Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Methods 203 (2022) 584–593

Contents lists available at ScienceDirect

Methods

journal homepage: www.elsevier.com/locate/ymeth

Data-driven case fatality rate estimation for the primary lineage of SARS-CoV-2 in Poland

Krzysztof Gogolewski [a], Błażej Miasojedow [a], Małgorzata Sadkowska-Todys [b], Małgorzata Stepień [b], Urszula Demkow [c], Agnieszka Lech [d], Ewa Szczurek [a], Daniel Rabczenko [e], Magdalena Rosińska [b], Anna Gambin [a,*]

[a] Faculty of Mathematics, Informatics and Mechanics, University of Warsaw, Warsaw, Poland
[b] Department of Infectious Disease Epidemiology and Surveillance, National Institute of Public Health NIH – National Research Institute, Warsaw, Poland
[c] Department of Laboratory Diagnostics and Clinical Immunology of Developmental Age, Medical University of Warsaw, Warsaw, Poland
[d] Department of Medical Microbiology, Medical University of Warsaw, Warsaw, Poland
[e] Department for Monitoring and Analysis of Population Health Status, National Institute of Public Health NIH – National Research Institute, Warsaw, Poland

ARTICLE INFO

Keywords:
COVID-19
Case fatality rate
Mortality in Poland
Original SARS-CoV-2 lineage

MSC:
92D30
62N02
62–07

ABSTRACT

After more than one and a half year since the COVID-19 pandemics outbreak the scientific world is constantly trying to understand its dynamics. In this paper of the case fatality rates (CFR) for COVID-19 we study the historic data regarding mortality in Poland during the first six months of pandemic, when no SARS-CoV-2 variants of concern were present among infected. To this end, we apply competing risk models to perform both uni- and multivariate analyses on specific subpopulations selected by different factors including the key indicators: age, sex, hospitalization. The study explores the case fatality rate to find out its decreasing trend in time. Furthermore, we describe the differences in mortality among hospitalized and other cases indicating a sudden increase of mortality among hospitalized cases at the end of the 2020 spring season. Exploratory and multivariate analysis revealed the real impact of each variable and besides the expected factors indicating increased mortality (age, comorbidities) we track more non-obvious indicators. Recent medical care as well as the identification of the source contact, independently of the comorbidities, significantly impact an individual mortality risk. As a result, the study provides a twofold insight into the COVID-19 mortality in Poland. On one hand we explore mortality in different groups with respect to different variables, on the other we indicate novel factors that may be crucial in reducing mortality. The later can be coped, e.g. by more efficient contact tracing and proper organization and management of the health care system to accompany those who need medical care independently of comorbidities or COVID-19 infection.

1. Introduction

As the epidemic of the coronavirus disease 2019 (COVID-19) caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) unfolds, the case fatality rates (CFRs) of COVID-19 are still being debated and research is ongoing on how to best predict severe course of the disease. The first estimates provided by the World Health Organization (WHO) [1] and China [2] suggested that the CFR ranges between 3.4% and 5.8%. However, it now became quite clear that the observed CFR may differ considerably not only by country [3,4] but also by the variants of SARS-CoV-2 [5,6].

In 2020, for the original SARS-CoV-2 strain, a couple of varying CFR estimates were presented on the national level for the first phase of outbreaks, e.g. estimates made for Italy, indicated a much higher fatality rate than in China [7], but CFR observed in Germany was already lower (1.2%) [8]. Some of these differences were subsequently attributed to the different age distributions in Asian and European countries as the age specific CFRs appeared to be more in line [9,10]. However, in 2021 over longer time span, the fatality rate of the SARS-CoV-2 in China was reported to decrease to 0.7% (95% CI 0.4–1.0%) [11], while German update showed an important age-dependent increase to approximately 7.5% during the epidemic peak [12]. Finally, a study for hospitalized cases in Brazil presented a decreasing trend over time of CFR from 31.8% (March 2020) to 18.2% (October 2020) [13].

In addition, as many cases are in fact asymptomatic or paucisymptomatic they may go undetected, depending on local testing

* Corresponding author.

https://doi.org/10.1016/j.ymeth.2022.01.006
Received 14 August 2021; Received in revised form 5 December 2021; Accepted 18 January 2022
Available online 24 January 2022

policies [14]. In fact, in a large multi-country overview the observed CFRs were negatively correlated with the testing frequency [15]. As already exemplified, adjusting the CFR to account for undetected cases may substantially lower the estimates [16,17].

As opposed to the fatality rate among the diagnosed cases, the rate of deaths among all infected individuals, including those who never developed symptoms, was termed infection fatality rate (IFR). IFR estimates usually require additional data on the true rate of infections, such as data coming from seroprevalence studies. The IFR estimates based on this method in 2020 were reported of order 0.3%–0.7% [18,19], while later the aforementioned German study estimated IFR between 1.5% and 3.5%. [12]. Of note, even if the IFR represents more closely the biological process, still heterogeneity between countries is expected relating for e.g. to age distribution of cases or efficiency of health care system provision [18]. Studying individual level predictors of IFR is often difficult due to lack of representative samples of infected individuals and therefore more evidence is available at the level of predictors of the outcomes among the diagnosed cases. While a number of predictors of the severe course of SARS-CoV-2 have been identified, there still remain open questions. Both IFR and CFR increase with age and are impacted by coexisting chronic conditions [20,21]. The relation with age is consistent across practically all studies and the risk is slightly higher among men. Deaths are sporadic among children and young adults, but increase from the age of about 50 and the IFR is estimated to reach 8% - 15% at the age of 80 or above [22,23,16]. Among chronic conditions the strongest predictors of poor outcomes tend to be the cardiovascular diseases - especially coronary heart disease, congestive heart failure and cerebrovascular disease, but also hypertension, diabetes and obesity [21,24,20,25]. Other diseases found to deteriorate the clinical course include chronic respiratory diseases, chronic kidney diseases and malignancies [26].

Finally, there remains a matter of genetic diversity of SARS-CoV-2. As of August 2021 four variants of concern (VoC) are spread around the world [27]. Their transmissibility is reported to increase in contrast to the primary virus linage [28,29] and the effectiveness of vaccines in use is lower than against the original strain of the virus [30]. Even though the spread of SARS-CoV-2 B.1.1.7 (Alpha) and B.1.617.2 (Delta) VoC is observed worldwide, their severity is constantly being studied. Latest reports, however, consequently prove that the severity of COVID-19 caused by the original (non-VoC) SARS-CoV-2 strain [5,31]. Importantly for our study, it was confirmed that the first non-outbreak of the Alpha VoC was detected in the UK on September 2020, while its outbreak subvariant was detected on December 2020 [32].

In addition to the dependency on the variant, severity of the disease may also depend on co-infections with other viral and bacterial respiratory pathogens. Bacterial co-infections appear to worsen the prognosis of viral pneumonias and there is already some evidence that this is also the case in COVID-19 pneumonia [33,34]. Worse prognosis was also observed in influenza co-infected patients [35]. The frequency of co-infections noted so far was not high [33], but some of these infections display strong seasonality and their impact may be higher in the winter season.

We aim at estimating CFR for Poland over first six months after the outbreak, when only the original strain of SARS-CoV-2 was present in the population. We investigate the possible trends by the time period, differences by age and other key factors influencing CFR levels. During the period of analysis the testing strategy remained stable, targeting testing people who were in contact with a confirmed or probable case as well as individuals presenting with COVID-19 compatible symptoms, especially with severe acute respiratory infection. In addition, screening during outbreaks was also in place. Other non-pharmaceutical measures were variable over the time period [36,37]. Therefore, in order to understand the possible confounding factors in terms of temporal changes of the affected populations we characterize the cases by the infection context. In addition we include a separate analysis in the subgroup of hospitalized cases among whom the testing for SARS-CoV-2 is considered more complete. The next section describes all the methods used in the presented work. Moreover, for convenience of the reader and more readability they are also visualized as a workflow diagram in Fig. 1.

2. Methods

2.1. Data source

SARS-CoV-2 genomic data. The analyzed genomic data comes from GISAID database (as of July 21, 2021) [38], where also our data collected between May 16, 2020 and February 12, 2021 from 31 patients from children's hospital are available. Out of more than a million coronavirus genome sequences from 172 different countries and territories available in GISAID, in our study we consider all samples collected in Poland before January 1, 2021 (see supplementary file SarsCov2Poland.tsv). Finally, we use 546 samples in the study, out of which 13 come from our cohort.

