

# A Framework to Recommend Interventions for 30-Day Heart Failure Readmission Risk

Rui Liu, Kiyana Zolfaghar, Si-chi Chin, Senjuti Basu Roy, Ankur Teredesai  
 Center for Data Science, Institute of Technology, University of Washington Tacoma  
 Email: {rui.liu,kiyana,scchin,senjutip,ankurt}@uw.edu

**Abstract**—In this paper, we describe a novel framework to recommend personalized intervention strategies to minimize 30-day readmission risk for heart failure (HF) patients, as they move through the provider’s cardiac care protocol. We design principled solutions by learning the structure and parameters of a multi-layer hierarchical Bayesian network from underlying high-dimensional patient data. Next, we generate and summarize the rules leading to personalized interventions which can be applied to individual patients as they progress from admit to discharge. We present comprehensive experimental results as well as interesting case studies to demonstrate the effectiveness of our proposed framework using large real-world patient datasets on Microsoft Azure for Research platform.

## I. INTRODUCTION

Heart Failure (henceforth referred to as HF) is one of the leading causes of hospitalization, and studies [1], [2] show that many of these admissions are readmissions within a short window of time. Readmission can result from a variety of reasons, including early discharge of patients, improper discharge planning, and poor care transitions. In particular, studies [1], [2] have shown that targeted interventions during pre-discharge [3] and post discharge phases, like home based follow up [1], patient education [4], or administering appropriate procedures during the hospital stay can reduce the readmission rates considerably and improve the health outcome of the patients. Such factors that could be externally controlled (or administered) are construed as *interventions* and are applicable at different phases of a patients life cycle.

We, for the first time, attempt to go beyond risk prediction for HF, and focus on actionable intervention recommendation to aid clinicians in designing improved quality of care protocols to reduce the 30-day<sup>1</sup> readmission risk for HF. To recommend intervention strategies, we consider a multitude of factors, such as, socio-demographic factors, co-morbidities<sup>2</sup> and other diagnoses, and procedures by understanding the complex interplay between these factors and how they contribute to the 30-day readmission risk.

As an example, consider an elderly black female patient with long-standing hypertension who has developed shortness of breath and significant dependent lower extremity swelling. She undergoes an echocardiogram and other outpatient testing. It is determined that she has evidence of diastolic heart failure

and a treatment plan is initiated. Unfortunately, her symptoms continue to aggravate and she is admitted to the local hospital for further intervention. However, an alternate care pathway might have included the patient being referred to a heart failure specialty clinic after the initial diagnosis of diastolic heart failure was made. We intend to investigate all these alternatives and recommend that care pathway which is likely to minimize her readmission risk. Therefore, the problem we address is:

*Problem Definition 1:* Make personalized intervention recommendations to minimize 30-day HF readmission risk.

The underpinning of our proposed framework relies on the following four steps: (1) Since we deal with very high dimensional data involving several hundreds of factors, we first attempt to *learn the structure of the network automatically* from the data itself using different approaches. (2) Once the structure is defined, we use parameter learning [6] techniques to compute probability distribution. (3) Third, we propose novel algorithm to generate a set of intervention rules. (4) Finally, for a given patient the generated rules are summarized to offer personalized intervention recommendations.

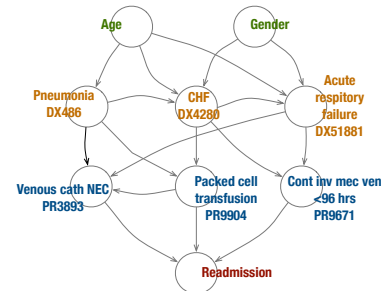


Fig. 1: A learned structure for Example 1.

The contributions of our work are:

- We initiate the study for recommending personalized interventions to minimize 30-day HF readmission risk.
- We formalize the intervention recommendation task as a hierarchical Bayesian Structure Learning problem. Furthermore, we propose multiple algorithms as solutions and summarize rules for personalized intervention.
- We present comprehensive experimental results on Microsoft Azure, as well as case studies to demonstrate the effectiveness of our proposed techniques.

Our proposed intervention recommendation framework is proposed in Section II. Sections III and IV contain our experimen-

<sup>1</sup>30-day is considered *clinically meaningful* by different healthcare services and standards [5].

