

## Research and Applications

# Development and validation of a machine learning model predicting illness trajectory and hospital utilization of COVID-19 patients: A nationwide study

Michael Roimi<sup>1,\*</sup>, Rom Gutman<sup>2,\*</sup>, Jonathan Somer<sup>2</sup>, Asaf Ben Arie<sup>3</sup>, Ido Calman, Yaron Bar-Lavie<sup>1</sup>, Udi Gelbshtein<sup>4</sup>, Sigal Liverant-Taub<sup>4</sup>, Arnona Ziv<sup>5</sup>, Danny Eytan<sup>6,7</sup>, Malka Gorfine<sup>3,†</sup>, and Uri Shalit<sup>2,†</sup>

<sup>1</sup>Department of Critical Care Medicine, Rambam Health Care Campus, Haifa, Israel, <sup>2</sup>Faculty of Industrial Engineering and Management, Technion – Israel Institute of Technology, Haifa, Israel, <sup>3</sup>Department of Statistics and Operations Research, Tel Aviv University, Tel Aviv, Israel, <sup>4</sup>Division of General Medicine, Medical Directorate, Israel Ministry of Health, Jerusalem, Israel, <sup>5</sup>Gertner Institute for Epidemiology and Health Policy Research, Chaim Sheba Medical Center, Tel HaShomer, Israel, <sup>6</sup>Pediatric Intensive Care, Rambam Health Care Campus, Haifa, Israel, and <sup>7</sup>Faculty of Medicine, Technion – Israel Institute of Technology, Haifa, Israel

\*Joint first authors.

†Joint last authors.

\*Corresponding Author: Malka Gorfine, PhD, Department of Statistics and Operations Research, Tel Aviv University Ramat Aviv, Tel Aviv, 6997801, Israel; gorfine@tauex.tau.ac.il. OR Uri Shalit, PhD, Technion—Israel Institute of Technology, Kiryat HaTechnion, Haifa, 3200003, Israel; urishalit@technion.ac.il

Received 18 November 2020; Editorial Decision 4 January 2021; Accepted 12 January 2021

### ABSTRACT

**Objective:** The spread of coronavirus disease 2019 (COVID-19) has led to severe strain on hospital capacity in many countries. We aim to develop a model helping planners assess expected COVID-19 hospital resource utilization based on individual patient characteristics.

**Materials and Methods:** We develop a model of patient clinical course based on an advanced multistate survival model. The model predicts the patient's disease course in terms of clinical states—critical, severe, or moderate. The model also predicts hospital utilization on the level of entire hospitals or healthcare systems. We cross-validated the model using a nationwide registry following the day-by-day clinical status of all hospitalized COVID-19 patients in Israel from March 1 to May 2, 2020 ( $n = 2703$ ).

**Results:** Per-day mean absolute errors for predicted total and critical care hospital bed utilization were  $4.72 \pm 1.07$  and  $1.68 \pm 0.40$ , respectively, over cohorts of 330 hospitalized patients; areas under the curve for prediction of critical illness and in-hospital mortality were  $0.88 \pm 0.04$  and  $0.96 \pm 0.04$ , respectively. We further present the impact of patient influx scenarios on day-by-day healthcare system utilization. We provide an accompanying R software package.

**Discussion:** The proposed model accurately predicts total and critical care hospital utilization. The model enables evaluating impacts of patient influx scenarios on utilization, accounting for the state of currently hospitalized patients and characteristics of incoming patients. We show that accurate hospital load predictions were possible using only a patient's age, sex, and day-by-day clinical state (critical, severe, or moderate).

**Conclusions:** The multistate model we develop is a powerful tool for predicting individual-level patient outcomes and hospital-level utilization.

**Key words:** COVID-19, hospital utilization, survival analysis, healthcare facilities, multistate model

## INTRODUCTION

The coronavirus disease 2019 (COVID-19) pandemic is taking its toll on healthcare systems around the world, with some patients requiring lengthy general and intensive care.<sup>1-4</sup>

Given the danger of unprecedented burden on healthcare systems due to COVID-19, there is a need for tools helping decision makers plan resource allocation on the unit, hospital and national levels. COVID-19 hospitalization times are often lengthy and can vary substantially among patients.<sup>5</sup> Therefore, in this study, we aimed to develop a model of hospitalization trajectories of COVID-19 patients, in order to accurately predict the number of hospitalized and critical care patients. The predictions are based on the clinical state of each of the currently hospitalized patients, and projections of hospital patient influx.

We validated the model using a nationwide hospitalization registry, which includes the day-by-day hospitalization record of all confirmed COVID-19 patients in Israel between March 1 and May 2, 2020, totaling 2703 patients (see Table 1).

To facilitate use of the model, we provide an R (version 3.5.1; R Foundation for Statistical Computing, Vienna, Austria)<sup>6</sup> software package (<https://github.com/JonathanSomer/covid-19-multi-state-model>), enabling anyone with similar data to develop a model tailored to specific patient and healthcare system characteristics, and a Web application (<https://covid19-hospitalcourse.net/>), taking as input the patient characteristics and predicting the probabilities of different disease courses. Finally, we share an anonymized version of the dataset used to develop the tool.

## MATERIALS AND METHODS

The main idea behind our proposed model is tracking the manner in which hospitalized COVID-19 patients transitions between different

clinical states. Specifically, we assume a hospitalized patient is in 1 of 3 clinical states: moderate, severe, or critical; the exact definition of the states for the validation cohort was given by the Israeli Ministry of Health (MOH) and is detailed subsequently.

As an example, a patient might be hospitalized in a severe state, deteriorate into a critical state after 5 days, spend 10 days in a critical state, and then recover and spend 3 days in a severe state and 2 more days in a moderate state before being discharged from the hospital. In Table 2, we show the distribution of all transitions observed in the Israeli COVID-19 registry, including patients dying and patients being discharged from hospital (here, we merged moderate and severe, as explained subsequently). In general, the transition process is non-Markovian, and observations are often right censored or left truncated. We therefore developed a multistate model that can account for all these properties.

