BMJ Open

Published models that predict hospital readmission: a critical appraisal

Lisa Grossman Liu [ORCID],[1] James R Rogers,[1] Rollin Reeder,[2,3] Colin G Walsh,[2,3,4] Devan Kansagara,[5] David K Vawdrey,[1,6] Hojjat Salmasian [ORCID] [7,8,9]

For numbered affiliations see end of article.

Correspondence to
Dr Lisa Grossman Liu;
lvg2104@cumc.columbia.edu

ABSTRACT

Introduction The number of readmission risk prediction models available has increased rapidly, and these models are used extensively for health decision-making. Unfortunately, readmission models can be subject to flaws in their development and validation, as well as limitations in their clinical usefulness.

Objective To critically appraise readmission models in the published literature using Delphi-based recommendations for their development and validation.

Methods We used the modified Delphi process to create *Critical Appraisal of Models that Predict Readmission* (CAMPR), which lists expert recommendations focused on development and validation of readmission models. Guided by CAMPR, two researchers independently appraised published readmission models in two recent systematic reviews and concurrently extracted data to generate reference lists of eligibility criteria and risk factors.

Results We found that published models (n=81) followed 6.8 recommendations (45%) on average. Many models had weaknesses in their development, including failure to internally validate (12%), failure to account for readmission at other institutions (93%), failure to account for missing data (68%), failure to discuss data preprocessing (67%) and failure to state the model's eligibility criteria (33%).

Conclusions The high prevalence of weaknesses in model development identified in the published literature is concerning, as these weaknesses are known to compromise predictive validity. CAMPR may support researchers, clinicians and administrators to identify and prevent future weaknesses in model development.

Strengths and limitations of this study

► This study appraises readmission models, of which dozens are published and which have limited regulatory oversight, to ensure high quality and usefulness for health decision-making.

► Recommendations are specific to readmission models and were developed using a modified Delphi process. Critical appraisal was undertaken independently by two researchers.

► Recommendations are limited to model development and validation, and future revision will be necessary as the field advances and developers adopt more modern techniques.

2018.[4] Readmissions have received considerable attention due to this financial burden, their impact on patient care and their use as a performance indicator.[5 6] Consequently, efforts to research[7 8] and market[9 10] readmission models have increased rapidly, and these models are used extensively for health decision-making.

However, uncritically accepting the results of any single model has risks. These models can be subject to flaws in their development and validation, as well as limitations in their clinical usefulness. Given the hundreds of readmission models now available, distinguishing the highest quality, most clinically useful models can be challenging for clinicians, researchers and healthcare administrators. In this study, we address this gap for readmission models in the published literature by critically appraising them. To conduct the critical appraisal, we developed *Critical Appraisal of Models that Predict Readmission* (CAMPR), which lists 15 Delphi-based expert recommendations for high-quality, clinically useful readmission models. CAMPR focuses on the unique considerations of readmission modelling, such as purpose, competing risks, outcome timing, risk factor definitions, data sources and thresholding. This manuscript discusses the expert recommendations and subsequent critical appraisal in detail, and

INTRODUCTION

In 2013, the Centers for Medicare and Medicaid Services (CMS) Hospital Readmission Reduction Program (HRRP) began to financially penalise US hospitals with excessive 30-day readmission rates, with the goal of improving patient care. Subsequently, research on readmission risk prediction models increased exponentially,[1 2] with two distinct goals: (1) to identify high-risk patients for targeted interventions, and (2) to standardise institutions' readmission rates for use as a performance indicator. Preventable hospital readmissions cost CMS $17 billion each year,[3] and CMS penalties for subpar readmission rates totalled $566 million in

provides reference lists of eligibility criteria and risk factors to consider when assessing readmission models.

METHODS

The Delphi method is a well-established structured communication technique for systematically seeking expert opinion on a specific topic.[11–13] Traditionally, the first round uses open-ended questions to generate ideas from participants, whereas subsequent rounds rely on more structured communication to achieve consensus. In this study, we conducted two rounds of online expert surveys, the first open-ended and the second semistructured, and a third round consisting of expert application to the published literature.[11]

Round 1: development of CAMPR (open-ended survey)
Survey content

A Delphi expert used an iterative process to develop the initial survey in collaboration with four physician experts in readmission models. The survey collected personal information on the respondents' institution(s) and relevant expertise, as well as information on models at the respondents' institution(s). Then, the survey assessed perceived barriers to model development and implementation, as well as strategies to overcome barriers and recommendations to improve models. The complete survey is available as online supplemental appendix A.

Data collection

To ensure rapidity and anonymity, provide convenience and recruit individuals from diverse backgrounds and geographical locations, we invited experts to participate electronically using Qualtrics Survey Software. Expert panels for Delphi studies traditionally have 12–20 members.[14] Electronic participation enabled us to include more than 20 members, as desired due to the complex nature of the problem and the probable diversity of opinions. We distributed our survey via personalised, individual emails to all corresponding authors of readmission prediction studies from two recent systematic reviews (2011, 2016).[1 2] Additionally, we publicly distributed it to members of the American Medical Informatics Association.

Eligibility criteria

We included both model developers and implementers in our expert panel to capture a broad range of perspectives on the readmission prediction literature. We required that participants speak English and self-report involvement in (1) the development of one or more readmission models, or (2) the implementation of one or more readmission models at one or more institutions.

Data analysis

Two researchers conducted thematic analysis in NVivo V.11 (QSR International). First, the researchers independently read each response and defined codes in a dictionary for the remaining analysis. Then, the

Table 1 Quality dimensions of readmission risk prediction models

Dimension	Key domain	Description of barrier
Development	Validation	Inadequate validation causing poor model discrimination when generalised
	Features	Over-reliance on biomedical features from administrative and claims data
	Timeframe	Possibility that the 30-day timeframe is not optimal for accurate prediction
	Data access	Dependence on inaccessible, low quality, outdated or manually entered data
Implementation	Resources	Insufficient personnel, statistical expertise or financial resources
	Vision	Lack of leadership, purpose or policy priorities related to readmission prediction
	Clinical relevance	Unclear clinical utility, actionability or relevance
	Workflow integration	Poor integration into the clinical workflow
	Maintenance	Inadequate maintenance or continuous improvement

researchers independently coded all responses using codes corresponding to the dictionary and summarised themes that emerged. Together, the researchers reviewed and named common themes that emerged, and resolved conflicts by discussion. To enhance confirmability, we shared summaries of coded data with three participants and asked for their confirmation or revisions to interpretation.

Round 2: further development of CAMPR (semistructured survey)
Preliminary version

Based on the thematic analysis in round 1, the study team identified 48 preliminary recommendations to operationalise two quality dimensions of readmission models: (1) development and (2) implementation. Each preliminary recommendation addressed one of four key thematic domains for development and five key thematic domains for implementation, identified via expert consensus (table 1).

Survey content and data collection

For each preliminary recommendation, the second survey asked participants to score the usefulness and content validity. Free-text fields enabled participants to add additional comments on each individual recommendation, as well as CAMPR in its entirety. The study team

reviewed the survey before electronic delivery using Qualtrics. We distributed the survey via personalised emails to all previous eligible respondents who had agreed to additional contact. The complete survey is available as online supplemental appendix B.

Data analysis

We conducted a quantitative descriptive analysis of usefulness and content validity using R-3.3.3. Preliminary recommendations with a usefulness and content validity below predetermined thresholds (<50% useful or valid) were excluded, unless the free-text commentary indicated that the usefulness and content validity would greatly improve with revision. The study team reviewed all free-text commentary and refined, reworded or combined recommendations accordingly.

Round 3: application of CAMPR
Modified preliminary version

After refinement and reduction in round 2, the modified preliminary version contained 34 recommendations. We identified 23 development-related recommendations for inclusion in CAMPR, which primarily reflect the four key development domains identified in round 1 (validation, features, timeframe, data access). The remaining 11 implementation-related recommendations, which primarily reflect the five key implementation domains (resources, vision, clinical relevance, workflow integration and maintenance), will be reported on separately.

Iterative validation

Two researchers applied CAMPR to all published readmission prediction models identified in all known systematic reviews (2011, 2016).[1 2] First, the researchers independently applied CAMPR to one-third of studies, then revised it to improve clarity, resolve discrepancies in application and combine redundant recommendations, which reduced the total number to 15. Then, the researchers independently applied the finalised version to all studies (detailed in online supplemental appendix C) and assessed inter-rater reliability. Conflicts were resolved by discussion.

