Generalizable Prognostic Models for Patient-Centered Decisions in COVID-19: Transportability over Time and Space

NCT ID: *Not yet assigned*

Version date: 12/1/2020

PROTOCOL

ENHANCEMENT TO PCORI CONTRACT ME-1606-35555

ENHANCEMENT PROJECT TITLE: Generalizable Prognostic Models for Patient-Centered Decisions in COVID-19: Transportability over Time and Space

Principal Investigator: David M. Kent, MD[1]

Co-Investigators: Negin Hajizadeh, MD, MPH[2]; Hester Lingsma, PhD[3]; Thomas McGinn, MD[2]; Jessica Paulus, ScD[1]; David van Klaveren, PhD[3]; Theodoros P. Zanos, PhD[2]

[1]Tufts Medical Center, [2]Northwell Health, [3]Erasmus Medical Center

Background

Approximately 20% of patients hospitalized with COVID-19 require intensive care and possibly invasive mechanical ventilation (MV)[1,2]. Patient preferences with COVID-19 for MV may be different, because intubation for these patients is often prolonged (for several weeks), is administered in settings characterized by social isolation and is associated with very high average mortality rates[1,3]. Supporting patients facing this decision requires providing an accurate forecast of their likely outcomes based on their individual characteristics.

A number of COVID-19 clinical prediction models (CPMs) have appeared in the literature, but none developed with high methodological rigor[4]. Colleagues have developed models tailored to their own systems[5], but the validity and generalizability to other settings is unknown, and the variables used are not uniformly and reliably obtained across systems. More generally, most hospitals (and systems) do not have a sufficient number of cases (and outcomes) to develop models fit to their local population. Therefore, pooling data resources and assessing generalizability across different sites is urgently needed.

In addition to examining geographic transportability, since best practices are rapidly evolving over time (e.g., proning, minimizing paralytics, lung-protective volumes, remdesivir, dexamethasone or other treatments), updating and recalibrating these CPMs over time is crucially important. Changes in what patients are admitted to the hospital over time can also have important effects on outcomes and also on the consistency of variable effects[6]. Models developed on abundant first wave data may have little generalizability to the current situation. Thus this project will also focus specifically on temporal validation and updating.

In our current PCORI Methods project, we developed a CPM evaluation and updating framework including both conventional and novel performance measures. We performed independent model validations and updating on over 180 unique CPM-database pairs. We will use this framework to evaluate COVID-19 prognostic models in the largest cohort of COVID-19 patients examined to date, spanning 2 datasets from very different settings. These models will be directly applied to support and improve patient centered decision making for patients potentially requiring MV. Augmenting our original team of internationally recognized modelers with experts in clinical decision science, we will also lay the groundwork for a data-rich COVID modeling collaborative poised to address subsequent challenges. As the COVID-19 pandemic affects different regions, with subsequent waves expected, identifying the most accurate, robust and generalizable prognostic tools is needed to guide patient-centered decision making across diverse populations, settings and phases of the epidemic.

Aims: We therefore aim to:

1) *Develop* 3 CPMs in each of 2 hospital systems (i.e. 6 distinct models) to predict:
    i) the need for MV in patients hospitalized with COVID-19;
    ii) mortality in patients receiving MV;
    iii) length of stay in the ICU.

2) *Evaluate* the geographic and temporal transportability of these models and examine updating approaches.

a. To evaluate geographic transportability, we will apply the evaluation and updating framework developed (in our parent PCORI grant) to assess CPM validity and generalizability across the different datasets.
b. To evaluate temporal transportability, we will examine both the main effect of calendar time and also examine calendar time as an effect modifier.

3) *Engage* stakeholders to facilitate best use of these CPMs in the care of patients with COVID-19.

