Predicting Readmission at Early Hospitalization Using Electronic Clinical Data

An Early Readmission Risk Score

Ying P. Tabak, PhD*, Xiaowu Sun, PhD*, Carlos M. Nunez, MD*,† Vikas Gupta, PharmD, BCPS,* and Richard S. Johannes, MD, MS*,‡

Background: Identifying patients at high risk for readmission early during hospitalization may aid efforts in reducing readmissions. We sought to develop an early readmission risk predictive model using automated clinical data available at hospital admission.

Methods: We developed an early readmission risk model using a derivation cohort and validated the model with a validation cohort. We used a published Acute Laboratory Risk of Mortality Score as an aggregated measure of clinical severity at admission and the number of hospital discharges in the previous 90 days as a measure of disease progression. We then evaluated the administrative data–enhanced model by adding principal and secondary diagnoses and other variables. We examined the c-statistic change when additional variables were added to the model.

Results: There were 1,195,640 adult discharges from 70 hospitals with 39.8% male and the median age of 63 years (first and third quartile: 43, 78). The 30-day readmission rate was 11.9% (n = 142,211). The early readmission model yielded a graded relationship of readmission and the Acute Laboratory Risk of Mortality Score and the number of previous discharges within 90 days. The model c-statistic was 0.697 with good calibration. When administrative variables were added to the model, the c-statistic increased to 0.722.

Conclusions: Automated clinical data can generate a readmission risk score early at hospitalization with fair discrimination. It may have applied value to aid early care transition. Adding administrative data increases predictive accuracy. The administrative data–enhanced model may be used for hospital comparison and outcome research.

Key Words: readmission risk score, electronic health record (EHR), laboratory data, predictive model, decision support, Clinical Classification System (CCS), Comorbidity Software (CS)

(Med Care 2017;55: 267–275)

Readmission shortly after discharge is associated with clinical and financial burden to patients and society.1 It is used as a publicly reported metric with reimbursement implications to hospitals by the Centers for Medicare and Medicaid Services (CMS).2 For hospitals with excess risk-standardized readmission rates, CMS started lowering the reimbursement in FY2013. As of FY2015, the penalty is up to 3% lower reimbursement for certain clinical conditions [acute myocardial infarction (AMI), heart failure (HF), pneumonia, chronic obstructive pulmonary disease, total hip arthroplasty, and total knee arthroplasty].

Many factors are associated with readmissions. Some are intrinsic, attributable to the reduced patient reserve due to disease progression and severity at each admission.1,3–5 Some may relate to the clinical planning and care coordination while patients are still in the hospital. Others may relate to postdischarge care and other social factors.6,7

Current readmission risk adjustment models mostly rely on administrative data submitted to the payers after patients are discharged.8–10 Models relying on administrative data cannot be applied to the real-time patient care environment because the discharge diagnoses are not available until after patients are discharged. Most current models incorporated variables for comorbidity based on secondary diagnoses and use of prior medical services, but lack clinical data to assess the impact of physiological function or acute illness severity.1,8–10 The limited number of studies that attempted to incorporate physiological variables focused on a predefined clinical condition such as congestive HF.11 Such disease-specific models have applied value when the electronic health.

From the *Medical Informatics, Becton, Dickinson and Company; †The Biomedical Informatics Research Center at San Diego State University, San Diego, CA; and |Harvard Medical School and Brigham and Women’s Hospital, Boston, MA.
The preliminary data were presented in part at the Academy Health Annual Research Meeting, June 2016, Boston, MA.
The authors are current employees of Becton, Dickinson & Company.

Reprints: Ying P. Tabak, PhD, Scientific Research/Biostatistics, Medical Informatics, Becton Dickinson & Company, 3750 Torrey View Court, San Diego, CA 92130, E-mail: ying.tabak@bd.com.

Supplemental Digital Content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal’s Website, www.lww-medicalcare.com.

Copyright © 2017 The Author(s). Published by Wolters Kluwer Health, Inc. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

ISSN: 0025-7079/17/5503-0267

Medical Care • Volume 55, Number 3, March 2017

www.lww-medicalcare.com | 267
record (EHR) system is highly advanced and able to correctly classify patients into specific disease groups at admission. However, unlike discharge diagnoses that are standardized using codes from the International Classification of Diseases (ICD), admission diagnoses are not standardized across US hospitals. Often at the time of admission, patients display multiple signs and symptoms suggesting more than one clinical condition, which makes it difficult to unambiguously classify them into a single clinical category.

From the hospital comparison perspective, while it is helpful to assess readmission rates for patient groups with specific conditions, these conditions account for only a small proportion of total readmissions. According to the nationally representative data reported by the Agency for Healthcare Research and Quality (AHRQ), the 30-day all-cause readmission volume of AMI, HF, and pneumonia in 2012 accounted for just 9.8% of total readmission volume. In fact, the CMS has begun exploring hospital-wide (all-condition) 30-day risk-standardized readmission measure and risk adjustment models. Hence, understanding the full picture of hospital readmission may have practical implications to hospitals, all payers, and the governmental agencies.

