Thick Data Analytics (TDA):

An Iterative and Inductive Framework for Algorithmic Improvement

Minh Nguyen[a], Tiffany Eulalio[a], Ben Marafino[a], Christian Rose[b], Jonathan H Chen[c], and Michael Baiocchi[d]*

[a]*Department of Biomedical Data Science, Stanford University, School of Medicine,*

Stanford, CA, USA;

[b]*Department of Emergency Medicine, Stanford University, School of Medicine, Stanford, CA, USA;*

[c]*Department of Medicine, Center for Biomedical Informatics Research, Stanford University, Stanford, CA, USA;*

[d*]*Department of Epidemiology and Population Health, Stanford University, Stanford, CA, USA. Email: baiocchi@stanford.edu*

Supplementary Materials

The software code needed to replicate the experiments in this study is available at

https://github.com/Minh084/TriageTD

Section 1: Medical context

Most hospital courses start in the ED, where a patient is initially evaluated and managed, then either discharged home or admitted to a short-stay unit (less than 24 hours) versus an inpatient unit for a longer stay (Gonzalez Morganti et al. 2013; Kelen et al. 2021). Based on their perceived acuity level at the time of admission, a patient is admitted to either a general ward (low acuity) or an ICU (high acuity). Admission triage decisions largely depend on human judgment with few points of observation in a high-stakes and evolving ED environment, where the patient's clinical course is most variable and symptoms often undifferentiated. As a result, ED triage decisions often are highly variable, based on limited communication and information (Levin et al. 2018; Mistry et al. 2018; Nates et al. 2016). Accurately prioritizing patients for ICU admission based on care need and timeliness remain major triage challenges (Tang et al. 2021). The initial triage decision can have numerous downstream effects, including delays in care in cases of under-triage, and suboptimal resource allocation for cases of over-triage (Marquet et al., 2015). Predicting ICU admission risk essentially requires prediction from a short observation period near the time of admission. This prediction is commonly done at the initial admission time (t_0). Alternatively, due to the uncertain nature of the ED environment, a prediction at some point within 24 hours following t_0 could be informative to determine a patient's acuity level for their initial admission.

Section 2: Machine Learning Modeling

2.1 Method

We split data by time to account for feature drift (Barddal et al. 2017). The training, validation, and test dataset include all patient admissions from 2015 to 2017, 2018, and from 2019 to March 2020, respectively. To accommodate sparsity and the wide intra- and inter-feature variances, continuous features (lab results and vital signs) were also turned into counts. To accomplish this, we quantized feature distributions into decile bins, then assigned values to bins, and finally counted the frequencies within bins. This method naturally handles missingness by yielding count vectors of zeros over all bins if a particular numerical feature is not available (Rajkomar et al. 2018). There were only small percentages of missing values for ESI (3.9%), height (3.1%), and weight (0.67%). The single predictive mean matching method was used to impute these missing data. Indicators for missingness were also added as features.

In our previous work (XXX et al. 2021), we compared four machine learning algorithms: (1) elastic net regularized logistic regression, (2) random forests, (3) gradient-boosted trees, and (4) four-layer perceptron feed-forward neural networks. These algorithms are well-understood and have been shown to perform well on EHR data (Uddin et al. 2019). In this extended work, we only used the gradient-boosted tree algorithm to train and tune hyperparameters as we had found this algorithm consistently outperformed the others. For model evaluation, in addition to the Area Under the Receiver Operating Characteristic (AUROC) curve, we also used the Area Under the Precision-Recall Curve (AUPRC) to account for data imbalance. We calculated 95% confidence intervals by bootstrapping with 2000 iterations.

2.2 Results

Our original cohort consisted of 43,980 distinct patient visits for 30,451 unique patients. In the holdout test set of 12,418 visits, 80.8% of patients were not in the training and validation sets. There were 11% positive labels at t_0 and 9.5% at t_{24}. During the first 24 hours, there were 912 (2.1%) patients transferred from non-ICUs to ICUs and 1587 (3.6%) of patients transferred from ICUs to non-ICUs. For the prediction at t_0, the model achieved an Area under Receiving Curve (AUROC) of 0.91 (95% CI: 0.90 - 0.92) and an Area Under the Precision Recall Curve (AUPRC) of 0.64 (95% CI: 0.62 - 0.67). For the prediction at t_{24}, the model achieved an AUROC of 0.85 (95% CI: 0.84 - 0.86) and an AUPRC of 0.49 (95% CI: 0.47 - 0.53). Supplementary figure 3 shows the calibration plots on test data, showing how consistent the predicted probabilities are with observed ICU admission percentages. It was expected that the model performance for outcomes measured at t_0 are better than that at t_{24} since all inputs are extracted prior to t_0. Although the model with prediction window t_0 achieved higher AUROC and AUPRC, calibration for the model at t_{24} appears to be slightly better on average. We obtained 16,484 new patient visits from April 2020 to September 2021 to be used for the redesign of model developments.

