BMJ Open

Covidogram as a simple tool for predicting severe course of COVID-19: population based study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2020-045442
Article Type:	Original research
Date Submitted by the Author:	01-Oct-2020
Complete List of Authors:	Jarkovský, Jiří; Masaryk University Institute of Biostatistics and Analyses; Institute of Health Information and Statistics of the Czech Republic Benešová, Klára; Institute of Health Information and Statistics of the Czech Republic; Institute of Health Information and Statistics of the Czech Republic Cerny, Vladimir; Masaryk Hospital in Usti nad Labem, Anesth. and ICU; Dalhousie University, Cerny Vladimir Razova, Jarmila; Ministry of Health of the Czech Republic Kala, petr; Masaryk University; University Hospital Brno Dolina, Jiri; University Hospital Brno; Masaryk University Majek, Ondrej; Institute of Health Information and Statistics of the Czech Republic; Masaryk University Institute of Biostatistics and Analyses Sebestova, Silvie; Institute of Health Information and Statistics of the Czech Republic Bezdekova, Monika; Institute of Health Information and Statistics of the Czech Republic Melicharova, Hana; Institute of Health Information and Statistics of the Czech Republic Snajdrova, Lenka; Institute of Health Information and Statistics of the Czech Republic; Masaryk University Institute of Biostatistics and Analyses Dusek, Ladislav; Institute of Health Information and Statistics of the Czech Republic; Masaryk University Institute of Biostatistics and Analyses Parenica, Jiri; University Hospital Brno, Internal and Cardiology Department; Masaryk University
Keywords:	COVID-19, Gastroduodenal disease < GASTROENTEROLOGY, Organisation of health services < HEALTH SERVICES ADMINISTRATION & MANAGEMENT

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Covidogram as a simple tool for predicting severe course of COVID-19: population based study

Jiri Jarkovsky[1,2], Klara Benesova[1,2], Vladimir Cerny[3,4], Jarmila Razova[5], Petr Kala[6,7], Jiri Dolina[7,8], Ondrej Majek[1,2], Silvie Sebestova[1], Monika Bezdekova[1], Hana Melicharova[1], Lenka Snajdrova[1,2], Ladislav Dusek[1,2] and Jiri Parenica[1,6,7]

[1] Institute of Health Information and Statistics of the Czech Republic, Prague, Czech Republic

[2] Institute of Biostatistics and Analysis, Faculty of Medicine, Masaryk University, Brno, Czech Republic

[3] Department of Anaesthesiology, Perioperative Medicine and Intensive Care, J.E. Purkinje University, Masaryk Hospital, Usti nad Labem, Czech Republic

[4]Department of Anesthesia, Pain Management and Perioperative Medicine, Dalhousie University, Halifax, Canada

[5] Ministry of Health, Czech Republic

[6] Internal and Cardiology Department, University Hospital Brno, Czech Republic

[7] Faculty of Medicine, Masaryk University, Brno, Czech Republic

[8] Gastroenterology and Internal Department, University Hospital Brno, Czech Republic

Key words: Covid-19; severe course; prognostic score; proton-pump inhibitors

A funding statement This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors

A competing interests statement All authors have disclosed that they do not have any conflicts of interest (no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work; no other relationships or activities that could appear to have influenced the submitted work)

Word count 2211 (without abstract, tables, references)

Contributors JJ, LD, VC and JR designed the study and wrote the research plan. OM, SS, MB, HMextracted the data used for the study from the databases. JJ and KB undertook statistical analysis with feedback from LD. JP, JJ, PK, JD interpreted the results and wrote the first draft of the manuscript with critical comments and revision from VC, LD and LS.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication The analysis was done on data from the National Health Information System (NHIS), which was supplemented with data from the Information System of Infectious Diseases (ISID). Data in the ISID are collected in compliance with Act No. 258/2000 Coll. on Protection of Public Health, data in

the NHIS are collected – and interconnected with data from ISID – in accordance with Act No. 372/2011 Coll., on Health Services and Conditions of Their Provision. Due to this legal mandate, the retrospective analyses did not require either approval by an ethics committee or informed consents from participants.

Data availability statement The anonymized data available upon reasonable request. The data are deidentified participant data, and available from the first author JJ (jarkovsky@uzis.cz). The reuse of the data subset is permitted only for revalidation of the results.

Corresponding author:

Jiri Parenica, MD, MSc, PhD, Professor,

Internal and Cardiology Department, University Hospital Brno,

Jihlavska 20, 625 00 Brno, Czech Republic

e-mail: jiri.parenica@atlas.cz

Telephone: +420 53223 2654; Fax: +420 532232907

Abstract

Objectives COVID-19 might either be entirely asymptomatic or manifest itself with a large variability of disease severity. It is beneficial to identify early patients with a high risk of severe course. The aim of the analysis was to determine factors for the development of the severe acute respiratory infection.

Design A population based study.

Setting Czech Republic.

Participants The first 7,455 consecutive COVID-19 patients who were identified by PCR testing from 1 March 2020 to 17 May 2020.

Interventions None

Primary outcome Severe course of Covid-19.

Result Of a total 6.2% of patients developed a severe course of COVID-19. Age, male sex, chronic kidney disease, chronic obstructive pulmonary disease, recent history of cancer, chronic heart failure, acid-related disorders and diabetes mellitus were found to be independent negative prognostic factors. (AUC was 0.893). The results were visualised by risk heat maps and we called this diagram a “covidogram”. Acid-related disorders, predominantly treated with proton-pump inhibitors, might represent a negative prognostic factor.

Conclusion We developed a very simple prediction model called “covidogram”, which is based on elementary independent variables (age, male sex and the presence of several chronic diseases) and represents a tool that makes it possible to identify – with a high reliability – patients who are at risk of a severe course of COVID-19. Obtained results open clinically relevant question about the role of acid-related disorders treated by proton pump inhibitors as predictor for severe course of COVID-19.

Strengths and limitations of the study

- Age, male sex, chronic kidney disease, chronic obstructive pulmonary disease, recent history of cancer, chronic heart failure, acid-related disorders and diabetes mellitus represent independent negative prognostic factors of severe course of COVID-19 in consecutive patients.
- The proposed prediction model “covidogram” is a simple tool that makes it possible to identify – with a high reliability (AUC 0.893) – patients who are at risk of a severe course of COVID-19.
- Acid-related disorders treated with proton-pump inhibitors might be predictor of a severe course of the disease.
- Due to the retrospective nature of this study, which is based on data of administrative registries, results of laboratory, clinical and X-ray examinations were not available. Conclusions regarding the influence of comorbidities and the consumption of medicinal products should be interpreted with caution and will require further validation.

Introduction

The coronavirus disease 2019 (COVID-19) is caused by betacoronavirus SARS-CoV-2, which enters human cells via the membrane-bound angiotensin-converting enzyme 2 (ACE2)[1]. The presence of infection might be entirely asymptomatic[2], or manifest itself with a large variability of disease severity, a number of unspecific clinical symptoms (fever, fatigue, myalgia) and various degrees of organ dysfunction. Most frequently, the disease affects the respiratory system (manifested as dry cough, dyspnoea, haemoptysis, pneumonia or ARDS) and the cardiovascular system (presented as myocardial injury or myocarditis, ventricular arrhythmias, haemodynamic instability or deep vein thrombosis), while other organ systems (such as the central nervous, gastrointestinal system or kidneys) are affected less frequently. In a number of patients, there is a risk of multiple organ failure, ultimately leading to death[3–7]. The management of COVID-19 patients depends on the disease severity: patients with a mild course of the disease can be treated in their home environment[8]. However, the clinical course might progress over time[5], and it is therefore clinically relevant to identify patients with a high risk of severe course of the disease as early as possible [9]. It was repeatedly demonstrated that older age, chronic obstructive pulmonary disease, diabetes mellitus,

chronic renal disease and cardiovascular disease are related to a high-risk course of the disease[5,10,11].

In the Czech Republic, a prospective population-based and centralised collection of data on COVID-19 patients was developed at the beginning of the pandemic, with a possibility to interconnect these data with those recorded in other population-based registries of the National Health Information System (NHIS), and thus to obtain information on each patient's history and management.

The aim of this study was to determine factors for the development of the severe acute respiratory infection, defined by the necessity of intensive care being provided in the Intensive care units (ICU), that include mechanical ventilation, extracorporal membrane oxygenation (ECMO) support and/or death.

Methods

Population of patients

The analysis is based on data from a population-based registry containing records of all consecutive COVID-19 patients in the Czech Republic who were identified by PCR testing and validated by the National Institute of Public Health.

The monitored cohort consisted of patients who were recorded in the National Information System of Infectious Diseases (ISID) between 1 March 2020 and 17 May 2020.

As of 17 May 2020, a total of 356,515 tests were performed in the Czech Republic, and the diagnosis of COVID-19 was confirmed in 8,475 cases. Of the confirmed COVID-19 cases, 464 patients with still unknown history in the National Health Information System (NHIS) or the patients being foreign nationals with unknown medical history – were excluded from the analysis, and further 556 patients with a follow-up period shorter than 14 days. were excluded as well. 90% of events occur within 14 days (Figure S1). Analysis without censoring and with a fixed follow-up length was chosen with the objective to simplify visualisation and interpretation of results of the analysis for its practical application. On top of that, characteristics of the cohort diagnosed with COVID-19 were compared to those of the population of the Czech Republic (10.6 million).

Diagnosis

The diagnosis of COVID-19 was based on the detection of SARS-CoV-2 by real-time reverse transcription polymerase chain reaction (RT-PCR). At first, the analyses were performed in the National Reference Laboratory of the National

Institute of Public Health (NIPH); other certified laboratories were later appointed to carry out PCR testing as well.

Systematic collection of data

The analysis was done on data from the National Health Information System (NHIS), which was supplemented with data from the Information System of Infectious Diseases (ISID). Data in the ISID are collected in compliance with Act No. 258/2000 Coll. on Protection of Public Health, whereas data in the NHIS are collected – and interconnected with data from ISID – in accordance with Act No. 372/2011 Coll., on Health Services and Conditions of Their Provision. Due to this legal mandate, the retrospective analyses did not require either approval by an ethics committee or informed consents from participants; moreover, it is a population-based analysis of all COVID-19 cases in the Czech Republic.

The latest data on COVID-19 patients, the severity of their condition as well as the necessity of hospitalisation in an ICU, including the use of mechanical ventilation or ECMO, together with information on death, have been entered into the ISID in real time. Apart from that, data on COVID-19 patients have been enriched with information on their comorbidities: this information is available in the National Register of Reimbursed Health Services (NRHZS), which contains data on all healthcare reported within the public health

insurance system (accounting for almost 100% of healthcare provided in the Czech Republic). Comorbidities are determined from combinations of reported diagnoses (ICD-10), ATC codes of drugs, and procedures defined in code lists used by Czech health insurance companies. Only diseases and conditions with a higher prevalence in the population or those identified in literature[5,7,12,13] as potential predictors of a severe course of COVID-19 were evaluated, with the aim to assess their potential influence on the resulting model.

Statistical analysis

Standard descriptive statistics were used to describe the data: categorical variables were described by absolute and relative frequencies, whereas continuous variables were described by means and standard deviations. The Fisher's exact test (for categorical variables) and Mann–Whitney U test (for continuous variables) were used to compare characteristics between groups depending on the monitored endpoint, unless stated otherwise. The predictive power of patient characteristics with regard to the analysed endpoint was evaluated by univariate and multivariate logistic regression and described by odds ratios, their 95% confidence intervals and statistical significance; a backward stepwise algorithm was used to choose the optimum model, and a ROC analysis was employed to evaluate the overall predictive power of the model. The results of the model were expressed by a

nomogram in the form of a risk heat map taking account of the patients' age, sex and comorbidities. A 10-fold cross-validation was performed to reduce the likelihood of model overfitting. The analysis was computed using the Vertica database and a MS SQL Server for data pre-processing and SPSS 25.0.0.1. for the statistical analysis of data. The level of statistical significance was set at $\alpha=0.05$ for all analyses.

Results

Of the total of 7,455 evaluated patients, 1,182 (15.9%) were hospitalised, 465 of them (6.2%) developed a severe course of the disease (i.e. reached the primary endpoint), 174 patients (2.3%) required mechanical ventilation, 11 patients (0.1%) were provided ECMO, and 287 patients (3.8%) died. Patients with the monitored endpoint were older (74.8 ± 13.4 vs 45.4 ± 20.2 years), more frequently of male sex and suffered at least one of all monitored comorbidities (Table 1, $p < 0.001$ for all parameters). Older age was determined by the multivariate logistic regression analysis to be the most significant predictor: the risk of a severe course of the disease increases progressively from the age of 40 years onwards (Table 2). Male sex, chronic kidney disease, chronic obstructive pulmonary disease, recent history of cancer (in the last five years), chronic heart failure, acid-related disorders and diabetes

mellitus were other significant predictors; the latter six conditions are hereinafter referred to as ***prognostically significant comorbidities*** (Table 2). The overall predictive power of the model, evaluated by the C-statistic and expressed by the AUC, was 0.893 (95% CI: 0.880–0.907; sensitivity 85.8% and specificity 80.3%). After performing the 10-fold cross-validation to validate the results, the average AUC of 0.891 (in the range 0.856–0.943) was obtained. For the purpose of an easier interpretation in clinical practice, a simplified version of the model was developed, taking into consideration the number of prognostically significant comorbidities obtained from the previous model (Table S1). The results were visualised by risk heat maps for men and women separately (Figure 1), and we called this diagram a "***covidogram***". The diagram shows how the risk increases progressively with age and with the number of prognostically significant comorbidities.

It is obvious from the comparison of basic patient characteristics (Table 1) and the results of the multivariate analysis (Table 2) that although a number of conditions occur more frequently in the group of patients with a severe course of the disease, not all of them are independent predictors (coronary artery disease, history of stroke, atrial fibrillation, hypertension or treatment by ACEIs and ARBs).

The comparison of characteristics of patients with confirmed COVID-19 to those of the entire population of the Czech Republic showed that COVID-19 patients are slightly older and have the monitored comorbidities slightly more frequently. (Table S2).

Discussion

New findings about COVID-19

We have described prognostically significant factors that increase the risk of a severe course of COVID-19 in consecutive patients with positive COVID-19 PCR test. Age is the most significant factor, and the risk increases progressively from the age of 40 years onwards. To our knowledge, this is the first study to suggest that acid-related disorders treated by proton pump inhibitors might be independent risk predictors as well. By contrast, not all cardiovascular diseases (such as uncomplicated hypertension) increase the risk of a severe course of COVID-19. We have developed a simple tool called “covidogram” for an early identification of risk of a severe course of the disease. This model has a very good predictive power (AUC 0.893).

The “covidogram” was designed as a model to assess the risk of unfavourable development of the patient’s condition based on his/her history of chronic disease, and can serve as a tool to estimate the number of severe

cases of COVID-19 in a population. When assessing the risk for an individual patient in clinical practice, it is certainly necessary to take into consideration also other pieces of information on the current condition of that patient (respiratory rate, peripheral oxygen saturation, level of consciousness, urea level, C reactive protein[14], procalcitonin, aspartate aminotransferase[13], high temperature[12], elevation of cardiac markers, lung infiltrates >50%[7]) as well as obesity, which can also increase the risk of a severe course of COVID-19 [15].

Acid-related disorders

Surprisingly, our analysis revealed that the presence of acid-related disorders might be an independent predictor of a severe course of COVID-19. Patients were predominantly treated with proton-pump inhibitors (1,175 patients in total, out of which 706 were treated with omeprazole and 402 with pantoprazole as the two most frequently used drugs); only a small proportion of them were treated with H_2-receptor antagonists (30 patients). The main indications for treatment with these drugs generally involve gastro-oesophageal reflux disease, functional dyspepsia, gastric and duodenal ulcers, gastric acid hypersecretory states as well as gastroprotection in patients using non-steroidal anti-inflammatory drugs or dual antiplatelet therapy. The effect of inhibition of hydrochloric acid secretion is followed by an increase in the

intragastric pH (to a value above 2–4), which might hypothetically decrease the physiological bactericidal/virucidal effect of gastric acid and decrease the activity of lysosomal enzymes. Published data showed that long-term use of proton-pump inhibitors could slightly increase the risk of pneumonia[16,17] and enteric infections[18].