The objective of this study was 2-fold. First, we sought to develop an early readmission risk score (RRS) for all hospitalized adult patients, using clinical data at the time of admission, that were widely available in the EHR systems. Second, we expanded the early RRS model to include administrative data that are available after patients are discharged. We hypothesized that (1) clinical severity and frequency of recent hospital stays, captured in the real-time EHR system, increase the risk of readmission; (2) type of diseases and comorbidities captured in the administrative data may enhance the predictive accuracy. For applications, the early RRS may serve as a potential real-time decision support tool for clinicians to target high-risk patients for discharge planning and other readmission preventive interventions. The administrative data enhanced postdischarge model may be used for retrospective interhospital comparison of risk adjusted readmission rates.

**METHODS**

**Data**

We used one of the Clinical Research Databases from Becton Dickinson & Company (Franklin Lakes, NJ). This database has been used for research for the past 2 decades and the data collection system has been fully described elsewhere. We used data from the EHR systems of 70 acute care hospitals for all consecutively hospitalized adult patients (age 18 years or older) from 2006 through 2008. The laboratory data included numeric laboratory test results and collection time. A total of 96% of patients had laboratory data on the day of admission. For patients with multiple laboratory assessments on the admission day, we used the first reported value. For patients who did not have laboratory data on admission day, the laboratory data on the next day were used. Missing laboratory test results were treated as not clinically indicated and therefore were grouped into the reference group, an approach used by APACHE and other researchers.

The database also included hospital administrative data comprising demographics, admission and discharge date, discharge disposition, principal diagnosis, and secondary diagnosis codes.

**Outcome**

The outcome variable was all-cause readmission within 30 days from index hospital discharge. Each live index hospital discharge date served as the starting date for the 30-day readmission window. A patient could have multiple admissions during the study period. The analytic unit was an inpatient admission episode. This floating hospital discharge index approach is used by the Healthcare Cost and Utilization Project, a widely used nation-wide database from the AHRQ, as well as other researchers.

**Readmission Risk Candidate Variables**

For the first model (RRS), we restricted the candidate variables to information in the EHR system that were available at or soon after the hospital admission to generate a predictive model applicable shortly after patients were hospitalized. We opted to use a published Acute Laboratory Risk of Mortality Score (ALaRMS) as an aggregated measure of clinical severity. ALaRMS assesses the clinical severity of hospitalized patients using demographics and admission laboratory test results. It included weighted age, sex, and a total of 23 numeric laboratory test results: (a) serum chemistry: albumin, aspartate transaminase, alkaline phosphatase, blood urea nitrogen, calcium, creatinine, glucose, potassium, sodium, and total bilirubin; (b) hematology and coagulation parameters: bands, hemoglobin, partial thromboplastin time, prothrombin time international normalized ratio, platelets, and white blood cell count; (c) arterial blood gas: partial pressure of carbon dioxide, partial pressure of oxygen, and pH value; (d) cardiac markers: brain natriuretic peptide, creatine phosphokinase MB, probrain natriuretic peptide, and troponin I.

We included the number of admissions in the previous 90 days as the second dimension of readmission risk. The rationale is that both progression of disease and disease severity reduce patients’ intrinsic reserve. The more frequent hospitalization, the more likely they will be readmitted. This variable also takes into account the nature of many chronic illnesses that can be managed but not necessarily cured. Hence, a certain rate of readmission may be expected.

For the administrative data–enhanced model, in addition to the RRS, we tested the Medicaid status as a surrogate for low social economic status, length of stay, admission source, discharge disposition, and discharge diagnoses of index hospital discharge. We used the AHRQ’s principal diagnosis-based Clinical Classifications Software (CCS) and secondary diagnosis-based Comorbidity Software (CS) as standard disease classifications. The CCS collapses over 14,000 ICD-9 diagnosis codes into 285 clinically meaningful categories. The CS grouped selected secondary diagnosis codes into 30 comorbidity categories.
Model Development and Validation

Early Readmission Risk Model

We randomly split study population into 70% as the derivation cohort and 30% as the validation cohort. To derive the early readmission risk model, we fit a multivariable logistic regression model. We first conducted univariable analysis on the relationship of the ALaRMS score (the aggregated clinical severity) and the readmission rate. We examined the distribution of the ALaRMS scores and the corresponding readmission rates and categorized the continuous variable accordingly for easier application. We fit logistic regression models to compare the model fit. Similarly, we examined and categorized the number of inpatient discharges during the 90 days before the index discharge and its corresponding readmission rate. Then, we added the categorized number of prior 90-day admissions into the logistic regression model. We converted the final multivariable logistic regression model to an integer score system (early RRS), using methods described by Sullivan et al.\(^2\) Specifically, we identified the variable with the smallest coefficient in the final multivariable model and applied it as the denominator. Then we divided each of the remaining regression coefficients in the model by this denominator and rounded the resulting quotient to the nearest whole number (integer), which formed the score weight for that variable. Then we calculated each person’s overall RRS by summing the points across all variables present. Converting model coefficients into a score system makes the risk adjustment model easy to understand and implement. We validated the RRS model using the validation cohort. We assessed model discrimination using the c-statistic.\(^2\)

Administrative Data–Enhanced Model

For the administrative data–enhanced model, we added candidate variables available from standard administrative data, including admission source, discharge disposition, payer, index hospital length of stay, principal diagnosis-based CCS, and secondary diagnosis-based CS. We used \(P<0.05\) as the model variable retention criterion. We examined the c-statistic change.

