CHARLES UNIVERSITY
FACULTY OF SOCIAL SCIENCES
Institute of Economic Studies

The effect of SSRIs on the severity of COVID-19

Bachelor's thesis

Author: Anna Marie Břicháčková

Study program: Economics and Finance

Supervisor: PhDr. Mgr. Jana Votápková, Ph.D.

Year of defense: 2024

Declaration of Authorship

The author hereby declares that he or she compiled this thesis independently, using only the listed resources and literature, and the thesis has not been used to obtain any other academic title.

While preparing the thesis, the author utilized AI tools to improve the writing and stylistic form required for its completion. Following the use of these tools, the author cautiously reviewed and edited the content and takes full responsibility for the final publication. The AI tools were employed in compliance with principles of academic integrity.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis in whole or in part and agrees with the thesis being used for study and scientific purposes.

Prague, July 23, 2024

Anna Marie Břicháčková

Abstract

Recent medical research suggested that antidepressants, particularly Selective Serotonin Reuptake Inhibitors (SSRIs), might potentially exhibit antiviral properties against COVID-19. The opinions about repurposing antidepressants as a form of COVID-19 treatment vary markedly among scientific and medical professionals, especially when one considers the wide range of side effects that antidepressants may induce. The aim of this thesis is to examine the effect of SSRIs on the severity of COVID-19. Our analysis will specifically target the individuals who use antidepressants actively and regularly, i.e., those who were not prescribed SSRIs intentionally due to COVID-19. To evaluate the impact of SSRIs, we will perform logistic regression and utilize the zero-inflated negative binomial model. The results reveal a significant association between the use of SSRIs and increased probability of both hospitalisation and death due to COVID-19. The effect is rather small, however, we find the effect is statistically significant. Additionally, our analysis discovered no significant evidence that SSRIs affect the length of hospital stay. Our results thereby do not support the hypothesis that SSRIs provide protective effects against COVID-19 or function as a form of long-term preventive antiviral pharmaceuticals.

Keywords antidepressants, COVID-19, SSRIs, hospitalisation, death of COVID-19

Title The effect of SSRIs on the severity of COVID-19

Abstrakt

V průběhu pandemie COVID-19 vznikly na základě odborných vědeckých publikací hypotézy, které naznačují, že by antidepresiva, zejména pak selektivní inhibitory zpětného vychytávání serotoninu (SSRI), mohli vykazovat antivirové účinky. Názory na použití antidepresiv jako jednu z možných forem léčby COVID-19 se mezi vědeckými a lékařskými profesionály výrazně liší. Diverzita názorů je dále podněcována častým výskytem vedlejších účinků spojených s užíváním antidepresiv. Cílem této práce je zanalyzovat vliv SSRI na závažnost COVID-19. Naše analýza se konkrétně zaměří na jedince, kteří antidepresiva užívají aktivně a pravidelně. K posouzení efektu antidepresiv využíváme ekonometrické metody, konkrétně logistickou regresi a zero-inflated negative binomial model. Výsledky analýzy ukazují, že SSRI antidepresiva signifikantně zvyšují pravděpodobnost úmrtí i hospitalizace v důsledku COVID-19. Tento efekt je statisticky významný, avšak jeho velikost je relativně malá. Analýza dále naznačuje, že vliv SSRI na očekávaný počet dnů strávených v nemocnici je statisticky nevýznamný. Naše výsledky tedy neprokazují, že by antidepresiva měla z dlouhodobého hlediska ochranné učinky proti COVID-19.

Klíčová slova antidepresiva, SSRI, COVID-19, hospitalizace, úmrtí

Název práce Vliv SSRI na závažnost Covidu-19

Acknowledgments

The author is grateful especially to PhDr. Mgr. Jana Votápková, Ph.D., whose advice, guidance and support have been invaluable throughout the writing of this thesis. Their expertise and insightful feedback have greatly contributed to the completion of this work.

This output was supported by the NPO "Systemic Risk Institute" number LX22NPO5101, funded by European Union - Next Generation EU (Ministry of Education, Youth and Sports, NPO: EXCELES).

Typeset in LaTeXusing the IES Thesis Template.

Bibliographic Record

Brichackova, Anna Marie: *The effect of SSRIs on the severity of COVID-19*. Bachelor's thesis. Charles University, Faculty of Social Sciences, Institute of Economic Studies, Prague. 2024, pages 81. Advisor: PhDr. Mgr. Jana Votápková, Ph.D.

Contents

List of Tables

List of Figures

Acronyms

ÚZIS Institute of Health Information and Statistics of the Czech Republic

ICD-10 International Classification of Diseases and Related Health Problems, Tenth Revision

WHO World Health Organisation

ATC Anatomical Therapeutic Chemical

NRHZS National Register of Reimbursed Health Services

COVID-19

DDD Daily Defined Dosage

MZČR Ministerstvo zdravotnictví České republiky

CCI Charlson Comorbidity Index

SSRIs Selective Serotonin Reuptake Inhibitors

SSRI Selective Serotonin Reuptake Inhibtor

ER endoplasmic reticulum

GDP Gross domestic product

LR likelihood ratio

AUC area under the receiver operating characteristic curve

df degree-of-freedom

cdf cumulative distribution function

SNRIs Serotonin-Norepinephrine Reuptake Inhibitors

TCAs Tricyclic Antidepressants

FIASMAs Functional Inhibitors of Acid SphingomyelinaseCOVID-19

AME Average Marginal Effect

COPD Chronic Obstructive Pulmonary Disease

MLE Maximum Likelihood Estimation

GVIF Generalised Variance Inflation Factor

LPM Linear Probability Model

IHD Ischemic Heart Disease

AIC Akaike Information Criterion

ZINB Zero-Inflated Negative Binomial Model

ZIP Zero-Inflated Poisson Model

ICU intensive unit care

OCD obsessive-compulsive disorder

Chapter 1

Introduction

The COVID-19 pandemic was labeled as one of the most devastating pandemics of all times, causing more than 7 million officially reported deaths worldwide (Giménez-Llort et al., 2022; WHO, 2021b). The pandemic posed a threat to both the economic and health states of the population. The primary global objective shifted towards rapidly combating the pandemic and consequently mitigating costs to society as much as possible. In the early stages of the pandemic, prior to the vaccine development, there was a need for a temporary remedy that could have protected the most vulnerable groups of individuals. A wide range of pharmaceutical interventions has been proposed, including antiviral drugs (Beigel et al., 2020; Keshta et al., 2021). Additionally, there have been suggestions for unconventional candidates like antidepressants (Nakhaee et al., 2022).

Antidepressants represent an effective pharmacological treatment for one of the most prevalent mental disorders, i.e., the major depressive disorder (WHO, 2022a). The most frequently prescribed types of antidepressants are called Selective Serotonin Reuptake Inhibitors (SSRIs) (Gautam et al., 2017). Due to their specific mechanisms, SSRIs were suggested to prevent COVID-19 from progressing into a more severe form of the disease. From an economic point of view, antidepressants represented a very convenient way for treatment owing to their affordability and wide accessibility (Sukhatme et al., 2021). Yet, the idea of repurposing antidepressants as a form of the COVID-19 treatment remains rather contentious, especially when one considers the wide range of side effects that antidepressants may induce (Kelly et al., 2008).

Numerous studies have already been conducted to analyze the effect of antidepressants on COVID-19, however, they yield no unified conclusion. Some

researchers found a significant association between antidepressants and reduced risk of severe COVID-19 (Hoertel et al., 2021; Fritz et al., 2022; Lenze et al., 2020). On the other hand, others assert that the use of antidepressants either increases the probability of a severe outcome (Stauning et al., 2023; Bliek-Bueno et al., 2021) or shows no significant effect on such outcomes (Rauchman et al., 2022). In the studies, individuals with severe form of COVID-19 are predominantly identified as those requiring hospitalisation, emergency room admission, supplemental oxygen, or who died of COVID-19 (Deng et al., 2023). The diversity of opinions among scientists and medical professionals highlights the need for continued research in this area.

The objective of this thesis is to examine the impact of SSRIs on the severity of COVID-19. The severity will be represented either by the need for hospitalisation or the death due to COVID-19. We will specifically focus on individuals who use the antidepressants actively and regularly. This includes those who were not prescribed SSRIs intentionally due to COVID-19. By doing so, we aim to analyze the long-term protective effects of antidepressants. The study sample consists of individuals who tested COVID-19 positive once during the early stage of the pandemic (01/03/2020-27/12/2020). Logistic regression will be implemented to estimate the probability of hospitalisation (or death) among the individuals with a confirmed positive COVID-19 test. We will account for potential confounding factors, such as chronic diseases of individuals, their demographic characteristics or the specific waves of the pandemic. Numerous models incorporating different independent variables will be created to robustly verify our results. Later, we will use the zero-inflated negative binomial model to estimate the expected days patients spent in hospital during our study period. The binary part of the model will serve as a form of sensitivity analysis for the logistic regressions. Finally, we will recreate the models by replacing SSRIs with the aggregated group of all the antidepressants, to examine their effects irrespective of the specific pharmacological class.

The thesis is structured as follows: Chapter 2 will briefly introduce the theoretical background to our topic. First, we will delve into the field of mental health, including the use of antidepressants for addressing mental health issues. Second, we will introduce the COVID-19 pandemic and the idea of repurposing antidepressants as an antiviral treatment. In Chapter 3, we will interpret results from existing studies that analyzed the effect of antidepressants on COVID-19. At the end of the chapter, our hypothesis will be introduced. Chapter 4 will describe the datasets and variables used in our analysis. We will explain

the process of how our raw medical data were reshaped to more appropriate format. Later, we will describe the characteristics of the final sample. Chapter 5 comprises the description of econometric methods used to test our hypothesis. Chapter 6 will present results stemming from our regressions. We will further discuss the potential rationale behind the results.

Chapter 2

Theoretical background

The aim of this chapter is to introduce a theoretical context associated with the topic of this thesis. Section 2.1 will briefly introduce the field of mental health, including the use of antidepressants for addressing mental health issues. Section 2.2 will present the background on the COVID-19 pandemic including tools used for its treatment. Lastly, Section 2.3 will explain the plausible rationale behind the antiviral mechanism of antidepressants.

2.1 Mental health and antidepressants

In the present dynamic world, the concept of mental health has evolved beyond just the absence of mental illness (WHO, 2022b). It comprises the overall comfort and prosperity of individuals. The World health organisation (WHO) (2022b, para.1) defines mental health as

> "a state of mental well-being that enables people to cope with the stresses of life, realize their abilities, learn well and work well, and contribute to their community."

This definition highlights the obvious relationship between individual resilience and societal functioning. Indeed, scientists suggest that mental health stands as the principal factor in enhancing the productivity of individuals, thus empowering the society as a whole (de Oliveira et al., 2023). The total economic costs associated with mental health problems are estimated to be more than 4% of GDP across the 28 EU countries (OECD & European Union, 2018). In total, it sums up to approximately 600 billion EUR per year. The COVID-19

pandemic certainly did not help to ease the situation, with new cases of anxiety and depression[1] increasing by 25% (Rathod et al., 2023).

Antidepressants represent an effective pharmacological treatment for one of the most prevalent psychiatric diseases, major depressive disorder (WHO, 2022a). It is estimated that around 12.9% of adults worldwide live with depression (Lim et al., 2018). The presence of depression is associated with high rate of both mortality and morbidity as indicated by estimated loss of 28.9 years in quality-adjusted life expectancy (Jia et al., 2015). Besides serving as the main medication for depression, the antidepressants are an effective tool in managing the spectrum of mental health conditions. These include social phobia, obsessive-compulsive disorder (OCD), panic disorder, post-traumatic stress disorder or anxiety disorder (Sheffler et al., 2023). Fluoxetine, a representative of antidepressant drugs, has also proven to be a helpful means in curing eating disorders (Edinoff et al., 2021).

Traditional theory states that people suffering from depression face a lack of certain monoamine neurotransmitters[2] in the brain; specifically serotonin, norepinephrine and dopamine (Cui et al., 2024). Antidepressants work on the basis of increasing the level of these neurotransmitters (Cui et al., 2024). Antidepressants can be divided into several groups, based on the processes through which they affect neurotransmitters: Selective Serotonin Reuptake Inhibitors (SSRIs), Serotonin-Norepinephrine Reuptake Inhibitors (SNRIs), Tricyclic Antidepressants (TCAs), atypical antidepressants, etc. (Sheffler et al., 2023). Doctors prescribing medication to treat major depressive disorder or anxiety tend to first opt for novel SSRIs (Gautam et al., 2017). One of the reasons is their potentially lower frequency of side effects while maintaining the similar efficacy to other antidepressants (Chae et al., 2004).

It is, however, widely acknowledged that no antidepressant is devoid of side effects (Kelly et al., 2008). Ramic et al. (2020) conclude that side effects occurred in up to 23% people taking antidepressants. The most common side effects of SSRIs include fatigue, dizziness, nausea and vomiting, weight loss or contrarily weight gain, headaches and sexual dysfunction (Braund et al., 2021; Wang et al., 2018). SSRIs can bear even more dangerous adverse effects. Warnings have been issued regarding the potential for suicidal thoughts and

[1] Anxiety and depression are the two most common mental health problems globally; described in more detail below (WHO, 2022a).

[2] "Neurotransmitters are endogenous chemicals that allow neurons to communicate with each other throughout the body. They enable the brain to provide a variety of functions, through the process of chemical synaptic transmission" (Sheffler et al., 2023, para. 1).

consequently increased risk of suicidality, especially for adults and adolescents under 24 years (Edinoff et al., 2021; Braund et al., 2021). Although SSRIs were previously believed to be a safer option when considering heart function than older antidepressant classes (TCAs, etc.), it has been recently suggested that cardiovascular difficulties, indeed, occur in SSRI users as well (Wang et al., 2018).

2.2 COVID-19 pandemic

The COVID-19 pandemic has had extensive social, health and economic consequences for people worldwide. Immediately, the pandemic posed a threat to both the physical and mental health of individuals. According to the WHO (2021b), the total number of deaths associated with COVID-19 exceeded 3 million[3] in 2020. In April 2024, this figure further surpassed 7 million globally, labeling COVID-19 pandemic as one of the most devastating pandemics of all time (Giménez-Llort et al., 2022; World Health Organization, 2024). Even though the most prevalent symptoms of COVID-19 were mild or moderate in nature[4], still a substantial part of those infected, around 10% to 15%, suffered from severe illness[5] (WHO, 2023). Reportedly, critical outcomes[6] were experienced by 5% of those infected (WHO, 2023).

The primary global objective shifted towards mitigating the pandemic as quickly as possible. Vaccination was generally regarded as the most efficient way to accomplish this objective (Valizadeh et al., 2023). It is broadly acknowledged that the process of pharmaceutical development for a vaccine takes a substantial amount of time, often a decade or more (Wouters et al., 2021). This is why scientists sought temporary solutions that could have positively impacted the severity of COVID-19 for those at the highest risk. Even though the exceptional circumstances of the pandemic accelerated vaccine development to less than 12 months, there still remained a period without an adequate protection against COVID-19 (WHO, 2020). Moreover, it was unclear whether the vaccine would be effective against evolving viral strains (El-Shabasy et al., 2022). Thus, alternative strategies to the vaccine were ardently welcomed.

[3]The estimates include predictions of cases not officially reported.

[4]The most common mild or moderate symptoms comprise fever, cough, fatigue, shortness of breath, etc. (WHO, 2023).

[5]Severe outcome was defined as pneumonia with the need for oxygen therapy.

[6]Critical outcome was defined as COVID-19 accompanied by the conditions such as respiratory failure, septic shock, thromboembolism or organ failure.

Antiviral drugs, such as remdesivir or favipiravir, were the first choice for examination as a potential treatment. Remdesivir is a drug that possess antiviral activity against many viruses, including coronavirus SARS-CoV-1 (Beigel et al., 2020). Given that COVID-19 is caused by a genetically similar virus SARS-CoV-2, remdesivir was a target of many scientists (Beigel et al., 2020; Keshta et al., 2021). Indeed, Beigel et al. (2020) identified remdesivir as a potentionally effective choice for the rapid improvement and recovery of the hospitalised patients with COVID-19. Despite remdesivir eventually becoming licensed for the related use of COVID-19, Whittington et al. (2022) claim that remdesivir's exceptionally high price outweighed its efficacy. For instance, the manufacturer Gilead arranged the price for 5 days of therapy at 2 340 USD (Dal-Ré et al., 2021).

It has been suggested that some SSRIs may exhibit antiviral properties and can potentially reduce the risk of severe COVID-19. This idea represented a rather unexpected, however, fairly convenient treatment approach since antidepressants are affordable and widely accessible drugs (Sukhatme et al., 2021).

2.3 Mechanism of action of SSRIs

There are a myriad hypotheses available explaining how SSRIs work against COVID-19. One theory builds on the fact that SSRIs act as sigma-1 receptor agonists[7]. The sigma-1 receptor is a protein placed in the endoplasmic reticulum (ER), which can be coordinated by sigma-1 agonists and antagonists (Facente et al., 2021; Hashimoto, 2021). Thanks to its chaperone activity[8], a sigma-1 receptor manages protein folding and ER stress responses[9] (Vela, 2020).

