Ranking Rule-Based Automatic Explanations for Machine Learning Predictions on Asthma Hospital Encounters in Patients with Asthma: Retrospective Cohort Study

Xiaoyi Zhang, MS; **Gang Luo**, PhD
Department of Biomedical Informatics and Medical Education, University of Washington, UW Medicine South Lake Union, 850 Republican Street, Building C, Box 358047, Seattle, WA 98195, USA
xyzhang7@uw.edu, luogang@uw.edu

Corresponding author:
Gang Luo, PhD
Department of Biomedical Informatics and Medical Education, University of Washington, UW Medicine South Lake Union, 850 Republican Street, Building C, Box 358047, Seattle, WA 98195, USA
Phone: 1-206-221-4596
Fax: 1-206-221-2671
Email: luogang@uw.edu

Abstract

Background: Asthma hospital encounters impose a heavy burden on the health care system. To improve preventive care and outcomes for asthma patients, we recently developed a black-box machine learning model to predict whether an asthma patient will have 1 or more asthma hospital encounters in the succeeding 12 months. Our model is more accurate than previous models. However, black-box machine learning models do not explain their predictions, forming a barrier to widespread clinical adoption. To solve this issue, we previously developed a method to automatically provide rule-based explanations for the model's predictions and to suggest tailored interventions without sacrificing model performance. For an average patient correctly predicted by our model to have future asthma hospital encounters, our explanation method generated over 5,000 rule-based explanations, if any. However, the user of the automated explaining function, often a busy clinician, wants to quickly obtain the most useful information for a patient by viewing just the top few explanations. Therefore, we need a methodology to appropriately rank the explanations generated for a patient. This is currently an open problem.

Objective: This work aims to develop a method to appropriately rank the rule-based explanations that our automated explaining method generates for a patient.

Methods: We developed a ranking method that struck a balance among multiple factors. Through a secondary analysis of 82,888 data instances of adults with asthma from the University of Washington Medicine between 2011 and 2018, we demonstrated our ranking method on the test case of predicting asthma hospital encounters in asthma patients.

Results: For each patient predicted to have asthma hospital encounters in the succeeding 12 months, the top few explanations returned by our ranking method typically have high quality and low redundancy. Many top-ranked explanations give useful insights on various aspects of the patient's situation, which cannot be easily obtained by viewing the patient's data in the current electronic health record system.

Conclusions: The explanation ranking module is an essential component of the automated explaining function, which addresses the interpretability issue that deters the widespread adoption of machine learning predictive models in clinical practice. In the next few years, we plan to test our explanation ranking method on predictive modeling problems addressing other diseases as well as on data from other health care systems.

International Registered Report Identifier (IRRID): PRR2-10.2196/5039

Keywords: Asthma; clinical decision support; machine learning; patient care management; forecasting

Introduction

Background

Around 7.7% of Americans and over 339 million people worldwide have asthma [1,2]. Asthma incurs a total medical cost of U.S. $50 billion [3], 1,564,440 emergency department (ED) visits, and 182,620 inpatient stays annually in the United States [1]. A primary goal of asthma management is to decrease the number of asthma hospital encounters, namely ED visits and inpatient stays. The state-of-the-art approach for reaching this goal is to deploy a predictive model to identify patients at high risk of having poor outcomes in the future. Once identified, the patient is placed into a care management program. The program will assign a care manager to regularly contact the patient for assessing asthma control status, adjusting asthma medications when needed, and helping schedule appointments for health and other relevant services. Many health plans, including those in 9 of 12 metropolitan communities [4], and many health care systems, such as the University of Washington Medicine (UWM), Intermountain Healthcare, and Kaiser Permanente Northern California, currently employ this approach [5]. When used correctly, this approach prevents up to 40% of future asthma hospital encounters [4,6-9].

Due to limited capacity, a care management program can serve at most 3% of the patients [10]. To maximize the effectiveness of these programs, we should use an accurate predictive model to identify the highest risk patients. For this purpose, we recently developed a machine learning model powered by extreme gradient boosting (XGBoost) [11] on UWM data to predict which asthma patents will have asthma hospital encounters in the succeeding 12 months [12]. Compared with previous models [5,13-26], this model is more accurate and improves the area under the receiver operating characteristic curve (AUC) by ≥0.09. In addition, we previously developed a method to automatically explain the model's predictions in the form of rules and to suggest tailored interventions without sacrificing model performance [27,28]. Our method works for any black-box machine learning predictive model built on tabular data, and addresses the interpretability issue that deters the widespread adoption of machine learning predictive models in clinical practice. Among all of the published automated explaining methods for machine learning predictions [29,30], only our method can automatically recommend tailored interventions. For an average patient whom our UWM model correctly predicted to have future asthma hospital encounters, our method generated over 5,000 rule-based explanations, if any [27]. The amount of non-redundant information in these explanations is usually 2 orders of magnitude less than the number of explanations, as multiple explanations often share some common components. The user of the automatic explanation function wants to quickly obtain the most useful information for a patient by viewing just the top few explanations.

Therefore, we need to appropriately rank the explanations generated for a patient. Currently an open problem, procedures for appropriately ranking explanations are particularly important for adoption of our automated explaining method in a busy clinical environment.

Objectives

To fill the gap, this study aims to develop a method to appropriately rank the rule-based explanations that our automated explaining method [27,28] generates for a patient. We demonstrated our explanation ranking method on a test case that predicts asthma hospital encounters in asthma patients.

Methods

We reused the following items from our prior papers [12,27]: patient cohort, prediction target (a.k.a. the dependent variable), features (a.k.a. independent variables), data set, data pre-processing method, predictive model, cutoff threshold for binary classification, and automated explaining method.

Ethics approval

UWM's institutional review board approved this secondary analysis retrospective cohort study.

Patient cohort

In Washington state, UWM is the largest academic health care system. Its enterprise data warehouse stores clinical and administrative data from 3 hospitals and 12 clinics for adults. The patient cohort contained all adult asthma patients (age≥18) who obtained care at any of those UWM facilities between 2011 and 2018. In a specific year, a patient was considered asthmatic if the patient had 1 or more asthma diagnosis codes (International Classification of Diseases, Tenth Revision [ICD-10]: J45.x; International Classification of Diseases, Ninth Revision [ICD-9]: 493.0x, 493.1x, 493.8x, 493.9x) documented in the encounter billing database during the year [13,31,32]. We excluded those patients who passed away during that year.

Prediction target

Given a patient deemed asthmatic in an index year, we wanted to predict in the succeeding 12 months, whether the patient would experience any asthma hospital encounter at UWM, i.e., any ED visit or inpatient stay at UWM with asthma (ICD-10: J45.x; ICD-9: 493.0x, 493.1x, 493.8x, 493.9x) as its principal diagnosis. In predictive model training and testing, the patient's outcome in the succeeding 12 months was predicted using the patient's data until the end of the year.

Data set

We used a structured administrative and clinical data set retrieved from UWM's enterprise data warehouse. This data set contained information recorded for the visits by the patient cohort to the 12 clinics and 3 hospitals of UWM over the 9-year span of 2011-2019. Since the prediction target was for the following 12 months, the effective data in the data set spanned across the 8-year period of 2011-2018.

The training and test set split

We used the data between 2011 and 2017 as the training set to train the predictive model and to mine the association rules used by our automated explaining method. We used the data of 2018 as the test set to demonstrate our ranking method for the rule-based explanations generated by our automated explaining method.

Predictive model and features

Our UWM model used the XGBoost classification algorithm [11] and 71 features to predict the prediction target. As our UWM model was built on a single computer whose memory could hold the entire data set, the exact greedy algorithm was used to find the best split for tree learning in XGBoost [11]. These 71 features are listed in Table 2 in the online Appendix of our paper [12]. They were constructed based on the structured attributes in our data set and describe various aspects of the patient's situation, such as demographics, encounters, diagnoses, laboratory tests, procedures, vital signs, and medications. An example feature is the patient's mean length of stay for an ED visit in the past year. Every input data instance to our predictive model includes these 71 features. Features that are the same as or similar to these 71 features were formerly used to predict asthma hospital encounters in asthma patients and to give automatic explanations on Intermountain Healthcare data as well as on Kaiser Permanente Southern California data [28,33-35]. For the binary classification, we set the cutoff threshold at the top 10% of patients predicted to be at the highest risk. Our prior paper [12] showed that on the test set, our model reached an AUC of 0.902, an accuracy of 90.60% (13,268/14,644), a sensitivity of 70.18% (153/218), a specificity of 90.91% (13,115/14,426), a positive predictive value of 10.45% (153/1,464), and a negative predictive value of 99.51% (13,115/13,180).

Review of our automated explaining method

Success stories

Our automated explaining method [27,28] was designed as a general method that works for any machine learning predictive model built on tabular data. We initially demonstrated our method on predicting the diagnosis of type 2 diabetes [36]. Later, we successfully applied our method to predict asthma hospital encounters in asthma patients on Intermountain Healthcare data [28], on UWM data [27], and on Kaiser Permanente Southern California data [34]. Other researchers also successfully applied our method to project lung transplantation or death in cystic fibrosis patients [37], to project cardiac death in cancer patients, and to using projections to manage heart transplant waiting list, post-transplant follow-ups, and preventive care in patients having cardiovascular diseases [38].

Main idea

Our automated explaining method [27,28] uses class-based association rules [39,40] mined from historical data to explain a model's predictions and to recommend tailored interventions. As shown in Figure 1, the association rules are constructed separately from the predictive model and are used solely to give explanations rather than to make predictions. Thus, our automated explaining method can work with any machine learning predictive model built on tabular data with no performance penalty. That is, our method falls into the category of model-agnostic explaining methods, which are widely used to automatically explain machine learning predictions [29,30].

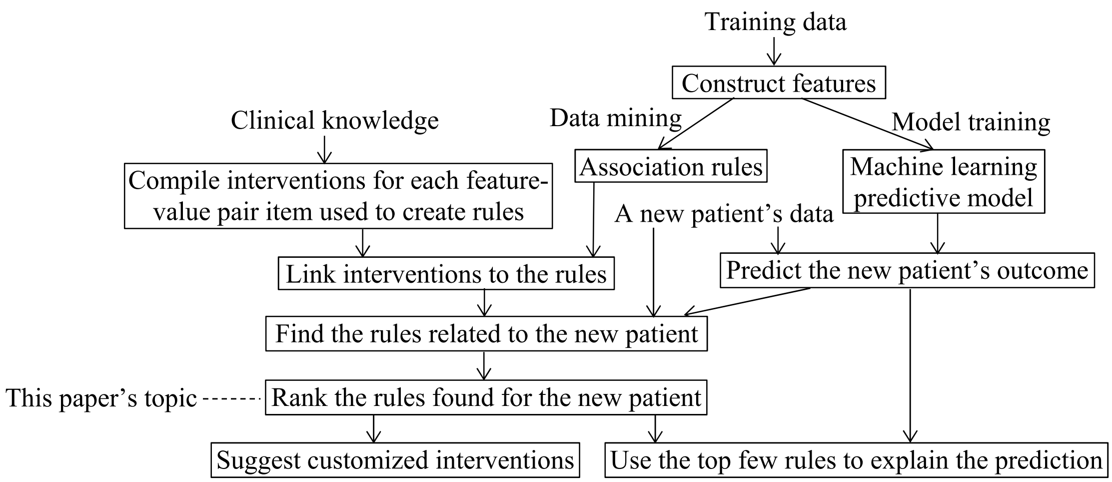

Figure 1. The flow diagram of our automated explaining method coupled with our explanation ranking method.

