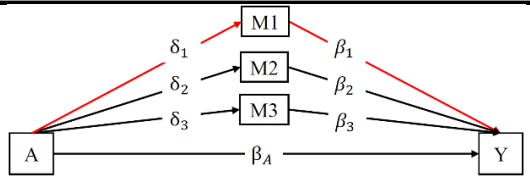
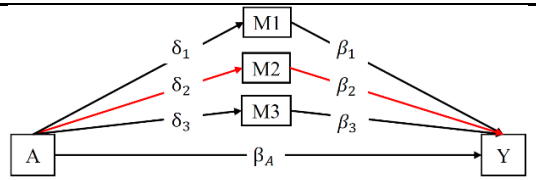
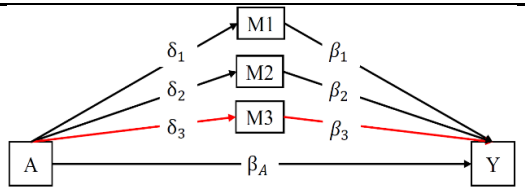
| Interventional effects | Estimate | Bootstrap SE | 95% CI | Mediation path diagram | Interventional effects definition |
|---|---|---|---|---|---|
| Total effect (TE) | 0.00653 | 0.00054 | 0.00546 - 0.07619 |  | $TE = \beta_A + \sum_{s=1}^{3} \delta_s \beta_s$ |
| Direct effect (DE) | 0.00554 | 0.00054 | 0.00449 - 0.00662 |  | $DE = \beta_A = TE - \sum_{s=1}^{3} \delta_s \beta_s$ |
| Joint indirect effect (JIE) | 0.00099 | 0.00008 | 0.00082 - 0.00115 |  | $JIE = \sum_{s=1}^{3} \delta_s \beta_s = TE - \beta_A$ |

**Table 17. Continued**

| Interventional effects | Estimate | Bootstrap SE | 95% CI | Mediation path diagram | Interventional effects definition |
|---|---|---|---|---|---|
| Indirect effect via M1 (vaccination) ($IE_{M1}$) | 0.00062 | 0.00005 | 0.00054 - 0.00071 |  | $IE_{M1} = \delta_1 \beta_1$ |
| Indirect effect via M2 (underlying health conditions) ($IE_{M2}$) | 0.00038 | 0.00005 | 0.00029 - 0.00046 |  | $IE_{M2} = \delta_2 \beta_2$ |
| Indirect effect via M3 (working in healthcare) ($IE_{M3}$) | -0.00001 | 0.00005 | -0.00012 - 0.00008 |  | $IE_{M3} = \delta_3 \beta_3$ |

*Interventional (in)direct effect models are adjusted for age, gender, province, and migration background. The numbering of the mediators is arbitrary and does not imply an order of causality between the multiple mediators. Standard error and 95% CI were constructed using 500 non-parametric bootstrap samples. 'A' denotes the exposure (i.e., education). 'Y' denotes the outcome (i.e., COVID-19 hospitalization). Red lines denote causal paths through which (in)direct interventional effects pass.*

**Table 18. Interventional (in)direct effect estimates, bootstrap standard errors (SEs), and 95% confidence intervals (CIs) for the mediating effect of vaccination, underlying health conditions, and working in healthcare on the association between education and COVID-19 hospitalization for Omicron-predominant period, along with mediation path diagrams and interventional effects definition.**

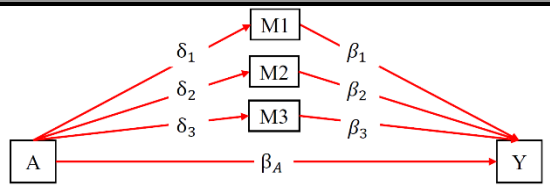
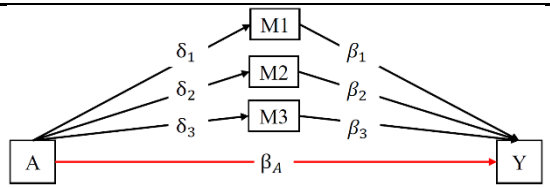
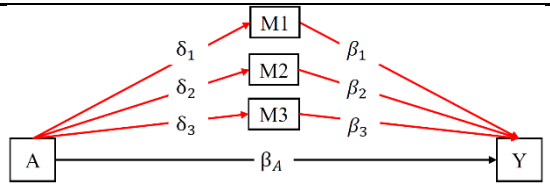
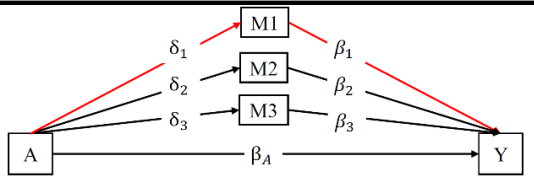
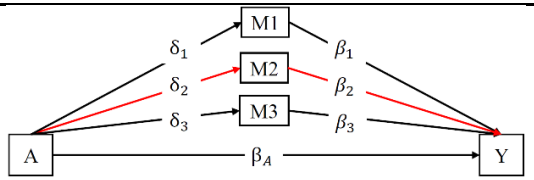
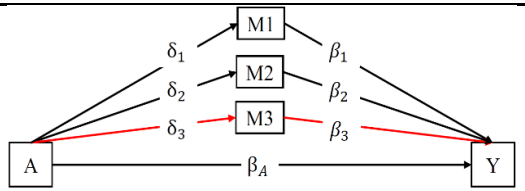
| Interventional effects | Estimate | Bootstrap SE | 95% CI | Mediation path diagram | Interventional effects definition |
|---|---|---|---|---|---|
| Total effect (TE) | 0.00516 | 0.00024 | 0.00469 - 0.00565 |  | $TE = \beta_A + \sum_{s=1}^{3} \delta_s \beta_s$ |
| Direct effect (DE) | 0.00484 | 0.00027 | 0.00437 - 0.00529 |  | $DE = \beta_A = TE - \sum_{s=1}^{3} \delta_s \beta_s$ |
| Joint indirect effect (JIE) | 0.00032 | 0.00003 | 0.00028 - 0.00039 |  | $JIE = \sum_{s=1}^{3} \delta_s \beta_s = TE - \beta_A$ |

**Table 18. Continued**

| Interventional effects | Estimate | Bootstrap SE | 95% CI | Mediation path diagram | Interventional effects definition |
|---|---|---|---|---|---|
| Indirect effect via M1 (vaccination) ($IE_{M1}$) | 0.00006 | 0.00001 | 0.00005 - 0.00008 |  | $IE_{M1} = \delta_1 \beta_1$ |
| Indirect effect via M2 (underlying health conditions) ($IE_{M2}$) | 0.00021 | 0.00002 | 0.00018 - 0.00025 |  | $IE_{M2} = \delta_2 \beta_2$ |
| Indirect effect via M3 (working in healthcare) ($IE_{M3}$) | 0.00005 | 0.00002 | 0.00002 - 0.00010 |  | $IE_{M3} = \delta_3 \beta_3$ |

*Interventional (in)direct effect models are adjusted for age, gender, province, and migration background. The numbering of the mediators is arbitrary and does not imply an order of causality between the multiple mediators. Standard error and 95% CI were constructed using 500 non-parametric bootstrap samples. 'A' denotes the exposure (i.e., education). 'Y' denotes the outcome (i.e., COVID-19 hospitalization). Red lines denote causal paths through which (in)direct interventional effects pass.*

### 3.3. Sensitivity Analyses

A first sensitivity analysis was computed by including exposure-mediator, mediator-mediator, and exposure-mediator-mediator interactions in the outcome model. The resulting estimated interventional (in)direct effects and 95% CI showed similar trends compared to the main analysis, except for the estimated interventional indirect effect via working in healthcare during the Omicron period which is no longer significant (**Supplementary Table 11**). A second sensitivity analysis was computed accounting for the variation of the indirect effect through each mediator due to the labelling assigned to the mediators, presumed to be arbitrary. The results obtained from analyzing each of all possible permutations of mediator indices (3!=6) yielded similar estimates for both the maximum and minimum values of the (in)direct interventional effects (**Supplementary Table 12**).

The second sensitivity analyses assessing whether a potential bias due to outcome-dependent sampling was present, yielded results consistent with the primary analysis, both without interaction (**Supplementary Tables 13** and **14**) and with exposure-mediator, mediator-mediator, and exposure-mediator-mediator interactions (**Supplementary Tables 15** and **16**). These findings suggest that outcome-dependent test selection did not substantially affect our results.

## 4. Discussion

Overall, 1,059 (0.71%) and 1,828 (0.43%) individuals with a confirmed SARS-CoV-2 infection were recorded as hospitalized for COVID-19 during Delta- and Omicron-predominant periods, respectively, in the Clinical Hospital Survey database, a non-exhaustive reporting system in Belgium. During both VOC-predominant periods, the majority of hospitalized COVID-19 patients had low education (56.37% and 66.30% for Delta and Omicron periods, respectively). In comparison, 32% of the total Belgian adult population (aged 25 years and older) had a low educational level (up to lower secondary education) in 2021 [261]. This highlights the overrepresentation of adults with a low educational level among COVID-19 hospitalized patients compared to the general adult population in Belgium.

By using interventional (in)direct effect models, our findings revealed a higher probability of being hospitalized for COVID-19 among individuals with low education compared to those with moderate or high education, with an increase of 0.653% and 0.517% during Delta and Omicron periods, respectively. During both VOC periods, the association between education and COVID-19 hospitalization was partially mediated by vaccination (9.49% and 1.16% during Delta and Omicron periods, respectively) and underlying health conditions (5.82% and 4.06% during Delta and Omicron periods, respectively). During the Delta period, working in healthcare did not significantly mediate the association between education and

COVID-19 hospitalization. During the Omicron period, working in healthcare significantly mediated the association between education and COVID-19 hospitalization, however, when taking into account exposure-mediator, mediator-mediator, and exposure-mediator-mediator interactions, the interventional indirect effect via working in healthcare was no longer significant.

Our findings showed a higher rate of COVID-19 hospitalization among individuals with low education, aligning with several international studies [266–272]. Among all confirmed COVID-19 cases in Sweden by mid-September 2020, Bergman et al. found that, compared to individuals with primary education, the probability of hospitalization decreased by 13% for those with secondary education, 21% for those with post-secondary education (<3 years), and 27% for those with post-secondary education (≥3 years) [270]. Similarly, in Switzerland, individuals residing in neighborhoods with the highest SE position had a 44% lower risk of COVID-19 hospitalization compared to those in the lowest SE position [271]. At the European level, a two-sample Mendelian randomization study showed a 0.540 higher odds of severe COVID-19 per one standard deviation increase in years of schooling [272].

The mediation role of chronic comorbidities in overall respiratory infectious diseases has already been shown by Ye et al. reporting that cardiovascular diseases and diabetes mediated the association between SE status and respiratory infectious diseases by 6.8% and 2.2%, respectively [273]. This is in line with our results

showing a mediating effect of underlying health conditions on the association between education and COVID-19 hospitalization. Research on SE disparities in COVID-19 hospitalization, including comorbidities as clinical confounders, found that having comorbidities (e.g. diabetes, cardiovascular diseases, obesity, chronic obstructive pulmonary disorders, renal failure) largely increased the probability of being hospitalized for COVID-19 [268–270]. To our knowledge, the extent to which comorbidities mediate this association has not been investigated yet. However, the association of comorbidities with both SE status and COVID-19 hospitalization is well known, making this mediating role plausible. Indeed, individuals from lower SE groups present higher rates of comorbidities [103] due to multiple underlying reasons such as a decreased access to healthcare resources and services or unhealthy lifestyle habits (e.g. smoking, alcohol consumption, lower physical activity) [274]. Having comorbidities has also been shown to increase the risk of COVID-19 hospitalization [268–270] through several biological mechanisms such as an excessive pro-inflammatory response or a reduced tolerance to respiratory failure [124].

Similar mediation mechanisms can be hypothesized regarding vaccination in light of its known association with both SE status and COVID-19 hospitalization [64,232,244,275,276]. A lower COVID-19 vaccine uptake was found among SE disadvantaged groups [232,244,276]. This may be due to a lack of belief in vaccine

efficacy and a fear of side effects [235], a restricted access to healthcare and resources [232], or a lack of health literacy exacerbating fear and negative preconceptions about the vaccine [233]. However, the vaccine has been shown to be effective against COVID-19 hospitalization [64,277]. In the UK, Gram et al. reported a vaccine effectiveness (VE) against COVID-19 hospitalization 14 to 30 days since last vaccination among those who received two or three doses of mRNA vaccine of 98.1% or above for Delta and 95.5% or above for Omicron [276]. In Belgium, for the Delta variant, VE against COVID-19 hospitalization 0 to 50 days since last vaccination was 93% and 94% for primary and booster vaccination, respectively. For the Omicron variant, the VE was 59% and 87% for primary and booster vaccination, respectively [64]. While the specific outcome may vary, a previous study conducted in the United States has also shown evidence of the mediating effect of vaccination on the association between SE status and severe COVID-19 outcomes by demonstrating that 37.6% of the impact of SE status on COVID-19 mortality was mediated by the vaccination coverage rate [278]. This highlights the importance of variations in vaccination coverage in exacerbating social inequalities in COVID-19 severe outcomes, and advocates achieving the most equitable vaccination coverage possible.

Our results showed that vaccination mediated 9.49% of the association between education and COVID-19 hospitalization during the Delta period while this

percentage dropped to 1.16% during the Omicron period. This difference in the mediating effect of the vaccine by VOC period could be due to a higher VE against severe SARS-CoV-2 infection in individuals infected with Delta compared to Omicron combined with the development of a less severe form of SARS-CoV-2 infection among individuals infected with Omicron in comparison to Delta [59,60,64,279–281]. Waning immunity from vaccines over time, along with natural immunity from prior SARS-CoV-2 infections, could also influence the reduced efficacy of the vaccines against COVID-19 hospitalization during the Omicron period compared to the Delta period, and explain the lower mediating effect of the vaccine during the Omicron period [282,283].

To the best of our knowledge, this is the first study investigating the mediation role of vaccination, underlying health conditions, and working in healthcare in the association between education and COVID-19 hospitalization using interventional effect models allowing to consider multiple mediators without assuming any causal structure between them.

However, our study has several limitations. First, the CHS database is not exhaustive and recorded only 59.3% and 42.3% of the total COVID-19 hospital admissions during Delta and Omicron periods, respectively, leading to a sampling bias and significant underestimations of hospitalization rates. In addition to the non-exhaustive nationwide hospital coverage, variation in hospital reporting rate were

found by province. To counter this latter sampling bias, we used the province of residence as confounders but an additional limitation lies under the assumption that an individual was hospitalized in her or his province of residence which may not be true. However, we have assumed that the variation in the hospital reporting rate did not depend on the SE status of the province, which would have further increased the sampling bias. Indeed, during our study period, the province of Hainaut had the highest reporting rate (±70%) while the province of Liege had the lowest (±15%). Average annual tax income per capita in 2020 was higher in the province of Liege (18,233 euros) compared to the province of Hainaut (17,118 euros) [259] and the unemployment rate in 2022 was 10.6% for the province of Hainaut and 8.4% for the province of Liege [260].

Second, a proxy based on the date of diagnosis was used to determine whether the infection was due to Delta or Omicron variant due to the few sequenced-laboratory confirmed infection during the study period. As such, certain infections will undoubtedly have been misclassified, but as we assume the required proportion to consider a period as dominant for a certain variant to be 80%, we expect the impact of this to be limited.

Third, to define the mediator 'Working in healthcare', we used a proxy variable identifying individuals with a degree in health, the rationale being that individuals working in healthcare are expected to have a higher level of health literacy that we

hypothesized to lie on the causal pathway between education and COVID-19 hospitalization. However, having a healthcare degree does not necessarily imply working in the healthcare sector or applying the health knowledge acquired during training. This is therefore a crude measure which may have impacted our results. In addition, other factors than the possession of a degree in healthcare contribute to health knowledge.

Fourth, the mediators included in our model were not sufficient to explain the entire association between education and COVID-19 hospitalization, the direct effect of education remaining significant. We were unable to consider certain other unknown factors that could play an additional mediating effect between education and COVID-19 hospitalization. Future research could consider investigating other factors (e.g., living and working conditions, access to healthcare, cultural practices and beliefs, mental health) that may mediate the association between SE determinants and COVID-19 hospitalization.

## 5. Conclusion

Underlying health conditions and vaccination partially mediate the relationship between educational level and COVID-19 hospitalization rates. These findings hold important implications for crisis management, indicating that, to reduce socioeconomic disparities in COVID-19 related hospitalizations, public health strategies should focus on preventing underlying health conditions and enhancing

vaccination rates by promoting health education, especially within the lowest socioeconomic groups, to increase vaccine knowledge and confidence.

## 6. Declarations

### 6.1. Ethics approval and consent to participate

The protocol of the LINK-VACC project was approved by the Medical Ethics Committee from the Vrije Universiteit of Brussels on 3 February 2021 (B.U.N 1432020000371) and obtained authorization from the Information Security Committee (ISC) Social Security and Health (reference number: IVC/KSZG/21/034).

As confirmed by sections 23 and 24 of the *Guidelines 03/2020 on the processing of data concerning health for the purpose of scientific research in the context of the COVID-19 outbreak* of the European Data Protection Board (V1.0 of 21 April 2020), this survey falls under Article 6 §1(e) and Article 9 §2(i) of the General Data Protection Regulation (GDPR). In appliance with these GDPR legal grounds of data processing, no informed consent had to be signed by the patients. Clinical trial number: not applicable

### 6.2. Consent for publication

Not applicable

### 6.3. Availability of data and materials

The data used in this study are pseudonymized individual data obtained from several national registers hosted by Healthdata.be, an independent institution within Sciensano aiming to bring together all the data currently stored in multiple health registers on a single web-based platform. Due to the General Data Protection Regulation (GDPR) legislations in Belgium, these data are not publicly available. Data request must be addressed to the Information Security Committee (ISC).

### 6.4. Competing interests

The authors have no competing interests to declare that are relevant to the content of this article.

### 6.5. Funding

This study was supported by the FOD/SPF Public Health of Belgium and the Belgian federated entities through funding for the LINK-VACC project on vaccine surveillance. In addition, Belgian Science Policy Office (BELSPO) funded this research within the BRAIN-be 2.0 framework supporting pillar 3 Federal societal challenges (grant number B2/202/P3/HELICON). The sponsors did not have a role in the conduction of the study, the writing of the manuscript, or in the decision to submit it for publication.

### 6.6. Authors' contributions

LC reviewed the literature; LC, BD, JG, J A.F. L, LB, NS, and RP conceived the study; LC, BD, JG, J A.F. L, LB, NS, and RP selected the study population; LC, BM, TL, and RP designed the statistical methodology; LC and RP conducted the statistical analyses; LC, BM, TL, and RP interpret the findings; LC wrote the first draft of the paper. All authors revised the text. All authors approved the final version of this manuscript and accepted responsibility for its submission for publication.

### 6.7. Acknowledgments

We would like to thanks the Walloon and Brussels regions who gave us the authorization to use the LINKVACC data to perform this study.

## 7. Supplementary materials

### 7.1. Table of content

#### *7.1.1. Supplementary Tables*

- **Supplementary Table 11.** Interventional (in)direct effect estimates, bootstrap standard errors (SEs), and 95% confidence intervals (CIs) for the mediating effect of vaccination, underlying health conditions, and working in healthcare on the association between education and COVID-19 hospitalization accounting for exposure-mediator, mediator-mediator, and exposure-mediator-mediator interactions. All results are stratified by VOC-predominant period.

- **Supplementary Table 12.** Minimum (Min.) and maximum (Max.) values of the interventional (in)direct effect estimates and lower and upper bounds of the 95% confidence interval (CIs) across all 3!=6 possible permutations of mediator indices stratified by VOC-predominant period.
- **Supplementary Table 13.** Interventional (in)direct effect estimates, bootstrap standard errors (SEs), and 95% confidence intervals (CIs) for the mediating effect of vaccination, underlying health conditions, and working in healthcare on the association between education and COVID-19 hospitalization for Delta-predominant period, along with mediation path diagrams and interventional effects definition, using a random positive test per person to address outcome-dependent sampling bias.
- **Supplementary Table 14.** Interventional (in)direct effect estimates, bootstrap standard errors (SEs), and 95% confidence intervals (CIs) for the mediating effect of vaccination, underlying health conditions, and working in healthcare on the association between education and COVID-19 hospitalization for Omicron-predominant period, along with mediation path diagrams and interventional effects definition, using a random positive test per person to address outcome-dependent sampling bias
- **Supplementary Table 15.** Interventional (in)direct effect estimates, bootstrap standard errors (SEs), and 95% confidence intervals (CIs) for the mediating effect of vaccination, underlying health conditions, and working in healthcare

on the association between education and COVID-19 hospitalization accounting for exposure-mediator, mediator-mediator, and exposure-mediator-mediator interactions stratified by VOC-predominant period, using a random positive test per person to address outcome-dependent sampling bias.

- **Supplementary Table 16.** Minimum (Min.) and maximum (Max.) values of the interventional (in)direct effect estimates and lower and upper bounds of the 95% confidence interval (CIs) across all 3!=6 possible permutations of mediator indices stratified by VOC-predominant period, using a random positive test per person to address outcome-dependent sampling bias.

**Supplementary Table 11. Interventional (in)direct effect estimates, bootstrap standard errors (SEs), and 95% confidence intervals (CIs) for the mediating effect of vaccination, underlying health conditions, and working in healthcare on the association between education and COVID-19 hospitalization accounting for exposure-mediator, mediator-mediator, and exposure-mediator-mediator interactions. All results are stratified by VOC-predominant period.**

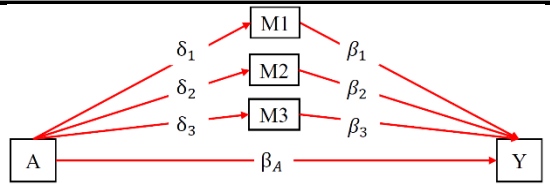
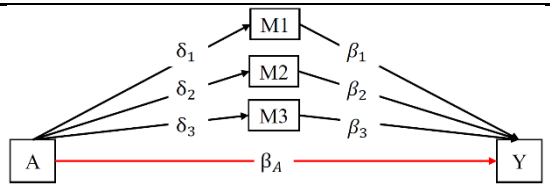
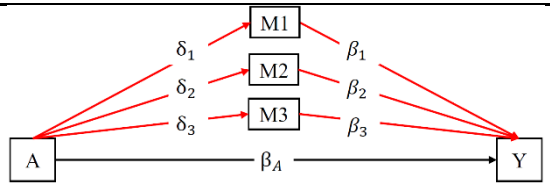
| | Delta-predominant period | | | Omicron-predominant period | | |
|---|---|---|---|---|---|---|
| **Interventional effect** | **Estimate** | **Bootstrap SE** | **95% CI** | **Estimate** | **Bootstrap SE** | **95%CI** |
| Total effect (TE) | 0.00653 | 0.00054 | 0.00551 - 0.00758 | 0.00517 | 0.00024 | 0.00469 - 0.00562 |
| Direct effect (DE) | 0.00509 | 0.00052 | 0.00415 - 0.00611 | 0.00435 | 0.00023 | 0.00388 - 0.00477 |
| Indirect effect via M1 (vaccination) ($IE_{M1}$) | 0.00097 | 0.00010 | 0.00078 - 0.00119 | 0.00011 | 0.00003 | 0.00007 - 0.00016 |
| Indirect effect via M2 (underlying health conditions) ($IE_{M2}$) | 0.00091 | 0.00010 | 0.00071 - 0.00112 | 0.00058 | 0.00004 | 0.00049 - 0.00067 |
| Indirect effect via M3 (working in healthcare) ($IE_{M3}$) | 0.00010 | 0.00021 | -0.00033 - 0.00050 | 0.00014 | 0.00009 | -0.00006 - 0.00032 |
| Indirect effect via mutual dependence | -0.00057 | 0.00023 | -0.00099 - -0.00011 | -0.00014 | 0.00007 | -0.00027 - 0.00003 |

*Interventional (in)direct effect models are adjusted for age, gender, province, and migration background. The numbering of the mediator is arbitrary and does not imply an order of causality between the multiple mediators. Standard error and 95% CI were constructed using 500 non-parametric bootstrap samples.*

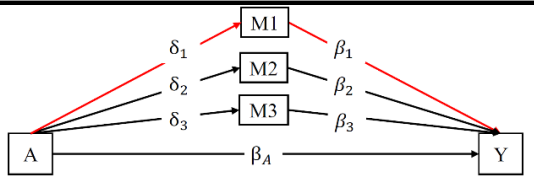
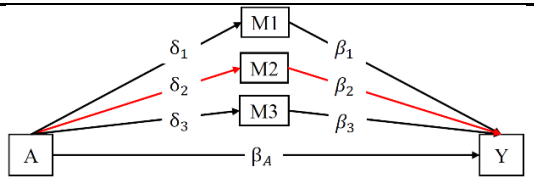
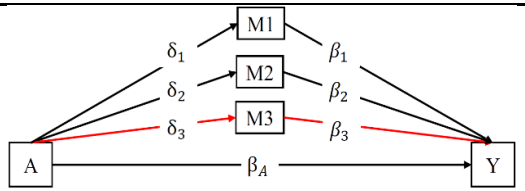
**Supplementary Table 12. Minimum (Min.) and maximum (Max.) values of the interventional (in)direct effect estimates and lower and upper bounds of the 95% confidence interval (CIs) across all 3!=6 possible permutations of mediator indices stratified by VOC-predominant period.**

| | Delta-predominant period | | | | Omicron-predominant period | | | |
|---|---|---|---|---|---|---|---|---|
| **Interventional indirect effect** | **Min.** | **Max.** | **Min. (lower)** | **Max. (upper)** | **Min.** | **Max.** | **Min. (lower)** | **Max. (upper)** |
| Vaccination | 0.00081 | 0.00099 | 0.00065 | 0.00119 | 0.00009 | 0.00011 | 0.00005 | 0.00016 |
| Underlying health conditions | 0.00079 | 0.00105 | 0.00059 | 0.00127 | 0.00055 | 0.00059 | 0.00045 | 0.00067 |
| Working in healthcare | 0.00010 | 0.00028 | -0.00033 | 0.00072 | 0.00014 | 0.00019 | -0.00007 | 0.00041 |

**Supplementary Table 13. Interventional (in)direct effect estimates, bootstrap standard errors (SEs), and 95% confidence intervals (CIs) for the mediating effect of vaccination, underlying health conditions, and working in healthcare on the association between education and COVID-19 hospitalization stratified for Delta-predominant period, along with mediation path diagrams and interventional effects definition, using a random positive test per person to address outcome-dependent sampling bias.**

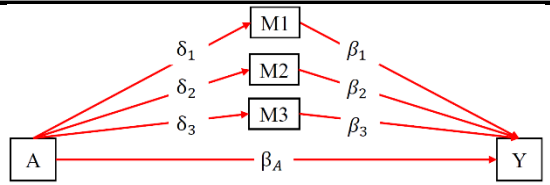
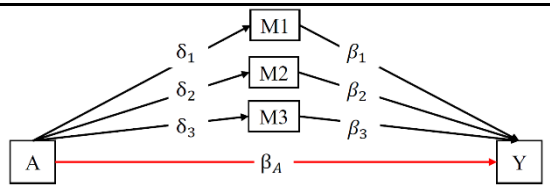
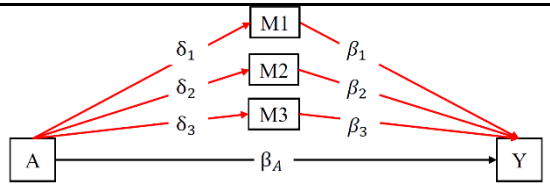
| Interventional effects | Estimate | Bootstrap SE | 95% CI | Mediation path diagram | Interventional effects definition |
|---|---|---|---|---|---|
| Total effect (TE) | 0.00694 | 0.00066 | 0.00561 - 0.00819 |  | $TE = \beta_A + \sum_{s=1}^{3} \delta_s \beta_s$ |
| Direct effect (DE) | 0.00587 | 0.00068 | 0.00452 - 0.00722 |  | $DE = \beta_A = TE - \sum_{s=1}^{3} \delta_s \beta_s$ |
| Joint indirect effect (JIE) | 0.00109 | 0.00009 | 0.00090 - 0.00128 |  | $JIE = \sum_{s=1}^{3} \delta_s \beta_s = TE - \beta_A$ |

**Supplementary Table 13. Continued**

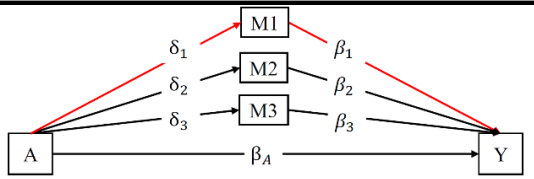
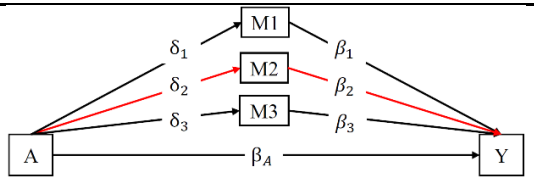
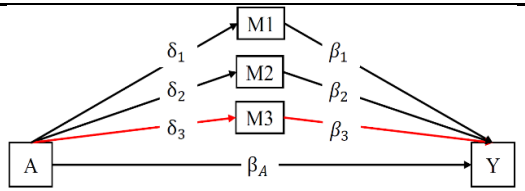
| Interventional effects | Estimate | Bootstrap SE | 95% CI | Mediation path diagram | Interventional effects definition |
|---|---|---|---|---|---|
| Indirect effect via M1 (vaccination) ($IE_{M1}$) | 0.00070 | 0.00006 | 0.00059 - 0.00081 |  | $IE_{M1} = \delta_1 \beta_1$ |
| Indirect effect via M2 (underlying health conditions) ($IE_{M2}$) | 0.00044 | 0.00006 | 0.00032 - 0.00057 |  | $IE_{M2} = \delta_2 \beta_2$ |
| Indirect effect via M3 (working in healthcare) ($IE_{M3}$) | -0.00005 | 0.00005 | -0.00015 - 0.00090 |  | $IE_{M3} = \delta_3 \beta_3$ |

*Interventional (in)direct effect models are adjusted for age, gender, province, and migration background. The numbering of the mediators is arbitrary and does not imply an order of causality between the multiple mediators. Standard error and 95% CI were constructed using 500 non-parametric bootstrap samples. 'A' denotes the exposure (i.e., education). 'Y' denotes the outcome (i.e., COVID-19 hospitalization). Red lines denote causal paths through which (in)direct interventional effects pass.*

**Supplementary Table 14. Interventional (in)direct effect estimates, bootstrap standard errors (SEs), and 95% confidence intervals (CIs) for the mediating effect of vaccination, underlying health conditions, and working in healthcare on the association between education and COVID-19 hospitalization stratified for Omicron-predominant period, along with mediation path diagrams and interventional effects definition, using a random positive test per person to address outcome-dependent sampling bias.**

| Interventional effects | Estimate | Bootstrap SE | 95% CI | Mediation path diagram | Interventional effects definition |
|---|---|---|---|---|---|
| Total effect (TE) | 0.00547 | 0.00028 | 0.00491 - 0.00603 |  | $TE = \beta_A + \sum_{s=1}^{3} \delta_s \beta_s$ |
| Direct effect (DE) | 0.00509 | 0.00029 | 0.00450 - 0.00529 |  | $DE = \beta_A = TE - \sum_{s=1}^{3} \delta_s \beta_s$ |
| Joint indirect effect (JIE) | 0.00038 | 0.00003 | 0.00032 - 0.00045 |  | $JIE = \sum_{s=1}^{3} \delta_s \beta_s = TE - \beta_A$ |

**Supplementary Table 14. Continued**

| Interventional effects | Estimate | Bootstrap SE | 95% CI | Mediation path diagram | Interventional effects definition |
|---|---|---|---|---|---|
| Indirect effect via M1 (vaccination) ($IE_{M1}$) | 0.00008 | 0.00001 | 0.00006 - 0.00010 |  | $IE_{M1} = \delta_1 \beta_1$ |
| Indirect effect via M2 (underlying health conditions) ($IE_{M2}$) | 0.00027 | 0.00002 | 0.00022 - 0.00032 |  | $IE_{M2} = \delta_2 \beta_2$ |
| Indirect effect via M3 (working in healthcare) ($IE_{M3}$) | 0.00004 | 0.00002 | 0.00000 - 0.00007 |  | $IE_{M3} = \delta_3 \beta_3$ |

*Interventional (in)direct effect models are adjusted for age, gender, province, and migration background. The numbering of the mediators is arbitrary and does not imply an order of causality between the multiple mediators. Standard error and 95% CI were constructed using 500 non-parametric bootstrap samples. 'A' denotes the exposure (i.e., education). 'Y' denotes the outcome (i.e., COVID-19 hospitalization). Red lines denote causal paths through which (in)direct interventional effects pass.*

**Supplementary Table 15. Interventional (in)direct effect estimates, bootstrap standard errors (SEs), and 95% confidence intervals (CIs) for the mediating effect of vaccination, underlying health conditions, and working in healthcare on the association between education and COVID-19 hospitalization accounting for exposure-mediator, mediator-mediator, and exposure-mediator-mediator interactions stratified by VOC-predominant period, using a random positive test per person to address outcome-dependent sampling bias.**

| | Delta-predominant period | | | Omicron-predominant period | | |
|---|---|---|---|---|---|---|
| **Interventional effect** | **Estimate** | **Bootstrap SE** | **95% CI** | **Estimate** | **Bootstrap SE** | **95%CI** |
| Total effect (TE) | 0.00695 | 0.00063 | 0.00572 - 0.00812 | 0.00547 | 0.00027 | 0.00493 - 0.00599 |
| Direct effect (DE) | 0.00541 | 0.00056 | 0.00439 - 0.00645 | 0.00454 | 0.00024 | 0.00404 - 0.00503 |
| Indirect effect via M1 (vaccination) ($IE_{M1}$) | 0.00108 | 0.00013 | 0.00086 - 0.00133 | 0.00011 | 0.00003 | 0.00005 - 0.00077 |
| Indirect effect via M2 (underlying health conditions) ($IE_{M2}$) | 0.00107 | 0.00014 | 0.00081 - 0.00135 | 0.00067 | 0.00005 | 0.00056 - 0.00077 |
| Indirect effect via M3 (working in healthcare) ($IE_{M3}$) | 0.00002 | 0.00014 | -0.00028 - 0.00027 | 0.00009 | 0.00007 | -0.00005 - 0.00022 |
| Indirect effect via mutual dependence | -0.00086 | 0.00018 | -0.00123 - -0.00053 | -0.00004 | 0.00005 | -0.00015 - 0.00007 |

*Interventional (in)direct effect models are adjusted for age, gender, province, and migration background. The numbering of the mediator is arbitrary and does not imply an order of causality between the multiple mediators. Standard error and 95% CI were constructed using 500 non-parametric bootstrap samples.*

**Supplementary Table 16. Minimum (Min.) and maximum (Max.) values of the interventional (in)direct effect estimates and lower and upper bounds of the 95% confidence interval (CIs) across all 3!=6 possible permutations of mediator indices stratified by VOC-predominant period, using a random positive test per person to address outcome-dependent sampling bias.**

| | Delta-predominant period | | | | Omicron-predominant period | | | |
|---|---|---|---|---|---|---|---|---|
| **Interventional indirect effect** | **Min.** | **Max.** | **Min. (lower)** | **Max. (upper)** | **Min.** | **Max.** | **Min. (lower)** | **Max. (upper)** |
| Vaccination | 0.00087 | 0.00108 | 0.00067 | 0.00134 | 0.00009 | 0.00011 | 0.00004 | 0.00017 |
| Underlying health conditions | 0.00094 | 0.00127 | 0.00065 | 0.00162 | 0.00059 | 0.00069 | 0.00048 | 0.00079 |
| Working in healthcare | 0.00001 | 0.00023 | -0.00027 | 0.00050 | 0.00009 | 0.00019 | -0.00006 | 0.00034 |

# *Chapter 3. Social health inequalities in COVID-19 mortality*

***Adapted from***

**Lisa Cavillot**, Laura Van den Borre, Katrien Vanthomme, Aline Scohy, Patrick Deboosere, Brecht Devleesschauwer, Niko Speybroeck & Sylvie Gadeyne. **Unravelling demographic and socioeconomic patterns of COVID-19 death and all other causes of death: results of an individual-level analysis of exhaustive cause of death data in Belgium, 2020.** *Arch Public Health* 82, 209 (2024).

## Abstract

**Background** : the COVID-19 pandemic led to significant excess mortality in 2020 in Belgium. By using microlevel cause-specific mortality data for the total adult population in Belgium in 2020, three outcomes were considered in this study aiming at predicting sociodemographic (SD) and socioeconomic (SE) patterns of COVID-19 specific death and all other causes of death (OCOD), respectively.

**Methods** : two complementary statistical methods were used. First, multivariable logistic regression models providing odds ratios and 95% confidence intervals were fitted for the two study outcomes. In addition, we computed conditional inference tree (CIT) algorithms, a non-parametric class of classification trees, to identify and rank by significance level the strongest predictors of the two study outcomes.

**Results**: older individuals, males, individuals living in collectivities, first-generation migrants, and deprived SE groups experienced higher odds of dying from COVID-19; living in care homes was identified by the CIT as the strongest predictor followed by age and sex. Overall, similar patterns were observed for all OCOD except for first- and second-generation migrants having lower odds of all OCOD; age group was identified by the CIT as the strongest predictor.

**Conclusions**: this study identified important SD and SE disparities in COVID-19 mortality, with living in collectivities highlighted as the strongest predictor. This underlines the importance of implementing preventive measures, particularly within the most vulnerable populations, in infectious disease pandemic preparedness to reduce virus circulation and the resulting lethality.

## 1. Background

The Coronavirus Disease 2019 (COVID-19), caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has spread rapidly around the world, resulting in a global pandemic as announced by the WHO in March 2020 [284]. The COVID-19 pandemic had major health consequences. A systematic review including 191 countries identified that the pandemic had led to a global excess mortality rate of 120.3 deaths per 100,000 persons over the period from January 1st 2020 to December 31st 2021 [285]. In Belgium, 126,850 deaths were recorded in 2020 of which 19,801 were due to COVID-19 [25,286]. In the same year, 18,765 excess deaths were recorded leading to an excess mortality of 17.5% [33].

The COVID-19 pandemic was swiftly characterized as a syndemic pandemic, highlighting the interplay between the virus biological aspects and the social determinants of health, further intensifying existing social inequalities deeply rooted in our society [101,105]. Indeed, beyond age and sex, which are well-established risk factors for COVID-19 mortality [287,288], some international studies, based on either excess mortality or COVID-19 specific mortality, showed that ethnic minorities and deprived socioeconomic (SE) groups (e.g., with lower income and education levels, living in overcrowded households) experienced higher COVID-19 mortality or excess mortality during the pandemic [289–294]. In

Belgium, research based on excess mortality identified similar social patterns during the first COVID-19 wave [120,211,212,295].

Due to delays in obtaining cause-specific mortality data so far, analyses on the SD and SE patterns of COVID-19 specific mortality have not yet been published in Belgium. Our paper aims to bridge this void by offering a nuanced and deeper understanding of the SD and SE inequalities in COVID-19 mortality using nationwide microlevel cause-specific mortality data for the entire Belgian adult population. This population-based study will use individual cause-specific mortality data for 2020 in Belgium to predict the SD and SE patterns of COVID-19 specific deaths and all other causes of death (OCOD), respectively.

## 2. Materials and Methods

### 2.1. Data sources and study population

Pseudonymized individual level data for the entire Belgian population (over 11 million individuals) were provided by Statistics Belgium (Statbel) who performed a record linkage between four exhaustive data sources linked at the individual level using the Belgian social security number: 1) the Belgian national register providing yearly stock files including demographic indicators and mortality data for all individuals officially residing in Belgium as of 1st January of each year; 2) the administrative census 2011 providing data on education and housing; 3) the tax

register providing data on yearly net taxable personal income; and 4) death certificates providing data on causes of death.

As cause-specific mortality data on COVID-19 were available from January 1st 2020 to December 31st 2020 only, we limited our analysis to 2020, using the yearly stock file of 2020 and the individual cause-specific mortality data for that year. In order to determine the SE characteristics properly, the study population consisted of 8,254,632 adults aged 25 years and older officially residing in Belgium. A flowchart demonstrating the study population selection process is available in **Supplementary Figure 2**.

## 3. Variables

### 3.1. Mortality outcomes

Two mortality outcomes, occurring over the period from January 1st 2020 to December 31st 2020, were considered in our study: COVID-19 specific death and all other causes of death (OCOD).

The individual cause-specific mortality data, classified according to the International Standard Classification of Disease and Related Health Problems 10th Revisions (ICD-10), enabled the identification of COVID-19 specific death using the ICD-10 codes U07.1 (*COVID-19, virus identified*) and U07.2 (*COVID-19, virus not identified*). We considered as COVID-19 deaths, deaths with one of these two codes

as underlying cause of death. OCOD were identified by any ICD-10 codes not falling within the aforementioned COVID-19 specific ICD-10 categories.

**Figure 24** shows the daily number of COVID-19 deaths and all other causes of death during the year 2020 by age groups in Belgium. The first COVID-19 death recorded in Belgium occurred on 11th March 2020.

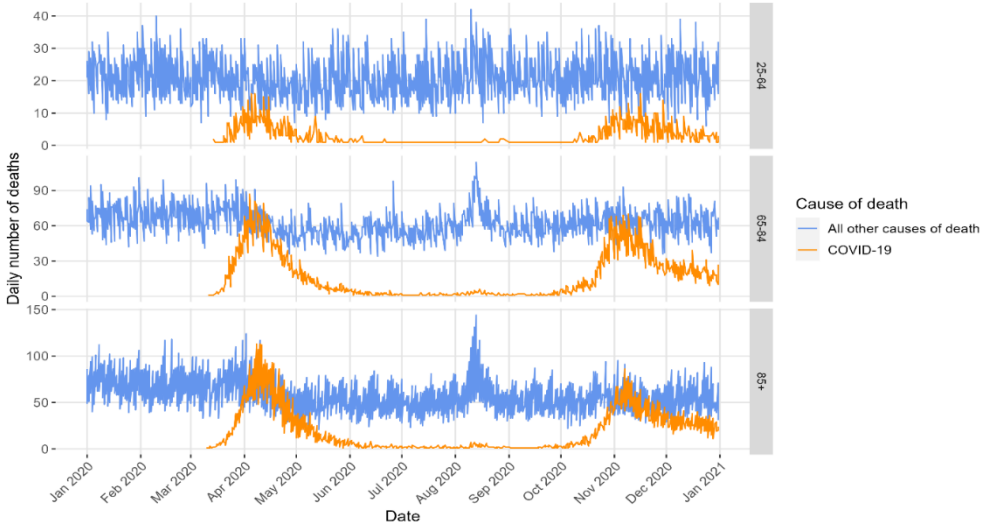



***Figure 24. Daily number of COVID-19 deaths and all other causes of death by age groups in 2020 in Belgium.***

## 3.2. Sociodemographic and socioeconomic characteristics

To identify correlates of COVID-19 specific death and all OCOD, we selected a set of relevant SD and SE variables regarding their association, previously demonstrated in several national and international studies, with COVID-19 mortality, excess mortality during the pandemic or all-cause mortality [120,211,212,296–299].

Age group, sex, migration background, and living situation were included as SD characteristics. Migration background was based on the individual's country of birth, nationality of origin and nationality of citizenship, as well as the country of birth of the parents. Individuals born outside Belgium have been classified as 'First-generation migrants'. Individuals born in Belgium with a nationality of origin and a nationality of citizenship equivalent to Belgian, and whose parents were both born in Belgium were classified as 'Belgian natives'. Individuals born in Belgium with at least one nationality other than Belgian, or at least one parent born outside Belgium have been classified as 'Second-generation migrants'. The living situation, partially reflecting the social environment, was categorized into five groups: 'With a partner', 'Without a partner', 'Care homes', 'Other collectivities (no care home)' (e.g., individuals living in prison), and 'Other' (e.g., adult children still living with their parents, other co-residents, individuals living in atypical households).

As SE predictors, we included education and income levels. Education level was classified into five categories based on the International Standard Classification of Education (ISCED): 'Primary or less' (ISCED 0 to ISCED 1), 'Lower secondary' (ISCED 2), 'Upper secondary' (ISCED 3 to ISCED 4), 'Higher education' (ISCED 5 to ISCED 8), and 'Missing'. Income data was available as the deciles of the yearly net taxable income per person and distinguished: 'Low income' (deciles 1 to 4), 'Middle income' (deciles 5 to 7), 'High income' (deciles 8 to 10), and 'Missing'.

## 4. Statistical analyses

First, the descriptive analysis shows the distribution (frequencies and percentages) of all SD and SE characteristics in the total study population among individuals who died from COVID-19, from all OCOD, or survived as of 31st December 2020. **Supplementary Table 17** shows the distribution (frequencies and percentages) of all SD characteristics in the study population with missing SE variables.

Second, the methodology includes two complementary statistical methods allowing a more in-depth exploration of the data: multivariable logistic regression models on the one hand and conditional inference tree (CIT) algorithms on the other hand.

### 4.1. Multivariable logistic regression models

In a first step, adjusted odds ratios (aORs) and Wald 95% confidence intervals (CIs) were estimated by fitting multivariable logistic regression models. A first logistic regression model was fitted considering COVID-19 specific deaths as cases and survivors and individuals who died from all OCOD as controls. A second regression model was fitted considering all OCOD as cases and survivors and individuals who died from COVID-19 as controls. The multivariable logistic regression models were adjusted for all SD and SE covariates included in our study.

4.2. Conditional inference tree algorithm

In a second step, CIT algorithms were computed, complementing logistic regression models. CIT algorithm, a non-parametric class of classification tree, is a valuable machine learning method to assess the predictive association between a dependent variable and a set of predictors. This method relies on the concept of statistical significance, employing p-values. Specifically, we used conditional recursive partitioning tree algorithms, a subtype of CIT. This method uses binary recursive partitioning within a conditional inference framework to model a regression relationship. The p-values used in the algorithm construction are obtained from conditional permutation tests [190,191,193]. The algorithm works as follows [190,191,194]:

1) The algorithm initiates with the entire sample and seeks to identify the optimal binary initial split to form the root node. For this purpose, it assesses the correlation ($\rho_{X_jY}$) between the response variable (*Y*) and any random variable from the set of predictors ($X_j$). The null hypothesis ($H_0$*:* $\rho_{X_jY} = 0$) between *Y* and $X_j$ is assessed using a statistical test (i.e. conditional permutation test). If $H_0$ is rejected ($H_1$: $\rho_{X_jY} \neq 0$), the predictors $X_j$ with the strongest association (i.e. the lowest p-value) to *Y* is selected and constitutes the root node.

2) For each subset created, the algorithm recursively applies the conditional permutation tests for all predictors $X_j$ and selects the one that maximizes the pre-specified significance level α for splitting that subset. The algorithm stops when the stopping criteria are met.

3) The algorithm checks whether the specific stopping criteria are met at each recursive step. When the stopping criteria are met, the algorithm stops splitting the subset and the current subsets become the terminal nodes of the tree. The stopping criterium can be, for example, that $H_0$ can no longer be rejected at a pre-specified level α, or that a specific tree depth or a minimum number of observations in a terminal node has been achieved.

Logistic regression and CIT offer a comprehensive examination of the data, each method possessing distinct advantages. Logistic regression is useful to quantify the association between a response variable and covariates, providing probability levels and confidence intervals for regression coefficients [300]. However, multicollinearity is not handled and the inclusion of interactions can become complicated to interpret. Classification tree, serving as a valuable complement to standard regression methods, addresses these limitations by handling complex, non-linear interactions between variables, in which subgroups are formed through optimal splitting variables, leading to the natural emergence of interactions as new starting population [193]. The multicollinearity issues are managed by the

algorithm by selecting the most important of the two colinear variables and dealing with the non-selected variables by calculating importance scores, indicating their role as a substitute for primary divisions [188]. While some classification trees are prone to overfitting (e.g., Classification and Regression Tree) [188,191], the CIT algorithm mitigates the overfitting issue by employing a statistical testing approach at each node. Instead of growing the tree without constraints, the CIT algorithm used the significance of a test (i.e., permutation tests) to determine whether a split is statistically significant, preventing the algorithm from creating overly complex structures that may not generalize well to new data. This approach helps maintain a more accurate and less overfitted model, enhancing the reliability of predictions [190,191]. Finally, the classification tree provides a clear and easy hierarchical understanding of the relationship between the predictors and the outcome that can help policymakers target their interventions [193].

In our analyses, three CIT algorithms were computed to identify and rank in order of importance the SD and SE predictors of cause-specific mortality. Following exactly the same approach applied to the multivariable logistic regression models, the first CIT considered COVID-19 deaths as cases and survivors and individuals who died from all OCOD as controls. The second CIT considered all OCOD as cases and survivors and individuals who died from COVID-19 as controls.

By applying these two complementary statistical methods to each outcome, this methodological approach allows to unravel the SD and SE patterns associated with COVID-19 death and all OCOD, respectively. This approach deepens the analysis and contributes to a more robust understanding of the multi-faceted dynamics underlying mortality during the early stages of the COVID-19 pandemic.

### 4.3. Sensitivity analyses

To test the robustness of our results, we estimated adjusted subdistribution hazard ratios (aSHRs) and 95% CIs using Fine and Gray competing risk models to identify the SD and SE predictors of COVID-19 death and all OCOD. Two competing risk models were performed. The first model considered COVID-19 death as the event of interest and all OCOD as the competing event. The second model considered all OCOD as the event of interest and COVID-19 death as the competing event. Due to the use of large datasets leading to computation constraints, the competing risk models could only be performed on a 50% random subset representative of the total study population. The cumulative incidence curves by levels of all SD and SE variables, along with the Gray's test p-values, have also been computed on the full total study population.

In our analyses, we set a significance level $\alpha$ of 0.05 and imposed a maximum tree depth of four levels. A four-level tree allowed to maintain accuracy in the information provided by the CIT while avoiding an excessive number of terminal

nodes making visual interpretation and the message to be conveyed to policymakers more complex. For each CIT, the prevalence of cases in the study population of interest was used as the predictive threshold. CIT were trained on 70% of the data, 30% of the data was used to validate the performance of the algorithm. CIT performances were measured by the means of sensitivity [95%CI] and specificity [95%CI]. All analyses were performed in R version 4.2.0. [221]. The package "Partykit" was used for computing the CIT algorithms [301].

## 5. Results

### 5.1. Description of the study population

Our study population was composed of 8,254,632 individuals aged 25 years and over officially residing in Belgium as of 1st January 2020 (see **Supplementary Figure 2**).

**Table 19** characterizes the study population in detail. As of January 1st 2020, the study population included 8,254,632 individuals officially residing in Belgium aged 25 years and over. Individuals who died from COVID-19 accounted for 0.27% (N=21,941) of the total study population, while individuals who died from all OCOD accounted for 1.25% (N=103,591) of the total study population. Individuals who survived as of December 31st 2020 accounted for 98.48% (N=8,129,100) of the total study population.

**Supplementary Table 17** characterizes the study population including individuals with missing SE information. The majority of individuals with missing SE information was aged 25 to 64 years (93.58%), was females (49.32%), lived with a partner (45.05%), and a first-generation migration background (89.61%).

**Table 19. Characterization of the study population, January 1st 2020 – December 31st 2020, Belgium**

| | All n=8 254 632 (100%) | COVID-19 specific death n=21 941 (0.27%) | All other causes of death n=103 591 (1.25%) | Survivors n=8 129 100 (98.48%) |
|---|---|---|---|---|
| Age groups, n(%) | | | | |
| 25-64 | 6 050 325 (73.30) | 1 421 (6.48) | 15 204 (14.68) | 6 033 699 (74.22) |
| 65-84 | 1 869 239 (22.64) | 9 483 (43.22) | 45 823 (44.23) | 1 813 933 (22.31) |
| 85+ | 335 069 (4.06) | 11 037 (50.30) | 42 564 (41.09) | 281 468 (3.46) |
| Sex, n (%) | | | | |
| Females | 4 247 959 (51.46) | 11 253 ( 51.29 ) | 52 889 ( 51.06 ) | 4 183 817 ( 51.47 ) |
| Males | 4 006 673 (48.54) | 10 688 ( 48.71 ) | 50 702 ( 48.94 ) | 3 945 283 ( 48.53 ) |
| Living situation, n (%) | | | | |
| With partner | 5 226 987 (63.32) | 6 713 ( 30.6 ) | 39 082 ( 37.73 ) | 5 181 192 ( 63.74 ) |
| Without partner | 2 166 454 (26.25) | 6 307 ( 28.75 ) | 37 106 ( 35.82 ) | 2 123 041 ( 26.12 ) |
| Care homes | 96 464 (1.17) | 7 283 ( 33.19 ) | 19 825 ( 19.14 ) | 69 356 ( 0.85 ) |
| Other collectivities (no care home) | 35 857 (0.43) | 655 ( 2.99 ) | 2 021 ( 1.95 ) | 33 181 ( 0.41 ) |
| Other | 728 870 (8.83) | 983 ( 4.48 ) | 5 557 ( 5.36 ) | 722 330 ( 8.89 ) |
| Migration background, n (%) | | | | |
| Belgian natives | 5 668 832 (68.67) | 17 175 (78.28) | 83 628 (80.73) | 5 568 029 (68.50) |
| Second-generation migrants | 923 599 (11.19) | 1 630 (7.43) | 8 034 (7.76) | 913 935 (11.24) |
| First-generation migrants | 1 662 201 (20.14) | 3 136 (14.29) | 11 929 (11.52) | 1 647 136 (20.26) |

**Table 19. Continued**

| | All n=8 254 632 (100%) | COVID-19 specific death n=21 941 (0.27%) | All other causes of death n=103 591 (1.25%) | Survivors n=8 129 100 (98.48%) |
|---|---|---|---|---|
| Income, n (%) | | | | |
| Low | 2 733 998 (33.12) | 9 992 ( 45.54 ) | 46 624 ( 45.01 ) | 2 677 382 ( 32.94 ) |
| Middle | 2 319 805 (28.10) | 8 814 ( 40.17 ) | 40 005 ( 38.62 ) | 2 270 986 ( 27.94 ) |
| High | 2 632 537 (31.89) | 2 534 ( 11.55 ) | 13 864 ( 13.38 ) | 2 616 139 ( 32.18 ) |
| Missing | 568 292 (6.88) | 601 ( 2.74 ) | 3 098 ( 2.99 ) | 564 593 ( 6.95 ) |
| Education, n (%) | | | | |
| Primary or less | 934 378 (11.32) | 8 581 ( 39.11 ) | 34 985 ( 33.77 ) | 890 812 ( 10.96 ) |
| Lower secondary | 1 731 746 (20.98) | 5 279 ( 24.06 ) | 26 069 ( 25.17 ) | 1 700 398 ( 20.92 ) |
| Upper secondary | 2 470 213 (29.93) | 3 188 ( 14.53 ) | 18 309 ( 17.67 ) | 2 448 716 ( 30.12 ) |
| Higher education | 1 981 898 (24.01) | 2117 ( 9.65 ) | 12 437 ( 12.01 ) | 1 967 344 ( 24.2 ) |
| Missing | 1 136 397 (13.77) | 2776 ( 12.65 ) | 11 791 ( 11.38 ) | 121 830 ( 13.8 ) |

## 5.2. Predictors of COVID-19 specific death and other causes of death vs. survival, respectively

### *5.2.1. Multivariable logistic regression model results*

**Table 20** shows the fully aORs and 95% CIs estimated resulting from the multivariable logistic regression models to identify the association of SD and SE variables with COVID-19 death and all OCOD compared to survival, respectively.

A higher odds of dying from COVID-19 was found among the middle-aged (aOR 18.93 [17.86 – 20.06]) and the elderly (aOR 65.47 [61.55 – 69.64]), compared to the youngest age group. Men had a two times higher mortality from COVID-19 (aOR 2.06 [2.00 – 2.12]) than women. Individuals living with a partner have lower odds of dying from COVID-19 compared to individuals in any other type of living situation; individuals living in care homes presenting the highest odds of dying from COVID-19 (aOR 10.22 [9.82 – 10.65]). Compared to Belgian natives, first-generation migrants had higher odds of dying from COVID-19 (aOR 1.23 [1.18 – 1.29]). A higher odds of dying from COVID-19 was identified among individuals with low (aOR 1.32 [1.25 – 1.38]) and middle (aOR 1.18 [1.13 – 1.24]) income, compared to individuals with high income. Compared to individuals with higher education, individuals with primary or less education (aOR 1.72 [1.63 – 1.81]), lower secondary education (aOR 1.40 [1.33 – 1.48]), and upper secondary education (aOR 1.18 [1.11 – 1.25]) had

higher odds of dying from COVID-19, highlighting a clear gradient in COVID-19 mortality by education level.

Overall, the SD and SE patterns identified in all OCOD were similar to those observed in COVID-19 mortality. An important difference was reflected by the much lower ORs for individuals aged 65 to 84 and aged 85 and over (aOR 8.45 [8.29 – 8.62] and aOR 31.71 [31.01 – 32.42] compared to individuals aged 25 to 64, respectively), as well as for individuals living in care homes (aOR 5.75 [5.63 – 5.87]) compared to individuals living with a partner (**Table 20**). Another difference between all OCOD and COVID-19 death was observed for the migration background variable. Compared to Belgian natives, second-generation migrants (aOR 0.79 [0.77 – 0.81]) and first-generation migrants (aOR 0.92 [0.90 – 0.94]) had lower odds of dying from OCOD.

