The clinical spectrum of COVID-19: A population-based cohort study in Iceland

<table>
<thead>
<tr>
<th>Journal</th>
<th>BMJ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>BMJ-2020-061220</td>
</tr>
<tr>
<td>Article Type</td>
<td>Research</td>
</tr>
<tr>
<td>BMJ Journal</td>
<td>BMJ</td>
</tr>
<tr>
<td>Date Submitted by the Author</td>
<td>12-Aug-2020</td>
</tr>
<tr>
<td>Complete List of Authors</td>
<td>Eythorsson, Elias; Landspitali, Internal Medicine Services Helgason, Dadi; Landspitali, Internal Medicine Services Ingvarsson, Ragnar; Landspitali, Internal Medicine Services Bjornsson, Helgi; Landspitali, Internal Medicine Services Olafsdottir, Lovisa; Landspitali Bjarnadottir, Valgerdur; Landspitali, Internal Medicine Services Runolfsdottir, Hrafnhildur; Landspitali, Internal Medicine Services Bjarnadottir, Solveig; University of Iceland School of Health Sciences, Faculty of Medicine Agustsson, Arnar; Landspitali Oskarsdottir, Kristin; Landspitali Thorvaldsson, Hrafn; Landspitali Kristjansdottir, Gudrun; University of Iceland School of Health Sciences, Faculty of Medicine Armanndottir, Brynja; Landspitali Bjarnason, Agnar; Landspitali Johannsson, Birgir; Landspitali Gudlaugsson, Olafur; Landspitali Gottfreðsson, Magnús; University of Iceland School of Health Sciences, Faculty of Medicine Sigurðsson, Martin; Landspitali, Division of Anaesthesia and Intensive Care Medicine; University of Iceland School of Health Sciences, Faculty of Medicine Indridason, Olafur; Landspitali Palsson, Runolfur; Landspitali, Internal Medicine Services; University of Iceland School of Health Sciences, Faculty of Medicine</td>
</tr>
<tr>
<td>Keywords</td>
<td>SARS-CoV-2, Covid-19</td>
</tr>
</tbody>
</table>

https://mc.manuscriptcentral.com/bmj
The clinical spectrum of COVID-19: A population-based cohort study in Iceland

Elias Eythorsson M.D., Ph.D.¹, Dadi Helgason M.D., Ph.D.¹, Ragnar Freyr Ingvarsson M.D.¹, Helgi K Bjornsson M.D.¹, Lovisa Bjork Olafsdottir M.D.¹, Valgerdur Bjarnadottir M.D.¹, Hrafnhildur Linnet Runolfsdottir M.D.¹, Solveig Bjarnadottir B.Sc.¹,², Arnar Snaer Agustsson M.D.¹, Kristin Oskarsdottir M.D.¹, Hrafn Hliddal Thorvaldsson M.D.¹, Gudrun Kristjansdottir B.Sc.¹,², Brynja Armannsdottir M.D., M.P.H.¹, Agnar Bjarnason M.D., Ph.D.¹,², Birgir Johannsson M.D.¹, Olafur Gudlaugsson M.D.¹, Magnus Gottfredsson M.D., Ph.D.¹,², Martin I Sigurdsson M.D., Ph.D.¹,², Olafur S Indridason M.D., M.H.S.¹, Runolfur Palsson M.D.¹,²

¹Landspitali–The National University Hospital of Iceland
²Faculty of Medicine, School of Health Sciences, University of Iceland

Corresponding author is Runolfur Palsson M.D.
Address: Landspitali–The National University Hospital of Iceland, Reykjavik, Iceland
Email: runolfur@landspitali.is
Telephone: +354-824-5827
Contributorship statements:

Elias Eythorsson – Conception and design, acquisition of data, analysis and interpretation of data and drafting of the manuscript.

Dadi Helgason - Conception and design, acquisition of data, interpretation of data and drafting of the manuscript

Ragnar Freyr Ingvarsson - Conception and design, and drafting of the manuscript

Helgi K Bjornsson - Conception and design, acquisition of data and drafting of the manuscript

Lovisa Bjork Olafsdottir - Conception and design, acquisition of data and drafting of the manuscript

Valgerdur Bjarnadottir - Acquisition of data and revision of manuscript

Hrafnhildur Linnet Runolfsdottir - Acquisition of data and revision of manuscript

Solveig Bjarnadottir - Acquisition of data and revision of manuscript

Arnar Snaer Agustsson - Acquisition of data and revision of manuscript

Kristin Oskarsdottir - Acquisition of data and revision of manuscript

Hrafn Hliddal Thorvaldsson - Acquisition of data and revision of manuscript

Gudrun Kristjansdottir - Acquisition of data and revision of manuscript

Brynja Armannsdottir - Conception and design and revision of manuscript

Agnar Bjarnason - Conception and design and drafting of the manuscript

Birgir Johannsson - Conception and design and revision of manuscript

Olafur Gudlaugsson - Conception and design and revision of manuscript

Magnus Gottfredsson - Conception and design and revision of manuscript

Martin I Sigurdsson - Conception and design, interpretation of data and revision of manuscript

Olafur S Indridason - Conception and design, interpretation of data and revision of manuscript

Runolfur Palsson - Conception and design, interpretation of data and revision of manuscript

Elias Eythorsson and Runolfur Palsson are guarantours of the study.

All authors gave final approval of the submitted manuscript.

Conflict of interest statement:

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; and no other relationships or activities that could appear to have influenced the submitted work.

Copyright statement:

The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, an exclusive licence (or non exclusive for government employees) on a worldwide basis to the BMJ Publishing Group Ltd to permit this article (if accepted) to be published in BMJ editions and any other BMJ PGL products and sublicences such use and exploit all subsidiary rights, as set out in our licence.
Transparency declaration:
The lead author affirms that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

Details of ethical approval:
The study was approved by the National Bioethics Committee (VSN-20-078).

Details of funding and the role of the study sponsors:
No specific funding was provided for the study. A statement of independence from the study funders is therefore not applicable.
Key points

Question: What is the frequency and progression of various symptoms experienced by patients with COVID-19?

Findings: In this population-based cohort study that included all SARS-CoV-2-positive patients in Iceland, most patients (67.5%) had mild symptoms throughout their disease course. At the time of diagnosis, 5.3% were asymptomatic, of whom roughly half developed symptoms during follow-up. Common presenting symptoms included myalgia (55%), headache (51%), and non-productive cough (49%). At diagnosis, 13.8% and 22.3% did not meet the Centers for Disease Control and Prevention and World Health Organization case definitions for suspected COVID-19, respectively.

Meaning: In the setting of broad access to diagnostic testing, the majority of SARS-CoV-2-positive patients were found to have mild symptoms and almost one-fifth did not meet published clinical criteria for RT-PCR testing.
Abstract

Background: Previous studies on the epidemiology and clinical characteristics of COVID-19 have generally been limited to hospitalized patients. The aim of this study was to describe the complete clinical spectrum of COVID-19, based on a nationwide cohort with extensive diagnostic testing and a rigorous contact tracing approach.

Methods: A population-based cohort study examining symptom progression using prospectively recorded data on all individuals with a positive test (RT-PCR) for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) who were enrolled in a telehealth monitoring service provided to all identified cases in Iceland. Symptoms were systematically monitored from diagnosis to recovery.

Results: From January 31 to April 30, 2020, a total of 45,105 individuals (12% of the Icelandic population) were tested for SARS-CoV-2, of whom 1797 were positive, yielding a population incidence of 5 per 1000 individuals. The most common presenting symptoms were myalgia (55%), headache (51%), and non-productive cough (49%). At the time of diagnosis, 5.3% of cases reported no symptoms and 3.1% remained asymptomatic during follow-up. In addition, 216 patients (13.8%) and 349 patients (22.3%) did not meet the case definition of the Centers for Disease Control and Prevention and the World Health Organization, respectively. The majority (67.5%) of patients had mild symptoms throughout the course of the disease.

Conclusion: In the setting of broad access to diagnostic testing, the majority of SARS-CoV-2-positive patients were found to have mild symptoms. Fever and dyspnea were
less common than previously reported. A substantial proportion of patients did not meet recommended case definitions at the time of diagnosis.
Introduction

On December 31, 2019, the first cases of an atypical pneumonia of unidentified etiology were reported in Wuhan, China.¹ One week later, a novel betacoronavirus, later named severe acute respiratory syndrome coronavirus (SARS-CoV-2), was identified as the causative pathogen²,³, and the disease subsequently termed coronavirus disease 2019 (COVID-19). The World Health Organization (WHO) declared the COVID-19 outbreak a pandemic on March 11, 2020.⁴

COVID-19 has a wide range of clinical manifestations, ranging from an asymptomatic state or mild respiratory symptoms to severe viral pneumonia and acute respiratory distress syndrome.⁵-⁷ Previous publications have suggested that approximately 81% of patients have mild symptoms, 14% have severe symptoms, and 5% become critically ill.⁸ Besides respiratory symptoms, dysosmia, dysgeusia, abdominal pain, diarrhea, and rash have been described.⁵,⁹ Most published studies on the clinical characteristics of COVID-19 have been retrospective⁵,⁸,¹⁰,¹¹ and limited to inpatients¹²,¹³ and therefore do not capture the full clinical spectrum of the disease.

The first case of COVID-19 in Iceland was diagnosed on February 27, 2020.¹⁴ Icelandic health authorities responded immediately by isolating patients and instituting systematic contact tracing and quarantine of exposed individuals.¹⁵ Broad access to diagnostic testing became available in Iceland early in the course of the pandemic, allowing the highest rate of SARS-CoV-2 testing in the world.¹⁶ Approximately one month after the first case was identified, the incidence of undetected cases was found to be only 0.6% using random population screening.¹⁴ All SARS-CoV-2-positive individuals were actively
monitored at a newly established COVID-19 outpatient clinic at Landspitali–The National University Hospital (LUH). The contact tracing and containment strategies implemented by the Icelandic authorities rapidly curbed the epidemic, with only 4 new cases diagnosed between May 1 and 15.

