Clinical paper

A novel methodological framework for multimodality, trajectory model-based prognostication

Jonathan Elmer [a,b,c,*]**, Bobby L. Jones** [d]**, Vladimir I. Zadorozhny** [e]**,
Juan Carlos Puyana** [a]**, Kate L. Flickinger** [a]**, Clifton W. Callaway** [a]**,
Daniel Nagin** [f]

[a] Department of Emergency Medicine, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA

[b] Department of Critical Care Medicine, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA

[c] Department of Neurology, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA

[d] Western Pennsylvania Institute and Clinic, UPMC, Pittsburgh, PA, USA

[e] Department of Informatics and Networked Systems, University of Pittsburgh School of Computing and Information, Pittsburgh, PA, USA

[f] Heinz College, Carnegie Mellon University, Pittsburgh, PA, USA

Abstract

Introduction: Prognostic tools typically combine several time-invariant clinical predictors using regression models that yield a single, time-invariant outcome prediction. This results in considerable information loss as repeatedly or continuously sampled data are aggregated into single summary measures. We describe a method for real-time multivariate outcome prediction that accommodates both longitudinal data and time-invariant clinical characteristics.

Methods: We included comatose patients treated after resuscitation from cardiac arrest who underwent $\geq 6\,h$ of electroencephalographic (EEG) monitoring. We used Persyst v13 (Persyst Development Corp, Prescott AZ) to generate quantitative EEG (qEEG) features and calculated hourly summaries of whole brain suppression ratio and amplitude-integrated EEG. We randomly selected half of subjects as a training sample and used the other half as a test sample. We applied group-based trajectory modeling (GBTM) to the training sample to group patients based on qEEG evolution, then estimated the relationship of group membership and clinical covariates with awakening from coma and surviving to hospital discharge using logistic regression. We used these parameters to calculate posterior probabilities of group membership (PPGMs) in the test sample, and built three prognostic models: adjusted logistic regression (no GBTM), unadjusted GBTM (no clinical covariates) and adjusted GBTM (all data). We compared these models performance characteristics.

Results: We included 723 patients. Group-specific outcome estimates from a 7-group GBTM ranged from 0 to 75%. Compared to unadjusted GBTM, adjusted GBTM calibration was significantly improved at 6 and 12 h, and time to an outcome estimate $<10\%$ and $<5\%$ were significantly shortened. Compared to simple logistic regression, adjusted GBTM identified a substantially larger proportion of subjects with outcome probability $<1\%$.

Conclusions: We describe a novel methodology for combining GBTM output and clinical covariates to estimate patient-specific prognosis over time. Refinement of such methods should form the basis for new avenues of prognostication research that minimize loss of clinically important information.

Keywords: Cardiac arrest, Prognostication, Electroencephalography, Data, Analytics, Precision medicine

* Corresponding author at: Iroquois Building, Suite 400A, 3600 Forbes Avenue, Pittsburgh, PA 15213, USA.
  E-mail address: elmerjp@upmc.edu (J. Elmer).

Introduction

Neurological prognostication after resuscitation from cardiac arrest is challenging.[1,2] Most patients who survive to hospital care after cardiac arrest are initially comatose.[3] Some awaken and enjoy favorable neurological recovery but most are ultimately deemed to have devastating brain injury and die after withdrawal of life-sustaining therapy based on perceived poor neurological prognosis (WLST-N).[4] Because clinical signs and prognostic test results are imperfect and may change over time, guidelines recommend WLST-N be delayed 72 h or longer after initial resuscitation or be based on multiple prognostic modalities.[1,2] Indeed, using multiple prognostic modalities in conjunction improves the prediction in this population.[5–7]

Traditional prognostic models combine one or more independent categorical (e.g. initial rhythm) or continuous (e.g. age) predictors into a regression model to make a single outcome prediction once all data are available.[3,8–11] This results in considerable information loss as repeatedly or continuously sampled data are massively down sampled into single discrete predictors. Furthermore, these models are only useful after all information is available. We examined electroencephalography (EEG) data as a specific example of continuously measured, prognostically important data. After cardiac arrest, quantitative EEG (qEEG) measures evolve over time and group-based trajectory modeling (GBTM) can describe this evolution.[6] We have previously reported that although most subjects' initial EEG tracings are quite suppressed, resolution of this suppression over the first 36 h of monitoring is associated with an 8-fold increase in the odds of survival to hospital discharge compared to persistent suppression.[6] The 36-h qEEG trajectory outperforms use of baseline and 24-hour measures in isolation as a prognostic marker.[6]

