Supplementary Material

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

Michael Roimi*, Rom Gutman†, Jonathan Somer,
Asaf Ben Arie, Ido Calman, Yaron Bar-Lavie,
Udi Gelbshtein, Sigal Liverant-Taub, Arnona Ziv, Danny Eytan,
Malka Gorfine‡, Uri Shalit§

*joint first authors
†joint first authors
‡joint last authors and corresponding authors, gorfinem@tauex.tau.ac.il
§joint last authors and corresponding authors, urishalit@technion.ac.il

# Contents

# S1 Models and Methods

## S1.1 Introduction

The hospitalization course of each patient is described as a multi-state process, depicted in Figure S1. A patient enters the hospital at one of the following three clinical states: moderate, severe or critical. During the course of hospitalization a patient can move among the transient clinical states: Critical, denoted by $C$; Moderate or Severe, denoted by $M/S$; and Discharged, denoted by $Di$. Our multi-state model combines Moderate and Severe, during hospitalization, due to the small number of observed transitions from and to each of these states, separately. The state $Di$ is considered as a transient state rather than a terminal state since frequently patients in a milder state were released from hospital to a dedicated quarantine for COVID-19 patients, with some later experiencing deterioration leading to re-hospitalization; in Figure S1 we see there were 102 transitions from state $Di$ to state $M/S$. The terminal state Deceased is denoted by $De$.

Our multi-state model allows the following six transitions

$$C \to M/S \quad C \to De \quad M/S \to De \quad M/S \to C \quad M/S \to Di \quad Di \to M/S\,.$$

Three transitions were excluded from the model due to small sample size: The dataset includes 10 records of transition $C \to Di$, no records of $Di \to De$, and 2 records of $Di \to C$. Hence, these three transitions were excluded from the multi-state model. Each possible transition is characterized by a transition-specific Cox proportional hazard model with an unspecified transition-specific baseline hazard function, $\lambda_{0\cdot,\cdot}$ and a transition-specific vector $\beta_{\cdot,\cdot}$ of regression coefficients. Specifically, for $t > 0$, the corresponding Cox proportional hazard functions are

$$\lambda_{C,M/S}(t|Z) = \lambda_{0C,M/S}(t)\exp(\beta_{C,M/S}^T Z)\,, \tag{1}$$

$$\lambda_{C,De}(t|Z) = \lambda_{0C,De}(t)\exp(\beta_{C,De}^T Z)\,, \tag{2}$$

$$\lambda_{M/S,Di}(t|Z) = \lambda_{0M/S,Di}(t)\exp(\beta_{M/S,Di}^T Z)\,, \tag{3}$$

$$\lambda_{M/S,C}(t|Z) = \lambda_{0M/S,C}(t)\exp(\beta_{M/S,C}^T Z)\,, \tag{4}$$

$$\lambda_{M/S,De}(t|Z) = \lambda_{0M/S,De}(t)\exp(\beta_{M/S,De}^T Z)\,, \tag{5}$$

and

$$\lambda_{Di,M/S}(t|Z) = \lambda_{0Di,M/S}(t)\exp(\beta_{Di,M/S}^T Z)\,, \tag{6}$$

where $Z$ is a vector of covariates, possibly with time-dependent covariates. For simplicity of notation, we use $Z$ instead of $Z(t)$ whenever confusion is unexpected. Although $Z$ is shared by

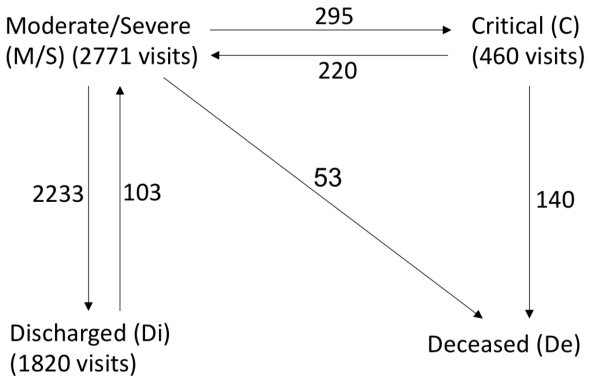


Figure S1: Multi-State Model: Data as of May 2, 2020, $n = 2675$. We model a COVID-19 patients disease course as moving between 4 possible clinical states: (i) moderate or severe, (ii) critical, (iii) discharged and (iv) deceased. We combined the two clinical states moderate and severe into a single model state due to statistical considerations; however, we keep a distinction between the two 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 clinical state transitions, and may visit a transient clinical state more than once. The dataset includes 10 records of transitions $C \to Di$, no records of $Di \to De$, and 2 records of $Di \to C$. Hence, these three transitions were excluded from the multi-state model.

the six models above, it does not imply that identical covariates must be used in these models, since the regression coefficient vectors, $\beta_{\cdot,\cdot}$, are transition dependent, and one can set any specific coefficient to 0 in order to exclude the corresponding covariate.

The covariates included in the Cox models were: age, sex, state at time of hospitalization (three categories: Moderate, Severe, or Critical), a binary variable equal to 1 if the patient previously was in a critical state and 0 otherwise, the cumulative number of days in hospital at entry time to the current state, and the interaction between age and each of the other covariates listed above. The models of transitions $M/S \to De$ and $Di \to M/S$ are slightly different due to only a few events with a critical clinical state at time of hospital admission and a few events with previous visits in critical state. Therefore, for these two models, the binary covariate taking the value of 1 if the patient previously visited the critical state is excluded, and the covariate for clinical state at time of hospitalization was redefined as a binary covariate indicating Moderate versus Severe/Critical.

## S1.2 Estimation

Estimating the above hazard functions (1)–(6) involves several major issues (beside right censoring), which we describe below: multi-state process, left truncation, competing risks and recurrent events.

**Multi-state process:** We describe the hospitalization path of each patient as a multi-state process, starting at states $C$ or $M/S$. Each patient may visits a transient state ($C$, $M/S$ or $Di$) multiple times before reaching a terminal state ($De$).

**Left truncation:** Consider, for example, a patient who entered the hospital at state $M/S$ and moved to state $C$ at the 10th day of hospitalization. The contribution of such a patient to transitions from state $C$ (back to $M/S$ or to $De$) starts only at the 10th day of hospitalization. Hence, when estimating the model of a certain transition from origin state (e.g. $C$) to a target state (e.g. $De$), those who entered the origin state during the course of hospitalization are left-truncated by their entering day to that origin state.

**Competing risks (competing events):** Given that a patient is, for example, at state $C$, there are two possible transitions: $C \to M/S$ and $C \to De$. Since only one of these transitions can occur at each point in time, $M/S$ and $De$ are competing events in the sense that at each time point, the occurrence of one type of event will prevent the occurrence of the other. Similarly, given that a patient is at state $M/S$, the events $Di$, $C$ and $De$ are competing events.

**Recurrent events:** A patient may visit states $C$, $M/S$ and $Di$ multiple times. For example, 13 patients had the following hospitalization path

$$M/S \to C \to M/S \to C \to M/S \to Di\,,$$

and 68 patients had

$$M \to Di \to M/S \to Di\,.$$

All the observed paths and their frequencies are provided in Table S1. When the event of interest can occur more than once in a patient, the events are termed recurrent events.

We overcame all the above challenges and provided consistent estimators (under mild regularity conditions) for the six Cox proportional hazard models. Specifically, by extending the approach of Andersen et al. (1991), it can be shown that maximizing the likelihood function in terms of the six Cox models can be done by maximizing the likelihood of each transition separately, while using the risk-set correction for dealing with left truncation (Klein & Moeschberger, 2006); treating competing transitions as right censoring (Kalbfleisch & Prentice, 2011); and adopting Andersen-Gill approach for dealing with recurrent events so the robust standard errors account for correlated outcomes within a patient (Andersen & Gill, 1982).

## S1.3 Predictions

Based on our multi-state model, we accurately predict at the **patient level**, given age, sex and state at time of hospitalization, the following quantities:

1. The chance of in-hospital mortality (state $De$).
2. The chance of being at a critical state (state $C$).
3. The total length of stay (LOS) in hospital (not including time in a dedicated out-of-hospital quarantine).
4. The total length of stay in critical state (LOSCS).

The above quantities can be predicted at the first day of hospitalization and also during the course of hospitalization, while correctly taking into account the accumulated hospitalization history.

Weighted estimators of the area under the Receiver Operating Characteristic curve (AUROC) were used to evaluate death and critical-visit predictions for binary classifications (i.e. yes/no in-hospital mortality and yes/no previously visiting critical state). The weights eliminate the bias due to exclusion of censored observations, and were defined as 1 over the probability of being uncensored (Robins & Finkelstein, 2000). The weights were estimated by a Kaplan-Meier estimator of the censoring survival function. Weighted AUROC estimates were calculated by the R package WeightedROC (Hocking, 2020). Brier scores were used as measures of prediction accuracy.

Based on the above predictions, we go one step further and provide predictions at the **hospital level**, in the following manner:

1. **Snapshot:** Assume that at a given calendar day, we are given the current state and hospitalization history of all the COVID-19 patients currently at a specific hospital. Beyond predicting the above quantities for each patient, we also predict the total number of patients at the hospital, and at a critical clinical state in particular, for each day over the next 8 weeks. Namely, we provide predictions on the total occupancy on a calendar scale which are only due to the currently hospitalized patients.
2. **Arrival:** Given as input the arrival process of patients to the hospital at each day, including the number of arriving patients, their age, sex and state at time of hospitalization, we predict the total number of patients at hospital, and at critical state in particular, for each day of the next 8 weeks. Here we provide a prediction for the total occupancy on a calendar scale, for any possible hypothetical arrival scenario.
3. **Arrival plus Snapshot:** At a given calendar day, we are given the current state and hospitalization history of all the COVID-19 patients currently at the hospital along with

an arrival process of the patients to be hospitalized starting the next day up to a pre-specified time period. Again, we predict the total number of patients at hospital, and at critical state in particular, on a calendar scale.

## S1.4 Monte Carlo Estimator of Length of Stay

Since our hospitalization model consists of a multi-state model with recurring events (i.e. a patient can visit a transient state multiple times) the closed form marginal probabilities required for predictions are intractable. Instead we use a Monte Carlo (MC) approach for estimating all the required quantities listed above (Section S1.3).