**Table 20. Fully adjusted odds ratios (aORs) and their 95% confidence intervals (CIs) for the sociodemographic and socioeconomic characteristics associated with COVID-19 death and all other causes of death compared to survival, respectively.**

| | COVID-19 death* | | All other causes of death* | |
|---|---|---|---|---|
| | **aOR [95%CI]** | **P-value** | **aOR [95%CI]** | **P-value** |
| Age group (in years) | | | | |
| 25-64 | 1 | 1 | 1 | 1 |
| 65-84 | 18.93 [17.86 – 20.06] | <0.001 | 8.45 [8.29 – 8.62] | <0.001 |
| 85+ | 65.47 [61.55 – 69.64] | <0.001 | 31.71 [31.01 – 32.42] | <0.001 |
| Sex | | | | |
| Females | 1 | 1 | 1 | 1 |
| Males | 2.06 [2.00 – 2.12] | <0.001 | 1.80 [1.77 – 1.82] | <0.001 |
| Living situation | | | | |
| With partner | 1 | 1 | 1 | 1 |
| Without partner | 1.56 [1.50 – 1.62] | <0.001 | 1.62 [1.59 – 1.64] | <0.001 |
| Care homes | 10.22 [9.82 – 10.65] | <0.001 | 5.75 [5.63 – 5.87] | <0.001 |
| Other collectivities (no care home) | 6.54 [6.01 – 7.11] | <0.001 | 3.72 [3.54 – 3.90] | <0.001 |
| Other | 1.64 [1.53 – 1.76] | <0.001 | 1.47 [1.42 – 1.51] | <0.001 |
| Migration background | | | | |
| Belgian natives | 1 | 1 | 1 | 1 |
| Second-generation migrants | 1.02 [0.97 – 1.08] | 0.410 | 0.79 [0.77 – 0.81] | <0.001 |
| First-generation migrants | 1.23 [1.18 – 1.29] | <0.001 | 0.92 [0.90 – 0.94] | <0.001 |

**Table 20. Continued**

| | COVID-19 death* | | All other causes of death* | |
|---|---|---|---|---|
| | **aOR [95%CI]** | **P-value** | **aOR [95%CI]** | **P-value** |
| Income | | | | |
| High | 1 | 1 | 1 | 1 |
| Low | 1.32 [1.25 – 1.38] | <0.001 | 1.52 [1.49 – 1.56] | <0.001 |
| Middle | 1.18 [1.13 – 1.24] | <0.001 | 1.29 [1.27 – 1.32] | <0.001 |
| Missing | 1.27 [1.15 – 1.40] | <0.001 | 1.42 [1.36 – 1.48] | <0.001 |
| Education | | | | |
| Higher education | 1 | 1 | 1 | 1 |
| Primary or less | 1.72 [1.63 – 1.81] | <0.001 | 1.47 [1.43 – 1.50] | <0.001 |
| Lower secondary | 1.40 [1.33 – 1.48] | <0.001 | 1.25 [1.22 – 1.28] | <0.001 |
| Upper secondary | 1.18 [1.11 – 1.25] | <0.001 | 1.10 [1.07 – 1.12] | <0.001 |
| Missing | 1.49 [1.40 – 1.59] | <0.001 | 1.27 [1.23 – 1.31] | <0.001 |

** All ORs were fully adjusted for all SD and SE variables.*

### *5.2.2. Conditional recursive partitioning tree algorithm results*

**Figure 25** presents the CIT algorithm considering COVID-19 specific deaths as cases and survivors and individuals who died from all OCOD as controls, to identify and rank in order of importance the SD and SE predictors of COVID-19 mortality.

'Living situation', 'Age group', 'Sex', and 'Migration background' emerged as the strongest predictors of COVID-19 death, ranked according to their statistical significance. The variable 'Living situation' was assigned to the root node and constituted the first split by dividing the total study population in two subgroups (i.e., individuals living in care homes vs. those not living in care homes). The tree continued to grow while 15 terminal nodes were determined. The proportion of COVID-19 deaths relative to the sample size of each terminal node is indicated in each of them. COVID-19 mortality was predicted when the proportion of COVID-19 deaths in a terminal node exceeded its prevalence in the study population (i.e., 0.27%).

The left branch of the tree, composed of individuals not living in care homes, was secondly partitioned based on 'Age group' between individuals aged 25 to 84 and those aged 85 and over. The subgroup composed of individuals aged 25 to 84 was thirdly partitioned between those aged 25 to 64 and those aged 65 to 84 while the subgroup composed of individuals aged 85 and over was thirdly partitioned based on 'Living situation' between individuals living in collectivities (no care homes) and

those not living in collective settings. ‘Living situation’ finally partitioned the subgroups of individuals aged 25 to 64 and 65 to 84 years in four terminal nodes, with the terminal nodes including individuals living in collectivities (no care home) predicting the highest proportion of COVID-19 death (0.81%). ‘Sex’ finally partitioned the subgroups of individuals aged 85 and over living in either collectivities (no care homes) or not in collectivities in four terminal nodes. Those four terminal nodes predicted COVID-19 death with the highest proportion of COVID-19 deaths among males living in collectivities (no care home) (8.77% of COVID-19 deaths).

The right branch of the tree, composed of individuals living in care home, was secondly partitioned between males and females. The subgroup composed of females was thirdly partitioned based on ‘Age group’ between individuals aged 25 to 84 and those aged 85 and over. The subgroup composed of individuals aged 25 to 84 was finally partitioned between those aged 25 to 64 and those aged 65 to 84, with the highest proportion of COVID-19 death among the latter subgroup (5.53% of COVID-19 deaths). The subgroup composed of individuals aged 85 and over was finally partitioned in two terminal nodes based on ‘Migration background’ with the terminal node including first-generation migrants predicting the highest proportion of COVID-19 deaths (9.52% of COVID-19 deaths). The subgroup composed of males was thirdly partitioned between individuals aged 25 to 64, directly leading to a

terminal node predicting COVID-19 death (2.57% of COVID-19 deaths), and individuals aged 65 and over. This latter subgroup was finally partitioned in two terminal nodes between individuals aged 65 to 84 and those aged 85 and over, with the terminal node including individuals aged 85 and over predicting the highest proportion of COVID-19 death (12.73% of COVID-19 deaths).

The sensitivity [95%] and specificity [95%] of the CIT were 0.945 [0.941 - 0.954] and 0.738 [0.732 - 0.751], respectively.

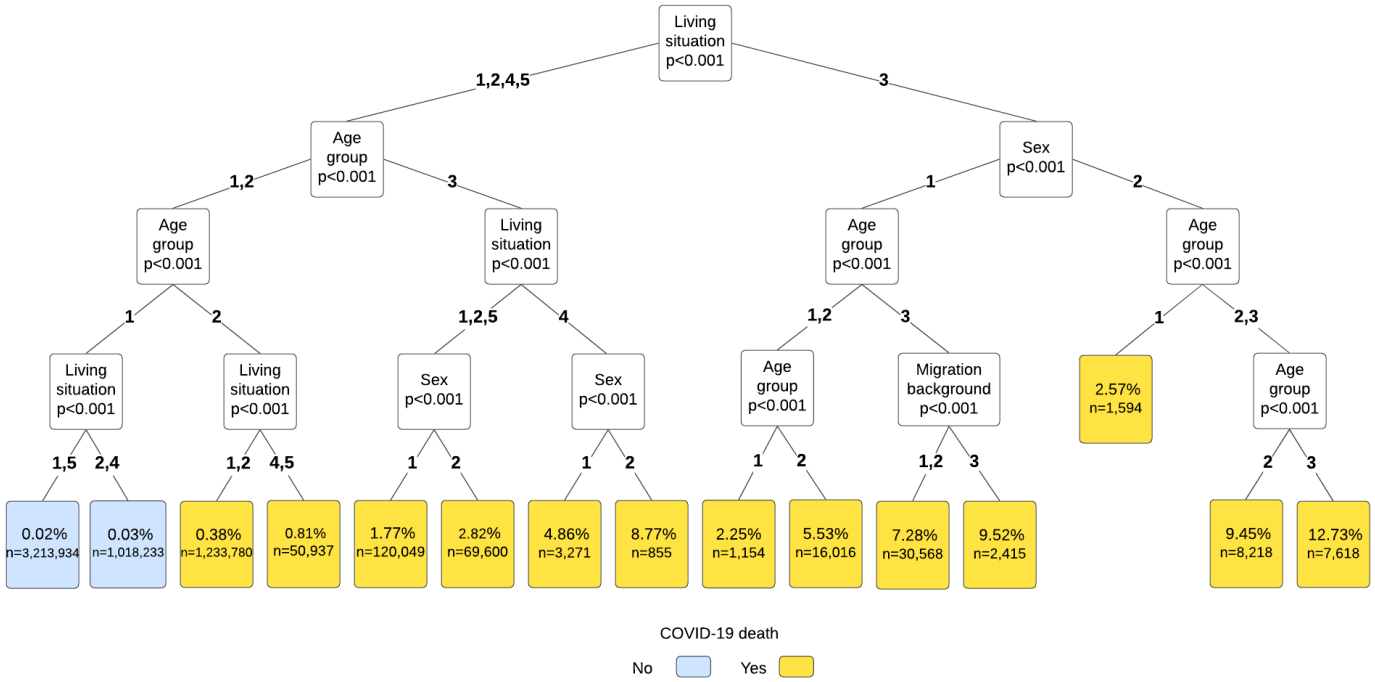



***Figure 25. Conditional recursive partitioning tree algorithm considering COVID-19 specific death as cases.*** *The variable 'Living situation' was categorized as follows: 1='With partner', 2='Without partner', 3='Care homes', 4='Other collectivities (no care home)', 5='Other'. The variable 'Age group' (in years) was categorized as follow: 1='25-64', 2='65-84', 3='85+'. The variable 'Sex' was categorized as follow: 1='Females', 2='Males'. The variable 'Migration background' was categorized as follows: 1='Belgian natives', 2='Second-generation migrants', 3='First-generation migrants'. The percentages in the terminal nodes indicate the percentage of COVID-19 deaths relative to the sample size of each terminal node, denoted 'n'. A terminal node predicted COVID19 death when the proportion of COVID-19 deaths in the terminal node exceeded its prevalence in the study population (i.e., 0.27%).*

**Figure 26** presents the CIT algorithm considering all OCOD as cases and survivors and individuals who died from COVID-19 as controls, to identify and rank in order of importance the SD and SE predictors of all OCOD.

'Age group', 'Living situation', and 'Sex' emerged as the three strongest predictors of all OCOD, ranked according to their statistical significance. The variable 'Age group' was assigned to the root node and constituted the first split by dividing the study population into two subgroups (i.e., individuals aged 25 to 84 years vs. individuals aged 85 years and over). The tree continued to grow resulting in 16 terminal nodes. The proportion of all OCOD relative to the sample size of each terminal node is indicated in each of them. All other causes mortality was predicted when the proportion of all OCOD in a terminal node exceeded its prevalence in the study population (i.e., 1.25%).

The left branch of the tree, composed of individuals aged 25 to 84, was secondly partitioned between those living with a partner or in the category 'Other' and those living without a partner or in collectivities. The subgroup of individuals living with a partner or in the category 'Other' was thirdly partitioned in four terminal nodes based on 'Age group' and divided between those aged 25 to 64 and those aged 65 to 84. Among individuals aged 25 to 64, the variable 'Education' performed the final split, with a three-fold higher proportion of all OCOD among those with the lowest level of education (0.47% of all OCOD), compared to those with highest levels of

education (0.17% of all OCOD). Among individuals aged 65 to 84, the final split was performed by the variable 'Sex' with the terminal nodes composed of males and females predicting all OCOD (2.47% and 1.38% of all OCOD, respectively). The subgroup of individuals living without a partner or in collectivities was further partitioned based on 'Living situation' and 'Age group' in four terminal nodes. Out of these four nodes, three predicted all OCOD, with the highest proportion of all OCOD identified among individuals living in care homes aged 65 to 84 (15.29% of all OCOD).

The right branch of the tree, composed of individuals aged 85 and over, was secondly partitioned between individuals living in collectivities and those not living in collectivities. The subgroup composed of individuals not living in collectivities was further partitioned into four terminal nodes based on 'Sex' and 'Living situation'. Those four terminal nodes predicted all OCOD with the highest proportions of all OCOD identified among males living without a partner or in other types of household (13.23% of all OCOD), followed by females living in other types of household as well (13.48% of all OCOD). The subgroup composed of individuals living in collectivities was further partitioned based on 'Sex' and 'Living situation' into four terminal nodes, all predicting all OCOD, with the highest proportion of all OCOD among males living in care homes (29.05% of all OCOD).

The sensitivity [95%CI] and specificity [95%CI] were 0.712 [0.705 - 0.714] and 0.823 [0.815 - 0.825], respectively.

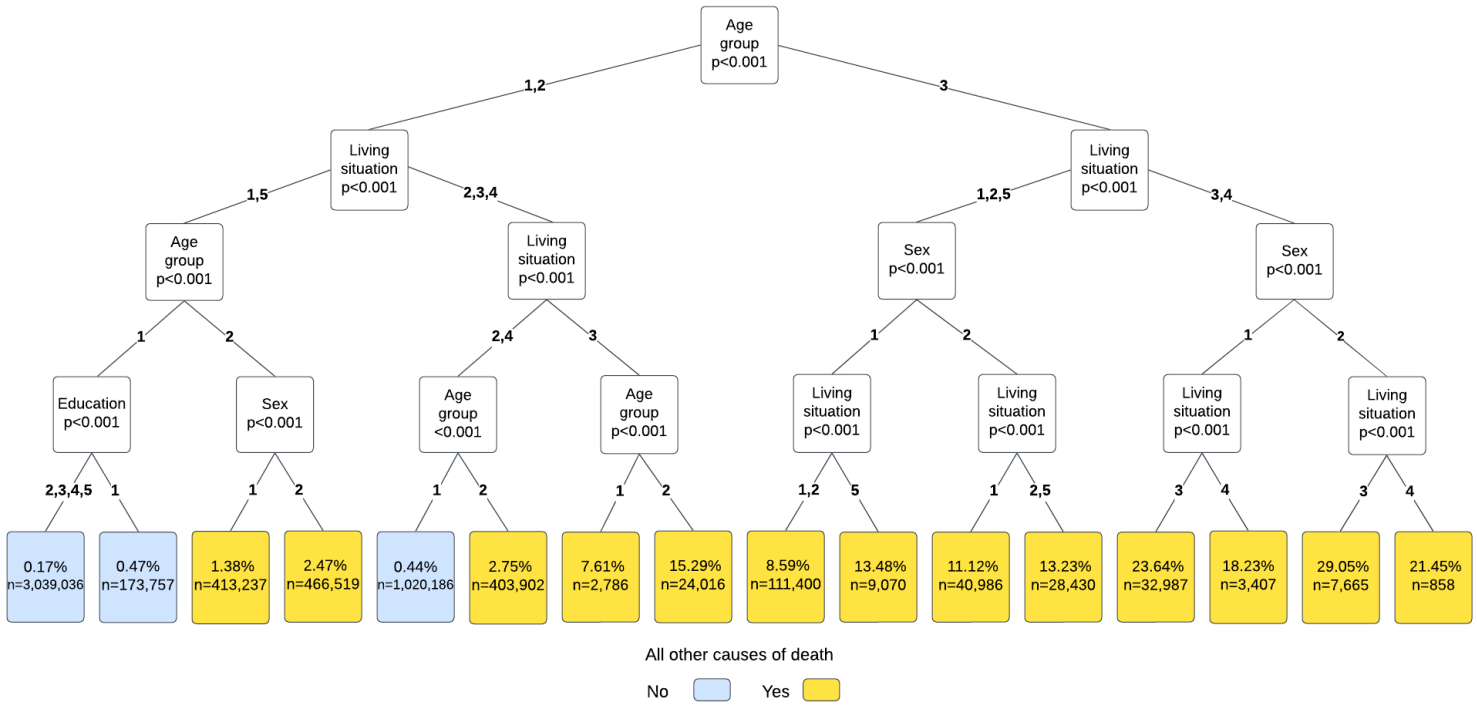



***Figure 26. Conditional recursive partitioning tree algorithm considering all other causes of death (OCOD) as cases.*** *The variable 'Age group' (in years) was categorized as follows: 1='25-64', 2='65-84', 3='85+'. The variable 'Living situation' was categorized as follow: 1='With partner', 2='Without partner', 3='Care homes', 4='Other collectivities (no care home)', 5='Other'. The variable 'Sex' was categorized as follow: 1='Females', 2='Males'. The variable 'Education' was categorized as follows: 1='Primary or less', 2='Lower secondary', 3= 'Upper secondary', 4='Higher education', 5='Missing'. The percentages in the terminal nodes indicate the percentage of all OCOD relative to the sample size of each terminal node, denoted 'n'. A terminal node predicted all OCOD when the proportion of all OCOD in the terminal node exceeded its prevalence in the study population (i.e., 1.25%).*

5.3. Sensitivity analyses

**Supplementary Table 18** shows the aSHRs and 95% CIs from Fine and Gray competing risk models identifying SD and SE predictors of COVID-19 death, accounting for all OCOD as competing events, and vice versa. These models were performed on a 50% random subset representative of the total study population. The results show similar trends to those observed in the logistic regression models (**Table 20**), reinforcing the robustness of our main results. **Supplementary Figures 3** to **8** provide the cumulative incidences of COVID-19 death accounting for all OCOD as competing events, and vice versa, stratified by SD and SE variables.

## 6. Discussion

Thanks to the implementation of two complementary statistical methods, i.e., multivariable logistic regression models and CIT algorithms, the identification and ranking of SD and SE predictors of COVID-19 death and all OCOD were performed, providing more reliable estimates and an easy hierarchical visual way of understanding the patterns characterizing mortality during the early phases of the COVID-19 pandemic in Belgium.

Compared to individuals not living in collectivities, our results showed higher odds of dying from COVID-19 among individuals living in collectivities (approximatively 10-fold higher and 7-fold higher for individuals living in care homes and those living in other type of collectivities, respectively). These results were confirmed by the CIT

algorithm, highlighting that living in care homes was the strongest predictor of COVID-19 death across all age groups. This aligns with several studies reporting important excess mortality during the COVID-19 pandemic in collectivities, whether in care homes or other settings (e.g. prisons) [120,297,302,303]. This higher rate of COVID-19 deaths can be attributed on one hand to a higher comorbidity burden among residents of certain collectivities, such as prisons [304] and especially care homes, given the frailty profile of the residents [303,305,306]. On the other hand, higher virus transmission in collective settings in addition to pandemic-induced shortage of medical staff and resources may have led to higher COVID-19 related lethality [303,307,308] especially during the first year of the pandemic characterized by shortage masks, hydroalcoholic gels, and the unavailability of the vaccine at that time [44,309].

Following the living situation, the CIT algorithm reveals that age and gender are the most important predictors of COVID-19 mortality. Age and gender are well known risk factors for COVID-19 death [287,288,310] and for mortality in general [311]. Ageing increases the risk of mortality following COVID-19 via many biological and behavioral factors such as: the overexpression of Angiotensin converting enzyme-2 (ACE-2) (receptor enzyme allowing the SARS-CoV-2 to enter the cell and replicate, leading to an acceleration of the replication of the virus in the lungs and thus a more severe infection) [298]; the presence of aging-associated comorbidities (e.g.,

diabetes, cardiovascular diseases) which are determinants in the progression towards a severe form of COVID-19 [312]; a change in lifestyle habits (malnutrition, decreased physical activity) also contributing to a more severe progression of the disease [298,312]. The variation in susceptibility to infectious diseases between men and women can be explained by hormonal factors (e.g., higher plasma levels of ACE-2 in males) [296], genetic factors (e.g., high density of immune-related genes are located on the X chromosome giving women stronger immune response) [296], and behavioral factors (e.g., high-risk behaviors among men such as decreased perception of risks, higher rate of tobacco and alcohol use) [313].

Our multivariable logistic regression model results also showed a 32% higher COVID-19 mortality among individuals with the lowest level of income compared to those with the highest, and a 72% higher COVID-19 mortality rate among those with the lowest level of education compared to the highest. Numerous studies documented social disparities in COVID-19 mortality or excess mortality during the pandemic [120,211,212,314–319]. In Italy, Bello et al. identified a protective role of education on excess mortality during the first major COVID-19 outbreak. Their results showed that when the proportion of residents with at least secondary education increased by 10%, deaths per month decreased by 0.426 in the Northern region and by 0.081 in the Southern region [315]. In Canada, van Ingen et al. found an adjusted relative COVID-19 mortality risk two times higher in the neighborhood

with a high proportion of individuals with less than high school education [320]. In Belgium, similar SE patterns were observed in excess mortality during the first COVID-19 wave with a 23.4% and 21.3% excess mortality among females with the lowest level of education not living in collectivities aged 25 to 64 and 65 to 84 years, respectively [120].

Regarding income inequalities, a study conducted in the Eastern Mediterranean region found a 3.9% increase in COVID-19 death per million population with a one unit increase in the Gini coefficient (a measure of income inequalities) [321]. In Belgium, during the first COVID-19 wave, similar patterns were observed by Decoster et al. who found a significant negative income gradient in excess mortality among people aged over 65 years, for both men and women [211]. In the Netherlands, Wouterse et al. found that, across all-age groups, COVID-19 mortality was more concentrated among low-income groups, compared to all other causes of mortality [322].

A study conducted by Albani et al. identified that inequalities in transmission and in vulnerability were the two main factors explaining the highest proportion of COVID-19 mortality by deprivation [323]. Indeed, SE disadvantaged groups are more at risk of SARS-CoV-2 virus transmission and related severe COVID-19 outcomes due to increased vulnerability, susceptibility, exposure, and transmission [105]. SE disadvantaged groups are more vulnerable in terms of health conditions,

as they are experiencing a higher burden of comorbidities [103], which has been shown to increase COVID-19 mortality [104]. They are more susceptible to be infected and developed severe related COVID-19 outcomes due to poor living conditions, isolation, and chronic stress weakening the immune system, even without preexisting underlying health conditions [105,324]. Working in essential sectors (e.g., social sectors, education, defense, logistic and transportation, manufacturing, facilities) during the pandemic has been shown to led to higher rates of SARS-CoV-2 infection and COVID-19 death [325,326]. Individuals with lower levels of education are more likely to hold essential jobs or work in lower-skilled occupations, limiting their ability to work remotely during the pandemic. This reduced access to remote work created obstacles to implementing preventive measures, potentially contributing to higher exposure to the virus among SE disadvantaged groups [105,109,327]. In addition, a lower adoption of preventive measures has been identified as highly correlated with the level of education due to decreased perception of the risk and trust in the effectiveness of preventive measures [106,107]. Finally, SE disadvantaged groups experienced an increased transmission by living in overcrowded households and neighborhood with higher population density with restricted access to outside spaces [105,108,109].

We identified in the multivariable logistic regression model a lower risk of all OCOD compared to survival up to 21% among first- and second-generation migrants. The

lower migrant mortality, or *migrant mortality advantage*, is often explained by a selection effect (i.e., the healthiest individuals are able to immigrate) and healthier lifestyle habits (e.g., the Mediterranean diet leads to lower rates of chronic diseases) [328–330], leading to lower mortality in the destination country, at least in the early stages of immigration [328]. Conversely, we identified in the multivariable logistic regression model results a 23% higher COVID-19 mortality among first-generation migrants, even when controlling for SE characteristics. This is in line with numerous studies reporting a higher risk of COVID-19 death among ethnic minorities [98,212,289,292,331,332].

The observation of a "migrant mortality advantage" in our results for all OCOD, but the observation of the opposite trends for COVID-19 mortality, is consistent with previous findings in Belgium [333]. This pattern may be explained by poorer migrant's living conditions leading to a higher SARS-CoV-2 virus transmission and resulting in higher mortality from COVID-19 [334,335]. Indeed, two studies reported that living in area with high density population, favoring virus exposure and infection, was the main factor explaining up to 60% of the disparities in COVID-19 mortality among ethnic minorities [292,331]. Migrant's working conditions may also contribute to these disparities, as they are disproportionately represented in essential and public-facing jobs, which increases their exposure to the virus [336]. In addition, once infected, difficulties in accessing health care services among

migrants may have arisen due to language and cultural barriers as well as a lack of support system or digital literacy [337–339].

Our study has several limitations. Firstly, we were unable to adjust for the health status prior to COVID-19 outbreak (i.e., comorbidities), an essential factor to control for when analyzing cause-specific mortality data. Adjusting for some mediating factors of the socioeconomic disparities in COVID-19 severe outcomes such as the type of occupation, the adoption of preventive measures (e.g., mask wearing), and access to healthcare resources and services would have better captured the underlying pathways of cause-specific mortality during the COVID-19 pandemic. Second, the dataset used in this study contains individuals officially residing in Belgium. We therefore missed unregistered individuals, a more vulnerable part of the population (e.g., undocumented migrants), who may have had higher exposure to the SARS-CoV-2 virus and higher related COVID-19 mortality leading to an underestimation of cause-specific mortality for certain SD and SE groups. Third, the exact country of origin was not available in the dataset used. Yet, evidence from Belgian census and registry data has shown that country of origin influences the risk of mortality from both chronic and infectious diseases. For example, migrants from sub-Saharan Africa faced up to a 19-fold higher risk of death from infectious diseases compared to Belgian natives, whereas this risk was not significant for migrants from Morocco. Conversely, migrants from Morocco

presented a lower risk of death from chronic conditions, whereas this risk was not significant for migrants from sub-Saharan Africa [333]. The lack of detailed country-of-origin data in our study therefore limited the precision of our findings and related tailored policy recommendations. However, it should be noted that another Belgian study conducted by Vanthomme et al. [295] analyzed excess mortality among migrants during the first wave using more granular country-of-origin data. Fourth, the exact cause of death of individuals who died abroad were not available in the death certificate, deaths recorded abroad were therefore automatically categorized as all OCOD. This has potentially led to an underestimation of COVID-19 deaths among our study sample given that 21,792 deaths were recorded abroad, or 17% of the total number of deaths recorded in 2020. Fifth, the lack of testing at the start of the pandemic, and the fact that healthcare professionals might be still unfamiliar with the symptomatology of the COVID-19, which could potentially be confused with another respiratory virus, may have also led to an underestimation of COVID-19 deaths. Conversely, an overestimation of COVID-19 deaths during the first and second COVID-19 waves can be expected in care homes. Finally, the use of the income variable, which is based on taxes records, may not accurately represent certain population groups, such as those working in the informal economy, individuals earning below certain eligibility thresholds, or wealthier people benefiting from special tax arrangements. However, the direction and magnitude of this bias on the results remain uncertain.

Our study has several strengths. First, using an exhaustive dataset on the Belgian adult population allows to generalize our results at the national level. Second, the use of microlevel cause-specific mortality data allowed to distinguish COVID-19 specific death from all OCOD in our analyses. Belgium recorded COVID-19 related deaths with fairly good accuracy compared with other European countries, despite limited testing capacity during the first wave of the COVID-19 pandemic and difficulties in determining cause of death due to overlap with comorbidities [340]. Third, the use of two complementary methods provided a more in-depth investigation of the data. Thanks to the CIT algorithm, we were able to account for multicollinearity and interactions between the different SD and SE predictors, leading to robust conclusions (also with regard to the good performances of the CIT algorithms measured in terms of sensitivity and specificity).

As the vaccination campaign started on 28 December 2020 in Belgium, the effect of the vaccine on cause-specific mortality is beyond the scope of this article. Given the SD and SE disparities identified in the COVID-19 vaccine uptake in Belgium [244], future research should focus on the effect of the vaccination campaign on SD and SE disparities in COVID-19 specific mortality during the pandemic.

## 7. Conclusion

In conclusion, we identified important sociodemographic and socioeconomic disparities in COVID-19 mortality, with living in care homes being the key predictor. Our results are useful for future pandemic preparedness and suggest that policymakers should implement political measures and communication strategies promoting preventive behaviors adapted and targeted to the most vulnerable populations more exposed to the virus.

## 8. Declarations

### 8.1. Ethics approval and consent to participate

This research adhere to the ethical code of scientific research in Belgium, see: http://www.belspo.be/belspo/organisation/publ/pub_ostc/Eth_code/ethcode_nl.pdf.

### 8.2. Consent for publication

Not applicable

### 8.3. Availability of data and materials

The analyses are based on data from a census-linked mortality follow-up study and cannot be made available due to privacy issues. Researchers can gain full access to the data by submitting an application to the Privacy Commission Belgium. In order to get permission to use data from the Belgian population register linked to census

data an authorization request (in Dutch or French) needs to be submitted to the Belgian Data Protection Authority. The authorization request includes an application form and additional forms regarding data security. The necessary forms for the authorization request can be downloaded from the Data Protection Authority website (https://www.dataprotectio nauthority.be). Next to information on the applicant and a list of requested data, the authorization request should specify why the data from the population register are necessary, for which time span data will be stored, and who will have access to the data.

### 8.4. Competing interests

Brecht Devleesschauwer is the editor in chief of Archives of Public Health. The other authors have no competing interests to declare that are relevant to the content of this article.


### 8.5. Funding

The Belgian Science Policy Office (BELSPO) funded this research within the BRAIN-be 2.0 framework supporting pillar 3 Federal societal challenges (grant number B2/202/P3/HELICON). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.


### 8.6. Authors' contributions

LC reviewed the literature; LC, LVdB, KV, AS, BD, NK, and SG conceived the study; LC, LVdB, KV, AS, BD, NK, and SG selected the population; LC, LVdB, KV, AS, and SG

reviewed all available data; LC, LVdB, KV, AS, BD, NK, and SG designed the statistical methodology; LC conducted the statistical analyses; LC, LVdB, KV, AS, BD, NK, and SG interpret the findings; LC wrote the first draft of the paper. All authors revised the text. LC, LVdB, KV, AS, and SG have directly access to the data of the present study. All authors approved the final version of this manuscript and accepted responsibility for its submission for publication.

### 8.7. Acknowledgments

We would like to thank Patrick Lusyne and Elias Neirynck from Statistic Belgium for their valuable advices. In addition, we would like to thank Johan Surkyn, the BRISPO data manager, who requested the data.

## 9. Supplementary Materials

### 9.1. Table of content

#### *9.1.1. Supplementary Figures*

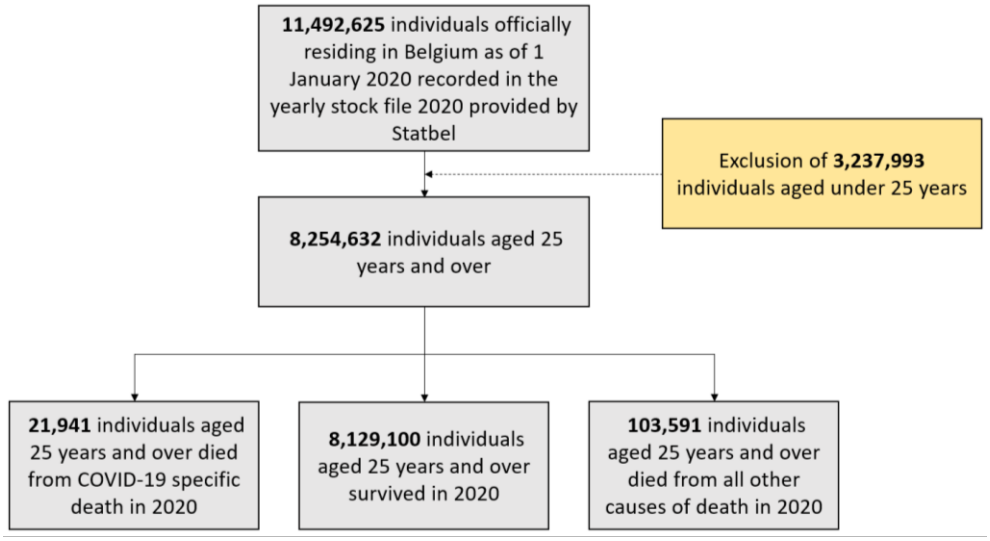
- **Supplementary Figure 2.** Flowchart demonstrating the study population selection process.
- **Supplementary Figure 3.** Cumulative incidences by age groups of (A) COVID-19 death considering survival as competing events; and (B) all other causes of death (OCOD) considering survival as competing events.

- **Supplementary Figure 4.** Cumulative incidences by gender of (A) COVID-19 death considering survival as competing events; and (B) all other causes of death (OCOD) considering survival as competing events.
- **Supplementary Figure 5.** Cumulative incidences by living situation of (A) COVID-19 death considering survival as competing events; and (B) all other causes of death (OCOD) considering survival as competing events.
- **Supplementary Figure 6.** Cumulative incidences by migration background of (A) COVID-19 death considering survival as competing events; and (B) all other causes of death (OCOD) considering survival as competing events.
- **Supplementary Figure 7.** Cumulative incidence by income level of (A) COVID-19 death considering survival as competing events; and (B) all other causes of death (OCOD) considering survival as competing events.
- **Supplementary Figure 8.** Cumulative incidences by education level of (A) COVID-19 death considering survival as competing events; and (B) all other causes of death (OCOD) considering survival as competing events.

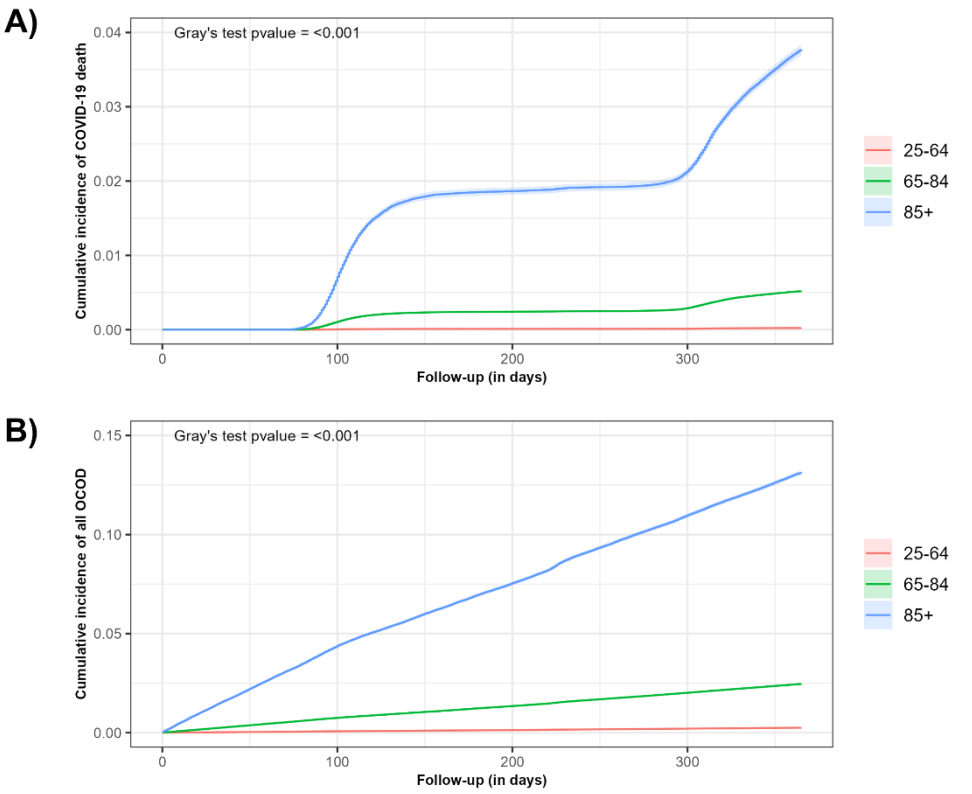
*9.1.2. Supplementary Tables*

- **Supplementary Table 17.** Characterization of the study population with missing socioeconomic information, January 1st 2020 – December 31st 2020, Belgium.
- **Supplementary Table 18.** Adjusted subdistribution hazard ratios (aSHRs) and their 95% confidence intervals (CIs) from Fine and Gray competing risk

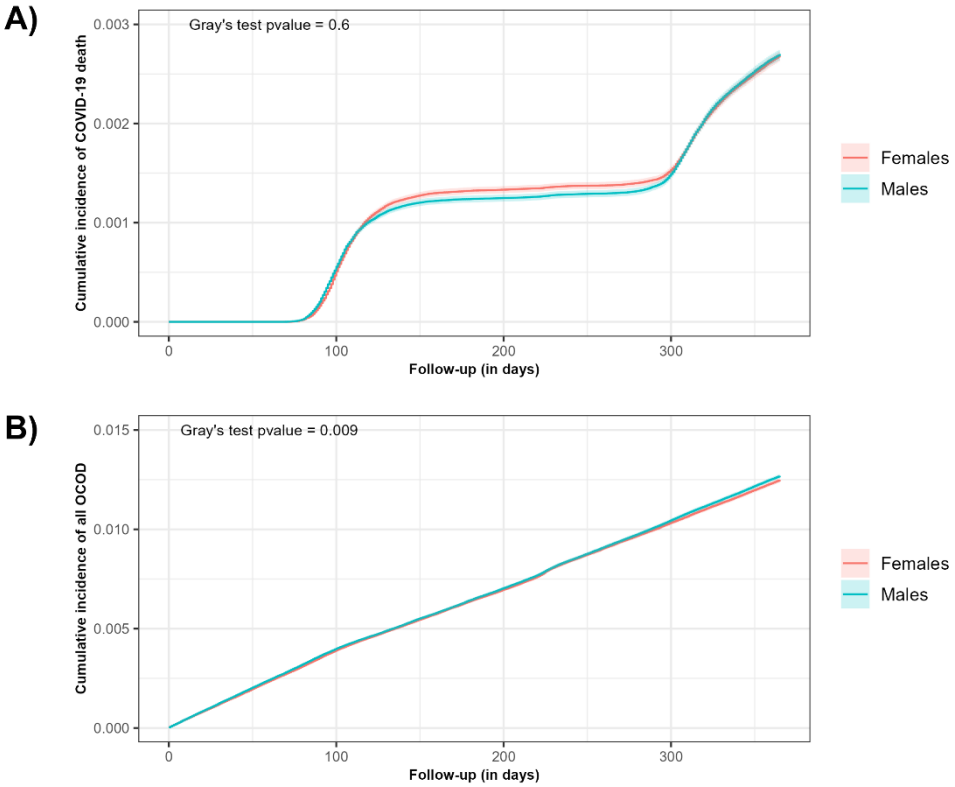
models for demographic and socioeconomic predictors of (1) COVID-19 death accounting for all other causes of death as competing events; and (2) all other causes of death accounting for COVID-19 death as competing events, computed on a 50% representative random sample of the total study population.



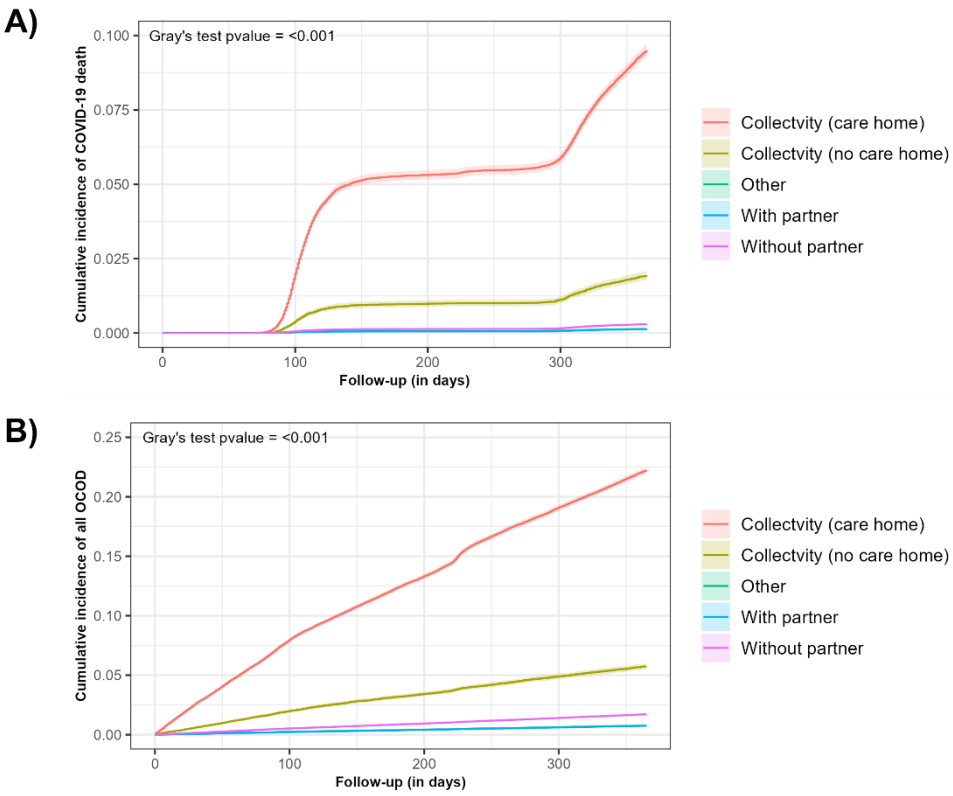

**Supplementary Figure 2. Flowchart demonstrating the study population selection process**



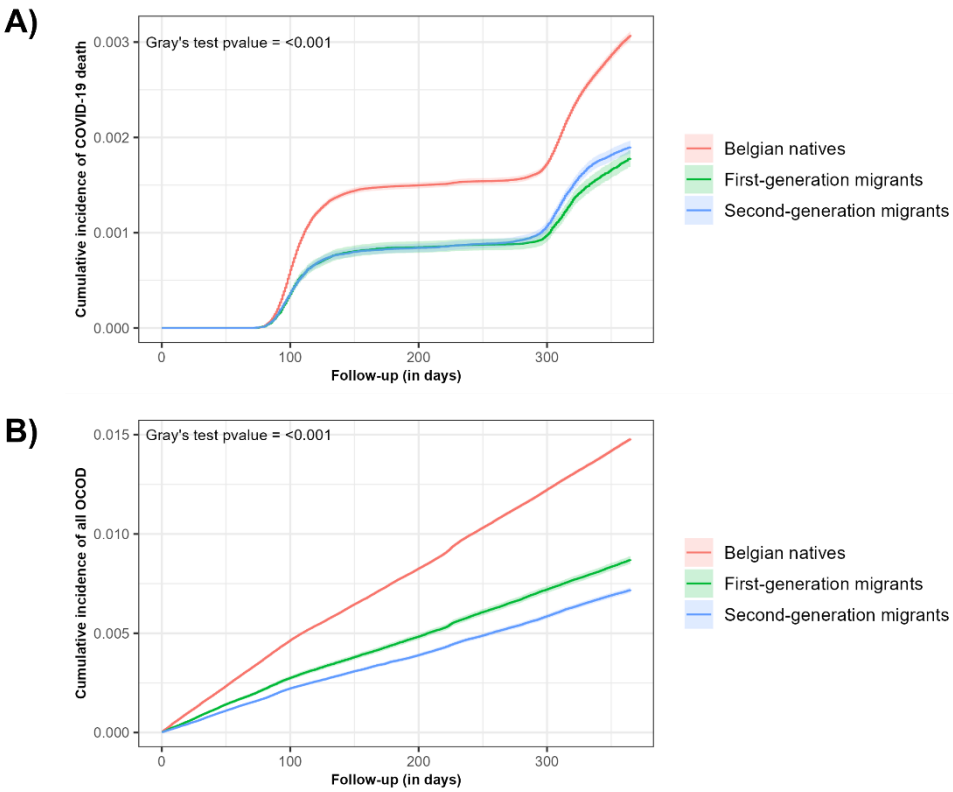

**Supplementary Figure 3. Cumulative incidences by age groups of (A) COVID-19 death considering all other causes of death (OCOD) as competing events; and (B) all OCOD considering COVID-19 death as competing events.**



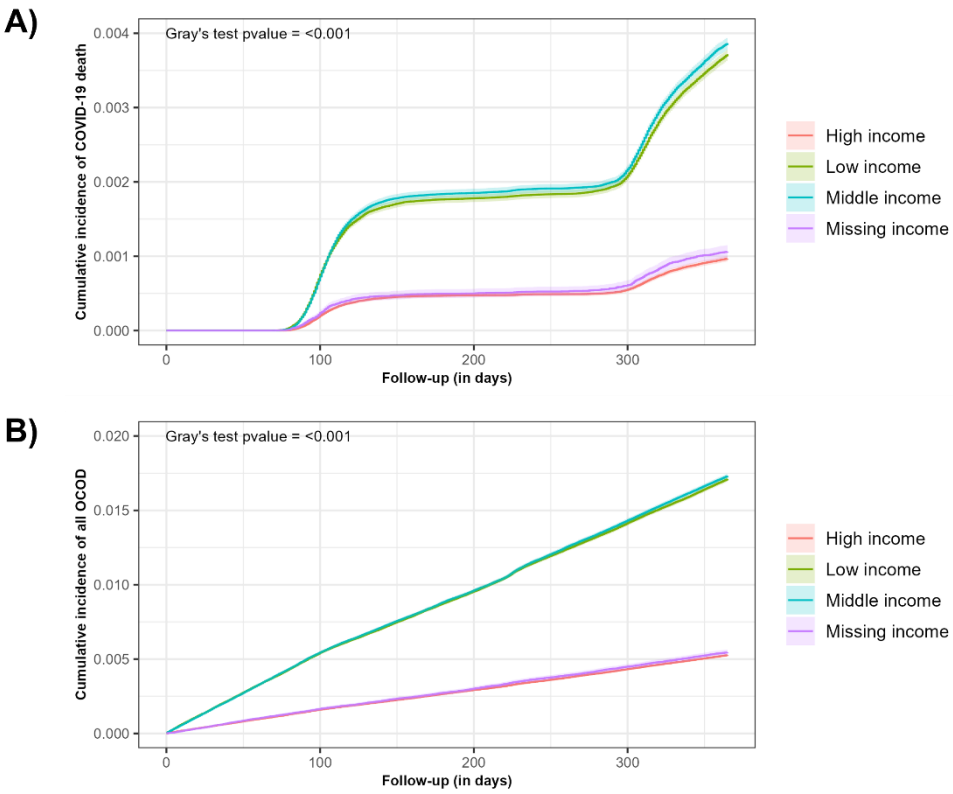

**Supplementary Figure 4. Cumulative incidences by gender of (A) COVID-19 death considering all other causes of death (OCOD) as competing events; and (B) all OCOD considering COVID-19 death as competing events.**



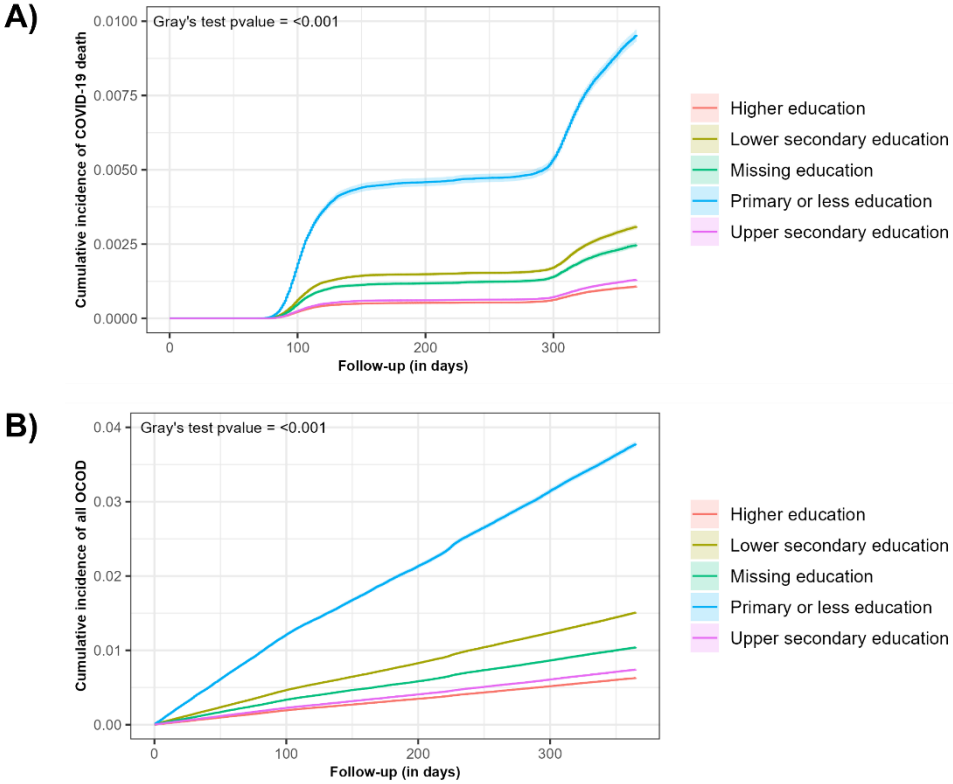

**Supplementary Figure 5. Cumulative incidences by living situation of (A) COVID-19 death considering all other causes of death (OCOD) as competing events; and (B) all OCOD considering COVID-19 death as competing events.**




**Supplementary Figure 6. Cumulative incidences by migration background of (A) COVID-19 death considering all other causes of death (OCOD) as competing events; and (B) all OCOD considering COVID-19 death as competing events.**




**Supplementary Figure 7. Cumulative incidences by income level of (A) COVID-19 death considering all other causes of death (OCOD) as competing events; and (B) all OCOD considering COVID-19 death as competing events.**




**Supplementary Figure 8. Cumulative incidences by education level of (A) COVID-19 death considering all other causes of death (OCOD) as competing events; and (B) all OCOD considering COVID-19 death as competing events.**

**Supplementary Table 17. Characterization of the study population with missing socioeconomic information, January 1st 2020 – December 31st 2020, Belgium**

| | All individuals with missing income and education levels n= 338 599 (100%) | COVID-19 specific death n= 299 (0.09%) | All other causes of death n=1 363 (0.40%) | Survivors n= 336 937 (99.51%) |
|---|---|---|---|---|
| Age groups, n(%) | | | | |
| 25-64 | 316 863 (93.58) | 122 (40.80) | 651 (47.76) | 316 090 (93.81) |
| 65-84 | 19 645 (5.80) | 120 (40.13) | 490 (35.95) | 19 035 (5.65) |
| 85+ | 2 091 (0.62) | 57 (19.06) | 222 (16.29) | 1 812 (0.54) |
| Sex, n (%) | | | | |
| Females | 166 984 (49.32) | 140 (46.82) | 616 (45.19) | 166 228 (49.34) |
| Males | 171 615 (50.68) | 159 (53.18) | 747 (54.81) | 170 709 (50.66) |
| Living situation, n (%) | | | | |
| With partner | 152 526 (45.05) | 51 (17.06) | 270 (19.81) | 152 205 (45.17) |
| Without partner | 108 124 (31.93) | 78 (16.09) | 423 (31.03) | 107 623 (31.94) |
| Care homes | 1 165 (0.34) | 55 (18.39) | 130 (9.54) | 980 (0.29) |
| Other collectivities (no care home) | 7 032 (2.08) | 32 (10.70) | 84 (6.16) | 6 916 (2.05) |
| Other | 69 752 (20.60) | 83 (27.76) | 456 (33.46) | 69 213 (20.54) |
| Migration background, n (%) | | | | |
| Belgian natives | 24 831 (7.33) | 92 (30.77) | 427 (31.33) | 24 312 (7.22) |
| Second-generation migrants | 10 358 (3.06) | 22 (7.36) | 95 (6.97) | 10 241 (3.04) |
| First-generation migrants | 303 410 (89.61) | 185 (61.87) | 841 (61.70) | 302 384 (89.74) |

**Supplementary Table 18. Adjusted subdistribution hazard ratios (aSHRs) and their 95% confidence intervals (CIs) from Fine and Gray competing risk models for demographic and socioeconomic predictors of (1) COVID-19 death accounting for all other causes of death as competing events; and (2) all other causes of death accounting for COVID-19 death as competing events, computed on a 50% representative random sample of the total study population.**

| | COVID-19 specific death* | | All other causes of death** | |
|---|---|---|---|---|
| | **aSHR [95%CI]** | **P-value** | **aSHR [95%CI]** | **P-value** |
| Age group (in years) | | | | |
| 25-64 | 1 | 1 | 1 | 1 |
| 65-84 | 18.26 [16.83 – 19.81] | <0.001 | 8.35 [8.05 – 8.66] | <0.001 |
| 85+ | 63.06 [57.57 – 69.09] | <0.001 | 29.78 [28.53 – 31.07] | <0.001 |
| Sex | | | | |
| Females | 1 | 1 | 1 | 1 |
| Males | 1.99 [1.92 - 2.08] | <0.001 | 1.76 [1.72 – 1.81] | <0.001 |
| Living situation | | | | |
| With partner | 1 | 1 | 1 | 1 |
| Without partner | 1.56 [1.49 – 1.65] | <0.001 | 1.61 [1.56– 1.65] | <0.001 |
| Care homes | 9.58 [9.00 – 10.21] | <0.001 | 5.04 [4.84 – 5.23] | <0.001 |
| Other collectivities (no care home) | 6.67 [5.95 – 7.48] | <0.001 | 3.57 [3.28 – 3.88] | <0.001 |
| Other | 1.55 [1.41 – 1.72] | <0.001 | 1.46 [1.38 – 1.54] | <0.001 |
| Migration background | | | | |
| Belgian natives | 1 | 1 | 1 | 1 |
| Second-generation migrants | 0.98 [0.92 – 1.06] | 0.078 | 0.92 [0.89 – 0.97] | 0.001 |
| First-generation migrants | 1.25 [1.18 – 1.33] | <0.001 | 0.82 [0.78 – 0.85] | <0.001 |

**Supplementary Table 18. Continued**

| | COVID-19 specific death* | | All other causes of death** | |
|---|---|---|---|---|
| | **aSHR [95%CI]** | **P-value** | **aSHR [95%CI]** | **P-value** |
| Income | | | | |
| High | 1 | 1 | 1 | 1 |
| Low | 1.30 [1.22 – 1.39] | <0.001 | 1.47 [1.42 – 1.53] | <0.001 |
| Middle | 1.19 [1.12 – 1.28] | <0.001 | 1.28 [1.23 – 1.33] | <0.001 |
| Missing | 1.25 [1.10 – 1.44] | 0.011 | 1.45 [1.34 – 1.56] | <0.001 |
| Education | | | | |
| Higher education | 1 | 1 | 1 | 1 |
| Primary or less | 1.63 [1.52 – 1.76] | <0.001 | 1.46 [1.39 – 1.52] | <0.001 |
| Lower secondary | 1.36 [1.27 – 1.48] | <0.001 | 1.25 [1.19 – 1.30] | <0.001 |
| Upper secondary | 1.17 [1.08 – 1.27] | <0.001 | 1.10 [1.05 – 1.15] | <0.001 |
| Missing | 1.42 [1.30 – 1.55] | <0.001 | 1.25 [1.18 – 1.32] | <0.001 |

** Individuals who died from COVID-19 were considered as cases (N=10 973) and survival (N=4 064 475) was considered as the competing event. Individuals who died from all OCOD (N=51 868) were excluded. All SHRs were fully adjusted for all SD and SE variables.*

*** Individuals who died from all OCOD were considered as cases (N=51 868) and all OCOD (N=10 973) were considered as the competing event. Individuals who survived (N=10 973) were excluded. All SHRs were fully adjusted for all SD and SE variables.*

# SECTION III

# What are the long-term direct health impacts of severe COVID-19 and related social patterns?

# *Chapter 4. Long-term organ complications following severe COVID-19 and related social health inequalities*

***Adapted from***

**Lisa Cavillot**, Laura Van den Borre, Jinane Ghattas, Elin Boiy, Pierre Hubin, Delphine De Smedt, Joris van Loenhout, Robby De Pauw, Niko Speybroeck & Brecht Devleesschauwer. **Post-acute organ complications following COVID-19 hospitalization and related socioeconomic inequalities.** Submitted to *Nature Communications* (May 2025).

## Abstract

Post-acute organ complications following COVID-19 hospitalization and potential socioeconomic inequalities therein are understudied. Here, we use individual-level data from three national health and social registries to assess whether COVID-19 hospitalization increases the risk of one-year post-acute organ complications in Belgian hospitalized adults without preexisting conditions affecting the specific organ system under study at baseline. In addition, we identify socioeconomic patterns in the development of these organ complications. Overlap propensity score-weighted odds ratios and adjusted odds ratios were estimated. All analyses were stratified between severe and critical COVID-19, with the latter defined by intensive care unit admission and/or the onset of acute respiratory distress syndrome. The results highlight significant cardiovascular and pulmonary damage within one-year following COVID-19 hospitalization compared to non-COVID-19 hospitalized controls, increasing among critical COVID-19 patients. Among severe COVID-19 patients, those with low income, compared to those with high income, had higher odds of post-acute pulmonary complications. Long-term care should prioritize monitoring organ complications, especially after critical illness.

## 1. Background

The COVID-19 pandemic, caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has had an unprecedent impact on our healthcare system with more than 4.8 million cases, 150,000 hospitalizations, and 34,000 deaths recorded in Belgium by 1st January 2024 [25].

The virus mainly affects the upper respiratory tract and causes cold-like symptoms in most individuals, with several experiencing mild to moderate illness and developing symptoms such as fever, diarrhea, headaches, coughing, and difficulty breathing. When SARS-CoV-2 virus replicates in the lower respiratory tract, pulmonary alveoli are infected which can lead to a severe form of the disease characterized by dyspnea and hypoxemia, progressively resulting in respiratory failure among critical cases [15,341]. While the primary cause of COVID-19 is viral replication in the lungs, inflammation triggered by the virus can also lead to other acute organ failures [1], such as acute myocardial, renal, and liver injuries requiring hospital care [342–345].

Beyond acute complications, some individuals may develop post-acute health issues several months following SARS-CoV-2 infection. The pathogenesis of post-acute COVID-19-related health complications is not yet fully understood. Proposed pathophysiological mechanisms include direct viral damage, inflammation, autoimmune responses, and endothelial injury caused by critical SARS-CoV-2

infection [76,78]. These processes can affect multiple organ systems, leading to mid- or long-term sequelae, commonly known as '*Post-acute sequelae of COVID-19*' (PASC) or *'Long COVID'*. The World Health Organization (WHO) defined PASC as "persistent or new symptoms three months after initial SARS-CoV-2 infection, lasting at least two months and with no other underlying explanation" [68].

Symptoms such as fatigue, post-exertional malaise, cognitive impairment, chest pain, anxiety/depression, and headache can persist up to several months following the infection and are particularly prevalent among individuals who developed a critical form of the disease [71,72]. Indeed, in a meta-analysis on 194 studies, the pooled prevalence of PASC among COVID-19 hospitalized survivors at 126 days was 52.6%, compared to 34.5% among non-hospitalized COVID-19 survivors [72]. Several studies suggest that, compared to SARS-CoV-2 test-negative controls, COVID-19 hospitalized patients faced increased risks of post-acute cardiovascular (CV) (e.g., dysrhythmias, inflammatory heart diseases, or thrombotic disorders) and pulmonary (e.g., decreased diffusion capacity) complications up to six months after COVID-19 exposure, even in the absence of preexisting CV or pulmonary conditions at baseline [346–348]. However, one-year follow-up studies focusing on COVID-19 hospitalized patients remain limited, particularly those assessing post-acute renal and liver complications, which are significant acute health complications following SARS-CoV-2 infection [349,350]. This study is particularly relevant in the Belgian

context as Belgium was heavily impacted by COVID-19 [25], has an important proportion of aging population vulnerable to long-term complications [351], and a substantial burden of CV diseases [29].

Although COVID-19 is a respiratory disease caused by a viral infection, the social component of the pandemic were quickly put forward by qualifying it as a syndemic, highlighting both biological and social interactions [105]. Indeed, beyond underlying health conditions and age, which are primary risk factors for COVID-19 critical outcomes [352], sociodemographic (SD) and socioeconomic (SE) characteristics were identified as additional risk factors. Men, individuals from migrant descent and those experiencing SE hardship (e.g., higher poverty level, living in densely populated areas, lower education) presented higher COVID-19 mortality and hospitalization rates [98,353–355].

The latter might be at higher risk of critical COVID-19 outcomes due to higher vulnerability (e.g., pre-existing health issues may increase COVID-19 severity), susceptibility (e.g., poor living conditions and chronic stress weaken the immune system), exposure (e.g., higher rates of individuals working in essential sectors during the pandemic limiting remote work, lower rates of preventive measures adoption including vaccination), transmission (e.g., higher rates of individuals living in overcrowded households and neighborhoods), and reduced access to and quality of healthcare [105,356]. Despite a clear social pattern identified in COVID-19

hospitalization, there is a dearth of knowledge on SE inequalities in the development of PASC among hospitalized COVID-19 survivors.