Approach

Study Populations

We will include consecutive COVID-19-positive patients admitted to the hospital, excluding those who were placed on MV prior to presentation. The sites include: (1) Erasmus MC (a partner in the parent PCORI grant) who is compiling COVID admissions from 5-6 large Dutch public hospitals including at least 4600 COVID-19 patients; (2) The Northwell Health System with *the largest reported COVID cohort in the US*[1], now with over 13,000 COVID-19-positive patients from 13 hospitals in the New York City area. A preliminary CPM has been developed using 6 features on a subset of this cohort (11,095 patients)[7].

Aim 1: Model development

Three models will be developed at each of 2 sites to: 1) predict the probability of requiring MV; 2) predict mortality *if* placed on MV; 3) predict the length of stay in the ICU with a survival model. We will use variables that are reliably obtainable across sites, starting with a broad set including: patient characteristics (e.g., age, BMI, smoking, co-morbidities), clinical parameters (e.g., O_2 saturation, blood pressure, HR, RR), and lab values (e.g., CRP, LDH, lymphocyte count, d-dimer, high-sensitivity troponin, ferritin, lactate, hemoglobin and albumin). The models will be developed "semi-independently"—i.e., coordinated to facilitate cross-system testing, as a main finding from the parent PCORI grant was the ubiquity of model-database incompatibility for validation.

The Northwell Health site will use LASSO regression with L1-norm regularization to identify the subset of EHR measurements that, when linearly combined, best predict the primary outcomes[8]. The Erasmus MC (Dutch site) will use a multi-state regression model[9,10] to predict health state transitions as a function of an individual participant's baseline characteristics including times to intubation and death (including discharge to hospice) among intubated patients, with variable selection based on likelihood ratio test statistics. The models will initially be cross-validated across hospitals within the datasets. Across sites, missingness will be accommodated by missing indicator variables or via multiple imputation (MI). The advantages and disadvantages of these two approaches will be explored.

As of this writing, there are 34 models predicting death or progression to severe disease.[4] Table 1 shows the 17 variables included in at least 3 of these models. Of these the following variables will not be considered because they are not well ascertained in one of the two sites: CT features and the presence of comorbidities (e.g., hypertension, diabetes, cardiovascular disease, chronic respiratory disease, cancer). Interaction terms will be considered on the basis of clinical reasoning. In general, predictors in both datasets were ascertained from routinely collected data from the electronic health records. Continuous variables will be modeled continuously. Because some variables are available at the Northwell site and not the Erasmus site (including predictor variables collected after hospital admission within 48 hours of intensive care unit (ICU) admission, not available at Erasmus), we will consider an unrestricted Northwell model using a larger array of variables and a more restricted Erasmus-compatible model that can be externally validated.

Table 1. Appearance of Variables in COVID-19 Models.

Variable	Number of models

Age	25
CRP	14
Presence of any comorbidity (e.g., hypertension, diabetes, cardiovascular disease, chronic respiratory disease, cancer)	13
Sex	12
Lymphocyte	10
LDH	7
CT features	6
Respiratory rate	5
Neutrophil	5
Platelets	5
Albumin	5
Creatinine	5
Heart rate	4
Body temperature	4
RDW	4
SPO_2	3
D-dimer	3

Aim 2: Model evaluation

To evaluate the geographic transportability of the models, the Tufts team will incorporate these models into their PCORI-supported evaluation framework to externally validate each model on the external datasets. This framework includes both conventional and novel measures and was tested across 180 model-database pairs in the parent project. Models will be assessed for discrimination (c-statistic), calibration (plots, Harrell's E) and net benefit (using decision curve analysis). Changes in discrimination will be evaluated by examining model-based c-statistics[11] – developed during the parent award– to assess the degree to which any decrement is due to model invalidity versus case mix differences. As models are intended to be used to align decisions with patients' own values and preferences, measures of calibration (and net benefit) will be prioritized. We will also assess performance on external databases after updating procedures (i.e., updating of intercepts; of intercepts and slope and finally parameter re-estimation). This provides critical information on CPM robustness, how they generalize and how best to update models over time and in new settings. In addition, we will compare the performance of our COVID-specific CPMs to promising models in the literature and to generic (non-COVID) prediction models[12 13].