## Outcomes

The primary outcome was prediction of total and critical care bed occupancy on a calendar scale, where future occupancy is due to currently hospitalized patients staying in the hospital, and due to newly arriving patients. We predict occupancy by predicting for each patient their day-by-day clinical state, including days in which the patient is in a critical state, discharged, or possibly died. We further use the day-by-day clinical state predictions for predicting the risk for a single patient of entering a critical state at some point throughout hospitalization, risk of in-hospital mortality, expected hospital length of stay (LOS) and expected LOS in critical state.

## Statistical analysis methods

We modeled the way patients move between different clinical states over time by a multistate Cox regression-based survival analysis with right censoring, competing events, recurrent events, left truncation, and time-dependent covariates.<sup>7-10</sup> The multistate model has 4

**Table 1.** Demographics and clinical characteristics of patients in the Israeli COVID-19 registry who were hospitalized between March 1 and May 2

Characteristic	Total	Critical by May 2	In-Hospital Mortality by May 2	Hospitalized on May 2
<b>Patients</b>	2675 (100) <sup>a</sup>	437 (16.34)	200 (7.48)	311 (11.63)
<b>Female</b>	1171 (43.78)	146 (33.41)	89 (44.5)	130 (42)
<b>Age, y</b>	55.3 ± 21.7	71 ± 16.35	80.66 ± 12.78	65.5 ± 20.01
<b>Age</b>				
<20 y	106 (3.96)	3 (0.69)	0 (0)	8 (2.57)
20-29 y	316 (11.81)	4 (0.92)	0 (0)	15 (4.82)
30-39 y	272 (10.17)	14 (3.2)	2 (1)	20 (6.43)
40-49 y	330 (12.34)	19 (4.35)	3 (1.5)	15 (4.82)
50-59 y	401 (15)	57 (13.04)	6 (3)	36 (11.58)
60-69 y	458 (17.12)	79 (18.08)	20 (10)	56 (18)
70-79 y	412 (15.4)	118 (27)	50 (25)	80 (25.72)
80+ y	380 (14.21)	143 (32.72)	119 (59.5)	81 (26.05)
<b>Initial state</b>				
Moderate	2048 (76.56)	113 (25.8)	50 (25)	164 (52.73)
Severe	432 (16.14)	129 (29.5)	66 (33)	83 (26.69)
Critical	195 (7.29)	195 (44.6)	84 (42)	64 (20.58)

Values are n (%) or mean ± SD.

COVID-19: coronavirus disease 2019.

<sup>a</sup>Patients who were hospitalized at least 1 day. Excluded 28 patients with missing age or sex information.

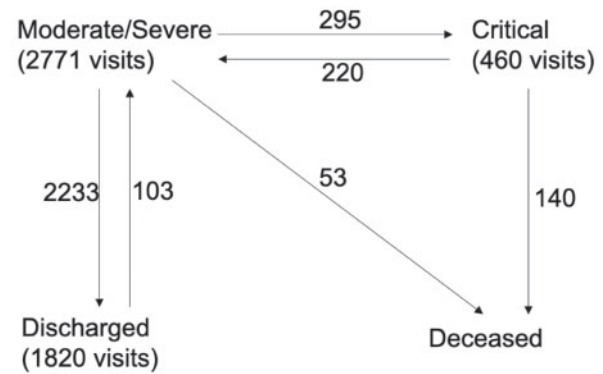
**Table 2.** Summary of the observed hospitalization course (observed paths):

	Path	Frequency
1	M/S	148
2	M/S Di	1977
3	M/S Di M/S	19
4	M/S Di M/S Di	68
5	M/S Di M/S Di M/S	1
6	M/S Di M/S Di M/S Di	5
7	M/S Di M/S Di M/S Di M/S Di	1
8	M/S Di M/S C	2
9	M/S Di M/S De	2
10	M/S Di C	1
11	M/S Di C M/S Di	1
12	M/S C	49
13	M/S C Di	4
14	M/S C M/S	25
15	M/S C M/S Di	61
16	M/S C M/S Di M/S Di	1
17	M/S C M/S C	8
18	M/S C M/S C M/S	4
19	M/S C M/S C M/S Di	13
20	M/S C M/S C M/S Di M/S	1
21	M/S C M/S C M/S C	1
22	M/S C M/S C M/S C M/S	1
23	M/S C M/S C M/S C M/S C	1
24	M/S C M/S C M/S C De	1
25	M/S C M/S C De	3
26	M/S C M/S De	2
27	M/S C De	64
28	M/S De	44
29	C	42
30	C Di	6
31	C M/S	12
32	C M/S Di	33
33	C M/S Di M/S	1
34	C M/S C	3
35	C M/S C M/S	2
36	C M/S C M/S Di	6
37	C M/S C M/S C	1
38	C M/S C M/S C M/S	2
39	C M/S C M/S C M/S Di	2
40	C M/S C M/S C M/S C	1
41	C M/S C M/S C De	2
42	C M/S C M/S De	1
43	C M/S C De	3
44	C M/S De	4
45	C De	74

A patient enters the hospital at a moderate, severe, or critical clinical state and can move among the transient clinical states during the course of hospitalization. The longest observed path consists of 9 transitions.

C: critical; De: deceased; Di: discharged; M/S: moderate/severe.

states, following the clinical states defined by the Israeli MOH detailed subsequently in the Model Validation section: (1) moderate or severe, (2) critical, (3) discharged, and (4) deceased. We merged the moderate and severe clinical states into a single model state due to sample size considerations. This multistate model consists of 6 Cox regression models, 1 for each possible state-to-state transition, shown in Figure 1; some transitions were excluded as either clinically implausible or due to few observed transitions. Details can be found in Supplementary Appendix S1.1.



**Figure 1.** We model a COVID-19 (coronavirus disease 2019) patient’s disease course as moving between 4 possible states: (1) moderate or severe, (2) critical, (3) discharged, and (4) deceased. We combined the 2 clinical states moderate and severe into a single model state due to statistical considerations; however, we emphasize that we keep a distinction between the 2 by a covariate indicating whether the patient first entered at mild/moderate clinical state or at a severe clinical state. Numbers next to arrows indicate number of observed transitions; each patient can make several state transitions, and may visit a transient state more than once.