Data extraction

Manual data extraction occurred concurrently with the application of CAMPR, to better assess characteristics of included studies. Importantly, we extracted eligibility criteria and risk factors for each readmission model, to generate reference lists for future developers (available tables 2 and 3 in the Results section). Examples of other types of extracted data include the readmission timeframe used (relevant to recommendation #6) and the validation technique used (relevant to recommendation #13). Two researchers developed the data extraction tool based on the initial review of one-third of studies. One researcher extracted data from each study, and another reviewed all extractions for completeness and accuracy.

Data Analysis

All analyses were performed in R V.3.6.3. We used Cohen's kappa and percentage agreement to measure inter-rater reliability. We stratified literature into recent (2011–2015) and early (1985–2010), using the year of CMS HRRP (2011) as our cut-off. We conducted bivariate analyses to assess whether adherence to each recommendation differed between recent and early literature, using an unequal variances t-test. Furthermore, we used Spearman's rank correlation to examine whether overall adherence to recommendations differed by publication year. When classifying risk factors, we used the same classification as the first systematic review,[1] with the added category 'institution-related' as suggested by expert consensus.

RESULTS
Development of CAMPR
Round 1

We successfully contacted 75 out of 81 corresponding authors who developed unique readmission models published from 1985 to 2015, of whom 14 (19%) completed our survey. An additional 49 respondents completed our survey after we publicly distributed it, of which we included 40 who had experience implementing readmission models. The final 54 eligible experts (14 developers and 40 implementers, characterised in online supplemental appendix D) represented 20 unique models, including well-known models such as LACE and CMS-endorsed models. Of 14 developers, only 7 (50%) reported that any institution currently used their model in any capacity. Table 4 reports expert-identified barriers to developing, validating and implementing readmission models, as well as strategies to overcome barriers.

Rounds 2 and 3

We had permission to reconnect with 22 previous respondents, of whom 5 (23%) completed our second survey.

Application of CAMPR

We included 81 published readmission models in our critical appraisal.[15–95] We found that published models followed 6.8 out of 15 recommendations (45%) on average. Fifty-five out of 81 (68%) followed less than half the recommendations, and no study followed every recommendation, suggesting an opportunity for improvement. Table 5 presents the percentages of published readmission models following each recommendation, stratified by publication year. Models published recently (2011–2015, n=55, 68%) followed significantly more recommendations than models published earlier (1985–2010, n=26, 32%) (7.1 vs 6.1, p=0.03), and publication year weakly correlated with recommendations followed (ρ=0.27, p=0.02), suggesting slight improvement in model quality over time as the field developed. Model types included regression (77, 95%), random forest (3, 4%), neural network (3, 4%), decision tree (2, 2%), discriminant analysis (2, 2%), support vector machine

Table 2 Barriers to developing and implementing readmission models

Domain	Barriers	Evidence*	Example†	Strategies to overcome barriers
Validation	Poor generalisability	Substantial	"[We are] questioning the generalization of the model to our population." (33)	▶ More advanced modelling strategies than regression
	Low discriminative value	Limited	"[The biggest barrier is] having a good model to start with." (48)	▶ Better access to data to improve prediction ▶ External rather than internal validation and calibration ▶ Account for different clinical subpopulations
Features	SDH not included	Extensive	"We need to look well beyond SES, etc. to social support and healthcare beliefs and behaviors." (2)	▶ Consider SDH and HAF when developing models ▶ Place SDH and HAF in structured data fields
	HAF not included	Substantial	"We need to look beyond academic status and for-profit status … to understand processes." (2)	▶ Record better measures of SDH and HAF ▶ Improve natural language processing and extraction ▶ More research on the factors leading to readmission
Timeframe	Timeframe not optimised	Limited	"30 days is probably too long to provide an accurate prediction." (48)	▶ Assess the possibility of using shorter timeframes
Data access	Barriers to data access	Extensive	"No matter how complex and good the model is, it is only as good as the data it has." (20)	▶ Federal incentives for data sharing ▶ Federal sanctions for data blocking ▶ Aggregate data sets from multiple sources
	Inadequate interoperability	Substantial	"We need access to databases, especially linked primary and secondary care ones." (9)	▶ Include sources with multiple data types ▶ Better documentation practices and training ▶ Human resources for data entry and validation
	Insufficient data	Substantial	"We don't have the necessary data and we don't know what the necessary data even are." (27)	▶ Better tools or interfaces for data collection ▶ Implement better structured data capture processes
	Poor quality data	Substantial	"When we use routinely collected data in EHRs, the quality is less reliable." (38)	▶ Reduce unnecessary documentation burdens on healthcare providers, so the quality can improve
	Lacks current information	Limited	"If and when the factors change, we don't know what they are - the case managers do." (34)	
Resources	Lacks personnel or expertise	Substantial	"[We lack] staffing resources for adequate capture of data and analytics of accumulated data." (43)	▶ More support from the hospital administration ▶ Better financial support for model development
	Financial barriers	Substantial	"[There is] reluctance to make the necessary investments to access the EHR's back end." (1)	▶ Better packages in statistical software to help developers move beyond logistic regression
Vision	Competing priorities	Limited	"It [the model] is not a priority… they [the administration] have competing priorities. (14)	▶ Engage key stakeholders and senior decision-makers ▶ Establish governance over model implementation
	Lack of leadership	Very limited	"[There is] no operational leadership, so the model hasn't been implemented." (21)	▶ Select a clinical or administrative champion ▶ Implement financial and other incentives for hospitals
Clinical relevance	Poor perceived relevance	Extensive	"Risk score doesn't necessarily flag the patients in whom we can most usefully intervene." (9)	▶ Models must determine how providers can intervene ▶ More research on hospital processes and needs
	Unclear usefulness	Substantial	"[Models must] fit into a workflow where an intervention can be made." (16)	▶ Invest more human resources to prevent readmission ▶ Feedback for providers if readmission averted (or not)
	Poor perceived accuracy	Substantial	"I've found that it [the model] is not accurate at the individual patient level." (6)	▶ Validate and calibrate model in the new clinical setting ▶ Ensure users receive the most up-to-date predictions
Workflow integration	Poor workflow integration	Extensive	"Getting it inserted into the EHR in a way that requires little provider effort is tough." (48)	▶ Consider who will receive the results, when, and how ▶ Tie integration to physical risk-reduction interventions
	Alert fatigue	Very limited	"Clinicians get alert fatigue and stop paying attention to the results." (33)	▶ Train, orient, and support potential users of the model ▶ Plan to iteratively improve on integration over time
Maintenance	Antiquated model or interface	Limited	"Our commercial partner no longer supports the front end they developed [for our model]." (9)	▶ Integrate within system with long-term IT support ▶ Designate responsibility of updating model annually

Table 2	Continued			
Domain	Barriers	Evidence*	Example†	Strategies to overcome barriers

*Extensive evidence (≥8 mentions); substantial evidence (4–7 mentions); limited evidence (2–3 mentions), very limited evidence (1 mention).
†Numbers in parentheses refer to participant study ID.
EHR, electronic health record; HAF, healthcare-associated factors; IT, information technology; SDH, social determinants of health; SES, socioeconomic status.

(1, 1%) and unclear (1, 1%). We found moderate-to-high inter-rater reliability for applying CAMPR (Cohen's kappa=0.76, agreement=88%). Here, we summarise each recommendation in CAMPR and present the critical appraisal results. Additional results are in online supplemental appendix D. The complete dataset is available on request. CAMPR is available as online supplemental appendix E.

Recommendation #1: is the model's purpose and eligibility criteria explicitly stated?