In this study, we describe a method for outcome prediction that combines information from baseline clinical characteristics and continuously sampled longitudinal data, and which can provide real-time prognostic information before all data are available. We compare the performance characteristics of one such model to two established methodologies: adjusted logistic regression without using GBTM,[12] and prediction based on GBTM without adjustment for time-invariant baseline characteristics.[13]

Methods

Patients and setting

The University of Pittsburgh Human Research Protection Office (formerly Institutional Review Board) approved all aspects of this work with a waiver of informed consent. We included patients admitted to a single academic medical center from May 2010 to March 2018 after resuscitation from cardiac arrest. We excluded patients who were awake within 6 h of arrival; those who arrested due to a primary neurological event or trauma; patients with death in the first 24 h post-arrest from rearrest, refractory shock, withdrawal due to preexisting advanced directives; patients with <6 h of EEG data available; and, those transferred to our facility >12 h post-arrest. During the study period, it was our standard of care to perform at least 24 h of continuous EEG monitoring on all comatose post-arrest patients except those with non-survivable illness (e.g. refractory shock or cerebral edema with herniation) present on arrival or advanced directives limiting provision of aggressive care.

Our Post-Cardiac Arrest Service (PCAS) coordinates care through the entire post-arrest course including initial resuscitation, critical care and neurological prognostication. We have previously described roles of the service in detail,[14,15] including our standardized bundle of sedation and antiepileptic drug therapy.[6] We routinely obtained baseline brain imaging using computerized tomography (CT) on arrival, and estimated the burden of cerebral edema by calculating the ratio of the density of grey matter to white matter Hounsfield units (GWR).[16] We actively managed patients' temperature to 33 °C (prior to November, 2014) or either 33 °C or 36 °C (after November, 2014) for 24 h. Thereafter, we warmed patients at 0.25 °C/h to 37 °C, and actively maintained normothermia until 72 h post-arrest or awakening from coma. We typically delayed neurological prognostication for at least 72 h after arrest, and based the decision for WLST-N on multiple prognostic modalities including daily neurological examination, initial brain imaging, continuous EEG monitoring, somatosensory evoked potentials, and occasionally magnetic resonance imaging of the brain. Raw EEG data used in this work were available to the treating clinician, but qEEG and trajectory model results were not.

EEG acquisition and processing

Our hospital is staffed around-the-clock by EEG technologists who initiated continuous EEG monitoring on admission to the intensive care unit, a mean of 6–8 hours post-arrest.[6] We used 22 gold-plated cup electrodes in the standard 10–20 International System of Electrode Placement and recorded EEG data using XLTech Natus Neuroworks digital video/EEG systems (Natus Medical, Inc.). We deidentified and exported each patient's EEG data and used Persyst Version 13 (Persyst Development Corp., Prescott AZ) to generate second-by-second qEEG measures, which we imported into BrainFlux– a custom-built large time-series data warehouse implemented using InfluxDB time-series database (InfluxData, San Francisco, CA). We considered any hour with <15 min of EEG data to be missing. We identified these epochs during the down-sampling process when data were aggregated from a frequency of 1 Hz to hourly summaries. For the present analysis, we summarized two qEEG features in hourly epochs. First, we calculated each patient's whole-brain suppression ratio (SR). We used Persyst to calculate SR in 10-s epochs in each electrode pair of a standard longitudinal bipolar montage by determining the proportion of each epoch suppressed below a voltage threshold of 2 uV and used the software's algorithm for artifact reduction. We used BrainFlux to calculate mean SR across leads in 10-s epochs, and then calculated the median of each whole-brain average each hour. We also used Persyst to calculate amplitude-integrated EEG (aEEG) values, which we similarly summarized first across space then over time to obtain hourly summaries for each patient. Because high-amplitude polyspike burst-suppression may have particular prognostic significance after cardiac arrest,[14,17] for aEEG we considered only the non-suppressed portions of the recording (i.e. took the median aEEG value conditional on aEEG ≥2) for patients with suppression-burst calculated aEEG only during bursts of activity. If no periods exceeded this voltage threshold during a given hour we defined aEEG as 0 for that hour. Data were exported from BrainFlux on September 15th, 2018 and reflect the most accurate information available for each patient as of that date.

Covariates and outcomes

We maintain a prospective registry including consecutive post-arrest patient's clinical characteristics and outcomes. From this, we

abstracted patients' age; gender; initial illness severity measured using Pittsburgh Cardiac Arrest Category;[3,18] arrest location in-hospital versus out-of-hospital; GWR on initial brain CT, which we categorized as <1.2, 1.2 to <1.3, 1.3 to <1.4, ≥1.4 or missing (CT not done or uninterpretable); initial arrest rhythm, which we dichotomized as shockable or nonshockable; whether or not each patient awakened from coma, which we defined as following verbal commands; and survival to hospital discharge. We combined these into a composite binary outcome of awake and alive at discharge for prognostication.

