What Are They Worth? Six 30-Day Readmission Risk Scores for Medical Inpatients Externally Validated in a Swiss Cohort

Tristan Struja, MD¹ , Ciril Baechli, MD¹, Daniel Koch, MSc¹, Sebastian Haubitz, MD¹, Andreas Eckart, MD¹, Alexander Kutz, MD, MSc¹, Martha Kaeslin, PhD¹, Beat Mueller, MD^{1,2}, and Philipp Schuetz, MD, MPH^{1,2}

¹Kantonsspital Aarau, Medical University Clinic, Aarau, Switzerland; ²Medical Faculty of the University of Basel, Basel, Switzerland.

BACKGROUND: Several clinical risk scores for unplanned 30-day readmission have been published, but there is a lack of external validation and head-to-head comparison.

OBJECTIVE: Retrospective replication of six clinical risk scores (LACE, HOSPITAL, SEMI, RRS, PARA, Tsui et al.).

DESIGN: Models were fitted with the original intercept and beta coefficients as reported. Otherwise, a logistic model was refitted (SEMI and Tsui et al.). We performed subgroup analyses on main admission specialty. This report adheres to the TRIPOD statement for reporting of prediction models.

PARTICIPANTS: We used our prospective cohort of 15,639 medical patients from a Swiss tertiary care institution from 2016 through 2018.

MAIN MEASURES: Thirty-day readmission rate and area under the curve (AUC < 0.50 worse than chance, > 0.70 acceptable, > 0.80 excellent)

CONCLUSIONS: Among several readmission risk scores, HOSPITAL, PARA, and the score from Tsui et al. showed the best predictive abilities and have high potential to improve patient care. Interventional research is now needed to understand the effects of these scores when used in clinical routine.

KEY RESULTS: Among the six risk scores externally validated, calibration of the models was overall poor with overprediction of events, except for the HOSPITAL and the PARA scores. Discriminative abilities (AUC) were as follows: LACE 0.53 (95% CI 0.50–0.56), HOSPITAL 0.73 (95% CI 0.72–0.74), SEMI 0.47 (95% CI 0.46–0.49), RRS 0.64 (95% CI 0.62–0.66), PARA 0.72 (95% CI 0.72–0.74), and the score from Tsui et al. 0.73 (95% CI 0.72–0.75). Performance in subgroups did not differ from the overall performance, except for oncology patients in the PARA score (0.57, 95% CI 0.54–0.60), and nephrology patients in the SEMI index (0.25, 95% CI 0.18–0.31), respectively.

J Gen Intern Med 35(7):2017–24

DOI: 10.1007/s11606-020-05638-z

© Society of General Internal Medicine 2020

INTRODUCTION

An inclusive in-hospital patient management with appropriate resource allocation is the current challenge of healthcare systems.^{1, 2} Improved diagnostic and therapeutic measures have increased life expectancy, but often do not achieve long-lasting cure.³ These patients are at high risk for an increased length of stay (LOS) and planning of post-acute care is neglected during hospitalization, increasing risk for readmission.⁴ Thus, early and efficient post-acute care planning—integrated into an accurate risk stratification⁵—has the potential to optimize acute care resources by reducing LOS while preventing functional disability by preventing frequent patient transfers.^{6–8}

Several algorithms for prediction of 30-day readmission risk have been proposed.⁹ Their purpose is to identify early patients at risk for unplanned hospital readmission to potentially prevent these events. Yet, the majority of these scores have not been externally validated.¹⁰ Before the publication of the “Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis” (TRIPOD) statement, there was no consensus how to report prediction modeling studies.¹⁰ Although risk scores perform best in the environment where they were developed, researchers and clinicians might be interested in the merits of a score in a different setting. This can guide adaptation or recalibration of an existing score to one’s own requisites or highlight the need for developing a new score.

In this study, we validated six risk scores previously published, and compared their prognostic performance head-to-head in a large, contemporary, and independent cohort of medical inpatients in a Swiss tertiary care center. This report adheres to the TRIPOD statement for reporting of prediction models.

METHODS

Information on the Studies Validated

Tsui et al. reported a risk score based on insurance data from Hong Kong on 2.3 million admissions.¹¹

Gildersleeve et al. used their own rural-based hospital database (16,889 admissions).¹² Zapatero et al. generated their

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s11606-020-05638-z>) contains supplementary material, which is available to authorized users.

Received September 11, 2019

Revised November 24, 2019

Accepted January 3, 2020

Published online January 21, 2020

SEMI Index based on a nationwide Spanish insurance registry encompassing almost 1 million internal medicine admissions.¹³ The PARA was established out of a Swiss tertiary care hospital cohort with 11,074 stays.¹⁴ The LACE score was derived on data derived from 11 hospitals in Ontario (4812 admissions) and externally validated in an insurance cohort with 1 million admissions.¹⁵ Donzé et al. based their HOSPITAL score on 10,731 admissions stemming from insurance data of 3 hospitals located in the Boston (USA) area.¹⁶ For further information, please consult Table 2 and the corresponding section in the [supplementary material](#).

