

Effect of a Real-Time Risk Score on 30-day Readmission Reduction in Singapore

Christine Xia Wu¹ Ernest Suresh² Francis Wei Loong Phng¹ Kai Pik Tai¹ Janthorn Pakdeethai²
 Jared Louis Andre D'Souza² Woan Shin Tan³ Phillip Phan^{4,5} Kelvin Sin Min Lew¹
 Gamaliel Yu-Heng Tan⁶ Gerald Seng Wee Chua² Chi Hong Hwang¹

¹Quality, Innovation and Improvement, Ng Teng Fong General Hospital, Singapore

²Department of Medicine, Ng Teng Fong General Hospital, Singapore

³Health Services and Outcomes Research, National Healthcare Group, Singapore

⁴Department of Medicine, Johns Hopkins University, Baltimore, Maryland, United States

⁵Department of Medicine, National University of Singapore, Singapore

⁶Group Medical Informatics Office, National University Health System, Singapore

Address for correspondence Christine Xia Wu, MSc, Quality, Innovation and Improvement, Ng Teng Fong General Hospital, 1 Jurong East St 21, Singapore 609606, Singapore (e-mail: christine_wu@nuhs.edu.sg).

Appl Clin Inform 2021;12:372–382.

Abstract

Objective To develop a risk score for the real-time prediction of readmissions for patients using patient specific information captured in electronic medical records (EMR) in Singapore to enable the prospective identification of high-risk patients for enrolment in timely interventions.

Methods Machine-learning models were built to estimate the probability of a patient being readmitted within 30 days of discharge. EMR of 25,472 patients discharged from the medicine department at Ng Teng Fong General Hospital between January 2016 and December 2016 were extracted retrospectively for training and internal validation of the models. We developed and implemented a real-time 30-day readmission risk score generation in the EMR system, which enabled the flagging of high-risk patients to care providers in the hospital. Based on the daily high-risk patient list, the various interfaces and flow sheets in the EMR were configured according to the information needs of the various stakeholders such as the inpatient medical, nursing, case management, emergency department, and postdischarge care teams.

Results Overall, the machine-learning models achieved good performance with area under the receiver operating characteristic ranging from 0.77 to 0.81. The models were used to proactively identify and attend to patients who are at risk of readmission before an actual readmission occurs. This approach successfully reduced the 30-day readmission rate for patients admitted to the medicine department from 11.7% in 2017 to 10.1% in 2019 ($p < 0.01$) after risk adjustment.

Conclusion Machine-learning models can be deployed in the EMR system to provide real-time forecasts for a more comprehensive outlook in the aspects of decision-making and care provision.

Keywords

- ▶ machine learning
- ▶ data modeling
- ▶ implementation and deployment
- ▶ testing and evaluation
- ▶ electronic health records and systems

received
 October 29, 2020
 accepted after revision
 May 12, 2021

© 2021. Thieme. All rights reserved.
 Georg Thieme Verlag KG,
 Rüdigerstraße 14,
 70469 Stuttgart, Germany

DOI <https://doi.org/10.1055/s-0041-1726422>.
 ISSN 1869-0327.

Background and Significance

Hospital readmissions are commonly associated with increasing health care costs. This poses additional burden on the health care system and exerts pressure on finite health care resources.^{1–3} Patients who are frequently admitted to the hospital contribute to bed shortage and may also experience significant psychological stress and financial burden.⁴ Tusó et al posited two key reasons for multiple readmissions: complex underlying medical conditions and complex social issues.⁵

Studies suggest that up to a third of readmissions were preventable, thus offering a significant opportunity for the application of effective interventions.^{6–12} Inadequate preparation for discharge and insufficient follow-up after discharge were among the reasons contributing to preventable readmissions.^{13,14} Thus, the identification of high-risk patients for proactive transitional care interventions across hospital and community settings may help to reduce readmissions.^{15,16} The results of a systematic review and meta-analysis of randomized trials by Leppin et al showed that such interventions had a consistent and beneficial effect on 30-day readmissions, with those consisting of multiple components being more effective.¹⁷ A systematic review conducted by Hansen et al corroborated with this viewpoint as he posited that isolated interventions showed limited efficacy.¹⁸

With the implementation of electronic medical records (EMR) systems, patient data and temporal clinical information are available at the point of care. The volume of data available in EMR has grown exponentially in recent years. Many machine learning algorithms had been developed to identify patterns and associations to aid in predictions and make useful sense of the wealth of data.^{19–27} These technologies serve as drivers for clinical decision support systems that enable informed data-driven decision-making, support clinical research, and improve quality of care.^{28,29} Many health systems have developed 30-day readmission models to identify patients with high risk of readmissions for risk stratification and predictive modeling.^{30–35} These tools relied on retrospective or real-time administrative data, with c-statistics ranging from 0.55 to 0.72.³⁰ However, very few models have incorporated social determinants, although it is well known that such variables may disproportionately influence the risk of readmissions in socioeconomically disadvantaged populations.^{30,35–44}

While existing models perform well, few studies have considered the impact of deploying a real-time model in tandem with risk stratified interventions. By adopting the Canadian LACE score (c-statistic: 0.68) together with a transitional care bundle, Kaiser Permanente had shown significant reduction in their 30-day readmission rates.^{5,33} The LACE model had been validated with a c-statistic of 0.70 in Singapore.⁴⁵ The aim of this work was to develop and implement a model by using patient specific information captured in EMR, which can be updated systematically in real time to prospectively identify high-risk patients and address their needs early during their hospitalization.^{45–47}

Methods

Study Design

The key to reducing readmission was developing a system that determines high-risk patients during the index admission that could be standardized and integrated into our EMR system. Patient health records were retrieved retrospectively from the hospital EMR to construct a “reference standard” dataset by using our national standard as a reference.⁴⁸ A split-sample design was then used for the model training and internal validation. The developed model was transformed into a risk score and embedded into the hospital EMR with automated generation of the 30-day readmission risk score in September 2017. Using data derived from the EMR, the tool automatically generates a “readmission risk score” based on patient information by day 2 of the patient’s index admission. To determine the accuracy, a case manager manually reviewed the patients’ data and readmissions for 6 months after deployment of the model.

We engaged the various user groups for feedback on their day-to-day use of the EMR system to determine the optimal placement of the risk flag. A real-time high-risk flag was configured in EMR storyboards and flow sheets to meet the needs of the inpatient medical, nursing, case management, ED, and postdischarge care teams.^{5,29} The list of 30-day readmission high-risk patients was flagged to clinicians and case managers in real time for early interventions during the patients’ stay and postdischarge (→ Fig. 1C). In addition to the stratified predicted risk of 30-day all-cause hospital readmission, the report also includes other important information, including other predictive variables that can facilitate the care team’s workup for each patient. The 30-day readmission rates were risk-adjusted based on patient’s characteristics, and we compared the risk-adjusted rates between pre- and postimplementation of the risk score generation in the hospital EMR.