This study aims to address these gaps by considering the two following objectives: 1) identify whether COVID-19 hospitalization is associated with the development of one-year post-acute CV, pulmonary, renal, and liver complications compared to non-COVID-19 hospitalized controls, both without preexisting conditions affecting the specific organ system under study at baseline; and 2) identify the social patterns of one-year post-acute CV and pulmonary complications following hospitalization among COVID-19 patients without preexisting conditions affecting the specific organ system under study at baseline. Those post-acute organ complications were chosen as these anatomical areas are well-known to be affected during the acute phase of SARS-CoV-2 infection, as previously established in the international literature [342–345].

## 2. Materials and Methods

### 2.1. Study design

This study relies on data from the HELICON project, a population data linkage set up by Sciensano, the Belgian National Institute for Health, to study the long term health impact of the COVID-19 crisis in Belgium [148]. The HELICON project contains selected variables from three national health and social registries linked at the individual level using the pseudonymized Belgian social security number.

To build the HELICON cohort, Statistics Belgium (Statbel) provided SD and SE data for a 10% random sample of the adult population legally residing in Belgium as of 1st January 2020. Next, a link was made with the Clinical Hospital Survey (CHS) database, containing patient-level data on approximately 50% of the patients admitted with a confirmed COVID-19 diagnosis across all 103 Belgian hospitals as of 1st July 2020. Statbel then completed the sample by adding SD and SE data for patients registered in the CHS database but not included in the initial 10% random sample. Finally, data from the InterMutualistic Agency (IMA) were linked for both the random population sample and the hospitalized patients. IMA provided two datasets: one including information on reimbursed healthcare services and medications dispensed in public pharmacies over a year; and one recording all hospitalization stays per patient. As the health insurance system is compulsory in Belgium, IMA covers more than 99% of the total Belgian population [139]. Within the HELICON project, IMA datasets were available for the reference years 2020 and 2021. The protocol of the HELICON population data linkage has been published elsewhere [148].

### 2.2. Study population and study period

**Figure 27** shows the flowchart of the study population within the HELICON project. Between 1 July 2020 (start of COVID-19 hospital case identification) and 31 December 2020, a total of 108,655 individuals were hospitalized with a confirmed

COVID-19 diagnosis, as recorded in the CHS database. Based on IMA data, individuals hospitalized during the same period but not recorded in the CHS database were considered as controls (N=64,535). This time frame was selected to enable the assessment of one-year post-acute organ complications using IMA data for the reference year 2021. Specifically, we selected hospitalized individuals without COVID-19 as the control group to isolate the effect of COVID-19 itself on the development of post-acute organ complications, independent of the impact of hospitalization. Selecting community controls from Statbel could have resulted in attributing post-acute complications to the hospitalization itself, rather than isolating the specific effects of COVID-19 on our outcomes of interest.

Among patients hospitalized with a confirmed COVID-19 diagnosis, we excluded those with missing week and year of admission, those not tested for COVID-19 symptoms (e.g., systematic screening), those readmitted or transferred, those who died in the hospital, those whose diagnosis was not confirmed by a positive PCR or antigen test, those discharged the same day as admission, and those with missing values for the admission in intensive care unit (ICU) and the onset of acute respiratory distress syndrome (ARDS). This resulted in 8,024 unique hospitalized COVID-19 survivors meeting the in- and exclusion criteria.

Next, both COVID-19 hospitalized patients and non-COVID-19 hospitalized controls were linked to Statbel data to obtain SD and SE information for the reference year

2020. In addition, only COVID-19 hospitalized patients and controls who were still alive in 2021 were included to avoid misclassifying death as the absence of post-acute organ complications in 2021 following hospitalization in 2020.

To accurately assess the onset of one-year post-acute organ complications, we created subcohorts of both COVID-19 hospitalized patients and non-COVID-19 hospitalized controls without preexisting CV, pulmonary, renal, and liver conditions at baseline, respectively, using IMA data for the reference year 2020.

Given the association between disease severity and the onset of post-acute organ complications, each subcohort of COVID-19 hospitalized patients was stratified into:

- *Critical COVID-19 patients:* defined as any COVID-19 hospitalized patients admitted in ICU and/or with ARDS;
- *Severe COVID-19 patients*: defined as any COVID-19 hospitalized patients not admitted in ICU and without ARDS.

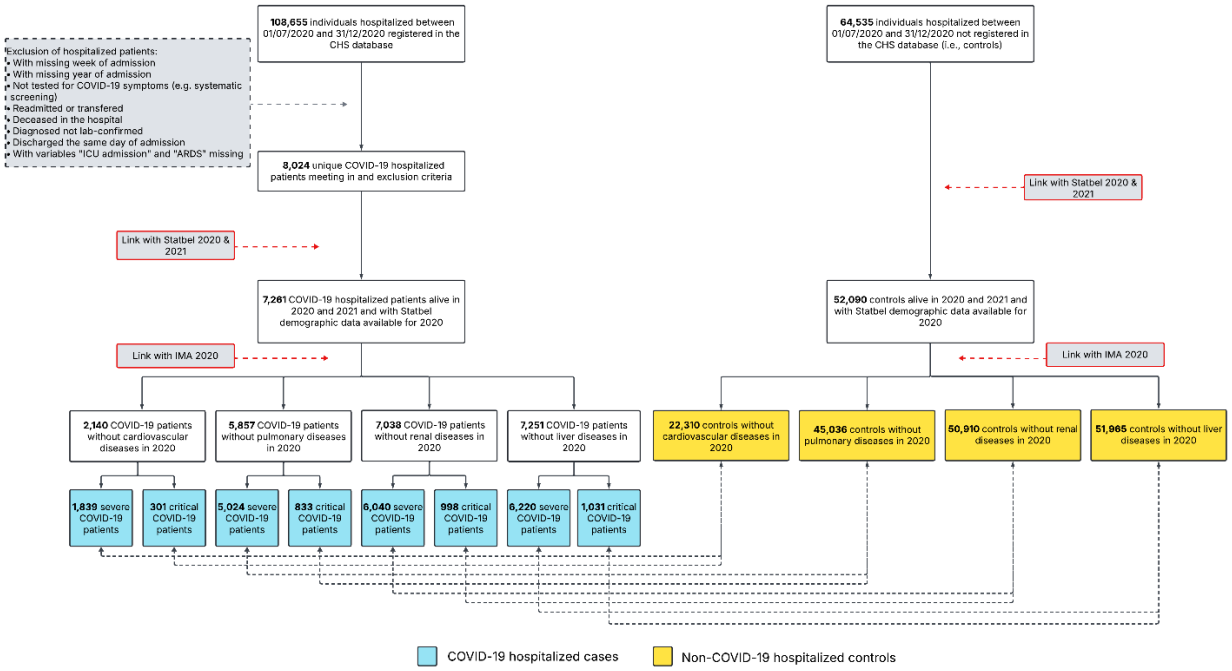



***Figure 27. Flowchart of the study population building on the HELICON population data linkage to assess the development of one-year post-acute organ complications following COVID-19 hospitalization among severe and critical COVID-19 hospitalized patients compared non-COVID-19 hospitalized controls, both without those organ conditions at baseline.*** *Critical COVID-19 hospitalized patients were defined as any patients admitted in intensive care unit (ICU) and/or with acute respiratory distress syndrome (ARDS). Severe COVID-19 patients were defined as any COVID-19 hospitalized patients not admitted in ICU and without ARDS. The dotted lines represent the comparison between the subcohorts of COVID-19 patients and their respective controls. Abbreviations: CHS = Clinical Hospital Survey; IMA = InterMutualistic Agency; Statbel = Statistic Belgium.*

## 2.3. Variables

### *2.3.1. Outcomes*

The study outcomes consisted of the onset of post-acute CV, pulmonary, renal, and liver complications, as recorded in the IMA database for 2021, among COVID-19 hospitalized patients compared to non-COVID-19 hospitalized controls. For each organ system analyzed, hospitalized patients with preexisting conditions affecting that specific system at baseline (as recorded in the IMA database for 2020) were excluded, while preexisting conditions in other organ systems were retained.

The post-acute organ complications considered in this study were based on pseudopathology groups defined by the National Institute for Health and Disability Insurance (RIZIV-INAMI). These groups are created on the basis of Anatomical Therapeutic Chemical (ATC) codes established by the World Health Organization [141]. An individual is assigned to a pseudopathology group when the total number of defined daily doses (estimated quantity of active ingredient for a 70 kg adult per day of a drug prescribed for its main action) for all ATC codes in this group is greater than or equal to 90.

A detailed description of the pseudopathologies, along with the related ATC codes and substances used to define the four post-acute organ complications considered in this study, is available in **Supplementary Table 19**. The methodology used to build the pseudopathology groups has been described by IMA elsewhere [140].

### *2.3.2. Covariates*

SD covariates included age group, sex, region of residence, migration background, and living situation. The Region of residence distinguished between Brussels Capital Region, the Flemish Region, and the Walloon Region and was included as confounder in the analyses because of the difference in hospitalization coverage by regions in the CHS database, as well as the known association between SE status and Regions in Belgium [259,261]. Migration background was based on a combination of the first nationality and the parent's country of origin and distinguished between 'Belgian natives', 'Second-generation migrants', 'First-generation European migrants', and 'First-generation non-European migrants'. The living situation gave a partial picture of the social environment and distinguished between 'Couples with children', 'Couples without children', 'Single parents', 'Single persons household', 'Collective households' (e.g., nursing homes, prisons), and 'Other' (e.g., adult children still living with their parents, other co-residents).

SE covariates included education and income levels. Education level was classified into five categories based on the International Standard Classification of Education (ISCED): 'Primary or less' (ISCED 0 to ISCED 1), 'Lower secondary' (ISCED 2), 'Upper secondary' (ISCED 3 to ISCED 4), 'Higher education' (ISCED 5 to ISCED 8), and 'Missing'. Income data was available as the deciles of the yearly net taxable income

per person and grouped as: ‘Low income’ (deciles 1 to 4), ‘Middle income’ (deciles 5 to 7), ‘High income’ (deciles 8 to 10), and ‘Missing’.

Health-related covariates consisted of comorbidities at baseline, defined based on the pseudopathology groups available in the IMA database for the reference year 2020, as described above. The following comorbidities were considered: ‘Cardiovascular comorbidities’, ‘Pulmonary comorbidities’, ‘Renal comorbidities’, ‘Liver comorbidities’, and ‘Other comorbidities’. The latter category included the following pseudopathologies: mucoviscidosis, morbid obesity, organ transplantation, immunodeficiency (HIV and no HIV), neurological disorders (e.g., Parkinson, multiple sclerosis, Alzheimer), psychosis, diabetes, and inflammatory and autoimmune disorders (e.g., rheumatoid arthritis, Crohn’s diseases).

A description of the data holders and time frame for all SD, SE, and health-related variables considered in this study is available in **Supplementary Table 20**.

*2.3.3. Exposure*

In the first objective, which aimed to determine whether COVID-19 hospitalization led to the onset of post-acute organ complications compared to non-COVID-19 hospitalization, the exposure was defined as hospitalization with a confirmed COVID-19 diagnosis based on a positive PCR or antigen test between 1 July 2020 and 31 December 2020 in Belgium.

In the second objective, aiming to identify SE inequalities in the development of post-acute organ complications among COVID-19 patients hospitalized during the same period, education and income were considered as the SE exposures of interest. In both objectives, COVID-19 hospitalization was stratified between severe and critical COVID-19 hospitalization, as defined above.

**Figure 28** presents a schematic overview of the outcome, exposure, covariates, and study population according to each objective considered in this study.

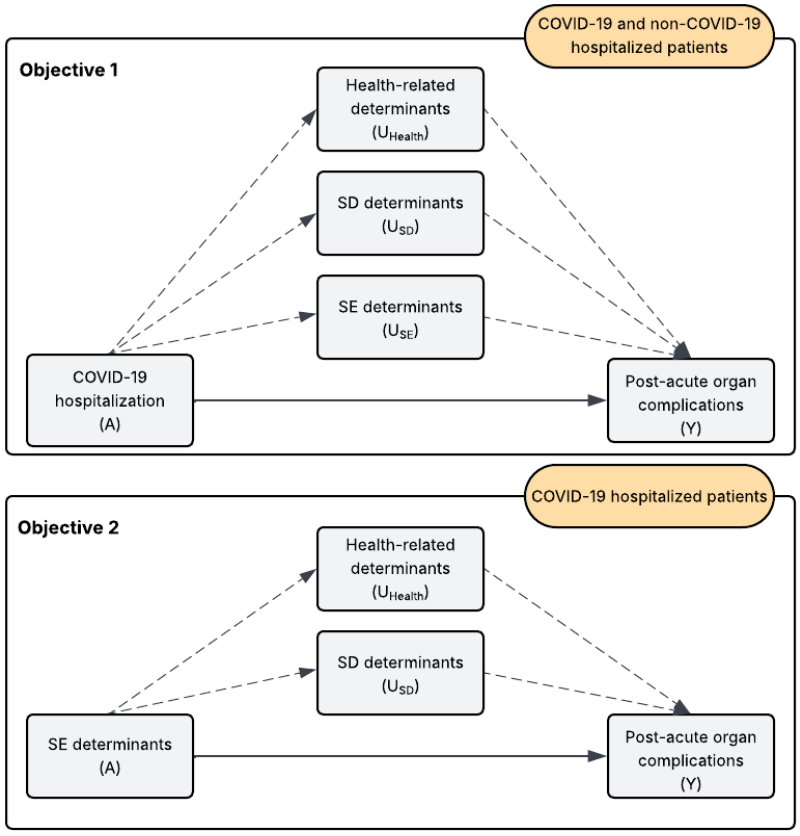



***Figure 28.*** **Schematic overview of the outcome (Y), exposure (A), covariates (U), and study population considered in each objective.** *Abbreviation: SD = sociodemographic ; SE = socioeconomic*

## 2.4. Statistical analyses

### *2.4.1. Objective 1: Impact of COVID-19 hospitalization on the onset of post-acute organ complications*

First, severe and critical COVID-19 hospitalized patients, along with non-COVID-19 hospitalized controls, were described with the frequency rates and percentages per categorical variables.

Second, propensity scores (PS) with overlap weights (OW) were applied to address imbalances in the characteristics between exposed and unexposed individuals. For each individual, the PS model was estimated as follows:

$$logit\ P(A \mid U_{SD}, U_{SE}, U_{Health}) = \beta_0 + \beta_{U_{SD}} U_{SD} + \beta_{U_{SE}} U_{SE} + \beta_{U_{Health}} U_{Health} \quad \textbf{\textit{(Eq. 13 )}}$$

where *A* is COVID-19 hospitalization, and $U_{SD}$, $U_{SE}$ and $U_{Health}$ denote the SD, SE, and health-related covariates, respectively.

In our study population, the distribution of PS among COVID-19 hospitalized patients and non-COVID-19 hospitalized controls showed evidence of extreme values across all patient subcohorts (see **Supplementary Figures 9** to **12** for patient subcohorts without CV, pulmonary, renal, and liver diseases at baseline, respectively).

This is a common limitation of balancing weight methods based on PS, as extreme values can inflate variance and bias estimates when using the standard inverse probability of treatment weighting (IPTW) method [180,181,357]. To address this issue, OW were applied, which are known to provide an accurate estimate of the exposure effect and its variance even in the presence of extreme PS values [180,181,357]. OW, denoted $ow_i$, are defined as $(1 - PS)$ for exposed individuals and $PS$ for unexposed individuals. This approach down-weights observations in the tails of the PS distribution and assigns greater weight to those in regions of overlap, creating a weighted pseudo-population where confounders are balanced [180,181]. The effectiveness of OW in achieving covariate balance was evaluated using weighted standardized mean difference (SMD) [358]. Importantly, OWs focus the causal estimates on the overlap population, meaning the results primarily apply to individuals who are sufficiently similar to plausibly belong to either treatment group, rather than to the entire study population [185].

Overlap PS-weighted odds ratios (ORs) and 95% confidence intervals (CIs) were then estimated using multivariable binary logistic regression models. The outcome model was defined as follows:

$$logit\ P(Y_i \mid A) = \beta_{0i} + \beta_{Ai}A \quad \textbf{\textit{(Eq. 14)}}$$

where $Y_i$ denotes the occurrence of a specific post-acute organ complication ($i$=CV, pulmonary, renal, or liver), and each observation was weighted by $ow_i$.

To address potential PS model misspecification, we conducted an adjusted analysis on the weighted sample (**Eq. 15**), which literature suggests can enhance robustness beyond weighting alone as long as either the outcome model or the PS model was correctly specified [359].

$$logit\ P(Y_i \mid A, U_{SD}, U_{SE}, U_{Health}) = \beta_{0i} + \beta_{Ai}A + \beta_{Ui_{SD}}U_{SD} + \beta_{Ui_{SE}}U_{SE} + \beta_{Ui_{Health}}U_{Health} \quad \textbf{\textit{(Eq. 15)}}$$

with each observation weighted by $ow_i$. However, this approach does not guaranty doubly robust estimation, which requires specific methods, such as fitting a canonical link GLM fit via maximum likelihood with weighted score equations, followed by regression standardization [187]. Variance estimation was performed using robust sandwich variance estimates, shown to produce valid estimates even in the case of model misspecification [360].

Third, overlap PS-weighted excess burdens of one-year post-acute organ complications per 1,000 individuals due to COVID-19 hospitalization were computed by estimating the difference in incidence rate of post-acute organ complications between COVID-19 hospitalized patients and non-COVID-19 hospitalized controls.

All analyses described above were applied separately to severe COVID-19 patients and critical COVID-19 patients and their controls.

### *2.4.2. Objective 2: Socioeconomic patterns in the development of post-acute organ complications among COVID-19 hospitalized patients*

The second objective aimed to identify SE patterns in the development of post-acute CV and pulmonary complications among COVID-19 hospitalized patients.

Post-acute renal and liver complications were excluded due to insufficient event rates in certain covariate and exposure categories preventing reliable convergence of estimates when limiting the analysis to COVID-19 patients.

Adjusted odds ratios (aORs) and 95% CIs were estimated using multivariable binomial logistic regression models to assess the association between post-acute CV and pulmonary complications and the SE exposures of interest (i.e., education, income) among COVID-19 hospitalized patients, respectively. The analyses described above were stratified between stable and critical COVID-19 patients.

### *2.4.3. Sensitivity analyses*

As sensitivity analyses, we aimed to identify whether the impact of SE determinants on the development of post-acute CV and pulmonary complications differed among COVID-19 hospitalized patients compared to non-COVID-19 hospitalized controls, respectively.

First, COVID-19 patients and controls were stratified and aORs and 95% CIs were estimated using multivariable binary logistic regression models. Next, multivariable

logistic regression models were applied on the full study population, including both COVID-19 patients and controls. In these models, an interaction term between COVID-19 hospitalization and the SD, SE, and health-related covariates was introduced. This allowed to extract the p-value for interaction testing $H_0$ that the effect of covariates on the outcomes is similar in both groups. These regression models with an interaction term were applied separately to severe COVID-19 patients and controls, as well as to critical COVID-19 patients and controls.

### 2.5. Analytical approach and tools

All analyses were performed in R (version 4.3.2) [221]. In our analyses, we set a significance level α of 5%. The package "tableone" was used to calculate the SMD [358]. The package "sandwich" was used to compute robust variance estimates [361].

## 3. Results

### 3.1. Baseline characteristics

We included 1,839, 5,024, 6,040, and 6,220 severe COVID-19 hospitalized patients without CV, pulmonary, renal, and liver conditions at baseline, respectively. Regarding critical COVID-19 hospitalized patients, 301, 833, 998, and 1,031 without CV, pulmonary, renal, and liver conditions at baseline were included, respectively. The corresponding control groups consisted of 22,310, 45,036, 50,910, and 51,965

non-COVID-19 hospitalized controls without CV, pulmonary, renal, and liver diseases at baseline, respectively.

The SD, SE, and health-related characteristics of severe and critical COVID-19 hospitalized patients and controls before weighting, along with the SMD assessing the effectiveness of OWs in achieving covariate balance, are provided in **Supplementary Tables 21** to **24**. The weighted SMDs showed small differences (i.e., <0.001) in baseline characteristics after weighting for both severe and critical COVID-19 patients and their respective controls (**Supplementary Tables 21** to **24**).

3.2. Objective 1 : Impact of COVID-19 hospitalization on the onset of post-acute organ complications

**Table 21** shows overlap PS-weighted ORs and doubly robust overlap PS-weighted ORs, along with their 95% CIs, estimated using binomial logistic regression models, assessing the development of one-year post-acute organ complications among severe COVID-19 hospitalized patients compared to non-COVID-19 hospitalized controls.

One-year post-acute organ complications were more common in severe COVID-19 patients than in controls, with 14.8%, 6.5%, 0.7%, and 0.21% of patients developing post-acute CV, pulmonary, renal and liver complications within one year following COVID-19 hospitalization, respectively, compared to 8.7%, 3%, 0.6%, and 0.20% of controls over the same duration.

Compared to non-COVID-19 hospitalized controls, severe COVID-19 hospitalized patients had significantly higher risks of developing post-acute CV (overlap PS-weighted OR 1.19, 95% CI 1.03 – 1.37) and pulmonary (overlap PS-weighted OR 2.05, 95% CI 1.80 – 2.34) complications within one year following hospitalization. However, the risks of developing one-year post-acute renal (overlap PS-weighted OR 0.96, 95% CI 0.68 – 1.35) and liver (overlap PS-weighted OR 0.84, 95% CI 0.46 – 1.52) complications following severe COVID-19 hospitalization were not statistically significant. Overlap PS-weighted aORs, adjusting for potential misspecification in the PS model that could bias the estimation of the exposure's effect on the outcome, yielded similar conclusions (**Table 21**). These overlap PS-weighted causal estimates focus on the subgroup of patients with comparable baseline characteristics who could plausibly have been hospitalized with or without COVID-19. As such, the results are generalizable to this overlap population, rather than to the entire study cohort.

**Table 21. Overlap propensity score (PS)-weighted odds ratios (ORs) and overlap PS-weighted adjusted OR (aOR), estimated using binomial logistic regression models to assess the development of one-year post-acute cardiovascular, pulmonary, renal and liver complications, respectively, among severe COVID-19 hospitalized patients compared to non-COVID-19 hospitalized controls, both without preexisting conditions affecting the specific organ system under study at baseline.**

| Post-acute health outcomes | N severe COVID-19 patients | Severe COVID-19 patients with outcome N (%) | N controls | Controls with outcome N (%) | overlap PS-weighted OR [95% CI] | Overlap PS-weighted aOR [95% CI] |
|---|---|---|---|---|---|---|
| Cardiovascular complications [a] | 1,839 | 272 (14.79) | 22,310 | 1,937 (8.68) | **1.19 [1.03 – 1.37]** | **1.20 [1.04 – 1.39]** |
| Pulmonary complications [b] | 5,024 | 326 (6.49) | 45,036 | 1,332 (2.96) | **2.05 [1.80 – 2.34]** | **2.06 [1.81 – 2.35]** |
| Renal complications [c] | 6,040 | 40 (0.66) | 50,910 | 302 (0.59) | 0.96 [0.68 – 1.35] | 0.67 [0.95 – 1.33] |
| Liver complications [d] | 6,220 | 13 (0.21) | 51,965 | 102 (0.20) | 0.84 [0.46 – 1.52] | 0.86 [0.47 – 1.56] |

*Doubly robust overlap PS-weighted ORs are adjusted for sociodemographic (age groups, sex, region of residence, living situation, migration background), socioeconomic (education, income) and health-related (cardiovascular comorbidities, pulmonary comorbidities, renal comorbidities, liver comorbidities, other comorbidities) variables. Severe COVID-19 patients were defined as any COVID-19 patients not admitted in ICU and without ARDS. Significant results are highlighted in bold. [a] The weighted logistic regression model were applied on severe COVID-19 patients and controls both without preexisting cardiovascular conditions at baseline (N=24,149). [b] The weighted logistic regression model were applied on severe COVID-19 patients and controls both without preexisting pulmonary conditions at baseline (N=50,060). [c] The weighted logistic regression models were applied on severe COVID-19 patients and controls both without preexisting renal conditions at baseline (N=56,950). [d] The weighted logistic regression model were applied on severe COVID-19 patients and controls both without preexisting liver conditions at baseline (N=58,185).*

**Table 22** shows overlap PS-weighted ORs and overlap PS-weighted aORs, along with their 95% CIs, estimated using binomial logistic regression models, assessing the development of one-year post-acute organ complications among critical COVID-19 hospitalized patients compared to non-COVID-19 hospitalized controls.

Overall, most post-acute organ complications were more common among critical COVID-19 hospitalized patients than among controls, with 21.7%, 7%, 1%, and 0.1% of critical COVID-19 patients developing post-acute CV, pulmonary, renal, and liver complications, respectively, compared to 8.7%, 3%, 0.6% and 0.2% of controls.

Compared to non-COVID-19 hospitalized controls, critical COVID-19 patients had significantly higher risks of developing post-acute CV (overlap PS-weighted OR 1.93, 95% CI 1.36 – 2.42) and pulmonary (overlap PS-weighted OR 2.72, 95% CI 1.98 – 3.79) complications within one year following hospitalization. However, the risks of developing post-acute renal (overlap PS-weighted OR 1.54, 95% CI 0.81 – 2.94) and liver (overlap PS-weighted OR 0.41, 95% CI 0.06 – 2.96) complications following critical COVID-19 hospitalizing were not statistically significant. Overlap PS-weighted aORs gave similar conclusions (**Table 22**). These overlap PS-weighted causal estimates focus on the subgroup of patients with comparable baseline characteristics who could plausibly have been hospitalized with or without COVID-19. As such, the results are most generalizable to this overlap population, rather than to the entire study cohort.

**Table 22. Overlap propensity score (PS)-weighted odds ratios (ORs) and overlap PS-weighted adjusted OR (aOR), along with their 95% confidence intervals (CI), estimated using binomial logistic regression models to assess the development of one-year post-acute cardiovascular, pulmonary, renal, and liver complications, respectively, among critical COVID-19 hospitalized patients compared to non-COVID-19 hospitalized controls, both without preexisting conditions affecting the specific organ system under study at baseline.**

| Post-acute health outcomes | N severe COVID-19 patients | Severe COVID-19 patients with outcome N (%) | N controls | Controls with outcome N (%) | overlap PS-weighted OR [95% CI] | Overlap PS-weighted aOR [95% CI] |
|---|---|---|---|---|---|---|
| Cardiovascular complications [a] | 301 | 65 (21.69) | 22,310 | 1,937 (8.68) | **1.93 [1.36 – 2.42]** | **1.41 [1.88 – 2.51]** |
| Pulmonary complications [b] | 833 | 58 (6.96) | 45,036 | 1,332 (2.95) | **2.72 [1.98 – 3.79]** | **2.71 [1.94 – 3.85]** |
| Renal complications [c] | 998 | 10 (1.00) | 50,910 | 302 (0.59) | 1.54 [0.81 – 2.94] | 1.54 [0.81 – 2.94] |
| Liver complications [d] | 1,031 | 1 (0.10) | 51,965 | 102 (0.20) | 0.41 [0.06 – 2.96] | 0.41 [0.06 – 2.97] |

*Doubly robust overlap PS-weighted ORs are adjusted for sociodemographic (age groups, sex, region of residence, living situation, migration background), socioeconomic (education, income) and health-related (cardiovascular comorbidities, pulmonary comorbidities, renal comorbidities, liver comorbidities, other comorbidities) variables. Severe COVID-19 patients were defined as any COVID-19 patients not admitted in ICU and without ARDS. Significant results are highlighted in bold. [a] The weighted logistic regression model were applied on severe COVID-19 patients and controls both without preexisting cardiovascular conditions at baseline (N=24,149). [b] The weighted logistic regression model were applied on severe COVID-19 patients and controls both without preexisting pulmonary conditions at baseline (N=50,060). [c] The weighted logistic regression models were applied on severe COVID-19 patients and controls both without preexisting renal conditions at baseline (N=56,950). [d] The weighted logistic regression model were applied on severe COVID-19 patients and controls both without preexisting liver conditions at baseline (N=58,185).*

**Figure 29** shows the overlap PS-weighted excess burdens per 1,000 persons of one-year post-acute organ complications attributable to severe and critical COVID-19 hospitalization, respectively.

The weighted excess burden of post-acute CV complications was substantially higher for critical COVID-19 hospitalization (weighted excess burden 84.83, 95% CI 36.89 – 132.77) compared to severe COVID-19 hospitalization (weighted excess burden 20.16, 95% CI 2.34 – 37.98). Similarly, a slightly higher excess burden of post-acute pulmonary complications was observed following critical COVID-19 hospitalization (weighted excess burden 38.7, 95% CI 21.35 – 56.05) compared to severe COVID-19 hospitalization (weighted excess burden 32.36, 95% CI 25.15 – 39.57).

However, the excess burdens of post-acute renal complications were not statistically significant following either severe (weighted excess burden -0.28, 95% CI -2.54 – 1.98) or critical (weighted excess burden 3.52, 95% CI -2.77 – 9.81) COVID-19 hospitalizations. Similarly, the excess burdens of post-acute liver complications due to either severe (weighted excess burden -0.4, 95% CI -1.67 – 0.87) or critical (weighted excess burden -1.45, 95% CI -3.63 – 0.63) COVID-19 hospitalizations were not statistically significant (**Figure 29**).

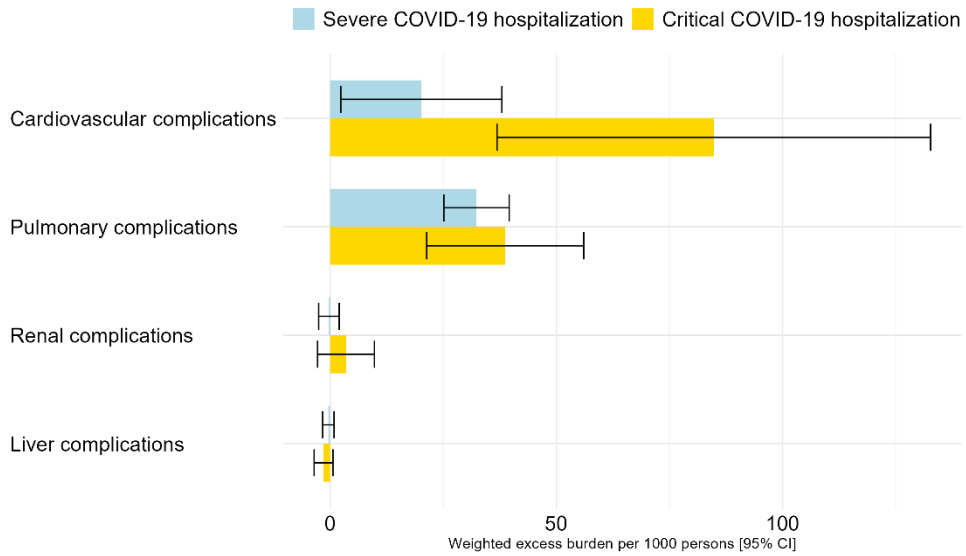



***Figure 29. Overlap propensity score-weighted excess burdens per 1,000 persons and 95% confidence intervals (CI) of one-year post-acute cardiovascular, pulmonary, renal, and liver complications among severe and critical COVID-19 patients compared to non-COVID-19 hospitalized controls, both without preexisting conditions affecting the specific organ system under study at baseline.*** *Critical COVID-19 patients were defined as any COVID-19 hospitalized patients admitted in ICU and/or with ARDS. Severe COVID-19 patients were defined as any COVID-19 hospitalized patients not admitted in ICU and without ARDS.*

## 3.3. Objective 2: socioeconomic patterns in the development of post-acute organ complications among COVID-19 hospitalized patients

**Table 23** shows the aORs and 95% CIs for the development of post-acute CV and pulmonary complications among severe COVID-19 hospitalized cases according to the SE exposures of interest (i.e., education and income).

Overall, no significant SE patterns were identified in the development of post-acute CV complications among severe COVID-19 patients. However, severe COVID-19 patients aged 45 to 64 years (aOR 1.94, 95% CI 1.34 – 2.86) and 65 to 84 years (aOR 2.97, 95% CI 1.90 – 4.69), compared to those aged 18 to 44 years, as well as those with 'other comorbidities' (aOR 2.00, 95% CI 1.48 – 2.68), had significantly increased odds of developing post-acute CV complications. Conversely, living in Brussels compared to Flanders (aOR 0.56, 95% CI 0.32 – 0.96) and having a migration background (aOR 0.61, 95% CI 0.38 – 0.97 for second generation migrants; aOR 0.54, 95% CI 0.30 – 0.93 for first-generation EU migrants; aOR 0.61, 95% CI 0.38 – 0.97 for first-generation non-EU migrants) were associated with decreased odds of developing post-acute CV complications (**Table 23**).

A SE pattern was identified in the development of post-acute pulmonary complications, with severe COVID-19 patients with low income having higher odds of developing such complications compared to those with high income (aOR 1.53, 95% CI 1.05 – 2.25). Other SD and health-related variables were not significantly associated with the development of post-acute pulmonary complications among severe COVID-19 patients (**Table 23**).

**Table 23. Fully adjusted odds ratios (aORs) and 95% confidence intervals (CIs) for the association between socioeconomic determinants and the development of post-acute cardiovascular (CV) and pulmonary complications among severe COVID-19 hospitalized patients without preexisting conditions affecting the specific organ system under study at baseline.**

| | Severe COVID-19 patients without preexisting CV diseases at baseline (N=1,839) | | | Severe COVID-19 patients without preexisting pulmonary diseases at baseline (N=5,024) | | |
|---|---|---|---|---|---|---|
| | **N** | **N (%) with post-acute CV complications** | **Post-acute CV complication aOR [95%CI]** | **N** | **N (%) with post-acute pulmonary complications** | **Post-acute pulmonary complications aOR [95%CI]** |
| **Socioeconomic exposures** | | | | | | |
| Education | | | | | | |
| High (ref.) | 414 | 50 (12.08) | 1 | 870 | 45 (5.17) | 1 |
| Middle | 512 | 72 (14.06) | 1.21 [0.81 – 1.83] | 1,110 | 75 (6.76) | 1.20 [0.82 – 1.79] |
| Low | 593 | 112 (18.89) | 1.46 [0.98 – 2.20] | 2,331 | 164 (7.04) | 1.20 [0.84 – 1.76] |
| Missing | 320 | 38 (11.88) | 1.20 [0.70 – 2.04] | 713 | 42 (5.89) | 0.75 [0.18 – 2.16] |
| Income | | | | | | |
| High (ref.) | 639 | 86 (13.46) | 1 | 1,292 | 71 (5.50) | 1 |
| Middle | 529 | 72 (13.61) | 0.99 [0.68 – 1.44] | 1,512 | 95 (6.28) | 1.18 [0.84 – 1.68] |
| Low | 574 | 94 (16.38) | 1.21 [0.80 – 1.84] | 1,903 | 138 (7.25) | **1.53 [1.05 – 2.25]** |
| Missing | 97 | 20 (20.62) | 1.94 [0.82 – 4.25] | 317 | 22 (6.94) | 0.75 [0.18 – 2.16] |
| **Sociodemographic covariates** | | | | | | |
| Age group (in years) | | | | | | |
| 18 – 44 (ref.) | 538 | 42 (7.81) | 1 | 600 | 36 (6.00) | 1 |
| 45 – 64 | 895 | 134 (14.97) | **1.94 [1.34 – 2.86]** | 1,766 | 105 (5.95) | 0.92 [0.61 – 1.41] |
| 65 – 84 | 358 | 86 (24.02) | **2.97 [1.90 – 4.69]** | 2,128 | 145 (6.81) | 0.92 [0.58 – 1.48] |

**Table 23. Continued**

| | Severe COVID-19 patients without preexisting CV diseases at baseline (N=1,839) | | | Severe COVID-19 patients without preexisting pulmonary diseases at baseline (N=5,024) | | |
|---|---|---|---|---|---|---|
| | **N** | **N (%) with post-acute CV complications** | **Post-acute CV complication aOR [95%CI]** | **N** | **N (%) with post-acute pulmonary complications** | **Post-acute pulmonary complications aOR [95%CI]** |
| 85+ | 48 | 10 (20.83) | 2.36 [0.99 – 5.27] | 530 | 40 (7.55) | 0.93 [0.53 – 1.65] |
| Gender | | | | | | |
| Males (ref.) | 1,042 | 168 (16.12) | 1 | 2,696 | 187 (6.94) | 1 |
| Females | 797 | 104 (13.05) | 0.78 [0.59 – 1.03] | 2,328 | 139 (5.97) | 0.80 [0.63 – 1.01] |
| Region of residence | | | | | | |
| Flanders (ref.) | 905 | 132 (14.59) | 1 | 428 | 18 (4.21) | 1 |
| Brussels | 202 | 18 (8.91) | **0.56 [0.32 – 0.96]** | 2,421 | 169 (6.98) | 0.65 [0.38 – 1.07] |
| Wallonia | 732 | 122 (16.67) | 1.12 [0.84 – 1.49] | 2,175 | 139 (6.39) | 0.86 [0.68 – 1.10] |
| Migration status | | | | | | |
| Belgian natives (ref.) | 1,002 | 178 (17.76) | 1 | 3,295 | 226 (6.86) | 1 |
| Second-generation migrants | 255 | 25 (9.80) | **0.61 [0.38 – 0.97]** | 465 | 34 (7.31) | 1.18 [0.78 – 1.74] |
| First-generation European migrants | 174 | 18 (10.34) | **0.54 [0.30 – 0.93]** | 434 | 27 (6.22) | 0.91 [0.58 – 1.39] |
| First-generation non-European migrants | 408 | 51 (12.50) | **0.61 [0.38 – 0.97]** | 830 | 39 (4.70) | 0.67 [0.44 – 1.00] |

**Table 23. Continued**

| | Severe COVID-19 patients without preexisting CV diseases at baseline (N=1,839) | | | Severe COVID-19 patients without preexisting pulmonary diseases at baseline (N=5,024) | | |
|---|---|---|---|---|---|---|
| | N | N (%) with post-acute CV complications | Post-acute CV complication aOR [95%CI] | N | N (%) with post-acute pulmonary complications | Post-acute pulmonary complications aOR [95%CI] |
| Household type | | | | | | |
| Couples with children (ref.) | 824 | 102 (12.38) | 1 | 1,442 | 87 (6.03) | 1 |
| Couples without children | 428 | 85 (19.86) | 1.05 [0.72 – 1.52] | 1,711 | 116 (6.78) | 0.87 [0.61 – 1.23] |
| One person household | 180 | 18 (10.00) | 0.66 [0.37 – 1.13] | 362 | 19 (5.25) | 0.79 [0.45 – 1.31] |
| Single parents | 324 | 53 (16.36) | 0.87 [0.55 – 1.36] | 1,188 | 76 (6.40) | 0.77 [0.51 – 1.17] |
| Collective households | 41 | 4 (9.76) | 0.56 [0.16 – 1.49] | 108 | 9 (8.33) | 1.30 [0.58 – 2.56] |
| Other | 42 | 10 (23.81) | 0.59 [0.19 – 1.79] | 213 | 19 (8.92) | 2.04 [0.61 – 9.38] |
| **Health-related covariates** | | | | | | |
| Cardiovascular comorbidities | | | | | | |
| No (ref.) | 1,839 | 272 (14.79) | 1 | 1,625 | 87 (5.35) | 1 |
| Yes | 0 | 0 (0.00) | NA | 3,399 | 239 (7.03) | 1.28 [0.95 – 1.73] |
| Pulmonary comorbidities | | | | | | |
| No (ref.) | 1,625 | 232 (14.28) | 1 | 5,378 | 326 (6.49) | 1 |
| Yes | 214 | 40 (18.69) | 1.09 [0.73 – 1.59] | 0 | 0 (0.00) | NA |
| Renal comorbidities | | | | | | |
| No (ref.) | 1,833 | 270 (14.73) | 1 | 4,880 | 322 (6.60) | 1 |
| Yes | 6 | 2 (33.33) | 2.68 [0.36 – 14.27] | 144 | 4 (2.78) | 0.36 [0.11 – 0.86] |

**Table 23. Continued**

| | Severe COVID-19 patients without preexisting CV diseases at baseline (N=1,839) | | | Severe COVID-19 patients without preexisting pulmonary diseases at baseline (N=5,024) | | |
|---|---|---|---|---|---|---|
| | **N** | **N (%) with post-acute CV complications** | **Post-acute CV complication aOR [95%CI]** | **N** | **N (%) with post-acute pulmonary complications** | **Post-acute pulmonary complications aOR [95%CI]** |
| Liver comorbidities | | | | | | |
| No (ref.) | 1,839 | 272 (14.79) | 1 | 5,015 | 326 (6.5) | 1 |
| Yes | 0 | 0 (0.00) | NA | 9 | 0 (0.00) | NA |
| Other comorbidities | | | | | | |
| No (ref.) | 1,446 | 181 (12.52) | 1 | 3,121 | 190 (6.09) | 1 |
| Yes | 393 | 91 (23.16) | **2.00 [1.48 – 2.68]** | 1,903 | 136 (7.15) | 1.13 [0.89 – 1.44] |

*All ORs are fully adjusted for sociodemographic, socioeconomic, and health-related variables. Severe COVID-19 patients are defined as any COVID-19 patients not admitted in ICU and without ARDS. Significant results are highlighted in bold.*

**Table 24** shows the aORs and 95% CIs for the development of post-acute CV and pulmonary complications among critical COVID-19 hospitalized patients, according to the SE exposures of interest (i.e., education and income).

No social major social patterns were identified in the development of either post-acute CV or pulmonary complications among critical COVID-19 patients. However, having missing income, compared to high income, was associated with increased odds of developing post-acute pulmonary complications (aOR 5.25, 95% CI 1.15 – 23.24). **Supplementary Table 25**, which characterizes critical COVID-19 patients with missing income by SD, SE, and health-related variables relative to the total number of critical COVID-19 patients without pulmonary diseases at baseline, showed an overrepresentation of individuals aged 85 years and over, males, first-generation non-EU migrants, and individuals with CV comorbidities, renal comorbidities, and 'other comorbidities' among those with missing income.

**Table 24. Fully adjusted odds ratios (aORs) and 95% confidence intervals (CIs) for the association between socioeconomic determinants and the development of post-acute cardiovascular (CV) and pulmonary complications among critical COVID-19 hospitalized patients without preexisting conditions affecting the specific organ system under study at baseline.**

| | Critical COVID-19 patients without preexisting CV diseases at baseline (N=301) | | | Critical COVID-19 patients without preexisting pulmonary diseases at baseline (N=833) | | |
|---|---|---|---|---|---|---|
| | **N** | **N (%) with post-acute CV complications** | **Post-acute CV complication aOR [95%CI]** | **N** | **N (%) with post-acute pulmonary complications** | **Post-acute pulmonary complications aOR [95%CI]** |
| **Socioeconomic exposures** | | | | | | |
| Education | | | | | | |
| High (ref.) | 69 | 16 (23.19) | 1 | 144 | 10 (6.94) | 1 |
| Middle | 88 | 21 (23.86) | 0.96 [0.43 – 2.16] | 214 | 10 (4.67) | 0.55 [0.21 – 1.42] |
| Low | 85 | 21 (24.71) | 0.87 [0.36 – 2.08] | 353 | 25 (7.08) | 0.84 [0.37 – 2.01] |
| Missing | 59 | 7 (11.86) | 0.55 [0.15 – 1.85] | 122 | 13 (10.66) | 1.17 [0.39 – 3.48] |
| Income | | | | | | |
| High (ref.) | 118 | 23 (19.49) | 1 | 249 | 18 (7.23) | 1 |
| Middle | 91 | 23 (25.27) | 1.32 [0.61 – 2.85] | 275 | 15 (5.45) | 0.87 [0.39 – 1.93] |
| Low | 77 | 18 (23.38) | 1.47 [0.58 – 3.69] | 264 | 20 (7.58) | 1.50 [0.66 – 3.47] |
| Missing | 15 | 1 (6.67) | 0.61 [0.03 – 4.44] | 45 | 5 (11.11) | **5.25 [1.15 – 23.24]** |
| **Sociodemographic covariates** | | | | | | |
| Age group (in years) | | | | | | |
| 18 – 44 (ref.) | 75 | 12 (16.00) | 1 | 85 | 2 (2.35) | 1 |
| 45 – 64 | 175 | 39 (22.29) | 1.42 [0.67 – 3.20] | 388 | 37 (9.54) | 3.54 [0.99 – 22.72] |
| 65 – 84 | 47 | 14 (29.79) | 1.62 [0.56 – 4.77] | 336 | 18 (5.36) | 1.82 [0.44 – 12.55] |

**Table 24. Continued**

| | Critical COVID-19 patients without preexisting CV diseases at baseline (N=1,839) | | | Critical COVID-19 patients without preexisting pulmonary diseases at baseline (N=5,024) | | |
|---|---|---|---|---|---|---|
| | **N** | **N (%) with post-acute CV complications** | **Post-acute CV complication aOR [95%CI]** | **N** | **N (%) with post-acute pulmonary complications** | **Post-acute pulmonary complications aOR [95%CI]** |
| 85+ | 4 | 0 (0.00) | NA | 24 | 1 (4.17) | 2.45 [0.10 – 30.49] |
| Gender | | | | | | |
| Males (ref.) | 217 | 49 (22.58) | 1 | 568 | 42 (7.39) | 1 |
| Females | 84 | 16 (19.05) | 0.81 [0.39 – 1.62] | 265 | 16 (6.04) | 0.85 [0.44 – 1.57] |
| Region of residence | | | | | | |
| Flanders (ref.) | 43 | 10 (23.26) | 1 | 83 | 7 (8.43) | 1 |
| Brussels | 144 | 27 (18.75) | 1.49 [0.53 – 4.03] | 399 | 24 (6.02) | 1.25 [0.45 – 3.12] |
| Wallonia | 114 | 28 (24.56) | 1.22 [0.64 – 2.34] | 351 | 27 (7.69) | 1.24 [0.67 – 2.28] |
| Migration status | | | | | | |
| Belgian natives (ref.) | 154 | 41 (26.62) | 1 | 492 | 28 (5.69) | 1 |
| Second-generation migrants | 42 | 10 (23.81) | 0.85 [0.35 – 1.99] | 91 | 9 (9.89) | 1.70 [0.70 – 3.84] |
| First-generation European migrants | 38 | 7 (18.42) | 0.60 [0.19 – 1.67] | 97 | 8 (8.25) | 0.92 [0.33 – 2.31] |
| First-generation non-European migrants | 67 | 7 (10.45) | 0.35 [0.11 – 1.00] | 153 | 13 (8.50) | 0.93 [0.38 – 2.16] |

**Table 24. Continued**

| | Critical COVID-19 patients without preexisting CV diseases at baseline (N=1,839) | | | Critical COVID-19 patients without preexisting pulmonary diseases at baseline (N=5,024) | | |
|---|---|---|---|---|---|---|
| | **N** | **N (%) with post-acute CV complications** | **Post-acute CV complication aOR [95%CI]** | **N** | **N (%) with post-acute pulmonary complications** | **Post-acute pulmonary complications aOR [95%CI]** |
| Household type | | | | | | |
| Couples with children (ref.) | 164 | 28 (17.07) | 1 | 283 | 21 (7.42) | 1 |
| Couples without children | 69 | 20 (28.99) | 1.47 [0.66 – 3.24] | 328 | 28 (8.54) | 1.13 [0.55 – 2.35] |
| One person household | 16 | 3 (18.75) | 0.74 [0.14 – 3.04] | 45 | 2 (4.44) | 0.48 [0.07 – 1.89] |
| Single parents | 46 | 13 (28.26) | 1.53 [0.61 – 3.80] | 142 | 6 (4.23) | 0.34 [0.11 – 0.99] |
| Collective households | 1 | 1 (100) | NA | 10 | 1 (10.00) | 1.46 [0.07 – 9.80] |
| Other | 5 | 0 (0.00) | NA | 25 | 0 (0.00) | NA |
| Health-related covariates | | | | | | |
| Cardiovascular comorbidities | | | | | | |
| No (ref.) | 301 | 65 (21.59) | 1 | 260 | 15 (5.77) | 1 |
| Yes | 0 | 0 (0.00) | NA | 573 | 43 (7.50) | 1.47 [0.74 – 3.05] |
| Pulmonary comorbidities | | | | | | |
| No (ref.) | 260 | 52 (20.00) | 1 | 833 | 58 (6.96) | 1 |
| Yes | 41 | 13 (31.71) | 1.39 [0.59 – 3.13] | 0 | 0 (0.00) | NA |
| Renal comorbidities | | | | | | |
| No (ref.) | 294 | 64 (21.55) | 1 | 806 | 57 (7.07) | 1 |
| Yes | 4 | 1 (25.00) | 1.10 [0.05 – 10.81] | 27 | 1 (3.70) | 0.41 [0.02 – 2.27] |

**Table 24. Continued**

| | Critical COVID-19 patients without preexisting CV diseases at baseline (N=1,839) | | | Critical COVID-19 patients without preexisting pulmonary diseases at baseline (N=5,024) | | |
|---|---|---|---|---|---|---|
| | **N** | **N (%) with post-acute CV complications** | **Post-acute CV complication aOR [95%CI]** | **N** | **N (%) with post-acute pulmonary complications** | **Post-acute pulmonary complications aOR [95%CI]** |
| Liver comorbidities | | | | | | |
| No (ref.) | 4 | 65 (21.59) | 1 | 833 | 58 (6.96) | 1 |
| Yes | 301 | 0 (0.00) | NA | 0 | 0 (0.00) | NA |
| Other comorbidities | | | | | | |
| No (ref.) | 245 | 47 (19.18) | 1 | 542 | 36 (6.64) | 1 |
| Yes | 56 | 18 (32.14) | 1.86 [0.87 – 3.92] | 291 | 22 (7.56) | 1.22 [0.66 – 2.22] |

*All ORs are fully adjusted for sociodemographic, socioeconomic, and health-related variables. Critical COVID-19 patients are defined as any COVID-19 patients admitted in ICU and/or having developed an ARDS. Significant results are highlighted in bold.*

### *3.3.1. Sensitivity analyses*

**Supplementary Tables 26** and **27** present aORs and 95% CIs assessing the association between SE variables and the development of post-acute CV and pulmonary complications among severe and critical COVID-19 patients and non-COVID-19 hospitalized controls, along with p-values for interaction. We did not identify any SE variables with a stronger association with the development of CV and pulmonary complications in either severe COVID-19 hospitalized patients compared to controls, or critical COVID-19 patients compared to controls.

## 4. Discussion

To our knowledge, this study is the first in Belgium to investigate the development of one-year post-acute organ complications following COVID-19 hospitalization and related SE inequalities using a unique population data linkage within the HELICON project.

Our findings showed a 1.19 times higher odds and 20.16 excess burden per 1,000 persons of post-acute CV complications among severe COVID-19 patients compared to non-COVID-19 hospitalized controls both without CV diseases at baseline, further increasing when considering critical COVID-19 patients. These results are in line with those from international studies [346,347,362,363]. In the United States (U.S.), COVID-19 hospitalization was associated with significant burden differences per 1,000 persons at 12 months ranging from 5.84 for

dysrhythmia to 80.86 for inflammatory heart diseases. These burden differences were reported to increase with ICU admission. A subgroup analysis of non-hospitalized infected cases without preexisting CV conditions, compared to non-infected cases, yielded similar findings [346]. In England, Ayoubkhani et al. [363] observed a 1.5 times higher rate of major adverse CV events (MACE) – a composite of non-fatal stroke, non-fatal myocardial infarction, and CV death – among COVID-19 patients compared to matched population controls. A similar although slightly smaller effect size was observed when considering COVID-19 patients without preexisting CV conditions at baseline.

Post-acute pulmonary complications, including pulmonary fibrosis and reduced lung function, have been recognized in several international studies as significant PASC [348,363–366]. This is in line with our results showing a 2.05 and 2.72 higher odds and a 32.36 and 38.7 excess burdens per 1,000 persons of post-acute pulmonary complications among severe and critical COVID-19 patients compared to controls both without pulmonary diseases at baseline, respectively. In England, compared to a matched control group, COVID-19-related post-acute respiratory complications were reported at rates six times higher in COVID-19 patients with preexisting pulmonary conditions, and 27 times higher in those without preexisting pulmonary conditions [363]. Our results, suggesting higher odds and excess burdens of post-acute pulmonary complications according to disease severity, align

with some studies showing that long-term pulmonary sequelae increase with the severity of COVID-19 [348,364].

We did not identify any significant odds or excess burden of post-acute renal complications following either severe or critical COVID-19 hospitalization, compared to non-COVID-19 hospitalization, which does not align with other international studies [366,367]. Similarly, we did not identify any significant association between either severe or critical COVID-19 hospitalization and the development of post-acute liver complications. Evidence on renal and especially liver function impairment during the post-acute COVID-19 phase remains limited. COVID-19 associated acute kidney injury (AKI), common acute complication in COVID-19 hospitalized patients, has been linked to post-AKI kidney function impairment and elevated serum creatinine [368–370]. A 6-month follow-up cohort study reported a 1.36 and 25.24 increased risk and excess burden of chronic kidney disease (CKD) following COVID-19 hospitalization, respectively, further increasing with ICU admission [367]. Regarding post-acute liver complications, findings generally indicate no significant liver function impairment post-COVID-19 hospitalization [371], although critical COVID-19 cases may exhibit elevated liver enzymes and markers of inflammation [349].

Research on SE inequalities in PASC remains limited, especially among hospitalized COVID-19 survivors, with inconsistent findings and the use of varying definitions of PASC, making it difficult to identify clear socioeconomic patterns therein [372]. Several population-based studies from the UK showed that the less deprived individuals, compared to the least deprived, present higher odds of COVID-19-related persistent symptoms at least four weeks following COVID-19 diagnosis [373–376]. In the U.S., Hirschtick et al. [377] identified a 1.5 higher prevalence of 30-days COVID-19-related symptoms following SARS-CoV-2 infection among individuals with lower income levels. However, several population-based studies, of smaller sample sizes, found no significant association between PASC and SE factors such as income or education levels [378–380].

Although a SE gradient has been observed in COVID-19 hospitalization, established as one of the strongest predictor of PASC [381], long-term follow-up studies on SE inequalities in post-acute complications following COVID-19 hospitalization are scarce. Huang et al. [348] did not identify any significant association between education level and diffusion impairment or percentage change of CT score among COVID-19 patients discharged from the hospital. Unexpectedly, a study conducted in China showed higher odds of lung diffusion impairment at 2-years among COVID-19 hospitalized survivors with higher levels of education, compared to those with lower levels of education [365]. While we did not identify any association between

education and post-acute CV complications, our results indicate a 1.48 higher odds of post-acute pulmonary complications among severe COVID-19 patients with low income, compared to those with high income. However, among critical COVID-19 patients, post-acute pulmonary complications show no SE patterns.

Similar unclear social patterns are observed in international studies on racial and ethnic disparities in PASC, with some research reporting a lower risk of PASC following infection among Black or Asians individuals compared to white individuals [373,374], and some showing no significant association between ethnic minority status and PASC [378,382]. In our study, compared to Belgian native, individuals with a migration background presented lower odds of developing post-acute CV complications following severe COVID-19 hospitalization. As the study outcomes were based on drug prescriptions, these results may be explained by the lower healthcare access and utilization among migrants in Europe, which is often influenced by factors such as language and communication barriers, lack of insurance, lower SES, limited GP's consultation, and high out-of-pocket healthcare costs [383,384].

Despite scarce and mixed findings regarding SE patterns of post-acute health complications, especially among COVID-19 hospitalized patients, factors such as being a female, older age, smoking, and preexisting comorbidities (i.e., high BMI, anxiety/depression, chronic obstructive pulmonary disease, diabetes,

immunosuppression, and ischemic heart disease) have been associated with a higher likelihood of developing PASC in a meta-analysis of 41 studies conducted by Tsampasian et al. [381]. Preexisting comorbidities were identified in our results as the strongest risk factor for post-acute CV complications following severe COVID-19 hospitalization. Older age was identified in our study as increasing the odds of developing post-acute CV complications among severe COVID-19 patients. We did not identify any gender disparities in the development of post-acute organ complications, similarly to other studies [72,377,379]. Smoking rates are generally higher in SE disadvantaged groups, and smoking-related social health inequalities are widening in many countries [385]. Although we could not adjust for smoking as a confounder in our model, it may have contributed to our findings that low-income severe patients faced higher risk of post-acute pulmonary complications compared to high-income severe patients.

Although clinical factors, such as disease severity and preexisting health conditions, seem to play a dominant role in long-term outcomes following COVID-19 hospitalization, SE inequalities in those outcomes remain present. SE disparities in PASC can be attributed to several factors [386]: 1) SE inequalities in preexisting health conditions, shown to be a strong risk factor for PASC [381] ; 2) SE inequalities in COVID-19 infection and disease severity leading to and increasing PASC, as well

as SE inequalities in vaccination, reducing prevalence of PASC [238,244,381,387]; and 3) a lower access to and quality of healthcare among deprived SE groups [388].

Our study has several limitations. First, to isolate the effect of COVID-19 on post-acute organ complications, we limited our study population to hospitalized patients with and without COVID-19. While this approach allowed us to control for the effect of hospitalization itself on the development of post-acute organ complications, it may introduce selection bias in the interpretation of both COVID-19 outcomes and social inequalities, potentially leading to paradoxical findings [389]. For example, individuals with higher socioeconomic status (SES) may only be hospitalized when severely ill, whereas those with lower SES might be hospitalized even with moderate illness. As a result, hospitalized patients with higher SES could appear sicker on average, creating an observed 'survival advantage' for patients with lower SES [389]. However, we controlled for baseline health characteristics and applied overlap weighting to achieve covariate balance, strengthening confidence in our findings. Nonetheless, because our analysis included only hospitalized patients, our conclusions regarding social inequalities in post-acute organ complications are limited to this subgroup and may not reflect broader disparities among all COVID-19 infections. Second, we lacked data on key mediators of SE disparities in COVID-19 critical outcomes, such as the type of occupation, the adoption of preventive measures (e.g., mask wearing, vaccination) and healthcare access, who would have

better capture underlying pathways of SE inequalities in post-acute organ complications. Future research on the Belgian population should consider the roles of these factors in the causal pathway between SE position and PASC to better capture its SE patterns. Third, the CHS database covers about 50% of hospitalized COVID-19 cases in Belgium, likely leading to an underestimation of our results, as some non-COVID-19 hospitalized controls could have been COVID-19 cases not registered in the CHS database. Third, the IMA pseudopathology groups, based on medication consumption, did not allow for identifying specific post-acute complications. As a result, we had to exclude individuals with preexisting conditions at baseline for our analyses. Moreover, while the IMA pseudopathology groups provide an acceptable proxy for baseline chronic diseases, these indicators have been shown to have a high rate of false negatives, resulting in low sensitivity. This underestimation may stem from two main factors: the stringent case definition (≥ 90 DDD per year), and the fact that some medications lack specificity to a single pseudopathology. To improve accuracy in estimating the prevalence of pseudopathologies, IMA data should ideally be supplemented with clinical or diagnostic data [390]. Additionally, IMA lacks data on several well-documented PASC commonly observed in hospitalized patients, such as fatigue, brain fog, and depression [72]. As a result, the analysis is restricted to broader organ-related complications, preventing a comprehensive assessment of post-acute outcomes

following severe COVID-19. Fourth, because each pseudopathology group provided by IMA included only a very small number of severe and especially critical COVID-19 patients, we were not able to investigate the SE patterns of post-acute renal and liver complications, and we had to merge a set of comorbidities into one single variable to allow estimates to converge properly. Future studies with larger sample sizes are needed to explore the SE patterns of post-acute renal and liver complications, as well as to identify the impact of comorbidities on PASC following COVID-19 hospitalization in more detail. Finally, our study period was primarily influenced by the Alpha variant, which was less severe than Omicron and less transmissible than Delta, potentially leading to an underestimation of our findings. Future research should consider the impact of variants, vaccination (which began in Belgium on 28 December 2020, near the end of our study period), and reinfection on PASC.