Sars-CoV-2 virus replication is closely linked to the ER (Vela, 2020). As a response to the SARS-CoV-2 infection, the body naturally produces cytokines (Costela-Ruiz et al., 2020). However, in some cases, Sars-CoV-2 induces excessive cytokine release and ER stress (Facente et al., 2021). In this situation, the immune system becomes overwhelmed, which may result in deterioration of the health status of the patient (Hashimoto, 2021). When SSRIs attach to a

[7]Agonist is "a drug or other chemical agent that binds to a particular receptor and produces a physiological effect" (American Psychological Association, 2018, para. 1).

[8]"Chaperones are a functionally related group of proteins assisting protein folding in the cell under physiological and stress conditions"(Beissinger & Buchner, 1998, Abstract).

[9]ER stress arises when ER is unable to adequately cope with the excessive workload of protein folding (Lin et al., 2008).

sigma-1 receptor, their agonist property triggers the chaperone activity of the sigma-1 receptor, which can help to modulate the ER stress (Vela, 2020).

Alternative theories, for the use of antidepressants as an effective treatment for COVID-19, rest on the properties of a class of drugs called Functional Inhibitors of Acid Sphingomyelinase (FIASMAs). FIASMAs are proposed to be able to impair viruses like SARS-CoV-2 from invading cells (Wang et al., 2023). Certain antidepressants, including those from both SSRI and non-SSRI categories (TCAs, etc.), are classified as FIASMAs. Therefore, researchers posit that non-SSRI antidepressants might be associated with lower risk of severe COVID-19 as well (Wang et al., 2023). On the other hand, opposite perspectives regarding the effects of antidepressant medication exist. For instance, McKeigue et al. (2021) suggest that anticholinergic effects of antidepressants might increase the risk of severe COVID-19.

Chapter 3

Literature review

This chapter intends to provide the comprehensive summary of existing research which analysed the impact of antidepressants on the severity of COVID-19. Section 3.1 will describe studies of both experimental and observational nature that were designed to assess the efficacy of fluvoxamine. Fluvoxamine has been considered as one of the strongest sigma-1 receptor agonists of all SSRIs, that is why it was chosen for testing purposes in early clinical trials (Sukhatme et al., 2021). Section 3.2 will present the results from studies examining medical data retrospectively. Finally, Section 3.3 will introduce our hypothesis that will be tested in this thesis.

3.1 Fluvoxamine

A meta-analysis by Nakhaee et al. (2022) found compelling evidence indicating that fluvoxamine exhibits protective effects against severe COVID-19. An illustration of such positive phenomenon can be found in the randomized controlled trial conducted by Lenze et al. (2020). This study examined whether administering fluvoxamine as an early intervention to people with mild COVID-19 could potentially prevent the progression of disease into the more severe case. A severe outcome in the study was identified as satisfying two conditions: firstly, the pneumonia[1] or the dyspnea[2] (simply the presence of dyspnea or dyspnea resulting in hospitalization); and secondly, a reduction in oxygen levels (saturation level[3] less than 92%) or the need for extra oxygen to maintain a saturation

[1]"Pneumonia is a form of acute respiratory infection" (WHO, n.d., para. 1).

[2]Dyspnea means shortness of breath (Lenze et al., 2020).

[3]"Oxygen saturation level is a measure of how much haemoglobin is currently bound to oxygen compared to how much haemoglobin remains unbound" (WHO, 2021a, para.2).

level of at least 92%. People who experienced symptoms within 7 days after the detection of the COVID-19 positivity were randomly assigned to receive either fluvoxamine or placebo for 15 days. The progression into a severe case of COVID-19 registered 6 in 72 patients in the placebo group, however, none from the 80 individuals in the fluvoxamine (treatment) group.

Calusic et al. (2022) infer that fluvoxamine users demonstrated lower mortality as found in their prospective cohort trial. Besides receiving the standard therapy, the hospitalised patients in the study were either selected to receive fluvoxamine for 15 days or not. Based on the health status specifics from patients, standard therapy comprised of either conventional oxygen therapy or mechanical ventilator support. As for pharmacotherapy, the hospitalised received remdesivir and corticosteroids. The authors claim that fluvoxamine users (exposed group) exhibited a lower risk of mortality. Aside from this finding, no significant difference was discovered when comparing the exposed and non-exposed groups with regard to the total time spent in the intensive unit care (ICU) or hospital, as well as the time spent on ventilator support.

3.2 Antidepressants in observational studies

Retrospective observational studies examining the effect of antidepressants on the severity of COVID-19 demonstrate very diverse conclusions. Vast amount of studies (Firouzabadi et al., 2022; Fritz et al., 2022; Wang et al., 2023) point out the beneficial effects of antidepressants, while other (Rauchman et al., 2022; Bliek-Bueno et al., 2021; Stauning et al., 2023) present contradictory results and claim that there is either none or totally opposite effects associated with antidepressants and severity of COVID-19. The severe outcome is predominantly represented by a hospital or an emergency room admission, death, ventilator support or the need for supplemental oxygen (Deng et al., 2023).

The retrospective cohort study conducted by Wang et al. (2023) suggests that antidepressants, both SSRIs and non-SSRIs, were linked to lower probability of health deterioration in patients hospitalised due to COVID-19. An admission to ICU or death among hospitalised patients were used as endpoints in a Cox proportional hazard model. Only those patients who had a confirmed positive COVID-19 test and were treated with antidepressants within 10 days before the test to 7 days after the test were included in the exposed group. The hazard ratios, which were adjusted for the confounding variables, were 0.260 for ICU admission, 0.872 for inpatient death and 0.769 for the combination of

both events for SSRI users. For patients taking non-SSRI antidepressants, it was 0.401, 0.846 and 0.790, respectively. The similar results were obtained from the retrospective observational study by Hoertel et al. (2021), which focused on hospitalised patients as well. The authors infer that a lower risk of the intubation or death due to COVID-19 was apparent solely in patients who received the antidepressants during their hospital visit. Those who had been prescribed antidepressants three months prior to admission but did not take them during hospitalization were considered to be at a higher risk compared to those who took antidepressants while hospitalized.

Although Fritz et al. (2022) did not examine only medical records of hospitalized patients, their findings are in line with the results from abovementioned studies. Their study aimed to investigate whether there was significant difference in the probability of visiting a hospital in 30 days following a positive SARS-CoV-2 test between users and non-users of SSRIs. The authors incorporated all the medication from individuals' home medication lists. The lists were filled by the individuals themselves and were obtained from the electronic records of all hospital visits preceding the time of the individual's first positive COVID-19 test. Taking the duration of the study into account, the medication could have been taken from the list even over a year before a positive COVID-19 test. The authors propose that those, who had antidepressants on their home medication list, were less likely to visit a hospital or an emergency department within 30 days after the COVID-19 positive test.

As far as the contradictory findings are concerned, Stauning et al. (2023) argue that COVID-19 positive individuals using SSRIs were found to be at significantly greater risk of death. Their large retrospective cohort study focused on Danish residents who were detected to be COVID-19 positive during the first two years of the pandemic. Sensitivity analysis in their research supported the significant association between SSRI use and increased mortality even across different sub-groups and follow-up periods. The impact of sertraline, citalopram, and paroxetine was considered particularly noteworthy. The notable difference between the study by Stauning et al. (2023) and Fritz et al. (2022) may stem from the distinct design of both studies. While Fritz et al. (2022) focused on all the hospital encounters after a COVID-19 test and incorporated every antidepressant from the individuals' home medication lists, Stauning et al. (2023) examined all-cause mortality and severe acute respiratory syndrome following a confirmed COVID-19 test, and they included the antidepressants from national registers with a sufficient daily dosage covering

the date of the positive test. Furthermore, Fritz et al. (2022) found a beneficial effect only for higher doses of SSRIs, while Stauning et al. (2023) did not consider different daily doses.

The study of Bliek-Bueno et al. (2021) supports the conclusion of Stauning et al. (2023). Their results revealed a significant association between increased mortality and the use of antidepressants. The authors' main objective was to identify COVID-19 risk factors in the early stages of the pandemic. They did so by evaluating the medication that significantly influenced all-cause mortality examining COVID-19 positive individuals from 4 March to 17 April 2020 were examined. Amongst other drugs, antidepressants were associated with higher odds of dying.

Numerous other academic works undermine the favourable actions of SSRIs. The retrospective cohort study by Rauchman et al. (2022) did not reveal any significant difference in the odds of mortality for patients taking SSRIs and those not on SSRIs. The metanalysis conducted by Firouzabadi et al. (2022) ascertain that the impact of SSRIs on hospitalisation was insignificant. They even acknowledged that hospitalised patients with SSRIs experienced extended duration of their hospitalisation. Contrariwise, the results of their analysis indicated a significant association between lower mortality and SSRIs.

Although studies investigating the effect of antidepressants on the severity of COVID-19 vary widely in their designs, they bear no consistent outcomes. Both the diverging opinions of researchers and the lack of scientifically verified explanation behind the mechanism of antidepressants demand further research.

3.3 Hypothesis and contribution

In this thesis, we aim to analyse the plausible antiviral effects of antidepressants in active users of SSRIs. Our study will differentiate itself from clinical trials and other studies of observational nature that examined short-term antiviral properties of antidepressants. In those trials, participants typically received antidepressants only during their hospital stay as a potential antiviral treatment. In contrast, our research specifically targets the consistent users of SSRIs, intending to evaluate the underlying long-term preventive effects of antidepressants.

Building on the existing research outlined in Chapter 3, we hypothesize that individuals who take SSRIs regularly were less likely to be hospitalised due to health issues associated with COVID-19. Furthermore, we speculate that

active SSRI use reduced the probability of COVID-19 mortality. This thesis will contribute to the stream of diverging opinions by analysing administrative hospital data coming from the Czech Republic. The primary strength of this thesis lies in its large sample size, which is not usual in medical sciences. Additionally, rather than examining all-cause mortality, we focus on mortality specifically attributable to COVID-19 by implementing ICD-10 codes created for COVID-19 purposes, thus ensuring greater precision. Our research will not only prove useful in the case of an epidemic with virus of similar properties to Sars-CoV-2 breaks out but will also enrich the medical research on repurposing drugs for uses beyond their original intentions.

Chapter 4

Data

The following chapter will focus on data characteristics. We will describe datasets and variables used in our analysis. Furthermore, we will explain the process of data preparation including a detailed derivation of each variable. In Section 4.5, we will describe the characteristics of our sample.

4.1 Data source

In this thesis, six datasets[1] coming from the Institute of Health Information and Statistics of the Czech Republic (ÚZIS) will be analysed. The primary data source for these datasets were Czech insurance companies. The datasets encompass people who tested COVID-19 positive at least once between January 2019 and May 2023. There are more than 4 million people in total that have been detected positive during this period, representing nearly half of the Czech population. The datasets can be merged through patient's unique ID, which is consistent across all datasets.

Infections contains information about the date of the individual's COVID-19 positive test. Due to the fact that some people were detected positive more than once, each observation also includes corresponding serial number of the test. *Deaths* includes the date of the patient's death and details the cause of the death from the death certificate. *Diagnoses* consists of the clinical diagnosis of diseases from which the individuals suffer and the date the diagnoses were reported. The data on diagnoses are available even 5 years prior the start of the pandemic. *Hospitalisations* comprises the data about the date of the hospital

[1]Individual datasets will be referred to as: *Infections*, *Deaths*, *Diagnoses*, *Hospitalisations*, *Medication* and *Demographics*.

admission, the date of the hospital release, the primary and secondary causes of the admission, the severity of the case and the clinical speciality of the admitting health service. The records of hospitalisations come from the National Register of Reimbursed Health Services (NRHZS), which gathers details about reported health services covered by public health insurance. *Medication* considers medical information reported by health care providers. The relevant attributes for our study are ATC codes (described in detail in Section 4.3), the number of packages collected and the date the drug was withdrawn from the pharmacy. The dataset concerning medications contains only drugs covered by public health insurance. The last dataset, *Demographics*, includes demographic information of the individual, such as gender and the age group represented.

4.2 Data preparation

For the study period, we chose the time interval between the start of the pandemic[2] and the introduction of the vaccination[3], i.e. the early stage of the pandemic. Thus, we selected only people that tested COVID-19 positive between 1 March 2020 and 27 December 2020. We opted for such period, because we are interested in the pure effect of antidepressants, separating it from the effects of vaccination or immunisation. The early stage of the pandemic was the focus of many other researchers, such as Lenze et al. (2020) or Fritz et al. (2022). We excluded people that tested positive more than once during our study period, as we intend to analyse one's initial Sars-Cov2 infection and its subsequent impact on health. The datasets track the characteristics of individuals within specific point in time (during our study period), therefore, we deal with cross-sectional data.

[2]WHO declared the start of the pandemic on 11 March 2020 (WHO, 2020a). First positive tests in our datasets appeared on 1 March 2020.

[3]The deployment of the vaccine began in December 2020 (European Centre for Disease Prevention and Control, 2023) with the first dose in our sample being administered on 27 December 2020.

4.3 Description of variables

4.3.1 Dependent variables

Hospitalisation

The variable *hospitalisation* is a dummy variable that is assigned a value of 1 in case the patient was hospitalised due to COVID-19 during our study period. It equals 0 otherwise, meaning that the patient was not hospitalised. To ensure precision, we identified hospitalizations as attributable to COVID-19 by implementing ICD-10 codes. Following the commands by ÚZIS (2021) and the methodology of analytical report by Jarkovský et al. (2020), we defined hospitalisation as causally related to COVID-19 if (for the explanation of ICD-10 codes, see Table 4.1)

(i) the ICD-10 code U07.1 was at the position of the primary cause of the hospitalisation or,

(ii) the ICD-10 code A08.3 at the position of primary cause and U07.1 at the position of secondary cause or,

(iii) the ICD-10 code B34.8 at the position of primary cause and U07.1 at the position of secondary cause or,

(iv) the ICD-10 code Z22.8 at the position of primary cause and U07.1 at the position of secondary cause or,

(v) the ICD-10 codes starting with J (J00-J99) at the position of primary cause and U07.1 at the position of secondary cause.

Table 4.1: ICD-10 codes

ICD-10 code	**Description**
U07.1	COVID-19, virus identified
A08.3	Other viral enteritis
B34.8	Other viral infections of unspecified site
Z22.8	Carrier of other infectious diseases
J00-J99	Diseases of the respiratory system

Source: WHO, n.d.

Furthermore, only hospitalisations occurring between 30 days before and 2 days after the respective COVID-19 positive test were identified as relevant for our analysis. The lower threshold of 30 days was established to assure a direct association between the hospitalization and the positive test. The similar approach can also be seen for example in the study by Faes et al. (2020) or Fritz et al. (2022). The upper limit of 2 days was selected to distinguish between people who contracted the virus during their hospital stay and those who were already admitted with COVID-19. The hospital environment certainly increased the risk of exposure to the virus, especially during the active stage of the pandemic (Hatfield et al., 2023). Consequently, analyzing preexisting risk factors, that influence the probability of hospitalisation, may yield biased results if individual's virus was contracted during the hospital stay[4]. Given that the time for a positive result to appear on the test is approximately 3 to 5 days after the exposure to the virus (Kucirka et al., 2020; Washington State Department of Health, 2023), the test conducted 2 days after admission would still not be able to detect the infection contracted the first day in hospital. Therefore, the positive result of the test conducted 2 days after the admission signalises the presence of an infection contracted prior to the hospital encounter.

Death

The dummy variable *death_covid* denotes whether the patient died due to COVID-19. It is equal to one if the event happened, zero otherwise. Adhering to the recommendation for coding by the WHO (2020b), we extracted the deaths whose underlying cause of death was identified by the ICD-10 code U07.1 (see Table 4.1). The same document was followed by ÚZIS when making instructions on how to report a mortality (see ÚZIS, 2020).

Days spent in hospital

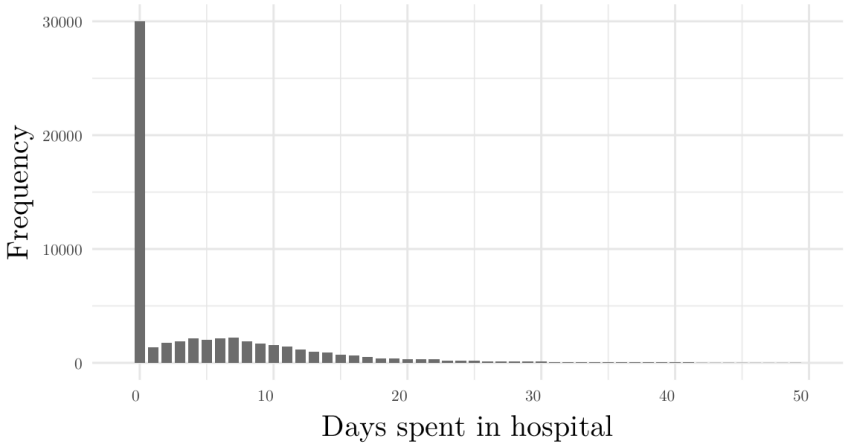
Dependent variable *days_in_hosp* considers the number of days spent in hospital by each individual. The distribution of *days_in_hosp* is displayed in Figure 4.1[5]. One can notice the excess of zeros that *days_in_hosp* exhibits.