Epidemiological data. We used data collected as part of routine mandatory surveillance of COVID-19 in Poland. Cases of COVID-19 are reported to the State Sanitary Inspection (SSI) by clinicians who diagnose the disease as well as by laboratories, who obtain a positive result from the RT-PCR (reverse-transcription polymerase chain reaction) or antigen test. Based on the notifications, the SSI performs contact tracing, issues quarantine decisions in case of close contact and collects basic demographic and clinical information as well as contact tracing information. The cases are followed until death or recovery. The resulting information is documented in the central database of the Epidemiological Investigations Registration System operated by the Department of Epidemiology of Infectious Diseases and Surveillance National Institute of Public Health NIH - National Research Institute [39]. The system is comprehensive, but substantial delays in data registration are noted. For the purpose of this study we extracted data as of September 3, 2020. This date coincides with substantial change to the COVID-19 testing criteria in Poland, which could significantly impact the observed case fatality rates. Moreover, it ensures that all studied cases were infected with non-VoC SARS-CoV-2 strains.

2.2. Data preprocessing

SARS-CoV-2 genomic data. Genomes of SARS-CoV-2 collected in Poland in 2020 were analyzed for the presence of characteristic variant substitutions in the spike protein and possible classification as any of variant of interest/concern. Out of 546 genomes only two were classified as B.1.1.7 VoC by the Pangolin tool [40]. However, one of the samples had to be wrongly described since its collection date is set to April 22, 2020 which contradicts with a common agreement on the first occurrence of alpha variant on September 2020 [32]. The other sample had the collection date set to December 22, 2020 and was used in the analysis. All other samples are neither of interest nor of concern variants and were composed mostly of B.1 lineage (524 samples, 96.15%) from which 453 (83.12%) of sublineage B.1.1.

To assess the genetic distance among samples a phylogenetic tree was built for randomly selected set of sequences form Poland and other European countries. The Bayesian phylogenetic tree was created using BEAST2 software (R wrapper `babette`) using the MCMC method, the HKY nucleotide substitution model, strict molecular clock, and the *a priori* Coalescent Exponential Population model [41].

Epidemiological data. For the purpose of this work, the initial date for calculation of survival was defined as the date of onset of symptoms for symptomatic cases and the date of laboratory positive test results for asymptomatic cases. In case the onset date was unknown, the earlier of the date of the first positive laboratory test result and the date of clinical visit was recorded as the initial date. We defined the follow-up time as the time until death of any cause or the date, when the recovery was registered. For unresolved cases the censoring date was set as the last

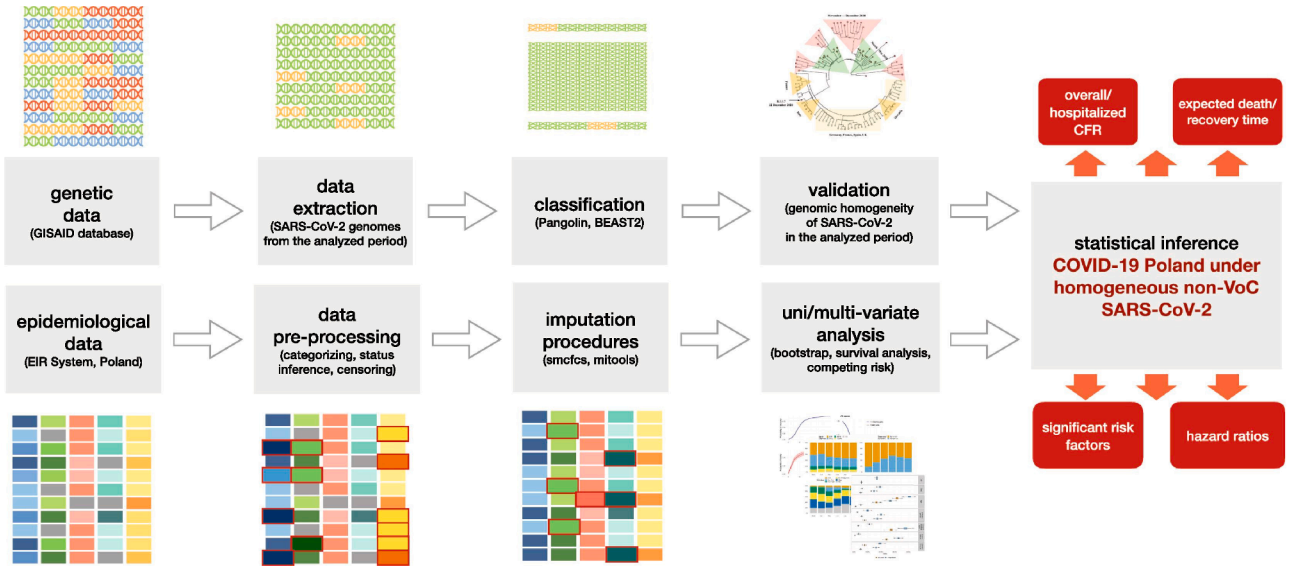

Fig. 1. Workflow diagram visualizing all of the steps that were taken to obtain the final results in the article.

date the status was recorded or, if this date was missing, the date the case was interviewed or the date the record was last modified. All pre-processing transformations of the raw data that were made to perform the statistical analyses are listed in the Table ST1.

Additionally, it should be emphasized, that the clinical symptoms were collected through check-boxes. Therefore, if the field was not completed (a symptom was not reported), but the patient was not coded as asymptomatic or the patient was hospitalized or the patient died – the symptoms were re-encoded as missing (not available, NA) values. This correction was introduced in order to better understand the distribution of symptoms, however, we cannot preclude any further misclassification.

2.3. Statistical analysis

We decided to use the competing risk model for the survival analysis. The multi-state models after one event occurs another may occur according to a given transition probability. On the contrary, in our situation we deal with two possible outcomes (death and recovery) and we observe the time only the first of them. One way of modeling competing risks employs standard methods of survival analysis for all risks separately treating other events as censored. This approach, however, is not relevant in most applications as it requires independence of risks. Hence we applied the methodology from [42,43], i.e. the generalized Cox model of competing risks. We assume a proportional hazards model for the subdistribution of the competing risk, which was in our case death. Therefore, the advantages of our approach are twofold. Not only we bypass the independence of risks, but also we investigate the covariates' effect on the cumulative incidence function and, as a result, we infer the absolute risk of the studied outcome over time. This approach was postulated as the most appropriate for investigating COVID-19 CFR [44–47].

In case the for some covariate patterns the observation time is shorter than the expected time needed to leave the system the estimated probability of death will be underestimated. In our study this is especially true for the most recent time periods. To overcome this issue we introduce a correction in the spirit of Ghani [48].

While Ghani uses two separate models for two outcomes treating the other one as a censoring event, we estimate them in the competing risk model. However, as the next step we introduce the similar correction as follows. Using this method the CFR is calculated with the formula:

$$\mathrm{CFR} = \frac{\widehat{\theta_0}}{\widehat{\theta_0} + \widehat{\theta_1}},$$

where $\widehat{\theta_0}$ and $\widehat{\theta_1}$ are cumulative hazards for death and recovery at the maximum observation time, respectively. Both of these were estimated as cumulative incidence functions from competing risks data [49]. Confidence intervals for both probability of fatality, probability of recovery and CFR are estimated using non-parametric bootstrap with 1000-fold resampling [50]. Competing risks proportional hazard regression [43] was used to investigate factors related to fatality rates.

The models were applied to all cases data and separately to the hospitalized cases only. The idea of the second analysis was that the registration and data collection of cases requiring hospitalisation should be more complete regardless of testing strategies or local shortages

Some of possibly important covariates suffered from a large proportion of missing values (NAs). We therefore used multiple imputations of missing covariates by substantive model compatible fully conditional specification. This method is a modification of multiple imputations by chained equations developed by Bartlett et al. [51], in particular for competing risks models, and implemented in R package `smcfcs` [52]. Finally, the results are aggregated to determine the multivariate risk values based on the methods implemented in R package `mitools` [53]. Alongside, the univariate risk values were calculated using functions provided by Fine and Gray [42] and implemented in the R `cmprsk` package [54].