In this paper, we describe the analysis of prospectively collected data on all SARS-CoV-2-positive patients in Iceland and characterize the epidemiology and full clinical spectrum of COVID-19 in a nationwide cohort.
Methods

Study population and design

This population-based cohort study included all patients who tested positive for SARS-CoV-2 between February 27 and April 30, 2020, and were actively monitored at LUH. All individuals in Iceland who tested positive for SARS-CoV-2 were immediately contacted, instructed to isolate, and enrolled in a telehealth monitoring service. Monitoring involved frequent telephone interviews by a nurse or physician, through which the patient’s clinical status was evaluated. From February 27 to March 16, 2020, the content and documentation of these interviews was at the discretion of the nurse or physician making the call. On March 17, a standardized data entry form was built directly into the national electronic medical record system, facilitating a structured approach to the clinical evaluation of COVID-19 patients. The study was approved by the National Bioethics Committee (VSN-20-078).

Virological testing

Real-time reverse-transcriptase polymerase chain reaction (RT-PCR) was used for detection of SARS-CoV-2 RNA in nasopharyngeal and oropharyngeal swabs, based on the WHO recommended protocol from Charité, Berlin or using a commercial kit (TaqMan 2019-nCoV Assay from Thermo Fisher Scientific) as previously described. Three testing protocols were implemented: targeted testing, open-invitation population screening, and random population screening. Targeted testing was performed at the Department of Clinical Microbiology at LUH, whereas population screening was carried out by deCODE genetics, a biopharmaceutical company based in Reykjavik. Targeted
testing began on January 31, 2020, and included clinically suspected cases and
individuals at high risk of exposure. Open-invitation population screening began on
March 13, 2020, and was available to all Icelandic residents who were not currently
quarantined and did not have symptoms that would prompt targeted testing. Finally, a
randomly chosen sample of 6782 Icelanders was offered testing via telephone text
message on March 31 and April 1, 2020, of whom 2283 were included.

Data collection

To confirm the completeness of telehealth enrollment, results of all SARS-CoV-2 testing
were obtained from databases of LUH and deCODE genetics. Negative samples were
used for denominator calculations. Baseline characteristics of SARS-CoV-2-positive
patients and longitudinal data on symptom progression were obtained from the
standardized data entry forms used by the COVID-19 Clinic. Data on clinical outcomes
were extracted from LUH database. Data were linked using government issued national
identification numbers. Population demographics were obtained from Statistics Iceland
(www.statice.is).

The telehealth monitoring service

The initial patient interviews were conducted by a physician who informed the patients
of the diagnosis, evaluated their health, and instructed them to self-isolate at home. A
checklist of 19 specific symptoms was presented to each patient during the initial and all
subsequent interviews (Table 1). Patients were asked whether they had experienced
any of the 19 symptoms from the time of symptom onset to the time of the interview.
Additionally, the patient’s baseline characteristics were documented, including past medical history, medication use, and social history.

Based on the symptoms documented during the interviews, patients were classified into one of three categories of clinical severity: low severity, defined as mild and improving symptoms; moderate severity, defined as mild dyspnea, cough or fever for less than five days; and high severity, defined as worsening dyspnea, worsening cough, high or persistent fever for five days or longer, or severe fatigue. The frequency of follow-up interviews ranged from daily to every fourth day, depending on clinical severity, age, and underlying conditions. Patients with concerning symptoms were referred for evaluation at the COVID-19 Clinic.

Patients were discharged from telehealth monitoring when they met both of the following criteria: 14 days had passed since the diagnosis of COVID-19, and they had been asymptomatic for the past seven consecutive days.

**Case definitions**

Patients were considered symptomatic due to COVID-19 using three different definitions: 1) Reporting any of the 19 symptoms; 2) WHO case definition of suspected COVID-19, which included fever and at least one other symptom of acute respiratory infection, e.g. cough or shortness of breath\(^{19}\); 3) Centers for Disease Control and Prevention (CDC) interim case definition of COVID-19, which included either two of the following symptoms: fever, rigor, myalgia, headache, sore throat, dysosmia or dysgeusia, or one of the following symptoms: cough, shortness of breath or difficulty breathing.\(^{20}\) Patients who did not fulfill these definitions at the time of diagnosis or
during follow-up, were determined to be asymptomatic. Otherwise, they were considered presymptomatic at the time of diagnosis.

Statistical analysis

The incidence of COVID-19 was calculated by age and sex, using both the Icelandic population and all SARS-CoV-2 tested individuals as denominators. Patients were followed until discharge from the telehealth monitoring service, hospital admission, death or end of the study period (May 22, 2020). Patients enrolled prior to the implementation of the standardized data entry form (March 17, 2020) were excluded from the analysis of symptoms and symptom progression.

The progression of COVID-19 symptoms was analyzed using parametric cure-mixture models and logistic regression. The time frame used in the analyses was days from symptom onset to discharge from telehealth monitoring. The data were interval censored due to the intermittent schedule of the interviews. To account for this, the cumulative incidence of symptom occurrence was estimated with the non-parametric Turnbull estimator and a parametric cure-mixture model using the log-logistic distribution. The proportion of patients experiencing a specific symptom per day was estimated employing logistic regression that allowed for non-linear effects using a three-knot restricted cubic spline. Missing information between interviews was addressed by multiple imputation using chained equations, performed by additive regression, bootstrapping and predictive mean matching procedure. The number of imputations equaled the highest proportion of missing data for any variable, multiplied by 100. The added uncertainty due to imputation was fully accounted for in the logistic regression
models. The result was compared to a complete-case analysis and naive up-down filling procedure.

All statistical analyses were performed in R version 3.6.3.
Results

Demographics and clinical characteristics

A total of 45,105 individuals (12% of the Icelandic population) underwent 47,800 tests for SARS-CoV-2 from January 31 to April 30, 2020. Of those, 18,023 (37.7%) were carried out as part of targeted testing and 29,785 (62.3%) were done as part of population screening. Altogether, 1797 patients experienced 1911 positive tests. The population incidence of COVID-19 was 4.9 cases per 1000 individuals, and 40.0 per 1000 tests performed were positive. This rate differed between targeted testing (1703 of 18,023 samples positive or 94.4 per 1000 tests performed) and population screening (208 of 29,785 samples positive or 7.0 per 1000 tests performed). Among tested individuals, the proportion of females (58%) was higher than in the general population (49%). However, the sex distribution was equal among patients who tested positive (Supplemental Figure 1).

All patients who tested positive were enrolled in telehealth monitoring or admitted to hospital, with no exceptions. The clinical progression from symptom onset for the 1797 positive individuals is shown in Figure 1. Thirty-two of those patients were diagnosed after being admitted to hospital of whom one was diagnosed post-mortem (Figure 2). In total, 101 (5.7%) patients were hospitalized, 27 (1.5%) were admitted to intensive care and 16 (0.9%) required mechanical ventilation. The median length of hospital stay was 8 days (IQR, 3.5-19) and 10 (0.6%) patients died.
Symptoms at diagnosis

Of the 1797 SARS-CoV-2-positive patients, 41.5% had experienced fever at diagnosis. Other symptoms such as cough (59.2%), dyspnea (30.4%), and gastrointestinal symptoms (32.3%) were also commonly reported. A total of 100 (5.7%) patients reported no symptoms at diagnosis, which was more common among those who were diagnosed through population screening (22.0%) as compared to targeted testing (5.5%). The total number of documented symptoms at diagnosis was lower among patients diagnosed before the implementation of standardized data collection compared to those diagnosed following the implementation (Supplemental Figure 2). Therefore, this latter group was used to study the symptom development and progression in detail (Figure 2).

Of the 1564 (87%) patients who were followed using the standardized data entry form, 791 (51%) were female and the median age was 40 years (IQR, 26-53; range, 0-103). The distribution of age and sex is shown in Supplemental Figure 1. Among these patients, 1055 (67.5%) were classified as having low disease severity throughout their follow-up, 55 (3.5%) were admitted to hospital and 13 (0.8%) required intensive care, six (0.4%) of whom received mechanical ventilation. Two (0.1%) of these patients died. The patients were followed for a median of 15 days (IQR, 14-18) and contributed 69.2 person-years of follow-up time. A total of 1509 patients completed telehealth follow-up from diagnosis until discharge. The observation time of the remaining 55 patients who were hospitalized was censored at the time of admission, after a median follow-up of 4 days (IQR, 2.5-8). Demographic information by subgroup of included and excluded patients is presented in Supplemental Table 1.
Of the 1564 patients with standardized symptom documentation, 42.7% had experienced fever at diagnosis. Cough (60.1%), dyspnea (31.7%), and gastrointestinal symptoms (35.7%) were also commonly reported. Eighty-three (5.3%) patients reported no symptoms at diagnosis, which was more commonly observed among those diagnosed through population screening (24.5%) as compared with patients diagnosed by targeted testing (3.2%). Of the patients who were asymptomatic at diagnosis, 49 (59.0%) remained asymptomatic throughout their telehealth monitoring, while the remaining 34 (41.0%) patients developed symptoms after a median of 3 days (IQR, 3-4.75).

Using the CDC case definition, 216 (13.8%) patients would be categorized as not having symptoms consistent with COVID-19 at the time of diagnosis. This was more common among patients diagnosed through population screening (40.6%) than by targeted testing (10.9%). Seventy-two (4.6%) of those patients developed symptoms compatible with the CDC case definition at a median of 5 days (IQR, 3-6) from diagnosis, whereas the other 144 (9.2%) never met the CDC criteria. Similarly, 349 (22.3%) patients did not fulfill the WHO case definition at the time of diagnosis, a finding that was more commonly observed among patients diagnosed through population screening (45.8%) than by targeted testing (19.7%). The WHO criteria were later met by 115 (7.4%) patients at a median of 4 days (IQR, 3-6) from diagnosis, while the remaining 234 (15.0%) patients never fulfilled the criteria.

The cumulative incidence and proportion of patients meeting CDC and WHO criteria by number of days from symptom onset is shown in Supplemental figure 3. Among the 216 individuals who did not meet CDC criteria at diagnosis, four (1.9%) were hospitalized...
later in the course of their disease, and one required mechanical ventilation. Similarly, four (1.1%) of the 349 individuals who did not fulfill the WHO criteria at diagnosis were admitted to hospital for illness related to COVID-19, two of whom required intensive care.

Symptom development and progression assessed by multiple imputation

The median time from symptom onset until RT-PCR diagnosis and enrollment interview was four days (IQR, 2-7) and a median of six (IQR, 4-8) interviews were conducted per patient during telehealth monitoring. The median time between interviews was two days (IQR, 1-3). No data were missing for the days during which interviews occurred. Symptoms occurring on days during which interviews were not conducted were imputed, and the imputation procedure was repeated 92 times. The proportion of patients with missing information on symptoms was highest during the first days following symptom onset (Supplemental Table 2). The proportion of patients experiencing specific symptoms per day from symptom onset was calculated using multiple imputation logistic regression, which produced acceptable results as compared with complete-case analysis and a naïve up-down filling procedure (Supplemental Figures 4-7).