## 5. Conclusion

Our study identified significant risks of one-year post-acute cardiovascular and pulmonary complications following COVID-19 hospitalization, increasing with disease severity, among patients without preexisting health conditions in these organ systems at baseline. Having underlying health conditions at baseline was identified as the strongest risk factor for the development of post-acute cardiovascular complications among COVID-19 hospitalized patients. While no

socioeconomic patterns were observed for post-acute cardiovascular complications, post-acute pulmonary complications were significantly more frequent among severe COVID-19 patients with lower income levels. These findings highlight the need for long-term follow-up care for patients discharged after COVID-19 hospitalization, particularly those who experienced critical illness and with underlying health conditions, while also taking into consideration socioeconomic groups.

## 6. Declarations

### 6.1. Ethics approval and consent to participate

The protocol of the HELICON project was approved by the Medical Ethics Committee from the University Hospital of Ghent and obtained authorization from the Information Security Committee (ISC) Social Security and Health (reference number: CSI/CSSS/22/034).

As confirmed by sections 23 and 24 of the *Guidelines 03/2020 on the processing of data concerning health for the purpose of scientific research in the context of the COVID-19 outbreak* of the European Data Protection Board (V1.0 of 21 April 2020), this survey falls under Article 6 §1(e) and Article 9 §2(i) of the General Data Protection Regulation (GDPR). In appliance with these GDPR legal grounds of data processing, no informed consent had to be signed by the patients. Clinical trial number: not applicable

### 6.2. Consent for publication

Not applicable

### 6.3. Availability of data and materials

The data used in this study are pseudonymized individual data obtained from several national registers hosted by Healthdata.be, an independent institution within Sciensano aiming to bring together all the data currently stored in multiple health registers on a single web-based platform. Due to the General Data Protection Regulation (GDPR) legislations in Belgium, these data are not publicly available. Data request must be addressed to the Information Security Committee (ISC).

### 6.4. Competing interests

The authors have no competing interests to declare that are relevant to the content of this article.


### 6.5. Funding

The Belgian Science Policy Office (BELSPO) funded this research within the BRAIN-be 2.0 framework supporting pillar 3 Federal societal challenges (grant number B2/202/P3/HELICON). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

### 6.6. Authors' contributions

LC reviewed the literature; LC, NS, LVdB, DDS, KV, SG, RDP, and BD conceived the study; LC, NS, LVdB, RDP and BD selected the study population; LC, NS, LVdB, RDP, and BD designed the statistical methodology; LC conducted the statistical analyses; LC, NS, and BD interpret the findings; LC wrote the first draft of the paper. All authors revised the text. All authors approved the final version of this manuscript and accepted responsibility for its submission for publication.

## 7. Supplementary materials

### 7.1. Table of content

#### *7.1.1. Supplementary Figures*

- **Supplementary Figure 9.** Density distribution of propensity scores in severe (A) and critical (B) COVID-19 hospitalized patients and non-COVID-19 hospitalized patients, both without cardiovascular diseases at baseline.
- **Supplementary Figure 10.** Density distribution of propensity scores in severe (A) and critical (B) COVID-19 hospitalized patients and non-COVID-19 hospitalized patients, both without pulmonary diseases at baseline.
- **Supplementary Figure 11.** Density distribution of propensity scores in severe (A) and critical (B) COVID-19 hospitalized patients and non-COVID-19 hospitalized patients, both without renal diseases at baseline.

- **Supplementary Figure 12.** Density distribution of propensity scores in severe (A) and critical (B) COVID-19 hospitalized patients and non-COVID-19 hospitalized patients, both without liver diseases at baseline.

*7.1.2. Supplementary Tables*

- **Supplementary Table 19.** Description of the substances, Anatomical Therapeutic Chemical (ATC) subcodes and ATC codes used to build the pseudopathologies available in the Intermutualistic Agency (IMA) database allowing to define post-acute cardiovascular, pulmonary, renal, and liver complications.
- **Supplementary Table 20.** Data holders and time frame for all sociodemographic, socioeconomic, and health-related variables considered as covariates in the statistical analyses.
- **Supplementary Table 21.** Characterization of severe and critical COVID-19 patients and non-COVID-19 hospitalized controls, all without cardiovascular diseases at baseline, along with weighted standardized mean differences (SMD).
- **Supplementary Table 22.** Characterization of severe and critical COVID-19 patients and non-COVID-19 hospitalized controls, all without pulmonary diseases at baseline, along with weighted standardized mean differences (SMD).

- **Supplementary Table 23.** Characterization of severe and critical COVID-19 patients and non-COVID-19 hospitalized controls, all without renal diseases at baseline, along with weighted standardized mean differences (SMD).
- **Supplementary Table 24.** Characterization of severe and critical COVID-19 patients and non-COVID-19 hospitalized controls, all without liver diseases at baseline, along with weighted standardized mean differences (SMD).
- **Supplementary Table 25.** Characterization of critical COVID-19 patients with missing income relative to the total study population composed of critical COVID-19 patients without pulmonary diseases at baseline.
- **Supplementary Table 26.** Fully adjusted odds ratio (aORs) and 95% confidence intervals (CIs) estimating the sociodemographic, socioeconomic, and health-related variables associated with the development of post-acute cardiovascular (CV) and pulmonary complications among severe COVID-19 hospitalized patients and non-COVID-19 hospitalized controls, along with p-values for interaction, respectively.
- **Supplementary Table 27.** Fully adjusted odds ratio (aORs) and 95% confidence intervals (CIs) estimating the sociodemographic, socioeconomic, and health-related variables associated with the development of post-acute cardiovascular (CV) and pulmonary complications among critical COVID-19

hospitalized patients and non-COVID-19 hospitalized controls, along with p-values for interaction, respectively.

**Supplementary Table 19. Description of the substances, Anatomical Therapeutic Chemical (ATC) subcodes, and ATC codes used to build the pseudopathologies available in the Intermutualistic Agency (IMA) database allowing to define post-acute cardiovascular, pulmonary, renal and liver complications.**

| Study outcomes | Pseudopathologies | ATC codes | System | ATC subcodes | Substance |
|---|---|---|---|---|---|
| Post-acute cardiovascular diseases | Cardiovascular diseases (general) | C | Cardiovascular system | C01 | Cardiac therapy |
| | | | | C02 | Antihypertensives |
| | | | | C03 | Diuretics |
| | | | | C04 | Peripheral vasodilatators |
| | | | | C05 | Vasoprotectives |
| | | | | C06 | Beta blocking agents |
| | | | | C07 | Calcium channel blockers |
| | | | | C08 | Agents acting on the renin-angiotensin system |
| | | | | C09 | Lipid modifying agents |
| | Thrombosis | B | Blood and blood forming organs | B01A | Anthithrombotic agents |
| | Cardiovascular diseases | C | Cardiovascular system | C01 | Cardiac therapy |

**Supplementary Table 19. Continued**

| Study outcomes | Pseudopathologies | ATC codes | System | ATC subcodes | Substance |
|---|---|---|---|---|---|
| Post-acute pulmonary complications | Chronic obstructive pulmonary disease | R | Respiratory system | R03BB | Anticholinergics |
| | | | | R03DA04 | Theophylline |
| | | | | R03A (if age > 50) | Adrenergics, inhalants |
| | | | | R03BA (if age > 50) | Glucocorticoids |
| Post-acute renal complications | Renal failure | A | Alimentary tract and metabolism | A11CC03 | Alfracalcidol |
| | | | | A11CC04 | Calcitriol |
| | | | | A11CC06 | Calcifediol |
| | | | | A12AA12 | Calcium |
| | | V | Various | V03AE02 | Sevelamer |
| | | | | V03AE03 | Lanthanum carbonate |
| | | | | V03AE04 | Calcium acetate and magnesium |
| | | | | V03AE01 | Polystyrene sulfonate |

**Supplementary Table 19. Continued**

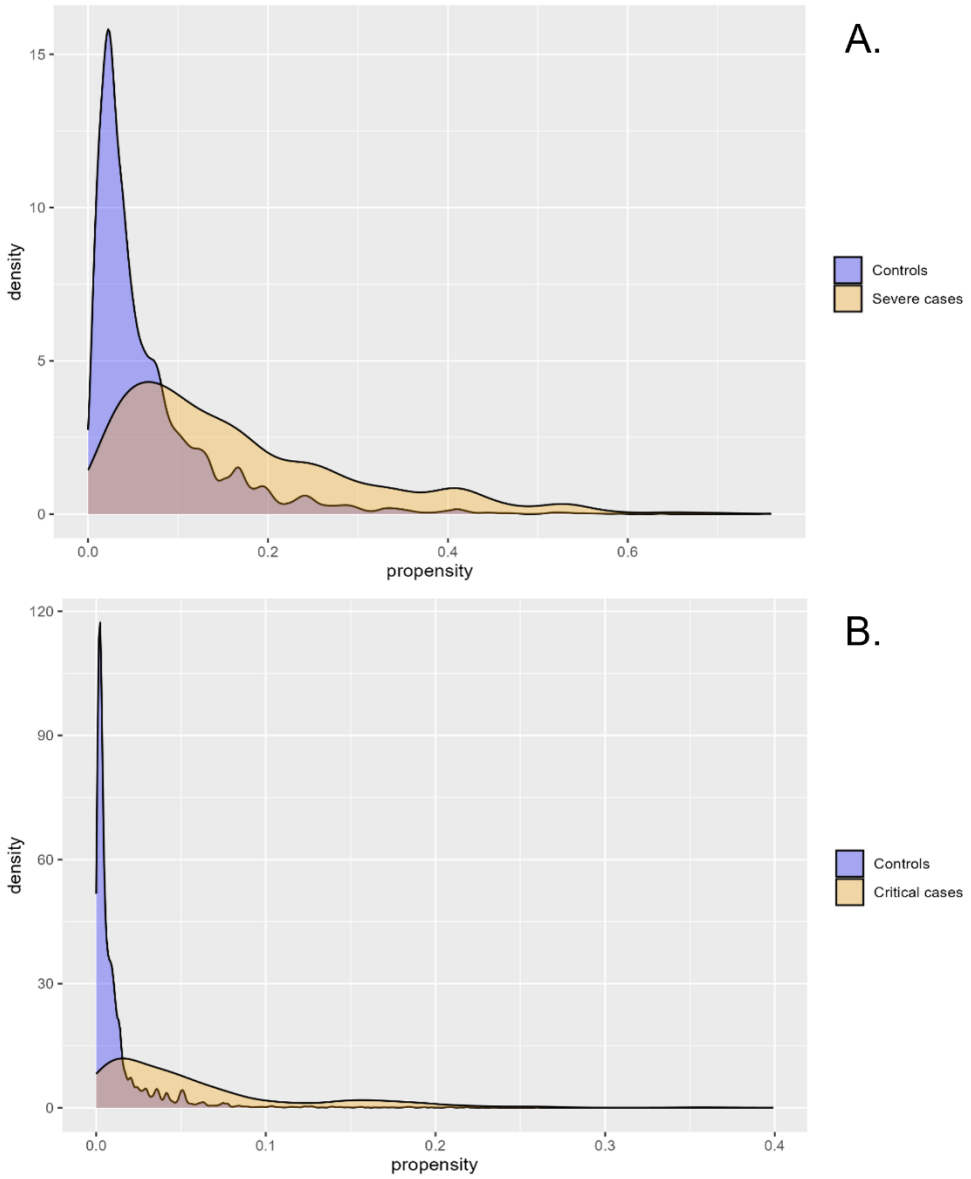
| Study outcomes | Pseudopathologies | ATC codes | System | ATC subcodes | Substance |
|---|---|---|---|---|---|
| | | H | Systemic hormonal preparation (exlusion of sexual hormones and insulins) | H05BX01 | Cinacalcet |
| Post-acute liver complications | Chronic hepatitis B and C | L | Antineoplastic and immunomodulating agents | L03AB04 | Inteferon alfa-2a |
| | | | | L03AB05 | Interferon alfa-2b |
| | | | | L03AB09 | Interferon alfacon-1 |
| | | | | L03AB10 | Peginterferon alfa-2b |
| | | | | L03AB11 | Albinterferon alfa-2b |
| | | J | Anti-infective for systemic use | J05AF08 | Adefovir dipivoxil |
| | | | | J05AF10 | Entecavir |
| | | | | J05AE11 | Protease inhibitors |
| | | | | J05AE12 | Protease inhibitors |
| | | | | J05AF08 | Adefovir dipivoxil |
| | | | | J05AF10 | Entecavir |
| | | | | J05AE11 | Protease inhibitors |

**Supplementary Table 19. Continued**

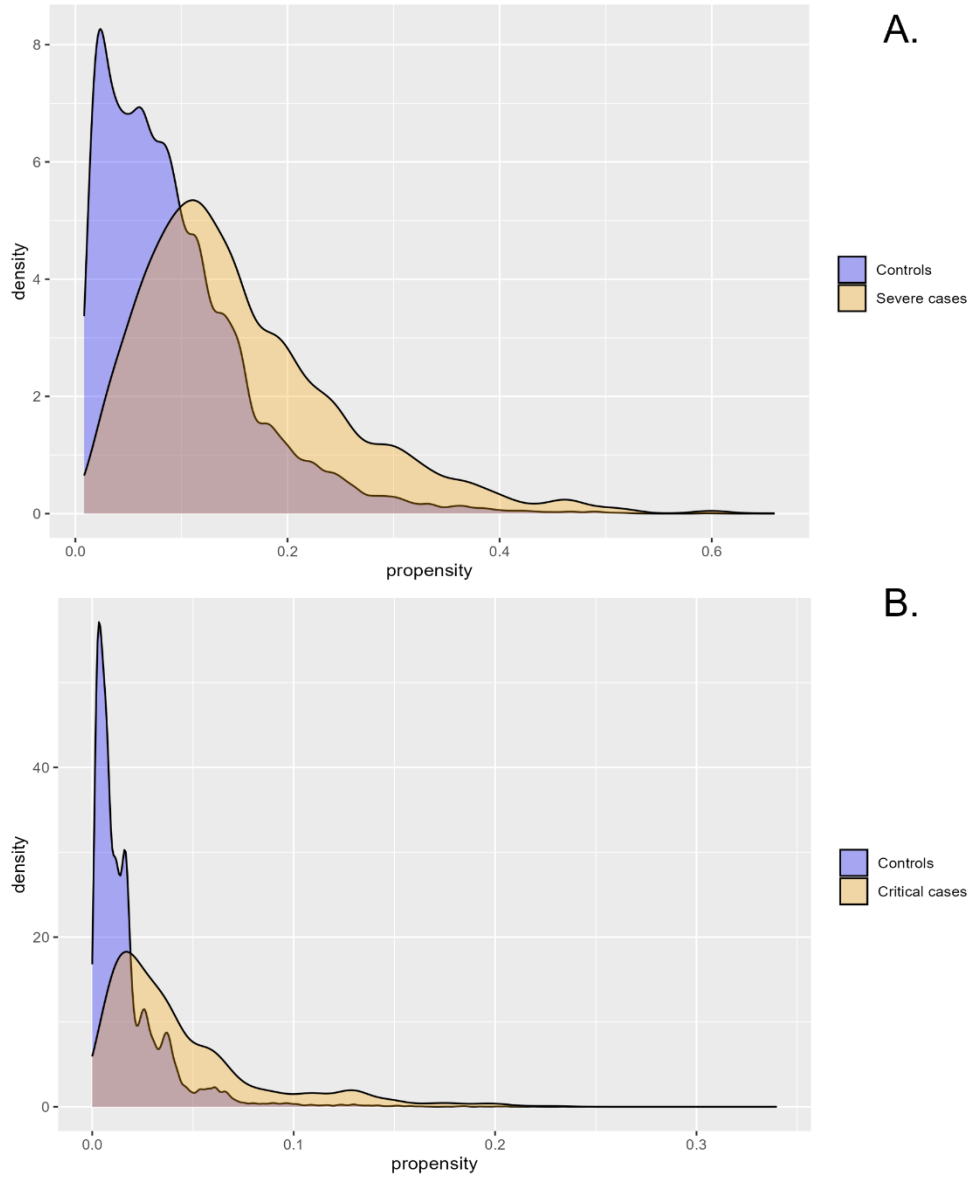
| Study outcomes | Pseudopathologies | ATC codes | System | ATC subcodes | Substance |
|---|---|---|---|---|---|
| | | | | J05AE12 | Protease inhibitors |
| | | | | J05AE14 | Protease inhibitors |
| | | | | J05AX15 | Other antivirals |
| | | | | J05AX65 | Other antivirals |

**Supplementary Table 20. Data holders and time frame for all sociodemographic, socioeconomic, and health-related variables considered as covariates in the statistical analyses**

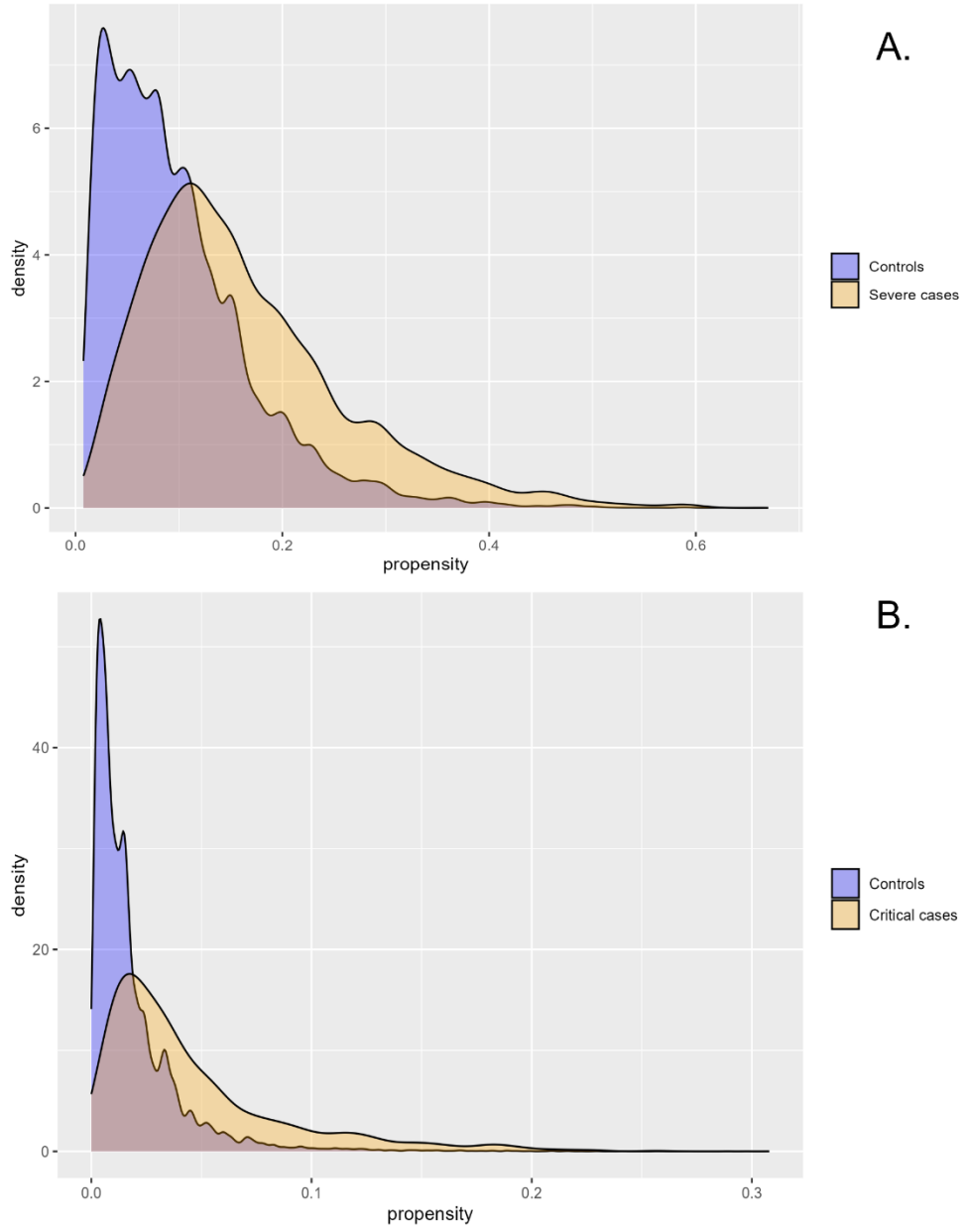
| Variables | Data holder | Time frame |
|---|---|---|
| **Sociodemographic variables** | | |
| Age groups | National register | 1st of January 2020 |
| Sex | National register | 1st of January 2020 |
| Region of residence | National register | 1st of January 2020 |
| Migration background | National register | 1st of January 2020 |
| Household type | National register | 1st of January 2020 |
| **Socioeconomic variables** | | |
| Income | Statistic Belgium | 2018 (fiscal year 2019) |
| Education | Statistics Belgium | Census 2017 |
| Employment status | Statistics Belgium | Year 2019 |
| **Health-related variables** | | |
| Cardiovascular comorbidities | Intermutualistic Agency | Year 2020 |
| Pulmonary comorbidities | Intermutualistic Agency | Year 2020 |
| Renal comorbidities | Intermutualistic Agency | Year 2020 |
| Liver comorbidities | Intermutualistic Agency | Year 2020 |
| Other comorbidities | Intermutualistic Agency | Year 2020 |



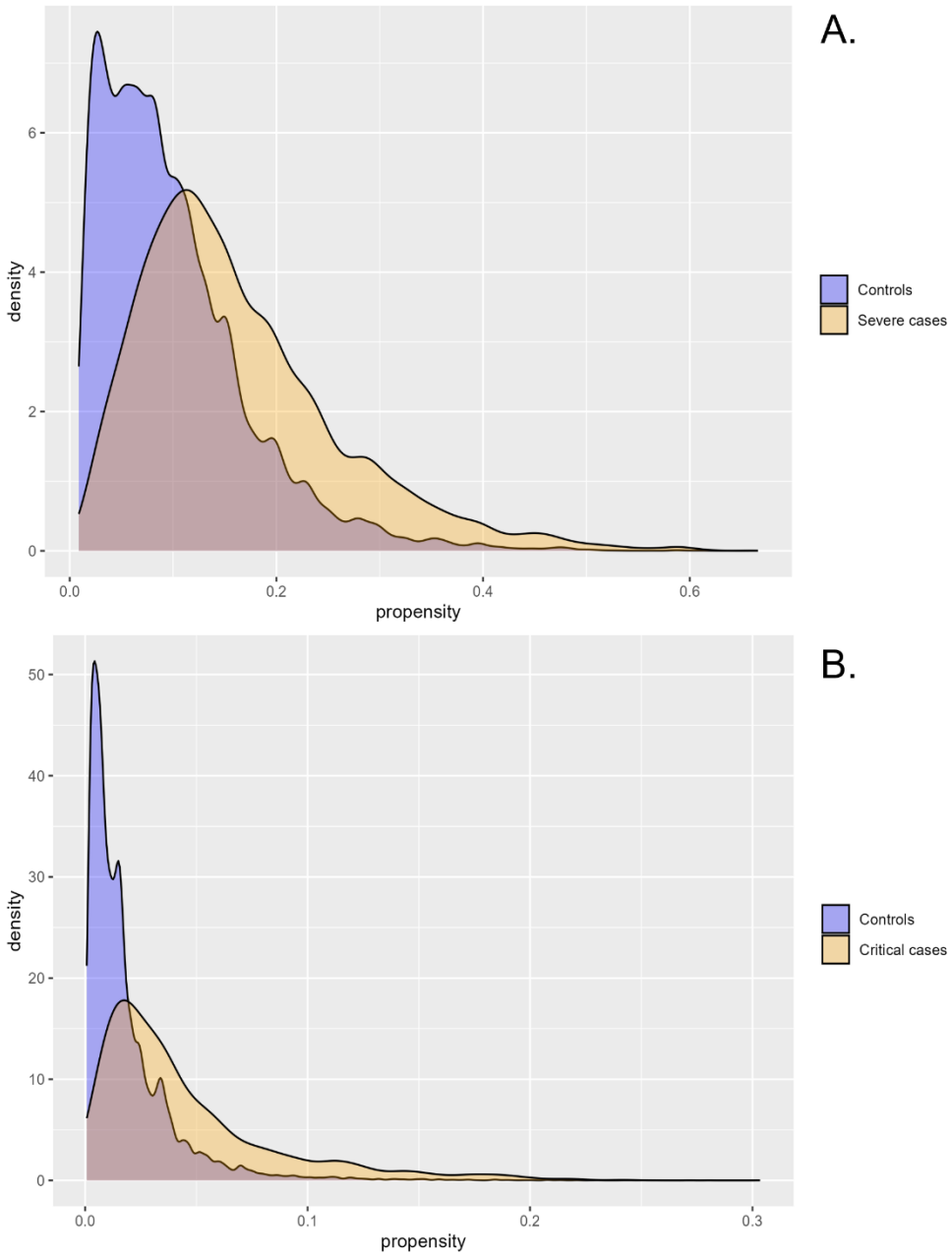

***Supplementary Figure 9. Density distribution of propensity scores in severe (A) and critical (B) COVID-19 hospitalized patients and non-COVID-19 hospitalized patients, both without cardiovascular diseases at baseline.***




***Supplementary Figure 10. Density distribution of propensity scores in severe (A) and critical (B) COVID-19 hospitalized patients and non-COVID-19 hospitalized patients, both without pulmonary diseases at baseline.***




***Supplementary Figure 11. Density distribution of propensity scores in severe (A) and critical (B) COVID-19 hospitalized patients and non-COVID-19 hospitalized patients, both without renal diseases at baseline.***




***Supplementary Figure 12. Density distribution of propensity scores in severe (A) and critical (B) COVID-19 hospitalized patients and non-COVID-19 hospitalized patients, both without liver diseases at baseline.***

**Supplementary Table 21. Characterization of severe and critical COVID-19 patients and non-COVID-19 hospitalized controls, all without cardiovascular (CV) diseases at baseline, along with weighted standardized mean differences (SMD).**

| | Non-COVID-19 hospitalized controls without CV diseases at baseline N=22,310 | Severe COVID-19 patients without CV diseases at baseline N=1,839 | | Critical COVID-19 patients without CV diseases at baseline N=301 | |
|---|---|---|---|---|---|
| | n (%) | n (%) | Weighted SMD | n (%) | Weighted SMD |
| **Demographic characteristics** | | | | | |
| Age group (in years) | | | <0.001 | | < 0.001 |
| 18-44 | 13,777 (61.8) | 538 (29.3) | | 75 (24.9) | |
| 45-64 | 6,366 (28.5) | 895 (48.7) | | 175 (58.1) | |
| 65-84 | 1,999 (9.0) | 358 (19.5) | | 47 (15.6) | |
| 85+ | 168 (0.8) | 48 (2.6) | | 4 (1.3) | |
| Sex | | | <0.001 | | <0.001 |
| Males | 7,691 (34.5) | 1,042 (56.7) | | 217 (72.1) | |
| Females | 14,619 (65.5) | 797 (43.3) | | 84 (27.9) | |
| Region of residence | | | <0.001 | | <0.001 |
| Brussels | 2,261 (10.1) | 202 (11.0) | | 43 (14.3) | |
| Flanders | 13,174 (59.0) | 905 (49.2) | | 144 (47.8) | |
| Wallonia | 6,875 (30.8) | 732 (39.8) | | 114 (37.9) | |

**Supplementary Table 21. Continued**

| | Non-COVID-19 hospitalized controls without CV diseases at baseline N=22,310 | Severe COVID-19 patients without CV diseases at baseline N=1,839 | | Critical COVID-19 patients without CV diseases at baseline N=301 | |
|---|---|---|---|---|---|
| | n (%) | n (%) | Weighted SMD | n (%) | Weighted SMD |
| Migration status | | | <0.001 | | <0.001 |
| Belgian natives | 14,910 (66.8) | 1,002 (54.5) | | 154 (51.2) | |
| Second-generation migrants | 3,395 (15.2) | 255 (13.9) | | 42 (14.0) | |
| First-generation European migrants | 1,476 (6.6) | 174 (9.5) | | 38 (12.6) | |
| First-generation non-European migrants | 2,529 (11.3) | 408 (22.2) | | 67 (22.3) | |
| Household type | | | <0.001 | | <0.001 |
| Couples with children | 9,842 (44.1) | 824 (44.8) | | 164 (54.5) | |
| Couples without children | 5,380 (24.1) | 428 (23.3) | | 69 (22.9) | |
| One person household | 2,714 (12.2) | 180 (9.8) | | 16 (5.3) | |
| Single parents | 3,708 (16.6) | 324 (17.6) | | 46 (15.3) | |
| Collective households | 450 (2.0) | 41 (2.2) | | 1 (0.3) | |
| Other | 216 (1.0) | 42 (2.3) | | 5 (1.7) | |

**Supplementary Table 21. Continued**

| | Non-COVID-19 hospitalized controls without CV diseases at baseline N=22,310 | Severe COVID-19 patients without CV diseases at baseline N=1,839 | | Critical COVID-19 patients without CV diseases at baseline N=301 | |
|---|---|---|---|---|---|
| | n (%) | n (%) | Weighted SMD | n (%) | Weighted SMD |
| **Socioeconomic characteristics** | | | | | |
| Education | | | <0.001 | | <0.001 |
| High | 6,252 (28.0) | 414 (22.5) | | 69 (22.9) | |
| Middle | 6,992 (31.3) | 512 (27.8) | | 88 (29.2) | |
| Low | 6,134 (27.5) | 593 (32.2) | | 85 (28.2) | |
| Missing | 2,932 (13.1) | 320 (17.4) | | 59 (19.6) | |
| Income | | | <0.001 | | <0.001 |
| High | 8,395 (37.6) | 639 (34.7) | | 118 (39.2) | |
| Middle | 6,566 (29.4) | 529 (28.8) | | 91 (30.2) | |
| Low | 6,431 (28.8) | 574 (31.2) | | 77 (25.6) | |
| Missing | 918 (4.1) | 97 (5.3) | | 15 (5.0) | |
| **Clinical characteristics** | | | | | |
| Cardiovascular comorbidities | | | <0.001 | | <0.001 |
| No | 22,310 (100) | 1,839 (100) | | 301 (100) | |
| Yes | 0 (0.0) | 0 (0.0) | | 0 (0.0) | |

**Supplementary Table 21. Continued**

| | Non-COVID-19 hospitalized controls without CV diseases at baseline N=22,310 | Severe COVID-19 patients without CV diseases at baseline N=1,839 | | Critical COVID-19 patients without CV diseases at baseline N=301 | |
|---|---|---|---|---|---|
| | n (%) | n (%) | Weighted SMD | n (%) | Weighted SMD |
| Pulmonary comorbidities | | | <0.001 | | <0.001 |
| No | 20,624 (92.4) | 1,625 (88.4) | | 260 (86.4) | |
| Yes | 1,686 (7.6) | 214 (11.6) | | 41 (13.6) | |
| Renal comorbidities | | | <0.001 | | <0.001 |
| No | 22,259 (99.8) | 1,833 (99.7) | | 297 (98.7) | |
| Yes | 51 (0.2) | 6 (0.3) | | 4 (1.3) | |
| Liver comorbidities | | | <0.001 | | <0.001 |
| No | 22,288 (99.9) | 1,839 (100) | | 301 (100) | |
| Yes | 22 (0.1) | 0 (0.0) | | 0 (0.0) | |
| Other comorbidities | | | <0.001 | | <0.001 |
| No | 12,846 (57.6) | 1,446 (78.6) | | 245 (81.4) | |
| Yes | 9,464 (42.4) | 393 (21.4) | | 56 (18.6) | |

**Supplementary Table 22. Characterization of severe and critical COVID-19 patients and non-COVID-19 hospitalized controls, all without pulmonary diseases at baseline, along with weighted standardized mean differences (SMD).**

| | Non-COVID-19 hospitalized controls without pulmonary diseases at baseline N=45,036 | Severe COVID-19 patients without pulmonary diseases at baseline N=5,024 | | Critical COVID-19 patients without pulmonary diseases at baseline N=833 | |
|---|---|---|---|---|---|
| | n (%) | n (%) | Weighted SMD | n (%) | Weighted SMD |
| **Demographic characteristics** | | | | | |
| Age group (in years) | | | <0.001 | | <0.001 |
| 18-44 | 14,955 (33.2) | 600 (11.9) | | 85 (10.2) | |
| 45-64 | 13,645 (30.3) | 1,766 (35.2) | | 388 (46.6) | |
| 65-84 | 13,825 (30.7) | 2,128 (42.4) | | 336 (40.3) | |
| 85+ | 2,611 (5.8) | 530 (10.5) | | 24 (2.9) | |
| Sex | | | <0.001 | | <0.001 |
| Males | 19,212 (42.7) | 2,696 (53.7) | | 568 (68.2) | |
| Females | 25,824 (57.3) | 2,328 (46.3) | | 265 (31.8) | |
| Region of residence | | | <0.001 | | <0.001 |
| Brussels | 3,670 (8.1) | 428 (8.5) | | 83 (10.0) | |
| Flanders | 27,275 (60.6) | 2,421 (48.2) | | 399 (47.9) | |
| Wallonia | 14,091 (31.3) | 2,175 (43.3) | | 351 (42.1) | |

**Supplementary Table 22. Continued**

| | Non-COVID-19 hospitalized controls without pulmonary diseases at baseline N=45,036 | Severe COVID-19 patients without pulmonary diseases at baseline N=5,024 | | Critical COVID-19 patients without pulmonary diseases at baseline N=833 | |
|---|---|---|---|---|---|
| | n (%) | n (%) | Weighted SMD | n (%) | Weighted SMD |
| Migration status | | | <0.001 | | <0.001 |
| Belgian natives | 33,581 (74.6) | 3,295 (65.6) | | 492 (59.1) | |
| Second-generation migrants | 4,832 (10.7) | 465 (9.3) | | 91 (10.9) | |
| First-generation European migrants | 2,855 (6.3) | 434 (8.6) | | 97 (11.6) | |
| First-generation non-European migrants | 3,768 (8.4) | 830 (16.5) | | 153 (18.4) | |
| Household type | | | <0.001 | | <0.001 |
| Couples with children | 13,690 (30.4) | 1,442 (28.7) | | 283 (34.0) | |
| Couples without children | 15,344 (34.1) | 1,711 (34.1) | | 328 (39.4) | |
| One person household | 3,977 (8.8) | 362 (7.2) | | 45 (5.4) | |
| Single parents | 10,362 (23.0) | 1,188 (23.6) | | 142 (17.0) | |
| Collective households | 852 (1.9) | 108 (2.1) | | 10 (1.2) | |
| Other | 811 (1.8) | 213 (4.2) | | 25 (3.0) | |

**Supplementary Table 22. Continued**

| | Non-COVID-19 hospitalized controls without pulmonary diseases at baseline N=45,036 | Severe COVID-19 patients without pulmonary diseases at baseline N=5,024 | | Critical COVID-19 patients without pulmonary diseases at baseline N=833 | |
|---|---|---|---|---|---|
| | n (%) | n (%) | Weighted SMD | n (%) | Weighted SMD |
| **Socioeconomic characteristics** | | | | | |
| Education | | | <0.001 | | <0.001 |
| High | 10,549 (23.4) | 870 (17.3) | | 144 (17.3) | |
| Middle | 12,549 (27.9) | 1,110 (22.1) | | 214 (25.7) | |
| Low | 17,105 (38.0) | 2,331 (46.4) | | 353 (42.4) | |
| Missing | 4,833 (10.7) | 713 (14.2) | | 122 (14.6) | |
| Income | | | <0.001 | | <0.001 |
| High | 13,470 (29.9) | 1,292 (25.7) | | 249 (29.9) | |
| Middle | 13,963 (31.0) | 1,512 (30.1) | | 275 (33.0) | |
| Low | 15,779 (35.0) | 1,903 (37.9) | | 264 (31.7) | |
| Missing | 1,824 (4.1) | 317 (6.3) | | 45 (5.4) | |
| **Clinical characteristics** | | | | | |
| Cardiovascular comorbidities | | | <0.001 | | <0.001 |
| No | 20,624 (45.8) | 1,625 (32.3) | | 260 (31.2) | |
| Yes | 24,412 (54.2) | 3,399 (67.7) | | 573 (68.8) | |

**Supplementary Table 22. Continued**

| | Non-COVID-19 hospitalized controls without pulmonary diseases at baseline N=45,036 | Severe COVID-19 patients without pulmonary diseases at baseline N=5,024 | | Critical COVID-19 patients without pulmonary diseases at baseline N=833 | |
|---|---|---|---|---|---|
| | n (%) | n (%) | Weighted SMD | n (%) | Weighted SMD |
| Pulmonary comorbidities | | | <0.001 | | <0.001 |
| No | 45,036 (100) | 5,024 (100) | | 833 (100) | |
| Yes | 0 (0.0) | 0 (0.0) | | 0 (0.0) | |
| Renal comorbidities | | | <0.001 | | <0.001 |
| No | 44,125 (98.0) | 4,880 (97.1) | | 806 (96.8) | |
| Yes | 911 (2.0) | 144 (2.9) | | 27 (3.2) | |
| Liver comorbidities | | | <0.001 | | <0.001 |
| No | 44,940 (99.8) | 5,015 (99.8) | | 833 (100) | |
| Yes | 96 (0.2) | 9 (0.2) | | 0 (0.0) | |
| Other comorbidities | | | <0.001 | | <0.001 |
| No | 22,832 (50.7) | 3,121 (62.1) | | 542 (65.1) | |
| Yes | 22,204 (49.3) | 1,903 (37.9) | | 291 (34.9) | |

**Supplementary Table 23. Characterization of severe and critical COVID-19 patients and non-COVID-19 hospitalized controls, all without renal diseases at baseline, along with weighted standardized mean differences (SMD).**

| | Non-COVID-19 hospitalized controls without renal diseases at baseline N=50,910 | Severe COVID-19 patients without renal diseases at baseline N=6,040 | | Critical COVID-19 patients without renal diseases at baseline N=998 | |
|---|---|---|---|---|---|
| | n (%) | n (%) | Weighted SMD | n (%) | Weighted SMD |
| **Demographic characteristics** | | | | | |
| Age group (in years) | | | <0.001 | | <0.001 |
| 18-44 | 15,898 (31.2) | 682 (11.3) | | 97 (9.7) | |
| 45-64 | 15,674 (30.8) | 2,070 (34.3) | | 463 (46.4) | |
| 65-84 | 16,442 (32.3) | 2,657 (44.0) | | 408 (40.9) | |
| 85+ | 2,896 (5.7) | 631 (10.4) | | 30 (3.0) | |
| Sex | | | <0.001 | | <0.001 |
| Males | 22,033 (43.3) | 3,220 (53.3) | | 675 (67.6) | |
| Females | 28,877 (56.7) | 2,820 (46.7) | | 323 (32.4) | |
| Region of residence | | | <0.001 | | <0.001 |
| Brussels | 4,050 (8.0) | 488 (8.1) | | 105 (10.5) | |
| Flanders | 30,576 (60.1) | 2,923 (48.4) | | 465 (46.6) | |
| Wallonia | 16,284 (32.0) | 2,629 (43.5) | | 428 (42.9) | |

**Supplementary Table 23. Continued**

| | Non-COVID-19 hospitalized controls without renal diseases at baseline N=50,910 | Severe COVID-19 patients without renal diseases at baseline N=6,040 | | Critical COVID-19 patients without renal diseases at baseline N=998 | |
|---|---|---|---|---|---|
| | n (%) | n (%) | Weighted SMD | n (%) | Weighted SMD |
| Migration status | | | <0.001 | | <0.001 |
| Belgian natives | 38,268 (75.2) | 4,030 (66.7) | | 595 (59.6) | |
| Second-generation migrants | 3,219 (6.3) | 542 (9.0) | | 105 (10.5) | |
| First-generation European migrants | 4,071 (8.0) | 529 (8.8) | | 112 (11.2) | |
| First-generation non-European migrants | 5,352 (25.6) | 939 (15.5) | | 186 (18.6) | |
| Household type | | | <0.001 | | <0.001 |
| Couples with children | 14,892 (29.3) | 1,693 (28.0) | | 338 (33.9) | |
| Couples without children | 17,685 (34.7) | 2,108 (34.9) | | 386 (38.7) | |
| One person household | 4,411 (8.7) | 421 (7.0) | | 57 (5.7) | |
| Single parents | 12,019 (23.6) | 1,430 (23.7) | | 179 (17.9) | |
| Collective households | 957 (1.9) | 127 (2.1) | | 12 (1.2) | |
| Other | 946 (1.9) | 261 (4.3) | | 26 (2.6) | |

**Supplementary Table 23. Continued**

| | Non-COVID-19 hospitalized controls without renal diseases at baseline N=50,910 | Severe COVID-19 patients without renal diseases at baseline N=6,040 | | Critical COVID-19 patients without renal diseases at baseline N=998 | |
|---|---|---|---|---|---|
| | n (%) | n (%) | Weighted SMD | n (%) | Weighted SMD |
| **Socioeconomic characteristics** | | | | | |
| Education | | | <0.001 | | <0.001 |
| High | 11,357 (22.3) | 988 (16.4) | | 167 (16.7) | |
| Middle | 14,022 (27.5) | 1,331 (22.0) | | 255 (25.6) | |
| Low | 20,107 (39.5) | 2,899 (48.0) | | 431 (43.2) | |
| Missing | 5,424 (10.7) | 822 (13.6) | | 145 (14.5) | |
| Income | | | <0.001 | | <0.001 |
| High | 14,732 (28.9) | 1,518 (25.1) | | 297 (29.8) | |
| Middle | 15,851 (31.1) | 1,822 (30.2) | | 327 (32.8) | |
| Low | 18,206 (35.8) | 2,323 (38.5) | | 324 (32.5) | |
| Missing | 2,121 (4.2) | 377 (6.2) | | 50 (5.0) | |
| **Clinical characteristics** | | | | | |
| Cardiovascular comorbidities | | | <0.001 | | <0.001 |
| No | 22,259 (43.7) | 1,833 (30.3) | | 297 (29.8) | |
| Yes | 28,651 (56.3) | 4,207 (69.7) | | 701 (70.2) | |

**Supplementary Table 23. Continued**

| | Non-COVID-19 hospitalized controls without renal diseases at baseline N=50,910 | Severe COVID-19 patients without renal diseases at baseline N=6,040 | | Critical COVID-19 patients without renal diseases at baseline N=998 | |
|---|---|---|---|---|---|
| | n (%) | n (%) | Weighted SMD | n (%) | Weighted SMD |
| Pulmonary comorbidities | | | <0.001 | | <0.001 |
| No | 44,125 (86.7) | 4,880 (80.8) | | 806 (80.8) | |
| Yes | 6,785 (13.3) | 1,160 (19.2) | | 192 (19.2) | |
| Renal comorbidities | | | <0.001 | | <0.001 |
| No | 50,910 (100) | 6,040 (100) | | | |
| Yes | 0 (0.0) | 0 (0.0) | | | |
| Liver comorbidities | | | <0.001 | | <0.001 |
| No | 50,791 (99.8) | 6,032 (99.9) | | 998 (100) | |
| Yes | 119 (0.2) | 8 (0.1) | | 0 (0.0) | |
| Other comorbidities | | | <0.001 | | <0.001 |
| No | 24,848 (48.8) | 3,703 (61.3) | | 998 (100) | |
| Yes | 26,062 (51.2) | 2,337 (38.7) | | 0 (0.0) | |

**Supplementary Table 24. Characterization of severe and critical COVID-19 patients and non-COVID-19 hospitalized controls, all without liver diseases at baseline, along with weighted standardized mean differences (SMD).**

| | Non-COVID-19 hospitalized controls without liver diseases at baseline N=51,965 | Severe COVID-19 patients without liver diseases at baseline N=6,220 | | Critical COVID-19 patients without liver diseases at baseline N=1,031 | |
|---|---|---|---|---|---|
| | n (%) | n (%) | Weighted SMD | n (%) | Weighted SMD |
| **Demographic characteristics** | | | | | |
| Age group (in years) | | | <0.001 | | <0.001 |
| 18-44 | 15,965 (30.7) | 684 (11.0) | | 100 (9.7) | |
| 45-64 | 15,901 (30.6) | 2,100 (33.8) | | 471 (45.7) | |
| 65-84 | 17,093 (32.9) | 2,785 (44.8) | | 429 (41.6) | |
| 85+ | 3,006 (5.8) | 651 (10.5) | | 31 (3.0) | |
| Sex | | | <0.001 | | <0.001 |
| Males | 22,610 (43.5) | 3,324 (53.4) | | 695 (67.4) | |
| Females | 29,355 (56.5) | 2,896 (46.6) | | 336 (32.6) | |
| Region of residence | | | <0.001 | | <0.001 |
| Brussels | 4,118 (7.9) | 497 (8.0) | | 107 (10.4) | |
| Flanders | 31,341 (60.3) | 3,034 (48.8) | | 485 (47.0) | |
| Wallonia | 16,506 (31.8) | 2,689 (43.2) | | 439 (42.6) | |

**Supplementary Table 24. Continued**

| | Non-COVID-19 hospitalized controls without liver diseases at baseline N=51,965 | Severe COVID-19 patients without liver diseases at baseline N=6,220 | | Critical COVID-19 patients without liver diseases at baseline N=1,031 | |
|---|---|---|---|---|---|
| | n (%) | n (%) | Weighted SMD | n (%) | Weighted SMD |
| Migration status | | | <0.001 | | <0.001 |
| Belgian natives | 39,132 (75.3) | 4,162 (66.9) | | 620 (60.1) | |
| Second-generation migrants | 5,393 (10.4) | 551 (8.9) | | 108 (10.5) | |
| First-generation European migrants | 3,280 (6.3) | 539 (8.7) | | 114 (11.1) | |
| First-generation non-European migrants | 4,160 (8.0) | 968 (15.6) | | 189 (18.3) | |
| Household type | | | <0.001 | | <0.001 |
| Couples with children | 15,055 (29.0) | 1,723 (27.7) | | 342 (33.2) | |
| Couples without children | 18,138 (34.9) | 2,181 (35.1) | | 402 (39.0) | |
| One person household | 4,486 (8.6) | 428 (6.9) | | 60 (5.8) | |
| Single parents | 12,351 (23.8) | 1,490 (24.0) | | 185 (17.9) | |
| Collective households | 981 (1.9) | 127 (2.0) | | 14 (1.4) | |
| Other | 954 (1.8) | 271 (4.4) | | 28 (2.7) | |

**Supplementary Table 24. Continued**

| | Non-COVID-19 hospitalized controls without liver diseases at baseline N=51,965 | Severe COVID-19 patients without liver diseases at baseline N=6,220 | | Critical COVID-19 patients without liver diseases at baseline N=1,031 | |
|---|---|---|---|---|---|
| | n (%) | n (%) | Weighted SMD | n (%) | Weighted SMD |
| **Socioeconomic characteristics** | | | | | |
| Education | | | <0.001 | | <0.001 |
| High | 11,488 (22.1) | 1,008 (16.2) | | 169 (16.4) | |
| Middle | 14,237 (27.4) | 1,350 (21.7) | | 261 (25.3) | |
| Low | 20,708 (39.8) | 3,019 (48.5) | | 452 (43.8) | |
| Missing | 5,532 (10.6) | 843 (13.6) | | 149 (14.5) | |
| Income | | | <0.001 | | <0.001 |
| High | 14,914 (28.7) | 1,540 (24.8) | | 300 (29.1) | |
| Middle | 16,168 (31.1) | 1,878 (30.2) | | 336 (32.6) | |
| Low | 18,748 (36.1) | 2,408 (38.7) | | 342 (33.2) | |
| Missing | 2,135 (4.1) | 394 (6.3) | | 53 (5.1) | |
| **Clinical characteristics** | | | | | |
| Cardiovascular comorbidities | | | <0.001 | | <0.001 |
| No | 22,288 (42.9) | 1,839 (29.6) | | 301 (29.2) | |
| Yes | 29,677 (57.1) | 4,381 (70.4) | | 730 (70.8) | |

**Supplementary Table 24. Continued**

| | Non-COVID-19 hospitalized controls without renal diseases at baseline N=50,910 | Severe COVID-19 patients without renal diseases at baseline N=6,040 | | Critical COVID-19 patients without renal diseases at baseline N=998 | |
|---|---|---|---|---|---|
| | n (%) | n (%) | Weighted SMD | n (%) | Weighted SMD |
| Pulmonary comorbidities | | | <0.001 | | <0.001 |
| No | 44,940 (86.5) | 5,015 (80.6) | | 833 (80.8) | |
| Yes | 7,025 (13.5) | 1,205 (19.4) | | 198 (19.2) | |
| Renal comorbidities | | | <0.001 | | <0.001 |
| No | 50,791 (97.7) | 6,032 (97.0) | | 998 (96.8) | |
| Yes | 1,174 (2.3) | 188 (3.0) | | 33 (3.2) | |
| Liver comorbidities | | | <0.001 | | <0.001 |
| No | 51,965 (100) | 6,220 (100) | | 1,031 (100) | |
| Yes | 0 (0.0) | 0 (0.0) | | 0 (0.0) | |
| Other comorbidities | | | <0.001 | | <0.001 |
| No | 25,159 (48.4) | 3,758 (60.4) | | 654 (63.4) | |
| Yes | 26,806 (51.6) | 2,462 (39.6) | | 377 (36.6) | |

**Supplementary Table 25. Characterization of critical COVID-19 patients with missing income relative to the total study population composed of critical COVID-19 patients without pulmonary diseases at baseline.**

| Variables | Total (N=833) | Missing income (n=45) | At least one missing information to the total population |
|---|---|---|---|
| | n | n | %row |
| Age group (in years) | | | |
| 18-44 | 85 | 3 | 3.53 |
| 45-64 | 388 | 20 | 5.15 |
| 65-84 | 336 | 16 | 4.76 |
| 85+ | 24 | 6 | 25.00 |
| Sex | | | |
| Males | 568 | 23 | 4.05 |
| Females | 265 | 22 | 8.30 |
| Region of residence | | | |
| Brussels | 83 | 7 | 8.43 |
| Flanders | 399 | 13 | 5.76 |
| Wallonia | 351 | 15 | 4.27 |
| Migration status | | | |
| Belgian natives | 492 | 21 | 4.27 |
| Second-generation migrants | 91 | 4 | 4.40 |
| First-generation European migrants | 97 | 8 | 7.84 |
| First-generation non-European migrants | 153 | 12 | 8.25 |
| Household type | | | |
| Couples with children | 283 | 2 | 0.71 |
| Couples without children | 328 | 7 | 2.13 |
| One person household | 45 | 2 | 4.44 |
| Single parents | 142 | 9 | 6.34 |
| Collective households | 10 | 0 | 0.00 |
| Other | 25 | 25 | 100 |

**Supplementary Table 25. Continued**

| Variables | Total (N=833) | Missing income (n=45) | At least one missing information to the total population |
|---|---|---|---|
| | n | *n* | *%row* |
| **Socioeconomic characteristics** | | | |
| Education | | | |
| High | 144 | 4 | 2.78 |
| Middle | 214 | 8 | 3.74 |
| Low | 353 | 11 | 3.12 |
| Missing | 122 | 22 | 18.03 |
| **Health-related characteristics** | | | |
| Cardiovascular comorbidities | | | |
| No | 260 | 14 | 5.38 |
| Yes | 573 | 31 | 5.41 |
| Pulmonary comorbidities | | | |
| No | 833 | 788 | 94.59 |
| Yes | 0 | 45 | 0.00 |
| Renal comorbidities | | | |
| No | 806 | 43 | 5.33 |
| Yes | 27 | 2 | 7.41 |
| Liver comorbidities | | | |
| No | 833 | 45 | 5.40 |
| Yes | 0 | 0 | 0.00 |
| Other comorbidities | | | |
| No | 542 | 22 | 4.06 |
| Yes | 291 | 23 | 7.90 |

**Supplementary Table 26. Fully adjusted odds ratio (aOR) and 95% confidence intervals (CIs) estimating the sociodemographic, socioeconomic, and health-related variables associated with the development of post-acute cardiovascular (CV) and pulmonary complications among severe COVID-19 hospitalized patients and non-COVID-19 hospitalized controls, along with p-values for interaction, respectively.**

| | Post-acute CV complications | | | Post-acute pulmonary complications | | |
|---|---|---|---|---|---|---|
| | **Severe COVID-19 patients without CV diseases at baseline** | **Controls without CV diseases at baseline** | | **Severe COVID-19 patients without pulmonary diseases at baseline** | **Controls without pulmonary diseases at baseline** | |
| | *aOR [95%CI]* | *aOR [95%CI]* | *P-value for interaction* | *aOR [95%CI]* | *aOR [95%CI]* | *P-value for interaction* |
| **Socioeconomic exposures** | | | | | | |
| *Education* | | | | | | |
| High (ref.) | 1 | 1 | 1 | 1 | 1 | 1 |
| Middle | 1.21 [0.81 – 1.83] | **1.27 [1.11 – 1.46]** | 0.833 | 1.20 [0.82 – 1.79] | 0.98 [0.82 – 1.17] | 0.350 |
| Low | 1.46 [0.98 – 2.20] | **1.28 [1.11 – 1.47]** | 0.536 | 1.20 [0.84 – 1.76] | 1.16 [0.99 – 1.38] | 0.865 |
| Missing | 1.20 [0.70 – 2.04] | 1.08 [0.76 – 1.50] | 0.743 | 0.75 [0.18 – 2.16] | **1.53 [1.22 – 1.91]** | 0.277 |
| *Income* | | | | | | |
| High (ref.) | 1 | 1 | 1 | 1 | 1 | 1 |
| Middle | 0.99 [0.68 – 1.44] | 1.00 [0.88 – 1.14] | 0.975 | 1.18 [0.84 – 1.68] | 1.17 [0.98 – 1.38] | 0.950 |
| Low | 1.21 [0.80 – 1.84] | **1.24 [1.06 – 1.45]** | 0.927 | **1.53 [1.05 – 2.25]** | **1.46 [1.22 – 1.76]** | 0.833 |
| Missing | 1.94 [0.82 – 4.25] | 1.08 [0.76 – 1.50] | 0.192 | 0.75 [0.18 – 2.16] | **2.46 [1.75 – 3.41]** | 0.060 |

**Supplementary Table 26. Continued**

| | Post-acute CV complications | | | Post-acute pulmonary complications | | |
|---|---|---|---|---|---|---|
| | **Severe COVID-19 patients without CV diseases at baseline** | **Controls without CV diseases at baseline** | | **Severe COVID-19 patients without pulmonary diseases at baseline** | **Controls without pulmonary diseases at baseline** | |
| | *aOR [95%CI]* | *aOR [95%CI]* | *P-value for interaction* | *aOR [95%CI]* | *aOR [95%CI]* | *P-value for interaction* |
| **Sociodemographic covariates** | | | | | | |
| Age group (in years) | | | | | | |
| 18 – 44 (ref.) | 1 | 1 | 1 | 1 | 1 | 1 |
| 45 – 64 | **1.94 [1.34 – 2.86]** | **3.66 [3.25 – 4.12]** | **<0.05** | 0.92 [0.61 – 1.41] | **1.36 [1.14 – 1.62]** | 0.086 |
| 65 – 84 | **2.97 [1.90 – 4.69]** | **6.05 [5.18 – 7.07]** | **< 0.05** | 0.92 [0.58 – 1.48] | **1.28 [1.04 – 1.57]** | 0.202 |
| 85+ | 2.36 [0.99 – 5.27] | **14.71 [10.48 – 20.56]** | **<0.001** | 0.93 [0.53 – 1.65] | **1.44 [1.10 – 1.88]** | 0.176 |
| Gender | | | | | | |
| Males (ref.) | 1 | 1 | 1 | 1 | 1 | 1 |
| Females | 0.78 [0.59 – 1.03] | **0.84 [0.76 – 0.93]** | 0.618 | 0.80 [0.63 – 1.01] | 0.99 [0.88 – 1.11] | 0.115 |
| Region of residence | | | | | | |
| Flanders (ref.) | 1 | 1 | 1 | 1 | 1 | 1 |
| Brussels | **0.56 [0.32 – 0.96]** | 0.96 [0.80 – 1.16] | 0.071 | 0.65 [0.38 – 1.07] | **1.24 [1.00 – 1.53]** | **<0.05** |
| Wallonia | 1.12 [0.84 – 1.49] | **1.23 [1.11 – 1.37]** | 0.542 | 0.86 [0.68 – 1.10] | **1.34 [1.19 – 1.51]** | **<0.05** |

**Supplementary Table 26. Continued**

| | Post-acute CV complications | | | Post-acute pulmonary complications | | |
|---|---|---|---|---|---|---|
| | **Severe COVID-19 patients without CV diseases at baseline** | **Controls without CV diseases at baseline** | | **Severe COVID-19 patients without pulmonary diseases at baseline** | **Controls without pulmonary diseases at baseline** | |
| | *aOR [95%CI]* | *aOR [95%CI]* | *P-value for interaction* | *aOR [95%CI]* | *aOR [95%CI]* | *P-value for interaction* |
| Migration status | | | | | | |
| Belgian natives (ref.) | 1 | 1 | 1 | 1 | 1 | 1 |
| Second-generation migrants | **0.61 [0.38 – 0.97]** | 1.05 [0.90 – 1.23] | **< 0.05** | 1.18 [0.78 – 1.74] | 1.21 [1.01 – 1.44] | 0.913 |
| First-generation European migrants | **0.54 [0.30 – 0.93]** | 1.11 [0.89 – 1.38] | **<0.05** | 0.91 [0.58 – 1.39] | 0.86 [0.68 – 1.09] | 0.830 |
| First-generation non-European migrants | **0.61 [0.38 – 0.97]** | 1.04 [0.85 – 1.26] | 0.141 | 0.67 [0.44 – 1.00] | **0.66 [0.51 – 0.85]** | 0.943 |

