

# Man vs. Machine: Comparing Physician vs. Electronic Health Record–Based Model Predictions for 30-Day Hospital Readmissions



Oanh Kieu Nguyen, MD, MAS<sup>1,2,3</sup>, Colin Washington, MD, MPH<sup>1</sup>, Christopher R. Clark, MPA<sup>4</sup>, Michael E. Miller, MS<sup>2</sup>, Vivek A. Patel, MD<sup>1</sup>, Ethan A. Halm, MD, MPH, MBA<sup>1,2</sup>, and Anil N. Makam, MD MAS<sup>1,2,3</sup>

<sup>1</sup>Department of Internal Medicine, UT Southwestern, Dallas, TX, USA; <sup>2</sup>Department of Population and Data Sciences, UT Southwestern, Dallas, TX, USA; <sup>3</sup>Division of Hospital Medicine at San Francisco General Hospital, University of California, San Francisco, San Francisco, CA, USA; <sup>4</sup>Department of Research Administration, Parkland Health and Hospital System, Dallas, TX, USA.

**BACKGROUND:** Electronic health record (EHR)-based readmission risk prediction models can be automated in real-time but have modest discrimination and may be missing important readmission risk factors. Clinician predictions of readmissions may incorporate information unavailable in the EHR, but the comparative usefulness is unknown. We sought to compare clinicians versus a validated EHR-based prediction model in predicting 30-day hospital readmissions.

**METHODS:** We conducted a prospective survey of internal medicine clinicians in an urban safety-net hospital. Clinicians prospectively predicted patients' 30-day readmission risk on 5-point Likert scales, subsequently dichotomized into low- vs. high-risk. We compared human with machine predictions using discrimination, net reclassification, and diagnostic test characteristics. Observed readmissions were ascertained from a regional hospitalization database. We also developed and assessed a "human-plus-machine" logistic regression model incorporating both human and machine predictions.

**RESULTS:** We included 1183 hospitalizations from 106 clinicians, with a readmission rate of 20.8%. Both clinicians and the EHR model had similar discrimination (C-statistic 0.66 vs. 0.66,  $p=0.91$ ). Clinicians had higher specificity (79.0% vs. 48.9%,  $p<0.001$ ) but lower sensitivity (43.9 vs. 75.2%,  $p<0.001$ ) than EHR model predictions. Compared with machine, human was better at reclassifying non-readmissions (non-event NRI +30.1%) but worse at reclassifying readmissions (event NRI –31.3%). A human-plus-machine approach best optimized discrimination (C-statistic 0.70, 95% CI 0.67–0.74), sensitivity (65.5%), and specificity (66.7%).

**CONCLUSION:** Clinicians had similar discrimination but higher specificity and lower sensitivity than EHR model predictions. Human-plus-machine was better than either alone. Readmission risk prediction strategies should incorporate clinician assessments to optimize the accuracy of readmission predictions.

**KEY WORDS:** patient readmission; logistic models; electronic health records; safety-net providers; hospitalization.

J Gen Intern Med 36(9):2555–62  
DOI: 10.1007/s11606-020-06355-3  
© Society of General Internal Medicine 2021

## INTRODUCTION

Up to 1 in 5 hospitalized adults experiences an unplanned readmission within 30 days, accounting for \$26 billion of annual Medicare costs.<sup>1–3</sup> Since 2012, hospitals have been subject to financial penalties under the Centers for Medicare and Medicaid Services Hospital Readmissions Reduction Program (HRRP) for excessive all-cause 30-day readmissions among patients with an index hospitalization for an increasing number of targeted medical and surgical conditions, including congestive heart failure, acute myocardial infarction, pneumonia, and chronic obstructive pulmonary disease.<sup>4,5</sup> These federal penalties have stimulated intense efforts to develop readmission reduction intervention strategies, which are highly resource-intensive but have been only modestly effective when indiscriminately applied to all hospital inpatients.<sup>6–9</sup>

Identifying hospitalized patients at high risk for readmission before they are discharged can enable interventions to be targeted to those at the highest risk and therefore most likely to benefit.<sup>10</sup> Electronic health record (EHR)-based risk prediction models incorporating granular clinical data (i.e., vital signs, hospital-acquired conditions, laboratory results, etc.), are superior to approaches using claims-based administrative data, but are limited by the types of information documented and encoded in the EHR.<sup>11–18</sup> A growing body of evidence suggests that social, functional, and behavioral factors are associated with increased risk of readmission, and that incorporating this information into prediction models improves readmission risk prediction across a variety of conditions.<sup>19–30</sup> However, at present, this information is not uniformly available in EHRs.<sup>31</sup>

Clinician perceptions of readmission risk are readily ascertainable and may incorporate valuable information on severity and complexity of patient illness, as well as information on

**Meeting Presentations** Preliminary findings from this project were presented at the Society of General Internal Medicine Annual Meeting in April 2018 in Denver, Colorado.

Received July 15, 2020

Accepted November 19, 2020

Published online January 14, 2021

social, functional, and behavioral factors unavailable in the EHR, but the comparative accuracy of physician predictions for 30-day readmissions is not well established. Prior research in this area consists of only two studies, which were both limited by small sample sizes and modest to poor discrimination of readmissions by both clinicians and comparator risk tools. One study, conducted prior to HRRP, compared provider predictions (attending and resident physicians, nurses, and case managers) with a standardized risk tool for a cohort of 164 patients and found that both providers and the risk tool had uniformly poor discrimination (C-statistic 0.50–0.59 for providers vs. 0.56 for the risk tool).<sup>32</sup> A more recent study compared clinician predictions (attending and resident physicians and nurses) with the LACE index (length of stay, acuity, comorbidities, emergency department use) for 359 patients and found that clinicians and the LACE index had similarly modest discrimination (C-statistic 0.63–0.69 for clinicians vs. 0.62 for LACE).<sup>33</sup>

Thus, we sought to conduct a head-to-head comparison of the performance of physicians versus a validated EHR-based prediction model in predicting 30-day hospital readmissions in a large, diverse cohort of hospitalized adults with a wide range of clinical, social, functional, and behavioral needs. We hypothesized that physicians would be better able to predict 30-day readmissions due to increased awareness of social context and other patient characteristics influencing readmissions not captured as well in EHR data.