Our comparison of patients with and without acid-related disorders (Table S3) showed that patients with these disorders are markedly older and have prognostically significant comorbidities more frequently. Our analysis cannot determine whether there is any causal relationship between the presence of acid-related disorders and a severe course of COVID-19 or whether it is just a coincidence. At the same time, it must be stressed out that the vast majority of patients were treated with proton-pump inhibitors, not with H_2-receptor agonists. Recently, Almario et al. demonstrated association between using proton pump inhibitors and odds of a positive COVID-19 test [19]. Similar trend was reported by Tarlow [20].

Strengths and limitations

This study is based on a fully integrated national health information system covering the entire population of a country – which proposed a prediction model estimating individually-based risk of a severe course of COVID-19. Because this model uses data readily available in health and

administrative registries, it can be easily used for the prediction of intensive care use in the context of decision-making at the national level.

On the other hand, our analysis has a number of limitations. Results of laboratory, clinical and X-ray examinations performed at the time of patient admission to hospitals were not available, and these very important pieces of information could therefore not be analysed; instead, our analysis is based on administrative data, with the exception of endpoints. Furthermore, analytical processing of a cohort of patients cannot capture the risk of less frequent conditions that might increase the risk of a severe course of COVID-19 (e.g. patients with immunodeficiencies, those after organ transplantation, or those undergoing immunosuppressive therapy or biological therapy). Due to the retrospective nature of this study, which is based on data of administrative registries and is focused on the development of a prediction model, any conclusions regarding the influence of comorbidities and the consumption of medicinal products should be interpreted with caution and will require further validation.

Conclusion

The proposed prediction model “covidogram” is based on elementary independent variables (age, male sex and the presence of chronic disease) and

represents a simple tool that makes it possible to identify – with a high reliability (AUC 0.893) – patients who are at risk of a severe course of COVID-19.

Finally, the analysis has shown, for the first time, that acid-related disorders treated with proton-pump inhibitors might also be predictor of a severe course of the disease.

References

1 Hoffmann M, Kleine-Weber H, Schroeder S, *et al.* SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* 2020;**181**:271-280.e8. doi:10.1016/j.cell.2020.02.052

2 Sutton D, Fuchs K, D'Alton M, *et al.* Universal Screening for SARS-CoV-2 in Women Admitted for Delivery. *N Engl J Med* Published Online First: 13 April 2020. doi:10.1056/NEJMc2009316

3 Giacomelli A, Pezzati L, Conti F, *et al.* Self-reported olfactory and taste disorders in SARS-CoV-2 patients: a cross-sectional study. *Clin Infect Dis Off Publ Infect Dis Soc Am* Published Online First: 26 March 2020. doi:10.1093/cid/ciaa330

4 Han C, Duan C, Zhang S, *et al.* Digestive Symptoms in COVID-19 Patients With Mild Disease Severity: Clinical Presentation, Stool Viral RNA Testing, and Outcomes. *Am J Gastroenterol* Published Online First: 15 April 2020. doi:10.14309/ajg.0000000000000664

5 Huang C, Wang Y, Li X, *et al.* Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet Lond Engl* 2020;**395**:497–506. doi:10.1016/S0140-6736(20)30183-5

6 Kim I-C, Kim JY, Kim HA, *et al.* COVID-19-related myocarditis in a 21-year-old female patient. *Eur Heart J* Published Online First: 13 April 2020. doi:10.1093/eurheartj/ehaa288

7 Wang D, Hu B, Hu C, *et al.* Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus–Infected Pneumonia in Wuhan, China. *JAMA* 2020;**323**:1061–9. doi:10.1001/jama.2020.1585

8 Gandhi RT, Lynch JB, Rio C del. Mild or Moderate Covid-19. *N Engl J Med* Published Online First: 24 April 2020. doi:10.1056/NEJMcp2009249

9 Berlin DA, Gulick RM, Martinez FJ. Severe Covid-19. *N Engl J Med* 2020;**0**:null. doi:10.1056/NEJMcp2009575

10 Mehra MR, Desai SS, Kuy S, *et al.* Cardiovascular Disease, Drug Therapy, and Mortality in Covid-19. *N Engl J Med* Published Online First: 1 May 2020. doi:10.1056/NEJMoa2007621

11 Wang T, Tang C, Chen R, *et al.* Clinical Features of Coronavirus Disease 2019 Patients With Mechanical Ventilation: A Nationwide Study in China. *Crit Care Med* 2020;**Online First**. doi:10.1097/CCM.0000000000004473

12 Wu C, Chen X, Cai Y, *et al.* Risk Factors Associated With Acute Respiratory Distress Syndrome and Death in Patients With Coronavirus Disease 2019 Pneumonia in Wuhan, China. *JAMA Intern Med* Published Online First: 13 March 2020. doi:10.1001/jamainternmed.2020.0994

13 Chen R, Liang W, Jiang M, *et al.* Risk Factors of Fatal Outcome in Hospitalized Subjects With Coronavirus Disease 2019 From a Nationwide Analysis in China. *Chest* Published Online First: 15 April 2020. doi:10.1016/j.chest.2020.04.010

14 Knight SR, Ho A, Pius R, *et al.* Risk stratification of patients admitted to hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: development and validation of the 4C Mortality Score. *BMJ* 2020;**370**. doi:10.1136/bmj.m3339

15 Malik P, Patel U, Patel K, *et al.* Obesity a predictor of outcomes of COVID-19 hospitalized patients- A systematic Review and Meta-Analysis. *J Med Virol* Published Online First: 25 September 2020. doi:10.1002/jmv.26555

16 Eom C-S, Jeon CY, Lim J-W, *et al.* Use of acid-suppressive drugs and risk of pneumonia: a systematic review and meta-analysis. *CMAJ Can Med Assoc J J Assoc Medicale Can* 2011;**183**:310–9. doi:10.1503/cmaj.092129

17 Zirk-Sadowski J, Masoli JA, Delgado J, *et al.* Proton-Pump Inhibitors and Long-Term Risk of Community-Acquired Pneumonia in Older Adults. *J Am Geriatr Soc* 2018;**66**:1332–8. doi:10.1111/jgs.15385

18 Moayyedi P, Eikelboom JW, Bosch J, *et al.* Safety of Proton Pump Inhibitors Based on a Large, Multi-Year, Randomized Trial of Patients Receiving Rivaroxaban or Aspirin. *Gastroenterology* 2019;**157**:682-691.e2. doi:10.1053/j.gastro.2019.05.056

19 Almario CV, Chey WD, Spiegel BMR. Increased Risk of COVID-19 Among Users of Proton Pump Inhibitors. *Am J Gastroenterol* Published Online First: 25 August 2020. doi:10.14309/ajg.0000000000000798

20 Tarlow B, Gubatan J, Khan MA, *et al.* Are Proton Pump Inhibitors Contributing to SARS-COV-2 Infection? *Off J Am Coll Gastroenterol ACG* 2020;**Publish Ahead of Print**. doi:10.14309/ajg.0000000000000933

Figure legend

Figure 1. COVIDOGRAM – predicted probability of severe course of Covid-19 for men and women – visualization of simplified multivariate logistic regression model (for more details, see Supplemental Table 1)

Table 1. Characteristics of COVID-19 patients according to endpoint

	With severe course N = 465	Without severe course N = 6,990
Basic characteristics		
Men	260 (55.9%)	3,221 (46.1%)
Age, mean ± SD	74.8 ± 13.4	45.4 ± 20.2
History		
Hypertension	258 (55.5%)	1,266 (18.1%)
Atrial fibrillation	75 (16.1%)	179 (2.6%)
History of stroke	55 (11.8%)	131 (1.9%)
History of MI or PCI	32 (6.9%)	106 (1.5%)
Chronic heart failure	90 (19.4%)	157 (2.2%)
Chronic kidney disease	99 (21.3%)	215 (3.1%)
Diabetes mellitus	127 (27.3%)	519 (7.4%)
Chronic obstructive pulmonary disease	102 (21.9%)	643 (9.2%)
Acid-related disorders	206 (44.3%)	999 (14.3%)
Recent history of cancer (≤5 years)	48 (10.3%)	169 (2.4%)
Rheumatoid arthritis	14 (3.0%)	75 (1.1%)
Treatment		
ACE inhibitors	182 (39.1%)	947 (13.5%)
ARBs	85 (18.3%)	543 (7.8%)
Calcium channel blockers	121 (26.0%)	584 (8.4%)
Beta-blockers	212 (45.6%)	884 (12.6%)
Diuretics	214 (46.0%)	814 (11.6%)
Anticoagulants / antithrombotic agents	143 (30.8%)	502 (7.2%)
Statins	169 (36.3%)	882 (12.6%)

ACE = angiotensin-converting enzyme, ARBs = angiotensin II receptor blockers, MI = myocardial infarction, PCI = percutaneous coronary intervention

Table 2. Multivariate logistic regression model using backward stepwise algorithm for selection of independent predictors of severe course of COVID-19. Age, sex and comorbidities from Table 1 were entered into the model.

Predictors		**OR (95% CI)**	**P-value**
Sex	Women	Reference category	
	Men	2.10 (1.68; 2.62)	**< 0.001**
Age	< 40 years	Reference category	
	40–49 years	3.41 (1.37; 8.48)	**0.008**
	50–59 years	11.92 (5.30; 26.81)	**< 0.001**
	60–69 years	30.68 (14.04; 67.04)	**< 0.001**
	70–79 years	60.89 (27.93; 132.73)	**< 0.001**
	80–89 years	112.68 (51.48; 246.63)	**< 0.001**
	90+ years	200.12 (86.50; 462.97)	**< 0.001**
Comorbidities	Chronic kidney disease	1.97 (1.45; 2.68)	**< 0.001**
	Chronic obstructive pulmonary disease	1.55 (1.17; 2.05)	**0.002**
	Recent history of cancer (≤5 years)	1.54 (1.05; 2.25)	**0.026**
	Chronic heart failure	1.50 (1.09; 2.08)	**0.014**
	Acid-related disorders	1.47 (1.16; 1.85)	**0.001**
	Diabetes mellitus	1.38 (1.07; 1.78)	**0.012**

Overall predictive power: AUC (95% CI): 0.893 (0.880; 0.907); sensitivity 85.8% and specificity 80.3%.

Figure 1.

Men – predicted probability (95% CI)

Age	Morbidity level 0	1	2	3	4+
< 40	**0.3** (0.2; 0.7)	**0.5** (0.2; 1.1)	**0.8** (0.4; 1.7)	-	-
40–49	**1.1** (0.7; 1.9)	**1.7** (1.0; 2.9)	**2.7** (1.5; 4.5)	**4.1** (2.3; 7.2)	-
50–59	**3.7** (2.7; 5.2)	**5.7** (4.1; 7.8)	**8.7** (6.2; 12.0)	**12.9** (9.0; 18.3)	-
60–69	**8.9** (7.0; 11.3)	**13.3** (10.8; 16.3)	**19.4** (15.7; 23.7)	**27.3** (21.7; 33.8)	**37.0** (28.8; 46.1)
70–79	**16.4** (13.1; 20.3)	**23.5** (19.7; 27.8)	**32.5** (27.8; 37.6)	**43.0** (36.7; 49.5)	**54.1** (46.0; 62.0)
80–89	**26.7** (21.6; 32.5)	**36.3** (30.8; 42.2)	**47.2** (41.0; 53.4)	**58.3** (51.1; 65.1)	**68.6** (60.7; 75.6)
90+	**39.9** (30.5; 50.2)	**51.0** (41.1; 60.8)	**62.0** (52.1; 70.9)	**71.8** (62.3; 79.7)	**80.0** (71.2; 86.5)

Women – predicted probability (95% CI)

Age	Morbidity level 0	1	2	3	4+
< 40	**0.2** (0.1; 0.3)	**0.3** (0.1; 0.5)	**0.4** (0.2; 0.8)	-	-
40–49	**0.5** (0.3; 0.9)	**0.8** (0.5; 1.4)	**1.3** (0.7; 2.2)	**2.0** (1.1; 3.6)	-
50–59	**1.8** (1.3; 2.6)	**2.8** (2.0; 3.9)	**4.3** (3.0; 6.2)	**6.6** (4.4; 9.7)	-
60–69	**4.5** (3.4; 5.9)	**6.8** (5.3; 8.7)	**10.2** (7.9; 13.2)	**15.2** (11.4; 19.9)	**21.8** (15.9; 29.2)
70–79	**8.5** (6.6; 11.0)	**12.7** (10.3; 15.7)	**18.6** (15.2; 22.5)	**26.4** (21.4; 32.0)	**35.9** (28.6; 43.9)
80–89	**14.8** (11.7; 18.5)	**21.3** (17.7; 25.5)	**29.8** (25.1; 34.9)	**39.9** (33.6; 46.5)	**50.9** (42.7; 59.1)
90+	**24.0** (17.6; 31.8)	**33.1** (25.5; 41.7)	**43.6** (34.8; 52.9)	**54.8** (44.8; 64.4)	**65.5** (54.8; 74.8)

0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%

Table S1. Simplified multivariate logistic regression model based on predictors selected by a backward stepwise algorithm

Predictors		**OR (95% CI)**	**P-value**
Sex	Women	Reference category	
	Men	2.10 (1.69; 2.62)	**< 0.001**
Age	< 40 years	Reference category	
	40–49 years	3.37 (1.36; 8.37)	**0.009**
	50–59 years	11.69 (5.20; 26.30)	**< 0.001**
	60–69 years	29.59 (13.56; 64.59)	**< 0.001**
	70–79 years	59.27 (27.22; 129.05)	**< 0.001**
	80–89 years	109.93 (50.31; 240.20)	**< 0.001**
	90+ years	200.66 (86.84; 463.63)	**< 0.001**
Morbidity level*	1-point increase	1.56 (1.42; 1.72)	**< 0.001**

*One point for each comorbidity: chronic kidney disease, chronic obstructive pulmonary disease, recent history of cancer (≤5 years), chronic heart failure, acid-related disorders, diabetes mellitus (maximum of 4 points per person, additional points were not considered due to a low prevalence in our cohort).

Overall predictive power: AUC (95% CI): 0.893 (0.879; 0.907); sensitivity 85.8% and specificity 80.3%.