Symptoms at disease onset

As shown in Table 1, the most common symptoms at the onset of COVID-19 were myalgia, headache, and non-productive cough, observed in 51% (95%CI, 47%-55%), 49% (95%CI, 45%-53%), and 55% (95%CI, 51%-60%), respectively. However, 82% (95%CI, 80%-85%) of patients experienced at least one generalized symptom and 63% (95%CI, 59%-67%) at least one upper respiratory symptom. Compared with patients
diagnosed by targeted testing, those diagnosed through population screening were less likely to have experienced cardinal symptoms of COVID-19 at disease onset, including fever (30% vs. 43%), cough (47% vs. 59%), dyspnea (18% vs. 26%) and gastrointestinal symptoms (23% vs. 31%). Other symptoms according to testing protocol are shown in Supplemental Figures 8-11.

The proportion of patients experiencing specific symptoms by sex and age group is displayed in Supplemental Figures 12-15. The initial presentation of COVID-19 varied only slightly between the sexes. The proportion experiencing fever or gastrointestinal symptoms at onset was similar between age groups, but cough and dyspnea were more common among older individuals (Supplemental Figure 14).

Differences in symptoms at onset between hospitalized and non-hospitalized patients are shown in Supplementary Figures 16-19. Several symptoms at disease onset were more common among patients who were later admitted to hospital, including generalized symptoms (94% vs 82%), lower respiratory symptoms (71% vs 62%) and gastrointestinal symptoms (48% vs 29%). However, the proportion of patients who experienced upper respiratory symptoms was lower, 48% compared with 64% of those who were never hospitalized.

**Progression of symptoms**

By day 21 from disease onset, the most commonly experienced symptoms were lethargy, headache, and productive or non-productive cough, noted in 74% (95%CI, 72%-77%), 73% (95%CI, 70%-75%) 73% (95%CI 70%-75%), respectively. Overall, 93% (95%CI, 91%-94%) had experienced at least one generalized symptom, 87%
(95%CI, 85%-89%) at least one upper respiratory symptom, and 80% (95%CI, 78%-82%) at least one lower respiratory symptom. The cumulative incidence of fever, dyspnea, and gastrointestinal symptoms were 49%, 52%, and 53%, respectively (Figure 3). The proportion of patients experiencing each symptom by days from symptom onset is shown in Figure 4. Of the 741 patients who experienced fever at any time during the course of the disease, 630 (85%) had done so by day 3. Furthermore, 902 of 1140 patients (79%) had developed cough by day 3, 427 of 826 (52%) had developed dyspnea and 485 of 788 (62%) had developed any gastrointestinal symptom by day 3 (Figure 3). These proportions were comparable for both sexes and all age groups (Supplemental Figures 20-31).

Of the 19 symptoms, only dysosmia and dysgeusia were more common later in the disease course than at symptom onset. Both symptoms peaked on day eight from diagnosis. The trend was most pronounced among patients aged 25 to 55 years and was more marked among females (Supplemental Figure 13). Other symptoms attributed to COVID-19 were most prevalent during the onset of the disease. No symptom exhibited a bimodal pattern (Supplemental Figures 12-15). The cumulative incidence of each specific symptom was lower among patients diagnosed through population screening compared to targeted testing, except for rhinorrhea and vomiting (Supplemental Figures 32-35). By day 21, a large proportion of patients who were ever admitted to hospital for COVID-19 had experienced fever, dyspnea and and gastrointestinal symptoms, or 91% (95%CI, 71%-95%), 83% (95%CI, 64%-91%), and 84% (95%CI, 66%-91%), respectively (Supplemental Figures 36-39).
Discussion

In this study we present the clinical characteristics of COVID-19 in a national population-based cohort. Due to aggressive contact tracing and widespread virological testing it is likely that the cohort includes the majority of symptomatic cases in the population. This assumption is supported by the low prevalence of the disease detected by random population screening (0.6%).\(^\text{15}\) Our prospectively collected data regarding symptoms and disease progression among patients who tested positive for SARS-CoV-2 in Iceland revealed that 49% of patients experienced fever, 73% cough, and 52% dyspnea. At the time of diagnosis, 5% were completely asymptomatic, 13% did not meet the CDC case definition, and 22% did not fulfill the WHO criteria.

The comprehensive, nationwide characterization of COVID-19 symptoms was facilitated by broad access to diagnostic testing in Iceland. The Icelandic healthcare system is a single-payer system with a universal government-run health insurance provider. The SARS-CoV-2 RT-PCR test was free of charge for both targeted testing and population screening, resulting in over 12% of the population tested, which was higher than in any other country during the study period.\(^\text{16}\) As a result, we were able to describe the true spectrum of COVID-19, while previous studies were largely based on hospitalized cohorts or cases identified in the setting of more restrictive testing.\(^\text{6,12,22}\) Furthermore, subgroup analysis allowed us to quantify the degree by which cohorts with only hospitalized patients might overestimate the presence of specific symptoms, for instance fever, which was considerably more common at symptom onset in patients
who were later admitted to hospital (74%) than in patients who never were hospitalized (40%).

The proportion of patients with COVID-19 who were completely asymptomatic has been a focus of interest during the pandemic with implications for the risk of disease dissemination. A recent report from Iceland examining an overlapping cohort suggested that 43% of patients were asymptomatic at the time of sampling. This estimate was derived from 43 out of 100 RT-PCR positive cases reporting no symptoms at the time of diagnosis among 13,080 participants in a population screening program, which included both open-invitation and random invitation screening. Follow-up information on the patients’ symptoms was not available to the authors, possibly resulting in an overestimation of asymptomatic individuals. In the current study, 83 patients reported no symptoms at the time of diagnosis, approximately half of whom developed symptoms in the ensuing days. Thus only 3.1% of diagnosed cases remained completely symptom-free during follow-up. However, as some degree of suspicion of COVID-19 was needed to prompt an individual to be tested, symptomatic patients are likely to be overrepresented in our sample. Estimating the true proportion of SARS-CoV-2-positive patients who never develop symptoms is difficult, as asymptomatic patients are unlikely to be diagnosed outside of random population screening, and differentiating pre-symptomatic from asymptomatic patients requires longitudinal follow-up. Gudbjartsson et al. found that 13 out of 2283 randomly sampled Icelanders were SARS-CoV-2-positive, of which 7 (53.8%) reported no symptoms at the time of diagnosis. We have shown that 59.0% of patients who are symptom free at diagnosis, never develop
symptoms. Based on these observations, the rough estimate of the proportion of truly asymptomatic SARS-CoV-2-positive Icelanders is approximately 30%.

Of symptomatic patients, most experienced only minor symptoms. Only 22% of patients developed moderate symptoms, 8% severe symptoms, and 3.5% were hospitalized.

The standardized prospective recording of clinical symptoms made it possible to evaluate the sensitivity of the widely used CDC\textsuperscript{20} and WHO\textsuperscript{19} case definitions for the diagnosis of COVID-19 throughout the course of the disease. By applying these definitions, we demonstrate that a substantial number of cases would have been missed; approximately 9% by the CDC criteria and 15% by the WHO criteria. The identification of additional 4% and 7% of cases would have been delayed by a median of 5 and 4 days, respectively. These are concerning results with immediate implications for current efforts to curtail the pandemic. Our data show that most patients have mild symptoms that may not have prompted the consideration of COVID-19 by either patients or health care providers in more resource-limited settings, and indicates a need for revising and widening the CDC and WHO case definitions to increase their sensitivity.

Symptoms observed among patients with mild forms of COVID-19 have previously been examined in a multicenter European study of 1420 RT-PCR positive patients who answered a questionnaire.\textsuperscript{10} Severely ill patients were excluded and the remaining cohort was predominantly female (68%), young (94% were <60 years of age), and biased towards healthcare workers (31% of the group).\textsuperscript{10} While these results are not easily generalizable to the entire population, the investigators found that only 7% of patients required hospitalization\textsuperscript{10} compared to 3.5% in the present study. Headache,
loss of smell, and nasal obstruction were the most common symptoms identified.\(^\text{10}\) Although these symptoms were also frequently identified in our cohort, we found cough and myalgias to be more common. The predominance of loss of smell identified in the aforementioned study agrees with our observation that olfactory symptoms are most common in younger age groups.\(^\text{10}\)

We found that slightly less than 50% of patients developed fever during the course of the disease, already present in 85% of those by day 3. This is consistent with the study by Lechien et al. in which fever $\geq 38.0^\circ$C was reported in 45.4% of cases\(^\text{10}\), while it is higher than was reported by Guan et al. (21.7%)\(^\text{5}\) and Goyal et al. (25%)\(^\text{12}\). A meta-analysis by Sun et al. found that 89% of COVID-19 patients had a fever $\geq 37.3^\circ$C\(^\text{6}\), but this definition of fever is rarely used in clinical practice. Over the follow-up period, 70% of patients experienced cough which is consistent with the findings of Lechien et al. who observed cough in 63% of cases\(^\text{10}\). In total, 52% of patients reported any dyspnea and only 13% reported dyspnea at rest during the disease. The reported incidence of dyspnea ranges from 22-49%\(^\text{7,10,23}\) but most previous studies do not differentiate between dyspnea at rest and on exertion. Gastrointestinal symptoms were common, reported by almost half of patients at some point during the first 14 days. Abdominal pain (22%) and diarrhea (28%) were frequent as in previous studies.\(^\text{12,24}\) Interestingly, although one-fourth of patients experienced nausea, vomiting was rare.

Our findings indicate a lower rate of hospital admissions and mortality in Iceland compared with many other countries. The reasons for these disparate outcomes are likely multifactorial. Iceland has a relatively young population, with 85.8% younger than
65 years, compared with 77.1% in Italy, 80.6% in Spain, 81.6% in the United Kingdom, and 83.5% in the USA. This, in addition to a strong emphasis on limiting exposure of elderly and multimorbid individuals, resulted in a low median age of confirmed COVID-19 cases of 40 years (IQR, 26-53) in Iceland compared to 51 years (IQR, 36-65) among all cases reported to the WHO and 48 years (IQR, 33-63) in the USA. Different rates of other recognized risk factors for worse outcomes in COVID-19 are unlikely to explain this difference as they have a similar distribution in Iceland as in other countries, for instance 28.8% of Icelandic adults have hypertension and 27% are obese.