Patient Settings

Singapore’s public health care system was organized into three integrated clusters, namely National University Health System (NUHS), National Healthcare Group (NHG), and Singapore Health Services.⁴⁹ Ng Teng Fong General Hospital (NTFGH) is a 700-bedded hospital on the western side of Singapore residing under NUHS which started operations in June 2015. In April 2017, our hospital’s 30-day readmission rate was as high as 15%, which was noticeably higher than the national average. Thus, we were tasked to develop and validate the readmission predictive model and chose our study period to be January 2016 to December 2016 for the model development to provide sufficient runway to ensure a robust validation of the model’s performance for year 2017 and onwards. Our cohort consisted of 25,472 patients discharged from the medicine department in NTFGH during this study period.

The patient health records were retrospectively extracted from the hospital EMR based on inpatient, outpatient, and emergency department (ED) settings. Patients were excluded if they were transferred to other acute hospitals, continuing care facilities or if they died during their inpatient stay.

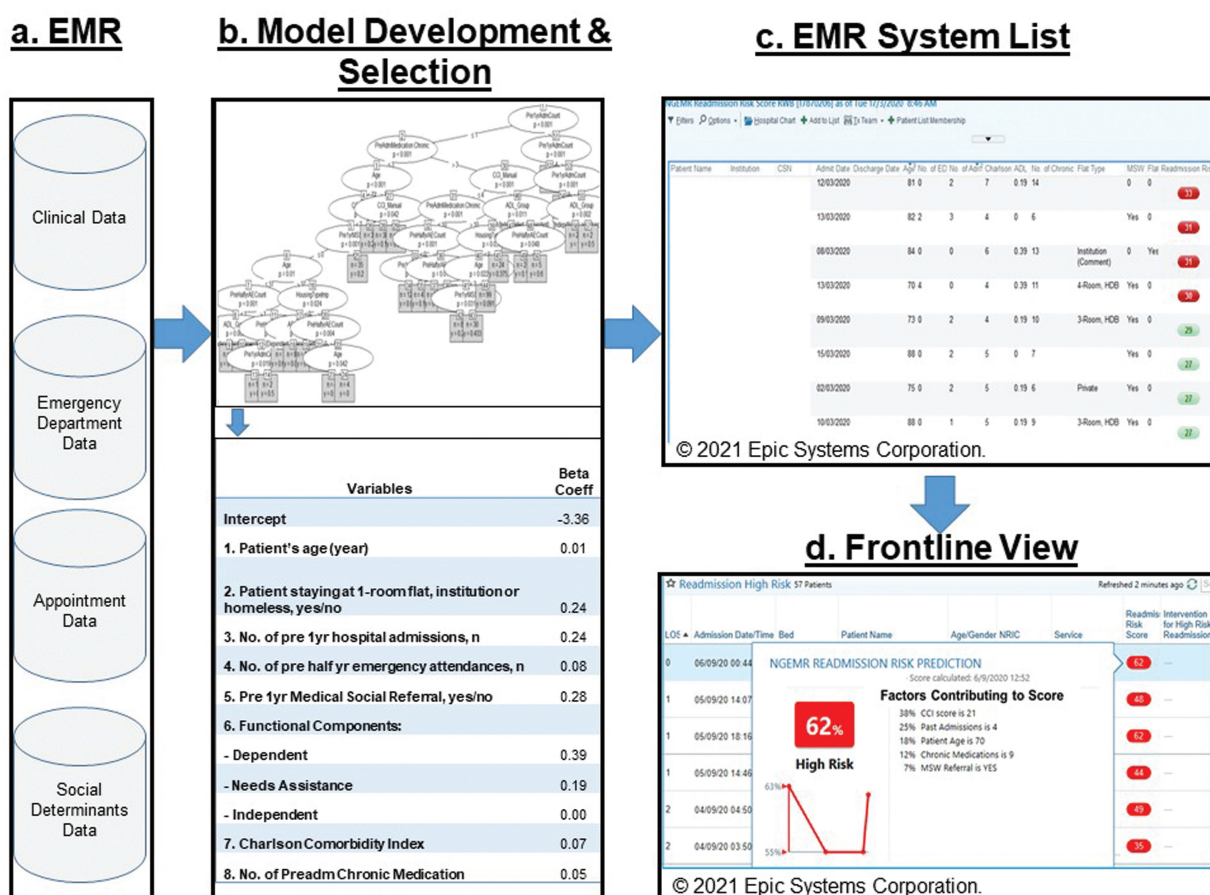


Fig. 1 Development and usage of risk score.

Predictor Variables

In the development of the model, we aimed to include predictors with clinical significance which were either based on advice from local clinicians or from the literature to generate actionable insights for the local setting. As we sought to build a model to identify high-risk patients by day 2 of their admission, we chose variables which would be available at the time of admission. To construct a comprehensive range of possible variables, we researched predictors included in literature such as the Canadian LACE score,³³ obtained inputs from clinicians and conducted feature engineering.^{50–54} The selected variables comprising of patient demographics, medical conditions, socioeconomic status, functional scores, and health care utilization data obtained from the EMR (→ **Fig. 1A**) are described in → **Supplementary Table S1** (available in the online version).

Socioeconomic predictors were chosen based on locally known socioeconomic determinants. Public rental housing, which is heavily subsidized to cater to lower income households, had been validated as a sensitive indicator of area-level socioeconomic status in Singapore.⁵² Admission ward classes A, B1, B2, and C are categorized according to different levels of government subsidies.^{7,55} Medical social referral is a proxy predictor indicating the presence of unresolved social issues that may influence health care utilization, which includes the aspects of casework management, counseling and therapy, financial assistance,

suicide assessment and intervention, interpersonal violence intervention, psychological trauma support, home visits, and support groups.⁵⁶

Statistical Methods

Supervised machine learning models for binary classification were built to estimate the probability of readmission within 30 days (→ **Fig. 1B**). Multivariable logistic regression was used to measure the effect of predictors associated with 30-day readmission.⁵⁷ Gradient boosting machines method builds the model in a stage-wise boosting method and generalizes them by allowing optimization of an arbitrary differentiable loss function.⁵⁸ Conditional decision trees and generalized linear model trees are decision tree methods by recursive partitioning of dependent variables.^{59–61} Random Forest is an ensemble learner that generates many decision trees and uses majority “voting” of all the trees’ outcomes to decide on the binary classification.^{62,63} Goodness-of-fit tests such as Hosmer–Lemeshow test and out-of-bag crossvalidation were conducted to examine the model calibration of the corresponding machine learning models.