Table S2. Characteristics of COVID-19 patients in comparison to general population of the Czech Republic

	COVID-19 patients N = 7,455	General population N = 10.6 million	P-value
Basic characteristics			
Men	46.7%	49.1%	**< 0.001**
Age, mean ± SD	47.2 ± 21.1	42.5 ± 23.1	**< 0.001**
Medical history			
Hypertension	20.4%	17.5%	**< 0.001**
Atrial fibrillation	3.4%	2.3%	**< 0.001**
History of stroke	2.5%	1.4%	**< 0.001**
History of MI or PCI	1.9%	1.5%	**0.011**
Chronic heart failure	3.3%	1.6%	**< 0.001**
Chronic kidney disease	4.2%	2.4%	**< 0.001**
Diabetes mellitus	8.7%	7.2%	**< 0.001**
Chronic obstructive pulmonary disease	10.0%	8.4%	**< 0.001**
Acid-related disorders	16.2%	11.7%	**< 0.001**
Recent history of cancer (≤5 years)	2.9%	2.1%	**< 0.001**
Rheumatoid arthritis	1.2%	0.9%	**0.003**
Treatment			
ACE inhibitors	15.1%	14.5%	0.100
ARBs	8.4%	7.6%	**0.009**
Calcium channel blockers	9.5%	8.8%	0.066
Beta-blockers	14.7%	12.0%	**< 0.001**
Diuretics	13.8%	11.6%	**< 0.001**
Anticoagulants / antithrombotic agents	8.7%	6.1%	**< 0.001**
Statins	14.1%	13.9%	0.701

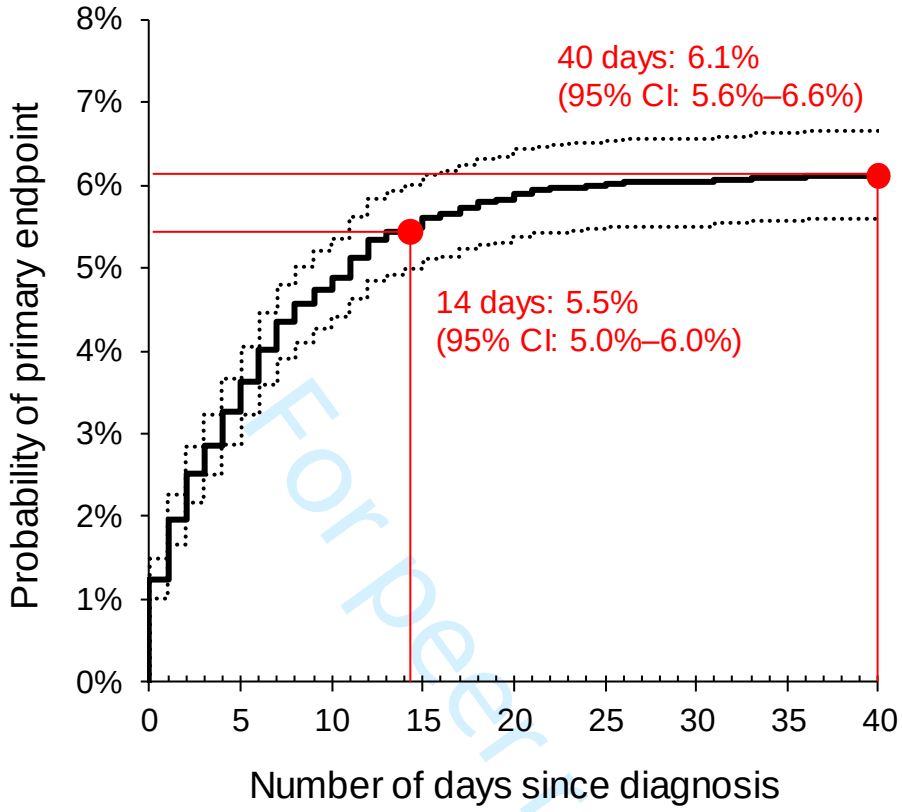
P-value of binomial test (categorical variables) and P-value of one-sample t-test (continuous variables).
ACE = angiotensin converting enzyme, ARB = angiotensin II receptor blockers, MI = myocardial infarction, PCI = percutaneous coronary intervention

Table S3. Characteristics of COVID-19 patients by using drugs for acid-related disorders

	Using drugs N = 1,205	Not using drugs N = 6,250	P-value
Basic characteristics			
Men	524 (43.5%)	2 957 (47.3%)	**0.015**
Age, mean ± SD	62.6 ± 18.3	44.2 ± 20.2	**< 0.001**
Medical history			
Hypertension	542 (45.0%)	982 (15.7%)	**< 0.001**
Atrial fibrillation	125 (10.4%)	129 (2.1%)	**< 0.001**
History of stroke	91 (7.6%)	95 (1.5%)	**< 0.001**
History of MI or PCI	69 (5.7%)	69 (1.1%)	**< 0.001**
Chronic heart failure	138 (11.5%)	109 (1.7%)	**< 0.001**
Chronic kidney disease	158 (13.1%)	156 (2.5%)	**< 0.001**
Diabetes mellitus	222 (18.4%)	424 (6.8%)	**< 0.001**
Chronic obstructive pulmonary disease	271 (22.5%)	474 (7.6%)	**< 0.001**
Recent history of cancer (≤5 years)	93 (7.7%)	124 (2.0%)	**< 0.001**
Rheumatoid arthritis	39 (3.2%)	50 (0.8%)	**< 0.001**
Treatment			
ACE inhibitors	348 (28.9%)	781 (12.5%)	**< 0.001**
ARBs	217 (18.0%)	411 (6.6%)	**< 0.001**
Calcium channel blockers	253 (21.0%)	452 (7.2%)	**< 0.001**
Beta-blockers	422 (35.0%)	674 (10.8%)	**< 0.001**
Diuretics	399 (33.1%)	629 (10.1%)	**< 0.001**
Anticoagulants / antithrombotic agents	270 (22.4%)	375 (6.0%)	**< 0.001**
Statins	369 (30.6%)	682 (10.9%)	**< 0.001**

P-value of Fisher's exact test (categorical variables) and P-value of Mann-Whitney U test (continuous variables).
ACE = angiotensin-converting enzyme, ARBs = angiotensin II receptor blockers, MI = myocardial infarction, PCI = percutaneous coronary intervention

Figure S1. Time to analysed endpoint (severe condition)



STROBE Statement—Checklist of items that should be included in reports of ***cohort studies***

	Item No	Recommendation	Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	4
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	6
Objectives	3	State specific objectives, including any prespecified hypotheses	7
Methods			
Study design	4	Present key elements of study design early in the paper	4
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	7,8
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	7
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	8,9
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	8,9
Bias	9	Describe any efforts to address potential sources of bias	9
Study size	10	Explain how the study size was arrived at	8,9
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	9,10
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	9,10
		(*b*) Describe any methods used to examine subgroups and interactions	
		(*c*) Explain how missing data were addressed	
		(*d*) If applicable, explain how loss to follow-up was addressed	
		(*e*) Describe any sensitivity analyses	
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	7,8,10
		(b) Give reasons for non-participation at each stage	NA
		(c) Consider use of a flow diagram	-
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	10, Table 1
		(b) Indicate number of participants with missing data for each variable of interest	8
		(c) Summarise follow-up time (eg, average and total amount)	7
Outcome data	15*	Report numbers of outcome events or summary measures over time	10

Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included (*b*) Report category boundaries when continuous variables were categorized (*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	10 Table 2
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	11, Tabel S1,2
Discussion			
Key results	18	Summarise key results with reference to study objectives	12
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	15,16
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	13,14
Generalisability	21	Discuss the generalisability (external validity) of the study results	16
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	17

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at http://www.strobe-statement.org.

BMJ Open

Covidogram as a simple tool for predicting severe course of COVID-19: population based study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2020-045442.R1
Article Type:	Original research
Date Submitted by the Author:	21-Dec-2020
Complete List of Authors:	Jarkovský, Jiří; Masaryk University Institute of Biostatistics and Analyses; Institute of Health Information and Statistics of the Czech Republic Benešová, Klára; Institute of Health Information and Statistics of the Czech Republic; Institute of Health Information and Statistics of the Czech Republic Cerny, Vladimir; Masaryk Hospital in Usti nad Labem, Anesth. and ICU; Dalhousie University, Cerny Vladimir Razova, Jarmila; Ministry of Health of the Czech Republic Kala, petr; Masaryk University; University Hospital Brno Dolina, Jiri; University Hospital Brno; Masaryk University Majek, Ondrej; Institute of Health Information and Statistics of the Czech Republic; Masaryk University Institute of Biostatistics and Analyses Sebestova, Silvie; Institute of Health Information and Statistics of the Czech Republic Bezdekova, Monika; Institute of Health Information and Statistics of the Czech Republic Melicharova, Hana; Institute of Health Information and Statistics of the Czech Republic Snajdrova, Lenka; Institute of Health Information and Statistics of the Czech Republic; Masaryk University Institute of Biostatistics and Analyses Dusek, Ladislav; Institute of Health Information and Statistics of the Czech Republic; Masaryk University Institute of Biostatistics and Analyses Parenica, Jiri; University Hospital Brno, Internal and Cardiology Department; Masaryk University
Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Epidemiology, Gastroenterology and hepatology
Keywords:	COVID-19, Gastroduodenal disease < GASTROENTEROLOGY, Organisation of health services < HEALTH SERVICES ADMINISTRATION & MANAGEMENT

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Covidogram as a simple tool for predicting severe course of COVID-19: population based study

Jiri Jarkovsky[1,2], Klara Benesova[1,2], Vladimir Cerny[3,4], Jarmila Razova[5], Petr Kala[6,7], Jiri Dolina[7,8], Ondrej Majek[1,2], Silvie Sebestova[1], Monika Bezdekova[1], Hana Melicharova[1], Lenka Snajdrova[1,2], Ladislav Dusek[1,2] and Jiri Parenica[1,6,7]

[1] Institute of Health Information and Statistics of the Czech Republic, Prague, Czech Republic

[2] Institute of Biostatistics and Analysis, Faculty of Medicine, Masaryk University, Brno, Czech Republic

[3] Department of Anaesthesiology, Perioperative Medicine and Intensive Care, J.E. Purkinje University, Masaryk Hospital, Usti nad Labem, Czech Republic

[4]Department of Anesthesia, Pain Management and Perioperative Medicine, Dalhousie University, Halifax, Canada

[5] Ministry of Health, Czech Republic

[6] Internal and Cardiology Department, University Hospital Brno, Czech Republic

[7] Faculty of Medicine, Masaryk University, Brno, Czech Republic

[8] Gastroenterology and Internal Department, University Hospital Brno, Czech Republic

Key words: Covid-19; severe course; prognostic score; proton-pump inhibitors

A funding statement This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors

A competing interests statement All authors have disclosed that they do not have any conflicts of interest (no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work; no other relationships or activities that could appear to have influenced the submitted work)

Word count 2522 (without abstract, tables, references)

Contributors JJ, LD, VC and JR designed the study and wrote the research plan. OM, SS, MB, HMextracted the data used for the study from the databases. JJ and KB undertook statistical analysis with feedback from LD. JP, JJ, PK, JD interpreted the results and wrote the first draft of the manuscript with critical comments and revision from VC, LD and LS.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication The analysis was done on data from the National Health Information System (NHIS), which was supplemented with data from the Information System of Infectious Diseases (ISID). Data in the ISID are collected in compliance with Act No. 258/2000 Coll. on Protection of Public Health, data in

the NHIS are collected – and interconnected with data from ISID – in accordance with Act No. 372/2011 Coll., on Health Services and Conditions of Their Provision. Due to this legal mandate, the retrospective analyses did not require either approval by an ethics committee or informed consents from participants.

Data availability statement The anonymized data available upon reasonable request. The data are deidentified participant data, and available from the first author JJ (jarkovsky@uzis.cz). The reuse of the data subset is permitted only for revalidation of the results.

Corresponding author:

Jiri Parenica, MD, MSc, PhD, Professor,

Internal and Cardiology Department, University Hospital Brno,

Jihlavska 20, 625 00 Brno, Czech Republic

e-mail: jiri.parenica@atlas.cz

Telephone: +420 53223 2654; Fax: +420 532232907

Abstract

Objectives COVID-19 might either be entirely asymptomatic or manifest itself with a large variability of disease severity. It is beneficial to identify early patients with a high risk of severe course. The aim of the analysis was to develop a prognostic model for the prediction of the severe course of acute respiratory infection.

Design A population based study.

Setting Czech Republic.

Participants The first 7,455 consecutive COVID-19 patients who were identified by RT-PCR testing from 1 March 2020 to 17 May 2020.

Interventions None

Primary outcome Severe course of Covid-19.

Result Of a total 6.2% of patients developed a severe course of COVID-19. Age, male sex, chronic kidney disease, chronic obstructive pulmonary disease, recent history of cancer, chronic heart failure, acid-related disorders treated with proton-pump inhibitors and diabetes mellitus were found to be independent negative prognostic factors. (AUC was 0.893). The results were visualised by risk heat maps and we called this diagram a “covidogram”. Acid-related disordersreatedwith proton-pump inhibitors might represent a negative prognostic factor.

Conclusion We developed a very simple prediction model called “covidogram”, which is based on elementary independent variables (age, male sex and the presence of several chronic diseases) and represents a tool that makes it possible to identify – with a high reliability – patients who are at risk of a severe course of COVID-19. Obtained results open clinically relevant question about the role of acid-related disorders treated by proton pump inhibitors as predictor for severe course of COVID-19.

Strengths and limitations of the study

- The majority of consecutive patients diagnosed with COVID-19 in the Czech Republic were included in the analysis, regardless of whether they were hospitalized or not.
- The cohort covers also asymptomatic and oligosymptomatic patients identified thanks to epidemiological monitoring.
- The cohort does not include strictly all COVID-19 cases in the Czech Republic because some patients are asymptomatic and have not been tested.
- The proposed prediction model is a simple tool that makes it possible to identify – with a high reliability (AUC 0.893) – patients who are at risk of a severe course of COVID-19.
- Due to the retrospective nature of this study, which is based on data of administrative registries, results of laboratory, clinical and X-ray examinations were not available. Conclusions regarding the influence of comorbidities and the consumption of medicinal products should be interpreted with caution and will require further validation.

Introduction

The coronavirus disease 2019 (COVID-19) is caused by betacoronavirus SARS-CoV-2, which enters human cells via the membrane-bound angiotensin-converting enzyme 2 (ACE2)[1]. The presence of infection might be entirely asymptomatic[2], or manifest itself with a large variability of disease severity, a number of unspecific clinical symptoms (fever, fatigue, myalgia) and various degrees of organ dysfunction. Most frequently, the disease affects the respiratory system (manifested as dry cough, dyspnoea, haemoptysis, pneumonia or ARDS) and the cardiovascular system (presented as myocardial injury or myocarditis, ventricular arrhythmias, haemodynamic instability or deep vein thrombosis), while other organ systems (such as the central nervous [3], gastrointestinal system orkidneys [4]) are affected less frequently. In a number of patients, there is a risk of multiple organ failure, ultimately leading to death[5–9]. According to the report of World Health Organization, as 12 November 2020, the rate of mortality among COVID-19 patients is 2.28%[10]. The management of COVID-19 patients depends on the disease severity: patients with a mild course of the disease can be treated in their home environment[11]. However, the clinical picture of COVID-19 patients can quickly turn into an unfavourable clinical course [7], and it is therefore clinically relevant to identify patients with a high risk of severe course of the disease as

early as possible [12]. It was repeatedly demonstrated that older age > 65, cancer, chronic obstructive pulmonary disease, moderate-to-severe asthma, diabetes mellitus, chronic renal disease, immunocompromised state, obesity (BMI>30), pregnancy, sickle cell disease, smoking and cardiovascular disease are related to a high-risk course of the disease[7,13–15].

In the Czech Republic, a prospective population-based and centralised collection of data on COVID-19 patients was developed at the beginning of the pandemic, with a possibility to interconnect these data with those recorded in other population-based registries of the National Health Information System (NHIS), and thus to obtain information on each patient's history and management.

The aim of this study was to develop a prognostic model for the prediction of the severe course of acute respiratory infection, defined by the necessity of intensive care being provided in the Intensive care units (ICU), that include mechanical ventilation, extracorporal membrane oxygenation (ECMO) support and/or death.

Methods

Population of patients

The analysis is based on data from a population-based registry containing records of all consecutive COVID-19 patients in the Czech Republic who were identified by RT-PCR testing and validated by the National Institute of Public Health.

The monitored cohort consisted of patients who were recorded in the National Information System of Infectious Diseases (ISID) between 1 March 2020 and 17 May 2020.

As of 17 May 2020, a total of 356,515 tests were performed in the Czech Republic, and the diagnosis of COVID-19 was confirmed in 8,475 cases. Of the confirmed COVID-19 cases, 464 patients with still unknown history in the National Health Information System (NHIS) or the patients being foreign nationals with unknown medical history – were excluded from the analysis, and further 556 patients with a follow-up period shorter than 14 days. were excluded as well. 90% of events occur within 14 days (Figure S1). Analysis without censoring and with a fixed follow-up length was chosen with the objective to simplify visualisation and interpretation of results of the analysis for its practical application. The basic characteristics of patients (age, gender) were provided for all patients, data on comorbidities were available for all patients with match between COVID19 and health insurance companies datasets; patients without match between datasets were excluded from the

analysis, no other missing data handling was necessary. On top of that, characteristics of the cohort diagnosed with COVID-19 were compared to those of the population of the Czech Republic (10.6 million).

Diagnosis

The diagnosis of COVID-19 was based on the detection of SARS-CoV-2 by real-time reverse transcription polymerase chain reaction (RT-PCR). At first, the analyses were performed in the National Reference Laboratory of the National Institute of Public Health (NIPH); other certified laboratories were later appointed to carry out RT-PCR testing as well.

Systematic collection of data

The analysis was done on data from the National Health Information System (NHIS), which was supplemented with data from the Information System of Infectious Diseases (ISID). Data in the ISID are collected in compliance with Act No. 258/2000 Coll. on Protection of Public Health, whereas data in the NHIS are collected – and interconnected with data from ISID – in accordance with Act No. 372/2011 Coll., on Health Services and Conditions of Their Provision. Due to this legal mandate, the retrospective analyses did not require either approval by an ethics committee or informed consents from participants; moreover, it is a population-based analysis of all diagnosed COVID-19 cases in the Czech Republic. The cohort covers also

asymptomatic and oligosymptomatic patients identified thanks to epidemiological monitoring.

