

Forecasting immune effector cell-associated neurotoxicity syndrome after chimeric antigen receptor t-cell therapy

Yalda Amidi ^{1,2,3}, Christine A Eckhardt,^{1,2,3} Syed A Quadri,^{1,2,3} Preeti Malik,^{1,2,3} Marcos Santana Firme,^{1,2,3} Daniel K Jones,^{1,2,3} Aayushee Jain,^{2,3} Husain H Danish,^{1,4} Daniel B Rubin,^{1,2,3,5} Caron A Jacobson,⁶ Sydney S Cash,^{1,2,3} Jong Woo Lee,^{1,4} Jorg Dietrich,^{1,2} M Brandon Westover^{1,2,3,5}

To cite: Amidi Y, Eckhardt CA, Quadri SA, *et al.* Forecasting immune effector cell-associated neurotoxicity syndrome after chimeric antigen receptor t-cell therapy. *Journal for ImmunoTherapy of Cancer* 2022;**10**:e005459. doi:10.1136/jitc-2022-005459

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/jitc-2022-005459>).

YA and CAE are joint first authors.

Accepted 31 October 2022



© Author(s) (or their employer(s)) 2022. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹Harvard Medical School, Boston, Massachusetts, USA

²Department of Neurology, Massachusetts General Hospital, Boston, Massachusetts, USA

³Clinical Data Animation Center (CDAC), Massachusetts General Hospital, Boston, MA, USA

⁴Department of Neurology, Brigham and Women's Hospital, Boston, MA, USA

⁵Cancer Center for Brain Health, Massachusetts General Hospital, Boston, MA, USA

⁶Dana-Farber Cancer Institute, Boston, Massachusetts, USA

Correspondence to

Dr Yalda Amidi;
yamidi@mgh.harvard.edu

ABSTRACT

Background Immune effector cell-associated neurotoxicity syndrome (ICANS) is a clinical and neuropsychiatric syndrome that can occur days to weeks following administration chimeric antigen receptor (CAR) T-cell therapy. Manifestations of ICANS range from encephalopathy and aphasia to cerebral edema and death. Because the onset and time course of ICANS is currently unpredictable, prolonged hospitalization for close monitoring following CAR T-cell infusion is a frequent standard of care.

Methods This study was conducted at Brigham and Women's Hospital from April 2015 to February 2020. A cohort of 199 hospitalized patients treated with CAR T-cell therapy was used to develop a combined hidden Markov model and lasso-penalized logistic regression model to forecast the course of ICANS. Model development was done using leave-one-patient-out cross validation.

Results Among the 199 patients included in the analysis 133 were male (66.8%), and the mean (SD) age was 59.5 (11.8) years. 97 patients (48.7%) developed ICANS, of which 59 (29.6%) experienced severe grades 3–4 ICANS. Median time of ICANS onset was day 9. Selected clinical predictors included maximum daily temperature, C reactive protein, IL-6, and procalcitonin. The model correctly predicted which patients developed ICANS and severe ICANS, respectively, with area under the curve of 96.7% and 93.2% when predicting 5 days ahead, and area under the curve of 93.2% and 80.6% when predicting the entire future risk trajectory looking forward from day 5. Forecasting performance was also evaluated over time horizons ranging from 1 to 7 days, using metrics of forecast bias, mean absolute deviation, and weighted average percentage error.

Conclusion The forecasting model accurately predicts risk of ICANS following CAR T-cell infusion and the time course ICANS follows once it has begun. Cite Now

INTRODUCTION

Chimeric antigen receptor (CAR) T-cell therapy (CAR T-cell) has transformed the treatment of hematological malignancies, with primary approvals of CD19-directed or BCMA-directed CAR T-cell therapy for relapsed/refractory large B-cell lymphoma,

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ A prior multivariable logistic regression model predicted the occurrence of immune effector cell-associated neurotoxicity syndrome (ICANS) on days 3–5 with lower accuracy than our model.

WHAT THIS STUDY ADDS

⇒ Data from 199 patients receiving chimeric antigen receptor-T cell therapy were used to develop a forecasting model. The model provides well calibrated probabilities of ICANS onset and accurately predicts ICANS time course 1–7 days ahead.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ With further external validation, our forecasting model may be useful for triaging and resource allocation and may allow many patients to be safely discharged from the hospital earlier that is currently possible.

B cell acute lymphoblastic leukemia, follicular lymphoma, mantle cell lymphoma, and multiple myeloma. However, CAR T-cell therapy carries risk for complications. These include cytokine release syndrome (CRS) and immune effector cell-associated neurotoxicity syndrome (ICANS), which remain significant causes of CAR T-cell related morbidity and mortality.¹ Among trials of the five CD19-targeted CAR T-cell products, 20%–70% of patients experienced grade 1 or higher ICANS which may present as encephalopathy, aphasia, focal weakness, numbness, apraxia, seizures, or in rare cases, cerebral edema, and death.^{2–9} Previous studies have shown that ICANS is only partially responsive to available treatments, and patients often require prolonged hospitalization and supportive care.^{10–13} Steroids remain the mainstay of treatment for ICANS, which may affect the efficacy of CAR T-cell therapy¹⁴ and carry their own profile of adverse effects.

Prediction of which patients will develop ICANS, when, and what course ICANS will follow, remains challenging. The incidence and onset timing of ICANS is variable among cell products and malignancies, although it usually occurs 3–10 days after cell infusion. Accurate prediction of which patients will develop ICANS and the onset time and subsequent time course have the potential to not only affect management with steroids, but also to reduce unnecessary medical expenses and service, since those who do not develop ICANS may be hospitalized for an unnecessarily long duration.⁹ Prior research has uncovered risk factors associated with ICANS, including pre-existing neurological and medical disease, high malignancy burden, elevated intensity of lymphodepletion, high peak CAR T-cell expansion, and early and severe CRS.^{2 13} Several inflammatory biomarkers have been associated with increased occurrence and severity of ICANS, including C reactive protein (CRP), IL-6, and ferritin, but have limited specificity given their co-occurrence with CRS.^{2 9 11–13 15–18} Multivariable models have been developed to predict the risk of ICANS using both clinical factors and biomarkers,^{9 19} but do not attempt to predict onset timing or time course. In the present study, we addressed this gap by developing a statistical model to predict the risk and day of onset of ICANS, and to forecast ICANS scores up to 7 days into the future.