Assume a prediction is desired for a new patient with baseline covariates (i.e. age, sex, and clinical state at time of hospitalization) denoted by $X$. The MC-based prediction procedure can be summarized as follows. Given $X$, sample a large number (e.g. 20,000) of hospitalization paths, and use these paths for estimating the required quantities. Specifically, the probability of death is estimated by the proportion of paths ended at state $De$; the probability of visiting state $C$ is estimated by the proportion of paths visited state $C$; the expected total length of stay is estimated by the mean length of the paths (not including time at state $Di$); and the expected length of stay in state $C$ is estimated by the mean time spent in state $C$ over all the paths. The next subsection provides a detailed description of path sampling.

## S1.5 Path Sampling - Technical Details

Let $J_C$ and $J_N$ denote the current and next states, respectively. Assume a patient entered the hospital at state $J_C = j^*$ with a vector of baseline covariates $X$. The goal is to provide a MC estimator of the length of hospitalization given $X$ and $j^*$. Let $Z(t)$ be a time-dependent vector of covariates such that $Z(t) = (X^T, \tilde{X}(t)^T)^T$, where $\tilde{X}(t)$ is a time-dependent vector of covariates that are known at the entrance to the new state. Details of $\tilde{X}(t)$ in our setting are provided at the end of Section S1.1. Assume $K_{j^*}$ possible transitions from state $j^*$. For each state $j$, $j = 1, \ldots, K_{j^*}$,

$$\begin{aligned}
&\Pr(T \le t, J_N = j | J_C = j^*, Z(0) = Z) \\
&\quad = \int_0^t \exp\left(\beta_{j^*,j}^T Z\right) \lambda_{0j^*,j}(u) \exp\left\{ -\sum_{k=1}^{K_{j^*}} \Lambda_{0j^*,k}(u^-) \exp\left(\beta_{j^*,k}^T Z\right) \right\} du\,,
\end{aligned}$$

where $\beta_{j^*,j}$, $\lambda_{0j^*,j}$ and $\Lambda_{0j^*,j}$ are the vector of regression coefficients, the baseline hazard function and the cumulative baseline hazard function of transition $j^* \to j$, respectively. Then,

$$\begin{aligned}&\Pr(J_N = j | J_C = j^*, Z(0) = Z) \\ &\quad = \int_0^\infty \exp\left(\beta_{j^*,j}^T Z\right) \lambda_{0j^*,j}(u) \exp\left\{ -\sum_{k=1}^{K_{j^*}} \Lambda_{0j^*,k}(u^-) \exp\left(\beta_{j^*,k}^T Z\right) \right\} du\end{aligned}$$

and

$$\begin{aligned}&\Pr(T \le t | J_N = j, J_C = j^*, Z(0) = Z) \\ &\quad = \frac{\int_0^t \exp\left(\beta_{j^*,j}^T Z\right) \lambda_{0j^*,j}(u) \exp\left\{ -\sum_{k=1}^{K_{j^*}} \Lambda_{0j^*,k}(u^-) \exp\left(\beta_{j^*,k}^T Z\right) \right\} du}{\int_0^\infty \exp\left(\beta_{j^*,j}^T Z\right) \lambda_{0j^*,j}(u) \exp\left\{ -\sum_{k=1}^{K_{j^*}} \Lambda_{0j^*,k}(u^-) \exp\left(\beta_{j^*,k}^T Z\right) \right\} du}.\end{aligned}$$

We start by describing the sampling procedure of the next state. Let $\tau_{j^*,j}$ be the largest observed event time of transition $j^* \to j$. Then, the next state is sampled from a $K_{j^*}$ multinomial distribution with probabilities $p_{j|j^*,Z}$ where, for $j = 1, \ldots, K_{j^*}$,

$$\begin{aligned}&\widehat{\Pr}(J_N = j | J_C = j^*, Z(0) = Z) \\ &\quad = \sum_{t_m \le \tau_{j^*,j}} \exp\left(\widehat{\beta}_{j^*,j}^T Z\right) \widehat{\lambda}_{0j^*,j}(t_m) \exp\left\{ -\sum_{k=1}^{K_{j^*}} \widehat{\Lambda}_{0j^*,k}(t_{m-1}) \exp\left(\widehat{\beta}_{j^*,k}^T Z\right) \right\},\end{aligned}$$

the summation is over the distinct observed event times of transition $j^* \to j$ and

$$p_{j|j^*,Z} = \frac{\widehat{\Pr}(J_N = j | J_C = j^*, Z(0) = Z)}{\sum_{j'=1}^{K_{j^*}} \widehat{\Pr}(J_N = j' | J_C = j^*, Z(0) = Z)}. \tag{7}$$

Once we sampled the next state, denoted by $j'$, the time to be spent at state $j^*$ should be sampled based on

$$\begin{aligned}&\widehat{\Pr}(T \le t | J_N = j', J_C = j^*, Z(0) = Z) \\ &\quad = \frac{\sum_{t_m \le t} \exp\left(\widehat{\beta}_{j^*,j'}^T Z\right) \widehat{\lambda}_{0j^*,j'}(t_m) \exp\left\{ -\sum_{k=1}^{K_{j^*}} \widehat{\Lambda}_{0j^*,k}(t_{m-1}) \exp\left(\widehat{\beta}_{j^*k}^T Z\right) \right\}}{\sum_{t_m \le \tau_{j^*,j'}} \exp\left(\widehat{\beta}_{j^*,j'}^T Z\right) \widehat{\lambda}_{0j^*,j'}(t_m) \exp\left\{ -\sum_{k=1}^{K_{j^*}} \widehat{\Lambda}_{0j^*,k}(t_{m-1}) \exp\left(\widehat{\beta}_{j^*,k}^T Z\right) \right\}}.\end{aligned} \tag{8}$$

This could be done by sampling $U \sim Uniform[0,1]$, equating

$$U = \widehat{\Pr}(T \le t | J_N = j', J_C = j^*, Z(0) = Z)$$

and solving for $t$. Denote the sampled time by $t'$ and update $Z(t')$. In case $j' = De$, the sampling path ends here. Otherwise, the current state is updated to $J_C = j'$, and the following state is sampled by $p_{j|j',Z}$, where for $j = 1, \ldots, K_{j'}$

$$p_{j|j',Z} = \frac{\sum_{t'<t_m\leq\tau_{j',j}} \exp\left(\widehat{\beta}_{j',j}^T Z\right) \widehat{\lambda}_{0j',j}(t_m) \exp\left\{-\sum_{k=1}^{K_{j'}} \widehat{\Lambda}_{0j',k}(t_{m-1}) \exp\left(\widehat{\beta}_{j',k}^T Z\right)\right\}}{\sum_{j^{**}=1}^{K_{j'}} \sum_{t'<t_m\leq\tau_{j',j^{**}}} \exp\left(\widehat{\beta}_{j',j^{**}}^T Z\right) \widehat{\lambda}_{0j',j^{**}}(t_m) \exp\left\{-\sum_{k=1}^{K_{j'}} \widehat{\Lambda}_{0j',k}(t_{m-1}) \exp\left(\widehat{\beta}_{j',k}^T Z\right)\right\}}. \tag{9}$$

An exceptional state is $J_C = Di$, where one can either move back to $M/S$ or stay at $Di$.

Given the new sampled state, denoted by $\breve{j}$, the time to be spent at $j'$ is sampled by

$$\begin{aligned}&\widehat{\Pr}(T \leq t | J_N = \breve{j}, J_C = j', Z(t') = Z) \\ &= \frac{\sum_{t'<t_m\leq t} \exp\left(\widehat{\beta}_{j',\breve{j}}^T Z\right) \widehat{\lambda}_{0j',\breve{j}}(t_m) \exp\left\{-\sum_{k=1}^{K_{j'}} \widehat{\Lambda}_{0j',k}(t_{m-1}) \exp\left(\widehat{\beta}_{j',k}^T Z\right)\right\}}{\sum_{t'<t_m\leq\tau_{j',\breve{j}}} \exp\left(\widehat{\beta}_{j',\breve{j}}^T Z\right) \widehat{\lambda}_{0j',\breve{j}}(t_m) \exp\left\{-\sum_{k=1}^{K_{j'}} \widehat{\Lambda}_{0j',k}(t_{m-1}) \exp\left(\widehat{\beta}_{j',k}^T Z\right)\right\}}\end{aligned} \tag{10}$$

and then by solving for $U = \widehat{\Pr}(T \leq t | J_N = \breve{j}, J_C = j', Z(t') = Z)$. The sampled path is completed once state $De$ is sampled or state $Di$ is sampled and the next sampled state is again $Di$, or after sampling 9 states, whichever comes first. We set the maximum number of transitions to 9 as observed in our dataset (see TableS1).

## S1.6 Predictions at the Patient Level - Technical Details

We consider the following three types of predictions. Type A is prediction at time of hospitalization, Type B is prediction at the entrance to any new clinical state, and prediction Type C is done during the stay at a certain state. Let $Z(t)$ consist of age, sex, clinical state at time of hospitalization, the cumulative number of days in hospital **up to the entrance to the current state**, an indicator variable equal to 1 if previously visited Critical state and 0 otherwise, and the interactions of each of these covariates with age.

**Type A.** Given age $= a$, sex $= s$ ($s$ is either 1 or 0) and state at time of hospitalization $j^*$, the vector of covariates for prediction at time of hospitalization is given by

$$Z(t) = Z(0) = \{a, s, I(j^* = \text{M}), I(j^* = \text{S}), 0, 0, a \cdot s, a \cdot I(j^* = \text{M}), a \cdot I(j^* = \text{S}), 0, 0\}$$

and Eq's (7) and (8) provide the estimated probabilities of the next state $j$ and the probability of transitioning by day $t$ given the transition $j^* \to j$ and $Z(0)$.

**Type B.** Given age $= a$, sex $= s$, state at hospitalization $j^*$, the patient now entered state $j'$, the cumulative number of days in hospital up to the entrance to the current state equals $t'$, the

vector of covariates for predictions is

$$Z(t') = \{a, s, I(j^* = \mathrm{M}), I(j^* = \mathrm{S}), t', W, a \cdot s, a \cdot I(j^* = \mathrm{M}), a \cdot I(j^* = \mathrm{S}), a \cdot t', a \cdot W\}$$

where $W = 1$ if previously visited $C$, and 0 otherwise. Hence, Eq's (9) and (10) provide the estimated probabilities of the next state $j$ and the probability of transitioning by day $t$ given the transition $j^* \to j'$ and $Z(t')$.