3. Results

As of September 3, 2020, there were 32,447 registered cases in the surveillance database, which constituted 47% of the 69,129 cases officially reported in Poland by this date. Of the total number, the status (death, recovery, unresolved, with the last status contributing to the censored cases in the further crr analysis) was not recorded for 882 (2.72%). Further, for 2,796 cases (8.86%) the dates were inconsistent resulting in negative follow-up time. Finally, we included 28,769 cases in the analysis, of whom 984 (3.42%) died, 11,880 (41.29%) recovered and the disease outcome was unresolved at the censoring time for 15,905 (55.29%). The mean observation time available per case was 14.6 days (+/- 13.8), median 13 days (IQR 2–22). There were comparable numbers of males and females in the study group and the majority of the cases (81.8%) occurred among people younger than 65. Nearly 40% of infections could be attributed to documented household contact

or contact at work, and if cases with missing data are excluded this percent rises to 57%. Comorbidities were reported for 13.9% of all cases (19.8% excluding cases with missing data), but in 71.8% of those who died (89.3% excluding cases with missing data) (See Table 1).

Based on the provided genomic data we confirm that the all of the surveillance data gathered in our study were infected by the non-VoC SARS-CoV-2 strain. Fig. 2 represents a phylogenetic relationship among samples collected from polish patients before January 1, 2021, where only one is classified as B.1.1.7 SARS-CoV-2 VoC. The sample was collected on December 22, 2020. The clade of polish samples clearly separates from randomly selected European samples classified as B.1.1.7 variant, which was the first known variant of concern. Therefore, all presented results regarding CFR, time to death/recovery and all other indicators in the first six month after outbreak in Poland concern the non-VoC SARS-CoV-2 strain.

The characteristics of infections were not stable over time. Initially, a substantial share of cases occurred among health care workers, which declined by week 19. During weeks 19 – 22 there was a substantial increase in the proportion of cases among coal mine workers. This increase coincided with a larger proportion of asymptomatic cases and the smallest proportion of elderly people. The proportion of cases, for whom no contact with a confirmed case was identified increased over time and so did the proportion of cases infected in the contexts other than household or the working place (see Supplementary Fig. S1).

Majority of cases who died reported the typical COVID-19 symptoms such as fever, cough and dyspnea (Fig. 3). Among cases who recovered,

Table 1
Characteristics of the COVID-19 cases included in the analysis: overall, among hospitalized cases and death outcome. (N.D. – no data available, RMC – Recent Medical Care).

		All cases		Hospitalized cases		Deaths	
		N	%	N	%	N	%
Total		28769	100	7190	100	984	100
Sex	Female	14606	50.8	3750	52.2	454	46.1
	Male	14159	49.2	3437	47.8	530	53.9
	N.D.	4	0	3	0	0	0
Age group (years)	0–17	2187	7.6	253	3.5	1	0.1
	18–44	11837	41.1	1654	23	21	2.1
	45–64	9531	33.1	2317	32.2	154	15.7
	65–74	2498	8.7	1306	18.2	265	26.9
	75+	2492	8.7	1593	22.2	536	54.5
	N.D.	224	0.8	67	0.9	7	0.7
Residence type	City 100,000 or larger	6324	22	1470	20.4	211	21.4
	City 50,000–99,000	2400	8.3	724	10.1	110	11.2
	City 20,000–49,000	3467	12.1	862	12	106	10.8
	City < 20,000	4096	14.2	1109	15.4	142	14.4
	Rural	12390	43.1	3004	41.8	414	42.1
	N.D.	92	0.3	21	0.3	1	0.1
RMC	Yes	1544	5.4	980	9.4	320	32.5
	No	17248	59.9	3932	37.7	384	39.0
	N.D.	9976	34.7	5517	52.9	280	28.5
Occupation	Child (not working)	2367	8.2	274	3.8	1	0.1
	Health care worker	3683	12.8	592	8.2	4	0.4
	Coal mine worker	2101	7.3	38	0.5	1	0.1
	Other	7572	26.3	1596	22.2	41	4.2
	Not working	6504	22.6	3318	46.1	750	76.2
	N.D.	6542	22.7	1372	19.1	187	19
Hospital-ization	No	18339	63.7	0	0	43	4.4
	Yes	7190	25	7190	100	903	91.8
	N.D.	3240	11.2	0	0	38	3.8
Infection period	March-April	6268	21.8	2409	33.5	386	39.2
	May-June	12206	42.4	2792	38.8	478	48.6
	July-August	10294	35.8	1988	27.7	120	12.2
	N.D.	0	0	0	0	0	0
Contact source	Travel to affected country	1077	3.7	478	6.6	12	1.2
	Household contact	6098	21.2	1474	20.5	163	16.6
	Contact at work	5327	18.5	512	7.1	15	1.5
	Contact in health care (including long term facilities)	1303	4.5	882	12.3	279	28.4
	Other	1876	6.5	403	5.6	43	4.4
	Not identified	4429	15.4	1534	21.3	251	25.5
	N.D.	8659	30.1	1907	26.5	221	22.5
Comorb-idities	No	16226	56.4	2807	39	85	8.6
	Yes. including:	4005	13.9	2448	34	707	71.8
	- Cardiovascular diseases including hypertension	2191	7.6	1378	19.2	433	44
	- Chronic pulmonary diseases including asthma and COPD	407	1.4	264	3.7	103	10.5
	- Diabetes	644	2.2	438	6.1	123	12.5
	N.D.	8538	29.7	1935	26.9	192	19.5
Disease outcome	Ongoing	11880	41.3	2066	28.7	0	0
	Recovery	15905	55.3	4221	58.7	0	0
	Death (of COVID-19)	788	2.7	750	10.4	788	80.1
	Death (not of COVID-19)	157	0.5	120	1.7	157	16
	Death (unknown cause)	39	0.1	33	0.5	39	4

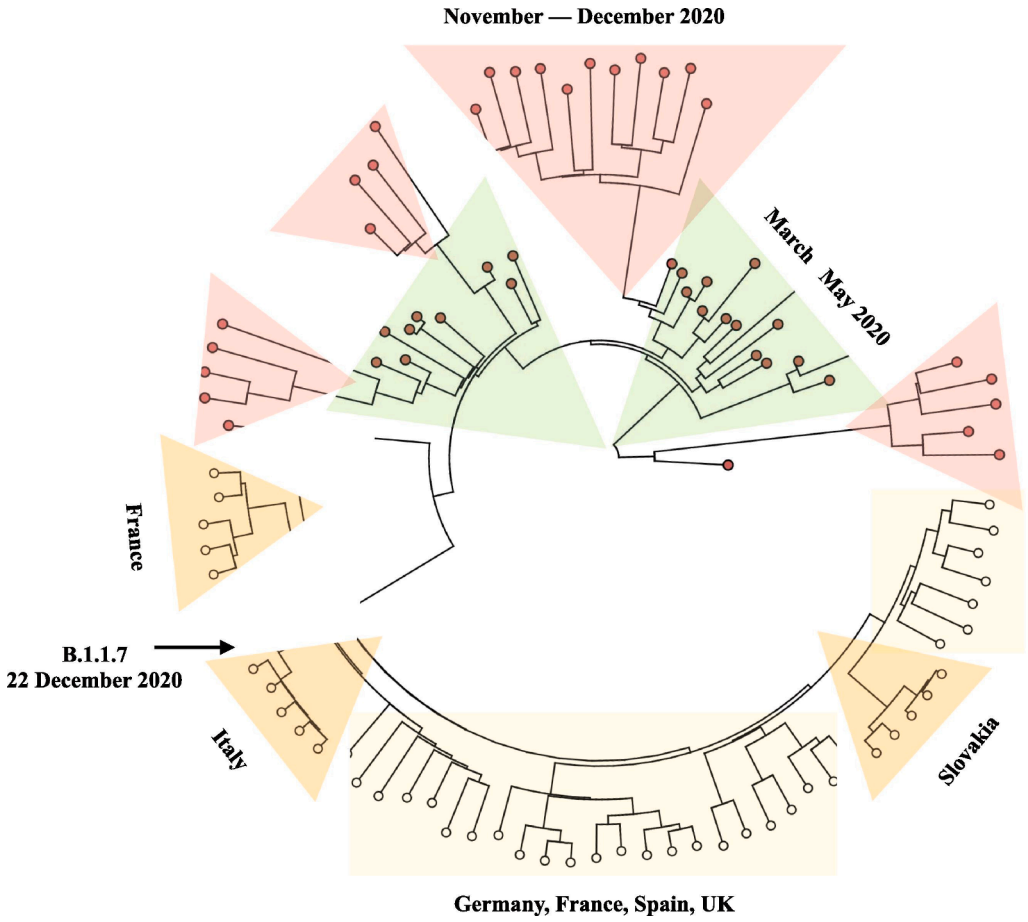

Fig. 2. Phylogenetic tree for 50 randomly selected samples collected in Poland during the year 2020 (red circles). Samples from other European countries are randomly selected from the B.1.1.7 lineage.