This study does have some limitations. In order to accurately describe symptom progression, cases diagnosed before the implementation of the standardized clinical data entry form were excluded from the analysis of symptom development, representing 11% of all SARS-CoV-2-positive cases in Iceland. The date of implementation of the standardized data entry form was not influenced by the clinical characteristics of the patients being diagnosed, and therefore should not introduce bias. Furthermore, the demographics and clinical characteristics of excluded cases were largely comparable to those that were included in the study. Another limitation is that daily standardized documentation of symptoms was not available during hospital admission. This could conceivably lead to an underestimation of severe symptoms such as dyspnea. However, only 3.5% of the included patients were hospitalized, and symptoms prior to hospitalization were included in the analysis. It is important to note that the data were based on self-reported symptoms via telephone calls. This shortcoming is mitigated by the fact that experienced nurses and physicians conducted the interviews.
strength of the study is its population-based approach, which included all confirmed cases in the country during the study period, regardless of their need for medical care.

### Conclusion

This study describes the symptomatology and clinical severity of the initial phase of COVID-19 in Iceland. The incidence of COVID-19 was high due to extensive testing of both symptomatic and asymptomatic individuals, while disease severity was lower than previously reported. Symptoms such as fever and dyspnea were less frequent than has been observed in earlier studies. Our findings suggest that both the CDC and WHO case definitions of COVID-19 lack sensitivity and miss a substantial proportion of patients, including cases who later develop severe disease.
References


30. OECD/European Observatory on Health Systems and Policies. Iceland: country health profile 2019, state of health in the EU.
# Tables

## Table 1. Proportion of SARS-CoV-2-positive patients who experienced specific symptoms and symptom constellations on days 1, 3, 7, and 14 from symptom onset.

The cumulative incidence of the specific symptoms by day 14 is shown in the last column. All values are reported as percentages with 95% confidence intervals (CI).

<table>
<thead>
<tr>
<th>Constellation of symptoms</th>
<th>Specific symptoms</th>
<th>Proportion of patients experiencing a symptom per day (%, 95%CI)</th>
<th>Cumulative incidence by day 14 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Symptom onset</td>
<td>Day 3</td>
<td>Day 7</td>
</tr>
<tr>
<td>Generalized</td>
<td>82 (80-85)</td>
<td>78 (75-80)</td>
<td>65 (64-67)</td>
</tr>
<tr>
<td>-</td>
<td>Fever ≥38°C</td>
<td>42 (36-48)</td>
<td>33 (29-37)</td>
</tr>
<tr>
<td>-</td>
<td>Rigor, chills</td>
<td>32 (27-38)</td>
<td>24 (21-28)</td>
</tr>
<tr>
<td>Symptom</td>
<td>Percentage (95% CI)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------------</td>
<td>---------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Headache</td>
<td>51 (47-56)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>46 (42-49)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>35 (34-37)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>21 (20-22)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>73 (70-75)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Myalgia</td>
<td>55 (51-60)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>45 (42-49)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>27 (26-29)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>10 (10-11)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>62 (59-64)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lethargy</td>
<td>38 (34-42)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>38 (34-41)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>36 (35-38)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>27 (26-29)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>74 (72-77)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Loss of appetite</td>
<td>24 (20-28)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>23 (20-26)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>20 (19-22)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>11 (11-12)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>46 (43-49)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upper respiratory</td>
<td>63 (59-67)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>60 (57-63)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>54 (53-56)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>37 (36-38)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>87 (85-89)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rhinorrhea</td>
<td>33 (28-38)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>28 (26-31)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>21 (20-22)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>12 (11-13)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>55 (53-58)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sore throat</td>
<td>33 (29-38)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>27 (25-30)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>18 (17-19)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>10 (9-11)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>45 (43-48)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Condition</td>
<td>Dysosmia</td>
<td>Dysgeusia</td>
<td>Lower respiratory</td>
</tr>
<tr>
<td>-----------</td>
<td>----------</td>
<td>-----------</td>
<td>-------------------</td>
</tr>
<tr>
<td></td>
<td>21 (18-25)</td>
<td>24 (21-26)</td>
<td>28 (27-29)</td>
</tr>
<tr>
<td></td>
<td>24 (21-28)</td>
<td>26 (24-29)</td>
<td>30 (29-32)</td>
</tr>
<tr>
<td></td>
<td>62 (58-65)</td>
<td>60 (58-63)</td>
<td>56 (55-58)</td>
</tr>
<tr>
<td></td>
<td>49 (45-53)</td>
<td>45 (42-48)</td>
<td>37 (36-39)</td>
</tr>
<tr>
<td></td>
<td>13 (11-16)</td>
<td>14 (12-17)</td>
<td>16 (15-17)</td>
</tr>
<tr>
<td></td>
<td>25 (22-29)</td>
<td>26 (24-29)</td>
<td>28 (26-29)</td>
</tr>
<tr>
<td></td>
<td>5.0 (3.8-6.7)</td>
<td>5.0 (4.0-6.2)</td>
<td>4.8 (4.2-5.5)</td>
</tr>
<tr>
<td></td>
<td>Gastrointestinal</td>
<td>30 (26-35)</td>
<td>27 (25-30)</td>
</tr>
<tr>
<td>----------------------</td>
<td>------------------</td>
<td>------------</td>
<td>------------</td>
</tr>
<tr>
<td>Nausea</td>
<td></td>
<td>13 (10-17)</td>
<td>12 (10-14)</td>
</tr>
<tr>
<td>Vomiting</td>
<td></td>
<td>3.3 (1.9-5.4)</td>
<td>2.6 (1.7-3.9)</td>
</tr>
<tr>
<td>Abdominal pain</td>
<td></td>
<td>11 (9-14)</td>
<td>9.3 (7.9-11)</td>
</tr>
<tr>
<td>Diarrhea</td>
<td></td>
<td>14 (11-18)</td>
<td>13 (11-15)</td>
</tr>
</tbody>
</table>
Total number tested for SARS-CoV-2 (n = 45,105)
- targeted testing (n = 16,750)
- population screening (n = 30,686)

SARS-CoV-2-positive patients (n = 1797)
- targeted testing (n = 1619)
- population screening (n = 178)

Enrolled in telehealth monitoring (n = 1765)
- targeted testing (n = 1587)
- population screening (n = 178)

Included in analysis of symptom progression (n = 1564)

Hospitalized at LUH at the time of diagnosis (n = 31)
Diagnosed post mortem (n = 1)

Enrolled in telehealth before implementation of standardized data entry form (n = 200)
RT-PCR positive several weeks after exposure and symptom resolution (n = 1)

Targeted testing (n = 1409)
Population screening (n = 155)
List of Figures

S1 The age and sex distribution of SARS-CoV-2-positive and negative patients. The standardized clinical data entry form was implemented on March 17, 2020. Distributions are shown for patients tested before and after March 17. Age distribution was largely comparable between the four groups, though SARS-CoV-2-positive patients diagnosed before March 17, were slightly older and the distribution more concentrated between 45 and 65 years of age. Among SARS-CoV-2-negative individuals, the proportion of females was consistently higher. Sex distribution was equal among SARS-CoV-2-positive patients.

S2 The total number of documented symptoms by date of diagnosis. The y-axis ranges from 0 to 19 symptoms, and the x-axis from February 27 to May 1, 2020. The median number of documented symptoms by date is depicted with a red point and the date at which the standardized data entry form was implemented (March 17) is illustrated with a dotted vertical line. The relationship between date of diagnosis and the total number of documented symptoms is further illustrated using a generalized additive model with integrated smoothness estimation, which is depicted as a blue line.

S3 The cumulative incidence and daily proportion of SARS-CoV-2-positive patients who fulfilled the WHO and CDC case definition criteria for suspected COVID-19 by age group and days from symptom onset. The x-axis ranges from symptom onset to day 21, and the y-axis shows the proportion or cumulative incidence as a percentage. The range of the y-axis varies between panels. The logistic regression estimate of the daily proportion of patients meeting criteria (panels A and B) are illustrated as lines with 95% confidence intervals depicted as shaded areas. The cumulative incidence of patients meeting criteria (panels C and D) are illustrated both as points (representing the Turnbull estimate) and as lines, estimated using parametric cure-mixture models using a log-logistic distribution.

S4 The proportion of SARS-CoV-2-positive patients experiencing generalized symptoms by day from symptom onset, estimated using different methods. The x-axis ranges from symptom onset until day 21, and the y-axis shows the proportion as a percentage. The range of the y-axis varies between panels. Four methods are shown. The daily proportion based on cases with no missing data (complete cases) is depicted with a closed circle, and a similar daily proportion based on a naive up-down filling procedure is illustrated as a closed triangle. Estimates based on logistic regression using only cases with no missing data (blue, broken line) and logistic regression with multiple imputation (red, solid line) are shown with 95% confidence intervals illustrated as a shaded area.

S5 The proportion of SARS-CoV-2-positive patients experiencing upper respiratory symptoms by day from symptom onset, estimated using different methods. The x-axis ranges from symptom onset until day 21, and the y-axis shows the proportion as a percentage. The range of the y-axis varies between panels. Four methods are shown. The daily proportion based on cases with no missing data (complete cases) is depicted with a closed circle, and a similar daily proportion based on a naive up-down filling procedure is illustrated as a closed triangle. Estimates based on logistic regression using only cases with no missing data (blue, broken line) and logistic regression with multiple imputation (red, solid line) are shown with 95% confidence intervals illustrated as a shaded area.

S6 The proportion of SARS-CoV-2-positive patients experiencing lower respiratory symptoms by day from symptom onset, estimated using different methods. The x-axis ranges from symptom onset until day 21, and the y-axis shows the proportion as a percentage. The range of the y-axis varies between panels. Four methods are shown. The daily proportion based on cases with no missing data (complete cases) is depicted with a closed circle, and a similar daily proportion based on a naive up-down filling procedure is illustrated as a closed triangle. Estimates based on logistic regression using only cases with no missing data (blue, broken line) and logistic regression with multiple imputation (red, solid line) are shown with 95% confidence intervals illustrated as a shaded area.