**Supplementary Table 26. Continued**

| | Post-acute CV complications | | | Post-acute pulmonary complications | | |
|---|---|---|---|---|---|---|
| | **Severe COVID-19 patients without CV diseases at baseline** | **Controls without CV diseases at baseline** | | **Severe COVID-19 patients without pulmonary diseases at baseline** | **Controls without pulmonary diseases at baseline** | |
| | *aOR [95%CI]* | *aOR [95%CI]* | *P-value for interaction* | *aOR [95%CI]* | *aOR [95%CI]* | *P-value for interaction* |
| Household type | | | | | | |
| Couples with children (ref.) | 1 | 1 | 1 | 1 | 1 | 1 |
| Couples without children | 1.05 [0.72 – 1.52] | **1.19 [1.04 – 1.36]** | 0.540 | 0.87 [0.61 – 1.23] | 1.08 [0.91 – 1.28] | 0.273 |
| One person household | 0.66 [0.37 – 1.13] | 1.08 [0.90 – 1.29] | 0.100 | 0.79 [0.45 – 1.31] | **1.34 [1.07 – 1.66]** | 0.071 |
| Single parents | 0.87 [0.55 – 1.36] | 1.02 [0.86 – 1.21] | 0.506 | 0.77 [0.51 – 1.17] | 1.08 [0.89 – 1.32] | 0.830 |
| Collective households | 0.56 [0.16 – 1.49] | 1.30 [0.91 – 1.82] | 0.147 | 1.30 [0.58 – 2.56] | 1.22 [0.80 – 1.80] | 0.892 |
| Other | 0.59 [0.19 – 1.79] | 1.14 [0.68 – 1.90] | 0.285 | 2.04 [0.61 – 9.38] | 0.76 [0.47 – 1.21] | 0.169 |

**Supplementary Table 26. Continued**

| | Post-acute CV complications | | | Post-acute pulmonary complications | | |
|---|---|---|---|---|---|---|
| | **Severe COVID-19 patients without CV diseases at baseline** | **Controls without CV diseases at baseline** | | **Severe COVID-19 patients without pulmonary diseases at baseline** | **Controls without pulmonary diseases at baseline** | |
| | *aOR [95%CI]* | *aOR [95%CI]* | *P-value for interaction* | *aOR [95%CI]* | *aOR [95%CI]* | *P-value for interaction* |
| **Health-related covariates** | | | | | | |
| Cardiovascular comorbidities | | | | | | |
| No (ref.) | 1 | 1 | 1 | 1 | 1 | 1 |
| Yes | NA | NA | NA | 1.28 [0.95 – 1.73] | **1.44 [1.24 – 1.68]** | 0.475 |
| Pulmonary comorbidities | | | | | | |
| No (ref.) | 1 | 1 | 1 | 1 | 1 | 1 |
| Yes | 1.09 [0.73 – 1.59] | **1.39 [1.19 – 1.62]** | 0.252 | NA | NA | NA |
| Renal comorbidities | | | | | | |
| No (ref.) | 1 | 1 | 1 | 1 | 1 | 1 |
| Yes | 2.68 [0.36 – 14.27] | **5.91 [3.24 – 10.68]** | 0.398 | 0.36 [0.11 – 0.86] | 1.28 [0.92 – 1.73] | **<0.005** |

**Supplementary Table 26. Continued**

| | Post-acute CV complications | | | Post-acute pulmonary complications | | |
|---|---|---|---|---|---|---|
| | **Severe COVID-19 patients without CV diseases at baseline** | **Controls without CV diseases at baseline** | | **Severe COVID-19 patients without pulmonary diseases at baseline** | **Controls without pulmonary diseases at baseline** | |
| | *aOR [95%CI]* | *aOR [95%CI]* | *P-value for interaction* | *aOR [95%CI]* | *aOR [95%CI]* | *P-value for interaction* |
| Liver comorbidities | | | | | | |
| No (ref.) | 1 | 1 | 1 | 1 | 1 | 1 |
| Yes | NA | **4.61 [1.88 – 11.60]** | NA | NA | 2.15 [0.95 – 4.21] | NA |
| Other comorbidities | | | | | | |
| No (ref.) | 1 | 1 | 1 | 1 | 1 | 1 |
| Yes | **2.00 [1.48 – 2.68]** | **1.22 [1.10 – 1.34]** | **<0.001** | 1.13 [0.89 – 1.44] | **1.36 [1.22 – 1.53]** | 0.175 |

**Supplementary Table 27. Fully adjusted odds ratio (OR) and 95% confidence intervals (CI) estimating the sociodemographic, socioeconomic, and health-related variables associated with the development of post-acute cardiovascular (CV) and pulmonary complications among critical COVID-19 hospitalized patients and non-COVID-19 hospitalized controls, along with p-values for interaction, respectively.**

| | Post-acute CV complications | | | Post-acute pulmonary complications | | |
|---|---|---|---|---|---|---|
| | **Critical COVID-19 patients without CV diseases at baseline** | **Controls without CV diseases at baseline** | | **Critical COVID-19 patients without pulmonary diseases at baseline** | **Controls without pulmonary diseases at baseline** | |
| | *aOR [95%CI]* | *aOR [95%CI]* | *P-value for interaction* | *aOR [95%CI]* | *aOR [95%CI]* | *P-value for interaction* |
| **Socioeconomic exposures** | | | | | | |
| Education | | | | | | |
| High (ref.) | 1 | 1 | 1 | 1 | 1 | 1 |
| Middle | 0.96 [0.43 – 2.16] | **1.27 [1.11 – 1.46]** | 0.495 | 0.55 [0.21 – 1.42] | 0.98 [0.82 – 1.17] | 0.451 |
| Low | 0.87 [0.36 – 2.08] | **1.28 [1.11 – 1.47]** | 0.394 | 0.84 [0.37 – 2.01] | 1.16 [0.99 – 1.38] | 0.232 |
| Missing | 0.55 [0.15 – 1.85] | 1.09 [0.88 – 1.34] | 0.287 | 1.17 [0.39 – 3.48] | **1.53 [1.22 – 1.91]** | 0.640 |
| Income | | | | | | |
| High (ref.) | 1 | 1 | 1 | 1 | 1 | 1 |
| Middle | 1.32 [0.61 – 2.85] | 1.00 [0.88 – 1.14] | 0.481 | 0.87 [0.39 – 1.93] | 1.17 [0.98 – 1.38] | 0.476 |
| Low | 1.47 [0.58 – 3.69] | **1.24 [1.06 – 1.45]** | 0.744 | 1.50 [0.66 – 3.47] | **1.46 [1.22 – 1.76]** | 0.956 |
| Missing | 0.61 [0.03 – 4.44] | 1.08 [0.76 – 1.50] | 0.625 | **5.25 [1.15 – 23.24]** | **2.46 [1.75 – 3.41]** | 0.335 |

**Supplementary Table 27. Continued**

| | Post-acute CV complications | | | Post-acute pulmonary complications | | |
|---|---|---|---|---|---|---|
| | **Critical COVID-19 patients without CV diseases at baseline** | **Controls without CV diseases at baseline** | | **Critical COVID-19 patients without pulmonary diseases at baseline** | **Controls without pulmonary diseases at baseline** | |
| | *aOR [95%CI]* | *aOR [95%CI]* | *P-value for interaction* | *aOR [95%CI]* | *aOR [95%CI]* | *P-value for interaction* |
| **Sociodemographic covariates** | | | | | | |
| Age group (in years) | | | | | | |
| 18 – 44 (ref.) | 1 | 1 | 1 | 1 | 1 | 1 |
| 45 – 64 | 1.42 [0.67 – 3.20] | **3.66 [3.25 – 4.12]** | **<0.05** | 3.54 [0.99 – 22.72] | **1.36 [1.14 – 1.62]** | 0.211 |
| 65 – 84 | 1.62 [0.56 – 4.77] | **6.05 [5.18 – 7.07]** | **<0.05** | 1.82 [0.44 – 12.55] | **1.28 [1.04 – 1.57]** | 0.667 |
| 85+ | NA | **14.71 [10.48 – 20.56]** | NA | 2.45 [0.10 – 30.49] | **1.44 [1.10 – 1.88]** | 0.689 |
| Gender | | | | | | |
| Males (ref.) | 1 | 1 | 1 | 1 | 1 | 1 |
| Females | 0.81 [0.39 – 1.62] | **0.84 [0.76 – 0.93]** | 0.931 | 0.85 [0.44 – 1.57] | 0.99 [0.88 – 1.11] | 0.651 |
| Region of residence | | | | | | |
| Flanders (ref.) | 1 | 1 | 1 | 1 | 1 | 1 |
| Brussels | 1.49 [0.53 – 4.03] | 0.96 [0.80 – 1.16] | 0.398 | 1.25 [0.45 – 3.12] | **1.24 [1.00 – 1.53]** | 0.794 |
| Wallonia | 1.22 [0.64 – 2.34] | 1.23 [1.11 – 1.37] | 0.987 | 1.24 [0.67 – 2.28] | **1.34 [1.19 – 1.51]** | 0.889 |

**Supplementary Table 27. Continued**

| | Post-acute CV complications | | | Post-acute pulmonary complications | | |
|---|---|---|---|---|---|---|
| | **Critical COVID-19 patients without CV diseases at baseline** | **Controls without CV diseases at baseline** | | **Critical COVID-19 patients without pulmonary diseases at baseline** | **Controls without pulmonary diseases at baseline** | |
| | *aOR [95%CI]* | *aOR [95%CI]* | *P-value for interaction* | *aOR [95%CI]* | *aOR [95%CI]* | *P-value for interaction* |
| Migration status | | | | | | |
| Belgian natives (ref.) | 1 | 1 | 1 | 1 | 1 | 1 |
| Second-generation migrants | 0.85 [0.35 – 1.99] | 1.05 [0.90 – 1.23] | 0.623 | 1.70 [0.70 – 3.84] | 1.21 [1.01 – 1.44] | 0.441 |
| First-generation European migrants | 0.60 [0.19 – 1.67] | 1.11 [0.89 – 1.38] | 0.265 | 0.92 [0.33 – 2.31] | 0.86 [0.68 – 1.09] | 0.889 |
| First-generation non-European migrants | 0.35 [0.11 – 1.00] | 1.04 [0.85 – 1.26] | 0.054 | 0.93 [0.38 – 2.16] | **0.66 [0.51 – 0.85]** | 0.465 |

**Supplementary Table 27. Continued**

| | Post-acute CV complications | | | Post-acute pulmonary complications | | |
|---|---|---|---|---|---|---|
| | **Critical COVID-19 patients without CV diseases at baseline** | **Controls without CV diseases at baseline** | | **Critical COVID-19 patients without pulmonary diseases at baseline** | **Controls without pulmonary diseases at baseline** | |
| | *aOR [95%CI]* | *aOR [95%CI]* | *P-value for interaction* | *aOR [95%CI]* | *aOR [95%CI]* | *P-value for interaction* |
| Household type | | | | | | |
| Couples with children (ref.) | 1 | 1 | 1 | 1 | 1 | 1 |
| Couples without children | 1.47 [0.66 – 3.24] | **1.19 [1.04 – 1.36]** | 0.607 | 1.13 [0.55 – 2.35] | 1.08 [0.91 – 1.28] | 0.905 |
| One person household | 0.74 [0.14 – 3.04] | 1.08 [0.90 – 1.29] | 0.620 | 0.48 [0.07 – 1.89] | **1.34 [1.07 – 1.66]** | 0.204 |
| Single parents | 1.53 [0.61 – 3.80] | 1.02 [0.86 – 1.21] | 0.387 | 0.34 [0.11 – 0.99] | 1.08 [0.89 – 1.32] | **<0.05** |
| Collective households | NA | 1.30 [0.91 – 1.82] | 0.977 | 1.46 [0.07 – 9.80] | 1.22 [0.80 – 1.80] | 0.878 |
| Other | NA | 1.14 [0.68 – 1.90] | 0.958 | NA | 0.76 [0.47 – 1.21] | 0.907 |

**Supplementary Table 27. Continued**

| | Post-acute CV complications | | | Post-acute pulmonary complications | | |
|---|---|---|---|---|---|---|
| | **Critical COVID-19 patients without CV diseases at baseline** | **Controls without CV diseases at baseline** | | **Critical COVID-19 patients without pulmonary diseases at baseline** | **Controls without pulmonary diseases at baseline** | |
| | *aOR [95%CI]* | *aOR [95%CI]* | *P-value for interaction* | *aOR [95%CI]* | *aOR [95%CI]* | *P-value for interaction* |
| **Health-related covariates** | | | | | | |
| Cardiovascular comorbidities | | | | | | |
| No (ref.) | 1 | 1 | 1 | 1 | 1 | 1 |
| Yes | NA | NA | NA | 1.47 [0.74 – 3.05] | **1.44 [1.24 – 1.68]** | 0.958 |
| Pulmonary comorbidities | | | | | | |
| No (ref.) | 1 | 1 | 1 | 1 | 1 | 1 |
| Yes | 1.39 [0.59 – 3.13] | **1.39 [1.19 – 1.62]** | 0.990 | NA | NA | NA |
| Renal comorbidities | | | | | | |
| No (ref.) | 1 | 1 | 1 | 1 | 1 | 1 |
| Yes | 1.10 [0.05 – 10.81] | **5.91 [3.24 – 10.68]** | 0.192 | 0.41 [0.02 – 2.27] | 1.28 [0.92 – 1.73] | 0.300 |

**Supplementary Table 27. Continued**

| | Post-acute CV complications | | | Post-acute pulmonary complications | | |
|---|---|---|---|---|---|---|
| | **Critical COVID-19 patients without CV diseases at baseline** | **Controls without CV diseases at baseline** | | **Critical COVID-19 patients without pulmonary diseases at baseline** | **Controls without pulmonary diseases at baseline** | |
| | *aOR [95%CI]* | *aOR [95%CI]* | *P-value for interaction* | *aOR [95%CI]* | *aOR [95%CI]* | *P-value for interaction* |
| Liver comorbidities | | | | | | |
| No (ref.) | 1 | 1 | 1 | 1 | 1 | 1 |
| Yes | NA | **4.61 [1.88 – 11.60]** | NA | NA | 2.15 [0.95 – 4.21] | NA |
| Other comorbidities | | | | | | |
| No (ref.) | 1 | 1 | | 1 | 1 | 1 |
| Yes | 1.86 [0.87 – 3.92] | **1.22 [1.10 – 1.34]** | | 1.22 [0.66 – 2.22] | **1.36 [1.22 – 1.53]** | 0.719 |

# PART IV

# General discussion

The frequency of (re-)emerging infectious diseases (EIDs) with pandemic potential has been rising over the past decades [391]. The COVID-19 pandemic clearly illustrated how such EIDs can pose serious threats to global public health. Although COVID-19 has been extensively studied worldwide, most research has primarily focused on biomedical aspects, including viral transmission pathways, morbidity and mortality rates, and the effectiveness of non-pharmaceutical interventions (NPIs) (e.g., social distancing, mask-wearing, handwashing) and pharmaceutical interventions (PIs) (e.g., vaccines) [391,392].

However, previous (re-)EID pandemics, such as the H1N1 influenza and the Spanish flu, have provided clear evidence of social inequalities in transmission, infection, and disease progression, patterns that were similarly observed from the early stages of the COVID-19 pandemic [105,391]. The relative lack of attention to social health disparities during the COVID-19 pandemic is therefore striking, especially in light of the well-established links between SDOH, vulnerability to EIDs, and comorbidity burden (a key risk factor for severe COVID-19 outcomes). These interactions lie at the core of the syndemic nature of the COVID-19 pandemic, a concept detailed earlier (see "*Why is COVID-19 a syndemic*?", p. 25) [391,393].

The main objective of this thesis was therefore to **contribute to knowledge of social inequalities in EIDs** by unraveling the association between SDOH and various COVID-19 health outcomes during the pandemic and its aftermath in Belgium.

The first part of this general discussion summarizes the **main findings** presented across the four chapters of this dissertation, each discussed in detail in its respective discussion section (see ***Chapters 1*** to ***4***). Second, the discussion dives into the **underlying mechanisms** that give rise to social inequalities in EIDs, aiming to understand **how** and **why** such public health crises can exacerbate social inequalities already deeply rooted in our society. Third, the **main limitations** of the thesis are discussed. Finally, the **implications of the findings for policy and research** are discussed, along with **recommendations** to address social health inequalities in future EID pandemic preparedness and responses.

# Summary of main findings

- **Social health inequalities in COVID-19-related health behaviors**

***Chapter 1*** aimed to contribute to the understanding of social health inequalities in COVID-19-related health behaviors. Specifically, we examined the association between individual sociodemographic (SD) and socioeconomic (SE) factors and the uptake of a first COVID-19 vaccine dose. Using multivariable logistic regression models, we found lower vaccine uptake among younger individuals, women, second- and first-generation migrants, single parents, one-person households, individuals with lower levels of income and education, and those without healthcare qualifications.

In Belgium, the vaccination rollout was managed independently by each region. Regional differences were therefore accounted for in our multivariable models. The results showed that individuals residing in Brussels and Wallonia were less likely to be vaccinated compared to those residing in Flanders. Nevertheless, similar SD and SE disparities were observed within each Belgian region.

Although a limitation of this study was that the study population included only individuals who were tested for SARS-CoV-2, Hubin et al. [394] have demonstrated the representativeness of this sample in relation to the Belgian population.

- **Social health inequalities in COVID-19-related morbidity and mortality**

***Chapter 2*** investigated whether educational level was associated with the likelihood of COVID-19 hospitalization and which variables mediated this association, by period of variant of concern (VOC) predominance. Given the social disparities in vaccine uptake identified in ***Chapter 1***, as well as well-documented social inequalities in comorbidities and health literacy, these three variables were chosen as potential mediators. Interventional effect models were used, allowing the consideration of multiple mediators without assuming any causal structure among the mediators. We observed significantly higher COVID-19 hospitalization rates among individuals with lower education levels. The interventional direct effect (IDE) of education explained 85% and 95% of the total effect during Delta and Omicron periods, respectively. During the Delta period, lack of vaccination mediated 9% of the association, and comorbidities 6%; for Omicron, comorbidities mediated 4% and lack of vaccination 1%. Health literacy (proxied by having a healthcare degree) was not a significant mediator once interactions between mediators and the exposure were considered. However, the IDE remained significant, suggesting that unmeasured additional factors may also contribute to this association (e.g., factors on transmission, exposure, or access to and quality of healthcare).

***Chapter 3*** explored the social patterns of COVID-19 mortality and the extent to which they differ from regular mortality. Using multivariable logistic regression models and conditional inference tree (CIT) algorithms, we identified and ranked predictors of COVID-19 death and all other causes of death (OCOD). Key predictors of COVID-19 death included older age, male sex, residence in collective settings, first-generation migration background, and lower SE status (i.e., lower levels of income and education). CIT analysis further highlighted living in care homes as the key predictor of COVID-19 death. Similar patterns were observed for all OCOD; however, both first- and second-generation migrants showed lower odds of all OCOD, reflecting the well-documented *migrant mortality advantage*, typically observed in non-communicable diseases (NCDs).

- **Long-term health impacts of severe COVID-19 and related social patterns**

***Chapter 4*** investigated post-acute health complications following severe COVID-19 and related social patterns. Specifically, two research questions were addressed: 1) Are patients hospitalized for COVID-19 at greater risk of developing post-acute organ complications compared to non-COVID-19 hospitalized controls, both free of such conditions at baseline? 2) Among COVID-19 hospitalized patients, are these complications influenced by SE factors? All analyses were stratified between critical (ICU admission and/or ARDS) and severe (neither ICU admission nor ARDS) COVID-19 hospitalized patients. To answer the first question, we used binomial logistic

regression models with overlap weights based on propensity scores. Results showed that COVID-19 hospitalization significantly increased the risk of post-acute cardiovascular (CV) and pulmonary complications within one year, overall increasing with disease severity. For the second research question, multivariable logistic regression models indicated that, among patients with severe COVID-19, those with low income had higher odds of developing post-acute pulmonary complications compared to those with high income. However, no consistent SE patterns were observed for other organ complications across severity groups.

# Mechanisms of emerging infectious diseases-related health inequalities

This thesis contributes to knowledge of social health inequalities in EIDs, using the COVID-19 pandemic as a case study. Following the summary of the main findings, this section explores the **underlying mechanisms driving EID-related health inequalities**. It aims to shed light on how SDOH influence these inequalities by shaping the emergence and severity of EID health outcomes, ultimately exacerbating pre-existing social health inequalities.

## 1. Pathways to social health inequalities in emerging infectious diseases

**Figure 30** illustrates the main pathways through which health inequalities in EIDs arise in relation to SDOH.

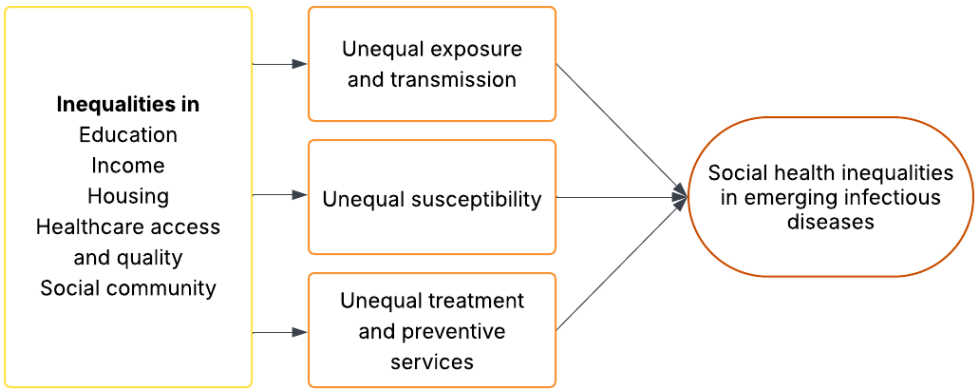



***Figure 30. Pathways through social health inequalities in emerging infectious diseases.*** *Sources: adapted from Bambra (2022) [391] and Blumenshine et al. (2008) [395].*

### 1.1. Unequal exposure and transmission

The disproportionate burden of EID-related morbidity and mortality among disadvantaged SE groups, as identified in ***Chapter 2*** and ***Chapter 3*** respectively, can be partly attributed to **unequal exposure** to the virus, driven by disparities in **working conditions**, and **unequal SARS-CoV-2 transmission**, driven by disparities in **housing conditions** [105,395,396].

In terms of housing, individuals with lower SES (socioeconomic status) and migrant populations are more likely to reside in **densely populated urban areas**, facilitating EID transmission [397–399]. In Belgium, the Brussels Capital Region has the highest population density (7,501 inhabitants/km$^2$) and the largest proportion of migrant and socially deprived populations [400–402], and recorded the highest age-standardized COVID-19 mortality rates from 2020 to 2022 compared to the Flemish and Walloon regions [29]. Even within densely populated areas, further disparities emerge as individuals with lower SES and migrants are more likely to live in **overcrowded** or **multigenerational households**, further increasing the risk of viral transmission [403–405].

Regarding working conditions, individuals with lower SES and migrants are overrepresented in **essential, in-person jobs**, such as cleaning, manufacturing, or transportation, where remote work is not an option. This increased occupational

exposure has been associated with higher rates of infection, morbidity, and mortality during the COVID-19 pandemic [406,407]. Socially deprived and migrant populations also disproportionately used **public transport** during the pandemic (in part due to their limited capacity for remote work), which was associated with higher SARS-CoV-2 infection rates [408–411].

Moreover, the **economic fallout** of the pandemic has had a particularly severe impact on socially deprived and migrant populations, many of whom lost their jobs or faced significant financial hardship [412,413]. Unemployment and reduced income forced many into shared housing arrangements and homelessness, limiting social distancing and access to hygiene facilities, further hindering preventive behaviors such as handwashing [414–416].

### 1.2. Unequal vulnerability

The **burden of NCDs**, such as diabetes, CV diseases, or obesity, is higher among socially deprived and migrant populations, due to its association with SDOH such as education, income, poor living and working conditions, or access to and quality of healthcare [105,417–419]. In Belgium, age-adjusted data from 2018 revealed that individuals with lower SES were 49% more likely to suffer from multimorbidity compared to those with higher SES, particularly from conditions such as hypertension, COPD, diabetes, and asthma [420]. In the Brussels Capital Region, Racape et al. [421] investigated the clinical and social profiles of patients

hospitalized for COVID-19 during the first wave in two hospitals located in deprived and multiethnic areas. Their findings showed a higher prevalence of obesity, diabetes, and hypertension among migrants from the Middle-East, Sub-Saharan Africa, and North Africa compared to Belgian natives, respectively.

Since these chronic conditions are well-established risk factors for COVID-19-related morbidity and mortality [422,423], their higher prevalence among socially disadvantaged and migrant populations partly explains the association between SDOH and EID-related health outcomes. This mechanism was confirmed in ***Chapter 2***, where the presence of at least one underlying health condition was found to partially mediate the association between education and COVID-19 hospitalization.

A particularly striking point regarding migrant populations is the "disappearance" of the *migrant mortality advantage* typically observed in NCDs, such as cancer or CV diseases, in the context of EIDs, as identified in ***Chapter 3*** [424]. While migrants often exhibit lower mortality rates for NCDs, morbidity within these populations remains high, due to many health issues that stem from lower SES and poor living and working conditions, thereby increasing their vulnerability to EID-related morbidity and mortality [333,425].

Beyond physical underlying health conditions, higher levels of **chronic stress** among socially deprived and migrant populations, closely related to social deprivation (e.g., poverty, discrimination), may further increase their vulnerability to EIDs

[391,395,396,426]. Chronic stress negatively affects immune system functioning, weakening protective immunity and increasing vulnerability to infection [427,428]. During the COVID-19 pandemic, this stress gap between the most and the least socially deprived populations widened, fueled by factors such as social isolation, financial insecurity, and job loss [324,429].

### 1.3. Unequal treatment and preventive services

The third mechanism conducting to EID-related health inequalities is the **unequal access to, and uptake of, timely and appropriate treatments and preventive services** [105,395].

As shown in ***Chapter 1***, socially deprived and migrant populations were less likely to be vaccinated against COVID-19. ***Chapter 2*** further demonstrated that COVID-19 vaccine uptake partially mediated the association between lower SES and increased COVID-19-related morbidity. **Unequal vaccination coverage** therefore partly contributes to the observed social gradient in COVID-19 incidence, morbidity, and mortality [278,425,430–433]. These disparities in vaccine uptake are mainly explained by higher levels of **vaccine hesitancy** in socially vulnerable populations [434,435]. According to the WHO, vaccine hesitancy is defined as "*the delay in acceptance or refusal of vaccine despite availability of vaccination services*" and is influenced primarily by three factors: confidence, complacency, and convenience (**Figure 31**) [436].

**Confidence** refers to **trust in the effectiveness and safety** of the vaccine, the vaccination providers, and the authorities responsible for vaccination decisions [220,436]. **Complacency** relates to the **perceived risk** of disease and whether vaccination is considered necessary [220,436]. **Convenience** refers to **structural** and **practical barriers** to access vaccination, including physical availability, affordability, geographical accessibility, and the ability to understand and seek vaccination services, challenges that may be heightened by language barriers, low health literacy, or limited digital skills [436,437].

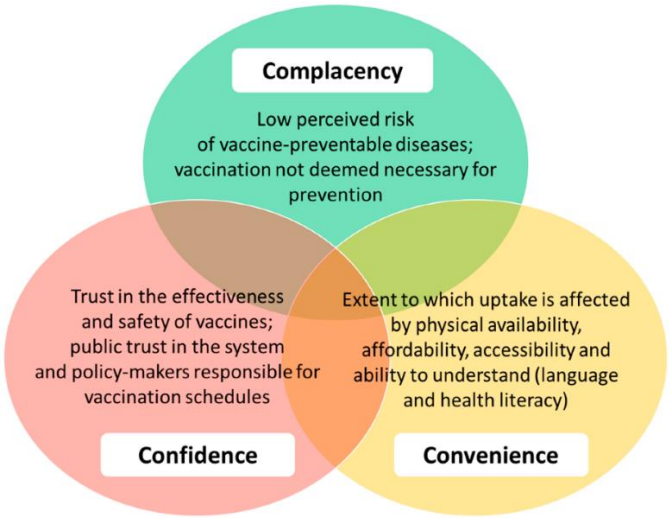



***Figure 31. Main mechanisms influencing vaccine hesitancy.*** *Source: Thompson et al. (2023) [438].*

Beyond vaccination, inequalities were also observed in access to **diagnostic testing** and **personal protective equipment** (PPE) during the COVID-19 pandemic [439–441]. These inequalities, combined with unequal vaccine uptake, interacted to exacerbate social inequalities in COVID-19-related health outcomes [431].

Another key aspect of unequal treatment is the tendency for socially deprived populations to **delay seeking medical care**. Such delays are influenced by a combination of SD and SE factors (e.g., gender, education, income), institutional barriers (e.g., financial or logistical access to care), individual factors (e.g., perceived severity), and health beliefs (e.g., disease knowledge) (**Figure 32**) [442,443]. Delayed care-seeking behaviors increase the risk of severe complications and reduces the likelihood of timely treatment following respiratory infection [396,444].

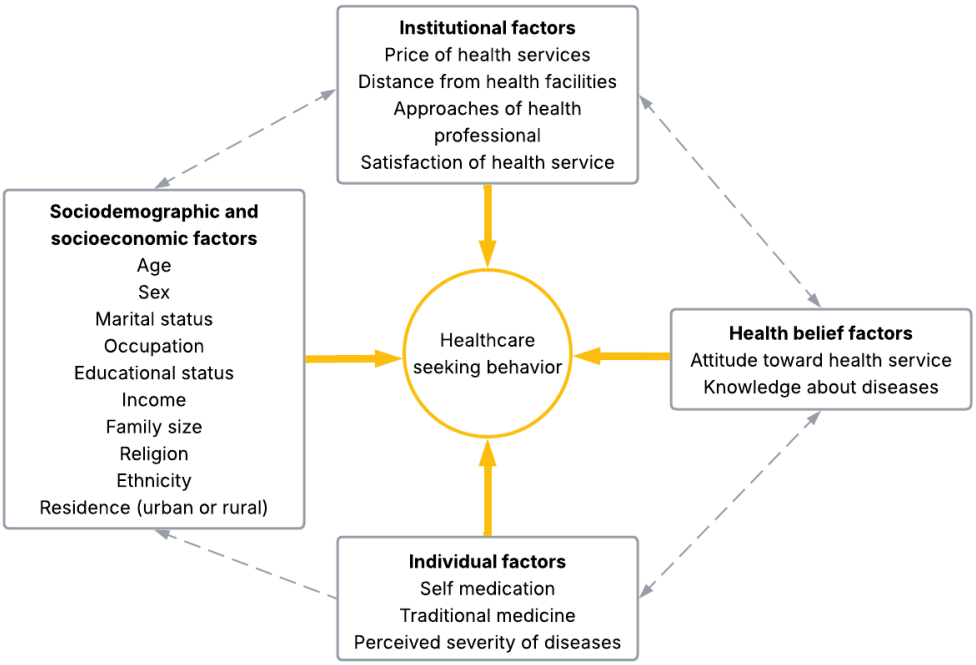



***Figure 32. Conceptual framework of healthcare seeking behavior.*** *Source: Bojola et al. (2018) [443].*

Furthermore, individuals with lower SES are **less likely to have a GP and to consult specialists or paramedical professionals** (e.g., physiotherapists), resulting in inequalities across the entire spectrum of preventive care, from primary to quaternary prevention, and contributing to the widening of health disparities in EIDs [445,446]. In Belgium, a survey conducted by the European Social Observatory in 2017 found that 6.7% of individuals in the lowest income quintile reported unmet medical needs due to financial reasons, compared to zero in the highest income quintile, one of the largest gaps within the EU [445]. This highlights the persistent inequalities in access to healthcare, even within a system like Belgium's, which provides mandatory health insurance and generally high-quality care services [445,447].

Finally, the **COVID-19 pandemic** also disrupted access to care for NCDs, with **delays in diagnosis and treatment** [448]. While **telemedicine** was introduced to mitigate these challenges, it may have further deepened existing social disparities, due to unequal digital literacy and limited internet access [146,449,450].

# Main limitations of the thesis

## 1. Availability of individual-level data on unequal transmission and exposure

The first main limitation of this thesis is the lack of data to account for unequal exposure and transmission pathways, which are key drivers of social health inequalities in EID's health outcomes, as discussed in the section "Pathways to social health inequalities in emerging infectious diseases" (p. 365).

A study conducted in England found that inequalities in SARS-CoV-2 transmission factors (e.g., overcrowded households, air quality, rurality, care homes) explained 73% of the effect of deprivation on COVID-19 mortality, followed by inequalities in vulnerability factors (i.e., chronic baseline conditions), accounting for 49% [323]. Although ***Chapters 2*** to ***4*** of this thesis, examining social inequalities in COVID-19 health outcomes, include variables such as employment status and household type, they lack key individual-level granular data on exposure and transmission, such as the type of occupation, household overcrowding, population density, use of public transport, or the ability to work remotely. As a result, the analyses may only partially capture the impact of social inequalities on COVID-19 health outcomes, given that exposure and transmission pathways are crucial within the syndemic framework of the pandemic.

This limitation was particularly evident in ***Chapter 2***, where vaccine uptake and baseline chronic conditions, reflecting unequal access to preventive services and vulnerability, respectively, played only a moderate mediating role in social disparities in COVID-19 hospitalization. The availability of mediating factors accounting for unequal transmission and exposure pathways would likely have provided a more comprehensive explanation of social disparities in severe COVID-19, and helped identify the most critical factors to target in order to reduce them. Therefore, by including mediating factors that account for the four pathways driving social inequalities in EIDs within the syndemic framework, more comprehensive and effective policy recommendations could have been developed.

## 2. Availability of individual-level data on access to treatment and preventive service

Another key pathway driving social inequalities in EID's health outcomes is unequal access to treatment and preventive services, as discussed in the section "Pathways to social health inequalities in emerging infectious diseases" (p. 365). The second main limitation of this thesis is therefore the lack of data to adequately capture this pathway in the analyses.

For instance, in ***Chapter 4***, vaccination status could not be included as the Belgian vaccine register was not available within the HELICON population data linkage. This represents a critical limitation, as evidence shows the effectiveness of the COVID-

19 vaccine against long COVID [451], and social inequalities in vaccination coverage [244]. In addition, individual-level data on access to treatment and preventive services, such as GP density per capita, health insurance coverage, distance to the nearest hospital, or mask wearing were not available in the population data linkages used in this thesis. The absence of these variables hinders the ability to fully capture a key pathway in the syndemic framework: unequal and untimely access to treatment and preventive services. Consequently, this limitation reduces the comprehensiveness of the findings and related policy recommendations.

### 3. Availability of more granular individual-level data on migrant groups

The third major limitation of this thesis is the lack of granularity in the data on migrant groups in the population data linkages used. In ***Chapters 1*** to ***4***, migration background was derived from a combination of parental country of origin and individual country of birth, both classified into three categories: Belgium, EU-countries, and non-EU countries. This resulted in a distinction between Belgian natives, second-generation migrants, first-generation EU migrants, and first-generation non-EU migrants. While this classification allows for a broad distinction based on generation, it masks substantial heterogeneity within each group.

Depending of their country of birth, migrants may face markedly different risks of severe COVID-19 outcomes compared to native individuals, largely due to disparities in occupational exposure, housing conditions, prevalence of

comorbidities, and access to public health messaging and the health system [425,452–455]. The lack of more detailed country-of-birth data therefore limits the capacity to identify within-group disparities and may result in missed opportunities for tailored interventions.

This limitation is particularly relevant in Belgium, and especially in Brussels, where 186 nationalities are represented [456]. Grouping such a diverse population into broad categories risks overlooking the specific needs and barriers experienced by particular migrant communities. As culturally and linguistically adapted interventions have been shown to be more effective in reducing health inequalities among migrant populations [454], the lack of granularity in the data constrained the ability of this thesis to support more tailored public health recommendations.

### 4. Availability of data on the deregistered population

The fourth main limitation of this thesis is the lack of health and social data on the deregistered population in Belgium. The data sources used in this thesis were all linked through the national social security number (NISS), which is assigned exclusively to individuals with legal residency status. Consequently, people without legal residency, most notably undocumented migrants and people experiencing homelessness, were systematically excluded from the analyses. This deregistered population is estimated at approximately 100,000 individuals in Belgium, which is about 1% of the Belgian population [457].

Evidence shows that undocumented migrants and homeless populations face disproportionate risks of infection, morbidity, and mortality during the COVID-19 pandemic, largely due to inequalities in the SDOH, including poorer housing conditions, economic precarity, and limited access to healthcare [458–460]. Their systematic exclusion from the analyses likely results in an underestimation of both the overall burden of the COVID-19 pandemic on population health and the extent of related social health inequalities in Belgium. Furthermore, the absence of this vulnerable part of the population in the data linkages prevents a comprehensive assessment of how they were affected by the pandemic, thereby limiting the evidence base needed to design tailored, equitable, and inclusive public health strategies for pandemic preparedness and response.

# Implications and recommendations

Inequalities in transmission, exposure, vulnerability, and access to treatment and preventive services interact and compound one another, generating social health inequalities associated with EIDs for years. Already back in the early 2000s, the following question arose concerning the incorporation of social equity into pandemic preparedness plans for Influenza: "*In the absence of social, political, and economic equality, racial and ethnic minorities and individuals of low socioeconomic status are left extremely vulnerable to every threat that becomes apparent. How do we build a fair distribution of benefits during a pandemic in a society so burdened by inequality*?" [461]. However, despite over two decades of calls to incorporate equity-focused strategies into pandemic response planning, it has largely been overlooked [391,393,396,461,462]. The persistence of these disparities during the COVID-19 pandemic underscores a failure to apply lessons learned from previous (re-)EID outbreaks, allowing well-documented patterns of unequal impact to persist.

This section examines the ongoing challenge by exploring the implications of key findings and offering recommendations for policy and research to promote a more equitable approach to future pandemic preparedness and response. The time for action is particularly urgent in view of the upcoming rise in EIDs with pandemic potential, poverty, and migratory movements due to wars and climate change.

## *For policy*

The findings of this dissertation imply the following lines of action for policies:

- Promoting social equity in vaccine access and uptake;
- Tackling social inequalities in the burden of comorbidities;
- Tackling the high burden of COVID-19 mortality in care homes;
- Tackling the long-term burden of COVID-19;
- Enhancing timely access to comprehensive individual-level health and social data.

Below, these action areas are detailed and corresponding policy recommendations are outlined.

### 1. Promoting social equity in vaccine access and uptake

This thesis provides important evidence on unequal COVID-19 vaccination coverage (***Chapter 1***) and how increasing equal vaccination rates can help reduce inequalities in COVID-19-related morbidity, albeit with moderate mediating effect (***Chapter 2***). These findings are particularly relevant for policymakers, highlighting the need to increase vaccine uptake among socially deprived and migrant populations in future EID vaccination campaigns.

In Belgium, the COVID-19 vaccination rollout was organized in three phases based on clinical risk and occupational exposure: *Phase 1a* included residents and staff in

care institutions and frontline healthcare workers; *Phase 1b* targeted people aged 65 years and over, those aged 45 to 64 with comorbidities, and essential social and economic function workers; and *Phase 2* extended vaccination to the general adult population [463]. However, **this approach overlooked key SDOH**, which play a critical role in shaping vaccine uptake and subsequent vaccine-preventable health outcomes.

To move towards more equitable vaccination campaigns in future (re-)EID outbreaks, policy could apply an approach based on **proportionate universalism**, which advocates for universal actions, but delivered with a scale and intensity proportionate to the level of social disadvantage, to flatten the social gradient [464]. Applied to vaccination, this approach recognizes that while vaccination should be accessible to all, populations experiencing greater barriers, such as migrants and socially disadvantaged groups, require additional, tailored support to achieve equitable coverage.

However, this approach carries a risk of stigmatization and presents challenges in implementation. It is therefore crucial to carefully identify the populations who struggle to benefit from universal resources, not only on the basis of SE indicators such as income, education, or geographic area, but also by assessing their specific needs to minimize the risk of stigma [465].

**Box 1** provides some recommendations which can help to ensure proportionate universal vaccination allocation resources.

**Box 1. Recommendations to promote equitable vaccination coverage through proportionate universal vaccination allocation resources.**

- **Strengthen community participation through co-created vaccination programs**

Engage local organizations and community leaders representing socially deprived and migrant populations in the design and delivery of vaccination strategies. Such participatory approaches increase cultural relevance and enhance trust in public health authorities, thereby reducing vaccine hesitancy [454,461].

- **Strengthen health and digital literacy to improve access to, understanding of, and use of reliable health information**

Health literacy is a crucial enabler of informed and equitable health decisions [466,467]. Health literacy is generally lower among ethnic minorities, immigrants, and socially deprived populations, yet it plays a crucial role in adopting preventive behaviors and in finding, understanding, and applying government messaging about COVID-19 [468].

**Box 1. Continued**

In 2021, among Organization for Economic Co-operation and Development (OECD) countries, Belgium was one of the few that did not implement specific policies to improve health and digital literacy to address inequities in preventive behavior uptake and the recognition of misinformation [468].

Strengthening health literacy among vulnerable groups can empower individuals to critically evaluate information, identify misinformation, and assess personal health risks. This can be achieved through the following actions:

- Providing timely, accurate, and accessible COVID-19 information in all relevant languages. Public health messaging should be clear, actionable, and accessible (i.e., adapted to varying literacy levels and available in multiple formats to take into account those with no internet connection or low digital literacy) [454,469] ;
- Developing culturally competent education and prevention campaigns in collaboration with community organizations and NGOs [468] ;
- Improving health professionals' communication skills to create supportive environments that enable vulnerable populations to understand and apply health information effectively [468].

**Box 1. Continued**

- Strengthening coordination between scientific and political sectors to ensure consistent messaging and avoid an overwhelming flow of (mis)information leading to public confusion and mistrust [470,471].

- **Address practical barriers to <u>vaccine access</u>**

Adapt vaccination services to people's lived realities by addressing logistical obstacles such as lack of transportation, inflexible work hours, and child care responsibilities. Solutions may include mobile vaccination units, extended service hours, workplace-based delivery, or even door-to-door outreach [454,461].

## 2. Tackling the high burden of COVID-19 mortality in care homes

As demonstrated in ***<u>Chapter 3</u>***, living in care homes was the key predictor of COVID-19 mortality among the Belgian adult population, aligning with many national [120,472] and international findings [473]. In Belgium, 44.7% of all notified COVID-19 deaths between the first and seventh waves occurred among care home residents [474]. During the first wave, excess mortality in these facilities was nearly double that observed in the community among individuals of the same age and sex

[475], underscoring that residing in a care home constituted a distinct risk environment.

Several factors contributed to this disproportionate mortality burden. At the individual level, care home residents were highly vulnerable due to advanced age, frailty, and comorbidities. At the institutional level, care home settings heightened exposure risks, primarily through staff serving as vectors of transmission from the community, particularly when employees worked across multiple facilities [115,476,477]. At the systemic level, public authorities focused primarily on preserving hospital capacity, often leaving care homes to function as improvised medical units without sufficient personal protective equipment (PPE), diagnostic capacity, or epidemic management expertise in closed settings [478]. These gaps induced early and rapid viral spread, facilitated by delayed government measures [475].

**Box 2** outlines some recommendations for strengthening pandemic preparedness and response in Belgian care homes to reduce excess mortality in future health crises.

**Box 2. Recommendations to tackle the high burden of COVID-19 mortality in care homes**

- **Implement systematic and regular testing of staff and residents, including asymptomatic screening**

Undetected introductions of SARS-CoV-2 were a major driver of outbreaks in care homes [479]. Evidence from Belgium and other European countries demonstrated a strong correlation between COVID-19 deaths in the community and in care homes, underscoring the importance of community transmission [473]. At the beginning of the pandemic, only symptomatic residents were tested in Belgium [480]. Expanded testing should encompass systematic and regular screening of staff and residents, with intensified frequency during high community incidence to inform IPC, isolation, and cohorting strategies [478].

- **Guarantee sufficient and adapted infection prevention and control (IPC) training**

During the first wave of COVID-19, Belgian care homes experienced acute shortages of staff and personal protective equipment (PPE), which severely hampered efforts to reduce viral transmission [115,478,481].

**Box 2. Continued**

IPC measures were often insufficient or poorly adapted, largely due to limited access to PPE, inadequate training in epidemic management among care-home staff, and the absence of clear, practical guidelines. To address these shortcomings, authorities should establish dedicated PPE stockpiles for long-term care facilities, supported by efficient distribution systems to prevent shortages during health emergencies. In parallel, ongoing and practical IPC training must be provided to all staff, including auxiliary and temporary personnel, to bridge gaps in knowledge and practice revealed during the pandemic, and to ensure the consistent application of protocols [478].

- **Provide external support to care homes in need**

Targeted support for care homes during crises is crucial. The intervention of *Médecins Sans Frontières* in Belgium demonstrated that mobile teams providing on-site coaching, psychological first aid, and practical guidance were highly valued, as they strengthened staff confidence and enhanced the overall crisis response. Building on this model, authorities should establish dedicated mobile units with the capacity for rapid deployment, ensuring timely assistance to care homes in need during future health emergencies [478].

## 3. Tackling social inequalities in the burden of comorbidities

Our findings highlight that comorbidities play a moderate mediating role in social disparities in COVID-19 morbidity (***Chapter 2***) and are among the strongest predictors of post-acute COVID-19-related organ complications following severe COVID-19 (***Chapter 4***). Reducing health inequalities in EID-related health outcomes therefore requires targeted policies aimed at reducing the burden of comorbidities, particularly among individuals with lower SES.

Belgium's healthcare system generally performs well: in 2018, 74% of the adult population reporting being in good health, compared to the EU average of 69% [118]. The system operates under a compulsory health insurance managed by RIZIV-INAMI, covering 99% of the population, with broad benefits and direct access to healthcare providers due to the absence of gatekeeping. Funded through social insurance principles, the system ensures both horizontal solidarity (between healthy and diseased individuals) and vertical solidarity (ensuring higher reimbursement rates based on income and social status) [118,447].

However, **health disparities persist in Belgium** [118,445]. In 2019, 59% of individuals in the lowest income quintile reported being in good health, compared to 87% in the highest, representing the largest gap among Western EU countries [118]. Indeed, chronic conditions are more prevalent among individuals in the lowest compared to the highest income quintile (39% vs. 16%), mainly due to

**unequal healthcare access** and greater **exposure to risk factors** [118]. In 2019, 4% of individuals in the lowest income group reported **unmet medical needs**, primarily for **financial reasons**, compared to 0.2% in the highest quintile [118]. **Behavioral risk factors**, such as smoking, unhealthy diet, and inactivity, are also more prevalent among individuals with lower SES, largely contributing to inequalities in health and life expectancy [118,482–484].

**Box 3** outlines policy recommendations for future pandemic preparedness plans aiming to reduce the disproportionate burden of comorbidities that drive unequal vulnerability to infection and increase morbidity and mortality rates among individuals with lower SES.

**Box 3. Recommendations to reduce social inequalities in the burden of comorbidities**

- **Strengthen targeted <u>prevention</u> programs**

In 2019, Belgium allocated 1.6% of total health spending to prevention, compared to the EU average of 2.9%, and reported higher rates of preventable mortality than many Western EU countries. Increasing health spending on prevention, particularly for socially vulnerable populations, could help reduce the burden of behavioral risk factors, which disproportionally affects socially vulnerable groups and accounted for one-third of all deaths in 2019 [118].

**Box 3. Continued**

- **Strengthen access to primary healthcare in deprived areas**

Community health workers (CHWs), deployed following the COVID-19 crisis, have proven effective in reaching socially vulnerable groups [485]. Sustained funding and stronger integration of CHWs in the formal healthcare system are crucial to overcome barriers such as mistrust, administrative complexity, and language or legal issues [486,487].

Additionally, healthcare centers should be equitably distributed geographically and accessible to individuals with reduced mobility, to ensure timely access to local primary care in deprived areas [486].

- **Ensure equitable access to teleconsultation**

While teleconsultation improved healthcare access during the pandemic [118], digital exclusion remains a barrier for many socially disadvantaged communities [486]. Policies should promote digital literacy, provide affordable devices, and ensure access to reliable high-speed internet to guarantee equal access to teleconsultation in future health crises.

## 4. Reducing the long-term burden of severe COVID-19

In ***Chapter 4***, significant post-acute CV and pulmonary complications were identified following COVID-19 hospitalization. These findings provide important evidence for policymakers regarding the long-term health impacts of severe COVID-19 and underscore the need to implement effective measures to address them. **Box 4** outlines several recommendations to help mitigate the long-term burden of severe COVID-19, that can be applied in response to future EIDs with similar characteristics.

**Box 4. Recommendations to reduce the long-term burden of severe COVID-19**

- **Strengthen general practitioners' knowledge of long COVID and promote multidisciplinary care**

In 2022, most GPs reported insufficient knowledge of long COVID and a lack of scientific guidance on its diagnosis and management [488]. It is essential to enhance their understanding of potential post-acute organ complications and emphasize the importance of targeted follow-up for socially vulnerable populations, who are more prone to comorbidities and face greater barriers to care. Given the complexity of long COVID, promoting multidisciplinary collaboration, recognized as important by 93% of GPs but not consistently implemented, is also critical [488].

**Box 4. Continued**

- **Raise awareness among healthcare providers about the personalized long COVID care pathway**

In Belgium, patients suffering from long COVID, regardless of hospitalization status, can benefit from a personalized care pathway established by RIZIV-INAMI [489]. Since July 2022, and updated most recently in July 2024, this pathway offers full reimbursement for a range of frontline services (e.g., physiotherapy, psychology, dietetics), provided the condition is certified by a GP or specialist [489]. However, awareness of this program among healthcare providers remains low, leading to its underutilization [490]. Improving communication about the availability and implementation of this care pathway is essential to support patient recovery.

- **Strengthen post-hospital follow-up for COVID-19 patients**

Several initiatives have been implemented in Belgium to support post-hospitalization follow-up for COVID-19 patients, including home-based telemonitoring and/or intermediate care structures [489,491]. However, the proportion of patients who have effectively benefited from these services remains limited [492].

**Box 4. Continued**

Strengthening and institutionalizing these follow-up mechanisms would help ensure optimal care for patients recovering from severe COVID-19.

## 5. Enhancing timely access to comprehensive individual-level health and social data

Although Belgium has many high-quality health and social data sources, researchers often face significant administrative, technical, and ethical barriers that impede timely access and use of comprehensive individual-level data. This is a critical issue, particularly in the context of public health emergencies, where rapid, evidence-based policy decisions are essential.

The HELICON population data linkage exemplifies these challenges. Following ethical approval, it took nearly three years to access the necessary linked datasets, resulting in major delays in conducting analyses and delivering findings to support decision-making. Due to technical and administrative difficulties, the project was also unable to include data from testing and vaccination registries, an omission that constrained the application of the syndemic framework of the pandemic, and therefore, the ability to fully address important research questions. Additionally, HELICON's restricted access to detailed data formats, such as exact dates of death or hospital admission, created further methodological challenges, constraining the

use of appropriate statistical techniques like survival analysis. Also, the restricted access to granular data, for example on the type of occupation or the exact country of origin, further restricted the precision of our conclusions and prevented related tailored recommendations.

**Box 5** outlines policy recommendations to address the limited and delayed access to health and social data in Belgium, which remains a major barrier to timely responses during public health crises.

**Box 5. Recommendations to enhance timely access to comprehensive individual-level health and social data**

- **Support the collection and use of individual-level data within the syndemic framework**

To fully understand and effectively address social health inequalities during EID outbreaks, it is essential to collect and make available comprehensive data within the syndemic framework, accounting for the four critical pathways driving social inequalities in EIDs: transmission, exposure, vulnerability, and access to treatment and preventive services. Policymakers should prioritize facilitating researchers' (timely) access to such multidimensional data, enabling analyses that reflect the full complexity of the syndemic framework.

**Box 5. Continued**

This is crucial for identifying the root causes of social disparities, informing tailored interventions, and ultimately effectively reducing inequities in EID-related health outcomes.

- **Support the collection and use of more granular individual-level country-of-origin data**

To ensure tailored recommendations and further health equity, the collection and use of individual-level data on cases, hospitalizations, deaths, and vaccination by migrant status with the specific country on origin should be supported [454].

- **Facilitate the automatic and systematic creation of a NISS Bis number for individuals not officially registered in Belgium**

The NISS Bis number provides a unique identifier without requiring official registration in the national register. Expanding and facilitating its use would both improve access to essential services for unregistered individuals (e.g., the BIS number gave access to COVID-19 vaccines) and allow to include unregistered populations in social and health databases through this unique identifier. This, in turn, strengthens the evidence base for policy decisions and helps design interventions that reflect the health and social realities of the entire population.

**Box 5. Continued**

- **Support the development of a sustainable longitudinal population-based Belgian cohort**

The COVID-19 pandemic highlighted the importance of longitudinal population-based cohorts in answering urgent policy questions, such post-infection sequelae, vaccine effectiveness, and societal impacts [493–495]. These cohorts also enabled comparisons between pre- and post-pandemic situations [496]. Although Belgium has the potential to support this type of system, through the NISS number and secure linkage platforms (i.e., Healthdata.be), such infrastructure does not yet exist [497]. Implementing it would allow for the tracking of individuals and events such as changes in employment, housing, and education, which are critical for understanding health inequalities and shaping preventive policies.

- **Move toward an integrated national health information system**

In the wake of the COVID-19 crisis, the Belgian government acknowledged the critical need for enhanced access to and availability of health and social data. In March 2023, the Belgian Health Data Agency (HDA) was launched as part of an integrated federal initiative to facilitate the (re-)use of routine health and social data for policy development and research [498].

**Box 5. Continued**

This national initiative provides a framework that will strengthen Belgium's response to future (re-)EIDs and improve overall pandemic preparedness.

## *For research*

The findings of this dissertation underscore the critical need for researchers to collect social indicators alongside clinical data in EID surveillance to promote equitable resource allocation and policy evaluation during pandemics. Furthermore, this thesis demonstrated the value of using causal inference methods, such as mediation analysis, to identify key factors that should be targeted to reduce social inequalities in EID health outcomes.

**Box 6** provides recommendations for future research in Belgium and beyond, aimed at further monitoring and understanding the scope, origins, and consequences of EID-related health inequalities.

**Box 6. Recommendations for future research**

- **Rethink emerging infectious disease routine surveillance system by systematically integrating social indicators**

In Belgium, this is already underway with the EPI-LINK project set up by Sciensano, which aims to couple clinical data with other administrative registers, such as vaccination, mortality, SD, and SE data. This project will enable routine surveillance of infectious diseases and support the documentation of responses to public health emergencies as part of pandemic preparedness [499].

**Box 6. Continued**

- **Continue investigating the causal pathways driving EID-related health inequalities within the syndemic framework**

Deepen research into transmission, exposure, vulnerability, and access to treatment and preventive services factors that influence the relationship between social inequalities and severe health outcomes during infectious disease pandemics, in order to identify the key risk factors that should be targeted to reduce these inequalities.

- **Apply intersectional approaches in epidemiological research**

Unravel how overlapping social determinants (e.g., ethnicity, gender, SE status) shape vulnerability during infectious disease outbreaks, to better tailor prevention and care strategies [500].

- **Apply mixed-methods designs**

Future research should apply mixed-methods designs to better understand social inequalities in vaccination . Quantitative methods can identify patterns of under vaccination across groups defined by age, migration background, socioeconomic status, or type of care facility. Qualitative methods can then explore underlying barriers, such as mistrust, language or administrative obstacles. Such evidence is crucial to design equitable strategies that improve vaccine uptake among the most vulnerable populations.

**Box 6. Continued**

- **Cross-country analysis of social inequalities in EIDs' outcomes**

Promote international comparative research on social inequalities in EID outcomes by harmonizing study populations, indicators, and methodologies across countries. Once disparities are mapped, further analysis could focus on the underlying drivers, such as differences in political or health systems, to inform evidence-based strategies that reduce inequities. The European Health Data Space can support this by providing standardized, cross-country health and social data, enabling robust comparisons and analysis of underlying drivers of EID outcome disparities [501].

- **Leverage artificial intelligence for early detection of emerging inequalities**

Use AI to automate data analysis, monitor trends, and quickly detect social disparities in real time during EID outbreaks [502]. Within the EU, the AI Act is the first legal framework aiming at fostering trustworthy AI in Europe [503]

# Conclusion

The findings of this dissertation call for urgent and sustained policy action to tackle persistent social health inequalities that have long shaped the impact of EIDs, as well as the long-term health direct impacts of COVID-19 on organ systems. Action is urgently needed, particularly in the Belgian context, where an aging population, ongoing migratory movements, a high prevalence of chronic conditions, and widening social inequalities amplify the challenges.

To move toward a more equitable health system and reduce the long-term consequences of COVID-19 and future similar vaccine-preventable viral respiratory EIDs, policymakers should prioritize the following actions:

- ➔ **Integrate social equity into pandemic preparedness and response;**
- ➔ **Enable real-time monitoring of social health inequalities in EIDs routine surveillance system;**
- ➔ **Strengthen primary healthcare and prevention in socially deprived areas;**
- ➔ **Ensure long-term follow-up after COVID-19 hospitalization.**

These policy directions are essential to reduce the inequitable impacts of EIDs and their long-term direct health consequences, with the aim of building a more resilient and equitable public health system for future health crises.

# Appendices

*Table of content*

- **Appendix A.** Variables list, description and data inclusion/removal of the Clinical survey admission and discharge forms.
- **Appendix B.** Description of the substances, Anatomical Therapeutic Chemical (ATC) subcodes and ATC codes used to build the pseudopathologies available in the Intermutualistic Agency (IMA) database.
- **Appendix C.** R code to reproduce the working example computing the density distribution of propensity scores, the relative contributions of all individuals in the overlap weighted pseudo-population and inverse probability of treatment weighted pseudo-population, along with the standardized mean difference by balancing weight method.
- **Appendix D.** R code to reproduce the working example implementing a conditional inference tree algorithm and assessing its predictive performance.
- **Appendix E.** R code to reproduce the working example illustrating the differences in estimating the incidence of COVID-19 mortality using the

Kaplan-Meier vs. competing risk method, in the presence of competing risks.

- **Appendix F.** R code to reproduce the working example illustrating the calculation of the cumulative incidence function by subgroup and the use of Gray's test to evaluate statistical differences between groups.