To examine the temporal transportability, we will test the effects of calendar time within each site (i.e., within the Dutch data and the Northwell data separately). This will be performed on data at both sites obtained prior to September 2020 (i.e., first wave data). The effect of secular trends by including calendar time as a continuous variable in a two variable model that also includes the previously-derived linear predictor. Nonlinear time effects will be examined with splines. In addition, since patient selection and new treatments may influence the consistency of variable effects, we will explore interactions between calendar time and each variable in a refit model. We anticipate obtaining second wave data (i.e., after September 2020) from both sites that can be used to test the models. Models will be tested by fixing calendar time to the most proximal time available in the development data set.

Aim 3: Stakeholder engagement

We will convene two multi-stakeholder panels of 6-8 members each. The first panel will include COVID-

19 survivors; family members of COVID-19 patients; and caregivers for COVID-19 patients. The second panel will include critical care physicians; palliative care physicians; hospitalists; nurses; respiratory therapists; leaders of our clinical ethics committees and pastoral care representatives. The goal of these discussions is to develop a full understanding of how accurate prognostic models mighit best support patients and clinicians in making these critical patient-centered decisions. The panel will be convened at regular intervals to elicit stakeholder feedback at each phase. We will leverage our experience using online collaborative tools to maximize feedback opportunities from each member.

REFERENCES CITED

Do not exceed 10 pages.

Bibliography

1. Richardson S, Hirsch JS, Narasimhan M, et al. Presenting Characteristics, Comorbidities, and Outcomes Among 5700 Patients Hospitalized With COVID-19 in the New York City Area. *JAMA.* 2020.
2. Zhou F, Yu T, Du R, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet (London, England).* 2020;395(10229):1054-1062.
3. Richardson S, Hirsch JS, Narasimhan M, et al. Clarification of Mortality Rate and Data in Abstract, Results, and Table 2. *JAMA.* 2020.
4. Wynants L, Van Calster B, Bonten MMJ, et al. Prediction models for diagnosis and prognosis of covid-19 infection: systematic review and critical appraisal. *BMJ (Clinical research ed).* 2020;369:m1328.
5. Levy TJ, Richardson S, Coppa K, et al. Development and Validation of a Survival Calculator for Hospitalized Patients with COVID-19. *medRxiv.* 2020:2020.2004.2022.20075416.
6. Griffith GJ, Morris TT, Tudball MJ, et al. Collider bias undermines our understanding of COVID-19 disease risk and severity. *Nat Commun.* 2020;11(1):5749.
7. Levy TJ, Richardson S, Coppa K, et al. Estimating Survival of Hospitalized COVID-19 Patients from Admission Information. *[Posted to medRxiv].* 2020.
8. Tibshirani R. Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society Series B (Methodological).* 1996;58(1):267-288.
9. Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: competing risks and multi-state models. *Statistics in medicine.* 2007;26(11):2389-2430.
10. de Wreede LC, Fiocco M, Putter H. The mstate package for estimation and prediction in non- and semi-parametric multi-state and competing risks models. *Computer methods and programs in biomedicine.* 2010;99(3):261-274.
11. van Klaveren D, Steyerberg EW, Gönen M, Vergouwe Y. The calibrated model-based concordance improved assessment of discriminative ability in patient clusters of limited sample size. *Diagn Progn Res.* 2019;3(11).
12. Jones AE, Trzeciak S, Kline JA. The Sequential Organ Failure Assessment score for predicting outcome in patients with severe sepsis and evidence of hypoperfusion at the time of emergency department presentation*. *Crit Care Med.* 2009;37(5):1649-1654.
13. Lim WS, van der Eerden MM, Laing R, et al. Defining community acquired pneumonia severity on presentation to hospital: an international derivation and validation study. *Thorax.* 2003;58(5):377-382.