The latest data on COVID-19 patients, the severity of their condition as well as the necessity of hospitalisation in an ICU, including the use of mechanical ventilation or ECMO, together with information on death, have been entered into the ISID in real time. Apart from that, data on COVID-19 patients have been enriched with information on their comorbidities: this information is available in the National Register of Reimbursed Health Services (NRHZS), which contains data on all healthcare reported within the public health insurance system (accounting for almost 100% of healthcare provided in the Czech Republic). Comorbidities are determined from combinations of reported diagnoses (ICD-10), ATC codes of drugs, and procedures defined in code lists used by Czech health insurance companies. Only diseases and conditions with a higher prevalence in the population or those identified in literature[7,9,15–17] as potential predictors of a severe course of COVID-19 were evaluated, with the aim to assess their potential influence on the resulting model.

Statistical analysis

Standard descriptive statistics were used to describe the data: categorical variables were described by absolute and relative frequencies, whereas continuous variables were described by means and standard

deviations. The Fisher's exact test (for categorical variables) and Mann–Whitney U test (for continuous variables) were used to compare characteristics between groups depending on the monitored endpoint, unless stated otherwise. The predictive power of patient characteristics with regard to the analysed endpoint was evaluated by univariate and multivariable logistic regression and described by odds ratios, their 95% confidence intervals and statistical significance; a backward stepwise algorithm was used to choose the optimum model, and a ROC analysis was employed to evaluate the overall predictive power of the model, Hosmer and Lemeshow test was adopted for the evaluation of goodness of fit of the model. The results of the model were expressed by a nomogram in the form of a risk heat map taking account of the patients' age, sex and comorbidities. A 10-fold cross-validation was performed to obtain estimates of model performance that are adjusted for in-sample optimism. A model was created in accordance with TRIPOD check list for prediction model development and validation [18]. The analysis was computed using the Vertica database and a MS SQL Server for data pre-processing and SPSS 25.0.0.1. for the statistical analysis of data. The level of statistical significance was set at $\alpha=0.05$ for all analyses.

Results

Of the total of 7,455 evaluated patients, 1,182 (15.9%) were hospitalised, 465 of them (6.2%) developed a severe course of the disease (i.e. reached the primary endpoint – the necessity of intensive care being provided in the Intensive care units (ICU), that include mechanical ventilation, ECMO support and/or death), 174 patients (2.3%) required mechanical ventilation, 11 patients (0.1%) were provided ECMO, and 287 patients (3.8%) died. Patients with the monitored endpoint were older (74.8 ± 13.4 vs 45.4 ± 20.2 years), more frequently of male sex and suffered at least one of all monitored comorbidities (Table 1, $p < 0.001$ for all parameters; univariable logistic regression results Table S1). Older age was determined by the multivariable logistic regression analysis to be the most significant predictor: the risk of a severe course of the disease increases progressively from the age of 40 years onwards (Table 2). Male sex, chronic kidney disease, chronic obstructive pulmonary disease, recent history of cancer (in the last five years), chronic heart failure, acid related disorders treated with proton-pump inhibitors and diabetes mellitus were other significant predictors; the latter six conditions are hereinafter referred to as ***prognostically significant comorbidities*** (Table 2). The overall predictive power of the model, evaluated by the C-statistic and expressed by the AUC, was 0.893 (95% CI: 0.880–0.907; sensitivity 85.8% and specificity

80.3%). After performing the 10-fold cross-validation to validate the results, the average AUC of 0.891 (in the range 0.856–0.943) was obtained. For the purpose of an easier interpretation in clinical practice, a simplified version of the model was developed, taking into consideration the number of prognostically significant comorbidities obtained from the previous model (Table S2). The results were visualised by risk heat maps for men and women separately (Figure 1), and we called this diagram a "***covidogram***". The diagram shows how the risk increases progressively with age and with the number of prognostically significant comorbidities.

It is obvious from the comparison of basic patient characteristics (Table 1) and the results of the multivariable analysis (Table 2) that although a number of conditions occur more frequently in the group of patients with a severe course of the disease, not all of them are independent predictors (coronary artery disease, history of stroke, atrial fibrillation, hypertension or treatment by ACEIs and ARBs).

The comparison of characteristics of patients with confirmed COVID-19 to those of the entire population of the Czech Republic showed that COVID-19 patients are slightly older and have the monitored comorbidities slightly more frequently. (Table S3).

Discussion

New findings about COVID-19

We have developed a prognostic model for the prediction of the severe course of COVID-19 in consecutive patients with positive COVID-19 RT-PCR test. This a simple tool called “covidogram” has a very good predictive power (AUC 0.893). Age is the most significant factor, and the risk increases progressively from the age of 40 years onwards. To our knowledge, this is the first study to suggest that acid related disorders treatedby proton pump inhibitors might be independent risk predictors as well. By contrast, not all cardiovascular diseases (such as uncomplicated hypertension or coronary artery disease) increase the risk of a severe course of COVID-19.

The “covidogram” was designed as a model to assess the risk of unfavourable development of the patient’s condition based on his/her history of chronic disease, and can serve as a tool to estimate the number of severe cases of COVID-19 in a population. When assessing the risk for an individual patient in clinical practice, it is certainly necessary to take into consideration also other pieces of information on the current condition of that patient (respiratory rate, peripheral oxygen saturation, level of consciousness, urea level, C reactive protein[19], procalcitonin, aspartate aminotransferase[17], high temperature[16], elevation of cardiac markers, lung infiltrates >50%[9]) as

well as obesity, which can also increase the risk of a severe course of COVID-19 [20].

Acid-related disorders

Surprisingly, our analysis revealed that the presence of acid-related disorders might be theoretically linked to a severe course of COVID-19. Patients were predominantly treated with proton-pump inhibitors (1,175 patients in total, out of which 706 were treated with omeprazole and 402 with pantoprazole as the two most frequently used drugs); only a small proportion of them were treated with H_2-receptor antagonists (30 patients). The main indications for treatment with these drugs generally involve gastro-oesophageal reflux disease, functional dyspepsia, gastric and duodenal ulcers, gastric acid hypersecretory states as well as gastroprotection in patients using non-steroidal anti-inflammatory drugs, dual antiplatelet therapy, biphophonates or same selective serotonin reuptake inhibitors (SSRIs). The effect of inhibition of hydrochloric acid secretion is followed by an increase in the intragastric pH (to a value above 2–4), which might hypothetically decrease the physiological bactericidal/virucidal effect of gastric acid and decrease the activity of lysosomal enzymes. Published data showed that long-term use of proton-pump inhibitors could slightly increase the risk of pneumonia[21,22] and enteric infections[23].

Our comparison of patients with and without acid-related disorders (Table S4) showed that patients with these disorders are markedly older and have prognostically significant comorbidities more frequently. Our analysis cannot determine whether there is any causal relationship between the presence of acid-related disorders and a severe course of COVID-19 or whether it is just a coincidence. At the same time, it must be stressed out that the vast majority of patients were treated with proton-pump inhibitors, not with H_2-receptor agonists. The observation is complicated also by the fact that some patients may not be adherent to their PPI regimen and there was great variability in the amount of time that they have been on PPIs (from one to twelfth month within 2019). Based on our analysis we are not able to decide whether severity of disease might be theoretically explained by pharmacology or by underlying pathology of acid related disorders. Recently, Almario et al. demonstrated association between using proton pump inhibitors and odds of a positive COVID-19 test [24]. Similar trend was reported by Tarlow [25].

Strengths and limitations

This study is based on a fully integrated national health information system covering the entire population of a country – which proposed a prediction model estimating individually-based risk of a severe course of COVID-19. Because this model uses data readily available in health and

administrative registries, it can be easily used for the prediction of intensive care use in the context of decision-making at the national level.

On the other hand, our analysis has a number of limitations. Results of laboratory, clinical and X-ray examinations performed at the time of patient admission to hospitals were not available, and these very important pieces of information could therefore not be analysed; instead, our analysis is based on administrative data, with the exception of endpoints. Furthermore, analytical processing of a cohort of patients cannot capture the risk of less frequent conditions that might increase the risk of a severe course of COVID-19 (e.g. patients with immunodeficiencies, those after organ transplantation, or those undergoing immunosuppressive therapy or biological therapy). The cohort does not include strictly all COVID-19 cases in the Czech Republic because some patients are asymptomatic and have not been tested. Older peoples with more comorbidities are probably more likely to have a symptomatic course of COVID-19. It could be also a reason why the population of patients diagnosed with COVID-19 is older and with more comorbidities in comparison with the Czech Republic population (Table S3). Due to the retrospective nature of this study, which is based on data of administrative registries and is focused on the development of a prediction model, any conclusions regarding the influence of

comorbidities and the consumption of medicinal products should be interpreted with caution and will require further validation.

Conclusion

The proposed prediction model “covidogram” is based on elementary independent variables (age, male sex and the presence of chronic disease) and represents a simple tool that makes it possible to identify – with a high reliability (AUC 0.893) – patients who are at risk of a severe course of COVID-19.

Finally, the analysis has shown, for the first time, that acid-related disorders treated with proton-pump inhibitors might also be theoretically associated with a severe course of the disease.

References

1 Hoffmann M, Kleine-Weber H, Schroeder S, *et al.* SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* 2020;**181**:271-280.e8. doi:10.1016/j.cell.2020.02.052

2 Sutton D, Fuchs K, D'Alton M, *et al.* Universal Screening for SARS-CoV-2 in Women Admitted for Delivery. *N Engl J Med* Published Online First: 13 April 2020. doi:10.1056/NEJMc2009316

3 Aghagoli G, Gallo Marin B, Katchur NJ, *et al.* Neurological Involvement in COVID-19 and Potential Mechanisms: A Review. *Neurocrit Care* 2020;:1–10. doi:10.1007/s12028-020-01049-4

4 Aghagoli G, Gallo Marin B, Soliman LB, *et al.* Cardiac involvement in COVID-19 patients: Risk factors, predictors, and complications: A review. *J Card Surg* Published Online First: 19 April 2020. doi:10.1111/jocs.14538

5 Giacomelli A, Pezzati L, Conti F, *et al.* Self-reported olfactory and taste disorders in SARS-CoV-2 patients: a cross-sectional study. *Clin Infect Dis Off Publ Infect Dis Soc Am* Published Online First: 26 March 2020. doi:10.1093/cid/ciaa330

6 Han C, Duan C, Zhang S, *et al.* Digestive Symptoms in COVID-19 Patients With Mild Disease Severity: Clinical Presentation, Stool Viral RNA Testing, and Outcomes. *Am J Gastroenterol* Published Online First: 15 April 2020. doi:10.14309/ajg.0000000000000664

7 Huang C, Wang Y, Li X, *et al.* Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet Lond Engl* 2020;**395**:497–506. doi:10.1016/S0140-6736(20)30183-5

8 Kim I-C, Kim JY, Kim HA, *et al.* COVID-19-related myocarditis in a 21-year-old female patient. *Eur Heart J* Published Online First: 13 April 2020. doi:10.1093/eurheartj/ehaa288

9 Wang D, Hu B, Hu C, *et al.* Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus–Infected Pneumonia in Wuhan, China. *JAMA* 2020;**323**:1061–9. doi:10.1001/jama.2020.1585

10 WHO. Coranvirus disease (COVID-19) pandemic. Available from https://www.who.int/emergencies/diseases/novel-coronavirus-2019. Accessed 12 Dec 2020.

11 Gandhi RT, Lynch JB, Rio C del. Mild or Moderate Covid-19. *N Engl J Med* Published Online First: 24 April 2020. doi:10.1056/NEJMcp2009249

12 Berlin DA, Gulick RM, Martinez FJ. Severe Covid-19. *N Engl J Med* 2020;**0**:null. doi:10.1056/NEJMcp2009575

13 Mehra MR, Desai SS, Kuy S, *et al.* Cardiovascular Disease, Drug Therapy, and Mortality in Covid-19. *N Engl J Med* Published Online First: 1 May 2020. doi:10.1056/NEJMoa2007621

14 Wang T, Tang C, Chen R, *et al.* Clinical Features of Coronavirus Disease 2019 Patients With Mechanical Ventilation: A Nationwide Study in China. *Crit Care Med* 2020;**Online First**. doi:10.1097/CCM.0000000000004473

15 CDC, Coronavirus-19. Afvailable from https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/people-with-medical-conditions.html.

16 Wu C, Chen X, Cai Y, *et al.* Risk Factors Associated With Acute Respiratory Distress Syndrome and Death in Patients With Coronavirus Disease 2019 Pneumonia in Wuhan, China. *JAMA Intern Med* Published Online First: 13 March 2020. doi:10.1001/jamainternmed.2020.0994

17 Chen R, Liang W, Jiang M, *et al.* Risk Factors of Fatal Outcome in Hospitalized Subjects With Coronavirus Disease 2019 From a Nationwide Analysis in China. *Chest* Published Online First: 15 April 2020. doi:10.1016/j.chest.2020.04.010

18 Collins Gary S., Reitsma Johannes B., Altman Douglas G., *et al.* Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD). *Circulation* 2015;**131**:211–9. doi:10.1161/CIRCULATIONAHA.114.014508

19 Knight SR, Ho A, Pius R, *et al.* Risk stratification of patients admitted to hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: development and validation of the 4C Mortality Score. *BMJ* 2020;**370**. doi:10.1136/bmj.m3339

20 Malik P, Patel U, Patel K, *et al.* Obesity a predictor of outcomes of COVID-19 hospitalized patients- A systematic Review and Meta-Analysis. *J Med Virol* Published Online First: 25 September 2020. doi:10.1002/jmv.26555

21 Eom C-S, Jeon CY, Lim J-W, *et al.* Use of acid-suppressive drugs and risk of pneumonia: a systematic review and meta-analysis. *CMAJ Can Med Assoc J J Assoc Medicale Can* 2011;**183**:310–9. doi:10.1503/cmaj.092129

22 Zirk-Sadowski J, Masoli JA, Delgado J, *et al.* Proton-Pump Inhibitors and Long-Term Risk of Community-Acquired Pneumonia in Older Adults. *J Am Geriatr Soc* 2018;**66**:1332–8. doi:10.1111/jgs.15385

23 Moayyedi P, Eikelboom JW, Bosch J, *et al.* Safety of Proton Pump Inhibitors Based on a Large, Multi-Year, Randomized Trial of Patients Receiving Rivaroxaban or Aspirin. *Gastroenterology* 2019;**157**:682-691.e2. doi:10.1053/j.gastro.2019.05.056

24 Almario CV, Chey WD, Spiegel BMR. Increased Risk of COVID-19 Among Users of Proton Pump Inhibitors. *Am J Gastroenterol* Published Online First: 25 August 2020. doi:10.14309/ajg.0000000000000798

25 Tarlow B, Gubatan J, Khan MA, *et al.* Are Proton Pump Inhibitors Contributing to SARS-COV-2 Infection? *Off J Am Coll Gastroenterol ACG* 2020;**Publish Ahead of Print**. doi:10.14309/ajg.0000000000000933

Figure legend

Figure 1. COVIDOGRAM – predicted probability of severe course of Covid-19 for men and women – visualization of simplified multivariable logistic regression model (for more details, see Supplemental Table 1)

Table 1. Characteristics of COVID-19 patients according to endpoint

	With severe course N = 465	Without severe course N = 6,990
Basic characteristics		
Men	260 (55.9%)	3,221 (46.1%)
Age, mean ± SD	74.8 ± 13.4	45.4 ± 20.2
History		
Hypertension	258 (55.5%)	1,266 (18.1%)
Atrial fibrillation	75 (16.1%)	179 (2.6%)
History of stroke	55 (11.8%)	131 (1.9%)
History of MI or PCI	32 (6.9%)	106 (1.5%)
Chronic heart failure	90 (19.4%)	157 (2.2%)
Chronic kidney disease	99 (21.3%)	215 (3.1%)
Diabetes mellitus	127 (27.3%)	519 (7.4%)
Chronic obstructive pulmonary disease	102 (21.9%)	643 (9.2%)
Acid-related disorders	206 (44.3%)	999 (14.3%)
Recent history of cancer (≤5 years)	48 (10.3%)	169 (2.4%)
Rheumatoid arthritis	14 (3.0%)	75 (1.1%)
Treatment		
ACE inhibitors	182 (39.1%)	947 (13.5%)
ARBs	85 (18.3%)	543 (7.8%)
Calcium channel blockers	121 (26.0%)	584 (8.4%)
Beta-blockers	212 (45.6%)	884 (12.6%)
Diuretics	214 (46.0%)	814 (11.6%)
Anticoagulants / antithrombotic agents	143 (30.8%)	502 (7.2%)
Statins	169 (36.3%)	882 (12.6%)

ACE = angiotensin-converting enzyme, ARBs = angiotensin II receptor blockers, MI = myocardial infarction, PCI = percutaneous coronary intervention

Table 2. Multivariable logistic regression model using backward stepwise algorithm for selection of independent predictors of severe course of COVID-19. Age, sex and comorbidities from Table 1 were entered into the model.