## METHODS

### Study Setting and Population

In this prospective cohort study, we recruited inpatient internal medicine physicians from Parkland Hospital, a large 862-bed, public safety-net hospital in Dallas, Texas, from September through October 2017. Parkland is a level 1 trauma center, the sole safety-net hospital in Dallas County, and is part of a large integrated county health system that cares for 1.5 million patients annually who represent a racially and ethnically diverse poor, underinsured and uninsured urban population.

We included all attending physicians, trainees, and advanced practice providers providing patient care on the inpatient internal medicine (6 teaching and 21 hospitalist direct care services) and cardiology services (4 teaching services), with a combined average daily census of about 300 patients. We obtained verbal consent from all participants. We recruited at least one team member from all teaching services. All recruited individuals other than two faculty on non-teaching services agreed to participate in the study. The UT Southwestern Institutional Review Board reviewed and approved this study prior to the onset of data collection.

### Clinicians' Predictions of Readmission Risk

We surveyed clinicians on their perceived risk of readmission for each patient under their care who was potentially being discharged that day, using a 3-item verbal questionnaire developed by the research team and administered in-person by a team member on the day of anticipated discharge (Appendix Figure 1). Clinicians were unaware of the included predictors and results of the EHR-based model prediction of readmission risk. For each patient, clinicians were asked to predict the risk of 30-day readmission and rate their own confidence in each prediction on 5-point Likert scales. They were also asked to select the main reason for potential 30-day readmission from three choices: medical issue, social issue, or mental health/substance abuse issue. All responses were directly recorded in REDCap.<sup>34</sup> For patients on a teaching service, we collected predictions from as many team members as possible and used the most senior team member's prediction (by descending level of experience: attending, resident, intern) to limit predictions to one per patient in the primary analysis. We identified individual clinician characteristics from information publicly available through the Texas Medical Board.<sup>35</sup>

### EHR-Based Model Predictions of Readmission Risk

After completion of the clinician surveys, we retrospectively extracted clinical data from the EHR (Epic Systems, Verona, WI) for hospitalized patients for whom we had at least one clinician survey, to calculate the risk of 30-day all-cause readmission using a previously validated EHR-based readmission risk prediction model.<sup>14</sup> The model included 24 EHR variables spanning demographic characteristics (age, sex, race/ethnicity, primary payer), utilization history (history of emergency department visits, hospitalizations in the prior year), clinical characteristics (comorbidities, laboratory values, vital signs) and hospital processes (hospital, acquired conditions and/or complications, length of stay, disposition status) (Appendix Table 1). We then categorized patients into quintiles of risk based on cutoffs from the original validation study.

### Ascertainment of Observed 30-Day Readmissions

We ascertained 30-day readmissions from October through November 2017 at the study hospital using EHR data and across 80 hospitals within a 100-mile radius using hospital claims from an all-payer regional hospitalization database (North Texas Health Information and Quality Collaborative), an approach used in our previous studies.<sup>14–16,36,37</sup>

### Statistical Analysis

**Comparison of Human vs. Machine Predictions of Readmission Risk.** We compared clinicians' predictions of readmission risk with the EHR model predictions by discrimination, agreement, net reclassification, and diagnostic test characteristics (sensitivity, specificity, and likelihood ratios).

We assessed discrimination using the area under the receiver operating curve (AUC or C-statistic) and conducted tests of equality of the AUCs to compare human vs. machine. We assessed agreement between human and machine predictions using unweighted kappa scores. For net reclassification and diagnostic test characteristics, we dichotomized clinician and EHR model risk predictions into “low risk” (Likert scores 1, 2, and 3 for clinician predictions; quintiles 1, 2, and 3 for EHR model predictions) and “high risk” (Likert scores 4 and 5 for clinician predictions, quintiles 4 and 5 for EHR model) risk in order to calculate net reclassification indices (NRIs), sensitivity, specificity, and positive and negative likelihood ratios.

We calculated the overall NRI for clinician predictions, using the EHR model as the reference standard. The NRI measures how many events and non-events (in this case, readmissions and non-readmissions) a new model (here, clinician perceptions) correctly predicts compared with a reference model (here, the EHR model). It is mathematically defined as the sum of the net proportions of correctly reclassified persons with and without the event of interest (i.e., the sum of event and non-event NRIs) compared with a reference model.<sup>38,39</sup> The theoretical range of values is  $-2$  to  $+2$  (or  $-200$  to  $+200$  when scaled to approximate percentages), with more positive values indicating that the new model outperforms the reference model.

We visualized diagnostic utility of predictions by clinicians and the EHR model using leaf plots, a data visualization tool that illustrates the usefulness of positive vs. negative test results in correctly ruling in or ruling out a readmission respectively across a range of pre-test probabilities.<sup>40</sup>

**Sensitivity and Subgroup Analyses.** We conducted two sensitivity analyses. First, we repeated the above analyses for a restricted cohort of inpatient hospitalizations without a preceding hospitalization in the 30 days prior to admission, to mirror the Centers for Medicare and Medicaid Services definition of individuals experiencing an index hospitalization.<sup>41</sup> Second, we repeated the above analyses including only hospitalizations for which attending predictions were available ( $n=1003$  hospitalizations). We found no meaningful differences in our findings for either analysis compared with the primary analysis presented here (data not shown; available upon request). We also conducted pre-specified subgroup analyses, repeating the above analyses stratified by clinician confidence (low, Likert score 1–3 vs. high, Likert score of 4–5) and reason for readmission (medical versus non-medical) to assess whether accuracy and discrimination of clinician predictions varied by these factors.