METHODS

Study setting, participants, and clinical data

We conducted a retrospective observational cohort study of patients undergoing CD-19 or BCMA targeted CAR T-cell therapy from April 2015 to February 2020 at Brigham and Women's Hospital (Boston, Massachusetts, USA). We used two existing cohorts, the first a group of patients at high risk for ICANS (patients who had undergone EEG monitoring from 2016 to 2020 because of concern for ICANS following CAR T-cell infusion), and the second a control group of patients who received CAR T-cell therapy at BWH from 2015 to 2019 that did not develop ICANS.

For all patients who developed ICANS during hospitalization, the ICANS grade was determined for each day of hospitalization through chart review by two independent graders. Each grader used oncology, neurology, and nursing notes to determine the Immune Effector Cell-Associated Encephalopathy score (0–10), which was combined with scores in four other neurological domains to determine the final ICANS grade (0–4, with 4 most severe) in accordance with American Society for Transplantation and Cellular Therapy (ASTCT) consensus grading.²⁰ All patients had ICANS=0 at the time of hospital discharge, and we assumed ICANS remained absent after discharge. Demographic information, oncological diagnosis, vital signs (max temperature per day) and laboratory values, where available (CRP, ferritin, white blood cell (WBC), IL-6, procalcitonin), were collected via review of the electronic medical record. CRS

incidence, day of onset, and peak grade were identified via chart review and by calculating CRS grade based on its constituent factors pulled from the electronic medical record (presence and degree of hypotension, temperature >38°C, need for supplementary oxygen and/or positive pressure ventilation) according to ASTCT consensus grading guidelines.²⁰

Forecasting model development

Leave-one-patient-out cross validation

We developed a Hidden Markov model (HMM) to forecast ICANS scores (figure 1). An HMM is a probabilistic model that has an internal state that cannot be observed directly, but is instead inferred through some probabilistic function of the observed data. To avoid overly optimistic (biased) estimates of model performance, we used a leave-one-patient-out cross validation approach for model training and evaluation. During model training, each of the 199 patients was 'held out', and an HMM was developed only using data from the remaining 198 patients. HMM training included estimating state transition and conditional observation probability matrices. During testing, the HMM model was tested on the held-out patient, and performance statistics were calculated by comparing the forecast results with the observed ICANS time course for that patient.

HMM structure and parameter estimation

To construct our HMM model we first defined four hidden substates s_1, s_2, \dots, s_N based on observable outputs (ICANS scores), as follows:

- ▶ $s_1 = 1$: pre-ICANS (ICANS=0).
- ▶ $s_2 = 2$: rising phase: ICANS>0 has begun and tends to rise until reaching a peak.
- ▶ $s_3 = 3$: falling phase: ICANS>0, but after the peak value.
- ▶ $s_4 = 4$: post-ICANS (ICANS=0 following the falling phase).

In terms of these substates and the most recent available ICANS observation, we defined the hidden state of the patient at time t as $Z_t = (S_t, y_{t-1})$, where $t = 1, 2, \dots, T$ and y_t stands for the ICANS score at time t . Z_t can take $M=20$ different values (figure 1B), z_1, z_2, \dots, z_M , and we denote the probability of transition from state z_m to state z_n as $a_{mn} \triangleq \Pr[Z_t = z_n | Z_{t-1} = z_m]$, $m, n = 1, 2, \dots, M$. Observations (ICANS scores) and hidden states are related through the conditional probability $b_{km} \triangleq \Pr[y_t = k | Z_t = z_m]$, where $k = 1, 2, 3, 4, 5$ are the five levels of ICANS scores.

With these definitions, the state transition matrix A and conditional observation probability matrix B dimensions are 20×20 and 20×5 , respectively. Values of these matrices were estimated nonparametrically from the training data,^{21 22} with the exception of the transition marking the initial onset of ICANS, which corresponds to the transition from $Z_{t-1} = 1$ to $Z_t = 6$ (see figure 1B). This transition was modeled parametrically, as described in the next section.