Before rule mining starts, an automated discretizing method based on the minimum description length principle [40,41] is first applied to the training set to convert continuous features to categorical features. The association rules are then mined from the training set using a standard method such as Apriori [39]. Each rule shows that a feature pattern links to an outcome value and has the form

p_1 AND p_2 AND ... AND $p_m \rightarrow v$.

Here, each item p_i ($1 \leq i \leq m$) is a feature-value pair (f, u). u is either the specific value of feature f or a range in which the value of f falls. For binary classification of a good vs. a poor outcome, v is the poor outcome value, such as the patient will have ≥1 inpatient stay or ED visit for asthma in the succeeding 12 months. For a patient fulfilling all of p_1, p_2, …, and p_m, the rule indicates that the patient's outcome is likely to be v. An example rule is given below:

The patient had ≥13 ED visits in the past year
AND the patient had ≥4 systemic corticosteroid prescriptions in the past year
→ the patient will likely have ≥1 inpatient stay or ED visit for asthma in the succeeding 12 months.

Constraints put on the association rules

Our automated explaining method puts several constraints on the employed association rules. In this section, we review some of these constraints that are relevant to our explanation ranking method. For an association rule

p_1 AND p_2 AND ... AND $p_m \rightarrow v$,

commonality measures its coverage in the context of v: among all of the data instances linking to v, commonality is the percentage of data instances fulfilling p_1, p_2, …, and p_m. Meanwhile, *confidence* measures its precision: among all of the data instances fulfilling p_1, p_2, …, and p_m, the confidence is the percentage of data instances linking to v. For every association rule used by our automated explaining method, we require its commonality to be ≥ a given minimum commonality threshold, such as 1%, its confidence to be ≥ a given minimum confidence threshold, such as 50%, and its left hand side to have ≤ a given number (e.g., 5) of feature-value pair items. As detailed in our previous papers [27,28], by setting the thresholds to these values, we can fulfill 3 goals concurrently. First, explanations can be given to most patients whom our UWM model correctly predicts to have ≥1 asthma hospital encounter in the succeeding 12 months. Second, the rule has a sufficiently high confidence for the user of the automated explaining function to trust the rule. Third, no rule is overly complex.

The explaining method

For each feature-value pair item used to create association rules, a clinician in the development team of the automated explaining function pre-compiles 0 or more interventions. An item linking to at least 1 intervention is called actionable. The interventions connecting to the actionable items on the left hand side of a rule are automatically linked to that rule. A rule linking to at least 1 intervention is called actionable.

For each patient predicted to have a poor outcome by the predictive model, the prediction is explained by the related association rules. For each such rule, the patient satisfies all of the feature-value pair items on its left hand side. The poor outcome value appears on its right hand side. Each rule delineates a reason of the patient's predicted poor outcome. Every actionable rule is displayed along with its linked interventions. The user of the automated explaining function can choose from these tailored interventions for the patient. The rules mined from the training set typically cover common reasons for having poor outcomes. Nonetheless, some patients could have poor outcomes due to rare reasons, such as the patient was ordered between 3 and 7 asthma medications during the past year AND the patient was ordered ≥11 distinct medications during the past year AND the patient has some drug or material allergy AND the patient had ≥1 active problem in the problem list during the past year. Hence, our explaining method usually explains the predictions for most, though not all, of the patients correctly predicted by the model to have poor outcomes.

Ranking the rule-based explanations generated by our automated explaining method

Overview

For an average patient whom the predictive model predicts to have a poor outcome, our automated explaining method finds many related association rules, if any. Multiple rules often share some common feature-value pair items on their left hand sides. To avoid overwhelming the user of the automated explaining function and to enable the user to quickly obtain the most useful information by viewing just the top few rules, we need to appropriately rank the rules found for a patient. Since a rule often has a long description, a standard computer screen can show only a few rules simultaneously. To reduce the burden on the user, we present the rules in a way similar to how a Web search engine presents its search results of a keyword query. We choose a small number n, such as 3. The user can opt to change the value of n, e.g., based on the size of the computer screen. If $\leq n$ rules are found for the patient, we display all of these rules. Otherwise, if $>n$ rules are found for the patient, we display the top n rules by default. If desired, the user can request to see more rules, e.g., by dragging a vertical scrollbar or by clicking the "next page" button.

The main idea of our association rule ranking method is to consider multiple factors in the ranking process. The procedure incorporates these factors into a rule scoring function that strikes a balance among them, and then ranks the rules found for a patient based on the scores computed for the rules in an iterative way. In each iteration, the scores of the remaining rules are re-computed and then a rule is chosen from them. In the following, we describe our rule ranking method in detail.

Factors considered in the association rule ranking process

When ranking the association rules found for a patient, we consider 5 factors:

1) *Factor 1*: All else being equal, a rule with a higher confidence is more precise and should rank higher.
2) *Factor 2*: All else being equal, a rule with a higher commonality covers a larger portion of patients with poor outcomes and should rank higher.
3) *Factor 3*: All else being equal, a rule with fewer feature-value pair items on its left hand side is easier to comprehend and should rank higher.
4) *Factor 4*: In information retrieval, search engine users want to see diversified search results [42-44]. Similarly, the user of the automated explaining function wants to see diversified information in the top-ranked rules. Hence, all else being equal, a rule whose left hand side has more items appearing in the higher-ranked rules should rank lower. The more times the items on the left hand side of this rule appear in those rules, the lower this rule should rank.
5) *Factor 5*: The user of the automated explaining function wants to find suitable interventions for the patient. Thus, all else being equal, an actionable rule should rank higher than a non-actionable rule.

The rule scoring function

We incorporate the 5 factors above into a rule scoring function to strike a balance among them. For an association rule

r: p_1 AND p_2 AND … AND $p_m \rightarrow v$,

its ranking score is a linear combination of 5 terms, 1 per factor:

$$score_r = w_c \cdot \mathrm{norm}(C_r) + w_s \cdot \mathrm{norm}(\log_{10}S_r) - w_n \cdot \mathrm{norm}(N_r) + w_d \cdot \mathrm{mean}(f(r)) + w_a \cdot \delta_{actionable}(r).$$

At a high level,

1) C_r denotes r's confidence. The term norm(C_r) has a weight w_c>0 and addresses Factor 1.
2) S_r denotes r's commonality. The term norm($\log_{10}S_r$) has a weight w_s>0 and addresses Factor 2.
3) N_r denotes the number of feature-value pair items on r's left hand side. The term norm(N_r) has a weight w_n>0 and addresses Factor 3.
4) The term $\mathrm{mean}(f(r)) \stackrel{\text{def}}{=} \sum_{i=1}^{m} f(d, p_i, r)/m$ has a weight w_d>0 and addresses Factor 4. For each i ($1 \le i \le m$), the function $f(d, p_i, r)$ is computed based on the number of times the item p_i appears in the higher-ranked rules. The value of $f(d, p_i, r)$ is always within [0, 1]. Consequently, the value of mean($f(r)$) is always within [0, 1].
5) The term $\delta_{actionable}(r)$ is the indicator function for whether r is actionable, has a weight w_a>0, and addresses Factor 5.

Let $v_r(x)$ denote the variable, such as confidence, whose value on the association rule r is x. min($v_r(x)$) and max($v_r(x)$) denote the minimum and the maximum value of $v_r(x)$ across all of the rules found for the patient, respectively. If max($v_r(x)$) ≠ min($v_r(x)$), the function $\mathrm{norm}(x) \stackrel{\text{def}}{=} [x - \min(v_r(x))] / [\max(v_r(x)) - \min(v_r(x))]$ normalizes x to a value within [0, 1]. If max($v_r(x)$) = min($v_r(x)$), all of the rules found for the patient have the same value of $v_r(x)$, and thus there is no need to consider $v_r(x)$ in ranking these rules. In this case, norm(x) is set to 0.

C_r, $\log_{10}S_r$, and N_r have different value ranges. To make C_r, $\log_{10}S_r$, and N_r comparable with each other, we use norm() to put them into the same range [0, 1]. mean($f(r)$) and $\delta_{actionable}(r)$ also fall into this range. To reflect that Factors 1, 2, and 3 are equally important, we set the default values of w_c, w_s, and w_n all to 1. To encourage the top-ranked rules to include diversified feature-value pair items, we wanted w_d's value to be >1 and set w_d's default value to 50. To strongly push the actionable rules to rank higher than the non-actionable rules, we wanted w_a's value to be ≫1 and set w_a's default value to 100. The value of w_a does not impact the score differences and hence the relative rankings among the actionable rules. When w_a is $>w_c+w_s+w_n+w_d$, the actionable rules always have larger scores than the non-actionable rules because norm(C_r), norm($\log_{10}S_r$), norm(N_r), and mean($f(r)$) all ∈[0, 1].

Detailed description of the 5 terms used in the rule scoring function

In this section, we sequentially describe the 5 terms used in the rule scoring function in detail.

Since norm() is a monotonically increasing function, all else being equal, the term norm(C_r) gives a larger ranking score to an association rule with a higher confidence C_r.

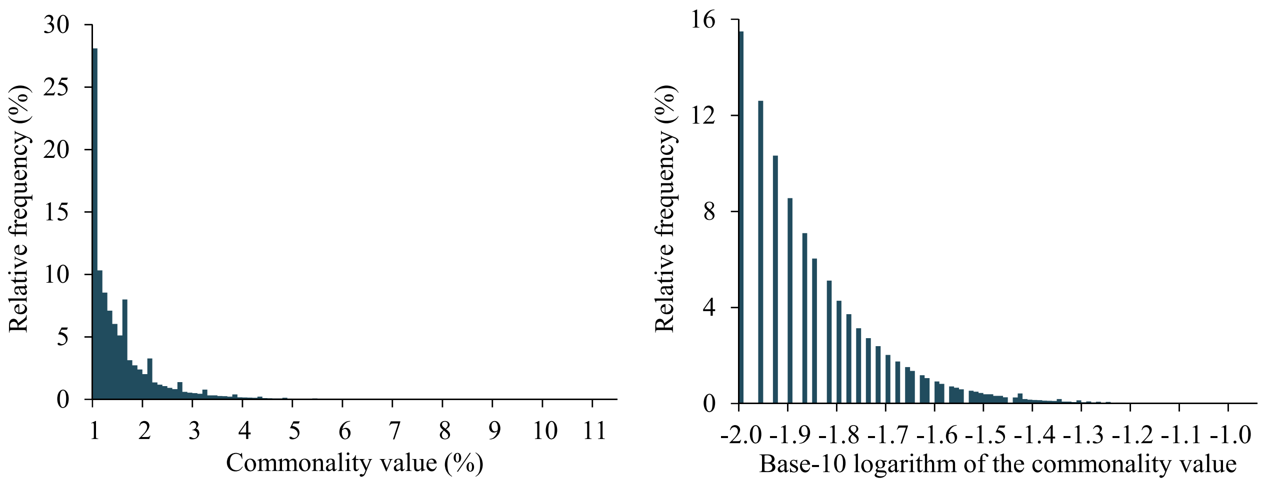

Figure 2. The distribution of the commonality values of all of the association rules used by our automated explaining method for predicting asthma hospital encounters in asthma patients at UWM.