5,959 (40.39%) were asymptomatic – had developed no symptoms of infection (59.61% had at least one symptom reported). Among these cases relatively fewer reported dyspnea and more reported disturbance of taste or smell. The symptom distribution was comparable among male and female cases as well as with respect to age and infection period (see Supplementary Figs. S2 and S3).

The overall CFR was estimated at 5.43% (95% CI: 5.08% – 5.73%), while only for hospitalized cases at 16.9% (CI: 15.8% - 17.9%) (see Supplementary Fig. S4). There was 0.91% of difference in CFR between male and female, and 4.3% in case of hospitalized. As expected, the CFR was strongly influenced by age both overall and for hospitalized cases. The values spanned between 1.19% (CI: 1.01%–1.36%) for younger than 65 years old up to 40.8% (CI: 37.9%–43.6%) for older than 75 years old, hospitalized cases. The CFR evolved in time. While larger values were observed in the spring especially early in March, the fatality in the summer was greatly decreased. The effect of the summer season on the CFR, resulted in a decrease to the level of 2.52% (CI: 2.06% - 3.02%) from 7.83% (CI: 7.08%–8.61%) in the first weeks after the outbreak. Interestingly, analogous decrease for the hospitalized cases from 17.3% (CI: 15.7%– 19.1%) to 10.2% (CI: 8.18% – 12.1%) was preceded by an increase up to 20.0% (CI: 18.4%–21.7%) during the May-June period. Additionally, three other factors were significant. Presence of comorbidities imply 23.2% increase in overall CFR and 29.4% for hospitalized cases. Finally, binary-valued indicators regarding: (i) the need of recent medical care up to 14 days before the infection, and (ii) complete contact source documentation were significantly differentiating the CFR level (see Fig. 4 and Supplementary Tab. ST2). Additionally, we report that the expected time to death was on average 14.8 days (median: 11 days, IQR: 31 days), with the time to recovery on average 23.9 days (median 22 days, IQR: 47 days).

Furthermore, we provide an uni- and multi-variate competing risk analysis based on the factors associated with a significant increase of CFR. All variables except from the infection period introduce an increased hazard ratio that is statistically significant for both all and hospitalized cases. Males, elder (proportionally to their age), comorbidities, those who needed medical care up to 14 days before the infection and undocumented source contact impose increased chances of fatality. Simultaneously, the infection period in an univariate analysis stays significant for both all and hospitalized cases. Nonetheless, multivariate analysis reduces its significance in favor of the remaining factors (see Table 2).

4. Discussion

We estimate the case fatality rate (CFR) in a large cohort of patients in Eastern Europe, demonstrating comparable values to the previously reported levels in other countries. We confirm the strong association of the CFR with the age group and co-existing conditions. In addition, we

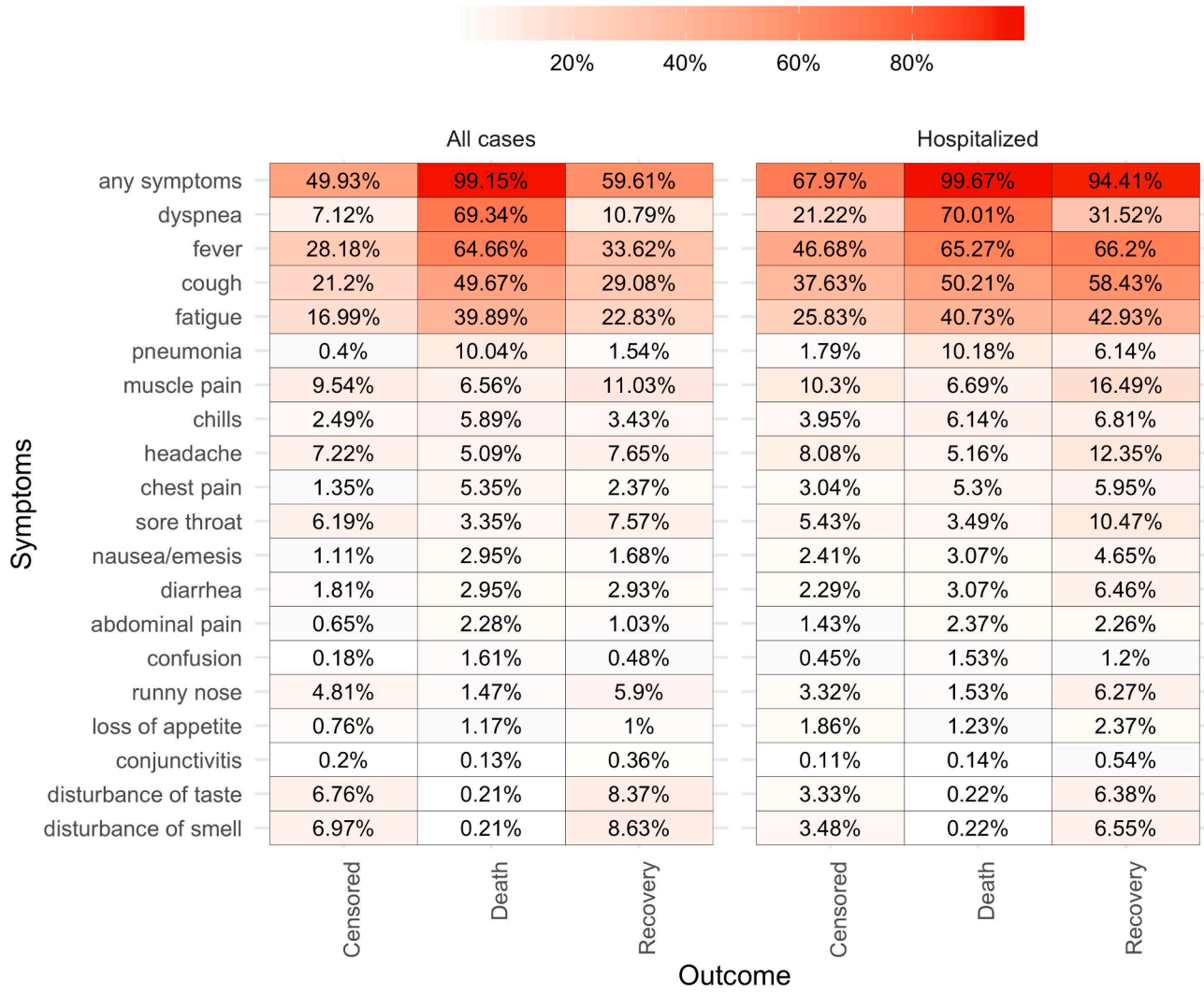

Fig. 3. Frequency of reported symptoms, by outcome (Recovery, Death, Censored) and by hospital treatment (All cases, Hospitalized). The percent in each cell corresponds to the fraction of patients of given sex and outcome that showed a given symptom, e.g. the value in the first row in the first column is understood as 49.93% of all cases who have censored outcome were symptomatic (showed some symptoms). The analogous figure describing symptoms among different age groups as well as onset months can be found in the Supplementary S2 and S3.

quantify the effect of the summer season on the CFR.

CFR shows significant differences between countries. It is influenced by such factors as: the age structure of the society, the wealth of the economy, the quality and efficiency of the healthcare system. Adequate measures taken by state authorities such as lockdown, massive testing and efficient strategies for contact tracking can significantly reduce CFR. In case of polish data, the estimated CFR value decreased from 7.83% to 2.52% with respect to infection period (March-April, July-August, respectively). Additionally, the determined decreasing trend is in accordance with the study by Ghayda et al. [55] were the time evolution of CFR in the several-month period after the outbreak is emphasized. Similarly, a study in Germany finds a decrease during the summer [12,56]. Interestingly, a study in Italy found no difference in CFR over the time [57]. Possible seasonality in the risk of severe course of COVID-19 is still not clear, but our findings as well as findings in Germany, are consistent with such hypothesis.