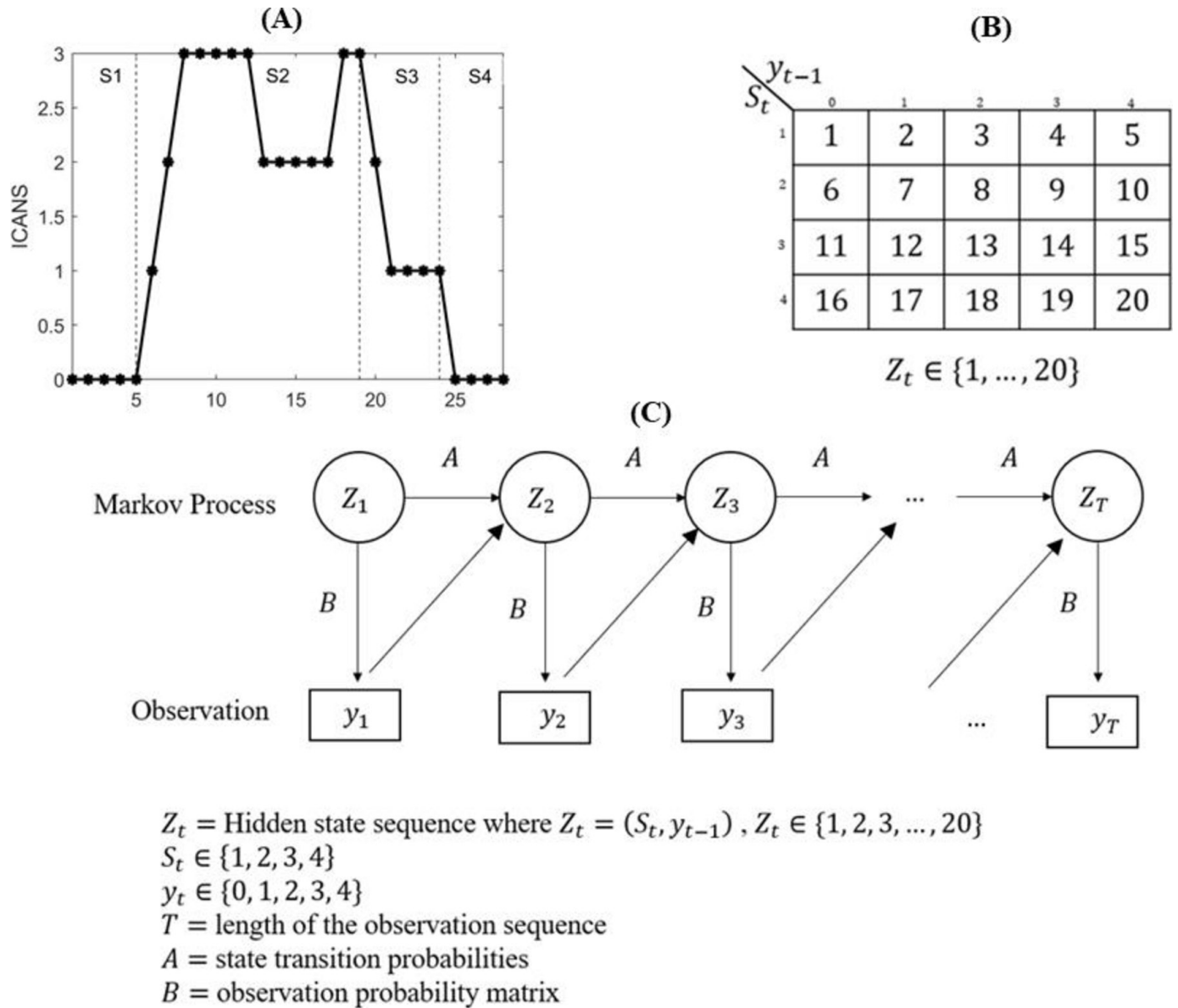


Figure 1 Hidden Markov Model diagram: (A) One example of observed ICANS score along with four hidden substates (B) The hidden state sequence Z_t contains 20 distinct values (C) Markov process and observation, (A) shows the state transition probability 20×20 matrix, and (B) shows observation probability 20×5 matrix. ICANS, immune effector cell-associated neurotoxicity syndrome.

Modeling ICANS onset

Prior studies suggest that the probability of developing ICANS, the time course, and severity of ICANS depend on baseline characteristics like age and clinical data like laboratory values or vital signs. Therefore, in developing the HMM model we modeled the initial transition from ICANS=0 to ICANS>0 ($Z_{t-1} = 1 \rightarrow Z_t = 6$) as a function of time using a pooled Logistic Regression (PLR) approach. Ten predictors (input features) were included in the model: functions of time since CAR T-cell infusion including t , t^2 , and $g(t)$, age, maximum daily body temperature, daily serum CRP, ferritin, IL-6, procalcitonin, and WBC. These clinical and laboratory features have been previously shown to consistently correlate with risk of neurotoxicity,^{9 10 12 13} and are routinely measured

in most patients who receive CAR T-cell treatment. We defined the PLR model to have the following form:

$$\Pr[Z_t = 6 | Z_{t-1} = 1] = \frac{1}{1 + \exp[-(w_0 + w_1 x_t^{(1)} + \dots + w_{10} x_t^{(10)})]} = \frac{1}{1 + \exp[-w^T x_t]} \quad (1)$$

$$\sum_{i=0}^{10} |W_i| \leq \lambda, \lambda > 0$$

where λ is a non-negative tuning parameter that controls the sparsity of the model (ie, the number of coefficients with a value of zero) and is selected by 10-fold internal cross validation. In this equation, w^T is the transposed vector of model coefficients, w_0 is a constant, and $x_t^{(i)}$ $i = 1, \dots, 10$ are the predictor variables described above. $x_t^{(1)}$ is a function of time that models the average profile of ICANS across the population. To model this

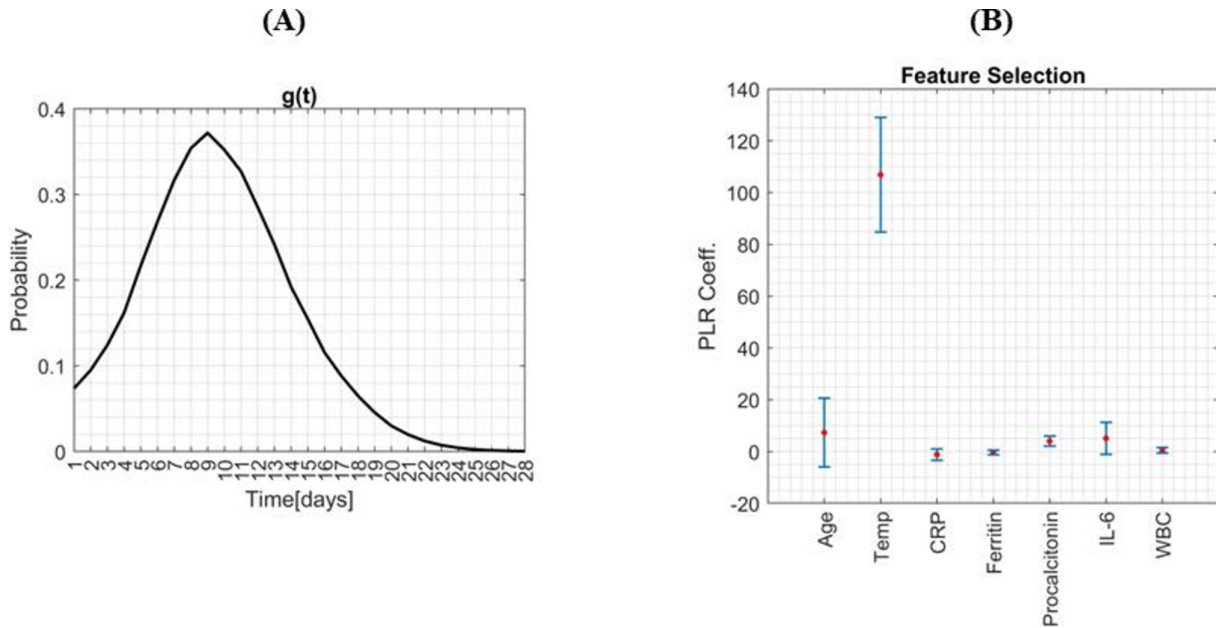


Figure 2 (A) Function of time used in PLR as a time dependent predictor variable, (B) Feature selection using PLR: the estimated coefficients for clinical predictors with the 95% CIs. CRP, C reactive protein; PLR, pooled logistic regression; WBC, white blood cell.

function, we first calculated the empirical probability of ICANS onset vs time in the training data (P_t^{rain}) and then fit a function $g(t)$ to it:

$$x_t^{(1)} \triangleq g(t) = \frac{1}{(1 + \exp(-f(t)))} \quad (2.a)$$

$$f(t) = a + bP_t^{rain} + c_1(t - t_b) + c_2(t - t_b)^2 \quad (2.b)$$

where $\{a, b, c_1, c_2\}$ are free parameters and $t_b = 9$ is the median time of ICANS onset in our cohort. The free parameters were estimated using the Levenberg-Marquardt algorithm.²³ According to the function $g(t)$, figure 2A, the risk of ICANS tends to increase over the time interval $t \in [0, 9]$ then decreases to zero at $t = 20$.