Statistical methods

GBTM is a specialized application of a finite mixture model that identifies groups of individuals following a similar trajectory of one or more measures of interest, in this case SR and aEEG, over time.[19,20] Let K denote the number of biomarkers defining each trajectory group, in our case 2 (SR and aEEG), Y_i^k denote a random vector of individual i's longitudinal sequence of measurements of the kth biomarker measured over T = 48 discrete time periods, and $P_k(Y_i^k|j; \beta_j^k)$ denote the distribution of that vector. This distribution is conditional on group membership and an unknown parameter vector β_j^k, which among other things determines the shape of the group-specific trajectory. GBTM assumes that, conditional on membership in the jth group, all Y_i^k are independently distributed with $P_j(Y_i^1, Y_i^2, .., Y_i^K) = P_{1j}(Y_i^1)P_{2j}(Y_i^2)..P_{Kj}(Y_i^K)$. Thus, the likelihood for each individual conditional on number of groups J may be written as: $P_j(Y_i^1, Y_i^2, .., Y_i^K) = \sum_{j=1}^{J} \pi_j \left[\prod_{k=1}^{K} P_k\left(Y_i^k|j; \beta_j^k\right) \right]$, where π_j is the probability of membership in group j.

We scaled both SR and aEEG to range from 0 to 1 by censoring then dividing through by the maximum biologically plausible values and specified both qEEG measures to follow a beta distribution the parameters of which we allowed to vary by trajectory group. We then estimated dual-GBTMs of the evolution of both measures over time that ranged from 3 to 8 groups. These GBTMs handled missing data in two ways. We considered intermittently missing qEEG data (i.e. data were available in at least one epoch before and one epoch after) to be missing at random. Data could also be missing due to nonrandom dropout (i.e. EEG monitoring was discontinued and never resumed). We did not impute any missing data and estimated models using maximum likelihood estimation, which yields asymptotically unbiased parameter estimates in the case of data missing at random. After dropout, patients did not contribute additional data to model estimation and patient-specific PPGMs were not updated further. We compared models using Bayesian and Akaike information criteria and their ability to identify prognostically important subgroups with predicted outcomes much lower (at or near 0%) or much higher (more than 2× the overall population average) than average.

We randomly selected 50% of our cohort as training data and withheld the other 50% of the cohort as test data. We estimated the GBTM parameters described above using the training cohort, then used these parameter estimates to calculate the posterior probability of group membership (PPGM) for each patient in the test data set. The PPGM is calculated according to Bayes' Rule. For each patient, it estimates the probability of membership in each group j given the individual's available data, in this case qEEG features available at or before the time of its calculation. We iteratively re-estimated the PPGM for each group j and patient i after each new epoch of data (i.e. each new hour of summarized qEEG

features). We then reversed the sample, training a new GBTM using data from subjects previously in the test dataset and calculated PPGMs for subjects formerly in the training dataset. We combined the two subsets of test data for subsequent diagnostic exercises. The first model for which we evaluated performance characteristics was patient outcome estimates based on the hourly PPGMs and training data outcomes for each group, which we will refer to as the unadjusted GBTM model.

Next, we generated covariate-adjusted GBTM prognostic estimates as follows. In our final model, two trajectory groups had no patients awake and alive at discharge, which we felt was clinically useful but also posed a challenge for regression-based prognostication because in these groups' covariates were irrelevant to outcome prediction. For the groups with non-zero outcomes, we normalized their PPGMs by dividing through by the sum of PPGMs for all non-zero outcome groups: $P_{norm}j(Y_{it}^K) = Pj(Y_{it}^K)/\sum_{1}^{J*} P\left(j Y_{it}^K\right)$, where J* is the number of the non-zero outcome groups. Using the training dataset, we then estimated a multivariable logistic regression model predicting awakening and survival at hospital discharge based on the normalized PPGMs for each non-zero outcome group for the predictor covariates age, GWR, initial rhythm, PCAC, and arrest location. Next, we applied the findings from the training cohort to the holdout sample. Based on the previously estimated GBTM parameters, we calculated each individual's PPGM for each trajectory group each hour, $Pj(Y_{it}^K)$ then calculated P_{norm} for each group at each hour. Then, we calculated each individual's probability of being awake and alive at hospital discharge conditional on membership in a non-zero outcome group, which we then rescaled back to account for the probability of membership in zero outcome groups as: $\hat{p}_{it} = \left[\left(e^{\beta_0 + \sum_{1}^{J*} \beta_{J*}P_{norm}J + \beta_{age}*age + ...} \right) / \left(1 + \left(e^{\beta_0 + \sum_{1}^{J*} \beta_{J*}P_{norm}J + \beta_{age}*age + ...} \right) \right) \right]$ $* \sum_{J*} P_{norm}(j)$ where β_{age} ... denotes the parameter estimate of each of the various adjustment covariates. We will refer to this model as the adjusted GBTM model. Finally, we used baseline covariates without any GBTM output to estimate a simple multivariable logistic regression model in each training cohort, and used this output to estimate each test subject's outcome. We will refer to this model as the logistic model.