As shown in Figure 2, the commonality values for the association rules used by our automated explaining method have a skewed distribution. Most of the commonality values are clustered in the lower value range. The commonality values of the rules that our automated explaining method generates for a patient are a sample from this distribution. We want the same weight w_s to work for different patients regardless of how the sample is taken from this distribution. Thus, for every patient, we want the variance of the terms computed on the corresponding rules' commonality values to have roughly the same scale. For this purpose, we use the $\log_{10}()$ function to transform the commonality values so that the resulting values are distributed more evenly than the raw values. Since both norm() and $\log_{10}()$ are monotonically increasing functions, norm($\log_{10}()$) is also a monotonically increasing function. All else being equal, the term norm($\log_{10}S_r$) gives a larger ranking score to a rule with a higher commonality S_r.

Since -norm() is a monotonically decreasing function, all else being equal, the term -norm(N_r) assigns a larger ranking score to an association rule with a smaller number N_r of feature-value pair items on its left hand side.

In the k-th iteration of the association rule ranking process, the top k-1 rules have already been decided. We work on identifying the k-th ranked rule. For each feature-value pair item p_i on the left hand side of a rule r that is found for the patient and whose rank has not yet been decided, we compute the exponential decay function $f(d, p_i, r) \stackrel{\text{def}}{=} \exp(-d \cdot t_i)$. Here, d>0 is the decay constant with a default value of 5. t_i is the number of times that p_i appears in the top k-1 rules. A larger value of t_i results in a smaller value of $f(d, p_i, r)$. Recall that the term $mean(f(r))$ is the mean of $f(d, p_i, r)$ over all of the items on r's left hand side. All else being equal, mean($f(r)$) assigns a smaller ranking score to a rule whose left hand side has more items appearing in the top k-1 rules.

$\delta_{actionable}(r)$ is =1 if the association rule r is actionable and is =0 if r is non-actionable. All else being equal, the term $\delta_{actionable}(r)$ assigns a larger ranking score to an actionable rule compared with that of a non-actionable rule.

The iterative association rule ranking process

If only 1 association rule is found for a patient, there is no need to rank the rule. If ≥2 rules are found for the patient, we rank these rules in an iterative way. In the k-th iteration, we compute the ranking score for every rule r that is found for the patient and whose rank has not yet been decided. Compared with the case in the previous iteration, the score needs to be updated if and only if the value of mean($f(r)$) changes, i.e., if and only if any feature-value pair item on r's left hand side also appears on the left hand side of the (k-1)-th ranked rule. Among all of the rules that are found for the patient and whose ranks have not yet been decided, we select the rule with the highest score as the k-th ranked rule. In case ≥2 of these rules have the same highest score, we choose 1 of them randomly as the k-th ranked rule.

For each association rule on display, sort the feature-value pair items on its left hand side

The same feature-value pair item could appear on the left hand sides of ≥2 top-ranked association rules. The user of the automated explaining function tends to read both the rules and the items on the left hand side of a rule in the display order. To help the user obtain the most useful information as quickly as possible, for each rule on display, we need to appropriately rank the items on its left hand side. For this purpose, we consider 2 factors:

1) *Factor 6*: The user wants to see new information as quickly as possible. Hence, all else being equal, an item for a rule that already appears in the higher-ranked rules should rank lower. As the number of times the item appears in higher ranked rules increases, the rank of the item should decrease.
2) *Factor 7*: The user wants to find suitable interventions for the patient. Thus, all else being equal, an actionable item should rank higher than a non-actionable item.

We incorporate the 2 factors above into an item scoring function to strike a balance between them. Consider the k-th ranked association rule. For each feature-value pair item p on its left hand side, p's ranking score is a linear combination of 2 terms, 1 per factor:

$$score_p = w_g \cdot \exp(-d \cdot t) + w_b \cdot \delta_{actionable}(p).$$

The terms in the equation above are further explained below.

1) In the equation for $score_p$ above, d is the same decay constant used in $f(d, p_i, r)$ in the rule scoring function. t is the number of times that p appears in the top k-1 rules. The larger the value of t, the smaller the value of the exponential decay function $\exp(-d \cdot t)$. Hence, all else being equal, the $\exp(-d \cdot t)$ term assigns a smaller ranking score to an item that appears more times in the top k-1 rules. This addresses Factor 6.
2) The term $\delta_{actionable}(p)$ is an indicator function for whether p is actionable. The term $\delta_{actionable}(p)$ is =1 if p is actionable and is =0 if p is non-actionable. All else being equal, the $\delta_{actionable}(p)$ term causes an actionable item to have a larger ranking score than that of a non-actionable item. This addresses Factor 7.

Both $\exp(-d \cdot t)$ and $\delta_{actionable}(p)$ fall into the range [0, 1]. For the weight w_g>0 of the term $\exp(-d \cdot t)$, we set its default value to 1. For the weight w_b>0 of the term $\delta_{actionable}(p)$, we set its default value to 2 that is >1. The value of w_b has no impact on the score differences and hence the relative ranking among the actionable items on the left hand side of the association rule. When w_b is >w_g, the actionable items always have larger scores than those of the non-actionable items because $\exp(-d \cdot t) \in [0, 1]$.

When the rank of an association rule is decided, we compute the ranking score for each feature-value pair item on the rule's left hand side. We then sort these items in descending order of their scores. Any items with the same score are randomly ordered and given consecutive ranks.

Computer coding implementation

We used the R programming language to implement our explanation ranking method.

Providing informative examples of the explanation ranking results

We want to demonstrate various aspects of the results produced by our explanation ranking method. For this purpose, we chose 8 asthma patients in the test set, each of whom our UWM model correctly predicted to have ≥1 asthma hospital encounter in 2019 and our automated explaining method could explain this prediction. For each patient, we show the top 3 explanations produced by our explanation ranking method. Each patient satisfied 1 or more of the following conditions and was an informative case:

1) *Condition 1*: The patient had numerous encounters, laboratory tests, or medication orders in 2018 reflecting a complex condition. In this case, we want to show how well the top 3 explanations capture and summarize the patient's key information related to asthma outcome prediction.
2) *Condition 2*: All or most of the asthma-related encounters that the patient had in 2018 were ED visits. Such a patient often had a poor asthma control due to poor treatment adherence. In this case, we want to show how well the interventions linking to the top 3 explanations address the poor asthma control.
3) *Condition 3*: For each of the top 3 association rules produced for the patient, its confidence value is close to the minimum confidence threshold. Its commonality value is close to the minimum commonality threshold. In this case, we want to illustrate these "borderline" rules. Recall that below either threshold, a rule will not be used by our automated explaining method.
4) *Condition 4*: The top 3 rules produced for the patient share several common feature-value pair items on their left hand sides. This could happen, e.g., when our automated explaining method finds only a few rules for the patient because the patient had only a small amount of information recorded in the electronic health record (EHR) system during the past year. In this case, we want to demonstrate the information redundancy in these rules.
5) *Condition 5*: A patient at high risk for future asthma hospital encounters often had ≥1 hospital encounter related to asthma during the past year. The patient being examined does not fall into this category. The patient had several feature values correlated with future asthma hospital encounters, but no hospital encounter related to asthma during the past year. In this case, we want to show how well the top 3 explanations capture these feature values.

Sensitivity analysis of the parameters used in the rule scoring function

The rule scoring function uses 6 parameters whose default values are: w_c=1, w_s=1, w_n=1, w_d=50, d=5, and w_a=100. To assess the impact of the 5 parameters w_c, w_s, w_n, w_d, and d on the association rule ranking results, we did 5 experiments. In each experiment, we changed the value of 1 of these 5 parameters and kept the other parameters at their default values. In comparison with the case of all parameters taking their default values, we measured the average percentage change in the unique feature-value pair items contained in the top min(3, q) rules for a patient, where q denotes the number of rules that our automated explaining method generated for the patient. The percentage change in the unique items was defined as 100 × the number of changed unique items divided by the number of unique items in the top min(3, q) rules. The average was taken over all patients in the test set, each of whom was predicted to have ≥1 asthma hospital encounter in 2019 and had at least 1 applicable rule (i.e., q≥1). Multiple rules often differ from each other by only 1 item on their left hand sides. Also, switching items among the top few rules for a patient has little impact on the total amount of information that the user of the automated explaining function obtains from these rules. Thus, we measured the number of changed unique items in the top few rules per patient instead of the number of changed top rules per patient or the number of changed items per top rule.

As explained before, when w_a is $>w_c+w_s+w_n+w_d$, the actionable rules always rank higher than the non-actionable rules. Meanwhile, the concrete value of w_a has no impact on the ranking of the actionable rules. All of the rules that our automated explaining method used on the UWM data set were actionable [27]. Thus, we did no sensitivity analysis on w_a. For a similar reason, we did no sensitivity analysis on the weights w_g and w_b used in the item scoring function.

Results

The demographic and clinical characteristics of our patient cohort

Each UWM data instance used in this study corresponds to a distinct (patient, index year) pair and is used to predict the patient's outcome in the succeeding 12 months. Tables 1 and 2 in the Appendix show our patient cohort's demographic and clinical characteristics during 2011-2017 and 2018 separately. These 2 sets of characteristics are similar to each other. During 2011-2017, 1.74% (1,184/68,244) of data instances were linked to asthma hospital encounters in the succeeding 12 months. During 2018, 1.49% (218/14,644) of data instances were linked to asthma hospital encounters in the succeeding 12 months. A detailed comparison of these 2 sets of characteristics is given in our paper [12].

Execution time

For an average asthma patient, our explanation ranking method took <0.01 second to produce the top 3 explanations. This is sufficiently fast for providing real-time clinical decision support.

Informative examples of the explanation ranking results

The test set included 134 asthma patients, each of whom our UWM model correctly predicted to have ≥1 asthma hospital encounter in 2019 and our automated explaining method could explain this prediction. To show the reader various aspects of the results produced by our explanation ranking method, we chose 8 of these patients who were informative cases. Tables 1-8 present the top 3 association rules that our explanation ranking method produced for each of those 8 patients, respectively. For each of the top 3 rules produced for the seventh selected patient, Table 9 lists the interventions linking to the rule.

Table 1. The top 3 association rules that our explanation ranking method produced for the first selected patient (*Patient 1*). This patient satisfied Condition 1.