A split-sample design was applied to derive and internally validate the predictive model. The dataset was randomly split into two 70% for model derivation and 30% for model validation. We used area under the receiver operating characteristic (AUC) with bootstrap method to evaluate the effectiveness of each prediction model.⁶⁴ AUC was determined by setting

thresholds according to the probability for prediction of readmission within 30 days calculated from the test data, then calculating true positive (TP), and false positive (FP) ratios sequentially while shifting the threshold level serially. Where the threshold level at which was minimal, the presence/absence of readmission within 30 days was labeled, followed by calculation of accuracy, precision, recall, and f1-measure. We decided to use bootstrap validation over k-fold cross-validation as the latter tends to overestimate the extra-sample error for regression tree algorithms.⁶⁵

The readmission rates were risk-adjusted by using our nationally adopted methodology for the Public Hospital Performance Report.⁴⁸ The risk adjustment factors used were age, gender, Charlson Comorbidity Index, diagnosis-related group type, number of hospitalizations in the past year, admission type, and principal diagnosis group of the initial episode. The reference population was all public health care institutions discharges in Singapore for the year 2012. All statistical analyses were conducted by using SPSS version 21 (SPSS Inc, Chicago Ill) and R 3.5.

Model Deployment

We used the Epic EMR storyboard interface tool to configure various interfaces and flow sheets to meet the differing information needs of the different health care professionals in their day-to-day work. There are eight major types of storyboards, namely inpatient, ED/urgent care, outpatient, operating room/procedure areas, therapy, link, patient access, and revenue, with variants that target specialty, role, and other contexts. We implemented the risk score flag into the inpatient and ED/urgent care storyboards, ensuring that they could see key information pertinent both to the encounter and to the user's specialty or role in an easy-to-navigate, centralized location. By hovering over the high-risk flag, additional information is provided and detailed information on the risk factors can be easily accessed simply through the use of clicks of a button⁶⁶ (►Fig. 1D).

Bundled Interventions According to Risk Stratification

Intervention care bundle elements were developed based on identified gaps and evidence pertaining to readmission reduction through a review of the literature.^{67–75} These evidence-based care elements were then refined for implementation in NTFGH. The 30-day readmission project workgroup reviews and approves the bundle elements for implementation in consideration of the intervention's effectiveness and our resource limitations. Eight care elements were implemented according to patients' risk stratification as shown in ►Table 1. The details of these elements are described in ►Supplementary Table S2 (available in the online version). Other interventions were also considered but excluded due to resource limitations, for example, "24-hour posthospital discharge hotline" was excluded as we did not have the resources to man a 24-hour hotline. Instead the discharge nurse's phone number or the ward's phone number would be provided to high-risk patients upon discharge.

Results

The model development cohort consisted of 25,472 patients index discharged from the medicine department in NTFGH from January 2016 to December 2016, and 15.2% of them were readmitted within 30 days of discharge. Their baseline characteristics are shown in ►Table 2.

We evaluated five machine learning models for this classification problem, and ►Table 3 shows the mean \pm standard deviation for the bootstrap yielded from evaluation with each technique.⁴⁵ Random forest exhibited the best performance (AUC: 0.806), but we chose logistic regression (AUC: 0.783) as the final model in consideration of the advantages of comparatively easy implementation, computational efficiency, transparency, and easy interpretability.^{75,76}

Stepwise elimination was conducted by using Akaike Information Criterion to retain the significant predictors. We noted that penalized regression methods are used extensively in

Table 1 Intervention bundle elements

S/N	Intervention bundle elements	Risk score	
		Low	High
1	Risk stratification with high-risk flag in EMR	X	X
2	Comprehensive and accurate discharge summary (to improve communication between hospital and primary care settings)	X	X
3	Medication reconciliation (to eliminate outdated medications)	X	X
4	Postdischarge phone call within 72 h of discharge		X
5	Posthospital visits at specialist clinic or polyclinic within 1 wk		X
6	Posthospital home visit		X
7	Initialization of Advance Care Planning (to assess patients' wishes regarding the future health care options)		X
8	Complex-case conference with multidisciplinary teams for patients staying more than 14 days and high risk of readmission		X

Abbreviation: EMR, electronic medical record.

Table 2 Baseline characteristics of model development cohort

Characteristics	Overall cohort	Derivation cohort	Validation cohort
	<i>n</i> = 25,472	<i>n</i> = 17,830	<i>n</i> = 7,642
Age (y), mean (SD)	63.2 (18.4)	63.4 (18.4)	62.8 (18.4)
Race, <i>n</i> (%)			
Chinese	15,940 (62.6)	11,180 (62.7)	4,760 (62.2)
Indian	2,767 (10.9)	1,937 (10.9)	830 (10.9)
Malay	4,178 (16.4)	2,882 (16.2)	1,296 (17.0)
Other	2,587 (10.1)	1,831 (10.2)	756 (9.9)
Gender (male), <i>n</i> (%)	13,638 (53.5)	9,513 (53.4)	4,125 (54.0)
Staying at rental flat, <i>n</i> (%)	1,308 (5.1)	922 (5.2)	386 (5.1)
Healthcare utilization history, mean (SD)			
Hospital admission previous 1 y	0.98 (2.1)	0.98 (2.1)	0.97 (2.1)
Emergency attendance previous half year	0.24 (1.4)	0.25 (1.4)	0.24 (1.4)
Medical social referral previous 1 y, <i>n</i> (%)	5,011 (19.7)	3,554 (19.9)	1,457 (19.1)
Subspecialty, <i>n</i> (%)			
Acute medicine	6,198 (24.3)	4,320 (24.2)	1,878 (24.6)
Endocrinology	1,180 (4.6)	817 (4.6)	363 (4.7)
Gastroenterology and hepatology	1,508 (5.9)	1,065 (6.0)	443 (5.8)
General cardiology	2,899 (11.4)	1,998 (11.2)	901 (11.8)
General medicine	4,419 (17.4)	3,107 (17.4)	1,312 (17.2)
Geriatric medicine	1,136 (4.5)	828 (4.6)	308 (4.0)
Hematology	1 (0.0)	1 (0.0)	0 (0.0)
Infectious diseases	1,829 (7.2)	1,310 (7.4)	519 (6.8)
Medical oncology	139 (0.5)	97 (0.5)	42 (0.5)
Neurology	1,912 (7.5)	1,322 (7.4)	590 (7.7)
Rehabilitation medicine	147 (0.6)	103 (0.6)	44 (0.6)
Renal medicine	1,150 (4.5)	816 (4.6)	334 (4.4)
Respiratory medicine	2,953 (11.6)	2,045 (11.5)	908 (11.9)
Rheumatology	1 (0.0)	1 (0.0)	0 (0.0)
Functional components, <i>n</i> (%)			
Dependent	1,433 (8.0)	632 (8.3)	2,065 (8.1)
Needs assistance	10,289 (57.7)	4,406 (57.6)	14,695 (57.7)
Independent	6,108 (34.3)	2,604 (34.1)	8,712 (34.2)
Emergency admission, <i>n</i> (%)	24,821 (97.4)	17,393 (97.6)	7,428 (97.2)
Admission or transfer to ICU in previous half year, mean (%)	951 (3.7)	662 (3.7)	289 (3.8)
Length of stay at index admission, mean (SD)	4.4 (6.9)	4.4 (6.3)	4.5 (8.0)
Total length of stay of previous 1-year admissions, <i>n</i> (%)			
0 d	15,637 (61.4)	10,946 (61.4)	4,691 (61.4)
1–10 d	6,180 (24.3)	4,301 (24.1)	1,879 (24.6)
11–20 d	1,847 (7.2)	1,298 (7.3)	549 (7.2)
21–30 d	805 (3.2)	574 (3.2)	231 (3.0)
31–40 d	419 (1.6)	292 (1.6)	127 (1.7)
41–50 d	233 (0.9)	169 (1.0)	64 (0.8)
51 d and above	351 (1.4)	250 (1.4)	101 (1.3)