Source of Data and Participants

Based on a systematic review on readmission risk scores from 2016,¹⁷ we identified studies on medical inpatients with a predefined overall area under the curve (AUC) of at least 0.7. We picked this cutoff as models with this discriminative ability are deemed useful.⁹ Additionally, we performed a literature search on PubMed for studies published after 2015 including the terms “readmission” and “risk factors,” while excluding studies with the terms “surgery” and “operation.”

Our data is based on a retrospective analysis of our prospective cohort studies on medical inpatients (i.e., TRIAGE and InHospiTOOL) of which the protocols have been published.^{18, 19} For this analysis, we used data from our suburban Swiss center only. From January 1, 2016, through February 28, 2019, we collected data on every patient aged 16 years or older being admitted to our medical wards. Exclusion criteria are depicted in Figure 1 and were adopted from the original studies. As part of our quality control, patients were contacted by telephone 30 days after admission to ask for quality measures (focused on their functional performance) and to identify readmissions. If patients or relatives could not be reached, we reached out to the primary care physician, post-acute care institutions, or the registry office at the place of residence.

As a retrospective, quality control study without any effects on individual patient outcomes, the ethical review board (EKNZ) waived the need for individual patient informed consent. The study has been conducted according to the principles of the Declaration of Helsinki.

Outcomes

The primary outcome was the first unplanned readmission within 30 days after the index admission. Unplanned readmission was defined as every readmission occurring within 30 days.

To provide a better overview of the predictive abilities on distinct patient populations, we performed subgroup analyses based on our main admission specialties.

Predictors

Our database comprises characteristics such as health insurance, demographics, and clinical features (e.g., diagnoses,

medications, and laboratory). Comorbidities were evaluated using the Charlson Comorbidity Index²⁰ based on information at hospital discharge. Hospitalization characteristics include LOS, number of admissions during the past year before index hospitalization, and visits to our emergency department (ED). Swiss procedure codes (CHOP codes) were used as indicators in terms of interventions during a stay. This system was originally based on the 2008 ICD-9-CM codes, but was later adapted to the Swiss healthcare system and is updated annually.²¹

This study is based on our TRIAGE/InHospiTOOL (not to be confused with the HOSPITAL score) study cohort.^{18, 19} Therein, we decided to explicitly store laboratory results from admission instead of discharge. Hence, we had to access laboratory results (i.e., sodium and hemoglobin) at discharge through our electronic health record system (EHR). We were not able to get discharge values before 2015 because the corresponding interface has not been instituted. As a proxy, we used admission values by carrying them forward.²²

In case of repeated hospitalizations more than 30 days apart, each admission was eligible to become a new index case. Data from the EHR are anonymized and entered into our data warehouse. In contrast to the SEMI index¹³ and the HOSPITAL score,¹⁶ we used ICD-10 codes instead of ICD-9 codes. Further, it was unclear how the diagnoses were determined by Tsui et al.¹¹; thus, we used ICD-10 diagnoses as appropriate.

Statistics, Sample Size, and Missing Data

We collected all information on variables out of the derivation studies. If information on β -coefficients and intercept were provided,¹⁶ we used them for replication of the score (probability $P = \frac{1}{1 + e^{\text{Intercept} - b \times \text{predictors}}}$). If coefficients or intercept were not reported, we emailed the corresponding authors.^{12, 14, 15} Otherwise, simplified scores were applied, and points were summarized as suggested by the authors.^{11, 13} We did not perform any model updating.

We assessed calibration graphically by plotting the observed risks on the y-axis against the predicted risks on the x-axis augmented by a lowess line. For comparison of discriminatory power, we calculated the AUC/C-statistic by a non-parametric model. All available admissions from our cohort were used without a priori sample size calculation.

As some of the data were gathered in clinical routine outside our TRIAGE study, we had missing data on the following predictors: number of medications, hemoglobin and sodium at discharge. As our database was still quite comprehensive, we primarily performed a complete case analysis (see Fig. 1). To gauge impact of missing data, we also performed an exploratory analysis based on multiple imputations (see [supplementary material](#)). Data are presented as medians and IQR [interquartile range], or as counts and frequencies as appropriate.

All significance tests were based on a two-sided α -error of 0.05. Statistical analysis was conducted using Stata software version 15.1 (Stata Corp., College Station, TX, USA).

Figure 1 Flowchart with exclusion of patients.

RESULTS

Out of 11,575 patients with 15,639 admissions, we recorded 1149 readmissions (7.3%) within 30 days of admission (see Fig. 1). The baseline characteristics among the patients included in the analysis compared with the overall cohort before application of exclusion criteria are presented in Table 1. Before application of exclusion criteria, the 30-day death rate was higher. On the other hand, LOS was longer.

Data on individuals with missing values per predictor and outcome, and comparison of characteristics of individuals with any missing value and those with completely observed data are provided in supplementary table 1.