# Appendix A.

**Variables list, description and data inclusion/removal of the Clinical survey admission and discharge forms. Source: Produced by Van Goethem N. and reused with permission.**

| Variable name | Description | Value set or terminology | Date inclusion or removal[1] |
|---|---|---|---|
| **Admission questionnaire** | | | |
| dt_reporting_admis | Declaration date | Date | |
| Hospital_name_admis | Hospital name | Dropdown list | |
| Hosp_patient_ID | Hospital record ID of the patient | Text | |
| National_nr | National registry number of the patient - pseudonimized | Text | From Sept. 14 |
| dob | Birth date of patient | Date | |
| gender | Gender or patient | Male/Female/Other/ Unknown | |
| Post_code | Post code of patient (in Belgium) | Numeric | |
| Country_resid | Country of residence | Belg./Unknown/Other | From June 8 |
| country_resid_other | Specification of country of residence, if 'other' | Text | From June 8 |
| Readmission | Previous hospitalization for COVID-19 and readmission | Yes/No | From April 22 |

[1] Date of inclusion (from …) / removal (up to …) of a variable in the questionnaire is only indicated if not collected from the initiation of the survey or if removed during the surveillance period up to 8th of June (last survey update).

| | | | |
|---|---|---|---|
| Prev_admis_same_hosp | If readmitted for covid, from same hospital? | Yes/No | From April 22, up to Sept. 14 |
| Prev_admis_hosp_patient_id | if same hospit, hospital patient id from previous hospital stay | Text | From April 22, up to Sept. 14 |
| Prev_admis_hospital_name | if readmitted after discharge from other hospital, hospital name | Dropdown list | From April 22, up to Sept. 14 |
| Prev_admis_dt_discharge | if readmitted after discharge from other hospital, date discharge | Date | From April 22, up to Sept. 14 |
| Expo_HCW | Patient is a health care worker | Yes/No/ Unknown | |
| Expo_retirement_home | Patient is a nursing home resident | Yes/No/Unknown | |
| Expo_collective_housing | Patient is a resident in a collective housing | No / Yes, revalidation centre / Yes, psychiatric centre / Yes, homeless housing / Yes, other / Unknown | From June 8, up to Sept. 14 |
| IndexCase | Index case in the HH (< 18)[2] | Patient / Yes, adult / Yes, child / Unknown | From June 8 |
| N_household | Number of persons living in the same household (<18) | Numeric | From June 8 |
| Ethnic | Ethnic group of the patient | European /Hispanic / Asian / Subsaharan African / North African / Mixed / Other / Unknown | From June 8 |
| Test_for_covid | Test for symptoms suggestive of COVID | Yes/No | From June 8 |
| Test_systematic | Systematic testing of an hospitalized patient | Yes/No | From June 8 |

[2] Questions asked only for patients below the age of 18 are indicated by (<18)

| | | | |
|---|---|---|---|
| Test_other | Test for other reason | Yes/No | From June 8 |
| Test_unknown | Unknown reason for testing | Yes/No | From June 8 |
| Expo_travel | Exposition – Travel to risk region | Yes/No | Up to June 8 |
| Expo_Contact_confirmed | Exposition – Contact with confirmed COVID-19 patient | Yes/No | Up to Sept. 14 |
| Expo_contact_probable | Exposition – Contact with probable COVID-19 patient | Yes/No | Up to Sept. 14 |
| Expo_nosocomial | Exposition – Suspicion of nosocomial infection[3] | Yes/No | Up to Sept. 14 |
| Expo_no_identified | Exposition – Not identified[4] | Yes/No | Up to Sept. 14 |
| Expo_unknown | Exposition – Unknown [5] | Yes/No | Up to Sept. 14 |
| Expo_other | Exposition - Other | Text field | Up to Sept. 14 |
| Dt_onset | Onset date of the symptoms | Date | |
| Dt_admission | Date of hospital admission | Date | |
| Reason_clinical_status | Reason 407odgkin407ization – Clinical status related to COVID-19 | Yes/No | Up to Sept. 14 |
| Reason_at_risk | Reason 407odgkin407ization – Person at risk for COVID-19 | Yes/No | Up to Sept. 14 |

[3] The official definition of 48 hours of hospitalization to consider an infection as nosocomial is not valid in the COVID-19 context, as SARS-CoV-2 has a longer incubation period. Thus, this reporting is based on the suspicion from the clinician, rather than the timing since hospital admission.

[4] If the patient has been queried regarding the exposure, but no exposure was identified.

[5] If the patient has not been queried regarding the exposure or there is no record of such an inquiry within the patient medical file.

| | | | |
|---|---|---|---|
| Reason_entourage | Reason 408odgkin408ization – Entourage at risk for COVID-19 | Yes/No | Up to Sept. 14 |
| Reason_imposs_at_home | Reason 408odgkin408ization – Isolation at home impossible | Yes/No | Up to Sept. 14 |
| Reason_noncov_hospit | Reason 408odgkin408ization – Hospitalisation for another reason no related to COVID | Yes/No | From June 8, up to Sept. 14 |
| Reason_transfer | Reason 408odgkin408ization – Transfer from other hospital | Yes/No | |
| Transfer_from | Transfer from which hospital | Dropdown list | |
| transfer_from_other | Specification of hospital name, if other | Text | |
| Reason_other | Reason hospitalision – Other | Yes/No | Up to Sept. 14 |
| Symptom_fever | Symptom at admission – Fever | Yes/No | |
| Symptom_viral | Symptom at admission – Viral syndrome: including previously collected weakness, headache, muscle & joint pain | Yes/No | From Sept. 14 |
| Symptom_weakness | Symptom at admission – General weakness | Yes/No | Up to Sept. 14 |
| Symptom_upper_respi | Symptom at admission of upper respiratory tract infection: including previously collected throat pain, runny nose | Yes/No | From Sept. 14 |
| Symptom_low_respi | Symptom at admission of low respiratory tract infection: including | Yes/No | From Sept. 14 |

| | | | |
|---|---|---|---|
| | previously collected cough, breathlessness, thoracic pain | | |
| Symptom_GI | Symptom at admission – gastro-intestinal: including previously collected diarrhea, nausea / vomiting, abdominal pain | | From Sept. 14 |
| Symptom_cough | Symptom at admission – Cough | Yes/No | Up to Sept. 14 |
| Symptom_throat_pain | Symptom at admission – Throat pain | Yes/No | Up to Sept. 14 |
| Symptom_runny_nose | Symptom at admission – Runny nose | Yes/No | Up to Sept. 14 |
| Symptom_anosmia | Symptom at admission – anosmia-agueusia | Yes/No | From march 23 |
| Symptom_breathless | Symptom at admission – Breathlessness | Yes/No | Up to Sept. 14 |
| Symptom_diarrhoea | Symptom at admission – Diarrhoea | Yes/No | Up to Sept. 14 |
| Symptom_nausea_vomit | Symptom at admission – Nausea | Yes/No | Up to Sept. 14 |
| Symptom_loss_apetite | Symptom at admission: reduced feeding (< 18) | Yes/No | From June 8 |
| Symptom_headache | Symptom at admission – Headache | Yes/No | Up to Sept. 14 |
| Symptom_mental | Symptom at admission – Irritability | Yes/No | Up to Sept. 14 |
| Symptom_pain | Symptom at admission – Pain | Yes/No | Up to Sept. 14 |
| Symptom_rash | Symptom at admission – Skin rash (< 18) | Yes/No | From June 8 |
| Symtom_erythema | Symptom at admission – Erythema multiforme (< 18) | Yes/No | From June 8 |
| Symptom_urtic | Symptom at admission – Urticaria (< 18) | Yes/No | From June 8 |
| Symptom_petechia | Symptom at admission – Petechia (< 18) | Yes/No | From June 8 |

| | | | |
|---|---|---|---|
| Pain_muscle | Pain location – Muscles | Yes/No | Up to Sept. 14 |
| Pain_abdomen | Pain location – Abdomen | Yes/No | Up to Sept. 14 |
| Pain_thorax | Pain location – Thorax | Yes/No | Up to Sept. 14 |
| Pain_joint | Pain location – Joint | Yes/No | Up to Sept. 14 |
| Pain_other | Pain location – Other | Yes/No | Up to Sept. 14 |
| Symptom_none | Symptom at admission – None | Yes/No | |
| Symptom_other | Symptom at admission – Other | Text field | |
| Clin_sign_pharyngitis | Clinical sign – Pharyngeal exsudate | Yes/No | From Sept. 14 for < 18 only |
| Clin_sign_conjonctivitis | Clinical sign – Conjonctival injection | Yes/No | From Sept. 14 for < 18 only |
| Clin_sign_convulsions | Clinical sign – Convulsions | Yes/No | From Sept. 14 for < 18 only |
| Clin_sign_coma | Clinical sign – Coma | Yes/No | From Sept. 14 for < 18 only |
| Clin_sign_dyspnea | Clinical sign – Dyspnea | Yes/No | From Sept. 14 for < 18 only |
| Clin_sign_lung_auscult_abnorm | Clinical signs – Abnormal pulmonary auscultation | Yes/No | From Sept. 14 for < 18 only |
| Clin_sign_lung_Xray_abnorm | Clinical signs – Abnormal pulmonary imaging compatible with viral pneumonia | Yes/No | From Sept. 14 for < 18 only |
| Clin_sign_none | Clinical signs – None | Yes/No | From Sept. 14 for < 18 only |

| | | | |
|---|---|---|---|
| Clin_sign_Other | Clinical signs – Other | Text field | From Sept. 14 for < 18 only |
| Temperature | Temperature | Numeric | From Sept. 14 for < 18 only |
| Freq_cardiac | Cardiac frequency (< 18) | Count/min – numeric | From June 8 |
| Freq_respi | Respiratory frequency (< 18) | Count/min – numeric | From June 8 |
| BP_syst | Systolic Blood pressure (< 18) | Numeric (mmHg) | From June 8 |
| BP_diast | Diastolic Blood pressure (< 18) | Numeric (mmHg) | From June 8 |
| Value_SpO2_admission | Value of SpO2 at hospital admission (< 18) | Numeric (%) | From June 8 |
| Comorbid_CVD | Comorbidities – Cardiovasular disease | | |
| Comorbid_HBP | Comorbidities – Arterial hypertension | Yes/No | |
| Comorbid_diabetes | Comorbidities – Diabetes | Yes/No | |
| Comorbid_renal_chronic | Comorbidities – Chronic renal disease | Yes/No | |
| Comorbid_liver_chronic | Comorbidities – Chronic liver disease | Yes/No | |
| Comorbid_neuro_chronic | Comorbidities – Chronic neurologic or neuromuscular disease (with the exception of cognitive troubles)[6] | Yes/No | |
| Comorbid_cognitive | Comorbidities – Cognitive troubles | Yes/No | From March 23 |

[6] Initially, it was not specified that cognitive conditions were excluded from the category "Comorbid_neuro_chronic"; from 23rd March on, the definition was clarified: "chronic neurologic conditions excluding cognitive disorders", and the category "Cognitive disorders" was added. Cognitive disorder is indicated as "N/A" for the cases reported before the 23rd March.

| | | | |
|---|---|---|---|
| Comorbid_immunodep | Comorbidities – Immunodepresion, including HIV | Yes/No | |
| Comorbid_lung_chronic | Comorbidities – Chronic lung disease | Yes/No | |
| Comorbid_solid_cancer | Comorbidities – Solid cancer | Yes/No | |
| Comorbid_hemato_cancer | Comorbidities – Hematological cancer | Yes/No | |
| Comorbid_transplant | Comorbidities – Solid organ transplantation | Yes/No | From June 8 |
| Comorbid_pregnancy | Comorbidities – Pregnancy | Yes/No | |
| Comorbid_postpartum | Comorbidities – Post partum (< 6 weeks) | Yes/No | |
| Comorbid_premature | Comorbidities – Prematurity | Yes/No | From Sept. 14 for < 18 only |
| Comorbid_obesity | Comorbidities – Obesity | Yes/No | From April 3 |
| Comorbid_none | Comorbidities – None | Yes/No | |
| Comorbid_other | Comorbidities – Other | Yes/No | From April 3 |
| Cardiac_failure | If CVD, cardiac failure? (< 18) | Yes/No | From June 8 |
| Lung_details | If comorbid_lung_chronic = “Yes”, details (< 18) | Cystic fibrosis / Severe asthma / Bronchiectasis / Ciliary diskynesia / tuberculosis / dysplasia / Other / Unknown | From June 8 |
| Immuno_details | If immunodepression = “Yes”, details (< 18) | HIV / primary immunodeficiency / immunosuppressive treatment / Other / Unknown | From June 8 |
| Transplant_organ | If comorbid_transplant = “Yes”, which organ? (< 18) | Kidney / Liver / Lung / Heart / Other | From June 8 |

| | | | |
|---|---|---|---|
| Dt_transplant | If comorbid_transplant = “Yes”, date of transplantation? (<18) | date | From June 8 |
| Solid_cancer_active | If comorbid_solid_cancer = “Yes”, active in last 5 years? | Yes/No/Unknown | From June 8 |
| S_cancer_treat6M | If comorbid_solid_cancer = “Yes”, treatment in the last 6 months? | Yes/No/Unknown | From June 8, up to Sept. 14 |
| Cancer_bladder | if comorbid_solid_cancer = “Yes”, cancer of bladder | Yes/No | From June 8, up to Sept. 14 |
| Cancer_breast | if comorbid_solid_cancer = “Yes”, cancer of breast | Yes/No | From June 8, up to Sept. 14 |
| Cancer_colon | if comorbid_solid_cancer = “Yes”, cancer of colon/rectum | Yes/No | From June 8, up to Sept. 14 |
| Cancer_endometrium | if comorbid_solid_cancer = “Yes”, cancer of endometrium | Yes/No | From June 8, up to Sept. 14 |
| Cancer_headneck | if comorbid_solid_cancer = “Yes”, cancer of head & neck | Yes/No | From June 8, up to Sept. 14 |
| Cancer_kidney | if comorbid_solid_cancer = “Yes”, cancer of kidney | Yes/No | From June 8, up to Sept. 14 |
| Cancer_liver | if comorbid_solid_cancer = “Yes”, cancer of liver | Yes/No | From June 8, up to Sept. 14 |
| Cancer_lung | if comorbid_solid_cancer = “Yes”, cancer of lung | Yes/No | From June 8, up to Sept. 14 |
| Cancer_melanoma | if comorbid_solid_cancer = “Yes”, melanoma | Yes/No | From June 8, up to Sept. 14 |

| | | | |
|---|---|---|---|
| Cancer_pancreas | if comorbid_solid_cancer = “Yes”, cancer of pancreas | Yes/No | From June 8, up to Sept. 14 |
| Cancer_prostate | if comorbid_solid_cancer = “Yes”, cancer of prostate | Yes/No | From June 8, up to Sept. 14 |
| Cancer_thyroid | if comorbid_solid_cancer = “Yes”, cancer of thyroid | Yes/No | From June 8, up to Sept. 14 |
| Cancer_ovarian | if comorbid_solid_cancer = “Yes”, cancer of ovary | Yes/No | From June 8, up to Sept. 14 |
| Cancer_other | if comorbid_solid_cancer = “Yes”, other cancer | Text field | From June 8, up to Sept. 14 |
| Cancer_unknown | if comorbid_solid_cancer = “Yes”, unknown location of cancer | Yes/No | From June 8, up to Sept. 14 |
| Hemato_cancer_active | If comorbid_hemato_cancer = “Yes”, active in last 5 years? | Yes/No/Unknown | From June 8 |
| H_cancer_treat6M | If comorbid_hemato_cancer = “Yes”, treatment in the last 6 months? | Yes/No/Unknown | From June 8, up to Sept. 14 |
| H_Cancer_hodgkin | if comorbid_solid_cancer = “Yes”, Hodgkin lymphoma | Yes/No | From June 8, up to Sept. 14 |
| H_Cancer_leukemia | if comorbid_solid_cancer = “Yes”, leukemia | Yes/No | From June 8, up to Sept. 14 |
| H_Cancer_myeloma | if comorbid_solid_cancer = “Yes”, multiple myeloma | Yes/No | From June 8, up to Sept. 14 |
| H_Cancer_non_hodgkin | if comorbid_solid_cancer = “Yes”, non 414odgkin lymphoma | Yes/No | From June 8, up to Sept. 14 |

| | | | |
|---|---|---|---|
| H_Cancer_other | if comorbid_solid_cancer = “Yes”, other | Text field | From June 8, up to Sept. 14 |
| H_Cancer_unknown | if comorbid_solid_cancer = “Yes”, unknown type of cancer | Yes/No | From June 8, up to Sept. 14 |
| Hemato_stemcell | if comorbid_hemato_cancer = “Yes”, patient received stem cell transplantation? (< 18) | Yes/No | From June 8 |
| Dt_stemcell | if hemato_stemcell = “Yes”, date of transplantation (< 18) | date | From June 8 |
| Prema_preg_weeks | if comorbid_premature = “Yes”, weeks of pregnancy at birth (< 18) | Numeric (weeks) | From June 8 |
| Prema_birth_weight | if comorbid_premature = “Yes”, weight at birth | Numeric (gr) | From June 8 |
| Pregnancy_trimester | Pregnancy trimester (< 18) | 1st/2nd/3rd/Unknown | |
| Current_smoker | Current smoker ? | Yes/No/Unknown | |
| Influenza_vaccine | Has the patient been vaccinated against influenza (season 2019-2020)? | Yes/No/Unknown | Up to Sept. 14 |
| Treat_IEC_Sartan | Is the patient under IEC and/or SARTAN treatment? | Yes, an IEC / Yes, a SARTAN / Yes, an IEC and a SARTAN / No / Unknown | Up to Sept. 14 |
| Treat_NSAI | Is the patient under usual treatment NSAI? (< 18) | Yes/No/Unknown | From June 8 |
| Treat_corticoid_PO | Is the patient under usual treatment with oral corticoids? (< 18) | Yes/No/Unknown | From June 8 |

| | | | |
|---|---|---|---|
| Cortic_dose | if Treat_corticoid_PO="Yes" then which dose. (< 18) | Numeric (/kg/day) | From June 8 |
| Treat_immunodep | Is the patient under usual treatment with immunodepressor? (< 18) | Yes/No/Unknown | From June 8 |
| Lab_name | Laboratory name | Text | Up to April 3 |
| Dt_sample | Date of sample for COVID-19 diagnosis | Date | Up to April 3 |
| Type_sample | Type of sample | Nasopharyngeal swab / nasopharyngeal aspirate / BAL / Other | Up to April 3 |
| Type_sample_other | Specification of type of sample, if other | | Up to April 3 |
| Dt_lab_confirmation | Date of lab confirmation | Date | From April 3 |
| Diag_method_PCR | Diagnostic method – Positive result by PCR? | Yes/No | From April 3 |
| Diag_method_Scan | Diagnostic method – Typical thoracic scan | Yes/No | From April 3 |
| Diag_method_Ag_rapid | Diagnostic method – Positive rapid antigen test | Yes/No | From April 3 |
| Diag_method_Unk | Diagnostic method – Unknown | Yes/No | From April 3 |
| dt_diagnosis | Date du diagnostic [7] | Date | |

[7] Initially it was asked to report all patients with a laboratory-confirmed diagnosis of COVID-19. From the 3rd April on, cases confirmed with CT-scan were also recorded (in line with the change of the official case definition by Sciensano). Thus, before the 3rd of April, the method of diagnosis was not asked and the date of laboratory confirmation was recorded as dt_lab_confirmation. From the 3rd April on, the method of diagnostic confirmation is now asked as: diag_method_PCR / diag_method_scan / diag_method_age_rapid / diag_method_unk and the date of diagnosis is recorded as dt_diagnosis (if it is confirmed by several methods, the date of the first confirmation is asked).

| | | | |
|---|---|---|---|
| Name_admis | Name of the person reporting the admission[8] | Text | |
| Phone_admis | Phone number of the person reporting the admission[8] | Text | |
| Email_admis | Email address of the person reporting the admission[8] | Text | |
| **Discharge questionnaire** | | | |
| Dt_reporting_disch | Declaration date | Date | |
| Hospital_name_disch | Hospital name | dropdown list | |
| Pneumonia_Xray | Pneumonia on medical imaging (X-ray or CT-scan)[9] | Yes/No**/**Not done**/**Unknown | |
| Pneumonia_localisation | if pneumonia_Xray="Yes", Location of pneumonia | Bilateral/Unilateral/Unknown | Up to Sept. 14 |
| X_ray_not_performed | X-ray: not perfomed (< 18) | Yes/No | From June 8 |
| X_ray_normal | X-ray: normal (< 18) | Yes/No | From June 8 |
| X_ray_inters_syndr | X-ray: interstitial syndrome (< 18) | Yes/No | From June 8 |
| X_ray_cardiomegaly | X-ray: cardiomegaly (< 18) | Yes/No | From June 8 |
| X_ray_pleural | X-ray: pleural effusion (< 18) | Yes/No | From June 8 |
| X_ray_consolidation | X-ray: alveolar consolidation (< 18) | Yes/No | From June 8 |
| X_ray_other | X-ray: other image | Yes/No | From June 8 |
| CTscan_not_performed | CT-scan: not perfomed (< 18) | Yes/No | From June 8 |

[8] Coordinates of the data providers are used only to recontact in case of question or communication related to the data but are not analyzed.
[9] Includes pneumonia observed at the CT-scan and/or at the thoracic X-ray.

| | | | |
|---|---|---|---|
| CT_scan_normal | CT-scan: normal (< 18) | Yes/No | From June 8 |
| CT_scan_ground_glass | CT-scan: Ground glass (< 18) | Yes/No | From June 8 |
| CT_scan_crazy_paving | CT-scan: crazy paving (< 18) | Yes/No | From June 8 |
| CT_scan_micronodule | CT-scan: micronodules (< 18) | Yes/No | From June 8 |
| CT_scan_pleural | CT-scan: pleural effusion (< 18) | Yes/No | From June 8 |
| CT_scan_consolidation | CT-scan: consolidation (< 18) | Yes/No | From June 8 |
| CT_scan_bronchiectasis | CT-scan:bronchiectasis (< 18) | Yes/No | From June 8 |
| CT_scan_other | CT-scan: other image (< 18) | Yes/No | From June 8 |
| Oxygen | Did the patient receive oxygen during his hospital stay? | Yes/No/Unknown | From June 8 |
| ICU_transfer | Was the patient transferred to ICU | Yes/No/Unknown | |
| Complic_none | Complication: None | Yes/No | From June 8 |
| ARDS | Complication: ARDS | Yes/No/Unknown | From June 8 |
| Complic_MOF | Complication: MOF | Yes/No | From June 8 |
| Complic_myocarditis | Complication: myocarditis | Yes/No | From June 8, up to Sept. 14 |
| Complic_sepsis | Complication: sepsis | Yes/No | From June 8 |
| Complic_bacterial_super | Complication bacterial superinfection | Yes/No | From June 8, up to Sept. 14 |
| Complic_bacterial_pulmonary | Pulmonary bacterial complication | Yes/No | From Sept. 14 |
| Complic_bacteremia | Complication bacteremia | Yes/No | From Sept. 14 |
| Complic_fungal | Complication: fungal infection | Yes/No | From June 8 |
| Complic_kidney | Complication: kidney failure | Yes/No | From June 8 |

| | | | |
|---|---|---|---|
| Complic_delirium | Complication: delirium | Yes/No | From June 8, up to Sept. 14 |
| Complic_conscious | Complication: altered consciousness | Yes/No | From June 8 |
| Complic_shock | Complication: shock | Yes/No | From June 8 |
| Complic_embol | Complication: pulmonary embolism | Yes/No | From June 8 |
| Complic_ACS | Complication: acute coronary syndrome | Yes/No | From June 8 |
| Complic_stroke | Complication: stroke | Yes/No | From June 8 |
| Complic_hep_hypoxia | Complication hepatic hypoxia | Yes/No | From June 8 |
| Complic_other | Complication: other | Text-field | From June 8 |
| Superinfection | Bacterial and/or fungal surinfection | Yes/No/Unknown | Up to June 8 |
| Kawasaki | Kawasaki/SIRS (only < 18) | Yes/No/Unknown | From June 8 |
| Kawasaki_immuno | Treatment for Kawasaki: Immunoglobulines (only < 18) | Yes/No/Unknown | From June 8 |
| Kawasaki_cortic | Treatment for Kawasaki: corticoides (only < 18) | Yes/No/Unknown | From June 8 |
| Kawasaki_AAS | Treatment for Kawasaki: AAS (only < 18) | Yes/No/Unknown | From June 8 |
| Value_PaO2_admission | Value of PaO2 at hospital admission | Numeric | |
| Value_lympho_admission | Value of Ph at hospital admission | Numeric | |
| Value_LDH_admission | Value of LDH at hospital admission | Numeric | |
| Value_CRP_admission | Value of CRP at hospital admission | Numeric | |
| Value_PCO2_admission | Value of PCO2 at hospital admission | Numeric | Up to June 8 |
| Value_Ph_admission | Value of Ph at hospital admission | Numeric | Up to June 8 |
| Value_lactate_admission | Value of lactate at hospital admission | Numeric | Up to June 8 |

| | | | |
|---|---|---|---|
| Creatinine_admission_mg | Value of creatinine in mg/dl at hospital admission | Numeric | Up to June 8 |
| Creatinine_admission_mol | Value of creatinine in µmol/l at hospital admission | Numeric | Up to June 8 |
| Value_PaO2_USI | Value of PaO2 at ICU admission | Numeric | Up to June 8 |
| Value_PCO2_USI | Value of PCO2 at ICU admission | Numeric | Up to June 8 |
| Value_Ph_USI | Value of Ph at ICU admission | Numeric | Up to June 8 |
| Creatinin_admission_USI_mg | Value of creatinine in mg/dl at ICU admission | Numeric | Up to June 8 |
| Creatinin_admission_USI_mol | Value of creatinine in µmol/l at ICU admission | Numeric | Up to June 8 |
| Value_lympho_USI | Value of Lymphocytes at ICU admission | Numeric | Up to June 8 |
| Value_LDH_USI | Value of LDH at ICU admission | Numeric | Up to June 8 |
| Serology_COVID | Serology was performed for the patient? | Yes/No/Unknown | From June 8 |
| Dt_serology | If performed, date of serology | Date | From June 8, up to Sept. 14 |
| Serology_result | Result of serology | Positive/Negative/Undetermined/Unknown | From June 8 |
| Hemoc | Hemoculture (only < 18) | Positive/negative/not performed/unknown | From June 8 |
| Hemoc_germ | if hemoc="positive", which germ was identified (only < 18) | Text field | From June 8 |
| Respi_virus | Coinfection with respiratory virus? | No/Yes/Not tested/Unknown | From June 8 |
| Virus_type | If yes, which virus(es) | Text field | From June 8 |

| | | | |
|---|---|---|---|
| Treatment1 | Specific COVID-19 treatment received? | Yes/No/Unknown | Up to June 8 |
| Name_Drug1 | Drug name treatment 1 | Text | Up to June 8 |
| Dt_init_drug1 | Start date treatment 1 | Date | Up to June 8 |
| Dt_stop_drug1 | Stop date treatment 1 | Date | Up to June 8 |
| Drug1_administration | Administration mode drug treatment 1 | Per os/IV/Unknown/Other | Up to June 8 |
| Drug1_admin_other | Specification of other administration mode of drug treatment 1, if 'other' | Text | Up to June 8 |
| Treatment2 | Specific COVID-19 treatment received? | Yes/No/Unknown | Up to June 8 |
| Name_Drug2 | Drug name treatment 2 | Text | Up to June 8 |
| Dt_init_drug2 | Start date treatment 2 | Date | Up to June 8 |
| Dt_stop_drug2 | Stop date treatment 2 | Date | Up to June 8 |
| Drug2_administration | Administration mode drug treatment 2 | Per os/IV/Unknown/Other | Up to June 8 |
| Drug2_admin_other | Specification of other administration mode of drug treatment 2, if 'other' | Text | Up to June 8 |
| Treatment3 | Specific COVID-19 treatment received? | Yes/No/Unknown | Up to June 8 |
| Name_Drug3 | Drug name treatment 3 | Text | Up to June 8 |
| Dt_init_drug3 | Start date treatment 3 | Date | Up to June 8 |
| Dt_stop_drug3 | Stop date treatment 3 | Date | Up to June 8 |
| Drug3_administration | Administration mode drug treatment 3 | Per os/IV/Unknown/Other | Up to June 8 |
| Drug3_admin_other | Specification of other administration mode of drug treatment 3, if 'other' | Text | Up to June 8 |
| HCQ | Hydoxychloroquine received? | Yes/No/Unknown | From June 8 |
| Dt_init_HCQ | Start date HCQ | Date | From June 8, up to Sept. 14 |

| | | | |
|---|---|---|---|
| Dt_stop_HCQ | Stop date HCQ | Date | From June 8, up to Sept. 14 |
| Remdesivir | Remdesivir received? | Yes/No/Unknown | From June 8 |
| Dt_init_Remdesivir | Start date remdesivir | Date | From June 8, up to Sept. 14 |
| Dt_stop_remdesivir | Stop date remdesivir | Date | From June 8, up to Sept. 14 |
| Lopi_ritonavir | Lopinavir/Ritonavir received? | Yes/No/Unknown | From June 8, up to Sept. 14 |
| Dt_init_Lopi_rito | Start date Lopinavir/Ritonavir | Date | From June 8, up to Sept. 14 |
| Dt_stop_Lopi_rito | Stop date Lopinavir/Ritonavir | Date | From June 8, up to Sept. 14 |
| Olsetamivir | Olsetamivir received? | Yes/No/Unknown | From June 8, up to Sept. 14 |
| Dt_init_Olsetamivir | Start date Olsetamivir | Date | From June 8, up to Sept. 14 |
| Dt_stop_Olsetamivir | Stop date Olsetamivir | Date | From June 8, up to Sept. 14 |
| Macrolides | Macrolides received? | Yes/No/Unknown | From June 8 |
| Dt_init_Macrolides | Start date Macrolides | Date | From June 8, up to Sept. 14 |
| Dt_stop_Macrolides | Stop date Macrolides | Date | From June 8, up to Sept. 14 |

| | | | |
|---|---|---|---|
| Tolicizumab | Tolicizumab received? | Yes/No/Unknown | From June 8, up to Sept. 14 |
| Dt_init_Tolicizumab | Start date Tolicizumab | Date | From June 8, up to Sept. 14 |
| Dt_stop_Tolicizumab | Stop date Tolicizumab | Date | From June 8, up to Sept. 14 |
| Other_treatment | Other medication specific for COVID-19 received? | Yes/No/Unknown | From June 8 |
| Name_other_treat | Name other treatment | Text | From June 8 |
| Dt_init_other_treat | Start date other treatment | Date | From June 8, up to Sept. 14 |
| Dt_stop_other_treat | Stop date other treatment | Date | From June 8, up to Sept. 14 |
| Clinical_trial | Is the patient part of a clinical trial | Yes/No/Unknown | From June 8, up to Sept. 14 |
| AB1 | Antibiotic received during hospitalization? (only < 18) | Yes/No/Unknown | From June 8 |
| AB1_indication | Indication AB 1 (only < 18) | Prophylaxis/Sepsis/Infection KTC /Other | From June 8 |
| AB2 | Second antibiotic received during hospitalization? (only < 18) | Yes/No/Unknown | From June 8 |
| AB2_indication | Indication AB 2 (only < 18) | Prophylaxis/Sepsis/Infection KTC /Other | From June 8 |
| Corticoid_treatment | Systemic corticoid treatment? | Yes/No/Unknown | From June 8 |

| | | | |
|---|---|---|---|
| LMWH_prevention | Prevention with low molecular weight heparin | Yes/No/Unknown | From June 8 |
| LMWH_dosis | Dosis of LMWH | Standard/High/Unknown | From June 8 |
| Parent_nutrition | Nutrition parentérale (only < 18) | Yes/No/Unknown | From June 8 |
| Status_discharge | Health status at discharge | Cured/Deceased/Transferred (hospital)/Other[10]/Unknown | |
| Death_covid | If dead, due to covid | Yes/No/Unknown | |
| Death_nocovid_cause | If dead not due to covid, cause | Text | |
| Status_discharge_other | Specification of health status at discharge, if 'other' | Text | |
| Transfer_out_hospit_name | Patient transferred to which hospital? | Dropdown list | |
| Transfer_out_hospname_other | Name of 'other' hospital the patient was transferred to | Text | |
| Dt_discharge | Date of discharge, transfer or death | Date | |
| Dt_last_test | Date of last COVID-19 test before discharge, if done | Date | Up to Sept. 14 |
| Result_last_test | Result of last COVID-19 test | Positive/Negative/Unknown | Up to Sept. 14 |

[10] When patients are sufficiently well to be discharged from the hospital, but need to continue their recovery/convalescence at home or at another place outside the hospital.

| | | | |
|---|---|---|---|
| Name_disch | Name of the person reporting the discharge[8] | Text | |
| Phone_disch | Phone number of the person reporting the discharge[8] | Text | |
| Email_disch | Email address of the person reporting the discharge[8] | Text | |
| **ICU questionnaire** | | | |
| Dt_reporting_ICU | Declaration date | Date | |
| Hospital_name_ICU | Hospital name | Dropdown list | |
| Dt_ICU_transfer | Date of transfer to ICU | Date | |
| Length_stay_ICU | Length of stay at ICU | Numeric | Up to June 8 |
| QuickSOFA_score | QuickSOFA score at admission in ICU | Numeric | Up to June 8 |
| SOFA_score | SOFA (or pedSOFA > 18 years) score at admission in ICU: TOTAL - if sedated use for neurological SOFA the last available GCS on admission or before intubation | Numeric (total) | |
| ICU_adm_reason | Reason for ICU admission | Dropdown list | From Sept. 14 |
| ICU_adm_reason_other | Reason for ICU admission, if other | Text | From Sept. 14 |
| Non_inv_vent_support | Non-invasive ventilatory support | Yes/No/Unknown | |
| Length_non_inv_vent | Number of days on non-invasive ventilatory support | Numeric (days) | From Sept. 14 |
| Non_inv_vent_extub | Non-invasive ventilation only after extubation and weaning from invasive ventilatory support | Yes/No/Unknown | From Sept. 14 |

| | | | |
|---|---|---|---|
| Inv_Vent_support | Invasive ventilatory support | Yes/No/Unknown | |
| Length_inv_vent | Number of days on invasive ventilatory support | Numeric (days) | From Sept. 14 |
| Weaning_inv_vent | Weaning of IVS | Yes (successfully weaned)/No(dead whilst on IVS)/Unknown | From Sept. 14 |
| ECMO | Oxigenation by ECMO | Yes/No/Unknown | |
| Length_ECMO | Number of days on ECMO/ECCO2R | Numeric (days) | From Sept. 14 |
| Weaning_ECMO | Weaning of ECMO/ECCO2R | Yes (successfully weaned)/No(dead whilst on IVS)/Unknown | From Sept. 14 |
| Vasopressor | Vasopressor received? | Yes/No/Unknown | From Sept. 14 |
| Inotropic | Inotropic agent received? | Yes/No/Unknown | From Sept. 14 |
| Prone_posit | Prone positioning? | Yes/No/Unknown | From Sept. 14 |
| Renal_replace | Renal replacement therapy? | Yes/No/Unknown | From Sept. 14 |
| Tracheostomy | Tracheostomy? | Yes/No/Unknown | From Sept. 14 |
| | | | |
| Value_lactate_USI | Value of lactate at ICU admission | Numeric | |
| Value_PaO2_FiO2_ratio_ICU | Lowest PaO2 / FiO2 on ICU admission (<=48h) | Numeric | From Sept. 14 |
| Value_CRP_USI | CRP value known at ICU admission? | Numeric | |
| Value_Ddimer_ICU | Value of ddimeres on ICU admission | Numeric | From Sept. 14 |
| Value_ferritin_ICU | Value of ferritin on ICU admission | Numeric | From Sept. 14 |
| ICU_outcome | Patient was weaned from mechanical ventilation | Alive/Dead | |
| Dt_ICU_discharge | Patient was weaned from ECMO | Date | |

| | | | |
|---|---|---|---|
| ICU_dead_cause | If dead during ICU stay, reason | Dropdown list | From Sept. 14 |
| ICU_dead_other | Other reason for ICU death | Dropdown list | From Sept. 14 |
| Email_ICU | Email address of the person reporting the ICU[8] | Text | |

# Appendix B.

**Description of the substances, Anatomical Therapeutic Chemical (ATC) subcodes and ATC codes used to build the pseudopathologies available in the Intermutualistic Agency (IMA) database.**

| Intermutualistic Agency (IMA) codes | Pseudopathologies available in IMA | Codes ATC | System | Sub-codes ATC | Corresponding drugs |
|---|---|---|---|---|---|
| FL_PSEUDOPATH_0101 | Cardiovascular diseases (general) | C | Cardiovascular system | C01 | Cardiac therapy |
| | | | | C02 | Antihypertensives |
| | | | | C03 | Diuretics |
| | | | | C04 | Peripheral vasodilatators |
| | | | | C05 | Vasoprotectives |
| | | | | C06 | Beta blocking agents |
| | | | | C07 | Calcium channel blockers |
| | | | | C08 | Agents acting on the renin-angiotensin system |
| | | | | C09 | Lipid modifying agents |
| FL_PSEUDOPATH_01A01 | Thrombosis | B | Blood and blood forming organs | B01A | Anthithrombotic agents |
| FL_PSEUDOPATH_0201 | Cardiovascular diseases | C | Cardiovascular system | C01 | Cardiac therapy |

| | | | | | |
|---|---|---|---|---|---|
| FL_PSEUDOPATH_0301 | Chronic obstructive pulmonary disease | R | Respiratory system | R03BB | Anticholinergics |
| | | | | R03DA04 | Theophylline |
| | | | | R03A (if age > 50) | Adrenergics, inhalants |
| | | | | R03BA (if age > 50) | Glucocorticoids |
| FL_PSEUDOPATH_2201 | Renal failure | A | Alimentary tract and metabolism | A11CC03 | Alfracalcidol |
| | | | | A11CC04 | Calcitriol |
| | | | | A11CC06 | Calcifediol |
| | | | | A12AA12 | Calcium |
| | | V | Various | V03AE02 | Sevelamer |
| | | | | V03AE03 | Lanthanum carbonate |
| | | | | V03AE04 | Calcium acetate and magnesium |
| | | | | V03AE01 | Polystyrene sulfonate |
| | | H | Systemic hormonal preparation (exlusion of sexual hormones and insulins) | H05BX01 | Cinacalcet |
| FL_PSEUDOPATH_1801 | Chronic hepatitis B and C | L | Antineoplastic and immunomodulating agents | L03AB04 | Inteferon alfa-2a |
| | | | | L03AB05 | Interferon alfa-2b |
| | | | | L03AB09 | Interferon alfacon-1 |
| | | | | L03AB10 | Peginterferon alfa-2b |
| | | | | L03AB11 | Albinterferon alfa-2b |

| | | | | | |
|---|---|---|---|---|---|
| | | J | Antiinfectives for systemic use | J05AF08 | Adefovir dipivoxil |
| | | | | J05AF10 | Entecavir |
| | | | | J05AE11 | Protease inhibitors |
| | | | | J05AE12 | Protease inhibitors |
| | | | | J05AE14 | Protease inhibitors |
| | | | | J05AX15 | Other antivirals |
| | | | | J05AX65 | Other antivirals |
| | | | | J05AB04 | Nucleosides and nucleotides (exclusion of reverse transcriptase inhibitors) |
| | | | | J05AF05 | Lamivudine |
| FL_PSEUDOPATH_0401 | Asthme | R | Respiratory system | R03DC01 | Zafirlukast |
| | | | | R03DC03 | Montelukast |
| | | | | R03DX05 | Omalizumab |
| | | | | R03A (if age ≤ 50) | Adrenergics, inhalants |
| | | | | R03BA (if age ≤ 50) | Glucocorticoids |
| FL_CANCER_MOC_YN | Based on reimbursement consultation oncologist over the year of reference | / | / | / | / |

| | | | | | |
|---|---|---|---|---|---|
| FL_CANCER_CHEMORT_YN | Based on reimbursed chemo- or radiotherapy | L | Antineoplastic and immunodulating agents | L01 | Antineoplastic agents |
| FL_PSEUDOPATH_2001 | Organ transplant | J | Antiinfectives for systematic use | J05AF05 | Lamivudine |
| | | | | J05AG | Non-nucleoside reverse transcriptase inhibitors |
| | | | | J05AE | Protease inhibitors |
| | | | | J05AR | Antivirals for treatment of HCV infections |
| | | | | J05AF | Nucleoside, and nucleotide reverse transcriptase inhibitors |
| | | | | J05AX | Other antivirals |
| FL_PSEUDOPATH_1701 | HIV | J | Antiinfectives for systematic use | J05AF05 | Lamivudine |
| | | | | J05AG | Non-nucleoside reverse transcriptase inhibitors |
| | | | | J05AR | Antivirals for treatment of HIV infections, combinations |
| | | | | J05AE | Protease inhibitors |
| FL_PSEUDOPATH_1501 | Parkinson | N | Nervous system | N04AB | Ethers chemically close to antihistamines |
| | | | | N04AC | Ethers of tropine or tropine derivatives |

| | | | | | |
|---|---|---|---|---|---|
| | | | | N04B | Dopaminergic agents |
| FL_PSEUDOPATH_1601 | Epilepsy and neurotic pain | N | Nervous system | N03 | Antiepileptics |
| FL_PSEUDOPATH_1901 | Multiple sclerosis | L | Antineoplastic and immunomodulating agents | L03AB07 | Interferon beta-1a |
| | | | | L03AB08 | Interferon beta-1b |
| | | | | L03AX13 | Glatiramer acetate |
| | | | | L05AA | Colony stimulating factors |
| | | N | Nervous system | N07XX09 | Other nervous system drugs |
| FL_PSEUDOPATH_2101 | Alzheimer | N | Nervous system | N06DX01 | Memantine |
| | | | | N06DA | Anticholinesterases |
| FL_PSEUDOPATH_1301 | Psychosis (< 70 years) | N | Nervous system | N05AA | Phenothiazines with aliphatic side-chain |
| | | | | N05AB | Phenothiazines with piperazine structure |
| | | | | N05AC | Phenothiazines with piperidine structure |
| | | | | N05AD | Butyrophenone derivatives |
| | | | | N05AE | Indole derivatives |
| | | | | N05AF | Thioxanthene derivatives |

| | | | | | |
|---|---|---|---|---|---|
| | | | | N05AG | Diphenylbutylpiperidine derivatives |
| | | | | N05AH | Diazpines, oxazepines, thiazepines and oxepines |
| | | | | N05AN | Lithium |
| | | | | N05AX | Other antipsychotics |
| | | | | N07XX06 | Tetrabenazine |
| FL_PSEUDOPATH_1401 | Psychosis (≥ 70 years) | N | Nervous system | N05AA | Phenothiazines with aliphatic side-chain |
| | | | | N05AB | Phenothiazines with piperazine structure |
| | | | | N05AC | Phenothiazines with piperidine structure |
| | | | | N05AD | Butyrophenone derivatives |
| | | | | N05AE | Indole derivatives |
| | | | | N05AF | Thioxanthene derivatives |
| | | | | N05AG | Diphenylbutylpiperidine derivatives |
| | | | | N05AH | Diazpines, oxazepines, thiazepines and oxepines |

| | | | | | |
|---|---|---|---|---|---|
| | | | | N05AN | Lithium |
| | | | | N05AX | Other antipsychotics |
| | | | | N07XX06 | Tetrabenazine |
| FL_PSEUDOPATH_0601 | Diabetes | A | Alimentary tract and metabolism | A10A | Insulins and analogues |
| | | | | A10B | Blood glucose lowering drugs (exclusion of insulins) |
| FL_PSEUDOPATH_0801 | Diabetes treated with insulin | A | Alimentary tract and metabolism | A10A | Insulins and analogues |
| | | | | A10B | Blood glucose lowering drugs (exclusion of insulins) |
| FL_PSEUDOPATH_0901 | Diabetes treated without insulin | A | Alimentary tract and metabolism | A10B | Blood glucose lowering drugs (exclusion of insulins) |
| FL_PSEUDOPATH_0501 | Mucovisidosis | R | Respiratory system | R05CB13 | Dornase alfa (desoxyribonuclease) |
| | | | | R07AX02 | Ivacaftor |
| | | A | Alimentary tract and metabolism | A09AA02 | Multienzymes (lipase, protease, etc.) |
| FL_PSEUDOPATH_1201 | Rhumatoid arthritis, Crohn's disases, ulcerative | L | Antineoplastic and immunomodulating agents | L04AA | Selective immunosupressants |
| | | | | L04AB01 | Etanercept |
| | | | | L04AB02 | Infliximab |

| | | | | | |
|---|---|---|---|---|---|
| | colitis, arthiritis psoriasis | | | L04AA24 | Abatacept |
| | | | | L04AB04 | Adalimumab |
| | | | | L04AB05 | Certolizumab pegol |
| | | | | L04AB06 | Golimumab |
| | | | | L04AB07 | Opinercept |
| | | A | Alimentary tract and metabolism | A07EC01 | Sulfasalazine |
| | | | | A07EC02 | Mesalazine |

# Appendix C.

**R code to reproduce the working example computing the density distribution of propensity scores, the relative contributions of all individuals in the overlap weighted pseudo-population and inverse probability of treatment weighted pseudo-population, along with the standardized mean difference by balancing weight method.**

```
# Load packages

library(survey)
library(tableone)
library(scales)
library(dplyr)
library(ggplot2)

# Simulated data based on the distribution of the variables
used in the study by Mehta et al. (2020)

set.seed(123)

simulate_exact_bin <- function(n, prop) {
  k <- round(n * prop)
  vec <- c(rep(1, k), rep(0, n - k))
  sample(vec)
}

generate_exact_group <- function(n, Age_mean, Age_sd,
                                 Male_prop, hta_prop, diab_p
rop,
                                 cad_prop, hf_prop, copd_pro
p,
                                 acei_value) {
  Age_raw <- rnorm(n, mean = Age_mean, sd = Age_sd)
  Age_corrected <- Age_raw + (Age_mean - mean(Age_raw))

  data.frame(
    Age = round(Age_corrected),
    Male = simulate_exact_bin(n, Male_prop),
    HTA = simulate_exact_bin(n, hta_prop),
```

```
    Diabetes = simulate_exact_bin(n, diab_prop),
    CAD = simulate_exact_bin(n, cad_prop),
    HF = simulate_exact_bin(n, hf_prop),
    COPD = simulate_exact_bin(n, copd_prop),
    ACEI = acei_value
  )
}

group_acei <- generate_exact_group(
  n = 7,
  Age_mean = 62, Age_sd = 16,
  Male_prop = 0.53,
  hta_prop = 0.94,
  diab_prop = 0.48,
  cad_prop = 0.29,
  hf_prop = 0.25,
  copd_prop = 0.23,
  acei_value = 1
)

group_non_acei <- generate_exact_group(
  n = 93,
  Age_mean = 48, Age_sd = 21,
  Male_prop = 0.39,
  hta_prop = 0.36,
  diab_prop = 0.17,
  cad_prop = 0.11,
  hf_prop = 0.09,
  copd_prop = 0.11,
  acei_value = 0
)

sim_data <- rbind(group_acei, group_non_acei)

# Compute propensity score model

ps_model <- glm(ACEI ~ Age + Male + HTA + Diabetes + CAD + H
F + COPD,
                data = sim_data, family = binomial())
sim_data$propensity_score <- predict(ps_model, type = "respo
nse")
```

```
# Plot PS distribution to visualize extreme PS

ggplot(sim_data, aes(x = propensity_score, fill = factor(ACE
I))) +
  geom_density(alpha = 0.5) +
  labs(title = "",
       x = "Propensity score",
       y="Density",
       fill = "") +
  scale_fill_manual(values = c("blue", "red"), labels = c("N
o ACEI", "ACEI")) +
  theme_minimal()+
  theme(
    panel.grid.major.x = element_line(color = "grey92"),
    panel.grid.major.y = element_line(color = "grey92"),
    panel.grid.minor = element_blank(),
    legend.title = element_blank()
  )

# Compute IPTW and OW

sim_data$ipw <- ifelse(sim_data$ACEI == 1, 1 / sim_data$prop
ensity_score, 1 / (1 - sim_data$propensity_score))
sim_data$ow <- ifelse(sim_data$ACEI == 1, 1 - sim_data$prope
nsity_score, sim_data$propensity_score)

# Normalize OW

sim_data <- sim_data %>%
  group_by(ACEI) %>%
  mutate(ow_norm = ow / sum(ow)) %>%
  ungroup()

# Comparison of the distribution of diabetes according to ag
e in weighted pseudo-populations and unweighted baseline pop
ulation

sim_data$diabetes_label <- factor(sim_data$Diabetes, levels
= c(0, 1),
                                  labels = c("No diabetes",
"Diabetes"))
```

```
## Unweighted population

ggplot(sim_data, aes(x = Age, y = diabetes_label, color = fa
ctor(ACEI))) +
  geom_jitter(height = 0.2, width = 0,size=3) +
  scale_color_manual(values = c("slateblue1", "orangered"),
labels = c("No ACEI", "ACEI")) +
  labs(
    x = "Age",
    y = "",
    color = ""
  ) +
  theme_minimal() +
  theme_classic()+
  theme(
    axis.title = element_text(color = "black",size=11),
    axis.text = element_text(color = "black",size=11),
    panel.grid.major.y = element_line(color = "grey80"),
    panel.grid.major.x = element_blank(),
    panel.grid.minor = element_blank(),
    legend.position = "bottom"
  )

## OW pseudo-population

ggplot(sim_data, aes(x = Age, y = diabetes_label, color = fa
ctor(ACEI), size = ow_norm)) +
  geom_jitter(height = 0.2, width = 0) +
  scale_color_manual(values = c("slateblue1", "orangered"),
labels = c("No ACEI", "ACEI")) +
  scale_size_continuous(name = "Weighting of each patient")
+
  labs(
    x = "Age",
    color = "",
    y=""
  ) +
  theme_minimal() +
  theme_classic()+
  theme(
    axis.title = element_text(color = "black",size=11),
    axis.text = element_text(color = "black",size=11),
    panel.grid.major.y = element_line(color = "grey80"),
```

```
    panel.grid.major.x = element_blank(),
    panel.grid.minor = element_blank(),
    legend.position = "bottom"
  )

## IPW pseudo-population

ggplot(sim_data, aes(x = Age, y = diabetes_label, color = fa
ctor(ACEI), size = ipw)) +
  geom_jitter(height = 0.2, width = 0,) +
  scale_color_manual(values = c("slateblue1", "orangered"),
labels = c("No ACEI", "ACEI")) +
  scale_size_continuous(name = "Weighting of each patient")
+
  labs(
    x = "Age",
    y = "",
    color = ""
  ) +
  theme_classic()+
  theme(
    axis.title = element_text(color = "black",size=11),
    axis.text = element_text(color = "black",size=11),
    panel.grid.major.y = element_line(color = "grey80"),
    panel.grid.major.x = element_blank(),
    panel.grid.minor = element_blank(),
    legend.position = "bottom"
  )

# Compute SMD across treatment groups

vars <- c("Age","Male","HTA","Diabetes","CAD","HF","COPD")

## SMD without weights
table_unadj <- CreateTableOne(vars = vars, strata = "ACEI",
data = sim_data, test = FALSE)
smd_unadj <- ExtractSmd(table_unadj)

## SMD without IPTW
design_ipw <- svydesign(ids = ~1, weights = ~ipw, data = sim
_data)
table_ipw <- svyCreateTableOne(vars = vars, strata = "ACEI",
data = design_ipw, test = FALSE)
```

```
smd_ipw <- ExtractSmd(table_ipw)

## SMD with OW
design_ow <- svydesign(ids = ~1, weights = ~ow_norm, data =
sim_data)
table_ow <- svyCreateTableOne(vars = vars, strata = "ACEI",
data = design_ow, test = FALSE)
smd_ow <- ExtractSmd(table_ow)

smd_df <- data.frame(
  Variable = rep(vars, 3),
  SMD = c(abs(smd_unadj), abs(smd_ipw), abs(smd_ow)),
  Method = rep(c("Unweighted", "IPW", "OW"), each = length(v
ars))
)

ggplot(smd_df, aes(x = Variable, y = SMD, color = Method)) +
  geom_point(position = position_dodge(width = 0.6), size =
3) +
  geom_hline(yintercept = 0.1, linetype = "dashed", color =
"red") + # seuil 0.1
  labs(title = "",
       y = "Absolute SMD",
       x = "") +
  scale_y_continuous(breaks = seq(0, 2, by = 0.5)) +
  theme_classic()+
  scale_color_manual(values = c("deepskyblue", "deeppink", "
orange1")) +
  coord_flip()+
  theme(
    axis.title = element_text(color = "black",size=11),
    axis.text = element_text(color = "black",size=11),
    panel.grid.major.x = element_line(color = "grey80"),
    panel.grid.major.y = element_blank(),
    panel.grid.minor = element_blank(),
    legend.title = element_blank()
  )
```

# Appendix D.

**R code to reproduce the working example implementing a conditional inference tree algorithm and assessing its predictive performance.**

```
# Install and load the necessary packages
library(party)

library(partykit) # To compute Conditonal Inference Trees

library(binom) # To compute 95% CI for sensitivity and speci
ficity

library(pROC) # To compute ROC curve (+95% CI) and AUC

# Fix the seed for reproducibility
set.seed(123)

# Number of observations
n <- 1000

# Simulate the data: 1000 individuals with binary predictors
'education' and 'income'
Education <- factor(sample(c("low", "high"), n, replace = TR
UE))
Income <- factor(sample(c("low", "high"), n, replace = TRUE)
)

# Define a probability for being diseases based on the predi
ctor values
prob <- ifelse(Education == "low" & Income == "high", 0.7,
               ifelse(Education == "low" | Income == "high",
0.4, 0.2))

# Generate health status using defined probability
status <- factor(ifelse(runif(n) < prob, "Diseased", "Health
y"))
```

```
# Cenerate dataset
data <- data.frame(Education, Income, status)

# Create the training (70%) and validation (30%) datasets at
random
train_indices <- sample(1:n, size = 0.7 * n)
train_data <- data[train_indices, ]
test_data <- data[-train_indices, ]

### Compute the conditional inference tree on the training d
ataset
ctree <- ctree(status ~ Education + Income, data = train_dat
a)
plot(ctree)

### Make predictions on the test dataset
predictions <- predict(ctree, newdata = test_data)

### Convert predictions and actual outcomes to a confusion m
atrix
confusion_matrix <- table(predictions,test_data$status)
print(confusion_matrix)

### Assess predictive performances


# Calculate sensitivity and specificity
sensitivity <- confusion_matrix[2, 2] / sum(confusion_matrix
[2, ])
specificity <- confusion_matrix[1, 1] / sum(confusion_matrix
[1, ])

# Print the results
print(paste("Sensitivity:", sensitivity))
print(paste("Specificity:", specificity))
```

```
# Calculate confidence intervals for sensitivity and specifi
city
sensitivity_conf_int <- binom.confint(confusion_matrix[2, 2]
, sum(confusion_matrix[2, ]), methods = "wilson")
specificity_conf_int <- binom.confint(confusion_matrix[1, 1]
, sum(confusion_matrix[1, ]), methods = "wilson")

# Print the results with confidence intervals
print(paste("Sensitivity:", round(sensitivity,3),
            ", 95% CI:", round(sensitivity_conf_int[1,5], 3)
, "-", round(sensitivity_conf_int[1,6], 3)))
print(paste("Specificity:", round(specificity,3),
            ", 95% CI:", round(specificity_conf_int[1,5], 3)
, "-", round(specificity_conf_int[1,6], 3)))

# Compute ROC and AUC

predictions <- as.data.frame(predict(ctree, newdata = test_d
ata, type = "p"))

# Create predicted class based on the threshold of 0.5
predict_ctree <- ifelse(predictions$Diseased > 0.5, 1, 0)

roc_ctree<-roc(test_data$status,predictions$Diseased,
               levels = c("Healthy", "Diseased"),
               smoothed = TRUE,
               # arguments for ci
               ci=TRUE, ci.alpha=0.95, stratified=FALSE,
               # arguments for plot
               plot=TRUE, max.auc.polygon=T, grid=T,
               print.auc=TRUE, show.thres=TRUE)
ci <- ci.se(roc_ctree)
plot(ci,type="shape",col="lightcyan")
```

# Appendix E.

**R code to reproduce the working example illustrating the differences in estimating the incidence of COVID-19 mortality using the Kaplan-Meier vs. competing risk method, in the presence of competing risks.**

```
library("survival")

library("cmprsk")

# Data
y <- c(1,2,3,4,5,6,7,8,9,10) # follow-up
d <- c(1,0,2,1,1,0,2,2,1,0) # 1= event of interest, 2=compet
ing event,0=censored

# Kaplan-Meier (KM) : we considered 2 as censored
d_km <- ifelse(d == 1, 1, 0)

#Fit KM estimator
km_fit <- survfit(Surv(y, d_km) ~ 1)

# Competing risks : estimation CIF for event k=1
cif <- cuminc(y, d, cencode = 0)
time_cif <- cif[["1 1"]]$time
est_cif <- cif[["1 1"]]$est

# Compare KM estimator and competing risk estimator on a gra
ph
plot(km_fit$time, 1 - km_fit$surv, type = "s", col = "orange
", lwd = 2,
     xlab = "Follow-up", ylab = "Cumulative probability",
     main = "",
     ylim = c(0, 1))
lines(time_cif, est_cif, type = "s", col = "red", lwd = 2)
legend("bottomright",
       legend = c("Kaplan-Meier", "Competing risk"),
       col = c("orange", "red"),
       lwd = 2)
grid()
```

# Appendix F.

**R code to reproduce the working example illustrating the calculation of the cumulative incidence function by subgroup and the use of Gray's test to evaluate statistical differences between groups.**

```
# Load packages

library(cmprsk)

library(ggplot2)

library(ggsurvfit)

library(tidycmprsk)

# Simulation of the dataset
set.seed(123)
N <- 1000
education <- sample(c("Low education", "High education"), N,
replace = TRUE)
prob_event <- c(covid = 0.2, other = 0.15, censored = 0.65)
event <- rep(0, N)
event[sample(1:N, size = N * prob_event["covid"])] <- 1
remaining <- which(event == 0)
event[sample(remaining, size = N * prob_event["other"])] <-
2
died_idx <- which(event %in% c(1, 2))
education[died_idx] <- sample(c("Low education", "High educa
tion"), length(died_idx), replace = TRUE, prob = c(0.8, 0.2)
)
censored_idx <- which(event == 0)
education[censored_idx] <- sample(c("Low education", "High e
ducation"), length(censored_idx), replace = TRUE, prob = c(0
.4, 0.6))
time <- rexp(N, rate = 0.2)
time <- pmin(time, 10)  # Tronquer à 10 mois
time <- round(time)
data <- data.frame(
  id = 1:N,
  time = time,
```

```
  status = event,
  education = factor(education, levels = c("Low education",
"High education"))
)

# Convert education in factor and define a refence level
data$education<-as.factor(data$education)
data$education<-relevel(data$education,ref = "Low education"
)

# Compute competing risk cumulative incidence
ci <- cuminc(Surv(time, as.factor(status)) ~ education, data
= data)

# Plot cumulative incidence of COVID-19 death according to t
he education level
ci %>%
  ggcuminc() +
  labs(
    x = "Follow-up (in months)",
    y = "Cumulative incidence of COVID-19 death"
  ) +
  scale_x_continuous(
    limits = c(1, 10),
    breaks = 1:10
  ) +
  add_confidence_interval()+
  theme(
    axis.title.x = element_text(size = 10, face = "bold"),
    axis.title.y = element_text(size = 10, face = "bold")

# Create table with 10-month cumulative incidence and obtain
the Gray test p-value
ci %>%
  tbl_cuminc(
    times = 10,  # since time is in months already
    label_header = "**{time}-month cuminc**"
  ) %>%
  add_p()
```

# References


1. Lamers MM, Haagmans BL. SARS-CoV-2 pathogenesis. Nat Rev Microbiol. 2022 May;20(5):270–84.

2. Centers for disease control and prevention (CDC). *About Middle East Respiratory Syndrome (MERS)*; 2024. Available from: https://www.cdc.gov/mers/about/index.html.