Table 2 (Continued)

Characteristics	Overall cohort	Derivation cohort	Validation cohort
	n = 25,472	n = 17,830	n = 7,642
Charlson Comorbidity Index, n (%)			
0	10,022 (39.4)	7,036 (39.5)	2,986 (39.1)
1	3,308 (13.0)	2,266 (12.7)	1,042 (13.6)
2+	12,142 (47.6)	8,528 (47.8)	3,614 (47.3)
Preadmission chronic medication, n (%)			
0	5,758 (22.6)	4,043 (22.7)	1,715 (22.4)
1–5	9,310 (36.6)	6,524 (36.6)	2,786 (36.5)
6+	10,404 (40.8)	7,263 (40.7)	3,141 (41.1)
Index discharge class, n (%)			
Class A	2,276 (8.9)	1,585 (8.9)	691 (9.0)
Class B1	1,098 (4.3)	765 (4.3)	333 (4.4)
Class B2	6,118 (24.0)	4,330 (24.3)	1,788 (23.4)
Class C	15,980 (62.8)	11,150 (62.5)	4,830 (63.2)

Abbreviations: ICU, intensive care unit; SD, standard deviation.

Table 3 Machine learning model comparison

Machine learning model	AUC mean \pm SD	Accuracy mean \pm SD	Precision mean \pm SD	Recall mean \pm SD	f1 mean \pm SD
Logistic regression	0.783 \pm 0.019	0.803 \pm 0.012	0.392 \pm 0.014	0.422 \pm 0.016	0.406 \pm 0.022
Conditional decision trees	0.791 \pm 0.016	0.805 \pm 0.011	0.378 \pm 0.018	0.413 \pm 0.021	0.395 \pm 0.020
Generalized linear model tree	0.774 \pm 0.030	0.801 \pm 0.013	0.375 \pm 0.018	0.409 \pm 0.018	0.391 \pm 0.022
Random forest	0.806 \pm 0.030	0.811 \pm 0.012	0.374 \pm 0.014	0.429 \pm 0.016	0.401 \pm 0.020
Gradient boosting	0.792 \pm 0.026	0.804 \pm 0.011	0.395 \pm 0.015	0.425 \pm 0.019	0.409 \pm 0.016

Abbreviations: AUC, area under the receiver operating characteristic; SD, standard deviation.

Table 4 Final eight variables for 30-day readmission risk score real-time generation in electronic medical record

Variables	β coefficient	Odds ratio (95% CI)
Intercept	−3.36	
Patient's age	0.01	1.02 (1.01–1.02)
Patient staying at 1-room flat, institution or homeless	0.24	1.39 (1.09–1.76)
Number of hospital admissions in previous 1 y	0.24	1.07 (1.01–1.14)
Number of emergency attendances in previous half year	0.08	1.16 (1.08–1.24)
Medical social referral in previous 1 y	0.28	1.96 (1.56–2.45)
Functional components:		
Dependent	0.39	1.50 (1.45–1.54)
Needs assistance	0.19	1.17 (1.09–1.24)
Charlson Comorbidities Index	0.07	1.11 (1.05–1.18)
Number of preadmission chronic medication	0.05	1.05 (1.03–1.07)

high-dimensional model selection problems. However, they are less interpretable than stepwise selection for their predictive performance in low-dimensional problems.^{76,77} Thus, we chose stepwise selection over penalized regression. Eight predictors were selected for the final logistic model, which

was transformed into a risk score by using the β coefficients (\rightarrow Table 4). For operational implementation with consideration of case management manpower, we focused on the top 15% of the cohort and set the cutoff of high-risk patients at risk score of 30 and above.