HMM FORECASTING

To forecast ICANS scores, we compute $Pr[Z_t | y_{1:t}]$ by applying Bayes rule sequentially using the HMM forward algorithm.^{21,24} The forward algorithm consists of two alternating steps: one-step ahead prediction and updating. The one-step ahead prediction density is:

$$Pr[Z_t = z_m | y_{1:t-1}] = \sum_{n=1}^M Pr[Z_t = z_m | Z_{t-1} = z_n] Pr[Z_{t-1} = z_n | y_{1:t-1}] \quad (3)$$

Then, in the update step we update the density based on data observed at time t :

$$\alpha_t(z_m) \triangleq Pr[Z_t = z_m | y_{1:t}] = Pr[Z_t = z_m | y_t, y_{1:t-1}] \frac{1}{N_t} Pr[Z_t = z_m | y_{1:t-1}] Pr[y_t | Z_t = z_m, y_{1:t-1}] \frac{1}{N_t} Pr[Z_t = z_m | y_{1:t-1}] Pr[y_t \vee Z_t = z_m], \quad (4)$$

where in the last line we have used the fact that y_t is independent of $y_{1:t-1}$ given Z_t , and the normalization constant N_t , is given by:

$$N_t \triangleq Pr[y_t | y_{1:t-1}] = \sum_{m=1}^M Pr[Z_t = z_m | y_{1:t-1}] Pr[y_t | Z_t = z_m] \quad (5)$$

Using equations (3) and (4), we can write one complete iteration of the predict-update cycle as:

$$\alpha_t(z_m) \propto Pr[y_t | Z_t = z_m] \sum_{n=1}^M Pr[Z_t = z_m | Z_{t-1} = z_n] \alpha_{t-1}(z_n) \quad (6)$$

To predict future states given past observations, we next compute $Pr[Z_{t+h} | y_{1:t}]$, where $h > 0$ is the prediction horizon. We perform this calculation by applying the transition matrix to the current distribution over hidden states $Pr[Z_t | y_{1:t}]$ as follows:

$$Pr[Z_{t+h} | y_{1:t}] = \sum_{z_{t+h-1}} \sum_{z_{t+h-2}} \dots \sum_{z_t} Pr[Z_{t+h} | Z_{t+h-1}] Pr[Z_{t+h-1} | Z_{t+h-2}] \dots Pr[Z_{t+1} | Z_t] Pr[Z_t | y_{1:t}] \quad (7)$$

Finally, we use the quantity $Pr[Z_{t+h} | y_{1:t}]$ to make predictions about future ICANS scores using:

$$Pr[y_{t+h} | y_{1:t}] = \sum_{z_{t+h}} Pr[y_{t+h} | Z_{t+h}] Pr[Z_{t+h} | y_{1:t}] \quad (8)$$

Evaluation forecast accuracy

To evaluate the quality of forecasts, we split ICANS scores of all patients into training and testing sets, using all but one patient's data for model training. We then run the forward algorithm (equations (3–8)) on data from the one patient left out to forecast ICANS scores h days ahead, where h is the forecast horizon. We considered values of h ranging from 1 to 7.

We quantified the quality of forecasts using the following standard metrics²⁵: forecast bias, mean average deviation (MAD), and weighted average percentage error (WAPE):

$$\text{Bias} \triangleq \frac{\sum_{i=1}^T (y_i - \hat{y}_i)}{T} \quad (9.a)$$

$$\text{MAD} \triangleq \frac{\sum_{i=1}^T (y_i - \hat{y}_i)}{T} \quad (9.b)$$

$$\text{WAPE} \triangleq \frac{\sum_{i=1}^T (y_i - \hat{y}_i)}{\sum_{i=1}^T |y_i|} \quad (9.c)$$

In equation (9.a), forecast bias is the difference between forecast (y_t) and observed ICANS scores. If the forecast overestimates, the forecast bias is positive; for underestimates the forecast bias is negative. MAD measures the magnitude, on average, of forecasting errors. WAPE, also referred to as the MAD/mean ratio, weights the error by the total true ICANS scores. Similar to MAD, a good forecasting method has smaller WAPE.²⁵

Evaluation of binary predictions of ICANS/severe ICANS

In addition to the probabilistic performance metrics described above, we also evaluated the performance of the model to make binary predictions about which patients will develop ICANS/severe ICANS. We did this in two ways.

(1) Predicting from day 5: Given data gathered on the first 5 days following CAR T-cell treatment, calculate $f_t = Pr(y_t > 0 | y_{1:5})$, for $t = 6, 7, \dots, 28$. We make a single binary prediction y for each patient by comparing the forecasts f_t for these 23 days with a threshold θ , and predicting ICANS will occur ($y = 1$) if $f_t > \theta$ for any of those days, and otherwise predicting that ICANS will not occur ($y = 0$). We predict severe ICANS in a similar way but using $f_t = Pr(y_t > 2 | y_{1:5})$.

(2) Predicting 5 days ahead: Beginning on day 0, calculate the probability that ICANS > 0 in 5 days, and continue this on days $t = 1, 2, \dots, 23$. That is, calculate $f_t = Pr(y_{t+5} > 0 | y_{1:t})$, which produces 19 probabilities (for days 5, 6, \dots , 23) for each patient. We make a single binary prediction y for each patient by comparing the forecasts f_t for these 19 days with a threshold θ , and predicting ICANS will occur ($y = 1$) if $f_t > \theta$ on any of those days, and otherwise predicting that ICANS will not occur ($y = 0$). We predicted severe ICANS in a similar way but using $f_t = Pr(y_{t+5} > 2 | y_{1:t})$.

For both methods (1) and (2), for predicting neurotoxicity (ICANS > 0) or severe neurotoxicity (ICANS > 2), each choice of the threshold value θ results in a binary prediction for each patient, which allows us to calculate an overall sensitivity and false positive rate for the algorithm. We generate receiver operator characteristic (ROC) curves by varying the threshold value and observing how the sensitivity and false positives rates vary.