Rank	Association rule	Confidence	Commonality
1	The patient had 2 or 3 ED visits related to asthma during the past year AND the patient was ordered between 7 and 11 distinct asthma medications during the past year AND the patient was ordered between 5 and 7 distinct asthma relievers during the past year AND the patient had ≥1 active problem in the problem list during the past year → the patient will likely have ≥1 inpatient stay or ED visit for asthma in the succeeding 12 months.	52% (24/46)	2.03% (24/1184)
2	The patient's mean length of stay of an ED visit during the past year was >0.205 day AND the patient was ordered ≥4 systemic corticosteroids during the past year AND the patient's most recent ED visit on asthma was from between 26 and 100 days ago AND the patient was ordered 2 distinct nebulizer medications during the past year AND the patient is not White → the patient will likely have ≥1 inpatient stay or ED visit for asthma in the succeeding 12 months.	50% (14/28)	1.18% (14/1184)
3	The patient had ≥8 nebulizer medication orders during the past year AND the patient had ≥5 no shows during the past year AND the patient had 2 or 3 ED visits related to asthma during the past year AND the patient's mean temperature during the past year was ≤98.09 Fahrenheit AND the patient is ≤54 years old → the patient will likely have ≥1 inpatient stay or ED visit for asthma in the succeeding 12 months.	56% (18/32)	1.52% (18/1184)

Table 2. The top 3 association rules that our explanation ranking method produced for the second selected patient (*Patient 2*). This patient satisfied Condition 1.

Rank	Association rule	Confidence	Commonality
1	The patient's most recent diagnosis of asthma with acute exacerbation or status asthmaticus was from ≤110 days ago AND the patient was ordered ≥10 short-acting beta-2 agonists during the past year AND the patient had no outpatient visit during the past year AND the patient's first encounter on asthma was from ≥1 year ago → the patient will likely have ≥1 inpatient stay or ED visit for asthma in the succeeding 12 months.	62% (54/87)	4.56% (54/1184)
2	The patient had ≥16 asthma medication orders during the past year AND the patient's mean respiratory rate during the past year was >16.89 breaths per minute AND the patient's most recent visit was an ED visit AND the patient is Black or an African American AND the patient was totally allowed between 1 and 33 medication refills during the past year → the patient will likely have ≥1 inpatient stay or ED visit for asthma in the succeeding 12 months.	56% (18/32)	1.52% (18/1184)
3	The patient had between 8 and 16 asthma diagnoses during the past year AND the patient's lowest peripheral capillary oxygen saturation (SpO_2) level during the past year was between 8.0% and 94.5% AND the patient's most recent ED visit on asthma was from between 26 and 100 days ago	51% (18/35)	1.52% (18/1184)

	AND the patient is not White AND the patient had ≤6 encounters during the past year → the patient will likely have ≥1 inpatient stay or ED visit for asthma in the succeeding 12 months.		

Table 3. The top 3 association rules that our explanation ranking method produced for the third selected patient (*Patient 3*). This patient satisfied Condition 1.

Rank	Association rule	Confidence	Commonality
1	The patient's most recent diagnosis of asthma with acute exacerbation or status asthmaticus was from ≤110 days ago AND the patient's most recent visit was an ED visit AND the patient had between 9 and 17 primary or principal asthma diagnoses during the past year → the patient will likely have ≥1 inpatient stay or ED visit for asthma in the succeeding 12 months.	62.2% (79/127)	6.67% (79/1184)
2	The patient had between 17 and 27 asthma diagnoses during the past year AND the patient's most recent visit was an ED visit AND the patient had no visit to the primary care provider during the past year → the patient will likely have ≥1 inpatient stay or ED visit for asthma in the succeeding 12 months.	56% (38/68)	3.21% (38/1184)
3	The patient was ordered ≥10 short-acting beta-2 agonists during the past year AND the highest severity of all asthma diagnoses of the patient during the past year was moderate or severe persistent asthma AND the patient was allowed ≥34 medication refills during the past year AND the patient is ≤54 years old → the patient will likely have ≥1 inpatient stay or ED visit for asthma in the succeeding 12 months.	50% (20/40)	1.69% (20/1184)

Table 4. The top 3 association rules that our explanation ranking method produced for the fourth selected patient (*Patient 4*). This patient satisfied Condition 2.

Rank	Association rule	Confidence	Commonality
1	The patient had ≥7 ED visits related to asthma during the past year AND the patient is single → the patient will likely have ≥1 inpatient stay or ED visit for asthma in the succeeding 12 months.	92% (34/37)	2.87% (34/1184)
2	The patient had between 9 and 17 primary or principal asthma diagnoses during the past year AND the patient's most recent outpatient visit on asthma was from ≥365 days ago → the patient will likely have ≥1 inpatient stay or ED visit for asthma in the succeeding 12 months.	62.9% (66/105)	5.57% (66/1184)
3	The patient had ≥28 asthma diagnoses during the past year AND the patient had no outpatient visit during the past year → the patient will likely have ≥1 inpatient stay or ED visit for asthma in the succeeding 12 months.	84% (16/19)	1.35% (16/1184)

Table 5. The top 3 association rules that our explanation ranking method produced for the fifth selected patient (*Patient 5*). This patient satisfied Condition 5.

Rank	Association rule	Confidence	Commonality
1	The patient had ≥20 diagnoses of asthma with acute exacerbation during the past year AND the patient was ordered ≥10 short-acting beta-2 agonists during the past year → the patient will likely have ≥1 inpatient stay or ED visit for asthma in the succeeding 12 months.	59% (48/82)	4.05% (48/1184)
2	The patient had ≥28 asthma diagnoses during the past year AND the patient had ≥8 nebulizer medication orders during the past year AND the patient had no outpatient visit to the primary care provider during the past year	67% (37/55)	3.13% (37/1184)

	→ the patient will likely have ≥1 inpatient stay or ED visit for asthma in the succeeding 12 months.		
3	The patient had ≥18 primary or principal asthma diagnoses during the past year AND the patient was ordered ≥8 distinct asthma relievers during the past year AND the patient's mean heart rate during the past year was >80 beats per minute → the patient will likely have ≥1 inpatient stay or ED visit for asthma in the succeeding 12 months.	50.0% (58/116)	4.90% (58/1184)

Table 6. The top 3 association rules that our explanation ranking method produced for the sixth selected patient (*Patient 6*). This patient satisfied Conditions 3 and 4.

Rank	Association rule	Confidence	Commonality
1	The patient had 2 or 3 ED visits related to asthma during the past year AND the patient's most recent outpatient visit on asthma was from ≤104 days ago AND the patient was ordered ≤2 inhaled corticosteroids during the past year AND the patient is ≤54 years old AND the patient's relative change of weight during the past year was ≤3% → the patient will likely have ≥1 inpatient stay or ED visit for asthma in the succeeding 12 months.	55% (22/40)	1.86% (22/1184)
2	The patient had between 3 and 8 diagnoses of asthma with (acute) exacerbation during the past year AND the patient had 2 or 3 ED visits related to asthma during the past year AND the patient is not White AND the patient was ordered ≤2 distinct asthma medications during the past year AND the patient is single → the patient will likely have ≥1 inpatient stay or ED visit for asthma in the succeeding 12 months.	56% (14/25)	1.18% (14/1184)
3	The patient's most recent outpatient visit on asthma was from ≤104 days ago AND the patient had 2 or 3 ED visits related to asthma during the past year AND the patient was ordered ≥1 unit of medications during the past year AND the patient had no public insurance on the last day of the past year AND the patient had between 1 and 13 outpatient visits during the past year → the patient will likely have ≥1 inpatient stay or ED visit for asthma in the succeeding 12 months.	50% (16/32)	1.35% (16/1184)

Table 7. The top 3 association rules that our explanation ranking method produced for the seventh selected patient (*Patient 7*). This patient satisfied Conditions 1 and 2.

Rank	Association rule	Confidence	Commonality
1	The patient had ≥7 ED visits related to asthma during the past year → the patient will likely have ≥1 inpatient stay or ED visit for asthma in the succeeding 12 months.	76% (39/51)	3.29% (39/1184)
2	The patient had between 17 and 27 asthma diagnoses during the past year AND the patient had no outpatient visit during the past year → the patient will likely have ≥1 inpatient stay or ED visit for asthma in the succeeding 12 months.	58% (28/48)	2.36% (28/1184)
3	The patient's mean length of stay of an ED visit during the past year was between 0.025 and 0.205 day AND the patient had ≥3 ED visits during the past year AND the patient was ordered ≥3 asthma relievers that are neither short-acting beta-2 agonists nor systemic corticosteroids during the past year AND the patient was ordered ≥4 systemic corticosteroids during the past year AND the patient is single → the patient will likely have ≥1 inpatient stay or ED visit for asthma in the succeeding 12 months.	50.0% (58/116)	4.90% (58/1184)

Table 8. The top 3 association rules that our explanation ranking method produced for the eighth selected patient (*Patient 8*). This patient satisfied Condition 5.

Rank	Association rule	Confidence	Commonality
1	The patient had between 9 and 17 primary or principal asthma diagnoses during the past year AND the patient had ≥16 asthma medication orders during the past year AND the patient had no outpatient visit to the primary care provider during the past year AND the patient is not White → the patient will likely have ≥1 inpatient stay or ED visit for asthma in the succeeding 12 months.	52% (45/87)	3.80% (45/1184)
2	For the patient's most recent visit, the time from making the request to the actual visit was ≤0.6 day AND the patient had ≥16 asthma medication orders during the past year AND the patient is Black or an African American AND the patient's first encounter on asthma was from ≥1 year ago AND the patient's lowest SpO_2 level during the past year was between 94.5% and 95.5% → the patient will likely have ≥1 inpatient stay or ED visit for asthma in the succeeding 12 months.	63% (12/19)	1.01% (12/1184)
3	The patient was ordered ≥12 distinct asthma medications during the past year AND the patient had ≥12 encounters during the past year AND the patient's most recent outpatient visit on asthma was from ≤104 days ago AND the patient had ≤82 laboratory tests during the past year AND the patient is not White → the patient will likely have ≥1 inpatient stay or ED visit for asthma in the succeeding 12 months.	63% (12/19)	1.01% (12/1184)

Table 9. For each of the top 3 association rules that our explanation ranking method produced for Patient 7, the interventions linking to the rule.

Rank	Association rule	Linked interventions
1	The patient had ≥7 ED visits related to asthma during the past year → the patient will likely have ≥1 inpatient stay or ED visit for asthma in the succeeding 12 months.	An intervention linked to the item "the patient had ≥7 ED visits related to asthma during the past year" is to carry out control strategies to prevent needing emergency care.
2	The patient had between 17 and 27 asthma diagnoses during the past year AND the patient had no outpatient visit during the past year → the patient will likely have ≥1 inpatient stay or ED visit for asthma in the succeeding 12 months.	An intervention linked to the item "the patient had between 17 and 27 asthma diagnoses during the past year" is to give the patient suggestions on how to improve asthma control. An intervention linked to the item "the patient had no outpatient visit during the past year" is to make sure that the patient has a primary care provider and to suggest the patient to regularly visit this provider.
3	The patient's mean length of stay of an ED visit during the past year was between 0.025 and 0.205 day AND the patient had ≥3 ED visits during the past year AND the patient was ordered ≥3 asthma relievers that are neither short-acting beta-2 agonists nor systemic corticosteroids during the past year AND the patient was ordered ≥4 systemic corticosteroids during the past year AND the patient is single → the patient will likely have ≥1 inpatient stay or ED visit for asthma in the succeeding 12 months.	An intervention linked to the items "the patient's mean length of stay of an ED visit during the past year was between 0.025 and 0.205 day" and "the patient had ≥3 ED visits during the past year" is to carry out control strategies to prevent needing emergency care. An intervention linked to the items "the patient was ordered ≥3 asthma relievers that are neither short-acting beta-2 agonists nor systemic corticosteroids during the past year" and "the patient was ordered ≥4 systemic corticosteroids during the past year" is to tailor the prescribed asthma medications, to help the patient adhere to asthma controllers, and to improve avoidance of triggers.