Out of the systematic review by Zhou et al.,⁹ we retrieved 13 studies reporting risk scores with an AUC > 0.7. Out of these, our data sufficed to replicate the following six risk scores: LACE Index, HOSPITAL Score, SEMI Index, RRS Score, PARA, and a score from Hong Kong by Tsui et al.^{11–16} The findings and cohort characteristics of these studies are summarized in Table 2. Scores for which we did not have sufficient data were mainly data-driven machine learning approaches based on North American hospital data,^{23–29} except one study from the UK.²⁹

Differences Between Validation and Development Cohorts

Readmission rate in our cohort was 7.3%, whereas it ranged from 7.0 to 14.8% in the derivation cohorts. Characteristics of the cohorts and readmission rates are summarized in Table 2. The cohorts of the LACE, HOSPITAL, and SEMI scores are based on insurance data. Our cohort, PARA, and RRS scores are based on hospital data, whereas the score from Hong Kong is based on a mixture of the two.

Calibration

Calibration was assessed graphically by tenths of the predicted risk (in the case of validation based on coefficients) or in the case of reported scores by the intervals of the score (see supplementary figure 1). Additionally, we provide a tabular overview (see supplementary table 2). Overall, calibration was rather poor, except for the HOSPITAL score and SEMI Index. Most models tended to overpredict the risk of readmission.

Discrimination

AUCs for the risk scores were as follows: LACE Index AUC 0.53 (95% CI 0.50–0.56), HOSPITAL Score AUC 0.73 (95% CI 0.72–0.74), SEMI Index AUC 0.47 (95% CI 0.46–0.49), RRS Score AUC 0.64 (95% CI 0.62–0.66), PARA AUC 0.73 (95% CI 0.72–0.74), and a score from Hong Kong by Tsui et al. AUC 0.73 (95% CI 0.72–0.75) (see Table 3). Subgroup analysis is presented in a forest plot (see Fig. 2). The plot is weighted on the number of patients for each subgroup. Performance in subgroups did not differ from the overall performance, except for oncology patients in the PARA score and nephrology patients in the SEMI index, respectively.

DISCUSSION

To our knowledge, this study is the first comprehensive comparison of risk scores head-to-head in one patient population. As EHR systems have become ubiquitous, it is tempting to use risk scores as clinical decision support tools. By the reason that such tools can have a strong influence on clinicians' decisions, they have been thoroughly assessed for their performance.

Table 1 Baseline Characteristics of Patients Before and After Application of Exclusion Criteria