RESULTS

Our cohort included 199 patients who underwent CAR T-cell therapy. Among these, 97 (48.7%) developed

ICANS. The median (IQR) day of ICANS onset was 9 (1, 20). The duration of neurotoxicity ranged from 2 to 65 days (mean (SD), 6.6 (10.9) days). Online supplemental table S1 summarizes the patient characteristics and online supplemental file 1 shows the time courses of ICANS scores for all patients.

Forecasting onset probability

We first investigated the model's ability to predict ICANS onset. Factors identified as conferring increased baseline risk were older age, higher body temperature, higher serum level of CRP, IL-6, and procalcitonin. These features were retained in the lasso-penalized logistic regression model. Figure 2B shows the model coefficients along with the 95% confidence intervals corresponding to the seven clinical baseline predictors that were included in the PLR model. As shown in figure 2B, the coefficients of ferritin and WBC are close to zero (0.1), and therefore, they are not strong predictors in our cohort. In figure 3, we show four examples of individual patients who underwent CAR T-cell infusion, including the observed ICANS score (orange) and the predicted probability of ICANS onset (blue) for each day of hospitalization. For two patients who developed ICANS (figure 3A,C), the ICANS onset probability initially increased for 9 and 7 days after cell infusion and then decreased. In contrast, for patients that did not develop ICANS (figure 3B,D), the probability of ICANS onset remained low (less than 0.2) throughout the window in which ICANS typically first develops.

Forecasting days with ICANS and severe ICANS

We next explored the model's ability to predict which days a patient would have any neurotoxicity (ICANS > 0), or severe ICANS (ICANS > 2) for predicting from day 5. Probabilities of each of these events as a function of time are shown in online supplemental figure S1. In online supplemental figure S2A, the first 53 patients have low probability of developing ICANS and online supplemental figure S2B shows the probability of developing severe ICANS. The model correctly predicts (mean absolute error of zero) the day of onset in 70.5% ($n = 134$) and 62.2% ($n = 109$) of cases respectively for ICANS and severe ICANS. In 9.6% and 5.7% the error is 1 day; the error is > 1 day in 19.8% and 32% of cases respectively for ICANS and severe ICANS (online supplemental file 1). As expected, the predicting of severe ICANS is more challenging and has more error.

In addition, the ROC curves of the binary predictions of ICANS/severe ICANS are displayed in online supplemental figure S2E,F. The model correctly predicts which patients develop ICANS/severe ICANS with an area under the curve (AUC) of 96.7% and 93.2% using 5 days ahead prediction and with an AUC of 93.2% and 80.6% using predicting from day 5. Online supplemental file 1 also shows the model behavior when forecasting future probability after $t = 1, 3, 5, 7$ days of observations of developing ICANS ($Pr[y_{t+h} > 0 | X_{1:t}, y_{1:t}]$) or severe ICANS ($Pr[y_{t+h} > 2 | X_{1:t}, y_{1:t}]$).

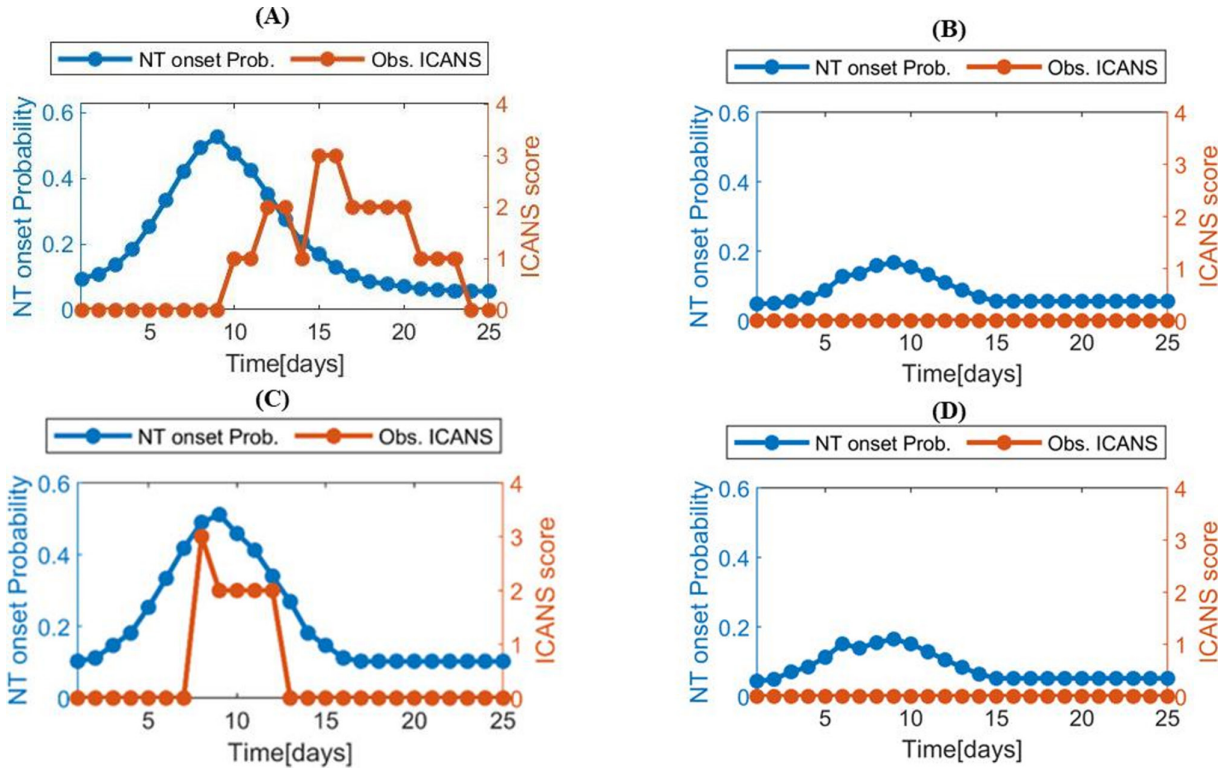


Figure 3 ICANS onset probability during days for four patients. The orange line shows the observed ICANS score, and the blue line shows the estimated ICANS onset probability. (A) ICANS onset probability for one patient who developed neurotoxicity, (B) ICANS onset probability for one patient who did not develop neurotoxicity after cell infusion, (C) ICANS onset probability for one patient who developed neurotoxicity, (D) ICANS onset probability for one patient who did not develop neurotoxicity after cell infusion. ICANS, immune effector cell-associated neurotoxicity syndrome.