**“Human-Plus-Machine” Model.** To assess whether incorporating information from both clinicians and the EHR model resulted in improved predictions, we developed a “human-plus-machine” model using logistic regression to predict 30-day readmissions. The model included the composite EHR model score as a single continuous predictor and clinician predictions of risk as a 5-level ordinal predictor.

We categorized human-plus-machine model scores into quintiles of risk to allow for direct comparison with clinician and EHR model predictions and repeated the above analyses for our human-plus-machine predictions.

We conducted all analyses using SAS (version 9.4, SAS Institute, Inc.) and Stata statistical software (version 16.0, Stata Corp).

## Role of the Funding Source

This study was supported by the Agency for Healthcare Quality and Research through the UT Southwestern Center for Patient-Centered Outcomes Research (AHRQ R24 HS022418), the National Heart, Lung, and Blood Institute (Nguyen, NHLBI 1K23HL133441), the National Institute on Aging (Makam, NIA 5K23AG052603), and the National Center for Advancing Translational Sciences of the National Institutes of Health (UL1TR001105). The funding sources had no role in the study design, data collection, analysis, preparation of the manuscript, or decision to submit the manuscript for publication.

## RESULTS

We surveyed 106 clinicians on 1183 hospitalizations among 1119 unique individuals. The observed 30-day readmission rate was 20.8% ( $n=246$ ). Mean age of hospitalized individuals was  $53.7 \pm 15.9$  years, with large proportions of individuals from minority racial and ethnic backgrounds and with charity health coverage, reflecting the safety-net study setting (Table 1, Appendix Table 2). Over two-thirds were hospitalized on non-teaching services. Surveyed clinicians were almost entirely physicians ( $n=104$ ), with two advanced practice providers in our cohort (Table 2). About half of clinicians were attending physicians, with a preponderance of hospitalists spanning a range of experience. Clinicians completed surveys for a median of 10 patients (range 1–50 patients) under their care.

## Prediction Performance for Human, Machine, and Human-Plus-Machine

**Discrimination.** Both clinician and EHR model predictions had similar discrimination (C-statistic 0.66 vs. 0.66, 95% CI 0.62–0.70 vs. 0.63–0.70,  $p=0.91$  for comparison, Fig. 1). However, there was only modest concordance between predictions (56.7% agreement, unweighted kappa 0.16), due to the differences in distribution of predictions. Clinician predictions were right-skewed towards more frequent predictions of lower risk while EHR predictions were left-skewed towards high risk (Appendix Figure 2). A human-plus-machine approach had the best discrimination, with a C-statistic of 0.70 (95% 0.67–0.74,  $p=0.001$  for comparison with clinician and  $p=0.002$  for comparison with EHR model predictions). The human-plus-machine model also effectively stratified individuals across the broadest range of readmission risk from 12.0% in the low-risk group to 34.0% in the high-risk group (Fig. 1).

Table 1 Patient Characteristics

Characteristic	All encounters* (n = 1183)
Age in years, mean (SD)	53.7 (15.9)
Female, n (%)	514 (43.5)
Race/ethnicity, n (%)	
White	213 (18.0)
Black	435 (36.8)
Hispanic	491 (41.5)
Other	44 (3.7)
Marital status, n (%)	
Single	585 (49.5)
Married	345 (29.2)
Other <sup>†</sup>	253 (21.4)
Primary payer, n (%)	
Private	60 (5.1)
Medicare	281 (23.8)
Medicaid	221 (18.7)
Charity, self-pay, or other <sup>‡</sup>	621 (52.5)
Hospitalizations in prior year, median (IQR)	1 (0–3)
Non-elective admission, n (%)	1171 (99.0)
Non-teaching service, n (%)	812 (68.6)
CMS HRRP-targeted medical conditions, n (%)	201 (17.0)
Congestive heart failure	115 (9.7)
Acute myocardial infarction	11 (0.9)
Pneumonia	53 (4.5)
Chronic obstructive pulmonary disease	22 (1.9)
Charlson comorbidity index, median (IQR)	2 (0–6)
Length of stay in days, median (IQR)	3 (2–6)
Vital sign instability <sup>§</sup> at discharge, n (%)	363 (30.7)
Discharge location, n (%)	
Home	1125 (95.1)
Post-acute care	51 (4.3)
Hospice	7 (0.6)

CMS, Centers for Medicare and Medicaid Services; ED, emergency department; HRRP, Hospital Readmissions Reduction Program; IQR, interquartile range; SD, standard deviation

\*Characteristics described across all hospital encounters, which occurred among 1119 unique individuals

<sup>†</sup>Includes widowed, separated/divorced, unknown status

<sup>‡</sup>Eligible individuals with “charity” and “self-pay” status had coverage through the Dallas County indigent health program which covers health services rendered at the study hospital

<sup>§</sup>Defined as having at least one of the following: temperature  $\geq 37.8$  °C, heart rate  $> 100$  beats/min, respiratory rate  $> 24$  breaths/min, systolic blood pressure  $\leq 90$  mmHg, or oxygen saturation  $< 90\%$

**Net Reclassification.** Compared with the EHR model, clinicians were better at reclassifying patients who were not readmitted (non-event NRI + 30.1%) but were also nearly equally worse at reclassifying patients who had a readmission (event NRI – 31.3%), with an overall NRI of – 1.2 (Fig. 2). Compared with the EHR model, a human-plus-machine approach was worse at reclassifying patients who were not readmitted (non-event NRI – 12.3%) but better at reclassifying patients who had a readmission (event NRI + 21.5%), with an overall NRI of + 9.3 (Fig. 2).