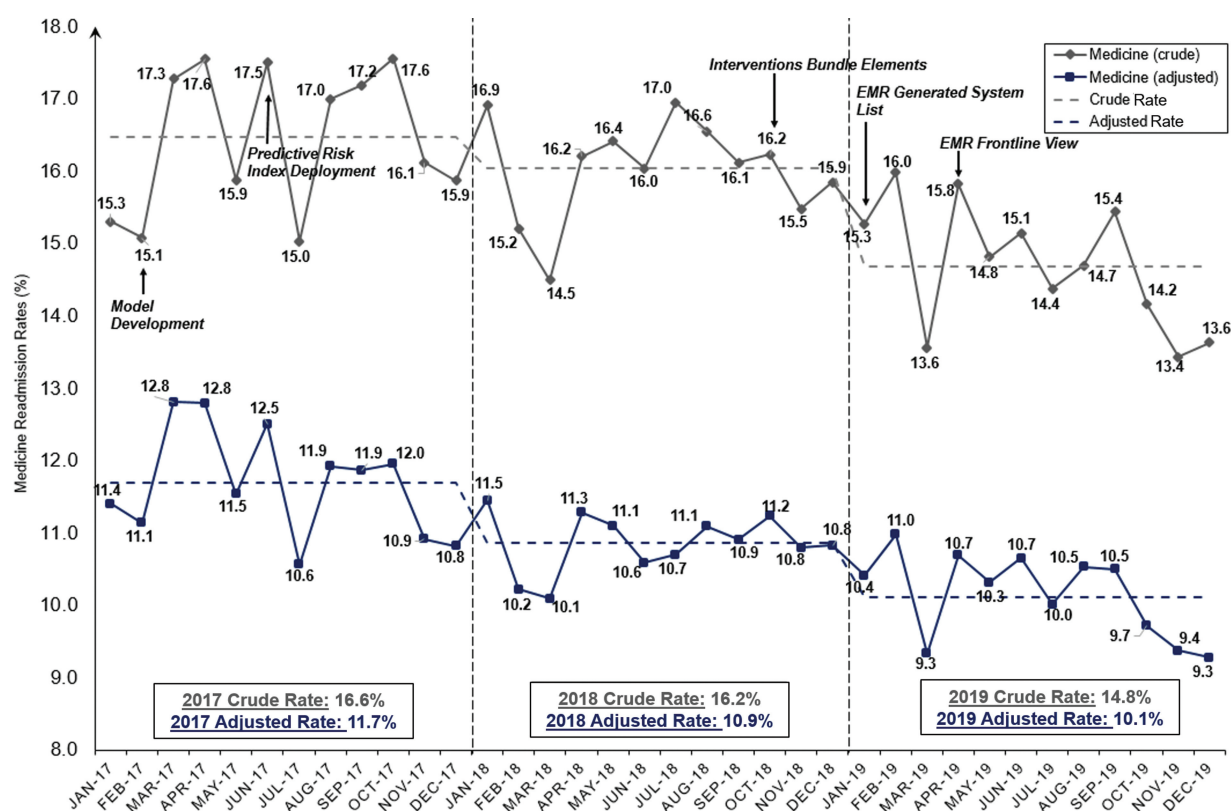


Fig. 2 Medicine 30-day readmission trend line.

30-Day Readmission Intervention Results

Since the deployment of our risk score and implementation of our readmission interventions during the last quarter of 2018, there has been significant improvements in both our crude and risk-adjusted readmission rates. The medicine 30-day readmission crude rate improved from 16.6 to 14.8% ($p < 0.01$), and the risk-adjusted rate improved from 11.7 to 10.1% ($p < 0.01$) between 2017 to 2019 (→ Fig. 2). Being able to identify patients with high risk for readmission based on the risk score enabled clinicians to address these risks during admission and to order a postdischarge care bundle for these at risk patients upon discharge. After implementation, these cumulative actions were associated with a 1.6% reduction in the readmission rates, which translates into 3,200 inpatient bed days savings annually.

Discussion

Our study presents a simple validated risk score for estimating the probability that a hospitalized patient would be readmitted within 30 days, with (1) good discriminative power for predicting the risk of 30-day readmission⁴⁵ and (2) intervention bundle care elements according to risk stratification. This score was successfully implemented in NTFGH's EMR system. The model was initially developed for the medicine cohort and showed good performance when expanded to include orthopedics and surgery patient cohorts.

Similar to other studies,^{78,79} the risk score was built to allow for the provision of continuously updated information independent of disease type or reasons for admission, using patient history, social, and clinical information available in real time

from the EMR system. This provides a longitudinal view, which may aid with earlier detection of acute events, discharge planning, and continuity of care.⁷⁹ Logistic regression models are easy to implement, computationally efficient, transparent, and easy to interpret. Thus, they are popular in clinical decision making and highly applicable in clinical practice.^{76,77} Using the eight predictors, the risk score flagged the top 15% of patients as high risk, which represents 40% risk of readmission within 30 days of discharge.

We found that significant reduction in 30-day readmission rates could be achieved with the risk score deployment and intervention care bundle elements.¹ On the ground, clinicians focus on the biological aspect of care and may not be as aware of the psychological and social aspects. The high-risk flag in the EMR system provides an early alert to the clinicians, as well as a view of the individual risk factors which are contributing to an individual patient's high-risk score. This enables early detection of patients at risk of readmission, for early referrals to the appropriate care team, such as medical social workers, case managers, pharmacists, or other allied health professionals. These patients would then be enrolled in tailored interventions to address the individual risk factors accordingly.⁵ Various interventions were identified for reducing readmissions at several fronts which have been assigned to either low or high-risk scores (→ Table 1). The results suggest that appropriate allocation of resources to the higher risk patients can result in meaningful outcomes. Risk stratification enabled the teams to better align resources to the higher risk patients and enable sustainability in the long term.

The risk score has several strengths to support its use. First, the outcome predicted by the risk score is not just important, but clinically relevant and reliably measured. The discrimination of the risk score was comparable if not better than those that are commonly used in other countries.^{24,32,33,80–82} The model can also be replicated by other hospitals in Singapore given that all the predictors can be obtained in a similar fashion upon admission at any acute hospital in Singapore. The risk score was developed on the NTFGH cohort and was adopted for implementation across the NUHS and NHG clusters' EMR system.

Second, we used parameters which were available at the time of admission for the development of our risk score model. This enables the provision of real-time forecasts of risks in terms of patient outcomes,^{6,28} and for high-risk patients to be flagged to the frontline clinical teams in bid of real-time decision-making such as inclusion of preventive interventions by day 2 of the patient's index admission.⁸³

Third, there is increased visibility on each of the components of the risk score, allowing the care team to see which risk factors are contributing to an individual patient's high-risk score for actionable insights and to tailor interventions around these risk factors accordingly.⁸⁴

The risk score is based on a fixed set of variables with fixed coefficients. The next step for research is a "smart" algorithm in which the risk score is continually evaluated against its own accuracy to update the list of variables and the coefficients to reflect case mix seasonality changes.

Prospectively, further research is required to externally validate the risk score and examine the effectiveness of such bundled interventions. We also only considered supervised machine-learning methods for our study and used parameters that were readily available in the EMR system at the point of the study. Additionally, our model was constructed by using a top-to-bottom approach. For future research, we may explore unsupervised learning methods²¹ and use a bottom-up approach for model construction. New parameters will be included in future as new types of data become readily available in EMR systems for use in prediction models.⁸⁵

Conclusion

We have developed a risk score which can provide a reliable basis for the prediction of 30-day unplanned hospital readmission in Singapore. The model presents the possibility for effective prediction of potential readmission cases and allows the stratification of patients according to risk of 30-day readmission to aid in decision-making at the point of care for better integrated care.

Clinical Relevance Statement

We developed a simple risk score predictive model using parameters available at the time of admission to enable real-time forecasts for decision-making and preventive interventions, showing promising results of using such a model. Through bundled interventions combined with risk

stratification, we aligned resource-intensive interventions to high-risk patients, leading to significant reductions in readmission rates. Our risk score model was accepted and extended for implementation by other institutions in their EMR systems.