The 6 semiparametric models each include a set of covariates, possibly with time-dependent covariates and different covariates for each model. In our data analysis with the Israeli cohort, we took in age, sex, and state at hospitalization as baseline covariates; for the latter, we kept a distinction between moderate and severe clinical states. We also added time-dependent covariates encoding the hospitalization history of the patient: cumulative days in hospital, and whether the patient had been in critical state before (see Supplementary Appendix S1.1).

Estimation of the 6 models involves several major issues, beyond right censoring: (1) multistate process, in which each patient may visit the states moderate or severe, critical, and discharged multiple times; (2) left truncation, in which a patient entering a new state after  $m$  days since hospitalization is left truncated by  $m$ ; (3) competing risks, in which in case of multiple possible transitions from a certain state, the occurrence of one type of transition at a certain point in time prevents the occurrence of other transitions at that time point; and (4) recurrent events, in which patients may visit states moderate or severe, critical, and discharged multiple times. We overcome all the previous challenges and provide consistent estimators of the 6 models (see Supplementary Appendix S1.2 for details).

Making predictions based on our proposed multistate model requires estimating the absolute risks, also known as the cumulative incidence functions. The absolute risks involve estimating the probabilities of moving between states, the time to be spent at each state and integrating over all possible combinations between any possible triplet of entry state, exit state and LOS. Because hospitalization consists of potentially multiple transitions between transient states, the absolute risks have no tractable analytic form. Thus, we performed Monte Carlo (MC) sampling from the multistate model, in order to obtain consistent predictions for individual patient and for cohorts of current and future patients. Each MC sample for a given patient consists of sampling a disease course, in terms of clinical states over time and how much time is spent in each one, conditioned on the patient’s history and covariates. MC sampling is the

keystone of our prediction tools (see [Supplementary Appendix S1.3](#) to [S1.5](#) for details including a description of path sampling).

Using this model, patient-level prediction is based on taking summary statistics of 20 000 MC paths for the patient based on the patient's covariates. Summary statistics include mean, median, and other quantiles of LOS and of LOS at critical state for a given patient.

A somewhat modified approach is required for predicting hospital load. The predictions on the cohort level are given for a set of patients consisting of (1) currently hospitalized patients and (2) patients coming in according to an arrival process of patients expected to be hospitalized in the next days. In short, 1 MC path is sampled for each patient based on a calendar time scale. Summary statistics are calculated over the paths of all patients during a prespecified period. Finally, these 2 steps are repeated 10 000 times and the final predictions are based on summaries of these repeats (see [Supplementary Appendix S1.6](#) and [S1.8](#) for technical details). Standard errors were obtained by weighted bootstrap (see [Supplementary Appendix Section S1.7](#)).

## Model validation: Hospital resource utilization and individual patient disease course prediction

### Dataset

We validate our model using the Israeli Ministry of Health COVID-19 hospitalized patient registry. The registry includes the patients' age and sex, dates and results of their SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2) polymerase chain reaction tests, dates of hospital admissions and discharge, daily clinical status during admission (moderate, severe, or critical, as detailed subsequently), and the death registry. COVID-19 confirmed diagnosis was defined as patients who had a positive SARS-CoV-2 polymerase chain reaction test.<sup>11</sup>

We included in the analyses all patients who were admitted between March 1 and May 2, 2020, and were hospitalized for at least 1 day. We excluded patients missing age or sex documentation. No data imputation was performed.

The first patient with COVID-19 in Israel was diagnosed on February 27, 2020. As of May 2, 2020, 16 137 patients had confirmed positive diagnosis. The median age of confirmed patients was 33 (interquartile range, 21-54) years; 44.5% were women. Of these 16 137 confirmed COVID-19 patients, 2703 (17%) were hospitalized by May 2 for at least 1 full day. Of these 2703 hospitalized patients, 28 had no documented age or sex covariates. For the remaining 2675 patients, median age was 58 (interquartile range, 39-73) years and 44.19% were female. The demographics and clinical status of hospitalized patients are shown in [Table 1](#).

Patients' clinical state during admission (moderate, severe, and critical) was defined by the Israeli MOH guidelines, which are related to National Institute for Health treatment guidelines.<sup>12</sup> A mild or moderate clinical state was defined as a patient with symptoms such as fever, cough, sore throat, malaise, headache, or muscle pain, without shortness of breath, dyspnea on exertion, or abnormal imaging; or as patients with clinical or imaging evidence for lower respiratory disease and oxygen saturation (SpO<sub>2</sub>) >90% on room air. For brevity, we refer to this state as moderate throughout this work. A severe clinical state was defined as a patient with respiratory rate higher than 30 breaths/min, SpO<sub>2</sub> <90% on room air, or ratio of the arterial partial pressure of oxygen to fraction of inspired oxygen (PaO<sub>2</sub>/FiO<sub>2</sub>) <300 mm Hg. A critical clinical state was defined as a state in which the patient suffers from respiratory failure, which

requires invasive or noninvasive mechanical ventilation, septic shock, or multiorgan dysfunction. In addition, we denote patients discharged from the hospital to their home or to out-of-hospital quarantine as discharged. We note that discharged patients might be readmitted upon deterioration, as can be seen in [Table 2](#).

In our dataset the patients' state was categorized by the Israeli MOH clinical progression scale which was in place in the first half of 2020. In June 2020 the World Health Organization published the WHO Clinical Progression Scale, which has 10 different stages, among them 6 for hospitalized patients.<sup>13</sup> The correspondence between the 2 scales is roughly as follows: the Israeli MOH mild and moderate clinical states correspond to WHO score 4, a severe clinical state corresponds to WHO score 5, and a critical clinical state corresponds to WHO scores 6 to 9.

### Evaluation method

We employed 8-fold cross-validation—fitting the model on seven-eighths of the data and evaluating performance on the remaining held-out one-eighth, repeated 8 times. Each held-out set consisted of 329 to 331 patients.