**Diagnostic Characteristics.** Compared with the EHR model, clinician predictions of readmission risk had much higher specificity (79.0% vs. 48.9%,  $p < 0.001$ ) but lower sensitivity (43.9% vs. 75.2%,  $p < 0.001$ ). Because of the high specificity, clinician predictions had a higher positive likelihood ratio (2.1 vs. 1.5) and were of higher diagnostic utility in ruling in readmissions than EHR model predictions (Fig. 3). Human-plus-machine predictions had the best balance between

Table 2 Clinician Characteristics (n = 106)

Characteristic	n (%)
Specialty*	
Hospital medicine	46 (43.4)
General internal medicine	6 (5.7)
Other medicine subspecialty	3 (2.8)
Internal medicine trainee	51 (48.1)
Clinician type	
Attending physician faculty	53 (50.0)
Advanced practice provider (NP or PA)	2 (1.9)
Resident or fellow <sup>†</sup>	20 (18.9)
Intern	31 (29.3)
Female	43 (41.0)
Years since graduation <sup>‡</sup>	
0–4	59 (55.7)
5–9	17 (16.0)
10–14	10 (9.4)
15–19	10 (9.4)
20+	10 (9.4)

\*Specialties assigned only for attending physicians and/or advanced practice providers

<sup>†</sup>Two fellows were included in this grouping

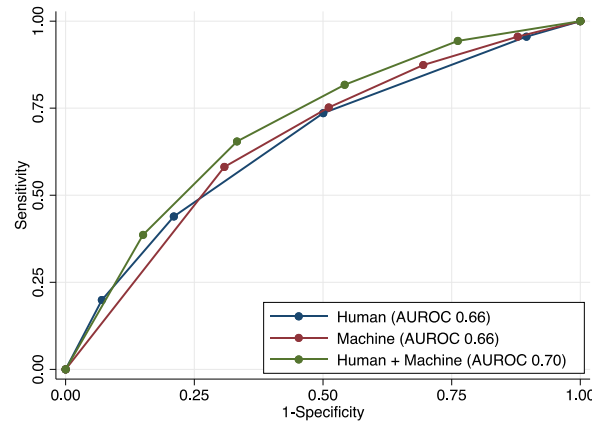
<sup>‡</sup>Denotes years since graduation from medical or professional school

maximizing sensitivity (65.5%, 95% CI 59.1–71.4%) and specificity (66.7, 95% CI 63.6–69.7%), and consequently had the best balance of diagnostic utility for both ruling in (positive likelihood ratio 2.0, 95% CI 1.7–2.2) and ruling out readmissions (negative likelihood ratio 0.5, 95% CI 0.4–0.6) (Fig. 3).

## Human Vs. Machine Predictions by Clinician Confidence and Reason for Readmission

We found no differences between clinician and EHR model predictions in terms of discrimination when predictions were stratified by high vs. low clinician confidence in predictions (Appendix Table 3). Both the clinician and EHR model had poor discrimination for low-confidence predictions (C-statistic 0.59 vs. 0.62, 95% CI 0.53–0.64 vs. 0.57–0.68,  $p = 0.35$ ), and improved but similarly modest discrimination for high-confidence predictions (C-statistic 0.70 vs. 0.70, 95% CI 0.65–0.74 vs. 0.65–0.75,  $p = 0.86$ ). The human-plus-machine model had better discrimination than either clinician or EHR model predictions across both high- and low-confidence clinician predictions (C-statistic 0.74, 95% CI 0.69–0.78 for high-confidence predictions,  $p = 0.006$  for comparisons to both human and machine; C-statistic 0.65, 95% CI 0.60–0.70 for low-confidence predictions,  $p = 0.01$  for comparison with human,  $p = 0.20$  for comparison with machine).

Clinicians identified medical illness as the main potential reason for readmission for 75% of discharges. In this subgroup ( $n = 887$ ), our findings were similar to the overall analysis (Appendix Table 4). However, when clinicians identified non-medical issues as the main potential reason for readmission ( $n = 296$ ), the EHR model had slightly better discrimination, similar accuracy, and slightly better sensitivity and specificity compared with clinician prediction. However, the human-plus-machine model still had the best discrimination in this subgroup, with a C-statistic of 0.74 (95% CI 0.68–0.79) and the best balance between sensitivity and specificity.



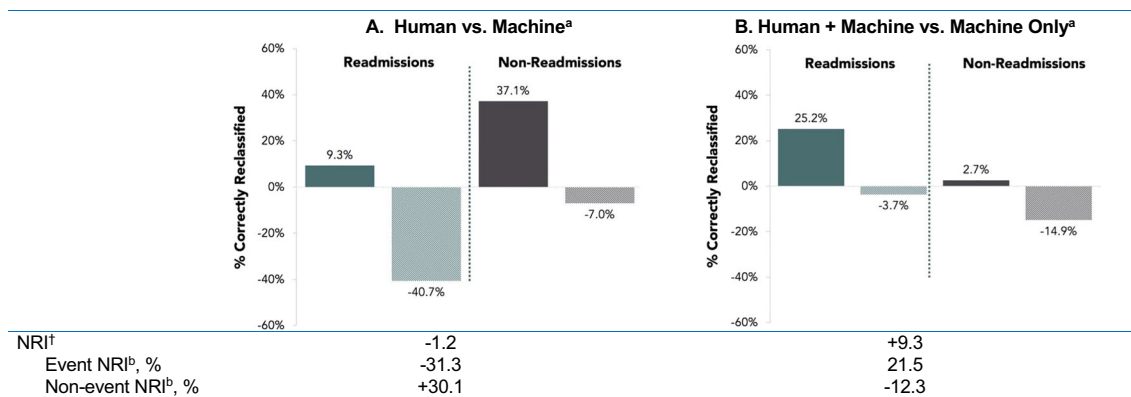
	Human (Clinician)	Machine (EHR Model)	Human + Machine
Prediction performance			
C-statistic (95% CI) <sup>a</sup>	0.66 (0.62-0.70)	0.66 (0.63-0.70)	0.70 (0.67-0.74)
Observed readmissions, %			
Low risk group <sup>b</sup>	15.7	12.1	12.0
High risk group <sup>b</sup>	35.4	28.5	34.0