		Total before exclusions (N = 33,027)	Total after exclusions (N = 15,639)	No readmission (N = 14,490)	Readmission (N = 1149)
Age, years (median, IQR)		65 (50–77)	69 (57–79)	69 (57–79)	70 (58–79)
Gender (n, %)	Female	14,948 (45.3%)	6815 (43.6%)	6322 (43.6%)	493 (42.9%)
Region of origin (n, %)	Europe	5489 (16.6%)	2215 (14.2%)	2058 (14.2%)	157 (13.7%)
	Outside Europe	1651 (5.0%)	542 (3.5%)	512 (3.5%)	30 (2.6%)
	Switzerland	25,887 (78.4%)	12,882 (82.4%)	11,920 (82.3%)	962 (83.7%)
Mother tongue (n, %)	Other	6973 (21.1%)	4205 (26.9%)	3855 (26.6%)	350 (30.5%)
	German	25,887 (78.4%)	11,434 (73.1%)	10,635 (73.4%)	799 (69.5%)
Civil status (n, %)	Divorced, widowed, or unmarried	26,201 (79.3%)	11,020 (70.5%)	10,087 (69.6%)	933 (81.2%)
	Married	6826 (20.7%)	4619 (29.5%)	4403 (30.4%)	216 (18.8%)
Insurance (n, %)	Public	28,986 (87.8%)	13,251 (84.7%)	12,294 (84.8%)	957 (83.3%)
	Private	4041 (12.2%)	2388 (15.3%)	2196 (15.2%)	192 (16.7%)
Admission type (n, %)	Emergency	28,133 (85.2%)	12,439 (79.5%)	11,551 (79.7%)	888 (77.3%)
	Elective	4630 (14.0%)	3043 (19.5%)	2792 (19.3%)	251 (21.8%)
Residency before admission (n, %)	Missing	264 (0.8%)	157 (1.0%)	147 (1.0%)	10 (0.9%)
	Home independent	14,850 (45.0%)	9369 (59.9%)	8814 (60.8%)	555 (48.3%)
	Home with assisting care	712 (2.2%)	438 (2.8%)	406 (2.8%)	32 (2.8%)
	Senior residence	2795 (8.5%)	1376 (8.8%)	1316 (9.1%)	60 (5.2%)
	Other	921 (2.8%)	421 (2.7%)	382 (2.6%)	39 (3.4%)
	Missing	13,749 (41.6%)	4035 (25.8%)	3572 (24.7%)	463 (40.3%)
Main diagnosis (n, %)	General internal medicine	15,180 (46.0%)	4655 (29.8%)	4390 (30.3%)	265 (23.1%)
	Gastroenterology	1562 (4.7%)	1060 (6.8%)	993 (6.9%)	67 (5.8%)
	Hematology/oncology	3370 (10.2%)	2277 (14.6%)	1917 (13.2%)	360 (31.3%)
	Cardiology	5304 (16.1%)	2890 (18.5%)	2689 (18.6%)	201 (17.5%)
	Nephrology	671 (2.0%)	495 (3.2%)	446 (3.1%)	49 (4.3%)
	Neurology	5755 (17.4%)	3488 (22.3%)	3333 (23.0%)	155 (13.5%)
	Pneumology	1185 (3.6%)	774 (4.9%)	722 (5.0%)	52 (4.5%)
	Anemia	2326 (7.0%)	1687 (10.8%)	1546 (10.7%)	141 (12.3%)
	Diabetes	2425 (7.3%)	1696 (10.8%)	1608 (11.1%)	88 (7.7%)
	Gastrointestinal disease	2886 (8.7%)	2055 (13.1%)	1915 (13.2%)	140 (12.2%)
Proportion with (n, %)	Hypertension	7814 (23.7%)	5301 (33.9%)	5058 (34.9%)	243 (21.1%)
	Malignant disease	3253 (9.8%)	2297 (14.7%)	2012 (13.9%)	285 (24.8%)
	Cardiac disease	10,727 (32.5%)	6948 (44.4%)	6625 (45.7%)	323 (28.1%)
	Neurologic disease	4085 (12.4%)	2454 (15.7%)	2352 (16.2%)	102 (8.9%)
	Chronic renal failure	3618 (11.0%)	2509 (16.0%)	2371 (16.4%)	138 (12.0%)
	No. of discharge medications (median, IQR)	7 (4–10)	8 (5–11)	8 (5–11)	8 (6–12)
	No. of admissions during last year (n, %)				
	0	27,195 (82.3%)	10,796 (69.0%)	10,596 (73.1%)	200 (17.4%)
	1–5	5629 (17.0%)	4663 (29.8%)	3753 (25.9%)	910 (79.2%)
	> 5	203 (0.6%)	180 (1.2%)	141 (1.0%)	39 (3.4%)
No. of ED visits during last year (n, %)	0	29,882 (90.5%)	15,400 (98.5%)	14,255 (98.4%)	1145 (99.7%)
	1–5	3046 (9.2%)	239 (1.5%)	235 (1.6%)	4 (0.3%)
	> 5	99 (0.3%)	0	0	0
Length of stay, days (median, IQR)		2 (0–6)	5 (3–8)	5 (3–8)	5 (3–9)
Charlson Comorbidity Index (median, IQR)		3 (1–4)	4 (2–5)	4 (2–5)	4 (2–6)
30-Day death rate (n, %)		1496 (4.5%)	249 (1.6%)	245 (1.7%)	4 (0.3%)

Authors of all risk scores performed validation by splitting their original cohort and reported an AUC > 0.7 in their validation cohort. We found that the performance of the scores applied to our Swiss, tertiary care, single-center, non-surgical dataset varied greatly with AUCs ranging from 0.47 to 0.73 and with rather poor calibration. Subgroup analysis showed that scores are consistent across various medical subspecialties. The two outliers (i.e., oncology in the PARA score, and nephrology in the SEMI index) might be explained by the low number of cases in these subgroups.

Any prediction score is prone to overestimate results in its original dataset by non-parsimonious addition of variables

(i.e., overfitting).^{10, 30} Thus, it is of utmost importance to test its predictive abilities in an independent cohort. The SEMI, PARA, RRS, and the score from Tsui et al. have been neither retrospectively validated nor prospectively assessed so far.

The HOSPITAL score has been prospectively validated by Donzé himself in Switzerland.³¹ We abstained from using the SQLape® algorithm³² to separate unavoidable and avoidable readmissions. Currently, there is no consensus how to classify readmissions, which carries the risk of introducing bias.¹⁹

His research group prospectively included 346 patients older than 50 years admitted to the general internal medicine wards of the cantonal hospital of Fribourg. A hospital that has

Table 2 Characteristics and Predictors of the Validated Scores (in Cases of Derivation and Internal Validations, Only Data for Validation Shown)