**Type C.** Given $Z(t')$ as in Type B, and given that the patient is at state $j'$ for already $d$ days, the following are the updated predictions where we take into account the $d$ days in current state but there is no change in the vector of covariates $Z(t')$. Specifically, the next state probability $p_{j|j',Z,d}$, $j = 1, \ldots, K_{j'}$, is given by

$$p_{j|j',Z,d} = \frac{\sum_{t'+d<t_m\le\tau_{j',j}} \exp\left(\widehat{\beta}_{j',j}^T Z\right) \widehat{\lambda}_{0j',j}(t_m) \exp\left\{-\sum_{k=1}^{K_{j'}} \widehat{\Lambda}_{0j',k}(t_{m-1}) \exp\left(\widehat{\beta}_{j',k}^T Z\right)\right\}}{\sum_{j^{**}=1}^{K_{j'}} \sum_{t'+d<t_m\le\tau_{j',j^{**}}} \exp\left(\widehat{\beta}_{j',j^{**}}^T Z\right) \widehat{\lambda}_{0j',j^{**}}(t_m) \exp\left\{-\sum_{k=1}^{K_{j'}} \widehat{\Lambda}_{0j',k}(t_{m-1}) \exp\left(\widehat{\beta}_{j',k}^T Z\right)\right\}},$$

where summations are over the observed event times of the respective transition. Given the new sampled state, denoted by $\breve{j}$, the time to be spent at $j'$ is predicted by

$$\begin{aligned}
&\widehat{\Pr}(T \le t | J_N = \breve{j}, J_C = j', Z(t') = Z, d) \\
&= \frac{\sum_{t'+d<t_m\le t} \exp\left(\widehat{\beta}_{j',\breve{j}}^T Z\right) \widehat{\lambda}_{0j',\breve{j}}(t_m) \exp\left\{-\sum_{k=1}^{K_{j'}} \widehat{\Lambda}_{0j',k}(t_{m-1}) \exp\left(\widehat{\beta}_{j',k}^T Z\right)\right\}}{\sum_{t'+d<t_m\le\tau_{j',\breve{j}}} \exp\left(\widehat{\beta}_{j',\breve{j}}^T Z\right) \widehat{\lambda}_{0j',\breve{j}}(t_m) \exp\left\{-\sum_{k=1}^{K_{j'}} \widehat{\Lambda}_{0j',k}(t_{m-1}) \exp\left(\widehat{\beta}_{j',k}^T Z\right)\right\}}.
\end{aligned}$$

## S1.7 Weighted Bootstrap Standard Error for Prediction at the Subject Level

Denote the total number of patients in the training data $n$. Our goal is estimating the standard error (SE) of predictions of a new patient with baseline covariate $X^o$. The following is a weighted bootstrap procedure (Kosorok, 2007) for SE estimation of our proposed MC-based estimators:

1. Sample $n$ weight values from an exponential distribution with mean 1, and assign a weight value to each patient in the training sample.
2. Estimate the six Cox PH models (1)–(6) with the weights sampled in Step 1.
3. Given $X^o$, sample 20,000 MC paths, and based on the 20,000 paths compute all the desired quantities according to Section S1.4.
4. Repeat Steps 1–3 $B$ times.

The empirical variances of the $B$ estimates are the weighted bootstrap SEs estimates. For

example, the empirical variance of the $B$ death probability estimates is the SE estimate of the estimated death probability of this new patient.

### S1.8 Prediction at the Health-System Level

As described above, for prediction at the patient level, for any given vector of covariates one should run a large number of MC paths, summarize the MC runs and obtain the distributions of LOS in hospital, of LOSCS, etc. Predictions at the hospital level for Snapshot, Arrival or Snapshot plus Arrival (see Section S1.3 for details) require a different MC approach, as follows:

1. Sample one hospitalization path for each patient at the starting date for Snapshot patients and at the patient hospitalization date for Arrival patients.

2. Summarize the paths over all the patients during the predefined period. For example, count the number of patients in the hospital at each day.

3. Repeat Steps 1–2 above a large number of times (e.g. 10,000), and summarize over these repeats. For example, get the mean or various quantiles of the number of patients in the hospital at each day.

## S2 Results

Data analysis was conducted in R (R Core Team, 2020). Table S2 – Table S4 provide the regression coefficients of each of the transition models (1)–(6). The following are summaries of the various MC predictions.

### S2.1 Prediction at the Patient Level

Table S5 provides death probabilities for various patient types defined by sex, age and state at time of hospitalization. Prediction is based on 20,000 MC paths for each patient type along with weighted-bootstrap 95% confidence intervals. As expected, death probability increases with age and with state severity at time of hospitalization. Table S6 gives the probability of being at critical state during hospitalization by patient types along with weighted-bootstrap 95% confidence intervals. The 10%, 25%, 50%, 75% and 90% quantiles of LOS in days and weighted-bootstrap standard errors, by patient type are given in Table S7. Quantiles of LOSCS in days, given being in critical state with weighted bootstrap standard errors are presented in Table S8 by patient type. Plots of the cumulative distribution function of LOS by patient types are provide in Figure S2. All the above results are based on 20,000 MC paths for the estimates of each patient type, and 500 weighted bootstrap samples, each with 20,000 MC paths, for the standard error estimates for each patient type.

The AUROC and Brier Score results, based on 8-fold cross validation, are presented in Table S10 (more details on the 8-fold cross validation study in the next section). The mean AUROC and Brier Scores estimates for death prediction over the eight held-out subsets are 0.955 (SE=0.035) and 0.043 (SE=0.011); the respective AUROC and Brier score of predicting critical clinical state visit are 0.880 (SE=0.040) and 0.049 (SE=0.013).

## S2.2 Prediction at the Health System Level - Random Subset of Patients 8-fold Cross Validation

The entire dataset was randomly partitioned into 8 groups of about 330 patients each. Each time the model was trained with one group omitted (held-out), and our prediction tool was used to provide predictions for the patients of the held-out group. We demonstrate the results of one held-out random group in Figure S3 for the arrival process and two snapshots, at April 1st and 15th. Evidently, our prediction tool performs very well in terms of absolute error. Similar conclusions are obtained from Figure S4 which provides the summary over the 8 held-out groups. Tables S9 presents the mean absolute errors (MAE) of each held-out group and over the eight groups. MAE is defined as the mean of the absolute daily-based differences between observed and predicted number of patients; the means are over 64, 32 and 18 days, for the Arrival and the two Snapshot settings, respectively. "ALL" (in Tables S9) refers to MAE between the mean curves of the estimated and predicted curves of the 8 held-out random groups, and "Mean" (SE) are the means and SEs of the 8 MAEs. The results show excellent prediction performances with small mean MAE. In particular, 4.72 (SE=1.07) and 1.68 (SE=0.40) for LOS and LOSCS of Arrival; 3.15 (SE=1.20) and 1.47 (SE=0.56) for LOS and LOSCS of April 1st Snapshot; and 3.13 (SE=1.07) and 1.98 (SE=0.93) for LOS and LOSCS of April 15th Snapshot.

## S2.3 Prediction at the Health System Level - Hospital Holdout

We further evaluate the model by training on a sample where patients from each hospital in turn are held-out and not included in the training dataset. Figure S5 and Figure S6 demonstrate our ability to predict for the two hospitals with the largest patient population in our sample, hospitals H5 and H7 (we were asked to avoid identifying the hospital names). Evidently, load predictions for H5 are satisfactory, but not so for H7. It shows that when using our model for predicting at the hospital level, it is preferable to include patients from the predicted hospital in the training data thus avoiding possible hospital-specific biases. Table S11 and Table S12 show the mean absolute error, ROC AUC and Brier Scores for predicting death and visiting critical state for each hospital. In some hospitals, the number of deceased or number of patients with visits at Critical are very small, so results should be interpreted with caution.

## S2.4 Prediction at the Health System Level - Hypothetical Scenarios

Assume we would like to predict the hospitalized load in a certain health-care system in a given period of time using the Snapshot plus Arrival approach (see Section S1.3 for details). Three hypothetical scenarios were constructed based on the observed arrival process of a random hold-out sample of size 330 (specifically, we used the first random subset described in Section S2.2 and Figure S3):

1. **"Younger":** patients of age 60 and above within the held-out sample were replaced by patients of ages between 40–50 with probability 2/3, and between 50–60 with probability 1/3. Given the age group of a new patient, the specific age, sex and clinical state at time of hospitalization were sampled based on the distribution of Israeli COVID-19 patients.

2. **"Milder":** patients at the critical clinical state at time of hospitalization were left at the critical state or replaced by Severe or Moderate patients, each with probabilities 1/3. Given the state at time of hospitalization of a new patient, the specific age and sex were sampled based on the distribution of Israeli COVID-19 patients.

3. **"Elder Care Nursing Home Outbreak":** On the 5th week from the beginning of the arrival process, and only for this week, the number of patients at age 70 and above was multiplied by four. Age and sex of the new patients were sampled based on the Israeli population distribution, and the state at time of hospitalization was sampled based on the distribution of Israeli COVID-19 patients.

The model was trained by the data not including the random held-out sample, and prediction was performed with 10,000 MC paths as described in Section S1.8. The results are presented in Figure S7 as well as in Figure 2 of the main paper. Table S13 (same as Table 3 of main paper) shows the number of observed death based on the actual data of the held-out sample, versus the expected number of deaths for each of the above hypothetical scenarios. The high similarity between the observed and expected number of deaths demonstrate that our proposed model is well calibrated. Under the "Younger" scenario the number of deaths decreases dramatically, the decrease is moderate under the "Milder" scenario, and as expected, an outbreak at elder care nursing home yields a substantial increase in number of deaths.