As illustrated by the cases shown in Tables 1-9, the top few explanations that our explanation ranking method produces for a patient offer 5 benefits for clinical decision support. We describe these 5 benefits sequentially as follows.

Benefit 1: The top few explanations give succinct summaries on a wide range of aspects of the patient's situation.

To make good clinical decisions on a patient, the clinician needs to understand the patient's situation well. For each of the 8 selected patients, the top 3 rule-based explanations produced by our explanation ranking method give succinct summaries on a wide range of aspects of the patient's situation, such as demographics, encounters, vital signs, laboratory tests, and medications. From these summaries, the user of the automated explaining function can quickly gain a comprehensive understanding of the patient's situation related to the prediction target. This saves the user a significant amount of time and effort. In comparison, currently to gain this understanding in a clinical setting, even if a clinician knows all of the features needed for this purpose, the clinician often needs to spend a significant amount of time laboriously checking many pages of information scattered in various places in the EHR system and doing manual calculations. For example, Patient 1 had a total of >1,000 encounters recorded in the EHR system at UWM over time. In 2018, this patient had 164 encounters, only 2 of which were related to asthma and both were ED visits. As Table 1 shows, the statistic of 2 ED visits related to asthma is reflected by the first item on the left hand side of the first association rule produced for this patient. As another example, in 2018, Patient 2 had 740 medication orders, 153 of which were asthma medication orders covering a total of 72 short-acting beta-2 agonists. As Table 2 shows, the statistic of 72 short-acting beta-2 agonists is reflected by the first item on the left hand side of the first rule produced for this patient. The statistic of 153 asthma medication orders is reflected by the first item on the left hand side of the second rule produced for this patient. The cases with the other items on the left hand sides of the top 3 rules produced for these 2 patients are similar.

To quickly gain a comprehensive understanding of a patient's situation, a clinician could ask the patient to describe his or her situation. However, the patient often cannot do this well. For example, Patients 1, 3, and 7 have severe mental disorders, which affect their memory and ability to describe their situation. This is a common scenario. Over 30% of asthma patients at UWM suffer from mental disorders. Moreover, when needing to make clinical decisions, the clinician does not always have direct access to the patient. For instance, when identifying candidate patients for care management, the care managers are sitting in a back office and cannot talk to patients. In either of these 2 cases, the summaries given by the top few rule-based explanations can help the clinician gain needed understanding on the patient.

Benefit 2: Showing the top few explanations can save the user of the automated explaining function from having to manually think of many features summarizing the patient's situation and compute their values.

Often, we need to use many features to adequately summarize a patient's situation related to the prediction target. In a busy clinical environment, a clinician cannot be expected to enumerate all of these features in a short amount of time. The top few rule-based explanations that our explanation ranking method produces for a patient cover the values of various features summarizing the patient's situation related to the prediction target. This saves the user of the automated explaining function from having to manually think of these features and to compute their values.

Benefit 3: The top few explanations can give information not easily obtainable from using the existing search and browsing functions of the EHR system to check the patient's data.

The EHR system provides some browsing and basic search functions. However, for certain important features summarizing a patient's situation related to the prediction target, we cannot easily obtain their values by using these functions to check the patient's EHR data. The top few rule-based explanations that our explanation ranking method produces for a patient cover the values of several such features. This saves the user of the automated explaining function a significant amount of work. For example, many different asthma medications exist. In 2018, Patient 2 had 740 medication orders. It is difficult and time-consuming to manually compute the number of asthma medication orders and the total number of short-acting beta-2 agonists ordered for this patient in 2018. In comparison, as mentioned before, these 2 statistics are directly reflected by the first and second rules produced for this patient. As a second example, in 2018, Patient 7 had 14 ED visits, 8 of which were related to asthma. For 2 of these 8 ED visits, asthma was not the primary diagnosis. To compute the patient's number of ED visits related to asthma in 2018, a clinician needs to find all of the patient's ED visits in 2018 and check each of them to see whether it has an asthma diagnosis code. This takes a non-trivial amount of time. In comparison, as Table 7 shows, the statistic of 8 ED visits related to asthma is directly reflected by the first item on the left hand side of the first rule produced for this patient. As a third example, in 2018, Patient 8 had 12 outpatient visits, none of which was to the patient's primary care provider. To compute the patient's number of outpatient visits to the primary care provider, a clinician needs to find all of the patient's outpatient visits in 2018 and manually check each of them to see whether it involved the patient's primary care provider. This takes a non-trivial amount of time. In comparison, as Table 8 shows, the third item on the left hand side of the first rule produced for this patient directly shows that the patient had 0 outpatient visit to the primary care provider in 2018.

Benefit 4: The top few explanations can help the user of the automated explaining function avoid overlooking certain important information of the patient and discover errors in the data recorded on the patient in the EHR system.

An asthma patient also often has several other diseases, which could distract the clinicians and cause them to pay insufficient attention to the patient's asthma and record incorrect data on the patient in the EHR system. For example, in 2018, the asthmatic Patient 3 also suffered from major depression disorder, anxiety, post-traumatic stress disorder, visual disturbance, chronic pain, and knee osteoarthritis. In the patient's problem list, these diseases were recorded as major problems, whereas asthma was recorded as a minor problem. However, the patient had 15 primary asthma diagnoses, some of which were severe persistent asthma and indicated that asthma was a major problem for the patient at that time. Only in 2020, asthma was first recorded as 2 major problems in the patient's problem list, 1 on asthma exacerbation and another on persistent asthma with status asthmaticus. As shown in Table 3, the first and third rules produced for the patient cover the patient's number of asthma diagnoses and the highest severity of these diagnoses in 2018, reflecting that the patient had severe persistent asthma at that time. This can help the user of the automated explaining function avoid overlooking this aspect and discover that asthma should be recorded as a major problem in the patient's problem list in 2018.

Benefit 5: The top few explanations can help the user of the automated explaining function identify certain problems of the patient not easily findable in the EHR system.

This can help the user of the automated explaining function identify suitable interventions for the patient. For example, as shown in Table 6, the first and second rules produced for Patient 6 show that this patient had quite a few ED visits related to asthma, but was ordered very few asthma medications in 2018. It turns out that this patient did not adhere to albuterol prescriptions due to some personal preference. Realizing this, the user could consider adopting the intervention of replacing albuterol with some other asthma medications that the patient is willing to take. As another example, as shown in Tables 4 and 7, for Patients 4 and 7, the top 3 rules produced for each patient reveal that the patient had many ED visits related to asthma, but no outpatient visit in 2018. It turns out that these 2 patients were homeless. With this information, the user could consider providing social resources to reduce the socioeconomic burden of homelessness that causes ineffective access to health care.

Description of 5 example patient cases, 1 case per each of Conditions 1-5

In this section, for each of Conditions 1-5, we choose 1 example patient satisfying it and show how this patient was an informative case.

As an example case for Condition 1, Patient 1 had 164 encounters and 644 medication orders in 2018. As shown in Table 1, the top 3 explanations produced for this patient effectively capture and summarize various aspects of the patient's key information related to future asthma hospital encounters.

As an example case for Condition 2, Patient 7 had 8 asthma-related encounters in 2018, all of which were ED visits. As shown in Table 7, the top 3 explanations produced for this patient reveal that the patient had many asthma diagnoses, had no outpatient visit, and was ordered ≥4 systemic corticosteroids during 2018, reflecting a poor asthma control. As shown in Table 9, the interventions linked to the top 3 explanations address various aspects related to the poor asthma control.

Patient 6 provides an example for Condition 3. As shown in Table 6, for each of the top 3 association rules produced for this patient, its confidence value is close to the minimum confidence threshold of 50%, and its commonality value is close to the minimum commonality threshold of 1%. These 3 rules cover a wide range of aspects of the patient's situation, including demographics, encounters, diagnoses, vital signs, and medications.

As an example case for Condition 4, Patient 6 had only 3 encounters and 1 medication order, and subsequently a small amount of information recorded in the EHR system in 2018. As shown in Table 6, the top 3 explanations produced for this patient share 3 common feature-value pair items on their left hand sides. Despite having moderate information redundancy, these explanations still cover a wide range of aspects of the patient's situation, including demographics, encounters, diagnoses, vital signs, and medications.

As an example case for Condition 5, Patient 8 had no hospital encounter related to asthma in 2018. As shown in Table 8, the top 3 explanations produced for this patient capture several feature values of the patient correlated with future asthma hospital encounters, such as the patient having between 9 and 17 primary or principal asthma diagnoses during the past year, the patient having ≥16 asthma medication orders during the past year, the patient having no outpatient visit to the primary care provider during the past year, and the patient having ≥12 encounters during the past year.

Sensitivity analysis results of the parameters used in the rule scoring function

We did 5 sensitivity analysis experiments, 1 for each of the 5 parameters w_c, w_s, w_n, w_d, and d used in the rule scoring function. In each experiment, we changed the corresponding parameter's value and kept the other parameters at their default values. In comparison with the case of all parameters taking their default values and for each of these 5 parameters, Figures 3 to 5 show the average percentage change in the unique feature-value pair items contained in the top min(3, q) association rules for a patient vs. the parameter's value. In each figure, a vertical dotted line shows the default value of the corresponding parameter.

For each parameter value tested, the average percentage change in the unique items is relatively small (<20%). The only exception is the case of either w_d=0 or d=0, where the average percentage change in the unique items is 43.57% (453.18/1040). In both cases, our explanation ranking method ignores the need for the top ranked rules to provide diversified information (Factor 4).

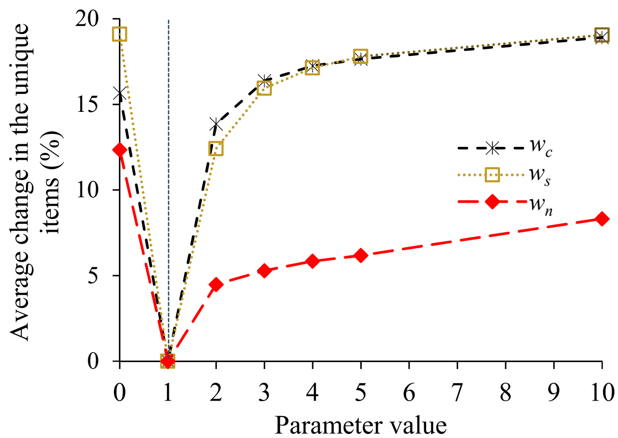

Figure 3. In comparison with the case of all parameters taking their default values and for each of the 3 parameters w_c, w_s, and w_n, the average percentage change in the unique feature-value pair items contained in the top min(3, q) association rules for a patient vs. the parameter's value.