Model	Country	Outcome	Readmission rate	Variables used in the corresponding model						
				Demographics	Laboratory	Comorbidities	Admission type	No. of admissions	ED usage	Various
LACE	Canada	30-days and death	8.0%	LOS		Charlson Index	ED vs. elective	In 1 year	Visits in 6 months	
HOSPITAL	USA	30 days	8.5%	LOS	Sodium < 135 mM, Hb < 12 g/dl	Discharge from oncology ward	ED vs. elective	In 1 year		Procedures ICD-9
PARA	Switzerland	30 days	7.0%	LOS > 11 days, ≥ 15 medications	Sodium < 135 mM	Charlson Index > 1		In 1 year		Cancer w or w/o metastasis and w or w/o SSRI
RRRS	USA	30 days	14.8%	LOS, sex, marital and living status		Charlson Index	ED vs. elective	In 1 year	Visits in 12 months	Medications at discharge and 2 days prior
SEMI INDEX	Spain	30 days	12.4%	LOS		“Charlson proxy”				Diagnosis ICD-9
Tsui et al.	Hong Kong	28 days	7.6%	LOS, insurance		“Charlson proxy”	ED vs. elective		Visits in 12 months	

Charlson proxy:

SEMI Index

age, sex, liver disease, COPD, malnutrition, chronic renal failure, congestive heart failure, anemia, diabetes, dementia, atrial fibrillation

Tsui et al.

age, sex, cancer, COPD, malnutrition, dialysis, congestive heart failure, chronic renal failure, diabetes, epilepsy, dementia, central nervous system disorders, Parkinson disease, ischemic heart disease, cerebrovascular disease, bronchiectasis, cirrhosis, gastrointestinal hemorrhage, chronic skin ulcer

Abbreviations: ED, emergency department; Hb, hemoglobin; LOS, length of stay; SSRI, selective serotonin reuptake inhibitor; w, with; w/o, without

roughly two-thirds the size of our own. They found an AUC of 0.70 (95% CI 0.62–0.79) with an 11.6% 30-day composite outcome (unplanned readmission and death rate). This AUC seems comparable with our result of 0.79 (95% CI 0.78–0.80), albeit we did not include death rate into our outcome which might explain the difference. Furthermore, we used CHOP codes instead of ICD-9-CM codes as predictor. Nevertheless, we do not think this has relevantly influenced our results. The same group published a simplification of the original HOSPITAL score which does not contain procedures as predictor,³³ and its AUC differs only by 2% from the original score. Additionally, this kind of adaptations would occur in every real-world application of a score, and as such rather strengthens our findings.

The LACE score has been analyzed in various retrospective and prospective studies. Many prospective studies used the LACE score merely for risk stratification. Consecutively, specific interventions were then taken to reduce the risk of readmission.^{34–37} Only one study specifically looked at the predictive ability of the LACE tool.³⁸ In 378 patients, Yazdan et al. dichotomized LACE points at different cutoffs and found AUCs ranging from 0.52 to 0.58. Another recent study retrospectively compared the LACE score against a machine learning approach.³⁹ Authors analyzed a cohort of 10,732 patients and reported AUCs of 0.66 for the LACE score, 0.63 for the HOSPITAL score, 0.64 for the Maxim/RightCare tool, and 0.81 for their own B score.

Besides risk scores, many different strategies (discharge checklists, follow-up phone calls, home visits by nurses, etc.) have been assessed and all have the potential to reduce readmissions.¹⁹

Strengths and Limitations

Compared with the derivation cohorts, our general readmission rate was at the lower end. This could be due to the fact that we assessed readmission 30 days after index admission. Taking our median LOS of 5 days into account, this shortens our observed time period by that amount. However, we do not think this has biased our results. First, our cohort is large, which should level out possible fluctuations over time. Second, our readmission rate is the same as the 7.0% observed in the other Swiss study by Uhlmann and colleagues,¹⁴ and comparable with the readmission rate of 5.45% reported by the five Swiss university hospitals in 2014.⁴⁰ The setting was also similar, as both hospitals provide tertiary care for a greater area, while providing basic care for patients living nearby.

Our 1149 readmissions occurred in 290 patients only, compared with 11,285 patients without a readmission. Because the other authors did not report these figures, we cannot directly compare our studies in this regard. But we do not have reason to believe that the other cohorts were dissimilar, and most readmissions are generated by high utilizers. An extensive Australian study in over 20,000 patients found that 80% of patients were being admitted only once.⁴¹ Hence, application

Table 3 Discriminative Ability of Each Score in the Original Publication and in Our Cohort

Model	Original N	Original AUC	Lower 95% CI	Upper 95% CI	Validation N	Validation AUC without MI	Lower 95% CI	Upper 95% CI	Validation AUC with MI	Lower 95% CI	Upper 95% CI
LACE	2419	0.69	0.65	0.73	15,639	0.53	0.50	0.56	0.52	0.51	0.54
HOSPITAL	3071	0.71	N/A	N/A	15,631	0.73	0.72	0.74	0.73	0.72	0.74
PARA	6729	0.70	N/A	N/A	15,639	0.73	0.72	0.74	0.74	0.73	0.75
RRS	8189	0.70	0.69	0.71	15,639	0.64	0.62	0.66	0.64	0.63	0.67
SEMI	510,588	0.88	N/A	N/A	15,639	0.47	0.46	0.49	0.46	0.44	0.47
Index*											
Tsui et al.*	1,176,482	0.82	N/A	N/A	15,639	0.73	0.72	0.75	0.74	0.72	0.75

*Complete validation not possible as no intercept was reported, and authors did not reply upon request

Abbreviations: N, number; N/A, not available in the original publication; AUC, area under the curve; CI, confidence interval; MI, multiple imputation

of a risk score makes even so much more sense, as it allows hospital staff to concentrate their efforts on patients being at high risk for readmission.