Table S1: Summary of the observed hospitalization course (observed paths): Di - Discharged, M/S - Moderate/Severe, C - Critical, De - Deceased. 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.

| | 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 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 |

Table S2: Results of Cox survival analysis models of transitions from Critical state: STH: state at time of hospitalization (Moderate/Severe/Critical); HistCrt equals 1 if visited Critical state previously; Cum. Days is the number of days in hospital at entry to current state. The reported values are the regression coefficients (robust SE, p-value if smaller than 0.05). Results are based on 460 state visits (a patient can visit the state multiple times). Next states: 140 to Deceased state, 220 to Moderate/Severe state, and 10 to Discharged state.

| | Transition to Deceased estimate (SE, pvalue) | Transition to Moderate/Severe estimate (SE, pvalue) |
|---|---|---|
| Age | 0.022 (0.023) | 0.006 (0.009) |
| Sex (Male) | -3.156 (1.329, 0.018) | 0.278 (0.614) |
| STH (Severe) | -0.799 (1.841) | -0.471 (0.775) |
| STH (Critical) | -1.410 (1.761) | 0.543 (0.758) |
| Cum. Days | 0.059 (0.253) | 0.057 (0.078) |
| HistCrt (yes) | 3.499 (3.921) | 3.180 (1.451, 0.028) |
| Age:sex (Male) | 0.037 (0.017, 0.028) | -0.011 (0.009) |
| Age:STH (Severe) | 0.012 (0.024) | 0.006 (0.011) |
| Age:STH (Critical) | 0.021 (0.022) | -0.011 (0.011) |
| Age:Cum. Days | -0.0008 (0.003) | -0.0008 (0.001) |
| Age:HistCrt (yes) | -0.055 (0.046) | -0.030 (0.019) |

Table S3: Results of Cox survival analysis models of transitions from Moderate/Severe state: STH: state at time of hospitalization (Moderate/Severe/Critical); HistCrt equals 1 if visited Critical state previously; Cum. Days is the number of days in hospital at entry to current state. The reported values are the regression coefficients (robust SE, p-value if smaller than 0.05). Results are based on 2771 state visits (a patient can visit in the state multiple times). Next states: 53 to Deceased state, 295 to Critical state, and 2233 to Discharged state. Some covariates are not used in some of the models, see S1.1.

| | Transition to Deceased estimate (SE, pvalue) | Transition to Critical estimate (SE, pvalue) | Transition to Discharged estimate (SE, pvalue) |
|---|---|---|---|
| Age | 0.103 (0.026, 0.0000) | 0.029 (0.006, 0.000) | -0.015 (0.002, 0.000) |
| Sex (Male) | -3.039 (2.351) | 0.045 (0.531) | -0.307 (0.145, 0.034) |
| STH (Severe/Critical) | 4.278 (2.544) | - | - |
| STH (Severe) | - | 1.945 (0.589, 0.001) | 0.205 (0.186) |
| STH (Critical) | - | 1.275 (1.710) | 0.598 (0.899) |
| Cum. Days | -0.872 (0.532) | 0.010 (0.070) | -0.010 (0.027) |
| HistCrt (yes) | - | 1.643 (1.301) | -0.183 (0.660) |
| Age:sex (Male) | 0.037 (0.028) | 0.005 (0.007) | 0.005 (0.002, 0.028) |
| Age:STH(Severe/Critical) | -0.037 (0.030) | - | - |
| Age:STH (Severe) | - | -0.010 (0.008) | -0.009 (0.003, 0.002) |
| Age:STH (Critical) | - | -0.003 (0.0M/S) | -0.0132 (0.014) |
| Age:Cum. Days | 0.010 (0.006) | 0.000 (0.0009) | 0.0002 (0.0004) |
| Age:HistCrt (yes) | - | -0.016 (0.018) | -0.003 (0.010) |

Table S4: Results of Cox survival analysis models of transition from Discharged state to Moderate/Severe state: STH: state at time of hospitalization (Moderate vs. Severe/Critical); Cum. Days is the number of days in hospital at entry to current state. The reported values are the regression coefficients (robust SE, p-value if smaller than 0.05). Results are based on 1820 state visits (a patient can visit in the state multiple times). Next states: 0 at Deceased state, 2 at Critical state, and 103 at Moderate/Severe state.

| | estimate (SE, pvalue) |
|---|---|
| Age | 0.043 (0.010, 0.0000) |
| Sex (Male) | 0.294 (0.763) |
| STH (Severe/Critical) | 0.648 (1.525) |
| Cum. Days | 0.287 (0.081, 0.0004) |
| Age:sex (Male) | 0.0017 (0.012) |
| Age:STH (Severe/Critical) | 0.007 (0.022) |
| Age:Cum. Days | 0.003 (0.001, 0.0099) |

Table S5: Death probability by patient type (state at time of hospitalization, age, sex) based on 20,000 MC paths for each patient type (weighted bootstrap 95% confidence interval).

| Patient Type | Male | Female |
|---|---|---|
| Moderate, 15 | 0.0003 (0.0002,0.0004) | 0.0018 (0.0014,0.0021) |
| Moderate, 25 | 0.0004 (0.0002,0.0005) | 0.0031 (0.0025,0.0036) |
| Moderate, 35 | 0.0012 (0.0008,0.0016) | 0.0039 (0.0030,0.0048) |
| Moderate, 45 | 0.0025 (0.0022,0.0036) | 0.0077 (0.0062,0.0091) |
| Moderate, 55 | 0.0065 (0.0055,0.0075) | 0.0119 (0.0103,0.0135) |
| Moderate, 65 | 0.021 (0.019,0.023) | 0.024 (0.022,0.027) |
| Moderate, 75 | 0.056 (0.053,0.058) | 0.047 (0.045,0.049) |
| Moderate, 85 | 0.147 (0.143,0.151) | 0.117 (0.106,0.127) |
| | | |
| Severe, 15 | 0.0009 (0.0005,0.0013) | 0.0059 (0.0048,0.0069) |
| Severe, 25 | 0.0024 (0.0018,0.0030) | 0.0107 (0.0092,0.0122) |
| Severe, 35 | 0.0054 (0.0044,0.0064) | 0.0212 (0.0192,0.0232) |
| Severe, 45 | 0.015 (0.013,0.016) | 0.038 (0.036,0.041) |
| Severe, 55 | 0.038 (0.035,0.041) | 0.069 (0.065,0.072) |
| Severe, 65 | 0.097 (0.094,0.100) | 0.116 (0.111,0.120) |
| Severe, 75 | 0.207 (0.202,0.213) | 0.201 (0.195,0.207) |
| Severe, 85 | 0.432 (0.421,0.444) | 0.376 (0.363,0.388) |
| | | |
| Critical, 15 | 0.004 (0.003,0.005) | 0.052 (0.041,0.063) |
| Critical, 25 | 0.009 (0.007,0.011) | 0.078 (0.065,0.091) |
| Critical, 35 | 0.023 (0.019,0.028) | 0.121 (0.106,0.136) |
| Critical, 45 | 0.062 (0.053,0.071) | 0.183 (0.171,0.195) |
| Critical, 55 | 0.139 (0.120,0.158) | 0.282 (0.276,0.288) |
| Critical, 65 | 0.303 (0.272,0.334) | 0.405 (0.390,0.420) |
| Critical, 75 | 0.551 (0.514,0.587) | 0.547 (0.513,0.582) |
| Critical, 85 | 0.826 (0.805,0.846) | 0.746 (0.701,0.790) |

Table S6: The probability of being at critical state during hospitalization by patient type (state at time of hospitalization, age, sex) based on 20,000 MC paths for each type (weighted bootstrap 95% confidence interval).

| Patient Type | Male | Female |
|---|---|---|
| Moderate, 15 | 0.010 (0.009,0.012) | 0.009 (0.008,0.010) |
| Moderate, 25 | 0.018 (0.016,0.020) | 0.013 (0.012,0.015) |
| Moderate, 35 | 0.025 (0.023,0.027) | 0.019 (0.018,0.021) |
| Moderate, 45 | 0.037 (0.034,0.040) | 0.029 (0.027,0.031) |
| Moderate, 55 | 0.054 (0.052,0.057) | 0.041 (0.039,0.044) |
| Moderate, 65 | 0.086 (0.082,0.090) | 0.064 (0.061,0.068) |
| Moderate, 75 | 0.125 (0.121,0.130) | 0.096 (0.090,0.102) |
| Moderate, 85 | 0.178 (0.173,0.183) | 0.135 (0.127,0.144) |
| | | |
| Severe, 15 | 0.056 (0.050,0.063) | 0.042 (0.036,0.047) |
| Severe, 25 | 0.081 (0.074,0.088) | 0.060 (0.054,0.067) |
| Severe, 35 | 0.121 (0.114,0.128) | 0.095 (0.088,0.101) |
| Severe, 45 | 0.170 (0.163,0.178) | 0.129 (0.124,0.134) |
| Severe, 55 | 0.237 (0.229,0.245) | 0.185 (0.181,0.190) |
| Severe, 65 | 0.321 (0.315,0.328) | 0.253 (0.243,0.262) |
| Severe, 75 | 0.404 (0.395,0.412) | 0.319 (0.302,0.335) |
| Severe, 85 | 0.473 (0.459,0.487) | 0.393 (0.369,0.417) |

Table S7: Quantiles of length of stay in days, by patient type (sex, age and state at time of hospitalization), based on 20,000 MC paths for each patient type (weighted bootstrap standard error).

