Supplementary Online Content


eAppendix
eTable 1. Model parameters
eTable 2. Internal and external validation
eFigure 1. State-transition diagram
eFigure 2. Face validation. Predicted survival by SOFA quartile
eFigure 3. Internal and external validation. Modeled vs. predicted survival for validation cohorts
eReferences

This supplementary material has been provided by the authors to give readers additional information about their work.
Methods

Identification of eligible subjects from MIMIC database

A full description of the MIMIC II database can be accessed at http://mimic.physionet.org/. MIMIC II, version 26, from which the derivation cohort came, included 32,536 patients. 3371 patients were identified by a database search as having cancer. Discharge summaries from these 3371 patients were manually reviewed by 3 attending level physicians (MGS, LAC, PSL) to verify a diagnosis of cancer as stated in the inclusion criteria, resulting in 920 subjects that were ultimately identified. Discharge summaries from these subjects were then reviewed in detail in order to verify baseline subject characteristics. Data was extracted into a standardized form on Epidata Entry Client (v1.4), and included the following variables: cancer location and type; solid vs. hematologic malignancy; cancer stage; whether the cancer was diagnosed during ICU admission; whether the patient was a surgical patient, was admitted after a procedure, or had a surgical complication; where the patient was admitted from; code status on admission; functional status (bedbound or not) on admission; recent chemotherapy administration prior to admission; whether the patient required intubation during the hospital admission, and, if so, whether extubation was successful; whether the patient required dialysis support during the admission, and, if so, whether dialysis could be stopped before discharge; whether the patient required vasopressor support during admission; whether the patient was made do-not-resuscitated, do-not-intubate, or comfort-measures-only (CMO), and, if the latter, whether the patient died after being made CMO; the discharge diagnosis and the location to which the patient was discharged. Data from each validation cohort were extracted in a similar fashion.
Model construction

We constructed a state-transition (Markov) model to simulate the clinical problem. The five mutually exclusive states any patient could enter were: aggressive care in the ICU, comfort measures, inpatient (non-ICU) care, home, and dead (eFigure 1). Because fewer than 2% of the patients in our cohort were readmitted to the ICU within 30 days after a discharge home, we assumed that patients who made it home remained home (ie “Home”, as well as “Dead”, were both absorbing states).

State-transition probabilities, baseline mortality rates, cutoffs for ICU discharge, and ICU mortality rates were calculated from MIMIC-II data using Kaplan-Meier time-to-event analysis, conditional on individual SOFA scores. SOFA scores were calculated for the 884 of these patients who had at least two consecutive SOFA scores recorded, on the day of and the day after ICU admission. Daily transition probabilities were modeled conditional on an individual’s SOFA score on that particular day. Using Kaplan-Meier time-to-event analysis, transition probabilities from each state to each possible destination state (eg, “Aggressive care” to “Inpatient”, or “Inpatient” to “Home”) were calculated from the MIMIC-II database. When data were missing from the MIMIC-II database, probabilities were extrapolated from patients within that state with neighboring SOFA scores. For patients remaining in the model past the first three days, the sample for these calculations became small. As such, transition probabilities from the first day were used. This has the potential of under- or overestimating survival for patients. As a result, sensitivity analyses were conducted on this assumption, as documented in the next section. Separate SOFA-specific transition probability matrices were calculated for each of the three non-absorbing states (“Aggressive Care,” “Comfort-measures only”, and “Inpatient Care”).

Patients obviously do not remain at the same SOFA score each day that they are in an ICU—their condition may improve or worsen, and capturing this improvement was necessary to construct a valid model. Changes in SOFA score were calculated on each day, and were conditional on the patient’s SOFA score the day before. The calculation proceeded as follows: each patient in the model faced a probability of improvement or worsening of their disease, based on their SOFA score the day before. Conditional on
improving or worsening, a new SOFA score was then drawn from a bootstrapped distribution of changes in SOFA score in all patients in the cohort who similarly improved or worsened. Daily state-transition probabilities were thus determined, using Kaplan-Meier time-to-event analyses and fitted polynomial regression curves with distributions around the probabilities determined from the bootstrapped samples. Patients were discharged from the ICU when their SOFA score had reached either a standard cutoff or had improved to a predetermined percentage of their admission SOFA score. Re-admission to the ICU was allowed during a given hospital stay; probabilities of re-admission were calculated from data available in MIMIC-II.