About the recommendation

Readmission models traditionally serve one of two purposes, or intended applications: (1) to identify patient candidates for targeted interventions to prevent readmission, or (2) to risk-adjust readmission rates for hospital quality comparison.[1] Developers should clearly state

Table 3 Critical Appraisal of Models that Predict Readmission (CAMPR)

Number	Recommendation	Studies following recommendation (n)			P value*
		Overall, n=81 (1985–2015)	Early, n=26 (1985–2010)	Recent, n=55 (2011–2015)	
#1	Is the model's purpose and eligibility criteria explicitly stated?	44 (54%)	11 (42%)	33 (60%)	0.14
#2	Does the model consider common patient-related and institution-related risk factors for readmission?	1 (1%)	1 (4%)	0 (0%)	0.33
#3	Does the model consider competing risks to readmission, particularly mortality?	68 (84%)	23 (88%)	45 (82%)	0.43
#4	Does the model identify how providers may intervene to prevent readmission?	4 (5%)	3 (12%)	1 (2%)	0.15
#5	Does the model consider recent changes in the patient's condition?	39 (48%)	6 (23%)	33 (60%)	0.001**
#6	Is the model's timeframe an appropriate trade-off between sensitivity and statistical power?	12 (15%)	4 (15%)	8 (15%)	0.92
#7	Does the model exclude either planned or unavoidable readmissions?	42 (52%)	12 (46%)	30 (55%)	0.49
#8	Is the model equipped to handle missing data and is missingness in the development datasets reported?	13 (16%)	4 (15%)	9 (16%)	0.91
#9	Is preprocessing discussed and does the model avoid problematic preprocessing, particularly binning?	27 (33%)	9 (35%)	18 (33%)	0.96
#10	Does the model make use of all available data sources to improve performance?	59 (73%)	15 (58%)	44 (80%)	0.05
#11	Does the model use electronically available data rather than relying on manual data entry?	55 (68%)	17 (65%)	38 (69%)	0.75
#12	Does the model rely on data available in sufficient quantity and quality for prediction?	58 (72%)	18 (69%)	40 (73%)	0.75
#13	Is the model internally validated using cross-validation or a similarly rigorous method?	4 (5%)	1 (4%)	3 (5%)	0.75
#14	Is the model's discrimination reported and compared with known models where appropriate?	74 (91%)	21 (81%)	53 (96%)	0.07
#15	Is the model calibrated if needed and is calibration reported?	47 (58%)	14 (54%)	33 (60%)	0.61

*Significant at p=0.05; **significant at p=0.001.
*Comparison is between models published earlier (1985–2010) and more recently (2011–2015).

Table 4 Eligibility criteria for readmission prediction models*

Criterion	Studies (n, %)	Citations (see online supplemental appendix E)
Inclusion criteria		
Age	48 (59%)	
>18 years (adults only)	25 (31%)	37 40 43 45 47 49–51 55 56 65 67 68 73 74 78 80–82 85 86 88 92 93 95
>65 years (elderly only)	19 (23%)	17–19 21 25 27 28 32–34 38 48 52 60 66 76 77 79 90
Other	5 (6%)	23 36 42 62 93
Condition specific	45 (56%)	
Heart failure	16 (20%)	18 21 26 29 33 46 47 67–69 71–75 86
AMI or other cardiovascular	11 (14%)	3 46–58 64–66 70
Pneumonia or other pulmonary	6 (7%)	19 46 47 62 76 77
Multiple	6 (7%)	23 24 28 31 54 63
Other	11 (14%)	80 84–93
Service specific	24 (30%)	
Medicine	14 (17%)	35 36 38 40 44 45 78–83 87
Surgery	7 (9%)	57 60 63 89–92
Other	3 (4%)	52 94 95
Beneficiaries	22 (27%)	
Medicare	15 (19%)	15 17–19 21 25 28 32 33 41 48 58 60 66 90
Veterans	7 (9%)	22 36 38 39 53 58 61
Exclusion criteria		
Disposition	28 (35%)	
Left against medical advice	11 (14%)	21 40 42 45 46 50 58 60 61 80 82
Skilled nursing or other care facility	10 (12%)	32 37 43 50 56 61 68 78 86 93
Hospice care	8 (10%)	32 43 49 61 78 85 86 88
Different healthcare system	4 (5%)	35 36 44 94
Transfers	26 (32%)	
Transfer-out (to other institution or service)	24 (30%)	17–22 27 33 40 42 45 46 56 58 61–63 68 72 74 76 77 80 93
Transfer-in (from other institution or service)	3 (4%)	47 53 61
Service or condition	24 (30%)	
Rehabilitation	10 (12%)	21 43 49–51 56 58 61–63 68 72 74 76 77 80 93
Psychiatric	8 (10%)	22 27 43 50 51 55 56 61
Obstetric/neonatal	8 (10%)	20 23 42 43 47 51 55 56 61
Transplant	4 (5%)	75 84–86
Other (ESRD, trauma, etc)	5 (6%)	15 34 36 42 86
Length of stay	22 (27%)	
<24 hours (same-day discharge or procedure)	16 (20%)	17 21 23 27 45–47 51 56 61 66 78 80 92 94
>30 days	5 (6%)	40 57 63 90 91
Hospitalisation type	9 (11%)	
Planned (aka elective)	3 (4%)	34 93
Non-acute (aka non-emergent)	3 (4%)	24 42 79
Observation only	2 (2%)	43 51
Study related	7 (9%)	
Unavailable for follow-up (no telephone, etc)	5 (6%)	36 37 40 70 89
Cannot consent (poor cognition)	2 (2%)	37 67

Continued

Table 4 Continued		
Criterion	Studies (n, %)	Citations (see online supplemental appendix E)
Non-English speaking	2 (2%)	67 73
Non-residents (outside the primary service area)	7 (9%)	20 33 36 46 65 67 93
No insurance coverage (for part or all of the data extraction period)	5 (6%)	42 56 60 66 77
Data-related (not recent or validated)	3 (4%)	32 33 52
Readmitted for different condition	3 (4%)	17–19

An extended version of this table is available in online supplemental appendix D.
*This table does not cover exclusions based on mortality (see recommendation #3) or missingness (see recommendation #8).
AMI, acute myocardial infarction; ESRD, end-stage renal disease.

which purpose their model serves, one or both. Developers should also define the target population by specifying eligibility criteria for patient inclusion in model development. Specifying eligibility criteria is critical to ensure implementers understand when each model applies, as unjustified application is a major reason why predictions fail.[96 97]

Critical appraisal results
Eighteen out of 81 studies (22%) did not define their model's purpose. Of the remaining models, 46 (57%) were for preventing readmission, 15 (19%) were for hospital quality comparison and 2 (2%) were for both. Table 4 provides an abbreviated reference list of eligibility criteria for published readmission models (the full reference list is available in online supplemental appendix D). Twenty-seven models (33%) did not specify their eligibility criteria.

Recommendation #2: does the model consider common patient-related and institution-related risk factors for readmission?
About the recommendation
Developers should show that they considered risk factors or features that were included in previous models. Notably, institution-related factors such as hospital name should not be used in models for hospital quality comparison, as they can mask differences in hospital quality.

Critical appraisal results
Table 5 provides an abbreviated reference list of known risk factors for readmission and their frequency of inclusion in published models (the full reference list is available as online supplemental appendix D). Based on expert consensus and the existing literature, we identified seven categories of factors.[1] Categories included (1) demographics (included in 75 models or 93%), (2) disease related (80, 99%), (3) functional ability (21, 26%), (4) healthcare utilisation (66, 81%), (5) medication related (33, 41%), (6) social determinants of health (53, 65%), (7) institution related (16, 23%). Five studies (out of 15, 33%) mistakenly used institution-related risk factors in models for hospital quality comparison.

Recommendation #3: does the model consider competing risks to readmission, particularly mortality?
About the recommendation
Death is a competing risk to readmission and may substantially impact readmission prediction depending on the target population.[63 67 68] A high mortality rate may reduce model discrimination because death and readmission share similar predictive features. Ignoring mortality may limit insight about risk factors, and unaccounted changes in mortality may cause model drift. Developers should indicate that they accounted for both in-hospital and post-discharge mortality, as well as other competing risks to readmission (eg, transfers).[28]

Critical appraisal results
Thirteen models (16%) did not account for mortality, 40 (49%) accounted for in-hospital mortality only, 5 (6%) accounted for post-discharge mortality only and 21 (26%) accounted for both.

Recommendation #4: does the model identify how providers may intervene to prevent readmission?
About the recommendation
The expert group recognised that building *actionable* models, which identify where providers can intervene on risk factors to prevent readmissions, is critical to clinical usefulness. An actionable model may (1) identify modifiable risk factors on the individual level,[36 62 90] or (2) identify which individuals will benefit most from intervention, which may not coincide with readmission risk. Notably, non-modifiable risk factors like age can obscure modifiable ones like polypharmacy or quality of care[60 98 99]; therefore, managing collinearity[100] is important. In the future, predicting benefit will become easier as options for intervention become more well researched.[101]

Critical appraisal results
Four published models (5%) identified modifiable factors on the individual level. No models have predicted which individuals would benefit most from intervention.