Section 3: Steps of Thick Data Analytics and Results

3.1 Detailed examination of EHR data for selected cases

We reviewed patients' discharge diagnoses, primary problems presented in the ED, and patients' chief complaints. Though these three structured variables contain important information about a patient at the time of admission, they are often not documented in the EHRs until much later in a patient's admission or even as late as at the time of discharge. Thus, these variables were not available in EHR data during our defined observation window before the prediction windows. Discharge diagnoses were available in 99.9% of our cohort while only 45.9% of the cohort had primary problems documented. Chief complaints appeared more often in the unstructured ED physician's notes. The discharge summary usually provides a reliable overview of a patient's clinical course during their hospitalization. The discharge summary allows us to understand a patient's big picture during the initial admission in highsight, which is helpful during record review.

In cases where discharge summaries are brief, other clinical notes could shed light on the patient's early presentation in the hospital. These included ED notes, the first progress note from the primary team, other notes by physicians, nurses, social workers, case managers, rapid response team, and consults. Physicians' notes are more curated, written with a higher level of clinical reasoning, filed at a later time. These notes synthesize clinical events in perspective and often in the context of a chronological narrative describing a patient's progression and treatment with rationales. On the other hand, reflecting more immediate snapshots, nurses' notes tend to be written and filed much closer to real time with short and factual descriptions of events. These notes complement each other for a record reviewer to understand how events unfold. Our observations are similar to those in Hsu et al. (2020), which also characterized the form of clinical notes. Examples of physician and nurse notes are included in section 4 of the supplemental materials.

3.2 Summary of results

The top 5 common discharge diagnoses are the same for the full cohort, as well as for the *discordance* group and *ICU transfer* subgroup. However, the top primary problems presented in the ED differed among these 3 cohorts. There were more unusual medical cases in the *discordance* group, whereas cases in the *ICU transfer* subgroup were more complex. Among the *discordance* group and *ICU transfer* subgroup, diabetic ketoacidosis (DKA) and cranial hemorrhage or hematoma were in the top 5 problems, which were not in the top 5 among the full cohort. Both the *discordance* group and *ICU transfer* subgroup stayed 1.3 hours fewer in the ED (3.6 vs 4.9 hours), compared to the full cohort. On average, the *discordance* group spent 2 days less in the hospital (4.0 vs 6.0).

Summarizing the discordance group

There were 67 patient admissions (57%) who had the same acuity level at both t_0 and t_{24}. The predicted risks appeared well-calibrated, correlating with the distribution of positive and negative labels in our cohort. The median length of stay for patients who remained in non-ICUs and ICUs were 3 and 4 days, respectively.

There were 50 patient admissions who transferred (43%) to a different acuity level at t_{24} compared to t_0. Out of the 3 admissions that were upgraded from non-ICUs to ICUs, one admission was from intermediate-ICU to ICU in 45 minutes, and another in 1 hour and 45 minutes. The third case was likely an order mistake. The patient presented with cardiac arrest, was shocked twice and intubated in the ED then transferred to ICU. Cases like this certainly require ICU admission. Another was transferred to ICU following a percutaneous coronary intervention procedure to open clogged coronary arteries to restore blood flow to the heart. This transfer is a common practice that should not require prediction.

Ketoacidosis was the predominant primary problem presented in the ED among 56 admissions (49%). Among 6 patients who had multiple admissions in the *discordance* group, DKA consistently as one of their main problems presenting in the ED. A patient presented in the ED with nausea and vomiting was found to be dehydrated and with an extremely high blood sugar. The patient was a parent, who was under distress because of the severe state of their young children, and possibly had issues with their home insulin pump. The patient was discharged the next day directly from the ICU. Other patients had challenges such as homelessness, drug use, non-compliance with medical treatments, low income, inadequate social support, impaired daily function, and physical and mental comorbidities. Two other common problems were cranial hemorrhage or hematoma and myocardial infarction (heart attack). These cases appeared to have more unpredictable progress where some patients deteriorated while others recovered quickly without further complications. Categories of primary clinical problems for all the discordant cases were summarized in table 3.