3. Centers for disease control and prevention (CDC) - *Office of Readiness and Response. Severe Acute Respiratory Syndrome (SARS)*; 2024. Available from: https://www.cdc.gov/orr/responses/sars.html.

4. 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 Feb 20;382(8):727–33.

5. Baylor College of Medicine (BCM). *Emerging Infectious Diseases*; 2023. Available from: https://www.bcm.edu/departments/molecular-virology-and-microbiology/emerging-infections-and-biodefense/emerging-infectious-diseases.

6. World Health Organization (WHO) - Regional Office for South-East Asia. *A brief guide to emerging infectious diseases and zoonoses*; 2014. Available from: https://iris.who.int/handle/10665/204722.

7. Fraser BJ, Beldar S, Seitova A, Hutchinson A, Mannar D, Li Y, et al. Structure and activity of human TMPRSS2 protease implicated in SARS-CoV-2 activation. Nat Chem Biol. 2022 Sep;18(9):963–71.

8. Jackson CB, Farzan M, Chen B, Choe H. Mechanisms of SARS-CoV-2 entry into cells. Nat Rev Mol Cell Biol. 2022 Jan;23(1):3–20.

9. Zhao MM, Yang WL, Yang FY, Zhang L, Huang WJ, Hou W, et al. Cathepsin L plays a key role in SARS-CoV-2 infection in humans and humanized mice and is a promising target for new drug development. Signal Transduct Target Ther. 2021 Mar 27;6(1):1–12.

10. Zhang H, Li HB, Lyu JR, Lei XM, Li W, Wu G, et al. Specific ACE2 expression in small intestinal enterocytes may cause gastrointestinal symptoms and injury after 2019-nCoV infection. Int J Infect Dis. 2020 Jul;96:19–24.

11. Lukassen S, Chua RL, Trefzer T, Kahn NC, Schneider MA, Muley T, et al. SARS-CoV-2 receptor ACE2 and TMPRSS2 are primarily expressed in bronchial transient secretory cells. EMBO J. 2020 May 18;39(10):e105114.

12. Pišlar A, Mitrović A, Sabotič J, Fonović UP, Nanut MP, Jakoš T, et al. The role of cysteine peptidases in coronavirus cell entry and replication: The therapeutic potential of cathepsin inhibitors. PLOS Pathog. 2020 Nov 2;16(11):e1009013.

13. World Health Organization (WHO). *Transmission of SARS-CoV-2: implications for infection prevention precautions*; 2020. Available from: https://www.who.int/news-room/commentaries/detail/transmission-of-sars-cov-2-implications-for-infection-prevention-precautions.

14. Government of Canada. *COVID-19 signs, symptoms and severity of disease: A clinician guide*; 2020. Available from: https://www.canada.ca/en/public-health/services/diseases/2019-novel-coronavirus-infection/guidance-documents/signs-symptoms-severity.html.

15. Xu Z, Li S, Tian S, Li H, Kong L quan. Full spectrum of COVID-19 severity still being depicted. The Lancet. 2020 Mar 21;395(10228):947–8.

16. Xu J, Lazartigues E. Expression of ACE2 in Human Neurons Supports the Neuro-Invasive Potential of COVID-19 Virus. Cell Mol Neurobiol. 2022;42(1):305–9.

17. Chen L, Li X, Chen M, Feng Y, Xiong C. The ACE2 expression in human heart indicates new potential mechanism of heart injury among patients infected with SARS-CoV-2. Cardiovasc Res. 2020 May 1;116(6):1097–100.

18. Ashraf UM, Abokor AA, Edwards JM, Waigi EW, Royfman RS, Hasan SAM, et al. SARS-CoV-2, ACE2 expression, and systemic organ invasion. Physiol Genomics. 2021 Feb 1;53(2):51–60.

19. Krishna BA, Metaxaki M, Sithole N, Landín P, Martín P, Salinas-Botrán A. Cardiovascular disease and covid-19: A systematic review. IJC Heart Vasc. 2024 Oct 1;54:101482.

20. Balcom EF, Nath A, Power C. Acute and chronic neurological disorders in COVID-19: potential mechanisms of disease. Brain. 2021 Aug 16;144(12):3576–88.

21. World Health Organization (WHO). *Statement on the second meeting of the International Health Regulations (2005) Emergency Committee regarding the outbreak of novel coronavirus (2019-nCoV)*; 2020. Available from: https://www.who.int/news/item/30-01-2020-statement-on-the-second-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-(2019-ncov).

22. World Health Organization (WHO). *WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020*; 2020. Available from: https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020.

23. World Health Organization (WHO). *COVID-19 cases | WHO COVID-19 dashboard*; 2024. Available from: https://data.who.int/dashboards/covid19/cases.

24. SPF Santé Publique. *Réunion des ministres de la Santé publique à propos du coronavirus*; 2020. Available from: https://www.health.belgium.be/fr/news/reunion-des-ministres-de-la-sante-publique-propos-du-coronavirus.

25. Sciensano. *Belgium COVID-19 Dashboard*; 2025. Available from: http://datastudio.google.com/reporting/c14a5cfc-cab7-4812-848c-0369173148ab/page/tpRKB?feature=opengraph.

26. Devleesschauwer B, Willem L, Jurčević J, Smith P, Scohy A, Wyper GMA, et al. The direct disease burden of COVID-19 in Belgium in 2020 and 2021. BMC Public Health. 2023 Sep 4;23(1):1707.

27. Bustos Sierra N, Bossuyt N, Braeye T, Leroy M, Moyersoen I, Peeters I, et al. All-cause mortality supports the COVID-19 mortality in Belgium and comparison with major fatal events of the last century. Arch Public Health. 2020 Nov 13;78(1):117.

28. Jurcevic J, Ekelson R, Nganda S, Bustos Sierra N, Vernemmen C. *Epidemiology of COVID-19 mortality in Belgium from wave 1 to wave 7 (March 2020 – 11 September 2022).* Brussels: Sciensano; 2023. Available from: https://www.sciensano.be/en/biblio/epidemiology-covid-19-mortality-belgium-wave-1-wave-7-march-2020-11-september-2022.

29. Sciensano. *Belgian National Burden of Disease Study (BeBOD)*. 2023; Available from: https://burden.sciensano.be/shiny/mortality/.

30. Sciensano. For a healthy Belgium. *Excess mortality during the COVID-19 crisis in 2020*; 2024. Available from: https://www.healthybelgium.be/en/health-status/factsheets/fc-excess2020.

31. Bustos Sierra N, Bossuyt N, Braeye T, Haarhuis F, Peeters I, Proesmans K, et al. Excess mortality during the first and second waves of the COVID-19 epidemic in Belgium. Brussels: Sciensano; 2021. Available from: https://www.sciensano.be/en/biblio/excess-mortality-during-first-and-second-waves-covid-19-epidemic-belgium.

32. Sciensano. *The Belgian Mortality Monitoring*; 2004. Available from: https://www.sciensano.be/fr/projets/belgian-mortality-monitoring.

33. Bustos Sierra N, Bossuyt N, Braeye T, Haarhuis F, Peeters I, Proesmans K, et al. Excess mortality during the first and second waves of the COVID-19 epidemic in Belgium. Brussels: Sciensano; 2021. Available from: https://www.sciensano.be/en/biblio/excess-mortality-during-first-and-second-waves-covid-19-epidemic-belgium.

34. European Medicine Agency (EMA). *Clinical Data Publication*. Available from: https://clinicaldata.ema.europa.eu/web/cdp/home.

35. Federal agency for medicines and health products (FAMHP). Vaccines; 2022. Available from: https://www.famhp.be/en/human_use/medicines/medicines/covid_19/vaccines.

36. European Medicines Agency (EMA). *COVID-19 medicines*. Available from: https://www.ema.europa.eu/en/human-regulatory-overview/public-health-threats/coronavirus-disease-covid-19/covid-19-medicines.

37. Agence fédérale des médicaments et des produits de santé (AFMPS). *Coronavirus : la Commission européenne délivre une autorisation de mise sur*

*le marché conditionnelle pour un vaccin contre la COVID-19*; 2020. Available from: https://www.afmps.be/fr/news/coronavirus_la_commission_europeenne_delivre_une_autorisation_de_mise_sur_le_marche.

38. Centers for disease control and prevention (CDC). *Understanding mRNA COVID-19 vaccines*; 2022. Available from: https://stacks.cdc.gov/view/cdc/119428.

39. Agence fédérale des médicaments et des produits de santé (AFMPS). *Coronavirus : la Commission européenne délivre une autorisation de mise sur le marché conditionnelle pour le vaccin d'AstraZeneca contre la COVID-19*; 2021. Available from: https://www.afmps.be/fr/news/coronavirus_la_commission_europeenne_delivre_une_autorisation_de_mise_sur_le_marche_1.

40. Agence fédérale des médicaments et des produits de santé (AFMPS). *Coronavirus: la Comission européenne délivre une autorisation de mise sur la marché conditonelle pour le vaccin Janssen contre la COVID-19*; 2021. Available from: https://www.afmps.be/fr/news/coronavirus_la_commission_europeenne_delivre_une_autorisation_de_mise_sur_le_marche_2.

41. Centers for disease control and prevention (CDC). *Understanding viral vector COVID-19 vaccines*; 2022. Available from: https://stacks.cdc.gov/view/cdc/115410.

42. European Medicines Agency (EMA). *EMA recommends Nuvaxovid for authorisation in the EU*; 2021. Available from: https://www.ema.europa.eu/en/news/ema-recommends-nuvaxovid-authorisation-eu.

43. Centers for disease control and prevention (CDC). *COVID-19 Vaccine Basics*; 2024. Available from: https://www.cdc.gov/covid/vaccines/how-they-work.html.

44. Catteau L, van Loenhout JAF, Stouten V, Billuart M, Hubin P, Haarhuis F, et al. Rapport thématique : Couverture vaccinale et impact épidémiologique de la campagne de vaccination COVID-19 en Belgique; 2021. Available: https://covid-19.sciensano.be/sites/default/files/Covid19/COVID-19_THEMATIC_REPORT_VaccineCoverageAndImpactReport_FR.pdf.

45. Kessels R, Luyten J, Tubeuf S. Willingness to get vaccinated against Covid-19 and attitudes toward vaccination in general. Vaccine. 2021 Jul 30;39(33):4716–22.

46. Dror AA, Eisenbach N, Taiber S, Morozov NG, Mizrachi M, Zigron A, et al. Vaccine hesitancy: the next challenge in the fight against COVID-19. Eur J Epidemiol. 2020 Aug 1;35(8):775–9.

47. European Centre for Disease Prevention and Control (ECDC). C*OVID-19 Vaccine Tracker*; 2023. Available from: https://vaccinetracker.ecdc.europa.eu/public/extensions/COVID-19/vaccine-tracker.html#uptake-tab.

48. Yang ZR, Jiang YW, Li FX, Liu D, Lin TF, Zhao ZY, et al. Efficacy of SARS-CoV-2 vaccines and the dose–response relationship with three major antibodies: a systematic review and meta-analysis of randomised controlled trials. Lancet Microbe. 2023 Apr 1;4(4):e236–46.

49. 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 May 19;22(1):42.

50. Wu N, Joyal-Desmarais K, Ribeiro PAB, Vieira AM, Stojanovic J, Sanuade C, et al. Long-term effectiveness of COVID-19 vaccines against infections, hospitalisations, and mortality in adults: findings from a rapid living

systematic evidence synthesis and meta-analysis up to December, 2022. Lancet Respir Med. 2023 May 1;11(5):439–52.

51. Xu K, Wang Z, Qin M, Gao Y, Luo N, Xie W, et al. A systematic review and meta-analysis of the effectiveness and safety of COVID-19 vaccination in older adults. Front Immunol. 2023 Mar 3;14:1113156.

52. Wang K, Wang L, Li M, Xie B, He L, Wang M, et al. Real-Word Effectiveness of Global COVID-19 Vaccines Against SARS-CoV-2 Variants: A Systematic Review and Meta-Analysis. Front Med. 2022 May 19;9:820544.

53. Ho NT, Hughes SG, Sekulovich R, Ta VT, Nguyen TV, Van Pham AT, et al. A randomized trial comparing safety, immunogenicity and efficacy of self-amplifying mRNA and adenovirus-vector COVID-19 vaccines. Npj Vaccines. 2024 Nov 23;9(1):1–10.

54. Bobrovitz N, Ware H, Ma X, Li Z, Hosseini R, Cao C, et al. Protective effectiveness of previous SARS-CoV-2 infection and hybrid immunity against the omicron variant and severe disease: a systematic review and meta-regression. Lancet Infect Dis. 2023 May 1;23(5):556–67.

55. Carabelli AM, Peacock TP, Thorne LG, Harvey WT, Hughes J, de Silva TI, et al. SARS-CoV-2 variant biology: immune escape, transmission and fitness. Nat Rev Microbiol. 2023 Mar;21(3):162–77.

56. Organisation mondiale de la santé (OMS). *Suivi des variants du SARS-CoV-2*; 2020. Available from: https://www.who.int/fr/emergencies/disease-outbreak-news/item/2020-DON297.

57. Organisation mondiale de la santé (OMS). *Suivi des variants du SARS-CoV-2*; 2020. Available from: https://www.who.int/fr/emergencies/disease-outbreak-news/item/2020-DON297.

58. Braeye T, Catteau L, Brondeel R, van Loenhout JAF, Proesmans K, Cornelissen L, et al. Vaccine effectiveness against transmission of alpha, delta and omicron SARS-COV-2-infection, Belgian contact tracing, 2021-2022. Vaccine. 2023 May 11;41(20):3292–300.

59. Van Goethem N, Chung PYJ, Meurisse M, Vandromme M, De Mot L, Brondeel R, et al. Clinical Severity of SARS-CoV-2 Omicron Variant Compared with Delta among Hospitalized COVID-19 Patients in Belgium during Autumn and Winter Season 2021–2022. Viruses. 2022 Jun 14;14(6):1297.

60. Relan P, Motaze NV, Kothari K, Askie L, Waroux OLP de, Kerkhove MDV, et al. Severity and outcomes of Omicron variant of SARS-CoV-2 compared to Delta variant and severity of Omicron sublineages: a systematic review and metanalysis. BMJ Glob Health. 2023 Jul 1;8(7):e012328.

61. Willett BJ, Grove J, MacLean OA, Wilkie C, De Lorenzo G, Furnon W, et al. SARS-CoV-2 Omicron is an immune escape variant with an altered cell entry pathway. Nat Microbiol. 2022 Aug;7(8):1161–79.

62. Jalali N, Brustad HK, Frigessi A, MacDonald EA, Meijerink H, Feruglio SL, et al. Increased household transmission and immune escape of the SARS-CoV-2 Omicron compared to Delta variants. Nat Commun. 2022 Sep 29;13(1):5706.

63. Sciensano - Epistat. *COVID-19 Monitoring*. Available from: https://epistat.sciensano.be/covid/.

64. Braeye T, Van Evercooren I, Van Loenhout J. *COVID-19 vaccine effectiveness against hospitalisation and death using electronic health records in eight European countries in the VEBIS monitoring network - October 2023 to April 2024*; 2024. Available from: https://www.sciensano.be/fr/biblio/covid-19-vaccine-effectiveness-against-hospitalisation-and-death-using-electronic-health-records.

65. Menni C, May A, Polidori L, Louca P, Wolf J, Capdevila J, et al. COVID-19 vaccine waning and effectiveness and side-effects of boosters: a prospective community study from the ZOE COVID Study. Lancet Infect Dis. 2022 Jul 1;22(7):1002–10.

66. Wong BKF, Mabbott NA. Systematic review and meta-analysis of COVID-19 mRNA vaccine effectiveness against hospitalizations in adults. Immunother Adv. 2024 Jan 1;4(1):ltae011.

67. Meurisse M, Catteau L, van Loenhout JAF, Braeye T, De Mot L, Serrien B, et al. Homologous and Heterologous Prime-Boost Vaccination: Impact on Clinical Severity of SARS-CoV-2 Omicron Infection among Hospitalized COVID-19 Patients in Belgium. Vaccines. 2023 Feb 7;11(2):378.

68. World Health Organization (WHO). *Post COVID-19 condition (Long COVID);* 2022. Available from: https://www.who.int/europe/news-room/fact-sheets/item/post-covid-19-condition.

69. Sk Abd Razak R, Ismail A, Abdul Aziz AF, Suddin LS, Azzeri A, Sha'ari NI. Post-COVID syndrome prevalence: a systematic review and meta-analysis. BMC Public Health. 2024 Jul 4;24(1):1785.

70. Hou Y, Gu T, Ni Z, Shi X, Ranney ML, Mukherjee B. Global Prevalence of Long COVID, its Subtypes and Risk factors: An Updated Systematic Review and Meta-Analysis. MedRxiv Prepr Serv Health Sci. 2025 Jan 6;2025.01.01.24319384.

71. Chen C, Haupert SR, Zimmermann L, Shi X, Fritsche LG, Mukherjee B. Global Prevalence of Post-Coronavirus Disease 2019 (COVID-19) Condition or Long COVID: A Meta-Analysis and Systematic Review. J Infect Dis. 2022 Nov 1;226(9):1593–607.

72. O'Mahoney LL, Routen A, Gillies C, Ekezie W, Welford A, Zhang A, et al. The prevalence and long-term health effects of Long Covid among hospitalised and non-hospitalised populations: a systematic review and meta-analysis. The Lancet. 2023 May 1. eClinicalMedicine, Volume 59, 101959.

73. Smith P, Charafeddine R, Drieskens S, De Pauw R, De Ridder K, Demarest S, et al. *COVIMPACT studie: Long COVID en de lichamelijke, psychische en sociale gevolgen – Resultaten van 3 maanden opvolging*. Brussels, Belgium: Sciensano; 2021 p. 20. Report No.: D/2021/14.440/85. Available from: https://doi.org/10.25608/gqcw-yk36.

74. Smith P, Proesmans K, Van Cauteren D, Demarest S, Drieskens S, De Pauw R, et al. Post COVID-19 condition and its physical, mental and social

implications: protocol of a 2-year longitudinal cohort study in the Belgian adult population. Arch Public Health. 2022 Jun 4;80(1):151.

75. NGAI JC, KO FW, NG SS, TO K, TONG M, HUI DS. The long-term impact of severe acute respiratory syndrome on pulmonary function, exercise capacity and health status. Respirol Carlton Vic. 2010 Apr;15(3):543–50.

76. Castanares-Zapatero D, Chalon P, Kohn L, Dauvrin M, Detollenaere J, Maertens de Noordhout C, et al. Pathophysiology and mechanism of long COVID: a comprehensive review. Ann Med. 2022;54(1):1473–87.

77. Peluso MJ, Deeks SG. Mechanisms of long COVID and the path toward therapeutics. Cell. 2024 Oct 3;187(20):5500–29.

78. Shirbhate E, Pandey J, Patel VK, Kamal M, Jawaid T, Gorain B, et al. Understanding the role of ACE-2 receptor in pathogenesis of COVID-19 disease: a potential approach for therapeutic intervention. Pharmacol Rep. 2021;73(6):1539–50.

79. Liew F, Efstathiou C, Fontanella S, Richardson M, Saunders R, Swieboda D, et al. Large-scale phenotyping of patients with long COVID post-hospitalization reveals mechanistic subtypes of disease. Nat Immunol. 2024 Apr;25(4):607–21.

80. Son K, Jamil R, Chowdhury A, Mukherjee M, Venegas C, Miyasaki K, et al. *Circulating anti-nuclear autoantibodies in COVID-19 survivors predict long COVID symptoms.* Eur Respir J. 2023 Jan 12;61(1):2200970.

81. Farshidfar F, Koleini N, Ardehali H. Cardiovascular complications of COVID-19. JCI Insight. 2021 Jul 8;6(13):e148980.

82. Lammi V, Nakanishi T, Jones SE, Andrews SJ, Karjalainen J, Cortés B, et al. *Genome-wide Association Study of Long COVID.* medRxiv. 2023 Jan 1;2023.06.29.23292056.

83. Copur S, Berkkan M, Basile C, Tuttle K, Kanbay M. Post-acute COVID-19 syndrome and kidney diseases: what do we know? J Nephrol. 2022;35(3):795–805.

84. World Health Organization (WHO). *Social determinants of health: Key concepts*; 2013. Available from: https://www.who.int/news-room/questions-and-answers/item/social-determinants-of-health-key-concepts.

85. Office of Disease Prevention and Health Promotion (OASH). *Social Determinants of Health*. Available from:

https://odphp.health.gov/healthypeople/priority-areas/social-determinants-health.

86. Centers for disease control and prevention (CDC). *What is Health Equity?;* 2025. Available from: https://www.cdc.gov/health-equity/what-is/index.html.

87. World Health Organization (WHO). *Uneven access to health services drives life expectancy gaps*; 2019. Available from: https://www.who.int/news/item/04-04-2019-uneven-access-to-health-services-drives-life-expectancy-gaps-who.

88. Otavova M, Masquelier B, Faes C, van den Borre L, Vandeninden B, de Clercq E, et al. Trends in socioeconomic inequalities in cause-specific premature mortality in Belgium, 1998–2019. BMC Public Health. 2024 Feb 14;24(1):470.

89. Whitehead M. A typology of actions to tackle social inequalities in health. J Epidemiol Community Health. 2007 Jun;61(6):473–8.

90. Johnson NPAS, Mueller J. Updating the accounts: global mortality of the 1918-1920 "Spanish" influenza pandemic. Bull Hist Med. 2002;76(1):105–15.

91. 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 Sep 1;12(9):687–95.

92. Grantz KH, Rane MS, Salje H, Glass GE, Schachterle SE, Cummings DAT. Disparities in influenza mortality and transmission related to sociodemographic factors within Chicago in the pandemic of 1918. Proc Natl Acad Sci. 2016 Nov 29;113(48):13839–44.

93. Basco S, Domènech J, Rosés JR. Socioeconomic mortality differences during the Great Influenza in Spain. Econ Hum Biol. 2024 Jan 1;52:101318.

94. Lowcock EC, Rosella LC, Foisy J, McGeer A, Crowcroft N. The Social Determinants of Health and Pandemic H1N1 2009 Influenza Severity. Am J Public Health. 2012 Aug;102(8):e51–8.

95. Tricco AC, Lillie E, Soobiah C, Perrier L, Straus SE. Impact of H1N1 on socially disadvantaged populations: summary of a systematic review. Influenza Other Respir Viruses. 2013 Sep;7(Suppl 2):54–8.

96. Rutter PD, Mytton OT, Mak M, Donaldson LJ. Socio-economic disparities in mortality due to pandemic influenza in England. Int J Public Health. 2012 Aug 1;57(4):745–50.

97. Sky news. *Coronavirus: "Virus does not discriminate" - Gove*; 2020. Available from: https://news.sky.com/video/coronavirus-virus-does-not-discriminate-gove-11964771.

98. Irizar P, Pan D, Kapadia D, Bécares L, Sze S, Taylor H, et al. *Ethnic inequalities in COVID-19 infection, hospitalisation, intensive care admission, and death: a global systematic review and meta-analysis of over 200 million study participants.* EClinicalMedicine. 2023 Mar;57:101877.

99. Ghattas J, Makovski TT, Monnier-Besnard S, Cavillot L, Ambrožová M, Vašinová B, et al. *The role of socio-economic determinants in SARS-CoV-2 health outcomes: systematic review of population-based studies.* medRxiv. 2024 Jan 1;2024.06.18.24309062.

100. Bambra C, Riordan R, Ford J, Matthews F. The COVID-19 pandemic and health inequalities. J Epidemiol Community Health. 2020 Jun 13;jech-2020-214401.

101. Horton R. Offline: COVID-19 is not a pandemic. The Lancet. 2020 Sep;396(10255):874.

102. Singer M, Bulled N, Ostrach B, Mendenhall E. Syndemics and the biosocial conception of health. Lancet Lond Engl. 2017 Mar 4;389(10072):941–50.

103. Pathirana TI, Jackson CA. Socioeconomic status and multimorbidity: a systematic review and meta-analysis. Aust N Z J Public Health. 2018;42(2):186–94.

104. Gold MS, Sehayek D, Gabrielli S, Zhang X, McCusker C, Ben-Shoshan M. COVID-19 and comorbidities: a systematic review and meta-analysis. Postgrad Med. 2020 Nov;132(8):749–55.

105. Bambra C, Lynch J, Smith KE. *The Unequal Pandemic: COVID- 19 and Health Inequalities.* Policy Press; 2021. Available from: https://library.oapen.org/handle/20.500.12657/51451.

106. Kaplan RM, Fang Z, Kirby J. Educational attainment and health outcomes: Data from the Medical Expenditures Panel Survey. Health Psychol Off J Div Health Psychol Am Psychol Assoc. 2017 Jun;36(6):598–608.

107. Pförtner TK, Dohle S, Hower KI. Trends in educational disparities in preventive behaviours, risk perception, perceived effectiveness and trust in

the first year of the COVID-19 pandemic in Germany. BMC Public Health. 2022 May 6;22(1):903.

108. Ahmad K, Erqou S, Shah N, Nazir U, Morrison AR, Choudhary G, et al. Association of poor housing conditions with COVID-19 incidence and mortality across US counties. PLoS ONE. 2020 Nov 2;15(11):e0241327.

109. Brandily P, Brébion C, Briole S, Khoury L. *A poorly understood disease? The impact of COVID-19 on the income gradient in mortality over the course of the pandemic*. Eur Econ Rev. 2021 Nov;140:103923.

110. SPF Justice. 25 FEVRIER 2018. - *Loi portant création de Sciensano*; 2018. Available from: https://www.ejustice.just.fgov.be/cgi_loi/change_lg.pl?language=fr&la=F&cn=2018022502&table_name=loi.

111. Sciensano. *Sciensano & Coronavirus*. Available from: https://www.sciensano.be/fr/sujets-sante/coronavirus/role.

112. Van Goethem N, Vilain A, Wyndham-Thomas C, Deblonde J, Bossuyt N, Lernout T, et al. Rapid establishment of a national surveillance of COVID-19 hospitalizations in Belgium. Arch Public Health Arch Belg Sante Publique. 2020 Nov 18;78(1):121.

113. Renard F, Scohy A, Heyden JV der, Peeters I, Dequeker S, Vandael E, et al. *Establishing an ad hoc COVID-19 mortality surveillance during the first epidemic wave in Belgium, 1 March to 21 June 2020.* Eurosurveillance. 2021 Dec 2;26(48):2001402.

114. Meurisse M, Lajot A, Dupont Y, Lesenfants M, Klamer S, Rebolledo J, et al. One year of laboratory-based COVID-19 surveillance system in Belgium: main indicators and performance of the laboratories (March 2020-21). Arch Public Health Arch Belg Sante Publique. 2021 Oct 27;79(1):188.

115. Vandael E, Latour K, Islamaj E, Panis LI, Callies M, Haarhuis F, et al. COVID-19 cases, hospitalizations and deaths in Belgian nursing homes: results of a surveillance conducted between April and December 2020. Arch Public Health. 2022 Dec;80(1):45.

116. Vlaanderen verbeelding werkt. *Vaccinatie Systeem (Vaccinnet)*; 2025. Available from: https://www.vaccinnet.be/Vaccinnet/welkom.do.

117. Linn ST, Guralnik JM, Patel KV. Disparities in Influenza Vaccine Coverage in the United States, 2008. J Am Geriatr Soc. 2010;58(7):1333–40.

118. OECD/European Observatory on Health Systems and Policies (2021), *Belgium: Country Health Profile 2021*, State of Health in the EU,

OECD Publishing, Paris. Available from: https://doi.org/10.1787/57e3abb5-en.

119. Molenberghs G, Faes C, Verbeeck J, Deboosere P, Abrams S, Willem L, et al. *Belgian COVID-19 Mortality, Excess Deaths, Number of Deaths per Million, and Infection Fatality Rates (9 March — 28 June 2020).* medRxiv. 2020 Jan 1;2020.06.20.20136234.

120. Gadeyne S, Rodriguez-Loureiro L, Surkyn J, Hemelrijck WV, Nusselder W, Lusyne P, et al. Are we really all in this together? The social patterning of mortality during the first wave of the COVID-19 pandemic in Belgium. Int J Equity Health. 2021 Dec 18;20(1):258.

121. Meurisse M, Lajot A, Devleesschauwer B, Van Cauteren D, Van Oyen H, Van den Borre L, et al. The association between area deprivation and COVID-19 incidence: a municipality-level spatio-temporal study in Belgium, 2020–2021. Arch Public Health. 2022 Apr 2;80(1):109.

122. Edouard Mathieu, Hannah Ritchie, Lucas Rodés-Guirao, Cameron Appel, Daniel Gavrilov, Charlie Giattino, Joe Hasell, Bobbie Macdonald, Saloni Dattani, Diana Beltekian, Esteban Ortiz-Ospina, and Max Roser (2020) - "Coronavirus (COVID-19) Deaths" Published online at OurWorldinData.org. Available from: https://ourworldindata.org/covid-deaths.

123. Johns Hopkins Coronavirus Resource Center. Mortality Analyses; 2023. Available from: https://coronavirus.jhu.edu/data/mortality.

124. Russell CD, Lone NI, Baillie JK. Comorbidities, multimorbidity and COVID-19. Nat Med. 2023 Feb;29(2):334–43.

125. Van Wilder L, Devleesschauwer B, Clays E, Van der Heyden J, Charafeddine R, Scohy A, et al. QALY losses for chronic diseases and its social distribution in the general population: results from the Belgian Health Interview Survey. BMC Public Health. 2022 Jul 7;22(1):1304.

126. Sciensano. La santé des belges reste globalement stable, mais les maladies chroniques augmentent; 2025. Available from: https://www.sciensano.be/fr/coin-presse/la-sante-des-belges-reste-globalement-stable-mais-les-maladies-chroniques-augmentent.

127. Statbel. *Statbel - l'office belge de statistique*. Available from: https://statbel.fgov.be/fr.

128. Statbel. *Demobel: nouvelle base de données démographique;* 2019. Available from: https://statbel.fgov.be/fr/nouvelles/demobel-nouvelle-base-de-donnees-demographique.

129. World Health Organization (WHO). *International Classification of Diseases and Related Health Problems (ICD)*. Available from: https://www.who.int/standards/classifications/classification-of-diseases.

130. Sciensano. *Linking of registers for COVID-19 vaccine surveillance*; 2024. Available from: https://www.sciensano.be/en/projects/linking-registers-covid-19-vaccine-surveillance.

131. Sciensano. *healthdata.be - data we care for*. Available from: https://healthdata.sciensano.be/fr.

132. Sciensano. *La collecte des données - healthdata.be.* Available from: https://healthdata.sciensano.be/fr/la-collecte-des-donn%C3%A9es.

133. Van Goethem N, Serrien B, Vandromme M, Wyndham-Thomas C, Catteau L, Brondeel R, et al. Conceptual causal framework to assess the effect of SARS-CoV-2 variants on COVID-19 disease severity among hospitalized patients. Arch Public Health. 2021 Oct 25;79(1):185.

134. Limesurvey. *Créer des sondages: LimeSurvey*. Available from: https://www.limesurvey.org/fr.

135. Sciensano. *The COVID-19 Test Results data collection*; 2023. Available from: https://docs.healthdata.be/fr/node/282.

136. Rambaut, A., Holmes, E.C., O'Toole, Á. *et al. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology*. Nat Microbiol 5, 1403–1407 (2020). https://doi.org/10.1038/s41564-020-0770-5

137. Agentschap Zorg & Gezonheid. *VACCINNET - Manuel d'utilisation Vaccinnet*; 2021. Available from: https://covid.aviq.be/sites/default/files/fichiers-upload/Manuel%20d%27utilisation%20Vaccinnet%20v3.12.pdf.

138. eHealth. *CoBRHA - Common Base Registry for HealthCare Actor.* Available from: https://www.ehealth.fgov.be/ehealthplatform/fr/service-cobrha-common-base-registry-for-healthcare-actor.

139. Gerkens S, Lefèvre M, Bouckaert N, Levy M, Maertens de Noordhout C, Obyn C, et al. Performance du système de santé belge : rapport 2024. Health Services Research (HSR). Bruxelles: Centre Fédéral d'Expertise des Soins de Santé (KCE); 2024. KCE Reports 376B.

140. Agence Intermutusaliste. *Metadata : db Population*. Available from: https://metadata.ima-aim.be/fr/app/bdds/Pp.

141. Norwegian Institute of Public Health (NIPH). *ATCDDD - ATC/DDD Index*; 2024. Available from: https://atcddd.fhi.no/atc_ddd_index/.

142. Agence Intermutualiste. *Echantillon Permanente Steekproef EPS R13 – FLAGS Release 20190201*; 2023. Available from: https://ima-aim.be/IMG/pdf/eps_r13_-_flags_release_20190201_nl_-_vs2.pdf.

143. HELICON. *Home*; 2020. Available from: https://www.brain-helicon.be/index.php?option=com_content&view=article&id=2:home&catid=8.

144. Khan Y, Verhaeghe N, De Pauw R, Devleesschauwer B, Gadeyne S, Gorasso V, et al. *Correction: Evaluating the health and health economic impact of the COVID-19 pandemic on delayed cancer care in Belgium: A Markov model study protocol.* PLoS One. 2023 Dec 12;18(12):e0295992.

145. Khan Y, Van den Borre L, Smedt D, Verhaeghe N, Devleesschauwer B, Deboosere P, Vanthomme K, Gadeyne S. *A Nationwide Exploration of Social Inequalities in Cancer Mortality Amidst the COVID-19 Pandemic in Belgium.* Cancer Med. 2025 Jan;14(1):e70487.

146. Khan Y, Verhaeghe N, Devleesschauwer B, Cavillot L, Gadeyne S, Pauwels N, et al. The impact of the COVID-19 pandemic on delayed care of cardiovascular diseases in Europe: a systematic review. Eur Heart J - Qual Care Clin Outcomes. 2023 Nov 2;9(7):647–61.

147. Sharp S, Poulaliou M, Thompson S, White I, Wood A. A Review of Published Analyses of Case-Cohort Studies and Recommendations for Future Reporting. PloS One. 2014 Jun 27;9:e101176.

148. De Pauw R, Van den Borre L, Baeyens Y, Cavillot L, Gadeyne S, Ghattas J, et al. Social inequalities and long-term health impact of COVID-19 in Belgium: protocol of the HELICON population data linkage. BMJ Open. 2023 May 18;13(5):e069355.

149. eHealth. *Missions*. Available from: https://www.ehealth.fgov.be/ehealthplatform/fr/mission.

150. Banque Carrefour de la Sécurité Sociale. *Accueil*. Available from: https://www.ksz-bcss.fgov.be/fr.

151. Holland PW. Statistics and Causal Inference. J Am Stat Assoc. 1986 Dec;81(396):945–60.

152. Pearl J. An Introduction to Causal Inference. Int J Biostat. 2010 Feb 26;6(2):7.

153. Little RJ, Rubin DB. Causal effects in clinical and epidemiological studies via potential outcomes: concepts and analytical approaches. Annu Rev Public Health. 2000;21:121–45.

154. Imbens GW, Rubin DB. *Causal Inference for Statistics, Social, and Biomedical Sciences: An introduction.* Cambridge University Press; 2015.

155. Stone R. The Assumptions on Which Causal Inferences Rest. J R Stat Soc Ser B Methodol. 1993;55(2):455–66.

156. Rubin DB. Estimating causal effects of treatments in randomized and nonrandomized studies. J Educ Psychol. 1974;66(5):688–701.

157. Gvozdenović E, Malvisi L, Cinconze E, Vansteelandt S, Nakanwagi P, Aris E, et al. Causal inference concepts applied to three observational studies in the context of vaccine development: from theory to practice. BMC Med Res Methodol. 2021 Feb 15;21(1):35.

158. Bzdok D, Altman N, Krzywinski M. Statistics versus machine learning. Nat Methods. 2018 Apr 1;15(4):233–4.

159. Imai K, Keele L, Tingley D. A general approach to causal mediation analysis. Psychol Methods. 2010;15(4):309–34.

160. Zhang Z, Zheng C, Kim C, Van Poucke S, Lin S, Lan P. Causal mediation analysis in the context of clinical research. Ann Transl Med. 2016 Nov;4(21):425.

161. Nguyen TQ, Schmid I, Stuart EA. Clarifying causal mediation analysis for the applied researcher: Defining effects based on what we want to learn. Psychol Methods. 2021;26(2):255–71.

162. Rijnhart JJM, Lamp SJ, Valente MJ, MacKinnon DP, Twisk JWR, Heymans MW. Mediation analysis methods used in observational research: a scoping review and recommendations. BMC Med Res Methodol. 2021 Oct 25;21(1):226.

163. VanderWeele TJ. Mediation Analysis: A Practitioner's Guide. Annu Rev Public Health. 2016;37(1):17–32.

164. Tennant PWG, Murray EJ, Arnold KF, Berrie L, Fox MP, Gadd SC, et al. Use of directed acyclic graphs (DAGs) to identify confounders in applied health research: review and recommendations. Int J Epidemiol. 2021 May 17;50(2):620–32.

165. Pearl J. Interpretation and identification of causal mediation. Psychol Methods. 2014 Dec;19(4):459–81.

166. MacKinnon, D. (2008). *Introduction to Statistical Mediation Analysis* (1st ed.). Routledge. Available from: https://doi.org/10.4324/9780203809556.

167. Vanderweele TJ, Vansteelandt S. Conceptual issues concerning mediation, interventions and composition. Stat Interface. 2009;2(4):457–68.

168. Daniel RM, De Stavola BL, Cousens SN, Vansteelandt S. Causal Mediation Analysis with Multiple Mediators. Biometrics. 2015 Mar 1;71(1):1–14.

169. Loh WW, Ren D. Improving causal inference of mediation analysis with multiple mediators using interventional indirect effects. Soc Personal Psychol Compass. 2022;16(10):e12708.

170. Loh WW, Moerkerke B, Loeys T, Vansteelandt S. Heterogeneous Indirect Effects for Multiple Mediators using Interventional Effect Models. Epidemiol Methods. 2020 Jan 28;9(1):20200023.

171. Loh WW, Moerkerke B, Loeys T, Vansteelandt S. Disentangling indirect effects through multiple mediators without assuming any causal structure among the mediators. Psychol Methods. 2022;27:982–99.

172. Vansteelandt S, Daniel RM. Interventional effects for mediation analysis with multiple mediators. Epidemiol Camb Mass. 2017 Mar;28(2):258–65.

173. Moreno-Betancur M, Carlin JB. Brief Report: Understanding Interventional Effects: A More Natural Approach to Mediation Analysis? Epidemiology. 2018;29(5):614–7.

174. Gürel S, Leite WL. Comparison of Propensity Score Weighting Methods to Remove Selection Bias in Average Treatment Effect Estimates. Uluslar Türk Eğitim Bilim Derg. 2023 Oct 31;2023(21):989–1031.

175. Fu EL, Groenwold RHH, Zoccali C, Jager KJ, van Diepen M, Dekker FW. Merits and caveats of propensity scores to adjust for confounding. Nephrol Dial Transplant Off Publ Eur Dial Transpl Assoc - Eur Ren Assoc. 2019 Oct 1;34(10):1629–35.

176. Rosenbaum PR, Rubin DB. *The central role of the propensity score in observational studies for causal effects.* Biometrika. 1983 Apr 1;70(1):41–55.

177. Chesnaye NC, Stel VS, Tripepi G, Dekker FW, Fu EL, Zoccali C, et al. An introduction to inverse probability of treatment weighting in observational research. Clin Kidney J. 2021 Aug 26;15(1):14–20.

178. Bishop CD, Leite WL, Snyder PA. Using Propensity Score Weighting to Reduce Selection Bias in Large-Scale Data Sets. J Early Interv. 2018 Dec;40(4):347–62.

179. Austin PC. An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. Multivar Behav Res. 2011 May;46(3):399–424.

180. Thomas LE, Li F, Pencina MJ. Overlap Weighting: A Propensity Score Method That Mimics Attributes of a Randomized Clinical Trial. JAMA. 2020 Jun 16;323(23):2417–8.

181. Li F, Thomas LE, Li F. Addressing Extreme Propensity Scores via the Overlap Weights. Am J Epidemiol. 2019 Jan 1;188(1):250–7.

182. Zhou Y, Matsouaka RA, Thomas L. Propensity score weighting under limited overlap and model misspecification. Stat Methods Med Res. 2020 Dec;29(12):3721–56.

183. Li F, Thomas LE, Li F. *Addressing Extreme Propensity Scores via the Overlap Weights.* Am J Epidemiol. 2019 Jan 1;188(1):250–7.

184. Thomas L, Li F, Pencina M. Using Propensity Score Methods to Create Target Populations in Observational Clinical Research. JAMA. 2020 Feb 4;323(5):466–7.

185. Vansteelandt S, Daniel R m. On regression adjustment for the propensity score. Stat Med. 2014;33(23):4053–72.

186. Mehta N, Kalra A, Nowacki AS, Anjewierden S, Han Z, Bhat P, et al. Association of Use of Angiotensin-Converting Enzyme Inhibitors and

Angiotensin II Receptor Blockers With Testing Positive for Coronavirus Disease 2019 (COVID-19). JAMA Cardiol. 2020 Sep 1;5(9):1020–6.

187. Gabriel EE, Sachs MC, Martinussen T, Waernbaum I, Goetghebeur E, Vansteelandt S, et al. Inverse probability of treatment weighting with generalized linear outcome models for doubly robust estimation. Stat Med. 2024 Feb 10;43(3):534–47.

188. Speybroeck N. Classification and regression trees. Int J Public Health. 2012 Feb 1;57(1):243–6.

189. SONG Y yan, LU Y. Decision tree methods: applications for classification and prediction. Shanghai Arch Psychiatry. 2015 Apr 25;27(2):130–5.

190. Hothorn T, Hornik K, Zeileis A. *ctree: Conditional Inference Trees*. Available from: https://cran.r-project.org/web/packages/partykit/vignettes/ctree.pdf.

191. Hothorn T, Hornik K, Zeileis A. Unbiased Recursive Partitioning: A Conditional Inference Framework. J Comput Graph Stat. 2006 Sep;15(3):651–74.

192. Shih YS. A note on split selection bias in classification trees. Comput Stat Data Anal. 2004 Apr 10;45(3):457–66.

193. Kuncheva LI. *Combining Pattern Classifiers: Methods and Algorithms*. 2nd ed. Hoboken (NJ): Wiley; 2014.

194. Rdocumentation. *ctree function*; 2020. Available from: https://www.rdocumentation.org/packages/partykit/versions/1.2-20/topics/ctree.

195. Monaghan TF, Rahman SN, Agudelo CW, Wein AJ, Lazar JM, Everaert K, et al. Foundational Statistical Principles in Medical Research: Sensitivity, Specificity, Positive Predictive Value, and Negative Predictive Value. Medicina (Mex). 2021 May 16;57(5):503.

196. Zou KH, O'Malley AJ, Mauri L. Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models. Circulation. 2007 Feb 6;115(5):654–7.

197. Columbia University Mailman School of Public Health. *Evaluating Risk Prediction with ROC Curves*; 2016. Available from: https://www.publichealth.columbia.edu/research/population-health-methods/evaluating-risk-prediction-roc-curves.

198. Austin PC, Lee DS, Fine JP. Introduction to the Analysis of Survival Data in the Presence of Competing Risks. Circulation. 2016 Feb 9;133(6):601–9.

199. Columbia University Mailman School of Public Health. *Competing Risk Analysis*; 2016. Available from: https://www.publichealth.columbia.edu/research/population-health-methods/competing-risk-analysis.

200. de Glas NA, Kiderlen M, Vandenbroucke JP, de Craen AJM, Portielje JEA, van de Velde CJH, et al. Performing Survival Analyses in the Presence of Competing Risks: A Clinical Example in Older Breast Cancer Patients. JNCI J Natl Cancer Inst. 2016 May 1;108(5):djv366.

201. Zhang MJ, Zhang X, Scheike TH. Modeling cumulative incidence function for competing risks data. Expert Rev Clin Pharmacol. 2008 May 1;1(3):391–400.

202. Lau B, Cole SR, Gange SJ. *Competing Risk Regression Models for Epidemiologic Data.* Am J Epidemiol. 2009 Jul 15;170(2):244–56.

203. Wolbers M, Koller MT, Stel VS, Schaer B, Jager KJ, Leffondré K, et al. Competing risks analyses: objectives and approaches. Eur Heart J. 2014 Nov 7;35(42):2936–41.

204. Satagopan JM, Ben-Porat L, Berwick M, Robson M, Kutler D, Auerbach AD. A note on competing risks in survival data analysis. Br J Cancer. 2004 Oct;91(7):1229–35.

205. Gray RJ. A Class of K-Sample Tests for Comparing the Cumulative Incidence of a Competing Risk. Ann Stat. 1988;16(3):1141–54.

206. Dignam JJ, Zhang Q, Kocherginsky MN. The Use and Interpretion of Competing Risks Regression Models. Clin Cancer Res. 2012 Apr 15;18(8):2301–8.

207. Austin PC, Fine JP. Practical recommendations for reporting Fine-Gray model analyses for competing risk data. Stat Med. 2017 Nov 30;36(27):4391–400.

208. Vandentorren S, Smaïli S, Chatignoux E, Maurel M, Alleaume C, Neufcourt L, et al. The effect of social deprivation on the dynamic of SARS-CoV-2 infection in France: a population-based analysis. Lancet Public Health. 2022 Mar 1;7(3):e240–9.

209. Abedi V, Olulana O, Avula V, Chaudhary D, Khan A, Shahjouei S, et al. Racial, Economic and Health Inequality and COVID-19 Infection in the United States. medRxiv. 2020 May 1;2020.04.26.20079756.

210. Magesh S, John D, Li WT, Li Y, Mattingly-app A, Jain S, et al. Disparities in COVID-19 Outcomes by Race, Ethnicity, and Socioeconomic Status: A Systematic Review and Meta-analysis. JAMA Netw Open. 2021 Nov 11;4(11):e2134147.

211. Decoster A, Minten T, Spinnewijn J. The Income Gradient in Mortality during the Covid-19 Crisis: Evidence from Belgium. J Econ Inequal. 2021 Sep;19(3):551–70.

212. Vanthomme K, Gadeyne S, Lusyne P, Vandenheede H. A population-based study on mortality among Belgian immigrants during the first COVID-19 wave in Belgium. Can demographic and socioeconomic indicators explain differential mortality? SSM - Popul Health. 2021 Jun 1;14:100797.

213. Haas EJ, Angulo FJ, McLaughlin JM, Anis E, Singer SR, Khan F, et al. Impact and effectiveness of mRNA BNT162b2 vaccine against SARS-CoV-2 infections and COVID-19 cases, hospitalisations, and deaths following a nationwide vaccination campaign in Israel: an observational study using national surveillance data. Lancet Lond Engl. 2021 May 15;397(10287):1819–29.

214. Sadoff J, Gray G, Vandebosch A, Cárdenas V, Shukarev G, Grinsztejn B, et al. Safety and Efficacy of Single-Dose Ad26.COV2.S Vaccine against Covid-19. N Engl J Med. 2021 Jun 10;384(23):2187–201.

215. Spetz M, Lundberg L, Nwaru C, Li H, Santosa A, Leach S, et al. The social patterning of Covid-19 vaccine uptake in older adults: A register-based cross-sectional study in Sweden. Lancet Reg Health - Eur. 2022 Apr;15:100331.

216. Williams AM, Clayton HB, Singleton JA. Racial and Ethnic Disparities in COVID-19 Vaccination Coverage: The Contribution of Socioeconomic and Demographic Factors. Am J Prev Med. 2022 Apr;62(4):473–82.

217. Faes C, Molenberghs G, Hens N, Van Bortel L, Vandeboel N, Pellens K, et al. Geographical variation of COVID-19 vaccination coverage, ethnic diversity and population composition in Flanders. Vaccine X. 2022 Aug;11:100194.

218. Luyten J, Beutels P. The Social Value Of Vaccination Programs: Beyond Cost-Effectiveness. Health Aff (Millwood). 2016 Feb;35(2):212–8.

219. Levin J, Bradshaw M. Determinants of COVID-19 skepticism and SARS-CoV-2 vaccine hesitancy: findings from a national population survey of U.S. adults. BMC Public Health. 2022 Dec;22(1):1047.

220. Razai MS, Oakeshott P, Esmail A, Wiysonge CS, Viswanath K, Mills MC. COVID-19 vaccine hesitancy: the five Cs to tackle behavioural and sociodemographic factors. J R Soc Med. 2021 Jun 1;114(6):295–8.

221. R Core Team. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing; 2020. Available from: https://www.R-project.org/.

222. Fox J, Weisberg S, Price B, Friendly M, Hong J, Andersen R, et al. effects: *Effect Displays for Linear, Generalized Linear, and Other Models*; 2022. Available from: https://CRAN.R-project.org/package=effects.

223. Perry M, Akbari A, Cottrell S, Gravenor MB, Roberts R, Lyons RA, et al. Inequalities in coverage of COVID-19 vaccination: A population register based cross-sectional study in Wales, UK. Vaccine. 2021 Oct 8;39(42):6256–61.

224. Barry V, Dasgupta S, Weller DL, Kriss JL, Cadwell BL, Rose C, et al. Patterns in COVID-19 Vaccination Coverage, by Social Vulnerability and Urbanicity - United States, December 14, 2020-May 1, 2021. MMWR Morb Mortal Wkly Rep. 2021 Jun 4;70(22):818–24.

225. Farah W, Breeher L, Shah V, Hainy C, Tommaso CP, Swift MD. Disparities in COVID-19 vaccine uptake among health care workers. Vaccine. 2022 Apr 26;40(19):2749–54.

226. Azamgarhi T, Hodgkinson M, Shah A, Skinner JA, Hauptmannova I, Briggs TWR, et al. BNT162b2 vaccine uptake and effectiveness in UK healthcare workers – a single centre cohort study. Nat Commun. 2021 Dec;12(1):3698.

227. Robertson E, Reeve KS, Niedzwiedz CL, Moore J, Blake M, Green MJ, et al. OP34 Ethnic and educational inequalities in COVID-19 vaccine hesitancy: cross-sectional analysis of the UK household longitudinal study. J Epidemiol Community Health. 2021 Sep 1;75(Suppl 1):A16–A16.

228. AlShurman BA, Khan AF, Mac C, Majeed M, Butt ZA. What Demographic, Social, and Contextual Factors Influence the Intention to Use COVID-19 Vaccines: A Scoping Review. Int J Environ Res Public Health. 2021 Sep 4;18(17):9342.

229. Paul E, Steptoe A, Fancourt D. Attitudes towards vaccines and intention to vaccinate against COVID-19: Implications for public health communications. Lancet Reg Health Eur. 2021 Feb;1:100012.

230. Wang MW, Wen W, Wang N, Zhou MY, Wang C yi, Ni J, et al. COVID-19 Vaccination Acceptance Among Healthcare Workers and Non-healthcare Workers in China: A Survey. Front Public Health. 2021 Aug 2;9:709056.

231. American Psychological Association. *Building Vaccine Confidence Through Community Engagement*; 2020. Available: http://doi.apa.org/get-pe-doi.cfm?doi=10.1037/e510882020-001.

232. Caspi G, Dayan A, Eshal Y, Liverant-Taub S, Twig G, Shalit U, et al. Socioeconomic disparities and COVID-19 vaccination acceptance: a nationwide ecologic study. Clin Microbiol Infect Off Publ Eur Soc Clin Microbiol Infect Dis. 2021 Oct;27(10):1502–6.

233. Montagni I, Ouazzani-Touhami K, Mebarki A, Texier N, Schück S, Tzourio C, et al. Acceptance of a Covid-19 vaccine is associated with ability to detect fake news and health literacy. J Public Health Oxf Engl. 2021 Dec 10;43(4):695–702.

234. Svendsen MT, Bak CK, Sørensen K, Pelikan J, Riddersholm SJ, Skals RK, et al. Associations of health literacy with socioeconomic position, health risk behavior, and health status: a large national population-based survey among Danish adults. BMC Public Health. 2020 Dec;20(1):565.

235. Rosenthal S, Cummings CL. Influence of rapid COVID-19 vaccine development on vaccine hesitancy. Vaccine. 2021 Dec 20;39(52):7625–32.

236. McGowan VJ, Bambra C. COVID-19 mortality and deprivation: pandemic, syndemic, and endemic health inequalities. Lancet Public Health. 2022 Nov 1;7(11):e966–75.

237. Fronteira I, Sidat M, Magalhães JP, de Barros FPC, Delgado AP, Correia T, Daniel-Ribeiro CT, Ferrinho P. *The SARS-CoV-2 pandemic: A syndemic perspective*. One Health. 2021 Jun;12:100228.

238. Khanijahani A, Iezadi S, Gholipour K, Azami-Aghdash S, Naghibi D. A systematic review of racial/ethnic and socioeconomic disparities in COVID-19. Int J Equity Health. 2021 Nov 24;20:248.

239. Álvarez-Gálvez J, Ortega-Martín E, Carretero-Bravo J, Pérez-Muñoz C, Suárez-Lledó V, Ramos-Fiol B. *Social determinants of multimorbidity patterns: A systematic review.* Front Public Health. 2023 Mar 27;11:1081518.

240. Nanda S, Chacin Suarez AS, Toussaint L, Vincent A, Fischer KM, Hurt R, et al. Body Mass Index, Multi-Morbidity, and COVID-19 Risk Factors as Predictors of Severe COVID-19 Outcomes. J Prim Care Community Health. 2021;12:21501327211018560.

241. Feikin DR, Higdon MM, Abu-Raddad LJ, Andrews N, Araos R, Goldberg Y, et al. Duration of effectiveness of vaccines against SARS-CoV-2 infection and COVID-19 disease: results of a systematic review and meta-regression. The Lancet. 2022 Mar 5;399(10328):924–44.

242. Mohammed H, Pham-Tran DD, Yeoh ZYM, Wang B, McMillan M, Andraweera PH, et al. A Systematic Review and Meta-Analysis on the Real-World Effectiveness of COVID-19 Vaccines against Infection, Symptomatic and Severe COVID-19 Disease Caused by the Omicron Variant (B.1.1.529). Vaccines. 2023 Feb;11(2):224.

243. Zapata-Moya ÁR, Willems B, Bracke P. The (re)production of health inequalities through the process of disseminating preventive innovations: the dynamic influence of socioeconomic status. Health Sociol Rev. 2019 May 4;28(2):177–93.

244. Cavillot L, van Loenhout JAF, Devleesschauwer B, Wyndham-Thomas C, Van Oyen H, Ghattas J, et al. Sociodemographic and socioeconomic disparities in COVID-19 vaccine uptake in Belgium: a nationwide record linkage study. J Epidemiol Community Health. 2023 Dec 26;jech-2023-220751.

245. Castro-Sánchez E, Chang PWS, Vila-Candel R, Escobedo AA, Holmes AH. Health literacy and infectious diseases: why does it matter? Int J Infect Dis IJID Off Publ Int Soc Infect Dis. 2016 Feb;43:103–10.

246. McCaffery KJ, Dodd RH, Cvejic E, Ayrek J, Batcup C, Isautier JM, et al. *Health literacy and disparities in COVID-19-related knowledge, attitudes, beliefs and behaviours in Australia.* Public Health Res Pract. 2020 Dec 9;30(4):30342012.

247. Lillebråten A, Todd M, Dimka J, Bakkeli NZ, Mamelund SE. Socioeconomic status and disparities in COVID-19 vaccine uptake in Eastern Oslo, Norway. Public Health Pract. 2023 May 28;5:100391.

248. Bajos N, Spire A, Silberzan L, Sireyjol A, Jusot F, Meyer L, Franck JE, Warszawski J; EpiCov study group. *When Lack of Trust in the Government and in Scientists Reinforces Social Inequalities in Vaccination Against COVID-19.* Front Public Health. 2022 Jul 20;10:908152.

249. Statbel. *Structure de la population.* Available from: https://statbel.fgov.be/fr/themes/population/structure-de-la-population#panel-13.

250. Sciensano. Belgium COVID-19 dashboard. Available from: http://datastudio.google.com/reporting/c14a5cfc-cab7-4812-848c-0369173148ab/page/tpRKB?feature=opengraph.

251. Zeng B, Gao L, Zhou Q, Yu K, Sun F. Effectiveness of COVID-19 vaccines against SARS-CoV-2 variants of concern: a systematic review and meta-analysis. BMC Med. 2022 May 23;20(1):200.

252. European Centre for Disease Prevention and Control (ECDC). *Surveillance definitions for COVID-19*; 2023. Available from:

https://www.ecdc.europa.eu/en/covid-19/surveillance/surveillance-definitions.

253. Lederer DJ, Bell SC, Branson RD, Chalmers JD, Marshall R, Maslove DM, et al. Control of Confounding and Reporting of Results in Causal Inference Studies. Guidance for Authors from Editors of Respiratory, Sleep, and Critical Care Journals. Ann Am Thorac Soc. 2019 Jan;16(1):22–8.

254. Robalo Q, De Mot L, Vandromme M, Van Goethem N, Gabrio A, Chung PYJ, et al. Association between COVID-19 Primary Vaccination and Severe Disease Caused by SARS-CoV-2 Delta Variant among Hospitalized Patients: A Belgian Retrospective Cohort Study. Vaccines. 2023 Jan;11(1):14.

255. Higdon MM, Wahl B, Jones CB, Rosen JG, Truelove SA, Baidya A, et al. A Systematic Review of Coronavirus Disease 2019 Vaccine Efficacy and Effectiveness Against Severe Acute Respiratory Syndrome Coronavirus 2 Infection and Disease. Open Forum Infect Dis. 2022 Jun 1;9(6):ofac138.

256. Reyna-Villasmil E, Caponcello MG, Maldonado N, Olivares P, Caroccia N, Bonazzetti C, et al. Association of Patients' Epidemiological Characteristics and Comorbidities with Severity and Related Mortality Risk of SARS-CoV-2 Infection: Results of an Umbrella Systematic Review and Meta-Analysis. Biomedicines. 2022 Oct;10(10):2437.

257. Ganguli S, Howlader S, Dey K, Barua S, Islam MdN, Aquib TI, et al. Association of comorbidities with the COVID-19 severity and hospitalization: A study among the recovered individuals in Bangladesh. Int J Health Sci. 2022;16(4):30–45.