Table 5 Risk factors included in readmission prediction models

Risk factor	Studies (n, %)		Citations (see online supplemental appendix E)
	Overall (n=81)	General* (n=12)	
Demographics			
Age	38 (47%)	6 (50%)	15–22 24 25 27 28 31 32 39 41 42 50 56 60–63 65–67 69 70 72 74 76 77 79 80 82 90 93 95
Sex/gender	30 (37%)	5 (42%)	15–21 25 27 28 31 41 50 53 56 60–63 65–67 69 76 77 79 83 85 90 93
Race/ethnicity	12 (15%)	2 (17%)	15 16 26 27 31 41 46 52 57 63 71 72
Disease related			
Comorbidities	50 (62%)	4 (33%)	
Cardiovascular (heart failure, infarction, etc)	39 (48%)	4 (33%)	17–21 23 24 26 27 31–34 36 38 41 42 46 53 57 60–67 69 76–79 82 83 86 90 93 94
Oncological (metastatic, leukaemia, etc)	30 (37%)	4 (33%)	17–21 23 24 27 32 34 41 42 45 46 53 54 61–63 66 76–80 82 83 90 91 93
Pulmonary (COPD, pneumonia, asthma, etc)	29 (36%)	3 (25%)	17–21 23 24 26 27 31 38 42 53 57 60 61 63 65 66 76–79 82 83 86 90 91 93
Endocrine (diabetes, etc)	26 (32%)	2 (17%)	17–19 21 23 24 26 31–34 41 42 53 57 60 62–64 66 70 77 82 83 90 93
Genitourinary (renal failure, etc)	26 (32%)	3 (25%)	17–21 23 24 26 27 31 42 53 57 60 62–64 66 76 78–80 83 86 90 91
Psychiatric (alcohol/drug use, psychosis, etc)	21 (26%)	1 (8%)	18–21 23 24 29 31 39 46 52 53 61 62 76 77 82 90 93–95
Haematological (fluid disorder, anaemia, etc)	20 (25%)	2 (17%)	17–21 23 24 27 31 53 57 62–64 66 83 86 90 92 93
Neurological (stroke, paralysis, etc)	17 (21%)	3 (25%)	17–21 23 27 31 32 41 42 46 66 82 86 89 93
Gastrointestinal (cirrhosis, obstruction, etc)	16 (20%)	3 (25%)	18–21 23 24 27 42 46 53 77 83 86 88 90 93
End of life (dementia, malnutrition, cachexia, etc)	14 (17%)	3 (25%)	17–21 23 27 42 46 63 66 77 83 90
Musculoskeletal-dermatological (injuries, etc)	12 (15%)	2 (17%)	19–21 23 27 31 41 53 61 62 86 93
Infectious (sepsis, shock, etc)	10 (12%)	–	17–19 21 23 53 63 66 70 90
Obesity related (BMI, sleep apnoea, etc)	4 (5%)	–	62 63 78 90
Obstetric-gynaecological (pregnancy, etc)	3 (4%)	1 (8%)	20 23 53
Severity scores	29 (36%)	8 (67%)	
Charlson Comorbidity Index	14 (17%)	7 (58%)	16 20 32 37 40 46 50 51 55 56 61 74 92 95
Other (Elixhauser, Tabak, SOI, PMC-RIS, etc)	10 (12%)	2 (17%)	22 25 28 29 38 47 52 53 56 80
Disease specific (GRACE, MELD, etc)	6 (7%)	–	70 84 85 90 91 93
Laboratory values	22 (27%)	1 (8%)	21 33 35 36 45 46 51 57 60 65 69–72 76 77 84–88 91
Complications (post-procedure, etc)	8 (10%)	–	57 63 85 87 89 90 92 94
Emergent admission (or acute admission)	6 (7%)	4 (33%)	37 38 42 55 56 95
Complexity (No of medical conditions)	6 (7%)	–	28 39 52 67 69 79
Signs or symptoms (dyspnoea, ascites, etc)	5 (6%)	–	23 63 73 88 90
Condition description (chronic, high risk, etc)	4 (5%)	1 (8%)	16 25 68
Functional ability			
Functional status (or assistance with ADL)	5 (6%)	–	32 36 40 63 90
Mental status (MMSE, etc)	3 (4%)	1 (8%)	34 39 77
Dependencies (ambulation, ventilator, etc)	3 (4%)	–	39 63 94
Recent falls	2 (2%)	–	41 52
Other (bedridden, incontinent, sedentary, etc)	2 (2%)	–	41 67
Healthcare utilisation			
Previous admissions	32 (40%)	7 (58%)	15 20 24 29 31–34 39–42 45–47 49 50 52 56 61 62 64 71 76 78–80 82 84 87 93 95
Length of stay	26 (32%)	6 (50%)	28 37 40 45–47 50 51 53–56 61–63 68 69 74 76 78 79 83 86 91 92 95
Previous emergency visits	11 (14%)	5 (42%)	35 37 46 49 50 55 56 74 79 87 95

Continued

Table 5 Continued

Risk factor	Studies (n, %)		Citations (see online supplemental appendix E)
	Overall (n=81)	General* (n=12)	
Previous outpatient visits (specialist or primary care)	7 (9%)	–	29 41 54 75 78 86 95
Previous procedures	5 (6%)	–	26 60 64 65 89
Current admission	25 (31%)	3 (25%)	
Disposition (home, SNF, rehab, home health, etc)	8 (10%)	–	26 41 57 63 85 86 92 94
Source (emergency, SNF, outpatient, transfer, etc)	7 (9%)	1 (8%)	16 54 61 63 65 79 86
Ward (medical, surgical, neurology, etc)	3 (4%)	1 (8%)	22 27 54
Reimbursement amount	3 (4%)	1 (8%)	15 16 88
Other (consultancies, timing, refusal, etc)	9 (11%)	2 (17%)	16 29 31 45 47 62 69 71 75
Current procedures	15 (19%)	2 (17%)	
Performed vs not (binary)	9 (11%)	2 (17%)	15 20 22 45 46 51 53 60 77
Type (arthroplasty, splenectomy, etc)	6 (7%)	1 (8%)	26 46 51 57 60 77
Characteristic (high risk, operation time, open, etc)	6 (7%)	1 (8%)	20 57 63 64 90 92
Status (urgent, emergent, elective, etc)	4 (5%)	–	57 60 63 90
Medication related			
Specific medications (in-hospital or home)	19 (23%)	1 (8%)	
Steroid	5 (6%)	–	46 57 62 63 90
ACE inhibitor (or ARB)	3 (4%)	–	62 68 69
Other cardiac (beta-blocker, nitrate, etc.)	5 (6%)	1 (8%)	34 46 60 70 82
Immunosuppressant	3 (4%)	–	60 76 77
Antibiotic	3 (4%)	–	46 62 88
Opioid	3 (4%)	–	80 82 86
Other (statin, anticoagulant, insulin, etc)	7 (9%)	–	46 62 70 82 86 90 92
No of medications (polypharmacy)	7 (9%)	2 (17%)	34 46 50 52 80 82 85
Social determinants of health			
Zip code (or home address)	11 (14%)	2 (17%)	
Distance from hospital	4 (5%)	–	22 64 78 87
Socioeconomic status (based on zip code)	4 (5%)	2 (17%)	16 24 29 42
Rurality (urban, suburban, rural, etc)	3 (4%)	–	24 64 69
Insurance status	10 (12%)	2 (17%)	15 26 27 29 32 40 50 65 84 87
Living arrangement (alone, SNF, homeless, etc)	7 (9%)	–	39 41 67 69 76 77 87
Marital status	7 (9%)	2 (17%)	29 34 39 40 46 50 76
Disability status	6 (7%)	–	
Disabled vs not (binary)	5 (6%)	–	22 31 32 41 78
Type (developmental, visual, cognitive, hearing, etc)	3 (4%)	–	31 32 41
Patient-generated health data	5 (6%)	–	32 36 40 41 61
Smoking status (current, former, etc)	3 (4%)	1 (8%)	34 63 90
Education level	2 (2%)	–	41 89
Annual income	2 (2%)	–	41 76
Institution related			
Rurality (urban, suburban, rural, etc)	2 (2%)	–	15 26

Continued

Table 5 Continued

Risk factor	Studies (n, %) Overall (n=81)	General* (n=12)	Citations (see online supplemental appendix E)
Standardised admission ratios	2 (2%)	1 (8%)	16 31
Identification code	2 (2%)	1 (8%)	42 52

We could not extract factors from seven studies, either due to poor reporting or lack of feature selection.[30 43 44 48 58 59 81] An extended version of table is available in online supplemental appendix D.