Summarizing the ICU transfers subgroup with less discordance

There were 36 admissions being upgraded to ICUs, with one patient having two different admissions. There were 13 septic, 9 cardiac, 5 respiratory, 3 neurological, 3 trauma, and 3 alcohol withdrawal problems. Half of the transfer orders were placed within 3 hours of admission, as soon as 3 minutes. One-third of these patients had high acuity and died within 2 years of their hospitalizations. There were more post-operative admissions to ICUs among the *ICU transfers* subgroup compared to the *discordance* group. Figure 3 shows the overall number of transfers in both directions over the first 24 hours following initial inpatient admission. While transfers from ICUs to non-ICUs rose after t_{12}, transfers from non-ICUs to ICUs rapidly dropped off in the first 10 hours and plateaued after t_{12}. At t_6, the number of transfers from both directions came closest together and their trends crossed at t_9.

3.3 Reflection of results

Some ICU admissions appeared precautionary for those following surgical operations, an observation in line with studies showing about 80% of post-operative patients are immediately admitted to ICUs. However, for hemodynamically stable postoperative patients, immediate ICU admission may be unnecessary (Shavadia et al. 2019). In one case, a postoperative patient was stable and transferred out of ICU after one day and discharged home after two days. Their predicted risks at t_0 and t_{24} were 0.7 and 0.3, respectively. Surgery-related cases should be removed as the destinations of these patients are often deterministic, and ICU placements do not necessarily indicate decompensation.

For DKA cases, there is a competing need in terms of nursing care vs. medical acuity for ICU admissions. While DKA patients are commonly treated in ICUs, the benefit of ICU care has not been shown for non-severe DKA patients. Some institutions manage these patients outside of the ICUs with different pathways and well-established treatment protocols in an attempt to reduce ICU burden (Edholm et al. 2020; Mendez et al. 2017). However, increased nursing tasks as frequent as every 15 minutes justify ICU care in other institutions. As opposed to the surgery-related cases, DKA cases indicate true discordance, and they could remain in the cohort. Alternatively, a more specific risk prediction modeling should be developed for patients coming to the ED with DKA.

There is still a use for the general risk prediction modeling as there is less time to collect information to determine specific medical conditions among ED patients. Other input features that might be useful include length of time spent in the ED, and number of notes per time spent in the ED before admission. In terms of outcome labels, there was a clear pattern of some rapid changes in admission orders within the first 3 hours. Our prediction problem could be reframed corresponding to a different prediction window. For example, outcomes at t_3 could be considered as the true labels at t_0. Other prediction windows such as 6, 9, or 12 hours are also reasonable and helpful to uncover temporal acuity changes in both directions as in main figure 3.

Desirable features that could be obtained for use in ED prediction algorithms are patient's baseline and current mobility, cognitive status, functional status, history of substance use, living situation prior to arrival, current support systems, circumstances that led to seeking ED care, and patient responses to provided treatments in the ED. While some of these features can be extracted from notes if available, some can be designed to be collected as structured data for future use.

3.4 Redesign of the models

There are multiple features that would be useful for modeling, but they are not available in structured data or largely missing. Some desirable features are baseline and current mobility, cognitive status, and functional status, history of substance use, living situation prior to arrival, current support systems, reasons for seeking ED care, and patient responses to provided treatments in the ED. While some of these features might be extracted from notes, some can be designed to be collected as structured data for future research. Without the availability of actual notes in our data, the following changes were possible for our re-designed model development:

(1) Features: Only 2 new features were added: time spent in the ED and number of notes written by primary nurses, physicians, and all care providers prior to admission.
(2) Outcomes: Labels were created at 3, 6, 9, and 12 hours after t_0 to reflect the trends in transfers in and out of the ICUs. The choice of which prediction window to use depends on the desired outcome for a prediction problem and its framing.
(3) Cohort: We removed 8116 cases where patients had any operating related events such as being in surgeries or cardiac catheterization. The destination units assigned to these patients are often deterministic. Including these cases in a "prediction algorithm" undermines clinical knowledge. The accumulation of predictions incompatible with deterministic protocols could erode users' trust in algorithmic models as the clinicians repeatedly override the algorithm.
(4) Models: A new time split was used with 2019 data added to the training set, 2020 data in the validation set, and 2021 data as the new test set.