Predictors	Regression coefficients	SE	OR (95% CI)	P-value
Sex				
Women			Reference category	
Men	0.742	0.112	2.10 (1.68; 2.62)	**< 0.001**
Age				
< 40 years			Reference category	
40–49 years	1.227	0.464	3.41 (1.37; 8.48)	**0.008**
50–59 years	2.478	0.414	11.92 (5.30; 26.81)	**< 0.001**
60–69 years	3.424	0.399	30.68 (14.04; 67.04)	**< 0.001**
70–79 years	4.109	0.398	60.89 (27.93; 132.73)	**< 0.001**
80–89 years	4.725	0.400	112.68 (51.48; 246.63)	**< 0.001**
90+ years	5.299	0.428	200.12 (86.50; 462.97)	**< 0.001**
Comorbidities				
Chronic kidney disease	0.679	0.157	1.97 (1.45; 2.68)	**< 0.001**
Chronic obstructive pulmonary disease	0.436	0.144	1.55 (1.17; 2.05)	**0.002**
Recent history of cancer (≤5 years)	0.432	0.194	1.54 (1.05; 2.25)	**0.026**
Chronic heart failure	0.408	0.166	1.50 (1.09; 2.08)	**0.014**
Acid-related disorders	0.382	0.118	1.47 (1.16; 1.85)	**0.001**
Diabetes mellitus	0.323	0.129	1.38 (1.07; 1.78)	**0.012**
Intercept	-6.448	0.386	-	**< 0.001**

Hosmer and Lemeshow test: $\chi^2 = 9.315$, df = 8, p = 0.316;
overall predictive power: AUC (95% CI): 0.893 (0.880; 0.907); sensitivity 85.8% and specificity 80.3%.

Figure 1.

Men – predicted probability (95% CI)

Age	Morbidity level 0	1	2	3	4+
< 40	**0.3** (0.2; 0.7)	**0.5** (0.2; 1.1)	**0.8** (0.4; 1.7)	-	-
40–49	**1.1** (0.7; 1.9)	**1.7** (1.0; 2.9)	**2.7** (1.5; 4.5)	**4.1** (2.3; 7.2)	-
50–59	**3.7** (2.7; 5.2)	**5.7** (4.1; 7.8)	**8.7** (6.2; 12.0)	**12.9** (9.0; 18.3)	-
60–69	**8.9** (7.0; 11.3)	**13.3** (10.8; 16.3)	**19.4** (15.7; 23.7)	**27.3** (21.7; 33.8)	**37.0** (28.8; 46.1)
70–79	**16.4** (13.1; 20.3)	**23.5** (19.7; 27.8)	**32.5** (27.8; 37.6)	**43.0** (36.7; 49.5)	**54.1** (46.0; 62.0)
80–89	**26.7** (21.6; 32.5)	**36.3** (30.8; 42.2)	**47.2** (41.0; 53.4)	**58.3** (51.1; 65.1)	**68.6** (60.7; 75.6)
90+	**39.9** (30.5; 50.2)	**51.0** (41.1; 60.8)	**62.0** (52.1; 70.9)	**71.8** (62.3; 79.7)	**80.0** (71.2; 86.5)

Women – predicted probability (95% CI)

Age	Morbidity level 0	1	2	3	4+
< 40	**0.2** (0.1; 0.3)	**0.3** (0.1; 0.5)	**0.4** (0.2; 0.8)	-	-
40–49	**0.5** (0.3; 0.9)	**0.8** (0.5; 1.4)	**1.3** (0.7; 2.2)	**2.0** (1.1; 3.6)	-
50–59	**1.8** (1.3; 2.6)	**2.8** (2.0; 3.9)	**4.3** (3.0; 6.2)	**6.6** (4.4; 9.7)	-
60–69	**4.5** (3.4; 5.9)	**6.8** (5.3; 8.7)	**10.2** (7.9; 13.2)	**15.2** (11.4; 19.9)	**21.8** (15.9; 29.2)
70–79	**8.5** (6.6; 11.0)	**12.7** (10.3; 15.7)	**18.6** (15.2; 22.5)	**26.4** (21.4; 32.0)	**35.9** (28.6; 43.9)
80–89	**14.8** (11.7; 18.5)	**21.3** (17.7; 25.5)	**29.8** (25.1; 34.9)	**39.9** (33.6; 46.5)	**50.9** (42.7; 59.1)
90+	**24.0** (17.6; 31.8)	**33.1** (25.5; 41.7)	**43.6** (34.8; 52.9)	**54.8** (44.8; 64.4)	**65.5** (54.8; 74.8)

0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%

Table S1. Univariable logistic regression models

	Regression coefficients	SE	OR (95% CI)	P-value
Basic characteristics				
Sex – men (ref. women)	0.395	0.096	1.48 (1.23; 1.79)	**< 0.001**
Age 40–49 years (ref. < 40 years)	1.260	0.464	3.52 (1.42; 8.75)	**0.007**
Age 50–59 years (ref. < 40 years)	2.553	0.413	12.84 (5.72; 28.85)	**< 0.001**
Age 60–69 years (ref. < 40 years)	3.716	0.396	41.09 (18.93; 89.20)	**< 0.001**
Age 70–79 years (ref. < 40 years)	4.584	0.391	97.92 (45.52; 210.66)	**< 0.001**
Age 80–89 years (ref. < 40 years)	5.096	0.392	163.44 (75.80; 352.40)	**< 0.001**
Age 90+ years (ref. < 40 years)	5.555	0.418	258.59 (113.92; 586.98)	**< 0.001**
History				
Hypertension	1.729	0.098	5.64 (4.65; 6.83)	**< 0.001**
Atrial fibrillation	1.990	0.147	7.32 (5.48; 9.76)	**< 0.001**
History of stroke	1.949	0.169	7.02 (5.05; 9.77)	**< 0.001**
History of MI or PCI	1.569	0.208	4.80 (3.19; 7.21)	**< 0.001**
Chronic heart failure	2.346	0.142	10.45 (7.90; 13.81)	**< 0.001**
Chronic kidney disease	2.143	0.133	8.52 (6.57; 11.06)	**< 0.001**
Diabetes mellitus	1.544	0.114	4.68 (3.75; 5.85)	**< 0.001**
Chronic obstructive pulmonary disease	1.020	0.119	2.77 (2.19; 3.51)	**< 0.001**
Acid-related disorders	1.562	0.099	4.77 (3.93; 5.80)	**< 0.001**
Recent history of cancer (≤ 5 years)	1.536	0.171	4.65 (3.32; 6.50)	**< 0.001**
Rheumatoid arthritis	1.052	0.295	2.86 (1.60; 5.10)	**< 0.001**
Treatment				
ACE inhibitors	1.412	0.101	4.10 (3.37; 5.00)	**< 0.001**
ARBs	0.977	0.128	2.66 (2.07; 3.41)	**< 0.001**
Calcium channel blockers	1.350	0.114	3.86 (3.08; 4.83)	**< 0.001**
Beta-blockers	1.756	0.100	5.79 (4.76; 7.04)	**< 0.001**
Diuretics	1.867	0.100	6.47 (5.31; 7.87)	**< 0.001**
Anticoagulants / antithrombotic agents	1.747	0.111	5.74 (4.62; 7.13)	**< 0.001**
Statins	1.375	0.103	3.95 (3.23; 4.84)	**< 0.001**

Table S21. Simplified multivariable logistic regression model based on predictors selected by a backward stepwise algorithm

Predictors	Regression coefficients	SE	OR (95% CI)	P-value
Sex				
Women			Reference category	
Men	0.744	0.112	2.10 (1.69; 2.62)	**< 0.001**
Age				
< 40 years			Reference category	
40–49 years	1.215	0.464	3.37 (1.36; 8.37)	**0.009**
50–59 years	2.459	0.414	11.69 (5.20; 26.30)	**< 0.001**
60–69 years	3.388	0.398	29.59 (13.56; 64.59)	**< 0.001**
70–79 years	4.082	0.397	59.27 (27.22; 129.05)	**< 0.001**
80–89 years	4.700	0.399	109.93 (50.31; 240.20)	**< 0.001**
90+ years	5.302	0.427	200.66 (86.84; 463.63)	**< 0.001**
Morbidity level*				
1-point increase	0.448	0.049	1.56 (1.42; 1.72)	**< 0.001**
Intercept	-6.453	0.386		**< 0.001**

Hosmer and Lemeshow test: $\chi^2 = 10.310$, df = 8, p = 0.244;
overall predictive power: AUC (95% CI): 0.893 (0.879; 0.907); sensitivity 85.8% and specificity 80.3%.

*One point for each comorbidity: chronic kidney disease, chronic obstructive pulmonary disease, recent history of cancer (≤ 5 years), chronic heart failure, acid-related disorders, diabetes mellitus (maximum of 4 points per person, additional points were not considered due to a low prevalence in our cohort).

Table S32. Characteristics of COVID-19 patients in comparison to general population of the Czech Republic

	COVID-19 patients N = 7,455	General population N = 10.6 million	P-value
Basic characteristics			
Men	46.7%	49.1%	**< 0.001**
Age, mean ± SD	47.2 ± 21.1	42.5 ± 23.1	**< 0.001**
Medical history			
Hypertension	20.4%	17.5%	**< 0.001**
Atrial fibrillation	3.4%	2.3%	**< 0.001**
History of stroke	2.5%	1.4%	**< 0.001**
History of MI or PCI	1.9%	1.5%	**0.011**
Chronic heart failure	3.3%	1.6%	**< 0.001**
Chronic kidney disease	4.2%	2.4%	**< 0.001**
Diabetes mellitus	8.7%	7.2%	**< 0.001**
Chronic obstructive pulmonary disease	10.0%	8.4%	**< 0.001**
Acid-related disorders	16.2%	11.7%	**< 0.001**
Recent history of cancer (≤5 years)	2.9%	2.1%	**< 0.001**
Rheumatoid arthritis	1.2%	0.9%	**0.003**
Treatment			
ACE inhibitors	15.1%	14.5%	0.100
ARBs	8.4%	7.6%	**0.009**
Calcium channel blockers	9.5%	8.8%	0.066
Beta-blockers	14.7%	12.0%	**< 0.001**
Diuretics	13.8%	11.6%	**< 0.001**
Anticoagulants / antithrombotic agents	8.7%	6.1%	**< 0.001**
Statins	14.1%	13.9%	0.701

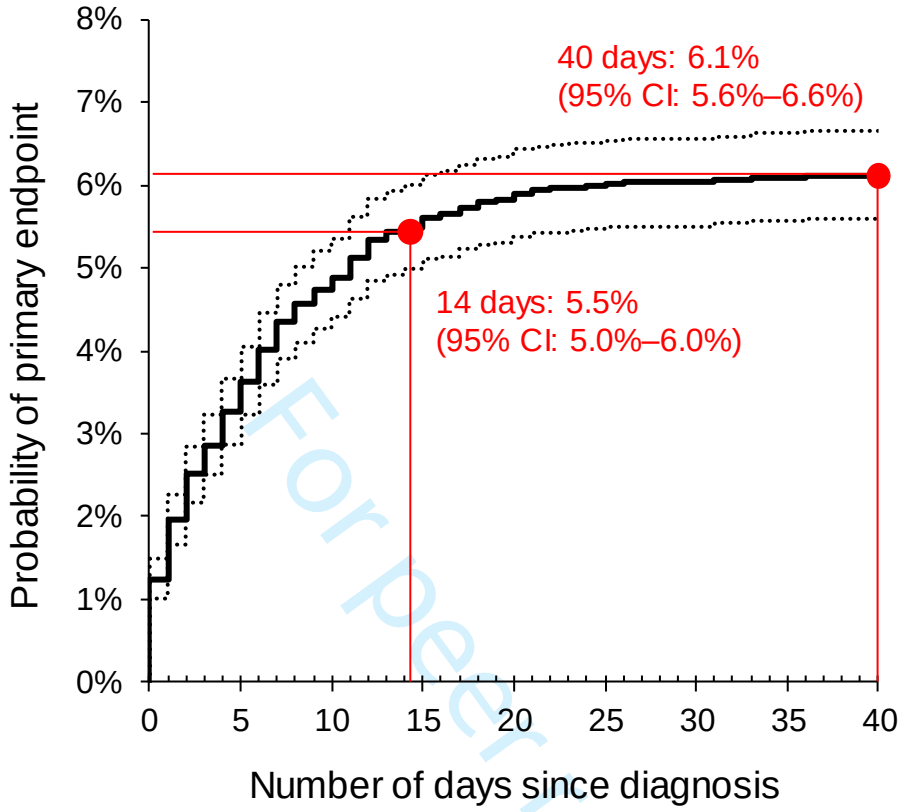
P-value of binomial test (categorical variables) and P-value of one-sample t-test (continuous variables).
ACE = angiotensin converting enzyme, ARB = angiotensin II receptor blockers, MI = myocardial infarction, PCI = percutaneous coronary intervention

Table S43. Characteristics of COVID-19 patients by using drugs for acid-related disorders

	Using drugs N = 1,205	Not using drugs N = 6,250	P-value
Basic characteristics			
Men	524 (43.5%)	2 957 (47.3%)	**0.015**
Age, mean ± SD	62.6 ± 18.3	44.2 ± 20.2	**< 0.001**
Medical history			
Hypertension	542 (45.0%)	982 (15.7%)	**< 0.001**
Atrial fibrillation	125 (10.4%)	129 (2.1%)	**< 0.001**
History of stroke	91 (7.6%)	95 (1.5%)	**< 0.001**
History of MI or PCI	69 (5.7%)	69 (1.1%)	**< 0.001**
Chronic heart failure	138 (11.5%)	109 (1.7%)	**< 0.001**
Chronic kidney disease	158 (13.1%)	156 (2.5%)	**< 0.001**
Diabetes mellitus	222 (18.4%)	424 (6.8%)	**< 0.001**
Chronic obstructive pulmonary disease	271 (22.5%)	474 (7.6%)	**< 0.001**
Recent history of cancer (≤5 years)	93 (7.7%)	124 (2.0%)	**< 0.001**
Rheumatoid arthritis	39 (3.2%)	50 (0.8%)	**< 0.001**
Treatment			
ACE inhibitors	348 (28.9%)	781 (12.5%)	**< 0.001**
ARBs	217 (18.0%)	411 (6.6%)	**< 0.001**
Calcium channel blockers	253 (21.0%)	452 (7.2%)	**< 0.001**
Beta-blockers	422 (35.0%)	674 (10.8%)	**< 0.001**
Diuretics	399 (33.1%)	629 (10.1%)	**< 0.001**
Anticoagulants / antithrombotic agents	270 (22.4%)	375 (6.0%)	**< 0.001**
Statins	369 (30.6%)	682 (10.9%)	**< 0.001**

P-value of Fisher's exact test (categorical variables) and P-value of Mann-Whitney U test (continuous variables).
ACE = angiotensin-converting enzyme, ARBs = angiotensin II receptor blockers, MI = myocardial infarction, PCI = percutaneous coronary intervention

Figure S1. Time to analysed endpoint (severe condition)



TRIPOD Checklist: Prediction Model Development and Validation

Section/Topic	Item		Checklist Item	Page
Title and abstract				
Title	1	D;V	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted.	1
Abstract	2	D;V	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	5
Introduction				
Background and objectives	3a	D;V	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	8
	3b	D;V	Specify the objectives, including whether the study describes the development or validation of the model or both.	9
Methods				
Source of data	4a	D;V	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.	10
	4b	D;V	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.	10
Participants	5a	D;V	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.	11
	5b	D;V	Describe eligibility criteria for participants.	10
	5c	D;V	Give details of treatments received, if relevant.	-
Outcome	6a	D;V	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	14
	6b	D;V	Report any actions to blind assessment of the outcome to be predicted.	-
Predictors	7a	D;V	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.	12
	7b	D;V	Report any actions to blind assessment of predictors for the outcome and other predictors.	-
Sample size	8	D;V	Explain how the study size was arrived at.	10
Missing data	9	D;V	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	10
Statistical analysis methods	10a	D	Describe how predictors were handled in the analyses.	12
	10b	D	Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation.	13
	10c	V	For validation, describe how the predictions were calculated.	13
	10d	D;V	Specify all measures used to assess model performance and, if relevant, to compare multiple models.	13
	10e	V	Describe any model updating (e.g., recalibration) arising from the validation, if done.	-
Risk groups	11	D;V	Provide details on how risk groups were created, if done.	-
Development vs. validation	12	V	For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.	-
Results				
Participants	13a	D;V	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	10
	13b	D;V	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.	T1
	13c	V	For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).	-
Model development	14a	D	Specify the number of participants and outcome events in each analysis.	14
	14b	D	If done, report the unadjusted association between each candidate predictor and outcome.	S4
Model specification	15a	D	Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point).	T2, S1
	15b	D	Explain how to the use the prediction model.	F1
Model performance	16	D;V	Report performance measures (with CIs) for the prediction model.	14
Model-updating	17	V	If done, report the results from any model updating (i.e., model specification, model performance).	-
Discussion				
Limitations	18	D;V	Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).	19
Interpretation	19a	V	For validation, discuss the results with reference to performance in the development data, and any other validation data.	-
	19b	D;V	Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.	19
Implications	20	D;V	Discuss the potential clinical use of the model and implications for future research.	20
Other information				
Supplementary information	21	D;V	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	3
Funding	22	D;V	Give the source of funding and the role of the funders for the present study.	2

*Items relevant only to the development of a prediction model are denoted by D, items relating solely to a validation of a prediction model are denoted by V, and items relating to both are denoted D;V. We recommend using the TRIPOD Checklist in conjunction with the TRIPOD Explanation and Elaboration document.