*'General' indicates models for the general hospitalised population, and is not restricted to any one target population.[16 20 27 34 37 42 47 49–51 55 56]

ACE, Angiotensin-converting enzyme; ADL, activities of daily living; ARB, Angiotensin receptor blocker; BMI, body mass index; COPD, chronic obstructive pulmonary disease; DNR, do not resuscitate; MMSE, Mini Mental State Exam; SNF, skilled nursing facility; SOI, severity of illness.

Recommendation #5: does the model consider recent changes in the patient's condition?

About the recommendation

A model that does not account for recent changes in the patient's condition may give an outdated prediction, limiting its clinical usefulness and eroding trust in its predictions. The expert group recommended that models that give predictions near hospital discharge (ie, most current models) should account for changes during hospitalisation, including treatment effects, hospital-acquired conditions and social support status.

Critical appraisal results

Thirty-nine models (48%) accounted for changes during hospitalisation.

Recommendation #6: is the model's timeframe an appropriate trade-off between sensitivity and statistical power?

About the recommendation

Researchers initially selected the 30-day timeframe as the optimal trade-off between statistical power and likelihood of association with the index admission.[28] As common data models and health information exchange support larger datasets for model development, shorter timeframes such as 7 days may enable greater sensitivity for readmissions associated with the index admission without loss in statistical power.[102] Therefore, developers should consider assessing prediction accuracy using multiple timeframes, as relevant to the clinical context and dataset size, to determine the best trade-off between sensitivity and statistical power. Timeframes should begin at discharge (as standardised by CMS)[103] to prevent immortal person–time bias.[63]

Critical appraisal results

Sixty-three models (78%) used the standardised 30-day timeframe adopted by CMS, while 2 (2%) used 7 days, 3 (4%) 28 days, 5 (6%) 60 days and 9 (11%) 1 year. Twelve studies (15%) considered more than one timeframe, of which 7 (9%) modelled readmission risk using hazard rates. Nine studies (11%) inappropriately defined timeframes as beginning at admission, rather than discharge,

and 14 (17%) did not specify when their timeframe began.

Recommendation #7: does the model exclude either planned or unavoidable readmissions?

About the recommendation

Planned readmission is defined as non-acute readmission for scheduled procedures. Planned readmissions should be excluded, as consistent with the standardised definition of all-cause readmission.[17–19 66] *Unavoidable* readmission is defined more broadly as readmission not preventable through clinical intervention.[28] As researchers develop standardised algorithms to more effectively identify unavoidable readmissions,[47 61 104] using the broader definition may enable greater sensitivity and improve the relevance of predictions to the clinical setting. Therefore, developers should consider excluding unavoidable readmissions if it is useful, such as in multiple sclerosis, where the disease inevitably progresses and later readmissions become increasingly unavoidable. Notably, exclusion criteria can be highly complex and require third-party processing (eg, Vizient). Ideally, developers should publish their code. If not, the readmission outcome should be sufficiently defined to ensure transparency and reproducibility.[105]

Critical appraisal results

Thirty-nine models (48%) did not explicitly exclude planned readmissions. The remaining models either excluded planned readmissions (38, 47%) or excluded unavoidable readmissions more broadly (4, 5%).

Recommendation #8: is the model equipped to handle missing data and is missingness in the development datasets reported?

About the recommendation

Developers should explicitly state whether their model handles missingness and how, such as designating a 'missing' category for categorical variables, or multiple regression imputation for continuous variables. Dropping individuals with excess missingness is problematic because it decreases models' generalisability to future individuals with excess missingness and falsely increases

model performance in cases of structural missingness. Developers should also report on missingness in the datasets used for model development, so that implementers can determine potential generalisability to their real-world datasets.

Critical appraisal results

Only 34 studies (42%) discussed how their model handled missingness. Of these, 20 (25%) used one or more inappropriate techniques, including (1) dropping individuals with excess missingness (17, 21%), and (2) binning or imputation which was done improperly (3, 4%).

Recommendation #9: is preprocessing discussed and does the model avoid problematic preprocessing, particularly binning?
About the recommendation

Developers should explain their data preprocessing methods, because problematic methods may produce models with less-than-optimal predictive performance.[96] One example of a problematic method is binning.[106] Originally intended to improve interpretability, binning can cause information loss, and is no longer justifiable given users' need for accurate predictions and modern interpretability techniques. In particular, manual or arbitrary binning, without clustering or splines, may decrease performance and introduce noise.[107]

Critical appraisal results

Only 27 studies (33%) discussed one or more data preprocessing techniques, despite mostly using regression models, which can be highly sensitive to small changes. Commonly discussed techniques included binning, interaction terms and transformations to resolve skewness, non-linearity and outliers.

Recommendation #10: does the model make use of all available data sources to improve performance?
About the recommendation

Developers should make use of publicly available data sources where possible and appropriate to the model's purpose, such as the Social Security Death Index to determine post-discharge mortality (see recommendation #3) or curated public datasets to externally validate (see recommendation #13). Other data sources such as health information exchanges can help assess readmission at multiple institutions, which is desirable to better estimate the true readmission rate. When considering data sources from multiple institutions, such as with health information exchanges, developers should account for hospital-level patterns and clustering of readmission risk, which may occur because quality of care and data collection practices vary between institutions.[36 61 68]

Critical appraisal results

In the literature, data sources included claims (19, 23%), administrative datasets (33, 41%), electronic health records (42, 52%), disease-specific registries (12, 15%), research datasets (11, 14%), death registries (9, 11%), health information exchanges or linkages (4, 5%), and

surveys or patient-generated health data (3, 4%). Seventy-five studies (93%) assessed readmission at only one institution, likely underestimating the true readmission rate. Thirty-nine studies (48%) used a single data source (administrative datasets: 15, 19%; electronic health records: 17, 21%).

Recommendation #11: does the model use electronically available data rather than relying on manual data entry?
About the recommendation

Developers should incorporate risk factors that will be available electronically at the time of prediction and avoid manual data entry by providers or research assistants. Manual data entry may inhibit widespread implementation, by consuming human resources and preventing automated generation of predictions.

Data extraction results

Twenty-six models (32%) relied on manual data entry.

Recommendation #12: does the model rely on data available in sufficient quantity and quality for prediction?
About the recommendation

Developers should indicate whether data included in their model can be accessed in sufficient amounts and quality for development and implementation. 'Sufficient' is subjective and requires consideration of real-world missingness. Automated quality assurance, which identifies erroneous entries (eg, age>120 years) and incorrect data combinations (eg, former smoker YES, never smoker YES), may help to improve quality.

Critical appraisal results

Twenty-three studies (28%) identified problems with either data quantity (17 out of 23, 74%) or quality (8 out of 23, 26%).

Recommendation #13: is the model internally validated using cross-validation or a similarly rigorous method?
About the recommendation

The importance of using repeated k-fold cross-validation or a similarly rigorous method is well established. Split-sample validation is insufficient and may cause unstable and suboptimal predictive performance.[108–110] If the model is intended for generalised use at more than one institution, developers or implementers should confirm external validity using one or more external, representative and independent datasets, from another institution or source. Internal validation alone is insufficient to ensure generalisability.[108]

Critical appraisal results

Ten models (12%) were not validated at all, 64 (79%) were internally validated only and 7 (9%) were internally and externally validated. For internal validation, 46 (57%) used random split-sample, 11 (14%) used split-sample by time, 12 (15%) used bootstrapping, 3 (4%) used cross-validation and 1 (1%) used out-of-bag estimates.

Recommendation #14: is the model's discrimination reported and compared with known models where appropriate?

About the recommendation

It is commonly accepted practice to prominently and clearly report discrimination using appropriate and well-known measures beyond just the concordance (c) statistic.[111] Where possible, comparison with an established baseline is essential, because so many models already exist. Developers should compare performance using statistical tests with cross-validation or another method, and only compare models with similar eligibility criteria.

Critical appraisal results

Seven models (8%) did not report discrimination. Commonly reported measures included the c statistic (47, 58%), sensitivity or specificity (23, 28%), area under the receiver operating characteristic curve (19, 23%), negative or positive predictive value (18, 22%), integrated discrimination improvement (5, 6%) and net reclassification improvement (3, 4%).