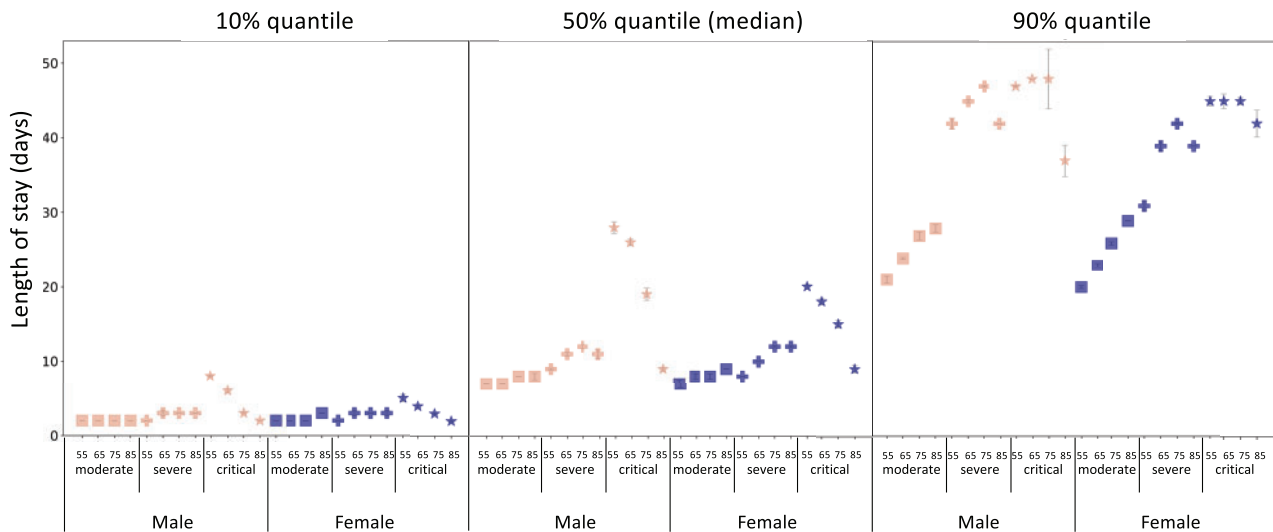
We validated predictions of hospital utilization by 2 methods: snapshot, in which we define a start date and predict future resource utilization for the set of all patients who were hospitalized on that date, without taking into account future incoming patients; and arrival process, in which we use the known hospitalization dates and characteristics of incoming patients between March 1 and May 2, 2020, to estimate utilization for the entire course of the “first wave” of COVID-19 in Israel (see [Supplementary Appendix S1.3](#) for description of both). For both validation methods, we estimate the mean absolute error (MAE) between the model's per-day utilization predictions and the actual number of hospitalized (or critical) patients on that day; the mean is over the number of days in the prediction window (see [Supplementary Appendix S1.3](#)).

We further validated the model's performance by testing its predictions for individual patients: we used data from the first day of a patient's admission to predict their probability of becoming critically ill (ie, in critical state) and probability of in-hospital mortality. For both, we report area under the receiver-operating characteristic curve (AUROC) with inverse weighting correction for censoring (see [Supplementary Appendix S1.3](#)). Finally, we validated the calibration of our predictions by tracking expected number of deaths vs actual number of deaths over time in an arrival+snapshot scenario.

## Using the model for prediction of hospital utilization under hypothetical scenarios

In order to illustrate how our model may be employed for utilization prediction, we focus on a single held-out cohort of 330 patients chosen at random. For this cohort, we predict total future hospital bed and critical hospital bed utilization up to 49 days ahead, starting from March 15, 2020. Utilization for this cohort is composed of patients among the 330 who were hospitalized at the starting date and remain at the hospital, as well as utilization by patients arriving after March 15.

We present the expected hospital utilization and number of deaths under 3 putative patient arrival scenarios: (1) younger—rate and state of incoming patients are the same as in Israel during the weeks from March 15 to May 2, but all patients ages 60+ are replaced with patients in their 40s and 50s; (2) milder—rate and age of incoming patients are the same as in Israel during the weeks from March 15 to May 2, but all incoming patients are in moderate or se-



**Figure 2.** Model estimates of quantiles of length of stay in days based on 20 000 Monte Carlo samples for each patient type. Error bars calculated by weighted bootstrap.

vere clinical state upon hospitalization, none at critical; and (3) eldercare nursing home outbreak, in which we assume that in addition to the arrival of patients as happened in Israel from March 15 to May 2 that there is a single week during which there are 4 times as many incoming patients 70+ years of age, arriving in various clinical states. Details of the scenarios are given in [Supplementary Appendix S2.4](#).

## RESULTS

### Hospitalized patient characteristics

We use the model to estimate the median, 10% and 90% quantiles of LOS for hospitalized patients, stratified by clinical state at time of admission. Results are shown in [Figure 2](#) and [Supplementary Table S7](#); in [Supplementary Table S8](#), we report expected LOS in critical state results.

[Table 3](#) presents the probability, estimated by the model, of patients entering a critical state and probability of in-hospital mortality, both stratified by age, sex, and clinical state at time of hospital admission. Both probabilities sharply increase with age; males of all ages tend to have a greater probability of becoming critically ill compared with females entering the hospital at the same clinical state, but hospitalized females over 75 years of age tend to have higher risk of mortality compared with hospitalized males entering at comparable age and clinical state. Results for younger ages are given in [Supplementary Tables S5 to S8](#), and the cumulative distribution function of LOS is in [Supplementary Figure S2](#).

### Cox models

The results for all 6 Cox models are given in [Supplementary Tables S2 to S4](#).