S7 The proportion of SARS-CoV-2-positive patients experiencing gastrointestinal symptoms by day from symptom onset, estimated using different methods. The x-axis ranges from symptom onset until day 21, and the y-axis shows the proportion as a percentage. The range of the y-axis varies between panels. Four methods are shown. The daily proportion based on cases with no missing data (complete cases) is depicted with a closed circle, and a similar daily proportion based on a naive up-down filling procedure is illustrated as a closed triangle. Estimates based on logistic regression using only cases with no missing data (blue, broken line) and logistic regression with multiple imputation (red, solid line) are shown with 95% confidence intervals illustrated as a shaded area.
S8 The proportion of SARS-CoV-2-positive patients experiencing generalized symptoms by diagnostic testing protocol. The x-axis ranges from symptom onset until day 21, and the y-axis shows the proportion as a percentage. The range of the y-axis varies between panels. Patients diagnosed through population screening are illustrated as a red line, and patients diagnosed by targeted testing are depicted with a blue line. The 95% confidence intervals are illustrated as a shaded area. 

S9 The proportion of SARS-CoV-2-positive patients experiencing upper respiratory symptoms by diagnostic testing protocol. The x-axis ranges from symptom onset until day 21, and the y-axis shows the proportion as a percentage. The range of the y-axis varies between panels. Patients diagnosed through population screening are illustrated as a red line, and patients diagnosed by targeted testing are depicted with a blue line. The 95% confidence intervals are illustrated as a shaded area. 

S10 The proportion of SARS-CoV-2-positive patients experiencing lower respiratory symptoms by diagnostic testing protocol. The x-axis ranges from symptom onset until day 21, and the y-axis shows the proportion as a percentage. The range of the y-axis varies between panels. Patients diagnosed through population screening are illustrated as a red line, and patients diagnosed by targeted testing are depicted with a blue line. The 95% confidence intervals are illustrated as a shaded area. 

S11 The proportion of SARS-CoV-2-positive patients experiencing gastrointestinal symptoms by diagnostic testing protocol. The x-axis ranges from symptom onset day 21, and the y-axis shows the proportion as a percentage. The range of the y-axis varies between panels. The logistic regression estimate of the proportion is illustrated with a colored line and the 95% confidence interval shown as a shaded area. 

S12 Proportion of SARS-CoV-2-positive patients experiencing generalized symptoms by age group, sex, and day from symptom onset. The x-axis ranges from symptom onset to day 21, and the y-axis shows the proportion as a percentage. The range of the y-axis varies between panels. The logistic regression estimate of the proportion is illustrated with a colored line and the 95% confidence interval shown as a shaded area. 

S13 Proportion of SARS-CoV-2-positive patients experiencing upper respiratory symptoms by age group, sex, and day from symptom onset. The x-axis ranges from symptom onset to day 21, and the y-axis shows the proportion as a percentage. The range of the y-axis varies between panels. The logistic regression estimate of the proportion is illustrated with a colored line and the 95% confidence interval shown as a shaded area. 

S14 Proportion of SARS-CoV-2-positive patients experiencing lower respiratory symptoms by age group, sex, and day from symptom onset. The x-axis ranges from symptom onset to day 21, and the y-axis shows the proportion as a percentage. The range of the y-axis varies between panels. The logistic regression estimate of the proportion is illustrated with a colored line and the 95% confidence interval shown as a shaded area. 

S15 Proportion of SARS-CoV-2-positive patients experiencing gastrointestinal symptoms by age group, sex, and day from symptom onset. The x-axis ranges from symptom onset to day 21, and the y-axis shows the proportion as a percentage. The range of the y-axis varies between panels. The logistic regression estimate of the proportion is illustrated with a colored line and the 95% confidence interval shown as a shaded area. 

S16 Proportion of SARS-CoV-2-positive patients experiencing generalized symptoms by hospitalization status. The x-axis ranges from symptom onset to day 21, and the y-axis shows the proportion as a percentage. The range of the y-axis varies between panels. The logistic regression estimate of the proportion is illustrated with a red or blue line the 95% confidence interval shown as a shaded area. 

S17 Proportion of SARS-CoV-2-positive patients experiencing upper respiratory symptoms by hospitalization status. The x-axis ranges from symptom onset to day 21, and the y-axis shows the proportion as a percentage. The range of the y-axis varies between panels. The logistic regression estimate of the proportion is illustrated with a red or blue line the 95% confidence interval shown as a shaded area. 

S18 Proportion of SARS-CoV-2-positive patients experiencing lower respiratory symptoms by hospitalization status. The x-axis ranges from symptom onset to day 21, and the y-axis shows the proportion as a percentage. The range of the y-axis varies between panels. The logistic regression estimate of the proportion is illustrated with a red or blue line the 95% confidence interval shown as a shaded area.
S19 Proportion of SARS-CoV-2-positive patients experiencing gastrointestinal symptoms by hospitalization status. The x-axis ranges from symptom onset to day 21, and the y-axis shows the proportion as a percentage. The range of the y-axis varies between panels. The logistic regression estimate of the proportion is illustrated with a red or blue line the 95% confidence interval shown as a shaded area.

S20 Cumulative incidence of generalized symptoms experienced by SARS-CoV-2-positive patients by day from symptom onset. The x-axis ranges from symptom onset to day 21, and the y-axis shows the cumulative incidence as a percentage. The range of the y-axis varies between panels. The non-parametric Turnbull estimate of the cumulative incidence is depicted with black points and the parametric cure-mixture estimate illustrated with a black line and the 95% confidence interval shown as a shaded grey area.

S21 Cumulative incidence of generalized symptoms experienced by SARS-CoV-2-positive patients by age group and day from symptom onset. The x-axis ranges from symptom onset to day 21, and the y-axis shows the cumulative incidence as a percentage. The range of the y-axis varies between panels. The non-parametric Turnbull estimate of the cumulative incidence is depicted with points and the parametric cure-mixture estimate illustrated with a line and the 95% confidence interval shown as a shaded area.

S22 Cumulative incidence of generalized symptoms experienced by SARS-CoV-2-positive patients by sex and day from symptom onset. The x-axis ranges from symptom onset to day 21, and the y-axis shows the cumulative incidence as a percentage. The range of the y-axis varies between panels. The non-parametric Turnbull estimate of the cumulative incidence is depicted with black points and the parametric cure-mixture estimate illustrated with a black line and the 95% confidence interval shown as a shaded grey area.

S23 Cumulative incidence of upper respiratory symptoms experienced by SARS-CoV-2-positive patients by day from symptom onset. The x-axis ranges from symptom onset to day 21, and the y-axis shows the cumulative incidence as a percentage. The range of the y-axis varies between panels. The non-parametric Turnbull estimate of the cumulative incidence is depicted with points and the parametric cure-mixture estimate illustrated with a line and the 95% confidence interval shown as a shaded area.

S24 Cumulative incidence of upper respiratory symptoms experienced by SARS-CoV-2-positive patients by age group and day from symptom onset. The x-axis ranges from symptom onset to day 21, and the y-axis shows the cumulative incidence as a percentage. The range of the y-axis varies between panels. The non-parametric Turnbull estimate of the cumulative incidence is depicted with points and the parametric cure-mixture estimate illustrated with a line and the 95% confidence interval shown as a shaded area.

S25 Cumulative incidence of upper respiratory symptoms experienced by SARS-CoV-2-positive patients by sex and day from symptom onset. The x-axis ranges from symptom onset to day 21, and the y-axis shows the cumulative incidence as a percentage. The range of the y-axis varies between panels. The non-parametric Turnbull estimate of the cumulative incidence is depicted with points and the parametric cure-mixture estimate illustrated with a line and the 95% confidence interval shown as a shaded area.

S26 Cumulative incidence of lower respiratory symptoms experienced by SARS-CoV-2-positive patients by day from symptom onset. The x-axis ranges from symptom onset to day 21, and the y-axis shows the cumulative incidence as a percentage. The range of the y-axis varies between panels. The non-parametric Turnbull estimate of the cumulative incidence is depicted with black points and the parametric cure-mixture estimate illustrated with a black line and the 95% confidence interval shown as a shaded grey area.

S27 Cumulative incidence of lower respiratory symptoms experienced by SARS-CoV-2-positive patients by age group and day from symptom onset. The x-axis ranges from symptom onset to day 21, and the y-axis shows the cumulative incidence as a percentage. The range of the y-axis varies between panels. The non-parametric Turnbull estimate of the cumulative incidence is depicted with points and the parametric cure-mixture estimate illustrated with a line and the 95% confidence interval shown as a shaded area.

S28 Cumulative incidence of lower respiratory symptoms experienced by SARS-CoV-2-positive patients by sex and day from symptom onset. The x-axis ranges from symptom onset to day 21, and the y-axis shows the cumulative incidence as a percentage. The range of the y-axis varies between panels. The non-parametric Turnbull estimate of the cumulative incidence is depicted with points and the parametric cure-mixture estimate illustrated with a line and the 95% confidence interval shown as a shaded area.
S29 Cumulative incidence of gastrointestinal symptoms experienced by SARS-CoV-2-positive patients by day from symptom onset. The x-axis ranges from symptom onset to day 21, and the y-axis shows the cumulative incidence as a percentage. The range of the y-axis varies between panels. The non-parametric Turnbull estimate of the cumulative incidence is depicted with black points and the parametric cure-mixture estimate illustrated with a line and the 95% confidence interval shown as a shaded area.

S30 Cumulative incidence of gastrointestinal symptoms experienced by SARS-CoV-2-positive patients by age group and day from symptom onset. The x-axis ranges from symptom onset to day 21, and the y-axis shows the cumulative incidence as a percentage. The range of the y-axis varies between panels. The non-parametric Turnbull estimate of the cumulative incidence is depicted with points and the parametric cure-mixture estimate illustrated with a line and the 95% confidence interval shown as a shaded area.

S31 Cumulative incidence of gastrointestinal symptoms experienced by SARS-CoV-2-positive patients by sex and day from symptom onset. The x-axis ranges from symptom onset to day 21, and the y-axis shows the cumulative incidence as a percentage. The range of the y-axis varies between panels. The non-parametric Turnbull estimate of the cumulative incidence is depicted with points and the parametric cure-mixture estimate illustrated with a line and the 95% confidence interval shown as a shaded area.

S32 Cumulative incidence of generalized symptoms experienced by SARS-CoV-2-positive patients by testing protocol. The x-axis ranges from symptom onset to day 21, and the y-axis shows the cumulative incidence as a percentage. The range of the y-axis varies between panels. The non-parametric Turnbull estimate of the cumulative incidence is depicted with points and the parametric cure-mixture estimate illustrated with a line and the 95% confidence interval shown as a shaded area.