BMJ Open

Covidogram as a simple tool for predicting severe course of COVID-19: population based study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2020-045442.R2
Article Type:	Original research
Date Submitted by the Author:	20-Jan-2021
Complete List of Authors:	Jarkovský, Jiří; Masaryk University Institute of Biostatistics and Analyses; Institute of Health Information and Statistics of the Czech Republic Benešová, Klára; Institute of Health Information and Statistics of the Czech Republic; Institute of Health Information and Statistics of the Czech Republic Cerny, Vladimir; Masaryk Hospital in Usti nad Labem, Anesth. and ICU; Dalhousie University, Cerny Vladimir Razova, Jarmila; Ministry of Health of the Czech Republic Kala, petr; Masaryk University; University Hospital Brno Dolina, Jiri; University Hospital Brno; Masaryk University Majek, Ondrej; Institute of Health Information and Statistics of the Czech Republic; Masaryk University Institute of Biostatistics and Analyses Sebestova, Silvie; Institute of Health Information and Statistics of the Czech Republic Bezdekova, Monika; Institute of Health Information and Statistics of the Czech Republic Melicharova, Hana; Institute of Health Information and Statistics of the Czech Republic Snajdrova, Lenka; Institute of Health Information and Statistics of the Czech Republic; Masaryk University Institute of Biostatistics and Analyses Dusek, Ladislav; Institute of Health Information and Statistics of the Czech Republic; Masaryk University Institute of Biostatistics and Analyses Parenica, Jiri; University Hospital Brno; Masaryk University Faculty of Medicine
Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Epidemiology, Gastroenterology and hepatology
Keywords:	COVID-19, Gastroduodenal disease < GASTROENTEROLOGY, Organisation of health services < HEALTH SERVICES ADMINISTRATION & MANAGEMENT

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Covidogram as a simple tool for predicting severe course of COVID-19: population based study

Jiri Jarkovsky[1,2], Klara Benesova[1,2], Vladimir Cerny[3,4], Jarmila Razova[5], Petr Kala[6,7], Jiri Dolina[7,8], Ondrej Majek[1,2], Silvie Sebestova[1], Monika Bezdekova[1], Hana Melicharova[1], Lenka Snajdrova[1,2], Ladislav Dusek[1,2] and Jiri Parenica[1,6,7]

[1] Institute of Health Information and Statistics of the Czech Republic, Prague, Czech Republic

[2] Institute of Biostatistics and Analysis, Faculty of Medicine, Masaryk University, Brno, Czech Republic

[3] Department of Anaesthesiology, Perioperative Medicine and Intensive Care, J.E. Purkinje University, Masaryk Hospital, Usti nad Labem, Czech Republic

[4]Department of Anesthesia, Pain Management and Perioperative Medicine, Dalhousie University, Halifax, Canada

[5] Ministry of Health, Czech Republic

[6] Internal and Cardiology Department, University Hospital Brno, Czech Republic

[7] Faculty of Medicine, Masaryk University, Brno, Czech Republic

[8] Gastroenterology and Internal Department, University Hospital Brno, Czech Republic

Key words: Covid-19; severe course; prognostic score; proton-pump inhibitors

A funding statement This research was supported by grant the Czech Republic Operational Programme eHealth and Rare Disease CZ.03.4.74/0.0/0.0/15_025/

Word count 2565 (without abstract, tables, references)

Contributors JJ, LD, VC and JR designed the study and wrote the research plan. OM, SS, MB and HM extracted the data used for the study from the databases. JJ and KB undertook statistical analysis with feedback from LD. JP, JJ, PK, JD interpreted the results and wrote the first draft of the manuscript with critical comments and revision from VC, LD and LS.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication The analysis was done on data from the National Health Information System (NHIS), which was supplemented with data from the Information System of Infectious Diseases (ISID). Data in the ISID are collected in compliance with Act No. 258/2000 Coll. on Protection of Public Health, data in the NHIS are collected – and interconnected with data from ISID – in accordance with Act No. 372/2011 Coll., on Health Services and Conditions of Their Provision. Due to this legal mandate, the retrospective analyses did not require either approval by an ethics committee or informed consents from participants.

Data availability statement The anonymized data available upon reasonable request. The data are deidentified participant data, and available from the first author JJ (jarkovsky@uzis.cz). The reuse of the data subset is permitted only for revalidation of the results.

Corresponding author:

Jiri Parenica, MD, MSc, PhD, Professor,

University Hospital Brno, Brno, Jihomoravský, CZ and Masaryk University Faculty of Medicine Brno, Jihomoravský, CZ

e-mail: jiri.parenica@atlas.cz

Telephone: +420 53223 2654; Fax: +420 532232907

Abstract

Objectives COVID-19 might either be entirely asymptomatic or manifest itself with a large variability of disease severity. It is beneficial to identify early patients with a high risk of severe course. The aim of the analysis was to develop a prognostic model for the prediction of the severe course of acute respiratory infection.

Design A population based study.

Setting Czech Republic.

Participants The first 7,455 consecutive COVID-19 patients who were identified by RT-PCR testing from 1 March 2020 to 17 May 2020.

Interventions None

Primary outcome Severe course of Covid-19.

Result Of a total 6.2% of patients developed a severe course of COVID-19. Age, male sex, chronic kidney disease, chronic obstructive pulmonary disease, recent history of cancer, chronic heart failure, acid-related disorders treated with proton-pump inhibitors and diabetes mellitus were found to be independent negative prognostic factors. (AUC was 0.893). The results were visualised by risk heat maps and we called this diagram a “covidogram”. Acid-related disorders treated with proton-pump inhibitors might represent a negative prognostic factor.

Conclusion We developed a very simple prediction model called “covidogram”, which is based on elementary independent variables (age, male sex and the presence of several chronic diseases) and represents a tool that makes it possible to identify – with a high reliability – patients who are at risk of a severe course of COVID-19. Obtained results open clinically relevant question about the role of acid-related disorders treated by proton pump inhibitors as predictor for severe course of COVID-19.

Strengths and limitations of the study

- The majority of consecutive patients diagnosed with COVID-19 in the Czech Republic were included in the analysis, regardless of whether they were hospitalized or not.
- The cohort covers also asymptomatic and oligosymptomatic patients identified thanks to epidemiological monitoring.
- The cohort does not include strictly all COVID-19 cases in the Czech Republic because some patients are asymptomatic and have not been tested.
- The proposed prediction model is a simple tool that makes it possible to identify – with a high reliability (AUC 0.893) – patients who are at risk of a severe course of COVID-19.
- Flexible calibration curves based on local regression confirm the predictive model is well-calibrated. The out of sample calibration is currently not available as data of large sample of patients from the second wave COVID-19 in the Czech Republic are still under preparation.
- Due to the retrospective nature of this study, which is based on data of administrative registries, results of laboratory, clinical and X-ray examinations were not available. Conclusions regarding the influence of

comorbidities and the consumption of medicinal products should be interpreted with caution and will require further validation.

Introduction

The coronavirus disease 2019 (COVID-19) is caused by betacoronavirus SARS-CoV-2, which enters human cells via the membrane-bound angiotensin-converting enzyme 2 (ACE2)[1]. The presence of infection might be entirely asymptomatic[2], or manifest itself with a large variability of disease severity, a number of unspecific clinical symptoms (fever, fatigue, myalgia) and various degrees of organ dysfunction. Most frequently, the disease affects the respiratory system (manifested as dry cough, dyspnoea, haemoptysis, pneumonia or ARDS) and the cardiovascular system (presented as myocardial injury or myocarditis, ventricular arrhythmias, haemodynamic instability or deep vein thrombosis), while other organ systems (such as the central nervous [3], gastrointestinal system or kidneys [4]) are affected less frequently. In a number of patients, there is a risk of multiple organ failure, ultimately leading to death[5–9]. According to the report of World Health Organization, as 12 November 2020, the rate of mortality among COVID-19 patients is 2.28%[10]. The management of COVID-19 patients depends on the disease severity: patients with a mild course of the disease can be treated in their home environment[11]. However, the clinical picture of COVID-19 patients can quickly turn into an unfavourable clinical course [7], and it is therefore clinically relevant to identify patients with a high risk of severe course of the disease as

early as possible [12]. It was repeatedly demonstrated that older age > 65, cancer, chronic obstructive pulmonary disease, moderate-to-severe asthma, diabetes mellitus, chronic renal disease, immunocompromised state, obesity (BMI>30), pregnancy, sickle cell disease, smoking and cardiovascular disease are related to a high-risk course of the disease[7,13–15].

In the Czech Republic, a prospective population-based and centralised collection of data on COVID-19 patients was developed at the beginning of the pandemic, with a possibility to interconnect these data with those recorded in other population-based registries of the National Health Information System (NHIS), and thus to obtain information on each patient's history and management.

The aim of this study was to develop a prognostic model for the prediction of the severe course of acute respiratory infection, defined by the necessity of intensive care being provided in the Intensive care units (ICU), that include mechanical ventilation, extracorporal membrane oxygenation (ECMO) support and/or death.

Methods

Population of patients

The analysis is based on data from a population-based registry containing records of all consecutive COVID-19 patients in the Czech Republic who were identified by RT-PCR testing and validated by the National Institute of Public Health.

The monitored cohort consisted of patients who were recorded in the National Information System of Infectious Diseases (ISID) between 1 March 2020 and 17 May 2020.

As of 17 May 2020, a total of 356,515 tests were performed in the Czech Republic, and the diagnosis of COVID-19 was confirmed in 8,475 cases. Of the confirmed COVID-19 cases, 464 patients with unknown history in the National Health Information System (NHIS) or the patients being foreign nationals with unknown medical history – were excluded from the analysis, and further 556 patients with a follow-up period shorter than 14 days (i. e. patients diagnosed after 3 May 2020) were excluded as well. 90% of events occur within 14 days (Figure S1). Analysis without censoring and with a fixed follow-up length was chosen with the objective to simplify visualisation and interpretation of results of the analysis for its practical application. The basic characteristics of patients (age, gender) were provided for all patients, data on comorbidities were available for all patients with match between ISID and NHIS datasets; patients without match between datasets were earlier excluded from the analysis, no

other missing data handling was necessary. On top of that, characteristics of the cohort diagnosed with COVID-19 were compared to those of the population of the Czech Republic (10.6 million).

Diagnosis

The diagnosis of COVID-19 was based on the detection of SARS-CoV-2 by real-time reverse transcription polymerase chain reaction (RT-PCR). At first, the analyses were performed in the National Reference Laboratory of the National Institute of Public Health (NIPH); other certified laboratories were later appointed to carry out RT-PCR testing as well.

Systematic collection of data

The analysis was done on data from the National Health Information System (NHIS), which was supplemented with data from the Information System of Infectious Diseases (ISID). Data in the ISID are collected in compliance with Act No. 258/2000 Coll. on Protection of Public Health, whereas data in the NHIS are collected – and interconnected with data from ISID – in accordance with Act No. 372/2011 Coll., on Health Services and Conditions of Their Provision. Due to this legal mandate, the retrospective analyses did not require either approval by an ethics committee or informed consents from participants; moreover, it is a population-based analysis of all diagnosed COVID-19 cases in the Czech Republic. The cohort covers also

asymptomatic and oligosymptomatic patients identified thanks to epidemiological monitoring.

The latest data on COVID-19 patients, the severity of their condition as well as the necessity of hospitalisation in an ICU, including the use of mechanical ventilation or ECMO, together with information on death, have been entered into the ISID in real time. Apart from that, data on COVID-19 patients have been enriched with information on their comorbidities: this information is available in the National Register of Reimbursed Health Services (NRRHS), which contains data on all healthcare reported within the public health insurance system (accounting for almost 100% of healthcare provided in the Czech Republic). Comorbidities are determined from combinations of reported diagnoses (ICD-10), ATC codes of drugs, and procedures defined in code lists used by Czech health insurance companies. Only diseases and conditions with a higher prevalence in the population or those identified in literature[7,9,15–17] as potential predictors of a severe course of COVID-19 were evaluated, with the aim to assess their potential influence on the resulting model.

Statistical analysis

Standard descriptive statistics were used to describe the data: categorical variables were described by absolute and relative frequencies, whereas continuous variables were described by means and standard

deviations. The Fisher's exact test (for categorical variables) and Mann–Whitney U test (for continuous variables) were used to compare characteristics between groups depending on the monitored endpoint, unless stated otherwise. The predictive power of patient characteristics with regard to the analysed endpoint was evaluated by univariable and multivariable logistic regression and described by odds ratios, their 95% confidence intervals and statistical significance; a backward stepwise algorithm was used to choose the optimal model, and a ROC analysis was employed to evaluate the overall predictive power of the model. A flexible calibration curve [18] was adopted for the evaluation of goodness of fit of the model. The results of the model were expressed by a risk heat map taking account of the patients' age, sex and comorbidities. A 10-fold cross-validation was performed to obtain estimates of model performance that are adjusted for in-sample optimism. A model was created in accordance with TRIPOD check list for prediction model development and validation [19]. The analysis was computed using the Vertica database and a MS SQL Server for data pre-processing and SPSS 25.0.0.1 and R-3.6.1 for the statistical analysis of data. The level of statistical significance was set at $\alpha=0.05$ for all analyses.

Results

Of the total of 7,455 evaluated patients, 1,182 (15.9%) were hospitalised, 465 of them (6.2%) developed a severe course of the disease (i.e. reached the primary endpoint – death or the necessity of intensive care being provided in the Intensive care units (ICU), that include mechanical ventilation and/or ECMO support): 174 patients (2.3%) required mechanical ventilation, 11 patients (0.1%) were provided ECMO, and 287 patients (3.8%) died. Patients with the monitored endpoint were older (74.8 ± 13.4 vs 45.4 ± 20.2 years), more frequently of male sex and suffered at least one of all monitored comorbidities (Table 1, $p < 0.001$ for all parameters; for univariable logistic regression results, see Table S1). Older age was determined by the multivariable logistic regression analysis to be the most significant predictor: the risk of a severe course of the disease increases progressively from the age of 40 years onwards (Table 2). Male sex, chronic kidney disease, chronic obstructive pulmonary disease, recent history of cancer (in the last five years), chronic heart failure, acid related disorders treated with proton-pump inhibitors and diabetes mellitus were other significant predictors; the latter six conditions are hereinafter referred to as ***prognostically significant comorbidities*** (Table 2). The overall predictive power of the model, evaluated by the ROC analysis and expressed by the AUC, was 0.893 (95% CI: 0.880–0.907; sensitivity 85.8% and

specificity 80.3%). After performing the 10-fold cross-validation to validate the results, the average AUC of 0.891 (in the range 0.856–0.943) was obtained. For the purpose of an easier interpretation in clinical practice, a simplified version of the model was developed, taking into consideration the number of prognostically significant comorbidities obtained from the previous model (Table 3). Both original and simplified models are well-calibrated, as is supported by calibration curves in the Supplementary Figures (Figure S2A for the original model and Figure S2B for the simplified model). The results of the simplified model were visualised by risk heat maps for men and women separately (Figure 1), and we called this diagram a "***covidogram***". The diagram shows how the risk increases progressively with age and with the number of prognostically significant comorbidities.

