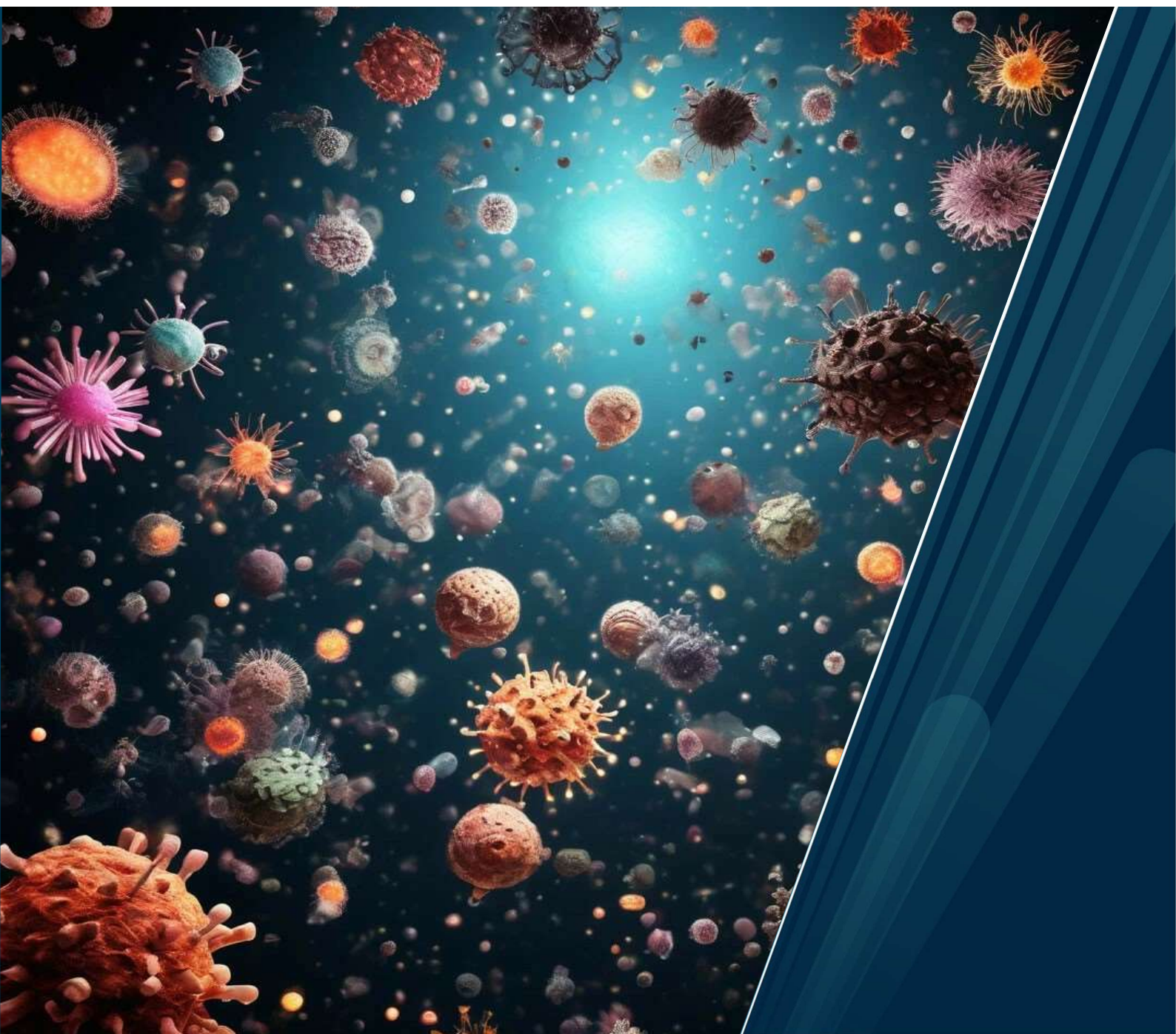![UiT The Arctic University of Norway]

Faculty of Health Sciences

COVID-19 morbidity, severity, and mortality in a population-based sample of adults in Arkhangelsk, Northwest Russia

Ekaterina Krieger

A dissertation for the degree of Philosophiae Doctor, August 2024

COVID-19 morbidity, severity, and mortality in a population-based sample of adults in Arkhangelsk, Northwest Russia

Ekaterina Krieger

A dissertation for the degree of Philosophiae Doctor, August 2024

Department of Community Medicine, Faculty of Health Sciences,

University of Tromsø –The Arctic University of Norway, Tromsø, Norway

Table of Contents

Preface

The issue resulting in this PhD project emerged in China in late December 2019. The COVID-19 pandemic changed the lives of many people around the world. At the time, I could not imagine how it would change mine.

I lost my father early in the pandemic. He always supported my steps in science and was pleased with my achievements, but he never knew that I was accepted into a PhD program.

A PhD position at UiT The Arctic University of Norway as part of the joint Arkhangelsk-Tromsø PhD program allowed me to pursue a PhD project in infectious diseases, a field of particular interest to me as a trained infectious disease physician. I am grateful to the UiT for this opportunity. Under the joint PhD program, the PhD scholarship was provided by the Northern State Medical University (NSMU) and the operating costs were covered by the UiT.

Due to COVID-19-related restrictions, my PhD journey lacked the usual get-togethers for educational and social activities. Most of the courses I attended were digital. Nevertheless, it was an interesting experience to observe the rapid and effective implementation of digital technologies in the educational process.

Arkhangelsk, Russia, August 2024   Ekaterina Krieger

Abstract (in English)

Objective. To investigate seroprevalence, spectrum of COVID-19 cases and factors associated with COVID-19 morbidity, severity, and mortality in a population-based sample of adults in Arkhangelsk, Northwest Russia.

Methods. Participants in the Know Your Heart study (2015-2017) were enrolled in a COVID-19 seroprevalence study (N=1348) in 2021 and followed up for COVID-19 diagnosis, vaccination, and all-cause mortality using registry data. Regression models were used to investigate factors associated with seropositivity, vaccination, adherence to non-pharmaceutical interventions (NPIs), symptomatic infection, hospitalization, and risk of death.

Results. One year after the start of the pandemic in Arkhangelsk, two-thirds of adults aged 40-74 years were seropositive, mainly due to infection. Those with infection-acquired immunity were more likely to be employed and less likely to smoke compared to seronegative individuals. Low adherence to NPIs was associated with male sex, employment, and low confidence in NPIs. Vaccination early in the pandemic was positively associated with older age and smoking, and negatively associated with low adherence to NPIs.

COVID-19 cases were 52.9% asymptomatic and 47.1% symptomatic, with 18.3% of symptomatic cases being hospitalized. Older age was associated with being symptomatic, whereas smoking was associated with being asymptomatic. Individuals older than 65 years and those with poor self-rated health were more likely to be hospitalized.

During the pandemic, the risk of death was 41.0% higher than in the pre-pandemic period. A greater increase in age-standardized all-cause mortality was observed in women than in men. Compared with the pre-pandemic period, women with obesity, angina, and kidney dysfunction and men with asthma and elevated cardiovascular biomarkers had a higher risk of death during the pandemic. Diabetes and smoking were factors associated with a higher risk of death in both sexes in both periods.

Conclusion. The results could contribute to the development of targeted prevention strategies to improve surveillance and health outcomes during future outbreaks.

Sammendrag (in Norwegian)

Formål. Å undersøke seroprevalens, spekteret av COVID-19-tilfeller og faktorer assosiert med COVID-19 sykelighet, alvorlighetsgrad og dødelighet i et befolkningsbasert utvalg av voksne i Arkhangelsk, Nordvest-Russland.

Metoder. Deltakere i "Know Your Heart"-studien (2015-2017) ble rekruttert til en COVID-19 seroprevalensstudie (N=1348) i 2021 og fulgt opp for COVID-19, vaksinasjon og mortalitet ved bruk av registerdata. Regresjonsmodeller ble brukt for å undersøke faktorer assosiert med seropositivitet, vaksinasjon, etterlevelse av ikke-farmasøytiske tiltak (NPIer), symptomatisk infeksjon, sykehusinnleggelse og risiko for død.

Resultater. Ett år etter pandemiens start i Arkhangelsk var to tredjedeler av voksne i alderen 40-74 år seropositive, hovedsakelig på grunn av infeksjon. De med immunitet oppnådd gjennom infeksjon var mer sannsynlig å være i arbeid og mindre sannsynlig å røyke sammenlignet med seronegative. Lav etterlevelse av NPIer var assosiert med mannlig kjønn, sysselsetting og lav tillit til NPIer. Vaksinasjon var positivt assosiert med høyere alder og røyking, og negativt assosiert med lav etterlevelse av NPIer.

COVID-19-tilfellene var 52.9% asymptomatiske og 47.1% symptomatiske, hvorav 18.3% av de symptomatiske tilfellene ble innlagt på sykehus. Høyere alder var assosiert med å være symptomatisk, mens røyking var assosiert med å være asymptomatisk. Individer eldre enn 65 år og de med dårlig selvrapportert helse hadde større sannsynlighet for å bli innlagt på sykehus. Under pandemien var risikoen for død 41.0% høyere enn i før-pandemisk periode. En større økning i aldersstandardiserte mortalitetsrater ble observert hos kvinner enn hos menn. Sammenlignet med før-pandemisk periode, hadde kvinner med fedme, angina og nyrefunksjonsfeil og menn med astma og forhøyede kardiovaskulære biomarkører en høyere risiko for død under pandemien. Diabetes og røyking var faktorer assosiert med høyere risiko for død i begge kjønn i begge periodene.

Konklusjon. Resultatene kan bidra til utviklingen av målrettede forebyggingsstrategier for å forbedre overvåkning og helseutfall under fremtidige utbrudd.

Abstract (in Russian)

Цель. Оценить серопревалентность, спектр клинических форм COVID-19 и факторы, связанные с заболеваемостью, тяжестью течения и смертностью от COVID-19 среди взрослого населения Архангельска, Северо-Запад России.

Методы. Участники исследования "Узнай свое сердце" (2015-2017 гг.) были включены в исследование серопревалентности к COVID-19 (N=1348) в 2021 году. Сведения о перенесённой COVID-19, вакцинации и смертности от всех причин собирались по данным регистров. Регрессионные модели использовались для изучения факторов, связанных с серопозитивностью, вакцинацией, приверженностью профилактическим мерам, симптоматической формой COVID-19, госпитализацией и риском смерти.

Результаты. Через год после начала пандемии в Архангельске две трети взрослого населения в возрасте 40-74 лет были серопозитивны к SARS-CoV-2, в основном за счет перенесённой инфекции. Постинфекционный иммунитет чаще наблюдался у трудоустроенных и реже у курящих. Мужской пол, трудоустроенность и неуверенность в эффективности профилактических мер были связаны с низкой приверженностью профилактическим мерам. Пожилые люди и курящие чаще вакцинировались в начале пандемии, а лица с низкой приверженность профилактическим мерам – вакцинировались реже. В 52.9% случаев COVID-19 протекала бессимптомно, в 47.1% - с симптомами, 18.3% участников с симптоматической формой COVID-19 были госпитализированы. Шансы симптоматического течения были выше у пожилых людей, а бессимптомного – у курящих. Лица старше 65 лет и участники с низкой самооценкой здоровья чаще госпитализировались. В период пандемии риск смерти от всех причин был на 41.0% выше, чем в предпандемическом периоде. У женщин наблюдалось более значительное увеличение стандартизованной по возрасту смертности в сравнении с мужчинами. Во время пандемии риск смерти был выше у женщин с ожирением, стенокардией и нарушением функции почек, а также у мужчин с бронхиальной астмой и повышенным уровнем биомаркеров сердечно-сосудистого риска. Диабет и курение были связанны с риском смерти, вне зависимости от пола, как в препандемический период, так и во время пандемии.

Заключение. Полученные результаты могут быть использованы для разработки стратегии оптимизации эпидемиологического надзора и минимизации последствий для здоровья населения при возникновении вспышек инфекционных заболеваний в будущем.

List of papers

The thesis is based on the following papers:

Krieger E, Kudryavtsev A, Sharashova E, Postoev V, Belova N, Shagrov L, et al. Seroprevalence of SARS-Cov-2 Antibodies in Adults, Arkhangelsk, Russia. Emerg Infect Dis.2022;28(2):463-465. doi: 10.3201/eid2802.211640 *(research letter)*

1. Krieger E, Sharashova E, Kudryavtsev AV, Samodova O, Kontsevaya A, Brenn T, Postoev V. COVID-19: seroprevalence and adherence to preventive measures in Arkhangelsk, Northwest Russia. Infect Dis (Lond). 2023;55(5):316-327

2. Krieger E, Kudryavtsev AV, Sharashova E, Samodova O, Kontsevaya A, Postoev V. Spectrum of COVID-19 cases in Arkhangelsk, Northwest Russia: findings from a population-based study linking serosurvey, registry data, and self-reports of symptoms. PlosOne, submitted.

3. Krieger E, Kudryavtsev AV, Sharashova E, Samodova O, Postoev V. Risk factors for all-cause mortality during the COVID-19 pandemic compared with the pre-pandemic period in an adult population of Arkhangelsk, Russia. Scientific Reports, submitted.

List of tables

List of figures

Terms and definitions

As the PhD thesis contains terms specific to infectious disease epidemiology, their definitions are explained below as intended in this thesis. The terms are listed in alphabetical order:

- ✓ attack rate – the proportion of susceptible individuals within a group of susceptible contacts of a primary (index) case who become infected within a given time period (1, 2);
- ✓ contact tracing (case finding) – the process of identifying, assessing, and managing individuals who have been exposed to an infected person (2);
- ✓ contagiousness – the probability that an infected person will transmit the infection to a susceptible person during contact or interaction, which depends on the infectiousness of the disease and factors related to human behavior, social interactions, and environmental conditions (2);
- ✓ cumulative incidence – the proportion of a group of people who experience the disease during a specified period of time (2);
- ✓ epidemic – the occurrence of disease, specific health behaviors, or other health-related events in a community or region that is significantly above normal levels (2);
- ✓ herd immunity – the resistance of a group or community to the spread of an infectious agent, achieved when a high proportion of individuals are immune to the disease, either through previous infection or vaccination (2);
- ✓ infectious period – the period of time during which infected individuals, whether symptomatic or not, shed a pathogen into the environment and new susceptible individuals can become infected (3);
- ✓ infectiousness – a characteristic of a disease that concerns the degree to which infected individuals can transmit the disease to others; refers to the presence and concentration of infectious agents in body fluids or secretions and depends on the type and stage of infection, viral load, and mode of transmission (2);
- ✓ infectivity – the ability of an infectious agent to cause a new infection in a susceptible individual (3);
- ✓ lockdown – a strict measure imposed by government authorities for a specified period of time during a pandemic that requires people to stay indoors and avoid or limit activities that involve public contact outside the home, with penalties for non-compliance (4, 5);
- ✓ movement restrictions – a measure that required people to stay at home, although they could leave for some circumstances, such as grocery shopping, emergency medical visits, brief

individual physical exercise, or walking the dog, but to return to their homes as soon as possible (5);

✓ non-pharmaceutical interventions – public health measures taken by governments to prevent the transmission of disease that are not based on medical products, such as drugs or vaccines (6);

✓ pandemic – an epidemic occurring over a very large area, crossing international boundaries, and usually affecting a large number of people (2);

✓ pathogen – the causative agent of an infectious disease (literally, a pathological process)(2);

✓ quarantine – the separation and restriction of movement of people who may have been exposed to a pathogen to reduce the risk of disease transmission (2);

✓ seroprevalence – the proportion of people having antibodies at a single time point or over a short period of time (7);

✓ social distancing – a measure that prevents infected people from having close physical contact with healthy people, thereby reducing the risk of disease transmission (8);

✓ susceptibility – the likelihood that an individual will develop an infectious disease after exposure to a pathogen (9);

✓ transmissibility – the ability of the pathogen to pass from one person to another (2);

✓ viral load – the amount of a virus present in a test sample, reflecting the replication of the virus in the infected individual (10);

✓ virulence (the degree of pathogenicity) – the ability of a pathogen to cause disease after having infected the host (11).

Abbreviations

CIs – confidence intervals

COVID-19 – the novel coronavirus disease 2019

DAGs – directed acyclic graphs

ELISA – enzyme-linked immunosorbent assay

ESSE-RF3 – the third nationwide survey "Epidemiology of Cardiovascular Diseases and their Risk Factors in Regions of the Russian Federation"

GGT – gamma-glutamyl transferase

HbA1c – glycated hemoglobin

HDL-C – high-density lipoprotein cholesterol

Hs-CRP – high-sensitivity C-reactive protein

Hs-TnT – high-sensitivity troponin T

HR – hazard ratio

ICD-10 – International Classification of Diseases 10th Revision

IgG – immunoglobulin G

KYH – Know Your Heart

LDL-C – low-density lipoprotein cholesterol

NPIs – non-pharmaceutical interventions

MIAC – Arkhangelsk Regional Medical Information Analytical Center

NSMU – Northern State Medical University

NT-proBNP – N-terminal pro-brain natriuretic peptide

OR – odds ratio

PCR – polymerase chain reaction

R_0 – basic reproductive number

R – effective reproductive number

SARS-CoV-2 – severe acute respiratory syndrome coronavirus 2

WHO – World Health Organization

1 Introduction

1.1 History of COVID-19

A novel pathogen emerged in Wuhan, the capital city of Hubei Province, China. By December 31, 2019, several cases of atypical respiratory disease of unknown etiology were reported, all linked to wild animals sold at the Huanan seafood wholesale market (southern China) (12). However, according to phylogenetic studies, the transmission of the novel disease to humans likely occurred earlier, between October and November 2019 (13, 14). On January 7, 2020, the Chinese Center for Disease Control and Prevention identified the causative agent of the novel infection as a coronavirus (15).

To date, more than 30 coronavirus strains have been identified, six of which have caused human infections, including two highly pathogenic coronaviruses: severe acute respiratory syndrome coronavirus (SARS-CoV) in 2002-2003 and Middle East respiratory syndrome coronavirus (MERS-CoV) in 2012 (16). The new virus that emerged at the end of 2019 was initially named 2019 novel coronavirus (2019-nCoV), but was later renamed severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) due to its genetic relationship to SARS-CoV (17).

SARS-CoV-2 is of zoonotic origin, with bats likely serving as natural reservoirs, but the precise role of animals in the disease transmission remains unclear (18, 19). According to the International Committee on Taxonomy of Viruses, SARS-CoV-2, the causative agent of COVID-19, belongs to the subgenus Sarbecovirus, the subfamily Orthocoronavirinae, the family Coronaviridae, and the order Nidovirales (17). Viruses in the family Coronaviridae were named for their crown-like appearance under the electron microscope, with club-shaped spikes covering their surface, resembling a crown or "corona" in Latin (20). The spikes on the surface of the virus include the spike (S) protein, which contains the receptor-binding domain responsible for attachment to the host cell via specific receptors, such as angiotensin-converting enzyme 2, allowing viruses to enter cells and transmit from cell to cell. Inside the viral envelope is a nucleocapsid, which consists of nucleocapsid (N) proteins bound to the single-stranded ribonucleic acid (RNA) genome. The S and N proteins are the primary targets of the antibody response following infection with SARS-CoV-2 (21). Mutations in the S protein have led to the emergence of new SARS-CoV-2 variants. Initially, SARS-CoV-2 variants were named based on the locations where they were identified. The

World Health Organization (WHO) later named variants using Greek letters, including Alpha (first identified in the United Kingdom), Beta (first identified in South Africa), Gamma (first identified in Brazil), Delta (first identified in India), and Omicron (first identified simultaneously in several countries). A more systematic approach was to name variants according to their Phylogenetic Assignment of Named Global Outbreak (PANGO) lineage: Alpha (B.1.1.7), Beta (B.1.351), Gamma (P.1), Delta (B.1.617.2), and Omicron (B.1.1.529) (22).

As SARS-CoV-2 was a novel virus to which everyone was susceptible, it spread rapidly from China to other countries around the world in early 2020. On January 30, 2020, the WHO declared the novel coronavirus disease 2019 (COVID-19) caused by the SARS-CoV-2 to be a public health emergency of international concern requiring the coordinated mobilization of resources by the international community (23). On March 11, 2020, the WHO declared the COVID-19 outbreak a pandemic (24). This is the second pandemic of the 21st century (following the 2009 influenza A (H1N1) pandemic) and the first pandemic caused by a coronavirus. The pandemic affected more than 200 countries on five continents. More than 700 million cases of COVID-19 and almost 7 million deaths were reported worldwide, ranking the COVID-19 pandemic fifth on the list of the deadliest epidemics and pandemics in history (25). Russia had the tenth highest number of COVID-19 cases in the world (26). During the three years of the pandemic, more than 21 million COVID-19 cases were registered in Russia, resulting in more than 380 thousand deaths (27). On May 5, 2023, the WHO declared the end of the COVID-19 pandemic (28). While the virus continues to circulate, the global public health emergency status has ended.

1.2 SARS-CoV-2 transmissibility and risk of infection

SARS-CoV-2, is mainly transmitted by respiratory droplets or direct contact. Isolation of the virus from blood samples and fecal swabs suggests other potential routes of transmission (29).

Transmissibility refers to the ability of the pathogen to pass from one person to another and is determined by the basic reproductive number (R_0). R_0 is the average number of new infections produced by an infectious individual in a 100% susceptible population (2). The transmissibility of SARS-CoV-2 is influenced by several factors, including the infectivity of the virus variant, the contagiousness of the infected individuals, the duration of

infectiousness, the susceptibility of the exposed individuals, the number and pattern of contacts between infected and exposed individuals, and the environmental factors affecting the virus during transmission (30).

R_0 can be calculated from the equation $R_0 = \beta \times k \times D$, which includes key parameters that affect SARS-CoV-2 transmissibility, such as the attack rate (β), which represents the proportion of contacts that result in new infections, the number of contacts during the infectious period (k), and the duration of the infectious period (D). The attack rate is often used as a proxy measure of infectiousness and susceptibility and depends on sociodemographic and environmental factors influencing transmission. The infectious period begins before the onset of symptoms and lasts up to 10-14 days. Infected individuals may be contagious regardless of the presence of symptoms, but the period of virus shedding is longer in patients with severe COVID-19 (30, 31). The risk of transmission is influenced by contact patterns, including duration of contact, frequency of contact, and proximity to the infected person. Close contacts of the infected individuals, such as household members, are at highest risk of infection, with attack rates ranging from 4,0% to 35,0% (32). The risk of contracting the virus is higher for individuals with large families and those who live and work in adverse conditions such as poorly ventilated and overcrowded environments. The number of contacts depends on the pattern of social interactions and may be higher in occupations with more social mixing and longer working hours, leading to an increased risk of exposure (33).

At the beginning of the pandemic, when there were no external interventions to control the spread of the virus, the WHO estimated the R_0 of the Wuhan SARS-CoV-2 variant to be 1.4 to 2.4 (34). Subsequent studies found that the R_0 ranged from 1.5 to 6.6, varying between countries (35).

An R_0 greater than 1 indicated exponential spread of infection in a susceptible population (36). As the pandemic progressed, some individuals acquired immunity, either through prior infection or vaccination. As a result, not all contacts resulted in infection, leading to a decrease in the average number of new cases per case of infection. The transmissibility of the infection during an epidemic is assessed using the effective reproductive number (R), which can be calculated as $R = R_0 \times X$, where X is a proportion of the susceptible population. R greater than 1 indicates an increasing number of cases, suggesting continuous and sustained transmission. R equal to 1 indicates a stable number of cases, while R less than 1 suggests a decrease in cases (36).

Predictions made early in the COVID-19 pandemic were that achieving population immunity

greater than 70% would lead to herd immunity and reduce the COVID-19 R below 1, thereby slowing the spread of infection (37). However, the protective role of antibodies against future infections remained uncertain, as SARS-CoV-2 was constantly mutating and developing new adaptation mechanisms (38). Variants with increased levels of transmissibility, infectivity, or virulence have been categorized as variants of concern. The WHO has identified five variants of concern: Alpha, Beta, Gamma, Delta, and Omicron (39). The R_0 of Delta is estimated to be between 3.2 and 8, with an average of 5.0, and was higher than the R_0 of any previously reported variant (40). The Omicron variants had the highest transmissibility, about three times higher than the Delta variant (41).

1.3 Non-pharmaceutical interventions

When the pandemic was declared, most countries implemented non-pharmaceutical interventions (NPIs) as the primary strategy to control the pandemic. Keeping the R below 1 with NPIs was an important goal until vaccines against SARS-CoV-2 became available. NPIs were applied at the global, national, or individual level, targeting key parameters of SARS-CoV-2 transmissibility, such as attack rate and number of contacts between people.
NPIs included travel-related measures, community NPIs, environmental NPIs, and personal NPIs (42). Travel-related measures were closure of international borders, testing and quarantine for incoming travelers. Early in the pandemic, countries advised against travel to China. As the pandemic progressed, COVID-19 cases were imported from various parts of the world. Some countries implemented testing and 14-day quarantines for travelers, initially targeting countries with high numbers of cases and eventually closing borders to travelers from all destinations (43).
Following the detection of the first COVID-19 case in a country, community NPIs were implemented to reduce person-to-person interactions, including lockdown, movement restrictions, modified work and educational arrangements, closure of non-essential services, isolation of infected persons, contact tracing, and quarantine (44, 45). Environmental NPIs such as increased air exchange, surface disinfection, and ultraviolet lighting were implemented to protect people from exposure to SARS-CoV-2. Wearing face masks in public places or on public transport, using hand sanitizers, and practicing social distancing provided individual-level protection.
Countries imposed combinations of NPIs with varying degrees of stringency, and

implementation varied over time, with periods of relaxation or reinforcement. Some countries, such as Greece and Norway, implemented most of the NPIs early (5). Others, such as Belgium, Italy, the United Kingdom, Canada, Russia, and the United States, implemented NPIs with a delay of several weeks after the first COVID-19 case in the country, while Brazil and Sweden did not implement most NPIs (5). Delayed implementation of NPIs had little effect on reducing the spread of COVID-19. Countries that delayed NPIs by more than three weeks after the first case had similar trends in COVID-19 cases to those that did not use NPIs (5).

The strictest restrictions were implemented by most countries in April-May 2020 (46). After the first wave of COVID-19, many countries relaxed their restrictions, leading in some cases to severe second waves (5). During the second waves in late 2020 and early 2021, some countries increased the stringency of NPIs (46).

The effectiveness of NPIs varied between countries, reflecting country-specific factors such as demographic structure, level of urbanization, population density, and the timing of NPIs implementation. Closing educational institutions, non-essential services, restricting movement and public gatherings have been identified as the most effective ways to reduce transmission (47, 48). Countries with younger populations, higher levels of urbanization, higher population density, and larger households tended to have lower effectiveness of NPIs (49). The effectiveness of NPIs also depended on public adherence and government monitoring of NPIs use (50, 51). Trust in authorities and understanding of the value of NPIs increased public confidence in the effectiveness of these measures (52). The mentality of the population, self-discipline, and social responsibility in the use of individual NPIs played an important role in controlling the spread of the virus. In China's neighboring countries, such as Taiwan and Japan, social distancing and personal hygiene measures were adopted without government enforcement due to previous experiences with SARS-CoV and MERS-CoV epidemics and a national mentality that encouraged the use of personal protective measures (5). In Norway, the government did not impose strict movement regulations because social distancing was already a Norwegian cultural behavior, in which people respect the personal space of others (5).

Adherence to NPIs was also influenced by factors such as availability of resources (e.g., face masks and hand sanitizers), perceived risk of infection, individual beliefs and attitudes, and perceived effectiveness of NPIs in controlling the disease transmission (53). Fear of contracting COVID-19 has been found to be the strongest predictor of desired behavior

change and adherence to NPIs (54). Higher adherence to NPIs has been demonstrated among the elderly, women compared to men, individuals with higher education and income, non-smokers, those living alone, and those with chronic diseases (55-57). There is some evidence to suggest that individuals who believed they had previously been infected with COVID-19 were less likely to adhere to NPIs because they asumed this would confer immunity and protection against further infection (58).

1.4 The spread of COVID-19 in Russia

During the first year of the pandemic, 3,159,297 cases of COVID-19 were recorded in Russia, with an incidence rate of 2152.63 per 100,000 (59). The first case of COVID-19 in Russia was detected on January 31, 2020 in a region bordering China. The virus was imported into the European part of the country from Italy on March 2, 2020 (59). In the Arkhangelsk region in northwestern Russia (population 1.1 million in January 2021) (60), which was the study site, the first case of COVID-19 occurred on March 17, 2020 (61). By April 16, 2020, COVID-19 cases were recorded in all regions of Russia (59).

Like most countries, Russia implemented NPIs as the primary strategy to control the transmission of infection. The NPIs were generally in line with international recommendations and were regulated by the Russian Federal Service for Surveillance on Consumer Rights Protection and Human Well-being (Rospotrebnadzor). These NPIs included travel restrictions, closure of educational institutions and non-essential services, restrictions on commercial activities, and mandatory use of face masks in public places and on public transport (62). The use of gloves for individual COVID-19 prevention has not been implemented worldwide, while in Russia the use of gloves was recommended by Rospotrebnadzor until February 4, 2022 (43, 63, 64).

From the end of March to the beginning of May 2020, nationwide paid non-working days were introduced across the country (65-67). During this period, movement restrictions, known as the "stay-at-home" regime, were introduced in most regions of Russia. Subsequently, individuals aged 65 years and older, people with chronic conditions such as diabetes, respiratory and cardiovascular diseases, chronic kidney disease, neoplasms, organ and tissue transplant recipients, and pregnant women, were advised to self-isolate (stay at home) (62). These NPIs made it possible to slow the virus transmission, thereby reducing the burden on the healthcare system, especially in areas with limited critical care capacity, and minimizing

deaths while waiting for effective vaccines or antiviral treatment. Other NPIs implemented in Russia and worldwide included isolation of infected individuals, contact tracing requiring testing of close contacts of confirmed COVID-19 cases, and quarantine for those exposed to the virus (44). Isolation and quarantine were mandatory measures enforced by government regulations, and failure to comply with these NPIs was considered an administrative offense. In the early stages of the pandemic in Russia, all detected COVID-19 patients were isolated by hospitalization (68). The increase in the number of cases over time led to changes in the hospitalization criteria, which became more focused on the severity of COVID-19 rather than isolation purposes (59).

The centralized NPIs policy was combined with regional variations, allowing regions to tailor NPIs to their specific local epidemiological and economic situation, in accordance with Presidential Decree #316 of May 11, 2020 (69). This decentralized approach was not unique. Some other countries, such as the United States, used a similar strategy (5). The severity of movement restrictions varied by region, with some regions (e.g., Moscow) implementing strict measures such as a lockdown for a period of time. Moscow was the only city in Russia to introduce IT technology to monitor the geolocation of confirmed COVID-19 cases and their close contacts through a mobile application. In central cities of Russia, wearing face masks and gloves in public places and on transport was enforced by governors' decrees under threat of fines.

In Arkhangelsk, the local government implemented NPIs to prevent COVID-19 transmission on March 18, 2020 (70). These measures were consistent with the nationwide NPIs implemented in Russia in accordance with Rospotrebnadzor's recommendations, as described above (62). In total, 43,679 COVID-19 cases (incidence rate – 3970.81 per 100 000) and 529 deaths (1.2%) were recorded in the Arkhangelsk region in 2020 (61). COVID-19 transmission in Arkhangelsk generally reflected the pattern observed throughout Russia (59, 61). Arkhangelsk experienced two waves of COVID-19. The first wave occurred from March 19, 2020 to July 2, 2020 (61). By the end of summer 2020, the incidence of COVID-19 decreased and stabilized. Beginning in September 2020, educational institutions resumed offline activities, and commercial and recreational activities were reinstated with strict safety measures, including maintaining a social distance (1.5-meter rule), wearing face masks, using hand sanitizers, and performing non-contact thermometry (71). The second wave of COVID-19 in Arkhangelsk (September 20, 2020 – March 3, 2021) followed the relaxation of NPIs and the emergence of the Delta strain, which replaced the Wuhan strain as the predominant

variant of SARS-CoV-2 during the first six months of the pandemic (61). The peak of the second wave was reached on December 10, 2020, with the highest daily incidence rate recorded at 37.3 per 100,000 population (71). During the second wave, Russia did not strengthen the NPIs, unlike other countries that experienced a significant increase in cases after relaxing the previously implemented NPIs (46).

In early 2021, the emergence of new highly transmissible virus variants, coupled with further relaxation of NPIs and increased availability of laboratory testing, led to an increase in COVID-19 cases in Arkhangelsk, as in all Russian regions (72, 73). In 2021, the Arkhangelsk region ranked ninth among all the federal subjects of Russia in the number of COVID-19 cases per 100,000 population (74).

1.5 Testing strategies, seroprevalence and factors associated with seropositivity

Testing strategies represent a range of approaches, including diagnostic testing to identify current infections and contact tracing, and serological testing to assess previous infections and immunity in the population. Polymerase chain reaction (PCR) tests are used to detect the presence of the virus in the nasopharyngeal cavity and identify those who are infectious, whether symptomatic or asymptomatic (75). The WHO interim guidance for global COVID-19 surveillance, released on January 31, 2020, categorizes COVID-19 cases as confirmed, probable, and suspected (76). Confirmed cases are individuals with a positive PCR test for the SARS-CoV-2 virus. Probable cases include those with compatible symptoms and exposure to confirmed cases but without confirmatory testing. Suspected cases are individuals with COVID-19 symptoms who are awaiting or undergoing testing for confirmation. The WHO introduced two emergency International Classification of Diseases, 10th Revision (ICD-10) codes on March 25, 2020, to classify COVID-19 cases worldwide. These are U07.1 "COVID-19 virus identified", which indicates infection confirmed by laboratory testing, and U07.2 "COVID-19 virus not identified", for cases diagnosed clinically or epidemiologically without laboratory testing (77).

In Russia, the Federal Registry of COVID-19 Patients (referred to as the COVID-19 case registry) was established nationwide for COVID-19 surveillance purposes, as regulated by Russian Government Decree #373 of March 31, 2020 (78). The registry collected data on

diagnosed COVID-19 cases from patients' electronic health records in the information systems of state non-military health services (79). The data accumulated in the case registry were based on positive PCR test results rather than clinical symptoms.

In the early stages of the pandemic, the availability of tests, criteria for testing, and priority groups for testing varied widely among countries (80). The first Russian diagnostic test systems for COVID-19 were developed and registered on February 11, 2020 (59). Eligibility for testing was limited, and the testing strategy was initially symptom-based, using symptoms such as fever, cough, and dyspnea as indications for testing (68). The symptom-based approach was only able to detect a fraction of the COVID-19 cases, mainly the more severe symptomatic cases (81). In addition, testing was initially prioritized for high-risk groups, such as those aged 65 years and older and those with chronic diseases. As a result, even symptomatic COVID-19 cases seeking medical care may have remained undiagnosed. As test availability increased, PCR screening expanded from symptom-based testing to include groups at high risk of exposure, such as travelers arriving in Russia from abroad and healthcare workers. Contact tracing was implemented in Russia, as in many countries, targeting PCR-positive cases, regardless of symptoms (68). PCR testing was performed by government laboratories and some private clinics (59, 61). Voluntary testing and mandatory testing for travelers arriving from abroad were not covered by national health insurance. Despite the availability of testing, test performance may affect the detection of COVID-19 cases. A recent meta-analysis showed a pooled sensitivity of 91.1% (95% CI: 88.9 to 93.2) and a pooled specificity of 95.6% (95% CI: 95.2 to 96.0) for the detection of SARS-CoV-2 in nasopharyngeal swabs using different PCR techniques (82). The accuracy of PCR testing may be influenced by the timeliness and expertise of specimen collection (83). Due to untimely or incorrect specimen collection and limited viral replication in upper respiratory epithelial cells, PCR testing could potentially be false negative, resulting in under-detection of COVID-19 cases (84).

The detection rate refers to the number of positive tests among all tests done and reflects the effectiveness of the testing strategy. During the first six months of the pandemic, the global detection rate was estimated to be 9.8% (85). The highest detection rates were in Australia (66.8%) and Iceland (60.3%), which implemented widespread testing strategies early in the pandemic. In Russia, the detection rate was estimated to be 25.4% (85).

Limited availability and prioritization of testing for vulnerable groups and severe cases, combined with high rates of asymptomatic and mild infection, as well as test accuracy issues,

have led to underestimation of COVID-19 spread (86). The WHO has recommended conducting serological surveys in representative samples of the general population to assist in the retrospective assessment of the spread of COVID-19 (87).

Serological tests detect specific antibodies to SARS-CoV-2, which serve as exposure biomarkers to identify those who have been previously infected, regardless of their symptomatic status (3). Seroprevalence, which refers to the proportion of people having antibodies at a single time point or over a short period of time, can provide more accurate estimates of the extent of COVID-19 infection than those based on the number of positive tests detected by the healthcare service (86, 88). Considering antibodies as a reliable measure of past infection, seroprevalence can serve as a proxy for the cumulative incidence of COVID-19 up to the time of serological survey (89).

During the first year of the COVID-19 pandemic, several random population serological surveys were conducted. By early 2021, 34.6% of the population sampled in England tested positive for antibodies to SARS-CoV-2, while the estimated seroprevalence in Norway and Iceland was less than 1.0% (90-92).

Few serological studies had been conducted in Russia until mid-2021, before the start of this PhD project. Between June and December 2020, the average COVID-19 seroprevalence in Russia was estimated to be 19.2%, with regional variations as follows: Irkutsk (Siberia) - 5.8%, Khabarovsk (Far East) - 19.6%, Moscow - 22.1%, Saint Petersburg - 26.0%, Astrakhan (south) - 27.3%, Murmansk (northwest) - 31.2%, and Kaliningrad (the westernmost part) - 50.2% (93). The study conducted in Chelyabinsk (Ural) from September to December 2020, targeting high-risk groups such as healthcare workers, education personnel, and supermarket employees, found that 25.0% were seropositive for SARS-CoV-2 (94).

Other studies found that serological status was associated with several demographic and behavioral factors. Male sex, living in a crowded household, using public transport, or having a high level of social interaction were associated with being seropositive for SARS-CoV-2 (50, 95, 96). People who were employed as essential workers were more likely to be infected with SARS-CoV-2 (97, 98). Conversely, older age groups and those with chronic diseases, may have taken precautions to reduce the risk of infection and were more likely to remain seronegative (90, 99, 100). Smokers may be at increased risk of infection due to frequent hand-to-mouth contact and the need to remove face masks to smoke, which reduces the effectiveness of the masks (101). Nevertheless, some studies have found that smoking is negatively associated with seropositive status (99, 102, 103).

1.6 Spectrum of COVID-19 severity and associated factors

SARS-CoV-2 is a respiratory pathogen that can colonize the respiratory tract and cause asymptomatic or symptomatic infection. Infected individuals are contagious regardless of symptomatic status (31, 104). Symptomatic COVID-19 cases have an average incubation period of 1-14 days, with symptoms typically appearing 3-14 days after exposure. SARS-CoV-2 causes non-specific signs and symptoms such as fever, myalgia, sore throat, rhinorrhea and cough. Anosmia (loss of sense of smell) and ageusia (loss of sense of taste) are observed in up to 60% of cases (105). As the disease progresses, the virus invades the lower respiratory tract, leading to pulmonary involvement characterized by pneumonia and dyspnea. Dyspnea typically coincides with the appearance of ground-glass opacities on chest X-rays and computed tomography scans. Approximately 20% of symptomatic patients may develop acute respiratory distress syndrome and 10% may develop multiple organ failure (106, 107).

The severity of COVID-19 depends on the SARS-CoV-2 variant, viral load, and host factors such as sex, age, and pre-existing chronic diseases (96, 108, 109). Chronic conditions such as cardiovascular diseases (hypertension, ischemic heart disease), metabolic disorders (diabetes, obesity), pulmonary diseases, and cancer can increase susceptibility and exacerbate the severity of COVID-19 (110, 111), leading to complications such as acute respiratory distress syndrome, sepsis, septic shock, and multiorgan failure (112). Conversely, COVID-19 may worsen pre-existing chronic conditions, particularly cardiovascular diseases, by potentially damaging cardiomyocytes and being pathogenically associated with thrombovascular events (113-115). SARS-CoV-2 may induce endothelial dysfunction by directly damaging endothelial cells or by inducing inflammation (113). Endothelial dysfunction increases the risk of clot-related complications such as myocardial infarction, stroke, and pulmonary embolism. Individuals with pre-existing chronic conditions are at higher risk for these life-threatening complications (116, 117).

In summary, the severity of COVID-19 is highly variable, ranging from asymptomatic or mild cases to severe illness and death (118). The infectious disease pyramid, a conceptual epidemiological model, can be used to illustrate the distribution of COVID-19 cases within a population (119)(Fig.1). It consists of several layers, each representing a group of individuals based on their infection status and degree of interaction with the healthcare system. Exposed individuals form the base of the pyramid and represent individuals susceptible to infection who have come into contact with the virus. As one moves up the pyramid, the size (number of

people) of the layers becomes smaller, reflecting the progression from exposure to infection, to clinical illness, to hospitalization, and ultimately to death. Asymptomatic cases are those who have been infected without symptoms, but can transmit the infection to others (31, 104). Symptomatic cases include individuals who develop a range of symptoms, from mild to severe, with more severe cases being more likely to seek healthcare. Hospitalized cases require medical attention due to the severity of their illness, while individuals who have died represent the top of the pyramid.
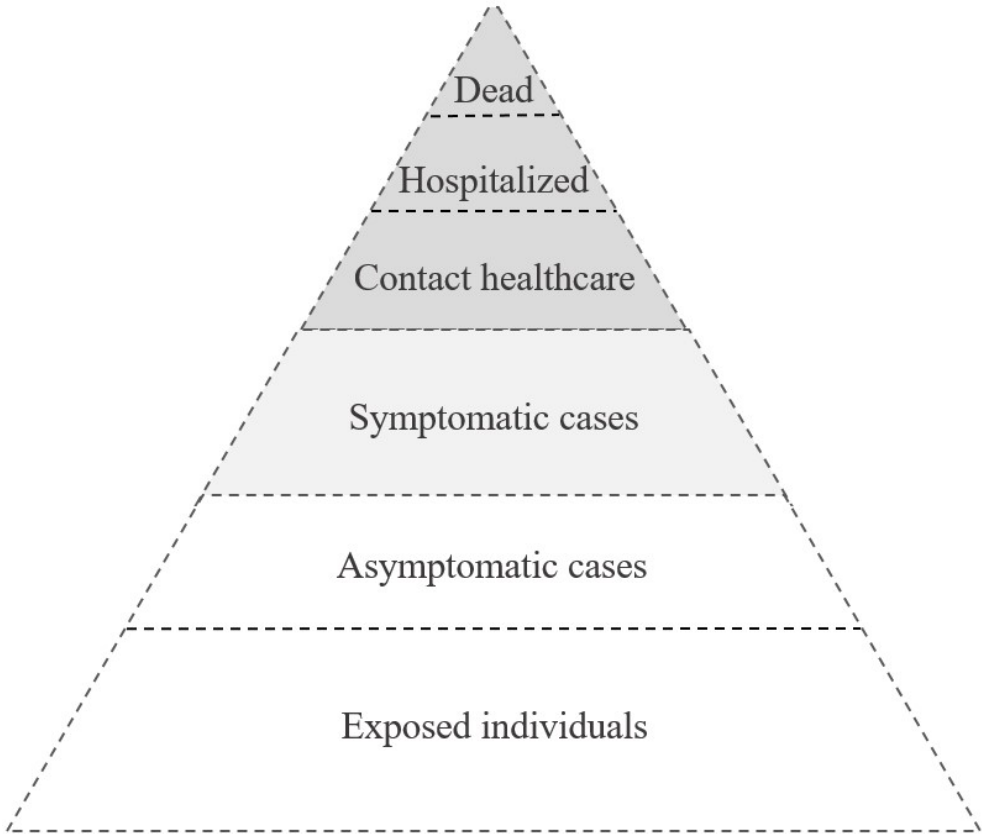
Figure 1. The infectious disease pyramid.

Early in the pandemic, most officially reported COVID-19 cases were symptomatic. Meta-analyses found that the proportion of asymptomatic COVID-19 cases detected by PCR varied significantly between countries depending on the testing strategy, with an overall proportion of 15.6-17.0% (55, 120). In Russia, the proportion of asymptomatic cases detected by the healthcare system was less than 7.0% (27, 59, 72). In Arkhangelsk, 14.1% of reported cases in 2020 were asymptomatic. Although the number of cases doubled in 2021, the proportion of asymptomatic cases decreased to 3.2% due to the cessation of contact tracing (73).

From 0.5 to 5.0% of COVID-19 patients develop severe or critical illness, depending on pre-

existing health conditions (121). In Russia, severe cases accounted for 3.4% of registered cases in 2020, 2.5% in 2021, and 1.1% in 2022, with the proportion of deaths not exceeding 2.0% (27, 59, 72). In Arkhangelsk, the proportion of deaths was 1.2% in 2020 and 3.3% in 2021 (61, 72).

1.7 Direct and indirect effects of the pandemic on mortality

The pandemic resulted in an overall increase in mortality, including causes directly related to COVID-19 and causes unrelated to infection that were indirect consequences of the pandemic (122, 123). The direct effects of the pandemic were deaths resulting from COVID-19 or its complications (124). The indirect effects were due to factors like limited access to health care, reduced screening activities, mandatory isolation, and avoidance of seeking care, which led to delayed diagnosis and treatment of some conditions and contributed to increased mortality from causes other than COVID-19 (125-127).

In 2020, the first year of the pandemic, Russia experienced more than 350,000 excess deaths, calculated as the difference between the observed and expected number of deaths for the year (128, 129). Previous research has shown that older people and individuals with chronic diseases were particularly vulnerable to both the direct and indirect adverse effects of the pandemic (122). They were at increased risk of death following a COVID-19 requiring hospitalization (112). According to the Federal State Statistics Service (Rosstat), the number of COVID-19 deaths (direct effect) in 2020 was approximately 145,000, accounting for about 40% of excess deaths (130). In the same year, the leading causes of death in Russia were cardiovascular diseases and neoplasms (59).

The WHO defines a COVID-19-related death as "*a death resulting from a clinically compatible illness, in a probable or confirmed COVID-19 case, unless there is a clear alternative cause of death that cannot be related to COVID-19 disease (e.g. trauma). There should be no period of complete recovery from COVID-19 between illness and death. A death due to COVID-19 may not be attributed to another disease (e.g. cancer) and should be counted independently of pre-existing conditions that are suspected of triggering a severe course of COVID-19*"(131).

The criteria for defining COVID-19-related deaths varied between countries due to differences in the implementation of international and national guidelines, resulting in limited

comparability in the reporting of COVID-19 deaths (80). Approaches to recording COVID-19 deaths varied depending on test availability and testing strategies, generally falling into one of two categories: clinical-based or test-based. The clinical-based approach, in which physicians can rely on their diagnostic expertise to certify a cause of death, was implemented in countries such as Belgium, France, Germany, and Indonesia. The test-based approach, which requires a positive laboratory test for SARS-CoV-2, was implemented in Austria, Italy, Korea, the Netherlands, Spain and the United Kingdom (80, 132). Although countries such as Cyprus, Greece, Romania, Russia, and Serbia included probable COVID-19 deaths in their definitions, a positive test result was still required in practice (80, 133). Countries that used a clinical-based approach were likely to record more COVID-19 deaths than those that relied on positive tests. Although the test-based approach is more precise and accurate in determining the cause of death, it may lead to underreporting by misclassifying untested individuals as non-COVID-19 deaths (117, 134).

The WHO has established guidelines for reporting COVID-19 deaths and a standardized format for death certificates to improve the accuracy and comparability of data (131). The certificate has two parts: Part 1 includes the underlying cause (c), pathophysiologic mechanism (b), and immediate cause of death (a), whereas Part 2 lists contributing conditions not directly related to the underlying cause. Deaths due to COVID-19 follow this order: COVID-19 as the underlying cause (c) leads to pneumonia (b), which ultimately leads to acute respiratory distress syndrome or sepsis (a) as the most common immediate cause of death (131). If chronic diseases of the deceased COVID-19 patient influenced the course of events and contributed to the fatal outcome, they should be reported in Part II of the death certificate.

The role of COVID-19 in the sequence of events leading to death in patients with comorbidities is complex and multifactorial (132, 135, 136). Chronic conditions may increase the severity of COVID-19, and COVID-19 may exacerbate these conditions, making it difficult to determine the primary cause of death (111, 117, 135). According to WHO guidelines, despite the strong association of COVID-19 with coagulopathy and thromboembolism, if a COVID-19 patient dies of myocardial infarction, the sequence leading to death includes ischemic heart disease as the underlying cause (c) and myocardial infarction as the immediate cause (a). In this scenario, COVID-19 is recorded as a contributing factor in Part 2 (131). A deceased patient with COVID-19 and progressive cancer may have cancer listed as the underlying cause of death, with COVID-19 considered a contributing factor

(131).

Statistical agencies count the underlying cause of death when reporting mortality from COVID-19 (137). When COVID-19 is listed as a contributing cause, it is not included in mortality statistics, potentially leading to underreporting and an incomplete understanding of the impact of the virus on mortality (138). Mortality statistics based on a single underlying cause of death may not fully capture the complexity of factors or processes leading to death in individuals with chronic diseases, including the role of the virus in this sequence (136, 139). Therefore, the impact of the COVID-19 pandemic on mortality is complex and determined by multiple factors directly and indirectly related to the infection. Given the multiple effects of different factors on mortality during the pandemic, as well as possible misclassification of causes of death, counting deaths from all causes together could provide a more comprehensive approach to measuring the impact of the pandemic on mortality, avoiding issues of attributing deaths specifically to COVID-19 (123).

1.8 Vaccination against SARS-CoV-2

Vaccination is a key public health strategy to control infectious diseases by inducing a specific immune response to pathogens, thereby preventing or reducing the severity of disease upon subsequent exposure to the corresponding pathogen. Development of vaccines against COVID-19 began immediately after the SARS-CoV-2 genome sequence became available. The WHO developed the Emergency Use Assessment and Listing Procedure to accelerate the availability of vaccines needed in public health emergency situations (140). Different types of COVID-19 vaccines have been developed worldwide, including inactivated, viral vector-based, protein-based, and nucleic acid-based vaccines (141). Each type has a unique structure, advantages and disadvantages regarding immunogenicity, safety, and efficacy.