Recommendation #15: is the model calibrated if needed and is calibration reported?

About the recommendation

Proper calibration is critical for sorting patients in descending order of readmission risk for making intervention decisions. It is commonly accepted practice to report calibration using calibration curves with no binning.[112 113] Reporting the Hosmer-Lemeshow (HL) goodness-of-fit statistic is insufficient, as a non-significant HL statistic does not imply the model is well calibrated, and the HL statistic is often insufficient to detect quadratic overfitting effects common to support vector machines and tree-based models.[112]

Critical appraisal results

Thirty-seven models (44%) did not assess calibration. Commonly reported measures included the HL statistic (29, 36%) and observed-to-expected ratios (17, 21%).

DISCUSSION

In this study, we critically appraised readmission models using 15 Delphi-based expert recommendations for development and validation. Interestingly, we found that many published readmission models did not follow the experts' recommendations. This included failure to internally validate (12%), failure to account for readmission at other institutions (93%), failure to account for missing data (68%), failure to discuss data preprocessing (67%) and failure to state the model's eligibility criteria (33%). The high prevalence of these weaknesses in model development identified in the published literature is concerning, because these weaknesses are known to compromise predictive validity. Identification of weaknesses in these domains should undermine confidence in a model's predictions.

In our expert surveys, several lessons emerged, most notably about improving models' relevance to clinical care and integration of predictions into clinicians' workflows. In particular, experts expressed concern that models identified the highest-risk patients rather than the patients who might benefit most from intervention, which led to recommendation #4. Experts also noted that the published literature in existing systematic reviews and therefore our critical appraisal is focused on development and internal validation. This suggests that literature on external validation and implementation is less common. Additional efforts to research external validation and implementation could improve readmission models, by making them more applicable to a broader patient population.

In the future, CAMPR may be a convenient teaching aid for model implementers and users at healthcare institutions, such as clinicians and healthcare administrators, as well as for model developers in academic and commercial research. CAMPR does not explain the detailed logic and methods of developing and implementing predictive models, and those looking for comprehensive advice should consult other resources. Finally, CAMPR is not intended as a reporting standard for academic studies, and responses to CAMPR recommendations should not be used to derive an overall score. An overall score may disguise critical weaknesses that should diminish confidence in model predictions. Rather than generating an overall score, consider the potential impact of failing to follow each recommendation, and how that may interact with the use of that model in the given patient population.

We developed CAMPR using a modified Delphi process consisting of two online rounds, which we found faster and more practical than conducting the traditional in-person meetings and three rounds. Beyond readmission modelling, other predictive modelling domains in healthcare (eg, sepsis risk, mortality risk, etc) could benefit from similar guidance. Thinking beyond better modelling techniques is essential, or model predictions will remain of limited clinical use. This includes thinking about how to generate better datasets, thinking about model drift and maintenance over time,[114] and thinking about how to clinicians should act on predictions.

Limitations

The study used a modified Delphi process, which may lack rigour compared with the traditional Delphi process. We used an 'opt-in' process to recruit experts, and this self-selection bias may have led to missed recommendations or opinions. Fewer participants responded to the second round than expected, although the number was sufficient for the Delphi process. Future revision of recommendations will likely be necessary as the field advances and developers adopt more modern techniques. CAMPR is not intended as a reporting standard, and a more formal evaluation of construct validity and generalisability would be needed before it could be used as such.

Author affiliations
¹Department of Biomedical Informatics, Columbia University, New York, New York, USA
²Department of Biomedical Informatics, Vanderbilt University, Nashville, Tennessee, USA
³Department of Medicine, Vanderbilt University, Nashville, Tennessee, USA
⁴Department of Psychiatry, Vanderbilt University, Nashville, Tennessee, USA
⁵Department of Medicine, Oregon Health and Science University and VA Portland Health Care System, Portland, Oregon, USA
⁶Steele Institute for Health Innovation, Geisinger, Danville, Pennsylvania, USA
⁷Division of General Internal Medicine, Brigham and Women's Hospital, Boston, Massachusetts, USA
⁸Harvard Medical School, Boston, Massachusetts, USA
⁹Mass General Brigham, Somerville, Massachusetts, USA

Twitter Lisa Grossman Liu @LisaGrossmanLiu

ORCID iDs
Lisa Grossman Liu http://orcid.org/0000-0002-1153-1590
Hojjat Salmasian http://orcid.org/0000-0002-9004-7149

REFERENCES

1 Kansagara D, Englander H, Salanitro A, *et al*. Risk prediction models for hospital readmission: a systematic review. *JAMA* 2011;306:1688–98.
2 Zhou H, Della PR, Roberts P, *et al*. Utility of models to predict 28-day or 30-day unplanned Hospital readmissions: an updated systematic review. *BMJ Open* 2016;6:e011060.
3 Agency for Healthcare Research and Quality. Statistical Brief #172: Conditions With the Largest Number of Adult Hospital Readmissions by Payer, 2011, 2014. Available: https://www.hcup-us.ahrq.gov/reports/statbriefs/sb172-Conditions-Readmissions-Payer.pdf
4 Centers for Medicare & Medicaid Services. Federal register volume 83, number 160, rules and regulations, 2018. Available: https://www.govinfo.gov/content/pkg/FR-2018-08-17/html/2018-16766.htm
5 Ody C, Msall L, Dafny LS, *et al*. Decreases in readmissions credited to Medicare's program to reduce hospital readmissions have been overstated. *Health Aff* 2019;38:36–43.
6 Wadhera RK, Joynt Maddox KE, Wasfy JH, *et al*. Association of the hospital readmissions reduction program with mortality among Medicare beneficiaries hospitalized for heart failure, acute myocardial infarction, and pneumonia. *JAMA* 2018;320:2542–52.
7 Lin Y-W, Zhou Y, Faghri F, *et al*. Analysis and prediction of unplanned intensive care unit readmission using recurrent neural networks with long short-term memory. *PLoS One* 2019;14:e0218942.
8 Maltenfort MG, Chen Y, Forrest CB. Prediction of 30-day pediatric unplanned hospitalizations using the Johns Hopkins adjusted clinical groups risk adjustment system. *PLoS One* 2019;14:e0221233.
9 Jvion. Jvion: readmission reduction with prescriptive analytics for preventable harm, 2017. Available: https://jvion.com/impact/readmissions
10 Medial EarlySign. New Study Shows EarlySign's Machine Learning Algorithm Can Predict Which Cardiac Patients are at High-Risk Following Discharge. *Cision PR Newswire*, 2019. Available: https://www.prnewswire.com/il/news-releases/new-study-shows-earlysigns-machine-learning-algorithm-can-predict-which-cardiac-patients-are-at-high-risk-following-discharge-300911407.html
11 Dalkey N, Helmer O. An experimental application of the Delphi method to the use of experts. *Manage Sci* 1963;9:458–67.
12 Goodman CM. The Delphi technique: a critique. *J Adv Nurs* 1987;12:729–34.
13 Katrak P, Bialocerkowski AE, Massy-Westropp N, *et al*. A systematic review of the content of critical appraisal tools. *BMC Med Res Methodol* 2004;4:22.
14 J. Skulmoski G, T. Hartman F, Krahn J. The Delphi method for graduate research. *JITE:Research* 2007;6:001–21.
15 Anderson GF, Steinberg EP. Predicting Hospital readmissions in the Medicare population. *Inquiry* 1985;22:251–8.
16 Bottle A, Aylin P, Majeed A. Identifying patients at high risk of emergency hospital admissions: a logistic regression analysis. *J R Soc Med* 2006;99:406–14.
17 Krumholz HM, Normand S-LT, Keenan PS. *Hospital 30-day acute myocardial infarction readmission measure methodology*, 2008.
18 Krumholz H, Normand S, Keenan P. *Hospital 30-day heart failure readmission measure methodology*, 2008.
19 Krumholz HM, Normand S-LT, Keenan PS. *Hospital 30-day pneumonia readmission measure methodology*, 2008.
20 Halfon P, Eggli Y, Prêtre-Rohrbach I, *et al*. Validation of the potentially avoidable Hospital readmission rate as a routine indicator of the quality of hospital care. *Med Care* 2006;44:972–81.
21 Hammill BG, Curtis LH, Fonarow GC, *et al*. Incremental value of clinical data beyond claims data in predicting 30-day outcomes after heart failure hospitalization. *Circ Cardiovasc Qual Outcomes* 2011;4:60–7.
22 Holloway JJ, Medendorp SV, Bromberg J. Risk factors for early readmission among Veterans. *Health Serv Res* 1990;25:213–37.
23 Holman C D'Arcy J, Preen DB, Baynham NJ, *et al*. A multipurpose comorbidity scoring system performed better than the Charlson index. *J Clin Epidemiol* 2005;58:1006–14.
24 Howell S, Coory M, Martin J, *et al*. Using routine inpatient data to identify patients at risk of hospital readmission. *BMC Health Serv Res* 2009;9:96.
25 Naessens JM, Leibson CL, Krishan I, *et al*. Contribution of a measure of disease complexity (complex) to prediction of outcome and charges among hospitalized patients. *Mayo Clin Proc* 1992;67:1140–9.
26 Philbin EF, DiSalvo TG. Prediction of hospital readmission for heart failure: development of a simple risk score based on administrative data. *J Am Coll Cardiol* 1999;33:1560–6.
27 Silverstein MD, Qin H, Mercer SQ, *et al*. Risk factors for 30-day Hospital readmission in patients ≥65 years of age. *Proc* 2008;21:363–72.
28 Thomas JW. Does risk-adjusted readmission rate provide valid information on hospital quality? *Inquiry* 1996;33:258–70 http://www.ncbi.nlm.nih.gov/pubmed/8883460
29 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;48:981–8.