We tested the performance characteristics of our final models in three ways. First, we compared the calibration of predicted versus actual outcome at 6, 12, 18, 24, 36 and 48 h of observation for all three models. Second, we compared the proportion of patients for whom adjustment (logistic vs adjusted GBTM and unadjusted GBTM vs adjusted GBTM) increased the accuracy of the predicted outcome. We considered accuracy increased if adjustment reduced $\hat{p}_{it}$ in patients who were deceased or unresponsive at discharge, or increased $\hat{p}_{it}$ among those awake and alive at discharge. Finally, we compared the time to potentially actionable predicted outcome threshold for WLST-N by using Cox proportional hazards regression to compare time to $\hat{p}_{it}$ less than 10%, less than 5%, less than 2% and less than 1% between unadjusted and adjusted models. We used Stata Version 14.2 (StataCorp, College Station, TX) for all statistical analyses and estimated GBTMs using the freely available *traj* package, which can be downloaded at https://www.andrew.cmu.edu/user/bjones/.

Results

During the study period, we cared for 2081 patients after presumed cardiac arrest, of whom 723 met inclusion and exclusion criteria for the
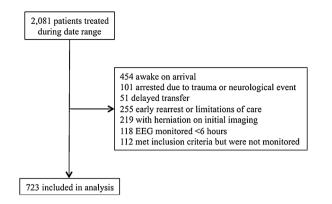
Fig. 1 – Cohort selection and exclusion criteria.

present analysis (Fig. 1). Mean (SD) age of included patients was 57 (16) years, 271 (37%) were female, 599 (83%) arrested out of hospital and 224 (31%) had an initial shockable rhythm. Most patients (58%) had PCAC IV (severe) illness severity, median GWR on initial brain CT was 1.33 [interquartile range (IQR) 1.27–1.39] and 121 (17%) had no baseline brain CT or uninterpretable images. At discharge, 198 (27%) patients were awake and alive. Overall, 1.4% of epochs had <15 min data available and were treated as missing. Another 1.6% had incomplete data (15–60 min). Remaining epochs reflected continuous periods of observation or appropriately missing after monitoring was discontinued. Using both training cohorts, we found a 7-group GBTM best fit the data (Fig. 2). Group-specific percentages awake and alive at hospital discharge ranged from 0% to 75% at hospital discharge (Table 1).

All models were generally well calibrated (Supplemental Fig. 1), except for significant miscalibration of the unadjusted GBTM at hour 6 (calibration curve β-coefficient 0.89, P for Wald test vs 1.0 (perfect calibration) = 0.04) and hour 12 (calibration curve β-coefficient 0.87, P = 0.01). At hour 48, adjustment of the GBTM improved the accuracy of predicted outcome in 574 (79%) of subjects compared to the unadjusted GBTM, and shifted the patient's predicted outcome probabilities by a median of 6.8 [IQR 0 to 17] percent (Fig. 3). These findings were comparable at earlier time points assessed (Fig. 4). Compared to the logistic model, the adjusted GBTM improved prediction accuracy in only 59% of subjects, but increased the proportion of subjects with <1% probability of favorable outcome from 0 to 233 (32% of the overall cohort). At a threshold of <10% and <5% probability of awakening and surviving to discharge, adjustment significantly reduced the time to actionable prognostic estimate (hazard ratio (HR) 0.48 (95%CI 0.37–0.64), P < 0.001 and HR 0.67 (95%CI 0.54–0.85), P = 0.001, respectively) (Fig. 5). There was no difference in time to an actionable estimate assessed at a threshold of <2% or <1% between unadjusted and adjusted models.