**Fig. 1** Prediction performance of human, machine and human-plus-machine predictions. CI, confidence interval; EHR, electronic health record; NRI, net reclassification index <sup>a</sup>*p* value for human vs. machine comparison = 0.91; *p* value for human vs. human + machine comparison = 0.001; *p* value for machine vs. human + machine comparison = 0.002 <sup>b</sup>For clinician predictions, “low risk” defined as Likert scale scores of 1–3 and “high risk” defined as scores of 4 or 5. For EHR model predictions, “low risk” defined as those in the lowest three quintiles of risk and “high risk” defined as those in the highest two quintiles of risk using cutoffs from our original validation study.<sup>14</sup>

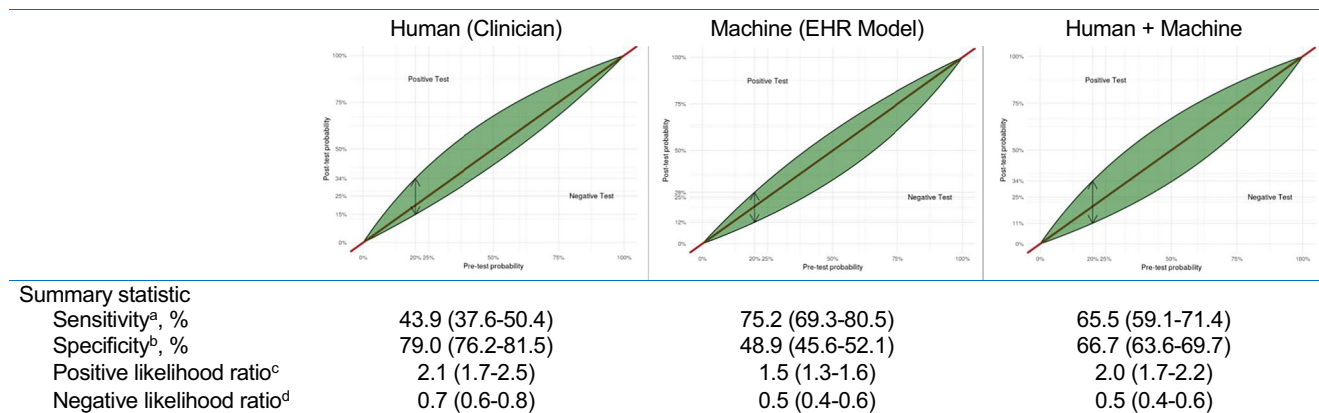
**DISCUSSION**

In a head-to-head comparison of clinician versus automated EHR-based predictions for 30-day hospital readmissions, we found that contrary to our hypothesis, both clinicians and the EHR model had similarly modest discrimination for readmissions, though each strategy had unique strengths and blind spots. The EHR model was better at predicting who would be readmitted but overestimated readmission risk

overall (i.e., high sensitivity but low specificity), while clinicians were better at predicting who would not be readmitted but underestimated readmission risk overall (i.e., high specificity but low sensitivity). A human-plus-machine approach incorporating clinician predictions as a variable in the EHR model had significantly better discrimination and also best optimized sensitivity and specificity. In other words, human-plus-machine was better than either alone in terms of predicting 30-day hospital readmissions.



**Fig. 2** Net reclassification for human and human-plus-machine predictions. NRI, net reclassification index. <sup>a</sup>Denotes the reference model for the described comparison. For panel A, clinician (“human”) predictions are compared with EHR model (“machine”) predictions as a reference; for panel B, a human-plus-machine predictions are compared with EHR model predictions as the reference standard. <sup>b</sup>Overall NRI is the sum of the event (readmissions) NRI and non-event (non-readmissions) NRI (i.e., the sum of the net proportions of correctly reclassified persons with and without the event of interest, respectively). For example, in panel A, 9.3% of readmissions (solid teal bar) were missed by the EHR model but correctly predicted by clinicians. However, 40.7% of readmissions were missed by clinicians but correctly predicted by EHR model. Thus, overall clinicians were 31.3% worse than the EHR model in correctly predicting readmissions (event NRI = 9.3 + -40.7 = -31.3). The black solid and hatched bars can be similarly interpreted to arrive at the non-event NRI, which indicates that clinicians were overall 30.1% better than the EHR model at predicting non-readmissions. The overall NRI is the sum of the event NRI (-31.3%) and the non-event NRI (+30.1%), which for panel A is equal to -1.2%. In other words, clinicians were overall about -1.2% worse at correctly classifying readmissions and non-readmissions compared with the EHR model, with most of this difference attributable to clinicians failing to predict readmissions that were correctly identified by the EHR model.



**Fig. 3** Diagnostic test characteristics of human, machine, and human-plus-machine predictions. CI, confidence interval; EHR, electronic health record; NRI, net reclassification index. Leaf plots illustrate the usefulness of positive vs. negative predictions by clinicians (human), the EHR model (machine), and a human-plus-machine prediction model in correctly “diagnosing” a readmission across a range of pre-test probabilities. The impact of a positive prediction is shown by the shaded area above the diagonal red line and the impact of a negative prediction is shown by the shaded area below the diagonal red line.<sup>40</sup> <sup>a</sup>Sensitivity defined as the probability that an individual who was readmitted had a “high-risk” prediction. <sup>b</sup>Specificity defined as the probability that an individual who was not readmitted had a “low-risk” prediction. <sup>c</sup>Positive likelihood ratio defined as the probability of a “high-risk” prediction in those who were readmitted divided by the probability of a “high-risk” prediction in those who were not readmitted—i.e., sensitivity/(1-specificity). <sup>d</sup>Negative likelihood ratio defined as the probability of a “low-risk” prediction in those who were readmitted divided by the probability of a “low-risk” prediction in those who were not readmitted—i.e., (1-sensitivity)/specificity.