30 Billings J, Mijanovich T. Improving the management of care for high-cost Medicaid patients. *Health Aff* 2007;26:1643–54.

31 Billings J, Dixon J, Mijanovich T, *et al*. Case finding for patients at risk of readmission to hospital: development of algorithm to identify high risk patients. *BMJ* 2006;333:327.

32 Coleman EA, Min S-joon, Chomiak A, *et al*. Posthospital care transitions: patterns, complications, and risk identification. *Health Serv Res* 2004;39:1449–65.

33 Krumholz HM, Chen YT, Wang Y, *et al*. Predictors of readmission among elderly survivors of admission with heart failure. *Am Heart J* 2000;139:72–7.

34 Morrissey EF, McElnay JC, Scott M, *et al*. Influence of drugs, demographics and medical history on hospital readmission of elderly patients. *Clin Drug Investig* 2003;23:119–28.

35 Smith DM, Norton JA, McDonald CJ. Nonelective readmissions of medical patients. *J Chronic Dis* 1985;38:213–24.

36 Smith DM, Katz BP, Huster GA, *et al*. Risk factors for nonelective Hospital readmissions. *J Gen Intern Med* 1996;11:762–4 http://www.ncbi.nlm.nih.gov/pubmed/9016426

37 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;182:551–7.

38 Burns R, Nichols LO. Factors predicting readmission of older general medicine patients. *J Gen Intern Med* 1991;6:389–93.

39 Evans RL, Hendricks RD, Lawrence KV, *et al*. Identifying factors associated with health care use: a hospital-based risk screening index. *Soc Sci Med* 1988;27:947–54.

40 Hasan O, Meltzer DO, Shaykevich SA, *et al*. Hospital readmission in general medicine patients: a prediction model. *J Gen Intern Med* 2010;25:211–9.

41 Boult C, Dowd B, McCaffrey D, *et al*. Screening elders for risk of hospital admission. *J Am Geriatr Soc* 1993;41:811–7.

42 Billings J, Blunt I, Steventon A, *et al*. Development of a predictive model to identify inpatients at risk of re-admission within 30 days of discharge (PARR-30). *BMJ Open* 2012;2:e001667–10.

43 Choudhry SA, Li J, Davis D, *et al*. A public-private partnership develops and externally validates a 30-day Hospital readmission risk prediction model. *Online J Public Health Inform* 2013;5:219.

44 Cotter PE, Bhalla VK, Wallis SJ, *et al*. Predicting readmissions: poor performance of the lace index in an older UK population. *Age Ageing* 2012;41:784–9.

45 Donzé J, Aujesky D, Williams D, *et al*. Potentially avoidable 30-day Hospital readmissions in medical patients: derivation and validation of a prediction model. *JAMA Intern Med* 2013;173:632–8.

46 Hebert C, Shivade C, Foraker R, *et al*. Diagnosis-specific readmission risk prediction using electronic health data: a retrospective cohort study. *BMC Med Inform Decis Mak* 2014;14:65.

47 Escobar GJ, Ragins A, Scheirer P, *et al*. Nonelective rehospitalizations and Postdischarge mortality: predictive models suitable for use in real time. *Med Care* 2015;53:916–23.

48 Yu S, Farooq F, van Esbroeck A, *et al*. Predicting readmission risk with institution-specific prediction models. *Artif Intell Med* 2015;65:89–96.

49 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;8:689–95.

50 Gildersleeve R, Cooper P. Development of an automated, real time surveillance tool for predicting readmissions at a community hospital. *Appl Clin Inform* 2013;4:153–69.

51 Kruse RL, Hays HD, Madsen RW. Risk factors for all-cause Hospital readmission within 30 days of hospital discharge. *JCOM* 2013;20:203–14.

52 Richmond DM. Socioeconomic predictors of 30-day Hospital readmission of elderly patients with initial discharge destination of home health care 2013.

53 Shulan M, Gao K, Moore CD. Predicting 30-day all-cause Hospital readmissions. *Health Care Manag Sci* 2013;16:167–75.

54 Lee EW. Selecting the best prediction model for readmission. *J Prev Med Public Health* 2012;45:259–66.

55 van Walraven C, Wong J, Forster AJ. Derivation and validation of a diagnostic score based on case-mix groups to predict 30-day death or urgent readmission. *Open Med* 2012;6:e90–100.

56 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:e80–90.

57 Iannuzzi JC, Chandra A, Kelly KN, *et al*. Risk score for unplanned vascular readmissions. *J Vasc Surg* 2014;59:1340–7.

58 Keyhani S, Myers LJ, Cheng E, *et al*. Effect of clinical and social risk factors on hospital profiling for stroke readmission: a cohort study. *Ann Intern Med* 2014;161:775–84.

59 Rana S, Tran T, Luo W, *et al*. Predicting unplanned readmission after myocardial infarction from routinely collected administrative hospital data. *Aust Health Rev* 2014;38:377–82.

60 Shahian DM, He X, O'Brien SM, *et al*. Development of a clinical registry-based 30-day readmission measure for coronary artery bypass grafting surgery. *Circulation* 2014;130:399–409.

61 Shams I, Ajorlou S, Yang K. A predictive analytics approach to reducing 30-day avoidable readmissions among patients with heart failure, acute myocardial infarction, pneumonia, or COPD. *Health Care Manag Sci* 2015;18:19–34.

62 Sharif R, Parekh TM, Pierson KS, *et al*. Predictors of early readmission among patients 40 to 64 years of age hospitalized for chronic obstructive pulmonary disease. *Ann Am Thorac Soc* 2014;11:685–94.

63 Lucas DJ, Haider A, Haut E, *et al*. Assessing readmission after general, vascular, and thoracic surgery using ACS-NSQIP. *Ann Surg* 2013;258:430–9.

64 Wallmann R, Llorca J, Gómez-Acebo I, *et al*. Prediction of 30-day cardiac-related-emergency-readmissions using simple administrative hospital data. *Int J Cardiol* 2013;164:193–200.

65 Wasfy JH, Rosenfield K, Zelevinsky K, *et al*. A prediction model to identify patients at high risk for 30-day readmission after percutaneous coronary intervention. *Circ Cardiovasc Qual Outcomes* 2013;6:429–35.

66 Krumholz HM, Lin Z, Drye EE, *et al*. An administrative claims measure suitable for profiling Hospital performance based on 30-day all-cause readmission rates among patients with acute myocardial infarction. *Circ Cardiovasc Qual Outcomes* 2011;4:243–52.

67 Betihavas V, Frost SA, Newton PJ, *et al*. An absolute risk prediction model to determine unplanned cardiovascular readmissions for adults with chronic heart failure. *Heart Lung Circ* 2015;24:1068–73.

68 Di Tano G, De Maria R, Gonzini L, *et al*. The 30-day metric in acute heart failure revisited: data from IN-HF outcome, an Italian nationwide cardiology registry. *Eur J Heart Fail* 2015;17:1032–41.

69 Huynh QL, Saito M, Blizzard CL, *et al*. Roles of nonclinical and clinical data in prediction of 30-day rehospitalization or death among heart failure patients. *J Card Fail* 2015;21:374–81.