Only certain parts of our current work were based on our prospective TRIAGE/InHospITool study series and provide Swiss data only. As we performed complete case analysis only, there is the potential bias of data not missing at random. Thus, we have up to 26% missing information on certain

predictors. In comparison (see supplementary table 1A), the amount of missing information on relevant predictors did not exceed 7% between patients with readmission and those without. Additionally, complete cases had slightly less medications at discharge and had lower hemoglobin values than incomplete cases (see supplementary table 1B). Otherwise, there were minor differences in the baseline characteristics before and after application of the exclusion criteria. Especially,

Figure 2 Forest plot depicting overall discriminative performance of each score and for distinct subgroups of patients.

median LOS rose from 2 to 5 days, death rate decreased from 4.9 to 1.6%, and the number of hospital admissions and ED visits changed. Although this could have introduced selection bias, it should not decrease comparability of our results as we applied the same inclusion and exclusion criteria as the original studies.

The strengths of our work are for one the large population in which we externally validated six scores. This allowed us to calculate precise predictions of the AUC and enough power to provide sensible subgroup analyses. Second, the utility of readmission risk scores lies in their daily use. As such, clinical data is needed to calculate a score. On the other hand, LACE, Tsui et al., and SEMI scores were mostly based on insurance data. Data from insurances can easily contain information that has not been known to a clinician when deciding the discharge. By using information available to physicians at discharge, we strengthen the validity of the models assessed. Furthermore, we compared scores developed in four different healthcare systems (i.e., Spain, North America, Hong Kong, Switzerland). Through such comparisons, confidence on the reliability of a score is strengthened, especially if performance across different healthcare settings remains similar. Also, a substantial part of our data was gathered through our prospective TRIAGE cohort, which strengthens confidence and integrity of our exposures and outcomes.

There is an abundance of retrospective trials on readmission risk prediction,⁹ but there is scant literature on prospective, multi-centric trials. A systematic review and meta-analysis looked on various forms of readmission prevention interventions.⁴² Authors identified 47 trials of which 22 were conducted in less than 200 patients and only 15 trials were multi-centric. Also, there was no trial that assessed a readmission risk prediction tool.

CONCLUSION

The comparison of six readmission risk scores revealed that there exists a substantial difference in performance in an external validation. HOSPITAL, PARA, and the score from Tsui et al. showed the best predictive abilities and have high potential to improve patient care. Our work highlights the importance of rigorous assessment of risk scores before they are being implemented into daily routine to improve patient care. Interventional research is now needed to understand the effects of these scores when used in clinical routine.

Acknowledgments:

Funding Information: The ongoing TRIAGE/InHospITool Project Series is supported by the Swiss National Research Foundation (No. 407440_167376), the Research Council of the Kantonsspital Aarau (1410.000.086), and the "Wissenschaft & Weiterbildung" (W&W) Fonds (140.000.495) of the Kantonsspital Aarau, the "Hugo und Elsa Isler Fonds."

This study was supported in part by the Swiss National Science Foundation (SNSF Professorship, PPO0P3_150531/1), the Research Council of the Kantonsspital Aarau (1410.000.044) on behalf of PS, and by the

"Elsa and Hugo Isler-Fonds" of the Argovian Department of Health and Social Affairs on behalf of TS.

Corresponding Author: Tristan Struja, MD; Kantonsspital Aarau, Medical University Clinic, Aarau, Switzerland (e-mail: tristan.struja@gmail.com).

Compliance with Ethical Standards:

As a retrospective, quality control study without any effects on individual patient outcomes, the ethical review board (EKNZ) waived the need for individual patient informed consent. The study has been conducted according to the principles of the Declaration of Helsinki.