| Patient Type | 10% | 25% | 50% | 75% | 90% |
|---|---|---|---|---|---|
| Male, 15, Moderate | 1 (0.00) | 3 (0.00) | 5 (0.00) | 8 (0.00) | 13 (0.00) |
| Male, 25, Moderate | 1 (0.00) | 3 (0.00) | 5 (0.00) | 9 (0.00) | 14 (0.00) |
| Male, 35, Moderate | 1 (0.00) | 3 (0.00) | 6 (0.00) | 10 (0.47) | 16 (0.00) |
| Male, 45, Moderate | 2 (0.34) | 3 (0.00) | 6 (0.00) | 11 (0.22) | 18 (0.00) |
| Male, 55, Moderate | 2 (0.00) | 4 (0.46) | 7 (0.00) | 12 (0.24) | 21 (0.48) |
| Male, 65, Moderate | 2 (0.00) | 4 (0.00) | 7 (0.00) | 13 (0.14) | 24 (0.14) |
| Male, 75, Moderate | 2 (0.00) | 4 (0.00) | 8 (0.00) | 15 (0.00) | 27 (0.52) |
| Male, 85, Moderate | 2 (0.00) | 4 (0.00) | 8 (0.45) | 16 (0.43) | 28 (0.49) |
| Male, 15, Severe | 1 (0.00) | 2 (0.00) | 5 (0.00) | 8 (0.42) | 13 (0.20) |
| Male, 25, Severe | 1 (0.10) | 3 (0.00) | 5 (0.00) | 9 (0.10) | 18 (0.41) |
| Male, 35, Severe | 2 (0.00) | 3 (0.00) | 6 (0.35) | 12 (0.50) | 25 (0.61) |
| Male, 45, Severe | 2 (0.00) | 4 (0.00) | 8 (0.00) | 15 (0.20) | 34 (0.96) |
| Male, 55, Severe | 2 (0.00) | 4 (0.14) | 9 (0.10) | 21 (0.56) | 42 (0.74) |
| Male, 65, Severe | 3 (0.38) | 5 (0.00) | 11 (0.24) | 25 (0.36) | 45 (0.22) |
| Male, 75, Severe | 3 (0.00) | 6 (0.00) | 12 (0.22) | 28 (0.53) | 47 (0.10) |
| Male, 85, Severe | 3 (0.00) | 6 (0.10) | 11 (0.55) | 23 (0.73) | 42 (0.00) |
| Male, 15, Critical | 6 (0.26) | 9 (0.20) | 17 (0.33) | 27 (0.20) | 41 (0.52) |
| Male, 25, Critical | 7 (0.31) | 11 (0.34) | 20 (0.43) | 31 (0.56) | 45 (0.20) |
| Male, 35, Critical | 8 (0.44) | 13 (0.55) | 23 (0.39) | 38 (0.50) | 47 (0.48) |
| Male, 45, Critical | 8 (0.38) | 14 (0.55) | 26 (0.43) | 42 (0.50) | 47 (0.20) |
| Male, 55, Critical | 8 (0.14) | 16 (0.25) | 28 (0.78) | 44 (0.46) | 47 (0.10) |
| Male, 65, Critical | 6 (0.20) | 13 (0.59) | 26 (0.43) | 44 (0.00) | 48 (0.00) |
| Male, 75, Critical | 3 (0.10) | 8 (0.20) | 19 (0.81) | 39 (1.87) | 48 (3.99) |
| Male, 85, Critical | 2 (0.00) | 4 (0.10) | 9 (0.22) | 21 (0.83) | 37 (2.11) |
| Female, 15, Moderate | 1 (0.00) | 2 (0.00) | 4 (0.00) | 7 (0.10) | 10 (0.10) |
| Female, 25, Moderate | 1 (0.00) | 2 (0.00) | 5 (0.00) | 8 (0.10) | 12 (0.34) |
| Female, 35, Moderate | 1 (0.00) | 3 (0.00) | 5 (0.00) | 9 (0.10) | 14 (0.00) |
| Female, 45, Moderate | 2 (0.10) | 3 (0.00) | 6 (0.00) | 10 (0.22) | 17 (0.42) |
| Female, 55, Moderate | 2 (0.00) | 4 (0.22) | 7 (0.49) | 12 (0.00) | 20 (0.20) |
| Female, 65, Moderate | 2 (0.00) | 4 (0.00) | 8 (0.39) | 13 (0.29) | 23 (0.30) |
| Female, 75, Moderate | 2 (0.00) | 4 (0.43) | 8 (0.37) | 16 (0.10) | 26 (0.22) |
| Female, 85, Moderate | 3 (0.00) | 5 (0.00) | 9 (0.00) | 18 (0.10) | 29 (0.00) |
| Female, 15, Severe | 1 (0.00) | 2 (0.00) | 4 (0.00) | 7 (0.10) | 10 (0.31) |
| Female, 25, Severe | 1 (0.00) | 2 (0.14) | 5 (0.39) | 8 (0.10) | 14 (0.45) |
| Female, 35, Severe | 2 (0.10) | 3 (0.00) | 6 (0.00) | 10 (0.52) | 19 (0.53) |
| Female, 45, Severe | 2 (0.00) | 4 (0.49) | 7 (0.00) | 13 (0.27) | 24 (0.34) |
| Female, 55, Severe | 2 (0.00) | 4 (0.00) | 8 (0.00) | 17 (0.17) | 31 (0.66) |
| Female, 65, Severe | 3 (0.38) | 5 (0.00) | 10 (0.00) | 21 (0.47) | 39 (0.67) |
| Female, 75, Severe | 3 (0.00) | 6 (0.00) | 12 (0.14) | 24 (0.50) | 42 (0.10) |
| Female, 85, Severe | 3 (0.36) | 6 (0.10) | 12 (0.10) | 25 (0.30) | 39 (0.00) |
| Female, 15, Critical | 5 (0.19) | 9 (0.20) | 16 (0.48) | 26 (0.00) | 40 (0.57) |
| Female, 25, Critical | 6 (0.10) | 9 (0.20) | 17 (0.44) | 28 (0.49) | 42 (0.69) |
| Female, 35, Critical | 6 (0.52) | 10 (0.44) | 19 (0.51) | 31 (0.72) | 44 (0.50) |
| Female, 45, Critical | 6 (0.10) | 10 (0.32) | 20 (0.40) | 33 (0.70) | 45 (0.60) |
| Female, 55, Critical | 5 (0.50) | 9 (0.30) | 20 (0.30) | 34 (0.75) | 45 (0.65) |
| Female, 65, Critical | 4 (0.20) | 8 (0.10) | 18 (0.39) | 34 (0.30) | 45 (0.98) |
| Female, 75, Critical | 3 (0.10) | 6 (0.35) | 15 (0.53) | 31 (0.80) | 45 (0.49) |
| Female, 85, Critical | 2 (0.10) | 5 (0.30) | 9 (0.51) | 24 (1.46) | 42 (1.81) |

Table S8: Quantiles of length of stay in critical state in days, given being in critical, by patient type (sex, age and state at time of hospitalization), based on 20,000 MC paths for each patient type (weighted bootstrap standard error).

| Patient Type | 10% | 25% | 50% | 75% | 90% |
|---|---|---|---|---|---|
| Male, 15, Moderate | 2 (0.43) | 4 (0.54) | 10 (0.84) | 18 (1.30) | 30 (1.75) |
| Male, 25, Moderate | 2 (0.14) | 4 (0.46) | 10 (0.61) | 19 (1.41) | 30 (1.48) |
| Male, 35, Moderate | 2 (0.42) | 6 (0.49) | 11 (0.64) | 21 (1.14) | 31 (1.23) |
| Male, 45, Moderate | 2 (0.16) | 5 (0.33) | 11 (0.65) | 22 (0.91) | 34 (0.11) |
| Male, 55, Moderate | 2 (0.20) | 5 (0.47) | 12 (0.58) | 22 (0.74) | 34 (0.65) |
| Male, 65, Moderate | 2 (0.40) | 5 (0.39) | 12 (0.48) | 22 (0.67) | 31 (0.64) |
| Male, 75, Moderate | 2 (0.48) | 4 (0.14) | 11 (0.47) | 20 (0.56) | 30 (0.50) |
| Male, 85, Moderate | 2 (0.00) | 4 (0.50) | 9 (0.30) | 18 (0.47) | 28 (0.53) |
| Male, 15, Severe | 3 (0.10) | 7 (0.38) | 14 (0.58) | 26 (1.03) | 37 (1.37) |
| Male, 25, Severe | 3 (0.24) | 7 (0.20) | 14 (0.67) | 25 (0.82) | 37 (0.91) |
| Male, 35, Severe | 3 (0.14) | 7 (0.42) | 14 (0.61) | 26 (0.76) | 37 (0.76) |
| Male, 45, Severe | 3 (0.10) | 7 (0.24) | 14 (0.33) | 26 (0.60) | 36 (0.62) |
| Male, 55, Severe | 3 (0.36) | 7 (0.20) | 14 (0.56) | 25 (0.44) | 34 (0.46) |
| Male, 65, Severe | 3 (0.00) | 6 (0.14) | 14 (0.41) | 25 (0.20) | 33 (0.29) |
| Male, 75, Severe | 3 (0.00) | 6 (0.14) | 13 (0.52) | 23 (0.28) | 32 (0.47) |
| Male, 85, Severe | 2 (0.00) | 4 (0.10) | 9 (0.10) | 19 (0.53) | 29 (0.49) |
| Male, 15, Critical | 2 (0.00) | 5 (0.10) | 11 (0.57) | 20 (0.30) | 30 (0.40) |
| Male, 25, Critical | 3 (0.10) | 6 (0.00) | 12 (0.20) | 23 (0.35) | 33 (0.59) |
| Male, 35, Critical | 3 (0.10) | 7 (0.00) | 14 (0.44) | 25 (0.10) | 37 (0.10) |
| Male, 45, Critical | 4 (0.30) | 8 (0.17) | 16 (0.00) | 28 (0.49) | 40 (0.10) |
| Male, 55, Critical | 4 (0.49) | 8 (0.00) | 17 (0.33) | 29 (0.49) | 42 (0.10) |
| Male, 65, Critical | 4 (0.10) | 8 (0.10) | 17 (0.48) | 30 (0.54) | 40 (0.10) |
| Male, 75, Critical | 3 (0.00) | 6 (0.34) | 14 (0.40) | 26 (0.65) | 37 (0.49) |
| Male, 85, Critical | 2 (0.00) | 4 (0.10) | 8 (0.20) | 18 (0.79) | 31 (0.60) |
| Female, 15, Moderate | 1 (0.32) | 4 (0.58) | 9 (0.94) | 19 (0.16) | 30 (2.84) |
| Female, 25, Moderate | 2 (0.40) | 4 (0.47) | 10 (0.65) | 18 (1.15) | 27 (1.94) |
| Female, 35, Moderate | 2 (0.31) | 4 (0.34) | 8 (0.62) | 16 (0.98) | 25 (1.59) |
| Female, 45, Moderate | 2 (0.33) | 4 (0.07) | 8 (0.49) | 17 (0.80) | 24 (1.30) |
| Female, 55, Moderate | 2 (0.46) | 3 (0.00) | 7 (0.52) | 14 (0.73) | 23 (0.91) |
| Female, 65, Moderate | 2 (0.48) | 3 (0.00) | 7 (0.47) | 14 (0.57) | 23 (0.61) |
| Female, 75, Moderate | 2 (0.10) | 3 (0.00) | 7 (0.10) | 14 (0.52) | 22 (0.53) |
| Female, 85, Moderate | 2 (0.00) | 3 (0.14) | 7 (0.44) | 13 (0.41) | 21 (0.53) |
| Female, 15, Severe | 3 (0.00) | 6 (0.14) | 14 (0.58) | 25 (0.75) | 38 (1.01) |
| Female, 25, Severe | 2 (0.00) | 5 (0.14) | 13 (0.34) | 24 (0.65) | 35 (0.93) |
| Female, 35, Severe | 2 (0.00) | 5 (0.50) | 11 (0.20) | 21 (0.61) | 33 (0.72) |
| Female, 45, Severe | 2 (0.14) | 4 (0.00) | 10 (0.43) | 18 (0.56) | 28 (0.61) |
| Female, 55, Severe | 2 (0.29) | 4 (0.00) | 9 (0.33) | 18 (0.42) | 27 (0.50) |
| Female, 65, Severe | 2 (0.00) | 4 (0.46) | 9 (0.00) | 16 (0.26) | 25 (0.40) |
| Female, 75, Severe | 2 (0.00) | 4 (0.00) | 8 (0.10) | 15 (0.10) | 23 (0.54) |
| Female, 85, Severe | 1 (0.00) | 3 (0.00) | 7 (0.00) | 14 (0.40) | 22 (0.41) |
| Female, 15, Critical | 3 (0.00) | 5 (0.10) | 11 (0.20) | 21 (0.42) | 32 (0.60) |
| Female, 25, Critical | 3 (0.46) | 5 (0.00) | 11 (0.10) | 22 (0.55) | 33 (0.10) |
| Female, 35, Critical | 3 (0.10) | 5 (0.00) | 11 (0.00) | 21 (0.17) | 33 (0.00) |
| Female, 45, Critical | 3 (0.10) | 5 (0.10) | 11 (0.42) | 21 (0.50) | 33 (0.57) |
| Female, 55, Critical | 2 (0.10) | 5 (0.10) | 11 (0.45) | 21 (0.29) | 32 (0.61) |
| Female, 65, Critical | 2 (0.10) | 5 (0.21) | 10 (0.20) | 20 (0.32) | 30 (0.65) |
| Female, 75, Critical | 2 (0.00) | 5 (0.50) | 9 (0.00) | 18 (0.38) | 28 (0.22) |
| Female, 85, Critical | 2 (0.00) | 4 (1.00) | 7 (2.00) | 16 (0.65) | 25 (0.41) |