S33 Cumulative incidence of upper respiratory symptoms experienced by SARS-CoV-2-positive patients by testing protocol. The x-axis ranges from symptom onset to day 21, and the y-axis shows the cumulative incidence as a percentage. The range of the y-axis varies between panels. The non-parametric Turnbull estimate of the cumulative incidence is depicted with points and the parametric cure-mixture estimate illustrated with a line and the 95% confidence interval shown as a shaded area.

S34 Cumulative incidence of lower respiratory symptoms experienced by SARS-CoV-2-positive patients by testing protocol. The x-axis ranges from symptom onset to day 21, and the y-axis shows the cumulative incidence as a percentage. The range of the y-axis varies between panels. The non-parametric Turnbull estimate of the cumulative incidence is depicted with points and the parametric cure-mixture estimate illustrated with a line and the 95% confidence interval shown as a shaded area.

S35 Cumulative incidence of gastrointestinal symptoms experienced by SARS-CoV-2-positive patients by testing protocol. The x-axis ranges from symptom onset to day 21, and the y-axis shows the cumulative incidence as a percentage. The range of the y-axis varies between panels. The non-parametric Turnbull estimate of the cumulative incidence is depicted with points and the parametric cure-mixture estimate illustrated with a line and the 95% confidence interval shown as a shaded area.

S36 Cumulative incidence of generalized symptoms experienced by SARS-CoV-2-positive patients by hospitalization status. The x-axis ranges from symptom onset to day 21, and the y-axis shows the cumulative incidence as a percentage. The range of the y-axis varies between panels. The non-parametric Turnbull estimate of the cumulative incidence is depicted with points and the parametric cure-mixture estimate illustrated with a line and the 95% confidence interval shown as a shaded area.

S37 Cumulative incidence of upper respiratory symptoms experienced by SARS-CoV-2-positive patients by hospitalization status. The x-axis ranges from symptom onset to day 21, and the y-axis shows the cumulative incidence as a percentage. The range of the y-axis varies between panels. The non-parametric Turnbull estimate of the cumulative incidence is depicted with points and the parametric cure-mixture estimate illustrated with a line and the 95% confidence interval shown as a shaded area.

S38 Cumulative incidence of lower respiratory symptoms experienced by SARS-CoV-2-positive patients by hospitalization status. The x-axis ranges from symptom onset to day 21, and the y-axis shows the cumulative incidence as a percentage. The range of the y-axis varies between panels. The non-parametric Turnbull estimate of the cumulative incidence is depicted with points and the parametric cure-mixture estimate illustrated with a line and the 95% confidence interval shown as a shaded area.
S39 Cumulative incidence of gastrointestinal symptoms experienced by SARS-CoV-2-positive patients by hospitalization status. The x-axis ranges from symptom onset to day 21, and the y-axis shows the cumulative incidence as a percentage. The range of the y-axis varies between panels. The non-parametric Turnbull estimate of the cumulative incidence is depicted with points and the parametric cure-mixture estimate illustrated with a line and the 95% confidence interval shown as a shaded area.
Figure S1: The age and sex distribution of SARS-CoV-2-positive and negative patients. The standardized clinical data entry form was implemented on March 17, 2020. Distributions are shown for patients tested before and after March 17. Age distribution was largely comparable between the four groups, though SARS-CoV-2-positive patients diagnosed before March 17, were slightly older and the distribution more concentrated between 45 and 65 years of age. Among SARS-CoV-2-negative individuals, the proportion of females was consistently higher. Sex distribution was equal among SARS-CoV-2-positive patients.
Table S1: Demographic information for included and excluded subgroups of the study population. Median age is shown with the interquartile range in parenthesis. The number of individuals in each subgroup who underwent targeted testing is shown, and the proportion of all tests given as a percentage within parenthesis.

<table>
<thead>
<tr>
<th>Condition</th>
<th>n</th>
<th>Age (years)</th>
<th>Sex (female)</th>
<th>Targeted testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>All individuals tested for SARS-CoV-2</td>
<td>46253</td>
<td>42 (28-56)</td>
<td>25933 (56)</td>
<td>16750 (36)</td>
</tr>
<tr>
<td>SARS-CoV-2-negative</td>
<td>44456</td>
<td>42 (28-56)</td>
<td>25029 (56)</td>
<td>15120 (34)</td>
</tr>
<tr>
<td>SARS-CoV-2-positive, diagnosed before March 17</td>
<td>200</td>
<td>47 (34-57)</td>
<td>93 (47)</td>
<td>178 (89)</td>
</tr>
<tr>
<td>SARS-CoV-2-positive, hospitalized at diagnosis</td>
<td>31</td>
<td>69 (75-82)</td>
<td>19 (61)</td>
<td>31 (100)</td>
</tr>
<tr>
<td>SARS-CoV-2-positive, diagnosed after March 17</td>
<td>1564</td>
<td>40 (26-53)</td>
<td>791 (51)</td>
<td>1409 (90)</td>
</tr>
<tr>
<td>- Asymptomatic at diagnosis</td>
<td>83</td>
<td>36 (20.5-57.0)</td>
<td>34 (41)</td>
<td>45 (54)</td>
</tr>
<tr>
<td>- Did not meet CDC criteria at diagnosis</td>
<td>216</td>
<td>31 (19-49.25)</td>
<td>102 (47)</td>
<td>153 (71)</td>
</tr>
<tr>
<td>- Did not meet WHO criteria at diagnosis</td>
<td>349</td>
<td>36 (23-49)</td>
<td>189 (54)</td>
<td>278 (80)</td>
</tr>
<tr>
<td>- Ever hospitalized</td>
<td>55</td>
<td>61 (54-70)</td>
<td>17 (31)</td>
<td>55 (100)</td>
</tr>
</tbody>
</table>

Figure S2: The total number of documented symptoms by date of diagnosis. The y-axis ranges from 0 to 19 symptoms, and the x-axis from February 27 to May 1, 2020. The median number of documented symptoms by date is depicted with a red point and the date at which the standardized data entry form was implemented (March 17) is illustrated with a dotted vertical line. The relationship between date of diagnosis and the total number of documented symptoms is further illustrated using a generalized additive model with integrated smoothness estimation, which is depicted as a blue line.
Table S2: The proportion of missing information per variable by day from symptom onset. No data were missing for any variable for any patient, during days on which interviews occurred. Information regarding the presence or absence of symptoms on days during which interviews did not occur, can be considered to be missing.

<table>
<thead>
<tr>
<th>Days from symptom onset</th>
<th>Age</th>
<th>Sex</th>
<th>Fever</th>
<th>Cough</th>
<th>Headache</th>
<th>Coughing</th>
<th>Sore throat</th>
<th>Loss of appetite</th>
<th>Rhinorrhea</th>
<th>Sore eye</th>
<th>Myalgia</th>
<th>Loss of taste</th>
<th>Dysosmia</th>
<th>Dry mouth</th>
<th>Dysphagia</th>
<th>Vomiting</th>
<th>Abdominal pain</th>
<th>Diarrhea</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0.96</td>
<td>0.96</td>
<td>0.96</td>
<td>0.96</td>
<td>0.96</td>
<td>0.96</td>
<td>0.96</td>
<td>0.96</td>
<td>0.96</td>
<td>0.96</td>
<td>0.96</td>
<td>0.96</td>
<td>0.96</td>
<td>0.96</td>
<td>0.96</td>
<td>0.96</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0.95</td>
<td>0.95</td>
<td>0.95</td>
<td>0.95</td>
<td>0.95</td>
<td>0.95</td>
<td>0.95</td>
<td>0.95</td>
<td>0.95</td>
<td>0.95</td>
<td>0.95</td>
<td>0.95</td>
<td>0.95</td>
<td>0.95</td>
<td>0.95</td>
<td>0.95</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0.89</td>
<td>0.89</td>
<td>0.89</td>
<td>0.89</td>
<td>0.89</td>
<td>0.89</td>
<td>0.89</td>
<td>0.89</td>
<td>0.89</td>
<td>0.89</td>
<td>0.89</td>
<td>0.89</td>
<td>0.89</td>
<td>0.89</td>
<td>0.89</td>
<td>0.89</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0.82</td>
<td>0.82</td>
<td>0.82</td>
<td>0.82</td>
<td>0.82</td>
<td>0.82</td>
<td>0.82</td>
<td>0.82</td>
<td>0.82</td>
<td>0.82</td>
<td>0.82</td>
<td>0.82</td>
<td>0.82</td>
<td>0.82</td>
<td>0.82</td>
<td>0.82</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0.73</td>
<td>0.73</td>
<td>0.73</td>
<td>0.73</td>
<td>0.73</td>
<td>0.73</td>
<td>0.73</td>
<td>0.73</td>
<td>0.73</td>
<td>0.73</td>
<td>0.73</td>
<td>0.73</td>
<td>0.73</td>
<td>0.73</td>
<td>0.73</td>
<td>0.73</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0.71</td>
<td>0.71</td>
<td>0.71</td>
<td>0.71</td>
<td>0.71</td>
<td>0.71</td>
<td>0.71</td>
<td>0.71</td>
<td>0.71</td>
<td>0.71</td>
<td>0.71</td>
<td>0.71</td>
<td>0.71</td>
<td>0.71</td>
<td>0.71</td>
<td>0.71</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0.66</td>
<td>0.66</td>
<td>0.66</td>
<td>0.66</td>
<td>0.66</td>
<td>0.66</td>
<td>0.66</td>
<td>0.66</td>
<td>0.66</td>
<td>0.66</td>
<td>0.66</td>
<td>0.66</td>
<td>0.66</td>
<td>0.66</td>
<td>0.66</td>
<td>0.66</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0.62</td>
<td>0.62</td>
<td>0.62</td>
<td>0.62</td>
<td>0.62</td>
<td>0.62</td>
<td>0.62</td>
<td>0.62</td>
<td>0.62</td>
<td>0.62</td>
<td>0.62</td>
<td>0.62</td>
<td>0.62</td>
<td>0.62</td>
<td>0.62</td>
<td>0.62</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>0</td>
<td>0.61</td>
<td>0.61</td>
<td>0.61</td>
<td>0.61</td>
<td>0.61</td>
<td>0.61</td>
<td>0.61</td>
<td>0.61</td>
<td>0.61</td>
<td>0.61</td>
<td>0.61</td>
<td>0.61</td>
<td>0.61</td>
<td>0.61</td>
<td>0.61</td>
<td>0.61</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0.62</td>
<td>0.62</td>
<td>0.62</td>
<td>0.62</td>
<td>0.62</td>
<td>0.62</td>
<td>0.62</td>
<td>0.62</td>
<td>0.62</td>
<td>0.62</td>
<td>0.62</td>
<td>0.62</td>
<td>0.62</td>
<td>0.62</td>
<td>0.62</td>
<td>0.62</td>
</tr>
<tr>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0.60</td>
<td>0.60</td>
<td>0.60</td>
<td>0.60</td>
<td>0.60</td>
<td>0.60</td>
<td>0.60</td>
<td>0.60</td>
<td>0.60</td>
<td>0.60</td>
<td>0.60</td>
<td>0.60</td>
<td>0.60</td>
<td>0.60</td>
<td>0.60</td>
<td>0.60</td>
</tr>
<tr>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
</tr>
<tr>
<td>13</td>
<td>0</td>
<td>0</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
</tr>
<tr>
<td>14</td>
<td>0</td>
<td>0</td>
<td>0.64</td>
<td>0.64</td>
<td>0.64</td>
<td>0.64</td>
<td>0.64</td>
<td>0.64</td>
<td>0.64</td>
<td>0.64</td>
<td>0.64</td>
<td>0.64</td>
<td>0.64</td>
<td>0.64</td>
<td>0.64</td>
<td>0.64</td>
<td>0.64</td>
<td>0.64</td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>0</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
</tr>
<tr>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0.64</td>
<td>0.64</td>
<td>0.64</td>
<td>0.64</td>
<td>0.64</td>
<td>0.64</td>
<td>0.64</td>
<td>0.64</td>
<td>0.64</td>
<td>0.64</td>
<td>0.64</td>
<td>0.64</td>
<td>0.64</td>
<td>0.64</td>
<td>0.64</td>
<td>0.64</td>
</tr>
<tr>
<td>17</td>
<td>0</td>
<td>0</td>
<td>0.66</td>
<td>0.66</td>
<td>0.66</td>
<td>0.66</td>
<td>0.66</td>
<td>0.66</td>
<td>0.66</td>
<td>0.66</td>
<td>0.66</td>
<td>0.66</td>
<td>0.66</td>
<td>0.66</td>
<td>0.66</td>
<td>0.66</td>
<td>0.66</td>
<td>0.66</td>
</tr>
<tr>
<td>18</td>
<td>0</td>
<td>0</td>
<td>0.68</td>
<td>0.68</td>
<td>0.68</td>
<td>0.68</td>
<td>0.68</td>
<td>0.68</td>
<td>0.68</td>
<td>0.68</td>
<td>0.68</td>
<td>0.68</td>
<td>0.68</td>
<td>0.68</td>
<td>0.68</td>
<td>0.68</td>
<td>0.68</td>
<td>0.68</td>
</tr>
<tr>
<td>19</td>
<td>0</td>
<td>0</td>
<td>0.67</td>
<td>0.67</td>
<td>0.67</td>
<td>0.67</td>
<td>0.67</td>
<td>0.67</td>
<td>0.67</td>
<td>0.67</td>
<td>0.67</td>
<td>0.67</td>
<td>0.67</td>
<td>0.67</td>
<td>0.67</td>
<td>0.67</td>
<td>0.67</td>
<td>0.67</td>
</tr>
<tr>
<td>20</td>
<td>0</td>
<td>0</td>
<td>0.67</td>
<td>0.67</td>
<td>0.67</td>
<td>0.67</td>
<td>0.67</td>
<td>0.67</td>
<td>0.67</td>
<td>0.67</td>
<td>0.67</td>
<td>0.67</td>
<td>0.67</td>
<td>0.67</td>
<td>0.67</td>
<td>0.67</td>
<td>0.67</td>
<td>0.67</td>
</tr>
<tr>
<td>21</td>
<td>0</td>
<td>0</td>
<td>0.69</td>
<td>0.69</td>
<td>0.69</td>
<td>0.69</td>
<td>0.69</td>
<td>0.69</td>
<td>0.69</td>
<td>0.69</td>
<td>0.69</td>
<td>0.69</td>
<td>0.69</td>
<td>0.69</td>
<td>0.69</td>
<td>0.69</td>
<td>0.69</td>
<td>0.69</td>
</tr>
</tbody>
</table>