[4]For instance, evaluating the preexisting factors of the individual who was hospitalised due other reason than COVID-19, however, caught COVID-19 during the hospital stay may distort our results.

[5]We set the limit for y axis as 30 000 to improve readability of the graph, the actual frequency of the count 0 is 583 479.

This feature can be attributable to the fact that the majority of individuals in our sample were not hospitalised.

Figure 4.1: Distribution of *days_in_hosp*



4.3.2 Explanatory variables

Age

We created a categorical variable, *age*, to be able to adjust for the potential influence of age on our outcomes. Age usually plays an important role in determining the effectiveness and speed of an individual's response to various health challenges. The adjustment for the higher age attains particular significance, since individuals 65 years old or above are considered as one of the most vulnerable groups to experience a severe course of COVID-19 (ÚZIS & MZČR, 2022). We defined following categories:

(i) individuals aged 15-30 years,

(ii) individuals aged 30-45 years,

(iii) individuals aged 45-60 years,

(iv) individuals aged 60-75 years,

(v) individuals aged 75 years and above.

Female

The dummy variable *female* was defined as 1 to denote the female gender. A value of 0 considers the male gender.

SSRIs

In our study, the primary interest lies in people who use SSRIs on regular basis. We thus established the value of the variable *SSRIs* as

$$SSRIs = \begin{cases} 1, & \text{if the individual takes SSRIs regularly,} \\ 0, & \text{if the individual does not use SSRIs regularly.} \end{cases}$$

The regular user of SSRIs at the time of the infection was defined as having at least one record of medication within 3 months before the positive test. According to Czech regulation n. 329/2019 Sb, the medication covered by public health insurance companies can be prescribed for a maximum of 3 months. Additionally, antidepressants are not available over the counter; they are prescribed by a specialist based on the diagnosis and are intended for a longer use. This implies that the prescription 3 months prior the test indicates active and regular use of the medication[6].

SSRIs were extracted from the database of pharmaceuticals based on ATC classification. Anatomical Therapeutic Chemical (ATC) system classifies drugs according to their chemical structure, therapeutic effects and intended medical use (WHO, 2019). The active substances in drugs are first sorted based on which organ system they influence the most and then categorized to form more specific subgroups. ATC codes starting with N correspond to medication influencing nervous system, including antidepressants as well as SSRIs (WHO, 2019). ATC code for the drug category of antidepressants begins with N06A, SSRIs are specifically coded as N06AB.

[6]The doctor can also prescribe a "repeat prescription", with the validity of the receipt 1 year at maximum (VZP, 2024). This allows individuals to potentially pick up their medication in advance, for a period of 6 or 12 months, for instance (SÚKL, 2018). Thus, we extended the former definition of regular user to also include individuals who had collected at least 3 packages within 6 months prior to testing positive or 6 packages within 12 months prior to testing positive. When developing this definition, we had to distinguish between individuals actively using their medication and individuals whose medication status is currently inactive. For instance, only 1 package picked up within 1 year would potentially indicate the inactive status. In our calculation, we followed the assumption that the daily dose of antidepressants ranges from 0.5 to 3 pills a day (Gautam et al., 2017). Given that the average antidepressant package in our sample contains 30 pills, we assume that a person using the medication actively would consume at least half a package per month.

Antidepressants

The analogy of dummy creation for SSRIs is applied for derivation of dummy variable *antidepressants*. We defined the variable *antidepressants* as

$$antidepressants = \begin{cases} 1, & \text{if the individual takes any antidepressants regularly,} \\ 0, & \text{if the individual does not use antidepressants.} \end{cases}$$

Serious medical conditions

To be able to account for the individual's health status that may contribute to the development of a severe outcome of COVID-19, we determined diseases viewed as risk factors of severe COVID-19. We based our selection of risk factors on the analysis conducted by ÚZIS & MZČR (2022). These risk factors enabled us to accurately assess the situation within our study sample, as we work with data from the Czech population. The risk factors are as follows:

(i) hypertension,

(ii) diabetes mellitus,

(iii) ischemic heart disease (IHD),

(iv) liver diseases,

(v) oncological treatment of malignant tumor in the last 5 years,

(vi) chronic respiratory diseases (asthma, COPD),

(vii) chronic kidney diseases,

(viii) obesity,

(ix) stroke.

For the derivation of each medical conditions, we used the combination of ATC and ICD-10 classification systems. The approach described by Sindet-Pedersen et al. (2018) was followed in our analysis. Conditions which were not described by Sindet-Pedersen et al. (2018), were derived from ICD-10 descriptions. We expect that chronic conditions, alongside their health consequences, influence the individual's health state for at least one year following their detection. Therefore, for the determination of relevant diagnoses and medication, we chose a 1-year "look-back" period from the individual's positive COVID-19

test[7]. This method is further supported by the studies of Fortin et al. (2017) and Preen et al. (2006), both of which recommend the use of a 1-year period for the inclusion of diagnoses. The enumeration of respective conditions, alongside their corresponding ATC and ICD-10 codes, can be found in Appendix A.

We set the risk factors as dummy variables, with a value of 1 indicating the presence of corresponding health condition.

CCI score

Each individual was evaluated based on the Charlson Comorbidity Index (CCI) score. CCI is a scoring system used in medical studies for predicting the mortality of individuals who may exhibit a spectrum of comorbid conditions[8] (Charlson et al., 1987). CCI examines selected serious conditions and assignes specific weights to each based on its severity. In practice, patients involved in a certain study are given a CCI score by summing the weight of each comorbidity exhibited.

CCI was developed by Mary E. Charlson in 1987 and since then the score has undergone many adjustments (Charlson et al., 2022; 1987). In our regression analyses, we use a *comorbidity package* provided by R that computes CCI score by implementing updated definition of CCI by Quan et al. (2011). The list of Charlson comorbidit conditions, along with the weights assigned to each of them by Quan et al. (2011), is displayed in Table 4.2.

The common and efficient approach is to use data on diagnoses from a period of up to one year prior to the study period (Metcalfe et al., 2019; Jürisson et al., 2017). Therefore, we used all the individuals' diagnoses starting from 2019 onward. Based on the summary statistic, we defined *cci_score* as a categorical variable with every individual classified into one of the following categories:

(i) having cci_score equal to 0,

(ii) having cci_score equal to 1 or 2,

(iii) having cci_score equal to 3 or 4,

(iv) having cci_score higher than 4.

[7]Oncological treatment was the only exemption for which 5 year period was explicitly stated by ÚZIS & MZČR (2022).

[8]Comorbidities are clinical conditions that co-occur with the medical condition of interest and that simultaneously influence the health state of the individual (Feinstein, 1970).

Due to the small proportion of individuals with CCI score higher than 4 in our dataset, we opted to aggregate them into one group. We argue that individuals with CCI score higher than 4 should all be considered as the most vulnerable. The same threshold of 4 was set for instance by Christensen et al. (2020).

Table 4.2: Charlson comorbid conditions and corresponding weight

Charlson comorbidity	**Weight**
Congestive heart failure	2
Dementia	2
Chronic pulmonary disease	1
Rheumatologic disease	1
Mild liver disease	2
Diabetes with chronic complications	1
Hemiplegia or paraplegia	2
Renal disease	1
Any malignancy, including leukemia and lymphoma	2
Moderate or severe liver disease	4
Metastatic solid tumor	6
AIDS/HIV	4
Maximum comorbidity score	24

Source: Quan et al. (2011)

Waves

We created categorical variable *waves* to be able to account for temporal variations in our regression. Different "waves" of the pandemic carried different levels of public health measures (lockdowns, social distancing, etc.) and changes in public behavior (Bali Swain et al., 2024). These measures impact the spread of the virus and, consequently, the hospitalization and mortality rates. Moreover, modified strains of virus might have developed over time, resulting in people being exposed to diverse viral characteristics at different points in time (Bali Swain et al., 2024). The evolution of new positive cases during our study period is displayed in Figure 4.2.

We chose the dates of individual waves based on the Bayesian change point detection, which is depicted in Figure 4.3. We used a *beast* function in R for identifying the trend breakpoints. The initial peak was at the beginning of April 2020, after the first public measures started (Czech Government, 2022). The second peak occurred at the end of October 2020, where the total number

Figure 4.2: Number of new COVID-19 positive cases

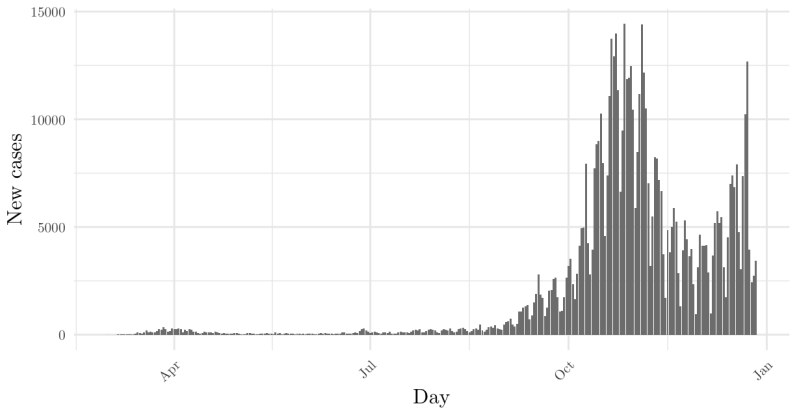
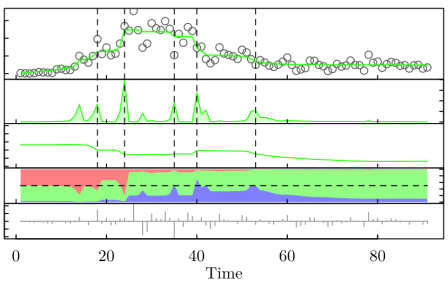
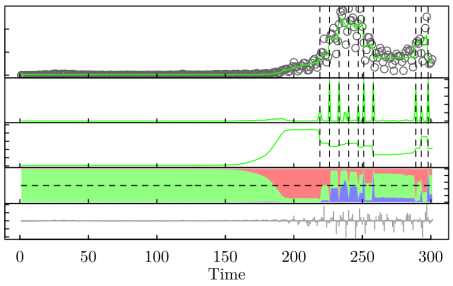


Figure 4.3: Beast decomposition and changepoint detection



(a) Wave 1



(b) Wave 2 and 3

of new cases registered a substantial and rapid rise. We decided to add one final wave, which occurred at the total end of our study period (December 2020).

The individual i in our study will be categorised into one of these groups:

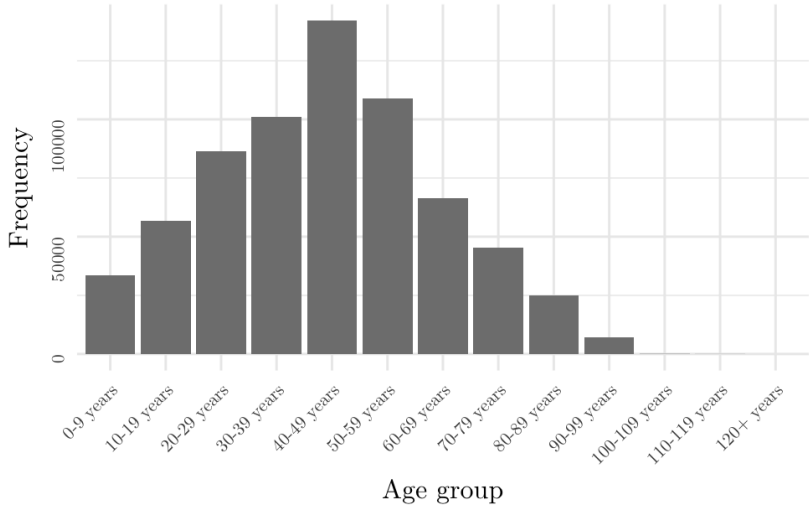
$$\begin{cases} wave\_1, & \text{if } i\text{'s first positive test was obtained from 25/03/2020 to 10/04/2020,} \\ wave\_2, & \text{if } i\text{'s first positive test was obtained from 13/10/2020 to 14/11/2020,} \\ wave\_3, & \text{if } i\text{'s first positive test was obtained from 15/12/2020 to 27/12/2020,} \\ no\_wave, & \text{if } i\text{'s first positive test was obtained at any other point.} \end{cases}$$

4.4 Data cleansing

To improve reliability of our analysis, the datasets underwent data cleansing procedures. We filtered out the drugs where the recorded amount of collected packages and the number of the pills in a package was negative. Incomplete data and negative number of packages, which is rationally not possible to withdraw, might signal reporting errors. We did not encounter any missing values.

In our datasets, we observed a category with ages even above 120 years, which is biologically unachievable. Figure 4.4 illustrates the age distribution of individuals involved in the study. Based on the distribution, we decided to filter out individuals above 100 years, since they appeared very rarely in the dataset and thus might indicate plausible data errors. We also excluded

Figure 4.4: Age distribution



children younger than 15 years age as they might distort our results. The alterations in children's dosage of medication may not align with the prescribed usage criteria derived from adult metrics (WHO, n.d.). Furthermore, the variable *cci_score* was derived from the adult patient-oriented research (Quan et al., 2011) and thus might yield biased estimates for children. In fact, many researchers studying the effects of antidepressants, such as Wang et al. (2023) or Stauning et al. (2023), included exclusively adult patients in their analysis.

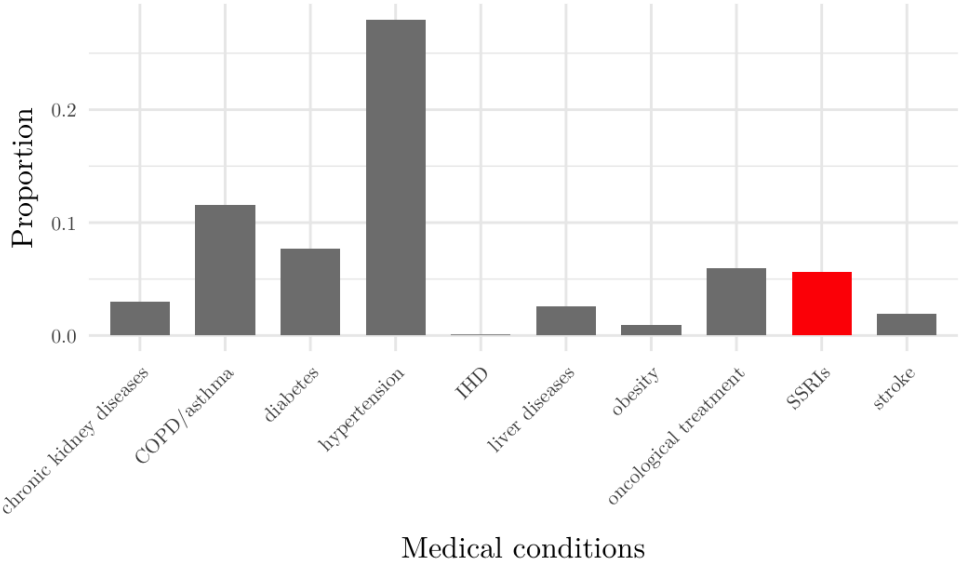
4.5 Descriptive statistics

Our final sample comprises 612 368 individuals that were once detected COVID-19 positive during our study period (from 01/03/2020 to 27/12/2020). A total

of 29 204 individuals were hospitalised due to COVID-19. Among the hospitalised patients, the average length of hospital stay was 9.9 days, with median value of 8 days. The range for the variable *days_spent_in_hosp* is 134 days, with minimum value of 0 and maximum value of 134. Since *days_in_hosp* includes an excessive number of zeros, $Q1 = Q3 = 0$[9].

In our study, there are 52.5% females and 47.5% males. The most frequent age bracket (12% of all individuals) comprises individuals between 45 and 50 years old. The high-risk group of 65 years and older represents 9.0% of all individuals, as displayed in Figure 4.4. There are 47 473 (7.7%) people who take antidepressants, from which 36 328 (5.6%) are SSRI users. This characteristic aligns with the statement of Gautam et al. (2017), demonstrating that SSRIs represent the most frequently prescribed class of antidepressants. Interestingly, more than 70% of people who actively use SSRIs are women.

Figure 4.5: Medical conditions



The prevalence of serious medical conditions in our sample is depicted in Figure 4.5, where the percentage of individuals taking SSRIs is highlighted in red. The proportion of each condition is roughly in line with the world estimates. Obesity, a condition which is hard to measure only via the medication and the diagnosis, might deviate from the reality the most. The treatment of this disease is often immeasurable as it primarily involves lifestyle changes. The CCI score ranges from 0 to 16, with 78% of individuals in 0 group. This implies

[9]Q1 is the lower quartile and Q3 is the upper quartile.

that the majority of individuals in our sample do not experience serious health complications, which aligns with our expectations. CCI score between 1 and 2 is observed in 18.1% of individuals, while there are only 3.1% of individuals with CCI score of 3 to 4, and 0.8% of individuals with CCI score higher than 4. Table 4.3 displays the precise description of our sample.