Multiple Choice Questions

1. Which socioeconomic variable had been validated as a sensitive indicator of socioeconomic status and was included in the risk score model?
 - a. Ward class
 - b. Activity of daily living
 - c. Rental housing
 - d. Income

Correct Answer: The correct answer is option c. A study had shown that public rental housing as an area-level measure of socioeconomic status is independently associated with increased readmission risk and being a frequent hospital admitter and ED user.

2. Why did we use parameters which were available at the time of admission for our risk score model?
 - a. Data are more accurate when gathered at admission
 - b. Enables the provision of real-time forecasts
 - c. Increased visibility on each of the components
 - d. Able to generalize to other hospitals

Correct Answer: The correct answer is option b. Using parameters which were available at the time of admission, the risk score could be generated early into the admission, which thus enables the provision of real-time forecasts for decision-making.

Author Contributions

C.X.W. provided the study concept and design and contributed to the critical review of the statistical analysis and manuscript revision. E.S. participated in the conception and design of the interventions and contributed to the manuscript revision. F.W.L.P. constructed the write-up of the manuscript and participated in the statistical analysis. K. P.T. built the risk score model and performed the statistical analysis. J.P. participated in constructing the intervention write-up and reviewed the manuscript for important intellectual content. J.L.A.D.S. contributed in building the risk score into the NTFGH Epic EMR system. W.S.T. provided advice to the team on the initial model development and reviewed the manuscript for important intellectual content. P.P. reviewed the manuscript for important intellectual content. K.S.M.L. provided inputs on the intervention write-up. G.Y.H.T. oversaw the implementation of the risk score into the hospital's EMR system. G.S.W.C. oversaw the conception and design of the interventions and provided advice on parameters with clinical significance. C.H.H. contributed to the critical review of the study design, statistical analysis, and reviewed the manuscript for important intellectual content. All the authors reviewed and approved the final manuscript.

Note

This study received ethical approval from the National Healthcare Group Domain Specific Review Board (NHG DSRB 2021/00284).

Conflict of Interest

None declared.

Acknowledgments

The authors thank Dr. Sua Jo Nie (Clinical Informatics, National University Polyclinics) and Dr. Sun Yan (Clinical Project Management & Planning, NHG) for their contributions in building the risk score into the NTFGH Epic EMR system and next generation EMR system, respectively. They also thank case managers, Fadilah Binte Ahmed and Kamala D/O Velu, for their contributions in the conception and design of the interventions and case management. They extend their gratitude to the NTFGH Quality, Innovation & Improvement, and Medical Informatics departments for their contributions to this project.

References

- Kripalani S, Theobald CN, Anctil B, Vasilevskis EE. Reducing hospital readmission rates: current strategies and future directions. *Annu Rev Med* 2014;65(01):471–485
- Hines AL (Truven Health Analytics), Barrett ML (ML Barrett, Inc), Jiang HJ (AHRQ), and Steiner CA (AHRQ) Conditions with the Largest Number of Adult Hospital Readmissions by Payer, 2011. HCUP Statistical Brief #172. April 2014. Rockville, MD: Agency for Healthcare Research and Quality
- Felix HC, Seaberg B, Bursac Z, Thostenson J, Stewart MK. Why do patients keep coming back? Results of a readmitted patient survey. *Soc Work Health Care* 2015;54(01):1–15
- Longman JMI, I Rolfe M, Passey MD, et al. Frequent hospital admission of older people with chronic disease: a cross-sectional survey with telephone follow-up and data linkage. *BMC Health Serv Res* 2012;12:373
- Tuso P, Huynh DN, Garofalo L, et al. The readmission reduction program of Kaiser Permanente Southern California—knowledge transfer and performance improvement. *Perm J* 2013;17(03):58–63
- Bates DW, Saria S, Ohno-Machado L, Shah A, Escobar G. Big data in health care: using analytics to identify and manage high-risk and high-cost patients. *Health Aff (Millwood)* 2014;33(07):1123–1131
- Kocher RP, Adashi EY. Hospital readmissions and the Affordable Care Act: paying for coordinated quality care. *JAMA* 2011;306(16):1794–1795
- van Walraven C, Bennett C, Jennings A, Austin PC, Forster AJ. Proportion of hospital readmissions deemed avoidable: a systematic review. *CMAJ* 2011;183(07):E391–E402
- Donzé J, Aujesky D, Williams D, Schnipper JL. Potentially avoidable 30-day hospital readmissions in medical patients: derivation and validation of a prediction model. *JAMA Intern Med* 2013;173(08):632–638
- Oddone EZ, Weinberger M, Horner M, et al. Classifying general medicine readmissions. Are they preventable? Veterans Affairs Cooperative Studies in Health Services Group on Primary Care and Hospital Readmissions. *J Gen Intern Med* 1996;11(10):597–607
- McInnes EG, Joshi DM, O'Brien TD. Failed discharges: setting standards for improvement. *Geriatr Med* 1988;18:35–42
- Balla U, Malnick S, Schattner A. Early readmissions to the department of medicine as a screening tool for monitoring quality of care problems. *Medicine (Baltimore)* 2008;87(05):294–300
- Gruneir A, Dhalla IA, van Walraven C, et al. Unplanned readmissions after hospital discharge among patients identified as being at high risk for readmission using a validated predictive algorithm. *Open Med* 2011;5(02):e104–e111
- Vest JR, Gamm LD, Oxford BA, Gonzalez MI, Slawson KM. Determinants of preventable readmissions in the United States: a systematic review. *Implement Sci* 2010;5:88
- Costantino ME, Frey B, Hall B, Painter P. The influence of a postdischarge intervention on reducing hospital readmissions in a Medicare population. *Popul Health Manag* 2013;16(05):310–316
- Low LL, Tay WY, Ng MJ, Tan SY, Liu N, Lee KH. Frequent hospital admissions in Singapore: clinical risk factors and impact of socioeconomic status. *Singapore Med J* 2018;59(01):39–43
- Leppin AL, Gionfriddo MR, Kessler M, et al. Preventing 30-day hospital readmissions: a systematic review and meta-analysis of randomized trials. *JAMA Intern Med* 2014;174(07):1095–1107
- Hansen LO, Young RS, Hinami K, Leung A, Williams MV. Interventions to reduce 30-day rehospitalization: a systematic review. *Ann Intern Med* 2011;155(08):520–528
- Dash S, Shakyawar SK, Sharma M, et al. Big data in healthcare: management, analysis and future prospects. *J Big Data* 2019;6:54
- McKinsey Global Institute (Dec 2016) The age of Analytics: Competing in a Data-Driven World. Accessed May 2020 at: <https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/the-age-of-analytics-competing-in-a-data-driven-world>
- Futoma J, Morris J, Lucas J. A comparison of models for predicting early hospital readmissions. *J Biomed Inform* 2015;56:229–238
- Frizzell JD, Liang L, Schulte PJ, et al. Prediction of 30-day all-cause readmissions in patients hospitalized for heart failure: comparison of machine learning and other statistical approaches. *JAMA Cardiol* 2017;2(02):204–209
- Mortazavi BJ, Downing NS, Bucholz EM, et al. Analysis of machine learning techniques for heart failure readmissions. *Circ Cardiovasc Qual Outcomes* 2016;9(06):629–640
- Billings J, Blunt I, Steventon A, Georgiou T, Lewis G, Bardsley M. Development of a predictive model to identify inpatients at risk of re-admission within 30 days of discharge (PARR-30). *BMJ Open* 2012;2(04):e001667
- 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(06):e011060
- 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
- Cholleti S, Post A, Gao J, et al. Leveraging derived data elements in data analytic models for understanding and predicting hospital readmissions. *AMIA Annu Symp Proc* 2012;2012:103–111
- Dagliati A, Tibollo V, Sacchi L. Big data as a driver for clinical decision support systems: a learning health systems perspective. *Front Digit Humanit* 2018. Doi: 3389/fdigh.2018.00008
- Romero-Brufau S, Wyatt KD, Boyum P, Mickelson M, Moore M, Cognetta-Rieke C. Implementation of artificial intelligence-based clinical decision support to reduce hospital readmissions at a regional hospital. *Appl Clin Inform* 2020;11(04):570–577
- Kansagara D, Englander H, Salanitro A, et al. Risk prediction models for hospital readmission: a systematic review. *JAMA* 2011;306(15):1688–1698
- King's Fund Predictive Risk Project [Internet]. King's Fund. 2005. Accessed May 2020 at: http://www.kingsfund.org.uk/sites/files/kf/field/field_document/predictive-risk-literature-review-june-2005.pdf
- Tsui E, Au SY, Wong CP, Cheung A, Lam P. Development of an automated model to predict the risk of elderly emergency medical