Our findings have several implications for hospitals and health systems developing workflows to identify and target hospitalized patients at risk for 30-day readmissions. Hospitals with limited care transitions resources or lacking real-time predictive analytic capabilities could opt for a “human-only” approach to reducing readmissions, where readmission reduction interventions would be targeted to patients who were identified as being at high-risk for readmission by their inpatient physicians using a 1-question screening tool that could be embedded directly into an EHR. Resources would thus be allocated to those most likely to potentially benefit from intervention, though many high-risk patients would be missed as a tradeoff of this approach. Hospitals for whom readmissions reduction is a high priority—and who have actionable, predictive analytic resources—could consider a “human-plus-machine” approach to target a larger number of patients for intervention, though some patients identified in this approach may be less likely to be readmitted, and thus may benefit less from intervention.

There are several potential reasons why clinicians were not better at predicting readmissions compared with the EHR model. First, clinicians may be overly optimistic when prognosticating readmission risk in general, as this phenomenon has been observed for prognostication in other areas.<sup>42-44</sup> Second, clinicians may have an optimism bias for the both the effectiveness of their care on readmission risk, and also for their own patients’ ability to avoid readmission compared with other patients in general.<sup>45,46</sup> Third, readmissions may be challenging to predict for both humans and machines due to the breadth and varying influence of social, behavioral, environmental, financial, medical and structural factors on readmission risk.<sup>12,22,23,25,28,30,47-57</sup> Lastly, clinicians’ framework for thinking about readmission risk may not incorporate social,

functional, and behavioral factors, since these readmission risk factors are novel and may not be well integrated in current clinical curricula and continuing medical education.<sup>58-60</sup> To this end, we found that clinicians performed somewhat worse than the EHR model at predicting readmissions if they identified non-medical issues as the main potential reason for readmission, versus medical illness where clinicians were more proficient prognosticators. This directly contradicted our initial study hypothesis that physicians would be better able to predict readmissions due to increased awareness of patient social, behavioral, and situational factors not captured as well in EHR data.

Our study had certain limitations. First, this was a single site study of patients hospitalized over 1 month. However, our cohort of both patients and clinicians is large and represents a diverse population with a wide range of clinical, social, functional, and behavioral needs. Additionally, we ascertained readmissions from hospitals within a 100-mile radius of the study hospital (not just same-site readmissions), and surveys were conducted for consecutive patients to mitigate potential sampling bias. Second, we were unable to explore clinicians’ rationale for their predictions, or for selecting medical vs. non-medical reasons for potential readmission due to the structured format of our questionnaire, since we designed the survey to be both brief and of low burden to busy practicing clinicians to facilitate recruitment. Future studies in this area would be valuable to not only improve readmission risk prediction but to also identify potentially modifiable social, functional, and behavioral risk factors to target for intervention. Finally, our human perceptions of risk only include the input of clinicians and not other members of the care team, such as nursing staff, pharmacists, social workers, or case managers. Future studies assessing predictions and perspectives from other members of

an interdisciplinary care team could be helpful to inform a more global perspective of readmission risk and potential targets for intervention.

## CONCLUSION

In conclusion, a human-plus-machine approach was better than either human or machine alone to predict 30-day hospital readmission. Readmission risk prediction strategies should seek to incorporate clinician assessments of readmission risk to optimize the accuracy of readmission predictions.

**Acknowledgments:** The authors thank Ida Sim, MD, PhD, for inspiring this study; Claudia Chavez, MBA, and Cynthia Ortiz, MPH, for their assistance with recruitment and data collection; and the faculty in the Section of Hospital Medicine at Parkland Health & Hospital System in Dallas, Texas, for their assistance with this study. Study data were collected in part using REDCap electronic data capture tools hosted at UT Southwestern.

**Corresponding Author:** Oanh Kieu Nguyen, MD, MAS; Division of Hospital Medicine at San Francisco General Hospital, University of California, San Francisco, San Francisco, CA, USA (e-mail: Oanh.Nguyen@ucsf.edu).

**Author Contributions** Dr. Nguyen had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

**Funding** This study was supported by the Agency for Healthcare Quality and Research through the UT Southwestern Center for Patient-Centered Outcomes Research (AHRQ R24 HS022418) and in part by the National Heart, Lung, and Blood Institute (Nguyen, NHLBI 1K23HL133441), the National Institute on Aging (Makam, NIA 5K23AG052603), and the National Center for Advancing Translational Sciences of the National Institutes of Health (Halm, UL1TR001105).

### Compliance with Ethical Standards:

**Conflict of Interest:** The authors declare that they do not have a conflict of interest.

**Role of the Funder/Sponsor:** The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; nor in the decision to submit the manuscript for publication.

**Supplementary Information:** The online version contains supplementary material available at <https://doi.org/10.1007/s11606-020-06355-3>.