In Russia, national vaccination against COVID-19 was initiated in December 2020, starting primarily with healthcare workers and education personnel. In January 2021, the vaccination program was extended to the entire population (59). In order to collect data on recipients of COVID-19 vaccines, the Federal Register of Persons Vaccinated against COVID-19 (referred to as the vaccination registry) was established nationwide in accordance with Decree of the Russian Government #373 of March 31, 2009 (90).

Three Russian vaccines were used: Gam-COVID-Vac (Sputnik V), EpiVacCorona and CoviVac. Gam-COVID-Vac (Sputnik V) is a viral vector-based vaccine that uses adenovirus

vectors carrying the gene for the S protein of SARS-CoV-2 to stimulate an immune response (142). The antibodies produced after vaccination primarily target the S protein, whereas the antibody response induced by natural infection can target multiple components of the virus, including the S and N proteins (21). Two doses are required to complete the series, administered intramuscularly with a 21-day interval. To provide more sustained immunity, Gam-COVID-Vac (Sputnik V) uses two different human adenovirus vectors: type 26 for the first dose and type 5 for the second dose (143). The single-dose Gam-COVID-Vac (Sputnik Light), based on adenovirus type 26, was recommended as a booster after the primary vaccination series or after COVID-19 infection (144). Gam-COVID-Vac (Sputnik V and Sputnik Light) was the most widely used vaccine in Russia (144). EpiVacCorona is a protein-based subunit vaccine containing synthetic antigenic components (S protein), while CoviVac contains an inactivated (killed) virus that retains the ability to stimulate an immune response (145, 146). Both vaccines require repeated doses and the use of an adjuvant to induce an adequate immune response.

The rapid pace of vaccine development raised concerns about the efficacy and safety of vaccines, which affected public confidence and led to vaccine hesitancy (147). Initially, the vaccination campaign in Russia progressed slowly, reaching 11.9% nationwide and 13.6% in Arkhangelsk by the end of June 2021 (when the serological survey was conducted as part of the PhD project) (72, 73, 148). By the end of 2021, the proportion of the vaccinated population increased to 46.0% in Russia and 43.6% in Arkhangelsk (72, 73, 148).

The results of short-term clinical trials demonstrated the safety and immunogenicity of the Gam-COVID-Vac (Sputnik V) (142). In contrast, two other vaccines used in Russia, EpiVacCorona and CoviVac, were not similarly effective against symptomatic COVID-19 (144, 149). A recent meta-analysis showed that vaccine effectiveness declined with the emergence of new virus variants of concern, although protection against severe COVID-19 remained high (150). Other studies have reported that vaccination reduces transmissibility of new viral variants, and disease severity, including thrombovascular complications, in both patients with and without pre-existing cardiovascular diseases (151, 152).

1.9 Motivation for the study

During the COVID-19 pandemic, the spread of the virus in the population was largely invisible due to mild or asymptomatic infection, with only a subset of infected cases being detected by the health care system. A year after the pandemic began in Arkhangelsk, when we initiated the seroprevalence survey, many questions remained unanswered: What proportion of the population was immune or seropositive? How many of them were identified by the healthcare system or included in the COVID-19 case registry? What factors were associated with being seropositive, having symptomatic infection, and being captured by the health care system? Who were the hidden spreaders, the asymptomatic cases? And who was at the highest risk for severe infection requiring hospitalization or leading to death?

This PhD project was an attempt to make the invisible visible by shedding light on previously obscured aspects of the COVID-19 pandemic, to understand the spread of the infection and its impact on population health, and to improve infection control strategies for managing future outbreaks.

2 Aim of the PhD thesis

The PhD thesis aimed to investigate the seroprevalence of SARS-CoV-2, the spectrum of COVID-19 cases, and factors associated with COVID-19 morbidity, severity, and mortality in a population-based sample of adults in Arkhangelsk, in the northwest of Russia.

Specific objectives were:

Paper I

✓ to assess the seroprevalence of SARS-CoV-2 in Arkhangelsk, a city in the northwest of Russia, in a year after the start of the pandemic

✓ to estimate the population's adherence to NPIs during the first year of the pandemic

✓ to investigate socioeconomic, behavioral and health-related characteristics associated with infection-acquired antibodies to SARS-CoV-2 and adherence to NPIs

Paper II

✓ to assess and describe the spectrum of COVID-19 cases in the sample of adult population one year after the start of the pandemic

✓ to investigate factors associated with symptomatic infection and hospitalization

Paper III

✓ to estimate all-cause and cause-specific mortality rates during the COVID-19 pandemic compared to the pre-pandemic period

✓ to investigate risk factors of death during the pandemic in a sample of adult population

3 Materials and methods

3.1 Study design and population

The PhD project comprises three parts: a COVID-19 seroprevalence study and factors associated with seropositivity (**Paper I**); an assessment of the clinical spectrum of COVID-19 cases and factors associated with symptomatic status and hospitalization (**Paper II**); an assessment of all-cause mortality and risk factors of death during the COVID-19 pandemic compared to the pre-pandemic period (**Paper III**). The study design and timeline are presented in Figure 2.
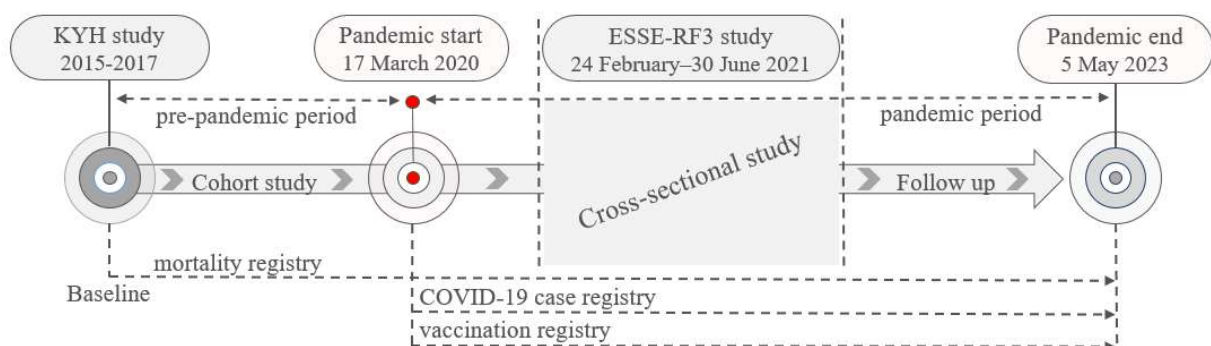
Figure 2. PhD project design and timeline.

All three parts of the PhD project involved participants from the earlier Know Your Heart (KYH) study of cardiovascular diseases. The KYH study sample comprised a population-based cohort of 2380 individuals aged 35 to 69 years at enrollment from 2015 to 2017 (153). Participants were randomly selected from four districts of Arkhangelsk based on anonymized address lists provided by the regional health insurance fund, with a participation rate of 68.2%. All KYH study participants were interviewed by trained interviewers, underwent a physical examination, and provided blood samples for further laboratory testing.

The COVID-19 seroprevalence study (**Papers I & II**) was conducted between February 24, 2021 and June 30, 2021 as a satellite study of the third nationwide survey "Epidemiology of Cardiovascular Diseases and their Risk Factors in Regions of the Russian Federation" (ESSE-RF3 study), recruiting adults aged 35-74 years (154). The PhD project did not include all participants of the ESSE-RF3 study sample in Arkhangelsk, but only those who had participated in the KYH study. Based on the informed consent obtained from the KYH

participants, they were invited in the ESSE-RF3 study, using their previously provided contact information. Most participants in the KYH study were 40 years or older at the time of the ESSE-RF3 study. After excluding individuals who did not consent to be contacted for new invitations (N=56), those who had died before the launch of ESSE-RF3 (N=61), and those who were older than 74 years at the start of ESSE-RF3 because they were beyond the age range of the study, the list of invitees comprised 2258 KYH participants (Fig. 3).
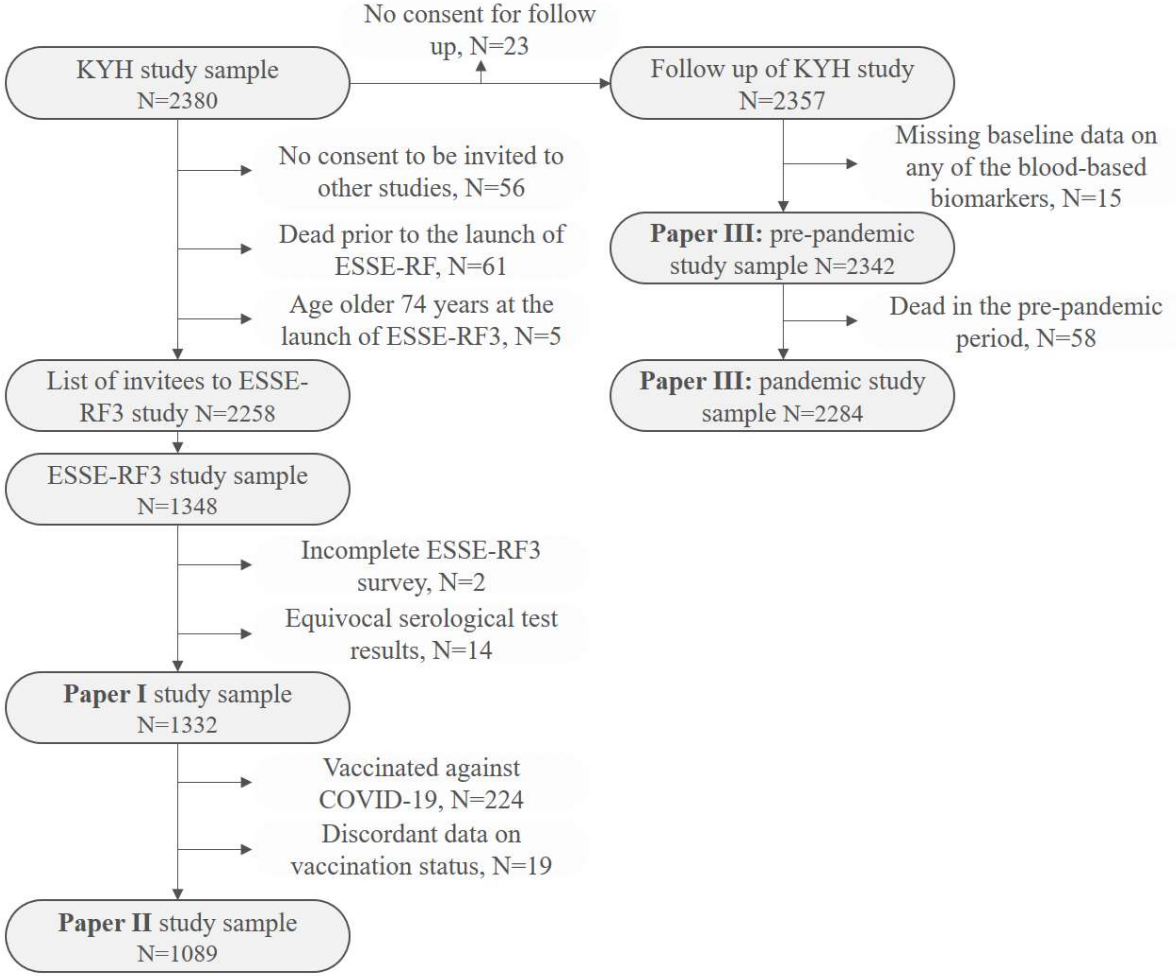
Fig. 3. Flow chart of the PhD project sample.

With a participation rate of 59.7%, 1348 KYH participants aged 40–74 years took part in the ESSE-RF3 study. The questionnaire on COVID-19 experience and symptoms was developed by the ESSE-RF3 team at the National Research Center for Therapy and Preventive Medicine, Moscow (154). The original Russian version and the English translation of the COVID-19 questionnaire are provided as supplementary materials to **Paper I**. The ESSE-RF3 study protocol in Arkhangelsk was expanded to include questions about vaccination against

COVID-19 and serological testing for immunoglobulin G (IgG) antibodies to SARS-CoV-2. The KYH and ESSE-RF3 study data were linked to the COVID-19 case registry and the vaccination registry (**Paper II**).

After excluding incomplete surveys and equivocal serological test results, **Paper I** included 1332 participants. When investigating infection-acquired immunity (**Paper I**), 242 participants who self-reported having been vaccinated against SARS-CoV-2 were excluded from the analysis. For **Paper II**, which focused on the spectrum of COVID-19 cases among unvaccinated individuals, vaccinated individuals as well as those with discrepancies between self-reported vaccination data and vaccination registry data were excluded (N=243). Thus, the **Paper II** study sample comprised 1089 unvaccinated participants with definitive serological test results.

At the time of enrollment in the KYH study, 2357 participants provided written consent to disclose their medical records for research purposes under confidentiality conditions, thereby initiating follow-up. From the time of enrollment in the KYH study, they were followed up for new clinical diagnoses and all-cause mortality using electronic healthcare records from the Regional segment of the Unified Healthcare Information System and the Arkhangelsk Regional Mortality Database (referred to as the mortality registry) at the Arkhangelsk Regional Medical Information Analytical Center (MIAC). In this study, a participant's follow-up period started on the date of the health examination in the KYH study (referred to as the baseline) and ended at the end of the COVID-19 pandemic (i.e., May 5,2023) or at the date of the participant's death. Fifteen individuals with missing baseline data were excluded, resulting in the **Paper III** pre-pandemic study sample of 2342 participants.All study participants alive at the start of the pandemic in Arkhangelsk (March 17, 2020) comprised the **Paper III** pandemic study sample (N=2284).

3.2 COVID-19-related data and case definitions

Our study combined COVID-19-related data from different sources to investigate the spectrum of COVID-19 cases by capturing individuals with different severities and outcomes of infection (119) (Fig. 4). The serological survey was used to detect individuals who were previously infected, regardless of their symptomatic status. To detect specific antibodies, the blood-based biomarkers of SARS-CoV-2 exposure, we used enzyme-linked immunosorbent assay (ELISA) test systems (D-5501 SARS-CoV-2-IgG-EIA-BEST, Russia) (155, 156). This

assay was a semi-quantitative test that allowed us to classify participants as having positive, negative, or equivocal results. The sensitivity of the assay was reported to be 72% within the first 12 days of infection and almost 100% at a later stage (155). An independent test-performance study has shown the assay sensitivity of 89% and the specificity of 100% based on the comparisons of test results in pre-pandemic samples (negative controls) and PCR-positive samples for SARS-CoV-2 (156).
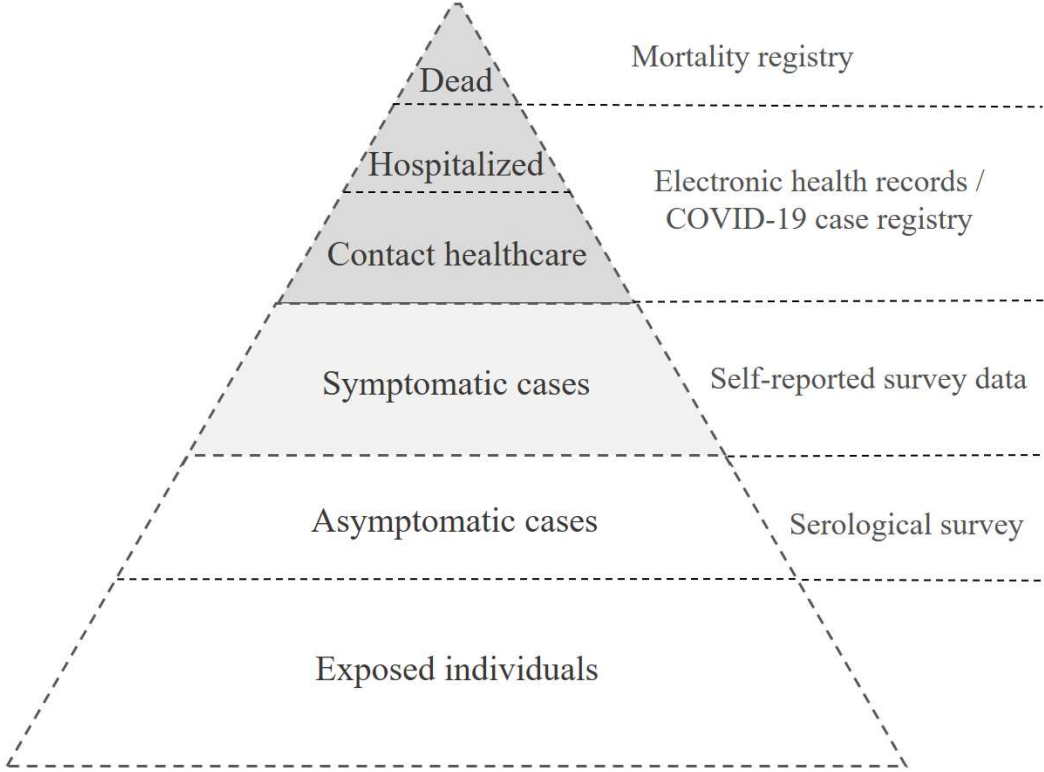
Figure 4. Data sources used to investigate the spectrum of COVID-19 cases.

Self-reported data on COVID-19 experiences and symptoms were collected to detect symptomatic cases that may have escaped the healthcare system, for example, due to testing limitations or individuals not seeking medical care. History of previous COVID-19 was assessed by asking participants, "Have you had COVID-19 in the past 12 months?" (yes/no/don't know). We treated "don't know" as a negative response. Those who self-reported having had COVID-19 were asked to report the date of onset and to answer the questions "Did you seek medical care?" and "Were you hospitalized?" with a yes/no response option. They were also asked to report the symptoms they experienced.
The COVID-19 case registry provided data on patients identified by the healthcare system,

including those who were hospitalized or died. The following information was collected from the COVID-19 case registry: all COVID-19-related visits (outpatient and inpatient) prior to participation in ESSE-RF3, including final diagnoses, date of disease onset, and outcome. Study participants were considered registered COVID-19 cases if they had records of COVID-19 diagnoses with ICD-10 codes U07.1 "COVID-19, virus identified" or U07.2 "COVID-19, virus not identified". Detailed information on cause of death was collected from the mortality registry, including date of death, immediate cause of death, associated pathological conditions, underlying cause of death, external cause of death, and other contributing conditions according to ICD-10.

To assess vaccination status, we used self-reported vaccination data and information from the vaccination registry. Self-reports of vaccination against COVID-19 were obtained by asking, "Have you received any vaccine against COVID-19?" Participants who responded positively were asked to provide information on the number of doses and dates of vaccination. The same information regarding the dates and the number of vaccine doses received by each participant was obtained from the vaccination registry.

Paper I combined the results of the serological survey and self-reported data on the presence of COVID-19 symptoms. Participants with positive test results for specific IgG antibodies were defined as seropositive, while those with negative results were defined as seronegative. Unvaccinated individuals who reported a positive PCR test for COVID-19 but had no symptoms of infection at the time of the positive test, and those who tested positive for SARS-CoV-2 antibodies but reported no history of COVID-19 infection, were classified as asymptomatic. In **Paper I**, we also estimated adherence to NPIs based on respondents' self-reports regarding five COVID-19 NPIs: self-isolation (stay-at-home order), social distancing, wearing face masks in public places or transport, wearing gloves, and using hand sanitizer (63). A "yes" response to adherence to each NPI during the pandemic was scored as 1, while a "no" response was scored as 0. The total score ranged from 0 to 5, which was then dichotomized by combining scores of 0-3 (low adherence) and 4-5 (high adherence). The threshold was chosen based on the median number of NPIs adhered to, which was 4. Since wearing gloves was not an international recommendation, we performed the analysis of factors associated with adherence to NPIs with and without considering gloves to ensure comparability of results. Excluding gloves, adherence scoring was as follows: 1 for low adherence (0 to 3 NPIs) and 0 for high adherence (4 NPIs).

The spectrum of COVID-19 cases (**Paper II**) was determined based on the agreement

between serological survey data, the COVID-19 case registry, and self-reported data on COVID-19 experience and symptoms. Unvaccinated study participants were categorized as previously infected or not previously infected. Previously infected participants were further classified as symptomatic or asymptomatic. Symptomatic cases were defined as individuals with positive serological tests or records in the COVID-19 case registry who also reported experiencing COVID-19 symptoms at the survey. Asymptomatic cases were those with positive serological tests or COVID-19 records in the registry, but who reported no COVID-19 experience or symptoms. Participants who had neither a positive serological test nor a record in the COVID-19 case registry were considered previously non-infected, regardless of the symptoms they reported. Infected cases were then divided into non-hospitalized and hospitalized cases. Hospitalized cases were defined as individuals recorded as inpatients in the COVID-19 case registry. Participants who reported a history of COVID-19 hospitalization with date and duration but had no record in the COVID-19 case registry were also considered hospitalized symptomatic cases.

In **Paper III**, we used mortality registry data to analyze all-cause and cause-specific mortality within the study sample during the pandemic. Due to the limited number of registered deaths with COVID-19 as the underlying cause and potential misclassification, we investigated demographic, behavioral, and health-related factors, including cardiometabolic biomarkers, associated with all-cause deaths during the pandemic without distinguishing between those related and unrelated to COVID-19. Since nearly all participants had evidence of exposure to the virus, confirmed by the presence of infection-acquired antibodies, at the end of the COVID-19 pandemic (May 5, 2023) (157, 158), we did not consider a history of COVID-19 as a factor associated with the risk of death.

3.3 Data sources and participant characteristics considered in the study

To achieve the objectives of the PhD project, we combined information from several data sources. Detailed information on participants' demographic, behavioral, and health-related characteristics, including blood-based cardiometabolic biomarkers, was collected in the KYH study (**Paper III**) and the ESSE-RF3 study (**Papers I & II**). Table 1 provides an overview of the data sources and details of the data collected from these sources that were considered in

the analyses.

Cardiometabolic biomarkers were measured in non-fasting blood samples at baseline (2015-2017) as part of the KYH study protocol (153). The following biomarker levels were considered abnormal: total cholesterol ≥5.2 mmol/L, low-density lipoprotein cholesterol (LDL-C) >3.0 mmol/L, high-density lipoprotein cholesterol (HDL-C) < 1.0 mmol/L for men and <1.3 mmol/L for women, triglycerides >1.7 mmol/L, glycated hemoglobin (HbA1C) ≥ 6.5%, gamma-glutamyl transferase (GGT) ≥40 U/L, high-sensitivity C-reactive protein (Hs-CRP) ≥2 mg/L, cystatin C ≥1.2 mg/L, N-terminal pro-b-type natriuretic peptide (NT-proBNP) ≥125 pg/mL, high-sensitivity troponin T (Hs-TnT) ≥6 ng/L.

3.4 Statistical analysis

Participant characteristics were presented as absolute numbers and percentages for categorical variables and as medians with first and third quartiles for continuous variables. Accordingly, Pearson's chi-squared test and the Mann-Whitney U test were used to compare groups on categorical and continuous characteristics. Seroprevalence (**Paper I**) was estimated as the number of SARS-CoV-2 IgG positive participants divided by the number of tested participants and reported as a percentage. Confidence intervals (CIs) for proportions were calculated using the Wilson method. Seroprevalence adjusted for test performance (89% sensitivity, 100% specificity) was estimated using the equation: (crude prevalence + test specificity-1)/(test sensitivity + test specificity - 1) (159). The 95% CIs for the adjusted estimates of seroprevalence were calculated by bootstrapping using the R package bootComb (version 4.1.1)(160).

Binary logistic regression was used to investigate factors associated with seropositive status and low adherence to NPIs (**Paper I**) as well as symptomatic infection and hospitalization (**Paper II**). Adherence to NPIs was re-assessed after exclusion of gloves to ensure comparability of the results.

The PhD thesis presents unpublished results on the investigation of factors associated with SARS-CoV-2 vaccination (1 - received at least one dose, 0 - unvaccinated) early in the pandemic before the start of the seroprevalence survey. Crude and adjusted odds ratios (ORs) with 95% CIs were calculated. A summary of the statistical analyses used in **Papers I-III** is presented in Table 2.

All-cause and cause-specific (defined by the ICD-10 chapters) mortality rates per 1000

Table 1. Data sources and variables considered in the study

Data sources	Paper I	Paper II	Paper III
KYH study	-	-	Demographic (age, education), behavioral (smoking, hazardous alcohol consumption: score of ≥8 on the Alcohol Use Disorders Identification Test – AUDIT(161)), obesity: body mass index ≥30 kg/m^2, health-related characteristics (hypertension, diabetes, angina, myocardial infarction, heart failure, asthma, chronic bronchitis, kidney disease, liver disease, neoplasms), blood-based cardiometabolic biomarkers
ESSE-RF3 study	Demographic (sex, age, education, marital status, number of persons in household, living with children (<18 years), occupation, income, behavioral (smoking, frequency of heavy drinking: ≥60 grams of pure alcohol on a single occasion (162), adherence to NPIs, confidence in efficiency of the NPIs), health-related characteristics (hypertension, diabetes, chronic pulmonary diseases, coronary heart diseases) Self-reported COVID-19 experience/ symptoms, and vaccination against COVID-19 with number of doses and dates	Demographic (sex, age, education), behavioral (smoking, frequency of heavy drinking: ≥60 grams of pure alcohol on a single occasion (162), health-related characteristics (hypertension, diabetes, abdominal, obesity, self-rated health on a 0-100 visual analog scale: ≤ median (poor health), > median (good health)) Self-reported COVID-19 experience/ symptoms and vaccination	-
Serological survey data	Positive/negative results (equivocal results excluded)	Positive/negative results (equivocal results excluded)	-
COVID-19 case registry	-	Diagnosis of COVID-19, hospitalization	-
Vaccination registry	-	Vaccination against COVID-19	Vaccination against COVID-19
Mortality registry	-	-	All-cause and cause-specific (as defined by the ICD-10 chapters) deaths

Table 2. Summary of statistical analysis

	Paper I	**Paper II**	**Paper III**
Study design	Cross-sectional study	Cross-sectional study	Cohort study
Research questions	Seroprevalence adjusted for test performance Factors associated with seropositive status and adherence to NPIs	Spectrum of COVID-19 cases Factors associated with symptomatic status and hospitalization	Mortality rates Risk factors for pre-pandemic and pandemic deaths
Statistical analysis	Binary logistic regression	Binary logistic regression	Cox proportional hazards regression
Outcomes and covariates	**Dependent variable (1)** Seropositive status: 1 – seropositive, 0 – seronegative **Independent variables (1)** were introduced stepwise *Block 1*: socio-demographic characteristics *Block 2*: behavioral (smoking, heavy drinking, adherence to NPIs), health-related characteristics (hypertension, diabetes, chronic pulmonary diseases, coronary heart diseases) **Dependent variable (2)** Adherence to COVID-19 NPIs: 1 – low (0-3 NPIs), 0 – high (4-5 NPIs) **Independent variables (2)** were introduced into the regression model using the enter option: sex, age, education, occupation, income, smoking and drinking habits, chronic health conditions, confidence in the efficiency of NPIs, vaccination against COVID-19	**Dependent variable (1)** Symptomatic COVID-19 case: 1 – symptomatic case, 0 – asymptomatic case **Dependent variable (2)** Hospitalization with COVID-19: 1 – hospitalized case, 0 – non-hospitalized case **Independent variables:** sex, age, higher education, hypertension, diabetes, abdominal obesity, self-rated health, smoking status, and frequency of heavy drinking	**Dependent variable** All-cause deaths **Time variable** Person-months of observation in the pre-pandemic period (date of KYH enrollment – March 16, 2020) and the pandemic period (March 17, 2020 - May 5, 2023) **Independent variables:** age, higher education, smoking, hazardous drinking, chronic health conditions (obesity, hypertension, diabetes, angina, history of myocardial infarction, chronic heart failure, asthma, chronic bronchitis, kidney disease, liver disease, neoplasms) and blood-based cardiometabolic biomarkers (total cholesterol, LDL-C, HDL-C, triglycerides, HbA1C, GGT, Hs-CRP, cystatin C, NT-proBNP, Hs-TnT)
Adjustment	Mutual adjustment	Demographic (age, sex, higher education) and behavioral (smoking, frequency of heavy drinking) factors	Demographic (age, higher education) and behavioral (smoking, hazardous drinking) factors

LDL-C – low-density lipoprotein cholesterol, HDL-C – high-density lipoprotein cholesterol, HbA1C – glycated hemoglobin, GGT – gamma-glutamyl transferase, Hs-CRP – high-sensitivity C-reactive protein, NT-proBNP – N-terminal pro-brain natriuretic peptide, Hs-TnT– high-sensitivity troponin T

person-years were calculated for both the pre-pandemic and pandemic periods (**Paper III**). Mortality rates for the pandemic period were age-standardized to the age distribution of the study population at baseline (direct standardization, 5-year bands). Mortality rates were presented with 95% CIs. Age-adjusted mortality ratios in the pandemic period were estimated as hazard ratios (HR) derived from Cox proportional hazards regression models of the studied death outcomes, with period (1 = pandemic period, 0 = pre-pandemic period) and age in years entered as covariates. We also used Cox proportional hazards regression models to investigate risk factors for pre-pandemic and pandemic deaths. For each covariate, the interaction with study period was assessed by comparing regression models with and without the interaction term using the likelihood ratio test. Based on the identified interaction between study period and sex, the analyses were stratified by sex. Effect estimates were reported as HRs adjusted for demographic and behavioral factors with corresponding 95% CIs.

In the PhD thesis, we also presented unpublished findings on the risk factors associated with the three major underlying causes of death during the pandemic in men and women combined (pooled analysis). For this analysis, we used Cox proportional hazards regression models and reported results as HRs adjusted for age and sex.

The Statistical Package for the Social Science SPSS version 24.0 (SPSS Inc, Chicago, Il) and Stata version 17.0 (StataCorp, College Station, TX, USA) were used for data analysis. A p-value < 0.05 was considered statistically significant.

3.5 Ethical considerations

The PhD project was conducted in compliance with the ethical standards of the 1964 Declaration of Helsinki and its later amendments. Ethical approval for the original KYH study was provided by the ethics committees of London School of Hygiene & Tropical Medicine (approval number 8808, received February 24, 2015) and NSMU, Arkhangelsk (approval number 01/01-15, received January 27, 2015). Ethical approval for follow-up of KYH participants using electronic health records was received from the ethics committee of NSMU, Arkhangelsk, Russia (approval number 01/04-19, received April 24, 2019). Ethical approval for ESSE-RF-3 was obtained from the ethics committee of the National Research Center for Therapy and Preventive Medicine Moscow, Russia (approval number 01-01/20, received February 04, 2020) and the ethics committee of NSMU, Arkhangelsk, Russia (approval number 01/02-21, received February 17, 2021).

Ethical approval for the sub-study of COVID-19-related issues was received from the ethics committee of NSMU, Arkhangelsk, Russia (approval number 01/02-21, received February 17, 2021). The PhD project was approved by the Regional Committee for Medical and Health Research Ethics (REK) in Norway (approval number 339397, received December 7, 2021). Participants included in the KYH study and the ESSE-RF3 study provided written informed consent to participate. At the time of enrollment in the KYH study, participants in the current study provided written consent to disclose their medical and other health-related records for research purposes under the condition of confidentiality and to be invited to participate in other studies. Data linkage between the KYH data, the COVID-19 case registry, the vaccination registry, and the regional mortality registry was performed by the MIAC in accordance with the NSMU-MIAC Confidentiality Agreement, based on the informed consent obtained from the participants, as well as legal and ethical approvals. Participants were anonymized using randomly assigned unique ID numbers. Subsequent data linkage was based on these depersonalized IDs, ensuring that no personal identifiers were present in the analyzed dataset.

4 Results

4.1 COVID-19 seroprevalence and adherence to preventive measures (Paper I + unpublished findings)

At 12-15 months after the onset of the pandemic in Arkhangelsk, 65.1% (95% CI: 62.5, 67.6) of the surveyed population aged 40-74 years (867/1332) were seropositive (**Paper I**). The seroprevalence rate increased to 73.0% (95% CI: 67.1, 85.7) after adjustment for test performance. Less than half (652/1332, 48.9%) of the study participants adhered to all five recommended NPIs, including self-isolation, social distancing, wearing facemasks in public places or transport, wearing gloves, and use of hand sanitizers. Altogether, 242 (18.2%) of 1332 study participants self-reported having been vaccinated against SARS-CoV-2, with 195 (14.6%) having received two doses and being fully immunized. Only 29 (61.7%) of the 47 individuals who had received one dose of the vaccine were seropositive, whereas 194 (99.5%) of the 195 participants who had received two doses were seropositive.

In the crude analysis, older age, male sex, former smoking status, and the presence of chronic health conditions were positively associated with receiving SARS-CoV-2 vaccination early in the pandemic before the start of the seroprevalence survey. Conversely, having a large family (≥4 persons in the household), living with children (<18 years), regular employment, and low adherence to NPIs were negatively associated with vaccination (Table 3). After adjustment for age and sex, both former and current smokers were more likely to be vaccinated, whereas participants with low adherence to NPIs were less likely to be vaccinated.

Of the 339 participants who self-reported having had COVID-19, 322 (95.0%) were seropositive. After excluding 16 individuals who received the vaccine after contracting the disease, 309 (95.7%) of the 323 participants were seropositive. Detailed information on the grouping of participants is shown in Figure 5.

Of the 1 090 unvaccinated study participants, 644 were seropositive, indicating infection-acquired immunity. The SARS-CoV-2 seroprevalence rate due to the infection was 59.1% (95% CI: 56.1; 62.0) and increased to 66.3% (95% CI: 58.1; 76.0) when adjusted for test performance.

Of the unvaccinated seropositive individuals, 309 self-reported having had COVID-19 (7 of 309 were asymptomatic cases with a positive PCR test), while 335 did not report having had COVID-19. A total of 342 (31.4%) asymptomatic infections were detected among the

unvaccinated study participants.

Table 3. Characteristics associated with vaccination against SARS-CoV-2 early in the pandemic, N=1332 (unpublished findings)

Characteristics	OR$_{crude}$ (95%CI)	OR$_{adj}$ (95%CI)[1]
Sex		
Female	reference	reference
Male	1.38 (1.04; 1.82)	1.53 (1.14; 2.06)*
Age		
40-54 years	reference	reference
55-64 years	1.88 (1.27; 2.78)	1.85 (1.25; 2.73)
65-74 years	6.14 (4.27; 8.82)	6.32 (4.39; 9.10)
Marital status		
Single	reference	reference
Married	1.13 (0.84; 1.52)	0.88 (0.73; 1.06)
Number of persons in household		
1	reference	reference
2-3	0.89 (0.62; 1.26)	1.10 (0.80; 1.70)
≥4	0.41 (0.24; 0.71)	1.00 (0.60; 1.80)
Living with children (<18 years)		
No	reference	reference
Yes	0.32 (0.21; 0.50)	0.62 (0.39; 1.00)
Education		
Secondary and lower	reference	reference
Specialized secondary	0.92 (0.56; 1.53)	1.14 (0.67; 1.94
Higher	0.78 (0.47; 1.32)	1.25 (0.72; 2.15)
Occupation		
Retired or unemployed	reference	reference
Regular employment	0.41 (0.31; 0.54)	0.88 (0.62; 1.24)
Income[2]		
Low	reference	reference
Middle	0.82 (0.61; 1.12)	1.04 (0.75; 1.44)
High	0.51 (0.30; 0.85)	0.99 (0.56; 1.74)
Adherence to NPIs		
high (4-5 NPIs)	reference	reference
low (0-3 NPIs)	0.58 (0.39; 0.88)	0.62 (0.40; 0.95)
Smoking		
Never smoker	reference	reference
Former smoker	1.41 (1.03; 1.94)	1.46 (1.01; 2.11)
Current smoker	1.18 (0.80; 1.75)	1.64 (1.06; 2.55)
Frequency of heavy drinking[3]		
Never	reference	reference
Once a week and less often	0.81 (0.60; 1.09)	0.93 (0.65; 1.33)
2 times a week and more often	0.61 (0.27; 1.38)	0.60 (0.25; 1.44)
Chronic health conditions[4]		
No	reference	reference
Yes	2.12 (1.49; 3.01)	1.22 (0.84; 1.79)

NPIs – non-pharmaceutical interventions
[1]adjusted for age and sex; [2]according to the income classification of the Federal State Statistics Service of Russia; [3]≥60 grams of pure alcohol on a single occasion (162); [4]having one or more of the following conditions: hypertension, diabetes, chronic pulmonary diseases, coronary heart diseases
*The association was attenuated after adjustment for smoking status.
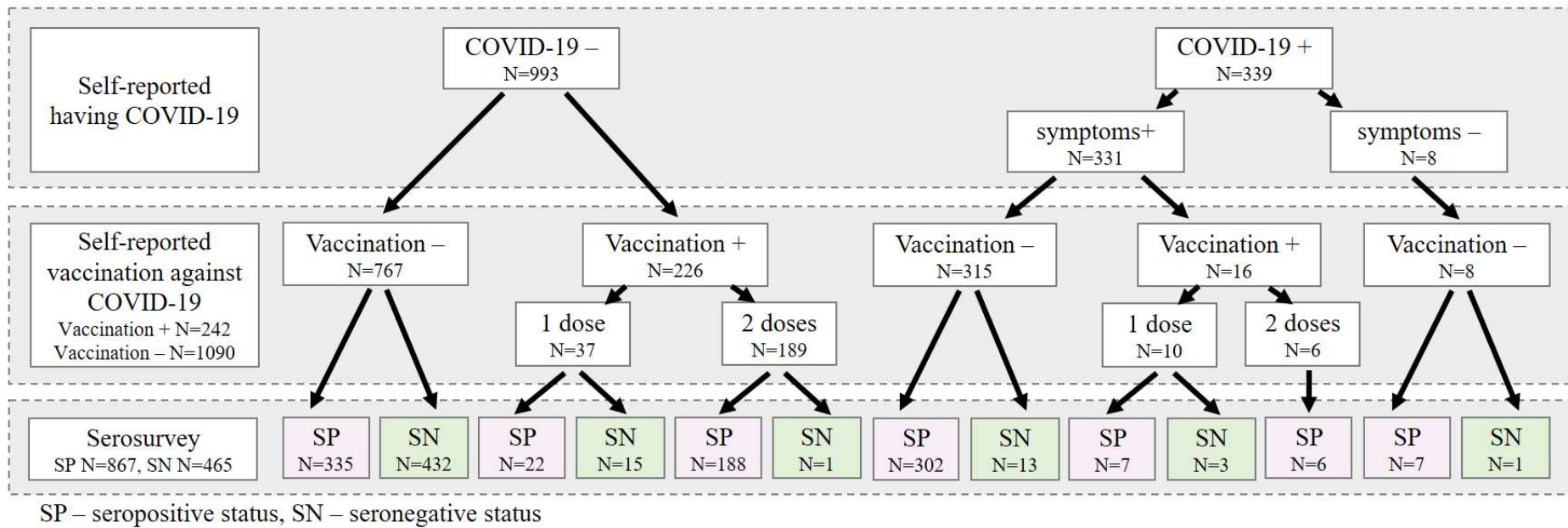
Figure 5. Classification of study participants based on self-reported COVID-19/vaccination status and serological survey results, N=1332.

Individuals with infection-acquired immunity were more likely to have regular employment (OR 2.06, 95% CI 1.50; 2.84) and less likely to be smokers (OR 0.37, 95% CI 0.25; 0.54). Adherence to NPIs was not found to be associated with contracting the virus. Male sex (OR 2.16, 95% CI 1.52; 3.09), low income (OR 2.22, 95% CI 1.20; 4.08), low confidence in the efficiency of NPIs (OR 2.73, 95% CI 2.02; 3.70), and heavy drinking twice a week or more often (OR 2.54, 95% CI 1.32; 4.88) were associated with low adherence to NPI, defined as following three or fewer recommended NPIs. After exclusion of glove use from the analysis (as wearing gloves was not recommended internationally), male sex (OR 1.72, 95% CI 1.27; 2.34), regular employment (OR 1.84, 95% CI 1.33; 2.54), and low confidence in the efficiency of NPIs (OR 1.98, 95% CI 1.51; 2.59) were associated with low adherence to internationally recommended NPIs.

4.2   Spectrum of COVID-19 cases (Paper II)

After excluding vaccinated participants and those with discrepancies between self-reported vaccination data and vaccination registry data (N=243), we classified the unvaccinated study participants (N=1089) as non-infected or previously infected (asymptomatic, non-hospitalized, and hospitalized symptomatic) cases based on agreement between serological survey data, the COVID-19 case registry, and self-reported COVID-19 status. Figure 6 provides detailed information about the classification of participants.

One and a half years into the pandemic in Arkhangelsk, 59.7% (95% CI: 56.7; 62.6) of the unvaccinated adult population aged 40-74 years were infected with SARS-CoV-2 (650/1089). Asymptomatic COVID-19 cases (344/1089) accounted for 31.6% (95% CI: 28.9; 34.5) of the study population.

More than half of the infected individuals (430/650, 66.2%) were not included in the healthcare-based registry, mainly with asymptomatic infection (331/430, 77.0%). Thus, 96.2% of all asymptomatic cases (331/344) and 32.3% of all symptomatic cases (99/306) were not recorded in the COVID-19 case registry.

The spectrum of infected cases was as follows: asymptomatic cases were 52.9% (344/650), symptomatic non-hospitalized cases were 38.5% (250/650), and symptomatic hospitalized cases were 8.6% (56/650). Therefore, less than half, 47.1% (306/650) of the infected individuals, were symptomatic.
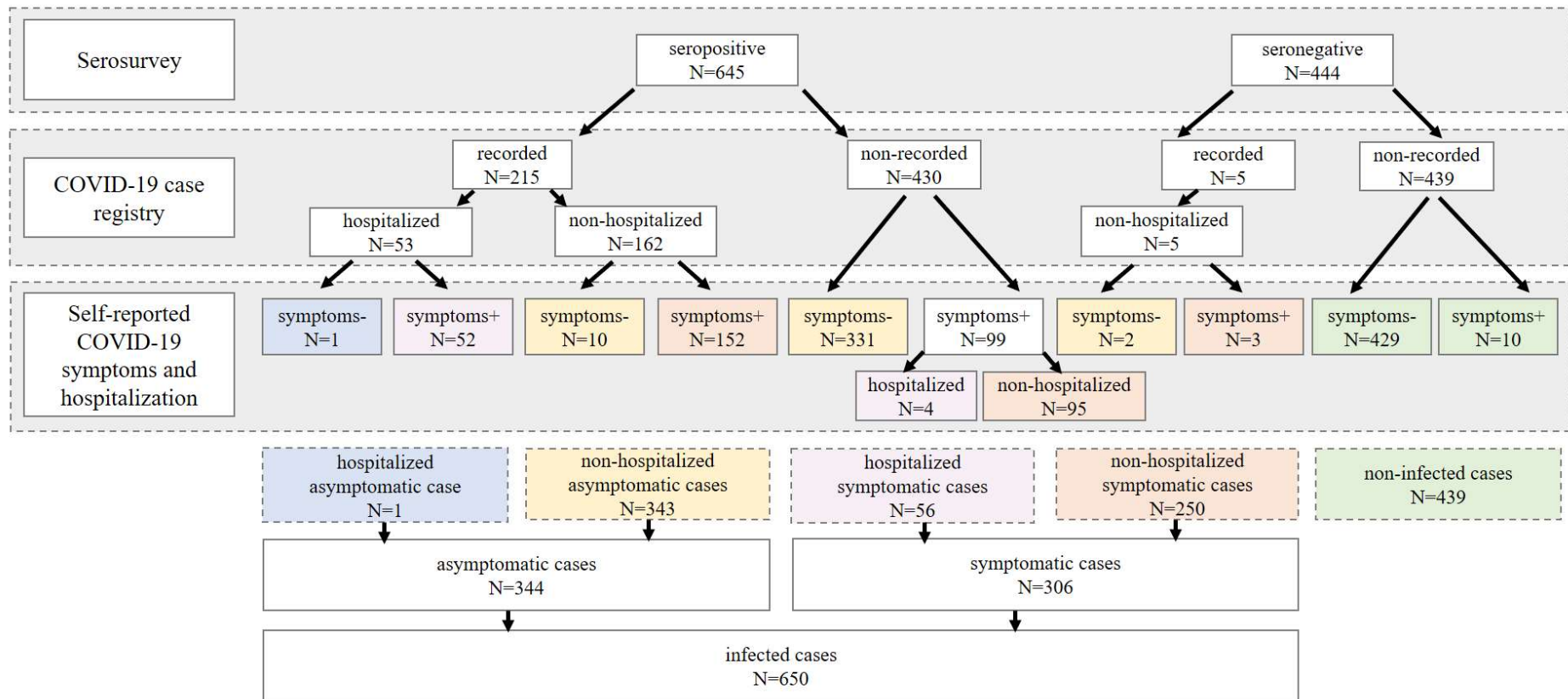
Figure 6. Classification of study participants based on linkage of serological survey, COVID-19 case registry data and self-reported survey data, N=1089.

Older age was positively associated (OR 1.93, 95% CI 1.10; 3.36 for 55-64 years, OR 2.96, 95% CI 1.58; 5.53 for 55-74 years when compared to 40-44 years), while smoking showed a negative association (OR 0.55, 95% CI 0.31; 0.97) with symptomatic COVID-19. Among symptomatic cases, 18.3% (56/306) were hospitalized. Individuals older than 65 years (OR 6.99, 95% CI 1.45; 33.72) and those with poor self-rated health (OR 2.51, 95% CI 1.23; 5.14) were more likely to be hospitalized.

4.3   Mortality rates and risk factors for death during the COVID-19 pandemic (Paper III + unpublished findings)

In the KYH study cohorts, there were 58 deaths in the pre-pandemic period and 92 deaths in the pandemic period, including 17 COVID-19-related deaths and 75 non-COVID-19 deaths. All-cause mortality increased from 7.24 (95% CI 5.59; 9.36) per 1000 person-years in the pre-pandemic period to 10.16 (95% CI 7.76; 12.57) per 1000 person-years in the pandemic period (standardized to the age distribution at baseline). The risk of death was 41.0% higher during the pandemic period than during the pre-pandemic period (HR 1.41 95% CI 1.00; 1.99). Cardiovascular diseases and neoplasms were the leading causes of death in both periods (Table 4).

Table 4. Underlying causes of death in study participants during the pre-pandemic and the pandemic periods (unpublished findings)

Underlying cause of death	Pre-pandemic period: date of KYH enrollment - March 17, 2020, N (%)	Pandemic period: March 17, 2020 - May 5, 2023, N (%)
Cardiovascular diseases I00-I99	28 (48.3)	31 (33.7)
Neoplasms C00-D48	13 (22.4)	28 (30.4)
COVID-19 U07	-	11 (12.0)
Respiratory diseases J00-J99	0 (0.0)	5 (5.4)
External causes V01-Y98	6 (10.3)	5 (5.4)
Digestive system diseases K00-K95	2 (3.5)	4 (4.3)
Diabetes E08-E13	4 (6.9)	3 (3.3)
Unattended death R 98	2 (3.5)	2 (2.2)
Kidney disease N03.3	1 (1.7)	1 (1.1)
Aneurism Q28.2	1 (1.7)	1 (1.1)
Viral hepatitis B18.9	1 (1.7)	1 (1.1)
Total	58 (100,0)	92 (100,0)

In 11 of 17 COVID-19-related deaths, COVID-19 was identified as the underlying cause. In the remaining six cases, COVID-19 was identified as a contributing factor. Of these six deceased patients, five had cancer as the underlying cause of death. The remaining case was

due to myocardial infarction, where the sequence of events leading to death included ischemic heart disease as the underlying cause and myocardial infarction as the immediate cause. During the pandemic, five deaths were attributed to pneumonia as the underlying cause. In addition, among those who died from cardiovascular disease, neoplasm, or diabetes (classified as underlying cause of death), there were nine cases in which pneumonia or acute respiratory distress syndrome were listed as the immediate cause. This raises questions about possible misclassification or undiagnosed COVID-19 cases. Table 5 presents risk factors for death from three major underlying causes, including cardiovascular diseases, neoplasms, and COVID-19.