- admissions within a month following an index hospital visit: a Hong Kong experience. *Health Informatics J* 2015;21(01):46–56
- 33 van Walraven C, Dhalla IA, Bell C, et al. Derivation and validation of an index to predict early death or unplanned readmission after discharge from hospital to the community. *CMAJ* 2010;182(06):551–557
 - 34 Canadian Institute for Health Information Early identification of people at-risk of hospitalization [internet]. Ontario: Queen's Printer for Ontario. 2013. Accessed May 2020 at: <http://www.hqontario.ca/Portals/0/Documents/qi/harp-a-new-tool-for-supporting-providers-and-patients.pdf>
 - 35 Eckert C, Nieves-Robbins N, Spieker E, et al. Development and prospective validation of a machine learning-based risk of readmission model in a large military hospital. *Appl Clin Inform* 2019;10(02):316–325
 - 36 Zhang Y, Zhang Y, Sholle E, et al. Assessing the impact of social determinants of health on predictive models for potentially avoidable 30-day readmission or death. *PLoS One* 2020;15(06):e0235064
 - 37 Bernheim SM, Spertus JA, Reid KJ, et al. Socioeconomic disparities in outcomes after acute myocardial infarction. *Am Heart J* 2007;153(02):313–319
 - 38 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(02):269–282
 - 39 Mahmoudi E, Kamdar N, Kim N, Gonzales G, Singh K, Waljee AK. Use of electronic medical records in development and validation of risk prediction models of hospital readmission: systematic review. *BMJ* 2020;369:m958
 - 40 Krumholz HM, Bernheim SM. Considering the role of socioeconomic status in hospital outcomes measures. *Ann Intern Med* 2014;161(11):833–834
 - 41 Hu J, Gonsahn MD, Nerenz DR. Socioeconomic status and readmissions: evidence from an urban teaching hospital. *Health Aff (Millwood)* 2014;33(05):778–785
 - 42 Filc D, Davidovich N, Novack L, Balicer RD. Is socioeconomic status associated with utilization of health care services in a single-payer universal health care system? *Int J Equity Health* 2014;13(01):115
 - 43 Nagasako EM, Reidhead M, Waterman B, Dunagan WC. Adding socioeconomic data to hospital readmissions calculations may produce more useful results. *Health Aff (Millwood)* 2014;33(05):786–791
 - 44 Feldman SS, Davlyatov G, Hall AG. Toward understanding the value of missing social determinants of health data in care transition planning. *Appl Clin Inform* 2020;11(04):556–563
 - 45 Tan SY, Low LL, Yang Y, Lee KH. Applicability of a previously validated readmission predictive index in medical patients in Singapore: a retrospective study. *BMC Health Serv Res* 2013;13(01):366
 - 46 Singh D, Ham Clmproving care for people with long-term conditions: a review of UK and international frameworks. The King's Fund Information and Library Service. 2015. Accessed May 2020 at: http://www.improvingchroniccare.org/downloads/review_of_international_frameworks_chris_hamm.pdf
 - 47 Demir EA decision support tool for predicting patients at risk of readmission: a comparison of classification trees, logistic regression, generalized additive models, and multivariate adaptive regression splines [internet]. University of Hertfordshire Business School. 2012. Accessed May 2020 at: <http://uhra.herts.ac.uk/bitstream/handle/2299/8985/s141.pdf?sequence=1>
 - 48 Ministry of Health Singapore. Technical Manual for Indicators in the Public Hospital Performance Report. 2018
 - 49 Moh.gov.sg. News highlights: Reorganisation of Healthcare System into three Integrated Clusters to better meet future healthcare needs. 2017. Accessed May 2020 at: <https://www.moh.gov.sg/news-highlights/details/reorganisation-of-healthcare-system-into-three-integrated-clusters-to-better-meet-future-healthcare-needs>
 - 50 Deyo RA, Cherkin DC, Ciol MA. Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. *J Clin Epidemiol* 1992;45(06):613–619
 - 51 Quan H, Sundararajan V, Halfon P, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. *Med Care* 2005;43(11):1130–1139
 - 52 Low LL, Wah W, Ng MJ, Tan SY, Liu N, Lee KH. Housing as a social determinant of health in Singapore and its association with readmission risk and increased utilization of hospital services. *Front Public Health* 2016;4:109
 - 53 Hdb.gov.sg. Public rental scheme. Accessed May 2020 at: <https://www.hdb.gov.sg/cs/infoweb/residential/renting-a-flat/renting-from-hdb/public-rental-scheme>
 - 54 Moh.gov.sg. Medifund, Ministry of Health [internet]. 2015. Accessed May 2020 at: https://www.moh.gov.sg/content/moh_web/home/costs_and_financing/schemes_subsidies/Medifund.html
 - 55 Moh.gov.sg. Hospital Services, Ministry of Health [Internet]. [cited 20 December 2019]. Accessed May 2020 at: <https://www.moh.gov.sg/home/our-healthcare-system/healthcare-services-and-facilities/hospital-services>
 - 56 Ng Teng Fong General Hospital. Medical Social Services. [cited 20 December 2019]. Accessed May 2020 at: https://www.ntfgh.com.sg/AlliedHealthDetail/Medical_Social_Services
 - 57 Sperandei S. Understanding logistic regression analysis. *Biochem Med (Zagreb)* 2014;24(01):12–18
 - 58 Natekin A, Knoll A. Gradient boosting machines, a tutorial. *Front Neurobot* 2013;7:21
 - 59 Ciampi A. Generalized regression trees. *Comput Stat Data Anal* 1991;12:57–78
 - 60 Kingsford C, Salzberg SL. What are decision trees? *Nat Biotechnol* 2008;26(09):1011–1013
 - 61 Song YY, Lu Y. Decision tree methods: applications for classification and prediction. *Shanghai Jingshen Yixue* 2015;27(02):130–135
 - 62 Breiman L. Random forests. *Mach Learn* 2001;45(01):5–32
 - 63 Antoniadi AM, Galvin M, Heverin M, Hardiman O, Mooney C. Prediction of caregiver burden in amyotrophic lateral sclerosis: a machine learning approach using random forests applied to a cohort study. *BMJ Open* 2020;10(02):e033109
 - 64 Kumar R, Indrayan A. Receiver operating characteristic (ROC) curve for medical researchers. *Indian Pediatr* 2011;48(04):277–287
 - 65 Borra S, Di Ciaccio A. Measuring the prediction error. A comparison of cross-validation, bootstrap and covariance penalty methods. *Comput Stat Data Anal* 2010;54:2976–2989
 - 66 Baillie CA, VanZandbergen C, Tait G, et al. The readmission risk flag: using the electronic health record to automatically identify patients at risk for 30-day readmission. *J Hosp Med* 2013;8(12):689–695
 - 67 Kripalani S, LeFevre F, Phillips CO, Williams MV, Basaviah P, Baker DW. Deficits in communication and information transfer between hospital-based and primary care physicians: implications for patient safety and continuity of care. *JAMA* 2007;297(08):831–841
 - 68 Gleason KM, McDaniel MR, Feinglass J, et al. Results of the Medications at Transitions and Clinical Handoffs (MATCH) study: an analysis of medication reconciliation errors and risk factors at hospital admission. *J Gen Intern Med* 2010;25(05):441–447
 - 69 Dudas V, Bookwalter T, Kerr KM, Pantilat SZ. The impact of follow-up telephone calls to patients after hospitalization. *Am J Med* 2001;111(9B):26S–30S
 - 70 Hernandez AF, Greiner MA, Fonarow GC, et al. Relationship between early physician follow-up and 30-day readmission among Medicare beneficiaries hospitalized for heart failure. *JAMA* 2010;303(17):1716–1722
 - 71 Nuhs.edu.sg. National University Health System. Accessed May 2020 at: <https://www.nuhs.edu.sg/Care-in-the-Community/Getting-Well/Pages/default.aspx>