We found a 67.8% relative decrease in the estimated CFR during the spring and especially summer months. The case fatality rate as opposed to the infection fatality rate may depend on the testing policy, which is reflected in the possibility of the system to detect all the infections, not only the severe ones. Testing strategies comprised testing of the persons with severe acute respiratory infection and other symptoms suggestive of COVID-19, testing of quarantined people, screening of selected health care workers as well as testing of exposed individuals during identified outbreaks, mainly at work places. However, although we note that the composition of occupational groups as well as of the main transmission risks vary considerably between the analysed periods, they do not display a clear tendency, which could explain the decrease in CFR, when corrected for the age of the cases. Moreover, in our analysis we correct for this effect by including in the analysis the variable identifying for which cases the source of infection was known. As expected the CFR for cases with known source was lower indicating better case ascertainment, especially for milder cases. In addition, a similar pattern was noted for cases that needed hospital admission, for whom the testing

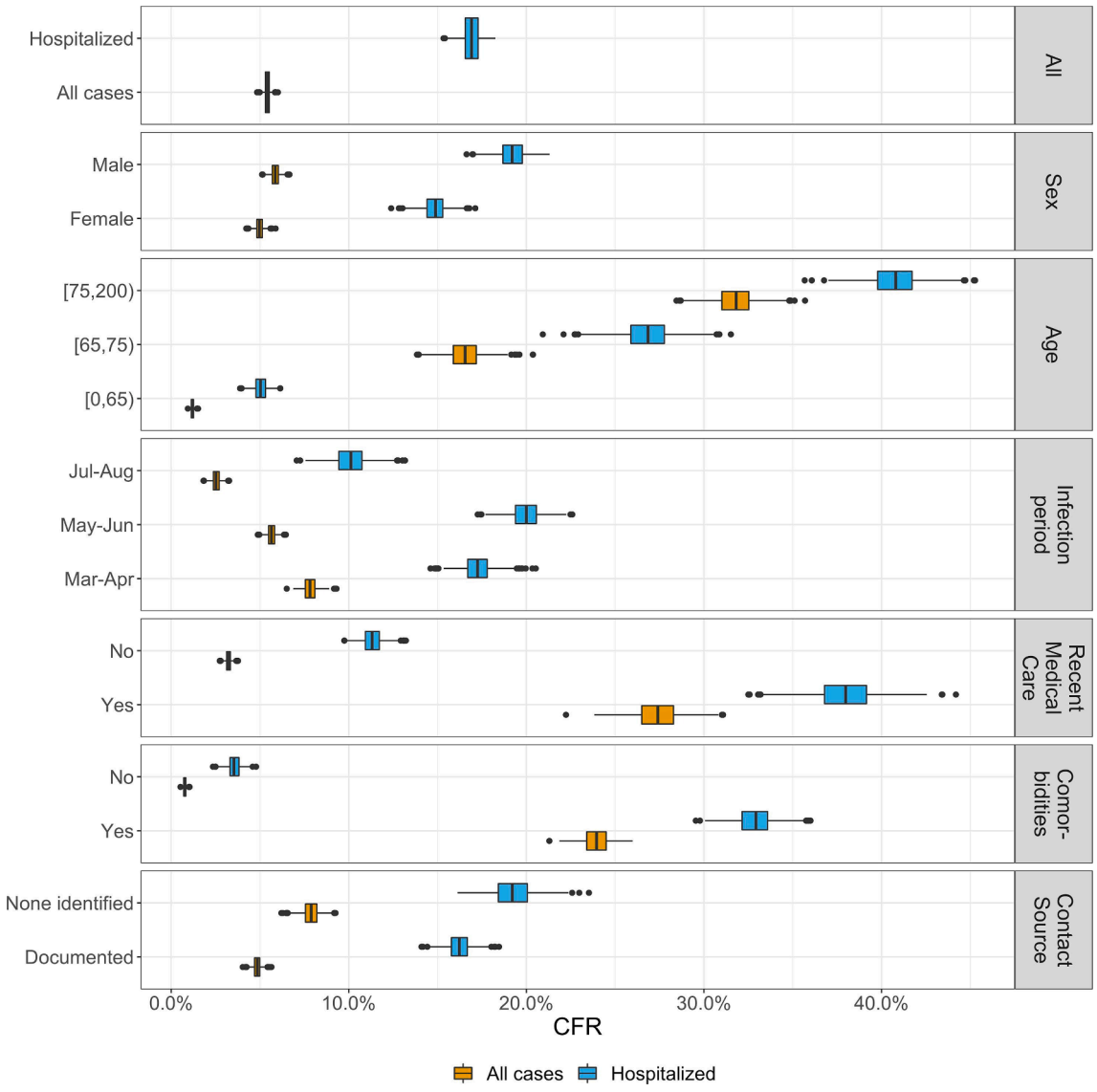

Fig. 4. The comparison of CFR values within different categories (left-hand side) determined by various factors(right-hand side). Each CFR is calculated for both all cases and hospitalized (orange and blue boxes, respectively). Each box corresponds to quartiles, the line to 95% confidence interval and points to outlier observations. Each horizontal line in the middle of the box corresponds to the median CFR value.

criteria were stable. Finally, the fatality rates decreased in multiple European countries, with different testing systems, making this explanation less likely. We hypothesize that the decrease could in fact be caused by lower virulence of the virus during the warmer periods of time which is in line with the findings of Spena et al. [58]. Other possible factor to take into account is lower incidence of other respiratory infections. A recent study shows that coinfection with another respiratory virus was a 2.7-fold (1.22–5.83) higher risk of fatal outcome [59].

Beside the possible CFR seasonality, improvement of the survival among hospitalized cases points to better clinical management. While initially all cases were hospitalized this lasted only during the initial 2 weeks of the epidemic. At latter stages home isolation was preferred for less severe cases. In addition the hospital resources during this first wave of epidemic were not exceeded. This leads to a conclusion that admission criteria remained comparable throughout the study period and could not explain the decreasing CFR. On the other hand the research on case management progressed over this time period more. Additionally, efficient organization and logistics in hospitals, including the establishment of hospital wards for COVID patients, might be an explanation for more favourable prognosis. Finally, clinical staff gained more experience with time and resulting physician learning curve might have also improve the situation of severe cases. Limitations. The data quality was not sufficient to run more elaborate analysis including the impact of comorbidities. Since the data was based on routinely collected data the information was not rich enough to attempt imputation of missing values for specific comorbidities. Disturbance of taste and smell was re-coded from text fields which could explain low percent of cases for whom this symptom was reported. Furthermore, we noted a very long time until recovery. The average of over 30 days is related to the strict criteria until the case may be considered recovered. During the time period included in the analysis 2 negative tests with 24 h were required before the case was considered recovered. Later evidence suggested that this requirement

Table 2
Estimated hazard ratios for factors associated with an increased CFR based on both multi- and uni-variate risk analysis performed on imputed data. The results are presented for: all (upper table) and hospitalized (bottom table) cases. Factors and their possible values are listed in the first two columns. Columns 4–6 describe the number (and percentage) of samples with a given outcome after 5-fold imputation run. The two last columns show the hazard ratio coefficients along with confidence interval in brackets and corresponding p-value. (RMC – Recent Medical Care, HR - Hazard Ratio).