Discussion

The evolution of continuous physiological data in critical illness is prognostic[6,21] but has not been considered in most models to predict outcome after cardiac arrest.[3,8–11] Conversely, previous studies using longitudinal data methods like GBTM have not leveraged the predictive capacity of this model or allowed covariates extrinsic to the trajectory model to alter outcome prediction.[22–25] In this work, we have described a method for combining both types of data to inform
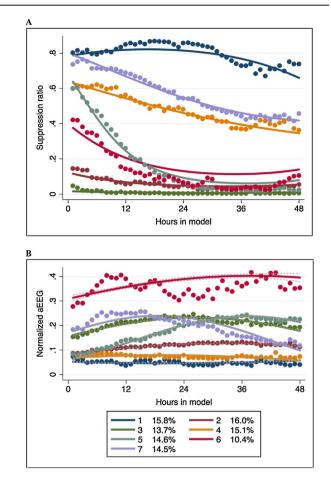

Fig. 2 – Final seven group multi-trajectory model of normalized suppression ratio (A) and aEEG (B) over time.

real-time, iteratively updated, statistically based prognostic estimates for individual patients. Compared to unadjusted GBTM-based prognostic estimates, adjustment improved early model calibration, improved prognostic accuracy in most patients, and shortened the time to an actionable prognostic estimate considered at a threshold of 10% or 5%. Compared to logistic regression based on time-invariant clinical predictors, addition of GBTM-derived information dramatically increased the number of subjects with near-zero probabilities of awakening and survival to discharge.

Comparing longitudinal models necessitated that we develop novel criteria to quantify their performance. To our knowledge, time to prediction is a performance metric not described previously. To a large extent, this is because most prior models provide a single time-invariant prognostic estimate. Our rationale for dichotomizing patients' estimated outcome probabilities at predefined thresholds is as follows. Generally speaking, predictive tools are easiest to understand at a population level. Among a group of 1000 individuals with 5% survival predicted by a well-calibrated model, on average 50 are expected to survive. Each individual patient will either live or die, and a clinician recommending WLST-N is not often interested in population-level estimates but rather the outcome of the individual patient in question.[26] With this in mind, an outcome prediction is only actionable if it falls below an individual physician or family's tolerance for error. This threshold likely varies across physicians, families and individual clinical cases, so we compared models at several potentially useful thresholds.

Table 1 – Trajectory group-specific clinical characteristics and outcomes.

Holdout cohort 1 (n = 360)

Characteristic	Group 1 (n = 59)	Group 2 (n = 66)	Group 3 (n = 57)	Group 4 (n = 47)	Group 5 (n = 47)	Group 6 (n = 37)	Group 7 (n = 50)
Age, years	61 ± 15	58 ± 18	58 ± 18	61 ± 15	62 ± 15	53 ± 19	57 ± 17
Female sex	32 (55)	17 (31)	16 (40)	16 (28)	16 (30)	22 (50)	23 (44)
Out-of-hospital arrest	49 (84)	43 (80)	32 (80)	46 (79)	43 (80)	38 (86)	49 (94)
Shockable rhythm	10 (17)	26 (28)	21 (52)	18 (31)	19 (35)	7 (16)	11 (21)
PCAC							
II	5 (8)	25 (46)	28 (10)	9 (16)	23 (43)	10 (23)	2 (4)
III	0 (0)	10 (19)	4 (10)	10 (17)	6 (11)	2 (5)	5 (10)
IV	47 (81)	13 (24)	5 (13)	36 (62)	23 (43)	30 (68)	41 (79)
Unknown	6 (10)	6 (11)	3 (8)	3 (5)	2 (4)	2 (5)	4 (8)
GWR							
<1.2	26 (45)	0 (0)	1 (3)	5 (9)	0 (0)	1 (2)	7 (13)
1.2 to <1.3	8 (14)	8 (15)	6 (15)	16 (28)	14 (26)	13 (30))	12 (23)
1.3 to <1.4	12 (21)	17 (31)	13 (33)	18 (31)	23 (43)	16 (36)	20 (38)
≥1.4	7 (12)	16 (30)	10 (25)	9 (16)	11 (20)	7 (16)	9 (17)
Missing	5 (9)	13 (24)	10 (25)	10 (17)	6 (11)	7 (16)	4 (8)
Awaked and survived	0 (0)	39 (59)	39 (68)	6 (13)	21 (45)	7 (19)	0 (0)

Holdout cohort 2 (n = 363)