Section 4. Example of Emergency Department's Physician and Nursing Notes

These are not actual notes but rather written by our ED physician and nurse on the team based on some real scenarios. These notes demonstrate how different the styles and contents are and each can offer a different view into a patient's clinical picture.

<u>Physician's notes:</u>

HPI

Sample Patient is a 40 y/o man with a PMH of DM1 (c/b nephropathy, retinopathy, gastroparesis), etoh abuse, pancreatitis, who presents with nausea, vomiting, diarrhea, and abdominal pain since drinking EtOH last night. Denies illicit drug use. Patient currently feels dehydrated, states he has had 12 episodes of NBNB emesis today. He endorses dizziness and lightheadedness. Reports he has been unable to tolerate PO today. Denies blood in stool. Denies changes in meds. No recent colds. Last FSBS was 210, per patient.

Review of Systems

Constitutional: Negative for chills and fever.
HENT: Negative for congestion and sore throat.
Eyes: Negative for pain.
Respiratory: Negative for cough and shortness of breath.
Cardiovascular: Negative for chest pain.
Gastrointestinal: Positive for abdominal pain, diarrhea, nausea and vomiting.
Genitourinary: Negative for dysuria.
Musculoskeletal: Negative for neck pain.
Skin: Negative for rash.
Neurological: Positive for dizziness. Negative for tingling, sensory change and focal weakness.
All other systems reviewed and are negative.

Physical Exam
Vitals signs and nursing notes reviewed.
Constitutional:
  General: Not in acute distress.
  Appearance: Normal appearance. Well-developed. Not ill-appearing.
  Comments: Vomiting, appears, uncomfortable
HENT:
  Head: Normocephalic and atraumatic.
Eyes:
  General: No scleral icterus.
  Conjunctiva/sclera: Conjunctiva normal.
Neck:
  Musculoskeletal: Normal range of motion and neck supple.
Cardiovascular:
  Rate and Rhythm: Normal rate and regular rhythm.
  Pulses: Normal pulses.
  Heart sounds: No friction rub.
Pulmonary:
  Effort: Pulmonary effort is normal.
  Breath sounds: Normal breath sounds. No wheezing.
Abdominal:
  Palpations: Abdomen is soft.
  Tenderness: There is no tenderness.
  Comments: Soft, mild epigastric tenderness on deep palpation.
Skin:
  General: Skin is warm and dry.
  Findings: No rash.
Neurological:
  Mental Status: Alert and oriented to person, place, and time.
Psychiatric:
  Behavior: Behavior normal.

Initial Ddx, assessment and plan:
40 y/o man with nausea, vomiting, diarrhea

Will obtain labs, ua

DDx given hx of T1DM will rule out DKA with trigger likely being vomiting related to alcohol use versus infectious etiology. Consider gastritis, gastroenteritis, no bright red emesis to suggest mallory-weiss tears. No chest wall crepitus or shortness of breath to suggest esophageal perforation.

Will give IVF and anti-emetics

Hyperglycemic with acidosis and elevated lactate
K hemolyzed
Pending rpt K for starting insulin drip
Switched to LR

Will start insulin drip, transfer to area with higher nursing ratio.

Summary of assessment:
DKA admitted to EDCC on insulin infusion

Disposition:
Diagnosis: Data Unavailable
Disposition: Data Unavailable
Admitting Attending: No admitting provider for patient encounter.
OR
Follow up: No follow-up provider specified.

<u>Bedside nurse's notes:</u>

- Assumed care of pt, fluids infusing, pt a&ox4, skin warm dry, denies pain, repeat labs K+ and blood gas
- Pt reported nausea with improved abd tenderness, continuous monitoring.
- Xray at bedside
- Spoke with md X, as per md hold kcl infusion at this time and reeval after lactate and bmp results. Will continue to monitor pt status.
- Bed assigned
- Patient resting comfortably in no apparent distress
- Pt ambulatory to restroom at this time with minimal assist, tolerated well.
- Physician at bedside

<u>Diabete education nurse note:</u> most difficult to manage diabetes is ETOH abuse

Supplementary Tables

Table 1. Summary of selected categorical variables

Caption: * includes publicly insured and uninsured patients

Variables		Count	Proportion
Gender			
	Female	21,033	47.8%
	Male	21,693	52.2%
Race			
	Asian	6,655	15.1%
	Black	3,104	7.1%
	Native American	187	0.4%
	Pacific Islander	902	2.1%
	White	22,580	51.3%
	Other	10,170	23.1%
	Unknown	382	0.9%
Insurance			
	*Public**	25,522	58.0%
	Private	18,458	42.0%
Language			
	English	37,024	84.2%
	Non-English	6,956	15.8%

Table 2. Summary of selected numerical variables

Caption:

- Diagnosis count: historical diagnosis from all prior visits, excluding current visits

- Medication, imaging order, procedure order, and lab order counts: orders within one year prior to admission time, including both current and past visits.