All cases

variable	levels	censored	died	alive	Multivariate HR	UNIVARIATE HR
Sex	Female	6066 (41.5%)	454 (3.1%)	8086 (55.4%)	1	1
	Male	5809 (41%)	530 (3.7%)	7819 (55.2%)	1.54 (1.35–1.75), p = 0	1.21 (1.06–1.37), p = 0.0035
Age Group	[0,65)	9984.4 (42.1%)	177 (0.7%)	13567 (57.2%)	1	1
	[65,75)	996.6 (39.5%)	268 (10.6%)	1256.6 (49.8%)	4.21 (3.43–5.18), p = 0	14.6 (12.07–17.67), p = 0
	[75,200]	894 (35.6%)	539 (21.4%)	1081.4 (43%)	7.68 (6.3–9.35), p = 0	29.41 (24.78–34.91), p = 0
Infection Period	Mar-Apr	1547 (24.7%)	386 (6.2%)	4331 (69.1%)	1	1
	May-Jun	4223 (34.6%)	478 (3.9%)	7505 (61.5%)	0.95 (0.83–1.09), p = 0.4689	0.77 (0.67–0.88), p = 0.0002
	Jul-Aug	6105 (59.3%)	120 (1.2%)	4069 (39.5%)	0.64 (0.51–0.8), p = 0.0001	0.4 (0.32–0.49), p = 0
RMC	No	11082 (42%)	540.4 (2%)	14782.8 (56%)	1	1
	Yes	793 (33.6%)	443.6 (18.8%)	1122.2 (47.6%)	2.01 (1.75–2.31), p = 0	9.04 (7.93–10.3), p = 0
Comor- bidities	No	9704.4 (42.3%)	99 (0.4%)	13111.4 (57.2%)	1	1
	Yes	2170.6 (37.1%)	885 (15.1%)	2793.6 (47.8%)	10.19 (7.92–13.11), p = 0	33.91 (27.01–42.56), p = 0
Contact Source	Known	9173.6 (41.2%)	654.2 (2.9%)	12412.8 (55.8%)	1	1
	Uniden- tified	2701.4 (41.4%)	329.8 (5.1%)	3492.2 (53.5%)	1.73 (1.49–2.0), p = 0	1.82 (1.55–2.14), p = 0

Hospitalized

variable	levels	censored	died	alive	Multivariate HR	UNIVARIATE HR
Sex	Female	1062 (28.3%)	415 (11.1%)	2273 (60.6%)	1	1
	Male	1000 (29.1%)	488 (14.2%)	1948 (56.7%)	1.37 (1.2–1.57), p = 0	1.28 (1.12–1.45), p = 0.0003
Age Group	[0,65)	1224.2 (28.7%)	159.2 (3.7%)	2875.2 (67.5%)	1	1
	[65,75)	397.4 (30.1%)	256.2 (19.4%)	666.6 (50.5%)	2.92 (2.37–3.6), p = 0	5.64 (4.63–6.88), p = 0
	[75,200]	440.4 (27.4%)	487.6 (30.3%)	679.2 (42.3%)	4.75 (3.92–5.75), p = 0	9.14 (7.63–10.95), p = 0
Infection Period	Mar-Apr	371 (15.4%)	359 (14.9%)	1676 (69.7%)	1	1
	May-Jun	667 (23.9%)	436 (15.6%)	1689 (60.5%)	1.01 (0.87–1.17), p = 0.8901	1.3 (1.13–1.5), p = 0.0002
	Jul-Aug	1024 (51.5%)	108 (5.4%)	856 (43.1%)	0.87 (0.7–1.09), p = 0.233	0.72 (0.58–0.89), p = 0.003
RMC	No	1701.8 (29.5%)	492 (8.5%)	3569.6 (61.9%)	1	1
	Yes	360.2 (25.3%)	411 (28.9%)	651.4 (45.8%)	1.92 (1.6–2.29), p = 0	3.57 (3.06–4.16), p = 0
Comor- bidities	No	1129.6 (29.8%)	92.4 (2.4%)	2564.2 (67.7%)	1	1
	Yes	932.4 (27.4%)	810.6 (23.8%)	1656.8 (48.7%)	4.58 (3.64–5.77), p = 0	9.95 (7.98–12.4), p = 0
Contact Source	Known	1426.6 (28.2%)	604 (11.9%)	3029.6 (59.9%)	1	1
	Uniden- tified	635.4 (29.9%)	299 (14.1%)	1191.4 (56%)	1.51 (1.3–1.75), p = 0	1.27 (1.09–1.49), p = 0.0029

was not necessary and the recommendations were adjusted accordingly. On the other hand, we only considered recovery from acute disease. Long COVID-19 was not considered.

5. Conclusions

In the presented work we provide an insight into the CFR levels as well as time to recovery/death estimates in Poland, in the first six months of COVID-19 epidemics, which was caused exclusively by the original strain of SARS-CoV-2. The survival analysis methodology used in the study is based on the competing risk regression along with corrections suggested by Ghani [48], which provide robust and reliable estimates of epidemiological COVID-19 risk factors. The work points out that the CFR does not only depend on already known factors i.e. age, sex or comorbidities, but also is influenced by the recent need of medical care and possibly identification of a contact source. Moreover, we show that the summer months are characterized by a significantly lower CFR than other seasons. The natural extension of this work is a consequence of novel variants of concern, which requires analogous analyses to be performed among patients infected by specific SARS-CoV-2 variants. Finally, the ongoing vaccination programs open a question of how the CFR levels change with an increasing percent of vaccinated people.

CRediT authorship contribution statement

Krzysztof Gogolewski: Data curation, Writing - original draft, Writing - review & editing. **Błażej Miasojedow:** Software, Validation. **Małgorzata Sadkowska-Todys:** Visualization, Investigation. **Małgorzata Stepień:** Visualization, Investigation. **Urszula Demkow:** Supervision. **Agnieszka Lech:** Visualization, Investigation. **Ewa Szczurek:** Conceptualization, Methodology, Software. **Daniel Rabczenko:** Data curation, Writing - original draft. **Magdalena Rosińska:** Conceptualization, Methodology, Software. **Anna Gambin:** Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.ymeth.2022.01.006.

References

[1] World Health Organization, Who director-general's opening remarks at the media briefing on covid-19 – 3 March 2020, WHO, Tech. rep., 2020 https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19-3-march-2020.

[2] Y.N. Mi, T.T. Huang, J.X. Zhang, Q. Qin, Y.X. Gong, S.Y. Liu, H.M. Xue, C.H. Ning, L. Cao, Y.X. Cao, Estimating the instant case fatality rate of COVID-19 in China, Int. J. Infect Dis. 97 (2020) 1–6.

[3] U. Patel, P. Malik, D. Mehta, D. Shah, R. Kelkar, C. Pinto, M. Suprun, M. Dhamoon, N. Hennig, H. Sacks, Early epidemiological indicators, outcomes, and interventions of COVID-19 pandemic: A systematic review, J. Global Health 10 (2).

[4] T. Ahammed, A. Anjum, M.M. Rahman, N. Haider, R. Kock, M.J. Uddin, Estimation of novel coronavirus (COVID-19) reproduction number and case fatality rate: A systematic review and meta-analysis, Health Sci. Rep. 4 (2) (2021), e274.

[5] P.H. England, Sars-cov-2 variants of concern and variants under investigation in england – technical briefing 18, Tech. rep., PHE (7 2021).https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1001358/Variants_of_Concern_VOC_Technical_Briefing_18.pdf.

[6] D.J. Grint, K. Wing, E. Williamson, H.I. McDonald, K. Bhaskaran, D. Evans, S.J. Evans, A.J. Walker, G. Hickman, E. Nightingale, A. Schultze, C.T. Rentsch, C. Bates, J. Cockburn, H.J. Curtis, C.E. Morton, S. Bacon, S. Davy, A.Y. Wong, A. Mehrkar, L. Tomlinson, I.J. Douglas, R. Mathur, P. Blomquist, B. MacKenna, P. Ingelsby, R. Croker, J. Parry, F. Hester, S. Harper, N.J. DeVito, W. Hulme, J. Tazare, B. Goldacre, L. Smeeth, R.M. Eggo, Case fatality risk of the SARS-CoV-2 variant of concern B.1.1.7 in England, 16 November to 5 February, Euro Surveill 26 (11).

[7] G. Onder, G. Rezza, S. Brusaferro, Case-Fatality Rate and Characteristics of Patients Dying in Relation to COVID-19 in Italy, JAMA 323 (18) (2020) 1775–1776, publisher: American Medical Association. doi:10.1001/jama.2020.4683.https://jamanetwork.com/journals/jama/fullarticle/2763667.

[8] N. Stafford, Covid-19: Why Germany's case fatality rate seems so low, BMJ 369, publisher: British Medical Journal Publishing Group Section: Feature. doi:10.1136/bmj.m1395.https://www.bmj.com/content/369/bmj.m1395.

[9] R. Porcheddu, C. Serra, D. Kelvin, N. Kelvin, S. Rubino, Similarity in Case Fatality Rates (CFR) of COVID-19/SARS-COV-2 in Italy and China, J. Infection Developing Countries 14 (02) (2020) 125–128, number: 02. doi:10.3855/jidc.12600.