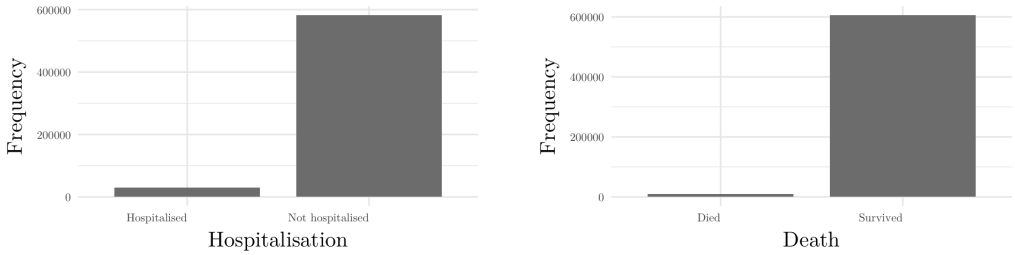
Table 4.3: Descriptive statistics of the study sample

Variable	**Number of observations**	**Proportion of sample (in %)**
female	321 576	52.5
hospitalised	29 204	4.8
age category		
15-19 years	30 481	5.0
20-29 years	86 334	14.1
30-39 years	101 013	16.5
40-49 years	141 978	23.2
50-59 years	108 916	17.8
60-69 years	66 320	10.8
70-79 years	45 269	7.4
80-89 years	24 928	4.1
90-99 years	7 129	1.2
antidepressants	47 473	7.8
SSRIs	34 328	5.6
hypertension	171 442	28.0
diabetes mellitus	46 883	7.7
IHD	415	0.1
liver diseases	15 668	2.6
oncological treatment	36 528	6.0
asthma/ COPD	70 893	11.6
obesity	5 755	0.9
chronic kidney diseases	18 424	3.0
stroke	11 738	1.9
CCI score		
CCI score 0	477 680	78.0
CCI score 1 or 2	110 957	18.1
CCI score 3 or 4	19 040	3.1
CCI score 4 plus	4 691	0.8

As far as the mortality is concerned, we identified deaths occurring 90 days after the respective COVID-19 positive test as relevant for our analysis. Deaths are well-documented events and serve as reliable endpoints in epidemiological studies. The cause of an individual's death is clearly stated on the death certifi-

cate, making it unlikely to result in interpretation or reporting errors. Thus, we chose a wider threshold compared to hospitalisation. Due to new assumptions posed on the dependent variable, the new sample size rose to 615 502 observations. During our study period, 8 975 deaths of COVID-19 occurred. The distribution of dependent variables *hospitalisation* and *death_covid* is displayed in Figure 4.6.

Figure 4.6: Distribution of dependent variables



Chapter 5

Methodology

This chapter will describe econometric models used in our analysis, specifically logistic regression and the zero-inflated negative binomial model.

5.1 Logistic regression

Logistic regression is an econometric method used for estimating the probability whether a certain event occurs or not (Boateng et al., 2019). Logistic regression is specifically designed for binary dependent variables, and therefore is convenient for scenarios where the outcome of interest has two possible values (for instance success/failure or yes/no). Models used for modeling binary variables are commonly referred to as binary response models (Wooldridge, 2013). The main objective of these models is to obtain the response probability

$$P(y = 1|x) = P(y = 1|x_1 + x_2 + \ldots + x_k), \tag{5.1}$$

where $y \in \{0, 1\}$ is dependent variable and $x \in \mathbb{R}$ is a set of independent variables (Wooldridge, 2013).

The alternative way to model binary dependent variable would be to use the Linear Probability Model (LPM). This model has, however, many imperfections. These imperfection include constant marginal effects of explanatory variables, inherently present heteroscedasticity and estimates of probability potentially outside the [0,1] interval (Wooldridge, 2013).

Fortunately, better models have been proposed for modelling binary variables, including the logistic regression model. The logistic regression model, or logit model, addresses the shortcomings of LPM by using a non-linear func-

tion to model the probability[1] (Wooldridge, 2013). This mechanism can be illustrated as

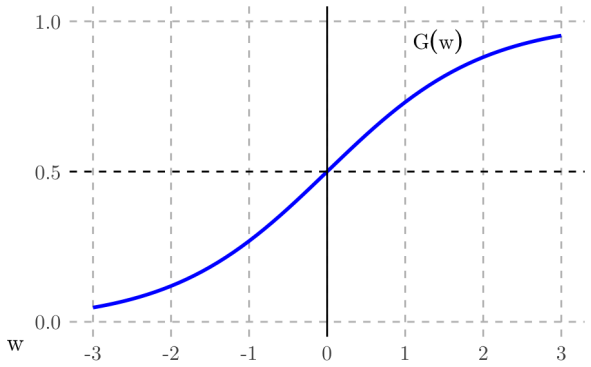
$$P(y = 1|x) = G(\beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k) = G(\beta_0 + x\beta), \tag{5.2}$$

where G is cumulative distribution function (cdf) of a standard logistic random variable and following property holds for G

$$G(w) = \frac{exp(w)}{1 + exp(w)} = \frac{exp(\beta_0 + x\beta)}{1 + exp(\beta_0 + x\beta)} \; ; w \in \mathbb{R}. \tag{5.3}$$

In logit, the dependent variable is assumed to have logistic distribution, thus G is called logistic function (Wooldridge, 2013). G is strictly increasing for all real numbers w, with values satisfying $0 < G(w) < 1$, which guarantees the desired probability within interval (0, 1). The logistic function is depicted in Figure 5.1.

Figure 5.1: Graph of the logistic function



Logit can be derived using the underlying latent (unobserved) variable y^* defined as

$$y^* = \beta_0 + x\beta + \varepsilon, \tag{5.4}$$

where ε is i.i.d. error term independent of x with logistic distribution. Furthermore, the binary variable y assumes that

$$y = \begin{cases} 1 & \text{if } y^* > 0, \\ 0 & \text{if } y^* \leq 0. \end{cases} \tag{5.5}$$

[1]One could use probit model instead of logit model. We will continue with explanation of logit, as it is widely used model in medical research and is suitable for our large dataset with imbalanced data (Stoltzfus, 2011; Wilkinson et al., 2022).

The formula for cdf of logit can be derived as follows

$$\begin{aligned} P(y=1|x) &= P(y^* > 0|x) = P(\beta_0 + x\beta + \varepsilon > 0|x) = P(\varepsilon < \beta_0 + x\beta|x) \\ &= P(\varepsilon \leq \beta_0 + x\beta) = G(\beta_0 + x\beta) = \frac{\exp(\beta_0 + x\beta)}{1 + \exp(\beta_0 + x\beta)}. \end{aligned} \tag{5.6}$$

Due to non-linear nature of logit, one must use the Maximum Likelihood Estimation (MLE) method to determine the parameters of the model. The probability density function g of y_i given x_i needed for MLE is defined as

$$g(y_i|x_i;\beta) = [G(x_i\beta)]^{y_i}[1 - G(x_i\beta)]^{1-y_i}. \tag{5.7}$$

The likelihood function is the joint probability function of y_i conditional on x_i. Since we assumed that ϵ_i is i.i.d., the probabilities are independent for each i. Thus, the joint probability function is the product of individual densities depicted in 5.7. By taking the logarithm of the likelihood, we derive the log-likelihood function. The log-likelihood of sample of size n is as follows

$$ln\mathcal{L}(\beta) = \sum_{i=1}^{n} y_i \ln[G(x_i\beta)] + (1 - y_i)\ln[1 - G(x_i\beta)], \tag{5.8}$$

where we applied a principle that the logarithm of a product (represented by the joint probability in our case) is a sum of logarithms. To obtain MLE of β, we need to maximize 5.8. The MLE of β, expressed by $\hat{\beta}$, is called the logit estimator. Since MLE considers the actual distribution of the dependent variable, heteroskedasticity in logistic regression is implicitly addressed during the estimation procedure (Wooldridge, 2013).

However, one disadvantage of logit exists; the nonlinear character of MLE makes the coefficients of logit troublesome for the interpretation. Even though the direction of the effect of a variable aligns with the sign of its coefficients, its exact numerical value does not have straightforward meaning. The effect of x_j on the probability that $y = 1$ depends also on x. The partial effect of x_j can be written as

$$\frac{\partial P(y=1|x)}{\partial x_j} = g(\beta_0 + x\beta)\beta_j. \tag{5.9}$$

For the accurate interpretation, one has to use alternative measures. One of the most convenient way is to calculate Average Marginal Effect (AME). AME is obtained by calculating the separate partial effects across the entire sample,

followed by their averaging. AME of x_j is calculated as

$$AME = \frac{\sum_{i=1}^{n} g\left(\hat{\beta}_0 + x_i\hat{\beta}\right)}{n}\beta_j. \tag{5.10}$$

AME for discrete variables or dummies is computed in the different manner (Wooldridge, 2013). Let x_k be discrete or dummy variable, then the effect of x_k on the probability resulting from change in x_k from c_k to $c_k + 1$ is

$$n^{-1}\sum_{i=1}^{n} G\left(\hat{\beta}_0 + \hat{\beta}_1\bar{x}_1 + \ldots + \hat{\beta}_k(c_k + 1)\right) - G\left(\hat{\beta}_0 + \hat{\beta}_1\bar{x}_1 + \ldots + \hat{\beta}_k c_k\right). \tag{5.11}$$

For discrete variables, 5.11 denotes the average change in the predicted probability for a unit change in the discrete variable, holding other variables constant. In case of dummy variables, 5.11 represents the predicted difference in probability that $y_i = 1$ when x_k changes from 0 to 1 (Wooldridge, 2013). As AME estimates the change in probability, it is interpreted in percentage points.

For the goodness-of-fit measure of logit model, one can use the pseudo R-squared (Wooldridge, 2013). Pseudo R-squared is defined as follows

$$R^2 = 1 - \frac{\mathcal{L}_{ur}}{\mathcal{L}_0}, \tag{5.12}$$

where $\mathcal{L}_{ur}$ is obtained from the log-likelihood function of the estimated model and $\mathcal{L}_0$ is obtained from the log-likelihood function of the model including only intercept, i.e., null model (Wooldridge, 2013). The alternative assessment of the overall fit of the model can be obtained by the likelihood ratio (LR) test (Boateng et al., 2019; Wooldridge, 2013), where

$$LR = 2(\mathcal{L}_{ur} - \mathcal{L}_0). \tag{5.13}$$

LR follows an approximate chi-square distribution under the null hypothesis that the estimated model demonstrates better fit to the data compared to the null model (Boateng et al., 2019). LR can also be used for selecting between the nested models.

Another convenient measure used for the comparison of nested models is called Akaike Information Criterion (AIC) (Portet, 2020). One should choose the model with the lowest AIC, since lower AIC indicates better fit for the data. AIC is defined as

$$AIC = -2\mathcal{L}_{ur} + 2k, \tag{5.14}$$

where k is the number of parameters in the estimated model (Portet, 2020).

5.2 Zero-inflated negative binomial model

Zero-Inflated Negative Binomial Model (ZINB) is a model used for estimating count data with excessive number of zeros (Zuur et al., 2012). ZINB assumes that zeros are generated by two separate processes. Individuals are either part of the "Always-0 Group" or the "Not Always-0 Group" (Long & Freese, 2006). The "Always-0 Group" (A_0) group is assumed to have probability of 1 that its count outcome is 0. On the other hand, individuals in the "Not Always-0 Group" (A_n) can have outcome equal to 0 but also have a positive probability to have a non-zero count outcome (Long & Freese, 2006). In our case[2], A_0 group comprises individuals who were not hospitalised and had all the predispositions not to be hospitalised. It implies that they were in the good health state and did not exhibit significant risk factors. A_n includes individuals who were either hospitalised or were not hospitalised but had nonzero probability to spent positive days in hospital. The latter case includes individuals who

(i) had certain predispositions to experience severe COVID-19, however, they were fortunate to avoid experiencing its severe form, or

(ii) went to the hospital but were not accepted due to capacity problems or

(iii) went to the hospital but were immediately released, since their condition could be managed at home with the prescribed medication.

ZINB uses binary model, usually logit, to estimate the group membership of an individual, i.e. being in A_0 or A_n. This membership is a latent (unobserved) variable (Long & Freese, 2006). Assume that $A = 1$ for A_0 members and $A = 0$ for A_n members, then probability of being in A_0 is

$$P(A_i = 1|z_i) = \psi_i = \frac{\exp(z_i\beta)}{1 + \exp(z_i\beta)}, \tag{5.15}$$

where z_i are called inflation variables. Negative binomial model is afterwards employed to predict the count process. Among A_n group members, the proba-

[2]Variable *days_spent_in_hospital* will serve as dependent variable in ZINB; discussed in detail in Chapter 6.

bility of each count (including 0) is as follows

$$P(y_i|x_i, A_i = 0) = \frac{\Gamma(y_i + \alpha^{-1})}{y_i!\Gamma(\alpha^{-1})} \left(\frac{\alpha^{-1}}{\alpha^{-1} + \mu_i}\right)^{\alpha^{-1}} \left(\frac{\mu_i}{\alpha^{-1} + \mu_i}\right)^{y_i}, \tag{5.16}$$

where $y_i \in \mathbb{N} \cup \{0\}$, Γ is the gamma function, α is the dispersion parameter and $\mu_i = \exp(x_i\beta)$ is the mean of y_i. Equation 5.16 is simply the parametrized probability mass function of negative binomial distribution. The dispersion parameter α determines the level of overdispersion in the predictions (Long & Freese, 2006).

Finally, the overall probability is defined as

$$P(Y_i = y_i|x_i, z_i) = \begin{cases} \psi_i + (1 - \psi_i)P(y_i = 0|x_i, A_i = 0), & \text{for } y_i = 0 \\ (1 - \psi_i)P(y_i|x_i, A_i = 0), & \text{for } y_i = 1,2,3,... \end{cases}$$

The expected counts are calculated as follows (Long & Freese, 2006)

$$E(Y_i = y_i|x_i, z_i) = 0 \cdot \mu_i + \mu_i(1 - \psi_i) = \mu_i(1 - \psi_i). \tag{5.17}$$

The poisson regression model can be used instead of the negative binomial to model the count part of the zero-inflated model. The resulting model is referred to as Zero-Inflated Poisson Model (ZIP). Both negative binomial model and poisson model are widely used methods for analysing count data. However, poisson model is inconvenient when the data exhibit overdispersion (Zuur et al., 2012). Overdispersion condition occurs when the variance of the variable is greater than the mean under the assumed poisson distribution (Long & Freese, 2006). Negative binomial model incorporates a parameter α (refer to Equation 5.16) to account for unobserved heterogeneity among observations. Therefore, when the count variable is overdispersed and excessive zeros are observed, ZINB is preferred model (Long & Freese, 2006). The same metrics described in Section 5.1, such as LR test or AIC, can be used for ZINB.

Chapter 6

Results

In this chapter, we will introduce specific models constructed to test our hypothesis and subsequently provide a comprehensive analysis of their results. The main analyses were conducted using the R programming language, with Python employed solely for data preparation of large datasets that exceeded handling capabilities of R. We defined statistical significance as p-value < 0.05.

6.1 Logistic regression

We aim to estimate the probability of hospitalization (or death) among the infected individuals and our dependent variables are binary, thus, logistic regression represents the most suitable method for our analysis. Indeed, logistic regression is widely used statistical approach in medical research (Bagley et al., 2001).

The Baseline Model with *hospitalisation* as dependent variable will be defined as follows

$$\begin{aligned} P(hospitalisation_i = 1|x_i) = G(&\beta_0 + \beta_1 female_i + \beta_2 age_i + \beta_3 SSRIs_i \\ &+ \beta_4 stroke_i + \beta_5 diabetes_i + \beta_6 IHD_i \\ &+ \beta_7 hypertension_i + \beta_8 liver\_diseases_i \\ &+ \beta_9 oncological\_treatment_i \\ &+ \beta_{10} chronic\_kidney\_diseases_i \\ &+ \beta_{11} copd\_asthma_i + \beta_{12} waves_i), \end{aligned} \tag{1}$$

where $G(\cdot)$ represents the function described in Section 5.1. The Baseline

Model passed LR test for overall significance, with p-value less than $2 \cdot 10^{-16}$. Furthermore, pseudo R-squared of 0.24 indicates very good overall fit of the estimated model (McFadden, 1977).

There are a few assumptions that must hold for logistic regression to yield reliable results (Boateng et al., 2019; Stoltzfus, 2011). First, a large sample size is required and the dependent variable must be binary. In our analysis, these assumption are inherently satisfied. Second, logit assumes independence of errors (Stoltzfus, 2011). Each individual appears uniquely in our dataset and exhibits characteristics that are independent of other individuals. Thus, the mutual independence across observations is satisfied. The third assumption concerns the linearity of the continuous independent variables and logit-transformed outcome (Stoltzfus, 2011). Since we use only categorical variables in our analysis, there is no need to check the validity of this assumption.