Because of the short time-course of the model, no discounting was applied to the outcomes. In keeping with the short cycle length, half-cycle correction was not applied.

Scenario and sensitivity analyses

The following model assumptions were explicitly examined in sensitivity analyses:

1. An assumption that daily state-transition probabilities only depended on a patient’s SOFA score, and not on how long the patient had been in the ICU.
2. An assumption that a one-day difference in mean survival was clinically significant to patients and families.
3. An assumption that patients transitioned to comfort-measures-only care immediately on the cessation of their trial.

How these assumptions were tested is discussed below.

In the base-case analysis, daily state-transition probabilities, conditional on disease severity, were assumed to be constant over the course of a patient’s ICU stay. To test the assumptions around the trajectory of a patient’s illness, adjustments on $p_i$, representing the probability of improving, and $p_w$, representing the probability of worsening, were performed. In the first sensitivity analysis, patients were more likely to worsen with prolonged ICU stays, so:
where $t$ represents the number of days the patient has spent so far in the ICU. The modified worsening probability, $p_w^*$, was bounded above at 1.0. To test the counter assumption, that the probability of any change in disease severity—for better or for worse—decreased with each successive day in the ICU, probabilities of improving and worsening were scaled as follows:

$$
p_i^* = p_i \left(1 - \frac{t}{30}\right)
$$

$$
p_w^* = \min\left[1, p_w \left(1 + \frac{t}{30}\right)\right]
$$

Changing criteria for equivalence did not require an alteration in the model structure.

Finally, to test assumptions around how patients transitioned out of a trial of ICU, we altered the construction of the model. In the base-case analysis, a patient undergoing a trial of ICU would stay in the ICU until she died, was discharged to the inpatient floor, or passed a predetermined number of days in ICU after which she was made CMO. In this sensitivity analysis, when the ICU trial expired, the patient’s severity was reassessed every day; the patient was made CMO if and only if she had not died, been discharged to the inpatient ward, or improved beyond a certain percentage of her admission SOFA score. If she had improved, she was kept on aggressive care. We tried cutoffs of 90%, 75%, 50%, and 25%—that is, at a cutoff of 90%, a patient who was admitted with a SOFA score of 20 had to improve to a SOFA score of 18 or better to remain in the ICU at the end of a trial.
Probabilistic sensitivity analysis was done for each possible trial length from 1 day to 29 days. Parameter uncertainty was modeled as in eTable 1. When probabilities were conditional on SOFA score, they were modeled as conjugate prior beta distributions, centered on the probabilities derived from the MIMIC dataset and with a wide variance. Each individual SOFA-specific probability is not shown in eTable 1.

For each possible trial length, 100 samples were taken from these distributions, and 10,000 patients per SOFA score were run for each of the 100 realizations of the distributions. Sample size was limited by computing power. Mean survival was calculated. With the posterior 95% confidence intervals generated by probabilistic sensitivity analysis, trials of ICU were considered equivalent to aggressive care when no statistically significant difference was found in mean survival between groups of patients treated under these two strategies.

Model validation

Face validation

As patients face increasingly severe disease, a well-constructed model would be expected to predict worsening 30-day survival. The model performs as expected in this case (eFigure 2).

Internal validation

Before further analysis was performed, the model underwent internal and external validation. Internal validation was performed on both the Aggressive Care and the CMO treatment strategies. For the former, Kaplan-Meier thirty-day overall survival curves were plotted for the patients for whom complete data were available. Predicted Kaplan-Meier survival curves were also constructed for an equally-sized simulated cohort of patients who underwent Aggressive Care in the model. These were compared, and p-for-difference values were calculated using the log-rank method and the Peto and Peto modification of the
Wilcoxon method. The CMO model strategy was validated by comparing the average modeled survival of a patient in CMO care with the survival reported in the literature.

External validation

External validation was performed by studying critically ill cancer patients meeting our inclusion criterion in three external cohorts; 349 patients admitted to BIDMC between 2008 and 2012, 158 patients admitted to Brigham and Women’s Hospital between 2007 and 2015, and 117 patients admitted to King Abdulaziz Medical City between 2002 and 2013. These represent two tertiary academic hospitals in Boston and one in Riyadh, Saudi Arabia, respectively. Sample size calculations showed that these validation cohorts had 80% power to detect a survival difference of 15.0%, 18.6%, and 15.5% respectively.