## REFERENCES

- Jencks SF, Williams MV, Coleman EA. Rehospitalizations among patients in the Medicare fee-for-service program. *N Engl J Med* 2009;360(14):1418-1428.
- Trudnak T, Kelley D, Zerzan J, Griffith K, Jiang HJ, Fairbrother GL. Medicaid admissions and readmissions: understanding the prevalence, payment, and most common diagnoses. *Health Aff (Millwood)* 2014;33(8):1337-1344.
- Ranasinghe I, Wang Y, Dharmarajan K, Hsieh AF, Bernheim SM, Krumholz HM. Readmissions after hospitalization for heart failure, acute myocardial infarction, or pneumonia among young and middle-aged adults: a retrospective observational cohort study. *PLoS Med* 2014;11(9):e1001737.
- McIvnenan CK, Eapen ZJ, Allen LA. Hospital readmissions reduction program. *Circulation*. 2015;131(20):1796-1803.
- The Patient Protection and Affordable Care Act. §6410. In:2010.
- Rennke S, Nguyen OK, Shoeb MH, Magan Y, Wachter RM, Ranji SR. Hospital-initiated transitional care interventions as a patient safety strategy: a systematic review. *Ann Intern Med* 2013;158(5 Pt 2):433-440.
- Goldman LE, Sarkar U, Kessel E, et al. Support from hospital to home for elders: a randomized trial. *Ann Intern Med* 2014;161(7):472-481.
- Hansen LO, Greenwald JL, Budnitz T, et al. Project BOOST: effectiveness of a multihospital effort to reduce rehospitalization. *J Hosp Med* 2013;8(8):421-427.
- Hansen LO, Young RS, Hinami K, Leung A, Williams MV. Interventions to reduce 30-day rehospitalization: a systematic review. *Ann Intern Med* 2011;155(8):520-528.
- Amarasingham R, Patel PC, Toto K, et al. Allocating scarce resources in real-time to reduce heart failure readmissions: a prospective, controlled study. *BMJ Qual Saf* 2013;22(12):998-1005.
- Amarasingham R, Moore BJ, Tabak YP, et al. An automated model to identify heart failure patients at risk for 30-day readmission or death using electronic medical record data. *Med Care* 2010;48(11):981-988.
- Kansagara D, Englander H, Salanitro A, et al. Risk prediction models for hospital readmission: a systematic review. *JAMA*. 2011;306(15):1688-1698.
- Amarasingham R, Velasco F, Xie B, et al. Electronic medical record-based multicondition models to predict the risk of 30 day readmission or death among adult medicine patients: validation and comparison to existing models. *BMC Med Inform Decis Making* 2015;15:39.
- Nguyen OK, Makam AN, Clark C, et al. Predicting all-cause readmissions using electronic health record data from the entire hospitalization: Model development and comparison. *J Hosp Med* 2016;11(7):473-480.
- Makam AN, Nguyen OK, Clark C, et al. Predicting 30-Day Pneumonia Readmissions Using Electronic Health Record Data. *J Hosp Med* 2017;12(4):209-216.
- Nguyen OK, Makam AN, Clark C, Zhang S, Das SR, Halm EA. Predicting 30-Day Hospital Readmissions in Acute Myocardial Infarction: The AMI "READMITS" (Renal Function, Elevated Brain Natriuretic Peptide, Age, Diabetes Mellitus, Nonmale Sex, Intervention with Timely Percutaneous Coronary Intervention, and Low Systolic Blood Pressure) Score. *J Am Heart Assoc*. 2018;7(8).
- Brungger B, Blozik E. Hospital readmission risk prediction based on claims data available at admission: a pilot study in Switzerland. *BMJ Open* 2019;9(6):e028409.
- Kabue S, Greene J, Kipnis P, et al. The Impact of Pharmacy-specific Predictors on the Performance of 30-Day Readmission Risk Prediction Models. *Med Care* 2019;57(4):295-299.
- Lindenauer PK, Lagu T, Rothberg MB, et al. Income inequality and 30 day outcomes after acute myocardial infarction, heart failure, and pneumonia: retrospective cohort study. *BMJ*. 2013;346:f521.
- Calvillo-King L, Arnold D, Eubank KJ, et al. Impact of social factors on risk of readmission or mortality in pneumonia and heart failure: systematic review. *J Gen Intern Med* 2013;28(2):269-282.
- Joynt Maddox KE, Reidhead M, Hu J, et al. Adjusting for social risk factors impacts performance and penalties in the hospital readmissions reduction program. *Health Serv Res* 2019;54(2):327-336.
- Weinreich M, Nguyen OK, Wang D, et al. Predicting the Risk of Readmission in Pneumonia. A Systematic Review of Model Performance. *Ann Am Thorac Soc* 2016;13(9):1607-1614.
- Smith LN, Makam AN, Darden D, et al. Acute Myocardial Infarction Readmission Risk Prediction Models: A Systematic Review of Model Performance. *Circ Cardiovasc Qual Outcomes* 2018;11(1):e003885.
- Jiang W, Siddiqui S, Barnes S, et al. Readmission Risk Trajectories for Patients With Heart Failure Using a Dynamic Prediction Approach: Retrospective Study. *JMIR Med Inform* 2019;7(4):e14756.
- Greysen SR, Stijacic Cenzer I, Auerbach AD, Covinsky KE. Functional impairment and hospital readmission in Medicare seniors. *JAMA Intern Med* 2015;175(4):559-565.
- Damiani G, Salvatori E, Silvestrini G, et al. Influence of socioeconomic factors on hospital readmissions for heart failure and acute myocardial infarction in patients 65 years and older: evidence from a systematic review. *Clin Interv Aging* 2015;10:237-245.