Table 5. Risk factors for death from three major underlying causes (cardiovascular diseases, neoplasms, COVID-19) during the pandemic period, N=2284 (unpublished findings)

Characteristics	Cardiovascular diseases (I00-I99), N=30[1]	Neoplasms (C00-D48), N=23[1]	COVID-19 (U07.1-U07.2), N=11
	HR_{adj} (95%CI)[2]	HR_{adj} (95%CI)	HR_{adj} (95%CI)
Age, years	1.06 (1.01; 1.10)	1.12 (1.06; 1.18)	1.15 (1.05; 1.26)
Male sex	3.50 (1.60; 7.63)	1.68 (0.74; 3.82)	1.30 (0.40; 4.28)
Higher education	0.39 (0.15; 1.01)	0.74 (0.29; 1.88)	0.81 (0.21; 3.06)
Smoking	2.09 (0.98; 4.48)	4.56 (1.89; 10.99)	1.93 (0.47; 7.90)
Hazardous drinking[3]	0.69 (0.27; 1.77)	0.92 (0.29; 2.92)	0.49 (0.06; 4.19)
Obesity[4]	1.45 (0.68; 3.09)	1.09 (1.45; 2.61)	5.37 (1.39; 20.75)
Hypertension[5]	2.47 (1.06; 5.78)	1.10 (0.45; 2.71)	0.89 (0.24; 3.16)
Myocardial infarction[5]	1.87 (0.63; 5.55)	2.12 (0.70; 6.43)	0.91 (0.11; 7.28)
Angina[5]	2.41 (1.06; 5.47)	1.42 (0.56; 3.58)	1.66 (0.46; 5.93)
Heart failure[5]	2.58 (1.11; 6.00)	1.04 (0.35; 3.12)	0.99 (0.21; 4.67)
Diabetes[5]	3.42 (1.42; 8.22)	1.14 (0.33; 3.91)	4.06 (1.14; 14.48)
Asthma[5]	1.22 (0.29; 5.19)	1.32 (0.30; 5.72)	4.99 (1.28; 19.53)
Chronic bronchitis[5]	0.26 (0.35; 1.93)	1.78 (0.66; 4.84)	0.60 (0.76; 4.70)
Kidney diseases[5]	3.41 (1.62; 7.16)	0.87 (0.30; 2.58)	0.87 (0.19; 4.07)
Liver disease[5]	1.61 (0.69; 3.77)	2.05 (0.83; 5.05)	0.43 (0.05; 3.41)
Neoplasms[5]	0.59 (0.08; 4.46)	3.38 (1.21; 9.44)	-
Total cholesterol, ≥5.2 mmol/L	1.10 (0.53; 2.27)	0.70 (0.27; 1.87)	0.59 (0.17; 1.97)
LDL-C, >3.0 mmol/L	1.06 (0.46; 2.48)	1.57 (0.45; 5.51)	1.41 (0.30; 6.56)
HDL-C, <1.0 mmol/L for men and <1.3 mmol/L for women	1.78 (0.78; 4.05)	2.34 (0.84; 6.48)	1.41 (0.36; 5.45)
Triglycerides, >1.7 mmol/L	1.57 (0.75; 3.25)	1.57 (0.60; 4.13)	0.84 (0.22; 3.20)
HbA1C, ≥6.5%	7.05 (3.03; 16.38)	3.85 (1.23; 12.02)	4.68 (1.21; 18.11)
GGT, ≥40 U/L	1.77 (0.84; 3.73)	2.24 (0.84; 5.98)	1.99 (0.57; 6.93)
Hs-CRP, ≥2 mg/L	2.66 (1.24; 5.72)	2.66 (0.93; 7.58)	1.95 (0.57; 6.70)
Cystatin C, ≥1.2 mg/L	1.91 (0.56; 6.53)	1.90 (0.55; 6.60)	4.44 (1.11; 17.72)
NT-proBNP, ≥125 pg/mL	6.74 (2.76; 16.46)	1.44 (0.52; 4.00)	1.17 (0.33; 4.10)
Hs-TnT, ≥6 ng/L	4.81 (1.24; 6.05)	2.46 (0.53; 11.45)	1.70 (0.34; 8.57)

LDL-C – low-density lipoprotein cholesterol, HDL-C – high-density lipoprotein cholesterol, HbA1C – glycated hemoglobin, GGT – gamma-glutamyl transferase, Hs-CRP – high-sensitivity C-reactive protein, NT-proBNP – N-terminal pro-brain natriuretic peptide, Hs-TnT– high-sensitivity troponin T
[1]deaths with COVID-19 as a contributing cause were excluded; [2]adjusted for age and sex; [3]score of ≥8 on the AUDIT(161);[4]body mass index ≥30 kg/m^2; [5]self-reported doctor-diagnosed diseases

Male sex, hypertension, angina, heart failure, diabetes, kidney diseases, elevated levels of HbA1C, Hs-CRP, NT-proBNP, and Hs-TnT were associated with a higher risk of death from cardiovascular causes during the pandemic period. Smoking and elevated levels of HbA1C were risk factors for death from neoplasms. Obesity, diabetes, asthma, elevated levels of HbA1C and cystatin C were associated with an increased risk of death from COVID-19. We repeated the analysis with COVID-19-related deaths (N=17), defined as any death with COVID-19 as an underlying (N=11) or contributing cause (N=6), as the outcome. Most of the factors associated with death from COVID-19 (N=11), except asthma, were also significantly associated with the risk of COVID-19-related death (N=17). This indicates the important role of the virus in the pathogenic mechanism that led to the deaths of the six patients with neoplasms and cardiovascular disease as underlying causes for which COVID-19 was recorded as a contributing cause.

In **Paper III**, to address the potential misclassification of deaths related and unrelated to COVID-19, we presented age-standardized all-cause mortality rates and risk factors for death during the pandemic period compared to the pre-pandemic period.

During the pandemic, the all-cause mortality rate in the KYH cohort remained higher in men than in women, while the mortality rate increased 2.32-fold in women (from 2.79 to 6.45 per 1000 person-years) but not in men (from 13.43 to 15.58 per 1000 person-years) compared with the pre-pandemic period. Women had a higher mortality rate from neoplasms (1.78 per 1000 person-years) than from cardiovascular diseases (1.39 per 1000 person-years) in the pandemic period, whereas the opposite pattern was observed in the pre-pandemic period. In men, cardiovascular disease remained the leading cause of death during the pandemic (6.34 per 1000 person-years), followed by neoplasms (4.64 per 1000 person-years). Mortality rates with COVID-19 as the underlying cause of death were similar in men (1.21 per 1000 person-years) and women (0.87 per 1000 person-years).

Older age (HR 1.09, 95% CI 1.05; 1.12 for men, HR 1.11, 95% CI 1.06; 1.16 for women), smoking (HR 2.80, 95% CI 1.64; 4.78 for men, HR 3.11, 95% CI 1.40; 6.94 for women), and diabetes (HR 3.41, 95% CI 1.63; 7.12 for men, HR 2.32, 95% CI 1.09; 4.92 for women) were associated with a higher risk of all-cause death in both sexes in both periods. In women, obesity (HR 2.19, 95% CI 1.11; 4.30), angina (HR 3.59, 95% CI 1.76; 7.30), and elevated cystatin C levels (HR 3.47, 95% CI 1.48; 8.12) were associated with a higher risk of death during the pandemic period, while higher education was associated with a lower risk (HR 0.30, 95% CI 0.11; 0.86). In men, asthma (HR 2.62, 95% CI 1.05; 6.61) and elevated levels of Hs-TnT (HR 2.79, 95% CI 1.17; 6.70) increased the risk of death during the COVID-19

pandemic, while elevated levels of hs-CRP (HR 2.42, 95% CI 1.27; 4.61 for the pre-pandemic period, HR 1.80, 95% CI 1.04; 3.12 for the pandemic) and NT-proBNP (HR 1.96 95% CI 1.01; 3.81 for the pre-pandemic period, HR 3.13, 95% CI 1.71; 5.75 for the pandemic) were associated with a higher risk of death in both periods. Vaccination against COVID-19 reduced all-cause mortality during the pandemic period in both sexes (HR 0.17, 95% CI 0.09; 0.32 for men, HR 0.19, 95% CI 0.09; 0.38 for women).

5   Discussion

5.1   Methodological considerations

In this chapter, we review the methodological challenges encountered during the PhD project and consider how they may have influenced the results.

5.1.1. Study design

The PhD thesis is based on data from participants (N=2380) in the KYH study, which was a cross-sectional study at baseline (2015-2017). More than half of them (N=1348) took part in a cross-sectional ESSE-RF3 study, which was conducted 12-15 months after the onset of the COVID-19 pandemic (2021). The KYH and ESSE-RF3 data were linked to the COVID-19 case and vaccination registries. In 2015-2023, the KYH sample was followed up as a cohort for all-cause mortality and new clinical diagnoses, and these longitudinal data were also used in this thesis.

Paper I and **Paper II** were based on data from the cross-sectional ESSE-RF3 data. The cross-sectional study design was relevant for assessing COVID-19 seroprevalence (**Paper I**) and the spectrum of COVID-19 cases by combining the serological survey data, the national COVID-19 case registry, and self-reported data on COVID-19 experience and symptoms (**Paper II**) (163).

The timing of the serological survey may affect seroprevalence estimates. Depending on the time of sampling, infection rates may change, and individuals in a target population may be at a higher or lower risk of COVID-19 than average (7). Seroprevalence estimates may be biased if individuals are enrolled at different stages of local epidemics, potentially leading to either overestimation or underestimation of prevalence. We conducted the serological survey (**Paper I**) after the peak of the second wave of COVID-19 in Arkhangelsk (September 20, 2020 - March 3, 2021). Blood samples for serological testing were collected over a period of four months as part of the data collection for the ESSE-RF3 study, 20 to 25 participants per day. We assessed a "period seroprevalence" reflecting the prevalence of IgG antibodies to SARS-CoV-2 in the KYH population during the period from February 24, 2021, to June 30, 2021, when the daily number of new COVID-19 cases was relatively stable. Therefore, the seroprevalence can be considered as an average estimate for the period studied.

We extended the cross-sectional design to explore the association between various factors

and outcomes of interest, including serological status, adherence to NPIs, symptomatic infection, and hospitalization (**Papers I & II**). As all the data used in these analyses were collected simultaneously, investigating associations of interest in a cross-sectional study presented challenges in determining the temporality of associations between exposure and outcome, making it difficult to draw causal inferences (84). Therefore, some potential for reverse causality, where the outcome occurred before the exposure, should also be considered. For example, this may partially explain the negative association between smoking and symptomatic COVID-19.

Paper III presented the results of a prospective cohort study of changes in mortality and risk factors associated with all-cause death in the KYH study sample in the pandemic period compared with the pre-pandemic period. Although the cohort study design was appropriate for assessing risk factors, it was necessary to clarify what was meant by the term "risk factor", as recommended elsewhere (164). In our study, we defined a risk factor as a factor associated with a change (increase or decrease) in the risk of the outcome, which is consistent with Porta's Dictionary of Epidemiology (2). We refrained from using the term "predictors" because our goal was not to predict the outcome or to construct the best predictive model. Instead, we estimated and compared factors influencing the risk of death in the pre-pandemic and pandemic periods to better understand the mechanisms of the impact of the COVID-19 pandemic on mortality.

5.1.2 Study validity

Observational studies are subject to both random and systematic (non-random) errors, which are universally present to some extent and may affect the accuracy of the results. Internal validity is the validity of the results as they apply to the study population, which refers to the absence of systematic errors, including selection bias, information bias, and confounding (165). The internal validity of the study presented in the PhD thesis refers to how applicable inferences are to the sampled population (the group of individuals studied). External validity reflects the extent to which study results can be generalized to a larger target population. Internal validity is a prerequisite for achieving external validity.

5.1.2.1 Selection bias and external validity

Selection bias occurs when individuals have unequal chances of being included in a study based on their exposure and outcome characteristics, resulting in a non-representative

sample (165). This bias can affect the external validity and generalizability of study findings.

Participants in the KYH study (**Paper III**) were randomly selected from the adult population of Arkhangelsk aged 35-69 years, with a participation rate of 68.2%, which is relatively high for population-based studies. Although participation rates may not be strongly related to non-response bias (166, 167), we cannot completely rule out the possibility that participants who agreed to participate may differ from those who did not. To assess the extent of selection bias due to non-response, the educational profile of the KYH study sample was compared with that of the Arkhangelsk population according to the 2010 Russian Census (153). The educational profile of those recruited to the study was similar to what would be expected based on the 2010 Russian Census data for the city (153).

External validity is critical for a seroprevalence study to ensure that the results accurately reflect the virus transmission by assessing the prevalence of having antibodies to SARS-CoV-2 in the population of interest (7). One year after the start of the COVID-19 pandemic, participants in the KYH study were invited to participate in a seroprevalence survey (**Paper I**) conducted as a satellite of the ESSE-RF3 study. The seroprevalence study sample represented a resurveyed subsample of the KYH study with a participation rate of 59.7% (1348 out of 2258 in the sampling frame). **Paper II** used the same subsample after excluding vaccinated individuals, who may differ from those remaining in the sample. Although participants in the seroprevalence study were recruited regardless of exposure or disease status, we cannot exclude the possibility that individuals who had not been exposed to COVID-19 may have avoided participation due to fear of infection. Conversely, those who believed they have had COVID-19 may have felt safer participating due to their perceived immunity. Individuals who refused to participate and avoided healthcare facilities may have been more likely to self-isolate and adhere more strictly to other NPIs, possibly leading to a higher likelihood of being seronegative.

Since the participation rate of 59.7% and the exclusion of some participants could be a source of selection bias, we compared the key demographic characteristics of the resurveyed subsample used in **Paper I** (N=1332) and **Paper II** (N=1089) with the characteristics of the 2380 individuals in the KYH sampling frame (Table 6).

We observed that the study samples for **Paper I** and **Paper II** were slightly skewed toward individuals who were younger at the time of inclusion in the KYH study and had a slightly higher proportion of individuals with higher education. At the time of the resurvey (the ESSE-RF study, 2021), the age of participants ranged from 40 to 74 years, as they aged over

time after inclusion in the KYH study (2015-2017). The resurveyed subsample did not include KYH participants who died between the studies or were unable to participate in the ESSE-RF study due to serious illness.

Table 6. Comparisons between current study participants and the original Know Your Heart study sample.

Characteristic	KYH study sample, N=2380 (%)	Paper I study sample, N=1332 (%)	p-values[1]	Paper II study sample, N=1089 (%)	p-value[1]
Age at the time of inclusion in KYH study, Me (Q1; Q3)	54 (45; 62)	52 (44; 61)	0.001	51 (44; 58)	<0.001
Sex					
Women	1391 (58.4)	789 (59.2)	0.640	665 (61.1)	0.145
Men	989 (41.6)	543 (40.8)		424 (38.9)	
Higher education[2]					
No	1455 (61.1)	769 (57.7)	0.043	630 (56.9)	0.019
Yes	925 (38.9)	563 (42.3)		469 (43.1)	

KYH study – Know Your Heart study; Me – median; Q1 – first quartile; Q3 – third quartile
[1]Each study sample was compared to the KYH study sample using Pearson's chi-squared test for categorical parameters, and the Mann–Whitney U-test for continuous characteristics.
[2]Education levels were assessed at the time of inclusion in the KYH study for the KYH study sample and at the time of inclusion in ESSE-RF3 for the Paper I and Paper II study samples.

A slightly higher proportion of participants in the resurveyed subsample had higher education compared to the sampling frame, possibly because older participants who died or dropped out had a lower proportion of individuals with higher education. Some individuals may have attained higher education between studies, and those with higher education may have increased their health awareness and willingness to participate in the study. These differences between the resurveyed subsample and the KYH sampling frame are unlikely to significantly affect the results and conclusions.

Nevertheless, the results of the seroprevalence study cannot be generalized to the entire population of Russia due to significant regional differences in socioeconomic levels, climate, implemented NPIs, and approaches to organizing medical care during the pandemic. Considering all of the above sampling characteristics, the seroprevalence study sample may not be fully representative of the target population of Arkhangelsk residents. As the study sample was restricted to a specific age group, the results cannot not be generalized

to the entire population of Arkhangelsk, as younger and older individuals may have a different risk of infection and serological profile compared to middle-aged individuals (7). As the KYH study participants were followed to investigate risk factors for all-cause death during the pandemic compared to the pre-pandemic period, the possibility of attrition bias could be considered. Attrition bias is a type of selection bias, also known as dropout bias or loss-to-follow-up bias (168). Some individuals may have relocated, making it impossible to record their outcomes. Consequently, they may have been misclassified as alive, potentially resulting in underestimated mortality rates during the pandemic period.

5.1.2.2 Information bias

Information bias results from misclassification of exposure and/or outcome status (165). The type and quality of data sources used in epidemiological studies could have led to some extent of information bias. Our study combines several sources of information, including registry data, self-reported survey data, seroprevalence survey results, and blood-based biomarkers. Below, we outline potential limitations associated with the use of each of these sources in terms of information bias.

Registry data

The KYH and ESSE-RF3 study data were linked to the COVID-19 case registry, the vaccination registry, and the regional mortality registry. The completeness and reliability of the registry data could potentially influence the study results (169). The COVID-19 case registry provided accurate information on COVID-19 cases detected by the healthcare system (79, 170). We found only four out of 57 hospitalized cases with discrepancies between self-reported and registry data on hospitalization. The case registry relied primarily on positive PCR test results rather than clinical symptoms. However, in the early stages of the pandemic, limited capacity for PCR testing and the heavy burden on the healthcare system resulted in inequitable access to COVID-19 testing. Individuals from high-risk groups (such as the elderly, those with chronic conditions, and healthcare workers) or those with severe COVID-19, who have received more medical attention, may have been more likely to be tested and included in the registry. Even as testing availability increased, individuals with symptomatic infections were more likely to seek testing or to be identified and tested as contacts than those with atypical or no symptoms. As with any case registry in any setting, the COVID-19 registry likely missed a large number of asymptomatic and mild cases (171).

Individuals with proactive health-seeking behaviors, such as women compared to men, and those committed to a healthy lifestyle may be more willing to be tested (172). However, even when PCR testing is performed, the reliability of test results may be compromised due to improper timing or specimen collection (55, 84). All of these factors may lead to potential misdiagnosis or missed diagnosis of COVID-19, resulting in non-differential misclassification of individuals who had COVID-19 and those who did not. Reliance on registry data alone may underestimate the prevalence of infection and the associations between different factors and symptomatic status (86). We addressed this issue by integrating the registry data with serological survey results and self-reported COVID-19-related information (**Paper II**), which allowed us to obtain reliable estimates of the proportion of infected cases as well as the proportions of asymptomatic and symptomatic COVID-19 cases.

Completeness of the vaccination registry was another challenge related to non-differential misclassification of vaccinated and unvaccinated participants. Exclusion of vaccinated individuals was necessary to ensure accurate assessment of the prevalence of infection-acquired immunity (**Paper I**) and the spectrum of COVID-19 cases (**Paper II**). To address potential completeness issues in the vaccination registry, we combined self-reported vaccination data with registry information. Individuals who reported receiving the SARS-CoV-2 vaccine and had a corresponding record in the vaccination registry were excluded, as were participants with discrepancies in their vaccination status. Therefore, there was low probability that vaccinated participants were included in the study of factors associated with infection-acquired immunity.

The mortality registry data used in our study are based on the causes of death listed on death certificates, which include the underlying cause of death, related pathological conditions, immediate cause of death, and other contributing conditions according to ICD-10. The accuracy of information in death certificates may affect cause-specific mortality rates. In Russia, death certificates were based on autopsies and recorded according to national guidelines that met WHO guidelines for recording causes of death during the pandemic, which made them reliable (131, 133). However, the lack of testing in the early stages of the pandemic may have led to an underreporting of COVID-19 deaths, as some cases might have gone untested and been misclassified as unrelated to COVID-19 (117, 134). Even among deceased individuals who tested positive for SARS-CoV-2, COVID-19 was not necessarily recoded as the underlying cause of death (137). Deaths of SARS-CoV-2 positive individuals can be classified according to the role of the virus in the sequence leading to death. If a death resulted from COVID-19 complications, such as pneumonia with acute respiratory distress

syndrome or sepsis, COVID-19 was recorded as the underlying cause (112). When a death was attributed to other underlying causes, COVID-19 was listed as a contributing factor on the death certificate (138). In SARS-CoV-2 positive deceased patients with chronic diseases, especially those with multiple comorbidities, determining the role of COVID-19 in the sequence of events leading to death may be challenging (136, 139). Patients with comorbidities were predisposed to severe COVID-19, and the virus may have contributed to a fatal outcome by being pathogenically associated with a condition that became the immediate cause of death. In these cases, COVID-19 was likely recorded as a contributing factor, potentially leading to an underestimation of COVID-19 deaths (113, 135).

Therefore, recording of causes of death during the COVID-19 pandemic posed significant challenges due to changes in death reporting practices over time, low availability of testing, and the complex role of comorbidities in the pathway to death. In **Paper III**, we analyzed all-cause mortality to address these challenges and provide a comprehensive assessment of the impact of the pandemic on mortality.

Self-reported data

The use of self-reported data can be a source of misclassification, as study participants may inaccurately report information about themselves due to recall bias, social desirability bias, or misunderstanding of questions. We used self-reported data on behavioral and health-related characteristics collected during the KYH and ESSE-RF3 studies. Self-reported data on chronic conditions in population-based studies are generally considered reliable, but withholding sensitive information about behavioral factors such as smoking and alcohol consumption may introduce social desirability bias to undervalue socially unwelcome attitudes (173, 174). Due to the potential influence of social desirability bias, self-reported smoking and alcohol consumption may have been underestimated, while adherence to NPIs may have been overestimated.

In our study, adherence to NPIs was assessed using a composite variable consisting of five items: self-isolation (stay-at-home order), social distancing, wearing face masks in public places or transport, wearing gloves, and using hand sanitizers. We assessed the correlation of the items (internal consistency) and estimated the measurement accuracy of the composite adherence score (reliability) by using the Cronbach's alpha coefficient, which was 0.67, indicating acceptable reliability (175). The term "self-isolation" was not explicitly defined in the ESSE-RF3 questionnaire, which may have led participants to interpret it as staying at home during paid non-working days in March-May 2020, being isolated due to a confirmed

COVID-19 case, or being under quarantine due to contact with a confirmed case (65-67). This ambiguity may obscure the relationship between self-isolation and SARS-CoV-2 seroprevalence. The ESSE-RF3 questions on NPIs adherence did not consider the consistency or thoroughness of NPIs use during the pandemic, such as the frequency or appropriateness of using face masks and hand sanitizers. Thus, inappropriate NPIs use (e.g., wearing face masks with the nose exposed, failure to change masks in a timely manner, etc.) could not be completely excluded among those who self-reported high adherence. Moreover, behaviors could have changed during the pandemic period, resulting in an underestimated association between serological status and adherence to NPIs due to non-differential misclassification of adherence status (176).

Self-reported survey data may better capture symptomatic COVID-19 cases, including those who did not seek medical care (87). However, retrospective self-report of COVID-19 symptoms may be subject to recall bias, as participants may not recall experiencing symptoms, potentially leading to non-differential misclassification of symptomatic status. Due to the difficulty in distinguishing COVID-19 from other respiratory infections, individuals who had symptoms that were not severe enough to prompt a hospital visit and testing may be unaware or unsure whether they had COVID-19 (177). Thus, some symptomatic cases who were not tested may misclassify themselves as non-infected, potentially leading to an underestimation of the number of symptomatic COVID-19 cases.

Seroprevalence survey data

Many serological tests have been developed to detect SARS-CoV-2 antibodies, with varying test performance characterized by sensitivity and specificity. Sensitivity is the probability of a positive result in infected individuals, while specificity is the probability of a negative result in non-infected individuals. Imperfect test performance may lead to misclassification and biased seroprevalence estimates (178, 179).

We used the semi-quantitative ELISA assay to detect IgG antibodies to SARS-CoV-2 S protein in blood serum, the only test available in Russia during the study period (February 24-June 30, 2021). This assay allowed us to classify participants as having positive, negative, or equivocal results, but did not allow us to present the distribution of quantitative values. Participants with equivocal results were excluded due to uncertainty in interpretation. Based on the results of an independent evaluation, the test had a sensitivity of 89% and a specificity of 100% (156). The use of an imperfectly sensitive test, if not adjusted for test performance, may underestimate the seroprevalence due to the presence of infected cases not

detected by the test (false negatives). False negative rates may be higher in individuals with mild or asymptomatic infections, in recently infected persons who have not yet developed antibodies, or in those infected long before testing due to waning antibody levels over time after infection (89).

To ensure comparability of our results with those of other studies that used different serological tests, we reported seroprevalence adjusted for test performance indicators with 95% confidence intervals (**Paper I**) (179, 180). We estimated the seroprevalence adjusted for test performance using the equation (crude prevalence + test specificity - 1) / (test sensitivity + test specificity - 1) (159). We calculated 95% CIs for the adjusted seroprevalence to account for uncertainty in the estimates by bootstrapping using the R package (160). Thus, based on 867 positive test results among 1332 individuals tested for antibodies to SARS-CoV-2 (**Paper I**), the unadjusted seroprevalence was 65.1% (95% CI: 62.5; 67.6). After adjustment for serological test performance, the seroprevalence estimate changed to 73.0% (95% CI: 67.1; 85.7).

The test used could not distinguish between previously infected individuals and those vaccinated against SARS-CoV-2 because it detected IgG to the S protein but not to the N-protein of the virus. Therefore, we assessed the seroprevalence due to infection (**Paper I**) and the spectrum of COVID-19 cases (**Paper II**) after excluding those who were vaccinated. We estimated the seroprevalence due to infection to be 59.1% (95% CI: 56.1, 62.0), which increased to 66.3% (95% CI: 58.1, 76.0) after adjustment. This adjusted seroprevalence rate may serve as an approximation of the cumulative incidence rate (89). In **Paper II**, no adjustments were made for test performance because the proportions of infected individuals were estimated using data from the COVID-19 case registry and self-reported data in addition to serological test results. However, the proportions of infected cases and seropositive participants were almost the same, with only five of 220 COVID-19 cases in the case registry testing negative for SARS-CoV-2 antibodies. Thus, the underestimation due to the imperfect test sensitivity may be very small.

Blood-based biomarkers

In **Paper III**, we examined the impact of blood-based biomarkers measured at the time of enrollment in the KYH study, six years before the onset of the pandemic, on all-cause mortality during both the pre-pandemic and pandemic periods. As the KYH study was designed to investigate cardiovascular diseases, the baseline data included predominantly cardiovascular biomarkers. Measuring blood-based biomarkers at a single point in time may

not provide a complete risk profile because it may not accurately reflect long-term exposure or average levels of these biomarkers, which may change over time. Changes in medication use may also have influenced biomarker levels and disease progression in terms of underlying pathophysiological mechanisms. For example, participants with an initially abnormal blood lipid profile who were not previously taking lipid-lowering medications may have subsequently started treatment. Changes in the use of lipid-lowering medications during the follow-up period may have affected lipid levels. Due to variations in biomarker levels over time and potential non-differential misclassification, associations between medication-dependent biomarkers and outcomes should be interpreted with caution.

5.1.2.3 Confounding and interaction

A confounder is a variable that is associated with both the exposure and the outcome, but is not on the pathway from the exposure to the outcome (165). Confounding occurs when there are differences between the exposed and unexposed groups regarding a third variable that influences the occurrence of the outcome (2). This results in a biased association between an exposure and an outcome, potentially indicating a false effect of the exposure or masking the true effect. Confounding can be reduced by applying statistical correction methods such as restriction, stratification/standardization, and multivariate regression (165).

A data-driven approach and Directed Acyclic Graphs (DAGs) are used to identify variables that should be adjusted for in statistical analyses to control for confounding and produce accurate estimates. In a data-driven approach, variables are selected for inclusion in the multivariate model based on their statistical significance in univariable analyses, which assess crude associations between various factors (exposures) and outcomes (181, 182). Variables are not included in the model if they are not significant and are not considered confounders. Significance is assessed at the 0.1-0.2 alpha level in the crude model (181). Confounding is assessed using the change-in-estimate procedure, which involves observing the effect of adding a potential confounder to a statistical model on the effect estimate of the exposure variable. If the effect estimate changes by 10-20% or more after adjustment, this indicates that the added variable is a confounder (181, 183).

Although a data-driven approach provides valuable insight into associations between variables, it does not consider the role of each variable in relation to exposure and outcome. Thus, selection based on significance may lead to inappropriate exclusion or inclusion in the model.

In contrast to the data-driven approach, DAGs are graphical tools used to visualize the

relationships among variables and determine the role of covariates in a statistical model, including the identification of confounders, mediators (variables in the pathway from exposure to outcome), and colliders (variables affected by both exposure and outcome) (Fig.7) (184).
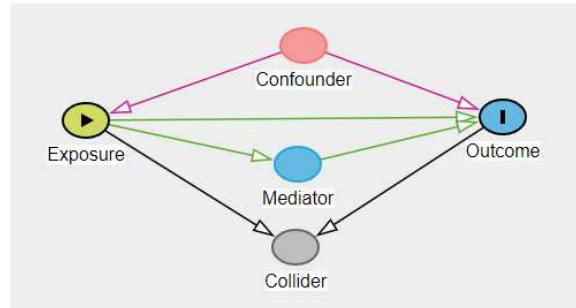
Figure 7. A directed acyclic graph representing the differences between confounders, mediators, and colliders.

In **Paper II**, we focused on associations between various factors and outcomes rather than on predicting outcomes. Therefore, adjustment for confounders specific to certain covariates should be considered to obtain interpretable effect estimates. Given the limited power of the study, we adopted a strategy of adjusting for potential confounders for all key variables (variables of interest) selected from the DAGs, including age, sex, higher education, smoking, and heavy drinking (Fig.8). Additional adjustment for variables specific to certain covariates did not significantly alter the results.

Similar results were obtained for confounder selection using the data-driven approach. The factors chosen as confounders for all key variables using DAGs were significantly associated with outcomes in the crude analysis. With this in mind, we used logistic regression models to investigate the associations between the studied variables and outcomes (symptomatic status, hospitalization), both in unadjusted analyses and after adjusting for the potential confounders. We presented unadjusted estimates alongside the confounder-adjusted estimates to show whether the observed associations could be explained by confounding.

In **Paper I**, we investigated factors associated with seropositive status and adherence to NPIs using the multivariate logistic regression models. The investigation of factors associated with seropositive status involved the stepwise introduction of blocks of variables into the model. To assess the isolated effect of each variable on adherence to NPIs, adjusted estimates were calculated by including all variables simultaneously in a regression model. Mutual adjustment does not take into account the relationships between variables and does not treat each independent variable separately as an exposure, but as any other covariate. This may lead to

the "Table 2 fallacy," which is the belief that effect estimates for all covariates derived from a single multivariable model have a similar interpretation, which is often not the case (185).

Symptomatic COVID-19   **Hospitalization status**

Hypertension

Diabetes

Abdominal obesity

Self-rated health
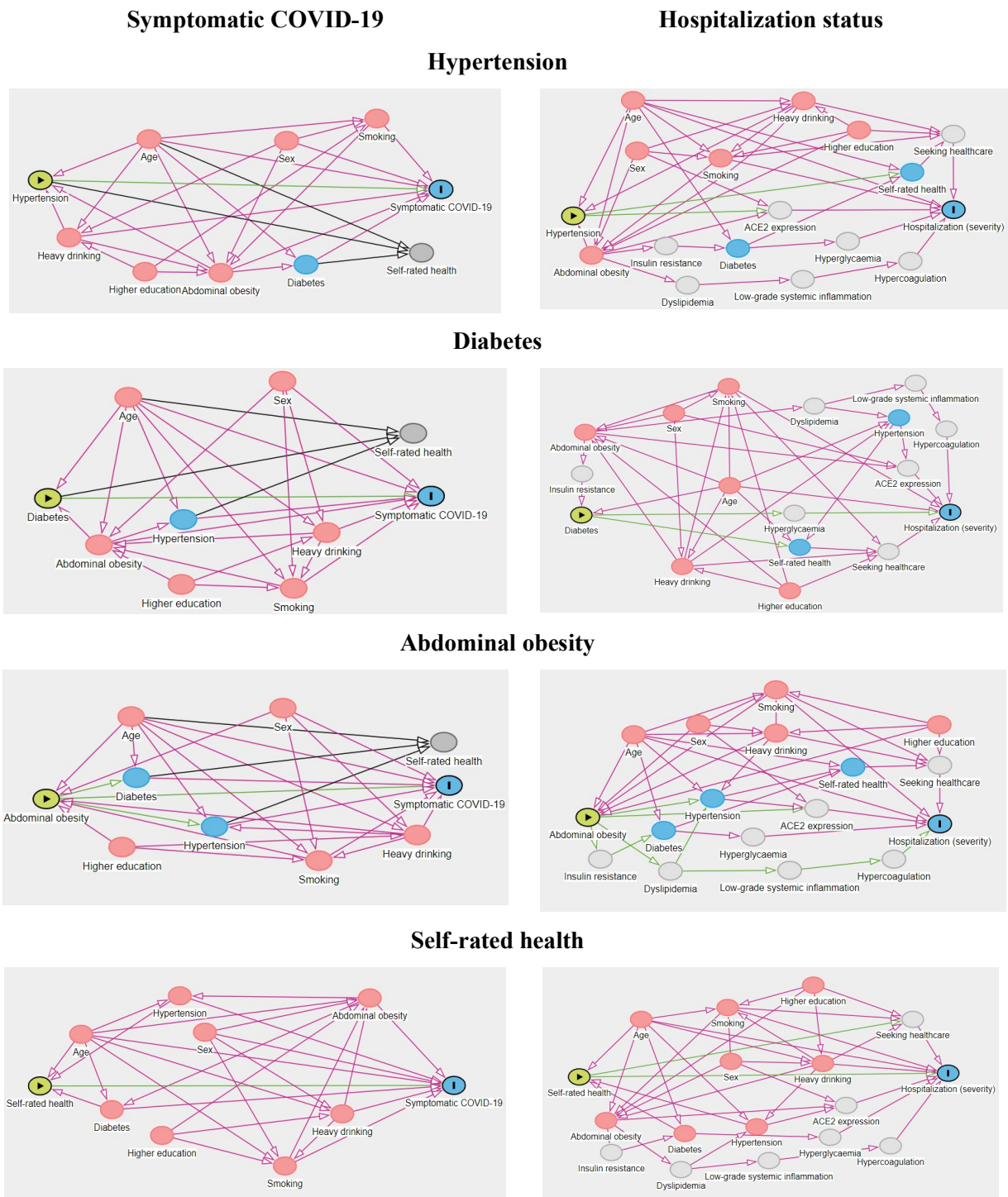
Figure 8. Directed acyclic graphs showing confounding in the estimation of associations between key variables and outcomes.

The DAGs constructed post hoc to account for the relationships between the variables

included in the analysis did not reveal any colliders, making collider bias unlikely (186). As some variables may be mediators, adjusting for them may lead to an underestimation of the total effect of some factors (191, 192). This should be taken into account when interpreting the results.

In **Paper III**, we standardized mortality rates to the age distribution of the study population in the pre-pandemic period (direct standardization, 5-year bands) to account for the aging of the sample over time following inclusion in the study. Cox proportional hazards regression models were fitted to investigate the factors associated with the change in risk of death during the pandemic compared with the pre-pandemic period. Effect estimates were reported as HRs adjusted for demographic (age, higher education) and behavioral (smoking, hazardous drinking) variables selected by DAGs as confounders for key covariates.

In **Paper III**, we identified an interaction between sex and study period, indicating that the average effect of study period on all-cause mortality depended on sex. Therefore, further analyses were stratified by sex. Failure to account for this interaction could have compromised the internal validity of the study results. Our findings confirmed this interaction by revealing sex differences in how the associations of chronic diseases and blood-based biomarkers with all-cause mortality changed during the pandemic compared with the pre-pandemic period.

5.1.3   Sample size and power calculation

The sample size should be sufficient to address the research questions with precision and to ensure adequate statistical power to obtain conclusive results. Since we used data collected in the previous KYH and ESSE-RF3 studies, the sample size was relatively small, which was a major limitation of the study. The calculation of the power achieved was performed using post hoc analysis. While investigating the association between demographic, behavioral, and health-related factors and seropositive status, we calculated the statistical power of a binary logistic regression model to identify ORs of 1.5 and greater for outcome prevalence of 10 to 90% and predictor prevalence of 10, 25, 50, 75, and 90% in the sample of 1332 participants (**Paper I**). Calculations showed that the sample has ≥80% power to identify factors that increase or decrease the odds of the outcome by 1.5-fold for all combinations of outcome prevalence in the range of 25% to 70% and predictor prevalence in the range of 30% to 70% (Table 7).

Table 7. Statistical power of binary logistic regression models to identify factors associated with seropositive status (G*Power 3.1.9.7)

Outcome prevalence	Predictor prevalence																
	0,1	0,15	0,2	0,25	0,3	0,35	0,4	0,45	0,5	0,55	0,6	0,65	0,7	0,75	0,8	0,85	0,9
0,1	0,32	0,41	0,48	0,53	0,57	0,6	0,62	0,62	0,62	0,61	0,59	0,57	0,53	0,47	0,4	0,32	0,23
0,15	0,4	0,52	0,6	0,66	0,7	0,73	0,75	0,76	0,76	0,76	0,74	0,71	0,67	0,62	0,54	0,44	0,32
0,2	0,46	0,59	0,68	0,74	0,79	0,82	0,83	0,84	0,84	0,84	0,83	0,8	0,77	0,72	0,64	0,53	0,39
0,25	0,51	0,64	0,73	0,8	0,84	0,86	0,88	0,89	0,89	0,89	0,88	0,86	0,83	0,78	0,71	0,6	0,45
0,3	0,53	0,68	0,77	0,83	0,87	0,89	0,91	0,91	0,92	0,91	0,9	0,89	0,86	0,82	0,75	0,64	0,5
0,35	0,55	0,7	0,79	0,85	0,88	0,91	0,92	0,93	0,93	0,93	0,92	0,91	0,88	0,84	0,78	0,67	0,53
0,4	0,56	0,71	0,8	0,86	0,89	0,92	0,93	0,94	0,94	0,94	0,93	0,91	0,89	0,85	0,79	0,7	0,55
0,45	0,56	0,71	0,8	0,86	0,9	0,92	0,93	0,94	0,94	0,94	0,93	0,92	0,9	0,86	0,8	0,71	0,56
0,5	0,55	0,7	0,79	0,85	0,89	0,91	0,93	0,94	0,94	0,94	0,93	0,92	0,89	0,86	0,8	0,71	0,56
0,55	0,53	0,68	0,78	0,84	0,88	0,91	0,92	0,93	0,93	0,93	0,92	0,91	0,89	0,85	0,79	0,7	0,55
0,6	0,5	0,65	0,76	0,82	0,86	0,89	0,91	0,92	0,92	0,92	0,91	0,89	0,87	0,83	0,77	0,68	0,54
0,65	0,47	0,62	0,72	0,79	0,84	0,87	0,88	0,9	0,9	0,9	0,89	0,87	0,85	0,81	0,75	0,65	0,52
0,7	0,42	0,57	0,67	0,75	0,8	0,83	0,85	0,86	0,87	0,87	0,86	0,84	0,81	0,77	0,71	0,62	0,48
0,75	0,37	0,51	0,61	0,69	0,78	0,78	0,8	0,81	0,82	0,82	0,81	0,79	0,76	0,72	0,65	0,57	0,44
0,8	0,31	0,43	0,53	0,6	0,66	0,7	0,72	0,74	0,75	0,75	0,74	0,72	0,69	0,65	0,59	0,5	0,39
0,85	0,24	0,34	0,42	0,49	0,54	0,58	0,61	0,63	0,64	0,64	0,63	0,62	0,59	0,54	0,49	0,42	0,33
0,9	0,16	0,23	0,3	0,35	0,4	0,43	0,46	0,48	0,49	0,49	0,49	0,47	0,45	0,42	0,38	0,33	0,26

For the sample of 1089 participants (**Paper II**), a binary logistic regression model had ≥80% power to detect factors that increased or decreased the odds of the outcome by 1.5-fold for all combinations of outcome prevalence between 30% and 60% and predictor prevalence between 35% and 65%. The relatively small sample size we used for analyses of factors associated with symptomatic status and hospitalization may not have allowed us to detect relatively weak associations, limiting the interpretation of the results.

For the sample of 2357 participants (**Paper III**), the Cox proportional hazards regression model has ≥80% power to detect a 1.5-fold increase or decrease in the risk of death associated with most of the characteristics studied. The relatively small number of deaths observed in the study sample due to the relatively young age of the participants may have limited the interpretation of the results regarding factors associated with risk of death.

5.2 Discussion of the main results

In this chapter, we have elaborated on the results presented in the papers and discussed the findings in relation to the objectives of the PhD project.

5.2.1 SARS-CoV-2 seroprevalence and spectrum of COVID-19 cases

Two-thirds of the population-based adult sample had specific SARS-CoV-2 antibodies 12-15 months after the onset of the pandemic in Arkhangelsk (**Paper I**). The level of population immunity was largely determined by the spread of infection, with only 14.6% being fully vaccinated. Among unvaccinated participants, 66.3% were seropositive (adjusted for test performance), which was considered to be infection-acquired immunity, regardless of whether they had symptoms or not. Despite a high proportion of the adult population having antibodies, the rate of new COVID-19 cases recorded in Arkhangelsk remained high in 2021 (73). This suggests that achieving herd immunity may be challenging due to the emergence of new SARS-CoV-2 variants to which people had partial or no immunity, coupled with waning antibody levels over time (187).

The COVID-19 case registry recorded only one in three infected cases (**Paper II**). Under-recording may have occurred because only a proportion of the infected individuals (47.1%) exhibited symptoms, and not all symptomatic cases sought medical care and were tested. According to the iceberg concept of infectious diseases, detected COVID-19 cases represent the visible tip of the iceberg, while undetected, mostly asymptomatic cases represent the invisible bottom (188) (Fig.9). In our study, one third of the participants had an asymptomatic infection prior to inclusion in the study; 96.2% of all asymptomatic cases were unaware that they had been infected.
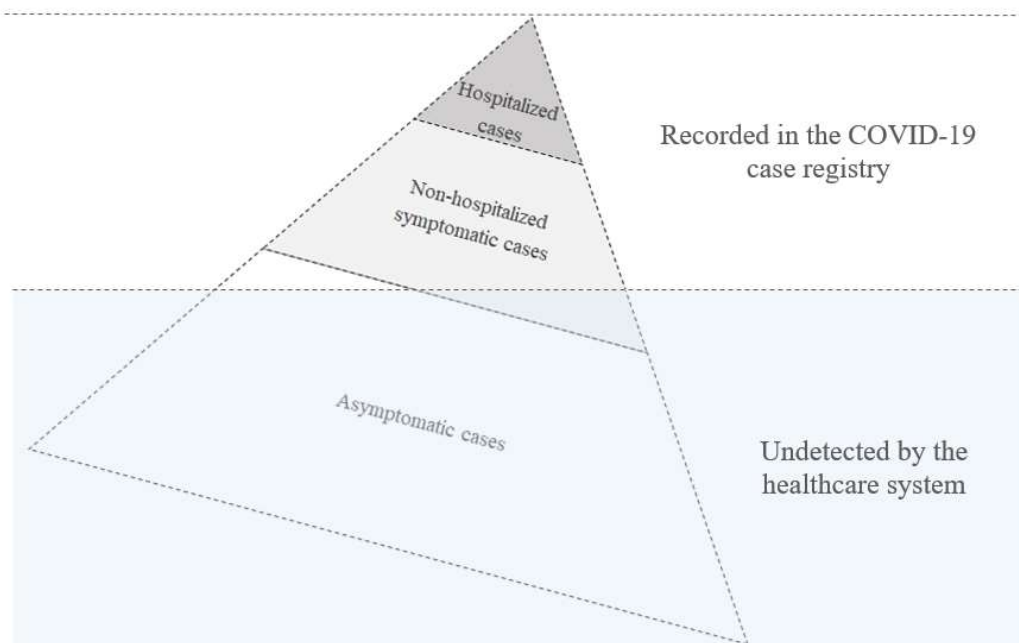

Figure 9. The iceberg concept of COVID-19 (adapted to COVID-19 from the paper published by Shubin M. et al.) (188).

The high proportion of symptomatic cases (18.3%) were hospitalized, with fever and dyspnea being more frequent symptoms in hospitalized cases, consistent with previous studies (189, 190). The most specific COVID-19 symptoms, anosmia (loss of sense of smell) and ageusia (loss of sense of taste), were more common in non-hospitalized symptomatic cases, consistent with findings showing a higher frequency in non-severe cases (191).

The high hospitalization rate for COVID-19 cases in our study may be partly explained by the common recommendation during the early stages of the pandemic in Russia to hospitalize all infected individuals for isolation (68). In support of this, a study in Saint Petersburg from March 2020 to April 2021 found higher hospitalization rates than our study. The rates for men ranged from 14.6% (under 49 years) to 56.9% (over 70 years), and for women from 11.5% (under 49 years) to 48.4% (over 70 years) (192). However, since only one of 57 hospitalized patients was asymptomatic, we considered hospitalization to be an acceptable indicator of COVID-19 severity in our study.

5.2.2 Factors associated with adherence to preventive measures, serological status, and COVID-19 severity

In our study (**Paper I**), less than half (48.9%) of the participants adhered to all five recommended NPIs, which included self-isolation (stay-at-home order), social distancing, wearing face masks in public places or on transport, wearing gloves, and using hand sanitizers. Self-reported high adherence to NPIs, defined as adherence to at least four out of the five measures, was not associated with contracting COVID-19. This may be partly explained by the fact that, in addition to the number of NPIs used, the risk of infection depended on the consistency and thoroughness of NPI use (176, 193). These behavioral factors are influenced by perceived risk of infection, self-discipline, and social responsibility, which could not be precisely measured (5).

Younger individuals had lower adherence to COVID-19 NPIs (**Paper I**) and a higher proportion of infected cases (**Paper II**) compared to older age groups, as reported in other studies (103). However, after adjustment for other factors, including regular employment, age was no longer associated with adherence to NPIs, consistent with other findings (194). Regular employment was independently associated with lower adherence to NPIs and with the presence of infection-acquired immunity (**Paper I**). Employed individuals were at increased risk of exposure to SARS-CoV-2 due to factors such as the inability to work remotely, reliance on public transport, and potential social interactions at work. Low

adherence may be due to the challenges of strict adherence to NPIs, including maintaining social distance, proper use of face masks, and their timely replacement, which further contributed to the increased risk of infection (195).

NPIs targeting individuals over 65 and those with chronic conditions, particularly self-isolation, resulted in higher adherence rates among older, typically unemployed or retired participants (62). This may have contributed to potentially lower overall COVID-19 severity. Nevertheless, older participants were more likely to experience symptomatic COVID-19 and require hospitalization (**Paper II**). The prioritization of testing during the early stages of the pandemic may have resulted in a higher likelihood of symptomatic cases being tested, diagnosed, and hospitalized for isolation or severity among individuals over 65 years. Nevertheless, our findings are consistent with the well-established association between older age and disease severity (196, 197).

Male sex was independently associated with low adherence to NPIs, in line with previous research (**Paper I**) (98, 198). However, the majority of infected men were asymptomatic (**Paper II**), whereas more than half of infected women were symptomatic. Given the significant association of smoking with symptomatic status and the substantial difference in smoking prevalence between men and women in Russia (199), the difference in symptomatic status between the sexes is likely attributable to smoking. This is supported by the disappearance of the association between symptomatic status and sex after adjustment for other demographic and behavioral factors, including smoking (**Paper II**).

Smokers were less likely to have infection-acquired antibodies (103, 200) and to be diagnosed with COVID-19, and they were underrepresented among hospitalized COVID-19 cases, consistent with findings from other studies (201-203). The suggestion that tobacco smoke compounds may reduce susceptibility and vulnerability to COVID-19 is largely speculative (200, 204-206). It is possible that smokers are less likely to develop immunity after COVID-19 infection as they may develop a lower antibody response and have a shorter duration of antibody circulation after infection (207). Smokers may have a lower likelihood of PCR testing or self-reporting of COVID-19 symptoms (200, 208), thereby increasing the likelihood of being misclassified as asymptomatic. The higher frequency of false negative PCR results associated with tobacco smoking may have further reduced the prevalence of diagnosed COVID-19 in smokers (200, 206, 209). We cannot completely exclude the reverse causality as a possible explanation for the reduced odds of symptomatic COVID-19 among smokers (**Paper II**) - some asymptomatic participants may have started smoking during the pandemic due to factors such as depression or isolation (210). However, a recent report showed that

smoking prevalence among middle-aged adults decreased during the pandemic, suggesting an increase in smoking cessation (211). Based on this, participants who had a symptomatic course of the infection may have quit smoking before or during their illness because it could have worsened COVID-19 symptoms. These individuals, classified in the study as former smokers, may have a higher prevalence of chronic diseases compared to those who continued to smoke. This may have reduced the overall severity of COVID-19 in current smokers by removing more vulnerable individuals from this group (212). In support of this, a recent meta-analysis showed that former smokers were at increased risk of severity, hospitalization, and mortality from COVID-19 (200). Consistent with the evidence linking smoking to mortality from both COVID-19 and non-COVID-19 causes, smokers may have had a higher risk of serious illness and death prior to the start of the ESSE-RF3 study (213-215). They may also have declined to participate in the study due to poor health or post-COVID symptoms. Smokers were more likely to be vaccinated early in the pandemic (due to their fear of severe infection) and were therefore excluded from the study of the spectrum of COVID-19 cases. As a result, they may be underrepresented in the resurveyed subsample, potentially attenuating the observed effect of smoking on disease severity.

Among symptomatic cases, men were more likely to be hospitalized than women, but this association was no longer present after adjustment for age, education level, smoking, and heavy drinking. This contradicts the recent meta-analysis showing that male sex was independently associated with severe COVID-19 (216). Symptomatic COVID-19 patients with poor self-rated health were more likely to require hospitalization, consistent with other studies suggesting that poor self-rated health and comorbidities exacerbate COVID-19 severity (111, 196, 197). In unadjusted analyses, diabetes was associated with higher odds of COVID-19 hospitalization, but this association disappeared after adjustment for demographic and behavioral factors. Compared to self-reported diagnosis of a specific disease, poor self-rated health reflects an individual's subjective perception of their overall health, which may be influenced by the severity of underlying pathogenic changes, reflecting the stage of the disease, and the degree of control with medication. Individuals with multiple chronic conditions may more often perceive their health as poor. Consequently, symptomatic COVID-19 patients with poor self-rated health were more likely to experience severe symptoms requiring hospitalization.

5.2.3 Mortality and risk factors for death during the pandemic

During the COVID-19 pandemic, the risk of death from all causes increased by more than 40.0% compared with the pre-pandemic period. Given the relatively young age of study participants (40-74 years at the pandemic onset), most deaths can be considered premature. The age-standardized mortality rate in women doubled during the pandemic, while the change in men was minor (**Paper III**). This is consistent with the results of a previous study showing that women in Russia had higher excess mortality (the difference between observed and expected deaths) than men, in contrast to most countries where excess mortality was skewed toward men (217, 218).

Cardiovascular diseases and neoplasms were the leading underlying causes of death. In women, mortality from neoplasms exceeded mortality from cardiovascular diseases, whereas the opposite was observed in the pre-pandemic period. In men, cardiovascular diseases remained the leading cause of death during the pandemic. In addition to the susceptibility and vulnerability of patients with cardiovascular disease and neoplasms to COVID-19, the restrictions implemented during the pandemic to reduce viral transmission may have unintentionally affected these patients. Several studies have shown that mortality from acute cardiovascular events has increased while hospitalizations for these conditions have decreased, suggesting that the increase in deaths is largely due to delayed diagnosis and treatment (219, 220). A reluctance to seek care, reduced capacity for screening and optimal investigations, and delays in specific treatment can have a negative impact on cancer patients, particularly on the presentation and stage of cancer at diagnosis, which in turn worsens prognosis (221). In Russia, the COVID-19 pandemic has limited access to preventive screening, including for neoplasms, resulting in a decrease in the number of newly diagnosed cases of breast, prostate, kidney, and thyroid cancers (222). Moreover, patients treated for neoplasms and those with other serious illnesses were more likely to become infected with SARS-CoV-2 because of their weakened immune systems and the need to visit health care facilities (222).