The model was explicitly not recalibrated to these data sets. A simulated cohort of the size and disease-severity distribution of the two external datasets was run through the model. Survival analysis proceeded as with the internal validation.

Results

Model validation

Modeled and actual survival curves for the internal and external validation cohorts are given in eFigure 3. Statistical analysis of differences in survival between the internal and external validation sets and modeled cohorts are shown in eTable 2. No statistically significant difference was seen between modeled survival estimates and actual survival in any of the four validation cohorts.

Sensitivity and scenario analysis

The model assumes that the probability of improving or worsening is only conditional on a patient’s SOFA score the day prior and not on how many days the patient has spent in the ICU. Relaxing that
assumption, and allowing time spent in ICU to affect probabilities of worsening and improving, decreased goodness of fit between predicted survival and actual survival.

Probabilistic sensitivity analysis was conducted, and optimal trial lengths were re-calculated as described above. For all admission SOFA scores, a statistically non-significant difference in survival between trial of ICU and aggressive care could only be established with trial lengths of longer than 10 days. When corrected for multiple comparisons, patients with a SOFA score of 20 or higher could be treated with a trial as short as nine days. All other levels of severity continued to require trials of at least 10 days to achieve non-differential mean survival.
### eTable 1. Model parameters.

Constant (non-severity-dependent) model parameters, with base-case (mean) values and distributions for probabilistic sensitivity analysis. See text for details. Daily probabilities come from analyses of the MIMIC-II database cohort.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Distribution</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Daily probability of improving given CMO status</td>
<td>Beta</td>
<td>0.00073</td>
<td>0.0000854</td>
</tr>
<tr>
<td>Daily probability of discharge home from an inpatient floor</td>
<td>Beta</td>
<td>0.0682</td>
<td>0.00252</td>
</tr>
<tr>
<td>Daily probability of return to ICU from an inpatient floor</td>
<td>Beta</td>
<td>0.00281</td>
<td>0.000167</td>
</tr>
</tbody>
</table>
eTable 2. Internal and external validation. This table contains *p*-values for difference in actual versus modeled thirty-day survival for one internal and three external validation datasets. The log-rank significance test puts more weight on survival differences occurring at 30 days, while the Peto and Peto modification to the Wilcoxon test puts more weight on differences occurring at day 1.

<table>
<thead>
<tr>
<th>Cohort</th>
<th><em>p</em>, log-rank (<em>ρ</em> = 0)</th>
<th><em>p</em>, Peto and Peto (<em>ρ</em> = 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIMIC II (design cohort, N = 884)</td>
<td>0.501</td>
<td>0.322</td>
</tr>
<tr>
<td>MIMIC III (N = 349)</td>
<td>0.386</td>
<td>0.377</td>
</tr>
<tr>
<td>King Abdulaziz (N = 117)</td>
<td>0.970</td>
<td>0.563</td>
</tr>
<tr>
<td>Brigham and Women’s (N = 159)</td>
<td>0.303</td>
<td>0.444</td>
</tr>
</tbody>
</table>
Patients entered the model in either the Aggressive Care state or the CMO state. Under the Trial of ICU strategy, patients transitioned from Aggressive Care to CMO if they were still in the ICU after the designated trial length elapsed. Solid arrows represent probabilistic transitions among states. See text for further details.
eFigure 2. Face validation. Predicted survival by SOFA quartile

ICU survival:
Modeled by quartile

Shaded region represents 95% CI.

© 2015 American Medical Association. All rights reserved.
Kaplan Meier curves based on observed data compared to model predictions. Data from a) 920 cancer patients from Beth Israel Deaconess Medical Center admitted between 2001 and 2007 (MIMIC II cohort) b) 349 cancer patients from Beth Israel Deaconess Medical Center admitted between 2008 and 2012 (MIMIC III cohort) c) 158 cancer patients from Brigham and Women’s Hospital admitted between 2007 and 2015, and d) 117 cancer patients from King Abdulaziz Medical City admitted between 2002 and 2013. No statistically significant differences in survival were noted between model predictions and observed survival in all cohorts.
eReferences