Variables	Mean	Standard Deviation
Age	58.3	18.8
Weight (kg)	76.6	229
Height (cm)	167.9	11.2
ESI	2.66	0.52

Medication count	43.4	49.3
Imaging count	7.4	6.8
Diagnosis count	74.5	79.5
Procedure count	2.7	2.3
Lab order count	22.7	20.8
Microbiology count	2.2	1.1

Table 3. Counts of primary problems among patients in the *discordance* subgroup.

Non-Transfers (67)		**Transfers (50)**	
ICUs (37)	***Non-ICUs (30)***	***ICUs to non-ICUs (43)***	***Non-ICUs to ICUs (3)***
Ketoacidosis, glycemia (13)	Ketoacidosis, glycemia (7)	Ketoacidosis, glycemia (26)	Cardiac arrest (1)
Respiratory problems (7)	Cardiac problems (7)	Neurological problems (11)	Sepsis (1)
Neurological problems (7)	Neurological problems (6)	Falls with fractures (3)	post Percutaneous Coronary Intervention (1)
Cardiac problems (5)	Falls with fractures (4)	Cardiac problems (2)	
Sepsis (2)	Sepsis (3)	Sepsis (2)	
Hemorrhagic epiglottitis (1)	Emergent tracheotomy	Asthma exacerbation (1)	
Fall with fractures (1)	Small bowel obstruction (1)	Opiate overdose (1)	
Burn (1)	Calciphylaxis (1)	Myxedema coma (1)	

Supplementary Figures

Figure 1. Calibration

Histogram of the difference in predicted risks from two models with prediction times at t_0 and t_{24}