Table S9: Random subset of patients holdout: Mean absolute error (MAE) for predicting number of hospitalized and number of critical patients over each held-out group and over the eight random groups together under the evaluation setups described in subsection S1.3.

| Group Number | Arrival All | Arrival Critical | Snapshot Apr 1st, All | Snapshot Apr 1st, Critical | Snapshot Apr 15st, All | Snapshot Apr 15th, Critical |
|---|---|---|---|---|---|---|
| 1 | 3.79 | 2.07 | 3.52 | 1.83 | 3.78 | 2.73 |
| 2 | 4.09 | 2.03 | 2.03 | 1.79 | 1.73 | 1.97 |
| 3 | 5.99 | 1.81 | 4.20 | 1.19 | 3.37 | 3.72 |
| 4 | 6.35 | 1.58 | 5.46 | 0.81 | 3.39 | 2.06 |
| 5 | 5.32 | 1.41 | 2.19 | 1.22 | 4.89 | 1.17 |
| 6 | 3.48 | 2.14 | 2.32 | 2.27 | 2.84 | 2.11 |
| 7 | 3.94 | 1.04 | 2.32 | 0.71 | 1.63 | 1.03 |
| 8 | 4.86 | 1.33 | 3.14 | 1.90 | 3.39 | 1.04 |
| All | 4.11 | 1.29 | 2.41 | 0.75 | 2.68 | 1.25 |
| Mean (SE) | 4.72 (1.07) | 1.68 (0.40) | 3.15 (1.20) | 1.47 (0.56) | 3.13 (1.07) | 1.98 (0.93) |

Table S10: Random subset of patients holdout: ROC AUC and Brier Score of death prediction and critical-state-visit prediction by held-out group. Critical-state-visit prediction is not including patients started at critical state. The mean ROC AUC and Brier Scores estimates for death prediction over the eight held-out subsets are 0.955 (SE=0.035) and 0.043 (SE=0.011); the respective numbers of visiting-critical prediction are 0.880 (SE=0.040) and 0.049 (SE=0.013).

| | | Death Prediction | | | Visiting-Critical Prediction | | |
|---|---|---|---|---|---|---|---|
| group number | group size | number of deceased | AUC | Brier | number in critical (started in critical) | AUC | Brier |
| 1 | 330 | 26 | 0.966 | 0.047 | 20 (29) | 0.899 | 0.034 |
| 2 | 331 | 29 | 0.936 | 0.056 | 23 (28) | 0.862 | 0.048 |
| 3 | 329 | 19 | 0.980 | 0.038 | 32 (22) | 0.958 | 0.030 |
| 4 | 329 | 20 | 0.978 | 0.028 | 30 (23) | 0.860 | 0.051 |
| 5 | 331 | 24 | 0.977 | 0.030 | 34 (23) | 0.828 | 0.053 |
| 6 | 330 | 28 | 0.958 | 0.057 | 33 (17) | 0.904 | 0.072 |
| 7 | 330 | 19 | 0.876 | 0.038 | 22 (20) | 0.849 | 0.049 |
| 8 | 330 | 28 | 0.971 | 0.052 | 32 (22) | 0.876 | 0.055 |

Table S11: Hospital holdout: Mean absolute error (MAE) for each held-out hospital for predicting number of hospitalized and number of critical patients under the evaluation setups described in subsection S1.3.

| Hospital | Number of Patients | All Patients (MAE) | Critical Patients (MAE) | Snapshot Apr 1st All (MAE) | Snapshot Apr 1st Critical (MAE) | Snapshot Apr 15st All (MAE) | Snapshot Apr 15st Critical (MAE) |
|---|---|---|---|---|---|---|---|
| H1 | 298 | 8.51 | 2.24 | 3.57 | 1.44 | 4.57 | 1.19 |
| H2 | 166 | 6.48 | 0.82 | 4.10 | 0.66 | 3.83 | 1.15 |
| H3 | 142 | 3.91 | 1.55 | 1.57 | 0.64 | 1.14 | 1.29 |
| H4 | 105 | 3.78 | 1.01 | 1.65 | 1.96 | 1.53 | 1.46 |
| H5 | 343 | 5.55 | 1.67 | 2.68 | 1.29 | 5.74 | 2.08 |
| H6 | 111 | 2.11 | 0.75 | 2.02 | 0.63 | 2.15 | 0.92 |
| H7 | 373 | 11.44 | 7.06 | 7.11 | 8.58 | 10.40 | 7.74 |
| H8 | 215 | 5.71 | 2.61 | 7.79 | 1.10 | 2.41 | 3.66 |

Table S12: Hospital holdout: ROC AUC and Brier Score of death prediction and critical-state visit prediction for each held-out hospital. Critical-state-visit prediction does not include patients who started hospitalization at critical state.

| | | Death Prediction | | | Visiting-Critical Prediction | | |
|---|---|---|---|---|---|---|---|
| Hospital | Number of Patients | Number of Deceased | AUC | Brier | Number in Critical (started in Critical) | AUC | Brier |
| H1 | 298 | 16 | 0.856 | 0.043 | 22 (18) | 0.699 | 0.045 |
| H2 | 166 | 5 | 0.986 | 0.020 | 7 (10) | 0.999 | 0.021 |
| H3 | 142 | 11 | 0.976 | 0.048 | 16 (11) | 0.917 | 0.092 |
| H4 | 105 | 8 | 0.946 | 0.053 | 8 (4) | 0.953 | 0.052 |
| H5 | 343 | 26 | 0.982 | 0.044 | 35 (34) | 0.892 | 0.059 |
| H6 | 111 | 8 | 0.913 | 0.047 | 12 (7) | 0.801 | 0.066 |
| H7 | 373 | 26 | 0.970 | 0.046 | 43 (27) | 0.897 | 0.046 |
| H8 | 215 | 23 | 0.959 | 0.039 | 29 (23) | 0.951 | 0.047 |

Table S13: Predicted number of deaths (in-hospital mortality) within a random subset of 330 held-out patients: based on Arrival plus Snapshot prediction results of hypothetical scenarios, prediction starts on the 15th day of the observed arrival process. The numbers are the observed and predicted number of deaths for each hypothetical scenario from hospitalization day up to day $t$ in hospital, $t = 5, 10, \ldots, 35$.

| Day $t$ | Observed | Expected | Younger | Milder | NH Outbreak |
|---|---|---|---|---|---|
| 5 | 7 | 6.5 | 0.6 | 3.7 | 8.2 |
| 10 | 16 | 16.6 | 1.9 | 11.7 | 22.6 |
| 15 | 20 | 20.4 | 2.5 | 15.0 | 28.3 |
| 20 | 23 | 22.8 | 3.0 | 17.3 | 32.1 |
| 25 | 24 | 25.4 | 3.4 | 19.7 | 36.5 |
| 30 | 25 | 25.9 | 3.5 | 20.1 | 37.0 |
| 35 | 26 | 26.6 | 3.7 | 20.6 | 37.7 |