Fourth, there should be an absence of influential outliers in the dataset (Stoltzfus, 2011). Identifying outliers in categorical variables is more complex than in continuous variables. Our dataset underwent data cleansing procedure (see Section 4.4) and we found no other data points to significantly deviate from other records. Finally, there must be no perfect colinearity among independent variables. We verified this assumption by calculating Generalised Variance Inflation Factor (GVIF)[1]. All GVIF values for our independent variables are around 1, demonstrating a small degree of multicollinearity. $GVIF$ for independent variables from the Baseline Model are presented in Table 6.1. The results for other models are presented in Table B.1.

An atypical attribute of our data involves the distribution of dependent variables (see Figure 4.6). Both dependent variables assume a disproportionately small number of ones relative to zeros. Abd Rahman & Ong (2020) addressed this imbalance in their simulations, where they demonstrated that logit estimates remain unbiased provided that the sample size is sufficiently large (4000 or more observations) and the proportion of ones exceeds 1% of the total observations. Both dependent variables in our dataset meet this criterion. In case of *hospitalisation*, there are 29 322 ones (4.8% of all observation) and 8 975 ones (1.5%), in case of *death_covid*.

[1]We chose GVIF because our model includes categorical independent variables that require more than 1 coefficient and thereby more than 1 degree-of-freedom (df). Fox & Monette (1992) suggest examining formula $GVIF^{\frac{1}{2df}}$, which is more convenient for the comparison across different dimensions.

Table 6.1: Results of GVIF

	GVIF	Df	$GVIF^{\frac{1}{2df}}$
female	1.055	1	1.027
age	1.423	4	1.045
SSRIs	1.053	1	1.026
diabetes	1.113	1	1.055
copd_asthma	1.012	1	1.006
hypertension	1.320	1	1.149
IHD	1.028	1	1.014
stroke	1.042	1	1.021
liver_diseases	1.005	1	1.003
oncological_treatment	1.052	1	1.026
chronic_kidney_diseases	1.072	1	1.036
obesity	1.007	1	1.004
waves	1.002	3	1.000

Results of the Baseline Model reveal that all selected variables are significant, except for *IHD*. Even though we believe that Ischemic Heart Disease (IHD) is an important risk factor and confounding variable, we decided to remove this variable from our model due to its small proportion in our dataset. Indeed, the model without IHD exhibits slightly lower AIC, indicating the better choice of model. For results of the Baseline Model, refer to the first column of Table 6.2. AME of individual variables are displayed in Table 6.4. All other serious medical conditions have a positive sign, meaning that the presence of one or more conditions increases the risk of being hospitalised. Inclusion of serious clinical conditions is crucial for our analysis as the (prior) health status is certainly an influential factor affecting individual's probability of hospitalisation (Ko et al., 2020). The presence of chronic respiratory disease or obesity have the largest average marginal effect on hospitalisation, increasing the probability by 2.7 and 2.5 percentage points, respectively.

Contrary to what we expected, variable *SSRIs* is positive and significant, even at the 1% significance level. It can be inferred that, other things being equal, being the long-term user of SSRIs increased the probability of hospitalisation by 0.4 percentage points. Although the effect is rather small, it is statistically significant. Age was one of the most significant predictors[2]. Being

[2]Age, when included as ordinal variable with all the age categories we have available in the dataset, did not significantly influence the results of our models. Refer to Table B.5 for results of model with age as ordinal variable.

75 years or above increased the probability of hospitalisation due COVID-19 by 43.6 percentage points in comparison to age group of 15 to 30 years. This implies that the elderly were a high-risk group during the pandemic. All other age groups had significantly higher probability than the base group comprising individuals 15 to 30 years old. Additionally, females exhibited a lower risk of hospitalisation compared to males. Having a female gender decreased the probability by 2.1 percentage points. This phenomenon illustrates the potential gender disparities in lifestyle factors.

The variable *waves* provides compelling insights into how the risk of hospitalisation altered over different periods of the pandemic. The chance of being hospitalised increased by 2.7 percentage points among people detected positive during the first wave compared to those who tested positive during the "no-wave" periods. Surprisingly, the second wave had the opposite effect to the first wave. Getting infection during the second wave decreased probability of the hospital visit by 0.7 percentage points in comparison to the infected in the "no-wave". Even though the second wave was substantially larger than the initial one, a smaller percentage of infected people experienced a severe outcome during this period (Meschiari et al., 2022). The final wave has the same effect as the second wave.

To verify our results based on specific risk factors by ÚZIS & MZČR (2022), *cci_score* is incorporated into the model instead of considering each medical condition separately. The results are displayed in the second column of Table 6.2. Variable SSRIs remains significant and positive even in this model, increasing the probability of hospitalisation by 0.2 percentage points when comparing users and non-users of SSRIs. The effect of variables *female*, *age* and *waves* persist approximately the same. Compared to the base group with the CCI score of 0, being the individual with CCI score between 1 and 2, 3 and 4, or greater than 4, increased the probability of hospitalisation by 2.4, 4.0, or 4.2 percentage points, respectively. It implies that the higher the CCI score, the higher the risk of a severe outcome. The larger AMEs for variable *cci_score* as compared to AMEs of individual conditions in the Baseline Model appear reasonable, since the CCI score groups serious conditions together. Our finding supports the outcomes of Christensen et al. (2020), who infer that people with higher CCI score were significantly more likely to incur a severe COVID-19 as well as had higher probability to die of COVID-19.

The model with *cci_score* also helped us to address the potential bias stemming from the low proportion of individual conditions in the Baseline Model.

To investigate the potential bias from another perspective, we gathered all chronic medical conditions into one group. The new variable *chronic1* equals one if the individual suffers from at least one chronic condition described in Section 4.3.2 and 0 otherwise. Afterwards, we created *chronic2*, which aggregates all chronic conditions except of *hypertension*. We included *hypertension* as a separate predictor into the model due its high prevalence in our sample. Both new variables are statistically significant and they changed neither the significance nor the effect of other variables. Furthermore, these variables enable us to account also for *IHD*, which was excluded from the Baseline Model. The results of the models incorporating *chronic1* and *chronic2* are displayed in the third and fourth column of Table 6.2, respectively.

The second part of the logit analysis revolves around the dependent variable *death_covid* while keeping the same independent variables as in the Baseline Model. The results of the model align with previous outcomes. Being a SSRI user increased the probability to die of COVID-19 by 0.3 percentage points. This probability is relatively small, yet, it remains statistically significant. The summary of logistic regressions is displayed in Table 6.3, the first column corresponds to the model with individual medical conditions and the second column to the model with *cci_score*. AMEs are displayed in Table 6.4. We slightly modified the age categories in the models, as most of the people who died of COVID-19 were elderly. Even though the overall effect of almost all variables remains the same, few changes are worth highlighting. Variable *IHD* becomes significant and the sign of *wave2* is opposite. The new significance of *IHD* should not raise much concern, as it potentially influences mortality more than hospitalisation.

We further suggest that despite individuals infected during the second wave having lower probability of hospitalisation compared to those infected during the "no-wave" period, they faced a higher risk of dying of COVID-19. Eventually, both the probability of death and hospitalisation were significantly lower during the third wave than in periods outside our defined waves. The gradual decline in risk of mortality (and hospitalization) across successive infection peaks, relative to the "no-peak" periods, can be attributed to an improved understanding of the virus over time. Hospitals attained better organisational structure and effective preventive measures were implemented outside the hospital (Caramello et al., 2022). Thanks to progress in research, the risk factors were known in advance and COVID-19 could be better categorized according to its stage and severity (Caramello et al., 2022). Moreover, as discussed in Sec-

tion 4.3.2, modified strains of virus with slightly different characteristics may have developed over time (Bali Swain et al., 2024). The evolved strains have been suggested to have varying degree of transmissibility as well as different impact on mortality and morbidity of COVID-19 (SeyedAlinaghi et al., 2021).

Finally, we recreated the presented models by replacing variable *SSRIs* with variable *antidepressants*. The goal is to further investigate the effect of aggregated group of all the antidepressants, regardless of which class they belong to (combining both SSRIs and non-SSRIs). Based on the results, we infer that antidepressant medications yield consistent outcomes, irrespective of their specific class. Refer to Table B.3 and Table B.4 for results of logistic regression incorporating *antidepressants*.

Table 6.2: Results of logistic regression with *hospitalisation*

	Dependent variable:			
	hospitalisation			
	(1)	(2)	(3)	(4)
female	−0.533***	−0.561***	−0.551***	−0.532***
	(0.013)	(0.013)	(0.013)	(0.013)
age				
30-45	1.179***	1.214***	1.146***	1.165***
	(0.065)	(0.065)	(0.065)	(0.065)
45-60	2.122***	2.285***	2.056***	2.089***
	(0.061)	(0.061)	(0.061)	(0.061)
60-75	3.455***	3.778***	3.451***	3.418***
	(0.061)	(0.060)	(0.061)	(0.061)
75plus	4.277***	4.573***	4.310***	4.230***
	(0.061)	(0.060)	(0.061)	(0.061)
SSRIs	**0.099****	**0.056****	**0.114****	**0.104****
	(0.021)	(0.021)	(0.021)	(0.021)
diabetes	0.482***			
	(0.016)			
COPD_asthma	0.594***			
	(0.016)			
stroke	0.193***			
	(0.026)			
oncological_treatment	0.212***			
	(0.018)			
chronic_kidney_diseases	0.389***			
	(0.023)			
liver_diseases	0.101**			
	(0.034)			
obesity	0.537***			
	(0.061)			
hypertension	0.442***			0.475***
	(0.016)			(0.016)
chronic2				0.644***
				(0.014)
chronic1			0.866***	
			(0.018)	
cci_score				
cci_1to2		0.553***		
		(0.015)		
cci_3to4		0.801***		
		(0.022)		
cci_4plus		0.814***		
		(0.038)		
waves				
wave1	0.561***	0.528***	0.545***	0.550***
	(0.066)	(0.065)	(0.065)	(0.065)
wave2	−0.182***	−0.173***	−0.182***	−0.180***
	(0.014)	(0.014)	(0.014)	(0.014)
wave3	−0.323***	−0.303***	−0.314***	−0.319***
	(0.023)	(0.022)	(0.022)	(0.022)
Constant	−5.762***	−5.724***	−5.832***	−5.785***
	(0.059)	(0.059)	(0.060)	(0.059)
Observations	612,368	612,368	612,368	612,368
Log Likelihood	−89,194	−90,439	−90,255	−89,615
AIC	178,423	180,903	180,532	179,254
McFadden R-squared	0.240	0.229	0.231	0.236
Note:				*p<0.05; **p<0.01; ***p<0.001

Table 6.3: Results of logistic regression with *death_covid*

	Dependent variable:			
	death_covid			
	(1)	(2)	(3)	(4)
female	−0.795***	−0.810***	−0.822***	−0.798***
	(0.024)	(0.024)	(0.023)	(0.024)
age				
65-75	2.828***	2.992***	2.850***	2.814***
	(0.047)	(0.044)	(0.046)	(0.046)
75-85	3.840***	3.973***	3.916***	3.844***
	(0.046)	(0.043)	(0.045)	(0.046)
85+	4.474***	4.512***	4.490***	4.458***
	(0.047)	(0.045)	(0.046)	(0.047)
SSRIs	**0.217****	**0.139****	**0.241****	**0.238****
	(0.032)	(0.032)	(0.032)	(0.032)
diabetes	0.422***			
	(0.025)			
COPD_asthma	0.291***			
	(0.028)			
IHD	0.335*			
	(0.159)			
stroke	0.432***			
	(0.034)			
liver_diseases	0.331***			
	(0.058)			
oncological_treatment	0.098***			
	(0.028)			
chronic_kidney_diseases	0.540***			
	(0.033)			
obesity	0.311*			
	(0.144)			
cci_score				
cci_1to2		0.601***		
		(0.028)		
cci_3to4		1.004***		
		(0.034)		
cci_4plus		1.254***		
		(0.050)		
hypertension	0.401***			0.450***
	(0.031)			(0.031)
chronic2				0.574***
				(0.025)
chronic1			0.899***	
			(0.041)	
waves				
wave1	0.462***	0.414***	0.449***	0.449***
	(0.108)	(0.108)	(0.107)	(0.107)
wave2	0.079***	0.089***	0.072**	0.078***
	(0.023)	(0.023)	(0.023)	(0.023)
wave3	−1.132***	−1.111***	−1.125***	−1.130***
	(0.056)	(0.056)	(0.056)	(0.056)
Constant	−6.360***	−6.298***	−6.539***	−6.428***
	(0.041)	(0.041)	(0.045)	(0.042)
Observations	615,497	615,497	615,497	615,497
Log Likelihood	−31,663	−31,726	−32,040	−31,869
AIC	63,362	63,476	64,101	63,761
McFadden R-squared	0.324	0.323	0.316	0.320

Note: *p<0.05; **p<0.01; ***p<0.001

Table 6.4: AME

	Dependent variable:			
	hospitalisation		death_covid	
female	-0.021^{***}	-0.023^{***}	-0.010^{***}	-0.010^{***}
age 30-45	0.064^{***}	0.068^{***}		
age 45-60	0.121^{***}	0.135^{***}		
age 60-75	0.260^{***}	0.306^{***}		
age 75 plus	0.436^{***}	0.492^{***}		
age 65-75			0.066^{***}	0.073^{***}
age 75-85			0.126^{***}	0.136^{***}
age 85 plus			0.198^{***}	0.202^{***}
SSRIs	$\mathbf{0.004}^{***}$	$\mathbf{0.002}^{**}$	$\mathbf{0.003}^{***}$	$\mathbf{0.002}^{**}$
diabetes	0.021^{***}		0.005^{***}	
copd_asthma	0.027^{***}		0.004^{***}	
hypertension	0.017^{***}		0.005^{***}	
stroke	0.008^{***}		0.006^{***}	
oncological_treatment	0.009^{***}		0.001^{***}	
chronic_kidney_diseases	0.017^{***}		0.008^{***}	
liver_diseases	0.004^{**}		0.008^{**}	
obesity	0.025^{***}		0.004	
IHD			0.005	
cci_1to2		0.024^{***}		0.008^{***}
cci_3to4		0.041^{***}		0.017^{***}
cci_4plus		0.042^{***}		0.025^{***}
wave1	0.027^{***}	0.025^{***}	0.007^{***}	0.006^{***}
wave2	-0.007^{***}	-0.007^{***}	0.001^{*}	0.001^{**}
wave3	-0.012^{***}	-0.011^{***}	-0.010^{***}	-0.010^{***}
Observations	612,368	612,368	615,497	615,497
Log Likelihood	−89,194	−90,439	−31,663	−31,726

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

6.2 Zero-inflated negative binomial model

The primary objective of the second analysis is to explore the effect of SSRIs while addressing the issue of excess zeros in the dataset. Variable *days_in_hosp* will be used as dependent variable. Since *days_in_hosp* is overdispersed (tested below) count variable with excess of zeros and we assume that zeros in our dataset are generated by two different processes, ZINB serves as the best model for our analysis. We chose the final predcitors for both parts of ZINB based on AIC measure. The resulting model contains aggregated chronic conditions *chronic1*, *female*, *age* and *waves* in the count part. For the logit part, the best fit represented the model with each medical condition included separately. Based on LR test of overall significance, we cannot reject the null hypothesis that our estimated model demonstrates improved form over the null model. Results of ZINB are shown in Table 6.5. The outcomes of alternative models with different independent variables are included in Table B.6.

The results of the *dispersiontest* in R confirmed the presence of overdispersion, validating the choice of ZINB over ZIP. In addition to overdispersion and excessive zeros, comparable assumptions to logit are required by ZINB (Zuur et al., 2012). There should be independence of observations, no outliers and no perfect collinearity (Zuur et al., 2012). Due to the excessive zeros[3] in *days_in_hosp*, one should conduct the outlier analysis only on the positive count part of the variable (Yang et al., 2011). We identified outliers as values exceeding $Q3 + IRQ \cdot k$, where $Q1$ is the first quantile, $Q3$ is the third quantile, $IRQ = Q3 - Q1$ is the inter-quartile range and k is set to 3 for the extreme outliers (Yang et al., 2011). The resulting value for *days_in_hosp* equals 42. However, we observed that most of the values above this threshold do not substantially deviate from other observations, do not appear rarely in our dataset and their values appear reasonable from a theoretical point of view. We decided against classifying all values above this threshold as outliers to avoid the potential loss of valuable information. Instead, we decided to remove the six highest observations, since they deviate from other observations the most[4]. We verified the remaining assumptions using the same approaches as in Section 6.1.

A necessary discussion within our analysis involves the problem of endogeneity. Scientists have suggested that hospitals often lengthen the hospital stay of patients even though they are already in the appropriate medical state

[3]For *days_in_hosp*, $Q1 = Q3 = 0$.