<sup>2</sup>Comorbidities are specific patient conditions that are secondary to the patient’s principal diagnosis and that require treatment during the stay.

tal setting, empirical evaluation, and case studies. Section V describes the related work, and we conclude in Section VI.

## II. INTERVENTION RECOMMENDATION FRAMEWORK

We model the intervention recommendation as a network learning task using the Bayesian network learning principles. Once we have the underlying structure and the conditional probabilities, we develop an algorithm to *generate recommendation rules* based on the fitted network learned from the given dataset. These three steps constitute the *training part* of the proposed framework. During testing (recommendation evaluation), we *summarize* the generated rules and validate the effectiveness of recommendation. Figure 2 describes the high level design of the framework.

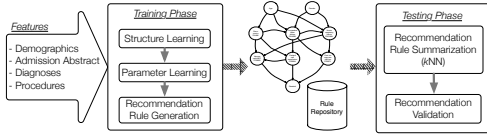


Fig. 2: Design of the proposed framework.

The running example is presented next to recommend interventions during hospitalization, but we note that the framework could be easily adapted to any other phase.

*Example 1:* Consider two socio-demographic factors : age (discretized appropriately), and gender(m/f); three binary diagnosis variables (Congestive Heart Failure (CHF) DX4280, Acute Respiratory Failure (ARF) DX51881, and Pneumonia (PN) DX486); and three binary procedure factors (Continuous Invasive Mechanical Ventilation < 96 hrs PR9671, Venous Cath NEC PR3893, Packed Cell Transfusion PR9904)<sup>3</sup>. These eight factors are predictors and we wish to learn how they relate to the likelihood of a 30-day HF readmission. The dependent variable “readmission” is a binary variable, where “Readmission=0” stands for 30 day readmission unlikely, and “Readmission=1” stands for highly-likely. For simplicity, we consider only the procedures that are intervenable.

**Bayesian Network and Relevant Notations:** Relevant notations and their interpretations are represented in Table I. A Bayesian network is a graphical representation of a probability distribution over a set of variables or factors  $\mathcal{U} = \{X_1, X_2, \dots, X_n\}$  [7], [8]. It consists of two components:

- A directed network structure as a DAG. Given Example 1, a possible structure is described in Figure 1.
- A set of probability distribution functions (PDFs), one on each node (variable), conditional on each value combination of the node’s parents. Together with the network structure, the PDFs are sufficient to represent the joint probability distribution of the domain.

$$Pr(X_1, X_2, \dots, X_n) = \prod_{i=1}^{i=n} Pr(X_i | P_a^{X_i}) \quad (1)$$

<sup>3</sup>Each diagnoses procedure has a unique code written after its name and these procedures are applicable during hospitalization.

TABLE I: Notations and Interpretations

Notation	Interpretation
$\mathcal{D}$	the dataset with $N$ data points
$X, Y, Z$	three variables
$x, y, z$	values of $X, Y, Z$ , respectively
$P_a^X$	a set of nodes that are parent of $X$
$\mathcal{X}$	the entire set of predictor variables(factors)

### A. Structure Learning

Our structure learning solution relies upon the *Causal Sufficiency Assumption* and the *Markov Assumption*[9]. We use *Constraint Based, Score-Based, and Hybrid* methods to learn the structure of the network.

1) *Constraint Based Methods:* These methods make use of the conditional independence tests using statistical tests on the data set. We use a computationally efficient algorithm, *Grow and Shrink* [10] which relies on detecting the *Markov Blanket* [11] of the variables to induce the network structure.

2) *Score-Based Methods:* Constraint-based algorithms suffer from poor “robustness”, i.e., a large effects on the output of the algorithm is observed, for small changes of the input. To overcome that shortcoming, we apply Score-based approach and apply hill climbing based greedy heuristics using Bayesian information criterion (BIC) [12] for approximation.

3) *Hybrid Approach:* We finally apply a hybrid approach to learn the network structure, namely the *max-min hill climbing algorithm* [8] which combines ideas from the both Score-Based approach and Constraint-Based Approach.