[10] C. Dudel, T. Riffe, E. Acosta, A. van Raalte, C. Strozza, M. Myrskylä, Monitoring trends and differences in COVID-19 case-fatality rates using decomposition methods: Contributions of age structure and age-specific fatality, PloS One 15 (9) (2020), e0238904.

[11] X.-S. Zhang, E. Vynnycky, A. Charlett, D.D. Angelis, Z. Chen, W. Liu, Transmission dynamics and control measures of COVID-19 outbreak in china: a modelling study, Sci. Rep. 11 (1). doi:10.1038/s41598-021-81985-z.

[12] C. Staerk, T. Wistuba, A. Mayr, Estimating effective infection fatality rates during the course of the COVID-19 pandemic in germany, BMC Public Health 21 (1). doi:10.1186/s12889-021-11127-7.

[13] I.R. Zimmermann, M.N. Sanchez, G.S. Frio, L.C. Alves, C.C. de Aguiar Pereira, R. T. de Sousa Lima, C. Machado, L.M.P. Santos, E.N. da Silva, Trends in COVID-19 case-fatality rates in brazilian public hospitals: A longitudinal cohort of 398,063 hospital admissions from 1st march to 3rd october 2020, PLOS ONE 16 (7) (2021), e0254633.

[14] G. Giordano, F. Blanchini, R. Bruno, P. Colaneri, A. Di Filippo, A. Di Matteo, M. Colaneri, Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy, Nature Medicine 26 (6) (2020) 855–860, number: 6 Publisher: Nature Publishing Group. doi:10.1038/s41591-020-0883-7.

[15] L.-L. Liang, C.-H. Tseng, H.J. Ho, C.-Y. Wu, COVID-19 mortality is negatively associated with test number and government effectiveness, Sci. Rep. 10 (1) (2020) 12567, https://doi.org/10.1038/s41598-020-68862-x, http://www.nature.com/articles/s41598-020-68862-x.

[16] T.W. Russell, J. Hellewell, C.I. Jarvis, K. v. Zandvoort, S. Abbott, R. Ratnayake, C. C.-. w. Group, S. Flasche, R.M. Eggo, W.J. Edmunds, A.J. Kucharski, Estimating the infection and case fatality ratio for coronavirus disease (COVID-19) using age-adjusted data from the outbreak on the Diamond Princess cruise ship, February 2020, Eurosurveillance 25 (12) (2020) 2000256, publisher: European Centre for Disease Prevention and Control. doi:10.2807/1560-7917.ES.2020.25.12.2000256. https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.2020.25.12.2000256.

[17] M. Lipsitch, Estimating case fatality rates of COVID-19, Lancet. Infect. Dis 20 (7) (2020) 775.

[18] G. Meyerowitz-Katz, L. Merone, A systematic review and meta-analysis of published research data on COVID-19 infection fatality rates, Int. J. Infectious Dis. 101 (2020) 138–148, https://doi.org/10.1016/j.ijid.2020.09.1464, https://linkinghub.elsevier.com/retrieve/pii/S1201971220321809.

[19] J.P.A. Ioannidis, Infection fatality rate of COVID-19 inferred from seroprevalence data, Bulletin of the World Health Organization.

[20] S. Katzenschlager, A.J. Zimmer, C. Gottschalk, J. Grafeneder, A. Seitel, L. Maier-Hein, A. Benedetti, J. Larmann, M.A. Weigand, S. McGrath, C.M. Denkinger, Can we predict the severe course of COVID-19 - a systematic review and meta-analysis of indicators of clinical outcome?, medRxiv: The Preprint Server for Health Sciences doi:10.1101/2020.11.09.20228858.

[21] V. Chidambaram, N.L. Tun, W.Z. Haque, M.G. Majella, R.K. Sivakumar, A. Kumar, A.T.-W. Hsu, I.A. Ishak, A.A. Nur, S.K. Ayeh, E.L. Salia, A. Zil-E-Ali, M.A. Saeed, A. P.B. Sarena, B. Seth, M. Ahmadzada, E.F. Haque, P. Neupane, K.-H. Wang, T.-M. Pu, S.M.H. Ali, M.A. Arshad, L. Wang, S. Baksh, P.C. Karakousis, P. Galiatsatos, Factors associated with disease severity and mortality among patients with COVID-19: A systematic review and meta-analysis, PloS One 15 (11) (2020), e0241541, https://doi.org/10.1371/journal.pone.0241541.

[22] P. Poletti, M. Tirani, D. Cereda, F. Trentini, G. Guzzetta, V. Marziano, S. Buoro, S. Riboli, L. Crottogini, R. Piccarreta, A. Piatti, G. Grasselli, A. Melegaro, M. Gramegna, M. Ajelli, S. Merler, Age-specific SARS-CoV-2 infection fatality ratio and associated risk factors, Italy, February to April 2020, Eurosurveillance 25 (31).

[23] R. Pastor-Barriuso, B. Perez-Gomez, M.A. Hernan, M. Perez-Olmeda, R. Yotti, J. Oteo, J.L. Sanmartin, I. Leon-Gomez, A. Fernandez-Garcia, P. Fernandez-Navarro, I. Cruz, M. Martin, C. Delgado-Sanz, N. Fernandez de Larrea, J.L. Paniagua, J.F. Munoz-Montalvo, F. Blanco, A. Larrauri, M. Pollan, M. Pollan, Infection fatality risk for SARS-CoV-2: a nationwide seroepidemiological study in the non-institutionalized population of Spain, medRxiv (2020) 2020.08.06.20169722.

[24] W. Tian, W. Jiang, J. Yao, C.J. Nicholson, R.H. Li, H.H. Sigurslid, L. Wooster, J.I. Rotter, X. Guo, R. Malhotra, Predictors of mortality in hospitalized COVID-19 patients: A systematic review and meta-analysis, J. Med. Virol. doi:10.1002/jmv.26050.

[25] G. Pugliese, M. Vitale, V. Resi, E. Orsi, Is diabetes mellitus a risk factor for COronaVIrus Disease 19 (COVID-19)? Acta Diabetol. 57 (11) (2020) 1275–1285.

[26] Y. Zhou, Q. Yang, J. Chi, B. Dong, W. Lv, L. Shen, Y. Wang, Comorbidities and the risk of severe or fatal outcomes associated with coronavirus disease, A systematic review and meta-analysis, Int. J. Infectious Dis. 99 (2020) (2019) 47–56.

[27] Centers for Disease Control and Prevention, Sars-cov-2 variant classifications and definitions, Tech. rep., CDC (2021).https://www.cdc.gov/coronavirus/2019-ncov/variants/variant-info.html.

[28] F. Campbell, B. Archer, H. Laurenson-Schafer, Y. Jinnai, F. Konings, N. Batra, B. Pavlin, K. Vandemaele, M.D.V. Kerkhove, T. Jombart, O. Morgan, O. le Polain de Waroux, Increased transmissibility and global spread of SARS-CoV-2 variants of concern as at June 2021, Eurosurveillance 26 (24).

[29] J. Dagpunar, Interim estimates of increased transmissibility, growth rate, and reproduction number of the covid-19 b.1.617.2 variant of concern in the united kingdom, medRxiv.

[30] L. Bian, F. Gao, J. Zhang, Q. He, Q. Mao, M. Xu, Z. Liang, Effects of SARS-CoV-2 variants on vaccine efficacy and response strategies, Expert Rev. Vaccines 20 (4) (2021) 365–373.

[31] R. Herlihy, W. Bamberg, A. Burakoff, N. Alden, R. Severson, E. Bush, B. Kawasaki, B. Berger, E. Austin, M. Shea, E. Gabrieloff, S. Matzinger, A. Burdorf, J. Nichols, K. Goode, A. Cilwick, C. Stacy, E. Staples, G. Stringer, Rapid increase in circulation of the SARS-CoV-2 b.1.617.2 (delta) variant — mesa county, colorado, april–june 2021, MMWR. Morbidity and Mortality Weekly Report 70 (32). doi:10.15585/mmwr.mm7032e2.

[32] European Centre for Disease Prevention and Control, Sars-cov-2 variants of concern as of 5 august 2021, Tech. rep., ECDC (2021).https://www.ecdc.europa.eu/en/covid-19/variants-concern.