https://mc.manuscriptcentral.com/bmj
Figure S3: The cumulative incidence and daily proportion of SARS-CoV-2-positive patients who fulfilled the WHO and CDC case definition criteria for suspected COVID-19 by age group and days from symptom onset. The x-axis ranges from symptom onset to day 21, and the y-axis shows the proportion or cumulative incidence as a percentage. The range of the y-axis varies between panels. The logistic regression estimate of the daily proportion of patients meeting criteria (panels A and B) are illustrated as lines with 95% confidence intervals depicted as shaded areas. The cumulative incidence of patients meeting criteria (panels C and D) are illustrated both as points (representing the Turnbull estimate) and as lines, estimated using parametric cure-mixture models using a log-logistic distribution.
Figure S4: The proportion of SARS-CoV-2-positive patients experiencing generalized symptoms by day from symptom onset, estimated using different methods. The x-axis ranges from symptom onset until day 21, and the y-axis shows the proportion as a percentage. The range of the y-axis varies between panels. Four methods are shown. The daily proportion based on cases with no missing data (complete cases) is depicted with a closed circle, and a similar daily proportion based on a naive up-down filling procedure is illustrated as a closed triangle. Estimates based on logistic regression using only cases with no missing data (blue, broken line) and logistic regression with multiple imputation (red, solid line) are shown with 95% confidence intervals illustrated as a shaded area.
Figure S5: The proportion of SARS-CoV-2-positive patients experiencing upper respiratory symptoms by day from symptom onset, estimated using different methods. The x-axis ranges from symptom onset until day 21, and the y-axis shows the proportion as a percentage. The range of the y-axis varies between panels. Four methods are shown. The daily proportion based on cases with no missing data (complete cases) is depicted with a closed circle, and a similar daily proportion based on a naive up-down filling procedure is illustrated as a closed triangle. Estimates based on logistic regression using only cases with no missing data (blue, broken line) and logistic regression with multiple imputation (red, solid line) are shown with 95% confidence intervals illustrated as a shaded area.
Figure S6: The proportion of SARS-CoV-2-positive patients experiencing lower respiratory symptoms by day from symptom onset, estimated using different methods. The x-axis ranges from symptom onset until day 21, and the y-axis shows the proportion as a percentage. The range of the y-axis varies between panels. Four methods are shown. The daily proportion based on cases with no missing data (complete cases) is depicted with a closed circle, and a similar daily proportion based on a naive up-down filling procedure is illustrated as a closed triangle. Estimates based on logistic regression using only cases with no missing data (blue, broken line) and logistic regression with multiple imputation (red, solid line) are shown with 95% confidence intervals illustrated as a shaded area.
Figure S7: The proportion of SARS-CoV-2-positive patients experiencing gastrointestinal symptoms by day from symptom onset, estimated using different methods. The x-axis ranges from symptom onset until day 21, and the y-axis shows the proportion as a percentage. The range of the y-axis varies between panels. Four methods are shown. The daily proportion based on cases with no missing data (complete cases) is depicted with a closed circle, and a similar daily proportion based on a naive up-down filling procedure is illustrated as a closed triangle. Estimates based on logistic regression using only cases with no missing data (blue, broken line) and logistic regression with multiple imputation (red, solid line) are shown with 95% confidence intervals illustrated as a shaded area.
Figure S8: The proportion of SARS-CoV-2-positive patients experiencing generalized symptoms by diagnostic testing protocol. The x-axis ranges from symptom onset until day 21, and the y-axis shows the proportion as a percentage. The range of the y-axis varies between panels. Patients diagnosed through population screening are illustrated as a red line, and patients diagnosed by targeted testing are depicted with a blue line. The 95% confidence intervals are illustrated as a shaded area.
Figure S9: The proportion of SARS-CoV-2-positive patients experiencing upper respiratory symptoms by diagnostic testing protocol. The x-axis ranges from symptom onset until day 21, and the y-axis shows the proportion as a percentage. The range of the y-axis varies between panels. Patients diagnosed through population screening are illustrated as a red line, and patients diagnosed by targeted testing are depicted with a blue line. The 95% confidence intervals are illustrated as a shaded area.
Figure S10: The proportion of SARS-CoV-2-positive patients experiencing lower respiratory symptoms by diagnostic testing protocol. The x-axis ranges from symptom onset until day 21, and the y-axis shows the proportion as a percentage. The range of the y-axis varies between panels. Patients diagnosed through population screening are illustrated as a red line, and patients diagnosed by targeted testing are depicted with a blue line. The 95% confidence intervals are illustrated as a shaded area.
Figure S11: The proportion of SARS-CoV-2-positive patients experiencing gastrointestinal symptoms by diagnostic testing protocol. The x-axis ranges from symptom onset until day 21, and the y-axis shows the proportion as a percentage. The range of the y-axis varies between panels. Patients diagnosed through population screening are illustrated as a red line, and patients diagnosed by targeted testing are depicted with a blue line. The 95% confidence intervals are illustrated as a shaded area.
Figure S12: Proportion of SARS-CoV-2-positive patients experiencing generalized symptoms by age group, sex and day from symptom onset. The x-axis ranges from symptom onset to day 21, and the y-axis shows the proportion as a percentage. The range of the y-axis varies between panels. The logistic regression estimate of the proportion is illustrated with a colored line and the 95% confidence interval shown as a shaded area.
Figure S13: Proportion of SARS-CoV-2-positive patients experiencing upper respiratory symptoms by age group, sex and day from symptom onset. The x-axis ranges from symptom onset to day 21, and the y-axis shows the proportion as a percentage. The range of the y-axis varies between panels. The logistic regression estimate of the proportion is illustrated with a colored line and the 95% confidence interval shown as a shaded area.
Figure S14: Proportion of SARS-CoV-2-positive patients experiencing lower respiratory symptoms by age group, sex and day from symptom onset. The x-axis ranges from symptom onset to day 21, and the y-axis shows the proportion as a percentage. The range of the y-axis varies between panels. The logistic regression estimate of the proportion is illustrated with a colored line and the 95% confidence interval shown as a shaded area.
Figure S15: Proportion of SARS-CoV-2-positive patients experiencing gastrointestinal symptoms by age group, sex and day from symptom onset. The x-axis ranges from symptom onset to day 21, and the y-axis shows the proportion as a percentage. The range of the y-axis varies between panels. The logistic regression estimate of the proportion is illustrated with a colored line and the 95% confidence interval shown as a shaded area.
Figure S16: Proportion of SARS-CoV-2-positive patients experiencing generalized symptoms by hospitalization status. The x-axis ranges from symptom onset to day 21, and the y-axis shows the proportion as a percentage. The range of the y-axis varies between panels. The logistic regression estimate of the proportion is illustrated with a red or blue line and the 95% confidence interval shown as a shaded area.
Figure S17: Proportion of SARS-CoV-2-positive patients experiencing upper respiratory symptoms by hospitalization status. The x-axis ranges from symptom onset to day 21, and the y-axis shows the proportion as a percentage. The range of the y-axis varies between panels. The logistic regression estimate of the proportion is illustrated with a red or blue line the 95% confidence interval shown as a shaded area.
Figure S18: Proportion of SARS-CoV-2-positive patients experiencing lower respiratory symptoms by hospitalization status. The x-axis ranges from symptom onset to day 21, and the y-axis shows the proportion as a percentage. The range of the y-axis varies between panels. The logistic regression estimate of the proportion is illustrated with a red or blue line the 95% confidence interval shown as a shaded area.
A. One or more of any gastrointestinal symptom