- 72 Howard M, Bernard C, Tan A, Slaven M, Klein D, Heyland DK. Advance care planning: let's start sooner. *Can Fam Physician* 2015;61(08):663–665
- 73 Kaboli PJ, Go JT, Hockenberry J, et al. Associations between reduced hospital length of stay and 30-day readmission rate and mortality: 14-year experience in 129 Veterans Affairs hospitals. *Ann Intern Med* 2012;157(12):837–845
- 74 Kuhle S, Maguire B, Zhang H, et al. Comparison of logistic regression with machine learning methods for the prediction of fetal growth abnormalities: a retrospective cohort study. *BMC Pregnancy Childbirth* 2018;18(01):333
- 75 Van Belle VM, Van Calster B, Timmerman D, et al. A mathematical model for interpretable clinical decision support with applications in gynecology. *PLoS One* 2012;7(03):e34312
- 76 Ehwerhemuepha L, Rakovski C. A comprehensive assessment of automatic logistic regression model selection methods. Accessed 2019 at: <https://www.researchsquare.com/article/rs-7597/v1>
- 77 Heinze G, Wallisch C, Dunkler D. Variable selection: a review and recommendations for the practicing statistician. *Biom J* 2018;60(03):431–449
- 78 Rothman MJ, Rothman SI, Beals J IV. Development and validation of a continuous measure of patient condition using the Electronic Medical Record. *J Biomed Inform* 2013;46(05):837–848
- 79 Cai X, Perez-Concha O, Coiera E, et al. Real-time prediction of mortality, readmission, and length of stay using electronic health record data. *J Am Med Inform Assoc* 2016;23(03):553–561
- 80 Hasan O, Meltzer DO, Shaykevich SA, et al. Hospital readmission in general medicine patients: a prediction model. *J Gen Intern Med* 2010;25(03):211–219
- 81 Cui Y, Metge C, Ye X, Moffatt M, Oppenheimer L, Forget EL. Development and validation of a predictive model for all-cause hospital readmissions in Winnipeg, Canada. *J Health Serv Res Policy* 2015;20(02):83–91
- 82 López-Aguilà S, Contel JC, Farré J, Campuzano JL, Rajmil L. Predictive model for emergency hospital admission and 6-month readmission. *Am J Manag Care* 2011;17(09):e348–e357
- 83 Escobar GJ, Ragins A, Scheirer P, Liu V, Robles J, Kipnis P. Nonelective rehospitalizations and postdischarge mortality: predictive models suitable for use in real time. *Med Care* 2015;53(11):916–923
- 84 Fu L, Li FM, Zhou J. Event prediction in healthcare analytics: beyond prediction accuracy. In: Cao HP, Li JY, Wang RL, eds. *Trends and Applications in Knowledge Discovery and Data Mining: PAKDD 2016 Workshops, BDM, MLSDA, PACC, WDMBF, Auckland, New Zealand, April 19, 2016, Revised Selected Papers*. Cham, Switzerland: Springer; 2016
- 85 Gottesman O, Kuivaniemi H, Tromp G, et al; eMERGE Network. The Electronic Medical Records and Genomics (eMERGE) Network: past, present, and future. *Genet Med* 2013;15(10):761–771