Characteristic	Group 1 (n = 58)	Group 2 (n = 54)	Group 3 (n = 40)	Group 4 (n = 58)	Group 5 (n = 54)	Group 6 (n = 44)	Group 7 (n = 52)
Age, years	60 ± 18	58 ± 14	49 ± 16	60 ± 18	58 ± 14	51 ± 15	56 ± 16
Female sex	22 (38)	20 (30)	22 (39)	13 (28)	15 (32)	16 (43)	21 (42)
Out-of-hospital arrest	54 (92)	49 (74)	47 (82)	32 (68)	36 (77)	35 (95)	46 (42)
Shockable rhythm	10 (17)	31 (47)	28 (49)	6 (13)	24 (51)	13 (35)	11 (22)
PCAC							
II	4 (7)	26 (39)	29 (51)	9 (19)	22 (47)	8 (23)	8 (16)
III	3 (5)	17 (26)	11 (9)	8 (17)	8 (17)	3 (8)	0 (0)
IV	49 (83)	19 (29)	10 (18)	25 (53)	15 (32)	25 (68)	37 (74)
Unknown	3 (5)	4 (6)	7 (12)	5 (11)	2 (4)	1 (3)	5 (10)
GWR							
<1.2	15 (25)	1 (2)	3 (5)	10 (21)	7 (15)	7 (19)	5 (10)
1.2 to <1.3	16 (27)	9 (14)	5 (9)	7 (15)	12 (26)	8 (22)	12 (24)
1.3 to <1.4	13 (22)	30 (45)	17 (30)	14 (30)	16 (34)	16 (43)	19 (38)
≥1.4	5 (8)	12 (18)	19 (33)	10 (21)	9 (19)	4 (11)	8 (16)
Missing	10 (17)	14 (21)	13 (23)	10 (21)	7 (15)	7 (19)	5 (10)
Awaked and survived	0 (0)	31 (57)	30 (75)	5 (9)	16 (30)	4 (9)	0 (0)

Data are presented as mean ± standard deviation or number (percent). Abbreviations: PCAC– Pittsburgh Cardiac Arrest Category; GWR– grey matter to white matter ratio of Hounsfield units.

The clinical covariates we included in our adjusted model were assessed at baseline and therefore known prior to GBTM start. We note, however, that it would be trivial to allow these covariates to vary over time. For example, if CT imaging were completed at hour 12 of the trajectory model, an individual's GWR could be categorized as "missing" during outcome estimation conducted prior to acquisition, then be updated to its measured value after imaging was acquired. This statistical flexibility is highly desirable when outcome prediction is updated as new data become available but the sequence in which new data are acquired varies because of real-world clinical considerations or random chance.

Although compared to the unadjusted GBTM the adjusted model improved the accuracy of most patients' outcome estimates, for 1 in 5 patients adjustment actually worsened model performance. This occurred when a patient had a set of clinical predictors (e.g. young age, shockable rhythm and low PCAC, all of which were associated with favorable outcomes) but an outcome that was discordant with these predictors. In these circumstances, prediction based on qEEG trajectories alone was more accurate. Such errors highlight the challenges of applying population-based estimates to individual patients, and clinically emphasize the continued need for delayed, multimodal prognostication. Indeed, we do not propose the adjusted model presented here as ready for routine clinical use. Rather, we describe the methodological framework that may be used to build progressively more sophisticated models that ultimately may reach a point of having clinical utility.

Our work has important limitations. First, we chose predictors to include in our current models based on biological plausibility and availability of data, and overall these predictors performed well. However, it is virtually certain that there are other variables that could be modeled using GBTM or included as covariates that perform better. An interesting area for future research will be to develop methods for screening and selection of optimal features to include in this type of prognostic model. Second, data were collected at a single academic medical center and our findings may not be generalizable to other centers. Again, our aim was to describe a framework for prognosticating based on both trajectories of continuous data and time-invariant factors, not to develop a tool ready for immediate clinical use.
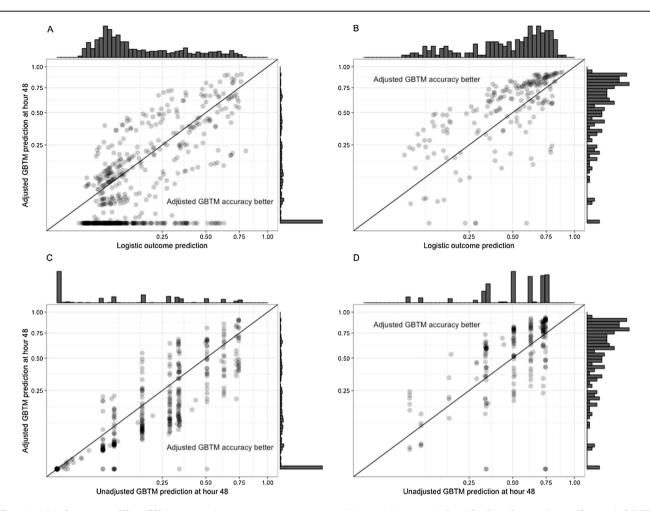
Fig. 3 – Patient-specific differences between outcome models at hour 48. Prediction from the adjusted GBTM was improved among subjects with unfavorable outcomes when it was lower than that of the logistic (A) or unadjusted GBTM (C) estimates, points falling below the reference line in these panels. By contrast, prediction from the adjusted GBTM was improved among subjects with favorable outcomes when it was higher than the logistic (B) or unadjusted GBTM (D), points falling above the reference line. All axes are shown in a square root scale.
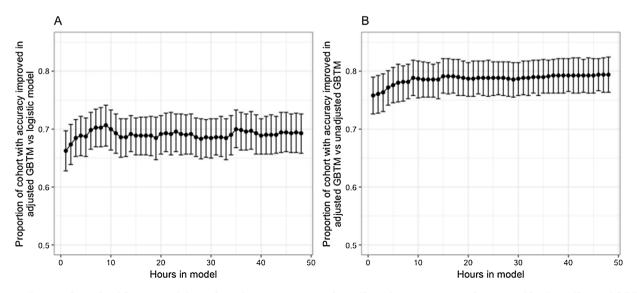