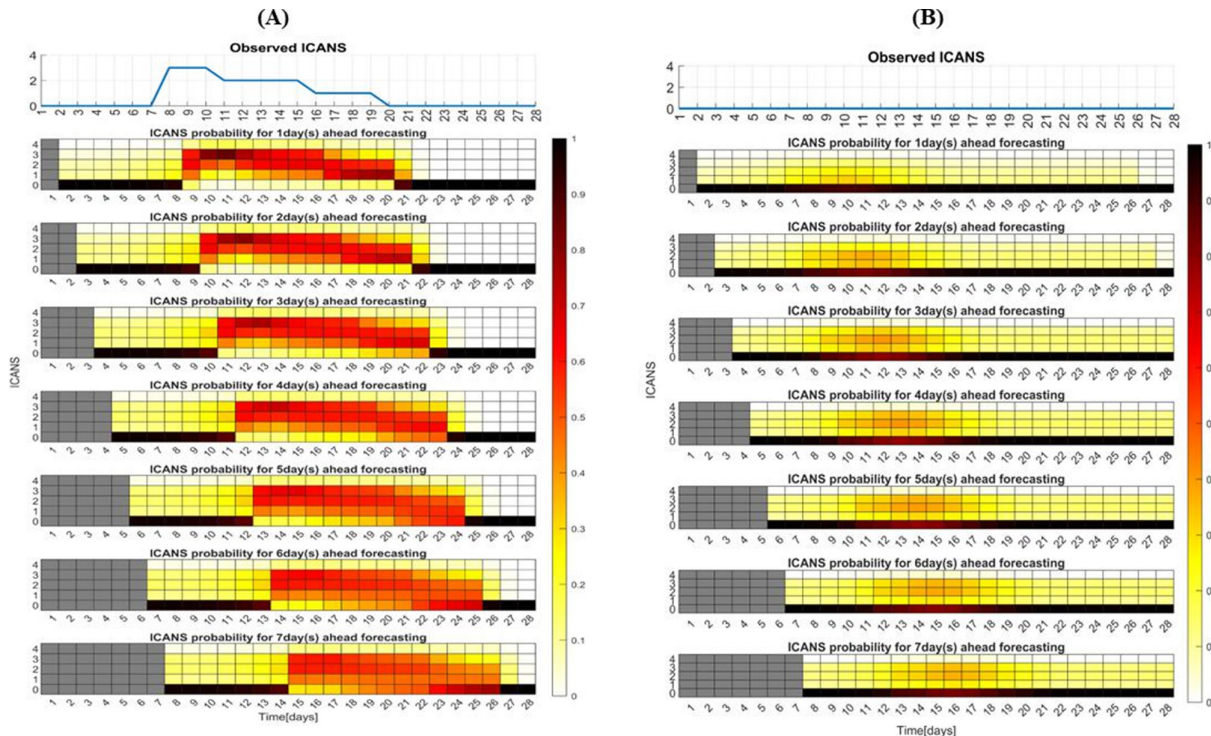


Figure 4 Two examples of 1–7 days ahead forecasting: (A) the patient who developed ICANS (B) the patient who did not develop ICANS. ICANS, immune effector cell-associated neurotoxicity syndrome.

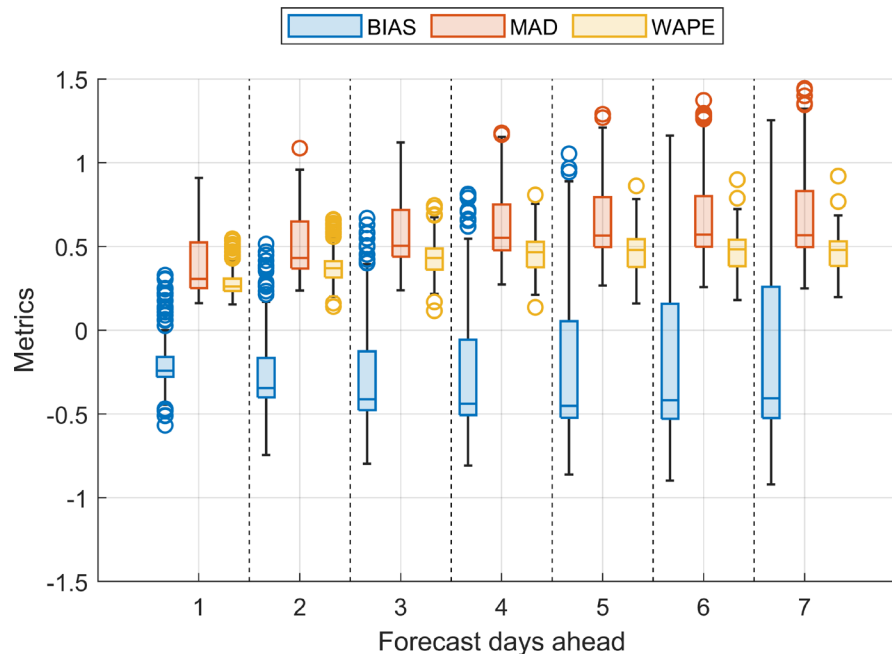


Figure 5 Three different metrics to assess the forecasting ICANS score using HMM+PLR method. HMM, Hidden Markov Model; ICANS, immune effector cell-associated neurotoxicity syndrome; MAD, mean average deviation; PLR, pooled logistic regression; WAPE, weighted average percentage error.

Forecasting ICANS scores one to seven days ahead

We next evaluated the model's ability to forecast ICANS scores between 1 and 7 days ahead. Results for two patients—one who developed ICANS, and another who did not—are illustrated in figure 4. As expected, predicted probabilities deviate more from the observed ICANS course as the forecasting horizon is extended, particularly beyond 3 days. We also note that, as we ask the model to forecast further into the future, there is corresponding shift in the forecasted onset time of risk for neurotoxicity. This is expected behavior. In figure 4A, the shift is predominantly due the onset of ICANS on day 8. When forecasting 1 day ahead, this change on day 8 is reflected in the day 9 forecast: the model predicts continuation of risk. In 7 days ahead forecasts, this same day 8 onset cannot be reflected until day 15. In figure 4B, when predicting 1 day ahead, we also observe an increased risk that peaks on day 10, due to the population average risk trajectory, but also modulated by the evolving laboratory values and vital signs. This peak also shifts to approximately day 16 when we increase the forecasting horizon by 7 days, because this is when this external information first enters the forecasting lookback window.