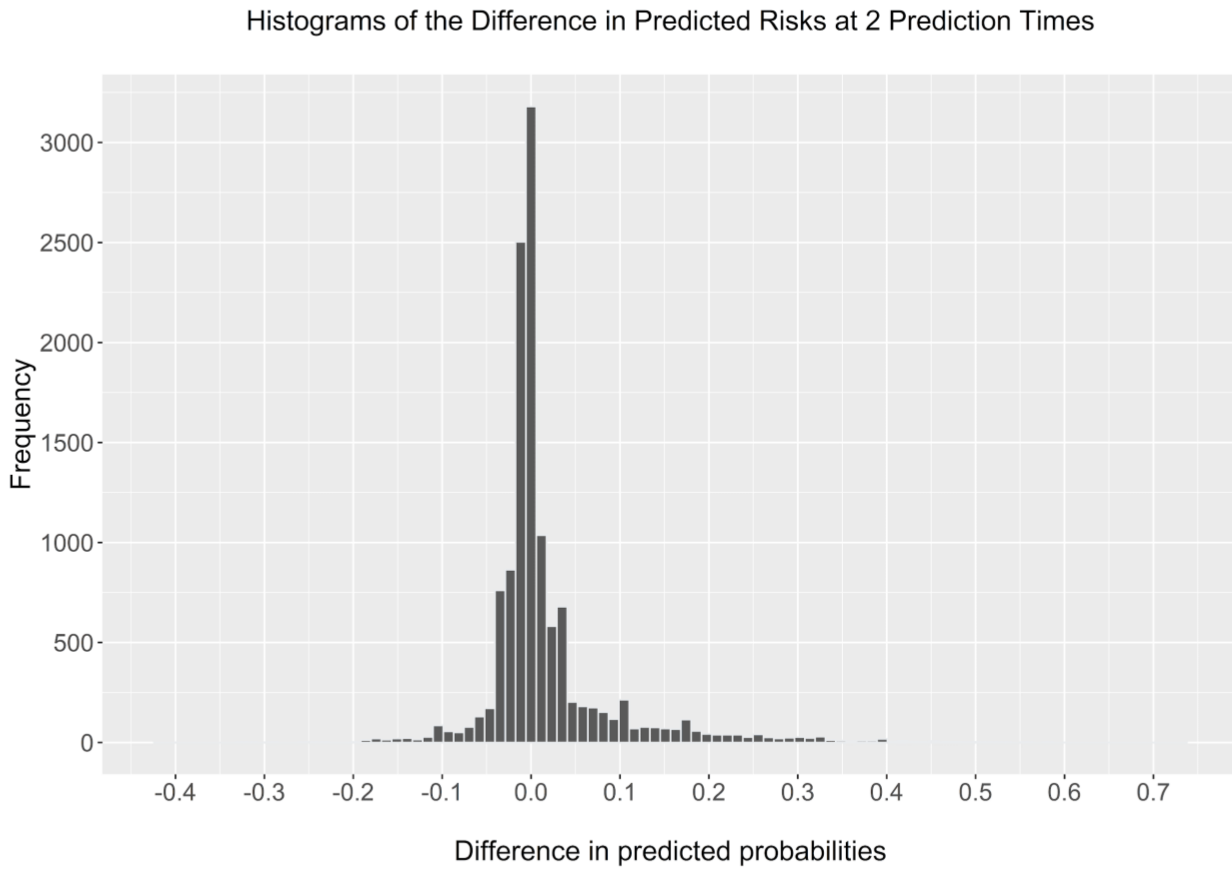


Figure 2. Overlapping histograms of predicted risks from two models with