Before the pandemic, women had a survival advantage over men, largely due to early diagnosis (at an early stage) and better adherence to treatment for chronic diseases (218, 223). Thus, the reduced availability of health care during the pandemic may have disproportionately affected women, who are generally more proactive in seeking health care (172). The higher prevalence of certain chronic conditions appears to have made women more vulnerable than men to potentially fatal complications directly related to the virus. This is supported by the

fact that most of the factors associated with an increased risk of all-cause death in women during the pandemic were also risk factors for death from COVID-19 (116, 224, 225).

In men, asthma was associated with a higher risk of all-cause death during the pandemic, probably due to poorer adherence to treatment compared to women. Other studies have shown that inhaled corticosteroids used to treat asthma may reduce the severity of COVID-19, leading to lower rates of hospitalization and death among users (226, 227). Elevated biomarkers of cardiovascular risk (hs-CRP, Hs-NT-proBNP and Hs-TnT) further increased the risk of death in men during the pandemic, suggesting their greater vulnerability to cardiovascular events related or unrelated to the virus.

The lack of sex differences in COVID-19 mortality may be due to underestimation of the impact of the virus and misclassification of COVID-19 deaths, particularly among women with a higher prevalence of diagnosed chronic diseases. Supporting this, another study suggested that the increase in deaths from cardiovascular diseases and neoplasms may be mainly associated with undetected deaths related to COVID-19 (221).

5.2.4 Public health implications of the findings

To our knowledge, this is the first population-based study in Russia to investigate the spectrum of COVID-19 cases and factors associated with seropositivity, symptomatic disease, and risk of death during the pandemic in the adult population.

Our study showed that much of the spread of SARS-CoV-2 went unnoticed by the healthcare system and that the use of multiple and complementary surveillance sources could ensure broader coverage of COVID-19 cases. We demonstrated that combining different sources of COVID-19-related data could make the COVID-19 surveillance system more sensitive to detect cases of varying severity representing different levels of the infectious disease pyramid. Integrating information from different sources could reduce the number of unidentified asymptomatic cases who, unaware of their disease status, could unknowingly transmit the virus to others. Estimating the proportion of asymptomatic COVID-19 cases also allows the calculation of key epidemiological characteristics, including the cumulative incidence of infection.

The use of population-based data collected several years before the pandemic allowed the identification of factors associated with risk of death during the pandemic in men and women. The study findings could be used to implement targeted prevention strategies for future outbreaks,including vaccine prioritization. It also increases the focus on facilitating a

sex-specific biomarker-based approach to effectively improve health outcomes and reduce mortality in high-risk groups. Assessing the proportion of infected cases and understanding the spectrum of COVID-19, including characteristics of asymptomatic, symptomatic, and hospitalizedcases, as well as risk factors for death, has facilitated the evaluation of pandemic control strategies and the overall impact of infectious disease outbreaks on population health. The results of the PhD project can inform health care planning, prioritize interventions to control the spread of infection, and identify areas for improvement in surveillance programs for future outbreaks.

6 Conclusions

SARS-CoV-2 spread rapidly through the population of Arkhangelsk, with two-thirds of adults aged 40-74 years (66.3%) being seropositive within one and a half years of the pandemic. Adherence to NPIs was not associated with serological status.

One third of the study participants had asymptomatic COVID-19, almost all of whom (96.2%) were unaware that they had been infected.

Of those infected, less than half (47.1%) had symptoms of COVID-19. A high proportion of symptomatic cases (18.3%) were hospitalized, mainly those over 65 years of age and those with poor self-rated health.

Smokers were less likely to have infection-acquired antibodies and more likely to be asymptomatic. However, the effect of smoking on COVID-19 severity could be underestimated and should be interpreted with caution.

During the pandemic, all-cause mortality was 41.0% higher than in the pre-pandemic period, with a greater increase in age-standardized mortality in women and minor change in men. Compared with the pre-pandemic period, women with obesity, angina, and kidney dysfunction and men with asthma and elevated cardiovascular risk biomarkers had an increased risk of all-cause death during the pandemic. Smoking and diabetes were associated with higher risk of death in both sexes in both periods.

7 Works cited

1. Liu Y, Eggo RM, Kucharski AJ. Secondary attack rate and superspreading events for SARS-CoV-2. Lancet (London, England). 2020;395(10227):e47.
2. Porta M et al. A dictionary of epidemiology. 6th ed. New York, NY: Oxford University Press, 2014.
3. Barreto ML, Teixeira MG, Carmo EH. Infectious diseases epidemiology. Journal of epidemiology and community health. 2006;60(3):192-5.
4. Woc-Colburn L, Godinez D. Lockdown as a public health measure: COVID-19 Pandemic. 2022:133-6.
5. Xylogiannopoulos KF, Karampelas P, Alhajj R. COVID-19 pandemic spread against countries' non-pharmaceutical interventions responses: a data-mining driven comparative study. BMC public health. 2021;21(1):1607.
6. Duval D, Evans B, Sanders A, Hill J, Simbo A, Kavoi T, et al. Non-pharmaceutical interventions to reduce COVID-19 transmission in the UK: a rapid mapping review and interactive evidence gap map. Journal of Public Health. 2024;46(2):279-93.
7. Accorsi EK, Qiu X, Rumpler E, Kennedy-Shaffer L, Kahn R, Joshi K, et al. How to detect and reduce potential sources of biases in studies of SARS-CoV-2 and COVID-19. European journal of epidemiology. 2021;36(2):179-96.
8. Murphy C, Lim WW, Mills C, Wong JY, Chen D, Xie Y, et al. Effectiveness of social distancing measures and lockdowns for reducing transmission of COVID-19 in non-healthcare, community-based settings. Philosophical transactions Series A, Mathematical, physical, and engineering sciences. 2023;381(2257):20230132.
9. Noppert GA, Hegde ST, Kubale JT. Exposure, Susceptibility, and Recovery: A Framework for Examining the Intersection of the Social and Physical Environments and Infectious Disease Risk. American journal of epidemiology. 2023;192(3):475-82.
10. Sender R, Bar-On YM, Gleizer S, Bernsthein B, Flamholz A, Phillips R, et al. The total number and mass of SARS-CoV-2 virions. Proceedings of the National Academy of Sciences of the United States of America. 2021;118(25):e2024815118.
11. Masana L, Correig E, Ibarretxe D, Anoro E, Arroyo JA, Jericó C, et al. Low HDL and high triglycerides predict COVID-19 severity. Scientific reports. 2021;11(1):7217.
12. Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. The New England journal of medicine. 2020;382(13):1199-207.
13. Andersen KG, Rambaut A, Lipkin WI, Holmes EC, Garry RF. The proximal origin of SARS-CoV-2. Nature medicine. 2020;26(4):450-2.
14. Pekar J, Worobey M, Moshiri N, Scheffler K, Wertheim JO. Timing the SARS-CoV-2 index case in Hubei province. Science (New York, NY). 2021;372(6540):412-7.
15. Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. Lancet (London, England). 2020;395(10223):507-13.
16. Tang G, Liu Z, Chen D. Human coronaviruses: Origin, host and receptor. Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology. 2022;155:105246.
17. The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. Nature microbiology. 2020;5(4):536-44.
18. Lau SKP, Luk HKH, Wong ACP, Li KSM, Zhu L, He Z, et al. Possible Bat Origin of Severe Acute Respiratory Syndrome Coronavirus 2. Emerg Infect Dis. 2020;26(7):1542-7.
19. Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, et al. A pneumonia outbreak

associated with a new coronavirus of probable bat origin. Nature. 2020;579(7798):270-3.

20.   Chorba T. The Concept of the Crown and Its Potential Role in the Downfall of Coronavirus: Emerg Infect Dis. 2020;26(9):2302-5.

21.   Ahmad A, Haq M, Rehman A, Haq NU. Anti-nucleocapsid IgG antibodies in SARS-CoV-2 recovered health care workers: One year follow-up study. International journal of immunopathology and pharmacology. 2023;37:3946320231187744.

22.   Pango Lineages: Latest epidemiological lineages of SARS-CoV-2. Available from: https://cov-lineages.org (accessed April 7, 2024).

23.   COVID-19 Public Health Emergency of International Concern (PHEIC) Global research and innovation forum https://www.who.int/publications/m/item/covid-19-public-health-emergency-of-international-concern-(pheic)-global-research-and-innovation-forum.

24.   WHO Director-General's Opening Remarks at the Media Briefing on COVID-19—11 March 2020. Available from: https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-oncovid-19---11-march-2020 (accessed on November 9, 2023).

25.   World Health Organization (WHO). Coronavirus (COVID-19) Dashboard. Available from: https://covid19.who.int/ (accessed on March 31, 2024).

26.   Ritchie, Hannah; Mathieu, Edouard; Rodés-Guirao, Lucas; Appel, Cameron; Giattino, Charlie; Ortiz-Ospina, Esteban; Hasell, Joe; Macdonald, Bobbie; Beltekian, Diana; Dattani, Saloni; Roser, Max (2020–2022). "Coronavirus Pandemic (COVID-19)". Our World in Data. Available from: https://ourworldindata.org/coronavirus (accessed on June 9, 2024).

27.   On the State of Sanitary and Epidemiological Well-being of the Population in the Russian Federation in 2022: State Report / Edited by R.V. Buzinov – Moscow, 2023. – 368 p.

28.   . WHO Director-General's opening remarks at the media briefing – 5 May 2023. Available from: https://wwwwhoint/news-room/speeches/item/who-director-general-s-opening-remarks-at-the-media-briefing---5-may-2023 (accessed on February 26, 2024).

29.   Li H, Wang Y, Ji M, Pei F, Zhao Q, Zhou Y, et al. Transmission Routes Analysis of SARS-CoV-2: A Systematic Review and Case Report. Frontiers in cell and developmental biology. 2020;8:618.

30.   Cevik M, Marcus JL, Buckee C, Smith TC. Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Transmission Dynamics Should Inform Policy. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2021;73(Suppl 2):170-6.

31.   Methi F, Madslien EH. Lower transmissibility of SARS-CoV-2 among asymptomatic cases: evidence from contact tracing data in Oslo, Norway. BMC medicine. 2022;20(1):427.

32.   Madewell ZJ, Yang Y, Longini IM, Jr., Halloran ME, Dean NE. Household Transmission of SARS-CoV-2: A Systematic Review and Meta-analysis. JAMA network open. 2020;3(12):e2031756.

33.   Weill JA, Stigler M, Deschenes O, Springborn MR. Social distancing responses to COVID-19 emergency declarations strongly differentiated by income. Proceedings of the National Academy of Sciences of the United States of America. 2020;117(33):19658-60.

34.   Statement on the meeting of the International Health Regulations (2005) Emergency Committee regarding the outbreak of novel coronavirus 2019 (n- CoV) on 23 January 2020. Available from: https://www.who.int/news-room/detail/23-01-2020-statement-on-themeeting-of-the-international-healthregulations-(2005)-emergency-committee-regarding-the-outbreak-of-novelcoronavirus-(2019-ncov) (accessed November 17, 2022).

35.   Liu Y, Gayle AA, Wilder-Smith A, Rocklöv J. The reproductive number of COVID-19 is higher compared to SARS coronavirus. Journal of travel medicine. 2020;27(2): taaa021.

36.   Rothman KJ, Greenland S, Lash TL. Modern epidemiology: Lippincott Williams & Wilkins; Philadelphia, 3rd edition, 2012.

37.   Kwok KO, Lai F, Wei WI, Wong SYS, Tang JWT. Herd immunity - estimating the level required to halt the COVID-19 epidemics in affected countries. The Journal of infection. 2020;80(6):32-3.

38.   Altmann DM, Douek DC, Boyton RJ. What policy makers need to know about COVID-19 protective immunity. Lancet (London, England). 2020;395(10236):1527-9.

39.   Tracking SARS-CoV-2 variants. Available from: https://www.who.int/activities/tracking-SARS-CoV-2-variants (accessed May, 17 2024).

40.   Ito K, Piantham C, Nishiura H. Relative instantaneous reproduction number of Omicron SARS-CoV-2 variant with respect to the Delta variant in Denmark. Journal of medical virology. 2022;94(5):2265-8.

41.   Xu A, Hong B, Lou F, Wang S, Li W, Shafqat A, et al. Sub-lineages of the SARS-CoV-2 Omicron variants: Characteristics and prevention. MedComm. 2022;3(3):e172.

42.   Chung JY, Lee C-K, Park Y-N. Trust in social non-pharmaceutical interventions and travel intention during a pandemic. Journal of Vacation Marketing. 2021;27(4):437-48.

43.   Guidelines for non-pharmaceutical interventions to reduce the impact of COVID-19 in the EU/EEA and the UK. 24 September 2020. ECDC: Stockholm; 2020.

44.   Balmford B, Annan JD, Hargreaves JC, Altoè M, Bateman IJ. Cross-Country Comparisons of Covid-19: Policy, Politics and the Price of Life. Environmental & resource economics. 2020;76(4):525-51.

45.   Adjodah D, Dinakar K, Chinazzi M, Fraiberger SP, Pentland A, Bates S, et al. Association between COVID-19 outcomes and mask mandates, adherence, and attitudes. PloS one. 2021;16(6):e0252315.

46.   Our World in Data. COVID-19: Stringency Index. https://ourworldindata.org/explorers/coronavirus-data-explorer (accessed on 8 July 2024).

47.   Sharma M, Mindermann S, Rogers-Smith C, Leech G, Snodin B, Ahuja J, et al. Understanding the effectiveness of government interventions against the resurgence of COVID-19 in Europe. Nature communications. 2021;12(1):5820.

48.   Liu Y, Morgenstern C, Kelly J, Lowe R, Jit M. The impact of non-pharmaceutical interventions on SARS-CoV-2 transmission across 130 countries and territories. BMC medicine. 2021;19(1):40.

49.   Banholzer N, Feuerriegel S, Vach W. Estimating and explaining cross-country variation in the effectiveness of non-pharmaceutical interventions during COVID-19. Scientific reports. 2022;12(1):7526.

50.   Middelburg RA, Rosendaal FR. COVID-19: How to make between-country comparisons. International journal of infectious diseases : IJID : official publication of the International Society for Infectious Diseases. 2020;96:477-81.

51.   Verelst F, Hermans L, Vercruysse S, Gimma A, Coletti P, Backer JA, et al. SOCRATES-CoMix: a platform for timely and open-source contact mixing data during and in between COVID-19 surges and interventions in over 20 European countries. BMC medicine. 2021;19(1):254.

52.   Al-Hasan A, Yim D, Khuntia J. Citizens' Adherence to COVID-19 Mitigation Recommendations by the Government: A 3-Country Comparative Evaluation Using Web-Based Cross-Sectional Survey Data. Journal of medical Internet research. 2020;22(8):e20634.

53.   Seale H, Dyer CEF, Abdi I, Rahman KM, Sun Y, Qureshi MO, et al. Improving the impact of non-pharmaceutical interventions during COVID-19: examining the factors that influence engagement and the impact on individuals. BMC infectious diseases. 2020;20(1):607.

54.   Harper CA, Satchell LP, Fido D, Latzman RD. Functional Fear Predicts Public Health Compliance in the COVID-19 Pandemic. International journal of mental health and addiction. 2021;19(5):1875-88.

55.   Byambasuren O, Cardona M, Bell K, Clark J, McLaws M-L, Glasziou P. Estimating the extent of asymptomatic COVID-19 and its potential for community transmission: systematic review and meta-analysis. medRxiv. 2020:2020.05.10.20097543.

56.   Ahmed MAM, Siewe Fodjo JN, Gele AA, Farah AA, Osman S, Guled IA, et al. COVID-19 in Somalia: Adherence to Preventive Measures and Evolution of the Disease Burden. Pathogens (Basel, Switzerland). 2020;9(9):735.

57.   Acurio-Páez D, Vega B, Orellana D, Charry R, Gómez A, Obimpeh M, et al. Seroprevalence of SARS-CoV-2 Infection and Adherence to Preventive Measures in Cuenca, Ecuador, October 2020, a Cross-Sectional Study. International journal of environmental research and public health. 2021;18(9):4657.

58.   Smith LE, Mottershaw AL, Egan M, Waller J, Marteau TM, Rubin GJ. The impact of believing you have had COVID-19 on self-reported behaviour: Cross-sectional survey. PloS one. 2020;15(11):e0240399.

59.   On the State of Sanitary and Epidemiological Well-being of the Population in the Russian Federation in 2020: State Report / Edited by R.V. Buzinov – Moscow, 2021. – 256 p.

60.   Federal State Statistics Service, 2021. Available from: https://engrosstatgovru/. (accessed May, 17 2024).

61.   On the State of Sanitary and Epidemiological Well-being of the Population in the Arkhangelsk Region in 2020: State Report / Edited by R.V. Buzinov – Arkhangelsk, 2021. – 144 p.

62.   Methodological Recommendations MP 3.1.0178-20 "Definition of a Set of Measures, as Well as Indicators That Are the Basis for the Phased Removal of Restrictive Measures in the Conditions of the Epidemic Spread of COVID-19". Available from: https://www.garant.ru/products/ipo/prime/doc/73904890/ (accessed on November, 25 2023).

63.   Decree of the Chief State Sanitary Doctor of the Russian Federation dated 22.05.2020 № 15 "On Approval of Sanitary and Epidemiological Rules SP 3.1.3597-20" Prevention of New Coronavirus Infection (COVID-19)".

64.   Decree of the Chief State Sanitary Doctor of the Russian Federation dated 4.02.2022 № 4 "On Approval of Sanitary and Epidemiological Rules SP 3.1.3597-20 Prevention of New Coronavirus Infection (COVID-19)".

65.   Decree of the Russian President of March 25, 2020 No. 206 "On the establishment of non-working days in the territory of the Russian Federation".

66.   Decree of the Russian President of April 2, 2020 No. 239 "On measures for ensuring sanitary and epidemiologic wellbeing of the population in the territory of the Russian Federation in connection with spread of new coronavirus infection (COVID-19)".

67.   Decree of the Russian President of April 28, 2020 No. 294. "On the extension of measures to ensure the sanitary and and epidemiologic wellbeing of the population in the territory of the Russian Federation in connection with spread of new coronavirus infection (COVID-19)".

68.   Prevention, diagnosis and treatment of new coronavirus infection (COVID-19): temporal methodological recommendations of the Ministry of health of Russian Federation. Version 3-5. 2021.

69.   Decree of the Russian President of May 11, 2020. No. 316. "On definition of the procedure for prolongation of measures to ensure sanitary-epidemiological well-being of the population in the subjects of the Russian Federation in connection with the spread of new coronavirus infection (COVID-19)" Moscow.

70.   Governor's Decree № 28-u, March 17, 2020, "On the set of restrictive and other measures to counteract the spread of new coronavirus infection (COVID-19) in the Arkhangelsk Region". "On the set of restrictive and other measures to counteract the spread of new coronavirus infection (COVID-19) in the Arkhangelsk Region".

71.   Decree of the Arkhangelsk Governor March 17, 2020. No. 28-u. "On the set of restrictive and other measures to counteract the spread of new coronavirus infection (COVID-19) in the Arkhangelsk Region".

72.   On the State of Sanitary and Epidemiological Well-being of the Population in the Russian Federation in 2021: State Report / Edited by R.V. Buzinov – Moscow, 2022. – 340 p.

73.   On the State of Sanitary and Epidemiological Well-being of the Population in the Arkhangelsk Region in 2021: State Report / Edited by T.I. Nosovsky – Arkhangelsk, 2022. – 146 p.

74.   Coronavirus Monitor. https://coronavirus-monitor.info/country/russia (accessed on November 27, 2023).

75.   Lv H, Wu NC, Tsang OT, Yuan M, Perera R, Leung WS, et al. Cross-reactive Antibody Response between SARS-CoV-2 and SARS-CoV Infections. Cell reports. 2020;31(9):107725.

76.   World Health Organization. WHO COVID-19 Case definition. https://www.who.int/publications/i/item/WHO-2019-nCoV-Surveillance_Case_Definition-2022.1 (accessed November 27, 2023).

77.   World Health Organization (WHO). International Statistical Classification of Diseases and Related Health Problems 10th Revision (ICD-10 Version:2019). Geneva: WHO; 2019 Available from: https://icd.who.int/browse10/2019/en (accessed 27 November 2023).

78.   Decree of the Russian Government of March 31, 2020 No. 373 "About approval of Provisional rules of accounting of information for the purpose of prevention of spread of new koronavirusny infection (COVID-19)".

79.   Clark-Boucher D, Boss J, Salvatore M, Smith JA, Fritsche LG, Mukherjee B. Assessing the added value of linking electronic health records to improve the prediction of self-reported COVID-19 testing and diagnosis. PloS one. 2022;17(7):e0269017.

80.   COVID-19 Health System Response Monitor (HSRM). https://eurohealthobservatory.who.int/monitors/hsrm/hsrm-countries

81.   Nguyen NNT, McCarthy C, Lantigua D, Camci-Unal G. Development of Diagnostic Tests for Detection of SARS-CoV-2. Diagnostics (Basel, Switzerland). 2020;10(11).

82.   Al-Shaibari KSA, Mousa HA, Alqumber MAA, Alqfail KA, Mohammed A, Bzeizi K. The Diagnostic Performance of Various Clinical Specimens for the Detection of COVID-19: A Meta-Analysis of RT-PCR Studies. Diagnostics (Basel, Switzerland). 2023;13(19): 3057.

83.   Böger B, Fachi MM, Vilhena RO, Cobre AF, Tonin FS, Pontarolo R. Systematic review with meta-analysis of the accuracy of diagnostic tests for COVID-19. American journal of infection control. 2021;49(1):21-9.

84.   Kucirka LM, Lauer SA, Laeyendecker O, Boon D, Lessler J. Variation in False-Negative Rate of Reverse Transcriptase Polymerase Chain Reaction-Based SARS-CoV-2 Tests by Time Since Exposure. Annals of internal medicine. 2020;173(4):262-7.

85.   Kuster AC, Overgaard HJ. A novel comprehensive metric to assess effectiveness of COVID-19 testing: Inter-country comparison and association with geography, government, and policy response. PloS one. 2021;16(3):e0248176.

86.   Byambasuren O, Dobler CC, Bell K, Rojas DP, Clark J, McLaws ML, et al. Comparison of seroprevalence of SARS-CoV-2 infections with cumulative and imputed COVID-19 cases: Systematic review. PloS one. 2021;16(4):e0248946.

87.   Public health surveillance for COVID-19. Interim guidance. Available from: https://www.who.int/publications/i/item/WHO-2019-nCoV-SurveillanceGuidance-2022.2 / (accessed on May 31, 2023).

88.   Krajewski R, Gołębiowska J, Makuch S, Mazur G, Agrawal S. Update on serologic testing in COVID-19. Clinica chimica acta; international journal of clinical chemistry. 2020;510:746-50.

89. Sethuraman N, Jeremiah SS, Ryo A. Interpreting Diagnostic Tests for SARS-CoV-2. Jama. 2020;323(22):2249-51.

90. Anda EE, Braaten T, Borch KB, Nøst TH, Chen SLF, Lukic M, et al. Seroprevalence of antibodies against SARS-CoV-2 virus in the adult Norwegian population, winter 2020/2021: pre-vaccination period. Euro Surveillance. 2022 Mar;27(13):2100376.

91. Coronavirus (COVID-19) Infection Survey. Infection Survey, antibody data for the UK. Available from: https://wwwonsgovuk (accessed on February 26, 2024).

92. Gudbjartsson DF, Norddahl GL, Melsted P, Gunnarsdottir K, Holm H, Eythorsson E, et al. Humoral Immune Response to SARS-CoV-2 in Iceland. The New England journal of medicine. 2020;383(18):1724-34.

93. Popova AY, Smirnov VS, Andreeva EE, Babura EA, Balakhonov SV, Bashketova NS, et al. SARS-CoV-2 Seroprevalence Structure of the Russian Population during the COVID-19 Pandemic. Viruses. 2021;13(8): 1648.

94. Zurochka A, Dobrinina M, Zurochka V, Hu D, Solovyev A, Ryabova L, et al. Seroprevalence of SARS-CoV-2 Antibodies in Symptomatic Individuals Is Higher than in Persons Who Are at Increased Risk Exposure: The Results of the Single-Center, Prospective, Cross-Sectional Study. Vaccines. 2021;9(6): 627.

95. Oliveira MS, Lobo RD, Detta FP, Vieira-Junior JM, Castro TLS, Zambelli DB, et al. SARS-Cov-2 seroprevalence and risk factors among health care workers: Estimating the risk of COVID-19 dedicated units. American journal of infection control. 2021;49(9):1197-9.

96. Sannigrahi S, Pilla F, Basu B, Basu AS, Molter A. Examining the association between socio-demographic composition and COVID-19 fatalities in the European region using spatial regression approach. Sustainable cities and society. 2020;62:102418.

97. Faria de Moura Villela E, López RVM, Sato APS, de Oliveira FM, Waldman EA, Van den Bergh R, et al. COVID-19 outbreak in Brazil: adherence to national preventive measures and impact on people's lives, an online survey. BMC public health. 2021;21(1):152.

98. Urbán R, Paksi B, Miklósi Á, Saunders JB, Demetrovics Z. Non-adherence to preventive behaviours during the COVID-19 epidemic: findings from a community study. BMC public health. 2021;21(1):1462.

99. Barchuk A, Skougarevskiy D, Titaev K, Shirokov D, Raskina Y, Novkunkskaya A, et al. Seroprevalence of SARS-CoV-2 antibodies in Saint Petersburg, Russia: a population-based study. Scientific reports. 2021;11(1):12930.

100. Batista SR, Souza ASS, Nogueira J, Andrade FB, Thumé E, Teixeira D, et al. Protective behaviors for COVID-19 among Brazilian adults and elderly living with multimorbidity: the ELSI-COVID-19 initiative. Cadernos de saude publica. 2020;36Suppl 3(Suppl 3):e00196120.

101. He Y, Sun J, Ding X, Wang Q. Mechanisms in Which Smoking Increases the Risk of COVID-19 Infection: A Narrative Review. Iranian journal of public health. 2021;50(3):431-7.

102. Günther F, Einhauser S, Peterhoff D, Wiegrebe S, Niller HH, Beileke S, et al. Higher Infection Risk among Health Care Workers and Lower Risk among Smokers Persistent across SARS-CoV-2 Waves-Longitudinal Results from the Population-Based TiKoCo Seroprevalence Study. International journal of environmental research and public health. 2022;19(24): 16996.

103. Iruretagoyena M, Vial MR, Spencer-Sandino M, Gaete P, Peters A, Delgado I, et al. Longitudinal assessment of SARS-CoV-2 IgG seroconversionamong front-line healthcare workers during the first wave of the Covid-19 pandemic at a tertiary-care hospital in Chile. BMC infectious diseases. 2021;21(1):478.

104. Sah P, Fitzpatrick MC, Zimmer CF, Abdollahi E, Juden-Kelly L, Moghadas SM, et al. Asymptomatic SARS-CoV-2 infection: A systematic review and meta-analysis. Proceedings of the National Academy of Sciences of the United States of America. 2021;118(34):

e2109229118.

105.   Struyf T, Deeks JJ, Dinnes J, Takwoingi Y, Davenport C, Leeflang MM, et al. Signs and symptoms to determine if a patient presenting in primary care or hospital outpatient settings has COVID-19 disease. The Cochrane database of systematic reviews. 2020;7(7):Cd013665.

106.   Meo SA, Alhowikan AM, Al-Khlaiwi T, Meo IM, Halepoto DM, Iqbal M, et al. Novel coronavirus 2019-nCoV: prevalence, biological and clinical characteristics comparison with SARS-CoV and MERS-CoV. European review for medical and pharmacological sciences. 2020;24(4):2012-9.

107.   Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet (London, England). 2020;395(10223):497-506.

108.   Çelik I, Öztürk R. From asymptomatic to critical illness: decoding various clinical stages of COVID-19. Turkish journal of medical sciences. 2021;51(Si-1):3284-300.

109.   Pijls BG, Jolani S, Atherley A, Derckx RT, Dijkstra JIR, Franssen GHL, et al. Demographic risk factors for COVID-19 infection, severity, ICU admission and death: a meta-analysis of 59 studies. BMJ open. 2021;11(1):e044640.

110.   Mal P, Mukherjee T, Upadhyay AK, Mohanty S, Pattnaik AK. Connecting the dots between inflammatory cascades of obesity and COVID-19 in light of mortal consequences-a review. Environmental science and pollution research international. 2022;29(38):57040-53.

111.   Mubarik S, Liu X, Eshak ES, Liu K, Liu Q, Wang F, et al. The Association of Hypertension With the Severity of and Mortality From the COVID-19 in the Early Stage of the Epidemic in Wuhan, China: A Multicenter Retrospective Cohort Study. Frontiers in medicine. 2021;8:623608.

112.   von Stillfried S, Bülow RD, Röhrig R, Boor P. First report from the German COVID-19 autopsy registry. The Lancet regional health Europe. 2022;15:100330.

113.   Huertas A, Montani D, Savale L, Pichon J, Tu L, Parent F, et al. Endothelial cell dysfunction: a major player in SARS-CoV-2 infection (COVID-19)? The European respiratory journal. 2020;56(1): 2001634.

114.   De Michieli L, Ola O, Knott JD, Akula A, Mehta RA, Hodge DO, et al. High-Sensitivity Cardiac Troponin T for the Detection of Myocardial Injury and Risk Stratification in COVID-19. Clinical chemistry. 2021;67(8):1080-9.

115.   Chen C, Chen C, Yan JT, Zhou N, Zhao JP, Wang DW. [Analysis of myocardial injury in patients with COVID-19 and association between concomitant cardiovascular diseases and severity of COVID-19]. Zhonghua xin xue guan bing za zhi. 2020;48(7):567-71.

116.   Bustos-Vázquez E, Padilla-González E, Reyes-Gómez D, Carmona-Ramos MC, Monroy-Vargas JA, Benítez-Herrera AE, et al. Survival of COVID-19 with Multimorbidity Patients. Healthcare (Basel, Switzerland). 2021;9(11): 1423.

117.   Nada KM, Hsu ES, Seashore J, Zaidan M, Nishi SP, Duarte A, et al. Determining Cause of Death During Coronavirus Disease 2019 Pandemic. Critical care explorations. 2021;3(4):e0419.

118.   Zeng H, Ma Y, Zhou Z, Liu W, Huang P, Jiang M, et al. Spectrum and Clinical Characteristics of Symptomatic and Asymptomatic Coronavirus Disease 2019 (COVID-19) With and Without Pneumonia. Frontiers in medicine. 2021;8:645651.

119.   Ryu S, Chun JY, Lee S, Yoo D, Kim Y, Ali ST, et al. Epidemiology and Transmission Dynamics of Infectious Diseases and Control Measures. Viruses. 2022;14(11): 2510.

120.   He J, Guo Y, Mao R, Zhang J. Proportion of asymptomatic coronavirus disease 2019: A systematic review and meta-analysis. Journal of medical virology. 2021;93(2):820-30.

121.   [An update on the epidemiological characteristics of novel coronavirus pneumonia (COVID-19)]. Zhonghua liu xing bing xue za zhi = Zhonghua liuxingbingxue zazhi.

2020;41(2):139-44.

122.   Karlinsky A, Kobak D. Tracking excess mortality across countries during the COVID-19 pandemic with the World Mortality Dataset. eLife. 2021;10: e69336.

123.   Strongman H, Carreira H, De Stavola BL, Bhaskaran K, Leon DA. Factors associated with excess all-cause mortality in the first wave of the COVID-19 pandemic in the UK: A time series analysis using the Clinical Practice Research Datalink. PLoS medicine. 2022;19(1):e1003870.

124.   Lee WE, Park SW, Weinberger DM, Olson D, Simonsen L, Grenfell BT, et al. Direct and indirect mortality impacts of the COVID-19 pandemic in the US, March 2020-April 2021. Elife. 2023;12:e77562.

125.   Cuschieri S, Mamo J. Taking care of the ordinary in extraordinary times-delayed routine care means more morbidity and pre-mature mortality. European journal of public health. 2021;31(Supplement_4):iv27-iv30.

126.   Order of the Ministry of Health of the Russian Federation № 198n of March 19, 2020. "On the temporary organization of the activities of medical organizations for the implementation of measures to prevent and reduce the risks of spreading the new coronavirus infection COVID-19"

127.   Resolution of the Government of the Russian Federation No. 432 of April 3, 2020. "On the features of implementing the basic program of compulsory medical insurance in conditions of the emergence of threats of the spread of diseases caused by the new coronavirus infection".

128.   Timonin S, Klimkin I, Shkolnikov VM, Andreev E, McKee M, Leon DA. Excess mortality in Russia and its regions compared to high income countries: An analysis of monthly series of 2020. SSM - population health. 2022;17:101006.

129.   Scherbov S, Gietel-Basten S, Ediev D, Shulgin S, Sanderson W. COVID-19 and excess mortality in Russia: Regional estimates of life expectancy losses in 2020 and excess deaths in 2021. PloS one. 2022;17(11):e0275967.

130.   Russian Federal State Statistics Service. ROSSTAT. Demography. Available from: https://rosstat.gov.ru/folder/12781. (accessed November, 27 2023).

131.   World Health Organization. International Guidelines for Certification and Classification (Coding) of COVID-19 as Cause of Death. https://www.who.int/publications/m/item/international-guidelines-for-certification-and-classification-(coding)-of-covid-19-as-cause-of-death (accessed November, 27 2023).

132.   Meilia PDI, Manela C, Yudy, Sawitri R, Syukriani YF, Fitrasanti BI, et al. Characteristics of deceased and quality of death certificates for cases subjected to Indonesia's management of the dead protocol for bodies with COVID-19. Forensic science, medicine, and pathology. 2022;18(1):45-56.

133.   The Ministry of Health of Russia. Guidelines for Coding and Selection of the Underlying Condition in Morbidity Statistics and the Initial Cause in Mortality Statistics Associated with COVID-19 (in Russian). The Ministry of Health of Russia; 2020. https://static-0.minzdrav.gov.ru/system/attachments/attaches/000/050/527/original/27052020_MR_STAT_1.pdf (accessed November, 27 2023).

134.   Nab L, Parker EPK, Andrews CD, Hulme WJ, Fisher L, Morley J, et al. Changes in COVID-19-related mortality across key demographic and clinical subgroups in England from 2020 to 2022: a retrospective cohort study using the OpenSAFELY platform. The Lancet Public health. 2023;8(5):e364-e77.

135.   Nogueira PJ, de Araújo Nobre M, Elias C, Feteira-Santos R, Martinho AC, Camarinha C, et al. Multimorbidity Profile of COVID-19 Deaths in Portugal during 2020. Journal of clinical medicine. 2022;11(7): 1898.

136. Harrison SL, Fazio-Eynullayeva E, Lane DA, Underhill P, Lip GYH. Comorbidities associated with mortality in 31,461 adults with COVID-19 in the United States: A federated electronic medical record analysis. PLoS medicine. 2020;17(9):e1003321.

137. World Health Organization. Cause of death. https://www.who.int/standards/classifications/classification-of-diseases/cause-of-death (accessed November, 27 2023).

138. Kim NY, Kim SS, Lee HJ, Kim DH, Ryu B, Shin E, et al. Risk factors for deaths associated with COVID-19 according to the cause of death classification in Republic of Korea. Osong public health and research perspectives. 2023;14(2):89-99.

139. Grippo F, Navarra S, Orsi C, Manno V, Grande E, Crialesi R, et al. The Role of COVID-19 in the Death of SARS-CoV-2-Positive Patients: A Study Based on Death Certificates. Journal of clinical medicine. 2020;9(11): 3459.

140. Singh JA, Upshur REG. The granting of emergency use designation to COVID-19 candidate vaccines: implications for COVID-19 vaccine trials. The Lancet Infectious diseases. 2021;21(4):103-9.

141. Shrotri M, Swinnen T, Kampmann B, Parker EPK. An interactive website tracking COVID-19 vaccine development. The Lancet Global health. 2021;9(5):590-2.

142. Logunov DY, Dolzhikova IV, Shcheblyakov DV, Tukhvatulin AI, Zubkova OV, Dzharullaeva AS, et al. Safety and efficacy of an rAd26 and rAd5 vector-based heterologous prime-boost COVID-19 vaccine: an interim analysis of a randomised controlled phase 3 trial in Russia. Lancet (London, England). 2021;397(10275):671-81.

143. Balakrishnan VS. The arrival of Sputnik V. The Lancet Infectious diseases. 2020;20(10):1128.

144. Barchuk A, Bulina A, Cherkashin M, Berezina N, Rakova T, Kuplevatskaya D, et al. Gam-COVID-Vac, EpiVacCorona, and CoviVac effectiveness against lung injury during Delta and Omicron variant surges in St. Petersburg, Russia: a test-negative case-control study. Respiratory research. 2022;23(1):276.

145. Kozlovskaya LI, Piniaeva AN, Ignatyev GM, Gordeychuk IV, Volok VP, Rogova YV, et al. Long-term humoral immunogenicity, safety and protective efficacy of inactivated vaccine against COVID-19 (CoviVac) in preclinical studies. Emerging microbes & infections. 2021;10(1):1790-806.

146. Kashte S, Gulbake A, El-Amin Iii SF, Gupta A. COVID-19 vaccines: rapid development, implications, challenges and future prospects. Human cell. 2021;34(3):711-33.

147. Wong JC, Lao CT, Yousif MM, Luga JM. Fast Tracking—Vaccine Safety, Efficacy, and Lessons Learned: A Narrative Review. Vaccines. 2022;10(8):1256.

148. Statista. Number of people vaccinated and fully vaccinated against COVID-19 per 100 population in Russia. https://www.statista.com/statistics/1239299/covid-19-vaccination-rate-in-russia/ (accessed on June 7, 2024).

149. Barchuk A, Bulina A, Cherkashin M, Berezina N, Rakova T, Kuplevatskaya D, et al. COVID-19 vaccines effectiveness against symptomatic SARS-CoV-2 during Delta variant surge: a preliminary assessment from a case-control study in St. Petersburg, Russia. BMC public health. 2022;22(1):1803.

150. Shao W, Chen X, Zheng C, Liu H, Wang G, Zhang B, et al. Effectiveness of COVID-19 vaccines against SARS-CoV-2 variants of concern in real-world: a literature review and meta-analysis. Emerging microbes & infections. 2022;11(1):2383-92.

151. Eyre DW, Taylor D, Purver M, Chapman D, Fowler T, Pouwels KB, et al. Effect of Covid-19 Vaccination on Transmission of Alpha and Delta Variants. The New England journal of medicine. 2022;386(8):744-56.

152. Fundora MP, Kamidani S, Oster ME. COVID Vaccination as a Strategy for Cardiovascular Disease Prevention. Current cardiology reports. 2023;25(10):1327-35.

153.   Cook S, Malyutina S, Kudryavtsev AV, Averina M, Bobrova N, Boytsov S, et al. Know Your Heart: Rationale, design and conduct of a cross-sectional study of cardiovascular structure, function and risk factors in 4500 men and women aged 35-69 years from two Russian cities, 2015-18. Wellcome open research. 2018;3:67.

154.   Kontsevaya A.V. SSA, Drapkina O.M. . ESSE-RF study: epidemiology and public health promotion. Cardiovascular Therapy and Prevention. 2021;20(5):2987.

155.   Kuvshinova IN, Livitskaya NI, Molodykh SV, et al. Sensitivity and specificity of reagent kits of JSC "Vector-Best" for the detection of immunoglobulins of different classes to SARSCoV-2. Spravochnik Zaveduyushchego KDL. 2021;10:27-32.

156.   Barchuk A, Shirokov D, Sergeeva M, Tursun Zade R, Dudkina O, Tychkova V, et al. Evaluation of the performance of SARS--CoV--2 antibody assays for a longitudinal population-based study of COVID--19 spread in St. Petersburg, Russia. Journal of medical virology. 2021;93(10):5846-52.

157.   Krieger E.A., Samodova O.V., Svitich O.A., Samoilikov R.V., Meremianina E.A., Ivanova L.V., Bebyakova N.A., Ilina E.N., Pavlenko A.V., Esin Y.I., Arkhipova A.L., Kovalchuk S.N., Kudryavtsev A.V. The impact of polymorphic variants of interferon receptor genes on COVID-19 severity and antibiotic resistance // Russian Journal of Infection and Immunity. - 2023. - Vol. 13. - N. 6. - P. 1027-1039.

158.   Krieger E, Sharashova E, Kudryavtsev AV, Samodova O, Kontsevaya A, Brenn T, et al. COVID-19: seroprevalence and adherence to preventive measures in Arkhangelsk, Northwest Russia. Infectious diseases (London, England). 2023;55(5):316-27.

159.   Sempos CT, Tian L. Adjusting Coronavirus Prevalence Estimates for Laboratory Test Kit Error. American journal of epidemiology. 2021;190(1):109-15.

160.   Henrion MY. bootComb—an R package to derive confidence intervals for combinations of independent parameter estimates. International Journal of Epidemiology. 2021;50(4):1071-6.

161.   World Health Organization; Babor, T. F., Higgins-Biddle, J. C., Saunders, J. B., Monteiro, M. G. AUDIT: The Alcohol Use Disorders Identification Test: Guidelines for Use in Primary Health Care, 2nd edn (World Health Organization, Geneva, Switzerland, 2001).

162.   WHO. Alcohol, heavy episodic drinking (15+) past 30 days (%), age-standardized with 95%CI. https://wwwwhoint/data/gho/data/indicators/indicator-details/GHO/alcohol-heavy-episodic-drinking-(15-)-past-30-days-(-)-age-standardized-with-95-ci. 2021.

163.   Kesmodel US. Cross-sectional studies - what are they good for? Acta obstetricia et gynecologica Scandinavica. 2018;97(4):388-93.

164.   Huitfeldt A. Is caviar a risk factor for being a millionaire? BMJ (Clinical research ed). 2016;355:i6536.

165.   Szklo M, Nieto FJ. Epidemiology: Beyond the Basics. 4th ed. Sudbury: Jones & Bartlett Learning, LLC; 2018.

166.   Hendra R, Hill A. Rethinking Response Rates: New Evidence of Little Relationship Between Survey Response Rates and Nonresponse Bias. Evaluation review. 2019;43(5):307-30.

167.   Meterko M, Restuccia JD, Stolzmann K, Mohr D, Brennan C, Glasgow J, et al. Response Rates, Nonresponse Bias, and Data Quality: Results from a National Survey of Senior Healthcare Leaders. Public Opinion Quarterly. 2015;79(1):130-44.

168.   Nunan D, Aronson J, Bankhead C. Catalogue of bias: attrition bias. BMJ evidence-based medicine. 2018;23(1):21-2.

169.   Binkheder S, Asiri MA, Altowayan KW, Alshehri TM, Alzarie MF, Aldekhyyel RN, et al. Real-World Evidence of COVID-19 Patients' Data Quality in the Electronic Health Records. Healthcare (Basel, Switzerland). 2021;9(12): 1648.

170.   Krantz SG, Rao A. Level of underreporting including underdiagnosis before the first

peak of COVID-19 in various countries: Preliminary retrospective results based on wavelets and deterministic modeling. Infection control and hospital epidemiology. 2020;41(7):857-9.

171.   Ricoca Peixoto V, Nunes C, Abrantes A. Epidemic Surveillance of Covid-19: Considering Uncertainty and Under-Ascertainment. Portuguese Journal of Public Health. 2020;38(1):23-9.

172.   Padidar S, Liao SM, Magagula S, Mahlaba TAM, Nhlabatsi NM, Lukas S. Assessment of early COVID-19 compliance to and challenges with public health and social prevention measures in the Kingdom of Eswatini, using an online survey. PloS one. 2021;16(6):e0253954.

173.   Latkin CA, Edwards C, Davey-Rothwell MA, Tobin KE. The relationship between social desirability bias and self-reports of health, substance use, and social network factors among urban substance users in Baltimore, Maryland. Addictive behaviors. 2017;73:133-6.

174.   Najafi F, Moradinazar M, Hamzeh B, Rezaeian S. The reliability of self-reporting chronic diseases: how reliable is the result of population-based cohort studies. Journal of preventive medicine and hygiene. 2019;60(4): 349-53.

175.   Tavakol M, Dennick R. Making sense of Cronbach's alpha. International journal of medical education. 2011;2:53-5.

176.   Crane MA, Shermock KM, Omer SB, Romley JA. Change in Reported Adherence to Nonpharmaceutical Interventions During the COVID-19 Pandemic, April-November 2020. Jama. 2021;325(9):883-5.

177.   Jiang C, Yao X, Zhao Y, Wu J, Huang P, Pan C, et al. Comparative review of respiratory diseases caused by coronaviruses and influenza A viruses during epidemic season. Microbes and infection. 2020;22(6-7):236-44.

178.   Meyer MJ, Yan S, Schlageter S, Kraemer JD, Rosenberg ES, Stoto MA. Adjusting COVID-19 seroprevalence survey results to account for test sensitivity and specificity. The American Journal of Epidemiology. 2022;191(4):681-688.

179.   Chen X, Chen Z, Azman AS, Deng X, Sun R, Zhao Z, et al. Serological evidence of human infection with SARS-CoV-2: a systematic review and meta-analysis. The Lancet Global health. 2021;9(5):598-609.

180.   Lai CC, Wang JH, Hsueh PR. Population-based seroprevalence surveys of anti-SARS-CoV-2 antibody: An up-to-date review. International journal of infectious diseases : IJID : official publication of the International Society for Infectious Diseases. 2020;101:314-22.

181.   Bursac Z, Gauss CH, Williams DK, Hosmer DW. Purposeful selection of variables in logistic regression. Source code for biology and medicine. 2008;3:17.

182.   Textor J, Liśkiewicz M. Adjustment criteria in causal diagrams: an algorithmic perspective. Proceedings of the Twenty-Seventh Conference on Uncertainty in Artificial Intelligence; Barcelona, Spain: AUAI Press; 2011. p. 681–8.

183.   Lee PH. Is a cutoff of 10% appropriate for the change-in-estimate criterion of confounder identification? Journal of epidemiology. 2014;24(2):161-7.

184.   Heinze G, Wallisch C, Dunkler D. Variable selection - A review and recommendations for the practicing statistician. Biometrical journal Biometrische Zeitschrift. 2018;60(3):431-49.

185.   Westreich D, Greenland S. The table 2 fallacy: presenting and interpreting confounder and modifier coefficients. American journal of epidemiology. 2013;177(4):292-8.

186.   Dekkers OM, Laugesen K, Groenwold RHH. Directed acyclic graphs in clinical research. European journal of endocrinology. 2024;190(4):5-7.

187.   Anderson RM, Vegvari C, Truscott J, Collyer BS. Challenges in creating herd immunity to SARS-CoV-2 infection by mass vaccination. Lancet (London, England). 2020;396(10263):1614-6.

188.   Shubin M, Virtanen M, Toikkanen S, Lyytikäinen O, Auranen K. Estimating the

burden of A(H1N1)pdm09 influenza in Finland during two seasons. Epidemiology and infection. 2014;142(5):964-74.

189.   Vahey GM, Marshall KE, McDonald E, Martin SW, Tate JE, Midgley CM, et al. Symptom Profiles and Progression in Hospitalized and Nonhospitalized Patients with Coronavirus Disease, Colorado, USA, 2020. Emerg Infect Dis. 2021;27(2):385-95.

190.   Schäfer E, Scheer C, Saljé K, Fritz A, Kohlmann T, Hübner NO, et al. Course of disease and risk factors for hospitalization in outpatients with a SARS-CoV-2 infection. Scientific reports. 2022;12(1):7249.

191.   Kim GU, Kim MJ, Ra SH, Lee J, Bae S, Jung J, et al. Clinical characteristics of asymptomatic and symptomatic patients with mild COVID-19. Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases. 2020;26(7):948.

192.   Akimkin V.G., Kuzin S.N., Kolosovskaya E.N., Kudryavtceva E.N., Semenenko T.A., Ploskireva A.A., et al. Assessment of the COVID-19 epidemiological situation in St. Petersburg // Journal of microbiology, epidemiology and immunobiology. 2021; 98 (5): 497-511.

193.   Schultze A NE, Evans D, Hulme W, Rosello A, Bates C, et al. . Mortality among Care Home Residents in England during the first and second waves of the COVID-19 pandemic: an analysis of 4.3 million adults over the age of 65. Lancet Regional Health - Europe. 2021.

194.   Nivette A, Ribeaud D, Murray A, Steinhoff A, Bechtiger L, Hepp U, et al. Non-compliance with COVID-19-related public health measures among young adults in Switzerland: Insights from a longitudinal cohort study. Social science & medicine (1982). 2021;268:113370.

195.   Ding X, Brazel DM, Mills MC. Factors affecting adherence to non-pharmaceutical interventions for COVID-19 infections in the first year of the pandemic in the UK. BMJ open. 2021;11(10):e054200.

196.   Li Y, Shi J, Xia J, Duan J, Chen L, Yu X, et al. Asymptomatic and Symptomatic Patients With Non-severe Coronavirus Disease (COVID-19) Have Similar Clinical Features and Virological Courses: A Retrospective Single Center Study. Frontiers in microbiology. 2020;11:1570.

197.   Singhal S, Kumar P, Singh S, Saha S, Dey AB. Clinical features and outcomes of COVID-19 in older adults: a systematic review and meta-analysis. BMC geriatrics. 2021;21(1):321.

198.   Abeya SG, Barkesa SB, Sadi CG, Gemeda DD, Muleta FY, Tolera AF, et al. Adherence to COVID-19 preventive measures and associated factors in Oromia regional state of Ethiopia. PloS one. 2021;16(10):e0257373.

199.   Shkolnikov VM, Churilova E, Jdanov DA, Shalnova SA, Nilssen O, Kudryavtsev A, et al. Time trends in smoking in Russia in the light of recent tobacco control measures: synthesis of evidence from multiple sources. BMC public health. 2020;20(1):378.

200.   Simons D, Shahab L, Brown J, Perski O. The association of smoking status with SARS-CoV-2 infection, hospitalization and mortality from COVID-19: a living rapid evidence review with Bayesian meta-analyses (version 7). Addiction (Abingdon, England). 2021;116(6):1319-68.

201.   Farsalinos K, Barbouni A, Poulas K, Polosa R, Caponnetto P, Niaura R. Current smoking, former smoking, and adverse outcome among hospitalized COVID-19 patients: a systematic review and meta-analysis. Therapeutic advances in chronic disease. 2020;11:2040622320935765.