Sensitivity Analysis

We conducted sensitivity analysis to assess the predictive accuracy (c-statistic) of the RRS when applied to patients by different type of hospital characteristics (teaching status, number of beds, rural/urban) as well as patient categories (medical vs. surgical and discharge principal diagnosis-based disease categories).

RESULTS

Patient Characteristics

Table 1 showed that the derivation cohort consisted of 836,992 discharges and 99,386 (11.9%) readmissions. Overall, 39.8% were male, 81.7% white, and 36.6% covered by Medicare. The overall median age was 63 years, interquartile range was 43, 78. The admission laboratory test results showed that patients with abnormal results had higher readmission rates. Patients with readmissions had higher ALaRMS scores than those without readmissions. The mean (SD) of the ALaRMS score was 43.9 (18.9) versus 33.8 (19.8), \(P<0.0001\) and the median (interquartile range) was 43 (32, 55) versus 34 (20, 46), \(P<0.0001\), respectively. The number of discharges during the previous 90 days was associated with the readmission rate in a graded fashion. The patient characteristics for the validation cohort were similar.

Early Readmission Model and Risk Score

The early readmission model yielded a graded relationship of the ALaRMS score and the elevated risk of readmission (Table 2). Compared with those with ALaRMS \(\leq 10\), those with ALaRMS between 11 and 20 would have a 53% increase in risk of readmission (odds ratio, 1.53; 95% confidence intervals, 1.47–1.60). In general, the higher the ALaRMS score, the higher the readmission risk. The number of discharges during the previous 90 days was a significant predictor of readmission. Compared with those with no hospital discharges in the previous 90 days, one previous discharge would double the risk of readmission (odds ratio, 2.08; 95% confidence interval, 2.05–2.12). This early readmission risk model had a c-statistic of 0.697 (Table 3). When administrative data were added in the RRS model, the c-statistic improved to 0.722. The Hosmer-Lemeshow goodness-of-fit test for the RRS model indicated good calibration for both derivation and validation cohorts (Appendix A, Supplemental Digital Content 1, http://links.lww.com/MLR/B299).

The RRS prevalence distribution and corresponding readmission rates by derivation and validation cohorts were presented in Figure 1. There were 12.9% discharges falling into the lowest RRS group (RRS = 0), whose readmission rate was 3.2%. In contrast, for those in the top 2 highest RRS groups (RRS = 9 or 10+), over 50% were readmitted. With each point increase of RRS, there was a nearly linear increase of readmission risk. The Cochran-Armitage trending test was significant (\(P<0.0001\)).

Sensitivity Analysis

The sensitivity analysis revealed that the RRS model had similar predictive power for the patient populations in the teaching versus nonteaching, small versus large, or rural versus urban hospitals, with a c-statistic ranging from 0.690 to 0.706 (Table 4). The RRS model also displayed a similar c-statistic for medical versus surgical patients (0.692 vs. 0.690). When patients were segregated into subgroups by major disease categories, 14 of 17 (82%) categories had c-statistic above 0.65. 2 (12%) between 0.60 and 0.64, only 1 category (complications of pregnancy, childbirth, and the puerperium) was below 0.60 due to very low readmission rate of 2.9%. The Hosmer-Lemeshow model calibration charts for all sensitivity analysis are presented in Appendix B1–B3 (Supplemental Digital Content 2, http://links.lww.com/MLR/B300; Supplemental Digital Content 3, http://links.lww.com/MLR/B301; Supplemental Digital Content 4, http://links.lww.com/MLR/B302).
TABLE 1. Patient Characteristics by Derivation and Validation Cohorts

| Variables | Derivation Cohort | Validation Cohort |
|-----------|-------------------|-------------------|
|           | Discharges (n)    | Readmission [n (%)] | Discharges (n) | Readmission [n (%)] |
| Total     | 836,992           | 99,386 (11.9)      | 358,648        | 42,825 (11.9)       |
| Demographics (ALaRMS variables) | | | | |
| Age (y)   |                   |                   |                   | |
| < 30      | 97,861            | 4652 (4.8)        | 41,570           | 1967 (4.7)          |
| 30–34     | 42,766            | 2220 (5.2)        | 18,339           | 936 (5.1)           |
| 35–39     | 38,357            | 2656 (6.9)        | 16,356           | 1106 (6.8)          |
| 40–44     | 39,137            | 3720 (9.5)        | 16,860           | 1610 (9.5)          |
| 45–49     | 48,359            | 4993 (10.3)       | 20,660           | 2127 (10.3)         |
| 50–54     | 55,226            | 6039 (10.9)       | 23,406           | 2653 (11.3)         |
| 55–59     | 58,781            | 7096 (12.1)       | 25,377           | 3015 (11.9)         |
| 60–64     | 62,193            | 8119 (13.1)       | 26,502           | 2574 (13.6)         |
| 65–69     | 63,709            | 8641 (13.6)       | 27,711           | 3759 (13.6)         |
| > 70      |                   |                   | 29,380           | 2299 (13.6)         |
| Male      | 333,362           | 45,429 (13.6)     | 142,368          | 19,552 (13.7)       |
| Admission laboratory test result (ALaRMS variables) | | | | |
| Albumin ≤2.4 g/dL | 22,182            | 5069 (22.9)       | 9527            | 2180 (22.9)         |
| Albumin 2.5–2.7 g/dL | 22,280            | 4583 (20.6)       | 9585            | 1875 (19.6)         |
| Albumin 2.8–3 g/dL | 39,678            | 7203 (18.2)       | 17,008          | 3085 (18.1)         |
| Albumin 3.1–3.3 g/dL | 57,962            | 9306 (16.1)       | 24,837          | 4064 (16.4)         |
| AST 31–40 U/L | 55,558            | 8156 (14.7)       | 23,781          | 3496 (14.7)         |
| AST 41–60 U/L | 36,898            | 5879 (15.9)       | 15,632          | 2559 (16.4)         |
| AST 61–100 U/L | 20,455            | 3406 (16.7)       | 8628            | 1489 (17.3)         |
| AST > 100 U/L | 21,044            | 3100 (14.7)       | 9084            | 1417 (15.6)         |
| Total bilirubin ≤1.5–2 mg/dL | 15,857            | 2374 (15.0)       | 6735            | 1048 (15.6)         |
| Total bilirubin > 2.0 mg/dL | 14,697            | 2655 (18.1)       | 6350            | 1225 (19.3)         |
| Calcium ≤7.9 mg/dL | 33,632            | 5520 (16.4)       | 14,433          | 2448 (17.0)         |
| Calcium 8–8.4 mg/dL | 78,666            | 11,936 (15.2)     | 33,423          | 5177 (15.5)         |
| Calcium > 10.1 mg/dL | 33,872            | 4749 (14.0)       | 14,631          | 2023 (13.8)         |
| Creatinine >2.0 mg/dL | 58,987            | 13,993 (23.0)     | 25,294          | 5852 (23.1)         |
| pro BNP 8001–18,000 pg/dL | 1761              | 401 (22.8)        | 743             | 204 (27.5)          |
| pro BNP > 18,000 pg/dL | 1069              | 275 (14.7)        | 431             | 105 (24.4)          |
| BNP 1201–2400 pg/mL | 8072              | 1844 (22.8)       | 3569            | 811 (22.7)          |
| BNP > 2400 pg/mL | 5672              | 1387 (24.5)       | 2300            | 603 (26.2)          |
| Glucose ≤70 mg/dL | 11,754            | 1996 (17.0)       | 4935            | 902 (18.3)          |
| Glucose 136–165 mg/dL | 96,757            | 13,634 (14.1)     | 41,719          | 5787 (13.9)         |
| Glucose > 165 mg/dL | 121,390           | 19,585 (16.1)     | 51,977          | 8413 (16.2)         |
| K ≤5.3 mEq/L | 37,439            | 5537 (14.8)       | 16,170          | 2459 (15.2)         |
| K > 5.3 mEq/L | 26,881            | 5037 (18.7)       | 11,469          | 2134 (18.6)         |
| Na 144–145 mEq/L | 15,952            | 2464 (15.4)       | 34,311          | 6249 (18.2)         |
| pH arterial ≤7.2 | 22,182            | 4652 (4.8)        | 41,570          | 1967 (4.7)          |
| pH arterial 7.21–7.3 | 30,381            | 6039 (10.9)       | 23,406          | 2653 (11.3)         |
| pH arterial 7.31–7.35 | 8399              | 1378 (16.4)       | 3642            | 644 (17.7)          |
| pH arterial > 7.48 | 4870              | 903 (18.5)        | 2095            | 391 (18.7)          |
| [PO2] ≤50 mm Hg | 4175              | 535 (12.8)        | 1830            | 239 (13.1)          |
| [PO2] > 55 mm Hg | 33,078            | 6121 (18.5)       | 14,036          | 2635 (18.8)         |
| pCO2 arterial ≤140.1 mm Hg | 15,952            | 2464 (15.4)       | 34,311          | 6249 (18.2)         |
| pCO2 arterial > 35 mm Hg | 18,703            | 3020 (16.1)       | 8123            | 1277 (15.7)         |
| pCO2 arterial > 50 mm Hg | 12,589            | 2394 (19.0)       | 5523            | 1101 (19.9)         |
| PTT ≤22 | 9761              | 1438 (14.7)       | 4284            | 643 (15.0)          |
TABLE 1. Patient Characteristics by Derivation and Validation Cohorts (continued)

| Variables | Derivation Cohort | Validation Cohort |
|-----------|------------------|------------------|
|           | Discharges (n)   | Readmission [n (%)] | Discharges (n)   | Readmission [n (%)] |
| PT 45.1–55 | 11,250           | 2140 (19.0)       | 4889            | 885 (18.1) |
| PT > 55    | 14,294           | 2483 (17.4)       | 6119            | 1058 (17.3) |
| PT INR 1.11–1.4 | 84,807   | 14,079 (16.6)     | 36,368          | 6074 (16.7) |
| PT INR 1.41–2 | 29,670         | 5554 (18.7)       | 12,667          | 2404 (19.0) |
| PT INR 2.1–5 | 41,481          | 7559 (18.2)       | 17,920          | 3294 (18.4) |
| PT INR > 5 | 6005             | 1174 (19.6)       | 2618            | 512 (19.6)  |
| Bands 7%–13% | 27,731         | 4135 (14.9)       | 11,944          | 1714 (14.4) |
| Bands 14%–32% | 18,341         | 2773 (15.1)       | 7910            | 1143 (14.5) |
| Bands >32% | 7649             | 1019 (13.5)       | 3199            | 449 (14.0)  |
| Hemoglobin ≤ 10 g/dL | 87,775     | 15,846 (18.1)    | 37,840          | 6928 (18.3) |
| Hemoglobin >18 g/dL | 2099       | 217 (10.3)        | 986             | 91 (9.2)    |
| Platelets ≤115 10^9/L | 30,257   | 6143 (20.3)       | 13,140          | 2780 (21.2) |
| Platelets 115.1–150 10^9/L | 47,704    | 6443 (13.5)       | 20,783          | 2796 (13.5) |
| Platelets >120 10^9/L | 38,561    | 7148 (18.5)       | 16,443          | 2979 (18.1) |
| WBC ≤4.3 × 10^10/mm$^3$ | 27,984   | 5339 (19.1)       | 12,087          | 2361 (19.5) |
| WBC > 11.4–14.1 × 10^10/mm$^3$ | 141,312  | 15,795 (11.2)    | 60,388          | 6747 (12.2) |
| WBC > 14.2–19.8 × 10^10/mm$^3$ | 85,213   | 10,478 (12.3)     | 36,095          | 4302 (11.9) |
| WBC > 19.8 × 10^10/mm$^3$ | 28,494   | 4577 (16.1)       | 12,109          | 1941 (16.0) |
| Troponin I 0.05–0.1 or CPK MB 3–5 ng/mL | 101,888  | 16,007 (15.7)     | 43,370          | 6865 (15.8) |
| Troponin I 0.11–0.2 or CPK MB 6–10 ng/mL | 29,352   | 4828 (16.4)       | 12,710          | 2139 (16.8) |
| Troponin I 0.21–0.3 or CPK MB 11–34 ng/mL | 7683    | 1328 (17.3)       | 3334            | 552 (16.0)  |
| Troponin I 0.51–0.7 or CPK MB >34 ng/mL | 17,576   | 2841 (16.2)       | 7480            | 1246 (16.7) |
| No. discharges during previous 90 d |       |                 |                 |                 |
| 0          | 673,214          | 59,951 (8.9)      | 288,330         | 25,867 (9.0) |
| 1          | 118,160          | 23,639 (20.0)     | 50,777          | 10,144 (20.0) |
| 2          | 31,653           | 9,553 (30.2)      | 13,527          | 4037 (29.8)  |
| 3          | 9471             | 3732 (39.4)       | 4073            | 1666 (44.0) |
| ≥ 4        | 4494             | 2511 (55.9)       | 1941            | 1111 (57.2) |

Additional information (from administrative data)

| Race/ethnicity | Derivation Cohort | Validation Cohort |
|----------------|------------------|------------------|
| White          | 683,564          | 82,438 (12.1)    | 292,977         | 35,421 (12.1) |
| Black          | 66,070           | 8155 (12.3)      | 28,428          | 3546 (12.5)  |
| Other          | 87,358           | 8793 (10.1)      | 37,243          | 3858 (10.4)  |
| Payor          |                  |                  |                 |                 |
| Medicare       | 306,470          | 47,584 (15.5)    | 131,255         | 20,495 (15.6) |
| Medicaid       | 86,094           | 8117 (9.4)       | 36,777          | 3622 (9.8)   |
| Private or other | 444,428       | 43,685 (9.8)     | 190,616         | 18,708 (9.8) |

For each laboratory variable in the table, only abnormal levels that were used for the ALaRMS score were presented. The reference group (lowest risk group), which included patients with reference range test results or with no test results (missing values), was omitted from the table to save the space.

ALaRMS indicates Acute Laboratory Risk of Mortality Score; Alkaline phos, alkaline phosphatase; AST, aspartate transaminase; BNP, brain natriuretic peptide; BUN, blood urea nitrogen; CPK MB, creatine phosphokinase MB; pCO$_2$, partial pressure of carbon dioxide in arterial blood; PO$_2$, partial pressure of oxygen in arterial blood; pro-BNP, pro-brain natriuretic peptide; PT INR, prothrombin time international normalized ratio; PTT, partial thromboplastin time; WBC, white blood cell count.

**DISCUSSION**

We developed a 30-day RRS for adult patients admitted to acute care hospitals based on objective clinical parameters available at the time of hospital admission. To our knowledge, this is the largest (over 1 million discharges) clinical database consisting of laboratory test results used for readmission model development and validation. The large sample size resulted in a precise estimate as evidenced by the tight confidence intervals of risk factors. We found a graded relationship between the RRS and the readmission. Although the validity of this RRS needs to be tested prospectively in an environment with electronic health care data automation, this scoring system could have potential use for prioritizing discharge planning and coordinating care transition and management early in the hospital stay. Identifying high risk cases at the time of admission could enable clinicians to make connections with the patient and start to build a relationship during the patient’s hospital stay, which could make postdischarge engagements more effective when undertaken by somebody the patient knew and felt comfortable with. Given the average hospital length of stay is approximately 3–4 days, identifying patients with high readmission risk at the early hospital stay has pragmatic value.

Our RRS model using objective data alone has the c-statistic of 0.697, compared with 0.693–5 reported for readmission models that used both clinical and administrative data. When we added discharge diagnosis based administrative variables, the model c-statistic increased from 0.697 to 0.722. When we applied the RRS model to the 3 disease groups being reported by the CMS, the c-statistics of our models were higher than the CMS readmission models: 0.63 versus 0.60 (HF), 8 0.65 versus 0.63 (AMI), 9 0.68 versus 0.63 (pneumonia), 10 respectively, albeit the definition of disease groups and patient population may not be directly comparable. In more recent CMS efforts using administrative data alone to predict 30-day all-cause readmissions hospital-wide,
TABLE 2. Early Readmission Model and Risk Score

| Variable                              | Multivariable Adjusted Estimate | Multivariable Adjusted Odds Ratio (95% Confidence Interval) | RRS Risk Score |
|---------------------------------------|---------------------------------|------------------------------------------------------------|----------------|
| ALaRMS score                          |                                 |                                                            |                |
| 0–10                                  | Reference                       | 0                                                          |                |
| 11–20                                 | 0.43                            | 1.53 (1.47–1.60)                                           | 1              |
| 21–30                                 | 0.74                            | 2.10 (2.02–2.17)                                           | 2              |
| 31–40                                 | 1.00                            | 2.72 (2.63–2.81)                                           | 2              |
| 41–50                                 | 1.27                            | 3.55 (3.43–3.67)                                           | 3              |
| 51–60                                 | 1.47                            | 4.34 (4.19–4.49)                                           | 3              |
| 61–70                                 | 1.57                            | 4.82 (4.64–5.01)                                           | 4              |
| 71–80                                 | 1.56                            | 4.78 (4.56–5.00)                                           | 4              |
| 81–90                                 | 1.56                            | 4.78 (4.50–5.07)                                           | 4              |
| >90                                   | 1.46                            | 4.31 (4.02–4.62)                                           | 3              |

No. discharges during previous 90 d

| Reference                            | 0.73                            | 2.08 (2.05–2.12)                                           | 2              |
| 1                                    | 1.23                            | 3.43 (3.34–3.52)                                           | 3              |
| 2                                    | 1.65                            | 5.20 (4.98–5.43)                                           | 4              |
| ≥4                                   | 2.42                            | 11.20 (10.54–11.90)                                         | 6              |

All P-value <0.0001.

ALaRMS indicates Acute Laboratory Risk of Mortality Score; RRS, early readmission risk score.

The reported model c-statistics ranged from 0.604 to 0.676, which are lower than our RRS model as well as other reported models that used both clinical and administrative data. Perhaps more importantly, incorporating objective clinical data increases the clinical plausibility.

Many factors affect readmission. Some are intrinsic. With disease progression, a patient’s physiological reserve weakens and the clinical condition deteriorates to the point where the need for more frequent acute care becomes necessary. The more physiological reserves are depleted, the more likely the readmission. This depletion can be objectively assessed by the laboratory test results and other clinical parameters. Our finding that the ALaRMS score, which consists of age, sex, and abnormal laboratory test results, was correlated in a graded fashion with readmission risk demonstrates that patients’ physiological reserve and clinical severity play important roles in rehospitalization. The 23 numeric laboratory results assessed the functions of major organ systems. The further the deviation from the reference range, the more severe the clinical conditions. Incorporating laboratory data in the readmission model has advantages. They are objective and less prone to variations in subjective judgment. They are quantitative, allowing more accurate depiction of the graded relationship between severity and the outcome. They are widely automated in EHR systems, allowing potential automated real-time application of algorithms.

The finding that admission frequency during the previous three months is a strong predictor of readmission is plausible, given that the more frequent hospitalization may potentially reflect disease progression, continued physiological decline, or near the end of life. This is consistent with previous studies. It is also corroborated with the findings that the health care expenditure for the final year of life accounts for approximately $40,000 per Medicare decedent, and the 5% of Medicare patients who die each year account for nearly 30% of payments.

Our study found that laboratory values, a measure of physiological status and clinical stability, predict readmission. This suggests the possibility that rapid improvement in laboratory values during hospitalization might lead to clinical and physiological stability. Those patients whose care process results in the largest improvements in the shortest time might have lower readmission risk. Consequently, improved clinical stability may be translated into fewer readmissions. Hence, clinical severity as indicated by deranged laboratory test results and disease progression could potentially be modifiable by better care, which could lead to lower readmission rates.

For retrospectively collected postdischarge administrative data, we found principal diagnosis-based disease group variables enhanced the model discrimination. Certain clinical conditions are more likely to require rehospitalization than others. For example, severe systemic or chronic diseases affecting major organ systems such as infections, neoplasms, circulatory, respiratory, endocrine, or other systems might be more likely to be associated with readmissions. Hence, adding the principal diagnosis-based disease groups improves model discrimination. Secondary diagnosis-based comorbidities added a small improvement (from 0.715 to 0.722) above and beyond what have been in the model. The small impact of comorbidity may be due to the objective assessment of physiological function by the clinical laboratory data that have already accounted for comorbidity to a large extent. This finding is consistent with a previous study that found limited contribution of comorbidity when clinical data are used in the model. As administrative variables are available after patients are discharged as standard billing data, the enhanced model may be used for retrospective hospital comparison and outcome research.

Limitations

We used all-cause readmission within 30 days as the outcome. We did not attempt to differentiate avoidable versus nonavoidable readmission. “Avoidable” is difficult to define. The SQLape, the ICD-10 code based system...
attempts to identify avoidable readmissions has been developed in Switzerland, but it was not validated in the United States. Studies on the “avoidable” readmission reported a wide range: 5.0%–78.9%. The large range may indicate the variation of the “avoidable” readmission definitions. Our conceptualization builds on the assumption that with the disease progression, patients’ physiological reserves weaken, which increases probability of readmission. Hence, we “expect” certain admission rates based on a given patient population. Adjusting physiological severity and primary diseases of major organ system may be more practical and fair than trying to define “avoidable” consistently in terms of implementing a hospital comparison metric. Likewise, we used existing data which did not contain information allowing us to identify “planned” readmissions. Continued effort in differentiating planned versus unplanned readmissions may further enhance the study in this field. For example, it may be possible that some “planned” readmissions may be identifiable if EHR systems get more integrated, in which the scheduling systems can specify planned readmissions.

Our patient population comprised consecutive patients from over 70 hospitals. Although this is an improvement in patient diversity, it may not represent the US patient population because our hospitals were primarily from the northeastern region. Some studies found that socioeconomic status (SES) or admission or discharge to a skilled nursing facility (SNF) or long-term care (LTC) facility to be associated with readmission. We assessed this using Medicaid as a surrogate for the low SES. The 3 variables all together added a 0.001 to the c-statistic (from 0.697 to 0.698) above the RRS model. This small improvement in model discrimination was in line with other studies. In our model, the small increase of cumulative c-statistic was also supported by a logistic model including just these three variables (Medicaid, SNF/LTC admission source, and SNF/LTC discharge disposition) without any other covariates or RRS, which yielded a c-statistic of 0.54, indicating low discrimination of these 3 variables in our study population. In addition, there is controversy regarding the use of SES as a risk adjuster for hospital comparison because the concern that the higher readmission of the low SES patients may be associated lower quality of care provided to the disadvantaged patient population.

Unlike predicting the mortality risk, which showed highly accurate predictive power when objective clinical data are carefully crafted, predicting readmission remains complex. In addition to the intrinsic clinical risk of the patients which can be measured objectively, extrinsic risk factors, including care transition planning and other social and environmental factors may play important roles. These external risk factors can be difficult to collect electronically without integrated health care systems but can be incorporated for quality improvement purposes as they become available electronically.

For the purpose of risk adjusted public reporting metrics, it may not even be desirable to adjust for process variables that are under the control of care providers. If substandard care processes, such as poor care coordination, are associated with higher readmission rates, then adding these variables in the model would give more credit (expecting more readmissions).
to those hospitals/systems doing a lesser job than those who do a better job in coordinating care. Nevertheless, with increasing automation and connectivity of health care systems, factors that influence readmissions, such as discharge transition planning, postdischarge care, social support, might be captured that influence readmissions, such as discharge transition plan.

**CONCLUSIONS**

Automated clinical laboratory data can generate a RRS early at hospital admission with fair discrimination. It may be implemented and tested in an electronic health care system to aid early care transition planning. Administrative data–enhanced model, created by adding discharge diagnosis data, improves readmission predictive accuracy. It may be used for retrospective risk adjusted hospital comparison and outcome research.

**REFERENCES**

1. Kansagara D, Englander H, Salanitro A, et al. Risk prediction models for hospital readmission: a systematic review. *JAMA*. 2011;306:1688–1698.
2. Centers for Medicare and Medicaid Services. Readmissions reduction program. 2015. Available at: www.cms.gov/Medicare/medicare-fee-for-service-payment/acuteinpatientPPS/readmissions-reduction-program.html. Accessed August 18, 2015.

3. Donze 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;173:632–638.
4. van Walraven C, Dhallia 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;182:551–557.
5. Amarasinha R, Velasco F, Xie B, et al. An electronic medical record-based multicondition models to predict the risk of 30 day readmission or death among adult medicine patients: validation and comparison to existing models. *BMC Med Inform Decis Mak*. 2015;15:39.
6. Carter JA, Carr LS, Collins J, et al. STAAR: improving the reliability of care coordination and reducing hospital readmissions in an academic medical centre. *BMJ Innov*. 2015;1:75–80.
7. Coleman EA, Parry C, Chalmers S, et al. The care transitions intervention: results of a randomized controlled trial. *Arch Intern Med*. 2006;166:1822–1828.
8. Keenan PS, Normand SL, Lin Z, et al. An administrative claims measure suitable for profiling hospital performance on the basis of 30-day all-cause readmission rates among patients with heart failure. *Circ Cardiovasc Qual Outcomes*. 2008;1:29–37.
9. 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;4:243–252.
10. Lindenauer PK, Normand SL, Drye EE, et al. Development, validation, and results of a measure of 30-day readmission following hospitalization for pneumonia. *J Hosp Med*. 2011;6:142–150.
11. Amarasinha R, Moore BJ, Tabak VP, 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;48:981–988.
12. Agency for Healthcare Research and Quality Healthcare Cost & Utilization Project. Available at: http://hcupnet.ahrq.gov/HCUPnet.jsp?id=E2CFDB82B9591CEA&Form=SelDB&JS=Y&Action=%3E%3E%3E%3E%3E%3E&DB=NIS12. Accessed October 4, 2016.

13. Horwitz L, Partovian C, Lin Z, et al. Hospital-wide (all-condition) 30-day risk-standardized readmission measure. 2011. Available at: www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/MMS/downloads/MMSHospital-WideAll-ConditionReadmissionRate.pdf. Accessed August 24, 2015.

14. Tabak YP, Sun X, Johannes RS, et al. Mortality and need for mechanical ventilation in acute exacerbations of chronic obstructive pulmonary disease: development and validation of a simple risk score. Arch Intern Med. 2009;169:1595–1602.

15. Tabak YP, Johannes RS, Silber JH. Using automated clinical data for risk adjustment: development and validation of six disease-specific mortality predictive models for pay-for-performance. Med Care. 2007;45:789–805.

16. Tabak YP, Sun X, Derby KG, et al. Development and validation of a disease-specific risk adjustment system using automated clinical data. Health Serv Res. 2010;45(pt 1):1815–1835.

17. Tabak YP, Sun X, Nunez CM, et al. Using electronic health record data to develop inpatient mortality predictive model: Acute Laboratory Risk of Mortality Score (ALaRMS). J Am Med Inform Assoc. 2014;21:455–463.

18. Knaus WA, Wagner DP, Draper EA, et al. The APACHE III prognostic system. Risk prediction of hospital mortality for critically ill hospitalized adults. Chest. 1991;100:1619–1636.

19. Horwitz LI, Lin Z, Herrin J, et al. Association of hospital volume with readmission rates: a retrospective cross-sectional study. BMJ. 2015;350:h447.

20. Elixhauser A, Steiner C, Palmer L. Clinical Classifications Software (CCS), 2010. US Agency for Healthcare Research and Quality, 2010. Available at: www.hcup-us.ahrq.gov/toolssoftware/ccs/CCSUsersGuide.pdf. Accessed January 19, 2010.

21. Agency for Healthcare Research and Quality. Comorbidity Software, version 3.5. 2010. Available at: www.hcup-us.ahrq.gov/toolssoftware/comorbidity/comorbidity.jsp. Accessed June 3, 2011.

22. Sullivan LM, Massaro JM, D’Agostino RB Sr. Presentation of multivariate data for clinical use: the Framingham Study risk score functions. Stat Med. 2004;23:1631–1660.

23. Hosmer DW, Lemeshow S. Applied Logistic Regression, 2nd ed. New York, NY: John Wiley & Sons, Inc.; 2000.

24. Gozalo P, Plotzke M, Mor V, et al. Changes in Medicare costs with the growth of hospice care in nursing homes. N Engl J Med. 2015;372:1823–1831.

25. Riley GF, Lubitz JD. Long-term trends in Medicare payments in the last year of life. Health Serv Res. 2010;45:565–576.

26. Halfon P, Eggli Y, van Melle G, et al. Measuring potentially avoidable hospital readmissions. J Clin Epidemiol. 2002;55:573–587.

27. van Walraven C, Bennett C, Jennings A, et al. Proportion of hospital readmissions deemed avoidable: a systematic review. CMAJ. 2011;183: E391–E402.

28. Eapen ZJ, McCoy LA, Fonarow GC, et al. Utility of socioeconomic status in predicting 30-day outcomes after heart failure hospitalization. Circ Heart Fail. 2015;8:473–480.

29. Blum AB, Egorova NN, Sosunov EA, et al. Impact of socioeconomic status measures on hospital profiling in New York City. Circ Cardiovasc Qual Outcomes. 2014;7:391–397.

30. Bernheim SM. Measuring quality and enacting policy: readmission rates and socioeconomic factors. Circ Cardiovasc Qual Outcomes. 2014;7:350–352.

31. Krumholz HM, Bernheim SM. Considering the role of socioeconomic status in hospital outcomes measures. Ann Intern Med. 2014;161:833–834.