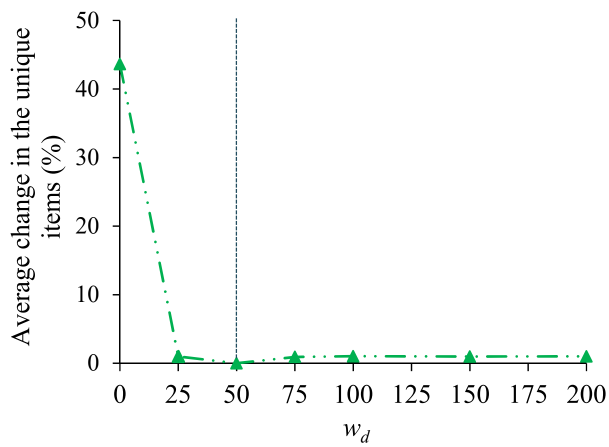

Figure 4. In comparison with the case of all parameters taking their default values, the average percentage change in the unique feature-value pair items contained in the top min(3, q) association rules for a patient vs. the value of the parameter w_d.

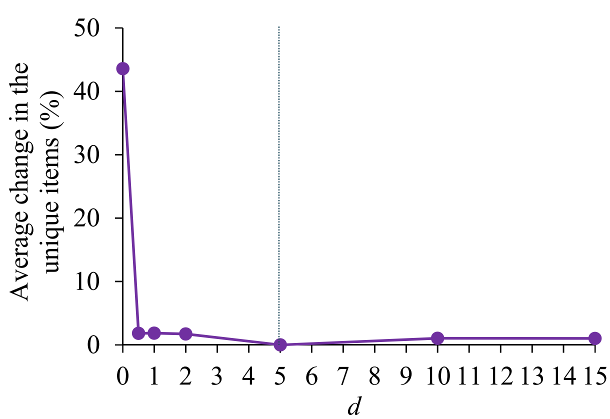

Figure 5. In comparison with the case of all parameters taking their default values, the average percentage change in the unique feature-value pair items contained in the top min(3, q) association rules for a patient vs. the value of the parameter d.

Discussion

Principal results

In a busy clinical environment, the explanation ranking module is essential for our automated explaining function for machine learning predictions to provide high-quality real-time decision support. For an average asthma patient correctly predicted by our UWM model to have future asthma hospital encounters, our automated explaining method generated over 5,000 rule-based explanations, if any. Within a negligible amount of time, our explanation ranking method can appropriately rank them and return the few highest ranked explanations. These few explanations typically have high quality and low redundancy. From them, the user of the automated explaining function can gain useful insights on various aspects of the patient's situation. Many of these insights cannot be easily obtained by viewing the patient's data in the current EHR system. With further improvement of model accuracy, our UWM model coupled with our automated explaining and explanation ranking methods could be deployed to better guide the use of asthma care management to save costs and improve patient outcomes.

Like our automated explaining method, our explanation ranking method is general purpose and does not rely on any specific property of a particular prediction target, disease, patient cohort, or health care system. Our automated explaining method coupled with our explanation ranking method can be used for any predictive modeling problem on any tabular data set. This gives a unique solution to the interpretability issue that deters the widespread adoption of machine learning predictive models in clinical practice.

In our sensitivity analysis, when we changed any parameter used in our explanation ranking method from its default value, the resulting average percentage change in the unique feature-value pair items contained in the top min(3, q) association rules for a patient was typically <20%. This is not a large change, as most (>80%) of the distinct feature-value pair items contained in these rules and subsequently most of the information seen by the user of the automated explaining function remain the same. For instance, if the top min(3, q) association rules contain 15 unique feature-value pair items, at most 3 of these feature-value pair items would vary due to the change of the parameter value while the other 12 or more remain the same as before. Thus, each parameter used in our explanation ranking method has a reasonably large stable range, within which the top few explanations produced by our method do not vary greatly as the parameter value changes. The default value of the parameter is in this stable range. According to our test results, the stable ranges are: [0, 10] for w_c, [0, 10] for w_s, [0, 10] for w_n, [25, 200] for w_d, and [0.5, 15] for d.

Adjusting certain parameters used in the rule scoring and the item scoring functions

Both the rule scoring and the item scoring functions have several parameters. Based on the preferences of the users of the automated explaining function and the specific needs of the particular health care application, the developer of the automated explaining function could change some of these parameters from their default values. In the UWM test case used in this paper, all association rules used by our automated explaining method were actionable. For some other predictive modeling problems, certain rules used by our automated explaining method are non-actionable [36]. In this case, if we want to allow some non-actionable rules to rank higher than some non-top-scored actionable rules on any patient, we need to reduce the weight w_a. Similarly, if we want to allow some non-actionable items to rank higher than some actionable items in any non-top-scored rule that our automated explaining method finds for any patient, we need to reduce the weight w_b.

Considerations on the threshold used to determine the top rules that will be displayed by default

Differing patients have different distributions of the ranking scores for the association rules found for the patient. No single threshold on the ranking score works for all patients. Thus, we use a threshold on the number of rules rather than a threshold on the ranking score to determine the top rules that will be displayed by default. This is similar to the case with a Web search engine such as Google. Google does not use any ranking score threshold to determine the search results that will be displayed on each search result page. Instead, by default Google displays 10 search results on each search result page. The user can request to see more search results by clicking the "next" button.

Considerations regarding potential clinical use

Understanding how a predictive model works requires global interpretation. Understanding a single prediction of a model requires only local interpretation [29,30]. Our automated explaining method provides local interpretations. For clinical applications, the user of the automated explaining function is frequently a clinician who has little or no background in machine learning, can see only the prediction results but not the internal of the machine learning predictive model, cares about understanding the prediction on an individual patient but not much about how the predictive model works internally, and possibly does not even know which predictive model is used because the model is often embedded in the clinical software. In this case, it does not matter whether the explanations provided by the automated explaining function match how the predictive model works internally, as long as the explanations can help the user understand the prediction for a specific patient. For a patient predicted to have a poor outcome, our automated explaining method will give the same set of explanations regardless of which machine learning model is used to make the prediction. In the case that a deep learning model built on longitudinal data is used to make predictions, we could use the method proposed in our paper [45] to extract temporal features from the deep learning model and longitudinal data, use these temporal features to convert longitudinal data to tabular data, and then apply our automated explaining method to a predictive model built on the tabular data.

To use our automated explaining method in clinical practice, we could implement our automated explaining method together with our explanation ranking method as a software library with an application programming interface. For any clinical decision support software that uses a machine learning predictive model, we could use the application programming interface to add the automated explaining function into the software to explain the model's predictions.

Related work

As surveyed in the book [29] and the papers [30,46-48], other researchers have proposed many automated methods to explain machine learning predictions. Some of these methods are for traditional machine learning algorithms, whereas others are specifically designed for deep learning algorithms [48]. The explanations given by most of these methods are not in rule form. A lot of these methods can handle only a specific machine learning algorithm or degrade the predictive model's performance measures. None of these methods can automatically suggest tailored interventions. Ribeiro *et al*. [49,50] used rules to automatically explain any machine learning model's predictions. However, automatically recommending tailored interventions is still beyond the reach of Ribeiro *et al*.'s methods, as the rules are not generated until prediction time. In comparison, our automated explaining method mines the association rules before prediction time, gives rule-based explanations, works for any machine learning predictive model built on tabular data, does not degrade model performance, and automatically recommends tailored interventions. Compared with other types of explanations, rule-based explanations can more directly recommend tailored interventions and are easier to understand.

As surveyed in the papers [39,51,52], association rules have been used in various applications to discover interesting patterns in the data and to make predictions. Various methods have been proposed to rank the rules mined from the data set for these purposes [39,51-55]. In comparison, we mine and rank association rules to automatically explain machine learning predictions and to recommend tailored interventions.

Limitations

This work has 3 limitations that are excellent areas for future work:

1) This work used data from a single health care system. In the future, it would be good to test our explanation ranking method on data from other health care systems.
2) This work tested our explanation ranking method on predicting 1 specific target in 1 disease. In the future, it would be good to test our method on predictive modeling problems addressing other prediction targets and diseases.
3) The data set used in this work contains no information on the patients' encounters outside of UWM. This forced us to limit the prediction target to asthma hospital encounters at UWM rather than asthma hospital encounters in any health care system. In addition, the features used in this study were computed solely on the data recorded for the patients' encounters at UWM. In the future, it would be worth investigating how the top few explanations produced by our explanation ranking method would differ if we have data on the patients' encounters in other health care systems.

Conclusions

In this paper, we developed a method to rank the rule-based explanations that our automated explaining method generates for machine learning predictions. Within a negligible amount of time, our explanation ranking method performs the ranking and returns the few highest ranked explanations. These few explanations typically have high quality and low redundancy. Many of them give useful insights on various aspects of the patient's situation, which cannot be easily obtained by viewing the patient's data in the current EHR system. Both our automated explaining method and our explanation ranking method are designed based on general computer science principles and rely on no special property of any specific disease, prediction target, patient cohort, or health care system. Although only tested on the case of predicting asthma hospital encounters in asthma patients, our explanation ranking method is general and can be used for any predictive modeling problem on any tabular data set. The explanation ranking module is an essential component of the automated explaining function, which addresses the interpretability issue that deters the widespread adoption of machine learning predictive models in clinical practice. In the next few years, we plan to test our explanation ranking method on predictive modeling problems addressing other diseases as well as on data from other health care systems.

Acknowledgments

We thank Brian Kelly for useful discussions. GL was partially supported by the National Heart, Lung, and Blood Institute of the National Institutes of Health under Award Number R01HL142503. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Authors' contributions

XZ participated in designing the study and conducting literature review, wrote the paper's first draft, did the computer coding implementation, and conducted experiments. GL conceptualized and designed the study, conducted literature review, and rewrote the whole paper. Both authors read and approved the final manuscript.

Conflicts of interest

None declared.

Abbreviations

AUC: Area Under the receiver operating characteristic Curve
ED: emergency department
EHR: electronic health record
ICD-10: International Classification of Diseases, Tenth Revision
ICD-9: International Classification of Diseases, Ninth Revision
SpO_2: peripheral capillary oxygen saturation
UWM: University of Washington Medicine
XGBoost: extreme gradient boosting

List of symbols

C_r: Confidence of the association rule r
d: Decay constant
$f(d, p_i, r)$: Exponential decay function computed for the feature-value pair item p_i on the left hand side of the association rule r
f: Feature
m: Number of feature-value pair items on the left hand side of an association rule
$\max(v_r(x))$: Maximum value of the variable $v_r(x)$ across all of the rules found for the patient
$\text{mean}(f(r))$: Mean of $f(d, p_i, r)$ over all of the feature-value pair items on the left hand side of the association rule r
$\min(v_r(x))$: Minimum value of the variable $v_r(x)$ across all of the rules found for the patient
n: Maximum number of top-ranked explanations that are allowed to be displayed initially
norm(): Normalization function
N_r: Number of feature-value pair items on the left hand side of the association rule r
p: Feature-value pair item
p_i: The i-th feature-value pair item on the left hand side of an association rule
q: Number of association rules that our automated explaining method generated for the patient
r: Association rule
$score_p$: Ranking score of the feature-value pair item p
$score_r$: Ranking score of the association rule r
S_r: Commonality of the association rule r
t, t_i: Number of times that a feature-value pair item appears in the higher ranked rules
u: A value or a range
v: Outcome value
$v_r(x)$: Variable whose value on the association rule r is x
w_a: Weight for the term $\delta_{actionable}(r)$ in the rule scoring function
w_b: Weight for the term $\delta_{actionable}(p)$ in the item scoring function
w_c: Weight for the term $\text{norm}(C_r)$ in the rule scoring function
w_d: Weight for the term $\text{mean}(f(r))$ in the rule scoring function
w_g: Weight for the term $\exp(-d \cdot t)$ in the item scoring function
w_n: Weight for the term $\text{norm}(N_r)$ in the rule scoring function
w_s: Weight for the term $\text{norm}(\log_{10} S_r)$ in the rule scoring function
x: Value
$\delta_{actionable}(p)$: Indicator function for whether the feature-value pair item p is actionable
$\delta_{actionable}(r)$: Indicator function for whether the association rule r is actionable

Appendix

Table 1. A summary of the demographic and clinical characteristics of the UWM asthma patients during 2011-2017.