202.   Hippisley-Cox J, Young D, Coupland C, Channon KM, Tan PS, Harrison DA, et al. Risk of severe COVID-19 disease with ACE inhibitors and angiotensin receptor blockers:

cohort study including 8.3 million people. Heart (British Cardiac Society). 2020;106(19):1503-11.

203.   Lee SC, Son KJ, Kim DW, Han CH, Choi YJ, Kim SW, et al. Smoking and the Risk of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Infection. Nicotine & tobacco research : official journal of the Society for Research on Nicotine and Tobacco. 2021;23(10):1787-92.

204.   Changeux JP, Amoura Z, Rey FA, Miyara M. A nicotinic hypothesis for Covid-19 with preventive and therapeutic implications. Comptes rendus biologies. 2020;343(1):33-9.

205.   Tizabi Y, Getachew B, Copeland RL, Aschner M. Nicotine and the nicotinic cholinergic system in COVID-19. The FEBS journal. 2020;287(17):3656-63.

206.   Farsalinos K, Barbouni A, Niaura R. Systematic review of the prevalence of current smoking among hospitalized COVID-19 patients in China: could nicotine be a therapeutic option? Internal and emergency medicine. 2020;15(5):845-52.

207.   Qiu F, Liang CL, Liu H, Zeng YQ, Hou S, Huang S, et al. Impacts of cigarette smoking on immune responsiveness: Up and down or upside down? Oncotarget. 2017;8(1):268-84.

208.   Li C, Sun J. The impact of current smoking, regular drinking, and physical inactivity on health care-seeking behavior in China. BMC health services research. 2022;22(1):52.

209.   Salerno M, Sessa F, Piscopo A, Montana A, Torrisi M, Patanè F, et al. No Autopsies on COVID-19 Deaths: A Missed Opportunity and the Lockdown of Science. Journal of clinical medicine. 2020;9(5).

210.   Schäfer AA, Santos LP, Quadra MR, Dumith SC, Meller FO. Alcohol Consumption and Smoking During Covid-19 Pandemic: Association with Sociodemographic, Behavioral, and Mental Health Characteristics. Journal of community health. 2022;47(4):588-97.

211.   Jackson SE, Tattan-Birch H, Shahab L, Beard E, Brown J. Have there been sustained impacts of the COVID-19 pandemic on trends in smoking prevalence, uptake, quitting, use of treatment, and relapse? A monthly population study in England, 2017-2022. BMC medicine. 2023;21(1):474.

212.   Gallus S, Scala M, Possenti I, Jarach CM, Clancy L, Fernandez E, et al. The role of smoking in COVID-19 progression: a comprehensive meta-analysis. European respiratory review : an official journal of the European Respiratory Society. 2023;32(167): 220191.

213.   Zhang H, Ma S, Han T, Qu G, Cheng C, Uy JP, et al. Association of smoking history with severe and critical outcomes in COVID-19 patients: A systemic review and meta-analysis. European journal of integrative medicine. 2021;43:101313.

214.   Gülsen A, Yigitbas BA, Uslu B, Drömann D, Kilinc O. The Effect of Smoking on COVID-19 Symptom Severity: Systematic Review and Meta-Analysis. Pulmonary medicine. 2020;2020:7590207.

215.   Reddy RK, Charles WN, Sklavounos A, Dutt A, Seed PT, Khajuria A. The effect of smoking on COVID-19 severity: A systematic review and meta-analysis. Journal of medical virology. 2021;93(2):1045-56.

216.   Peckham H, de Gruijter NM, Raine C, Radziszewska A, Ciurtin C, Wedderburn LR, et al. Male sex identified by global COVID-19 meta-analysis as a risk factor for death and ITU admission. Nature communications. 2020;11(1):6317.

217.   Nielsen J, Nørgaard SK, Lanzieri G, Vestergaard LS, Moelbak K. Sex-differences in COVID-19 associated excess mortality is not exceptional for the COVID-19 pandemic. Scientific reports. 2021;11(1):20815.

218.   Geldsetzer P, Mukama T, Jawad NK, Riffe T, Rogers A, Sudharsanan N. Sex differences in the mortality rate for coronavirus disease 2019 compared to other causes of death: an analysis of population-wide data from 63 countries. European journal of epidemiology. 2022;37(8):797-806.

219. Odone A, Delmonte D, Gaetti G, Signorelli C. Doubled mortality rate during the COVID-19 pandemic in Italy: quantifying what is not captured by surveillance. Public health. 2021;190:108-15.

220. Carey IM, Cook DG, Harris T, DeWilde S, Chaudhry UAR, Strachan DP. Risk factors for excess all-cause mortality during the first wave of the COVID-19 pandemic in England: A retrospective cohort study of primary care data. PloS one. 2021;16(12):e0260381.

221. Figueroa J, Brennan P, Theodoratou E, Poon M, Purshouse K, Din F, et al. Trends in excess cancer and cardiovascular deaths in Scotland during the COVID-19 pandemic 30 December – 20 April suggest underestimation of COVID-19 related deaths. medRxiv; 2020.

222. Sudarikov A. COVID-19 and Cancer Detection in Russia. Cancers. 2024;16(9): 1673.

223. Chen S, Prettner K, Kuhn M, Geldsetzer P, Wang C, Bärnighausen T, et al. Climate and the spread of COVID-19. Scientific reports. 2021;11(1):9042.

224. Zhang F, Xiong Y, Wei Y, Hu Y, Wang F, Li G, et al. Obesity predisposes to the risk of higher mortality in young COVID-19 patients. Journal of medical virology. 2020;92(11):2536-42.

225. Elliott J, Bodinier B, Whitaker M, Delpierre C, Vermeulen R, Tzoulaki I, et al. COVID-19 mortality in the UK Biobank cohort: revisiting and evaluating risk factors. European journal of epidemiology. 2021;36(3):299-309.

226. Halpin DMG, Faner R, Sibila O, Badia JR, Agusti A. Do chronic respiratory diseases or their treatment affect the risk of SARS-CoV-2 infection? The Lancet Respiratory medicine. 2020;8(5):436-8.

227. Hippisley-Cox J, Coupland CA, Mehta N, Keogh RH, Diaz-Ordaz K, Khunti K, et al. Risk prediction of covid-19 related death and hospital admission in adults after covid-19 vaccination: national prospective cohort study. BMJ (Clinical research ed). 2021;374:n2244.

Research letter

Krieger E., Kudryavtsev A., Sharashova E., Postoev V., Belova N., Shagrov L., Zvedina J., Drapkina O., Kontsevaya A., Shalnova S., Brenn T., Shkolnikov V., Eggo R., Leon D. (2022)

Seroprevalence of SARS-Cov-2 Antibodies in Adults, Arkhangelsk, Russia.

Seroprevalence of SARS-Cov-2 Antibodies in Adults, Arkhangelsk, Russia

Ekaterina Krieger, Alexander Kudryavtsev, Ekaterina Sharashova, Vitaly Postoev, Natalia Belova, Leonid Shagrov, Julia Zvedina, Oxana Drapkina, Anna Kontsevaya, Svetlana Shalnova, Tormod Brenn, Vladimir M. Shkolnikov, Rosalind M. Eggo, David A. Leon

Author affiliations: UiT The Arctic University of Norway, Tromso, Norway (E. Krieger, A. Kudryavtsev, E. Sharashova, T. Brenn, D.A. Leon); Northern State Medical University of the Ministry of Health of the Russian Federation, Arkhangelsk, Russia (E. Krieger, A. Kudryavtsev, V. Postoev, N. Belova, L. Shagrov, J. Zvedina); National Medical Research Centre for Therapy and Preventive Medicine, Moscow, Russia (O. Drapkina, A. Kontsev, S. Shalnova); Max-Planck-Institute for Demographic Research, Mecklenburg, Germany (V.M. Shkolnikov); National Research University Higher School of Economics, Moscow (V.M. Shkolnikov, D.A. Leon); London School of Hygiene & Tropical Medicine, London, UK (R.M. Eggo, D.A. Leon)

Population-based data on coronavirus disease in Russia and on the immunogenicity of the Sputnik V vaccine are sparse. In a survey of 1,080 residents of Arkhangelsk 40–75 years of age, 65% were seropositive for IgG. Fifteen percent of participants had been vaccinated; of those, 97% were seropositive.

Russia is one of the few countries to have produced a coronavirus (COVID-19) vaccine (*1*). It has also experienced substantial excess deaths during the pandemic (*2*). Few published estimates of antibody seroprevalence for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in Russia exist. A St. Petersburg survey in June 2020 used random-digit dialing to contact 66,250 residents; of those, 1,038 provided a blood sample, and the samples had 9%–10% seropositivity (*3*). A study conducted in Chelyabinsk (September 28–December 30, 2020) recruited 1,091 high-risk workers (healthcare workers, education staff, and supermarket employees) ≥18 years of age. Of the 882 screened, 25% were seropositive for IgG (*4*). We are not aware of any seroprevalence estimates from Russia based on samples collected in 2021 that have appeared in the scientific literature.

We interviewed and obtained blood samples from 1,080 adults 40–75 years of age who were residents of the city of Arkhangelsk in northwest Russia during February 24–May 28, 2021. We obtained participants for this study from 2,258 invitations sent to persons who had taken part in the Know Your Heart study (*5*) (2015–2018), which was based on a random sample of the city population (Appendix, https://wwwnc.cdc.gov/EID/article/28/2/21-1640-App1.pdf). The ethics committee of the Northern State Medical University approved our study proposal and protocol on February 17, 2021.

We used a Vector Best ELISA assay (D-5501 SARS-CoV-2-IgG-EIA-BEST; https://vector-best.ru) to analyze qualitatively detected IgG directed against SARS-CoV-2 in human blood serum samples. Data are limited on the performance of this immunoassay, in particular, on its sensitivity for infections that occurred >3 weeks previously. According to the manufacturer, the assay has a sensitivity of 72% when performed 6–12 days after infection and ≈100% at 13–20 days (*6*). An independent assessment of the Vector Best ELISA assay found a sensitivity of 89% and a specificity of 100%, derived from comparisons of test results in prepandemic samples (negative controls) and PCR positive samples for SARS-CoV-2 (*7*). We estimated seroprevalence adjusted for test performance (89% sensitivity, 100% specificity) using the equation (crude prevalence + test specificity − 1)/(sensitivity + specificity − 1) (*8*). We calculated 95% CIs for the adjusted estimates of seroprevalence using the R package bootComb (https://www.r-project.org).

Of the 1,080 samples (634 women, mean age 55 years), we excluded 13 who had an equivocal test result from analysis. Of the 1,067 remaining samples, 690 (65%) were seropositive for IgG (Table 1). Seroprevalence adjusted for test characteristics was 72.6% (95% CI 64.2%–83.1%).

Seroprevalence did not substantively differ by sex or by educational level. Of the 162 participants (15%) who reported having been vaccinated. 150 (93%) were seropositive. Among the 31 who received 1 dose, 20 (65%) were seropositive; of the 131 who had received 2 doses, 130 (99%) were seropositive. Of the 905 participants who said they had not been vaccinated, 256 said that they had previously been ill with COVID-19; of those, 248 (97%) were seropositive. Of those who stated they had not been vaccinated and did not report having previously been ill with COVID-19, 292 (45%) were seropositive, suggesting an appreciable level of unrecognized infection. Our overall estimates of seroprevalence (crude 65%, adjusted 72.6%) is appreciably higher than found in St Petersburg in

Table. Seroprevalence of severe acute respiratory syndrome coronavirus 2 in adults, Arkhangelsk, Russia

Characteristic	Unvaccinated		Vaccinated*		Total	
	No. seropositive/ total (%)	Adjusted seroprevalence, % (95% CI)†	No. seropositive/ total (%)	Adjusted seroprevalence, % (95% CI)†	No. seropositive/ total (%)	Adjusted seroprevalence, % (95% CI)†
Sex						
F	332/553 (60)	67.4 (58.4–77.9)	72/81 (89)	99.7 (87.1–99.9)	404/634 (64)	71.5 (62.6–82.3)
M	208/352 (59)	66.3 (56.5–77.3)	78/81 (96)	100 (93.2–100)	286/433 (66)	74.1 (64.5–85.6)
Age, y						
40–54	291/461 (63)	70.8 (61.4;81.8)	35/38 (92)	100 (84.8–100)	326/499 (65)	73.3 (64.0–84.6)
55–64	181/317 (57)	64.1 (54.1–75.0)	38/43 (88)	99.1 (82.6–100)	219/360 (61)	68.3 (58.4–79.4)
>65	68/127 (54)	60.1 (46.9–73.1)	77/81 (95)	100 (92.4–100)	145/208 (70)	78.2 (67.0–91.2)
Education						
Secondary and lower	26/47 (55)	62.1 (42.7–81.0)	9/9 (100)	100 (66.7–100)	35/56 (63)	70.1 (52.5–88.1)
Specialized secondary	253/433 (58)	65.6 (56.1–76.0)	81/87 (93)	100 (91.2–100)	334/520 (64)	72.1 (62.9–83.2)
Higher	261/425 (61)	68.9 (59.3–79.8)	60/66 (91)	100 (88.0–100)	321/491 (65)	73.3 (64.0–84.6)
Week of test						
7–14	395/651 (61)	68.1 (59.3–78,4)	49/58 (84)	94.8 (81.0–100)	444/709 (63)	70.3 (61.6–80.8)
15–21	145/254 (57)	64.0 (53.4–75.3)	101/104 (97)	100 (94.8–100)	246/358 (69)	77.1 (67.1–89.1)
Self-reported prior symptoms of infection						
No	172/477 (36)	40.5 (31.7–47.8)	133/143 (93)	100 (92.9–100)	305/620 (49)	55.2 (46.6–64.0)
Yes	248/256 (97)	100 (96.9–100)	8/9 (89)	99.7 (56.8–100)	256/265 (97)	100 (96.7–100)
Do not know	120/172 (70)	78.3 (66.5–91.6)	9/10 (90)	100 (60.4–100)	129/182 (71)	79.5 (68.1–92.8)
Total	540/905 (60)	66.9 (58.6–76.9)	150/162 (93)	100 (92.9–100)	690/1067 (65)	72.6 (64.2–83.1)

*Received >1 dose.
†Values >100% were rounded to 100%.
‡Weeks 7–14 are February 24–April 11 and weeks 15–21 are April 12–May 28, 2021.

June 2020 (*3*) (10%) or in Chelyabinsk (25%) in September–December 2020 (*4*). This result is consistent with the second wave of the pandemic in Russia (peak November–December 2020) being larger than the first (peak May–June 2020); our study started during the vaccination period.

Deployment of COVID-19 vaccine, mostly Sputnik V, in the Arkhangelsk region started in mid-January 2021; 11% of the population received >1 dose by May 30, 2021 (*9*). Our study covered an urban sample from the city of Arkhangelsk, the capital of the region. Our estimate of 15% coverage of the study population may be higher because the regional estimates included data from more dispersed communities in. Nevertheless, our vaccination rates were low compared with rates in most European Union and European Economic Area countries as reported in June 2021 by the European Centre for Disease Prevention and Control (*10*). Given the vaccination rate in the sample was 15% but the antibodies were present in 65% of participants, we suspect that most of the seropositive results were the result of acquired infection.

Russia is geographically the largest country in the world; its regions vary considerably in terms of socioeconomic level, climate, and healthcare provision. Our study results are restricted to an adult population and cannot be generalized to the total population of Arkhangelsk region or to Russia. The high levels of seroprevalence among vaccinated participants confirms the immunogenicity of the Sputnik vaccine and suggests that it can protect the population if the proportion vaccinated is increased substantially. We recommend further population-based seroprevalence studies, using World Health Organization–approved tests, for public health efforts in the COVID-19 pandemic.

About the Author

Dr. Krieger is a PhD student in UiT The Arctic University of Norway, Tromso, Norway. Her primary research interest is epidemiology, infectious diseases, and vaccination.

References

1. Logunov DY, Dolzhikova IV, Zubkova OV, Tukhvatulin AI, Shcheblyakov DV, Dzharullaeva AS, et al. Safety and immunogenicity of an rAd26 and rAd5 vector-based heterologous prime-boost COVID-19 vaccine in two formulations: two open, non-randomised phase 1/2 studies from Russia. Lancet. 2020;396:887–97. https://doi.org/10.1016/S0140-6736(20)31866-3

2. Karlinsky A, Kobak D. Tracking excess mortality across countries during the COVID-19 pandemic with the World Mortality Dataset. eLife. 2021;10:10. https://doi.org/10.7554/eLife.69336

3. Barchuk A, Skougarevskiy D, Titaev K, Shirokov D, Raskina Y, Novkunkskaya A, et al. Seroprevalence of SARS-CoV-2 antibodies in Saint Petersburg, Russia: a population-based study. Sci Rep. 2021;11:12930. https://doi.org/10.1038/s41598-021-92206-y

4. Zurochka A, Dobrinina M, Zurochka V, Hu D, Solovyev A, Ryabova L, et al. Seroprevalence of SARS-CoV-2 antibodies in symptomatic individuals is higher than in persons who are at increased risk exposure: the results of the single-center, prospective, cross-sectional study. Vaccines (Basel). 2021;9:627. https://doi.org/10.3390/vaccines9060627

5. Cook S, Malyutina S, Kudryavtsev A, et al. Know Your Heart: rationale, design and conduct of a cross-sectional study of cardiovascular structure, function, and risk factors in 4,500 men and women aged 35–69 years from two Russian cities, 2015–18. Wellcome Open Research 2018;3:67. https://doi.org/10.12688/wellcomeopenres.14619.3

6. Kuvshinova IN, Nekrasov BG, Livitskaya NI, Molodykh SV, Rukavishnikov M. Sensitivity and specificity of reagent kits of JSC "Vector-Best" for the detection of immunoglobulins of different classes to SARS-CoV-2 [in Russian]. Spravochnik Zaveduyushchego KDL. 2021;10:27–32.

7. Barchuk A, Shirokov D, Sergeeva M, Tursunzade R, Dudkina O, Tychkova V, et al. Evaluation of the performance of SARS-CoV-2 antibody assays for a longitudinal population-based study of COVID-19 spread in St. Petersburg, Russia. J Med Virol. 2021;93:5846–52. https://doi.org/10.1002/jmv.27126

8. Sempos CT, Tian L. Adjusting coronavirus prevalence estimates for laboratory test kit error. Am J Epidemiol. 2021;190:109–15. https://doi.org/10.1093/aje/kwaa174

9. Government of Russia. The number of people vaccinated against coronavirus in Arkhangelsk [in Russian]. 2021 [cited 2021 Sep 20]. https://gogov.ru/covid-v-stats/arkhangelsk

10. European Centre for Disease Prevention and Control. Overview of the implementation of COVID-19 vaccination strategies and deployment plans in the EU/EEA, 14 June 2021. 2021 [cited 2021 Sep 20]. https://www.ecdc.europa.eu/en/publications-data/overview-implementation-covid-19-vaccination-strategies-and-deployment-plans

Address for correspondence: David Leon, London School of Hygiene & Tropical Medicine, Keppel St, London, WC1E 7HT, UK; email: david.leon@lshtm.ac.uk

Seroprevalence of SARS-Cov-2 Antibodies in Adults, Arkhangelsk, Russia

Appendix

A potential limitation of this study is that the population we studied may not be fully representative of the target population of citizens of Arkhangelsk of the same age. This has 2 components. First, the sampling frame for the seroprevalence study was from the previous Know Your Heart (KYH) study conducted 2015–2018. KYH was itself based on a random sample of all persons 35–69 years of age residing in the city of Arkhangelsk. Although the response rate for this initial study was 53%, the educational profile of those who were recruited to the study was very similar to that expected on the basis of 2010 Russian Census data for the city (*1*). This finding, together with emerging evidence that response rates may not be as strongly related to nonresponse bias (*2*), suggests that although we sampled a particular age-range, the sampling frame is probably representative of the population of the city of Arkhangelsk.

The second issue of representativeness concerns the extent to which those participants in the recent seroprevalence study are similar in key respects to the sampling frame from the KYH study. Of the 2,380 KYH participants, we excluded 122 persons from consideration for the following reasons: 56 indicated at the KYH survey that they did not wish to be contacted to take part in further research, 61 had died before the study inception date, and 5 were ≥75 years of age. Overall, 2,258 people were invited to take part in 2021; a total of 1,080 (47.8%) provided blood samples for assessing seroprevalence.

We have compared the similarity of the 1,080 participants in the seroprevalence survey to the 2,380 persons in the sampling frame (Appendix Table). The sex and age distributions were very similar. However, the proportion of participants with higher education in the 2021 seroprevalence study was larger than in the sampling frame. However, we did not observe an association between education and seroprevalence, at least among the responders.

Our study had several limitations. Samples were taken over 4 months during February 24–May 28, 2021, because the fieldwork was nested within a much larger national multicenter

survey of the prevalence of risk factors of cardiovascular diseases in Russia (study ESSE-RF-3). Participants of this study underwent extensive tests as part of the ESSE-RF-3 protocol. Capacity limitations meant that we could invite a maximum of 25 participants per day. We noted, however, that the ESSE-RF study itself aimed to get a representative sample of the population and was in no way restricted to those who had cardiovascular problems. Although we collected the samples over an extended period during which the infection rates changed, we regard our estimate as an average of positivity over the period studied. Nevertheless, we could underestimate the seroprevalence due to low sensitivity in the 12 days following infection or sensitivity waning with increased time from the disease onset. Finally, a limitation of our findings was the small sample sizes we used for some of our analyses.

References

1. Cook S, Malyutina S, Kudryavtsev AV, Averina M, Bobrova N, Boytsov S, et al. Know Your Heart: rationale, design and conduct of a cross-sectional study of cardiovascular structure, function and risk factors in 4,500 men and women aged 35–69 years from two Russian cities, 2015-18 [version 2–referees: 3 approved]. Wellcome Open Research 2018;3 https://doi.org/10.12688/wellcomeopenres.14619.1

2. Hendra R, Hill A. Rethinking response rates: new evidence of little relationship between survey response rates and nonresponse bias. Eval Rev. 2019;43:307–30. https://doi.org/10.1177/0193841X18807719

Appendix Table. Comparisons of participants in study of seroprevalence of severe acute respiratory syndrome coronavirus 2 antibodies, Russia

Characteristic	Resurvey sample, N=1,067 (%)	Know Your Heart study, N=2,380 (%)	p value
Sex*			
F	629 (59.0)	1,377 (58.3)	0.72
M	438 (41.0)	985 (41.7)	
Median age at baseline, quartiles Q1 and Q3	51 (Q_1=44–Q_3=59)	54 (Q_1=45–Q_3=62)	<0.01†
Education*			<0.01
Secondary and lower	56 (5.2)	174 (7.4)	
Specialized secondary	520 (48.7)	1,269 (53.7)	
Higher	491 (46.0)	919 (38.9)	

*Determined by χ^2 test.
†Determined by Mann–Whitney U-test.

Paper 1

Krieger E., Sharashova E., Kudryavtsev A., Samodova O., Kontsevaya A., Brenn T., Postoev V. (2023).

COVID-19: seroprevalence and adherence to preventive measures in

Arkhangelsk, Northwest Russia.
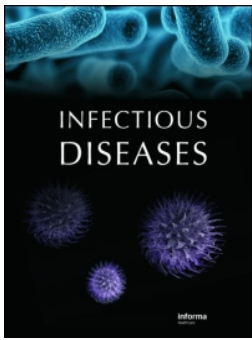
COVID-19: seroprevalence and adherence to preventive measures in Arkhangelsk, Northwest Russia

Ekaterina Krieger, Ekaterina Sharashova, Alexander V. Kudryavtsev, Olga Samodova, Anna Kontsevaya, Tormod Brenn & Vitaly Postoev

View supplementary material 🗗

Published online: 15 Mar 2023.

Submit your article to this journal 🗗

Article views: 21

View related articles 🗗

View Crossmark data 🗗

Taylor & Francis
Taylor & Francis Group

ORIGINAL ARTICLE

Check for updates

COVID-19: seroprevalence and adherence to preventive measures in Arkhangelsk, Northwest Russia

Ekaterina Krieger[a,b] , Ekaterina Sharashova[a] , Alexander V. Kudryavtsev[a,b] , Olga Samodova[c] ,
Anna Kontsevaya[d] , Tormod Brenn[a] and Vitaly Postoev[e]

[a]Department of Community Medicine, UiT The Arctic University of Norway, Tromsø, Norway; [b]International Research Competence
Centre, Northern State Medical University, Arkhangelsk, Russian Federation; [c]Department of Infectious Diseases, Northern State
Medical University, Arkhangelsk, Russian Federation; [d]Department of Public Health, National Medical Research Centre for Therapy
and Preventive Medicine, Moscow, Russian Federation; [e]Department of Research Methodology, Northern State Medical
University, Arkhangelsk, Russian Federation

ABSTRACT

Background: The published estimates of SARS-CoV-2 seroprevalence in Russia are few. The study aimed to assess the SARS-CoV-2 seroprevalence in Arkhangelsk (Northwest Russia), in a year after the start of the pandemic, to evaluate the population adherence to non-pharmaceutical interventions (NPIs), and to investigate characteristics associated with COVID-19 seropositive status.

Methods: We conducted a SARS-CoV-2 seroprevalence study between 24 February and 30 June 2021 involving 1332 adults aged 40–74 years. Logistic regression models were fit to identify factors associated with seropositive status and with adherence to NPIs.

Results: Less than half (48.9%) of study participants adhered all recommended NPIs. Male sex (odds ratio [OR] 1.7, 95% confidence intervals [CI] 1.3; 2.3), regular employment (OR 1.8, 95% CI 1.3; 2.5) and low confidence in the efficiency of the NPIs (OR 1.9, 95% CI 1.5; 2.5) were associated with low adherence to internationally recommended NPIs. The SARS-CoV-2 seroprevalence rate was 65.1% (95% CI: 62.5; 67.6) and increased to 73.0% (95% CI: 67.1; 85.7) after adjustment for test performance. Regular employment (OR 2.0, 95% CI 1.5; 2.8) and current smoking (OR 0.4, 95% CI 0.2; 0.5) were associated with being seropositive due to the infection.

Conclusions: Two third of the study population were seropositive in a year after the onset of the pandemic in Arkhangelsk. Individuals with infection-acquired immunity were more likely to have regular work and less likely to be smokers. The adherence to NPIs was not found associated with getting the virus during the first year of the pandemic.

Supplemental data for this article can be accessed online at https://doi.org/10.1080/23744235.2023.2179660.

Background

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) emerged in China in December 2019 and through the first half of 2020 developed into a global pandemic of coronavirus disease 2019 (COVID-19). The spectrum of infection varied from being asymptomatic to severe with hospitalization and death. The tendency of those who were asymptomatic or had only mild symptoms not to self-isolate will have contributed to further spread of infection [1].

The true cumulative incidence of COVID-19 in most populations remains uncertain and exceeds the number of reported cases [2]. Seroprevalence surveys of the general population can provide a less biased assessment of the extent of COVID-19 infection compared to those based on the number of positive tests reported by the health service. This is because the latter will differentially tend to exclude asymptomatic and mild cases and will also be a function of availability of testing facilities [2].

Several random population antibody surveys were performed during the first year of the COVID-19 pandemic. At the beginning of 2021, 34.6% of the population sample in England tested positive for antibodies against SARS-CoV-2, while the estimated seroprevalence in Norway was 0.9% [3,4]. In Russia, only a few serologic studies had been conducted by the middle of 2021 [5–7]. In June-December 2020, the average SARS-CoV-2 seroprevalence in Russia was estimated as 19.2% with variation between regions [7].

In the first year of the pandemic, studies tended to find seroprevalence to be associated with sociodemographic and behavioral characteristics. Male gender was shown to be positively associated with testing positive [8]. Having public-facing jobs, living in overcrowded households, using public transport, or having high social interaction for other reasons were also associated with testing positive for SARS-CoV-2 antibodies [9,10]. Lower seroprevalence has been reported in older populations and in people with chronic diseases who may have shielded themselves to reduce risks of getting infected [3,5,11–15].

Non-pharmaceutical interventions (NPIs), such as laws or regulations restricting face-to-face interactions, can significantly decrease the rate of viral transmission [16,17]. The varying speed of the infection spread in different countries depends on the public compliance with the related guidelines rather than the timeliness of implementing COVID-19 NPIs [10,18]. The level of adherence to COVID-19 restrictive measures can be influenced by public confidence in government and the information available to guide preventive behaviors [19]. Higher compliance with NPIs has been demonstrated in the elderly, women, those with higher educational level and income, non-smokers, people living alone and those having chronic diseases [1,20,21]. There is some evidence that some people who thought that they had had COVID-19 were less likely to adhere to NPIs due to their belief that this gave them protection from further infection [22].

In the Arkhangelsk Region in Northwest Russia (population 1.1 million in January 2021) [23], the first case of COVID-19 was registered on 17 March 2020 [24]. The period from 19 March to 2 July 2020 was associated with the first wave in the region, and the period from 20 September 2020 to 27 February 2021 was associated with the second wave of COVID-19 [24].

The vaccination campaign in the Arkhangelsk Region started in November 2020, mainly with Russian Sputnik V. As in other countries, therefore, levels of population immunity during the first year were driven largely by the spread of the infection.

As the disease can be asymptomatic and due to low access to testing, the true proportion of the Arkhangelsk population with immunity to COVID-19 acquired during the first year of the pandemic remains unknown. Beyond that, the lack of knowledge of the factors associated with the COVID-19 infection was an obstacle to preventing the further spread of COVID-19 and controlling other similar future pandemics.

This study aimed to assess the seroprevalence of SARS-CoV-2 in Arkhangelsk, a city in Northwest Russia, in a year after the start of the pandemic, to estimate the population adherence to NPIs during the first year, and to investigate socioeconomic, behavioral and health-related characteristics associated with infection-acquired antibodies against SARS-CoV-2 as well as with adherence to NPIs.

Materials and methods

Study design and participants

A cross-sectional study of SARS-CoV-2 seroprevalence was conducted in Arkhangelsk between 24 February and 30 June 2021 (50% of participants were enrolled by 5 April 2021). This was a sub-study of the third multi-center survey 'Epidemiology of Cardiovascular Diseases and their Risk Factors in Regions of the Russian Federation' (ESSE-RF3) [25]. The ESSE-RF study aimed to recruit a sample that was representative of the population of

Russia. In Arkhangelsk, this was done by inviting participants involved in an earlier study of cardiovascular diseases, Know Your Heart (KYH), which was conducted in 2015–2017 and included a random sample of Arkhangelsk population aged 35–74 years [26]. The KYH study participants were recruited from four districts of Arkhangelsk with an anonymized list of addresses of residents with compulsory medically insurance used as the sampling frame. The address list was provided by the regional health insurance fund, with each address supplemented by age and sex of the insured resident. Trained interviewers visited randomly selected addresses to invite persons of the corresponding age and sex to take part in the study. The participation rate was 68% out of the total invitees.

Of 2380 KYH study participants, we invited 2258 to ESSE-RF3. The exclusions ($N = 122$) were for the following reasons: 56 KYH participants had not consented to be contacted with invitations to other studies, 61 had died prior to the launch of ESSE-RF3 (4 of 61 deaths were related to COVID-19), and 5 had become older than the maximum age (74 years). Most of the participants in the KYH study were aged 40 years or older at the time of the seroprevalence survey. Of those invited, 1348 KYH study participants (60%) aged 40–74 years took part, all with signed informed consent.

The study procedure included a health check, blood sample collection (for biochemical assays and for SARS-CoV-2 antibodies) and an interview at the outpatient clinic of Northern State Medical University (NSMU), Arkhangelsk. Two participants who did not complete the questionnaire and 14 participants who had an equivocal serological test result were excluded from the seroprevalence sub-study sample. Therefore, the final analytic sample comprised 1332 participants.

Ethical considerations

All participants included in the seroprevalence study provided written informed consent to participate. The study was conducted in compliance with the ethical standards of the 1964 Declaration of Helsinki and its later amendments. Ethical approval for the original KYH study was given by the ethics committees of London School of Hygiene & Tropical Medicine (approval number 8808 received 24 February 2015) and NSMU, Arkhangelsk (approval number 01/01-15 received 27 January 2015). Ethical approval for ESSE-RF-3 was obtained from the Ethics Committee of the National Research Centre for Therapy and Preventive Medicine,

Moscow, Russia (approval number №01-01/20 received 04 February 2020) and the Ethics Committee of NSMU, Arkhangelsk, Russia (approval number 01/02-21 received 17 February 2021). Ethical approval for the sub-study on COVID-19 and health-related factors was received from the Ethics Committee of NSMU, Arkhangelsk, Russia (approval number 01/02-21 received 17 February 2021). All study procedures were approved by the Regional Committees for Medical and Health Research Ethics (REK) in Norway (approval number 339397 received 7 December 2021).

Measurements

Laboratory methods

Blood samples were tested for SARS-CoV-2 IgG antibodies using a Vector Best ELISA assay (D-5501 SARS-CoV-2-IgG-EIA-BEST) [27,28]. The assay is an enzyme linked immunosorbent assay (ELISA) test-systems (Russia) for the semi-quantitative detection of antibodies to the spike (S) protein of SARS-CoV-2 in human blood serum [27]. The sensitivity of this assay has been reported to be 72% within the first 12 days following the infection and close to 100% at a later stage [27]. An independent test-performance study has shown the assay sensitivity of 89% and the specificity of 100% based on the comparisons of test results in pre-pandemic samples (negative controls) and polymerase chain reaction positive samples for SARS-CoV-2 [28].

Questionnaire

Every participant completed an interviewer-administered questionnaire capturing data on demographics, health-related characteristics, history of COVID-19 symptoms and self-reported protective behaviors. The questionnaire on COVID-19 was developed by the ESSE-RF3 team at the National Research Centre for Therapy and Preventive Medicine, Moscow [25]. Additional questions were asked in Arkhangelsk to collect data on vaccination against COVID-19. The original Russian version and the English translation of the COVID-19 questionnaire are provided in Appendix A.

Adherence to NPIs estimates were based on respondents' self-reports with respect to five COVID-19 NPIs: self-isolation, social distancing, wearing facemasks in public settings or transport, wearing gloves, and use of hand sanitizers [29]. The use of gloves for the prevention of SARS-CoV-2 transmission was not implemented globally, while in Russia wearing gloves was

recommended by the Federal service on customers' rights protection and human well-being surveillance (Rospotrebnadzor) until 4 February 2022 [30,31].

The term 'self-isolation' was not described in the questionnaire, but was primarily considered as a measure required for those who were confirmed cases or had contacts with confirmed cases as well as those who were older than 65 years or had chronic diseases [32,33]. Besides, staying at home was recommended to everyone during paid non-working days (from March 30 to April 4, from April 4 to April 30, May 6 to May 8) established by the Presidential Decrees in 2020 [34–36].

The answer 'yes' to the question about adherence to NPIs during the pandemic was counted as 1 for each of the five interventions whereas the answer 'no' was counted as 0. The total gave a scale ranging from 0 to 5, which was dichotomized by combining scale values 0–3 (low adherence) and values 4–5 (high adherence). The threshold was chosen taking into account the median number of NPIs adhered, which is equal to 4.

The study participants assessed the efficiency of the five above-mentioned NPIs using a Likert scale (not effective at all = 0 to very effective = 5) (Appendix A). For each participant, we summed score assessments for all NPIs and this gave composite assessment score of NPIs ranging from 0 to 25. Values 0–19 were coded 1 (low confidence in the efficiency) and values 20–25 (at least 4 points for each assessment) were coded 0 (high confidence in the efficiency).

In order to identify factors associated with low adherence to NPIs and to investigate associations of the seroprevalence status with socioeconomic, behavioral and health-related characteristics, we used the questionnaire-based data on the following covariates: sex, age, education, marital status, number of persons in household, occupation, income, smoking, alcohol consumption and self-reported chronic non-communicable diseases (hypertension, diabetes mellitus, chronic pulmonary diseases, coronary heart diseases).

Age was categorized as: 40–54 years, 55–64 years and 65–74 years. The classification of educational levels was as follows: secondary or lower, specialized secondary and higher. Marital status was defined as single (including widowed and divorced) and married or living with partner. The number of persons in a household was analyzed as discreet variable. Living with children (<18 years) was categorized as yes or no. Occupation was classified as being in regular employment, or not. According to income categorization by the Federal State Statistics Service of Russia, participants earning less than

40,000 Rub a month were considered low-income, those earning between 40,000 and 100,000 Rub a month – middle income, and those earning more than 100,000 Rub a month – high-income [37,38]. With respect to tobacco smoking, the respondents were divided into three groups: never smokers, former smokers and current smokers. According to the definition of 'Heavy Episodic Drinking' suggested by the World Health Organization, heavy drinkers were defined as consumers of 60 or more grams of pure alcohol on a single occasion [39]. The frequency of heavy drinking was ascertained for previous 12 months and was classified as never, once a week or less often, and twice a week or more often.

Data on participant-reported chronic health conditions were collected in the ESSE-RF3 survey and included hypertension, diabetes mellitus, chronic pulmonary diseases and coronary heart diseases. Hypertension was defined as self-reported prior diagnosis and/or intake of antihypertensive drugs. Diabetes mellitus was defined as self-reported prior diagnosis of diabetes and/or diabetes medication use, chronic pulmonary disease – self-reported prior diagnosis of chronic bronchitis or bronchial asthma, coronary heart disease – prior diagnosis of angina pectoris and/or myocardial infarction and/or antianginal medication use.

Self-reports of vaccination against COVID-19 were used to classify participants either as those unvaccinated or those who had received one or two doses. Participants were classified as having COVID-19 during the first year of pandemic based on the answer to the question: 'Did you have COVID-19 during the last 12 months?' without further specification of symptoms or test results.

Study participants who reported having a positive test for COVID-19 but reported that they had had no symptoms of the infection as well as those who were tested positive for SARS-CoV2 antibodies but reported no COVID-19 infection were considered asymptomatic.

Statistical analysis

Continuous variables were summarized using descriptive statistics in terms of mean and standard deviation (SD). Seroprevalence was estimated as the number of SARS-CoV-2 IgG positive participants divided by the number of tested participants and reported in percentages. Confidence intervals (CIs) for unadjusted seroprevalence were calculated using Wilson's procedure. Seroprevalence adjusted for test performance (89%

sensitivity, 100% specificity) was estimated to improve comparability of the study findings with other studies as shown elsewhere [40,41]. The adjustment was performed using the equation: (crude prevalence + test specificity − 1)/(sensitivity + specificity − 1) [42]. The 95% CIs for the adjusted estimates of seroprevalence were calculated by bootstrapping procedure using R package bootComb (version 4.1.1) [43].

The unadjusted and adjusted seroprevalence of SARS-CoV-2 in subgroups were compared using 95% CIs. The Pearson Chi-squared test was used to analyze categorical data. Cronbach's alpha coefficient was used to measure the internal coherence/reliability of composite score for adherence to COVID-19 NPIs.

Binary logistic regression was used to investigate factors associated with low adherence to COVID-19 NPIs (1 – low adherence (0–3 NPIs), 0 – high adherence (4–5 NPIs)). Sex, age, education, occupation, income, smoking and drinking habits, chronic health conditions, vaccination against COVID-19 were introduced in regression model using enter option. Crude and adjusted odds ratios (ORs) with 95% CIs were calculated. Since wearing of gloves was not an international recommendation, we repeated the multivariable analysis with exclusion of the gloves wearing to ensure comparability of the results. In this analysis, the scoring for adherence was as follows: 1 – low adherence (0 to 3 NPIs), 0 – high adherence (4 NPIs).

The associations between sex, age, marital status, number of persons in household, living with children (<18 years), education, occupation, income, adherence to NPI, smoking and drinking habits, chronic health conditions (later referred as selected factors) and the seropositive status (1 – seropositive, 0 – seronegative) were investigated using binary logistic regression, with some analyses being stratified by vaccination status. The independent variables were grouped into two hierarchical blocks in relation to the dependent variable. Block 1 included socio-demographic characteristics (sex, age, marital status, number of persons in household, living with children (<18 years), education, occupation, income). Block 2 included behavioral factors (adherence to COVID-19 NPIs, smoking, heavy alcohol drinking) and chronic health conditions. The blocks of variables were introduced in the regression model in a stepwise manner. The independent variables associated with seropositive status after the exclusion of vaccinated individuals were interpreted as factors associated with infection-acquired immunity.

We calculated the statistical power of a binary logistic regression model to identify factors associated with seropositive status in a sample of 1332 observations. Calculations have shown that the sample gives a statistical power of ≥80% to identify factors which increase or reduce odds of the outcome by 1.5 times for all combinations of the outcome prevalence in the range from 25% to 70% and predictor prevalence in the range from 30% to 70%.

Results

The mean age of the study participants was 57 (SD 9.6) years. Women made up 59.7% of the sample (Table 1). Overall, 867 of the total 1332 participants were tested positive for SARS-CoV-2 IgG antibody corresponding to a seroprevalence rate of 65.1% (95% CI: 62.5; 67.6). Seroprevalence adjusted for test performance was 73.0% (95% CI: 67.1; 85.7).

Only 61.7% of 47 individuals who had received one dose of the vaccine were seropositive whereas 99.5% of 195 participants who had received two doses were seropositive. Among 339 participants who self-reported having had COVID-19, 95.0% were seropositive.

Among those unvaccinated ($N = 1090$), 335 (30.7%) did not report having had COVID-19 but were seropositive. In addition, seven of eight unvaccinated individuals who reported having had COVID-19 asymptomatically were also seropositive. Overall, the proportion of asymptomatic cases among unvaccinated study participants was 31.4% (342/1090).

Adjusted seroprevalence was lower in current smokers and was higher in vaccinated individuals and those who self-reported having COVID-19 (Table 1).

Almost all participants reported wearing facemasks in public settings or transport during the pandemic period (98.6%), maintaining social distancing (92.9%) and using hand sanitizers (91.5%). Over 90.0% of retired or unemployed study participants and 72.0% of regularly employed participants followed self-isolation guidelines during the COVID-19 pandemic, which may include not mixing socially during lockdowns, $p < .001$. More than half (59.6%) reported wearing gloves. The use of gloves was adhered by 55.7% of participants aged 40–54 years, 62.1% of those aged 55–64 years and 63.2% individuals older than 65 years, $p = .036$. In total, 48.9% adhered all five NPIs. The Cronbach alpha reliability coefficient for the composite adherence scale (0.67) was considered acceptable. The low-adherent participants were less

Table 1. Unadjusted and adjusted SARS-CoV-2 seroprevalence by selected participant characteristics, Arkhangelsk, Russia.

Variables	N (%)	Unadjusted seroprevalence[a] % (95% CI)	Seroprevalence[a] adjusted for test performance % (95% CI)
Sex			
Male	537 (40.3)	66.1 (62.0; 70.0)	74.2 (64.9; 85.4)
Female	795 (59.7)	64.4 (61.0; 67.7)	72.3 (63.7; 83.0)
Age			
40–54 years	575 (43.2)	66.3 (62.3; 70.0)	74.3 (65.2; 85.5)
55–64 years	420 (31.5)	60.7 (56.0; 65.3)	68.1 (58.7; 79.1)
65–74 years	337 (25.3)	68.6 (63.4; 73.3)	76.9 (66.9; 89.0)
Marital status			
Single (including widowed and divorced)	464 (34.8)	60.3 (55.8; 64.7)	67.7 (58.4; 78.4)
Married or living with partner	868 (65.2)	67.6 (64.5; 70.7)	75.9 (67.3; 87.0)
Number of persons in household			
1 (alone)	240 (18.0)	61.3 (55.0; 67.2)	68.7 (57.8; 80.6)
2–3	882 (66.2)	64.3 (61.1; 67.4)	72.1 (63.6; 82.7)
≥4	210 (15.8)	72.9 (66.5; 78.4)	81.7 (70.9; 94.5)
Living with children (<18 years)			
No	1019 (76.5)	63.7 (60.7; 66.6)	71.4 (63.0; 81.8)
Yes	313 (23.5)	69.7 (64.3; 74.4)	78.1 (67.9; 90.4)
Education			
Higher	563 (42.3)	64.3 (60.3; 68.2)	72.1 (63.0; 83.1)
Specialized secondary	661 (49.6)	66.0 (62.3; 69.5)	74.0 (65.1; 85.0)
Secondary or lower	108 (8.1)	63.9 (54.5; 72.3)	71.7 (57.9; 86.4)
Occupation			
Retired or unemployed	542 (40.7)	60.2 (56.0; 64.2)	67.5 (58.5; 78.0)
Regular employment	790 (59.3)	68.5 (65.2; 71.6)	76.8 (68.1; 88.0)
Income			
High	175 (13.1)	69.1 (61.9; 75.5)	77.6 (65.8; 90.8)
Middle	764 (57.4)	66.4 (62.9; 69.6)	74.4 (65.6; 85.3)
Low	393 (29.5)	60.8 (55.9; 65.5)	68.2 (58.5; 79.2)
Smoking			
Never smoker	742 (55.7)	67.8 (64.3; 71.1)	76.1 (67.2; 87.2)
Former smoker	366 (27.5)	67.8 (62.8; 72.3)	76.0 (66.3; 88.0)
Current smoker	224 (16.8)	51.8 (45.3; 58.3)	58.1 (47.2; 69.1)
Frequency of heavy drinking			
Never	825 (61.9)	66.2 (62.9; 69.3)	74.6 (65.6; 85.2)
Once a week or less often	453 (34.0)	64.5 (60.0; 68.7)	72.3 (62.9; 83.6)
Twice a week or more often	54 (4.1)	53.7 (40.6; 66.3)	72.3 (62.9; 83.6)
Chronic health conditions			
No	393 (29.5)	63.4 (58.5; 68.0)	71.1 (61.4; 82.4)
Yes	939 (70.5)	65.8 (62.7; 68.8)	73.8 (65.4; 84.7)
Self-reported having had COVID-19			
No	993 (74.5)	54.9 (51.8; 58.0)	61.6 (53.4; 70.7)
Yes	339 (25.5)	95.0 (92.1; 96.0)	100.0 (95.9; 100.0)
Vaccinated against COVID-19			
No	1090 (81.8)	59.1 (56.1; 62.0)	66.3 (58.1; 76.0)
1 dose	47 (3.5)	61.7 (47.4; 74.2)	69.2 (50.1; 88.3)
2 doses	195 (14.6)	99.5 (97.2; 99.9)	100.0 (97.2; 100.0)
Adherence to NPIs[b]			
High (4–5)	1082 (81.2)	65.3 (62.5; 68.1)	73.3 (65.0; 84.0)
Low (0–3)	250 (18.8)	64.0 (57.9; 69.7)	71.8 (60.9; 83.9)
Total	1332 (100.0)	65.1 (62.5; 67.6)	73.0 (64.9; 83.5)

CI: confidence interval; COVID-19: coronavirus disease 2019; NPI: non-pharmaceutical interventions.
[a]Including vaccinated individuals.
[b]The total number of NPIs adhered.

likely to have high confidence in NPIs (53.8%) compared to those who were highly adherent (77.2%), $p < .001$.

In the crude analysis, male sex, age 40–54 years compared to age 65–74 years, secondary or lower education compared to higher education, regular employment, smoking, heavy alcohol drinking, being unvaccinated and having low confidence in the efficiency of NPIs were associated with low adherence to COVID-19 NPIs. Those who were vaccinated were more likely to have high adherence (Table 2). Multivariable analysis showed that male sex, low income compared to high income,

low confidence in the efficiency of NPIs and heavy drinking twice a week or more often were associated with low adherence to NPI. After exclusion of wearing gloves from the analysis, low income and frequency of heavy drinking were no longer associated with low adherence to COVID-19 NPIs. Male sex, regular employment and low confidence in the efficiency of NPIs were associated with low adherence to NPIs recommended globally.

Among those unvaccinated, being in regular employment was associated with higher odds of being seropositive due to the infection (Table 3, Model 1). After

Table 2. Variables associated with low adherence to COVID-19 non-pharmaceutical interventions, Arkhangelsk, Russia (binary logistic regression analysis).

Characteristics	Crude OR (95% CI)[a] Low 0–3 / high 4–5	Adjusted[b] OR (95% CI) Low 0–3 / high 4–5[a]	Adjusted[b] OR (95% CI)[c] Low 0–3 / high 4[c]
Sex			
Female	Reference	Reference	Reference
Male	2.3 (1.8; 3.1)	2.2 (1.5; 3.1)	1.7 (1.3; 2.3)
Age			
40–54 years	1.9 (1.2; 2.7)	1.3 (0.8; 2.1)	1.3 (0.9; 2.0)
55–64 years	1.5 (1.0; 2.3)	1.2 (0.7; 1.9)	1.1 (0.7; 1.6)
65–74 years	Reference	Reference	Reference
Marital status			
Single (including widowed and divorced)	Reference	Reference	Reference
Married or living with partner	1.0 (0.8; 1.3)	0.9 (0.6; 1.2)	1.0 (0.7; 1.4)
Education			
Higher	Reference	Reference	Reference
Specialized secondary	1.3 (1.0; 1.7)	1.3 (0.9; 1.8)	1.1 (0.8; 1.5)
Secondary or lower	1.7 (1.1; 2.8)	1.6 (0.9; 2.8)	1.3 (0.8; 2.1)
Occupation			
Retired or unemployed	Reference	Reference	Reference
Regular employment	1.6 (1.2; 2.2)	1.4 (1.0; 2.0)	1.8 (1.3; 2.5)
Income			
High	Reference	Reference	Reference
Middle	1.4 (0.9; 2.2)	1.6 (1.0; 2.7)	1.3 (0.8; 1.9)
Low	1.5 (0.9; 2.5)	2.2 (1.2; 4.1)	1.7 (1.0; 2.8)
Smoking			
Never smoker	Reference	Reference	Reference
Former smoker	1.4 (1.0; 1.9)	0.9 (0.6; 1.3)	0.9 (0.7; 1.2)
Current smoker	1.8 (1.3; 2.6)	1.0 (0.7; 1.5)	1.1 (0.8; 1.6)
Frequency of heavy drinking			
Never	Reference	Reference	Reference
Once a week or less often	1.8 (1.4; 2.5)	1.3 (0.9; 1.8)	1.1 (0.8; 1.5)
Twice a week a week or more often	4.0 (2.3; 7.2)	2.5 (1.3; 4.9)	1.7 (0.9; 3.3)
Chronic health conditions			
No	Reference	Reference	Reference
Yes	1.0 (0.7; 1.3)	1.0 (0.7; 1.4)	1.1 (0.8; 1.7)
Confidence in the efficiency of NPIs			
Yes	Reference	Reference	Reference
No	2.9 (2.2; 3.8)	2.7 (2.0; 3.7)	1.9 (1.5; 2.5)
Vaccinated against COVID-19			
Yes	Reference	Reference	Reference
No	1.7 (1.1; 2.6)	1.4 (0.9; 2.1)	1.2 (0.8; 1.7)

Note: CI: confidence interval; COVID-19: coronavirus disease 2019; NPI: non-pharmaceutical interventions; OR: odds ratio.
[a]Low adherence meant to adhere 3 or less NPIs out of maximum 5 (self-isolation, social distancing, wearing facemasks in public places or transport, wearing gloves, and use of hand sanitizers).
[b]Adjusted for all independent variables included in the model.
[c]Wearing gloves was excluded from the analysis. Low adherence meant to adhere 3 or less NPIs out of maximum 4.

introducing behavioral factors and chronic health conditions, regular employment was associated with higher odds of being seropositive, while current smoking was associated with lower odds (Table 3, Model 2). The same analysis performed after including vaccinated individuals (N = 242) showed that regular employment and smoking had similar associations with seropositive status regardless of whether it was obtained *via* infection or *via* vaccination (Supplementary table 1, Appendix B).