Fig. 4 – Proportion of subjects each hour for whom accuracy of predicted outcome was improved in the adjusted GBTM compared to the logistic model (A) or unadjusted GBTM (B). Confidence bars show the exact binomial 95% confidence interval around the proportion point estimate.
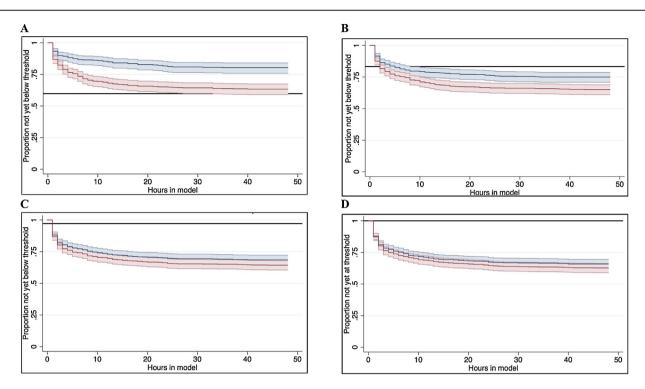
Fig. 5 – Compared to unadjusted models, with adjustment time to an actionable outcome estimate was shorter at a threshold of 10% (A) and 5% (B) but not 2% (C) or 1% (D). Time-invariant estimates from the logistic model are denoted with the solid reference line. Note, no subjects had a logistic model-based outcome estimate below the 1% threshold.

Broad clinical applicability would necessitate testing of a trained prognostic model on an external data set (for example, qEEG, clinical and outcomes data measured prospectively at multiple medical centers). From the perspective of real-time prediction, the technical infrastructure for qEEG signal processing, feature summarization and modeling would also need to be developed. For this reason, we focus on signals derived from clinically available software already used in routine practice, which can generate qEEG features in near real-time on standard beside monitors without significant computational needs. Nevertheless, near real-time analysis of biomedical data can be logistically challenging. Clinical acquisition of EEG in the intensive care unit is also resource intensive, typically requiring trained EEG technologists and physician interpreters, bedside monitoring equipment, and an informatics backbone and storage capable of transmitting and archiving the large files that are generated.[27] Whether or not a limited montage requiring less costly infrastructure, for example EEG signal acquired by a bispectral index monitor, would perform similarly to a full EEG montage is unknown and requires further investigation. Finally, and perhaps most importantly, because the clinicians at our center were not blinded to the data we modeled and also directed decisions about WLST-N, there is a real potential that our findings were influenced by self-fulfilling prophecies. As such, the estimate derived from our model cannot be interpreted as the probability that an individual could awaken and survive to discharge in the absence of WLST-N, but rather the probability of awakening and survival when WLST-N is applied similarly to that in this study.

In conclusion, we describe a novel methodology for real-time patient outcome prediction that accommodates information from both trajectories of continuous longitudinal data and other clinical covariates. We further propose several methods that allow head-to-head comparisons of the performance and clinical utility of these models. Future work will focus on refining methods for model building and external validation of our findings.

Disclosures

REFERENCES

1. Callaway CW, Donnino MW, Fink EL, et al. Part 8: post-cardiac arrest care: 2015 American Heart Association guidelines update for cardiopulmonary resuscitation and emergency cardiovascular care. Circulation 2015;(132):S465–82.
2. Sandroni C, Cariou A, Cavallaro F, et al. Prognostication in comatose survivors of cardiac arrest: an advisory statement from the European Resuscitation Council and the European Society of Intensive Care Medicine. Resuscitation 2014;85:1779–89.
3. Coppler PJ, Elmer J, Calderon L, et al. Validation of the Pittsburgh cardiac arrest category illness severity score. Resuscitation 2015;89:86–92.
4. Elmer J, Torres C, Aufderheide TP, et al. Association of early withdrawal of life-sustaining therapy for perceived neurological prognosis with mortality after cardiac arrest. Resuscitation 2016;102:127–35.
5. Youn CS, Callaway CW, Rittenberger JC, Post Cardiac Arrest S. Combination of initial neurologic examination and continuous EEG to predict survival after cardiac arrest. Resuscitation 2015;94:73–9.