It is obvious from the comparison of basic patient characteristics (Table 1) and the results of the multivariable analysis (Table 2) that although a number of conditions occur more frequently in the group of patients with a severe course of the disease, not all of them are independent predictors (coronary artery disease, history of stroke, atrial fibrillation, hypertension or treatment by ACEIs and ARBs).

The comparison of characteristics of patients with confirmed COVID-19 to those of the entire population of the Czech Republic showed that COVID-19

patients are slightly older and have the monitored comorbidities slightly more frequently. (Table S2).

Discussion

New findings about COVID-19

We have developed a prognostic model for the prediction of the severe course of COVID-19 in consecutive patients with positive COVID-19 RT-PCR test. This a simple tool called “covidogram” has a very good predictive power (AUC 0.893). Age is the most significant factor, and the risk increases progressively from the age of 40 years onwards. To our knowledge, this is the first study to suggest that acid related disorders treated by proton pump inhibitors might be independent risk predictors as well. By contrast, not all cardiovascular diseases (such as uncomplicated hypertension or coronary artery disease) increase the risk of a severe course of COVID-19.

The “covidogram” was designed as a model to assess the risk of unfavourable development of the patient’s condition based on his/her history of chronic disease, and can serve as a tool to estimate the number of severe cases of COVID-19 in a population. When assessing the risk for an individual patient in clinical practice, it is certainly necessary to take into consideration also other pieces of information on the current condition of that patient

(respiratory rate, peripheral oxygen saturation, level of consciousness, urea level, C reactive protein[20], procalcitonin, aspartate aminotransferase[17], high temperature[16], elevation of cardiac markers, lung infiltrates >50%[9]) as well as obesity, which can also increase the risk of a severe course of COVID-19 [21].

Acid-related disorders

Surprisingly, our analysis revealed that the presence of acid-related disorders might be theoretically linked to a severe course of COVID-19. Patients were predominantly treated with proton-pump inhibitors (1,175 patients in total, out of which 706 were treated with omeprazole and 402 with pantoprazole as the two most frequently used drugs); only a small proportion of them were treated with H_2-receptor antagonists (30 patients). The main indications for treatment with these drugs generally involve gastro-oesophageal reflux disease, functional dyspepsia, gastric and duodenal ulcers, gastric acid hypersecretory states as well as gastroprotection in patients using non-steroidal anti-inflammatory drugs, dual antiplatelet therapy, biphophonates or same selective serotonin reuptake inhibitors (SSRIs). The effect of inhibition of hydrochloric acid secretion is followed by an increase in the intragastric pH (to a value above 2–4), which might hypothetically decrease the physiological bactericidal/virucidal effect of gastric acid and decrease the

activity of lysosomal enzymes. Published data showed that long-term use of proton-pump inhibitors could slightly increase the risk of pneumonia[22,23] and enteric infections[24].

Our comparison of patients with and without acid-related disorders (Table S3) showed that patients with these disorders are markedly older and have prognostically significant comorbidities more frequently. Our analysis cannot determine whether there is any causal relationship between the presence of acid-related disorders and a severe course of COVID-19 or whether it is just a coincidence. At the same time, it must be stressed out that the vast majority of patients were treated with proton-pump inhibitors, not with H_2-receptor agonists. The observation is complicated also by the fact that some patients may not be adherent to their PPI regimen and there was great variability in the amount of time that they have been on PPIs (from one to twelfth month within 2019). Based on our analysis we are not able to decide whether severity of disease might be theoretically explained by pharmacology or by underlying pathology of acid related disorders. Recently, Almario et al. demonstrated association between using proton pump inhibitors and odds of a positive COVID-19 test [25]. Similar trend was reported by Tarlow [26].

Strengths and limitations

This study is based on a fully integrated national health information system covering the entire population of a country – which proposed a prediction model estimating individually-based risk of a severe course of COVID-19. Because this model uses data readily available in health and administrative registries, it can be easily used for the prediction of intensive care use in the context of decision-making at the national level.

On the other hand, our analysis has a number of limitations. Results of laboratory, clinical and X-ray examinations performed at the time of patient admission to hospitals were not available, and these very important pieces of information could therefore not be analysed; instead, our analysis is based on administrative data, with the exception of endpoints. Furthermore, analytical processing of a cohort of patients cannot capture the risk of less frequent conditions that might increase the risk of a severe course of COVID-19 (e.g. patients with immunodeficiencies, those after organ transplantation, or those undergoing immunosuppressive therapy or biological therapy). The cohort does not include strictly all COVID-19 cases in the Czech Republic because some patients are asymptomatic and have not been tested. Older peoples with more comorbidities are probably more likely to have a symptomatic course of COVID-19. It could be also a reason why the population of patients diagnosed with COVID-19 is older and with more comorbidities in comparison with the Czech

Republic population (Table S2). Due to the retrospective nature of this study, which is based on data of administrative registries and is focused on the development of a prediction model, any conclusions regarding the influence of comorbidities and the consumption of medicinal products should be interpreted with caution and will require further validation. The in sample callibration of the model was assessed by a by flexible calibration curves which confirmed that the predictive model is well calibrated. Out-of-sample calibration is currently not available as data of large sample of patients from the second wave COVID-19 in the Czech Republic are still under preparation.

Conclusion

The proposed prediction model "covidogram" is based on elementary independent variables (age, male sex and the presence of chronic disease) and represents a simple tool that makes it possible to identify – with a high reliability (AUC 0.893) – patients who are at risk of a severe course of COVID-19.

Finally, the analysis has shown, for the first time, that acid-related disorders treated with proton-pump inhibitors might also be theoretically associated with a severe course of the disease.

A competing interests statement All authors have disclosed that they do not have any conflicts of interest (no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work; no other relationships or activities that could appear to have influenced the submitted work).

References

1 Hoffmann M, Kleine-Weber H, Schroeder S, *et al.* SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* 2020;**181**:271-280.e8. doi:10.1016/j.cell.2020.02.052

2 Sutton D, Fuchs K, D'Alton M, *et al.* Universal Screening for SARS-CoV-2 in Women Admitted for Delivery. *N Engl J Med* Published Online First: 13 April 2020. doi:10.1056/NEJMc2009316

3 Aghagoli G, Gallo Marin B, Katchur NJ, *et al.* Neurological Involvement in COVID-19 and Potential Mechanisms: A Review. *Neurocrit Care* 2020;:1–10. doi:10.1007/s12028-020-01049-4

4 Aghagoli G, Gallo Marin B, Soliman LB, *et al.* Cardiac involvement in COVID-19 patients: Risk factors, predictors, and complications: A review. *J Card Surg* Published Online First: 19 April 2020. doi:10.1111/jocs.14538

5 Giacomelli A, Pezzati L, Conti F, *et al.* Self-reported olfactory and taste disorders in SARS-CoV-2 patients: a cross-sectional study. *Clin Infect Dis Off Publ Infect Dis Soc Am* Published Online First: 26 March 2020. doi:10.1093/cid/ciaa330

6 Han C, Duan C, Zhang S, *et al.* Digestive Symptoms in COVID-19 Patients With Mild Disease Severity: Clinical Presentation, Stool Viral RNA Testing, and Outcomes. *Am J Gastroenterol* Published Online First: 15 April 2020. doi:10.14309/ajg.0000000000000664

7 Huang C, Wang Y, Li X, *et al.* Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet Lond Engl* 2020;**395**:497–506. doi:10.1016/S0140-6736(20)30183-5

8 Kim I-C, Kim JY, Kim HA, *et al.* COVID-19-related myocarditis in a 21-year-old female patient. *Eur Heart J* Published Online First: 13 April 2020. doi:10.1093/eurheartj/ehaa288

9 Wang D, Hu B, Hu C, *et al.* Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus–Infected Pneumonia in Wuhan, China. *JAMA* 2020;**323**:1061–9. doi:10.1001/jama.2020.1585

10 WHO. Coranvirus disease (COVID-19) pandemic. Available from https://www.who.int/emergencies/diseases/novel-coronavirus-2019. Accessed 12 Dec 2020.

11 Gandhi RT, Lynch JB, Rio C del. Mild or Moderate Covid-19. *N Engl J Med* Published Online First: 24 April 2020. doi:10.1056/NEJMcp2009249

12 Berlin DA, Gulick RM, Martinez FJ. Severe Covid-19. *N Engl J Med* 2020;**0**:null. doi:10.1056/NEJMcp2009575

13 Mehra MR, Desai SS, Kuy S, *et al.* Cardiovascular Disease, Drug Therapy, and Mortality in Covid-19. *N Engl J Med* Published Online First: 1 May 2020. doi:10.1056/NEJMoa2007621

14 Wang T, Tang C, Chen R, *et al.* Clinical Features of Coronavirus Disease 2019 Patients With Mechanical Ventilation: A Nationwide Study in China. *Crit Care Med* 2020;**Online First**. doi:10.1097/CCM.0000000000004473

15 CDC, Coronavirus-19. Afvailable from https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/people-with-medical-conditions.html.

16 Wu C, Chen X, Cai Y, *et al.* Risk Factors Associated With Acute Respiratory Distress Syndrome and Death in Patients With Coronavirus Disease 2019 Pneumonia in Wuhan, China. *JAMA Intern Med* Published Online First: 13 March 2020. doi:10.1001/jamainternmed.2020.0994

17 Chen R, Liang W, Jiang M, *et al.* Risk Factors of Fatal Outcome in Hospitalized Subjects With Coronavirus Disease 2019 From a Nationwide Analysis in China. *Chest* Published Online First: 15 April 2020. doi:10.1016/j.chest.2020.04.010

18 Van Calster B, Nieboer D, Vergouwe Y, *et al.* A calibration hierarchy for risk models was defined: from utopia to empirical data. *J Clin Epidemiol* 2016;**74**:167–76. doi:10.1016/j.jclinepi.2015.12.005

19 Collins Gary S., Reitsma Johannes B., Altman Douglas G., *et al.* Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD). *Circulation* 2015;**131**:211–9. doi:10.1161/CIRCULATIONAHA.114.014508

20 Knight SR, Ho A, Pius R, *et al.* Risk stratification of patients admitted to hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: development and validation of the 4C Mortality Score. *BMJ* 2020;**370**. doi:10.1136/bmj.m3339

21 Malik P, Patel U, Patel K, *et al.* Obesity a predictor of outcomes of COVID-19 hospitalized patients- A systematic Review and Meta-Analysis. *J Med Virol* Published Online First: 25 September 2020. doi:10.1002/jmv.26555

22 Eom C-S, Jeon CY, Lim J-W, *et al.* Use of acid-suppressive drugs and risk of pneumonia: a systematic review and meta-analysis. *CMAJ Can Med Assoc J J Assoc Medicale Can* 2011;**183**:310–9. doi:10.1503/cmaj.092129

23 Zirk-Sadowski J, Masoli JA, Delgado J, *et al.* Proton-Pump Inhibitors and Long-Term Risk of Community-Acquired Pneumonia in Older Adults. *J Am Geriatr Soc* 2018;**66**:1332–8. doi:10.1111/jgs.15385

24 Moayyedi P, Eikelboom JW, Bosch J, *et al.* Safety of Proton Pump Inhibitors Based on a Large, Multi-Year, Randomized Trial of Patients Receiving Rivaroxaban or Aspirin. *Gastroenterology* 2019;**157**:682-691.e2. doi:10.1053/j.gastro.2019.05.056

25 Almario CV, Chey WD, Spiegel BMR. Increased Risk of COVID-19 Among Users of Proton Pump Inhibitors. *Am J Gastroenterol* Published Online First: 25 August 2020. doi:10.14309/ajg.0000000000000798

26 Tarlow B, Gubatan J, Khan MA, *et al.* Are Proton Pump Inhibitors Contributing to SARS-COV-2 Infection? *Off J Am Coll Gastroenterol ACG* 2020;**Publish Ahead of Print**. doi:10.14309/ajg.0000000000000933

Figure legend

Figure 1. COVIDOGRAM – predicted probability of severe course of Covid-19 for men and women – visualization of simplified multivariable logistic regression model (for more details, see Table 3)

Table 1. Characteristics of COVID-19 patients according to endpoint

	With severe course N = 465	Without severe course N = 6,990
Basic characteristics		
Men	260 (55.9%)	3,221 (46.1%)
Age, mean ± SD	74.8 ± 13.4	45.4 ± 20.2
History		
Hypertension	258 (55.5%)	1,266 (18.1%)
Atrial fibrillation	75 (16.1%)	179 (2.6%)
History of stroke	55 (11.8%)	131 (1.9%)
History of MI or PCI	32 (6.9%)	106 (1.5%)
Chronic heart failure	90 (19.4%)	157 (2.2%)
Chronic kidney disease	99 (21.3%)	215 (3.1%)
Diabetes mellitus	127 (27.3%)	519 (7.4%)
Chronic obstructive pulmonary disease	102 (21.9%)	643 (9.2%)
Acid-related disorders	206 (44.3%)	999 (14.3%)
Recent history of cancer (≤ 5 years)	48 (10.3%)	169 (2.4%)
Rheumatoid arthritis	14 (3.0%)	75 (1.1%)
Treatment		
ACE inhibitors	182 (39.1%)	947 (13.5%)
ARBs	85 (18.3%)	543 (7.8%)
Calcium channel blockers	121 (26.0%)	584 (8.4%)
Beta-blockers	212 (45.6%)	884 (12.6%)
Diuretics	214 (46.0%)	814 (11.6%)
Anticoagulants / antithrombotic agents	143 (30.8%)	502 (7.2%)
Statins	169 (36.3%)	882 (12.6%)

ACE = angiotensin-converting enzyme, ARBs = angiotensin II receptor blockers, MI = myocardial infarction, PCI = percutaneous coronary intervention

Table 2. Multivariable logistic regression model using backward stepwise algorithm for selection of independent predictors of severe course of COVID-19. Age, sex and comorbidities from Table 1 were entered into the model.

Predictors	Regression coefficients	SE	OR (95% CI)	P-value
Sex				
Women			Reference category	
Men	0.742	0.112	2.10 (1.68; 2.62)	**< 0.001**
Age				
< 40 years			Reference category	
40–49 years	1.227	0.464	3.41 (1.37; 8.48)	**0.008**
50–59 years	2.478	0.414	11.92 (5.30; 26.81)	**< 0.001**
60–69 years	3.424	0.399	30.68 (14.04; 67.04)	**< 0.001**
70–79 years	4.109	0.398	60.89 (27.93; 132.73)	**< 0.001**
80–89 years	4.725	0.400	112.68 (51.48; 246.63)	**< 0.001**
90+ years	5.299	0.428	200.12 (86.50; 462.97)	**< 0.001**
Comorbidities				
Chronic kidney disease	0.679	0.157	1.97 (1.45; 2.68)	**< 0.001**
Chronic obstructive pulmonary disease	0.436	0.144	1.55 (1.17; 2.05)	**0.002**
Recent history of cancer (≤ 5 years)	0.432	0.194	1.54 (1.05; 2.25)	**0.026**
Chronic heart failure	0.408	0.166	1.50 (1.09; 2.08)	**0.014**
Acid-related disorders	0.382	0.118	1.47 (1.16; 1.85)	**0.001**
Diabetes mellitus	0.323	0.129	1.38 (1.07; 1.78)	**0.012**
Intercept	-6.448	0.386	-	**< 0.001**

Overall predictive power: AUC (95% CI): 0.893 (0.880; 0.907); sensitivity 85.8% and specificity 80.3%.