Forecasting performance results are shown in figure 5. MAD and WAPE, respectively, ranged from 0.38 (0.16) to 0.66 (0.25) and from 0.28 (0.07) to 0.46 (0.10) for 1 to 7 day(s) ahead forecasting. All metrics showed reduced performance when forecasting 4 or more days ahead.

To further explore how the model behaves, we calculated forecasted ICANS trajectories after varying numbers (1, 3, 5, 7) of days of observations following CAR T-cell infusion. Results are shown in online supplemental figure S4 for the patient in figure 4. As expected, when

the observation period does not include any days with ICANS>0, then the probability of later developing ICANS is predicted to be relatively small for all future dates (online supplemental figure S4A-C). However, when the observation period includes nonzero ICANS observations (online supplemental file 1), the model forecasts a future trajectory lasting ~3 weeks with elevated risk of continuing to experience ICANS.

DISCUSSION

In this work, we developed statistical models to forecast ICANS scores, in order to predict not only day of onset and course of ICANS, but also which patients will develop ICANS and severe ICANS. Early and accurate detection of ICANS is of great clinical interest, because not all patients develop ICANS, but those who do require prolonged hospitalization and supportive care.⁹ As a result, statistical models such as these could aid clinical diagnosis and decision making.

We demonstrate high accuracy when predicting the day of onset of ICANS, which is an important clinical time point for management decisions such as steroid administration.¹² The examples shown of patients who developed and did not develop ICANS demonstrates the ability of the model to modulate the predicted onset probability using clinical and laboratory covariates from individuals with different baseline levels of risk.

Knowledge of which patients will develop any degree of ICANS and severe ICANS also has significant clinical implications, particularly affecting length of stay and level of inpatient care (floor vs Intensive Care Unit (ICU)¹²). Through the HMM+PLR method, we improve upon prior

studies predicting which patients develop ICANS of any grade or severe ICANS based on data in the first few days following CAR T-cell infusion. We correctly predict which patients developed ICANS and severe ICANS from day 5 with area under receiver operator characteristic curve 96.7% and 93.2% respectively. By comparison, a prior multivariable logistic regression model correctly predicted the occurrence of any ICANS from day 5 with an AUC of 74% and accuracy of 77%⁹. Another logistic regression model incorporating the modified endothelial activation and stress index score (m-EASIX) had an AUC of 73% in predicting severe ICANS during hospitalization based on data from day 3 following CAR T-cell infusion²⁶.

Our HMM+PLR model is the first to demonstrate forecasting of daily ICANS grades. The model shows excellent performance when predicting 1–3 days ahead, as reflected by MAD and WAPE metrics from 0.38 (0.16) to 0.5 (0.18) and 0.28 (0.07) to 0.38 (0.08). Performance decreases for predictions from day 4 to day 7, with greater divergence from observed values, as expected when forecasting over longer time periods. Nonetheless, the window of 1–3 days is still clinically meaningful to aid decision making, and the forecasting errors even beyond day 3 are relatively small and within clinically useful limits. According to the bias metric, the model tends to underestimate ICANS risk slightly, which may have arisen because the majority of patients do not typically develop ICANS.

The features in the model of onset probability were informed by prior studies showing associations between ICANS and clinical values, but our lasso regularization approach highlights the importance of a subset of these features in predicting ICANS occurrence. As in a prior study by Rubin *et al* involving many of the same patients included in our cohort, older age, fever, and serum CRP were notable predictors of ICANS. However, in contrast, our model also retained serum IL-6 and procalcitonin, an inflammatory cytokine and an acute phase reactant. Other predictors in Rubin *et al*, such as histological subtype, ferritin, minimum WBC, CRS severity, CRS onset day, and number of doses of tocilizumab, were not retained in our model, which could reflect missing values or indicate less importance of these features. Regardless, fewer features could ease implementation of the model, with both IL-6 and procalcitonin often standard more recently among tertiary care centers caring for these patients.

Our study has several limitations. First, ICANS grades were determined via chart review, and thus may contain noise. Second, our dataset has a lot of missing data, especially for procalcitonin and IL-6. We might achieve the higher forecasting accuracy with less missing data. Second, our dataset includes patients from 2015 to 2020, a time period in which management of these patients has evolved, reflecting greater understanding and familiarity with CAR T-cell therapy. Newer treatments that may have greater efficacy in treating ICANS, such as anakinra, are also under clinical trial. Second, the model does not directly account for treatments or clinical developments that may influence the time course of ICANS (eg,

development of cerebral edema or status epilepticus; or use of anesthetic drugs). Accounting for these aspects of ICANS might further improve model forecasts but would require fitting a more complex model and would thus require a larger dataset. The model also arises from patients at a single center, so may lack generalizability to other institutions. There also are differences in the rates of ICANS with different CAR T-cell products and malignancy diagnoses, which our model does not account for due to limitations in sample size. A larger dataset could allow for sub-analyses within malignancy subtypes and CAR T-cell products. In addition, our cohort included patients receiving two different CAR T-cell therapies, CD19 and BCMA, which may have somewhat different ICANS time courses. In future studies with sufficiently large cohorts taking each therapy, forecasting might be further improved by developing separate models for each therapy and by accounting for the effects on the course of ICANS of these additional events and interventions.

Future studies could improve performance through additional risk indicators that capture the dynamic nature of ICANS, such as EEG. Patients who develop ICANS exhibit marked EEG changes, such as increased delta and theta activity, generalized periodic discharge, and seizures, and accounting for these changes and their time course might further enhance forecasting. Whether there are EEG changes that precede ICANS or features of the EEG that might indicate increased risk for ICANS, is unknown, and warrants investigation.

CONCLUSIONS

The developed forecasting model in this study in addition to forecast the ICANS scores up to 7 days into the future, predicted which patients are likely to experience ICANS/severe ICANS and the time course ICANS/severe ICANS is likely to follow once it has begun. With further external validation, this model may be useful for triaging and resource allocation and may allow a large proportion of patients to be safely discharged from the hospital early.

Contributors CAE, YA, MBW designed the study. CAE, SAQ, PM, MSF, HHD and DBR collected data. CAJ provided patient and data resources. YA, CAE, MBW analyzed data, interpreted data, did the literature search, wrote the manuscript and created the figures. All authors reviewed and revised the manuscript.

Funding MBW received funding from the Glenn Foundation for Medical Research and American Federation for Aging Research (Breakthroughs in Gerontology Grant); American Academy of Sleep Medicine (AASM Foundation Strategic Research Award); NIH (R01NS102190, R01NS102574, R01NS107291, RF1AG064312, R01AG062989); and NSF (award SCH-2014431).