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

REFERENCES

1. Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. *Lancet*. 2012;380(9836):37–43. [https://doi.org/10.1016/S0140-6736\(12\)60240-2](https://doi.org/10.1016/S0140-6736(12)60240-2).
2. Mercer SW, Guthrie B, Furler J, Watt GC, Hart JT. Multimorbidity and the inverse care law in primary care. *BMJ*. 2012;344:e4152. <https://doi.org/10.1136/bmj.e4152>.
3. June CH, Sadelain M. Chimeric Antigen Receptor Therapy. *N Engl J Med*. 2018;379(1):64–73. <https://doi.org/10.1056/NEJMr1706169>.
4. Mabire C, Bula C, Morin D, Goulet C. Nursing discharge planning for older medical inpatients in Switzerland: A cross-sectional study. *Geriatr Nurs*. 2015;36(6):451–7. <https://doi.org/10.1016/j.gerinurse.2015.07.002>.
5. Louis Simonet M, Kossovsky MP, Chopard P, Sigaud P, Perneger TV, Gaspoz JM. A predictive score to identify hospitalized patients' risk of discharge to a post-acute care facility. *BMC health services research*. 2008;8:154. <https://doi.org/10.1186/1472-6963-8-154>.
6. Fox MT, Persaud M, Maimets I, Brooks D, O'Brien K, Tregunno D. Effectiveness of early discharge planning in acutely ill or injured hospitalized older adults: a systematic review and meta-analysis. *BMC Geriatrics*. 2013;13:70. <https://doi.org/10.1186/1471-2318-13-70>.
7. McMartin K. Discharge planning in chronic conditions: an evidence-based analysis. *Ontario Health Technol Assess Ser*. 2013;13(4):1–72.
8. Fox M. Nurse-led early discharge planning for chronic disease reduces hospital readmission rates and all-cause mortality. *Evid-Based Nurs*. 2016;19(2):62. <https://doi.org/10.1136/eb-2015-102197>.
9. 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. <https://doi.org/10.1136/bmjopen-2016-011060>.
10. Moons KG, Altman DG, Reitsma JB, Ioannidis JP, Macaskill P, Steyerberg EW, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med*. 2015;162(1):W1–73. <https://doi.org/10.7326/M14-0698>.
11. 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(1):46–56. <https://doi.org/10.1177/1460458213501095>.
12. Gildersleeve R, Cooper P. Development of an automated, real time surveillance tool for predicting readmissions at a community hospital. *Appl Clin Inform*. 2013;4(2):153–69. <https://doi.org/10.4338/ACI-2012-12-RA-0058>.
13. Zapatero A, Barba R, Marco J, Hinojosa J, Plaza S, Losa JE, et al. Predictive model of readmission to internal medicine wards. *Eur J Intern Med*. 2012;23(5):451–6. <https://doi.org/10.1016/j.ejim.2012.01.005>.
14. Uhlmann M, Lecureux E, Griesser AC, Duong HD, Lamy O. Prediction of potentially avoidable readmission risk in a division of general internal medicine. *Swiss Med Wkly*. 2017;147:w14470. <https://doi.org/10.4414/smw.2017.14470>.
15. van Walraven C, Dhalla IA, Bell C, Etchells E, Stiell IG, Zarnke K, et al. Derivation and validation of an index to predict early death or unplanned readmission after discharge from hospital to the community.