### Model validation

The results are all averaged over the 8 held-out validation cohorts, each including between 329 and 331 patients. Using snapshot evaluation with April 1 as start date, MAE for predicting the per-day number of hospitalized patients is  $3.15 \pm 1.20$  for total hospital bed utilization and  $1.47 \pm 0.56$  for critical care bed utilization. Us-

ing snapshot evaluation with April 15 start data, MAE predicting the per-day number of hospitalized patients is  $3.13 \pm 1.07$  for total hospital bed utilization and  $1.98 \pm 0.93$  for critical hospital bed utilization. Using arrival evaluation, MAE for predicting the per-day number of hospitalized patients is  $4.72 \pm 1.07$  for total hospital bed utilization and  $1.68 \pm 0.40$  for critical hospital bed utilization. See [Supplementary Table S9](#) and [Supplementary Figures S3 and S4](#) for further results.

Using only information from the first day of a patient's hospitalization, the AUROC for predicting in-hospital mortality and for predicting becoming critically ill (among patients who were not critically ill on their first day of admission) were  $0.96 \pm 0.04$  and  $0.88 \pm 0.04$ , respectively (see [Supplementary Table S10](#)). [Figure 3](#) (bottom) presents the number of deaths predicted by our model under the true patient influx process (Expected column), which matches very closely the observed number of deaths, showing the model is well calibrated.

### Predicting hospital bed utilization

In [Figure 3](#) (top), we show an example of utilization and mortality projections generated by the model under hypothetical scenarios. For example, our model can help planners assess when a new COVID-19 ward will need to open: Assuming that each COVID-19 regular and critical care wards can care for 30 and 15 patients, respectively, we show that the error our model makes in predicting the exact timing when total hospital bed utilization will hit such capacity thresholds is at most 1 day for total utilization, and 3 days for critical care bed utilization (see [Supplementary Figure S3](#)).

## DISCUSSION

One of the distinctive characteristics of COVID-19 is the way health systems are overwhelmed by a large number of patients.<sup>2-4,14</sup> Here, we report the development and validation of a flexible multistate survival analysis model of patient clinical course throughout admission, discharge, and possibly death. We applied our model to the complete set of COVID-19 patients in Israel, tracked day by day from March to May 2020.

**Table 3.** Probability of death and probability of becoming critical stratified by age and gender

Incoming State, Age	Probability of In-Hospital Mortality (%)		Probability of Becoming Critical (%)	
	Men	Women	Men	Women
Moderate, 55 y	0.65a (0.55-0.75)	1.2 (1-1.4)	5.4 (5.2-5.7)	4.1 (3.9-4.4)
Moderate, 65 y	2.1 (1.9-2.3)	2.4 (2.2-2.7)	8.6 (8.2-9)	6.4 (6.1-6.8)
Moderate, 75 y	5.6 (5.3-5.8)	4.7 (4.5-4.9)	12.5 (12.1-13)	9.6 (9-10.2)
Moderate, 85 y	14.7 (14.3-15.1)	11.7 (10.6-12.7)	17.8 (17.3-18.3)	13.5 (12.7-14.4)
Severe, 55 y	3.8 (3.5-4.1)	6.9 (6.5-7.2)	23.7 (22.9-24.5)	18.5 (18.1-19)
Severe, 65 y	9.7 (9.4-10)	11.6 (11.1-12)	32.1 (31.5-32.8)	25.3 (24.3-26.2)
Severe, 75 y	20.7 (20.2-21.3)	20.1 (19.5-20.7)	40.4 (39.5-41.2)	31.9 (30.2-33.5)
Severe, 85 y	43.2 (42.1-44.4)	37.6 (36.3-38.8)	47.3 (45.9-48.7)	39.3 (36.9-41.7)
Critical, 55 y	13.9 (12-15.8)	28.2 (27.6-28.8)	100	100
Critical, 65 y	30.3 (27.2-33.4)	40.5 (39-42)	100	100
Critical, 75 y	55.1 (51.4-58.7)	54.7 (51.3-58.2)	100	100
Critical, 85 y	82.6 (80.5-84.6)	74.6 (70.1-79)	100	100

Probabilities are based on Monte Carlo results, with weighted bootstrap 95% confidence interval.

We show that using simple and easily available patient characteristics, the multistate model we developed accurately predicts healthcare utilization for a given patient arrival process and can be used to simulate utilization under different patient influx scenarios. This can, in turn, be used to accurately plan resource allocation and the opening or closing of COVID-19 wards. We further provide an anonymized version of the dataset used to develop the model, a Web application for patient-level predictions, and an R software package to help planners fit a multistate model to their own data, or use the model we fit to the Israeli data.

Interestingly, on the one hand, we find that scenarios such as the arriving patients being much younger or in milder clinical state do not greatly affect total hospital utilization, possibly because some of these populations have longer hospitalization times; on the other hand, both scenarios affect critical care bed utilization. We further observe that an eldercare nursing home outbreak scenario leads to substantially higher total utilization and critical care utilization, underscoring the need to protect these communities not only in terms of preventing mortality, but also from the point of view of lowering the strain on hospital resources.