predictions time at t_0 and t_{24}

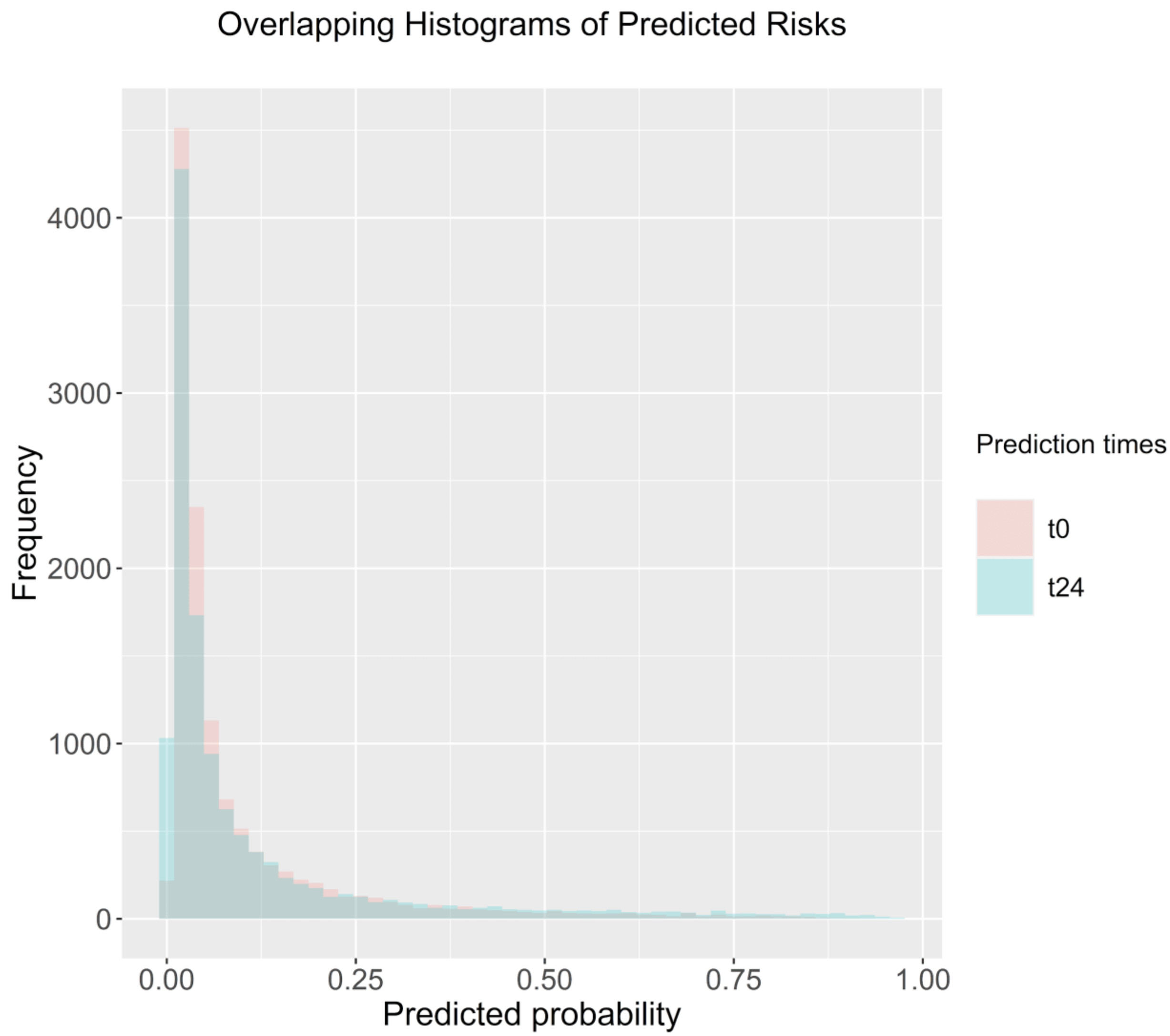
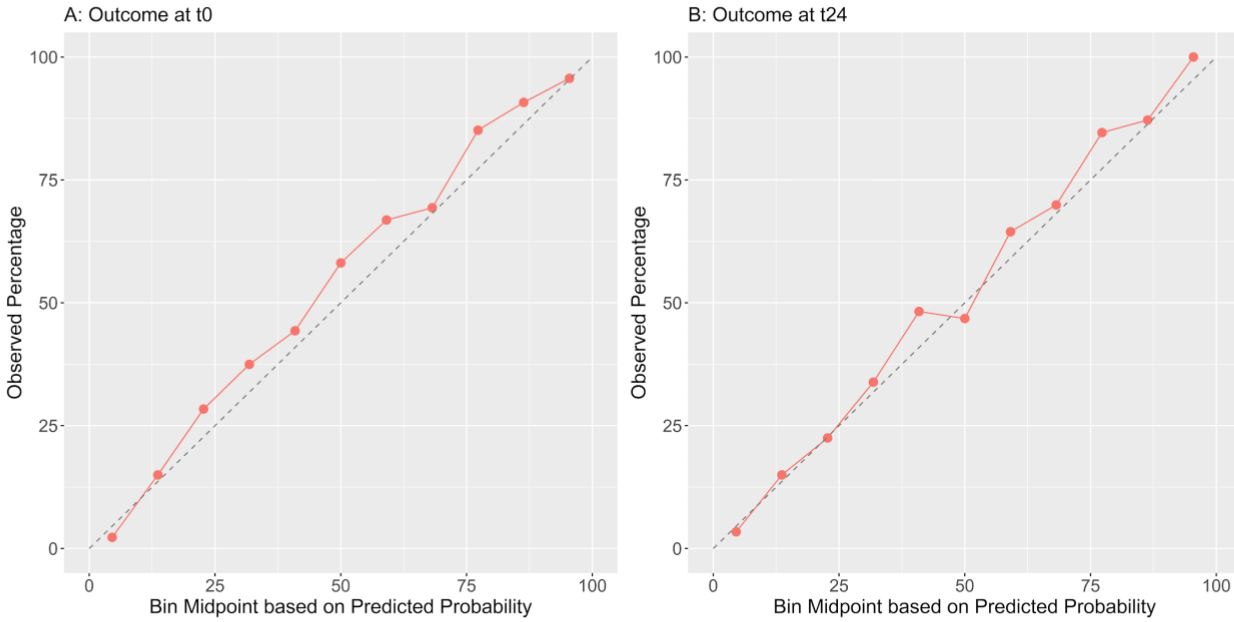