Competing interests MBW and SSC are co-founders of Beacon Biosignals, which played no role in this work. DBR served on the scientific advisory board for Celgene/Bristol Meyers Squib. JD received research support from Novartis, consulting fees from Amgen, Blue Earth Diagnostics, and Syndax and royalties from contributing to UpToDate. CAJ received consulting fees from Kite/Gilead, Novartis, BMS/Celgene, Bluebird Bio, Epizyme, Instill-Bio, Lonza, Ipsen, Abintus-Bio, Daiichi-Sankyo and research funding from Kite/Gilead and Pfizer. Other authors report no disclosures relevant to the manuscript.

Patient consent for publication Consent obtained directly from patient(s).

Ethics approval This study involves human participants and was approved by The study was conducted under a protocol approved by the Mass General Brigham

(MGB) Institutional Review Board using a waiver of written informed consent with the IRB protocol number 2013P001024. Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available in a public, open access repository.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iD

Yalda Amidi <http://orcid.org/0000-0001-6842-0034>

REFERENCES

- Holtzman NG, Xie H, Bentzen S, *et al.* Immune effector cell-associated neurotoxicity syndrome after chimeric antigen receptor T-cell therapy for lymphoma: predictive biomarkers and clinical outcomes. *Neuro Oncol* 2021;23:112–21.
- Gust J, Hay KA, Hanafi L-A, *et al.* Endothelial activation and blood-brain barrier disruption in neurotoxicity after adoptive immunotherapy with CD19 CAR-T cells. *Cancer Discov* 2017;7:1404–19.
- Neelapu SS, Locke FL, Bartlett NL, *et al.* Axicabtagene ciloleucel CAR T-cell therapy in refractory large B-cell lymphoma. *N Engl J Med* 2017;377:2531–44.
- Wang ML, Munoz J, Goy A, *et al.* KTE-X19, an anti-CD19 chimeric antigen receptor (CAR) T cell therapy, in patients (Pts) with relapsed/refractory (R/R) mantle cell lymphoma (MCL): results of the phase 2 ZUMA-2 study. *Blood* 2019;134:754.
- Munshi NC, Anderson LD, Shah N, *et al.* Idecabtagene vicleucel in relapsed and refractory multiple myeloma. *N Engl J Med* 2021;384:705–16.
- Abramson JS, Palomba ML, Gordon LI, *et al.* Lisocabtagene maraleucel for patients with relapsed or refractory large B-cell lymphomas (TRANSCEND NHL 001): a multicentre seamless design study. *Lancet* 2020;396: :839–52.
- Schuster SJ, Bishop MR, Tam CS, *et al.* Tisagenlecleucel in adult relapsed or refractory diffuse large B-cell lymphoma. *N Engl J Med* 2019;380: :45–56.
- Maude SL, Laetsch TW, Buechner J, *et al.* Tisagenlecleucel in children and young adults with B-cell lymphoblastic leukemia. *N Engl J Med* 2018;378:439–48.
- Rubin DB, Al Jarrah A, Li K, *et al.* Clinical predictors of neurotoxicity after chimeric antigen receptor T-cell therapy. *JAMA Neurol* 2020;77:1536–42.
- Karschnia P, Jordan JT, Forst DA, *et al.* Clinical presentation, management, and biomarkers of neurotoxicity after adoptive immunotherapy with CAR T cells. *Blood* 2019;133:2212–21.
- Neelapu SS, Tummala S, Kebriaei P, *et al.* Chimeric antigen receptor T-cell therapy — assessment and management of toxicities. *Nat Rev Clin Oncol* 2018;15:47–62.
- Rubin DB, Danish HH, Ali AB, *et al.* Neurological toxicities associated with chimeric antigen receptor T-cell therapy. *Brain* 2019;142:1334–48.
- Santomasso BD, Park JH, Salloum D, *et al.* Clinical and biological correlates of neurotoxicity associated with CAR T-cell therapy in patients with B-cell acute lymphoblastic leukemia. *Cancer Discov* 2018;8:958–71.
- Davila ML, Riviere I, Wang X, *et al.* Efficacy and toxicity management of 19-28z CAR T cell therapy in B cell acute lymphoblastic leukemia. *Sci Transl Med* 2014;6:224ra25.
- Gardner RA, Finney O, Annesley C, *et al.* Intent-to-treat leukemia remission by CD19 CAR T cells of defined formulation and dose in children and young adults. *Blood* 2017;129:3322–31.
- Turtle CJ, Hanafi L-A, Berger C, *et al.* Immunotherapy of non-Hodgkin's lymphoma with a defined ratio of CD8+ and CD4+ CD19-specific chimeric antigen receptor-modified T cells. *Sci Transl Med* 2016;8:355ra116.
- Gofshteyn JS, Shaw PA, Teachey DT, *et al.* Neurotoxicity after CTL019 in a pediatric and young adult cohort. *Ann Neurol* 2018;84:537–46.
- Lee DW, Gardner R, Porter DL, *et al.* Current concepts in the diagnosis and management of cytokine release syndrome. *Blood* 2014;124:188–95.
- Greenbaum U, Strati P, Saliba RM, *et al.* CRP and ferritin in addition to the EASIX score predict CAR-T-related toxicity. *Blood Adv* 2021;5:2799–806.
- Lee DW, Santomasso BD, Locke FL, *et al.* ASTCT consensus grading for cytokine release syndrome and neurologic toxicity associated with immune effector cells. *Biol Blood Marrow Transplant* 2019;25:625–38.
- Murphy KP. *Machine learning: a probabilistic perspective*. MIT press, 2012.
- Yousefi A *et al.* Real-time point process filter for multidimensional decoding problems using mixture models. *bioRxiv* 2018;505289.
- Moré JJ. *The Levenberg-Marquardt algorithm: implementation and theory, in numerical analysis*. Springer, 1978: 105–16.
- Rezaei MR *et al.* A comparison study of point-process filter and deep learning performance in estimating rat position using an ensemble of place cells. in 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) 2018. IEEE.
- Klimberg R *et al.* *Forecasting performance measures—what are their practical meaning?, in advances in business and management forecasting*. Emerald Group Publishing Limited, 2010.
- Pennisi M, Sanchez-Escamilla M, Flynn JR, *et al.* Modified EASIX predicts severe cytokine release syndrome and neurotoxicity after chimeric antigen receptor T cells. *Blood Adv* 2021;5:3397–406.