Characteristic	Data instances linked to no asthma hospital encounter in the succeeding 12 months (*N*=67,060), *n* (%)	Data instances linked to asthma hospital encounters in the succeeding 12 months (*N*=1,184), *n* (%)	Data instances (*N*=68,244), *n* (%)
Age			
<40	22,993 (34.29)	466 (39.36)	23,459 (34.38)
40 to 65	33,306 (49.67)	583 (49.24)	33,889 (49.66)
65+	10,761 (16.05)	135 (11.40)	10,896 (15.97)
Gender			
Male	23,647 (35.26)	551 (46.54)	24,198 (35.46)
Female	43,413 (64.74)	633 (53.46)	44,046 (64.54)
Race			
White	47,240 (70.44)	507 (42.82)	47,747 (69.97)
Black or African American	7,900 (11.78)	520 (43.92)	8,420 (12.34)
Asian	5,625 (8.39)	96 (8.11)	5,721 (8.38)
American Indian or Alaska native	1,326 (1.98)	32 (2.70)	1,358 (1.99)
Native Hawaiian or other Pacific islander	659 (0.98)	14 (1.18)	673 (0.99)
Unknown or not reported	4,310 (6.43)	15 (1.27)	4,325 (6.34)
Ethnicity			
Non-Hispanic	55,247 (82.38)	1,062 (89.70)	56,309 (82.51)
Hispanic	3,444 (5.14)	82 (6.93)	3,526 (5.17)
Unknown or not reported	8,369 (12.48)	40 (3.38)	8,409 (12.32)
Insurance			
Private	39,585 (59.03)	424 (35.81)	40,009 (58.63)
Public	28,031 (41.80)	756 (63.85)	28,787 (42.18)
Self-paid or charity	1,301 (1.94)	65 (5.49)	1,366 (2.00)
Number of years since the earliest asthma-related encounter in the data set			
≤3	59,887 (89.30)	986 (83.28)	60,873 (89.20)
>3	7,173 (10.70)	198 (16.72)	7,371 (10.80)
Asthma medication prescription			
Short-acting inhaled beta-2 agonist	46,798 (69.79)	1,010 (85.30)	47,808 (70.05)
Inhaled corticosteroid	28,263 (42.15)	626 (52.88)	28,889 (42.33)
Long-acting beta-2 agonist and inhaled corticosteroid combination	21,516 (32.08)	499 (42.15)	22,015 (32.26)
Systemic corticosteroid	18,085 (26.97)	614 (51.86)	18,699 (27.40)
Long-acting beta-2 agonist	11,919 (17.77)	374 (31.59)	12,293 (18.01)
Leukotriene modifier	7,970 (11.88)	201 (16.98)	8,171 (11.97)
Mast cell stabilizer	43 (0.06)	4 (0.34)	47 (0.07)
Comorbidity			
Anxiety or depression	19,513 (29.10)	372 (31.42)	19,885 (29.14)
Gastroesophageal reflux	12,053 (17.97)	238 (20.10)	12,291 (18.01)
Allergic rhinitis	11,277 (16.82)	172 (14.53)	11,449 (16.78)
Obesity	7,668 (11.43)	177 (14.95)	7,845 (11.50)
Sinusitis	7,172 (10.69)	89 (7.52)	7,261 (10.64)
Sleep apnea	4,468 (6.66)	88 (7.43)	4,556 (6.68)
Eczema	3,825 (5.70)	66 (5.57)	3,891 (5.70)
Chronic obstructive pulmonary disease	3,693 (5.51)	133 (11.23)	3,826 (5.61)
Cystic fibrosis	60 (0.09)	1 (0.08)	61 (0.09)
Premature birth	8 (0.01)	0 (0.00)	8 (0.01)
Bronchopulmonary dysplasia	1 (0.00)	0 (0.00)	1 (0.00)

Smoking status			
Former smoker	15,309 (22.83)	221 (18.67)	15,530 (22.76)
Current smoker	13,826 (20.62)	255 (21.54)	14,081 (20.63)
Never smoker or unknown	37,925 (56.55)	708 (59.80)	38,633 (56.61)

Table 2. A summary of the demographic and clinical characteristics of the UWM asthma patients in 2018.

Characteristic	Data instances linked to no asthma hospital encounter in the succeeding 12 months (*N*=14,426), *n* (%)	Data instances linked to asthma hospital encounters in the succeeding 12 months (*N*=218), *n* (%)	Data instances (*N*=14,644), *n* (%)
Age			
<40	4,746 (32.90)	77 (35.32)	4,823 (32.94)
40 to 65	6,683 (46.33)	111 (50.92)	6,794 (46.39)
65+	2,997 (20.78)	30 (13.76)	3,027 (20.67)
Gender			
Male	5,138 (35.62)	100 (45.87)	5,238 (35.77)
Female	9,288 (64.38)	118 (54.13)	9,406 (64.23)
Race			
White	10,103 (70.03)	110 (50.46)	10,213 (69.74)
Black or African American	1,491 (10.34)	79 (36.24)	1,570 (10.72)
Asian	1,307 (9.06)	18 (8.26)	1,325 (9.05)
American Indian or Alaska native	273 (1.89)	8 (3.67)	281 (1.92)
Native Hawaiian or other Pacific islander	129 (0.89)	2 (0.92)	131 (0.89)
Unknown or not reported	1,123 (7.78)	1 (0.46)	1,124 (7.68)
Ethnicity			
Non-Hispanic	12,370 (85.75)	196 (89.91)	12,566 (85.81)
Hispanic	830 (5.75)	20 (9.17)	850 (5.80)
Unknown or not reported	1,226 (8.50)	2 (0.92)	1,228 (8.39)
Insurance			
Private	10,692 (74.12)	108 (49.54)	10,800 (73.75)
Public	7,841 (54.35)	182 (83.49)	8,023 (54.79)
Self-paid or charity	459 (3.18)	25 (11.47)	484 (3.31)
Number of years since the earliest asthma-related encounter in the data set			
≤3	10,442 (72.38)	124 (56.88)	10,566 (72.15)
>3	3,984 (27.62)	94 (43.12)	4,078 (27.85)
Asthma medication prescription			
Short-acting inhaled beta-2 agonist	9,540 (66.13)	164 (75.23)	9,704 (66.27)
Inhaled corticosteroid	6,069 (42.07)	108 (49.54)	6,177 (42.18)
Long-acting beta-2 agonist and inhaled corticosteroid combination	4,425 (30.67)	83 (38.07)	4,508 (30.78)
Systemic corticosteroid	4,043 (28.03)	120 (55.05)	4,163 (28.43)
Long-acting beta-2 agonist	2,456 (17.02)	62 (28.44)	2,518 (17.19)
Leukotriene modifier	2,130 (14.77)	46 (21.10)	2,176 (14.86)
Mast cell stabilizer	13 (0.09)	1 (0.46)	14 (0.10)
Comorbidity			
Anxiety or depression	4,284 (29.70)	62 (28.44)	4,346 (29.68)
Gastroesophageal reflux	2,611 (18.10)	46 (21.10)	2,657 (18.14)
Allergic rhinitis	2,069 (14.34)	26 (11.93)	2,095 (14.31)
Obesity	1,579 (10.95)	25 (11.47)	1,604 (10.95)
Sinusitis	1,357 (9.41)	15 (6.88)	1,372 (9.37)
Sleep apnea	1,475 (10.22)	24 (11.01)	1,499 (10.24)
Eczema	732 (5.07)	11 (5.05)	743 (5.07)

Chronic obstructive pulmonary disease	902 (6.25)	30 (13.76)	932 (6.36)
Cystic fibrosis	17 (0.12)	0 (0.00)	17 (0.12)
Bronchopulmonary dysplasia	4 (0.03)	0 (0.00)	4 (0.03)
Premature birth	2 (0.01)	0 (0.00)	2 (0.01)
Smoking status			
Former smoker	3,453 (23.94)	41 (18.81)	3,494 (23.86)
Current smoker	3,193 (22.13)	49 (22.48)	3,242 (22.14)
Never smoker or unknown	7,780 (53.93)	128 (58.72)	7,908 (54.00)

References

1. Most Recent National Asthma Data. Centers for Disease Control and Prevention. https://www.cdc.gov/asthma/most_recent_national_asthma_data.htm, 2020. Accessed January 31, 2021.
2. Chronic respiratory diseases: Asthma. World Health Organization. https://www.who.int/news-room/q-a-detail/chronic-respiratory-diseases-asthma, 2020. Accessed January 31, 2021.
3. Nurmagambetov T, Kuwahara R, Garbe P. The economic burden of asthma in the United States, 2008-2013. Ann Am Thorac Soc 2018;15(3):348-356. PMID:29323930
4. Mays GP, Claxton G, White J. Managed care rebound? Recent changes in health plans’ cost containment strategies. Health Aff (Millwood). 2004;Suppl Web Exclusives:W4-427-436. PMID:15451964
5. Lieu TA, Quesenberry CP, Sorel ME, Mendoza GR, Leong AB. Computer-based models to identify high-risk children with asthma. Am J Respir Crit Care Med 1998;157(4 Pt 1):1173-1180. PMID:9563736
6. Caloyeras JP, Liu H, Exum E, Broderick M, Mattke S. Managing manifest diseases, but not health risks, saved PepsiCo money over seven years. Health Aff (Millwood) 2014;33(1):124-131. PMID:24395944
7. Greineder DK, Loane KC, Parks P. A randomized controlled trial of a pediatric asthma outreach program. J Allergy Clin Immunol 1999;103(3 Pt 1):436-440. PMID:10069877
8. Kelly CS, Morrow AL, Shults J, Nakas N, Strope GL, Adelman RD. Outcomes evaluation of a comprehensive intervention program for asthmatic children enrolled in Medicaid. Pediatrics 2000;105(5):1029-1035. PMID:10790458
9. Axelrod RC, Zimbro KS, Chetney RR, Sabol J, Ainsworth VJ. A disease management program utilizing life coaches for children with asthma. J Clin Outcomes Manag 2001;8(6):38-42.
10. Axelrod RC, Vogel D. Predictive modeling in health plans. Dis Manag Health Outcomes 2003;11(12):779-787. doi:10.2165/00115677-200311120-00003
11. Chen T, Guestrin C. XGBoost: A scalable tree boosting system. In: Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016 Presented at: KDD'16; August 13-17, 2016; San Francisco, CA p. 785-794. doi:10.1145/2939672.2939785
12. Tong Y, Messinger AI, Wilcox AB, Mooney SD, Davidson GH, Suri P, Luo G. Forecasting future asthma hospital encounters of patients with asthma in an academic health care system: predictive model development and secondary analysis study. J Med Internet Res 2021;23(4):e22796. PMID:33861206
13. Schatz M, Cook EF, Joshua A, Petitti D. Risk factors for asthma hospitalizations in a managed care organization: development of a clinical prediction rule. Am J Manag Care 2003;9(8):538-547. PMID:12921231
14. Grana J, Preston S, McDermott PD, Hanchak NA. The use of administrative data to risk-stratify asthmatic patients. Am J Med Qual 1997;12(2):113-119. PMID:9161058
15. Loymans RJ, Honkoop PJ, Termeer EH, Snoeck-Stroband JB, Assendelft WJ, Schermer TR, Chung KF, Sousa AR, Sterk PJ, Reddel HK, Sont JK, Ter Riet G. Identifying patients at risk for severe exacerbations of asthma: development and external validation of a multivariable prediction model. Thorax 2016;71(9):838-846. PMID:27044486
16. Eisner MD, Yegin A, Trzaskoma B. Severity of asthma score predicts clinical outcomes in patients with moderate to severe persistent asthma. Chest 2012;141(1):58-65. PMID:21885725
17. Sato R, Tomita K, Sano H, Ichihashi H, Yamagata S, Sano A, Yamagata T, Miyara T, Iwanaga T, Muraki M, Tohda Y. The strategy for predicting future exacerbation of asthma using a combination of the Asthma Control Test and lung function test. J Asthma 2009;46(7):677-682. PMID:19728204
18. Osborne ML, Pedula KL, O'Hollaren M, Ettinger KM, Stibolt T, Buist AS, Vollmer WM. Assessing future need for acute care in adult asthmatics: the Profile of Asthma Risk Study: a prospective health maintenance organization-based study. Chest 2007;132(4):1151-1161. PMID:17573515
19. Miller MK, Lee JH, Blanc PD, Pasta DJ, Gujrathi S, Barron H, Wenzel SE, Weiss ST; TENOR Study Group. TENOR risk score predicts healthcare in adults with severe or difficult-to-treat asthma. Eur Respir J 2006;28(6):1145-1155. PMID:16870656

