Seek COVER: Development and validation of a personalized risk calculator for COVID-19 outcomes in an international network

Ross D. Williams[1,†] MSc, Aniek F. Markus[1,†] MSc, Cynthia Yang[1] MSc, Talita Duarte Salles[2] PhD, Scott L. DuVall[3] PhD, Thomas Falconer[4] MS, Jitendra Jonnagaddala[5] PhD, Chungsoo Kim[6] PharmD, Dr. Yeunsook Rho[7] PhD, Andrew Williams[8] PhD, Amanda Alberga[9], MPH, Min Ho An[10] MD, María Aragón[2] PhD, Carlos Areia[11] MSc, Edward Burn[2,12] PhD, Young Hwa Choi[13] MD PhD, Iannis Drakos[14] PhD, Maria Tereza Fernandes Abrahão[15] PhD, Sergio Fernández-Bertolín[2] MSc, George Hripcsak[4] MD, Benjamin Skov Kaas-Hansen[16,17] MD, Prasanna L Kandukuri[18] MS, Jan A. Kors PhD[1], Kristin Kostka[19] MPH, Siaw-Teng Liaw[5] MBBS PhD, Kristine E. Lynch, PhD[3], Gerardo Machnicki[20] PhD, Michael E. Matheny[21], MD , Daniel Morales[22] PhD, Fredrik Nyberg[23] MD PhD, Rae Woong Park[24] MD, PhD, Albert Prats-Uribe[12] MPH, Nicole Pratt[25] PhD, Gowtham Rao[26] PhD MD PhD, Christian G. Reich[19] MD PhD, Marcela Rivera[27] PhD, Tom Seinen[1] MSc, Azza Shoaibi[26] MPH PhD, Matthew E Spotnitz[4] MD, Ewout W. Steyerberg[28,29] PhD, Marc A. Suchard[30] MD PhD, Seng Chan You[24] MD, Lin Zhang[31,32] MD PhD, Lili Zhou[18] PhD, Patrick B. Ryan[26] PhD, Daniel Prieto-Alhambra[12] MD PhD, Jenna M. Reps[26,&] PhD, Peter R. Rijnbeek[1,&,*] PhD

[1]Department of Medical Informatics, Erasmus University Medical Center, Rotterdam, The Netherlands

[2]Fundacio Institut Universitari per a la recerca a l'Atencio Primaria de Salut Jordi Gol i Gurina (IDIAPJGol)

[3]Department of Veterans Affairs, University of Utah, Salt Lake City, UT, US,

[4]Department of Biomedical Informatics, Columbia University, New York, NY

[5]School of Public Health and Community Medicine, UNSW Sydney

[6]Department of Biomedical Sciences, Ajou University Graduate School of Medicine, Suwon, Republic of Korea

[7]Department of Bigdata, Health Insurance Review & Assessment Service, Republic of Korea

[8]Tufts Institute for Clinical Research and Health Policy Studies, Boston, MA, 02111, USA

[9]Independent Epidemiologist, OHDSI

[10]So Ahn Public Health Center, Wando County Health Center and Hospital, Wando, Republic of Korea

[11]Nuffield Department of Clinical Neurosciences, University of Oxford

[12]Centre for Statistics in Medicine, NDORMS, University of Oxford

[13]Department of Infectious Diseases, Ajou University School of Medicine, Suwon, Republic of Korea

[14]Center for Surgical Science, Koege, Denmark

[15]Faculty of Medicine, University of Sao Paulo, Sao Paulo, Brazil

[16]Clinical Pharmacology Unit, Zealand University Hospital, Roskilde, Denmark

[17]NNF Centre for Protein Research, University of Copenhagen, Denmark

[18]Abbvie, Chicago, United States

[19]Real World Solutions, IQVIA, Cambridge, MA, United States

[20]Janssen Latin America, Buenos Aires, Argentina

[21]Department of Veterans Affairs, USA; Vanderbilt University, USA

[22]Division of Population Health and Genomics, University of Dundee, UK

[23]School of Public Health and Community Medicine, Institute of Medicine, Sahlgrenska Academy, University of Gothenburg Gothenburg, Sweden

[24]Department of Biomedical Informatics, Ajou University School of Medicine, Suwon, Republic of Korea

[25]Quality Use of Medicines and Pharmacy Research Centre, University of South Australia, Adelaide, Australia

[26]Janssen Research & Development, Titusville, NJ, USA

[27]Bayer Pharmaceuticals, Bayer Hispania, S.L., Barcelona, Spain

[28]Department of Public Health, Erasmus University Medical Center, Rotterdam, The Netherlands

[29]Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, The Netherlands [30]Department of Biostatistics, UCLA Fielding School of Public Health, University of California, Los Angeles, CA, USA

[31]School of Public Health, Peking Union Medical College, Beijing, China;

[32]Melbourne School of Public Health, The University of Melbourne, Victoria, Australia.

†These authors contributed equally as co-first authors.

&These authors contributed equally as co-last authors.

Abstract

Abstract text goes here, justified and in italics and be no more than 250 words. All submission types have the same format and length requirements. Use a minimum of 11pt font when writing all sections of this abstract.

Research Category (please highlight or circle which category best describes your research)

patient-level prediction

Introduction

COVID-19 is causing high mortality worldwide. Quantifying the risk of poor outcomes in COVID-19 infected patients could help develop strategies to shield the vulnerable during de-confinement[1].

Methods

We analyzed a federated network of electronic medical records and administrative claims data from 14 data sources and 6 countries, mapped to a common data model.

Participants

Model development used a patient population consisting of >2 million patients with a general practice (GP), emergency room (ER), or outpatient (OP) visit with diagnosed influenza or flu-like symptoms any time prior to 2020. The model was validated on patients with a confirmed COVID-19 diagnosis across five databases from South Korea, Spain and the United States.

Predictors and outcomes

Age, sex, historical conditions, and drug use prior to index date were considered as candidate predictors. Outcomes included i) hospitalization with pneumonia, ii) hospitalization with pneumonia requiring

intensive services or death, and iii) death in the 30 days after index date.

Results

Overall, 44,507 COVID-19 patients were included for model validation, after initial model development and validation using 6,869,127 patients with influenza or flu-like symptoms. We identified 7 predictors (history of cancer, chronic obstructive pulmonary disease, diabetes, heart disease, hypertension, hyperlipidemia, and kidney disease) which combined with age and sex could discriminate which patients would experience any of our three outcomes. The models achieved high performance in influenza. When transported to COVID-19 cohorts, the AUC ranges were, COVER-H: 0.69-0.81, COVER-I: 0.73-0.91, and COVER-F: 0.72-0.90. Calibration was overall acceptable, with overestimated risk in the most elderly and highest risk strata.

Conclusion

Three 9-predictor models perform well for COVID-19 patients for predicting hospitalization, intensive services and fatality. The models could aid in reassurance of low risk patients and shield high risk patients from COVID-19 during de-confinement.

References

1. Smith George Davey, Spiegelhalter David. Shielding from covid-19 should be stratified by risk *BMJ* 2020; 369 :m2063