Discussion

The SARS-CoV-2 seroprevalence 12–16 months after the beginning of the pandemic in Arkhangelsk was 65.1% (95% CI: 62.5; 67.6) and increased to 73.0% (95% CI: 67.1; 85.7) after adjustment for test performance. The individuals having regular employment had higher probability to be seropositive to SARS-CoV-2, while smokers were less likely to be seropositive. Low adherence to NPIs recommended globally during the first year of the pandemic was associated with male sex, regular employment and low confidence in the efficiency of NPIs.

SARS-CoV-2 seroprevalence

The SARS-CoV-2 seroprevalence in Arkhangelsk a year after the start of the pandemic was higher than found in other cities of Russia for which data had been reported [5–7,44]. Asymptomatic cases comprised 31.4% of unvaccinated participants. This percentage was twice as high as the asymptomatic proportion found in meta-analysis of COVID-19 studies [1,45]. This could be partly explained by the possibility of misdiagnosis or missed diagnosis of patients with COVID-19 due to limited capacity of testing and the pressure on the healthcare system in Arkhangelsk. People could

Table 3. Variables associated with being SARS-CoV-2 seropositive among those not immunized, Arkhangelsk, Russia (binary logistic regression).

Variables	Model 1 Adjusted[a] OR (95% CI)	Model 2 Adjusted[b] OR (95% CI)
Sex		
Female	Reference	
Male	0.8 (0.7; 1.1)	1.1 (0.8; 1.6)
Age		
40–54 years	0.8 (0.5; 1.2)	1.0 (0.7; 1.6)
55–64 years	0.8 (0.5; 1.1)	0.9 (0.6; 1.3)
65–74 years	Reference	Reference
Marital status		
Single	Reference	Reference
Married	1.2 (0.8; 1.6)	1.1 (0.8; 1.5)
Number of persons in household	1.2 (1.0; 1.4)	1.2 (1.0; 1.5)
Living with children (<18 years)		
No	Reference	Reference
Yes	1.0 (0.7; 1.6)	1.0 (0.6; 1.5)
Education		
Higher	Reference	Reference
Specialized secondary	1.1 (0.9; 1.6)	1.3 (1.0; 1.8)
Secondary or lower	1.1 (0.7; 1.8)	1.2 (0.7; 1.9)
Occupation		
Retired or unemployed	Reference	Reference
Regular employment	2.0 (1.5; 2.7)	2.0 (1.5; 2.8)
Income		
High	Reference	Reference
Middle	1.0 (0.7; 1.5)	1.0 (0.7; 1.5)
Low	1.0 (0.6; 1.6)	1.0 (0.6; 1.7)
Adherence to NPI[c]	–	
High (4–5)	–	Reference
Low (0–3)	–	1.0 (0.7; 1.4)
Smoking		
Never smoker	–	Reference
Former smoker	–	0.8 (0.6; 1.1)
Current smoker	–	0.4 (0.2; 0.5)
Frequency of heavy drinking		
Never	–	Reference
Once a week or less often	–	0.9 (0.7; 1.3)
Twice a week or more often	–	0.5 (0.3; 1.0)
Chronic health conditions		
No	–	Reference
Yes	–	1.1 (0.8; 1.4)

Note: CI: confidence interval; NPI: non-pharmaceutical interventions; OR: odds ratio.
[a]Adjusted for all sociodemographic factors in model 1.
[b]Adjusted for the factors in model 1, behavioral factors (adherence to preventive measures, smoking, heavy alcohol drinking) and chronic health conditions (model 2).
[c]The total number of NPIs adhered.

have had symptoms but were not tested for SARS-CoV2. Therefore, they did not realize that their symptoms were due to COVID-19. At the beginning of the pandemic only people arriving in Arkhangelsk from abroad and close contacts of confirmed cases were required to be tested for COVID-19 using polymerase chain reaction. By the middle of 2020, healthcare-workers as well as all patients with community-acquired pneumonia or other respiratory infection considered by a doctor as suspected COVID-19 cases were tested free of charge [32]. Voluntary testing and obligatory testing of those arriving from abroad were not covered by compulsory health insurance. Some enterprises provided COVID-19 testing for their employees for free. The number of people tested in the Arkhangelsk

Region during 2020 was 598,113 (690,209 tests), including voluntary testing in private clinics [24].

The proportion of fully vaccinated study participants (14.6%) was low as the vaccination campaign in Arkhangelsk, as in the rest of Russia, progressed slowly [5]. Nevertheless, the proportion of the study participants vaccinated against COVID-19 was higher than the officially reported percentage of vaccinated Arkhangelsk inhabitants, which varied during the study period from 3.3% in March 2021 to 13.6% in June 2021 [24]. The likely reason is that our study covered predominantly urban population with better access to vaccination. Another possible explanation is a higher willingness of vaccinated individuals to take part in the study.

We cannot exclude the possibility that participants who thought they had had COVID-19 were more likely to take part in the study. They might feel safer to 'go out' because of having recovered from the disease, while those who decided to avoid having a health check might be the people who had not had COVID-19 and preferred staying home because of fear of getting infected. Besides, those who refused to participate in the study and decided to avoid visiting the healthcare facility might have higher adherence to NPIs and higher probability to be seronegative.

Adherence to preventive measures

Less than half (48.9%) of study participants adhered all recommended NPIs, which corresponds to the results obtained by others [46,47]. In line with the results previously reported by other researchers, we found that people aged 65 years or older were more likely to have high adherence to NPIs as compared to a younger age group. After adjustment for all independent variables, adherence to NPIs was no longer different across age groups [48]. We did not find an association between smoking, chronic health conditions and adherence to COVID-19 NPIs [20,21,46]. Our results are in agreement with prior research, which has shown that males and heavy drinkers were less likely to comply with recommended NPIs [46,47,49,50]. Low confidence in NPIs could be associated with unwillingness to follow recommendations from the government.

Factors associated with SARS-CoV-2 seroprevalence

In unvaccinated individuals, seropositivity was associated with regular employment. These findings are in agreement with prior research, which has shown that

employed people had higher odds to be infected with SARS-CoV-2 [46,49].

In contrast with other studies, we did not find a negative association between adherence to COVID-19 NPIs and infection-acquired positive serological status of the participants [21]. This could be partly explained by common counterfeit compliance with NPIs (e.g. wearing facemask leaving nose exposed) which was reported as being adherent. This might lead to higher viral exposure compared to people who properly follow all restrictions [13,14]. Moreover, behaviors could have changed during the pandemic period, which may result in underestimated association between seroprevalence and adherence to NPIs due to non-differential misclassification of the adherence status.

In our study, adherence to NPIs was assessed as a composite variable made up of five variables (self-isolation, social distancing, wearing facemasks in public settings or transport, wearing gloves, and use of hand sanitizers) with acceptable reliability of 0.67. Since the term 'self-isolation' was not clearly defined in the questionnaire, the participants could consider staying at home during paid non-working days in March–May 2020 as being self-isolated. This could blur the positive association between self-isolation due to being a confirmed COVID-19 case or having contacts with confirmed cases and SARS-CoV-2 seroprevalence.

Smoking status was negatively associated with being tested positive for antibodies to SARS-CoV-2. Some authors reported similar results [5,51,52]. These findings require further research and should be interpreted with caution. The mechanisms that might underline this association suggested by others are largely speculative [53]. It remains possible that smokers develop a lower antibody response after the infection [54]. Other researchers demonstrated a higher expression of angiotensin converting enzyme-2 in smokers that might lead to greater susceptibility to COVID-19 [55]. Although the high expression of the angiotensin converting enzyme-2 had an inhibitory effect on virus replication and smokers might be more likely to have asymptomatic infection [55]. In our study, we found no evidence that would confirm this statement.

Strengths and limitations

This is the first population-based study in Russia estimating SARS-CoV-2 seroprevalence and exploring factors associated with seroprevalence, including the adherence to NPIs, that could play an important role in the COVID-19 pandemic and should be considered in preventing similar epidemics in the future. A strength of our study is that we adjusted seroprevalence estimates for laboratory test performance characteristics to prevent the bias associated with imperfect test performance and improve the comparability of the results [56–58].

Our findings should be interpreted in the light of some limitations. The study was limited to citizens aged 40–74 years in Arkhangelsk, Northwest Russia. We invited participants of the previous random population study (Know Your Heart), who may not have been a fully representative sample of the population. However, we have compared the socio-demographic characteristics of the 1332 seroprevalence survey participants to the 2380 persons in the sampling frame. There were not any significant differences in sex, age or education distributions between them [59]. This, together with the emerging evidence that response rates may not be as strongly related to non-response bias, suggests that the sampling frame is likely to be representative of the adult population aged 40–74 years of Arkhangelsk.

We did not verify the self-reported data on having COVID-19 or vaccination status with medical records, so they may not be completely accurate. Therefore, an element of informational bias could not be excluded.

Most of the participants reported that they followed COVID-19 NPIs during the pandemic period and considered preventive measures to be highly effective. These data can be compromised by socially desirable answer to the question regarding COVID-19-related restrictions introduced by the local government. Due to the possible influence of social desirability bias, the reported rates of adherence to NPIs may be overestimated. To improve reliability of the adherence assessment, we used composite scale with an acceptable value of Cronbach's alpha coefficient (0.67). The questions about adherence to NPIs did not specify whether the individual adherence was careful and permanent throughout the pandemic period, or formal and sporadic. For this reason, we did not take into account how frequently and appropriately facemasks, gloves, and sanitizers were used. We also cannot exclude the possibility of improper wearing of facemasks (covering only the mouth and leaving the nose exposed, and reusing disposable masks), which could be useless for the prevention of the infection. The definition of 'self-isolation' in the questionnaire was ambiguous, which might also influence the results.

We collected blood samples during four months between 24 February and 30 June 2021, when the infection rates were relatively steady. Regardless of the

extended period of sample collection, we can still consider the seroprevalence to be the average estimate over the period studied.

Individual antibody levels are highly dependent on the timing after exposure to the infection or vaccine. The seroprevalence could be underestimated due to low test sensitivity within two weeks following the infection or immunity waning with time passing after the disease onset or getting vaccine [27]. We cannot be certain that seronegative individuals were not previously exposed to the virus; their antibody levels may have declined with time to an undetectable level.

The analyses of predictors were performed with no adjustments for the test performance. The imperfect test performance could have attenuated the ORs toward unity because of the non-differential misclassification of the outcome status.

Finally, given the cross-sectional study design, the directions of the revealed associations cannot be interpreted unambiguously. For this reason, it was impossible to be clear on causality. Relatively small sample sizes may also limit the interpretation of our findings.

Public health importance of the findings

Level of antibodies correlating with antiviral protection as well as the proportion of the population immune to SARS-CoV-2 required to reach the herd immunity remain unknown [60]. Due to the individual-level infection-acquired or vaccine-induced immunity is short-lived and wanes rapidly over time, herd immunity might never be reached [61]. By the middle of 2021, over 70% of the population of Arkhangelsk became seropositive to SARS-CoV-2. Regardless of that, the rate of new cases registered in Arkhangelsk is still high. It can be due to the new strains of SARS-CoV-2 regularly appearing, while immunity wanes after both the infection and immunization. Nevertheless, regular seroprevalence studies should continue to be conducted in order to reveal changes in the proportion of the susceptible population [62].

We found a high rate of asymptomatic infection among unvaccinated study participants, which may play an important role in the ongoing pandemic. Previous studies showed that the proportion of asymptomatic cases could be even higher in younger adults and in children [63]. To stop the virus transmission by asymptomatic individuals, it is necessary to obtain high population coverage with vaccines, including the pediatric population [64].

Conclusion

Two third of the study population were SARS-CoV-2 seropositive in a year after the start of the pandemic in Arkhangelsk, Russia. Regular employment was positively associated with seropositive status, while smokers were less likely to be seropositive. Factors associated with low adherence to NPIs were male sex, employment and low confidence in the efficiency of NPIs. Seropositivity was not associated with adherence to NPIs during the first year of the pandemic.

ORCID

Ekaterina Krieger http://orcid.org/0000-0001-5179-5737
Ekaterina Sharashova http://orcid.org/0000-0001-9412-2065
Alexander V. Kudryavtsev http://orcid.org/0000-0001-8902-8947
Olga Samodova http://orcid.org/0000-0002-6730-6843
Anna Kontsevaya http://orcid.org/0000-0003-2062-1536
Tormod Brenn http://orcid.org/0000-0003-3717-8323
Vitaly Postoev http://orcid.org/0000-0003-4982-4169

Data availability statement

Researchers may apply for access to KYH and ESSE-RF3 data. See KYH data access regulations and instructions at https://metadata.knowyourheart.science/. Inquiries concerning ESSE-RF3 data access are to be sent to esserf2020@gmail.com. All data requests will be guided by protecting of personal information, confidentiality agreement with participants, and their informed consents.

References

[1] Byambasuren O, Cardona M, Bell K, et al. Estimating the extent of asymptomatic COVID-19 and its potential for community transmission: systematic review and meta-analysis. J Assoc Med Microbiol Infect Dis Can. 2020;5(4):223–234.

[2] Byambasuren O, Dobler CC, Bell K, et al. Comparison of seroprevalence of SARS-CoV-2 infections with cumulative and imputed COVID-19 cases: systematic review. PLoS One. 2021;16(4):e0248946.

[3] Anda EE, Braaten T, Borch KB, et al. Seroprevalence of antibodies against SARS-CoV-2 virus in the adult Norwegian population, winter 2020/2021: pre-vaccination period. Euro Surveill. 2022;27(13):2100376.

[4] Coronavirus (COVID-19) Infection Survey adftU. Infection Survey, antibody data for the UK. Available from: https://wwwonsgovuk

[5] Barchuk A, Skougarevskiy D, Titaev K, et al. Seroprevalence of SARS-CoV-2 antibodies in Saint Petersburg, Russia: a population-based study. Sci Rep. 2021;11(1):12930.

[6] Zurochka A, Dobrinina M, Zurochka V, et al. Seroprevalence of SARS-CoV-2 antibodies in symptomatic individuals Is higher than in persons who are at increased risk exposure: the results of the single-center, prospective, cross-sectional study. Vaccines. 2021;9(6):627.

[7] Popova AY, Smirnov VS, Andreeva EE, et al. SARS-CoV-2 seroprevalence structure of the Russian population during the COVID-19 pandemic. Viruses. 2021;13(8):1648.

[8] Oliveira MS, Lobo RD, Detta FP, et al. SARS-Cov-2 seroprevalence and risk factors among health care workers: estimating the risk of COVID-19 dedicated units. Am J Infect Control. 2021;49(9):1197–1199.

[9] Sannigrahi S, Pilla F, Basu B, et al. Examining the association between socio-demographic composition and COVID-19 fatalities in the European region using spatial regression approach. Sustain Cities Soc. 2020;62:102418.

[10] Middelburg RA, Rosendaal FR. COVID-19: how to make between-country comparisons. Int J Infect Dis. 2020;96:477–481.

[11] Batista SR, Souza ASS, Nogueira J, et al. Protective behaviors for COVID-19 among Brazilian adults and elderly living with multimorbidity: the ELSI-COVID-19 initiative. Cadernos De Saude Publica. 2020;36(Suppl 3):e00196120.

[12] Rogawski McQuade ET, Guertin KA, Becker L, et al. Assessment of seroprevalence of SARS-CoV-2 and risk factors associated With COVID-19 infection among outpatients in Virginia. JAMA Netw Open. 2021;4(2):e2035234.

[13] Birenbaum-Carmeli D, Chassida J. Covid-19 in Israel: socio-demographic characteristics of first wave morbidity in Jewish and Arab communities. Int J Equity Health. 2020;19(1):153.

[14] Schultze AN, Evans D, Hulme W, et al. Mortality among care home residents in England during the first and second waves of the COVID-19 pandemic: an analysis of 4.3 million adults over the age of 65. Lancet Regional Health Europe. 2022;14:100295.

[15] Kataria Y, Cole M, Duffy E, et al. Seroprevalence of SARS-CoV-2 IgG antibodies and risk factors in health care workers at an academic medical center in Boston, Massachusetts. Sci Rep. 2021;11(1):9694.

[16] Yatsyshina SB, Mamoshina MV, Elkina MA, et al. Prevalence of ARVI, influenza, and COVID-19 pathogens in individuals without symptoms of respiratory infection. J Microbiol Epidemiol Immunobiol. 2021;98(4):383–396.

[17] Adjodah D, Dinakar K, Chinazzi M, et al. Association between COVID-19 outcomes and mask mandates, adherence, and attitudes. PLoS One. 2021;16(6):e0252315.

[18] Verelst F, Hermans L, Vercruysse S, et al. SOCRATES-CoMix: a platform for timely and open-source contact mixing data during and in between COVID-19 surges and interventions in over 20 European countries. BMC Med. 2021;19(1):254.

[19] Al-Hasan A, Yim D, Khuntia J. Citizens' adherence to COVID-19 mitigation recommendations by the government: a 3-country comparative evaluation using web-based cross-sectional survey data. J Med Internet Res. 2020;22(8):e20634.

[20] Ahmed MAM, Siewe Fodjo JN, Gele AA, et al. COVID-19 in Somalia: adherence to preventive measures and evolution of the disease burden. Pathogens. 2020;9(9):735.

[21] Acurio-Páez D, Vega B, Orellana D, et al. Seroprevalence of SARS-CoV-2 infection and adherence to preventive measures in Cuenca, Ecuador, October 2020, a cross-sectional study. IJERPH. 2021;18(9):4657.

[22] Smith LE, Mottershaw AL, Egan M, et al. The impact of believing you have had COVID-19 on self-reported behaviour: cross-sectional survey. PLoS One. 2020;15(11):e0240399.

[23] Federal State Statistics Service.; 2021 Available from: https://engrosstatgovru/

[24] Buzinov eRV. The sanitary-epidemiological conditions in Arkhangelsk region in 2020: state report. 2021. p. 144.

[25] Kontsevaya AV, Shalnova SA, Drapkina OM. ESSE-RF study: epidemiology and public health promotion. Cardiovasc Ther Prev. 2021;20(5):2987.

[26] Cook S, Malyutina S, Kudryavtsev AV, et al. Know your heart: rationale, design and conduct of a cross-sectional study of cardiovascular structure, function and risk factors in 4500 men and women aged 35-69 years from two Russian cities, 2015-18. Wellcome Open Res. 2018;3:67.

[27] Kuvshinova IN, Livitskaya NI, Molodykh SV, et al. Sensitivity and specificity of reagent kits of JSC "Vector-Best" for the detection of immunoglobulins of different classes to SARS-CoV-2. Spravochnik Zaveduyushchego KDL. 2021;10:27–32.

[28] Barchuk A, Shirokov D, Sergeeva M, et al. Evaluation of the performance of SARS–CoV–2 antibody assays for a longitudinal population-based study of COVID–19 spread in St. Petersburg, Russia. J Med Virol. 2021;93(10):5846–5852.

[29] Decree of the Chief State Sanitary Doctor of the Russian Federation dated 22.052020. № 15 "On approval of sanitary and epidemiological rules SP 3.1.3597-20" prevention of new coronavirus infection (COVID-19). Russian. Moscow.

[30] Decree of the Chief State Sanitary Doctor of the Russian Federation dated 4.02 2022. № 4 "On approval of sanitary and epidemiological rules SP 3.1.3597-20" prevention of new coronavirus infection (COVID-19. Russian. Moscow.

[31] Guidelines for non-pharmaceutical interventions to reduce the impact of COVID-19 in the EU/EEA and the UK. Stockholm: ECDC; 2020.

[32] Prevention, diagnosis and treatment of new coronavirus infection (COVID-19): temporal methodological recommendations of the Ministry of health of Russian Federation. Version 3-5. Moscow: The Ministry of Health of Russian Federation; 2021.

[33] Governor's Decree № 28-u, March 17, 2020. "On the set of restrictive and other measures to counteract the spread of

new coronavirus infection (COVID-19) in the Arkhangelsk Region". "On the set of restrictive and other measures to counteract the spread of new coronavirus infection (COVID-19) in the Arkhangelsk Region." Arkhangelsk.

[34] Decree of the Russian President of March 25, 2020. No. 206 "On the establishment of non-working days in the territory of the Russian Federation." Moscow.

[35] Decree of the Russian President of April 2, 2020. No. 239 "On measures for ensuring sanitary and epidemiologic wellbeing of the population in the territory of the Russian Federation in connection with spread of new coronavirus infection (COVID-19)." Moscow.

[36] Decree of the Russian President of April 28, 2020. No. 294. "On the extension of measures to ensure the sanitary and epidemiologic wellbeing of the population in the territory of the Russian Federation in connection with spread of new coronavirus infection (COVID-19)." Moscow.

[37] Monitoring of incomes eacoRh. Moscow: HSE University; 2014. Available from: https://www.hse.ru/en/monitoring/income/

[38] Russia Household Income. 2022. Available from: https://wwwceicdatacom/en/russia/household-money-income/household-income

[39] WHO. Alcohol, heavy episodic drinking (15+) past 30 days (%), age-standardized with 95%CI; 2021. Available from: https://wwwwhoint/data/gho/data/indicators/indicator-details/GHO/alcohol-heavy-episodic-drinking-(15-)-past-30-days-(-)-age-standardized-with-95-ci.

[40] Lai CC, Wang JH, Hsueh PR. Population-based seroprevalence surveys of anti-SARS-CoV-2 antibody: an up-to-date review. Int J Infect Dis. 2020;101:314–322.

[41] Bergeri I, Whelan MG, Ware H, et al. Global SARS-CoV-2 seroprevalence from January 2020 to April 2022: a systematic review and meta-analysis of standardized population-based studies. PLoS Med. 2022;19(11):e1004107.

[42] Sempos CT, Tian L. Adjusting coronavirus prevalence estimates for laboratory test kit error. Am J Epidemiol. 2021;190(1):109–115.

[43] Henrion MY. bootComb—an R package to derive confidence intervals for combinations of independent parameter estimates. Int J Epidemiol. 2021;50(4):1071–1076.

[44] Popova AY, Andreeva EE, Babura EA, et al. Features of developing SARS-CoV-2 nucleocapsid protein population-based seroprevalence during the first wave of the COVID-19 epidemic in the Russian Federation. Russ J Infect Immun. 2021;11(2):297–323.

[45] Chen X, Huang Z, Wang J, et al. Ratio of asymptomatic COVID-19 cases among ascertained SARS-CoV-2 infections in different regions and population groups in 2020: a systematic review and meta-analysis including 130 123 infections from 241 studies. BMJ Open. 2021;11(12):e049752.

[46] Faria de Moura Villela E, López RVM, Sato APS, et al. COVID-19 outbreak in Brazil: adherence to national preventive measures and impact on people's lives, an online survey. BMC Public Health. 2021;21(1):152.

[47] Abeya SG, Barkesa SB, Sadi CG, et al. Adherence to COVID-19 preventive measures and associated factors in Oromia regional state of Ethiopia. PLoS One. 2021;16(10):e0257373.

[48] Nivette A, Ribeaud D, Murray A, et al. Non-compliance with COVID-19-related public health measures among young adults in Switzerland: insights from a longitudinal cohort study. Soc Sci Med. 2021;268:113370.

[49] Urbán R, Paksi B, Miklósi Á, et al. Non-adherence to preventive behaviours during the COVID-19 epidemic: findings from a community study. BMC Public Health. 2021;21(1):1462.

[50] Garnett C, Jackson S, Oldham M, et al. Factors associated with drinking behaviour during COVID-19 social distancing and lockdown among adults in the UK. Drug Alcohol Depend. 2021;219:108461.

[51] Emami A, Javanmardi F, Pirbonyeh N, et al. Prevalence of underlying diseases in hospitalized patients with COVID-19: a systematic review and meta-analysis. Arch Acad Emerg Med. 2020;8(1):e35.

[52] Vardavas CI, Nikitara K. COVID-19 and smoking: a systematic review of the evidence. Tob Induc Dis. 2020;18(March):20.

[53] Tizabi Y, Getachew B, Copeland RL, et al. Nicotine and the nicotinic cholinergic system in COVID-19. FEBS J. 2020;287(17):3656–3663.

[54] Qiu F, Liang CL, Liu H, et al. Impacts of cigarette smoking on immune responsiveness: up and down or upside down? Oncotarget. 2017;8(1):268–284.

[55] Wang J, Chang H, Qiao Y, et al. Angiotensin converting enzyme 2 (ACE2): Virus accomplice or host defender. bioRxiv. 2022.03.06.483197.

[56] Ma X, Li Z, Whelan MG, et al. Serology assays used in SARS-CoV-2 seroprevalence surveys worldwide: a systematic review and Meta-analysis of assay features, testing algorithms, and performance. Vaccines. 2022;10(12):2000.

[57] Meyer MJ, Yan S, Schlageter S, et al. Adjusting COVID-19 seroprevalence survey results to account for test sensitivity and specificity. Am J Epidemiol. 2022;191(4):681–688.

[58] Chen X, Chen Z, Azman AS, et al. Serological evidence of human infection with SARS-CoV-2: a systematic review and meta-analysis. Lancet Global Health. 2021;9(5):e598–e609.

[59] Krieger E, Ka Sharashova E, Postoev V, et al. Seroprevalence of SARS-Cov-2 antibodies in adults, Arkhangelsk, Russia. Emerg Infect Dis. 2022;28(2):463–465.

[60] Altmann DM, Douek DC, Boyton RJ. What policy makers need to know about COVID-19 protective immunity. Lancet. 2020;395(10236):1527–1529.

[61] Anderson RM, Vegvari C, Truscott J, et al. Challenges in creating herd immunity to SARS-CoV-2 infection by mass vaccination. Lancet. 2020;396(10263):1614–1616.

[62] Kadkhoda K. Herd immunity to COVID-19. Am J Clin Pathol. 2021;155(4):471–472.

[63] He J, Guo Y, Mao R, et al. Proportion of asymptomatic coronavirus disease 2019: a systematic review and meta-analysis. J Med Virol. 2021;93(2):820–830.

[64] Rane MS, Robertson MM, Westmoreland DA, et al. Intention to vaccinate children against COVID-19 among vaccinated and unvaccinated US parents. JAMA Pediatr. 2022;176(2):201.

Paper 2

Krieger E., Kudryavtsev A., Sharashova E., Samodova O., Kontsevaya A., Postoev V. (2024).

Spectrum of COVID-19 cases in Arkhangelsk, Northwest Russia: findings from a population-based study linking serosurvey, registry data, and self-reports of symptoms.

Spectrum of COVID-19 cases in Arkhangelsk, Northwest Russia: findings from a population-based study linking serosurvey, registry data, and self-reports of symptoms

Ekaterina Krieger[1,2¶*], Alexander V. Kudryavtsev[1,2¶], Ekaterina Sharashova[1¶], Olga Samodova[3&], Anna Kontsevaya[4&], Vitaly Postoev[5¶]

[1]Department of Community Medicine, UiT The Arctic University of Norway, Tromsø, 9019, Norway

[2]International Research Competence Centre, Northern State Medical University, Arkhangelsk, 163069, Russian Federation

[3]Department of Infectious diseases, Northern State Medical University, Arkhangelsk, 163069, Russian Federation

[4]Department of Public Health, National Medical Research Centre for Therapy and Preventive Medicine, Moscow, Russian Federation

[5]Department of Research Methodology, Northern State Medical University, Arkhangelsk, 163069, Russian Federation

*Corresponding author

E-mail: ekaterina.a.kriger@uit.no (EK)

[¶]These authors contributed equally to this work.

[&]These authors also contributed equally to this work.

Abstract

Introduction. The spectrum of COVID-19 manifestations makes it challenging to estimate the exact proportion of people who had the infection in a population, with the proportion of asymptomatic cases likely being underestimated. We aimed to assess and describe the spectrum of COVID-19 cases in a sample of adult population aged 40-74 years in Arkhangelsk, Northwest Russia, a year after the start of the pandemic.

Materials and methods. A population-based survey conducted between February 24, 2021 and June 30, 2021 with an unvaccinated sample aged 40–74 years (N=1089) combined a serological survey data, national COVID-19 case registry, and self-reported data on COVID-19 experience and symptoms. Based on the agreement between these sources, we classified the study participants as non-infected and previously infected (asymptomatic, non-hospitalized and hospitalized symptomatic) cases, and compared these groups regarding demographics, lifestyle and health characteristics.

Results. After a year of the pandemic in Arkhangelsk, 59.7% 95% confidence intervals (CI) (56.7; 62.6) of the surveyed population had had COVID-19. Among those who had been infected, symptomatic cases comprised 47.1% 95% CI (43.2; 51.0), with 8.6% 95% CI (6.6; 11.1) of them having been hospitalized. Of the asymptomatic cases, 96.2% were not captured by the healthcare system. Older age was positively associated, while smoking showed a negative association with symptomatic COVID-19. Individuals older than 65 years, and those with poor self-rated health were more likely to be hospitalized.

Conclusion. More than half of the infected individuals were not captured by the healthcare-based registry, mainly those with asymptomatic infections. COVID-19 severity was positively associated with older ageand poor self-rated health, and inversely associated with smoking. Combining different sources of surveillance data could reduce the number of unidentified asymptomatic cases and enhance surveillance for emerging infections.

Key words: COVID-19, SARS-CoV-2, asymptomatic infection, Russia

Introduction

The novel coronavirus disease 2019 (COVID-19) caused by the severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) is a highly contagious disease, which spread around the world despite the infection control efforts. Infected individuals may develop a range of COVID-19 manifestations, from asymptomatic or mild to severe disease and death. This range of the disease severity is termed as the spectrum of COVID-19 [1]. The COVID-19 severity depends on the SARS-CoV-2 variant, the viral load, and host factors such as sex, age, and chronic diseases [2-4].

The variety of individual responses to the infection makes it challenging to determine the exact proportion of the infected people in a population. Based on the iceberg concept of infectious diseases, detected COVID-19 cases represent only the emerging tip of the iceberg, while asymptomatic cases comprise the invisible base [5].

When the coronavirus pandemic was declared, most countries implemented non-pharmaceutical interventions to prevent further disease transmission [6]. During the initial stage of the pandemic in Russia, nationwide paid non-working days were established [7-9]. Subsequently, regional governments implemented and updated non-pharmaceutical interventions based on the local epidemiological situation, in accordance with Presidential Decree #316, dated May 11, 2020 [10]. These measures generally aligned with the recommendations of the Russian Federal Service for Surveillance on Consumer Rights Protection and Human Wellbeing (Rospotrebnadzor) and included mandatory wearing facemasks in public settings or transport, travel restrictions, limitations on commercial activities, and the closure of educational institutions and non-essential services [11]. Individuals aged 65 years and older, as well as those with chronic conditions such as diabetes, respiratory and cardiovascular issues, chronic kidney diseases, neoplasms in any location, organ and tissue transplant recipients, and pregnant women, were advised to self-isolate [11]. Other COVID-19 non-pharmaceutical interventions, implemented in Russia as well as worldwide, included testing strategies, contact tracing, and isolation of infected people [6].

The Polymerase Chain Reaction (PCR) test was used to detect ongoing infections by capturing SARS-CoV-2 genetic material in nasopharyngeal swabs [12-13]. Since infected individuals may transmit SARS-CoV-2 regardless of their symptomatic status, preventing the transmission to the susceptible population depends on a number of individuals with the ongoing infection who remained unidentified [14-15]. The detection proportion refers to the number of positive tests among all tests done and reflects testing strategy effectiveness. During the first six months of the

pandemic, the global detection proportion was estimated to be 9.8% (range 1.2-66.8%) [16]. The highest detection proportions were in Australia (66.8%) and Iceland (60.3%). In Russia, the detection proportion was estimated to be 25.4% [16].

The data about all diagnosed COVID-19 cases and COVID-19 vaccine recipients were accumulated in the Federal Registry of COVID-19 Patients (later in this paper referred to as the COVID-19 case registry or the case registry) and the Federal Registry of the Vaccinated Against COVID-19 (later in this paper referred to as the vaccination registry) respectively, as regulated by Russian Government Decree #373 dated March 31, 2020 [17]. The registries gathered COVID-19-related data from patients' electronic health records contained by information systems of governmental non-military health services. The data accumulated in the case registry were based on positive results of the PCR tests, rather than clinical symptoms [12, 18]. Patients with severe disease were more likely to be PCR-tested, while asymptomatic cases and those with minor symptoms might have failed to seek medical advice or undergo testing [19]. Due to a limited capacity of PCR testing and the high load on the healthcare system at the earlier stages of the pandemic, testing was restricted to patients with signs of pneumonia and individuals with other respiratory symptoms in the higher risk groups (healthcare workers, close contacts of a confirmed case, individuals older than 65 years) [20]. Voluntary testing and testing of travelers were not covered by compulsory health insurance. Thus, other symptomatic COVID-19 cases seeking healthcare might have remained undiagnosed. Besides, PCR testing could have been unreliable due to untimely or incorrect specimen collection and limited viral replication in the epithelial cells of the upper respiratory tract [21-22]. Therefore, recorded COVID-19 cases represented a subset of the actual number of the infected. Little is known about the spectrum of COVID-19 cases and the proportion of those who remained asymptomatic.

The World Health Organization recommended collecting self-reported COVID-19 history and performing a serological survey in a sample of the general population to support the retrospective assessment of the spread of the COVID-19 pandemic through the population [13]. Linkage of the data generated by different surveillance activities can lend insight into the spectrum of COVID-19 cases. The data accumulated in the case registry were described as providing accurate and reliable information on COVID-19 status in those captured by the healthcare system [12, 18]. In addition, self-reported survey data can provide a better capture of symptomatic cases, including those who did not seek medical advice [13]. A serological survey is used to detect those who were previously infected regardless of their symptomatic status and to assess the scale of the infection spread [23].

A previous population-based study in Arkhangelsk, Northwest Russia, conducted a year after the start of the pandemic, revealed a SARS-CoV-2 seroprevalence of 65.1% 95% CI (62.5; 67.6), with

associations found between seropositive status and regular employment as well as smoking [24]. The current study combined the serological survey results, the healthcare data on the recorded cases (COVID-19 case registry), and the self-reported survey data on COVID-19 experience and symptoms. The study **aimed** to assess and describe the spectrum of COVID-19 cases in the sample of adult population a year after the start of the pandemic.

Materials and methods

Study population

The study was a satellite of a national multi-center survey of the prevalence of cardiovascular risk factors among adults aged 35-74 years in Russian regions (ESSE-RF3). The Arkhangelsk part of ESSE-RF3 was conducted by the Northern State Medical University (NSMU) between February 24, 2021 and June 30, 2021, a year after the start of the COVID-19 pandemic. The ESSE-RF3 study sample in Arkhangelsk consisted of the participants of an earlier cross-sectional study of cardiovascular diseases – Know Your Heart (KYH) [25]. The KYH study was conducted in 2015-2017 on a random sample (N=2380) of Arkhangelsk population aged 35-69 years. The KYH participants were selected from four districts in Arkhangelsk using an anonymized list of addresses provided by the regional health insurance fund. Trained interviewers visited randomly chosen addresses, inviting individuals of the corresponding age and sex to participate. The overall participation rate was 68% among those invited. Based on the informed consent obtained from the KYH participants, they were invited to the ESSE-RF3 study by use of personal contact information. After the exclusion of those who had not consented to be contacted with new invitations (N=56), those who had died prior to the launch of ESSE-RF3 (N=61, including four deaths due to COVID-19), and those older than 74 years (N=5) at the launch of ESSE-RF3 as exceeding the age span of the study, the list of invitees included 2258 KYH participants aged 40-74 years. With response rate of 59.7%, 1348 KYH participants attended the ESSE-RF3 study and comprised the current study sample. Two participants with incomplete ESSE-RF3 survey, 14 participants with equivocal serological test results, 224 individuals who self-reported receiving at least one dose of vaccine against SARS-CoV-2 and had a record in the vaccination registry, and 19 participants with discordant data on their vaccination status were excluded (Fig 1). Therefore, the analyzed study sample comprised 1089 unvaccinated participants with definitive serological test results.
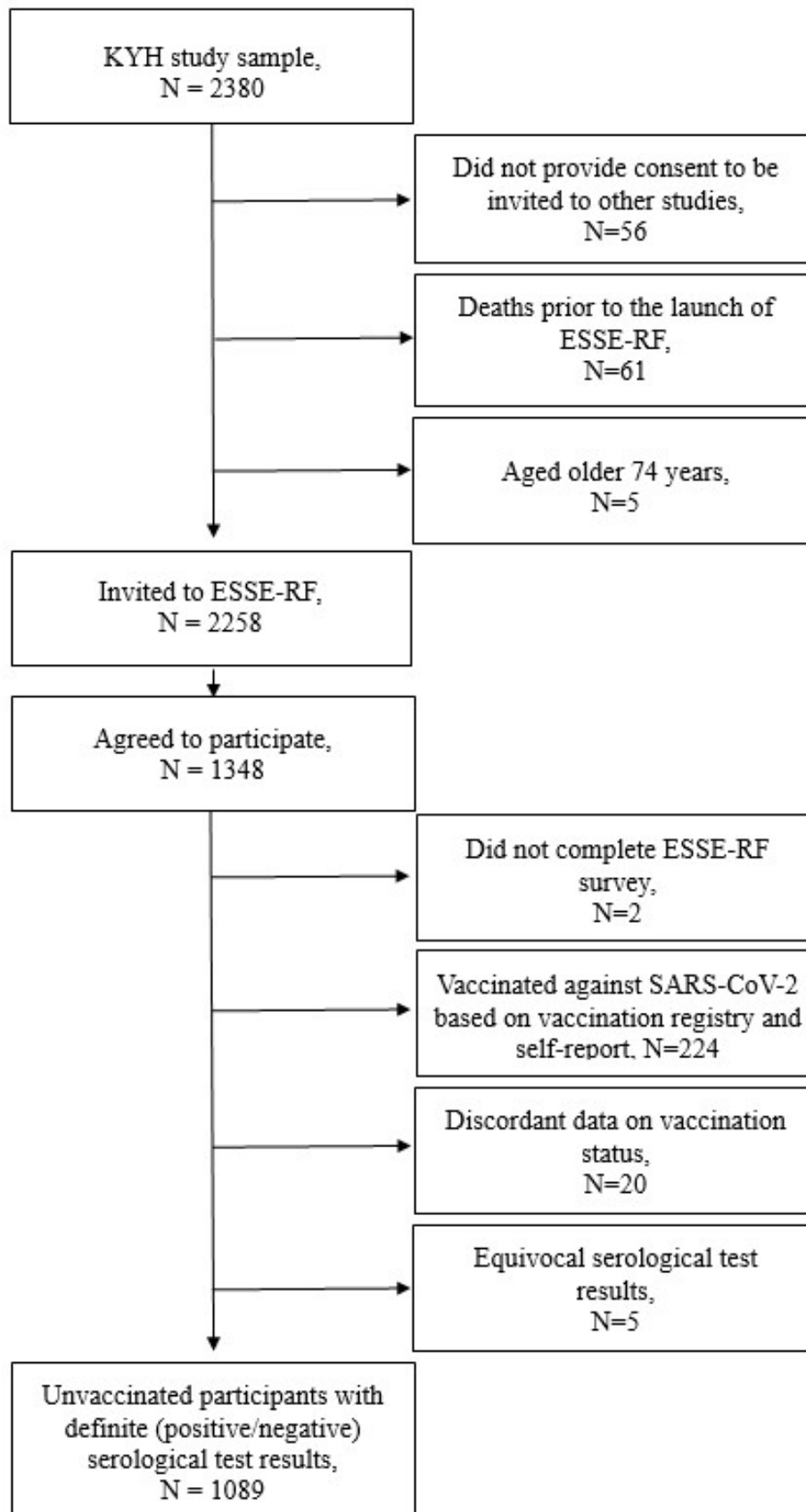
Fig.1. Flow chart of the study.

Ethics approval and consent to participate

Ethical approval for the KYH study was received from the ethics committees of the London School of Hygiene & Tropical Medicine (approval number 8808, February 24, 2015) and NSMU, Arkhangelsk (approval number 01/01-15, January 27, 2015). Ethical approval for the follow-up of the KYH using electronic health records was received from the ethics committee of NSMU, Arkhangelsk, Russia (approval number 01/04-19, April 24, 2019). Ethical approval for the ESSE-RF-3 was received from the ethics committee of the National Research Centre for Therapy and Preventive Medicine, Moscow, Russia (approval number 01-01/20, February 04, 2020) and the ethics committee of NSMU, Arkhangelsk, Russia (approval number 01/02-21, February 17, 2021). Ethical approval for the study of COVID-19-related issues was received from the ethics committees of NSMU, Arkhangelsk, Russia (approval number 01/02-21, February 17, 2021) and by the Regional Committee for Medical and Health Research Ethics (approval number 339397 received December 7, 2021).

All participants of the ESSE-RF study took part in the KYH study, and at the time of joining the KYH provided a written consent to disclose their medical and other health-related records for research purposes under the confidentiality condition and to be invited to other studies. The data linkage was performed by the Arkhangelsk Regional Medical Information Analytical Center (MIAC) in accordance with the NSMU-MIAC confidentiality agreement based on the informed consents obtained from the participants, as well as legal and ethical approvals. The dataset accessed and analyzed for this paper did not contain personal identifiers.

Data collection

The data collection for ESSE-RF3 in Arkhangelsk involved a standardized interviewer-administered questionnaire survey, blood sample collection, and health examination. Trained interviewers collected data on participants' demographic and lifestyle characteristics, self-reported diseases and vaccination. Trained physicians and nurses conducted health examination, including a medical interview (history and symptoms of diseases, medication use), blood sample taking, instrumental and functional measurements.

The following ESSE-RF3 variables were used in analyses: sex (male/female), age (5-year bands), higher education (yes/no), chronic health conditions (hypertension, diabetes, abdominal obesity), self-rated health, smoking, and frequency of heavy alcohol drinking. The participants with systolic blood pressure ≥ 140 mmHg and/or diastolic blood pressure ≥ 90 mmHg and/or using antihypertensive medications were treated as having hypertension. Diabetes was defined as self-reported diagnosis (including the type of diabetes) and/or self-reported use of antidiabetic

medications and/or having glycated hemoglobin (HbA1c) ≥6.5%. Abdominal obesity was defined as waist circumference ≥94 cm for men and ≥80 cm for women. Self-rated health was measured on a 0-100 visual analogue scale orientated upwards (best) from below (worst) and then divided into two levels: lower or equal to median (poor health) and greater than median (good health). Smoking status was classified as non-smoker, former, and current smoker. Heavy alcohol drinking was defined as consuming of 60 or more grams of pure alcohol on a single occasion [26]. The frequency of heavy drinking during the previous 12 months was classified as never, once a week or less often, and twice a week or more often.

The history of prior COVID-19 was assessed by asking the participants "Did you have COVID-19 during the previous 12 months?" (yes/no/don't know). We treated the "don't know" response as a negative answer. Those who self-reported having had COVID-19 were asked to indicate the date of the disease onset and to answer the questions "Did you seek any medical advice?" and "Were you hospitalized?" with a yes/no answer options. They were also asked whether they experienced the following symptoms: fatigue, fever, headache, myalgia / arthralgia, loss of smell and taste (anosmia / ageusia), cough, dyspnea, sore throat, rhinitis, diarrhea, nausea / vomiting, and rash. We coded each symptom as 1 (present) or 0 (absent).

As the data were collected after the start of the COVID-19 vaccination campaign, the question "Have you received a vaccine for COVID-19?" was added to the ESSE-RF3 survey. Those responding positively were asked to provide information on the number of the doses (one or two) and the vaccination dates.

For the semi-quantitative detection of immunoglobulins G to spike glycoprotein of SARS-CoV-2, blood serum samples were analyzed using Vector Best enzyme linked immunosorbent assay (D-5501 SARS-CoV-2-IgG-EIA-BEST, Russia), a method with the 89% sensitivity and the 100% specificity reported by an independent test-performance study [27].

The self-reported survey data on prior COVID-19 and serological test results were linked with data from the COVID-19 case registry and the vaccination registry [17]. The following information was collected from the COVID-19 case registry: all COVID-19-related visits (outpatient and inpatient) prior to the participation in ESSE-RF3, including final diagnoses, date of the disease onset and the outcome. The study participants were treated as registered COVID-19 cases if they had records of COVID-19 diagnoses with codes U07.1 (COVID-19, virus identified) (N=218, 96.5%) or U07.2 (COVID-19, virus not identified) (N=8, 3.5%) according to the International Classification of Diseases, 10th revision. The code U07.2 was assigned to a clinical or epidemiological diagnosis of COVID-19 where laboratory confirmation was inconclusive or not available.

We used the vaccination registry to obtain the dates and the number of vaccination doses received by each vaccinated participant. The vaccination registry data were compared with the self-reported vaccination status. Self-reported vaccination details matched the vaccination registry records for 224 participants. Two participants self-reported no vaccination but had it recorded in the registry, and 18 self-reported vaccinations but had no corresponding records in the registry.

Spectrum of COVID-19 cases

Based on the agreement between the serological test results, the COVID-19 case registry data and the self-reported survey data, the study participants were divided into those previously infected and non-infected (also referred to as 'infected cases' and 'non-infected cases' for short). The previously infected participants were classified as symptomatic or asymptomatic. Symptomatic cases were defined as having positive serological tests or records in the COVID-19 case registry and reporting COVID-19 symptoms at the survey. Asymptomatic cases were defined as having positive serological tests or COVID-19 record in the COVID-19 case registry, but reporting no symptoms of COVID-19 in the past. The participants having neither a positive serological test nor a record in the COVID-19 case registry were considered previously non-infected no matter what symptoms they reported. The infected cases were further grouped into non-hospitalized and hospitalized. The cases were defined as hospitalized if they were recorded as inpatients in the COVID-19 case registry (N=53). Four participants who had had the infection and reported COVID-19 hospitalization with the dates and duration of the hospital stay, but had no records in the COVID-19 case registry were treated as hospitalized symptomatic cases.

Statistical analysis

Absolute numbers and relative frequencies were presented for categorical data, and medians (first and third quartile) for continuous data. Confidence intervals (CIs) for proportions were calculated using Wilson's procedure. We used the Pearson Chi-squared test to compare the frequency of non-infected, asymptomatic, non-hospitalized and hospitalized symptomatic cases among different groups of the participants. Binary logistic regression was used to examine factors associated with being a symptomatic COVID-19 case on a subset of infected cases (N=650). Factors linked to hospitalization with COVID-19 were explored using binary logistic regression on a subset of symptomatic COVID-19 cases (N=306). We employed directed acyclic graphs (DAGs) to justify the selection of covariates for adjustment (32). Findings were presented as crude odds ratios (ORs) with 95% CIs and ORs adjusted for demographic variables (age, sex, higher education) and lifestyle characteristics (smoking, frequency of heavy drinking), all

selected from DAGs as potential confounders for all covariates (later in this paper referred to as the demographics and lifestyle factors). Adjusted ORs for smoking status were calculated comparing former smokers versus never-smokers, current smokers versus never-smokers, and ever-smokers (i.e., current and former smokers combined) versus never-smokers. The threshold for significance tests was 0.05. The Statistical Package for the Social Science SPSS version 24.0 (SPSS Inc, Chicago, Il) was used for the data analysis.

Results

The median age of the study sample was 55 (48; 63) years, 61.1% were women.
In total, 645 (59.2%) tested positive for immunoglobulins G to SARS-CoV-2, 220 (20.2%) had the infection documented in the case registry, and 316 (29.0%) self-reported COVID-19 symptoms.

According to the linkage of case data from the three sources, the proportion of the participants who had had the infection in the studied sample was 59.7% 95% CI (56.7; 62.6) (N=650). Asymptomatic COVID-19 cases comprised 52.9% 95% CI (49.0; 56.8) (N=344) of all the infected (31.6% 95% CI (28.9; 34.5) of the study population). Of them, 96.2% (N=331) were not captured by the case registry. The detailed information on the grouping of the participants is shown in Figure 2.
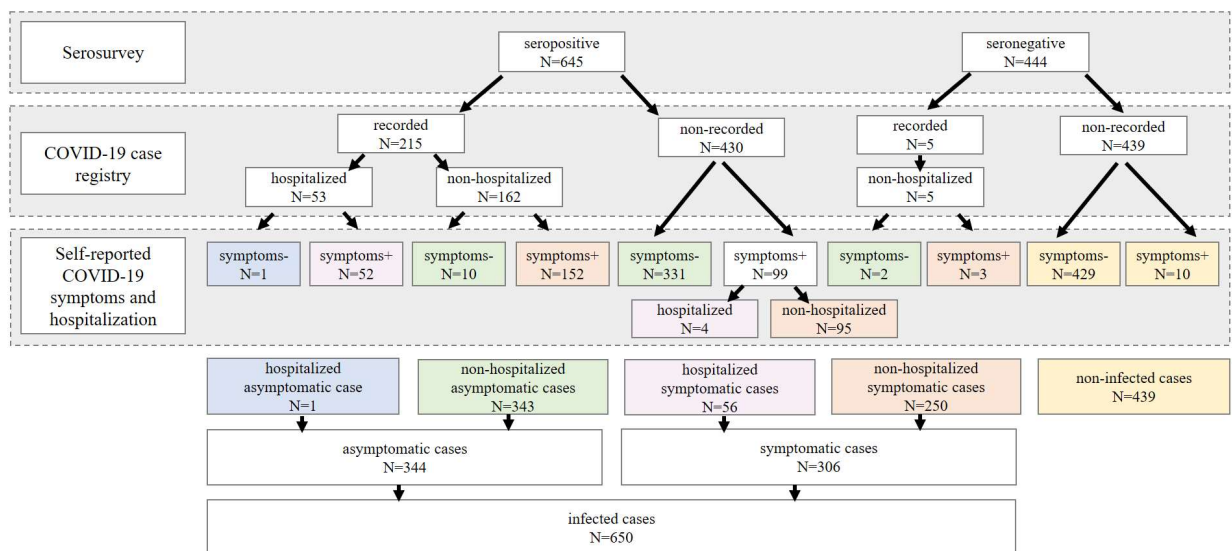
Fig. 2. Classification of the study participants based on the linkage of the serosurvey, the COVID-19 case registry data and the self-reported survey data, N=1089.