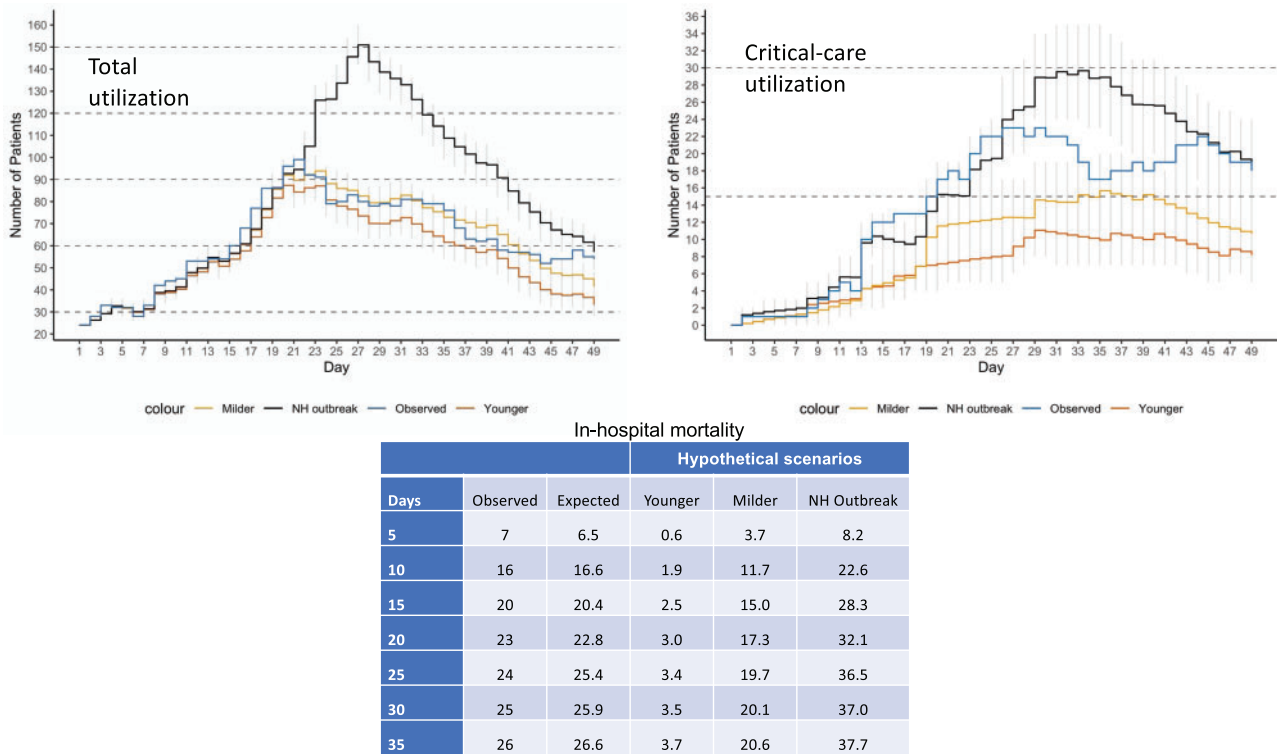
Many models exist for predicting the dynamics of COVID-19 case numbers and numbers of hospitalized patients. These models are usually based on extensions of the susceptible-infected-recovered model,<sup>15,16</sup> in which the number of hospitalized or critical patients are included as a component in the dynamic model.

Our model differs from these models in several aspects. First, settings in which the chance of experiencing one event is altered by the occurrence of other events are known as competing and semi-competing risks, and caution is needed in analyzing such data.<sup>10,17</sup> In this work, we use a multistate model as an excellent fit for the competing and semi-competing risks data of COVID-19 patient's

hospitalization course. Second, heterogeneity within and between patients matters, as overall bed utilization is partially determined by a long tail of some patients who require significantly longer stays than others. We model patients on an individual basis, taking into account how long each patient has already been in the hospital. Thus, we account for the case mix and heterogeneous histories of the patients currently hospitalized when making predictions about future utilization. Third, our model is different in scope: we do not aim to model the spread of the disease and the number of future infections. We focus on estimating hospital utilization under different patient arrival processes, while taking into account the load caused by currently hospitalized patients.

Recently, both Hazard et al<sup>18</sup> and Schmidt et al<sup>19</sup> proposed using a multistate model similar to ours, focusing only on patients admitted to the intensive care unit in the most severe of clinical states, requiring extracorporeal membrane oxygenation (ECMO); the cohorts in these studies are small (77 and 83 patients, respectively), and they are relevant only to the subset of COVID-19 patients requiring ECMO treatment. Moreover, these works did not include covariates, while our results demonstrate the importance of baseline and time-dependent covariates in the multistate models (see [Supplementary Tables S2 to S4](#)). To the best of our knowledge, there are no existing works that provide hospital load prediction on a daily basis while taking into account baseline patients' characteristics, clinical state, and time in hospital.

Other methods exist that might be used to model the patient transition process. For example, a Bayesian network, a hidden Markov model, or even a recurrent neural network; however, each of these has its drawbacks. None of them can deal in a straightforward manner with the issues of (semi)-competing risks that are prevalent in our data, leading to potentially substantial biases. Further, recur-



**Figure 3.** Observed and predicted total hospitalized (top left) and critical (top right) patients, and in-hospital mortality (bottom) under the following scenarios: (1) younger: rate and state of incoming patients are the same as in Israel during the weeks from March 15 to May 2, but with patients in their 50s and 60s instead of 60+ years of age; (2) milder: rate and age of incoming patients are the same as in Israel during the weeks from March 15 to May 2, but all patients incoming only in moderate and severe state, none at critical; and (3) nursing home (NH) outbreak, in which we assume that in addition to the arrival of patients as happened in Israel from March 15 to May 2, there is a single week during which there are 4 times as many incoming patients 70+ years of age, arriving in various clinical states. For in-hospital mortality, Expected is the model prediction assuming the patient arrival process in Israel during the weeks from March 15 to May 2, with no changes. For top left and top right figures, gray vertical lines are pointwise 10%-90% confidence predictions.

rent neural networks are not adapted to right-censoring, and left-truncation. Hidden Markov models are based on a Markovian assumption which might not agree with the nature of the disease trajectory. More advanced methods such as the method by Alaa and van der Schaar<sup>20</sup> do acknowledge censoring but their parametric approach is applicable when the transient states are unobservable; we are in an easier setting where the transient states are in fact observed. Very recent work adapting neural ordinary differential equations (ODEs) to multistate problems might be an interesting avenue in the future,<sup>21</sup> but currently this approach is limited to Markovian models. We adopted a semiparametric Cox model that covers non-Markovian models, at the price of the well-known proportional hazards assumption. However, under a short-term follow-up, as in COVID-19 hospitalization data, we believe that this assumption is reasonable, as evidenced in the model’s predictive performance. We have also experimented with using random survival forest for competing risks,<sup>22</sup> which is a fully nonparametric method; we found its performance to be similar or inferior to Cox models. Because our main goal is prediction based on sampling complete hospitalization trajectories, the semiparametric structure of Cox-based approach has a major computational advantage, which led us to focus on this method.