20. Peters D, Chen C, Markson LE, Allen-Ramey FC, Vollmer WM. Using an asthma control questionnaire and administrative data to predict health-care utilization. Chest 2006;129(4):918-924. PMID:16608939
21. Yurk RA, Diette GB, Skinner EA, Dominici F, Clark RD, Steinwachs DM, Wu AW. Predicting patient-reported asthma outcomes for adults in managed care. Am J Manag Care 2004;10(5):321-328. PMID:15152702
22. Loymans RJB, Debray TPA, Honkoop PJ, Termeer EH, Snoeck-Stroband JB, Schermer TRJ, Assendelft WJJ, Timp M, Chung KF, Sousa AR, Sont JK, Sterk PJ, Reddel HK, Ter Riet G. Exacerbations in adults with asthma: a systematic review and external validation of prediction models. J Allergy Clin Immunol Pract 2018;6(6):1942-1952.e15. PMID:29454163
23. Lieu TA, Capra AM, Quesenberry CP, Mendoza GR, Mazar M. Computer-based models to identify high-risk adults with asthma: is the glass half empty or half full? J Asthma 1999;36(4):359-370. PMID:10386500
24. Schatz M, Nakahiro R, Jones CH, Roth RM, Joshua A, Petitti D. Asthma population management: development and validation of a practical 3-level risk stratification scheme. Am J Manag Care 2004;10(1):25-32. PMID:14738184
25. Forno E, Fuhlbrigge A, Soto-Quirós ME, Avila L, Raby BA, Brehm J, Sylvia JM, Weiss ST, Celedón JC. Risk factors and predictive clinical scores for asthma exacerbations in childhood. Chest 2010;138(5):1156-1165. PMID:20472862
26. Xiang Y, Ji H, Zhou Y, Li F, Du J, Rasmy L, Wu S, Zheng WJ, Xu H, Zhi D, Zhang Y, Tao C. Asthma exacerbation prediction and risk factor analysis based on a time-sensitive, attentive neural network: retrospective cohort study. J Med Internet Res 2020;22(7):e16981. PMID:32735224
27. Tong Y, Messinger AI, Luo G. Testing the generalizability of an automated method for explaining machine learning predictions on asthma patients' asthma hospital visits to an academic healthcare system. IEEE Access 2020;8:195971-195979. PMID:33240737
28. Luo G, Johnson MD, Nkoy FL, He S, Stone BL. Automatically explaining machine learning prediction results on asthma hospital visits in asthmatic patients: secondary analysis. JMIR Med Inform 2020;8(12):e21965. PMID:33382379
29. Molnar C. Interpretable Machine Learning. Morrisville, NC: lulu.com; 2020. ISBN:0244768528
30. Guidotti R, Monreale A, Ruggieri S, Turini F, Giannotti F, Pedreschi D. A survey of methods for explaining black box models. ACM Comput Surv 2019;51(5):93. doi:10.1145/3236009
31. Desai JR, Wu P, Nichols GA, Lieu TA, O'Connor PJ. Diabetes and asthma case identification, validation, and representativeness when using electronic health data to construct registries for comparative effectiveness and epidemiologic research. Med Care 2012;50 Suppl:S30-S35. PMID:22692256
32. Wakefield DB, Cloutier MM. Modifications to HEDIS and CSTE algorithms improve case recognition of pediatric asthma. Pediatr Pulmonol 2006;41(10):962-971. PMID:16871628
33. Luo G, Nau CL, Crawford WW, Schatz M, Zeiger RS, Rozema E, Koebnick C. Developing a predictive model for asthma-related hospital encounters in patients with asthma in a large, integrated health care system: secondary analysis. JMIR Med Inform 2020;8(11):e22689. PMID:33164906
34. Luo G, Nau CL, Crawford WW, Schatz M, Zeiger RS, Koebnick C. Generalizability of an automatic explanation method for machine learning prediction results on asthma-related hospital visits in patients with asthma: quantitative analysis. J Med Internet Res 2021;23(4):e24153. PMID:33856359
35. Luo G, He S, Stone BL, Nkoy FL, Johnson MD. Developing a model to predict hospital encounters for asthma in asthmatic patients: secondary analysis. JMIR Med Inform 2020;8(1):e16080. PMID:31961332
36. Luo G. Automatically explaining machine learning prediction results: a demonstration on type 2 diabetes risk prediction. Health Inf Sci Syst 2016;4:2. PMID:26958341
37. Alaa AM, van der Schaar M. Prognostication and risk factors for cystic fibrosis via automated machine learning. Sci Rep 2018;8(1):11242. PMID:30050169
38. Alaa AM, van der Schaar M. AutoPrognosis: automated clinical prognostic modeling via Bayesian optimization with structured kernel learning. In: Proceedings of 35th International Conference on Machine Learning. 2018 Presented at: ICML'18; July 10-15, 2018; Stockholm, Sweden p. 139-148.
39. Thabtah FA. A review of associative classification mining. Knowledge Eng Review 2007;22(1):37-65. doi:10.1017/S0269888907001026
40. Liu B, Hsu W, Ma Y. Integrating classification and association rule mining. In: Proceedings of the 4th International Conference on Knowledge Discovery and Data Mining. 1998 Presented at: KDD'98; August 27-31, 1998; New York City, NY p. 80-86.
41. Fayyad UM, Irani KB. Multi-interval discretization of continuous-valued attributes for classification learning. In: Proceedings of the 13th International Joint Conference on Artificial Intelligence. 1993 Presented at: IJCAI'93; August 28-September 3, 1993; Chambéry, France p. 1022-1029.
42. Luo G, Thomas SB, Tang C. Automatic home medical product recommendation. J Med Syst 2012;36(2):383-398. PMID:20703712

43. Luo G, Tang C, Yang H, Wei X. MedSearch: a specialized search engine for medical information retrieval. In: Proceedings of the 17th ACM Conference on Information and Knowledge Management. 2008 Presented at: CIKM'08; October 26-30, 2008; Napa Valley, CA p. 143-152. doi:10.1145/1458082.1458104
44. Santos RLT, Macdonald C, Ounis I. Search result diversification. Foundations and Trends in Information Retrieval 2015;9(1):1-90. doi:10.1561/1500000040
45. Luo G. A roadmap for semi-automatically extracting predictive and clinically meaningful temporal features from medical data for predictive modeling. Glob Transit 2019;1:61-82. PMID:31032483
46. Du M, Liu, N, Hu X. Techniques for interpretable machine learning. Commun ACM 2020;63(1):68-77. doi:10.1145/3359786.
47. Gilpin LH, Bau D, Yuan BZ, Bajwa A, Specter M, Kagal L. Explaining explanations: an overview of interpretability of machine learning. In: Proceedings of the 5th IEEE International Conference on Data Science and Advanced Analytics. 2018 Presented at: DSAA'18; October 1-3, 2018; Turin, Italy p. 80-89. doi:10.1109/DSAA.2018.00018
48. Samek W, Montavon G, Lapuschkin S, Anders CJ, Müller K. Explaining deep neural networks and beyond: a review of methods and applications. Proc IEEE 2021;109(3):247-278. doi:10.1109/JPROC.2021.3060483
49. Ribeiro MT, Singh S, Guestrin C. Anchors: high-precision model-agnostic explanations. In: Proceedings of the 32nd AAAI Conference on Artificial Intelligence. 2018 Presented at: AAAI'18; February 2-7, 2018; New Orleans, LA p. 1527-1535.
50. Rudin C, Shaposhnik Y. Globally-consistent rule-based summary-explanations for machine learning models: application to credit-risk evaluation. In: Proceedings of INFORMS 11th Conference on Information Systems and Technology. 2019 Presented at: CIST'19; October 19-20, 2019; Seattle, WA p. 1-19. doi:10.2139/ssrn.3395422
51. Altaf W, Shahbaz M, Guergachi A. Applications of association rule mining in health informatics: a survey. Artif Intell Rev 2017;47(3):313-340. doi:10.1007/s10462-016-9483-9
52. Pazhanikumar K, Arumugaperumal S. Association rule mining and medical application: a detailed survey. International Journal of Computer Applications 2013;80(17):10-19. doi:10.5120/13967-1698
53. Yang G, Shimada K, Mabu S, Hirasawa K. A personalized association rule ranking method based on semantic similarity and evolutionary computation. In: Proceedings of the IEEE Congress on Evolutionary Computation. 2008 Presented at: CEC'08; June 1-6, 2008; Hong Kong, China p. 487-494. doi:10.1109/CEC.2008.4630842
54. Bouker S, Saidi R, Yahia SB, Nguifo EM. Ranking and selecting association rules based on dominance relationship. In: Proceedings of the IEEE 24th International Conference on Tools with Artificial Intelligence. 2012 Presented at: ICTAI'12; November 7-9, 2012; Athens, Greece p. 658-665. doi:10.1109/ICTAI.2012.94
55. Chen MC. Ranking discovered rules from data mining with multiple criteria by data envelopment analysis. Expert Syst Appl 2007;33(4):1110-1116. doi:10.1016/j.eswa.2006.08.007