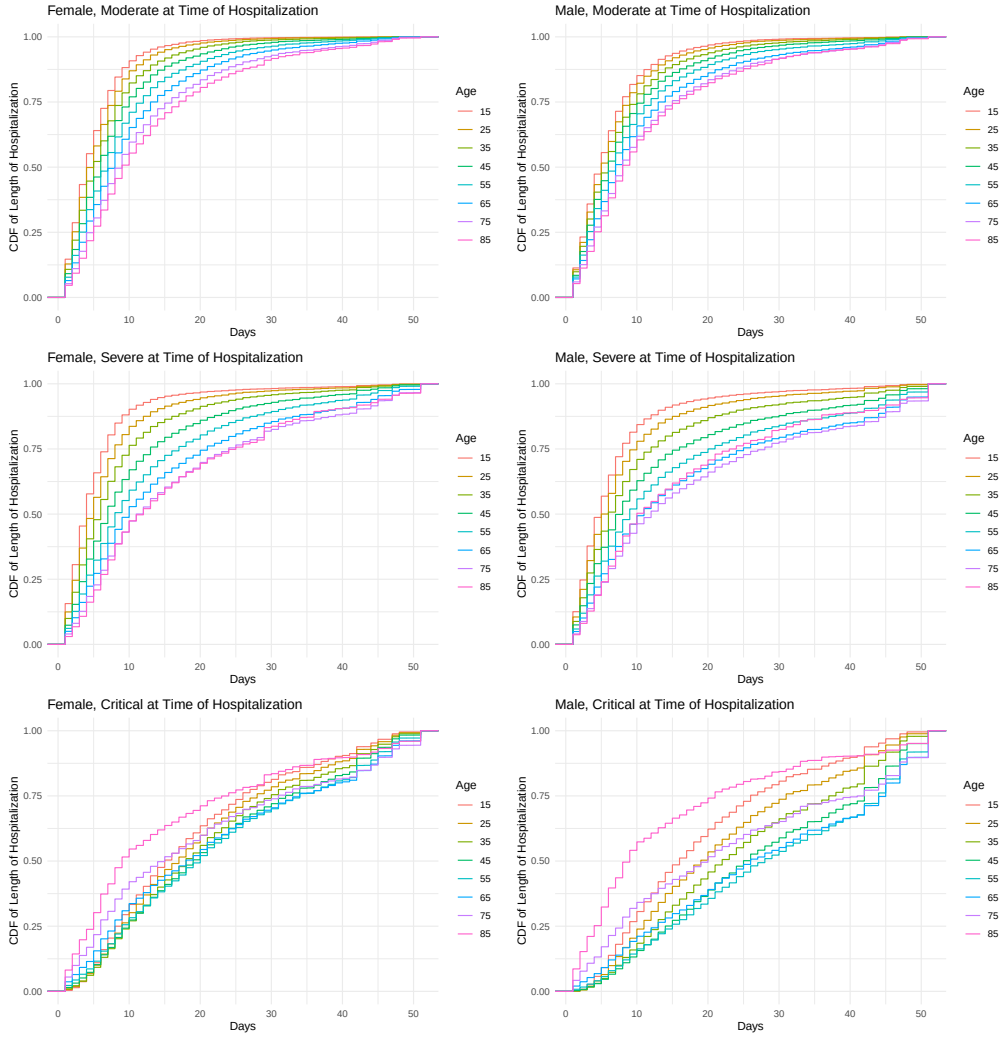


Figure S2: The cumulative distribution function of the length of hospitalization by patient types (sex, state at time of hospitalization and age). Each curve is based on 20,000 MC paths.

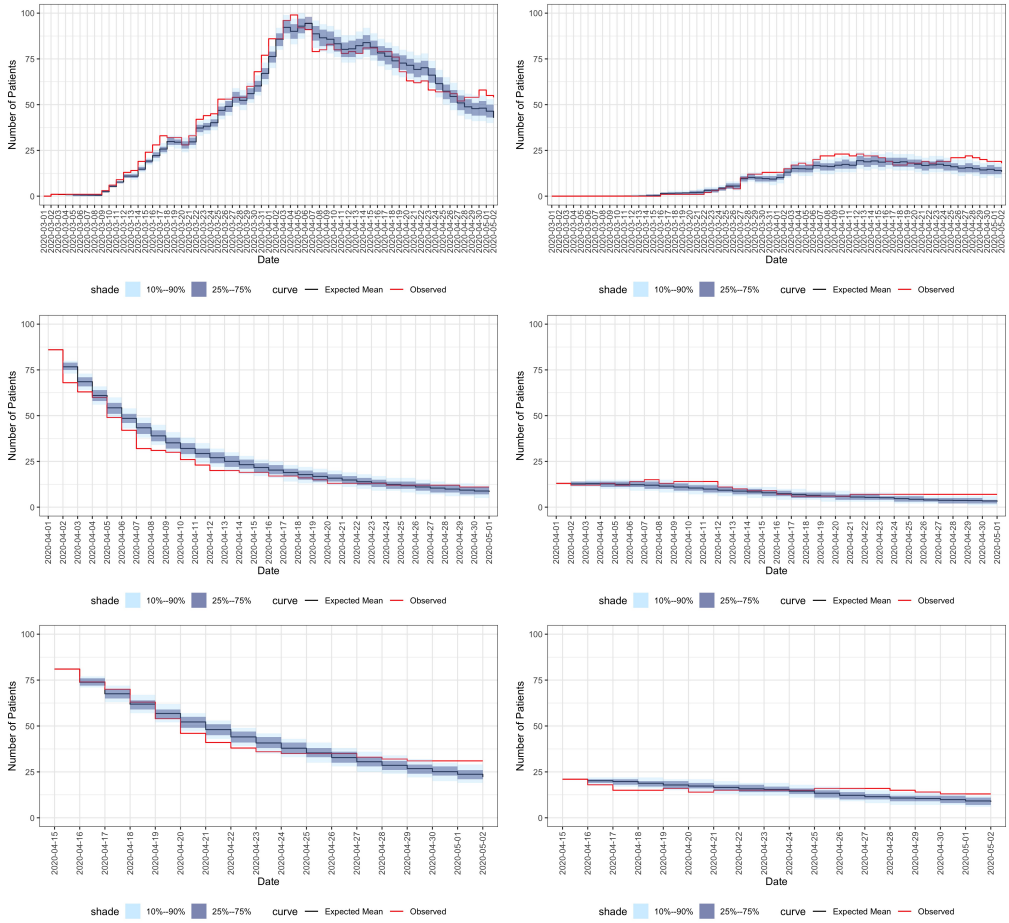


Figure S3: Prediction results of one held-out random group of patients. **Left figures:** utilization predictions for the entire held-out sample. **Right figures:** utilization predictions for critical patients among the held-out sample. **Top figures:** Arrival-type predictions of the entire held-out set based on the observed arrival process. **Middle figures:** Snapshot-type predictions for patients at the hospital on April 1st. **Bottom figures:** Snapshot predictions for patients in the hospital on April 15th. For description of Arrival and Snapshot see Section S1.3.

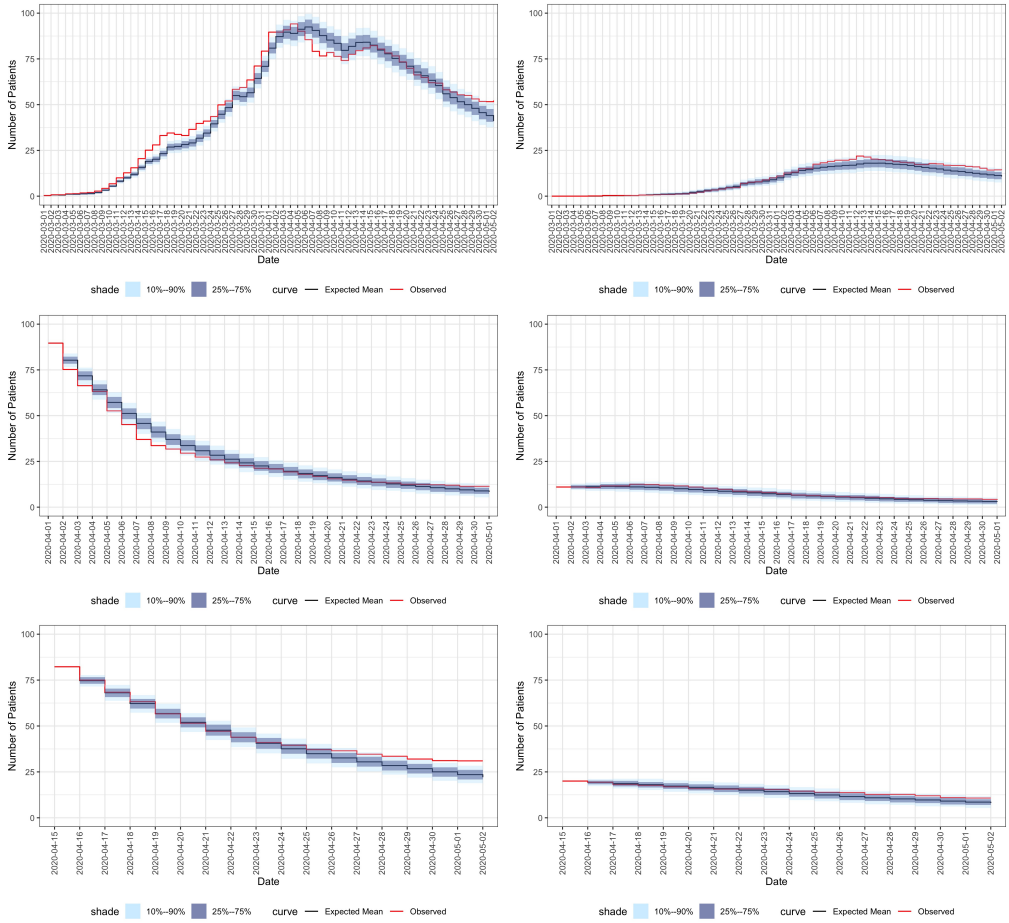


Figure S4: Prediction results summarized over all 8 held-out random groups of patients. **Left figures:** utilization predictions for the entire held-out sample. **Right figures:** utilization predictions for critical patients among the held-out sample. **Top figures:** Arrival-type predictions of the entire held-out set based on the observed arrival process. **Middle figures:** Snapshot-type predictions for patients at the hospital on April 1st. **Bottom figures:** Snapshot predictions for patients in the hospital on April 15th. For description of Arrival and Snapshot see Section S1.3.