[4]We excluded *days_in_hospital*={93,95,101,108,130,134}.

for discharge (Rojas-García et al., 2018). The non-clinical reasons for delayed discharge may involve operational issues (for instance limited hospital service on weekends), organisational delays (delayed inspections, assessments or administrative work) or financial incentives (potential for increased government funding stemming from each additional day a patient spends in hospital) (Cadel et al., 2021). Therefore, the extended hospital stay of a patient might not necessarily indicate that the patient required more time for recovery due to a more severe condition. The non-medical reasons for the increased number of days spent in hospital might lead to omitted variable bias. However, given the context of the pandemic, where hospital capacity was often exhausted (Setola et al., 2022), we assume that hospitals were not able to inefficiently prolong hospital stays. The primary goal was to release patients as soon as they were fit enough to continue recovery at home. This assumption allows us to mitigate concerns about endogeneity due to hospital inefficiencies.

The outcome of ZINB consists of two parts, negative binomial estimates and logit estimates. The interest of our analysis lies primarily in the second part of the model, the logit part, which predicts the probability of being in the A_0 ("certain zero") group. This part serves as a form of sensitivity analysis to Section 6.1. The coefficient for variable *SSRIs* is significant and negative. This implies that those taking SSRIs regularly had a lower probability to be among individuals with all predispositions to not be hospitalised. Among SARS-CoV-2 infected individuals, being a SSRI user decreased the probability of being in the "Always-0 Group" by 0.4 percentage points, holding other variables constant. The effect of antidepressants did not change even when we replaced *SSRIs* by *antidepressants*. Additionally, higher age or chronic diseases decreased the probability of being in the "certain zero" group. In contrast, female gender was associated with higher chance of being in the A_0 group. To conclude, the binary part of ZINB aligns with the results from Section 6.1.

The first part of ZINB, the count model, estimates the count of days spent in the hospital by COVID-19 positive individuals during our study period. This part reveals that the coefficient for SSRIs is insignificant, implying that there is not enough evidence to claim that the SSRI use influenced the length of the hospital stay. Contrarily, the individuals with chronic conditions exhibited significantly longer duration of the hospital stay compared to individuals without any chronic diseases. Among the individuals who had positive probability to be hospitalised, having at least one chronic disease increased the expected count of days spent in hospital by 0.89 days on average. Furthermore, being

Table 6.5: Results of ZINB

	Negative binomial part of ZINB	
	Coefficients	AME
female	-0.103^{***} (0.010)	-0.759
age 30-45	0.164^{**} (0.055)	0.961
age 45-60	0.367^{***} (0.052)	2.379
age 60-75	0.595^{***} (0.052)	4.370
age 75 plus	0.552^{***} (0.052)	3.957
SSRIs	**0.012** (0.015)	0.092
chronic1	0.120^{***} (0.014)	0.890
wave1	0.441^{***} (0.043)	4.186
wave2	-0.054^{***} (0.010)	-0.399
wave3	-0.011 (0.016)	-0.080
Constant	1.707^{***} (0.051)	
	Logit part of ZINB	
female	0.528^{***} (0.013)	0.021
age 30-45	-1.183^{***} (0.066)	-0.007
age 45-60	-2.112^{***} (0.062)	-0.023
age 60-75	-3.430^{***} (0.062)	-0.087
age 75 plus	-4.250^{***} (0.063)	-0.176
SSRIs	$\mathbf{-0.098}^{***}$ (0.021)	**-0.004**
diabetes	-0.482^{***} (0.016)	-0.022
COPD_asthma	-0.598^{***} (0.016)	-0.028
hypertension	-0.441^{***} (0.016)	-0.017
stroke	-0.192^{***} (0.026)	-0.008
oncological_treatment	-0.211^{***} (0.018)	-0.009
kidney_diseases	-0.388^{***} (0.023)	-0.017
liver_diseases	-0.100^{**} (0.034)	-0.004
obesity	-0.542^{***} (0.062)	-0.026
wave1	-0.542^{***} (0.066)	-0.028
wave2	0.184^{***} (0.014)	0.008
wave3	0.322^{***} (0.023)	0.013
Constant	5.716^{***} (0.061)	
Observations	612,362	
Log Likelihood	$-180{,}827$	
AIC	361,950	
Note:	$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$	

60 to 75 years old increased the expected stay in hospital by 4.37 days compared to those 15 to 30 years old. Variable *wave3* was insignificant, as well as some of the chronic conditions when included separately[5]. We argue that there is not enough evidence to claim that experiencing certain chronic conditions (specifically chronic respiratory disease, chronic liver disease or stroke) in 1 year prior to infection significantly affected the length of hospital stay for patients recovering from COVID-19. Diabetes, oncological treatment received within 5 years prior the infection and hypertension were significantly associated with the longer duration of the hospital visit in all models.

In the second model, we accounted for medication used for the treatment of severe form of COVID-19, specifically remdesivir, in the count part of ZINB. Results[6] of the model are displayed in Table 6.6. Since remdesivir is administered primarily to patients experiencing the most critical symptoms, this medication can serve as an indicator of the most severe cases of COVID-19. Consequently, the positive sign of *remdesivir* suggests that high-risk individuals who received remdesivir experienced longer hospital stay compared to those who did not receive the medication. The key observation is the persisting insignificance of *SSRIs* in the negative binomial part of the model and the negative coefficient in the logit part, confirming our previous results.

6.3 Discussion

The results of the analysis suggest a significant association between SSRIs and an increased probability of hospitalizations or deaths due to COVID-19, similar to other chronic diseases considered. One can speculate about plausible explanations of this finding. First, we cannot exclude the possibility of SSRI use being associated with specific behavioral patterns. People taking SSRIs are used to regular visits to healthcare facilities for medication adjustments and doctor's assessments. Therefore, they may be more prone to seek medical help even for minor symptoms of COVID-19 that most people would find tolerable. Furthermore, the pandemic related distress might have worsened the mental condition of people already using antidepressants (see Section 2.1). The

[5]For results, refer to Table B.6.

[6]Due to multicollinearity issues, we accounted for age in logit part of ZINB in form of dummy variable that distinguishes between high-risk individuals older than 65 years and those younger than 65 years.

Table 6.6: Results of ZINB

	Negative binomial part of ZINB	
	Coefficients	AME
female	-0.095^{***} (0.010)	-0.609
age 30-45	0.867^{***} (0.065)	3.232
age 45-60	1.144^{***} (0.063)	5.013
age 60-75	1.373^{***} (0.062)	6.905
age 75 plus	1.370^{***} (0.063)	6.873
SSRIs	**0.019** (0.015)	0.126
remdesivir	0.377^{***} (0.016)	2.930
chronic1	0.118^{***} (0.014)	0.756
wave1	0.500^{***} (0.044)	4.280
wave2	-0.066^{***} (0.010)	-0.419
wave3	-0.014 (0.017)	-0.089
Constant	0.876^{***} (0.063)	
	Logit part of ZINB	
female	0.484^{***} (0.013)	0.020
age 65 plus	-1.927^{***} (0.016)	-0.105
SSRIs	$\mathbf{-0.179}^{***}$ (0.021)	**-0.008**
diabetes	-0.499^{***} (0.016)	-0.023
COPD_asthma	-0.582^{***} (0.016)	-0.027
hypertension	-0.730^{***} (0.016)	-0.030
stroke	-0.266^{***} (0.026)	-0.012
oncological_treatment	-0.262^{***} (0.018)	-0.011
kidney_diseases	-0.416^{***} (0.023)	-0.019
liver_diseases	-0.115^{***} (0.035)	-0.005
obesity	-0.491^{***} (0.062)	-0.024
wave1	-0.525^{***} (0.066)	-0.028
wave2	0.158^{***} (0.014)	0.007
wave3	0.332^{***} (0.023)	0.013
Constant	3.975^{***} (0.016)	
Observations	612,362	
Log Likelihood	$-182{,}928$	
AIC	365,950	
Note:	$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$	

consequent factors like stress, lifestyle changes or a lack of respect for public regulations might have contributed to the development of severe outcome.

Second, the antiviral effects of antidepressants might not have been as intense in longer term users as in people receiving antidepressants as a short-term antiviral therapy. Stauning et al. (2023) hypothesised that the anti-inflammatory mechanisms of SSRIs may gradually diminish in regular users over time. Third, given that individuals were using SSRIs prior to their Sars-Cov2 infection, it is important to consider that their use may also indicate the presence of major depressive disorder. People with depression are believed to be more susceptible to a broad spectrum of infections plausibly due to immunological changes triggered by the mental illness (Andersson et al., 2015). Indeed, the depression is suggested to increase the risk of COVID-19 progressing into a severe outcome (Molero et al., 2023). Nakhaee et al. (2022) proposed that the effective treatment of depression, for instance in the form of SSRIs, could reduce this risk. However, our analysis proposes that the negative effect of underlying depression might have still outweighed the positive effect of antidepressants. Finally, our results support the theory by McKeigue et al. (2021), hypothesising that drugs with anticholinergic effects increase the probability of severe COVID-19 (mentioned in Section 2.3). This suggestion stems from the fact that both SSRIs and an aggregated group of all antidepressants yield similar results in our analysis.

The effect of SSRIs on the expected number of days spent in hospital proved insignificant. We infer that even though the consistent use of SSRIs increased the probability of hospitalisation, there is no significant evidence to claim that the subsequent length of hospital stay was influenced by their use.

Chapter 7

Conclusion

During the COVID-19 pandemic, the scientists proposed that antidepressants, especially SSRIs, might exhibit antiviral properties that could mitigate the severity of COVID-19 (Nakhaee et al., 2022). Yet, studies investigating the hypothesized protective effects of antidepressants yield conflicting results. The main purpose of this thesis is to examine whether the regular user of SSRIs was less likely to exhibit a severe form of COVID-19.

We implemented logistic regression to estimate the probability of hospitalisation (or death) due to COVID-19 among the infected individuals. The study sample consists of 612 368 individuals who tested COVID-19 positive once between 1 March 2020 and 27 December 2020. The results of the analysis suggest that the use of SSRIs was significantly associated with an increased probability of hospitalisation. Additionally, we found that SSRI users faced a higher risk to die of COVID-19 in comparison with individuals not using SSRIs. In both cases, the effect of antidepressants was rather small, however, statistically significant. We used various independent variables, such as age, gender, chronic diseases the individual suffered from or the "wave" of the pandemic in which the individual was infected, to account for potential confounding factors. The effect remained significant even when replacing SSRIs with the aggregated group of all antidepressants. This finding suggests that the antidepressant medication yield consistent outcomes, irrespective of its specific drug class. Afterwards, we employed the zero-inflated negative binomial model to estimate the expected days spent in hospital by each individual. The effect of SSRIs on the length of hospital stay proved insignificant. The insignificant effect persisted even after adjusting for remdesivir, which was used to treat the most severe cases of COVID-19.

The rationale behind our findings can be derived through several theories. First, the use of SSRIs may be potentially associated with specific behavioral patterns. Second, the intensity of "antiviral" effects of antidepressants may diminish over time (Stauning et al., 2023). Third, SSRIs could also act as an indicator of the underlying psychiatric condition. Indeed, major depressive disorder is suggested to increase the risk of COVID-19 progressing into severe outcome (Molero et al., 2023). Finally, the anticholinergic effects of antidepressants may contribute to the increased probability of severe COVID-19 (McKeigue et al., 2021).

Our thesis contributes to the stream of diverging opinions by analysing the large dataset coming from the Czech Republic, filling the research gap in Czechia. The strength of this thesis is in the large sample size, which is not usual in medical sciences. Furthermore, instead of all-cause mortality, we inspected specifically death attributable to COVID-19 by implementing ICD-10 codes created for COVID-19 purposes. Our study may be relevant in the case of an epidemic with a virus of similar properties to Sars-Cov-2 breaks out. Additionally, this study may have an implication in the medical field by enriching the research on repurposing drugs for uses beyond their original intentions. In further research, it might be interesting to investigate the effect of antidepressants on other contagious respiratory diseases.

Several limitations of our thesis should be highlighted. First, we could not include the factors associated with individual's behaviour and habits, such as smoking, alcohol consumption, highest education achieved or the type of behaviour the individual engages in (risk aversion, etc.), that could have contributed to the probability of experiencing severe COVID-19. Second, due to a lack of expertise, the current study did not account for the dose of prescribed antidepressants. Therefore, we cannot make a conclusion on whether the daily dosage plays a role in the effectiveness of antidepressants. Third, we focused solely on the early stage of the pandemic. We cannot rule out the possibility that the different strains of the virus, which could have evolved over time, interacted differently with SSRIs. Fourth, we cannot completely exclude the possibility of endogeneity discussed in Section 6.2.

To conclude, consistent with Stauning et al. (2023) and others, no beneficial protective effects of SSRIs against the severe form of COVID-19 was observed. Our results thereby disincentivize the use of antidepressants as a form of long-term preventive antiviral therapy against COVID-19.

Bibliography

Abd Rahman, H. A. & Ong, S.-H. (2020). Predictive performance of logistic regression for imbalanced data with categorical covariate. *Pertanika Journal of Science and Technology*, *28*.

American Psychological Association (2018). Agonist. Retrieved April 1, 2024 from https://dictionary.apa.org/agonist.

Andersson, N. W., Goodwin, R. D., Okkels, N., Gustafsson, L. N., Taha, F., Cole, S. W., & Munk-Jørgensen, P. (2015). Depression and the risk of severe infections: prospective analyses on a nationwide representative sample. *International Journal of Epidemiology*, *45*(1), 131–139.

Bagley, S. C., White, H., & Golomb, B. A. (2001). Logistic regression in the medical literature:: Standards for use and reporting, with particular attention to one medical domain. *Journal of Clinical Epidemiology*, *54*(10), 979–985.

Bali Swain, R., Lin, X., & Wallentin, F. Y. (2024). COVID-19 pandemic waves: Identification and interpretation of global data. *Heliyon*, *10*(3), e25090.

Beigel, J. H., Tomashek, K. M., Dodd, L. E., Mehta, A. K., Zingman, B. S., Kalil, A. C., Hohmann, E., Chu, H. Y., Luetkemeyer, A., Kline, S., de Castilla, D. L., Finberg, R. W., Dierberg, K., Tapson, V., Hsieh, L., Patterson, T. F., Paredes, R., Sweeney, D. A., Short, W. R., Touloumi, G., Lye, D. C., Ohmagari, N., don Oh, M., Ruiz-Palacios, G. M., Benfield, T., FÄ¤tkenheuer, G., Kortepeter, M. G., Atmar, R. L., Creech, C. B., Lundgren, J., Babiker, A. G., Pett, S., Neaton, J. D., Burgess, T. H., Bonnett, T., Green, M., Makowski, M., Osinusi, A., Nayak, S., & Lane, H. C. (2020). Remdesivir for the treatment of covid-19 – final report. *New England Journal of Medicine*, *383*(19), 1813–1826.

Beissinger, M. & Buchner, J. (1998). How chaperones fold proteins. [Abstract]. *Biological chemistry*, *379*(3), 245–259.

Bliek-Bueno, K., Mucherino, S., Poblador-Plou, B., González-Rubio, F., Aza-Pascual-Salcedo, M., Orlando, V., Clerencia-Sierra, M., Ioakeim-Skoufa, I., Coscioni, E., Carmona-Pírez, J., et al. (2021). Baseline drug treatments as indicators of increased risk of COVID-19 mortality in spain and italy. *International Journal of Environmental Research and Public Health*, *18*(22), 11786.

Boateng, E. Y., Abaye, D. A., et al. (2019). A review of the logistic regression model with emphasis on medical research. *Journal of data analysis and information processing*, *7*(04), 190.

Braund, T. A., Tillman, G., Palmer, D. M., et al. (2021). Antidepressant side effects and their impact on treatment outcome in people with major depressive disorder: an ispot-d report. *Translational Psychiatry*, *11*, 417.

Cadel, L., Guilcher, S. J. T., Kokorelias, K. M., Sutherland, J., Glasby, J., Kiran, T., & Kuluski, K. (2021). Initiatives for improving delayed discharge from a hospital setting: a scoping review. *BMJ Open*, *11*(2).

Calusic, M., Marcec, R., Luksa, L., Jurkovic, I., Kovac, N., Mihaljevic, S., & Likic, R. (2022). Safety and efficacy of fluvoxamine in COVID-19 ICU patients: An open label, prospective cohort trial with matched controls. *British Journal of Clinical Pharmacology*, *88*(5), 2065–2073.

Caramello, V., Catalano, A., Macciotta, A., Dansero, L., Sacerdote, C., Costa, G., Aprà, F., Tua, A., Boccuzzi, A., & Ricceri, F. (2022). Improvements throughout the three waves of COVID-19 pandemic: Results from 4 million inhabitants of north-west italy. *Journal of Clinical Medicine*, *11*(15).

Chae, J. H., Lee, K. U., Shin, Y. K., et al. (2004). Comparison of side effect profiles between mirtazapine and selective serotonin reuptake inhibitors; a naturalistic setting. *Clinical Psychopharmacology and Neuroscience*, *2*, 31–35.

Charlson, M. E., Carrozzino, D., Guidi, J., & Patierno, C. (2022). Charlson Comorbidity Index: A Critical Review of Clinimetric Properties. *Psychotherapy and Psychosomatics*, *91*(1), 8–35.