B. Nausea

C. Vomiting

D. Abdominal pain

E. Diarrhea

Admitted for COVID−19
Never admitted for COVID−19

Figure S19: Proportion of SARS-CoV-2-positive patients experiencing gastrointestinal symptoms by hospitalization status. The x-axis ranges from symptom onset to day 21, and the y-axis shows the proportion as a percentage. The range of the y-axis varies between panels. The logistic regression estimate of the proportion is illustrated with a red or blue line the 95% confidence interval shown as a shaded area.
Figure S20: Cumulative incidence of generalized symptoms experienced by SARS-CoV-2-positive patients by day from symptom onset. The x-axis ranges from symptom onset to day 21, and the y-axis shows the cumulative incidence as a percentage. The range of the y-axis varies between panels. The non-parametric Turnbull estimate of the cumulative incidence is depicted with black points and the parametric cure-mixture estimate illustrated with a black line and the 95% confidence interval shown as a shaded grey area.
Figure S21: Cumulative incidence of generalized symptoms experienced by SARS-CoV-2-positive patients by age group and day from symptom onset. The x-axis ranges from symptom onset to day 21, and the y-axis shows the cumulative incidence as a percentage. The range of the y-axis varies between panels. The non-parametric Turnbull estimate of the cumulative incidence is depicted with points and the parametric cure-mixture estimate illustrated with a line and the 95% confidence interval shown as a shaded area.
Figure S22: Cumulative incidence of generalized symptoms experienced by SARS-CoV-2-positive patients by sex and day from symptom onset. The x-axis ranges from symptom onset to day 21, and the y-axis shows the cumulative incidence as a percentage. The range of the y-axis varies between panels. The non-parametric Turnbull estimate of the cumulative incidence is depicted with points and the parametric cure-mixture estimate illustrated with a line and the 95% confidence interval shown as a shaded area.
Figure S23: Cumulative incidence of upper respiratory symptoms experienced by SARS-CoV-2-positive patients by day from symptom onset. The x-axis ranges from symptom onset to day 21, and the y-axis shows the cumulative incidence as a percentage. The range of the y-axis varies between panels. The non-parametric Turnbull estimate of the cumulative incidence is depicted with black points and the parametric cure-mixture estimate illustrated with a black line and the 95% confidence interval shown as a shaded grey area.
Figure S24: Cumulative incidence of upper respiratory symptoms experienced by SARS-CoV-2-positive patients by age group and day from symptom onset. The x-axis ranges from symptom onset to day 21, and the y-axis shows the cumulative incidence as a percentage. The range of the y-axis varies between panels. The non-parametric Turnbull estimate of the cumulative incidence is depicted with points and the parametric cure-mixture estimate illustrated with a line and the 95% confidence interval shown as a shaded area.
Figure S25: Cumulative incidence of upper respiratory symptoms experienced by SARS-CoV-2-positive patients by sex and day from symptom onset. The x-axis ranges from symptom onset to day 21, and the y-axis shows the cumulative incidence as a percentage. The range of the y-axis varies between panels. The non-parametric Turnbull estimate of the cumulative incidence is depicted with points and the parametric cure-mixture estimate illustrated with a line and the 95% confidence interval shown as a shaded area.
Figure S26: Cumulative incidence of lower respiratory symptoms experienced by SARS-CoV-2-positive patients by day from symptom onset. The x-axis ranges from symptom onset to day 21, and the y-axis shows the cumulative incidence as a percentage. The range of the y-axis varies between panels. The non-parametric Turnbull estimate of the cumulative incidence is depicted with black points and the parametric cure-mixture estimate illustrated with a black line and the 95% confidence interval shown as a shaded grey area.
Figure S27: Cumulative incidence of lower respiratory symptoms experienced by SARS-CoV-2-positive patients by age group and day from symptom onset. The x-axis ranges from symptom onset to day 21, and the y-axis shows the cumulative incidence as a percentage. The range of the y-axis varies between panels. The non-parametric Turnbull estimate of the cumulative incidence is depicted with points and the parametric cure-mixture estimate illustrated with a line and the 95% confidence interval shown as a shaded area.
Figure S28: Cumulative incidence of lower respiratory symptoms experienced by SARS-CoV-2-positive patients by sex and day from symptom onset. The x-axis ranges from symptom onset to day 21, and the y-axis shows the cumulative incidence as a percentage. The range of the y-axis varies between panels. The non-parametric Turnbull estimate of the cumulative incidence is depicted with points and the parametric cure-mixture estimate illustrated with a line and the 95% confidence interval shown as a shaded area.
Figure S29: Cumulative incidence of gastrointestinal symptoms experienced by SARS-CoV-2-positive patients by day from symptom onset. The x-axis ranges from symptom onset to day 21, and the y-axis shows the cumulative incidence as a percentage. The range of the y-axis varies between panels. The non-parametric Turnbull estimate of the cumulative incidence is depicted with black points and the parametric cure-mixture estimate illustrated with a black line and the 95% confidence interval shown as a shaded grey area.
Figure S30: Cumulative incidence of gastrointestinal symptoms experienced by SARS-CoV-2-positive patients by age group and day from symptom onset. The x-axis ranges from symptom onset to day 21, and the y-axis shows the cumulative incidence as a percentage. The range of the y-axis varies between panels. The non-parametric Turnbull estimate of the cumulative incidence is depicted with points and the parametric cure-mixture estimate illustrated with a line and the 95% confidence interval shown as a shaded area.
Figure S31: Cumulative incidence of gastrointestinal symptoms experienced by SARS-CoV-2-positive patients by sex and day from symptom onset. The x-axis ranges from symptom onset to day 21, and the y-axis shows the cumulative incidence as a percentage. The range of the y-axis varies between panels. The non-parametric Turnbull estimate of the cumulative incidence is depicted with points and the parametric cure-mixture estimate illustrated with a line and the 95% confidence interval shown as a shaded area.
Figure S32: Cumulative incidence of generalized symptoms experienced by SARS-CoV-2-positive patients by testing protocol. The x-axis ranges from symptom onset to day 21, and the y-axis shows the cumulative incidence as a percentage. The range of the y-axis varies between panels. The non-parametric Turnbull estimate of the cumulative incidence is depicted with points and the parametric cure-mixture estimate illustrated with a line and the 95% confidence interval shown as a shaded area.
Figure S33: Cumulative incidence of upper respiratory symptoms experienced by SARS-CoV-2-positive patients by testing protocol. The x-axis ranges from symptom onset to day 21, and the y-axis shows the cumulative incidence as a percentage. The range of the y-axis varies between panels. The non-parametric Turnbull estimate of the cumulative incidence is depicted with points and the parametric cure-mixture estimate illustrated with a line and the 95% confidence interval shown as a shaded area.
Figure S34: Cumulative incidence of lower respiratory symptoms experienced by SARS-CoV-2-positive patients by testing protocol. The x-axis ranges from symptom onset to day 21, and the y-axis shows the cumulative incidence as a percentage. The range of the y-axis varies between panels. The non-parametric Turnbull estimate of the cumulative incidence is depicted with points and the parametric cure-mixture estimate illustrated with a line and the 95% confidence interval shown as a shaded area.
Figure S35: Cumulative incidence of gastrointestinal symptoms experienced by SARS-CoV-2-positive patients by testing protocol. The x-axis ranges from symptom onset to day 21, and the y-axis shows the cumulative incidence as a percentage. The range of the y-axis varies between panels. The non-parametric Turnbull estimate of the cumulative incidence is depicted with points and the parametric cure-mixture estimate illustrated with a line and the 95% confidence interval shown as a shaded area.
Figure S36: Cumulative incidence of generalized symptoms experienced by SARS-CoV-2-positive patients by hospitalization status. The x-axis ranges from symptom onset to day 21, and the y-axis shows the cumulative incidence as a percentage. The range of the y-axis varies between panels. The non-parametric Turnbull estimate of the cumulative incidence is depicted with points and the parametric cure-mixture estimate illustrated with a line and the 95% confidence interval shown as a shaded area.
Figure S37: Cumulative incidence of upper respiratory symptoms experienced by SARS-CoV-2-positive patients by hospitalization status. The x-axis ranges from symptom onset to day 21, and the y-axis shows the cumulative incidence as a percentage. The range of the y-axis varies between panels. The non-parametric Turnbull estimate of the cumulative incidence is depicted with points and the parametric cure-mixture estimate illustrated with a line and the 95% confidence interval shown as a shaded area.
Figure S38: Cumulative incidence of lower respiratory symptoms experienced by SARS-CoV-2-positive patients by hospitalization status. The x-axis ranges from symptom onset to day 21, and the y-axis shows the cumulative incidence as a percentage. The range of the y-axis varies between panels. The non-parametric Turnbull estimate of the cumulative incidence is depicted with points and the parametric cure-mixture estimate illustrated with a line and the 95% confidence interval shown as a shaded area.
Figure S39: Cumulative incidence of gastrointestinal symptoms experienced by SARS-CoV-2-positive patients by hospitalization status. The x-axis ranges from symptom onset to day 21, and the y-axis shows the cumulative incidence as a percentage. The range of the y-axis varies between panels. The non-parametric Turnbull estimate of the cumulative incidence is depicted with points and the parametric cure-mixture estimate illustrated with a line and the 95% confidence interval shown as a shaded area.