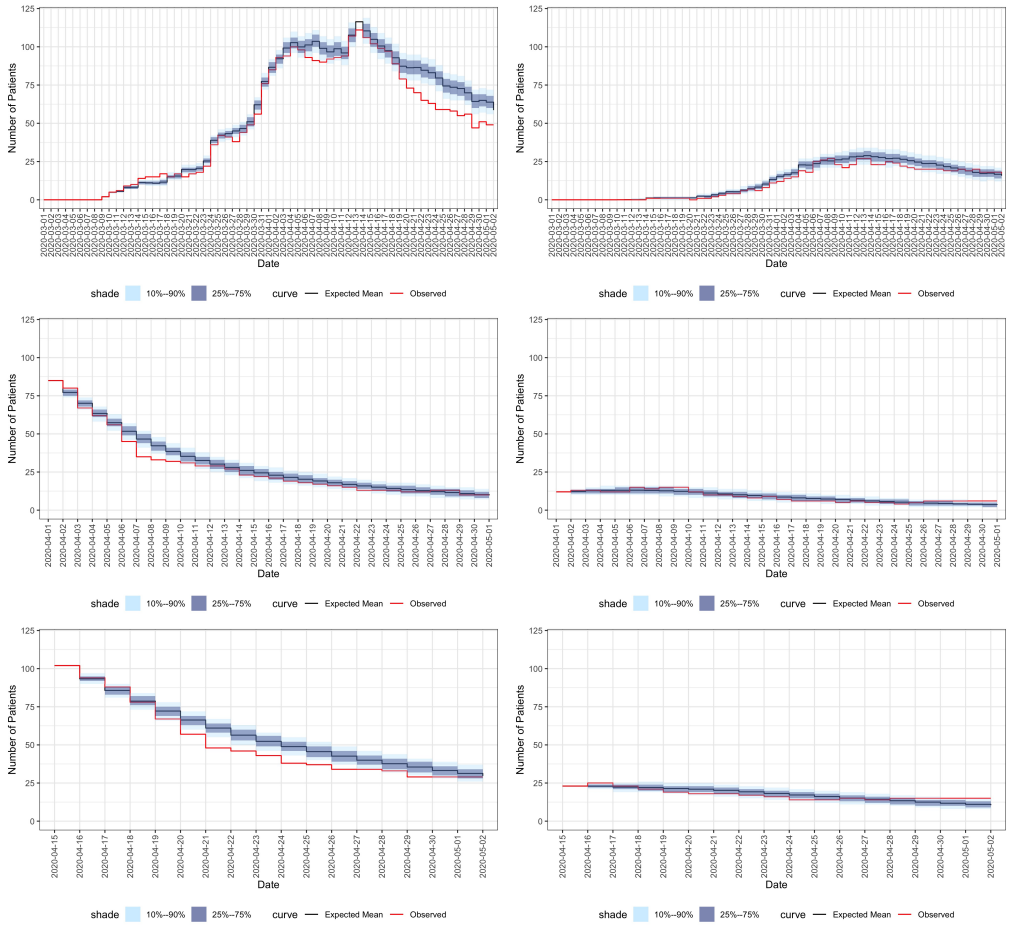


Figure S5: Prediction results - H5 Hospital is held out and predicted based on all other Israeli Hospitals. **Left figures:** utilization predictions for the entire held-out sample. **Right figures:** utilization predictions for critical patients among the held-out sample. **Top figures:** Arrival-type predictions of the entire held-out set based on the observed arrival process. **Middle figures:** Snapshot-type predictions for patients at the hospital on April 1st. **Bottom figures:** Snapshot predictions for patients in the hospital on April 15th. For description of Arrival and Snapshot see S1.3.

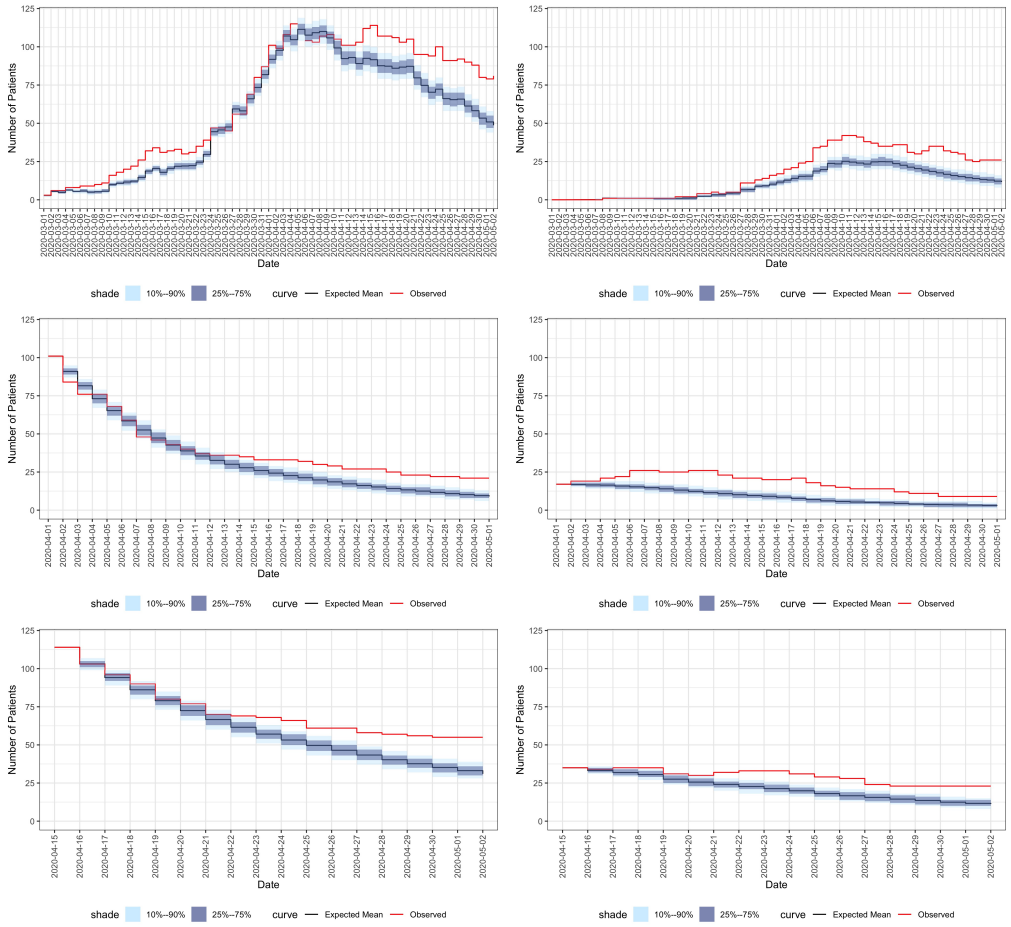


Figure S6: Prediction results - H7 Hospital is held out and predicted based on all other Israeli Hospitals. **Left figures:** utilization predictions for the entire held-out sample. **Right figures:** utilization predictions for critical patients among the held-out sample. **Top figures:** Arrival-type predictions of the entire held-out set based on the observed arrival process. **Middle figures:** Snapshot-type predictions for patients at the hospital on April 1st. **Bottom figures:** Snapshot predictions for patients in the hospital on April 15th. For description of Arrival and Snapshot see S1.3.

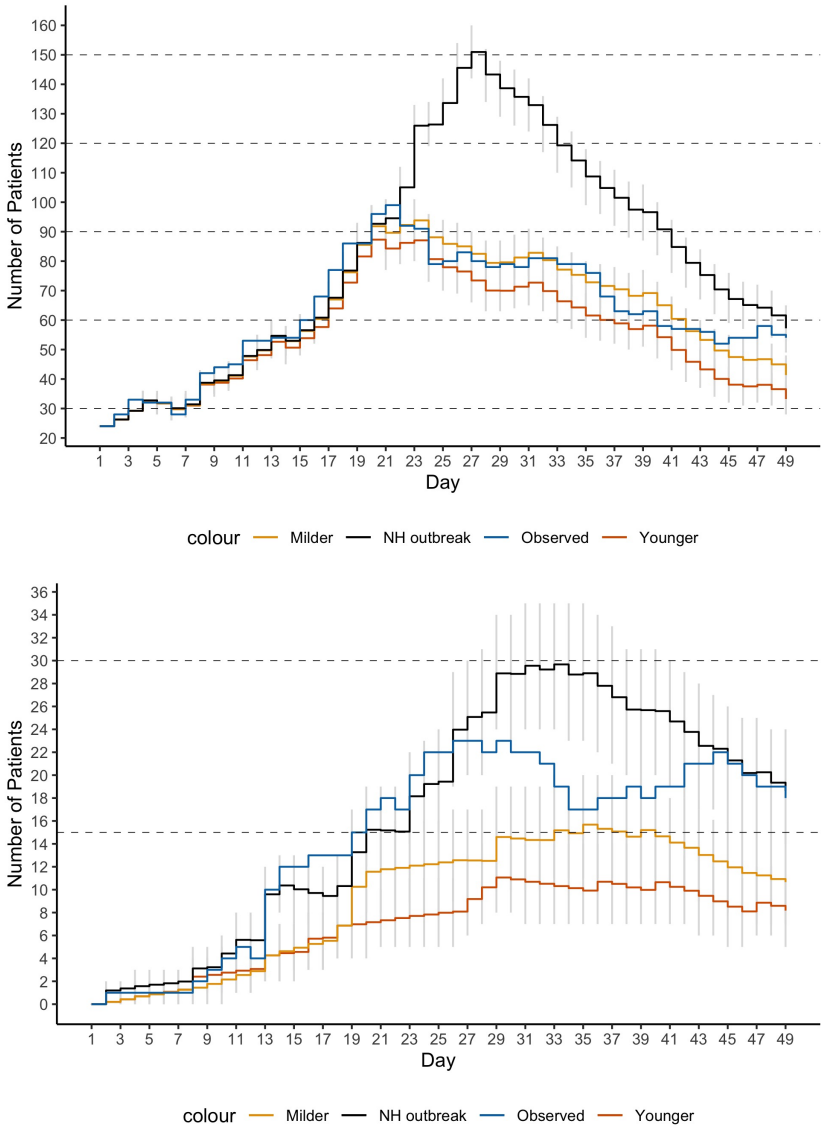


Figure S7: Arrival plus Snapshot prediction results of hypothetical scenarios: 330 random patients were held-out and prediction started on the 15th day of the observed arrival process (denoted as Day 1 in the figures). **Top figure:** utilization predictions based on the entire held-out sample. **Bottom figure:** utilization predictions for critical patients among the held-out sample. Gray vertical lines are point-wise 10%-90% confidence predictions.

# References


Andersen, P. K. & Gill, R. D. (1982). Cox's regression model for counting processes: a large sample study. *The annals of statistics* , 1100–1120.

Andersen, P. K., Hansen, L. S. & Keiding, N. (1991). Non-and semi-parametric estimation of transition probabilities from censored observation of a non-homogeneous markov process. *Scandinavian Journal of Statistics* , 153–167.

Hocking, T. D. (2020). *WeightedROC: Fast, Weighted ROC Curves*. R package version 2020.1.31.

Kalbfleisch, J. D. & Prentice, R. L. (2011). *The statistical analysis of failure time data*, vol. 360. John Wiley & Sons.

Klein, J. P. & Moeschberger, M. L. (2006). *Survival analysis: techniques for censored and truncated data*. Springer Science & Business Media.

Kosorok, M. R. (2007). *Introduction to empirical processes and semiparametric inference*. Springer Science & Business Media.

R Core Team (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Robins, J. M. & Finkelstein, D. M. (2000). Correcting for noncompliance and dependent censoring in an aids clinical trial with inverse probability of censoring weighted (ipcw) log-rank tests. *Biometrics* **56**, 779–788.