### B. Parameter Learning

After the structure of the network is constructed, the next step is to learn the parameters of the network, given the structure. Using Example 1, this step is analogous to creating pdfs to each node in the constructed network to create the conditional probability table at each node. In our implementation, we use Bayesian Parameter Estimation [13] to learn the parameter  $\theta$ . In this method, the prior distribution over  $\theta$  (i.e.,  $Pr(\theta)$ ) is known. Now the posterior distribution of  $\theta$  is calculated according to Bayes rule:

$$P(\theta|D) = \frac{Pr(D|\theta)Pr(\theta)}{\int Pr(D|\theta)Pr(\theta)d\theta}$$

### C. Recommendation Rule Generation

We make use of the inference learned by the network to perform recommendation. Using the constructed network, for each patient record  $d$ , we could compute the probability  $Pr(Readmit = 1|d)$  and  $Pr(Readmit = 0|d)$ . We describe next how to make use of these inference probabilities to generate a set of recommendation rules.

Without loss of generality, let us assume that a total of  $|\mathcal{X}'|$  of  $|\mathcal{X}|$  factors are non-interventionable, and the remaining set  $\{\mathcal{X}\} - \{\mathcal{X}'\}$  of factors could be recommended as interventions.

For each patient record  $d$  whose actual class label is 0 (i.e.,  $Readmit = 0$ ), we use only  $|\mathcal{X}'|$  attributes of record  $d$  (denoted as  $d(\mathcal{X}')$ ) and feed it through the constructed network to obtain the inference probability  $p_1$ . Then, we use

TABLE II: Attribute summary for heart failure cohort

Group	Category	Description
1	Demographics	Age, Gender, Race
2	Comorbidity;Diagnosis	21 comorbidity ;90 diagnosis
3	Utilization;Procedures	21 health service utilization flags ;70 procedures
4	Others	Length of stay, # diagnosis, # procedures, # chronic conditions

the entire patient record (with both interventionable and non-interventionable attributes, modulo the class label), and use that to make a second inference probability  $p_2$ .

$$p_1 = Pr(Readmit = 0|d(\mathcal{X}')), p_2 = Pr(Readmit = 0|d)$$

If  $p_2 > p_1$  (which indicates that our constructed model infers that the set of procedures associated with the patient input is effective in further bringing down her readmission risk), we store the set of procedures  $\{\mathcal{X}\} - \{\mathcal{X}'\}$  associated with  $d$  as the generated recommendation, given the values for the non-procedure attributes. Using Example 1, a recommendation rule in our case may look as follows:

- Rule-1: if Gender = Female & Age = 64 & diagnosis= PN & diagnosis= ARF & Readmit=0, recommended interventions (i.e, procedures) P1 (PR3893) = 1 & P2 (PR9904) = 0 & P3 (PR9671)= 1.

Similar check is performed to generate rules for patient records associated with actual class label of 1.

### III. EXPERIMENTAL RESULTS

Our empirical analyses are conducted on Microsoft Azure using 8 cores and each core with 56GB of RAM using R-studio and Python. All numbers are presented as the average of three runs.

#### A. Dataset Preparation

We use the State Inpatient Databases (SID) (<http://www.hcup-us.ahrq.gov/sidoverview.jsp>) of Washington State (referred as SID-WA for the rest of the paper) of year 2010 and 2011. SID of one year comprises four files (i.e. core file charges file, diagnosis and procedure groups file, and disease severity measures file) that provide 596 attributes in total for a single patient encounter.

We construct a heart failure cohort based on the initial dataset extracted from SID-WA. The cohort contains patients whose primary or secondary ICD9-CM diagnosis codes are listed in [14]. Initially, the cohort contains 3,908 distinct diagnosis codes and 2,049 procedure codes. In order to resolve the issue of sparsity and high dimensionality of the data, we perform *chi-square* feature selection to filter attributes that are less influential. Table II summarizes the 209 attributes used in the cohort. Unless otherwise stated, all 70 procedures and the attribute length of stay are construed as interventions.

The final heart failure cohort contains data extracted from SID-WA 2010 and SID-WA 2011. Our experiments use the 2010 data (67967 patients) for training and the 2011 data (52021 patients) for testing.