Less than half (47.1%, 95% CI (43.2; 51.0), N=306) of the participants who had had the infection reported having had symptoms. Of them, 18.3% (N=56) had been admitted to hospital (8.6% 95% CI (6.6; 11.1) of all previously infected).

The majority of the symptomatic cases (88.6%, N=272) occurred between September 2020 and March 2021 with a peak in December 2020 (Fig 3). Almost a third of these cases (32.2%, N=99) had no COVID-19 record in the case registry, despite the fact that most of them (88.9%, N=88) had sought medical advice according to the survey data. Four of the non-recorded symptomatic cases (4.0%) self-reported having been hospitalized and indicated the date of the disease onset and the duration of the hospital stay.
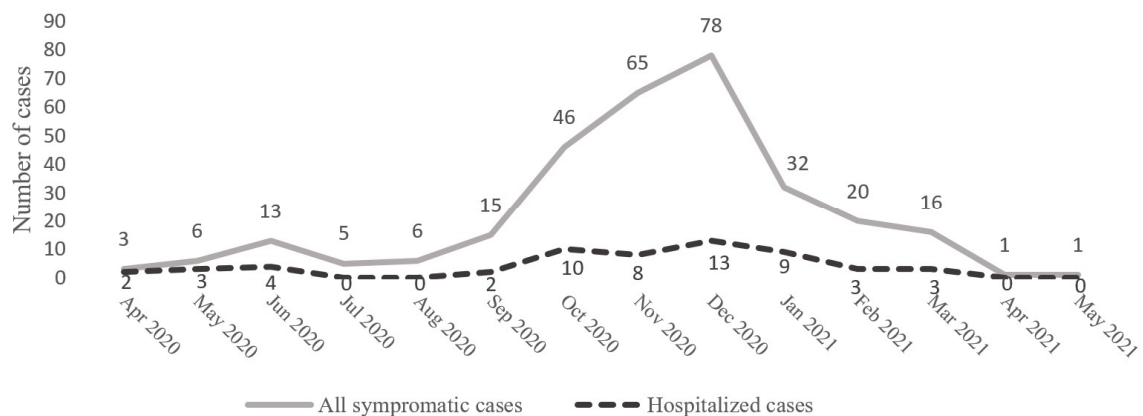
Fig. 3. Distribution of the symptomatic cases over time.

The number of symptoms reported by the symptomatic cases (N=306) varied from 1 to 12 with the median of 6 (5; 7). The most prevalent symptoms were fatigue (87.0%), fever (84.7%), headache (66.1%), myalgia / arthralgia (63.8%), anosmia / ageusia (63.5%), cough (51.1%), rhinitis (43.6%), dyspnea (41.7%), and sore throat (41.7%). Other symptoms included diarrhea (19.2%), nausea / vomiting (12.7%), and rash (3.3%). Most of the participants self-reported three or more symptoms (N=285, 92.8%), and the remaining 7.2% were oligosymptomatic.

The median number of self-reported symptoms was similar for the hospitalized and non-hospitalized symptomatic cases. Compared with the non-hospitalized patients, the hospitalized cases had higher frequencies of fever and dyspnea. Anosmia / ageusia and rhinitis were more often reported by the non-hospitalized symptomatic COVID-19 cases compared to the hospitalized ones (Fig 4).
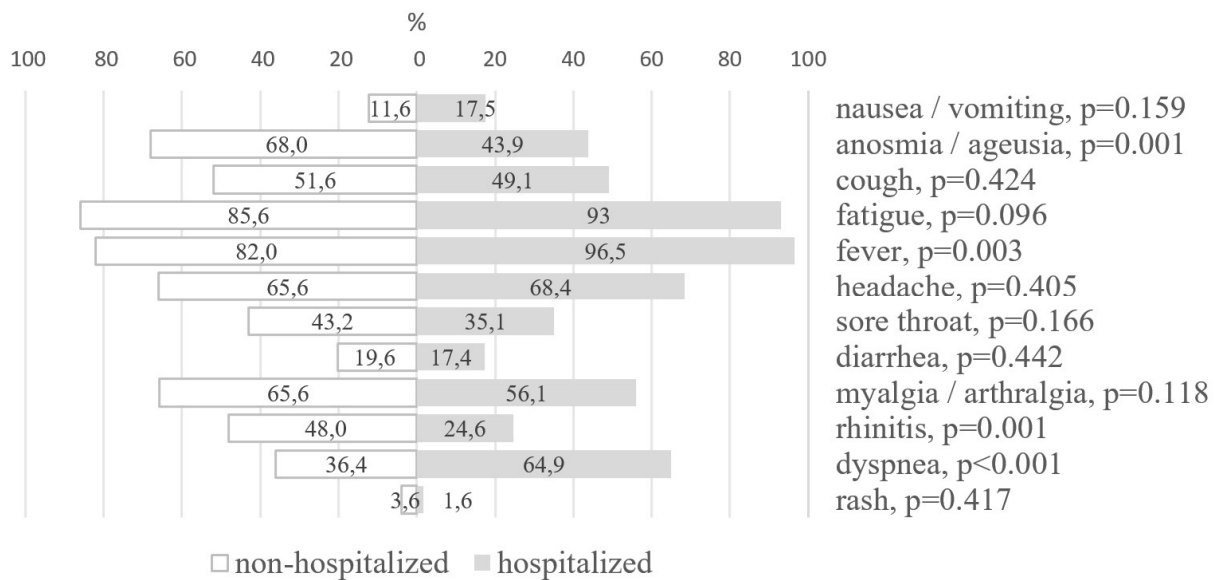
Fig. 4. Symptoms self-reported by the non-hospitalized and hospitalized symptomatic patients, N=306. p for Pearson's Chi-squared test.

Table 1 presents the demographics, lifestyle, and health characteristics of non-infected, asymptomatic and symptomatic COVID-19 cases. The proportions of both asymptomatic cases and hospitalized symptomatic cases were higher in men compared to women (Table 1). The proportion of infected cases was higher in younger age groups. Among infected individuals, the frequency of symptomatic cases as well as the percentage of hospitalizations increased with age. Symptomatic participants with poor self-rated health, and those with diabetes had a higher proportion of hospitalizations. Current smokers and individuals with heavy drinking habits were less likely infected and less likely symptomatic cases. A quarter of former smokers had been admitted to the hospital, while no current smokers were found among the hospitalized cases.

Table 1. Characteristics of non-infected and previously infected (asymptomatic and symptomatic COVID-19 cases)

Variable	Total N	Non-infected cases, N (%)	Asymptomatic cases, N (%)	Symptomatic non-hospitalized, N (%)	Symptomatic hospitalized cases, N (%)	p-value*
Sex						
Men	424	174 (41.0)	147 (34.6)	77 (18.1)	26 (6.3)	0.015
Women	665	265 (39.9)	197 (29.6)	173 (26.0)	30 (4.5)	
Age						
40-44 years	138	54 (39.1)	56 (40.6)	26 (18.8)	2 (1.5)	<0.001
45-54 years	386	132 (34.2)	142 (36.8)	102 (26.4)	10 (2.6)	
55-64 years	356	157 (44.1)	101 (28.4)	79 (22.2)	19 (5.3)	
65-74 years	209	96 (45.9)	45 (21.5)	43 (20.6)	25 (12.0)	

Higher education						
No	620	251 (40.5)	203 (32.7)	129 (20.8)	37 (6.0)	0.143
Yes	469	188 (40.0)	141 (30.1)	121 (25.8)	19 (4.1)	
Hypertension						
No	487	179 (36.8)	169 (34.7)	116 (23.8)	23 (4.7)	0.109
Yes	602	260 (43.2)	175 (29.1)	134 (22.2)	33 (5.5)	
Diabetes						
No	952	384 (40.3)	305 (32.0)	222 (23.4)	41 (4.3)	0.011
Yes	137	55 (40.1)	39 (28.5)	28 (20.4)	15 (11.0)	
Abdominal obesity						
No	476	189 (39.7)	153 (32.2)	112 (23.5)	22 (4.6)	0.867
Yes	613	250 (40.8)	191 (31.2)	138 (22.5)	34 (5.5)	
Self-rated health						
0-50%	212	88 (41.5)	64 (30.2)	39 (18.4)	21 (9.9)	0.003
51-100%	877	351 (40.0)	280 (31.9)	211 (24.1)	35 (4.0)	
Smoking status						
Non-smoker	621	226 (36.4)	196 (31.6)	164 (26.4)	35 (5.6)	<0.001
Former smoker	287	111 (38.7)	91 (31.7)	64 (22.3)	21 (7.3)	
Current smoker	181	102 (56.4)	57 (31.5)	22 (12.1)	0 (0.0)	
Frequency of heavy drinking						
Never	663	262 (39.5)	196 (29.6)	169 (25.5)	36 (5.4)	0.033
Once a week or less often	379	152 (40.1)	134 (35.4)	77 (20.3)	16 (4.2)	
Twice a week or more often	47	25 (53.2)	14 (29.8)	4 (8.5)	4 (8.5)	
Total	1089	439	344	250	56	

*p for Pearson's Chi-squared test

Male sex, current smoking, and frequency of heavy drinking (once a week or less often) showed negative associations with symptomatic COVID-19 in the unadjusted models (Table 2). Older participants were more likely to be symptomatic COVID-19 cases. After adjustment for demographic and lifestyle factors, age was positively associated and smoking was negatively associated with symptomatic COVID-19. Adjusted estimates for ever-smokers (i.e. current and former smokers combined) versus never-smokers were 0.95 (95% CI 0.66;1.36) for symptomatic status.

Table 2. Factors associated with being symptomatic COVID-19 cases, N=650

Variable	OR crude (95% CI)	OR adj. (95% CI)[1]
Sex		
Women	reference	reference
Men	0.68 (0.49; 0.94)	0.74 (0.50; 1.08)

Variable	OR crude (95% CI)	OR adj. (95% CI)[1]
Age		
40-44 years	reference	reference
45-54 years	1.58 (0.94; 2.65)	1.56 (0.92; 2.64)
55-64 years	1.94 (1.14; 3.30)	1.93 (1.10; 3.36)
65-74 years	3.02 (1.68; 5.45)	2.96 (1.58; 5.53)
Higher education		
No	reference	reference
Yes	1.21 (0.89; 1.66)	1.30 (0.93; 1.81)
Hypertension		
No	reference	reference
Yes	1.16 (0.85; 1.58)	1.14 (0.81; 1.59)
Diabetes		
No	reference	reference
Yes	1.28 (0.80; 2.03)	1.10 (0.68; 1.78)
Abdominal obesity		
No	reference	reference
Yes	1.03 (0.75; 1.40)	0.92 (0.66; 1.28)
Self-rated health		
0-50%	1.07 (0.72; 1.58)	0.92 (0.61; 1.40)
51-100%	reference	reference
Smoking status		
Non-smoker	reference	reference
Former smoker	0.92 (0.65; 1.31)	1.15 (0.78; 1.70)
Current smoker	0.38 (0.22; 0.65)	0.55 (0.31; 0.97)
Frequency of heavy drinking		
Never	reference	reference
Once a week or less often	0.66 (0.48; 0.92)	0.91 (0.62; 1.34)
Twice a week or more often	0.55 (0.22; 1.33)	0.88 (0.34; 2.32)

[1]adjusted for sex, age, higher education, smoking, frequency of heavy drinking

In univariate analyses, men, older participants, individuals with poor self-rated health, diabetes, and those with heavy drinking habits (two times a week or more) were more likely to be hospitalized (Table 3). After adjustment for the demographics and lifestyle factors, age older than 65 yearsand poor self-rated health were associated with hospitalization. Adjusted OR for ever-smokers (i.e. current and former smokers combined) versus never-smokers was 0.95 (95% CI 0.45;1.99) for hospitalization.

Table 3. Factors associated with being hospitalized symptomatic COVID-19 case, N=306

Variable	OR crude (95% CI)	OR adj. (95% CI)[1]
Sex		
Women	reference	reference
Men	1.95 (1.08; 3.51)	1.93 (0.88; 4.17)
Age		

40-44 years	reference	reference
45-54 years	1.28 (0.26; 6.18)	1.28 (0.26; 6.36)
55-64 years	3.13 (0.68; 14.34)	2.82 (0.60; 13.40)
65-74 years	7.56 (1.65; 34.57)	6.99 (1.45; 33.72)
Higher education		
No	reference	reference
Yes	0.55 (0.30; 1.00)	0.76 (0.38; 1.51)
Hypertension		
No	reference	reference
Yes	1.24 (0.69; 2.24)	0.66 (0.33; 1.32)
Diabetes		
No	reference	reference
Yes	2.90 (1.43; 5.90)	2.22 (0.99; 5.00)
Abdominal obesity		
No	reference	reference
Yes	1.25 (0.69; 2.27)	1.15 (0.59; 2.24)
Self-rated health		
0-50%	3.25 (0.71; 6.16)	2.51 (1.23; 5.14)
51-100%	reference	reference
Smoking status		
Non-smoker	reference	reference
Former smoker	1.54 (0.83; 2.84)	1.25 (0.58; 2.66)
Current smoker	-	-
Frequency of heavy drinking		
Never	reference	reference
Once a week or less often	0.97 (0.51; 1.96)	1.03 (0.46; 2.27)
Twice a week or more often	4.69 (1.12; 19.65)	3.56 (0.63; 20.18)

[1]adjusted for sex, age, higher education, smoking, frequency of heavy drinking

Discussion

In this study, we presented the spectrum and characteristics of COVID-19 cases in a sample of a population aged 40-74 in Arkhangelsk, Northwest Russia, a year after the start of the pandemic. The proportion of the participants who had had the infection was 59.7% 95% CI (56.7; 62.6) (N=650). Of them, 47.1% 95% CI (43.2; 51.0) (N=306) developed symptoms. Almost all of the symptomatic cases (96.4%, N=295) had sought medical advice and 56 (18.3%) of them had been admitted to hospital. More than half (52.9% 95% CI (49.0; 56.8)) of the participants who had had the infection had been asymptomatic. Almost all of the asymptomatic cases (96.2%) had been undiagnosed and had not been included in the COVID-19 case registry.

The first case of COVID-19 emerged in Arkhangelsk on 17 March 2020. During the following six months of the pandemic, the Wuhan strain of SARS-CoV-2 predominantly circulated in the

city. By the end of the summer of 2020, the incidence of COVID-19 had declined and stabilized [23]. Consequently, the local governmental restrictions, implemented on March 18, 2020, were relaxed starting from September 2020. Educational institutions resumed offline activities, and commercial and recreational activities were reinstated with the mandatory enforcement of a 1.5-meter distance rule, face mask wearing, hand antiseptic use, and non-contact thermometry [28]. The relaxation of restrictive measures coincided with the emergence of the Delta variant of SARS-CoV-2, with the first COVID-19 cases caused by the Delta variant registered in August 2020 [29]. Our findings show that the easing of restrictions, coupled with the Delta variant, boosted transmission (Fig. 3), which is in line with another study conducted in Saint Petersburg, Northwest Russia [30]. It is likely that the proportion of infected cases in Arkhangelsk was higher compared to other Russian cities at approximately the same time [30-31]. However, the differences could be attributed to the non-uniform sampling approaches used in the studies and the related variations in the characteristics of the studied samples [32].

In our study, 18.3% of the symptomatic cases (8.6% 95% CI (6.6; 11.1) of all previously infected) had been hospitalized. These findings are higher than those of Menachemi et al., who reported the hospitalization rate in the USA varied from 0.4% for those younger than 40 years to 9.2% for those older than 60 years [33]. In France, the proportion of the infected who needed hospitalization for COVID-19 from March 2020 to January 2021 was estimated to be 0.4% in individuals aged 20–29 years to 17.6% in those aged 70-89 years [34]. The high proportion of the participants who had been treated in hospital in our study could be partly explained by the common recommendation at early stages of the pandemic in Russia to isolate all detected COVID-19 patients by admitting them to hospital [20]. The percentage of hospitalized COVID-19 cases varies significantly across studies, depending on hospital admission policies in different countries at different time periods after the start of the pandemic and characteristics of the studied populations, such as age and comorbidities [20]. These inconsistencies make it difficult to compare the results of these studies.

In our study, the proportion of the hospitalized cases was higher in men compared to women, which is in line with other studies [35-38]. However, the association between sex and hospitalization did not persist after adjusting for age, higher education, smoking, frequency of heavy drinking. This suggests that the difference in hospitalization between men and women could be attributed to the varying prevalence of these factors between the sexes. Older age is a well-established factor associated with the disease severity [39-40]. In support of this, we found a higher proportion of symptomatic cases and hospitalized cases among older study participants. The symptomatic COVID-19 patients with poor self-rated health were more likely to be

hospitalized compared to those with good self-rated health. Our findings are in agreement with the results obtained by others, which demonstrated that comorbidities augmented the COVID-19 severity [41-43].

The highly specific COVID-19 symptom such as loss of smell and taste was more frequently reported by the non-hospitalized symptomatic cases compared to the hospitalized cases. Another study also found a higher frequency of the loss of smell and taste among individuals with mild COVID-19 [44], while other authors reported a positive correlation between these symptoms and COVID-19 severity [45]. Fever and dyspnea were more common in the hospitalized cases, which is in line with other studies [46-47].

The proportion of participants who remained non-infected a year after the start of the pandemic was higher among older age groups, possibly due to non-pharmaceutical interventions, primarily self-isolation, specifically targeted at older individuals and those with chronic diseases [11, 28]. This may have contributed to a potentially lower overall severity of symptomatic COVID-19.

The proportion of the asymptomatic cases (52.9% 95% CI (49.0; 56.8) among all the infected exceeded the overall estimates published in the recent systematic reviews (40.5-44.1%) [48-49]. Nevertheless, these estimates are influenced by the age groups included and should not be directly compared [32].

Limited testing access during the first months of the pandemic may have resulted in undiagnosed mild symptomatic infections among participants. Unaware of their COVID-19 status, these individuals might have erroneously reported having had no COVID-19 symptoms, potentially leading to misclassification as asymptomatic cases. The study design lacks symptom-related questions for those answering "No" to "Did you have COVID-19 in the previous 12 months?", possibly resulting to an underestimation of the number of symptomatic COVID-19 cases. Although COVID-19 symptoms are non-specific, and incorporating questions about previous cold-like symptoms would have relatively low specificity, even with sensitivity close to 100% [50]. Many of these additionally revealed symptomatic cases would be false positives, falling into the category of non-hospitalized cases, resulting in differential misclassification.

Our findings are in line with earlier reports of a lower probability of symptomatic cases in younger age groups [14, 51]. Men were more prone to self-report being symptom-free, which is consistent with a number of studies [52-53]. The associations between symptomatic status and sex disappeared after adjusting for age, education, smoking and frequency of heavy drinking. Thus, other factors, such as the varying prevalence of higher education, smoking and heavy drinking between the sexes, could contribute to the observed difference in symptomatic status. Like in other publications, we found a lower proportion of current smokers among symptomatic

cases [54-56]. Current smokers had half the odds of being symptomatic COVID-19 cases compared to non-smokers, with none of the current smokers being hospitalized. Former smokers and ever-smokers (i.e. current and former smokers combined) had similar odds of being symptomatic COVID-19 cases as well as being hospitalized compared to non-smokers, which contradicts recent meta-analyses showing higher odds of severity and hospitalization among former and ever-smokers [57]. Former smokers may have quit smoking before or during the pandemic due to chronic diseases known to be risk factors for severe infection. Consequently, the overall health status of former smokers may be worse compared to that of current smokers [57].

The cross-sectional study design, along with possible biases, may influence the observed associations between smoking and disease severity. Simultaneous measurement of exposure and outcome may result in reverse causality. However, it is unlikely that adults aged 40-74 years initiated smoking after contracting COVID-19 [58]. Additionally, the possibility of reverse causality has not been suggested by other reports. We classified the study participants as non-smokers, former smokers, and current smokers. A more precise measurement of smoking status, including duration (years) or frequency (number of cigarettes) of smoking, may shed light on the studied association. As participant classification relied on self-reported smoking status, it could be influenced by bias related to the desire to undervalue socially unwelcome behaviors (misclassification) [59-60]. Current smokers might be less prone to self-report symptoms and seek healthcare less frequently [61]. Consequently, they may be less likely to be tested for SARS-CoV-2 and to be captured in the COVID-19 case registry. Higher frequency of false-negative PCR tests in smokers may have contributed to the low prevalence of smokers among patients with COVID-19 [59, 62]. These factors could lead to an underrepresentation of current smokers among symptomatic COVID-19 cases.

As most studies [63-65] report a positive association between smoking and COVID-19 severity, the low frequency of symptomatic and hospitalized cases among current smokers in our study may also be explained by the possibility that they experienced severe illness and died before the initiation of the current study or refused to participate due to poor health conditions or post-COVID symptoms. The assumptions that tobacco smoke compounds reduce susceptibility to SARS-CoV-2 are speculative [66-68]. Further studies are required to properly explore how smoking status affects the course of COVID-19.

On the contrary to other studies, we found a lower proportion of symptomatic cases among those who reported frequent heavy drinking [35, 55]. However, after adjusting for age, sex, higher education, and smoking, frequency of heavy drinking was no longer associated with

symptomatic status. At the same time, it should be noted that that regular alcohol drinking as well as smoking may decrease the probability of seeking medical advice or reporting symptoms among adults who experienced a disease or physical discomfort, as was shown in the previous study [61].

To our knowledge, this is the first population-based study in Russia estimating the COVID-19 spectrum. Combining COVID-19-related data from different sources enabled us to make reliable estimates of the proportion of the infected cases as well as the proportions of asymptomatic and symptomatic COVID-19 cases. However, our findings should be interpreted in the light of some limitations.

First of all, the sample of the current study was a resurveyed subsample of the previous population-based KYH study aged 40-74 years and may not be fully representative of the target population of Arkhangelsk residents of the same age. The response rate of 59.7% among the invited KYH participants could be a source of a non-response bias. To address this issue, we compared the participants of the current study to the entire KYH-based sampling frame on the key demographic variables (sex, age, education) [69]. Compared to the original KYH sample, resurvey participants had similar proportions of men and were younger on average at the time of inclusion in the KYH study. This is because the current study sample did not include participants of older age who died between the studies or could not participate in the current study due to severe illnesses. A slightly higher proportion of participants in the resurvey had higher education, possibly due to older participants who died or dropped out having a lower proportion of individuals with higher education. Some individuals may have attained higher education between the studies, and those with higher education may have increased health awareness and willingness to engage in research. These differences between the current study sample and the original KYH sample were unlikely to significantly impact the main results and conclusions. We used the cross-sectional study design and therefore were not able to explore causal associations. Information bias could also impact the observed associations. Using data from multiple sources requires consideration of possible limitations and imperfect completeness of each of the sources used. For instance, our study results could be influenced by the registry data completeness and reliability [70]. Yet, we discovered only four cases showing discrepancy between the self-reported and the COVID-19 case registry data on hospitalization, which indicates only a negligible coverage deficiency of the registry with respect to hospitalized cases. Retrospective self-reported responses about symptoms may have been subjected to recall bias and non-differential misclassification of symptomatic status due to the difficulties of distinguishing between COVID-19 and other respiratory tract infections [71]. Those who had had COVID-19 but had not been tested might misclassify themselves as non-infected.

The number of the infected cases could have been underestimated due to the imperfect performance of the tests used as well as the antibodies waning over time after the infection [23]. To address this, the SARS-CoV-2 seroprevalence estimates described in our previous paper were adjusted for the serological test performance, and were increased by 7.9% compared to non-adjusted estimates [24]. Adjustments for test performance were not made in this study because the calculation of proportions of infected individuals considered data from various sources beyond the results of serological tests. However, the proportion of the infected cases and the proportion of the seropositive participants were almost the same in our study, with only five COVID-19 cases recorded in the case registry but tested negative for the SARS-CoV-2 antibodies. Thus, the underestimation might only be very small and could be due to the imperfect test sensitivity.

Alternately, we could overestimate the number of the seropositive people based on their cross-reactive immunity acquired after the infection caused by beta-coronaviruses with the molecular structure similar to SARS-CoV-2 [23, 72]. However, even though cross-reactivity was a concern early in the pandemic, there is limited evidence available. The standard procedure for validating the specificity of tests involves using pre-pandemic serum samples to ensure that the test accurately identifies true positives related to the virus and minimizes the likelihood of false positives caused by cross-reactivity [27]. Besides, some factors that could not be assessed in our study, such as the subtype of SARS-CoV-2 or the viral load, may influence the symptomatic status [73].

The study added to the knowledge of the COVID-19 spectrum in a Russian sample with a high proportion of the previously infected, but it may be of a limited novelty with respect to symptomatic COVID-19 cases. The proportion of the people who had been infected and remained asymptomatic was a key question in understanding the extent of COVID-19 in the study setting. Estimating the proportion of asymptomatic COVID-19 cases is critical for calculating key epidemiological characteristics, quantifying the cumulative incidence of the infection. An exact evaluation of the proportion and a better understanding of characteristics of asymptomatic cases could help to retrospectively assess the strategies which were implemented to control the COVID-19 pandemic. The identification of a considerable number of undetected cases underscores the potential lessons to be learned from the existing surveillance and contact tracking strategy, suggesting areas for improvement in infection control measures. Several studies have shown that asymptomatic individuals shed similar quantities of virus to the symptomatic persons [14-15]. Thus, being unaware of their disease status, the asymptomatic individuals could unknowingly transmit the virus to others [74].

Conclusion

A year after the start of the COVID-19 pandemic in Arkhangelsk, Northwest Russia, 59.7% 95% CI (56.7; 62.6) of the surveyed adult population aged 40-74 years had been infected by SARS-CoV-2. Symptomatic cases comprised 47.1% 95% CI (43.2; 51.0) of the total infected, and 8.6% 95% CI (6.6; 11.1) of those previously infected were hospitalized. Our findings indicated a high proportion of asymptomatic cases that remained undetected by the healthcare system. The asymptomatic COVID-19 cases were unaware that they had been infected and might have continued their usual activities spreading the infection to others. This could have resulted in the rapid COVID-19 transmission and unsuccessful disease control.

Since the asymptomatic COVID-19 patients are difficult to be diagnosed, a wider testing of high-risk populations should be performed regardless of symptoms to improve the control strategies. Combining different surveillance approaches could prevent future outbreaks by capturing silent infection spread.

Acknowledgements

References

1.   Zeng H, Ma Y, Zhou Z, Liu W, Huang P, Jiang M, et al. Spectrum and Clinical Characteristics of Symptomatic and Asymptomatic Coronavirus Disease 2019 (COVID-19) With and Without Pneumonia. Frontiers in medicine. 2021;8:645-651. doi: 10.3389/fmed.2021.645651.

2.   Çelik I, Öztürk R. From asymptomatic to critical illness: decoding various clinical stages of COVID-19. Turkish journal of medical sciences. 2021;51(Si-1):3284-3300. doi: 10.3906/sag-2107-137.

3.   Pijls BG, Jolani S, Atherley A, Derckx RT, Dijkstra JIR, Franssen GHL, et al. Demographic risk factors for COVID-19 infection, severity, ICU admission and death: a meta-analysis of 59 studies. BMJ open. 2021;11(1):e044640. doi: 10.1136/bmjopen-2020-044640.

4.   Sannigrahi S, Pilla F, Basu B, Basu AS, Molter A. Examining the association between socio-demographic composition and COVID-19 fatalities in the European region using spatial regression approach. Sustainable cities and society. 2020;62:102418. doi: 10.1016/j.scs.2020.102418.

5.   Shubin M., Virtanen M., Toikkanen S., Lyytikäinen O., & Auranen K. Estimating the burden of A(H1N1)pdm09 influenza in Finland during two seasons. Epidemiology & Infection. 2014; 142(5): 964-974. doi:10.1017/S0950268813002537

6. Balmford B, Annan JD, Hargreaves JC, Altoè M, Bateman IJ. Cross-Country Comparisons of Covid-19: Policy, Politics and the Price of Life. Environmental & resource economics. 2020;76(4):525-551. doi: 10.1007/s10640-020-00466-5.

7. Decree of the Russian President of March 25, 2020. No. 206 "On the establishment of non-working days in the territory of the Russian Federation." Moscow.

8. Decree of the Russian President of April 2, 2020. No. 239 "On measures for ensuring sanitary and epidemiologic wellbeing of the population in the territory of the Russian Federation in connection with spread of new coronavirus infection (COVID-19)." Moscow.

9. Decree of the Russian President of April 28, 2020. No. 294. "On the extension of measures to ensure the sanitary and epidemiologic wellbeing of the population in the territory of the Russian Federation in connection with spread of new coronavirus infection (COVID-19)." Moscow.

10. Decree of the Russian President of May 11, 2020. No. 316. "On definition of the procedure for prolongation of measures to ensure sanitary-epidemiological well-being of the population in the subjects of the Russian Federation in connection with the spread of new coronavirus infection (COVID-19)" Moscow.

11. Methodological Recommendations MP 3.1.0178-20 "Definition of a Set of Measures, as Well as Indicators That Are the Basis for the Phased Removal of Restrictive Measures in the Conditions of the Epidemic Spread of COVID-19".

12.   Krantz SG, Rao A. Level of underreporting including underdiagnosis before the first peak of COVID-19 in various countries: Preliminary retrospective results based on wavelets and deterministic modeling. Infection control and hospital epidemiology. 2020;41(7):857-859. doi: 10.1017/ice.2020.116

13.   Public health surveillance for COVID-19. Interim guidance. [cited 2023 May 31]. Available from: https://www.who.int/publications/i/item/WHO-2019-nCoV-SurveillanceGuidance-2022.2

14.   Sah P, Fitzpatrick MC, Zimmer CF, Abdollahi E, Juden-Kelly L, Moghadas SM, et al. Asymptomatic SARS-CoV-2 infection: A systematic review and meta-analysis. Proceedings of the National Academy of Sciences of the United States of America. 2021;118(34). doi: 10.1073/pnas.2109229118.

15.   Zou L, Ruan F, Huang M, Liang L, Huang H, Hong Z, et al. SARS-CoV-2 Viral Load in Upper Respiratory Specimens of Infected Patients. The New England journal of medicine. 2020;382(12):1177-1179. doi: 10.1056/NEJMc2001737.

16.   Kuster AC, Overgaard HJ. A novel comprehensive metric to assess effectiveness of COVID-19 testing: Inter-country comparison and association with geography, government, and policy response. PloS one. 2021;16(3):e0248176. doi: 10.1371/journal.pone.0248176.

17.   Decree of the Russian Government №373 of March 31, 2020. ″About approval of Provisional rules of accounting of information for the purpose of prevention of spread of new coronavirus infection (COVID-19)″.

18.   Clark-Boucher D, Boss J, Salvatore M, Smith JA, Fritsche LG, Mukherjee B. Assessing the added value of linking electronic health records to improve the prediction of self-reported COVID-19 testing and diagnosis. PloS one. 2022;17(7):e0269017. doi: 10.1371/journal.pone.0269017.

19.   Byambasuren O, Dobler CC, Bell K, Rojas DP, Clark J, McLaws ML, et al. Comparison of seroprevalence of SARS-CoV-2 infections with cumulative and imputed COVID-19 cases: Systematic review. PloS one. 2021;16(4):e0248946. doi: 10.1371/journal.pone.0248946.

20.   Prevention, diagnosis and treatment of new coronavirus infection (COVID-19): temporal methodological recommendations of the Ministry of health of Russian Federation. Version 3-5. Moscow: The Ministry of Health of Russian Federation; 2021.

21.   Böger B, Fachi MM, Vilhena RO, Cobre AF, Tonin FS, Pontarolo R. Systematic review with meta-analysis of the accuracy of diagnostic tests for COVID-19. American journal of infection control. 2021;49(1):21-29. doi: 10.1016/j.ajic.2020.07.011.

22.   Kucirka LM, Lauer SA, Laeyendecker O, Boon D, Lessler J. Variation in False-Negative Rate of Reverse Transcriptase Polymerase Chain Reaction-Based SARS-CoV-2 Tests by Time Since Exposure. Annals of internal medicine. 2020;173(4):262-7. doi: 10.7326/M20-1495.

23.   Krajewski R, Gołębiowska J, Makuch S, Mazur G, Agrawal S. Update on serologic testing in COVID-19. Clinica chimica acta; international journal of clinical chemistry. 2020;510:746-50. doi: 10.1016/j.cca.2020.09.015.

24.   Krieger E, Sharashova E, Kudryavtsev AV, Samodova O, Kontsevaya A, Brenn T, et al. COVID-19: seroprevalence and adherence to preventive measures in Arkhangelsk, Northwest Russia. Infectious diseases (London, England). 2023;55(5):316-327. doi: 10.1080/23744235.2023.2179660.

25.   Cook S, Malyutina S, Kudryavtsev AV, Averina M, Bobrova N, Boytsov S, et al. Know Your Heart: Rationale, design and conduct of a cross-sectional study of cardiovascular structure, function and risk factors in 4500 men and women aged 35-69 years from two Russian cities, 2015-18. Wellcome open research. 2018;3:67. doi: 10.12688/wellcomeopenres.14619.3.

26.   WHO. Alcohol, heavy episodic drinking (15þ) past 30 days (%), age-standardized with 95%CI; 2021. Available from: https://wwwwhoint/data/gho/data/indicators/indicatordetails/GHO/alcohol-heavy-episodic-drinking-(15-)-past-30-days-(-)-age-standardized-with-95-ci.

27.   Barchuk A, Shirokov D, Sergeeva M, Tursun Zade R, Dudkina O, Tychkova V, et al. Evaluation of the performance of SARS-CoV-2 antibody assays for a longitudinal population-based study of COVID-19 spread in St. Petersburg, Russia. Journal of medical virology. 2021;93(10):5846-5852. doi: 10.1002/jmv.27126.

28. Decree of the Arkhangelsk Governor March 17, 2020. No. 28-u. "On the set of restrictive and other measures to counteract the spread of new coronavirus infection (COVID-19) in the Arkhangelsk Region".

29."On the State of Sanitary and Epidemiological Well-being of the Population in the Arkhangelsk Region in 2021: State Report / Edited by T.I. Nosovsky – Arkhangelsk, 2022. – 146 p."

30.   Barchuk A, Skougarevskiy D, Kouprianov A, Shirokov D, Dudkina O, Tursun-Zade R, et al. COVID-19 pandemic in Saint Petersburg, Russia: Combining population-based serological study and surveillance data. PloS one. 2022;17(6):e0266945. doi: 10.1371/journal.pone.0266945.

31.   Popova AY, Smirnov VS, Andreeva EE, Babura EA, Balakhonov SV, Bashketova NS, et al. SARS-CoV-2 Seroprevalence Structure of the Russian Population during the COVID-19 Pandemic. Viruses. 2021;13(8). doi: 10.3390/v13081648.

32.   Accorsi, E.K., Qiu, X., Rumpler, E. et al. How to detect and reduce potential sources of biases in studies of SARS-CoV-2 and COVID-19. Eur J Epidemiol. 2021; 36: 179–196. doi:10.1007/s10654-021-00727-7

33.   Menachemi N, Dixon BE, Wools-Kaloustian KK, Yiannoutsos CT, Halverson PK. How Many SARS-CoV-2-Infected People Require Hospitalization? Using Random Sample Testing to Better Inform Preparedness Efforts. Journal of public health management and practice: JPHMP. 2021;27(3):246-50. doi: 10.1097/PHH.0000000000001331.

34.   Hozé N, Paireau J, Lapidus N, Tran Kiem C, Salje H, Severi G, et al. Monitoring the proportion of the population infected by SARS-CoV-2 using age-stratified hospitalisation and serological data: a modelling study. Lancet Public Health. 2021;6(6):e408-e415. doi: 10.1016/S2468-2667(21)00064-5.

35.   Malundo AFG, Abad CLR, Salamat MSS, Sandejas JCM, Planta JEG, Poblete JB, et al. Clinical characteristics of patients with asymptomatic and symptomatic COVID-19 admitted to a tertiary referral centre in the Philippines. IJID regions. 2022;2:204-211. doi: 10.1016/j.ijregi.2022.02.002.

36.   Conti P, Younes A. Coronavirus COV-19/SARS-CoV-2 affects women less than men: clinical response to viral infection. Journal of biological regulators and homeostatic agents. 2020;34(2):339-343. doi: 10.23812/Editorial-Conti-3.

37.   Abate BB, Kassie AM, Kassaw MW, Aragie TG, Masresha SA. Sex difference in coronavirus disease (COVID-19): a systematic review and meta-analysis. BMJ open. 2020;10(10):e040129. doi: 10.1136/bmjopen-2020-040129.

38.   Jin JM, Bai P, He W, Wu F, Liu XF, Han DM, et al. Gender Differences in Patients With COVID-19: Focus on Severity and Mortality. Frontiers in public health. 2020;8:152. doi: 10.3389/fpubh.2020.00152.

39.   Li Y, Shi J, Xia J, Duan J, Chen L, Yu X, et al. Asymptomatic and Symptomatic Patients With Non-severe Coronavirus Disease (COVID-19) Have Similar Clinical Features and Virological Courses: A Retrospective Single Center Study. Frontiers in microbiology. 2020;11:1570. doi: 10.3389/fmicb.2020.01570.

40. Singhal S, Kumar P, Singh S, Saha S, Dey AB. Clinical features and outcomes of COVID-19 in older adults: a systematic review and meta-analysis. BMC geriatrics. 2021;21(1):321. doi: 10.1186/s12877-021-02261-3.

41.   Mal P, Mukherjee T, Upadhyay AK, Mohanty S, Pattnaik AK. Connecting the dots between inflammatory cascades of obesity and COVID-19 in light of mortal consequences-a review. Environmental science and pollution research international. 2022;29(38):57040-53. doi: 10.1007/s11356-022-21461-x.

42.   Petersen A, Bressem K, Albrecht J, Thieß HM, Vahldiek J, Hamm B, et al. The role of visceral adiposity in the severity of COVID-19: Highlights from a unicenter cross-sectional pilot study in Germany. Metabolism: clinical and experimental. 2020;110:154317. doi:

43.   Mubarik S, Liu X, Eshak ES, Liu K, Liu Q, Wang F, et al. The Association of Hypertension With the Severity of and Mortality From the COVID-19 in the Early Stage of the Epidemic in Wuhan, China: A Multicenter Retrospective Cohort Study. Frontiers in medicine. 2021; 8: 623608. doi: 10.1016/j.metabol.2020.154317.

44.   Kim GU, Kim MJ, Ra SH, Lee J, Bae S, Jung J, et al. Clinical characteristics of asymptomatic and symptomatic patients with mild COVID-19. Clinical microbiology and infection: the official publication of the European Society of Clinical Microbiology and Infectious Diseases. 2020;26(7):948.e1-.e3. doi:10.1016/j.cmi.2020.04.040.

45.   Mazzatenta A, Neri G, D'Ardes D, De Luca C, Marinari S, Porreca E, et al. Smell and Taste in Severe CoViD-19: Self-Reported vs. Testing. Frontiers in medicine. 2020;7:589409. doi: 10.3389/fmed.2020.589409.

46.   Vahey GM, Marshall KE, McDonald E, Martin SW, Tate JE, Midgley CM, et al. Symptom Profiles and Progression in Hospitalized and Nonhospitalized Patients with Coronavirus Disease, Colorado, USA, 2020. Emerg Infect Dis. 2021;27(2):385-395. doi: 10.3201/eid2702.203729.

47.   Schäfer E, Scheer C, Saljé K, Fritz A, Kohlmann T, Hübner NO, et al. Course of disease and risk factors for hospitalization in outpatients with a SARS-CoV-2 infection. Scientific reports. 2022;12(1):7249. doi: 10.1038/s41598-022-11103-0.

48.   Ma Q, Liu J, Liu Q, Kang L, Liu R, Jing W, et al. Global Percentage of Asymptomatic SARS-CoV-2 Infections Among the Tested Population and Individuals With Confirmed COVID-19 Diagnosis: A Systematic Review and Meta-analysis. JAMA network open. 2021;4(12):e2137257-e. doi: 10.1001/jamanetworkopen.2021.37257.

49.   Wang B, Andraweera P, Elliott S, Mohammed H, Lassi Z, Twigger A, et al. Asymptomatic SARS-CoV-2 infection by age: A systematic review and meta-analysis. medRxiv. 2022:2022.05.05.22274697. doi: 10.1097/INF.0000000000003791.

50. Struyf T, Deeks JJ, Dinnes J, Takwoingi Y, Davenport C, et al., Cochrane COVID-19 Diagnostic Test Accuracy Group. Signs and symptoms to determine if a patient presenting in primary care or hospital outpatient settings has COVID-19. Cochrane Database of Systematic Reviews 2022, 5: CD013665. doi: 10.1002/14651858.CD013665.pub3

51.   Methi F, Madslien EH. Lower transmissibility of SARS-CoV-2 among asymptomatic cases: evidence from contact tracing data in Oslo, Norway. BMC medicine. 2022;20(1):427.

52.   Vachon CM, Norman AD, Prasad K, Jensen D, Schaeferle GM, Vierling KL, et al. Rates of Asymptomatic COVID-19 Infection and Associated Factors in Olmsted County, Minnesota, in the Prevaccination Era. Mayo Clinic proceedings Innovations, quality & outcomes. 2022;6(6):605-617. doi: 10.1016/j.mayocpiqo.2022.10.001.

53.   Pinto Saravia V. Sociodemographic Differences in COVID-19 Self-Reported Symptoms by Ethnicity and Older Adults in Bolivia. Journal of population ageing. 2022;15(3):811-41. doi: 10.1007/s12062-022-09383-5.

54.   Miyara M, Tubach F, Pourcher V, Morélot-Panzini C, Pernet J, Haroche J, et al. Lower Rate of Daily Smokers With Symptomatic COVID-19: A Monocentric Self-Report of Smoking Habit Study. Frontiers in medicine. 2021;8:668995. doi: 10.3389/fmed.2021.668995.

55.   Saurabh S, Verma MK, Gautam V, Kumar N, Jain V, Goel AD, et al. Tobacco, alcohol use and other risk factors for developing symptomatic COVID-19 vs asymptomatic SARS-CoV-2 infection: a case–control study from western Rajasthan, India. Transactions of The Royal Society of Tropical Medicine and Hygiene. 2021;115(7):820-831. doi: 10.1093/trstmh/traa172.

56. Naimeh M, Afsaneh A, Zahra A, Ali D, Abdullah B, Marzieh RR. Smoking as a Risk or Protective Factor in Developing and Severity of COVID-19? Journal of Acute Medicine. 2023;13(3):114-121. doi: 10.6705/j.jacme.202309_13(3).0003.

57. Gallus S, Scala M, Possenti I, Jarach CM, Clancy L, Fernandez E, et al. The role of smoking in COVID-19 progression: a comprehensive meta-analysis. European Respiratory Review. 2023;32(167):220191. doi: 10.1183/16000617.0191-2022.

58. Jackson SE, Tattan-Birch H, Shahab L, Beard E, Brown J. Have there been sustained impacts of the COVID-19 pandemic on trends in smoking prevalence, uptake, quitting, use of treatment, and relapse? A monthly population study in England, 2017-2022. BMC Medicine. 2023;21(1):474. doi: 10.1186/s12916-023-03157-2.

59. Farsalinos K, Barbouni A, Niaura R. Systematic review of the prevalence of current smoking among hospitalized COVID-19 patients in China: could nicotine be a therapeutic option? Intern Emerg Med 2020;15:845-852. doi: 10.1007/s11739-020-02355-7.

60. Latkin, C.A.; Edwards, C.; Davey-Rothwell, M.A.; Tobin, K.E. The relationship between social desirability bias and self-reports of health, substance use, and social network factors among urban substance users in Baltimore, Maryland. Addictive Behaviors. 2017; 73: 133–136. doi: 10.1016/j.addbeh.2017.05.005.

61. Li C, Sun J. The impact of current smoking, regular drinking, and physical inactivity on health care-seeking behavior in China. BMC health services research. 2022;22(1):52. doi: 10.1186/s12913-022-07462-z.

62. Salerno S, Zhao Z, Prabhu Sankar S, Salvatore M, Gu T, Fritsche LG, et al. Patterns of repeated diagnostic testing for COVID-19 in relation to patient characteristics and outcomes. Journal of Internal Medicine. 2021;289(5):726-737. doi: 10.1111/joim.13213.

63. Zhang H, Ma S, Han T, Qu G, Cheng C, Uy JP, et al. Association of smoking history with severe and critical outcomes in COVID-19 patients: A systemic review and meta-analysis. European journal of integrative medicine. 2021;43:101313. doi: 10.1016/j.eujim.2021.101313.

64. Gülsen A, Yigitbas BA, Uslu B, Drömann D, Kilinc O. The Effect of Smoking on COVID-19 Symptom Severity: Systematic Review and Meta-Analysis. Pulmonary medicine. 2020;2020:7590207. doi: 10.1155/2020/7590207.

65. Reddy RK, Charles WN, Sklavounos A, Dutt A, Seed PT, Khajuria A. The effect of smoking on COVID-19 severity: A systematic review and meta-analysis. Journal of medical virology. 2021;93(2):1045-1056. doi: 10.1002/jmv.26389.

66. Changeux JP, Amoura Z, Rey FA, Miyara M. A nicotinic hypothesis for Covid-19 with preventive and therapeutic implications. Comptes rendus biologies. 2020;343(1):33-39. doi: 10.5802/crbiol.8.

67. Liu W, Tao ZW, Wang L, Yuan ML, Liu K, Zhou L, et al. Analysis of factors associated with disease outcomes in hospitalized patients with 2019 novel coronavirus disease. Chinese medical journal. 2020;133(9):1032-1038. doi: 10.1097/CM9.0000000000000775.

68. Tanimoto K, Hirota K, Fukazawa T, Matsuo Y, Nomura T, Tanuza N, et al. Inhibiting SARS-CoV-2 infection in vitro by suppressing its receptor, angiotensin-converting enzyme 2, via aryl-hydrocarbon receptor signal. Scientific reports. 2021;11(1):16629. doi: 10.1038/s41598-021-96109-w.

69. Krieger E, Kudryavtsev A, Sharashova E, Postoev V, Belova N, Shagrov L, et al. Seroprevalence of SARS-Cov-2 Antibodies in Adults, Arkhangelsk, Russia. Emerg Infect Dis. 2022;28(2):463-465. doi: 10.3201/eid2802.211640.

70. Binkheder S, Asiri MA, Altowayan KW, Alshehri TM, Alzarie MF, Aldekhyyel RN, et al. Real-World Evidence of COVID-19 Patients' Data Quality in the Electronic Health Records. Healthcare (Basel, Switzerland). 2021;9(12). doi: 10.3390/healthcare9121648.

71. Jiang C, Yao X, Zhao Y, Wu J, Huang P, Pan C, et al. Comparative review of respiratory diseases caused by coronaviruses and influenza A viruses during epidemic season. Microbes and infection. 2020;22(6-7):236-244. doi: 10.1016/j.micinf.2020.05.005.

72. Le Bert N, Tan AT, Kunasegaran K, Tham CYL, Hafezi M, Chia A, et al. SARS-CoV-2-specific T cell immunity in cases of COVID-19 and SARS, and uninfected controls. Nature. 2020;584(7821):457-62. doi: 10.1038/s41586-020-2550-z.

73. Zhang S, Guo M, Wu F, Xiong N, Ma Y, Wang Z, et al. Factors associated with asymptomatic infection in health-care workers with severe acute respiratory syndrome coronavirus 2 infection in Wuhan, China: a multicentre retrospective cohort study. Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases. 2020;26(12):1670-1675. doi: 10.1016/j.cmi.2020.08.038.

74. Li G, Li W, He X, Cao Y. Asymptomatic and Presymptomatic Infectors: Hidden Sources of Coronavirus Disease 2019 (COVID-19). Clinical infectious diseases: an official publication of the Infectious Diseases Society of America. 2020;71(8):2018. doi: 10.1093/cid/ciaa418.

Paper 3

Risk factors for all-cause mortality during the COVID-19 pandemic compared with the pre-pandemic period in an adult population of Arkhangelsk, Russia.

Risk factors for all-cause mortality during the COVID-19 pandemic compared with the pre-pandemic period in an adult population of Arkhangelsk, Russia

Ekaterina Krieger[1,2*], Alexander V. Kudryavtsev[2], Ekaterina Sharashova[1], Olga Samodova[3], Vitaly Postoev[4]

[1]Department of Community Medicine, UiT The Arctic University of Norway, N-9037 Tromsø, Norway

[2]International Research Competence Centre, Northern State Medical University, Troitsky Ave., 51, Arkhangelsk, Russia 163069

[3]Department of Infectious Diseases, Northern State Medical University, Troitsky Ave., 51, Arkhangelsk, Russia 163069

[4]Department of Research Methodology, Northern State Medical University, Troitsky Ave., 51, Arkhangelsk, Russia 163069

*email: ekaterina.a.kriger@uit.no

Abstract

We investigated and compared mortality rates and risk factors for pre-pandemic and pandemic all-cause mortality in a population-based cohort of men and women in Arkhangelsk, Russia. A prospective cohort study enrolled 2,324 participants aged 35 to 69 years between 2015 and 2017. All participants were followed up for all-cause deaths using the mortality registry. Mortality rates per 1000 person-years were calculated for men and women in the pre-pandemic and pandemic periods. Cox regression models were used to investigate demographic, lifestyle, and health characteristics associated with increased risk of death in both periods. During the pandemic, age-standardized all-cause mortality increased in women, but minor change was observed in men. Older age, smoking, and diabetes were associated with a higher risk of all-cause death in both periods and for both sexes. In women, higher risk during the pandemic was associated with obesity, angina, and elevated cystatin C levels. In men, asthma and elevated hs-Troponin T levels increased the risk of death during the pandemic, while elevated hs-CRP and NT-proBNP levels were associated with higher risk in both periods. Targeted preventive interventions for men and women with specific risk factors can be implemented during potential future infectious disease outbreaks.