Charlson, M. E., Pompei, P., Ales, K. L., & MacKenzie, C. (1987). A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation. *Journal of Chronic Diseases*, *40*(5), 373–383.

Christensen, D. M., Strange, J. E., Gislason, G., Torp-Pedersen, C., Gerds, T., Fosbøl, E., & Phelps, M. (2020). Charlson comorbidity index score and risk of severe outcome and death in danish COVID-19 patients. *Journal of general internal medicine*, *35*(9), 2801–2803.

Costela-Ruiz, V. J., Illescas-Montes, R., Puerta-Puerta, J. M., Ruiz, C., & Melguizo-Rodríguez, L. (2020). Sars-cov-2 infection: The role of cytokines in COVID-19 disease. *Cytokine & Growth Factor Reviews*, *54*, 62–75.

Cui, L., Li, S., Wang, S., & et al. (2024). Major depressive disorder: hypothesis, mechanism, prevention and treatment. *Sig Transduct Target Ther*, *9*, 30.

Czech Government (2022). Measures adopted by the czech government against the coronavirus. Retrieved April 20, 2024 from https://vlada.gov.cz/en/media-centrum/aktualne/measures-adopted-by-the-czech-government-against-coronavirus-180545/.

Czech Republic (2019). Vyhláška č. 329/2019 sb., § 5 odst. 3 písm. d. *Sbírka zákonů* [Collection of Laws]. Retrieved April 2, 2024 from https://www.zakonyprolidi.cz/cs/2019-329.

Dal-Ré, R., Banzi, R., Georgin-Lavialle, S., Porcher, R., Sofat, R., Zeitlinger, M., & Rosendaal, F. R. (2021). Remdesivir for COVID-19 in europe: will it provide value for money? *The Lancet. Respiratory medicine*, *9*(2), 127–128.

de Oliveira, C., Saka, M., Bone, L., & et al. (2023). The role of mental health on workplace productivity: A critical review of the literature. *Applied Health Economics and Health Policy*, *21*, 167–193.

Deng, J., Rayner, D., Ramaraju, H. B., Abbas, U., Garcia, C., Heybati, K., Zhou, F., Huang, E., Park, Y.-J., & Moskalyk, M. (2023). Efficacy and safety of selective serotonin reuptake inhibitors in COVID-19 management: a systematic review and meta-analysis. *Clinical Microbiology and Infection*, *29*(5), 578–586.

Edinoff, A. N., Akuly, H. A., Hanna, T. A., Ochoa, C. O., Patti, S. J., Ghaffar, Y. A., Kaye, A. D., Viswanath, O., Urits, I., Boyer, A. G., et al. (2021). Se-

lective serotonin reuptake inhibitors and adverse effects: A narrative review. *Neurology International*, *13*(3), 387–401.

El-Shabasy, R. M., Nayel, M. A., Taher, M. M., Abdelmonem, R., Shoueir, K. R., & Kenawy, E. R. (2022). Three waves changes, new variant strains, and vaccination effect against COVID-19 pandemic. *International Journal of Biological Macromolecules*, *204*, 161–168.

European Centre for Disease Prevention and Control (2023). Overview of the implementation of COVID-19 vaccination strategies and deployment plans in the eu/eea.

Facente, S. N., Reiersen, A. M., Lenze, E. J., Boulware, D. R., & Klausner, J. D. (2021). Fluvoxamine for the early treatment of SARS-CoV-2 infection: A review of current evidence. *Drugs*, *81*(18), 2081–2089.

Faes, C., Abrams, S., Van Beckhoven, D., Meyfroidt, G., Vlieghe, E., Hens, N., & on COVID-19 Hospital Surveillance, B. C. G. (2020). Time between symptom onset, hospitalisation and recovery or death: statistical analysis of belgian covid-19 patients. *International journal of environmental research and public health*, *17*(20), 7560.

Feinstein, A. R. (1970). The pre-therapeutic classification of co-morbidity in chronic disease. *Journal of Chronic Diseases*, *23*(7), 455–468.

Firouzabadi, D., Kheshti, F., Abdollahifard, S., Taherifard, E., & Kheshti, M. R. (2022). The effect of selective serotonin and norepinephrine reuptake inhibitors on clinical outcome of COVID-19 patients: A systematic review and meta-analysis. *Health Science Reports*, *5*(6), e892.

Fortin, Y., Crispo, J. A., Cohen, D., McNair, D. S., Mattison, D. R., & Krewski, D. (2017). Optimal look back period and summary method for elixhauser comorbidity measures in a us population-based electronic health record database. *Open Access Medical Statistics*, *7*, 1–13.

Fox, J. & Monette, G. (1992). Generalized collinearity diagnostics. *Journal of the American Statistical Association*, *87*, 178–183.

Fritz, B. A., Hoertel, N., Lenze, E. J., Jalali, F., & Reiersen, A. M. (2022). Association between antidepressant use and ed or hospital visits in outpatients with sars-cov-2. *Translational psychiatry*, *12*(1), 341.

Gasparini, A. (2018). comorbidity: An r package for computing comorbidity scores. *Journal of Open Source Software*, *3*, 648.

Gautam, S., Jain, A., Gautam, M., Vahia, V. N., & Grover, S. (2017). Clinical practice guidelines for the management of depression. *Indian journal of psychiatry*, *59*(Suppl 1), S34–S50.

Giménez-Llort, L., Torres-Lista, V., Oghagbon, E. K., Pereira, H. V. F. S., Gijsberts, M.-J. H. E., & Invitto, S. (2022). Editorial: Death and mourning processes in the times of the coronavirus pandemic (COVID-19). *Frontiers in Psychiatry*, *13*.

Hashimoto, K. (2021). Repurposing of cns drugs to treat COVID-19 infection: targeting the sigma-1 receptor. *271*(2), 249–258.

Hatfield, K. M., Baggs, J., Maillis, A., Warner, S., Jernigan, J. A., Kadri, S. S., Klompas, M., & Reddy, S. C. (2023). Assessment of Hospital-Onset SARS-CoV-2 Infection Rates and Testing Practices in the US, 2020-2022. *JAMA Network Open*, *6*(8), e2329441–e2329441.

Hoertel, N., Sánchez-Rico, M., Vernet, R., Beeker, N., Jannot, A.-S., Neuraz, A., Salamanca, E., Paris, N., Daniel, C., Gramfort, A., Lemaitre, G., Bernaux, M., Bellamine, A., Lemogne, C., Airagnes, G., Burgun, A., Limosin, F., & AP-HP / Universities / INSERM COVID-19 Research Collaboration and AP-HP COVID CDR Initiative (2021). Association between antidepressant use and reduced risk of intubation or death in hospitalized patients with COVID-19: results from an observational study. *Molecular psychiatry*, *26*(9), 5199–5212.

Jarkovský, J., Pavlík, T., Dlouhý, P., Koziar Vašáková, M., Benešová, K., Šanca, O., Bortlíček, Z., Dušek, L., & Válek, V. (2020). Preventivní účinek orálních antivirotik proti úmrtí a těžkému průběhu COVID-19 v ČR v době výskytu varianty SARS-CoV-2 omikron. *COVID 19: Přehled aktuální situace v ČR. Onemocnění aktuálně*. [online]. Praha: Ministerstvo zdravotnictví ČR. Retrieved March 3, 2024 from https://onemocneni-aktualne.mzcr.cz/preventivni-ucinek-antivirotik. ISSN 2694-9423.

Jia, H., Zack, M. M., Thompson, W. W., Crosby, A. E., & Gottesman, I. I. (2015). Impact of depression on quality-adjusted life expectancy (qale) directly as well as indirectly through suicide. *Social Psychiatry and Psychiatric Epidemiology*, *50*(6), 939–949.

Jürisson, M., Raag, M., Kallikorm, R., et al. (2017). The impact of comorbidities on hip fracture mortality: a retrospective population-based cohort study. *Archives of Osteoporosis*, *12*(1), 76.

Kelly, K., Posternak, M., & Alpert, J. E. (2008). Toward achieving optimal response: understanding and managing antidepressant side effects. *Dialogues in clinical neuroscience*, *10*(4), 409–418.

Keshta, A. S., Mallah, S. I., Al Zubaidi, K., Ghorab, O. K., Keshta, M. S., Alarabi, D., Abousaleh, M. A., Salman, M. T., Taha, O. E., Zeidan, A. A., Elsaid, M. F., & Tang, P. (2021). COVID-19 versus sars: A comparative review. *Journal of infection and public health*, *14*(7), 967–977.

Ko, J. Y., Danielson, M. L., Town, M., Derado, G., Greenlund, K. J., Kirley, P. D., Alden, N. B., Yousey-Hindes, K., Anderson, E. J., Ryan, P. A., Kim, S., Lynfield, R., Torres, S. M., Barney, G. R., Bennett, N. M., Sutton, M., Talbot, H. K., Hill, M., Hall, A. J., Fry, A. M., Garg, S., Kim, L., & Team, C.-N. S. (2020). Risk Factors for Coronavirus Disease 2019 (COVID-19)-Associated Hospitalization: COVID-19-Associated Hospitalization Surveillance Network and Behavioral Risk Factor Surveillance System. *Clinical Infectious Diseases*, *72*(11), e695–e703.

Kucirka, L. M., Lauer, S. A., Laeyendecker, O., Boon, D., & Lessler, J. (2020). Variation in false-negative rate of reverse transcriptase polymerase chain reaction–based sars-cov-2 tests by time since exposure. *Annals of internal medicine*, *173*(4), 262–267.

Lenze, E. J., Mattar, C., Zorumski, C. F., Stevens, A., Schweiger, J., Nicol, G. E., Miller, J. P., Yang, L., Yingling, M., Avidan, M. S., & Reiersen, A. M. (2020). Fluvoxamine vs placebo and clinical deterioration in outpatients with symptomatic COVID-19: A randomized clinical trial. *JAMA*, *324*(22), 2292–2300.

Lim, G. Y., Tam, W. W., Lu, Y., Ho, C. S., Zhang, M. W., & Ho, R. C. (2018). Prevalence of depression in the community from 30 countries between 1994 and 2014. *Scientific reports*, *8*(1), 2861.

Lin, J. H., Walter, P., & Yen, T. B. (2008). Endoplasmic reticulum stress in disease pathogenesis. *Annual Review of Pathology: Mechanisms of Disease*, *3*(Volume 3, 2008), 399–425.

Long, J. S. & Freese, J. (2006). *Regression models for categorical dependent variables using Stata*, volume 7. Stata press.

McFadden, D. (1977). Quantitative methods for analysing travel behaviour of individuals: some recent developments. *Cowles Foundation Discussion Papers*, (707).

McKeigue, P., Kennedy, S., Weir, A., Bishop, J., Mcgurnaghan, S., McAllister, D., Robertson, C., Wood, R., Lone, N., Murray, J., Caparrotta, T., Smith-Palmer, A., Goldberg, D., Mcmenamin, J., Guthrie, B., Hutchinson, S., & Colhoun, H. (2021). Relation of severe COVID-19 to polypharmacy and prescribing of psychotropic drugs: the react-scot case-control study. *BMC Medicine*, *19*.

Meschiari, M., Cozzi-Lepri, A., Tonelli, R., Bacca, E., Menozzi, M., Franceschini, E., Cuomo, G., Bedini, A., Volpi, S., Milic, J., Brugioni, L., Romagnoli, E., Pietrangelo, A., Corradini, E., Coloretti, I., Biagioni, E., Busani, S., Girardis, M., Cossarizza, A., Clini, E., Guaraldi, G., & Mussini, C. (2022). First and second waves among hospitalised patients with COVID-19 with severe pneumonia: a comparison of 28-day mortality over the 1-year pandemic in a tertiary university hospital in italy. *BMJ Open*, *12*(1).

Metcalfe, D., Masters, J., Delmestri, A., Judge, A., Perry, D., Zogg, C., Gabbe, B., & Costa, M. (2019). Coding algorithms for defining charlson and elixhauser co-morbidities in read-coded databases. *BMC Medical Research Methodology*, *19*(1), 115.

Molero, P., Reina, G., Blom, J. D., Martínez-González, M. A., Reinken, A., de Kloet, E. R., & Molendijk, M. L. (2023). COVID-19 risk, course and outcome in people with mental disorders: a systematic review and meta-analyses. *Epidemiology and Psychiatric Sciences*, *32*. PMID: 37859501.

Nakhaee, H., Zangiabadian, M., Bayati, R., Rahmanian, M., Ghaffari Jolfayi, A., & Rakhshanderou, S. (2022). The effect of antidepressants on the severity of COVID-19 in hospitalized patients: A systematic review and meta-analysis. *PloS one*, *17*(10), e0267423.

OECD & European Union (2018). *Health at a Glance: Europe 2018*. OECD Publishing. https://doi.org/https://doi.org/10.1787/health_glance_eur-2018-en.

Portet, S. (2020). A primer on model selection using the Akaike Information Criterion. *Infectious Disease Modelling*, *5*, 111–128.

Preen, D. B., Holman, C. J., Spilsbury, K., Semmens, J. B., & Brameld, K. J. (2006). Length of comorbidity lookback period affected regression model performance of administrative health data. *Journal of Clinical Epidemiology*, *59*(9), 940–946.

Quan, H., Li, B., Couris, C. M., Fushimi, K., Graham, P., Hider, P., Januel, J.-M., & Sundararajan, V. (2011). Updating and Validating the Charlson Comorbidity Index and Score for Risk Adjustment in Hospital Discharge Abstracts Using Data From 6 Countries. *American Journal of Epidemiology*, *173*(6), 676–682.

Ramic, E., Prasko, S., Gavran, L., & Spahic, E. (2020). Assessment of the antidepressant side effects occurrence in patients treated in primary care. *Materia socio-medica*, *32*(2), 131–134.

Rathod, S., Pallikadavath, S., Graves, E., Rahman, M., Brooks, A., Rathod, P., Bhargava, R., Irfan, M., Aly, R., Mohammad Saleh Al Gahtani, H., Salam, Z., Chau, S., Paterson, T., Turner, B., Gorbunova, V., Klymchuk, V., & Phiri, P. (2023). Effects of cumulative COVID-19 cases on mental health: Evidence from multi-country survey. *World journal of psychiatry*, *13*(7), 461–477.

Rauchman, S. H., Mendelson, S. G., Rauchman, C., Kasselman, L. J., Pinkhasov, A., & Reiss, A. B. (2022). Ongoing use of SSRIs does not alter outcome in hospitalized COVID-19 patients: A retrospective analysis. *Journal of Clinical Medicine*, *11*(1), 70.

Rojas-García, A., Turner, S., Pizzo, E., Hudson, E., Thomas, J., & Raine, R. (2018). Impact and experiences of delayed discharge: A mixed-studies systematic review. *Health Expectations*, *21*(1), 41–56.

Setola, N., Naldi, E., Arnetoli, M. V., Marzi, L., & Bologna, R. (2022). Hospital responses to COVID-19: evidence from case studies to support future healthcare design research. *Facilities*, *40*(1/2), 131–145.

SeyedAlinaghi, S., Mirzapour, P., Dadras, O., et al. (2021). Characterization of sars-cov-2 different variants and related morbidity and mortality: a systematic review. *European Journal of Medical Research*, *26*, 51.

Sheffler, Z., Patel, P., & Abdijadid, S. (2023). Antidepressants. StatPearls [Internet]. Updated May 26, 2023.

Sheffler, Z. M., Reddy, V., & Pillarisetty, L. S. (2023). Physiology, neurotransmitters. StatPearls [Internet]. Updated 2023 May 1.

Sindet-Pedersen, C., Staerk, L., Lamberts, M., Gerds, T. A., Berger, J. S., Bonde, A. N., Pallisgaard, J. L., Hansen, M. L., Torp-Pedersen, C., Gislason, G. H., & Olesen, J. B. (2018). Use of oral anticoagulants in combination with antiplatelet(s) in atrial fibrillation. *Heart*, *104*(11), 912–920.

Stauning, M. A., Gür, D. J., Torp-Pedersen, C., & Tingleff, J. (2023). COVID-19 mortality among selective serotonin reuptake inhibitor users-results from a nationwide cohort. *Clinical Microbiology and Infection*, *29*(8), 1075–1082.

Stoltzfus, J. C. (2011). Logistic regression: A brief primer. *Academic Emergency Medicine*, *18*(10), 1099–1104.

Sukhatme, V. P., Reiersen, A. M., Vayttaden, S. J., & Sukhatme, V. V. (2021). Fluvoxamine: A review of its mechanism of action and its role in COVID-19. *Frontiers in pharmacology*, *12*, 652688.

SÚKL (2018). Často kladené otázky: Je možné opakovaný recept vybrat během jednoho dne? https://www.epreskripce.cz/faq/46-je-mozne-opakovany-recept-vybrat-behem-jednoho-dne.

Ústav zdravotnických informací a statistiky ČR (2020). Pokyny pro certifikaci a kódování mortality v souvislosti s onemocnění COVID-19.

Ústav zdravotnických informací a statistiky ČR (2021). Pokyny ke kódování onemocnění COVID-19 (aktualizace).