Another line of work related to ours are models for predicting individual patient outcomes.<sup>23</sup> Viewed through this lens, our model is distinctive in 2 ways: it provides time-to-event (discharge, deterioration, death) predictions and it is based on a very small number of covariates (age, sex, and 1 of 3 patient clinical states). In contrast,

Liang et al.<sup>24</sup> reported an AUROC of 0.88 for predicting critical illness or death using 10 covariates selected from 72 potential predictors. Bello-Chavolla et al.<sup>25</sup> reported a concordance of 0.83 using 7 covariates based mostly on comorbidities. We conjecture that the accuracy achieved by our model while using minimal, easily obtainable data as input might be explained by the fact that reported patient clinical states function as expert indicator variables summarizing more granular clinical measures and comorbidities. The fact that our model achieves high accuracy using only such basic covariates is encouraging: it implies that routinely available patient data could suffice for making accurate predictions, making adoption of our model easier in diverse settings across the world.

Our model has several limitations. First, its load predictions rely on estimation of the frequency and characteristics of future incoming patients. If arriving patient populations deviate significantly from the scenarios taken into account, the model’s predictions will be wrong. We thus recommend planners evaluate multiple hypotheticals for incoming patients, testing scenarios such as those presented in the Results, as well as predictions based on different scenarios for the rate of infection in weeks ahead of prediction time. Indeed, one could build a separate model predicting the future hospitalized population based on factors such as which nonpharmaceutical interventions are or will be in place.

A limitation of our validation strategy is that it uses only Israeli data from the first wave, in which we have no account of patients’ comorbidities.<sup>26–28</sup> We stress, however, that researchers can fit a similar model to their own data, and researchers with access to

patient-level comorbidity data can incorporate it into a multistate model using the software we provide. Furthermore, even within the same health system there might be considerable heterogeneity (see [Supplementary Appendix S2.3](#)) and in addition, changes might occur in the underlying distributions we estimate. The change can be due to changes in treatment strategies, seasonal changes, hospitalization policy changes, and more. We recommend continuously testing the model's prediction against reality using the methods we outlined previously, in order to detect such distribution shifts. A significant increase in error metrics compared with the errors on the original validation set might indicate a need to refit the model using only the more recent data available. Recently, a more advanced method has been proposed for correcting model drift in prediction models<sup>29</sup>; the method focuses on dichotomous outcomes and thus could be applied to those aspects of our model—predicting critical illness or death. Extending the previous method to the full spectrum of predictions given our multistate models is an interesting area for future work.

A further limitation is that in the data we used, the patients' clinical state was reported by the attending physician at the point of care, and individual physicians and medical centers have not adhered exactly to the Israeli MOH guidelines. Despite this possible ambiguity, empirically we find that the clinical state as reported is indeed highly predictive for individual patients.

We note that multistate models can be applied to data with more or less clinical states; for example, the states might be ventilated vs nonventilated, or alternately a more fine-grained spectrum of clinical states. Our model can also be used with any set of baseline or time-varying covariates, such as comorbidities or being on an ECMO machine. Finally, while we developed the model with COVID-19 patients as our main focus, multistate models such as the one we developed can be relevant for other diseases that are characterized by state transitions, especially if right censoring, left truncation, and state recurrence are in play.

## CONCLUSION

We developed and validated a multistate model aimed at modeling the trajectory of hospitalized COVID-19 patients. We found that focusing on the day-by-day tracking of patients' clinical state can yield accurate predictions of mortality, length-of-hospitalization, and critical illness even with a very basic set of measured covariates (age, sex, and patient being in 1 of 3 clinical states). We further show how these accurate predictions enable us to build a tool that lets healthcare managers accurately plan resource allocation for COVID-19 patient care in the face of potentially large patient surges. We believe that our model can be fruitfully adopted in healthcare systems struggling with the challenges of COVID-19 around the world.

## FUNDING

The work was funded by the Israeli Ministry of Health. MG received support from the U.S.-Israel Binational Science Foundation (2016126). US and RG received support from the Israel Science Foundation (1950/19).

## AUTHOR CONTRIBUTIONS

MR contributed to acquisition of data, study conception and design, writing of the first draft, and critical revision of the manuscript and the [Supplementary Appendix](#) for important intellectual content. RG contributed to acquisition of data, statistical analysis, and critical revision of the manuscript for

important intellectual content. JS, ABA, and IC contributed to programming required for the statistical data analysis. UG, SL-T, and AZ contributed to acquisition of data. YB-L contributed to a critical revision of the manuscript for important intellectual content. DE contributed to study conception and design and critical revision of the manuscript for important intellectual content. MG contributed to study conception and design, contributed to statistical analysis and interpretation of the results, wrote the [Supplementary Appendix](#), and contributed to critical revision of the manuscript for important intellectual content. US contributed to study conception and design, analysis and interpretation of data, writing of the first draft of the manuscript, and critical revision of the manuscript and of the [Supplementary Appendix](#) for important intellectual content.

## ETHICS APPROVAL

An exemption from institutional review board approval was determined by the Israeli Ministry of Health as part of an active epidemiological investigation, based on use of anonymous data only and no medical intervention.

## SUPPLEMENTARY MATERIAL

[Supplementary Appendix](#) is available at *Journal of the American Medical Informatics Association* online.

## ACKNOWLEDGMENTS

We thank Dr Amit Huppert and the biostatistics unit researchers at Gertner Institute for their insights and help in conducting this study. We thank Professor Orly Manor for her valuable comments on the manuscript. We further thank Professor Gadi Segal for insightful discussions. We thank the Medical Division and Information and Technologies Division of the Israeli Ministry of Health for their efforts in the gathering and organization of the clinical data from all the Israeli hospitals. We thank the information and technologies staff of medical centers across Israel for building the data infrastructure needed for the collection of the data used in this study.

## CONFLICT OF INTEREST STATEMENT

We declare no competing interests.

## DATA AVAILABILITY STATEMENT

An anonymized version of the data underlying this article will be made available in <https://github.com/JonathanSomer/covid-19-multi-state-model/tree/master/data>. The anonymization will include small random changes to first hospitalization dates, and binning of age data.