Key words: COVID-19, SARS-CoV-2, all-cause mortality, risk factors, Russia

Introduction

On December 31, 2019, China reported several cases of severe pneumonia with an unknown origin. The causative agent was identified on January 7, 2020, and officially named severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) [1]. On March 11, 2020, the World Health Organization (WHO) declared the novel coronavirus (COVID-19) outbreak as a pandemic. The pandemic resulted in an overall increase in mortality, encompassing causes both directly related to COVID-19 and unrelated to the infection, which represent indirect consequences of the pandemic [2, 3].

Pre-existing chronic diseases heighten susceptibility to SARS-CoV-2 and the risk of COVID-19 complications, such as acute respiratory distress syndrome, sepsis, septic shock, and multiorgan failure [4]. Conversely, COVID-19 can worsen pre-existing chronic conditions, particularly cardiovascular diseases, by potentially damaging myocardial cells and being pathogenically linked to thrombovascular events [5-7]. Death resulting from any of these complications could be attributed to the virus, considered as a direct effect of the pandemic [8].

The indirect consequences of the pandemic, including limited access to timely medical care due to overwhelmed healthcare systems, hindered routine and screening services, disproportionately affecting individuals with chronic diseases [2, 9]. Additionally, these individuals could hesitate to seek medical care for non-urgent conditions out of fear of contracting SARS-CoV-2 [10, 11]. People older than 65 years, as well as those with chronic health conditions, were informed about the risks to their health associated with COVID-19 and officially advised to self-isolate [12]. The limited access to healthcare, coupled with reduced screening activities, mandatory isolation, and avoidance of seeking healthcare, could result in delayed diagnoses and treatment of conditions such as acute cardiovascular events, neoplasms, or the progression of chronic diseases, potentially contributing to the increased mortality [9, 13, 14].

The mortality statistics, based on a single underlying cause of death, may not fully capture the complexity of factors or processes that lead to death in individuals with chronic diseases, including the role of the virus in this sequence [15, 16]. Limited access to testing for laboratory confirmation of infection could also introduce misclassification of deaths related and unrelated to COVID-19, especially during the early stages of the pandemic [17, 18].

Given the multiple effects of different factors on mortality during the pandemic, counting deaths from all causes combined provides a more comprehensive approach to measuring the overall mortality effects of the pandemic, avoiding issues of attributing deaths specifically to COVID-19 [2]. Assessment of the impact of demographic and lifestyle factors, pre-existing chronic diseases and the blood-based biomarkers on the risk of death during the pandemic compared with the pre-

pandemic period could shed light on pathways of how the pandemic affected mortality, enabling targeted preventive efforts.

The study **aimed** to investigate and compare mortality rates and risk factors for pre-pandemic and pandemic all-cause deaths within a population-based cohort of adults in Arkhangelsk, Russia.

Methods

Study design and study population

A prospective cohort study was carried out involving participants of Know Your Heart (KYH) study of cardiovascular diseases, which has been described previously [19]. KYH study sample is a population-based cohort of 2357 participants (68% response rate) aged 35 to 69 years at enrolment from 2015 to 2017 [19]. All participants underwent interviews conducted by trained interviewers, physical examinations and provided blood samples for further biomarker analysis. Follow-up period for a study participant started after the inclusion in the KYH study (later referenced as the baseline) and ended at the end of the COVID-19 pandemic (i.e. 5 May 2023) or at the date the participant died [20]. The total follow-up period was divided into the pre-pandemic period (date of KYH inclusion – 16 March 2020) and the pandemic period (17 March 2020 – 5 May 2023). Data collected at the baseline study were linked to the Arkhangelsk Regional Mortality Database (later referenced as the mortality registry). In this study, we included deaths from all causes among the participants recorded in the mortality registry during both periods.

Fifteen individuals with missing baseline data on any of the covariates were excluded, resulting in a pre-pandemic study sample of 2342 participants aged 35-69 years. All the study participants alive at the start of the pandemic in Arkhangelsk (17 March 2020) comprised the pandemic study sample of 2284 individuals aged 40-74 years. Data of the pandemic study sample were linked to the Federal Registry of the Vaccinated against COVID-19 (later referenced as the vaccination registry).

Outcomes and covariates

For each death, data from medical death certificates were collected, including the date of death, immediate cause of death, related pathological conditions, underlying cause of death, external cause of death, and other contributing conditions in accordance with the International Classification of Diseases, 10th revision (ICD-10).

The baseline data were collected in 2015-2017 years by questionnaire, physical examination, and laboratory tests by following the KYH study protocol as described by S. Cook et al. [19]. As the KYH study was designed to investigate cardiovascular diseases, the baseline data included predominantly cardiovascular biomarkers.

The following data variables were considered as covariates in analyses for this study: demographic factors: age (years), sex (male/female), higher education (yes/no); lifestyle characteristics: current smoking (yes/no), hazardous alcohol drinking (score of ≥8 on the Alcohol Use Disorders Identification Test – AUDIT [21]) (yes/no)); self-reported doctor-diagnosed diseases: hypertension (yes/no), diabetes (yes/no), angina (yes/no), history of myocardial infarction (yes/no), heart failure (yes/no), asthma (yes/no), chronic bronchitis (yes/no), kidney disease (yes/no), liver disease (yes/no), neoplasms (yes/no); obesity (measured at physical examination as body mass index ≥30 kg/m^2; yes/no); blood-based cardiac and metabolic biomarkers: lipid profile: total cholesterol ≥5.2 mmol/L (yes/no), low-density lipoprotein cholesterol (LDL-C) >3.0 mmol/L (yes/no), high-density lipoprotein cholesterol (HDL-C) < 1.0 mmol/L for men and <1.3 mmol/L for women (yes/no), triglycerides >1.7 mmol/L (yes/no); a biomarker of hyperglycemia: glycated hemoglobin (HbA1C) ≥ 6.5% (yes/no); a biomarker of alcohol consumption and cholestasis: Gamma-glutamyl transferase (GGT) ≥40 U/L (yes/no); a biomarker of systemic inflammation: high-sensitivity C-reactive protein (Hs-CRP) ≥2 mg/L (yes/no); a biomarker of kidney dysfunction: Cystatin C ≥1.2 mg/L (yes/no); a biomarker of cardiac wall stretch (heart failure): N-terminal pro-b-type natriuretic peptide (NT-proBNP) ≥125 pg/mL (yes/no); a biomarker of heart damage: high-sensitivity Troponin T (Hs-Troponin T) ≥6 ng/L (yes/no). For the pandemic period, we also considered data on vaccination against COVID-19 (yes/no) obtained the vaccination registry. We did not investigate the association between having had COVID-19 and risk of death during the pandemic because, according to the seroprevalence survey at the end of the pandemic, almost everyone in the sample had antibodies acquired through infection [22, 23].

Ethical approval

The study was conducted in accordance with the Declaration of Helsinki. The KYH study was approved by the ethics committees of the London School of Hygiene & Tropical Medicine, London, UK (approval number 8808, February 24, 2015) and Northern State Medical University (NSMU), Arkhangelsk, Russia (approval number 01/01-15, January 27, 2015). Ethical approval for the follow-up of the KYH participants was received from the ethics committee of NSMU (approval number 01/04-19, April 24, 2019). Ethical approval for the study of COVID-19 issues in the KYH cohort was obtained from the ethics committees of NSMU (approval number 01/02-

21, February 17, 2021) and by the Regional Committee for Medical and Health Research Ethics, Norway (approval number 339397 received December 7, 2021).

All participants provided a written consent to disclose their medical and other health-related records for research purposes under the confidentiality condition. The data linkage between KYH data, the vaccination registry and the regional mortality registry was performed by the Arkhangelsk Regional Medical Information Analytical Center (MIAC) in accordance with the NSMU-MIAC confidentiality agreement based on the informed consents obtained from the participants, legal and ethical approvals. The participants were anonymized by the randomly assigned unique ID numbers and the following data linkage were based on these depersonalized IDs, which has led to no personal identifiers in the analyzed dataset.

Statistical analysis

Participants characteristics were presented as medians with first and third quartiles for continuous variables, absolute numbers and percentages for categorical variables. Accordingly, Mann–Whitney U-test and Pearson's chi-squared test were used for comparing groups on continuous and categorical characteristics.

All-cause and cause-specific (defined by the ICD-10 chapters) mortality rates per 1000 person-years were calculated for men and women in both the pre-pandemic and pandemic periods. Mortality rates for the pandemic period were directly age-standardized to the age distribution of the study population in the pre-pandemic period using 5-year bands. Mortality rates were presented with 95% confidence intervals (CIs).

Age-adjusted mortality ratios in the pandemic period were estimated as hazard ratios (HR) derived from Cox proportional hazards regression models of the studied death outcomes, where period (1 = pandemic period, 0 = pre-pandemic period) and age in years were entered as covariates. Person-months of observation in the pre-pandemic (date of KYH inclusion -16 March 2020) and pandemic (17 March 2020-5 May 2023) periods were used as time variables in Cox models for the pre-pandemic and pandemic periods, respectively. When modelling each of death outcomes, interaction of the period with sex, likewise other interactions considered in this study, was assessed by comparing models with and without the interaction term using the likelihood ratio test. Based on the identified interaction of the period with sex, further analyses were stratified by sex. We also used Cox models to investigate the factors associated with a change (increase or decrease) in the risk of pre-pandemic and pandemic deaths. For each independent variable, interaction with the study period was assessed. The effect estimates were reported as HRs adjusted for demographic (age, higher education) and lifestyle (smoking, hazardous drinking) variables with the corresponding 95% CIs. All the statistical analyses were performed

in Stata version 17.0 (StataCorp, College Station, TX, USA). A p-value < 0.05 was considered statistically significant.

Results

The studied population comprised 980 men (41.8%) and 1362 women (58.2%). During a median follow-up period of 6.5 years (6.0; 7.0), a total of 150 deaths occurred among the study participants. In the pre-pandemic period, 45 men and 13 women died, leaving a sample of 935 men and 1349 women for the pandemic period. A total of 56 men and 36 women died in the pandemic period.

Characteristics of the study participants

The median age was 53 (44; 61) years at the time of inclusion in KYH study (at baseline) and 57 (48; 65) years at the onset of the pandemic (table 1). Men were more likely to be smokers and hazardous drinkers compared to women. Higher proportions of women compared to men had obesity, diabetes, heart failure, asthma, chronic bronchitis, kidney disease, liver disease, and neoplasms. Conversely, more men than women reported having prior myocardial infarction.

Characteristics[a]	Men, N=980	Women, N=1362	p[b]
Age, years	53 (44-61)	53 (44-61)	0.684[3]
Higher education	334 (34.1)	510 (37.4)	0.094
Smoking	338 (34.5)	207 (25.2)	<0.001
Hazardous drinking	373 (38.1)	91 (6.7)	<0.001
Obesity	238 (24.3)	460 (33.8)	<0.001
Hypertension	454 (46.3)	676 (49.6)	0.114
Myocardial infarction	69 (7.0)	47 (3.5)	<0.001
Angina	133 (13.6)	192 (14.1)	0.717
Heart failure	87 (8.9)	181 (13.3)	0.001
Diabetes	64 (6.5)	132 (9.7)	0.006
Asthma	39 (4.0)	109 (8.0)	<0.001
Chronic bronchitis	86 (8.8)	189 (13.9)	<0.001
Kidney diseases	124 (12.7)	305 (22.4)	<0.001
Liver disease	142 (14.5)	263 (19.3)	0.002
Neoplasms	27 (2.8)	105 (7.7)	<0.001

Table 1. Baseline demographic, lifestyle and self-reported health characteristics of the study population stratified by sex, N=2342. [a]Me – median, Q1 – the first quartile; Q3 – the third quartile for age, N (%) – for all other characteristics. [b]Pearson's chi-squared test for categorical parameters, Mann–Whitney U-test for continuous characteristics.

Women exhibited a higher proportion of elevated total cholesterol and decreased HDL-C levels, while increased triglycerides were more prevalent among men (table 2). Men were also more

likely to have elevated levels of GGT and Hs-Troponin T, whereas elevated levels of NT-proBNP were more commonly observed among women.

Characteristics [a]	Men, N=980	Women, N=1362	P[b]
Total cholesterol, ≥5.2 mmol/L, N (%)	498 (50.8)	810 (59.5)	<0.001
LDL-C, >3.0 mmol/L, N (%)	737 (75.2)	1035 (76.0)	0.661
HDL-C, <1.0 mmol/L for men and <1.3 mmol/L for women, N (%)	134 (13.7)	342 (25.1)	<0.001
Triglycerides, >1.7 mmol/L, N (%)	305 (31.1)	334 (24.5)	<0.001
HbA1C, ≥6.5%	45 (4.6)	68 (5.0)	0.655
GGT, ≥40 U/L, N (%)	349 (35.6)	216 (15.9)	<0.001
Hs-CRP, ≥2 mg/L, N (%)	405 (41.3)	571 (41.9)	0.772
Cystatin C, ≥1.2 mg/L, N (%)	47 (4.8)	50 (3.7)	0.178
NT-proBNP, ≥125 pg/mL, N (%)	255 (26.0)	513 (37.7)	<0.001
Hs-Troponin T, ≥6 ng/L, N (%)	672 (68.6)	632 (46.4)	<0.001

Table 2. Baseline blood-based biomarkers characteristics in the study population stratified by sex, N=2342. [a]N (%) – for all the characteristics. [b]Pearson's chi-squared test.

Vaccination against COVID-19 was recorded for 64.7% of men and 64.9% of women (p=0.938).

Mortality rates

The mortality rates were higher in men compared to women in both periods (table 3).

In women, age-standardized all-cause mortality rates increased from 2.79 per 1000 person-years in the pre-pandemic period to 6.45 per 1000 person-years in the pandemic period (a 2.32-fold increase), but they did not change significantly in men (p=0.047 for the interaction between sex and the study period). In the pandemic period, neoplasms were the leading cause of death in women, followed by cardiovascular diseases, whereas the opposite pattern was observed in the pre-pandemic period. In men, cardiovascular diseases were the leading cause of death in both periods, followed by neoplasms. Mortality rates from COVID-19 as underlying cause of death were similar in men and women.

Underlying cause of deaths	Pre-pandemic period: date of KYH inclusion – 17 March 2020		Pandemic period: 17 March 2020 – 5 May 2023			Age-adjusted mortality ratios (95%CI)[a]
	Women, N=1362		Women, N=1349			
	N	MR$_{crude}$ (95%CI)[b]	N	MR$_{crude}$ (95%CI)[b]	MR$_{stand}$ (95%CI)[bc]	
All causes	13	2.79 (1.62; 4.80)	36	8.75 (6.31; 12.13)	6.45 (3.93; 9.98)	2.32 (1.17; 4.61)
I00-I99 Cardiovascular diseases	5	1.07 (0.45; 2.57)	9	2.19 (1.14; 4.21)	1.39 (0.24; 2.54)	1.64 (0.49; 5.43)
C00-D48 Neoplasms	4	0.86 (0.32; 2.28)	12	2.92 (1.66; 5.14)	1.78 (0.46; 3.10)	1.77 (0.55; 5.67)
V01-Y98 External causes	1	0.21 (0.03; 1.52)	2	0.49 (0.12; 1.94)	0.54 (0.00; 1.29)	1.98 (0.17; 22.56)

		Men, N=980		Men, N=935		
		MR$_{crude}$ (95%CI)[b]		MR$_{crude}$ (95%CI)[b]	MR$_{stand}$ (95%CI)[bc]	
U07 COVID-19	-	-	6	1.46 (0.66; 3.25)	0.87 (0.00; 1.78)	-
Other causes	3	0.64 (0.21; 1.99)	7	1.70 (0.81; 3.57)	1.86 (0.47; 3.25)	2.76 (0.56; 13.54)
All causes	45	13.43 (10.03; 17.99)	56	20.03 (15.41; 26.03)	15.58 (10.93; 20.24)	1.16 (0.77; 1.75)
I00-I99 Cardiovascular diseases	23	6.86 (4.56; 10.33)	22	7.87 (5.18; 11.95)	6.34 (3.38; 9.29)	0.93 (0.50; 1.72)
C00-D48 Neoplasms	9	2.69 (1.40; 5.16)	16	5.72 (3.51; 9.34)	4.64 (2.07; 7.20)	1.45 (0.61; 3.46)
V01-Y98 External causes	5	1.49 (0.62; 3.58)	3	1.07 (0.35; 3.33)	1.11 (0.00; 2.36)	0.75 (0.16; 3.41)
U07 COVID-19	-	-	5	1.79 (0.74; 4.30)	1.21 (0.00; 2.52)	-
Other causes	8	2.39 (1.19; 4.77)	10	3.58 (1.92; 6.65)	2.29 (0.54; 4.03)	1.07 (0.41; 2.75)

Table 3. Crude and age-standardized mortality rates per 1000 person-years. [a]Hazard ratios derived from Cox regression models of the studied death outcomes, where period (1 = pandemic period, 0 = pre-pandemic period) and age in years were entered as covariates. [b]MR - mortality rate, 95% CI – 95% confidence intervals. [c]age-standardized to the age distribution of the study population in the pre-pandemic period.

Risk factors for pre-pandemic and pandemic all-cause deaths

After adjustment for demographic and lifestyle factors, older participants, smokers, and those with a self-reported diagnosis of diabetes had a significantly higher risk of all-cause death in both sexes in both periods (table 4). In men, neoplasms were associated with increased the risk of all-cause death only in the pre-pandemic period, whereas asthma increased the risk only in the pandemic period. In women, self-reported heart failure was associated with increased risk of death only in the pre-pandemic period, while obesity, and angina – only in the pandemic period. Higher education was associated with a decreased risk of death during the pandemic in women. Vaccination against COVID-19 reduced all-cause mortality during the pandemic period in both sexes. There were no significant interactions between the study period and demographic, lifestyle, and health characteristics in their effects on all-cause mortality.

Characteristics[a]	Men			Women		
	Pre-pandemic period[b], 45/980	Pandemic period[b], 56/935	p[c]	Pre-pandemic period[b], 13/1362	Pandemic period[b], 36/1349	p[c]
	HR$_{adj}$ (95%CI)[d]	HR$_{adj}$ (95%CI)[d]		HR$_{adj}$ (95%CI)[d]	HR$_{adj}$ (95%CI)[d]	
Age, years[e]	1.06 (1.03; 1.10)	1.09 (1.05; 1.13)	0.784	1.11 (1.03; 1.20)	1.11 (1.06; 1.16)	0.470
Higher education	0.50 (0.23; 1.09)	0.67 (0.37; 1.24)	0.474	0.21 (0.03; 1.61)	0.30 (0.11; 0.86)	0.736
Smoking	2.40 (1.32; 4.38)	2.80 (1.64; 4.78)	0.994	4.01 (1.13; 14.24)	3.11 (1.40; 6.94)	0.695
Hazardous drinking	1.28 (0.70; 2.34)	0.69 (0.38; 1.25)	0.104	1.07 (0.13; 8.82)	0.82 (0.19; 3.55)	0.764
Obesity	1.67 (0.88; 3.17)	1.07 (0.57; 2.01)	0.347	0.83 (0.27; 2.56)	2.19 (1.11; 4.30)	0.144

Hypertension	1.24 (0.67; 2.31)	1.32 (0.75; 2.30)	0.704	1.71 (0.45; 6.58)	1.81 (0.80; 4.11)	0.925
Myocardial infarction	1.31 (0.54; 3.18)	1.67 (0.83; 3.39)	0.434	3.77 (0.96; 14.88)	2.06 (0.77; 5.46)	0.495
Angina	1.78 (0.89; 3.58)	0.85 (0.42; 1.71)	0.309	2.90 (0.89; 9.47)	3.59 (1.76; 7.30)	0.757
Heart failure	1.30 (0.54; 3.13)	1.38 (0.66; 2.86)	0.668	3.48 (1.12; 10.82)	1.28 (0.59; 2.77)	0.171
Diabetes	4.68 (2.26; 9.69)	3.41 (1.63; 7.12)	0.579	3.83 (1.19; 12.35)	2.32 (1.09; 4.92)	0.530
Asthma	1.34 (0.32; 5.57)	2.62 (1.05; 6.61)	0.376	1.39 (0.31; 6.39)	1.61 (0.66; 3.88)	0.899
Chronic bronchitis	0.83 (0.30; 2.34)	0.49 (0.15; 1.58)	0.469	1.70 (0.51; 5.69)	0.95 (0.41; 2.20)	0.406
Kidney diseases	1.27 (0.59; 2.76)	1.41 (0.72; 2.75)	0.754	0.60 (0.13; 2.72)	1.64 (0.82; 3.29)	0.206
Liver disease	1.29 (0.60; 2.77)	1.78 (0.94; 3.37)	0.548	0.70 (0.15; 3.18)	0.88 (0.38; 2.02)	0.754
Neoplasms	3.79 (1.45; 9.91)	1.37 (0.42; 4.48)	0.299	2.69 (0.73; 9.92)	1.08 (0.38; 3.07)	0.314
Vaccination against COVID-19[f]	-	0.17 (0.09; 0.32)	-	-	0.19 (0.09; 0.38)	-

Table 4. Cox regressions describing the associations between demographic, health-related factors and all-cause mortality during pre-pandemic and pandemic periods stratified by sex. [a]baseline data collected in 2015-2017. [b]pre-pandemic period: date of KYH inclusion – 16 March 2020; pandemic period: 17 March 2020 – 5 May 2023. [c]p-value for interaction with the time period. [d] HR – hazard ratio with 95% CI – 95% confidence intervals adjusted for the demographic (age, higher education) and lifestyle factors (smoking, hazardous drinking). [e]age at the baseline for pre-pandemic period; age on 17 March 2020 for pandemic period. [f]COVID-19 – new coronavirus disease.

Elevated levels of HbA1C increased risk of death in both sexes during both periods (table 5). In men, during the pre-pandemic period, elevated levels of Hs-CRP, Cystatin C, and NT-proBNP were associated with an increased risk, while elevated levels of total cholesterol were associated with a lower risk of all-cause deaths. During the pandemic period, men had increased risks of death associated with elevated levels of Hs-CRP, NT-proBNP and Hs-Troponin T. The strength of the association between elevated NT-proBNP and risk of death increased during the pandemic period, although there was no significant interaction with the study period, p=0.177. The protective effect of elevated cholesterol decreased during the pandemic period (p for interaction with period = 0.002), while the effect of triglycerides increased (p for interaction with period = 0.033), although the effects of triglycerides were not significant in any of the periods. In women, elevated NT-proBNP levels were associated with a higher risk of pre-pandemic deaths, while elevated Cystatin C levels increased the risk during the pandemic (table 5).

Table 5. Risk factors (blood-base biomarkers) for all-cause mortality during pre-pandemic and pandemic periods stratified by sex

Characteristics[a]	Men		p[c]	Women		p[c]
	Pre-pandemic period[b], 45/980	Pandemic period[b], 56/935		Pre-pandemic period[b], 13/1362	Pandemic period[b], 36/1349	
	HR$_{adj}$ (95%CI)[d]	HR$_{adj}$ (95%CI)[d]		HR$_{adj}$ (95%CI)[d]	HR$_{adj}$ (95%CI)[d]	
Total cholesterol, ≥5.2 mmol/L	0.29 (0.14; 0.58)	1.18 (0.70; 2.01)	0.002	0.77 (0.25; 2.37)	0.48 (0.25; 0.92)	0.468

Biomarker						
LDL-C[e], >3.0 mmol/L	0.59 (0.32; 1.08)	1.20 (0.64; 2.23)	0.115	0.86 (0.24; 3.13)	1.05 (0.46; 2.40)	0.797
HDL-C[e], <1.0 mmol/L for men and <1.3 mmol/L for women	1.32 (0.61; 2.84)	1.06 (0.50; 2.24)	0.696	2.22 (0.28; 17.52)	3.77 (1.31; 10.83)	0.151
Triglycerides, >1.7 mmol/L	0.60 (0.29; 1.24)	1.58 (0.92; 2.71)	0.033	1.32 (0.43; 4.10)	1.22 (0.62; 2.44)	0.890
HbA1C[e], ≥6.5%	4.04 (1.76; 9.31)	5.19 (2.56; 10.52)	0.536	3.86 (1.02; 14.57)	3.73 (1.67; 8.32)	0.981
GGT[e], ≥40 U/L	1.84 (1.01; 3.37)	1.36 (0.78; 2.38)	0.277	2.15 (0.66; 7.00)	1.18 (0.52; 2.69)	0.423
Hs-CRP[e], ≥2 mg/L	2.42 (1.27; 4.61)	1.80 (1.04; 3.12)	0.479	2.34 (0.72; 7.63)	1.49 (0.77; 2.90)	0.495
Cystatin C, ≥1.2 mg/L	3.84 (1.82; 8.15)	1.36 (0.56; 3.27)	0.176	2.40 (0.51; 11.26)	3.47 (1.48; 8.12)	0.668
NT-proBNP[e], ≥125 pg/mL	1.96 (1.01; 3.81)	3.13 (1.71; 5.75)	0.177	5.14 (1.08; 24.52)	1.79 (0.86; 3.72)	0.219
Hs-Troponin T[e], ≥6 ng/L	1.47 (0.68; 3.17)	2.79 (1.17; 6.70)	0.165	2.26 (0.58; 8.78)	1.76 (0.81; 3.80)	0.785

Table 5. Cox regressions describing the associations between blood-base biomarkers and all-cause mortality during pre-pandemic and pandemic periods stratified by sex. [a]baseline data collected in 2015-2017. [b]pre-pandemic period: date of KYH inclusion – 16 March 2020; pandemic period: 17 March 2020 – 5 May 2023. [c]p-value for interaction with the time period. [d] HR – hazard ratio with 95% CI – 95% confidence intervals adjusted for the demographic (age, higher education) and lifestyle factors (smoking, hazardous drinking). [e]BMI – body mass index, LDL – low-density lipoprotein cholesterol, HbA1C – glycated hemoglobin, GGT – Gamma-glutamyl transferase, Hs-CRP – high-sensitivity C-reactive protein, NT-proBNP – N-terminal pro-brain natriuretic peptide, Hs-Troponin T – high-sensitivity Troponin T.

Discussion

The age-standardized mortality rates were higher in men during both periods, although the increase in mortality rates during the pandemic period was more pronounced in women than in men. Cardiovascular diseases and neoplasms were the leading causes of death in both periods and for both sexes. During the pandemic period, mortality from neoplasms exceeded mortality from cardiovascular diseases in women, while cardiovascular diseases remained the leading cause of death in men. Mortality from COVID-19 as the underlying cause were similar in both sexes. Compared with the pre-pandemic period, during the COVID-19 pandemic, women had an increased risk of death associated with obesity, angina, and elevated cystatin C levels, while men had an increased risk of death associated with asthma, elevated biomarkers of cardiovascular risk.

Previous studies have shown that during the COVID-19 pandemic, all-cause excess mortality, the difference between observed and expected deaths, was highest in Russia among European countries [3, 24, 25]. In Russia, women experienced higher excess mortality than men, while in most countries excess mortality was higher in men [26]. A population-based study involving 63

countries showed substantial sex differences in age-standardized COVID-19 mortality rates, which were greater than those for all-cause mortality, with higher mortality rates in men than in women [27]. On the contrary, women were more susceptible to COVID-19 than men in countries such as India, Nepal, Slovenia, and Vietnam [28, 29].

Several studies have reported an increase in mortality from cardiovascular diseases and neoplasms during the COVID-19 pandemic, suggesting that the increase in deaths is largely due to indirect effects of the pandemic, such as delayed diagnosis and treatment [30-32]. However, other authors have suggested that the increase in deaths from cardiovascular diseases and neoplasms may be primarily due to undetected deaths related to COVID-19 [33].

In our study, mortality increased significantly in women during the pandemic, with obesity, angina, and elevated cystatin C levels being associated with an increased risk of death. The fact that these factors are also risk factors for death from COVID-19 supports that women may be more vulnerable than men to complications directly related to the virus [34-36]. Except for asthma, the increased risk of death in men during the pandemic was associated with elevated biomarkers of cardiovascular risk. The lack of differences in COVID-19 mortality rates between the sexes may be due to an underestimation of the role of the virus and misclassification of COVID-19 deaths as non-COVID-19 deaths. Deaths from virus-related thrombovascular events may have been recorded as deaths from cardiovascular causes, or deaths in COVID-19 patients with advanced cancer may have been classified as deaths from neoplasms, especially when testing availability was limited [37].

Given that the age of participants at the start of the pandemic ranged from 40 to 74 years, deaths in the study sample were relatively seldom and most could be considered premature. Therefore, some well-known risk factors were not associated with the risk of death during the two relatively short study periods.

Higher education had a protective effect on the risk of death during the pandemic in women, but not in men. The association between higher education and mortality during the pandemic is complex and may involve factors such as employment status, work arrangements, social interactions, adherence to non-pharmaceutical interventions, healthcare-seeking and testing behaviors, and vaccination intentions [38, 39].

Our study found a positive association between smoking and mortality risk in both study periods for both sexes, consistent with evidence linking smoking to mortality from both COVID-19 and non-COVID-19 causes [40-42]. Despite recent findings suggesting an association between hazardous alcohol drinking and increased risk of death during COVID-19 pandemic due to poor health status and comorbidities, our study did not observe such an association [43-45].

Women with angina and men with elevated levels of Hs-CRP, Hs-Troponin T and NT-proBNP were at higher risk of death during the pandemic period, consistent with other research [6, 29, 46, 47]. Alongside the increased risk of encountering severe COVID-19 and life-threatening thrombovascular complications of COVID-19 during the acute phase of illness and for several months following infection [5, 36, 48], individuals with cardiovascular diseases experienced the indirect effects of disrupted healthcare services and delayed diagnostics and treatment [30, 31]. Patients with pre-existing cardiovascular disease had a higher risk of death during the pandemic than those with pre-existing chronic respiratory disease, even though SARS-CoV-2 primarily affects the respiratory system [49]. In our study, asthma appeared to be associated with death in men during the pandemic, contradicting findings published by other authors [50]. Sex differences in the association may be related to poorer asthma control in men compared with women, although there is no clear consensus on the association between sex and treatment compliance [51]. Another study found that inhaled corticosteroids used to treat asthma may reduce the severity of COVID-19, with lower rates of hospitalization and death reported among users [52, 53].

Obesity, a well-established risk factor for severe COVID-19, increased the risk of death in women during the pandemic, consistent with previous research [54]. Women in our study may be more vulnerable to complications directly related to the virus, possibly due to the higher prevalence of obesity compared to men [55]. Obesity exacerbates SARS-CoV-2-induced inflammation, leading to cytokine storms and increased risk of severe illness and mortality [34, 56, 57].

Elevated total cholesterol levels were paradoxically associated with a lower risk of all-cause death in the pre-pandemic period for men and in the pandemic period for women, possibly due to the use of lipid-lowering medications. Other authors have reported that in younger individuals, elevated total cholesterol levels increase the risk of death, whereas in older individuals with multiple comorbidities, total cholesterol levels decrease due to the use of lipid-lowering medications while the risk of death increases, leading to a negative association between elevated total cholesterol levels and death [58]. The effects of biomarkers influenced by medication use must be interpreted with caution, as medication status is likely to have changed over time.

Neoplasms were associated with death in men shortly after enrollment and before the pandemic. Since baseline data on neoplasms were not updated, the pandemic sample included long-term cancer survivors diagnosed over 5 years ago, potentially underestimating the impact of neoplasms on pandemic mortality.

We found no significant associations between chronic kidney disease or chronic liver disease and mortality in both study periods, despite previous research suggesting potential associations between these conditions and COVID-19-related mortality [29, 59, 60].

Elevated Cystatin C levels were associated with an increased risk of death in women in the pandemic. Although a biomarker of kidney dysfunction, elevated levels of cystatin C are also associated with an increased risk of cardiovascular events related to endothelial dysfunction due to atherosclerosis [61, 62] and are correlated with severe COVID-19 and COVID-19-related death [35, 63, 64].

Vaccination against COVID-19 reduced all-cause mortality during the pandemic period in both sexes. This finding is consistent with another study showing that vaccination protects against severe COVID-19, reduces mortality in individuals with pre-existing cardiovascular disease, and prevents long-term sequelae such as myocarditis, myocardial infarction, and stroke [65].

The study contributes to understanding how COVID-19 affected premature mortality in men and women, as well as identifying the risk factors associated with death during the COVID-19 pandemic compared to the pre-pandemic period in the Russian adult population. We analyzed all-cause mortality rather than cause-specific mortality data to indirectly assess the impact of the COVID-19 pandemic on overall mortality, accounting for potential misclassification between COVID-19 and non-COVID-19 deaths. In our analysis, we consider the timing of deaths to assess whether pre-existing chronic diseases and blood-based biomarkers exhibit varying associations with subsequent mortality over time. The relatively high response rate of 68% in the KYH study suggests that the study sample is representative of the population of Arkhangelsk, aged 35-69 years at the baseline.

Our study has several limitations. The baseline characteristics of the study participants, particularly blood-based biomarkers, were collected at a single time point more than five years before the pandemic period, and they might have changed over time. We focused primarily on cardiovascular biomarkers as the baseline data were collected for the KYH study of cardiovascular disease. The lack of data on dropouts is another limitation. Some individuals may have relocated, making it no longer possible to register their outcomes. The proportion of unaccounted dropouts likely increased over time, resulting in a higher number of dropouts during the pandemic period. Consequently, they may have been misclassified as alive, potentially resulting in underestimated mortality rates during the pandemic period. Cause-specific mortality rates may also be affected by the accuracy of information in death certificates. We did not investigate the association between having had COVID-19 and risk of death during the pandemic period due to the high prevalence of infection-acquired immunity in the study sample according to the seroprevalence survey [22, 23]. The relatively small number of deaths analyzed was due to

the relatively young age of the cohort, which may limit the interpretation of the results regarding factors associated with risk of death.

Conclusions

During the COVID-19 pandemic, mortality rates remained higher in men than in women, with cardiovascular disease being the leading cause of death. There was a two-fold increase in premature mortality in women, but a minor change in men. The increase in mortality in women during the pandemic could be explained by the effects of obesity, angina and elevated Cystatin C levels (indicating kidney dysfunction), with those without higher education being more vulnerable. Along with elevated biomarkers of cardiovascular risk, asthma emerged as a factor associated with increased risk of death during the pandemic in men. Targeted preventive measures for women and men with specific risk factors can be implemented during potential future infectious disease outbreaks.

Data availability

Data from the Know Your Heart Study are available upon reasonable request after contacting Alexander V. Kudryavtsev at alex.v.kudryavtsev@yandex.ru. See data access regulations and instructions at https://metadata.knowyourheart.science (Accessed on 7 June May 2024). All data requests will be guided by protecting of personal information, confidentiality agreement with participants, and their informed consents.

References

1.   Adão, R., Guzik, T.J. Inside the heart of COVID-19. *Cardiovascular research*. **116**, 59-61 (2020).
2.   Strongman, H., Carreira, H., De Stavola, B.L., Bhaskaran, K., Leon, D.A. Factors associated with excess all-cause mortality in the first wave of the COVID-19 pandemic in the UK: A time series analysis using the Clinical Practice Research Datalink. *PLoS medicine*. **19**, 1003870; 10.1371/journal.pmed.1003870 (2022).
3.   Karlinsky, A., Kobak, D. Tracking excess mortality across countries during the COVID-19 pandemic with the World Mortality Dataset. *eLife*. **10**, 69336; 10.7554/eLife.69336 (2021).
4.   von Stillfried, S., Bülow, R.D., Röhrig, R., Boor, P. First report from the German COVID-19 autopsy registry. *The Lancet regional health Europe*.**15**, 100330; 10.1016/j.lanepe.2022.100330 (2022).
5.   Huertas, A., Montani, D., Savale, L, Pichon J, Tu L, Parent F, et al. Endothelial cell dysfunction: a major player in SARS-CoV-2 infection (COVID-19)? The European respiratory journal. 2020;56(1).
6.   De Michieli, L. et al. High-Sensitivity Cardiac Troponin T for the Detection of Myocardial Injury and Risk Stratification in COVID-19. *Clinical chemistry*. **67**, 1080-1089 (2021).
7.   Chen, C. et al. [Analysis of myocardial injury in patients with COVID-19 and association between concomitant cardiovascular diseases and severity of COVID-19]. *Zhonghua xin xue guan bing za zhi*. **48**, 567-571 (2020).
8.   Lee, W.E. et al. Direct and indirect mortality impacts of the COVID-19 pandemic in the US, March 2020-April 2021. **12**, 77562; 10.7554/elife.77562 (2022).
9.   Cuschieri, S., Mamo, J. Taking care of the ordinary in extraordinary times-delayed routine care means more morbidity and pre-mature mortality. *European journal of public health*. **31**, 27-30 (2021).
10.   Medline, A. et al. Evaluating the impact of stay-at-home orders on the time to reach the peak burden of Covid-19 cases and deaths: does timing matter? BMC public health. **20**, 1750; 10.1186/s12889-020-09817-9 (2020).
11.   Mafham, M.M. et al. COVID-19 pandemic and admission rates for and management of acute coronary syndromes in England. *Lancet*. **396**, 381-389 (2020).
12.   Methodological Recommendations MP 3.1.0178-20 "Definition of a Set of Measures, as Well as Indicators That Are the Basis for the Phased Removal of Restrictive Measures in the Conditions of the Epidemic Spread of COVID-19" (in Russian) https://www.garant.ru/products/ipo/prime/doc/73904890/ (2020).
13.   Order of the Ministry of Health of the Russian Federation № 198n of March 19, 2020. "On the temporary organization of the activities of medical organizations for the implementation of measures to prevent and reduce the risks of spreading the new coronavirus infection COVID-19"
14.   Resolution of the Government of the Russian Federation No. 432 of April 3, 2020. "On the features of implementing the basic program of compulsory medical insurance in conditions of the emergence of threats of the spread of diseases caused by the new coronavirus infection".
15.   Harrison, S.L., Fazio-Eynullayeva, E., Lane, D.A., Underhill, P., Lip, G.Y.H. Comorbidities associated with mortality in 31,461 adults with COVID-19 in the United States: A federated electronic medical record analysis. *PLoS medicine*.**17**,1003321; 10.1371/journal.pmed.1003321 (2020).
16.   Grippo, F. et al. The Role of COVID-19 in the Death of SARS-CoV-2-Positive Patients: A Study Based on Death Certificates. *Journal of clinical medicine*. **9**, 3459; 10.3390/jcm9113459 (2020).
17.   Nada, K.M. et al. Determining Cause of Death During Coronavirus Disease 2019 Pandemic. *Critical care explorations*. **3**, 0419; 10.1097/CCE.0000000000000419 (2021).
18.   Nab, L. et al. Changes in COVID-19-related mortality across key demographic and clinical subgroups in England from 2020 to 2022: a retrospective cohort study using the OpenSAFELY platform. *The Lancet Public health*. **8**, 364-377 (2023).
19.   Cook, S. et al. Know Your Heart: Rationale, design and conduct of a cross-sectional study of cardiovascular structure, function and risk factors in 4500 men and women aged 35-69 years from two Russian cities, 2015-18. *Wellcome open research*. **3**, 67; 10.12688/wellcomeopenres.14619.3 (2018).
20.   WHO Director-General's opening remarks at the media briefing – 5 May 2023 https://wwwwhoint/news-room/speeches/item/who-director-general-s-opening-remarks-at-the-media-briefing---5-may-2023 (2023).

21.   World Health Organization; Babor, T. F., Higgins-Biddle, J. C., Saunders, J. B., Monteiro, M. G. *AUDIT: The Alcohol Use Disorders Identification Test: Guidelines for Use in Primary Health Care*, 2nd edn (World Health Organization, Geneva, Switzerland, 2001).

22.   Krieger, E.A. et al. The impact of polymorphic variants of interferon receptor genes on COVID-19 severity and antibiotic resistance. *Russian Journal of Infection and Immunity*. **13**, 1027-1039 (2023).

23.   Krieger, E. et al. COVID-19: seroprevalence and adherence to preventive measures in Arkhangelsk, Northwest Russia. *Infectious diseases*. **55**, 316-327 (2023).

24.   Timonin, S. et al. Excess mortality in Russia and its regions compared to high income countries: An analysis of monthly series of 2020. *SSM - population health*. **17**, 101006; 10.1016/j.ssmph.2021.101006 (2022).

25.   Islam, N. et al. Effects of covid-19 pandemic on life expectancy and premature mortality in 2020: time series analysis in 37 countries. *BMJ*. **375**, 066768; 10.1136/bmj-2021-066768 (2021).

26.   Nielsen, J., Nørgaard, S.K., Lanzieri, G., Vestergaard, L.S., Moelbak, K. Sex-differences in COVID-19 associated excess mortality is not exceptional for the COVID-19 pandemic. *Scientific reports*. **11**, 20815; 10.1038/s41598-021-00213-w (2021).

27.   Geldsetzer, P. et al. Sex differences in the mortality rate for coronavirus disease 2019 compared to other causes of death: an analysis of population-wide data from 63 countries. *European journal of epidemiology*. **37**, 797-806 (2022).

28.   Dehingia, N., Raj, A. Sex differences in COVID-19 case fatality: do we know enough? *The Lancet Global health*. **9**, 14-15 (2021).

29.   Yoshida, Y., Wang, J., Zu, Y. Sex differences in comorbidities and COVID-19 mortality-Report from the real-world data. *Frontiers in public health*. **10**, 881660; 10.3389/fpubh.2022.881660 (2022).

30.   Odone, A., Delmonte, D., Gaetti, G., Signorelli, C. Doubled mortality rate during the COVID-19 pandemic in Italy: quantifying what is not captured by surveillance. *Public health*. **190**,108-115 (2021).

31.   Carey, I.M. et al. Risk factors for excess all-cause mortality during the first wave of the COVID-19 pandemic in England: A retrospective cohort study of primary care data. *PloS one*. **16**, 0260381; 10.1371/journal.pone.0260381 (2021).

32.   Maringe, C. et al. The impact of the COVID-19 pandemic on cancer deaths due to delays in diagnosis in England, UK: a national, population-based, modelling study. *The Lancet Oncology*. **21**, 1023-1034 (2020).

33.   Figueroa, J. et al. Trends in excess cancer and cardiovascular deaths in Scotland during the COVID-19 pandemic 30 December – 20 April suggest underestimation of COVID-19 related deaths. Preprint at https://www.medrxiv.org/content/10.1101/2020.05.02.20086231v1 (2020).

34.   Zhang, F. et al. Obesity predisposes to the risk of higher mortality in young COVID-19 patients. *Journal of medical virology*. **92**, 2536-2542 (2020).

35.   Elliott, J. et al. COVID-19 mortality in the UK Biobank cohort: revisiting and evaluating risk factors. *European journal of epidemiology*. **36**, 299-309 (2021).

36.   Bustos-Vázquez, E. et al. Survival of COVID-19 with Multimorbidity Patients. *Healthcare*. **9**, 1423; 10.3390/healthcare9111423 (2021).

37.   World Health Organization. International Guidelines for Certification and Classification (Coding) of COVID-19 as Cause of Death. https://www.who.int/publications/m/item/international-guidelines-for-certification-and-classification-(coding)-of-covid-19-as-cause-of-death (2020).

38.   Vander Woude, C.A.et al. Differential care-seeking behaviors during the beginning of the COVID-19 pandemic in Michigan: a population-based cross-sectional study. *BMC public health*. **23**, 2101; 10.1186/s12889-023-16999-5 (2023).

39.   Tan, J., Yoshida, Y., Sheng-Kai, Ma. K., Mauvais-Jarvis, F. Gender Differences in Health Protective Behaviors During the COVID-19 Pandemic in Taiwan: An Empirical Study. *BMC Public Health*. **22**,1900; 10.1101/2021.04.14.21255448 (2022).

40.   Zurochka, A. et al. Seroprevalence of SARS-CoV-2 Antibodies in Symptomatic Individuals Is Higher than in Persons Who Are at Increased Risk Exposure: The Results of the Single-Center, Prospective, Cross-Sectional Study. *Vaccines*. **9**, 627; doi: 10.3390/vaccines9060627 (2021).

41.   Zhang, H. et al. Association of smoking history with severe and critical outcomes in COVID-19 patients: A systemic review and meta-analysis. *European journal of integrative medicine*. **43**,101313; 10.1016/j.eujim.2021.101313 (2021).

42.   Simons, D., Shahab, L., Brown, J., Perski, O. The association of smoking status with SARS-CoV-2 infection, hospitalization and mortality from COVID-19: a living rapid evidence review with Bayesian meta-analyses (version 7). *Addiction*. **116**,1319-1368 (2021).

43.   Testino, G., Pellicano, R. Alcohol consumption in the COVID-19 era. *Minerva gastroenterologica e dietologica*.**66**, 90-92 (2020).

44.   White AM, Castle IP, Powell PA, Hingson RW, Koob GF. Alcohol-Related Deaths During the COVID-19 Pandemic. *Jama*.**327**,1704-1706 (2022).

45.   Benzano, D. et al. Clinical vulnerability for severity and mortality by COVID-19 among users of alcohol and other substances. *Psychiatry research*. **300**,113915; 10.1016/j.psychres.2021.113915 (2021).

46.   Pareek, M. et al. Relation of Cardiovascular Risk Factors to Mortality and Cardiovascular Events in Hospitalized Patients With Coronavirus Disease 2019 (from the Yale COVID-19 Cardiovascular Registry). *The American journal of cardiology*.**146**, 99-106 (2021).

47.   Sabanoglu, C., Inanc, I.H., Polat, E., Peker, S.A. Long-term predictive value of cardiac biomarkers in patients with COVID-19 infection. *European review for medical and pharmacological sciences*. **26**, 6396-6403 (2022).

48.   Rezel-Potts, E. et al. Cardiometabolic outcomes up to 12 months after COVID-19 infection. A matched cohort study in the UK. *PLoS medicine*.**19**, 1004052; 10.1371/journal.pmed.1004052 (2022).

49.   Wu, Z., McGoogan, J.M. Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72 314 Cases From the Chinese Center for Disease Control and Prevention. *Jama*. **323**, 1239-1242 (2020).

50.   Dolby, T. et al. Relationship between asthma and severe COVID-19: a national cohort study. *Thorax*.**78**,120-127 (2023).

51.   Sundberg, R. et al. Asthma in men and women: treatment adherence, anxiety, and quality of sleep. *Respiratory medicine*. **104**, 337-344 (2010).

52.   Halpin, D.M.G., Faner, R., Sibila, O., Badia, J.R., Agusti, A. Do chronic respiratory diseases or their treatment affect the risk of SARS-CoV-2 infection? *The Lancet Respiratory medicine*. **8**, 436-438 (2020).

53.   Hippisley-Cox, J. et al. Risk prediction of covid-19 related death and hospital admission in adults after covid-19 vaccination: national prospective cohort study. BMJ (Clinical research ed). 374, 2244; 10.1136/bmj.n2244 (2021).

54.   Peters, S.A.E., MacMahon, S., Woodward, M. Obesity as a risk factor for COVID-19 mortality in women and men in the UK biobank: Comparisons with influenza/pneumonia and coronary heart disease. *Diabetes, obesity & metabolism*. **23**, 258-262 (2021).

55.   Kholmatova, K. et al. Obesity Prevalence and Associated Socio-Demographic Characteristics and Health Behaviors in Russia and Norway. *International journal of environmental research and public health*. **19**, 9428; 10.3390/ijerph19159428 (2022).

56.   Booth, A., Magnuson, A., Fouts, J., Foster, M.T. Adipose tissue: an endocrine organ playing a role in metabolic regulation. *Hormone molecular biology and clinical investigation*.**26**, 25-42 (2016).

57.   Aghili, S.M.M. et al. Obesity in COVID-19 era, implications for mechanisms, comorbidities, and prognosis: a review and meta-analysis. *International journal of obesity (2005)*. **45**, 998-1016 (2021).

58.   Anderson, K.M., Castelli, W.P., Levy, D. Cholesterol and mortality. 30 years of follow-up from the Framingham study. *Jama*. **257**, 2176-2180 (1987).

59.   Ge, E., Li, Y., Wu, S., Candido, E., Wei, X. Association of pre-existing comorbidities with mortality and disease severity among 167,500 individuals with COVID-19 in Canada: A population-based cohort study. *PloS one*. **16**, 0258154; 10.1371/journal.pone.0258154 (2021).

60.   Ssentongo, P., Ssentongo, A.E., Heilbrunn, E.S., Ba, D.M., Chinchilli, V.M. Association of cardiovascular disease and 10 other pre-existing comorbidities with COVID-19 mortality: A systematic review and meta-analysis. *PloS one*. **15**, 0238215; 10.1371/journal.pone.0238215 (2020).

61.   Shlipak, M.G. et al. Cystatin C and the risk of death and cardiovascular events among elderly persons. *The New England journal of medicine*. **352**,2049-2060 (2005).

62.   Helmersson-Karlqvist, J. et al. Addition of cystatin C predicts cardiovascular death better than creatinine in intensive care. *Heart (British Cardiac Society)*. **108**, 279-284 (2022).

63.   Ho, F.K. et al. Modifiable and non-modifiable risk factors for COVID-19, and comparison to risk factors for influenza and pneumonia: results from a UK Biobank prospective cohort study. *BMJ open*. **10**, 040402; 10.1136/bmjopen-2020-040402 (2020).

64.   Kermali, M., Khalsa, R.K., Pillai, K., Ismail, Z., Harky, A. The role of biomarkers in diagnosis of COVID-19 - A systematic review. *Life sciences*. **254**, 117788; 10.1016/j.lfs.2020.117788 (2020).

65.   Fundora MP, Kamidani S, Oster ME. COVID Vaccination as a Strategy for Cardiovascular Disease Prevention. *Current cardiology reports*. **25**,1327-1335 (2023).