70 Raposeiras-Roubín S, Abu-Assi E, Cambeiro-González C, *et al*. Mortality and cardiovascular morbidity within 30 days of discharge following acute coronary syndrome in a contemporary European cohort of patients: how can early risk prediction be improved? the six-month grace risk score. *Rev Port Cardiol* 2015;34:383–91.

71 Fleming LM, Gavin M, Piatkowski G, *et al*. Derivation and validation of a 30-day heart failure readmission model. *Am J Cardiol* 2014;114:1379–82.

72 Eapen ZJ, Liang L, Fonarow GC, *et al*. Validated, electronic health record deployable prediction models for assessing patient risk of 30-day rehospitalization and mortality in older heart failure patients. *JACC Heart Fail* 2013;1:245–51.

73 Zai AH, Ronquillo JG, Nieves R, *et al*. Assessing Hospital readmission risk factors in heart failure patients enrolled in a telemonitoring program. *Int J Telemed Appl* 2013;2013:305819.

74 Au AG, McAlister FA, Bakal JA, *et al*. Predicting the risk of unplanned readmission or death within 30 days of discharge after a heart failure hospitalization. *Am Heart J* 2012;164:365–72.

75 Watson AJ, O'Rourke J, Jethwani K, *et al*. Linking electronic health record-extracted psychosocial data in real-time to risk of readmission for heart failure. *Psychosomatics* 2011;52:319–27.

76 Mather JF, Fortunato GJ, Ash JL, *et al*. Prediction of pneumonia 30-day readmissions: a single-center attempt to increase model performance. *Respir Care* 2014;59:199–208.

77 Lindenauer PK, Normand S-LT, Drye EE, *et al*. Development, validation, and results of a measure of 30-day readmission following hospitalization for pneumonia. *J Hosp Med* 2011;6:142–50.

78 Shadmi E, Flaks-Manov N, Hoshen M, *et al*. Predicting 30-day readmissions with preadmission electronic health record data. *Med Care* 2015;53:283–9.

79 Tsui E, Au SY, Wong CP, *et al*. 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;21:46–56.

80 Donzé J, Lipsitz S, Schnipper JL. Risk factors for potentially avoidable readmissions due to end-of-life care issues. *J Hosp Med* 2014;9:310–4.

81 He D, Mathews SC, Kalloo AN, *et al*. Mining high-dimensional administrative claims data to predict early hospital readmissions. *J Am Med Inform Assoc* 2014;21:272–9.

82 Taha M, Pal A, Mahnken JD, *et al*. Derivation and validation of a formula to estimate risk for 30-day readmission in medical patients. *Int J Qual Health Care* 2014;26:271–7.

83 Zapatero A, Barba R, Marco J, *et al*. Predictive model of readmission to internal medicine wards. *Eur J Intern Med* 2012;23:451–6.

84 Singal AG, Rahimi RS, Clark C, *et al*. An automated model using electronic medical record data identifies patients with cirrhosis at high risk for readmission. *Clin Gastroenterol Hepatol* 2013;11:1335–41.

85 Volk ML, Tocco RS, Bazick J, *et al*. Hospital readmissions among patients with decompensated cirrhosis. *Am J Gastroenterol* 2012;107:247–52.

86 Perkins RM, Rahman A, Bucaloiu ID, *et al*. Readmission after hospitalization for heart failure among patients with chronic kidney disease: a prediction model. *Clin Nephrol* 2013;80:433–40.

87 Nijhawan AE, Clark C, Kaplan R, *et al*. An electronic medical record-based model to predict 30-day risk of readmission and death among HIV-infected inpatients. *J Acquir Immune Defic Syndr* 2012;61:349–58.

88 Whitlock TL, Tignor A, Webster EM, *et al*. A scoring system to predict readmission of patients with acute pancreatitis to the hospital within thirty days of discharge. *Clin Gastroenterol Hepatol* 2011;9:175–80.

89 Taber DJ, Palanisamy AP, Srinivas TR, *et al*. Inclusion of dynamic clinical data improves the predictive performance of a 30-day readmission risk model in kidney transplantation. *Transplantation* 2015;99:324–30.

90 Lawson EH, Hall BL, Louie R, *et al*. Identification of modifiable factors for reducing readmission after colectomy: a national analysis. *Surgery* 2014;155:754–66.

91 Iannuzzi JC, Fleming FJ, Kelly KN, *et al*. Risk scoring can predict readmission after endocrine surgery. *Surgery* 2014;156:1432–8. discussion 1438-40.

92 Mesko NW, Bachmann KR, Kovacevic D, *et al*. Thirty-day readmission following total hip and knee arthroplasty - a preliminary single institution predictive model. *J Arthroplasty* 2014;29:1532–8.

93 Moore L, Stelfox HT, Turgeon AF, *et al*. Derivation and validation of a quality indicator for 30-day unplanned Hospital readmission to evaluate trauma care. *J Trauma Acute Care Surg* 2014;76:1310–6.

94 Graboyes EM, Liou T-N, Kallogjeri D, *et al*. Risk factors for unplanned Hospital readmission in otolaryngology patients. *Otolaryngol Head Neck Surg* 2013;149:562–71.

95 Vigod SN, Kurdyak PA, Seitz D, *et al*. READMIT: a clinical risk index to predict 30-day readmission after discharge from acute psychiatric units. *J Psychiatr Res* 2015;61:205–13.

96 Kuhn M, Johnson K. Applied predictive modeling 2013.

97 Wiens J, Saria S, Sendak M, *et al*. Do no harm: a roadmap for responsible machine learning for health care. *Nat Med* 2019;25:1337–40.

98 Ancker JS, Kim M-H, Zhang Y, *et al*. The potential value of social determinants of health in predicting health outcomes. *J Am Med Inform Assoc* 2018;25:1109–10.

99 National Quality Forum. Measure evaluation criteria, 2012. Available: http://www.qualityforum.org/Measuring_Performance/Submitting_Standards/Measure_Evaluation_Criteria.aspx

100 Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. *Adv Neural Inf Process Syst* 2017:4765–74.

101 Kripalani S, Theobald CN, Anctil B, *et al*. Reducing Hospital readmission rates: current strategies and future directions. *Annu Rev Med* 2014;65:471–85.

102 Schneider EB, Hyder O, Brooke BS, *et al*. Patient readmission and mortality after colorectal surgery for colon cancer: impact of length of stay relative to other clinical factors. *J Am Coll Surg* 2012;214:390–8.

103 Centers for Medicare & Medicaid Services. CMS.gov: Hospital readmissions reduction program (HRRP), 2013. Available: https://www.cms.gov/medicare/medicare-fee-for-service-payment/acuteinpatientpps/readmissions-reduction-program.html

104 Goldfield NI, McCullough EC, Hughes JS, *et al*. Identifying potentially preventable readmissions. *Health Care Financ Rev* 2008;30:75–91.

105 Walsh C, Hripcsak G. The effects of data sources, cohort selection, and outcome definition on a predictive model of risk of thirty-day Hospital readmissions. *J Biomed Inform* 2014;52:418–26.

106 Lipton ZC. *The Mythos of model Interpretability*, 2016.

107 Austin PC, Brunner LJ. Inflation of the type I error rate when a continuous confounding variable is categorized in logistic regression analyses. *Stat Med* 2004;23:1159–78.

108 Steyerberg EW, Harrell FE. Prediction models need appropriate internal, internal-external, and external validation. *J Clin Epidemiol* 2016;69:245–7.

109 Austin PC, Steyerberg EW. Events per variable (EPV) and the relative performance of different strategies for estimating the out-of-sample validity of logistic regression models. *Stat Methods Med Res* 2017;26:796–808.

110 Molinaro AM, Simon R, Pfeiffer RM. Prediction error estimation: a comparison of resampling methods. *Bioinformatics* 2005;21:3301–7.

111 Harrell FE. *Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis*. 2nd Editio. Cham: Springer International Publishing, 2015.

112 Steyerberg EW, Vickers AJ, Cook NR, *et al*. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology* 2010;21:128–38.

113 Crowson CS, Atkinson EJ, Therneau TM. Assessing calibration of prognostic risk scores. *Stat Methods Med Res* 2016;25:1692–706.

114 Davis SE, Greevy RA, Fonnesbeck C, *et al*. A nonparametric updating method to correct clinical prediction model drift. *J Am Med Inform Assoc* 2019;26:1448–57.