Table 3. Simplified multivariable logistic regression model based on predictors selected by a backward stepwise algorithm

Predictors	Regression coefficients	SE	OR (95% CI)	P-value
Sex				
Women			Reference category	
Men	0.744	0.112	2.10 (1.69; 2.62)	**< 0.001**
Age				
< 40 years			Reference category	
40–49 years	1.215	0.464	3.37 (1.36; 8.37)	**0.009**
50–59 years	2.459	0.414	11.69 (5.20; 26.30)	**< 0.001**
60–69 years	3.388	0.398	29.59 (13.56; 64.59)	**< 0.001**
70–79 years	4.082	0.397	59.27 (27.22; 129.05)	**< 0.001**
80–89 years	4.700	0.399	109.93 (50.31; 240.20)	**< 0.001**
90+ years	5.302	0.427	200.66 (86.84; 463.63)	**< 0.001**
Morbidity level*				
1-point increase	0.448	0.049	1.56 (1.42; 1.72)	**< 0.001**
Intercept	-6.453	0.386		**< 0.001**

Overall predictive power: AUC (95% CI): 0.893 (0.879; 0.907); sensitivity 85.8% and specificity 80.3%.

*One point for each comorbidity: chronic kidney disease, chronic obstructive pulmonary disease, recent history of cancer (≤ 5 years), chronic heart failure, acid-related disorders, diabetes mellitus (maximum of 4 points per person, additional points were not considered due to a low prevalence in our cohort).

Figure 1.

Men – predicted probability (%)

Age	Morbidity level				
	0	1	2	3	4+
< 40	0.3	0.5	0.8	-	-
40–49	1.1	1.7	2.7	4.1	-
50–59	3.7	5.7	8.7	12.9	-
60–69	8.9	13.3	19.4	27.3	37.0
70–79	16.4	23.5	32.5	43.0	54.1
80–89	26.7	36.3	47.2	58.3	68.6
90+	39.9	51.0	62.0	71.8	80.0

Women – predicted probability (%)

Age	Morbidity level				
	0	1	2	3	4+
< 40	0.2	0.3	0.4	-	-
40–49	0.5	0.8	1.3	2.0	-
50–59	1.8	2.8	4.3	6.6	-
60–69	4.5	6.8	10.2	15.2	21.8
70–79	8.5	12.7	18.6	26.4	35.9
80–89	14.8	21.3	29.8	39.9	50.9
90+	24.0	33.1	43.6	54.8	65.5

0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%

Table S1. Univariable logistic regression models

	Regression coefficients	SE	OR (95% CI)	P-value
Basic characteristics				
Sex – men (ref. women)	0.395	0.096	1.48 (1.23; 1.79)	**< 0.001**
Age 40–49 years (ref. < 40 years)	1.260	0.464	3.52 (1.42; 8.75)	**0.007**
Age 50–59 years (ref. < 40 years)	2.553	0.413	12.84 (5.72; 28.85)	**< 0.001**
Age 60–69 years (ref. < 40 years)	3.716	0.396	41.09 (18.93; 89.20)	**< 0.001**
Age 70–79 years (ref. < 40 years)	4.584	0.391	97.92 (45.52; 210.66)	**< 0.001**
Age 80–89 years (ref. < 40 years)	5.096	0.392	163.44 (75.80; 352.40)	**< 0.001**
Age 90+ years (ref. < 40 years)	5.555	0.418	258.59 (113.92; 586.98)	**< 0.001**
History				
Hypertension	1.729	0.098	5.64 (4.65; 6.83)	**< 0.001**
Atrial fibrillation	1.990	0.147	7.32 (5.48; 9.76)	**< 0.001**
History of stroke	1.949	0.169	7.02 (5.05; 9.77)	**< 0.001**
History of MI or PCI	1.569	0.208	4.80 (3.19; 7.21)	**< 0.001**
Chronic heart failure	2.346	0.142	10.45 (7.90; 13.81)	**< 0.001**
Chronic kidney disease	2.143	0.133	8.52 (6.57; 11.06)	**< 0.001**
Diabetes mellitus	1.544	0.114	4.68 (3.75; 5.85)	**< 0.001**
Chronic obstructive pulmonary disease	1.020	0.119	2.77 (2.19; 3.51)	**< 0.001**
Acid-related disorders	1.562	0.099	4.77 (3.93; 5.80)	**< 0.001**
Recent history of cancer (≤ 5 years)	1.536	0.171	4.65 (3.32; 6.50)	**< 0.001**
Rheumatoid arthritis	1.052	0.295	2.86 (1.60; 5.10)	**< 0.001**
Treatment				
ACE inhibitors	1.412	0.101	4.10 (3.37; 5.00)	**< 0.001**
ARBs	0.977	0.128	2.66 (2.07; 3.41)	**< 0.001**
Calcium channel blockers	1.350	0.114	3.86 (3.08; 4.83)	**< 0.001**
Beta-blockers	1.756	0.100	5.79 (4.76; 7.04)	**< 0.001**
Diuretics	1.867	0.100	6.47 (5.31; 7.87)	**< 0.001**
Anticoagulants / antithrombotic agents	1.747	0.111	5.74 (4.62; 7.13)	**< 0.001**
Statins	1.375	0.103	3.95 (3.23; 4.84)	**< 0.001**

Table S2. Characteristics of COVID-19 patients in comparison to general population of the Czech Republic

	COVID-19 patients N = 7,455	General population N = 10.6 million	P-value
Basic characteristics			
Men	46.7%	49.1%	**< 0.001**
Age, mean ± SD	47.2 ± 21.1	42.5 ± 23.1	**< 0.001**
Medical history			
Hypertension	20.4%	17.5%	**< 0.001**
Atrial fibrillation	3.4%	2.3%	**< 0.001**
History of stroke	2.5%	1.4%	**< 0.001**
History of MI or PCI	1.9%	1.5%	**0.011**
Chronic heart failure	3.3%	1.6%	**< 0.001**
Chronic kidney disease	4.2%	2.4%	**< 0.001**
Diabetes mellitus	8.7%	7.2%	**< 0.001**
Chronic obstructive pulmonary disease	10.0%	8.4%	**< 0.001**
Acid-related disorders	16.2%	11.7%	**< 0.001**
Recent history of cancer (≤5 years)	2.9%	2.1%	**< 0.001**
Rheumatoid arthritis	1.2%	0.9%	**0.003**
Treatment			
ACE inhibitors	15.1%	14.5%	0.100
ARBs	8.4%	7.6%	**0.009**
Calcium channel blockers	9.5%	8.8%	0.066
Beta-blockers	14.7%	12.0%	**< 0.001**
Diuretics	13.8%	11.6%	**< 0.001**
Anticoagulants / antithrombotic agents	8.7%	6.1%	**< 0.001**
Statins	14.1%	13.9%	0.701

P-value of binomial test (categorical variables) and P-value of one-sample t-test (continuous variables).
ACE = angiotensin converting enzyme, ARB = angiotensin II receptor blockers, MI = myocardial infarction, PCI = percutaneous coronary intervention

Table S3. Characteristics of COVID-19 patients by using drugs for acid-related disorders

	Using drugs N = 1,205	Not using drugs N = 6,250	P-value
Basic characteristics			
Men	524 (43.5%)	2 957 (47.3%)	**0.015**
Age, mean ± SD	62.6 ± 18.3	44.2 ± 20.2	**< 0.001**
Medical history			
Hypertension	542 (45.0%)	982 (15.7%)	**< 0.001**
Atrial fibrillation	125 (10.4%)	129 (2.1%)	**< 0.001**
History of stroke	91 (7.6%)	95 (1.5%)	**< 0.001**
History of MI or PCI	69 (5.7%)	69 (1.1%)	**< 0.001**
Chronic heart failure	138 (11.5%)	109 (1.7%)	**< 0.001**
Chronic kidney disease	158 (13.1%)	156 (2.5%)	**< 0.001**
Diabetes mellitus	222 (18.4%)	424 (6.8%)	**< 0.001**
Chronic obstructive pulmonary disease	271 (22.5%)	474 (7.6%)	**< 0.001**
Recent history of cancer (≤5 years)	93 (7.7%)	124 (2.0%)	**< 0.001**
Rheumatoid arthritis	39 (3.2%)	50 (0.8%)	**< 0.001**
Treatment			
ACE inhibitors	348 (28.9%)	781 (12.5%)	**< 0.001**
ARBs	217 (18.0%)	411 (6.6%)	**< 0.001**
Calcium channel blockers	253 (21.0%)	452 (7.2%)	**< 0.001**
Beta-blockers	422 (35.0%)	674 (10.8%)	**< 0.001**
Diuretics	399 (33.1%)	629 (10.1%)	**< 0.001**
Anticoagulants / antithrombotic agents	270 (22.4%)	375 (6.0%)	**< 0.001**
Statins	369 (30.6%)	682 (10.9%)	**< 0.001**

P-value of Fisher's exact test (categorical variables) and P-value of Mann-Whitney U test (continuous variables).
ACE = angiotensin-converting enzyme, ARBs = angiotensin II receptor blockers, MI = myocardial infarction, PCI = percutaneous coronary intervention

Figure S1. Time to analysed endpoint (severe condition)

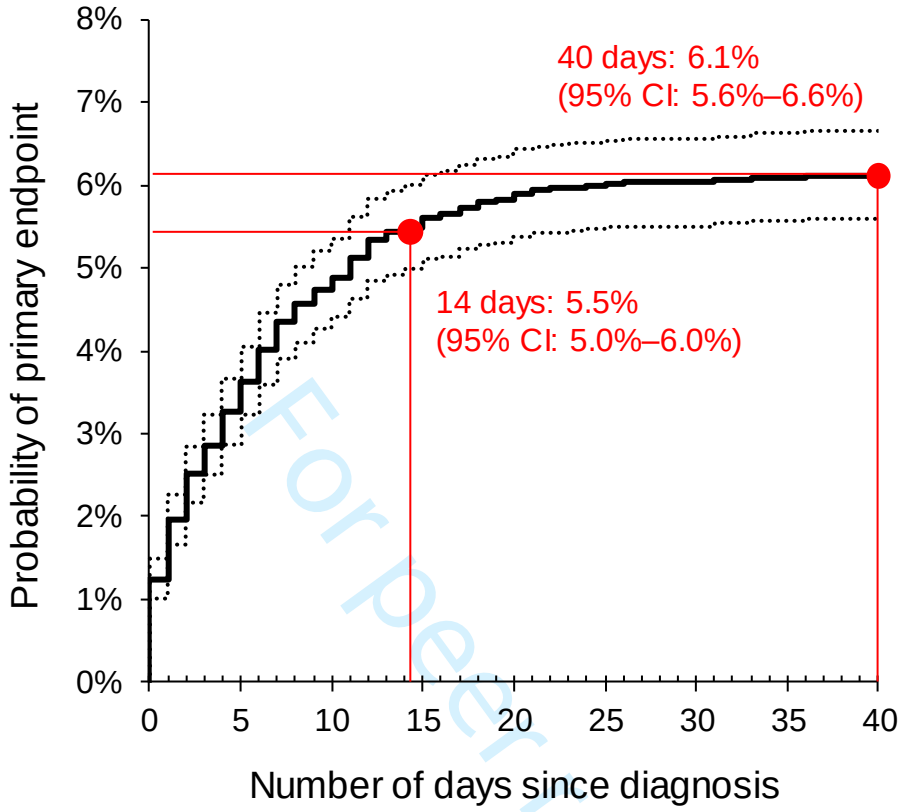
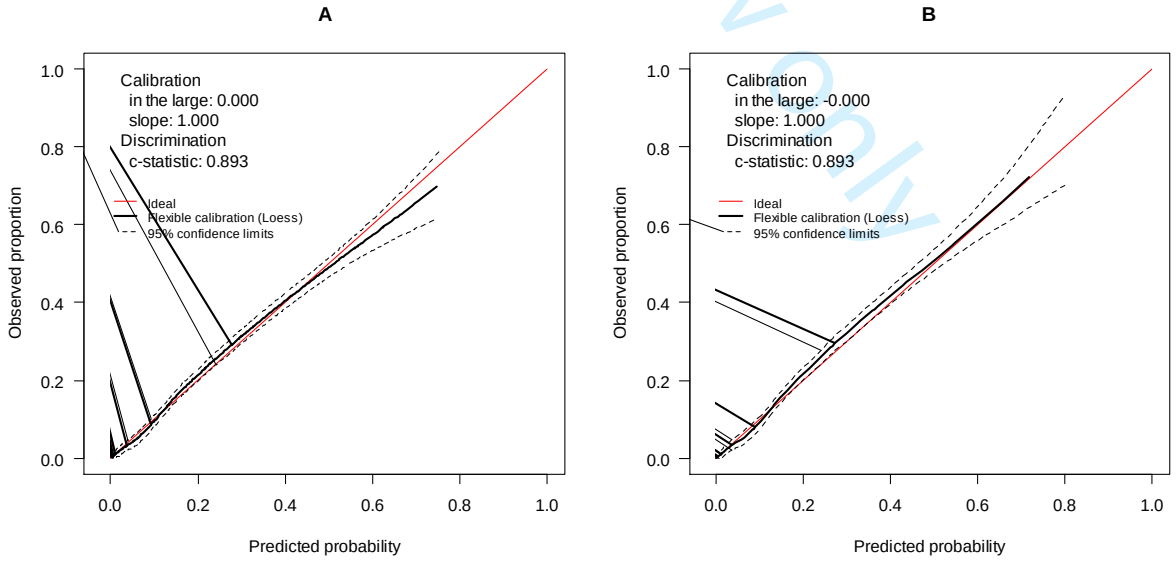


Figure S2. Calibration curves of the multivariable logistic regression model (A) and its simplified version (B)



TRIPOD Checklist: Prediction Model Development and Validation

Section/Topic	Item		Checklist Item	Page
Title and abstract				
Title	1	D;V	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted.	1
Abstract	2	D;V	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	5
Introduction				
Background and objectives	3a	D;V	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	8
	3b	D;V	Specify the objectives, including whether the study describes the development or validation of the model or both.	9
Methods				
Source of data	4a	D;V	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.	10
	4b	D;V	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.	10
Participants	5a	D;V	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.	11
	5b	D;V	Describe eligibility criteria for participants.	10
	5c	D;V	Give details of treatments received, if relevant.	-
Outcome	6a	D;V	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	14
	6b	D;V	Report any actions to blind assessment of the outcome to be predicted.	-
Predictors	7a	D;V	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.	12
	7b	D;V	Report any actions to blind assessment of predictors for the outcome and other predictors.	-
Sample size	8	D;V	Explain how the study size was arrived at.	10
Missing data	9	D;V	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	10
Statistical analysis methods	10a	D	Describe how predictors were handled in the analyses.	12
	10b	D	Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation.	13
	10c	V	For validation, describe how the predictions were calculated.	13
	10d	D;V	Specify all measures used to assess model performance and, if relevant, to compare multiple models.	13
	10e	V	Describe any model updating (e.g., recalibration) arising from the validation, if done.	-
Risk groups	11	D;V	Provide details on how risk groups were created, if done.	-
Development vs. validation	12	V	For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.	-
Results				
Participants	13a	D;V	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	10
	13b	D;V	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.	T1
	13c	V	For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).	-
Model development	14a	D	Specify the number of participants and outcome events in each analysis.	14
	14b	D	If done, report the unadjusted association between each candidate predictor and outcome.	S4
Model specification	15a	D	Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point).	T2, S1
	15b	D	Explain how to the use the prediction model.	F1
Model performance	16	D;V	Report performance measures (with CIs) for the prediction model.	14
Model-updating	17	V	If done, report the results from any model updating (i.e., model specification, model performance).	-
Discussion				
Limitations	18	D;V	Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).	19
Interpretation	19a	V	For validation, discuss the results with reference to performance in the development data, and any other validation data.	-
	19b	D;V	Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.	19
Implications	20	D;V	Discuss the potential clinical use of the model and implications for future research.	20
Other information				
Supplementary information	21	D;V	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	3
Funding	22	D;V	Give the source of funding and the role of the funders for the present study.	2

*Items relevant only to the development of a prediction model are denoted by D, items relating solely to a validation of a prediction model are denoted by V, and items relating to both are denoted D;V. We recommend using the TRIPOD Checklist in conjunction with the TRIPOD Explanation and Elaboration document.