6. Elmer J, Gianakas JJ, Rittenberger JC, et al. Group-based trajectory modeling of suppression ratio after cardiac arrest. Neurocrit Care 2016;25:415–23.

7. Elmer J, Jeong K, Abebe KZ, et al. Serum neutrophil gelatinase-associated lipocalin predicts survival after resuscitation from cardiac arrest. Crit Care Med 2016;44:111–9.

8. Chan PS, Spertus JA, Krumholz HM, et al. A validated prediction tool for initial survivors of in-hospital cardiac arrest. Arch Intern Med 2012;172:947–53.

9. Wang CH, Chang WT, Huang CH, et al. Validation of the cardiac arrest survival postresuscitation in-hospital (CASPRI) score in an East Asian population. PLoS One 2018;13:e0202938.

10. Maupain C, Bougouin W, Lamhaut L, et al. The CAHP (cardiac arrest hospital prognosis) score: a tool for risk stratification after out-of-hospital cardiac arrest. Eur Heart J 2016;37:3222–8.

11. Adrie C, Cariou A, Mourvillier B, et al. Predicting survival with good neurological recovery at hospital admission after successful resuscitation of out-of-hospital cardiac arrest: the OHCA score. Eur Heart J 2006;27:2840–5.

12. Bewick V, Cheek L, Ball J. Statistics review 14: logistic regression. Crit Care 2005;9:112–8.

13. Niyonkuru C, Wagner AK, Ozawa H, Amin K, Goyal A, Fabio A. Group-based trajectory analysis applications for prognostic biomarker model development in severe TBI: a practical example. J Neurotrauma 2013;30:938–45.

14. Elmer J, Rittenberger JC, Faro J, et al. Clinically distinct electroencephalographic phenotypes of early myoclonus after cardiac arrest. Ann Neurol 2016;80:175–84.

15. Rittenberger JC, Guyette FX, Tisherman SA, DeVita MA, Alvarez RJ, Callaway CW. Outcomes of a hospital-wide plan to improve care of comatose survivors of cardiac arrest. Resuscitation 2008;79:198–204.

16. Metter RB, Rittenberger JC, Guyette FX, Callaway CW. Association between a quantitative CT scan measure of brain edema and outcome after cardiac arrest. Resuscitation 2011;82:1180–5.

17. Hofmeijer J, Tjepkema-Cloostermans MC, van Putten MJ. Burst-suppression with identical bursts: a distinct EEG pattern with poor outcome in postanoxic coma. Clin Neurophysiol 2014;125:947–54.

18. Rittenberger JC, Tisherman SA, Holm MB, Guyette FX, Callaway CW. An early, novel illness severity score to predict outcome after cardiac arrest. Resuscitation 2011;82:1399–404.

19. Nagin D. Group-based modeling of development. Cambridge, Mass: Harvard University Press; 2005.

20. Nagin DS, Jones BL, Passos VL, Tremblay RE. Group-based multi-trajectory modeling. Stat Methods Med Res 2018;27:2015–23.

21. Chen L, Dubrawski A, Clermont G, Hravnak M, Pinsky MR. Modelling risk of cardio-respiratory instability as a heterogeneous process. AMIA Annu Symp Proc 2015;2015:1841–50.

22. Nagin DS, Land KC. Age, criminal careers, and population heterogeneity: specification and estimation of a nonparametric, mixed poisson model. Criminology 1993;31:327–62.

23. Burckhardt P, Nagin DS, Padman R. Multi-trajectory models of chronic kidney disease progression. AMIA Annu Symp Proc 2016;2016:1737–46.

24. Yeates KO, Taylor HG, Rusin J, et al. Longitudinal trajectories of postconcussive symptoms in children with mild traumatic brain injuries and their relationship to acute clinical status. Pediatrics 2009;123:735–43.

25. Nagin DS, Odgers CL. Group-based trajectory modeling in clinical research. Ann Rev Clin Psychol 2010;6:109–38.

26. Rose G. Sick individuals and sick populations. Int J Epidemiol 1985;14:32–8.

27. Herman ST, Abend NS, Bleck TP, et al. Consensus statement on continuous EEG in critically ill adults and children, part II: personnel, technical specifications, and clinical practice. J Clin Neurophysiol 2015;32:96–108.