- CMAJ: Can Med Assoc J. 2010;182(6):551–7. <https://doi.org/10.1503/cmaj.091117>.
16. **Donze 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(8):632–8. <https://doi.org/10.1001/jamainternmed.2013.3023>.
 17. **Zhao JG, Zeng XT, Wang J, Liu L.** Association Between Calcium or Vitamin D Supplementation and Fracture Incidence in Community-Dwelling Older Adults: A Systematic Review and Meta-analysis. *JAMA.* 2017;318(24):2466–82. <https://doi.org/10.1001/jama.2017.19344>.
 18. **Schuetz P, Hausfater P, Amin D, Haubitz S, Fassler L, Grolimund E, et al.** Optimizing triage and hospitalization in adult general medical emergency patients: the triage project. *BMC Emerg Med.* 2013;13:12. <https://doi.org/10.1186/1471-227X-13-12>.
 19. **Kutz A, Koch D, Conca A, Baechli C, Haubitz S, Regez K, et al.** Integrative hospital treatment in older patients to benchmark and improve outcome and length of stay - the In-HospITool study. *BMC Health Serv Res.* 2019;19(1):237. <https://doi.org/10.1186/s12913-019-4045-x>.
 20. **Charlson ME, Pompei P, Ales KL, MacKenzie CR.** A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation. *J Chron Dis.* 1987;40(5):373–83. [https://doi.org/10.1016/0021-9681\(87\)90171-8](https://doi.org/10.1016/0021-9681(87)90171-8).
 21. (BFS) BfS. Schweizerische Operationsklassifikation (CHOP) Systematisches Verzeichnis – Version 2019. Statistik der Schweiz; 2018.
 22. **Haukoos JS, Newgard CD.** Advanced statistics: missing data in clinical research—part 1: an introduction and conceptual framework. *Acad Emerg Med.* 2007;14(7):662–8. <https://doi.org/10.1197/j.aem.2006.11.037>.
 23. **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):916–23. <https://doi.org/10.1097/MLR.0000000000000435>.
 24. **van Walraven C, Wong J, Forster AJ.** LACE+ index: extension of a validated index to predict early death or urgent readmission after hospital discharge using administrative data. *Open Med.* 2012;6(3):e90–e100.
 25. **He D, Mathews SC, Kallou AN, Hutfless S.** Mining high-dimensional administrative claims data to predict early hospital readmissions. *J Am Med Inform Assoc.* 2014;21(2):272–9. <https://doi.org/10.1136/amiainjnl-2013-002151>.
 26. **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(4). <https://doi.org/10.1136/bmjopen-2012-001667>.
 27. **Shulan M, Gao K, Moore CD.** Predicting 30-day all-cause hospital readmissions. *Health Care Manag Sci.* 2013;16(2):167–75. <https://doi.org/10.1007/s10729-013-9220-8>.
 28. **Yu S, Farooq F, van Esbroeck A, Fung G, Anand V, Krishnapuram B.** Predicting readmission risk with institution-specific prediction models. *Artif Intell Med.* 2015;65(2):89–96. <https://doi.org/10.1016/j.artmed.2015.08.005>.
 29. **Choudhry SA, Li J, Davis D, Erdmann C, Sikka R, Sutariya B.** A public-private partnership develops and externally validates a 30-day hospital readmission risk prediction model. *Online J Public Health Inform.* 2013;5(2):219. <https://doi.org/10.5210/ojphi.v5i2.4726>.
 30. **Rajkomar A, Dean J, Kohane I.** Machine Learning in Medicine. *N Engl J Med.* 2019;380(14):1347–58. <https://doi.org/10.1056/NEJMra1814259>.
 31. **Aubert CE, Folly A, Mancinetti M, Hayoz D, Donze J.** Prospective validation and adaptation of the HOSPITAL score to predict high risk of unplanned readmission of medical patients. *Swiss Med Wkly.* 2016;146:w14335. <https://doi.org/10.4414/smw.2016.14335>.
 32. **Halfon P, Eggli Y, Pretre-Rohrbach I, Meylan D, Marazzi A, Burnand B.** Validation of the potentially avoidable hospital readmission rate as a routine indicator of the quality of hospital care. *Medical Care.* 2006;44(11):972–81. <https://doi.org/10.1097/01.mlr.0000228002.43688.c2>.
 33. **Aubert CE, Schnipper JL, Williams MV, Robinson EJ, Zimlichman E, Vasilevskis EE, et al.** Simplification of the HOSPITAL score for predicting 30-day readmissions. *BMJ Qual Saf.* 2017;26(10):799–805. <https://doi.org/10.1136/bmjqs-2016-006239>.
 34. **Lau D, Padwal RS, Majumdar SR, Pederson JL, Belga S, Kahlon S, et al.** Patient-Reported Discharge Readiness and 30-Day Risk of Readmission or Death: A Prospective Cohort Study. *Am J Med.* 2016;129(1):89–95. <https://doi.org/10.1016/j.amjmed.2015.08.018>.
 35. **Kahlon S, Pederson J, Majumdar SR, Belga S, Lau D, Fradette M, et al.** Association between frailty and 30-day outcomes after discharge from hospital. *CMAJ: Can Med Assoc J.* 2015;187(11):799–804. <https://doi.org/10.1503/cmaj.150100>.
 36. **Teh R, Janus E.** Identifying and targeting patients with predicted 30-day hospital readmissions using the revised LACE index score and early postdischarge intervention. *Int J Evid Based Healthc.* 2018;16(3):174–81. <https://doi.org/10.1097/XEB.0000000000000142>.
 37. **Low LL, Tan SY, Ng MJ, Tay WY, Ng LB, Balasubramaniam K, et al.** Applying the Integrated Practice Unit Concept to a Modified Virtual Ward Model of Care for Patients at Highest Risk of Readmission: A Randomized Controlled Trial. *PLoS One.* 2017;12(1):e0168757. <https://doi.org/10.1371/journal.pone.0168757>.
 38. **Yazdan-Ashoori P, Lee SF, Ibrahim G, Van Spall HG.** Utility of the LACE index at the bedside in predicting 30-day readmission or death in patients hospitalized with heart failure. *Am Heart J.* 2016;179:51–8. <https://doi.org/10.1016/j.ahj.2016.06.007>.
 39. **Morgan DJ, Bame B, Zimand P, Dooley P, Thom KA, Harris AD, et al.** Assessment of Machine Learning vs Standard Prediction Rules for Predicting Hospital Readmissions. *JAMA Netw Open.* 2019;2(3):e190348. <https://doi.org/10.1001/jamanetworkopen.2019.0348>.
 40. **Munzer T.** Models to predict readmission rates - trying to sand the wings of the boomerang. *Swiss Med Wkly.* 2017;147:w14493. <https://doi.org/10.4414/smw.2017.14493>.
 41. **Kirby SE, Dennis SM, Jayasinghe UW, Harris MF.** Patient related factors in frequent readmissions: the influence of condition, access to services and patient choice. *BMC Health Serv Res.* 2010;10:216. <https://doi.org/10.1186/1472-6963-10-216>.
 42. **Leppin AL, Gionfriddo MR, Kessler M, Brito JP, Mair FS, Gallacher K, et al.** Preventing 30-day hospital readmissions: a systematic review and meta-analysis of randomized trials. *JAMA Intern Med.* 2014;174(7):1095–107. <https://doi.org/10.1001/jamainternmed.2014.1608>.

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