258. Johnston DW, Lordan G, Shields MA, Suziedelyte A. Education and health knowledge: Evidence from UK compulsory schooling reform. Soc Sci Med. 2015 Feb 1;127:92–100.

259. Statbel. *Revenus fiscaux*; 2023. Available from: https://statbel.fgov.be/fr/themes/menages/revenus-fiscaux.

260. Statbel. *Emploi et chômage*; 2024. Available from; https://statbel.fgov.be/fr/themes/emploi-formation/marche-du-travail/emploi-et-chomage.

261. Statbel. *Niveau d'instruction*; 2024. Available from: https://statbel.fgov.be/fr/themes/census/education/niveau-dinstruction.

262. Gomila R. Logistic or linear? Estimating causal effects of experimental treatments on binary outcomes using regression analysis. J Exp Psychol Gen. 2021 Apr;150(4):700–9.

263. Hellevik O. Linear versus logistic regression when the dependent variable is a dichotomy. Qual Quant. 2009 Jan 1;43(1):59–74.

264. Loh WW, Moerkerke B, Loeys T, Vansteelandt S. Disentangling indirect effects through multiple mediators without assuming any causal structure among the mediators. Psychol Methods. 2022;27:982–99.

265. Rosseel Y, Jorgensen TD, Rockwood N, Oberski D, Byrnes J, Vanbrabant L, et al. lavaan: Latent Variable Analysis [Internet]. 2023 [cited 2023 Jun 10]. Available from: https://cran.r-project.org/web/packages/lavaan/index.html

266. Dellicour S, Linard C, Van Goethem N, Da Re D, Artois J, Bihin J, et al. Investigating the drivers of the spatio-temporal heterogeneity in COVID-19 hospital incidence—Belgium as a study case. Int J Health Geogr. 2021 Jun 14;20(1):29.

267. Patel AP, Paranjpe MD, Kathiresan NP, Rivas MA, Khera AV. Race, socioeconomic deprivation, and hospitalization for COVID-19 in English participants of a national biobank. Int J Equity Health. 2020 Jul 6;19(1):114.

268. Semenzato L, Botton J, Drouin J, Cuenot F, Dray-Spira R, Weill A, et al. Chronic diseases, health conditions and risk of COVID-19-related hospitalization and in-hospital mortality during the first wave of the

epidemic in France: a cohort study of 66 million people. Lancet Reg Health Eur. 2021 Sep;8:100158.

269. Zhang Y, Khullar D, Wang F, Steel P, Wu Y, Orlander D, et al. Socioeconomic variation in characteristics, outcomes, and healthcare utilization of COVID-19 patients in New York City. PloS One. 2021;16(7):e0255171.

270. Bergman J, Ballin M, Nordström A, Nordström P. Risk factors for COVID-19 diagnosis, hospitalization, and subsequent all-cause mortality in Sweden: a nationwide study. Eur J Epidemiol. 2021 Mar;36(3):287–98.

271. Riou J, Panczak R, Althaus CL, Junker C, Perisa D, Schneider K, et al. Socioeconomic position and the COVID-19 care cascade from testing to mortality in Switzerland: a population-based analysis. Lancet Public Health. 2021 Jul 10;6(9):e683–91.

272. Yoshikawa M, Asaba K. Educational Attainment Decreases the Risk of COVID-19 Severity in the European Population: A Two-Sample Mendelian Randomization Study. Front Public Health. 2021 Jun 3;9:673451.

273. Ye X, Wang Y, Zou Y, Tu J, Tang W, Yu R, et al. Associations of socioeconomic status with infectious diseases mediated by lifestyle, environmental

pollution and chronic comorbidities: a comprehensive evaluation based on UK Biobank. Infect Dis Poverty. 2023 Jan 30;12(1):5.

274. Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. Lancet Lond Engl. 2012 Jul 7;380(9836):37–43.

275. Gram MA, Emborg HD, Schelde AB, Friis NU, Nielsen KF, Moustsen-Helms IR, et al. Vaccine effectiveness against SARS-CoV-2 infection or COVID-19 hospitalization with the Alpha, Delta, or Omicron SARS-CoV-2 variant: A nationwide Danish cohort study. PLoS Med. 2022 Sep;19(9):e1003992.

276. Gram MA, Moustsen-Helms IR, Valentiner-Branth P, Emborg HD. Sociodemographic differences in Covid-19 vaccine uptake in Denmark: a nationwide register-based cohort study. BMC Public Health. 2023 Feb 24;23(1):391.

277. Fontán-Vela M, Kissling E, Nicolay N, Braeye T, Van Evercooren I, Holm Hansen C, et al. Relative vaccine effectiveness against COVID-19 hospitalisation in persons aged ≥ 65 years: results from a VEBIS network, Europe, October 2021 to July 2023. Euro Surveill Bull Eur Sur Mal Transm Eur Commun Dis Bull. 2024 Jan;29(1):2300670.

278. Chen Y, Zhang L, Li T, Li L. Amplified effect of social vulnerability on health inequality regarding COVID-19 mortality in the USA: the mediating role of vaccination allocation. BMC Public Health. 2022 Nov 19;22(1):2131.

279. Langlete P, Tesli M, Veneti L, Starrfelt J, Elstrøm P, Meijerink H. Estimated vaccine effectiveness against SARS-CoV-2 Delta and Omicron infections among health care workers and the general adult population in Norway, August 2021 – January 2022. Vaccine. 2023 Jun 13;41(26):3923–9.

280. Tamandjou Tchuem CR, Auvigne V, Vaux S, Montagnat C, Paireau J, Monnier Besnard S, et al. Vaccine effectiveness and duration of protection of COVID-19 mRNA vaccines against Delta and Omicron BA.1 symptomatic and severe COVID-19 outcomes in adults aged 50 years and over in France. Vaccine. 2023 Mar 24;41(13):2280–8.

281. Paul P, El-Naas A, Hamad O, Salameh MA, Mhaimeed N, Laswi I, et al. Effectiveness of the pre-Omicron COVID-19 vaccines against Omicron in reducing infection, hospitalization, severity, and mortality compared to Delta and other variants: A systematic review. Hum Vaccines Immunother. 19(1):2167410.

282. Suarez Castillo M, Khaoua H, Courtejoie N. Vaccine-induced and naturally-acquired protection against Omicron and Delta symptomatic infection and

severe COVID-19 outcomes, France, December 2021 to January 2022. Eurosurveillance. 2022 Apr 21;27(16):2200250.

283. de Gier B, Huiberts AJ, Hoeve CE, den Hartog G, van Werkhoven H, van Binnendijk R, et al. Effects of COVID-19 vaccination and previous infection on Omicron SARS-CoV-2 infection and relation with serology. Nat Commun. 2023 Aug 9;14(1):4793.

284. Cucinotta D, Vanelli M. WHO Declares COVID-19 a Pandemic. Acta Bio-Medica Atenei Parm. 2020 Mar 19;91(1):157–60.

285. Wang H, Paulson KR, Pease SA, Watson S, Comfort H, Zheng P, et al. Estimating excess mortality due to the COVID-19 pandemic: a systematic analysis of COVID-19-related mortality, 2020–21. The Lancet. 2022 Apr 16;399(10334):1513–36.

286. Statbel. *Décès*; 2023. Available from: https://statbel.fgov.be/fr/themes/population/mouvement-de-la-population/deces.

287. Pijls BG, Jolani S, Atherley A, Derckx RT, Dijkstra JIR, Franssen GHL, et al. Demographic risk factors for COVID-19 infection, severity, ICU admission

and death: a meta-analysis of 59 studies. BMJ Open. 2021 Jan 1;11(1):e044640.

288. Booth A, Reed AB, Ponzo S, Yassaee A, Aral M, Plans D, et al. Population risk factors for severe disease and mortality in COVID-19: A global systematic review and meta-analysis. PloS One. 2021;16(3):e0247461.

289. Drefahl S, Wallace M, Mussino E, Aradhya S, Kolk M, Brandén M, et al. A population-based cohort study of socio-demographic risk factors for COVID-19 deaths in Sweden. Nat Commun. 2020 Oct 9;11(1):5097.

290. Williamson EJ, Walker AJ, Bhaskaran K, Bacon S, Bates C, Morton CE, et al. Factors associated with COVID-19-related death using OpenSAFELY. Nature. 2020 Aug;584(7821):430–6.

291. Ribeiro KB, Ribeiro AF, Veras MA de SM, de Castro MC. Social inequalities and COVID-19 mortality in the city of São Paulo, Brazil. Int J Epidemiol. 2021 Jul 9;50(3):732–42.

292. Nafilyan V, Islam N, Mathur R, Ayoubkhani D, Banerjee A, Glickman M, et al. Ethnic differences in COVID-19 mortality during the first two waves of the Coronavirus Pandemic: a nationwide cohort study of 29 million adults in England. Eur J Epidemiol. 2021 Jun;36(6):605–17.

293. Seligman B, Ferranna M, Bloom DE. Social determinants of mortality from COVID-19: A simulation study using NHANES. PLoS Med. 2021 Jan;18(1):e1003490.

294. Krieger N, Waterman PD, Chen JT. COVID-19 and Overall Mortality Inequities in the Surge in Death Rates by Zip Code Characteristics: Massachusetts, January 1 to May 19, 2020. Am J Public Health. 2020 Dec;110(12):1850–2.

295. Vanthomme K, Gadeyne S, Devleesschauwer B, Van den Borre L. Excess mortality among native Belgians and migrant groups in Belgium during the first three COVID-19 waves: the evolving dynamics of social inequalities. J Public Health (Berl.) 2023 Dec 28.

296. Klein SL, Flanagan KL. Sex differences in immune responses. Nat Rev Immunol. 2016 Oct;16(10):626–38.

297. Davies B, Parkes BL, Bennett J, Fecht D, Blangiardo M, Ezzati M, et al. Community factors and excess mortality in first wave of the COVID-19 pandemic in England. Nat Commun. 2021 Jun 18;12:3755.

298. Wong LSY, Loo EXL, Kang AYH, Lau HX, Tambyah PA, Tham EH. Age-Related Differences in Immunological Responses to SARS-CoV-2. J Allergy Clin Immunol Pract. 2020;8(10):3251–8.

299. Balaj M, Henson CA, Aronsson A, Aravkin A, Beck K, Degail C, et al. Effects of education on adult mortality: a global systematic review and meta-analysis. Lancet Public Health. 2024 Mar 1;9(3):e155–65.

300. Stoltzfus JC. Logistic Regression: A Brief Primer. Acad Emerg Med. 2011 Oct;18(10):1099–104.

301. Hothorn T, Seibold H, Zeileis A. partykit: A Toolkit for Recursive Partytioning; 2023. Available from: https://cran.r-project.org/web/packages/partykit/index.html.

302. Gulliford MC, Prevost AT, Clegg A, Rezel-Potts E. Mortality of Care Home Residents and Community-Dwelling Controls During the COVID-19 Pandemic in 2020: Matched Cohort Study. J Am Med Dir Assoc. 2022 Jun;23(6):923-929.e2.

303. Hardy OJ, Dubourg D, Bourguignon M, Dellicour S, Eggerickx T, Gilbert M, et al. A world apart: Levels and determinants of excess mortality due to COVID-19 in care homes: The case of the Belgian region of Wallonia during the spring 2020 wave. Demogr Res. 2021 Nov 4;45:1011–40.

304. Ali K, Rao S, Berdine G, Test V, Nugent K. A Retrospective Analysis and Comparison of Prisoners and Community-Based Patients with COVID-19

Requiring Intensive Care During the First Phase of the Pandemic in West Texas. J Prim Care Community Health. 2020 Aug 29;11:2150132720954687.

305. Akner G. Analysis of multimorbidity in individual elderly nursing home residents. Development of a multimorbidity matrix. Arch Gerontol Geriatr. 2009 Nov 1;49(3):413–9.

306. Zhang X, Dou Q, Zhang W, Wang C, Xie X, Yang Y, et al. Frailty as a Predictor of All-Cause Mortality Among Older Nursing Home Residents: A Systematic Review and Meta-analysis. J Am Med Dir Assoc. 2019 Jun 1;20(6):657-663.e4.

307. Adams C, Chamberlain A, Wang Y, Hazell M, Shah S, Holland DP, et al. The Role of Staff in Transmission of SARS-CoV-2 in Long-term Care Facilities. Epidemiology. 2022 Sep;33(5):669.

308. Sugie NF, Turney K, Reiter K, Tublitz R, Kaiser D, Goodsell R, et al. Excess mortality in U.S. prisons during the COVID-19 pandemic. Sci Adv. 2023 Dec;9(48):eadj8104.

309. Jamart H, Van Maele L, Ferguson M, Drielsma P, Macq J, Van Durme T. La première vague de Covid-19 en Belgique et les soins primaires. Rev Med Suisse. 2020 Nov 4;713:2119–22.

310. Dessie ZG, Zewotir T. Mortality-related risk factors of COVID-19: a systematic review and meta-analysis of 42 studies and 423,117 patients. BMC Infect Dis. 2021 Aug 21;21(1):855.

311. GBD 2013 Mortality and Causes of Death Collaborators. *Global, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013*. Lancet. 2015 Jan 10;385(9963):117-71.

312. Farshbafnadi M, Kamali Zonouzi S, Sabahi M, Dolatshahi M, Aarabi MH. Aging & COVID-19 susceptibility, disease severity, and clinical outcomes: The role of entangled risk factors. Exp Gerontol. 2021 Oct 15;154:111507.

313. Griffith DM, Sharma G, Holliday CS, Enyia OK, Valliere M, Semlow AR, Stewart EC, Blumenthal RS. *Men and COVID-19: A Biopsychosocial Approach to Understanding Sex Differences in Mortality and Recommendations for Practice and Policy Interventions*. Prev Chronic Dis. 2020 Jul 16;17:E63.

314. Arceo-Gomez EO, Campos-Vazquez RM, Esquivel G, Alcaraz E, Martinez LA, Lopez NG. The income gradient in COVID-19 mortality and hospitalisation: An observational study with social security administrative records in Mexico. Lancet Reg Health Am. 2022 Feb;6:100115.

315. Bello P, Rocco L. Education and COVID-19 excess mortality. Econ Hum Biol. 2022 Dec;47:101194.

316. Di Girolamo C, Bartolini L, Caranci N, Moro ML. *Socioeconomic inequalities in overall and COVID-19 mortality during the first outbreak peak in Emilia-Romagna Region (Northern Italy).* Epidemiol Prev. 2020 Sep-Dec;44(5-6 Suppl 2):288-296.

317. Dukhovnov D, Barbieri M. County-level socio-economic disparities in COVID-19 mortality in the USA. Int J Epidemiol. 2022 May 9;51(2):418–28.

318. Feng Z. Spatiotemporal pattern of COVID-19 mortality and its relationship with socioeconomic and environmental factors in England. Spat Spatio-Temporal Epidemiol. 2023 Jun;45:100579.

319. Hawkins D. Social Determinants of COVID-19 in Massachusetts, United States: An Ecological Study. J Prev Med Pub Health. 2020 Jul;53(4):220–7.

320. van Ingen T, Akingbola S, Brown KA, Daneman N, Buchan SA, Smith BT. *Neighbourhood-level risk factors of COVID-19 incidence and mortality.* medRxiv. 2021 Jan 1;2021.01.27.21250618.

321. Alam MF, Wildman J, Rahim HA. Income inequality and its association with COVID-19 cases and deaths: a cross-country analysis in the Eastern Mediterranean region. BMJ Glob Health. 2023 Nov 1;8(11):e012271.

322. Wouterse B, Geisler J, Bär M, van Doorslaer E. Has COVID-19 increased inequality in mortality by income in the Netherlands? J Epidemiol Community Health. 2023 Apr;77(4):244–51.

323. Albani V, Welsh CE, Brown H, Matthews FE, Bambra C. Explaining the deprivation gap in COVID-19 mortality rates: A decomposition analysis of geographical inequalities in England. Soc Sci Med 1982. 2022 Oct;311:115319.

324. Mattos dos Santos R. Isolation, social stress, low socioeconomic status and its relationship to immune response in Covid-19 pandemic context. Brain Behav Immun - Health. 2020 Aug 1;7:100103.

325. Bonde JPE, Begtrup LM, Jensen JH, Flachs EM, Schlünssen V, Kolstad HA, et al. Occupational risk of SARS-CoV-2 infection: a nationwide register-based study of the Danish workforce during the COVID-19 pandemic, 2020–2021. Occup Environ Med. 2023 Apr;80(4):202–8.

326. Chen YH, Riley AR, Duchowny KA, Aschmann HE, Chen R, Kiang MV, et al. COVID-19 mortality and excess mortality among working-age residents in California, USA, by occupational sector: a longitudinal cohort analysis of mortality surveillance data. Lancet Public Health. 2022 Sep;7(9):e744–53.

327. Capasso A, Kim S, Ali SH, Jones AM, DiClemente RJ, Tozan Y. Employment conditions as barriers to the adoption of COVID-19 mitigation measures: how the COVID-19 pandemic may be deepening health disparities among low-income earners and essential workers in the United States. BMC Public Health. 2022 May 2;22(1):870.

328. Anson J. The Migrant Mortality Advantage: A 70 Month Follow-up of the Brussels Population. Eur J Popul Rev Eur Démographie. 2004 Sep 1;20(3):191–218.

329. Darmon N, Khlat M. An overview of the health status of migrants in France, in relation to their dietary practices. Public Health Nutr. 2001 Apr;4(2):163–72.

330. Deboosere P, Gadeyne S. La sous-mortalité des immigrés adultes en Belgique: une réalité attestée par les recensements et les registres. Popul Fr Ed. 2005;60(5/6):765–811.

331. Ayoubkhani D, Nafilyan V, White C, Goldblatt P, Gaughan C, Blackwell L, et al. Ethnic-minority groups in England and Wales-factors associated with the size and timing of elevated COVID-19 mortality: a retrospective cohort study linking census and death records. Int J Epidemiol. 2021 Jan 23;49(6):1951–62.

332. Mackey K, Ayers CK, Kondo KK, Saha S, Advani SM, Young S, et al. Racial and Ethnic Disparities in COVID-19-Related Infections, Hospitalizations, and Deaths : A Systematic Review. Ann Intern Med. 2021 Mar;174(3):362–73.

333. Deboosere P, Gadeyne S. Adult Migrant Mortality Advantage in Belgium: Evidence Using Census and Register Data. Population. 2005;60(5):655–98.

334. Hayward SE, Deal A, Cheng C, Crawshaw A, Orcutt M, Vandrevala TF, et al. *Clinical outcomes and risk factors for COVID-19 among migrant populations in high-income countries: A systematic review*. J Migr Health. 2021 Jan 1;3:100041.

335. Kjøllesdal M, Skyrud K, Gele A, Arnesen T, Kløvstad H, Diaz E, et al. The correlation between socioeconomic factors and COVID-19 among immigrants in Norway: a register-based study. Scand J Public Health. 2022 Feb 1;50(1):52–60.

336. Jaljaa A, Caminada S, Tosti ME, D'Angelo F, Angelozzi A, Isonne C, et al. Risk of SARS-CoV-2 infection in migrants and ethnic minorities compared with the general population in the European WHO region during the first year of the pandemic: a systematic review. BMC Public Health. 2022 Jan 20;22:143.

337. Kang SJ, Hyung JA, Han HR. Health literacy and health care experiences of migrant workers during the COVID-19 pandemic: a qualitative study. BMC Public Health. 2022 Nov 9;22(1):2053.

338. Knights F, Carter J, Deal A, Crawshaw AF, Hayward SE, Jones L, et al. Impact of COVID-19 on migrants' access to primary care and implications for vaccine roll-out: a national qualitative study. Br J Gen Pract. 2021 Aug 1;71(709):e583–95.

339. Gil-González D, Carrasco-Portiño M, Vives-Cases C, Agudelo-Suárez AA, Castejón Bolea R, Ronda-Pérez E. Is health a right for all? An umbrella review of the barriers to health care access faced by migrants. Ethn Health. 2015;20(5):523-41.

340. OCDE/Observatoire européen des systèmes et des politiques de santé (2022), *Belgique : Profils de santé par pays 2021*, Éditions OCDE, Paris. Available from: https://doi.org/10.1787/17c34ec7-fr.

341. Centers for disease control and prevention (CDC). Symptoms of COVID-19; 2024. Available from: https://www.cdc.gov/covid/signs-symptoms/index.html.

342. Berlin DA, Gulick RM, Martinez FJ. Severe Covid-19. N Engl J Med. 2020 Dec 17;383(25):2451–60.

343. Qiu H, Li J, Li J, Li H, Xin Y. COVID-19 and Acute Cardiac Injury: Clinical Manifestations, Biomarkers, Mechanisms, Diagnosis, and Treatment. Curr Cardiol Rep. 2023 Aug;25(8):817–29.

344. Tazarghi A, Bazoq S, Taziki Balajelini MH, Ebrahimi M, Hosseini SM, Razavi Nikoo H. Liver injury in COVID-19: an insight into pathobiology and roles of risk factors. Virol J. 2024 Mar 15;21(1):65.

345. Sabaghian T, Kharazmi AB, Ansari A, Omidi F, Kazemi SN, Hajikhani B, et al. COVID-19 and Acute Kidney Injury: A Systematic Review. Front Med. 2022 Apr 4;9:705908.

346. Xie Y, Xu E, Bowe B, Al-Aly Z. Long-term cardiovascular outcomes of COVID-19. Nat Med. 2022 Feb 7;1–8.

347. Lim JT, Liang En W, Tay AT, Pang D, Chiew CJ, Ong B, et al. Long-term Cardiovascular, Cerebrovascular, and Other Thrombotic Complications in

COVID-19 Survivors: A Retrospective Cohort Study. Clin Infect Dis. 2024 Jan 15;78(1):70–9.

348. Huang C, Huang L, Wang Y, Li X, Ren L, Gu X, et al. 6-month consequences of COVID-19 in patients discharged from hospital: a cohort study. Lancet Lond Engl. 2023;401(10393):e21–33.

349. Wong YJ, Tan M, Zheng Q, Li JW, Kumar R, Fock KM, et al. A systematic review and meta-analysis of the COVID-19 associated liver injury. Ann Hepatol. 2020;19(6):627–34.

350. Silver SA, Beaubien-Souligny W, Shah PS, Harel S, Blum D, Kishibe T, et al. The Prevalence of Acute Kidney Injury in Patients Hospitalized With COVID-19 Infection: A Systematic Review and Meta-analysis. Kidney Med. 2021 Jan 1;3(1):83-98.e1.

351. Statbel. *Age.* Available from: https://statbel.fgov.be/fr/themes/census/population/age#panel-14.

352. Du P, Li D, Wang A, Shen S, Ma Z, Li X. A Systematic Review and Meta-Analysis of Risk Factors Associated with Severity and Death in COVID-19 Patients. Can J Infect Dis Med Microbiol. 2021;2021(1):6660930.

353. Gustafsson PE, Sebastian MS, Fonseca-Rodriguez O, Connolly AMF. Inequitable impact of infection: social gradients in severe COVID-19 outcomes among all confirmed SARS-CoV-2 cases during the first pandemic wave in Sweden. J Epidemiol Community Health. 2022 Mar 1;76(3):261–7.

354. Soares R de CM, Mattos LR, Raposo LM. Risk Factors for Hospitalization and Mortality due to COVID-19 in Espírito Santo State, Brazil. Am J Trop Med Hyg. 2020 Sep;103(3):1184–90.

355. Cavillot L, Van den Borre L, Vanthomme K, Scohy A, Deboosere P, Devleesschauwer B, et al. Unravelling demographic and socioeconomic patterns of COVID-19 death and other causes of death: results of an individual-level analysis of exhaustive cause of death data in Belgium, 2020. Arch Public Health. 2024 Nov 13;82(1):209.

356. Abel ZDV, Roope LSJ, Duch R, Clarke PM. Access to healthcare services during the COVID-19 pandemic: a cross-sectional analysis of income and user-access across 16 economically diverse countries. BMC Public Health. 2024 Oct 1;24(1):2678.

357. Zajichek AM, Grunkemeier GL. Propensity scores used as overlap weights provide exact covariate balance. Eur J Cardiothorac Surg. 2024 Sep 1;66(3):ezae318.

358. Yoshida K. Standardized mean difference; 2022. Available from: https://cran.r-project.org/web/packages/tableone/vignettes/smd.html.

359. Funk MJ, Westreich D, Wiesen C, Stürmer T, Brookhart MA, Davidian M. Doubly Robust Estimation of Causal Effects. Am J Epidemiol. 2011 Apr 1;173(7):761–7.

360. Shu D, Young JG, Toh S, Wang R. Variance estimation in inverse probability weighted Cox models. Biometrics. 2021 Sep;77(3):1101–17.

361. Zeileis A, Lumley T, Graham N, Koell S. sandwich: Robust Covariance Matrix Estimators; 2024. Available from: https://cran.r-project.org/web/packages/sandwich/index.html.

362. Negreira-Caamaño M, Martínez-Del Río J, Águila-Gordo D, Mateo-Gómez C, Soto-Pérez M, Piqueras-Flores J. Cardiovascular events after COVID-19 hospitalization: long-term follow-up. Rev Esp Cardiol Engl Ed. 2022 Jan 1;75(1):100–2.

363. Ayoubkhani D, Khunti K, Nafilyan V, Maddox T, Humberstone B, Diamond I, et al. Post-covid syndrome in individuals admitted to hospital with covid-19: retrospective cohort study. BMJ. 2021 Mar 31;372:n693.

364. Han X, Fan Y, Alwalid O, Li N, Jia X, Yuan M, et al. Six-month Follow-up Chest CT Findings after Severe COVID-19 Pneumonia. Radiology. 2021 Apr;299(1):E177–86.

365. Huang L, Li X, Gu X, Zhang H, Ren L, Guo L, et al. *Health outcomes in people 2 years after surviving hospitalisation with COVID-19: a longitudinal cohort study*. Lancet Respir Med. 2022 Sep 1;10(9):863–76.

366. Raman B, McCracken C, Cassar MP, Moss AJ, Finnigan L, Samat AHA, et al. *Multiorgan MRI findings after hospitalisation with COVID-19 in the UK (C-MORE): a prospective, multicentre, observational cohort study*. Lancet Respir Med. 2023 Nov 1;11(11):1003–19.

367. Al-Aly Z, Xie Y, Bowe B. High-dimensional characterization of post-acute sequelae of COVID-19. Nature. 2021 Jun;594(7862):259–64.

368. Fabrizi F, Alfieri CM, Cerutti R, Lunghi G, Messa P. COVID-19 and Acute Kidney Injury: A Systematic Review and Meta-Analysis. Pathogens. 2020 Dec 15;9(12):1052.

369. Tan BWL, Tan BWQ, Tan ALM, Schriver ER, Gutiérrez-Sacristán A, Das P, et al. *Long-term kidney function recovery and mortality after COVID-19-*

*associated acute kidney injury: An international multi-centre observational cohort study*. EClinicalMedicine. 2023 Jan;55:101724.

370. Nugent J, Aklilu A, Yamamoto Y, Simonov M, Li F, Biswas A, et al. Assessment of Acute Kidney Injury and Longitudinal Kidney Function After Hospital Discharge Among Patients With and Without COVID-19. JAMA Netw Open. 2021 Mar 1;4(3):e211095.

371. Zhang Y, Zheng L, Liu L, Zhao M, Xiao J, Zhao Q. Liver impairment in COVID-19 patients: A retrospective analysis of 115 cases from a single centre in Wuhan city, China. Liver Int Off J Int Assoc Study Liver. 2020 Sep;40(9):2095–103.

372. Lammers N, Beese F, Hoebel J, Poethko-Müller C, Wachtler B. Social Inequalities in Long-Term Health Effects After COVID-19—A Scoping Review. Int J Public Health. 2024 Feb 7;69:1606739.

373. Thompson EJ, Williams DM, Walker AJ, Mitchell RE, Niedzwiedz CL, Yang TC, et al. Long COVID burden and risk factors in 10 UK longitudinal studies and electronic health records. Nat Commun. 2022 Jun 28;13(1):3528.

374. Whitaker M, Elliott J, Chadeau-Hyam M, Riley S, Darzi A, Cooke G, et al. Persistent COVID-19 symptoms in a community study of 606,434 people in England. Nat Commun. 2022 Apr 12;13(1):1957.

375. Subramanian A, Nirantharakumar K, Hughes S, Myles P, Williams T, Gokhale KM, et al. Symptoms and risk factors for long COVID in non-hospitalized adults. Nat Med. 2022 Aug;28(8):1706–14.

376. Shabnam S, Razieh C, Dambha-Miller H, Yates T, Gillies C, Chudasama YV, et al. Socioeconomic inequalities of Long COVID: a retrospective population-based cohort study in the United Kingdom. J R Soc Med. 2023 Aug;116(8):263–73.

377. Hirschtick JL, Titus AR, Slocum E, Power LE, Hirschtick RE, Elliott MR, et al. Population-Based Estimates of Post-acute Sequelae of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Infection (PASC) Prevalence and Characteristics. Clin Infect Dis Off Publ Infect Dis Soc Am. 2021 Dec 6;73(11):2055–64.

378. Wu Q, Ailshire JA, Crimmins EM. Long COVID and symptom trajectory in a representative sample of Americans in the first year of the pandemic. Sci Rep. 2022 Jul 8;12(1):11647.

379. Westerlind E, Palstam A, Sunnerhagen KS, Persson HC. Patterns and predictors of sick leave after Covid-19 and long Covid in a national Swedish cohort. BMC Public Health. 2021 May 31;21(1):1023.

380. 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 Sep 5;9(2):e002084.

381. Tsampasian V, Elghazaly H, Chattopadhyay R, Debski M, Naing TKP, Garg P, et al. Risk Factors Associated With Post-COVID-19 Condition: A Systematic Review and Meta-analysis. JAMA Intern Med. 2023 Jun 1;183(6):566–80.

382. Tenforde MW, Kim SS, Lindsell CJ, Billig Rose E, Shapiro NI, Files DC, et al. *Symptom Duration and Risk Factors for Delayed Return to Usual Health Among Outpatients with COVID-19 in a Multistate Health Care Systems Network - United States, March-June 2020.* MMWR Morb Mortal Wkly Rep. 2020 Jul 31;69(30):993–8.

383. Galanis P, Spyros K, Siskou O, Konstantakopoulou O, Angelopoulos G, Kaitelidou D. Healthcare services access, use, and barriers among migrants in Europe: a systematic review. medRxiv. 2022 Jan 1;2022.02.24.22271449.

384. Lebano A, Hamed S, Bradby H, Gil-Salmerón A, Durá-Ferrandis E, Garcés-Ferrer J, et al. Migrants' and refugees' health status and healthcare in Europe: a scoping literature review. BMC Public Health. 2020 Jun 30;20(1):1039.

385. Hitchman SC, Fong GT, Zanna MP, Thrasher JF, Chung-Hall J, Siahpush M. Socioeconomic Status and Smokers' Number of Smoking Friends: Findings from the International Tobacco Control (ITC) Four Country Survey. Drug Alcohol Depend. 2014 Oct 1;143:158–66.

386. Banerjee *A. Disparities by Social Determinants of Health: Links Between Long COVID and Cardiovascular Disease.* Can J Cardiol. 2024 Jun;40(6):1123-1134.

387. Fernández-Martínez NF, Ruiz-Montero R, Gómez-Barroso D, Rodríguez-Torronteras A, Lorusso N, Salcedo-Leal I, et al. Socioeconomic differences in COVID-19 infection, hospitalisation and mortality in urban areas in a region in the South of Europe. BMC Public Health. 2022 Dec 12;22(1):2316.

388. Nunes BP, Thumé E, Tomasi E, Duro SMS, Facchini LA. Socioeconomic inequalities in the access to and quality of health care services. Rev Saúde Pública. 2014 Oct;48(6):968–76.

389. Banack HR, Kaufman JS. The “obesity paradox” explained. Epidemiol Camb Mass. 2013 May;24(3):461–2.

390. S. Demarest, R. Charafeddine. Enquête de santé 2023-2024 : Méthodologie. Bruxelles, Belgique : Sciensano ; 2025. Numéro de rapport : D/2025.14.440/39. Available from : www.enquetesante.be.

391. Bambra C. Pandemic inequalities: emerging infectious diseases and health equity. Int J Equity Health. 2022 Jan 14;21(1):6.

392. Farmer P. Social inequalities and emerging infectious diseases. Emerg Infect Dis. 1996;2(4):259–69.

393. Mamelund SE, Dimka J. Social inequalities in infectious diseases. Scand J Public Health. 2021 Nov 1;49(7):675–80.

394. Hubin P, Van den Borre L, Braeye T, Cavillot L, Billuart M, Stouten V, et al. Area and individual level analyses of demographic and socio-economic disparities in COVID-19 vaccination uptake in Belgium. Vaccine X. 2024 Jun 1;18:100496.

395. Blumenshine P, Reingold A, Egerter S, Mockenhaupt R, Braveman P, Marks J. Pandemic Influenza Planning in the United States from a Health Disparities Perspective. Emerg Infect Dis. 2008 May;14(5):709–15.

396. Quinn SC, Kumar S. Health Inequalities and Infectious Disease Epidemics: A Challenge for Global Health Security. Biosecurity Bioterrorism Biodefense Strategy Pract Sci. 2014 Sep 1;12(5):263–73.

397. Lefebvre G, Haddad S, Moncion-Groulx D, Saint-Onge M, Dontigny A. Socioeconomic disparities and concentration of the spread of the COVID-19 pandemic in the province of Quebec, Canada. BMC Public Health. 2023 Jun 6;23(1):1096.

398. Kumar S, Piper K, Galloway DD, Hadler JL, Grefenstette JJ. Is population structure sufficient to generate area-level inequalities in influenza rates? An examination using agent-based models. BMC Public Health. 2015 Sep 23;15(1):947.

399. Alam MdZ. Is Population Density a Risk Factor for Communicable Diseases Like COVID-19? A Case of Bangladesh. Asia Pac J Public Health. 2021 Nov;33(8):949–50.

400. Statbel. *CensusInd - MS_POP_MOVED_RT*. Available from: https://app.luzmo.com/s/6b30b508-1ef8-4153-b378-9623ec5b104a.

401. Statbel. *P - Armoederisico's* 2019. Available from: https://app.luzmo.com/s/6ed4ec42-eaaf-4bfe-bfd1-55579853a128.

402. Statbel. *Densité de la population*; 2024. Available from: https://statbel.fgov.be/fr/themes/population/structure-de-la-population/densite-de-la-population.

403. Aldridge RW, Pineo H, Fragaszy E, Eyre MT, Kovar J, Nguyen V, et al. Household overcrowding and risk of SARS-CoV-2: analysis of the Virus Watch prospective community cohort study in England and Wales. Wellcome Open Res. 2021 Dec 15;6:347.

404. Herath S, Mansour A, Bentley R. Urban density, household overcrowding and the spread of COVID-19 in Australian cities. Health Place. 2024 Sep 1;89:103298.

405. Household crowding. In: WHO Housing and Health Guidelines [Internet]. World Health Organization; 2018 [cited 2025 Apr 16]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK535289/

406. Chen YH, Glymour M, Riley A, Balmes J, Duchowny K, Harrison R, et al. Excess mortality associated with the COVID-19 pandemic among Californians 18–65 years of age, by occupational sector and occupation: March through November 2020. PLOS ONE. 2021 Jun 4;16(6):e0252454.

407. Brigham E, Harris D, Carlsten C, Redlich CA. Occupational health disparities: The pandemic as prism and prod. J Allergy Clin Immunol. 2021 Nov;148(5):1148–50.

408. Ellingjord-Dale M, Kalleberg KT, Istre MS, Nygaard AB, Brunvoll SH, Eggesbø LM, et al. The use of public transport and contraction of SARS-CoV-2 in a large prospective cohort in Norway. BMC Infect Dis. 2022 Mar 14;22(1):252.

409. Liu T, Gong D, Xiao J, Hu J, He G, Rong Z, et al. Cluster infections play important roles in the rapid evolution of COVID-19 transmission: A systematic review. Int J Infect Dis IJID Off Publ Int Soc Infect Dis. 2020 Oct;99:374–80.

410. Cavalcante TF, Barboza W de S, Santos CA dos, Araújo AA de S, Quintans-Júnior LJ, Martins-Filho PR. Network analysis of socioeconomic disparities and public transport in COVID-19 spread: a case study in northeast Brazil. J Infect Dev Ctries. 2024 Dec 31;18(12.1):S305–9.

411. Brough R, Freedman M, Phillips DC. Understanding socioeconomic disparities in travel behavior during the COVID-19 pandemic. J Reg Sci. 2021 Sep;61(4):753–74.

412. Parker K, Minkin R, Bennett J. *Economic Fallout From COVID-19 Continues To Hit Lower-Income Americans the Hardest*. Washington (DC): Pew Research Center; 2020. Available from: https://www.pewresearch.org/social-trends/2020/09/24/economic-fallout-from-covid-19-continues-to-hit-lower-income-americans-the-hardest/.

413. Alhomsi A, Quintero SM, Ponce S, Mendez I, Stewart AL, Napoles AM, et al. Racial/Ethnic Disparities in Financial Hardship During the First Year of the Pandemic. Health Equity. 2023 Aug 30;7(1):453–61.

414. Atlanta (GA): Centers for Disease Control and Prevention (CDC) - National Center for Immunization and Respiratory Diseases (U.S.). Division of Viral Diseases. *Health equity considerations and racial and ethnic minority groups*; 2020. Available from: https://stacks.cdc.gov/view/cdc/91049.

415. Tsai J, Wilson M. COVID-19: a potential public health problem for homeless populations. Lancet Public Health. 2020 Apr;5(4):e186–7.

416. Boobis S, Albanese F. *The impact of COVID-19 on people facing homelessness and service provision across Great Britain*. London: Crisis; 2020. Available from: https://www.crisis.org.uk/ending-homelessness/homelessness-knowledge-hub/services-and-interventions/the-impact-of-covid-19-on-

people-facing-homelessness-and-service-provision-across-great-britain-2020/.

417. Commission on Social Determinants of Health. *Closing the gap in a generation : health equity through action on the social determinants of health : final report of the commission on social determinants of health.* Combler Fossé En Une Génér Instaur Léquité En Santé En Agissant Sur Déterm Sociaux Santé Rapp Final Comm Déterm Sociaux Santé. 2008;247.

418. Feinstein JS. The relationship between socioeconomic status and health: a review of the literature. Milbank Q. 1993;71(2):279–322.

419. World Health Organization (WHO) - Geneva. *Hidden in plain sight: the increasing burden of noncommunicable diseases among refugees and migrants*; 2023. Available from: https://www.who.int/news-room/events/detail/2023/03/08/default-calendar/hidden-in-plain-sight--the-increasing-burden-of-noncommunicable-diseases-among-refugees-and-migrants.

420. Sciensano. *Inequalities in non-communicable diseases*. Brussels: For a Healthy Belgium; 2022. Available from: https://www.healthybelgium.be/en/health-status/health-inequalities/inequalities-in-non-communicable-diseases.

421. Racape J, Noel AC, Lurel J, Dauby N, Coppieters Y, Goffard JC, et al. Social and clinical risk factors associated with hospitalized COVID-19 patients in Brussels's deprived and multiethnic areas. PLOS Glob Public Health. 2023 Jul 14;3(7):e0002039.

422. Ganaza-Domingues KLT, Ramos-Milaré ÁCFH, Lera-Nonose DSSL, Brustolin AÁ, de Oliveira LF, Rosa JS, et al. Effect of Comorbidities on the Mortality of Patients With COVID-19: A Systematic Review of Reviews and Meta-Analyses. Rev Med Virol. 2025 Mar;35(2):e70024.

423. Chenchula S, Vidyasagar K, Pathan S, Sharma S, Chavan MR, Bhagavathula AS, et al. Global prevalence and effect of comorbidities and smoking status on severity and mortality of COVID-19 in association with age and gender: a systematic review, meta-analysis and meta-regression. Sci Rep. 2023 Apr 19;13(1):6415.

424. Vanthomme K, Vandenheede H. Migrant mortality differences in the 2000s in Belgium: interaction with gender and the role of socioeconomic position. Int J Equity Health. 2019 Jun 20;18(1):96.

425. Rostila M, Cederström A, Wallace M, Brandén M, Malmberg B, Andersson G. Disparities in Coronavirus Disease 2019 Mortality by Country of Birth in

Stockholm, Sweden: A Total-Population-Based Cohort Study. Am J Epidemiol. 2021 Aug 1;190(8):1510–8.

426. Braveman, Paula, 'Stress Mediates the Health Effects of Many Social Determinants', *The Social Determinants of Health and Health Disparities* (New York, 2023; online edn, Oxford Academic, 23 Feb. 2023). Available from https://doi.org/10.1093/oso/9780190624118.003.0004.

427. Alotiby A. Immunology of Stress: A Review Article. J Clin Med. 2024 Oct 25;13(21):6394.

428. Morey JN, Boggero IA, Scott AB, Segerstrom SC. Current Directions in Stress and Human Immune Function. Curr Opin Psychol. 2015 Oct 1;5:13–7.

429. Bernardini F, Attademo L, Rotter M, Compton MT. *Social Determinants of Mental Health As Mediators and Moderators of the Mental Health Impacts of the COVID-19 Pandemic*. Psychiatr Serv. 2021 May 1;72(5):598-601.

430. Masterson JM, Luu M, Dallas KB, Daskivich LP, Spiegel B, Daskivich TJ. *Disparities in COVID-19 Disease Incidence by Income and Vaccination Coverage — 81 Communities, Los Angeles, California, July 2020–September 2021. MMWR Morb Mortal Wkly Rep*. 2023 Jun 30;72(26):728–731.

431. Manna A, Karsai M, Perra N. *Social inequalities in vaccine coverage and their effects on epidemic spreading.* medRxiv. 2024 Jan 1;2024.11.01.24316556.

432. Roel E, Raventós B, Burn E, Pistillo A, Prieto-Alhambra D, Duarte-Salles T. Socioeconomic Inequalities in COVID-19 Vaccination and Infection in Adults, Catalonia, Spain. Emerg Infect Dis. 2022 Nov;28(11):2243–52.

433. Goto R, Kawachi I, Kondo N, Inoue K. Contribution of vaccinations to reducing socioeconomic disparities in COVID-19 deaths across U.S. counties. Ann Epidemiol. 2023 Oct 1;86:65-71.e3.

434. Marques MD, Sheel M. Addressing the social inequities of vaccination: an imperative to close the gap. Lancet Glob Health. 2023 Feb 1;11(2):e173–4.

435. Bergen N, Kirkby K, Fuertes CV, Schlotheuber A, Menning L, Feely SM, et al. Global state of education-related inequality in COVID-19 vaccine coverage, structural barriers, vaccine hesitancy, and vaccine refusal: findings from the Global COVID-19 Trends and Impact Survey. Lancet Glob Health. 2023 Feb 1;11(2):e207–17.

436. MacDonald NE. Vaccine hesitancy: Definition, scope and determinants. Vaccine. 2015 Aug 14;33(34):4161–4.

437. Njoku A, Joseph M, Felix R. Changing the Narrative: Structural Barriers and Racial and Ethnic Inequities in COVID-19 Vaccination. Int J Environ Res Public Health. 2021 Sep 20;18(18):9904.

438. Thompson S, Meyer JC, Burnett RJ, Campbell SM. Mitigating Vaccine Hesitancy and Building Trust to Prevent Future Measles Outbreaks in England. Vaccines. 2023 Feb;11(2):288.

439. Geoffard PY, Jusot F, Sireyjol A, Warszawski J, Bajos N. Socio-economic inequalities in access to COVID-19 tests in France in 2020: evidence from the EPICOV socio-epidemiological cohort. Front Public Health. 2024;12:1434370.

440. Vandentorren S, Smaïli S, Chatignoux E, Maurel M, Alleaume C, Neufcourt L, et al. The effect of social deprivation on the dynamic of SARS-CoV-2 infection in France: a population-based analysis. Lancet Public Health. 2022 Mar 1;7(3):e240–9.

441. Shelton C, El-Boghdadly K, Appleby JB. The “haves” and “have-nots” of personal protective equipment during the COVID-19 pandemic: the ethics of emerging inequalities amongst healthcare workers. J Med Ethics. 2021 Dec 17;medethics-2021-107501.

442. Laboratory IS. Across The Spectrum of Socioeconomics: Issue IV. International Socioeconomics Laboratory; 2021. 291 p.

443. Dessu S. Assessment of Health Care Seeking Behavior among House Hold Heads in Dale Woreda, Sidama Zone, Southern Ethiopia, Ethiopia. Glob J Med Res. 2018 Jan 15;18(F1):19–29.

444. Biggerstaff M, Jhung MA, Reed C, Fry AM, Balluz L, Finelli L. Influenza-like illness, the time to seek healthcare, and influenza antiviral receipt during the 2010-2011 influenza season-United States. J Infect Dis. 2014 Aug 15;210(4):535–44.

445. Cès S, Baeten R. *Inequalities in access to healthcare in Belgium*. Brussels: European Social Observatory; 2020. Available from: https://www.ose.be/publication/inequalities-access-healthcare-belgium.

446. Sciensano. *Équité et inégalités de santé. Vers une Belgique en bonne santé;* 2024. Available from: https://www.belgiqueenbonnesante.be/fr/hspa/equite-et-inegalites-de-sante.

447. Gerkens S, Merkur S. Belgium: Health system review. Health Syst Transit. 2010;12(5):1–266, xxv.

448. Peacock HM, Tambuyzer T, Verdoodt F, Calay F, Poirel HA, De Schutter H, et al. Decline and incomplete recovery in cancer diagnoses during the COVID-19 pandemic in Belgium: a year-long, population-level analysis. ESMO Open. 2021 Aug;6(4):100197.

449. Pujolar G, Oliver-Anglès A, Vargas I, Vázquez ML. Changes in Access to Health Services during the COVID-19 Pandemic: A Scoping Review. Int J Environ Res Public Health. 2022 Feb 3;19(3):1749.

450. Baker E. *First Do No Harm: The Impact of Telemedicine on Health Disparities. Telehealth Med Today.* 2024. Available from: https://telehealthandmedicinetoday.com/index.php/journal/article/view/493.

451. European Centre for Disease Prevention and Control. Does COVID-19 vaccination reduce the risk and duration of post-COVID-19 condition? Rapid systematic literature review. Stockholm: ECDC: 2025.

452. Collard D, Stronks K, Harris V, Coyer L, Brinkman K, Beudel M, et al. Ethnic Differences in Coronavirus Disease 2019 Hospitalization and Hospital Outcomes in a Multiethnic Population in the Netherlands. Open Forum Infect Dis. 2022 Jun;9(6):ofac257.

453. Bruyneel A, Dauvergne JE, Dauby N, Goffard JC, Rea A, Racape J. Social health gradient and risk factors among patients hospitalized for COVID-19 and pre-pandemic respiratory infections. A linked national individual case-control study in Belgium. Front Public Health. 2024 Oct 28;12:1426898.

454. European Centre for Disease Prevention and Control. Reducing COVID-19 transmission and strengthening vaccine uptake among migrant populations in the EU/EEA – 3 June 2021. ECDC: Stockholm; 2021.

455. Costa R, de Valk HAG. Ethnic and Socioeconomic Segregation in Belgium: A Multiscalar Approach Using Individualised Neighbourhoods. Eur J Popul Rev Eur Démographie. 2018 Apr 6;34(2):225–50.

456. Institut bruxellois de statistiques et d'analyse. Population. Available from: https://ibsa.brussels/le-saviez-vous/186-nationalites-sont-representees-en-region-bruxelloise.

457. European migration network. 112,000 people without residence documents live in Belgium, according to new research from Vrije Universiteit Brussel; 2023. Available from: https://emnbelgium.be/news/112000-people-without-residence-documents-live-belgium-according-new-research-vrije.

458. Stafford A, Wood L. Tackling Health Disparities for People Who Are Homeless? Start with Social Determinants. Int J Environ Res Public Health. 2017 Dec;14(12):1535.

459. Genovese E, Page KR, Cailhol J, Jackson Y. Learning from the COVID-19 pandemic response to strengthen undocumented migrant-sensitive health systems: case studies from four countries. Lancet Reg Health Eur. 2023 Apr;27:100601.

460. Ogbonna O, Bull F, Spinks B, Williams D, Lewis R, Edwards A. The Impact of Being Homeless on the Clinical Outcomes of COVID-19: Systematic Review. Int J Public Health. 2023 Sep 15;68:1605893.

461. Kayman H, Ablorh-Odjidja A. Revisiting Public Health Preparedness: Incorporating Social Justice Principles Into Pandemic Preparedness Planning for Influenza. J Public Health Manag Pract. 2006 Aug;12(4):373.

462. Uscher-Pines L, Duggan PS, Garron JP, Karron RA, Faden RR. Planning for an Influenza Pandemic: Social Justice and Disadvantaged Groups. Hastings Cent Rep. 2007;37(4):32–9.

463. Taskforce opérationnalisation de la stratégie de vaccination. *Avis pour l'opérationnalisation de la Stratégie de vaccination COVID-19 pour la*

*Belgique*. Bruxelles: Commissariat Corona du Gouvernement; 2020. Available from: https://www.le-gbo.be/wp-content/uploads/2021/01/Note_TF_Strategy_Vaccination_FR_0312_post_press.pdf.

464. Marmot M, Bell R. Fair society, healthy lives. Public Health. 2012 Sep;126 Suppl 1:S4–10.

465. Guiheneuf C., Malengreaux S., Aujoulat I., Ferron C., Doumont D. Inscrire les actions dans une approche d'universalisme proportionné. Bruxelles : UCLouvain/IRSS-RESO, mai 2025, 6 p.

466. Hahn RA, Truman BI. Education Improves Public Health and Promotes Health Equity. Int J Health Serv Plan Adm Eval. 2015;45(4):657–78.

467. Khoury G, Ward JK, Mancini J, Gagneux-Brunon A, Nguyen LBL. Health Literacy and Health Care System Confidence as Determinants of Attitudes to Vaccines in France: Representative Cross-Sectional Study. JMIR Public Health Surveill. 2024 May 7;10(1):e45837.

468. Berchet, C., J. Bijlholt and M. Ando (2023), "Socio-economic and ethnic health inequalities in COVID-19 outcomes across OECD countries", *OECD*

*Health Working Papers*, No. 153, OECD Publishing, Paris. Available from: https://doi.org/10.1787/6c2a96c9-en.

469. Lee C, Rogers WA, Braunack-Mayer A. *Social Justice and Pandemic Influenza Planning: The Role of Communication Strategies.* Public Health Ethics. 2008 Nov 1;1(3):223–34.

470. Van den Broucke S. Why health promotion matters to the COVID-19 pandemic, and vice versa. Health Promot Int. 2020 Apr 1;35(2):181–6.

471. Diseases TLI. The COVID-19 infodemic. Lancet Infect Dis. 2020 Aug 1;20(8):875.

472. Molenberghs G, Faes C, Verbeeck J, Deboosere P, Abrams S, Willem L, et al. COVID-19 mortality, excess mortality, deaths per million and infection fatality ratio, Belgium, 9 March 2020 to 28 June 2020. Eurosurveillance. 2022 Feb 17;27(7):2002060.

473. International long-term care policy network. International data on deaths attributed to COVID-19 among people living in care homes; 2022. Available from: https://ltccovid.org/2022/02/22/international-data-on-deaths-attributed-to-covid-19-among-people-living-in-care-homes/.

474. Jurcevic J, Ekelson R, Nganda S, Bustos Sierra N, Vernemmen C. *Epidemiology of COVID-19 mortality in Belgium from wave 1 to wave 7 (March 2020 – 11 September 2022).* Brussels: Sciensano; 2023. Available from: https://www.sciensano.be/en/biblio/epidemiology-covid-19-mortality-belgium-wave-1-wave-7-march-2020-11-september-2022.

475. Lagasse R, Deboosere P. Evaluation épidémiologique de l'impact du COVID-19 en Belgique à la date du 15 juillet 2020, Rapport remis en date du 15 juillet 2020 à la Première Ministre, Madame Sophie Wilmès. Available from: https://esp.ulb.be/fr/les-actus/les-actualites-de-lesp/rapport-d-analyse-de-l-epidemie-covid-19-n-ii.

476. Ladhani SN, Chow JY, Janarthanan R, Fok J, Crawley-Boevey E, Vusirikala A, et al. Increased risk of SARS-CoV-2 infection in staff working across different care homes: enhanced CoVID-19 outbreak investigations in London care Homes. J Infect. 2020 Oct 1;81(4):621–4.

477. Rosello A, Barnard RC, Smith DRM, Evans S, Grimm F, Davies NG, et al. Impact of non-pharmaceutical interventions on SARS-CoV-2 outbreaks in English care homes: a modelling study. BMC Infect Dis. 2022 Apr 1;22(1):324.

478. Médecins sans frontières. Left behind in the time of COVID-19: a report into nursing care homes in Belgium; 2020. Available from: https://www.msf.org/left-behind-coronavirus-belgium-care-homes-report

479. Hoxha A, Wyndham-Thomas C, Klamer S, Dubourg D, Vermeulen M, Hammami N, et al. Asymptomatic SARS-CoV-2 infection in Belgian long-term care facilities. Lancet Infect Dis. 2021 Apr;21(4):e67.

480. RTL info. Coronavirus en Belgique: des milliers de tests vont être faits dans les maisons de repos; 2020. Available from: http://www.rtl.be/actu/coronavirus-en-belgique-des-milliers-de-tests-vont-etre-faits-dans-les-maisons/2020-04-08/article/300327.

481. Flaxman S, Mishra S, Gandy A, Unwin HJT, Mellan TA, Coupland H, et al. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. Nature. 2020 Aug;584(7820):257–61.

482. Sciensano. Vers une Belgique en bonne santé. Alimentation; 2020. Available from: https://www.belgiqueenbonnesante.be/fr/etat-de-sante/determinants-de-sante/alimentation.

483. Sciensano. Vers une Belgique en bonne santé. Activité physique; 2021. Available from: https://www.belgiqueenbonnesante.be/fr/etat-de-sante/determinants-de-sante/activite-physique.

484. Sciensano. Vers une Belgique en bonne santé. Consommation de tabac; 2023. Available from: https://www.belgiqueenbonnesante.be/fr/etat-de-sante/determinants-de-sante/consommation-de-tabac.

485. Intermut vzw. *Community Health Workers*. Bruxelles: Intermut. Available from: https://www.chw-intermut.be/homepage-fr/.

486. Masquillier C, Cosaert T. Facilitating access to primary care for people living in socio-economically vulnerable circumstances in Belgium through community health workers: towards a conceptual model. BMC Prim Care. 2023 Dec 20;24(1):281.

487. Apers H, Masquillier C. Facilitators and Barriers in Collaborations Between Community Health Workers with Primary and Well-Being Providers in Primary Healthcare in Belgium. Healthcare. 2024 Nov 24;12(23):2348.

488. Smith DP, Moreels S. L'épidémiologie de l'affection post-Covid-19 en Belgique. 2023;

489. Institut national d'assurance maladie-invalidité (INAMI). C*OVID long : Remboursement des soins en cas de symptômes COVID-19 persistants*. Available from: https://www.inami.fgov.be/fr/themes/soins-de-sante-cout-et-remboursement/maladies/covid-long-remboursement-des-soins-en-cas-de-symptomes-covid-19-persistants#:~:text=Pour%20pouvoir%20b%C3%A9n%C3%A9ficier%20du%20remboursement,m%C3%A9decin%20g%C3%A9n%C3%A9raliste%20ou%20m%C3%A9decin%20sp%C3%A9cialiste.

490. Institut national d'assurance maladie-invalidité (INAMI). *Evaluation trajet de soins COVID long : Remboursement des soins en cas de symptômes COVID-19 persistants*; 2023. Available from: https://www.inami.fgov.be/fr/actualites/evaluation-trajet-de-soins-covid-long-remboursement-des-soins-en-cas-de-symptomes-covid-19-persistants.

491. Institut national d'assurance maladie-invalidité (INAMI). Soins à distance : suivi des patients COVID-19 à domicile via le télémonitoring. Available from: https://www.inami.fgov.be/fr/theme/soins-a-distance-suivi-des-patients-covid-19-a-domicile-via-le-telemonitoring.

492. Cornelis J, Van Grootven B, Alvarez Irusta L, de Meester C, Christiaens W, van Durme T et al. Télémonitoring des patients COVID-19. Health Services Reserach (HSR). Bruxelles. Centre Fédéral d'Expertise des Soins de Santé (KCE); 2022. KCE Reports 354.

493. Vasileiou E, Simpson CR, Shi T, Kerr S, Agrawal U, Akbari A, et al. Interim findings from first-dose mass COVID-19 vaccination roll-out and COVID-19 hospital admissions in Scotland: a national prospective cohort study. Lancet Lond Engl. 2021 May 1;397(10285):1646–57.

494. Zhang Y, Li J, Feng L, Luo Y, Pang W, Qiu K, et al. A Population-Based Outcome-Wide Association Study of the Comorbidities and Sequelae Following COVID-19 Infection. J Epidemiol Glob Health. 2023 Dec;13(4):870–85.

495. Mc Intyre K, Lanting P, Deelen P, Wiersma HH, Vonk JM, Ori APS, et al. Lifelines COVID-19 cohort: investigating COVID-19 infection and its health and societal impacts in a Dutch population-based cohort. BMJ Open. 2021 Mar 17;11(3):e044474.

496. Rangul V, Sund ER, Ingul JM, Rimehaug T, Pape K, Kvaløy K. Exploring the Link Between Physical Activity, Sports Participation, and Loneliness in

Adolescents Before and Into the COVID-19 Pandemic: The HUNT Study, Norway. Int J Environ Res Public Health. 2024 Oct 25;21(11):1417.

497. Schutte N, Saelaert M, Bogaert P, De Ridder K, Van Oyen H, Van der Heyden J, et al. Opportunities for a population-based cohort in Belgium. Arch Public Health Arch Belg Sante Publique. 2022 Aug 11;80(1):188.

498. Health Data Agency (HAD). *About us*. Available from: https://www.hda.belgium.be/en/about_us.

499. Sciensano. *Couplage de données pour une meilleure surveillance épidémiologique des maladies infectieuses*; 2025. Available from: https://www.sciensano.be/fr/projets/couplage-de-donnees-pour-une-meilleure-surveillance-epidemiologique-des-maladies-infectieuses.

500. Mandelbaum J. Advancing health equity by integrating intersectionality into epidemiological research: applications and challenges. J Epidemiol Community Health. 2020 Sep 1;74(9):761–2.

501. European Commission. European Health Data Space (EHDS) Regulations. Available from: https://www.european-health-data-space.com/.

502. Panteli D, Adib K, Buttigieg S, Goiana-da-Silva F, Ladewig K, Azzopardi-Muscat N, et al. *Artificial intelligence in public health: promises, challenges,*

*and an agenda for policy makers and public health institutions.* Lancet Public Health. 2025 May 1;10(5):e428–32.

503. European Commission. AI Act - Shaping Europe's digital future; 2025. Available from: https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai.