Potentially modifiable risk factors for 30-day unplanned hospital readmission preventive intervention—A data mining and statistical analysis

Peng Zhao and Illhoi Yoo
University of Missouri, USA;

Abstract
Unplanned hospital readmissions have a high prevalence and substantial healthcare costs. Preventive intervention during hospitalization holds the potential for reducing readmission risk. However, it is challenging to develop individualized interventions during hospitalization because the causes of readmissions have not been clearly known and because patients are heterogeneous. This work aimed to identify potentially modifiable risk factors of readmission to help clinicians better plan and prioritize interventions for different patient subgroups during hospitalization. We performed the analysis of associations between the changes of potentially modifiable risk factors and the change of readmission status with association rule mining and statistical methods. Twenty-nine risk factors were identified from the association rules, and twenty-five of them were potentially modifiable. The association rules with potentially modifiable risk factors can be recommended to different patient subgroups to support the development of customized readmission preventive interventions.

Keywords
patient readmission, risk factors, unplanned, 30-day, data mining

Introduction
Unplanned hospital readmissions have imposed a significant economic burden on American society. In 2011, the total hospital costs for 30-day hospital readmissions reached $41.3 billion. According to a 2019 report by the Healthcare Cost and Utilization Project (HCUP), 17.1% of...
Medicare beneficiaries experienced 30-day all-cause hospital readmissions in 2016, followed by Medicaid beneficiaries (13.7%), uninsured patients (11.8%), and patients with private insurance (8.6%). As a result, reducing unplanned hospital readmissions has captivated policymakers as a goal that improves health care quality and reduces costs. Since 2009, the Centers for Medicare and Medicaid Services (CMS) has been publicly reporting hospital-level 30-day risk-standardized readmission rates on the Hospital Compare website, which allows patients to compare hospitals with government ratings. In 2012, the Affordable Care Act implemented the Hospital Readmission Reduction Program (HRRP) to financially penalize hospitals with excessive readmission rates. As of January 2020, the HRRP has considered readmissions following acute myocardial infarction (AMI), chronic obstructive pulmonary disease (COPD), heart failure (HF), pneumonia (PN), coronary artery bypass graft (CABG), and total hip or knee arthroplasty (THA/TKA) in the calculation of readmission penalties due to high cost, substantial morbidity and mortality, and marked performance variations across hospitals. The percentage of penalized hospitals and the amount of penalty have increased since the implementation of the HRRP. In fiscal year 2020, 83% of 3129 HRRP eligible hospitals will be penalized with an estimated $563 million reduction in Medicare payments.

Existing unplanned hospital readmission (hereafter, referred to as “readmission”) reduction programs tend to focus on care transition and post-discharge interventions. However, one limitation of this approach is that they are mostly initiated near or after discharge when clinicians are no longer impactful on inpatient care. Preventive intervention during hospitalization is an under-explored area that holds the potential for reducing readmission risk. However, there are two challenges to deliver interventions during hospitalization. First, it is difficult to foresee the causes of readmissions because patients can be readmitted for many reasons. The CMS uses index admissions’ principal diagnosis codes and procedure codes to define the six HRRP eligible cohorts, but imposes no restriction on the causes of readmissions. According to an analysis of 217,767 index admissions with readmissions by Rosen et al. about 60% of readmissions have different principal diagnoses, different diagnosis-related groups (DRG), or different procedures from the precedent index admissions. Second, it is challenging for clinicians to find a universal solution to reduce their readmission risk because patients are heterogeneous. Even if they share the same principal diagnoses and procedures in index admissions, they may have different combinations of comorbidities and social factors. It has been found that direct or indirect complications of patients’ comorbidities are the top causes of readmissions. Social factors, such as demographics and socioeconomic status, can also influence readmission risk. Therefore, interventions should be customized for different patients based on their specific clinical and social factors.

Modifiable risk factors are risk factors that can be potentially controlled or treated by interventions. It has been shown that the management of modifiable risk factors was effective in reducing the risk of undesired outcomes. Several works have investigated the identification or management of modifiable readmission risk factors. Stephens et al. conducted a literature review and found that multi-faceted interventions and post-discharge clinic visits may be effective for reducing readmissions in children with medical complexity. Maddocks et al. purposed to use rehabilitation to modify COPD readmission risk. Kelly et al. reported that discharge to a skilled nursing facility was a modifiable risk factor for postoperative readmissions in major gastrointestinal resections. Brodsky et al. found that the lack of a primary care provider and the number of hospital admissions caused by vaso occlusive pain episodes in the previous year were modifiable readmission risk factors for adults with sickle cell disease. Singh et al. reported that obesity, comorbidities, medical complications, and hospital factors were modifiable readmission risk factors for THA. Garcia-Aymerich et al. showed that a higher level of usual physical activity was a modifiable protective factor of COPD readmission. Boraiah et al. obtained eight modifiable risk factors
by performing a regression analysis on modifiable THA/TKA risk factors reported by other studies and clinical experience. Most of the reported modifiable risk factors are not timely, such as discharge disposition, medical history in last year. In addition, all of these modifiable risk factors were identified by population-based regression analyses and were not customizable for patient subgroups with different demographic and clinical background. As a result, they cannot be used in the delivery of interventions during hospitalization.

In this work, we purposed to identify patterns with potentially modifiable risk factors and recommend them to different patient subgroups to support the development of customized interventions. To identify potentially modifiable risk factors, we investigated a novel method to compare different index admissions of the same patients. Compared to traditional logistic regression-based approach which can only study associations between risk factors and readmission status, this method allows us to further explore associations between changes of modifiable risk factors and the change of readmission status. It can also automatically identify patient subgroups based on demographic and clinical factors. This makes it possible to customize interventions for different patient subgroups. We focused on medical services for AMI, COPD, HF, and PN because they are targeted by the HRRP. Surgical services were not included in this analysis because surgical readmissions tend to be related to postoperative complications, which are undesired outcomes of medical care. Surgical readmissions can be potentially reduced by minimizing the risk of complications during the care process.

Materials and methods

Study design

This study was a retrospective analysis of clinical data. We identified pairs of index admissions of the same patients and recorded the element-wise difference of each diagnosis’ existence and the difference of readmission status in each pair. For example, a patient has a pair of index admissions 1 and 2. This patient has diagnosis A in both index admissions and diagnosis B in index admission 1 only. Index admission 1 was followed with a readmission, and index admission 2 was not. The change of diagnosis in these two index admissions was “diagnosis B: true (T) → false (F)” and the change of readmission status was “readmission → no readmission.” Each patient has one pair of index admissions. We looked for associations between changes of diagnoses’ existence and the change of readmission status.

Data source and inclusion criteria

We purchased the 2014 Nationwide Readmissions Database (NRD) from the Agency for Healthcare Research and Quality (AHRQ). The NRD is a discharge-level database available for each calendar year from 2010 to 2016 (as of November 2019). We chose the 2014 NRD because this data was the latest database including comorbidity information. The 2015 and 2016 NRDs do not contain this information. The 2014 NRD is about 12 GB with information from about 15 million discharges. It was derived from inpatient hospitalization data of 2048 hospitals in 22 states, which accounted for 51.2% of the U.S. population and 49.3% of all hospitalizations. Patients can be tracked across different hospitals within a state by the verified and de-identified patient linkage numbers.

We constructed four separate cohorts for AMI, COPD, HF, and PN patients with the following criteria: (1) We captured index admissions with a discharge month from January to November in 2014. Index discharges in December were excluded because December was the last month available in the data and 30-day unplanned readmissions cannot be tracked. Eligible index admissions
had a principal discharge diagnosis of AMI, COPD, HF, or PN based on the International classification of diseases, ninth revision, clinical modification (ICD-9-CM) codes used by the CMS. The length of stay was longer than 1 day. (2) Patients were older than 18 at admission. They were not transferred to other hospitals and were alive at discharge. They did not leave against medical advice. (3) The gap between index admission discharge and readmission was between one and 30 days (inclusive). Patients could be readmitted for all causes except elective services or other scheduled care according to version 4.0 (ICD-9-CM) of the planned readmission identification algorithm used by the CMS.

**Data preprocessing and transformation**

The 2014 NRD contains information about admissions, patient demographics, hospital characteristics, diagnoses, and procedures. Diagnoses are available in both ICD-9-CM codes and Clinical Classification Software (CCS) categories. We only kept the CCS-level diagnoses since ICD codes are too granular for this data mining analysis. They are represented as one principal diagnosis (DX1) and up to 29 secondary diagnoses (DX2 to DX30). We reshaped diagnoses into a sparse matrix by using the CCS-level diagnoses as attributes and values of true (T) and false (F) representing the presence and the absence of the corresponding diagnoses. We captured pairs of index admissions of the same patients. For each pair of index admissions, we took the element-wise differences of the readmission status and CCS-level diagnoses. We also extracted patients’ characteristics from the first index admission of each pair, including age, sex, primary insurance payer, type of the patient’s residing county, and median household income in the patient’s zip code. We only used one pair of index admissions of each patient. Attributes of the derived dataset are shown in Table 1.

**Association rule mining**

We performed association rule mining to unearth association’s rules between changes of diagnoses and the change of readmission status. Association rule mining is an unsupervised data mining approach to discover associations between itemsets in the form of “IF {antecedent} THEN {consequent}.” We used the Apriori algorithm to identify association rules with the consequent being the change of readmission status (e.g. “readmission → no readmission”). To ensure the association rules were interesting and non-trivial, we used support, confidence, and lift to filter out trivial association rules. The support is the frequency of a pattern occurring in all transactions ranging from 0 to 1. The confidence measures the percentage of transactions with the consequent given it
contains the antecedent. The lift is the ratio of the observed support to expected support given the antecedent and the consequent are independent. A lift greater than 1 indicates the antecedent and the consequent are dependent on each other. Because the standard association rule mining may generate numerous spurious rules, we performed Fisher’s exact test to measure the statistical significance of the positive correlation between the antecedent and the consequent. We used a significance level of 0.05. A \( p \)-value greater than 0.05 indicates a spurious association rule occurring by chance. For all association rule mining experiments, we used 0.001, 0.75, and 1 as the minimum support, minimum confidence, and minimum lift, respectively. We also assessed each rule’s medical soundness and only kept rules with positive associations between a diagnosis change of “T→F” or “F→T” and “readmission→no readmission” or “no readmission→readmission.”

**Results**

**Readmission rate and time gap**

We found 70419, 313837, 349548, and 431149 eligible index admissions for the AMI, COPD, HF, and PN cohorts with the readmission rate of 14.1%, 18.6%, 19.9%, and 14.4%, respectively. Figure 1 shows distributions of gaps (in days) between discharges and readmissions for the four cohorts. It can be seen that as time goes towards the end of the monitoring window (30 days), the number of readmissions decreases. This indicates that readmissions tend to occur early after discharge.
Patient characteristics

We performed a Chi-square analysis to test the association between patients’ demographic information and readmission status. All of the demographic variables are significantly associated ($p < 0.05$) with readmission risk for the four cohorts except sex of the HF cohort ($p=0.886$). The result is shown in Table A1 in the appendix. We further identified 853, 10820, 14343, and 11275 pairs of index admissions from the AMI, COPD, HF, and PN cohorts, respectively. Each pair of index admissions belongs to a unique patient. Table 2 shows their demographic information. It can be seen that the majority of patients are older than 65 and are Medicare beneficiaries. There are slightly more female patients. Most patients live in large central metropolitans and neighborhoods with a median household income (zip code level) less than $40,000.

Association rule mining

We obtained 108, 72, 81, and 58 eligible association rules for AMI, COPD, HF, and PN cohorts, respectively, after applying the rule quality criteria. From each rule, we extracted the diagnosis with changed status (i.e. “T→F” or “F→T”) in the antecedent and quantified the strength of its association with readmission by odds ratio (OR). We found 29 diagnoses with a significantly
### Table 3. Risk factors of readmission of the four cohorts.

| Category                          | Description (CCS-level)                                                                 | Cohort |
|-----------------------------------|----------------------------------------------------------------------------------------|--------|
|                                   |                                                                                        | AMI    |
|                                   |                                                                                        | COPD   |
|                                   |                                                                                        | HF     |
|                                   |                                                                                        | PN     |
| **Medical conditions**            |                                                                                       | Y\(^{35}\) |
|                                   | Chronic ulcer of skin                                                                  | Y\(^{35}\) |
|                                   | Acute and unspecified renal failure                                                    | Y\(^{35}\) |
|                                   | Diabetes mellitus without complication                                                 | Y\(^{35,37-40}\) |
|                                   | Fluid and electrolyte disorders                                                       | Y\(^{42}\) |
|                                   | Hemorrhoids                                                                            | Y      |
|                                   | Hyperplasia of prostate                                                                | Y\(^{37}\) |
|                                   | Hypertension with complications and secondary hypertension                              | Y\(^{37}\) |
|                                   | Infective arthritis and osteomyelitis (except that caused by tuberculosis or sexually transmitted disease) | Y\(^{37}\) |
|                                   | Intestinal obstruction without hemia                                                   | Y\(^{43}\) |
|                                   | Late effects of cerebrovascular disease                                                | Y\(^{43}\) |
|                                   | Other connective tissue disease                                                       | Y      |
|                                   | Other diseases of bladder and urethra                                                  | Y      |
|                                   | Other disorders of stomach and duodenum                                                | Y\(^{44}\) |
|                                   | Other gastrointestinal disorders                                                       | Y      |
|                                   | Other hereditary and degenerative nervous system conditions                            | Y\(^{44}\) |
|                                   | Other nutritional; endocrine; and metabolic disorders                                  | Y\(^{44}\) |
|                                   | Pancreatic disorders (not diabetes)                                                    | Y      |
|                                   | Pneumonia (except that caused by tuberculosis or sexually transmitted disease)        | Y\(^{35}\) |
|                                   | Retinal detachments; defects; vascular occlusion; and retinopathy                      | Y      |
|                                   | Rheumatoid arthritis and related disease                                               | Y\(^{45}\) |
|                                   | Septicemia (except in labor)                                                          | Y\(^{35}\) |
|                                   | Skin and subcutaneous tissue infections                                                | Y\(^{35}\) |
| **Mental conditions/substance-related disorders** |                                                                                       | Y\(^{45}\) |
|                                   | Schizophrenia and other psychotic disorders                                            | Y\(^{35}\) |
|                                   | Substance-related disorders                                                            | Y\(^{35}\) |
|                                   | Adjustment disorders                                                                   | Y\(^{46}\) |
| **Medical care complications/adverse effects** |                                                                                       | Y\(^{44}\) |
|                                   | Complication of device; implant or graft                                               | Y\(^{44}\) |
|                                   | Complications of surgical procedures or medical care                                   | Y\(^{44}\) |
|                                   | Adverse effects of medical care                                                        | Y      |
| **External causes**                |                                                                                        | Y      |
|                                   | Superficial injury; contusion                                                          | Y      |

CCS: Clinical classification software.
positive association with readmission (risk factor, OR > 1, p < 0.05) and showed them in Table 3, where “Y” indicates the diagnosis is a risk factor of the cohort. For risk factors that have been reported by other studies, we put citations after “Y.” It can be seen from Table 3 that only 5/29 (17%) of risk factors are shared by two cohorts and no other risk factor is commonly found in three cohorts. This indicates that the readmission patterns or risk factors of AMI, COPD, HF, and PN cohorts are different. In Tables A1 to A4 in the appendix, we showed readmission rates of the four cohorts with and without each risk factor, OR, and 95% confidence interval (CI) of OR. For each risk factor that can be potentially modified, we showed one association rule with the highest support in Tables 4 to 7.

Table 4. Association rules of the AMI cohort.

| Association rules                                                                 | Support | Confidence | Lift  | p-Value |
|----------------------------------------------------------------------------------|---------|------------|-------|---------|
| IF {hypertension with complications and secondary hypertension = T → F; sex = female; age = 65+} THEN {readmission → no readmission} | 0.041 (35/853) | 0.897 | 1.162 | 0.036 |
| IF {other connective tissue disease = T → F; age = 65+} THEN {readmission → no readmission} | 0.035 (30/853) | 0.909 | 1.177 | 0.037 |
| IF {pneumonia (except that caused by tuberculosis or sexually transmitted disease) = T → F; income = $1–$39,999} THEN {readmission → no readmission} | 0.026 (22/853) | 0.957 | 1.238 | 0.019 |
| IF {diabetes mellitus without complication = T → F; income = $66,000+} THEN {readmission → no readmission} | 0.020 (17/853) | 1.000 | 1.294 | 0.012 |
| IF {acute and unspecified renal failure = T → F; chronic kidney disease = T → T} THEN {readmission → no readmission} | 0.014 (12/853) | 1.000 | 1.294 | 0.044 |
| IF {hyperplasia of prostate = F → T; disorders of lipid metabolism = T → T; hypertension with complications and secondary hypertension = T → T} THEN {no readmission → readmission} | 0.005 (4/853) | 1.000 | 4.397 | 0.011 |
| IF {chronic ulcer of skin = F → T; congestive heart failure; non-hypertensive = T → T; location = large central metropolitan} THEN {no readmission → readmission} | 0.005 (4/853) | 0.800 | 3.518 | 0.011 |
| IF {substance-related disorders = F → T; coronary atherosclerosis and other heart disease = T → T; congestive heart failure; non-hypertensive = T → T} THEN {no readmission → readmission} | 0.004 (3/853) | 0.750 | 3.298 | 0.011 |
| IF {septicemia (except in labor) = F → T; insurance = Medicare} THEN {no readmission → readmission} | 0.004 (3/853) | 1.000 | 4.397 | 0.003 |
| IF {other disorders of stomach and duodenum = F → T; sex = female; insurance = Medicare} THEN {no readmission → readmission} | 0.004 (3/853) | 0.750 | 3.298 | 0.011 |
| IF {retinal detachments; defects; vascular occlusion; and retinopathy = F → T; chronic kidney disease = T → T; sex = female} THEN {no readmission → readmission} | 0.004 (3/853) | 0.750 | 3.298 | 0.003 |
Zhao and Yoo

Discussion

Analysis of changes

To our knowledge, this is the first study to analyze changes of modifiable risk factors of readmission with a combination of data mining and statistical methods. Our method has two advantages. First, by comparing the different causes (in ICD code) of index admissions of the same patients, we can better understand readmission through associations between potentially modifiable risk factors and readmission. Second, our method can provide more information about the potential effect of risk factor modification. The traditional logistic regression-based risk factor identification method only evaluates the association between the presence of a factor and the response (e.g. readmission). However, this does not necessarily mean the modification of this risk factor is associated with the reduction of readmission. Our method moves a step further to directly test the association between the changes of risk factors and the change of readmission status.

| Association rules | Support   | Confidence | Lift   | p-Value |
|-------------------|-----------|------------|-------|---------|
| IF {other disorders of stomach and duodenum = T → F; sex = male} THEN {readmission → no readmission} | 0.002 (23/10,820) | 0.958 | 1.225 | 0.021 |
| IF {other nutritional; endocrine; and metabolic disorders = T → F; chronic obstructive pulmonary disease and bronchiectasis = T → T; cancer of bronchus; lung = T → T; income = $1–$39,999} THEN {readmission → no readmission} | 0.002 (22/10,820) | 0.957 | 1.223 | 0.026 |
| IF {late effects of cerebrovascular disease = T → F; chronic obstructive pulmonary disease and bronchiectasis = T → T; location = medium metropolitan} THEN {readmission → no readmission} | 0.002 (18/10,820) | 1.000 | 1.279 | 0.012 |
| IF {intestinal obstruction without hernia = T → F; coronary atherosclerosis and other heart disease = T → T; sex = male} THEN {readmission → no readmission} | 0.001 (13/10,820) | 1.000 | 1.279 | 0.041 |
| IF {skin and subcutaneous tissue infections = T → F; screening and history of mental health and substance abuse codes = T → T; congestive heart failure; non-hypertensive = T → T} THEN {readmission → no readmission} | 0.001 (13/10,820) | 1.000 | 1.279 | 0.041 |
| IF {other hereditary and degenerative nervous system conditions = T → F; respiratory failure; insufficiency; arrest (adult) = T → T; income = $1–$39,999; sex = female; age = 65+} THEN {readmission → no readmission} | 0.001 (11/10,820) | 0.786 | 3.607 | <0.001 |
| IF {diabetes mellitus without complication = F → T; other aftercare = T → T; other nervous system disorders = T → T; chronic obstructive pulmonary disease and bronchiectasis = T → T} THEN {no readmission → readmission} | 0.001 (11/10,820) | 0.786 | 3.607 | <0.001 |

Table 5. Association rules of the COPD cohort.
Table 6. Association rules of the HF cohort.

| Association rules | Support       | Confidence | Lift   | p-Value |
|-------------------|---------------|------------|--------|---------|
| IF {schizophrenia and other psychotic disorders = T → F} THEN {readmission → no readmission} | 0.005 (77/14,343) | 0.865  | 1.097  | 0.045  |
| IF {fluid and electrolyte disorders = T → F; cancer of prostate = T → T} THEN {readmission → no readmission} | 0.003 (41/14,343) | 0.976  | 1.238  | 0.001  |
| IF {infective arthritis and osteomyelitis (except that caused by tuberculosis or sexually transmitted disease) = T → F} THEN {readmission → no readmission} | 0.002 (27/14,343) | 0.931  | 1.181  | 0.038  |
| IF {rheumatoid arthritis and related disease = T → F; coronary atherosclerosis and other heart disease = T → T; sex = female} THEN {readmission → no readmission} | 0.001 (19/14,343) | 1.000  | 1.268  | 0.011  |
| IF {pancreatic disorders (not diabetes) = T → F; disorders of lipid metabolism = T → T} THEN {readmission → no readmission} | 0.001 (18/14,343) | 1.000  | 1.268  | 0.014  |
| IF {other diseases of bladder and urethra = T → F; disorders of lipid metabolism = T → T; coronary atherosclerosis and other heart disease = T → T; chronic kidney disease = T → T; insurance = Medicare} THEN {readmission → no readmission} | 0.001 (15/14,343) | 1.000  | 1.268  | 0.028  |
| IF {adjustment disorders = T → F; location = large central metropolitan; insurance = Medicare} THEN {readmission → no readmission} | 0.001 (15/14,343) | 1.000  | 1.268  | 0.028  |

Table 7. Association rules of the PN cohort.

| Association rules | Support       | Confidence | Lift   | p-Value |
|-------------------|---------------|------------|--------|---------|
| IF {skin and subcutaneous tissue infections = T → F; pneumonia (except that caused by tuberculosis or sexually transmitted disease) = T → T; age = 65+; insurance = Medicare} THEN {readmission → no readmission} | 0.005 (57/11,275) | 0.905  | 1.114  | 0.035  |
| IF {retinal detachments; defects; vascular occlusion; and retinopathy = T → F; sex = male} THEN {readmission → no readmission} | 0.005 (51/11,275) | 0.911  | 1.121  | 0.035  |
| IF {acute and unspecified renal failure = T → F; cancer of prostate = T → T; insurance = Medicare} THEN {readmission → no readmission} | 0.003 (30/11,275) | 0.938  | 1.154  | 0.045  |
| IF {other gastrointestinal disorders = T → F; pneumonia (except that caused by tuberculosis or sexually transmitted disease) = T → T; cancer of breast = T → T} THEN {readmission → no readmission} | 0.002 (28/11,275) | 0.966  | 1.189  | 0.018  |
| IF {hemorrhoids = T → F; essential hypertension = T → T; location = large fringe metropolitan} THEN {readmission → no readmission} | 0.001 (15/11,275) | 1.000  | 1.231  | 0.044  |
Potentially modifiable risk factors

By comparing the presence and the absence of the same diagnosis in two different index admissions of the same patient, we can computationally test if the change of the diagnosis is positively associated with the change of readmission status. It is noteworthy that the absence of a diagnosis in the index admission does not necessarily mean the patient is free of the condition, since it is possible that the condition is inactive and does not impact inpatient care. For example, although diabetes mellitus is a nearly incurable chronic condition, this disease can be well-controlled by medications and lifestyle adjustments and changes.

There are four main categories of these 29 risk factors, including medical conditions, mental conditions or substance-related disorders, medical care complications or adverse events, and external causes. We do not consider medical care complications or adverse events as potentially modifiable risk factors because they will only occur after medical care as an outcome. Their risk should be anticipated and minimized during the care process. Similarly, external causes are not patient factors and are not under the control of hospitals. The remaining 25 medical and mental risk factors can be potentially modified by interventions, such as medication, surgery, and psychotherapy. According to the Charlson comorbidity index, five of these 25 risk factors have a mortality risk score greater than zero, including acute and unspecified renal failure, chronic ulcer of skin, diabetes mellitus without complication, late effects of cerebrovascular disease, and other connective tissue disease. These conditions are more severe and should receive more attention in the development of interventions.

Recommendation by association rules

Tables 4 to 7 show that the antecedent of each association rule is composed of patients’ non-modifiable factors (e.g. age, sex) and the change of a potentially modifiable risk factor. Because each association rule represents a patient subgroup, it can be used to recommend the modification of readmission risk factors for patients falling into the subgroup. For example, the association rule of “IF {Hypertension with complications and secondary hypertension = T → F; Sex = Female; Age = 65+} THEN {Readmission = No readmission}” in the AMI cohort can be recommended for female AMI patients older than 65 having the comorbidity of hypertension with complications and secondary hypertension. Besides the treatment of the principal condition of AMI, the intervention of hypertension should be prioritized for these patients to minimize the readmission risk. For each association rule, we performed Fisher’s exact test to ensure the association between the antecedent and the consequent is significant. Here, we provide potentially modifiable risk factors instead of intervention plans (e.g. medications, treatment pathways) because we believe that they should be developed by clinicians based on their medical judgments.

Limitations

Our work has three potential limitations. First, because this is a retrospective analysis, there is no way to control the confounding effects. Although the association rules are represented in “IF-THEN” patterns, the relationship indicates an association not causality. This study cannot replace controlled experiments, such as case-control and prospective cohort studies. However, our results can potentially offer a data-driven hypothesis to a randomized controlled trial and guide other studies to truly disclose the causal relationships between the identified potentially modifiable risk factors and readmission. Second, we used the CCS-level diagnosis in the analysis because ICD-9-CM codes were too granular for data mining analysis. As a result, some diagnoses can only provide
very general information, such as “other connective tissue disease.” Nevertheless, the purpose of this work is to provide information about potentially modifiable risk factors. Clinicians can potentially map patients’ problems and needs into these risk factors and design specific interventions. Third, we relied on ICD codes for identifying the changes of potentially modifiable risk factors. Because ICD codes are mainly designed for billing purposes instead of research, it is possible that some chronic diseases were not recorded in medical records if they did not affect the primary condition. However, this problem would not significantly impact on our findings since our data, the 2014 NRD (derived from inpatient hospitalization data of 2048 hospitals in 22 states) is large enough to considerably offset the coding issue since we measured the statistical significance of our findings.

Conclusions

In this work, we combined data mining and statistical methods to analyze the associations between the changes of potentially modifiable risk factors and the change of readmission status for Acute myocardial infarction (AMI), COPD, HF, and PN cohorts. Compared to other studies,\textsuperscript{21–26} which used the traditional logistic regression method to study associations between risk factors and readmission status, our approach moved a step further to analyze associations between changes of modifiable risk factors and the change of readmission status. It is imperative to identify modifiable risk factors associated with readmission because those risk factors can be used to predict and reduce readmissions. In addition, because each identified association rule represents a patient subgroup, clinicians can use it to customize interventions for patients falling in the subgroup. From the association rules, we identified 25 potentially modifiable risk factors of readmission. Our results would facilitate clinical research to further understand the causes of readmission.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Illhoi Yoo \textsuperscript{1} https://orcid.org/0000-0003-2536-3660

Supplemental material

Supplemental material for this article is available online.

References

1. Hines AL, Barrett ML, Jiang J, et al. Conditions with the largest number of adult hospital readmissions by payer, 2011. \textit{Healthc Cost Util Proj} 2014; 363: 1–10.
2. Bailey MK, Weiss AJ, Barrett ML, et al. Characteristics of 30-day all-cause hospital readmissions, 2010–2016, 2019, https://www.hcupus.ahrq.gov/reports/statbriefs/sb248-Hospital-Readmissions-2010-2016.jsp (2019, accessed 1 December 2019).
3. Hansen LO, Young RS, Hinami K, et al. Interventions to reduce 30-day rehospitalization: a systematic review. \textit{Ann Intern Med} 2011; 155: 520–528.
4. Medicare. Hospital compare—unplanned hospital visits, https://www.medicare.gov/hospitalcompare/About/HospitalReturns.html (2009), accessed 1 December 2012).
5. Medicare. Hospital inpatient quality reporting program, https://www.cms.gov/medicare/quality-initiatives-patient-assessment-instruments/hospitalqualityinitits/hospitalrhqadapu.html (2003, accessed 1 December 2012).

6. The 111th United States Congress. The Patient Protection and Affordable Care Act, https://www.govinfo.gov/content/pkg/PLAW-111publ148/pdf/PLAW-111publ148.pdf (2010, accessed 1 December 2012).

7. Medicare. Readmissions reduction program (HRRP), https://www.cms.gov/medicare/medicare-fee-for-service-payment/acuteinpatientppps/readmissions-reduction-program.html (2012, accessed 1 December 2012).

8. The Centers for Medicare & Medicaid Services. 2014 Reporting of AMI Payment Measure (09/15/14)-FAQs: CMS publically reported risk-standardized outcome and payment measures, https://qualitynet.cms.gov/files/5d0d376f764be766b010154f?filename=2014_DryRunRprt_Rsrcs_AMI.zip (2014, accessed 1 December 2019).

9. Boccuti C and Casillas G. Aiming for fewer hospital U-turns: the medicare hospital readmission reduction program. Policy brief, Menlo Park, CA, 2017.

10. Rau J. New round of medicare readmission penalties hits 2,583 hospitals. Kaiser Health News, https://khn.org/news/hospital-readmission-penalties-medicare-2583-hospitals/ (2019, accessed 1 December 2019).

11. Rosen AK, Chen Q, Shin MH, et al. Medical and surgical readmissions in the Veterans Health Administration: what proportion are related to the index hospitalization? Med Care 2014; 52: 243–249.

12. Donzé J, Lipsitz S, Bates DW, et al. Causes and patterns of readmissions in patients with common comorbidities: retrospective cohort study. BMJ (Clin Res Ed) 2013; 347: f7171.

13. Magdeltijns FJH, Schepers L, Pijpers E, et al. Unplanned readmissions in younger and older adult patients: the role of healthcare-related adverse events. Eur J Med Res 2016; 21: 35.

14. Kind AJH, Jencks S, Brock J, et al. Neighborhood socioeconomic disadvantage and 30-day rehospitalization: a retrospective cohort study. Ann Intern Med 2014; 161: 765–774.

15. Emde RN. Risk, intervention and meaning. Psychiatry 1988; 51: 254–259.

16. Turan TN, Al Kasab S, Nizam A, et al. Relationship between risk factor control and compliance with a lifestyle modification program in the stenting aggressive medical management for prevention of recurrent stroke in intracranial stenosis trial. J Stroke Cerebrovasc Dis 2018; 27: 801–805.

17. Meng X, Brunet A, Turecki G, et al. Risk factor modifications and depression incidence: a 4-year longitudinal Canadian cohort of the Montreal Catchment Area Study. BMJ Open 2017; 7: e015156.

18. Chandrasiri A, Dissanayake A and De Silva V. Health promotion in workplaces as a strategy for modification of risk factors for Non Communicable Diseases (NCDs): a practical example from Sri Lanka. Work (Reading, MA) 2016; 55: 281–284.

19. Siegel D, Grady D, Browner WS, et al. Risk factor modification after myocardial infarction. Ann Intern Med 1988; 109: 213–218.

20. Stephens JR, Kimple KS, Steiner MJ, et al. Discharge interventions and modifiable risk factors for preventing hospital readmissions in children with medical complexity. Rev Recent Clin Trials 2017; 12: 290–297.

21. Maddocks M, Kon SSC, Singh SJ, et al. Rehabilitation following hospitalization in patients with COPD: can it reduce readmissions? Respirology 2015; 20: 395–404.

22. Kelly KN, Iannuzzi JC, Rickles AS, et al. Risk factors associated with 30-day postoperative readmissions in major gastrointestinal resections. J Gastrointest Surg 2014; 18: 35–44.

23. Brodsky MA, Rodeghier M, Sanger M, et al. Risk factors for 30-day readmission in adults with sickle cell disease. Am J Med 2017; 130: 601.e9–601.e15.

24. Singh JA, Bini SA, Namba RS, et al. Are there modifiable risk factors for hospital readmission after total hip arthroplasty in a US healthcare system? Clin Orthop Relat Res 2015; 473: 3446–3455.

25. Garcia-Aymerich J, Farrero E, Félez MA, et al. Risk factors of readmission to hospital for a COPD exacerbation: a prospective study. Thorax 2003; 58: 100–105.

26. Borahia S, Inneh IA, Rathod P, et al. Management of modifiable risk factors prior to primary hip and knee arthroplasty: a readmission risk assessment tool. J Bone Joint Surg Am 2014; 97: 1921–1928.

27. Nelder JA and Wedderburn RWM. Generalized linear models. JR Stat Soc Ser A (Gen) 1972; 135: 370.
28. Healthcare Cost and Utilization Project. Nationwide Readmissions Database (NRD) 2014, https://www.hcup-us.ahrq.gov/nrdoverview.jsp (accessed 1 December 2019).
29. Yale New Haven Health Services Corporation/Center for Outcomes Research & Evaluation. 2016 Condition-specific measures updates and specifications report hospital-level 30-day risk-standardized readmission measures, https://qualitynet.cms.gov/files/5d0d3ac7764be766b0104a88?filename=2016_Rdmsn_Msr_Resources.zip (2016, accessed 1 December 2019).
30. Healthcare Cost and Utilization Project. Clinical classification software (CCS) for ICD10-CM/PCS, https://www.hcup-us.ahrq.gov/toolssoftware/ccs10/ccs10.jsp (2018, accessed 1 December 2019).
31. Piatetsky-Shapiro G. Discovery, analysis, and presentation of strong rules. In: Piatetsky-Shapiro G and Frawley W (eds) Knowledge discovery in databases. AAAI/MIT Press, 1991, pp.229–248.
32. Agrawal R and Srikant R. Fast algorithms for mining association rules. In: Bocca JB, Jarke M and Zaniolo C (eds) VLDB ’94 proceedings of the 20th international conference on very large data bases, Santiago, Chile, 12–15 September 1994, pp.487–499.
33. Webb GI. Discovering significant patterns. Mach Learn 2007; 68: 1–33.
34. Fisher RA. On the interpretation of $\chi^2$ from contingency tables, and the calculation of P. J R Stat Soc 1922; 85: 87.
35. Khera R, Jain S, Pandey A, et al. Comparison of readmission rates after acute myocardial infarction in 3 patient age groups (18 to 44, 45 to 64, and $\geq$65 years) in the United States. Am J Cardiol 2017; 120: 1761–1767.
36. Koulouridis I, Price LL, Madias NE, et al. Hospital-acquired acute kidney injury and hospital readmission: a cohort study. Am J Kidney Dis 2015; 65: 275–282.
37. Kim LK, Yeo I, Cheung JW, et al. Thirty-day readmission rates, timing, causes, and costs after stsegment-elevation myocardial infarction in the United States: a national readmission database analysis 2010–2014. J Am Heart Assoc. Epub ahead of print September 2018. DOI: 10.1161/JAHA.118.009863.
38. Nguyen OK, Makam AN, Clark C, et al. Predicting 30-day hospital readmissions in acute myocardial infarction: the AMI ‘READMITS’ (renal function, elevated brain natriuretic peptide, age, diabetes mellitus, nonmale sex, intervention with timely percutaneous coronary intervention, and low systo). J Am Heart Assoc 2018; 7: 1–10.
39. Rodriguez-Padial L, Elola FJ, Fernández-Pérez C, et al. Patterns of inpatient care for acute myocardial infarction and 30-day, 3-month and 1-year cardiac diseases readmission rates in Spain. Int J Cardiol 2017; 230: 14–20.
40. Tisminetzky M, McManus DD, Erskine N, et al. Thirty-day hospital readmissions in patients with non-ST-segment elevation acute myocardial infarction. Am J Med 2015; 128: 760–765.
41. Jacobs DM, Noyes K, Zhao J, et al. Early hospital readmissions after an acute exacerbation of chronic obstructive pulmonary disease in the nationwide readmissions database. Ann Am Thorac Soc 2018; 15: 837–845.
42. Chamberlain RS, Sond J, Mahendraraj K, et al. Determining 30-day readmission risk for heart failure patients: the readmission after heart failure scale. Int J Gen Med 2018; 11: 127–141.
43. Gershon AS, Thiruchelvam D, Aaron S, et al. Socioeconomic status (SES) and 30-day hospital readmissions for chronic obstructive pulmonary (COPD) disease: a population-based cohort study. PLoS One 2019; 14: 1–16.
44. Smilowitz NR, Beckman JA, Sherman SE, et al. Hospital readmission after perioperative acute myocardial infarction associated with noncardiac surgery. Circulation 2018; 137: 2332–2339.
45. Stamp LK, Frampton C, Drake J, et al. Associations of gout and baseline serum urate level with cardiovascular outcomes: analysis of the coronary disease cohort study. Arthritis Rheumatol 2019; 71: 1733–1738.
46. Ahmedani BK, Solberg LI, Copeland LA, et al. Psychiatric comorbidity and 30-day readmissions after hospitalization for heart failure, AMI, and pneumonia. Psychiatr Serv 2015; 66: 134–140.
47. National Center for Health Statistics. Uniform Hospital Discharge Data: Minimum Data Set. In: Report of the National Committee on Vital and Health Statistic, Washington, DC, 1980.
48. Charlson ME, Pompei P, Ales KL, et al. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987; 40: 373–383.