Ústav zdravotnických informací a statistiky ČR (2023). MKN-10 2024. https://mkn10.uzis.cz.

Ústav zdravotnických informací a statistiky ČR & MZČR (2022). Datová a informační základna pro management pandemie COVID-19. Vývoj epidemie COVID-19 a efekt přijatých opatření. Akutní lůžková péče v době epidemie COVID-19. https://drg.uzis.cz/res/file/konference-cz-drg-2022/13-dusek.pdf.

Valizadeh, J., Boloukifar, S., Soltani, S., Jabalbarezi Hookerd, E., Fouladi, F., Andreevna Rushchtc, A., Du, B., & Shen, J. (2023). Designing an optimization model for the vaccine supply chain during the COVID-19 pandemic. *Expert Systems with Applications*, *214*, 119009.

Vela, J. M. (2020). Repurposing sigma-1 receptor ligands for COVID-19 therapy? *Frontiers in Pharmacology*, *11*.

Všeobecná zdravotní pojišťovna (2024). Opakovací recept. Retrived April 23, 2024 from https://www.vzp.cz/o-nas/tiskove-centrum/otazky-tydne/opakovaci-recept.

Wang, H., Wei, Y., Hung, C. T., Jiang, X., Li, C., Jia, K. M., Leung, E. Y. M., Yam, C. H. K., Chow, T. Y., Zhao, S., Guo, Z., Li, K., Wang, Z., Yeoh, E. K., & Chong, K. C. (2023). Relationship between antidepressants and severity of SARS-CoV-2 Omicron infection: a retrospective cohort study using real-world data. *The Lancet regional health. Western Pacific*, *34*, 100716.

Wang, S. M., Han, C., Bahk, W. M., Lee, S. J., Patkar, A. A., Masand, P. S., & Pae, C. U. (2018). Addressing the side effects of contemporary antidepressant drugs: A comprehensive review. *Chonnam medical journal*, *54*(2), 101–112.

Washington State Department of Health (2023). What to do if you were potentially exposed to someone with COVID-19. Guidance Document. Retrived April 14, 2024 from https://doh.wa.gov/sites/default/files/2022-02/COVIDexposed.pdf.

Whittington, M. D., Pearson, S. D., Rind, D. M., & Campbell, J. D. (2022). The cost-effectiveness of remdesivir for hospitalized patients with COVID-19. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research*, *25*(5), 744–750.

Wilkinson, J. D., Mamas, M. A., & Kontopantelis, E. (2022). Logistic regression frequently outperformed propensity score methods, especially for large datasets: a simulation study. *Journal of Clinical Epidemiology*, *152*, 176–184.

Wooldridge, J. (2013). *Introductory Econometrics: A Modern Approach*. South-Western Cengage Learning. 5th ed.

World Health Organisation (2019). Anatomical Therapeutic Chemical (ATC) Classification. Retrived April 14, 2024 from https://www.who.int/tools/atc-ddd-toolkit/atc-classification.

World Health Organisation (2020). International guidelines for certification and classification (coding) of COVID-19 as cause of death. Retrived April 15, 2024 from https://www.who.int/publications/m/item/international-guidelines-for-certification-and-classification-(coding)-of-covid-19-as-cause-of-death.

World Health Organisation (n.d.a). Defined Daily Dose (DDD). Retrived April 25, 2024 from https://www.who.int/tools/atc-ddd-toolkit/about-ddd.

World Health Organisation (n.d.b). International Classification of Diseases (ICD). Retrived April 14, 2024 from https://icd.who.int/browse10/2019/en.

World Health Organization (2020). Listings of who's response to COVID-19. Retrived April 3, 2024 from https://www.who.int/news/item/29-06-2020-covidtimeline.

World Health Organization (2021a). Oxygen. Retrived April 30, 2024 from https://www.who.int/news-room/questions-and-answers/item/oxygen.

World Health Organization (2021b). The true death toll of COVID-19: estimating global excess mortality. Retrived March 24, 2024 from https://www.who.int/data/stories/the-true-death-toll-of-covid-19-estimating-global-excess-mortality.

World Health Organization (2022a). Fact sheet: Mental disorders. Retrived March 15, 2024 from https://www.who.int/news-room/fact-sheets/detail/mental-disorders.

World Health Organization (2022b). Fact sheet: Mental health. Retrived January 15, 2024 from https://www.who.int/news-room/fact-sheets/detail/mental-health-strengthening-our-response.

World Health Organization (2023). Clinical management of COVID-19: Living guideline, 18 august 2023. Living Guideline WHO/2019-nCoV/clinical/2023.2, World Health Organization, Geneva. Licence: CC BY-NC-SA 3.0 IGO.

World Health Organization (2024). COVID-19 epidemiological update, edition 166, 12 april 2024. *World Health Organization*. Retrived April 20, 2024 from https://iris.who.int/handle/10665/376598.

World Health Organization (n.d.). Pneumonia. Retrived April 8, 2024 from https://www.who.int/health-topics/pneumoniatab=tab_1.

Wouters, O. J., Shadlen, K. C., Salcher-Konrad, M., Pollard, A. J., Larson, H. J., Teerawattananon, Y., & Jit, M. (2021). Challenges in ensuring global access to COVID-19 vaccines: production, affordability, allocation, and deployment. *397*(10278), 1023–1034.

Yang, J., Xie, M., & Goh, T. (2011). Outlier identification and robust parameter estimation in a zero-inflated poisson model. *Journal of Applied Statistics*, *38*, 421–430.

Zuur, A., Saveliev, A., & Ieno, E. (2012). *Zero Inflated Models and Generalized Linear Mixed Models with R.* Highland Statistics Ltd. book series. Highland Statistics Limited.

Appendix A

Risk factors

Table A.1: Coding algorithm for risk factors

Risk factor	**ICD-10 codes**	**ATC codes**
Diabetes mellitus		A10
Coronary artery disease	I20-I25	
Liver diseases	B15-B19, C22, D68A4, I982, K70-K77, Q618A, Z944	
Oncological treatment	D00-D09	L02
Hypertension		CO2A, CO2B, CO2C, CO2L, C03A, C03B, C03D, C03E, C03DA, C07B, C07C, C07D, C08G, C02DA, C09BA, C09DA, C02DB, C02DD, C02DG, C07A, C07B, C07C, C07D, C07F, C08, C09BB, C09DB, C09AA, C09BA, C09BB, C09CA, C09DA, C09DB, C09XA02, C09XA52
Asthma/COPD	J40-J47	R03A, R03BB, R03DC
Chronic kidney diseases	N02-NO8, N11, N12, N14, N18, N19, N26, N158, N159, N160, N162, N163, N164, N168, Q61, E102, E112, 132, E142	
Stroke	I63, I64	
Obesity	E66, U59	A08

Source: Sindet-Pedersen et al. (2018); ÚZIS (2023)

Appendix B

Results

Table B.1: GVIFs for Model with *cci_score*

	GVIF	Df	GVIF^(1/(2*Df))
female	1.049	1	1.024
age	1.210	4	1.024
SSRI	1.061	1	1.030
cci_1to2	1.224	1	1.106
cci_3to4	1.223	1	1.106
cci_4plus	1.068	1	1.033
waves	1.002	3	1.000

Table B.2: GVIFs for model with *chronic1*

	GVIF	Df	GVIF^(1/(2*Df))
female	1.049	1	1.024
age	1.257	4	1.029
SSRI	1.048	1	1.024
chronic1	1.204	1	1.097
waves	1.002	3	1.000

Table B.3: Logit results for model including antidepressants

	Dependent variable:	
	hospitalisation	
	(1)	(2)
female	−0.540*** (0.013)	−0.568*** (0.013)
age		
30-45	1.178*** (0.065)	1.213*** (0.065)
45-60	2.118*** (0.061)	2.283*** (0.061)
60-75	3.449*** (0.061)	3.774*** (0.060)
75plus	4.265*** (0.061)	4.563*** (0.060)
antidepressants	**0.148*** (0.018)	**0.105*** (0.018)
diabetes	0.482*** (0.016)	
COPD_asthma	0.592*** (0.016)	
hypertension	0.440*** (0.016)	
IHD	0.140 (0.125)	
stroke	0.188*** (0.026)	
oncological_treatment	0.212*** (0.018)	
chronic_kidney_diseases	0.382*** (0.023)	
liver_diseases	0.100** (0.034)	
obesity	0.533*** (0.061)	
cci_score		
cci_1to2		0.549*** (0.015)
cci_3to4		0.794*** (0.022)
cci_4plus		0.806*** (0.038)
waves		
wave1	0.561*** (0.066)	0.528*** (0.065)
wave2	−0.182*** (0.014)	−0.173*** (0.014)
wave3	−0.323*** (0.023)	−0.303*** (0.022)
Constant	−5.761*** (0.059)	−5.723*** (0.059)
Observations	612,368	612,368
Log Likelihood	−89,172	−90,426
Akaike Inf. Crit.	178,382	180,878
R-squared	0.240	0.230
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table B.4: Logit results for model including antidepressants

	Dependent variable:	
	death_covid	
	(1)	(2)
female	−0.809*** (0.024)	−0.822*** (0.024)
age		
65-75	2.819*** (0.047)	2.987*** (0.044)
75-85	3.826*** (0.046)	3.964*** (0.043)
85plus	4.453*** (0.047)	4.498*** (0.045)
antidepressants	**0.280**** (0.028)	**0.199**** (0.029)
diabetes	0.423*** (0.025)	
COPD_asthma	0.290*** (0.028)	
hypertension	0.401*** (0.031)	
IHD	0.337* (0.159)	
stroke	0.425*** (0.034)	
oncological_treatment	0.099*** (0.028)	
chronic_kidney_diseases	0.538*** (0.033)	
liver_diseases	0.330*** (0.058)	
obesity	0.303* (0.145)	
cci_score		
cci_1to2		0.594*** (0.028)
cci_3to4		0.993*** (0.034)
cci_4plus		1.241*** (0.050)
waves		
wave1	0.463*** (0.108)	0.414*** (0.108)
wave2	0.079*** (0.023)	0.089*** (0.023)
wave3	−1.131*** (0.056)	−1.110*** (0.056)
Constant	−6.364*** (0.041)	−6.298*** (0.041)
Observations	615,497	615,497
Log Likelihood	−31,638	−31,711
Akaike Inf. Crit.	63,312	63,447
R-squared	0.325	0.323
Note:		*p<0.1; **p<0.05; ***p<0.01

Table B.5: Logit results for model including *age* as ordinal variable

	Dependent variable:
	hospitalisation
female	-0.554^{***} (0.013)
ordinal_age	
25-29	0.527^{***} (0.118)
30-34	0.954^{***} (0.108)
35-39	1.298^{***} (0.101)
40-44	1.754^{***} (0.094)
45-49	2.077^{***} (0.092)
50-54	2.429^{***} (0.091)
55-59	2.709^{***} (0.090)
60-64	3.279^{***} (0.090)
65-69	3.806^{***} (0.089)
70-74	4.225^{***} (0.089)
75-79	4.537^{***} (0.090)
80-84	4.729^{***} (0.090)
85-89	4.712^{***} (0.091)
90-94	4.633^{***} (0.094)
95-99	4.453^{***} (0.114)
SSRIs	$\mathbf{0.076}^{***}$ (0.021)
diabetes	0.446^{***} (0.016)
COPD_asthma	0.582^{***} (0.016)
hypertension	0.359^{***} (0.016)
stroke	0.147^{***} (0.026)
oncological_treatment	0.158^{***} (0.018)
chronic_kidney_diseases	0.363^{***} (0.023)
liver_diseases	0.120^{***} (0.034)
obesity	0.600^{***} (0.062)
wave	
wave1	0.543^{***} (0.066)
wave2	-0.179^{***} (0.014)
wave3	-0.323^{***} (0.023)
Constant	-6.001^{***} (0.087)
Observations	612,368
Log Likelihood	−88,294.580
Akaike Inf. Crit.	176,647.200
Note:	*p<0.05; **p<0.01; ***p<0.001

Table B.6: Results of zero-inflated binomial models

	Negative binomial part		
	Dependent variable: days_in_hosp		
	(1)	(3)	(4)
female	−0.099*** (0.010)	−0.101*** (0.010)	−0.108*** (0.010)
age			
30-45	0.170** (0.055)	0.148** (0.055)	
45-60	0.377*** (0.052)	0.349*** (0.052)	
60-75	0.603*** (0.052)	0.578*** (0.052)	
75 plus	0.560*** (0.052)	0.538*** (0.052)	
65 plus			0.178*** (0.011)
SSRIs	**0.014** (0.015)	**0.014** (0.015)	**0.014** (0.015)
diabetes	0.044*** (0.011)		0.054*** (0.011)
COPD_asthma	0.018 (0.011)		0.022* (0.011)
hypertension	0.048*** (0.011)		0.073*** (0.011)
stroke	0.028 (0.018)		0.028 (0.018)
oncological_t.	0.033** (0.012)		0.037*** (0.012)
kidney_dis.	0.027* (0.015)		0.022 (0.015)
liver_dis.	0.002 (0.024)		0.006 (0.024)
obesity	0.044 (0.044)		0.036 (0.044)
wave1	0.441*** (0.043)		0.428*** (0.043)
wave2	−0.054*** (0.010)		−0.052*** (0.010)
wave3	−0.011 (0.016)		−0.006 (0.017)
chronic1		0.110*** (0.014)	
Constant	1.735*** (0.051)	1.711*** (0.051)	2.117*** (0.013)
	Logit part		
	(1)	(3)	(4)
female	0.528*** (0.013)	0.546*** (0.013)	0.484*** (0.013)
age			
30-45	−1.184*** (0.066)	−1.153*** (0.066)	
45-60	−2.113*** (0.062)	−2.049*** (0.062)	
60-75	−3.432*** (0.062)	−3.431*** (0.062)	
75 plus	−4.251*** (0.063)	−4.287*** (0.062)	
65 plus			−1.958*** (0.016)
SSRIs	**−0.098***** (0.021)	**−0.113***** (0.021)	**−0.182***** (0.021)
diabetes	−0.480*** (0.016)		−0.498*** (0.016)
COPD_asthma	−0.598*** (0.016)		−0.584*** (0.016)
hypertension	−0.442*** (0.016)		−0.756*** (0.016)
stroke	−0.190*** (0.026)		−0.264*** (0.026)
oncological_t.	−0.210*** (0.018)		−0.263** (0.018)
kidney_dis.	−0.386*** (0.023)		−0.414*** (0.023)
liver_dis.	−0.101** (0.034)		−0.122*** (0.035)
obesity	−0.541*** (0.062)		−0.490*** (0.062)
chronic1		−0.861*** (0.018)	
wave1	−0.543*** (0.066)	−0.551*** (0.066)	−0.536*** (0.066)
wave2	0.184*** (0.014)	0.187*** (0.014)	0.157*** (0.014)
wave3	0.322*** (0.023)	0.313*** (0.023)	0.331*** (0.023)
Constant	5.717*** (0.061)	5.785*** (0.061)	4.033*** (0.016)
Observations	612,362	612,362	612,362
Log Likelihood	−180,830.900	−181,961.100	−183,505.800

Note: *p<0.05; **p<0.01; ***p<0.001

Table B.7: Results of zero-inflated binomial model including *age* as ordinal variable

	Dependent variable: days_in_hosp	
	Count part	Binary part
female	−0.095*** (0.010)	0.484*** (0.013)
20-24	0.678*** (0.153)	
25-29	1.550*** (0.131)	
30-34	1.742*** (0.122)	
35-39	1.944*** (0.116)	
40-44	1.989*** (0.112)	
45-49	2.102*** (0.110)	
50-54	2.219*** (0.110)	
55-59	2.296*** (0.110)	
60-64	2.424*** (0.109)	
65-69	2.445*** (0.109)	
70-74	2.452*** (0.109)	
75-79	2.459*** (0.109)	
80-84	2.402*** (0.109)	
85-89	2.370*** (0.109)	
90-94	2.376*** (0.111)	
95-99	2.236*** (0.121)	
65 plus		-1.921*** (0.016)
SSRIs	**0.015** (0.015)	**−0.178***** (0.021)
diabetes	0.037*** (0.011)	−0.496*** (0.016)
COPD_asthma	0.015 (0.011)	−0.585*** (0.016)
hypertension	0.035*** (0.012)	−0.722*** (0.016)
stroke	0.026 (0.018)	−0.265*** (0.026)
oncological_treatment	0.027** (0.012)	−0.259*** (0.018)
kindey_diseases	0.030** (0.015)	−0.414*** (0.023)
liver_diseases	0.004 (0.025)	−0.113*** (0.035)
obesity	0.090** (0.045)	−0.494*** (0.062)
wave1	0.456*** (0.044)	−0.524*** (0.066)
wave2	−0.056*** (0.010)	0.162*** (0.014)
wave3	−0.012 (0.017)	0.333*** (0.023)
Constant	−0.098 (0.108)	3.961*** (0.016)
Observations	612,362	
Log Likelihood	−183,118.800	

Note: *p<0.05; **p<0.01; ***p<0.001