Figure 3. Calibration plots for two models predicting outcomes at the initial hospital admission (t_0) and 24 hours afterwards (t_{24})



References

Barddal, J. P., Gomes, H. M., Enembreck, F., and Pfahringer, B. (2017). A survey on feature drift adaptation: Definition, benchmark, challenges and future directions. *Journal of Systems and Software*, *127*, 278–294. https://doi.org/10.1016/j.jss.2016.07.005

Edholm, K., Lappé, K., Kukhareva, P., Hopkins, C., Hatton, N. D., Gebhart, B., Nyman, H., Signor, E., Davis, M., Kawamoto, K., and Johnson, S. A. (2020). Reducing Diabetic Ketoacidosis Intensive Care Unit Admissions Through an Electronic Health Record-Driven, Standardized Care Pathway. *Journal for Healthcare Quality: Official Publication of the National Association for Healthcare Quality*, *42*(5), e66–e74. https://doi.org/10.1097/JHQ.0000000000000247

Gonzalez Morganti, K., Bauhoff, S., Blanchard, J. C., Abir, M., Smith, A., Vesely, J., Okeke, E. N., Kellermann, A. L., and Iyer, N. (2013). *The Evolving Roles of Emergency Departments:* [Product Page]. RAND Corporation.

https://www.rand.org/pubs/research_briefs/RB9715.html

Hsu, C.-C., Karnwal, S., Mullainathan, S., Obermeyer, Z., and Tan, C. (2020). Characterizing the Value of Information in Medical Notes. *ArXiv:2010.03574 [Cs]*. http://arxiv.org/abs/2010.03574

Kelen, G. D., Wolfe, R., D'Onofrio, G., Mills, A. M., Diercks, D., Stern, S. A., Wadman, M. C., and Sokolove, P. E. (2021). Emergency Department Crowding: The Canary in the Health Care System. *Catalyst Non-Issue Content*, *2*(5). https://doi.org/10.1056/CAT.21.0217

Levin, S., Toerper, M., Hamrock, E., Hinson, J. S., Barnes, S., Gardner, H., Dugas, A., Linton, B., Kirsch, T., and Kelen, G. (2018). Machine-Learning-Based Electronic Triage More Accurately Differentiates Patients With Respect to Clinical Outcomes Compared With the Emergency Severity Index. *Annals of Emergency Medicine*, *71*(5), 565-574.e2. https://doi.org/10.1016/j.annemergmed.2017.08.005

Marquet, K., Claes, N., De Troy, E., Kox, G., Droogmans, M., Schrooten, W., Weekers, F., Vlayen, A., Vandersteen, M., and Vleugels, A. (2015). One Fourth of Unplanned Transfers to a Higher Level of Care Are Associated With a Highly Preventable Adverse Event: A Patient Record Review in Six Belgian Hospitals. *Critical Care Medicine*, *43*(5), 1053–1061. https://doi.org/10.1097/CCM.0000000000000932

Mendez, Y., Surani, S., and Varon, J. (2017). Diabetic ketoacidosis: Treatment in the intensive care unit or general medical/surgical ward? *World Journal of Diabetes*, *8*(2), 40–44. https://doi.org/10.4239/wjd.v8.i2.40

Mistry, B., Stewart De Ramirez, S., Kelen, G., Schmitz, P. S. K., Balhara, K. S., Levin, S., Martinez, D., Psoter, K., Anton, X., and Hinson, J. S. (2018). Accuracy and Reliability of Emergency Department Triage Using the Emergency Severity Index: An International Multicenter Assessment. *Annals of Emergency Medicine*, *71*(5), 581-587.e3.

https://doi.org/10.1016/j.annemergmed.2017.09.036

Nates, J. L., Nunnally, M., Kleinpell, R., Blosser, S., Goldner, J., Birriel, B., Fowler, C. S., Byrum, D., Miles, W. S., Bailey, H., and Sprung, C. L. (2016). ICU Admission, Discharge, and Triage Guidelines: A Framework to Enhance Clinical Operations, Development of Institutional Policies, and Further Research. *Critical Care Medicine*, *44*(8), 1553–1602. https://doi.org/10.1097/CCM.0000000000001856

Rajkomar, A., Oren, E., Chen, K., Dai, A. M., Hajaj, N., Hardt, M., Liu, P. J., Liu, X., Marcus, J., Sun, M., Sundberg, P., Yee, H., Zhang, K., Zhang, Y., Flores, G., Duggan, G. E., Irvine, J., Le, Q., Litsch, K., … Dean, J. (2018). Scalable and accurate deep learning with electronic health records. *Npj Digital Medicine*, *1*(1), Article 1. https://doi.org/10.1038/s41746-018-0029-1

Shavadia, J. S., Chen, A. Y., Fanaroff, A. C., de, L. J. A., Kontos, M. C., and Wang, T. Y. (2019). Intensive Care Utilization in Stable Patients With ST-Segment Elevation Myocardial Infarction Treated With Rapid Reperfusion. *JACC: Cardiovascular Interventions*, *12*(8), 709–717. https://doi.org/10.1016/j.jcin.2019.01.230

Tang, K. J. W., Ang, C. K. E., Constantinides, T., Rajinikanth, V., Acharya, U. R., and Cheong, K. H. (2021). Artificial Intelligence and Machine Learning in Emergency Medicine. *Biocybernetics and Biomedical Engineering*, *41*(1), 156–172. https://doi.org/10.1016/j.bbe.2020.12.002

Uddin, S., Khan, A., Hossain, M. E., and Moni, M. A. (2019). Comparing different supervised machine learning algorithms for disease prediction. *BMC Medical Informatics and Decision Making*, *19*(1), 281. https://doi.org/10.1186/s12911-019-1004-8