27. **Tonkikh O, Shadmi E, Flaks-Manov N, Hoshen M, Balicer RD, Zisberg A.** Functional status before and during acute hospitalization and readmission risk identification. *J Hosp Med* 2016;11(9):636-641.
28. **Meddings J, Reichert H, Smith SN, et al.** The Impact of Disability and Social Determinants of Health on Condition-Specific Readmissions beyond Medicare Risk Adjustments: A Cohort Study. *J Gen Intern Med* 2017;32(1):71-80.
29. **Garcia-Olmos L, Aguilar R, Lora D, et al.** Development of a predictive model of hospitalization in primary care patients with heart failure. *PLoS One* 2019;14(8):e0221434.
30. **Dodson JA, Hajduk AM, Murphy TE, et al.** Thirty-Day Readmission Risk Model for Older Adults Hospitalized With Acute Myocardial Infarction. *Circ Cardiovasc Qual Outcomes* 2019;12(5):e005320.
31. **Adler NE, Stead WW.** Patients in context—EHR capture of social and behavioral determinants of health. *N Engl J Med* 2015;372(8):698-701.
32. **Allaudeen N, Schnipper JL, Orav EJ, Wachter RM, Vidyarthi AR.** Inability of providers to predict unplanned readmissions. *J Gen Intern Med* 2011;26(7):771-776.
33. **Miller WD, Nguyen K, Vangala S, Dowling E.** Clinicians can independently predict 30-day hospital readmissions as well as the LACE index. *BMC Health Serv Res* 2018;18(1):32.
34. **Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG.** Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009;42(2):377-381.
35. Texas Medical Board. Licensee Database. 2018; <http://www.tmb.state.tx.us/>. Accessed October 17, 2018.
36. **Nguyen OK, Makam AN, Clark C, et al.** Vital Signs Are Still Vital: Instability on Discharge and the Risk of Post-Discharge Adverse Outcomes. *J Gen Intern Med* 2017;32(1):42-48.
37. **Nguyen OK, Vazquez MA, Charles L, et al.** Association of Scheduled vs Emergency-Only Dialysis With Health Outcomes and Costs in Undocumented Immigrants With End-stage Renal Disease. *JAMA Intern Med* 2019;179(2):175-183.
38. **Steyerberg EW, Vickers AJ, Cook NR, et al.** Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology*. 2010;21(1):128-138.
39. **Leening MJ, Vedder MM, Witteman JC, Pencina MJ, Steyerberg EW.** Net reclassification improvement: computation, interpretation, and controversies: a literature review and clinician's guide. *Ann Intern Med* 2014;160(2):122-131.
40. **Coulthard MG, Coulthard T.** The leaf plot: a novel way of presenting the value of tests. *Br J Gen Pract* 2019;69(681):205-206.
41. Centers for Medicare & Medicaid Services. *2017 All-Cause Hospital-Wide Measure Updates and Specifications Report: Hospital-Level 30-Day Risk-Standardized Readmission Measure — Version 6.0.* March 2017.
42. **Christakis NA, Lamont EB.** Extent and determinants of error in doctors' prognoses in terminally ill patients: prospective cohort study. *BMJ*. 2000;320(7233):469-472.
43. **Durstenfeld MS, Saybolt MD, Praestgaard A, Kimmel SE.** Physician predictions of length of stay of patients admitted with heart failure. *J Hosp Med* 2016;11(9):642-645.
44. **Glare P, Virik K, Jones M, et al.** A systematic review of physicians' survival predictions in terminally ill cancer patients. *BMJ*. 2003;327(7408):195-198.
45. **Sharot T.** The optimism bias. *Curr Biol* 2011;21(23):R941-945.
46. **Clyne W, McLachlan S, Mshelia C, et al.** "My patients are better than yours": optimistic bias about patients' medication adherence by European health care professionals. *Patient Prefer Adher* 2016;10:1937-1944.
47. **Graham KL, Auerbach AD, Schnipper JL, et al.** Preventability of Early Versus Late Hospital Readmissions in a National Cohort of General Medicine Patients. *Ann Intern Med* 2018;168(11):766-774.
48. **Jencks SF, Schuster A, Dougherty GB, Gerovich S, Brock JE, Kind AJH.** Safety-Net Hospitals, Neighborhood Disadvantage, and Readmissions Under Maryland's All-Payer Program: An Observational Study. *Ann Intern Med* 2019;171(2):91-98.
49. **Keeney T, Jette DU, Cabral H, Jette AM.** Frailty and Function in Heart Failure: Predictors of 30-Day Hospital Readmission? *J Geriatr Phys Ther*. 2019.
50. **Kind AJ, Jencks S, Brock J, et al.** Neighborhood socioeconomic disadvantage and 30-day rehospitalization: a retrospective cohort study. *Ann Intern Med* 2014;161(11):765-774.
51. **Navathe AS, Zhong F, Lei VJ, et al.** Hospital Readmission and Social Risk Factors Identified from Physician Notes. *Health Serv Res* 2018;53(2):1110-1136.
52. **Nijhawan AE, Metsch LR, Zhang S, et al.** Clinical and Sociobehavioral Prediction Model of 30-Day Hospital Readmissions Among People With HIV and Substance Use Disorder: Beyond Electronic Health Record Data. *J Acquir Immune Defic Syndr* 2019;80(3):330-341.
53. **Panagiotou OA, Kumar A, Gutman R, et al.** Hospital Readmission Rates in Medicare Advantage and Traditional Medicare: A Retrospective Population-Based Analysis. *Ann Intern Med* 2019;171(2):99-106.
54. **Shih SL, Gerrard P, Goldstein R, et al.** Functional Status Outperforms Comorbidities in Predicting Acute Care Readmissions in Medically Complex Patients. *J Gen Intern Med* 2015;30(11):1688-1695.
55. **Sieck C, Adams W, Burkhart L.** Validation of the BOOST Risk Stratification Tool as a Predictor of Unplanned 30-Day Readmission in Elderly Patients. *Qual Manag Health Care* 2019;28(2):96-102.
56. **Singh S, Goodwin JS, Zhou J, Kuo YF, Nattinger AB.** Variation Among Primary Care Physicians in 30-Day Readmissions. *Ann Intern Med* 2019;170(11):749-755.
57. **Zhou H, Della PR, Roberts P, Goh L, Dhaliwal SS.** Utility of models to predict 28-day or 30-day unplanned hospital readmissions: an updated systematic review. *BMJ Open* 2016;6(6):e011060.
58. **Daniel H, Bornstein SS, Kane GC.** Health, Public Policy Committee of the American College of P. Addressing Social Determinants to Improve Patient Care and Promote Health Equity: An American College of Physicians Position Paper. *Ann Intern Med* 2018;168(8):577-578.
59. **Doobay-Persaud A, Adler MD, Bartell TR, et al.** Teaching the Social Determinants of Health in Undergraduate Medical Education: a Scoping Review. *J Gen Intern Med* 2019;34(5):720-730.
60. **Gard LA, Peterson J, Miller C, et al.** Social Determinants of Health Training in U.S. Primary Care Residency Programs: A Scoping Review. *Acad Med* 2019;94(1):135-143.

**Publisher's Note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.