## REFERENCES

1. World Health Organization. *Coronavirus Disease 2019 (COVID-19): Situation Report, 51*. Geneva, Switzerland: World Health Organization; 2020.
2. Grasselli G, Pesenti A, Cecconi M. Critical care utilization for the COVID-19 outbreak in Lombardy, Italy. *JAMA* 2020; 323 (16): 1545–6.
3. Peters AW, Chawla KS, Turnbull ZA. Transforming ORs into ICUs. *N Engl J Med* 2020; 382 (19): e52.
4. Li R, Rivers C, Tan Q, Murray MB, Toner E, Lipsitch M. Estimated demand for US hospital inpatient and intensive care unit beds for patients with COVID-19 based on comparisons with Wuhan and Guangzhou, China. *JAMA Netw Open* 2020; 3 (5): e208297.
5. Rees EM, Nightingale ES, Jafari Y, et al. COVID-19 length of hospital stay: a systematic review and data synthesis. *BMC Med* 2020; 18 (1): 1–22. doi: 10.16/s12916-020-01726-3



6. R Core Team. R: A Language and Environment for Statistical Computing. R Found Stat Comput; 2018.
7. Andersen PK, Gill RD. Cox's regression model for counting processes: a large sample study. *Ann Statist* 1982; 10 (4): 1100–20.
8. Andersen PK, Hansen LS, Keiding N. Non-and semi-parametric estimation of transition probabilities from censored observation of a non-homogeneous Markov process. *Scand J Stat* 1991; 18 (2): 153–67.
9. Klein JP, Moeschberger ML. *Survival Analysis: Techniques for Censored and Truncated Data*. New York, NY: Springer Science & Business Media; 2006.
10. Kalbfleisch JD, Prentice RL. *The Statistical Analysis of Failure Time Data*. Vol. 360. Hoboken, NJ: Wiley; 2011.
11. Centers for Disease Control and Prevention. *ICD-10-CM Official Coding and Reporting Guidelines*. Atlanta, GA: Centers for Disease Control and Prevention; 2020.
12. National Institutes of Health. Management of COVID-19 | *Coronavirus Disease COVID-19*. Bethesda, MD: National Institutes of Health; 2020.
13. Marshall JC, Murthy S, Diaz J, et al. A minimal common outcome measure set for COVID-19 clinical research. *Lancet Infect Dis* 2020; 20 (8): E192–7. doi: 10.1016/S1473-3099(20)30483-7.
14. Condes E, Arribas JR; COVID-19 MADRID-S.P.P.M. group. Impact of COVID-19 on Madrid hospital system. *Enferm Infecc Microbiol Clin* 2020 Jun 25 [E-pub ahead of print]. doi: 10.1016/j.eimc.2020.06.005.
15. Weissman GE, Crane-Droesch A, Chivers C, et al. Locally informed simulation to predict hospital capacity needs during the COVID-19 pandemic. *Ann Intern Med* 2020; 173 (1): 21–8.
16. Moghadas SM, Shoukat A, Fitzpatrick MC, et al. Projecting hospital utilization during the COVID-19 outbreaks in the United States. *Proc Natl Acad Sci U S A* 2020; 117 (16): 9122–6.
17. Oulhaj A, Ahmed LA, Prattes J, et al. The competing risk between in-hospital mortality and recovery: a pitfall in COVID-19 survival analysis research. *medRxiv*, doi: <https://www.medrxiv.org/content/10.1101/2020.07.11.20151472v2>, 15 Jul 2020, preprint: not peer reviewed.
18. Hazard D, Kaier K, von Cube M, et al. Joint analysis of duration of ventilation, length of intensive care, and mortality of COVID-19 patients: a multistate approach. *BMC Med Res Methodol* 2020; 20 (1): 206.
19. Schmidt M, Hajage D, Lebreton G, et al. Extracorporeal membrane oxygenation for severe acute respiratory distress syndrome associated with COVID-19: a retrospective cohort study. *Lancet Respir Med* 2020; 8 (1): 1121–31.
20. Alaa AM, van der Schaar M. A hidden absorbing semi-Markov model for informatively censored temporal data: Learning and inference. *J Mach Learn Res* 2018; 19: 1–62.
21. Groha S, Schmon SM, Gusev A. Neural ODEs for multi-State survival analysis. *arXiv*, doi: <https://arxiv.org/abs/2006.04893>, 8 Jun 2020, preprint: not peer reviewed.
22. Ishwaran H, Lu M. Random survival forests. *Wiley StatsRef: Statistics Reference Online*. 2019. doi: 10.1002/9781118445112.stat08188.
23. 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* 2020; 369(8242): m1328.
24. Liang W, Liang H, Ou L, et al.; for the China Medical Treatment Expert Group for COVID-19. Development and validation of a clinical risk score to predict the occurrence of critical illness in hospitalized patients with COVID-19. *JAMA Intern Med* 2020; 180 (8): 1081–9.
25. Bello-Chavolla OY, Bahena-López JP, Ne A-V, et al. Predicting mortality due to SARS-CoV-2: a mechanistic score relating obesity and diabetes to COVID-19 outcomes in Mexico. *J Clin Endocrinol Metab* 2020; 105 (8): dgaa346. doi: 10.1210/clinem/dgaa346.
26. Atkins JL, Masoli JAH, Delgado J, et al. Preexisting Comorbidities Predicting COVID-19 and Mortality in the UK Biobank Community Cohort. *J Gerontol A Biol Sci Med Sci* 2020; 75 (11): 2224–30.
27. Petrilli CM, Jones SA, Yang J, et al. Factors associated with hospital admission and critical illness among 5279 people with coronavirus disease 2019 in New York City: prospective cohort study. *BMJ* 2020; 369 (8249): m1966.
28. Fang L, Karakiulakis G, Roth M. Are patients with hypertension and diabetes mellitus at increased risk for COVID-19 infection? *Lancet Respir Med* 2020; 8 (4): e21.
29. Davis SE, Greevy RA, Fonnesbeck C, Lasko TA, Walsh CG, Matheny ME. A nonparametric updating method to correct clinical prediction model drift. *J Am Med Inform Assoc* 2019; 26 (12): 1448–57. doi: 10.1093/jamia/ocz127.