[33] L. Lansbury, B. Lim, V. Baskaran, W.S. Lim, Co-infections in people with COVID-19: a systematic review and meta-analysis, J. Infection 81 (2) (2020) 266–275.

[34] C. Feldman, R. Anderson, The role of co-infections and secondary infections in patients with COVID-19, Pneumonia 13 (1).

[35] J. Stowe, E. Tessier, H. Zhao, R. Guy, B. Muller-Pebody, M. Zambon, N. Andrews, M. Ramsay, J. Lopez Bernal, Interactions between SARS-CoV-2 and Influenza and the impact of coinfection on disease severity: A test negative design, medRxiv (2020) 2020.09.18.20189647.

[36] Helsinki Foundation for Human Rights, Coronavirus covid-19 outbreak in the eu fundamental rights implications; country - poland, Tech. rep., HFHR (2020).https://fra.europa.eu/sites/default/files/fra_uploads/poland-report-covid-19-april-2020_en.pdf.

[37] M. Ganczak, O. Pasek, Ł. D. –. Duma, D. Świstara, M. Korzeń, Use of masks in public places in poland during SARS-cov-2 epidemic: a covert observational study, BMC Public Health 21 (1).

[38] S. Elbe, G. Buckland-Merrett, Data, disease and diplomacy: GISAID's innovative contribution to global health, Global Challenges 1 (1) (2017) 33–46.

[39] W. Kitowska, M. Stepien, M. Rosinska, K. Szmulik, M. Jamsheer-Bratkowska, P. Budziosz, M. Sadkowska-Todys, COVID-19 module in SRWE system – application and use in epidemiological surveillance and reporting to international institutions, Przegl. Epidemiol. (2020) 416–431, 10.32394%2Fpe.74.35.

[40] A. Rambaut, E.C. Holmes, Á. O'Toole, V. Hill, J.T. McCrone, C. Ruis, L. du Plessis, O.G. Pybus, Addendum: A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology, Nature Microbiol. 6 (3) (2021) 415.

[41] R. Bouckaert, T.G. Vaughan, J. Barido-Sottani, S. Duchêne, M. Fourment, A. Gavryushkina, J. Heled, G. Jones, D. Kühnert, N.D. Maio, M. Matschiner, F. K. Mendes, N.F. Müller, H.A. Ogilvie, L. du Plessis, A. Popinga, A. Rambaut, D. Rasmussen, I. Siveroni, M.A. Suchard, C.-H. Wu, D. Xie, C. Zhang, T. Stadler, A. J. Drummond, BEAST, 2.5: An advanced software platform for bayesian evolutionary analysis, PLOS Computat. Biol. 15 (4) (2019), e1006650, https://doi.org/10.1371/journal.pcbi.1006650.

[42] J.P. Fine, R.J. Gray, A Proportional Hazards Model for the Subdistribution of a Competing Risk, J. Am. Stat. Assoc. 94 (446) (1999) 496–509, publisher: Taylor & Francis. doi:10.1080/01621459.1999.10474144. https://amstat.tandfonline.com/doi/abs/10.1080/01621459.1999.10474144.

[43] L. Scrucca, A. Santucci, F. Aversa, Regression modeling of competing risk using R: an in depth guide for clinicians, Bone Marrow Transplant. 45 (9) (2010) 1388–1395, https://doi.org/10.1038/bmt.2009.359, http://www.nature.com/articles/bmt2009359.

[44] World Health Organization, Estimating mortality from covid-19, Tech. rep., WHO (8 2020).https://www.who.int/news-room/commentaries/detail/estimating-mortality-from-covid-19.

[45] A. Oulhaj, L.A. Ahmed, J. Prattes, A. Suliman, A. Alsuwaidi, R.H. Al-Rifai, H. Sourij, I.V. Keilegom, The competing risk between in-hospital mortality and recovery: A pitfall in COVID-19 survival analysis research, medRxiv, Cold Spring Harbor Laboratory Press, 2020 doi:10.1101/2020.07.11.20151472, https://www.medrxiv.org/content/10.1101/2020.07.11.20151472v2.

[46] Z.R. McCaw, L. Tian, J.L. Vassy, C.S. Ritchie, C.-C. Lee, D.H. Kim, L.-J. Wei, How to Quantify and Interpret Treatment Effects in Comparative Clinical Studies of COVID-19, Ann. Internal Med. 173 (8) (2020) 632–637, publisher: American College of Physicians.https://www.acpjournals.org/doi/10.7326/M20-4044.

[47] X. Wang, Z. Ma, Y. Ning, C. Chen, R. Chen, Q. Chen, H. Zhang, C. Li, Y. He, T. Wang, C. Tong, J. Wu, Y. Li, H. Ma, S. Zhang, H. Zhao, Estimating the case fatality ratio of the COVID-19 epidemic in China, medRxiv (2020) 2020.02.17.20023630Publisher: Cold Spring Harbor Laboratory Press. doi: 10.1101/2020.02.17.20023630.https://www.medrxiv.org/content/10.1101/2020.02.17.20023630v5.

[48] A.C. Ghani, C.A. Donnelly, D.R. Cox, J.T. Griffin, C. Fraser, T.H. Lam, L.M. Ho, W.S. Chan, R.M. Anderson, A.J. Hedley, G.M. Leung, Methods for Estimating the Case Fatality Ratio for a Novel, Emerging Infectious Disease, Am. J. Epidemiol. 162 (5) (2005) 479–486, publisher: Oxford Academic. doi:10.1093/aje/kwi230.

[49] L. Scrucca, A. Santucci, F. Aversa, Competing risk analysis using R: an easy guide for clinicians, Bone Marrow Transplant 40 (4) (2007) 381–387.

[50] B. Efron, R. Tibshirani, Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy, Stat. Sci. 1 (1) (1986) 54–75, publisher: Institute of Mathematical Statistics. doi:10.1214/ss/1177013815. https://projecteuclid.org/euclid.ss/1177013815.

[51] J.W. Bartlett, S.R. Seaman, I.R. White, Multiple imputation of covariates by fully conditional specification: Accommodating the substantive model, Stat. Methods Med. Res. 24 (4) (2014) 462–487.

[52] J. Bartlett, R. Keogh, smcfcs: Multiple Imputation of Covariates by Substantive Model Compatible Fully Conditional Specification, r package version 1.4.1 (2020). https://CRAN.R-project.org/package=smcfcs.

[53] T. Lumley, mitools: Tools for Multiple Imputation of Missing Data, r package version 2.4 (2019).https://CRAN.R-project.org/package=mitools.

[54] B. Gray, cmprsk: Subdistribution Analysis of Competing Risks, r package version 2.2-10 (2020).https://CRAN.R-project.org/package=cmprsk.

[55] R.A. Ghayda, K.H. Lee, Y.J. Han, S. Ryu, S.H. Hong, S. Yoon, G.H. Jeong, J. Lee, J. Y. Lee, J.W. Yang, M. Effenberger, M. Eisenhut, A. Kronbichler, M. Solmi, H. Li, L. Jacob, A. Koyanagi, J. Radua, J.I. Shin, L. Smith, Estimation of global case fatality rate of coronavirus disease(COVID-19) using meta-analyses: Comparison between calendar date and days since the outbreak of the first confirmed case, Int. J. Infectious Diseases 100 (2020) (2019) 302–308, https://doi.org/10.1016/j.ijid.2020.08.065.

[56] J. Oke, D. Howdon, C. Heneghan, Declining covid-19 case fatality rates across all ages: analysis of german data, Tech. rep., The Centre for Evidence-Based Medicine (9 2020).https://www.who.int/news-room/commentaries/detail/estimating-mortality-from-covid-19.

[57] C. Signorelli, A. Odone, Age-specific COVID-19 case-fatality rate: no evidence of changes over time, Int. J. Public Health 65 (8) (2020) 1435–1436.

[58] A. Spena, L. Palombi, M. Corcione, A. Quintino, M. Carestia, V.A. Spena, Predicting SARS-CoV-2 Weather-Induced Seasonal Virulence from Atmospheric Air Enthalpy, Int. J. Environ Res. Public Health 17 (23).

[59] N. Shafran, I. Shafran, H. Ben-Zvi, S. Sofer, L. Sheena, I. Krause, A. Shlomai, E. Goldberg, E.H. Sklan, Secondary bacterial infection in COVID-19 patients is a stronger predictor for death compared to influenza patients, Sci. Rep. 11 (1).