TABLE III: Number of edges and distinct rules discovered in each experiment setting

Alg	Number of edges;rules		
	30 diagnosis	60 diagnosis	90 diagnosis
HC	277; 58,864	306; 37,028	341; 51,479
GS	63; 46,501	67; 46,501	75; 46,531
HY	107; 42,465	108; 24,322	119; 24,322

#### B. Implemented Algorithms

We implement 3 different structure learning algorithms (Hill Climbing (HC), Grow-Shrink (GS), Hybrid (HY) and compare it with a baseline implementation using Logistic Regression (LR) technique. We vary the number of diagnosis attributes (30, 60, 90). The LR algorithm does not “discover” any causal relationship between the variables, but “learns” the association between different interventions and readmission risk. Based on that, it ranks the interventions and returns them. We note that other ranking based methods, such as, Odds Ratio [15] does not lend itself naturally to our problem settings. We also implement two natural baselines for summarization module (referred to Section III-D2). For brevity, we present a subset of results. The omitted ones are akin to the ones that are presented.

#### C. Learning Phase

During learning, in order to allow further flexibility and observe the relationship between procedures(i.e, interventions) and readmission, we white-list edges from influential procedure nodes (based on the results of feature selection) to readmission node. We use heuristics to construct the blacklist which contains out-links from readmission, in-links to Demographics attributes and in-links to any attributes in Comorbidity and Diagnosis coming from any attributes in Utilization and Procedures. Finally, we obtain 12657 edges in the black-list and 5 edges in the white-list. Table III shows the complexity of the networks and the number of distinct rules discovered for each experiment setting.

#### D. Validation Phase: Summarization & Scalability

The network structure is learned on Microsoft Azure using 10,000 records randomly sampled from the training data and the parameters are using all 68K records.

The validation phase uses the *summarization module*. The intuition is to be able to discover a set of patients “akin” to the given patient and aggregate their interventions to generate recommendation for her. Given an input patient record (only with non-procedure attribute), we find out a set of  $k$  rules which gives rise to the highest similarity with the input using K-Nearest Neighbor search (K-NN) [15]. Based on these  $k$ -rules, we create a summary rule, by taking the majority voting of the suggested recommendations.

1) *Evaluation Measures*: We design our experiments in order to assess the effect of the three structure learning algorithms described in Section II and the effect of various numbers of diagnosis attributes on the quality of evaluation. We use four metrics to evaluate our experiment results: 1) the number of exact matches from the rules of the test data (HIT); 2) the Jaccard index between the recommendation procedure vector and the actual observed procedure vector (JAC); 3)

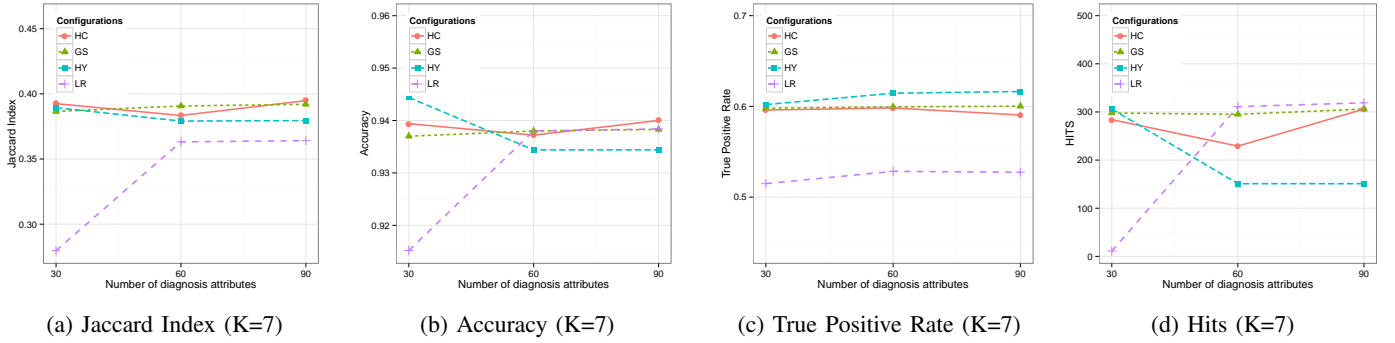


Fig. 3: Effectiveness of different algorithms for intervention recommendation; x-axis varies the number of diagnoses attributes and y-axis captures the average of a respective quality measure.

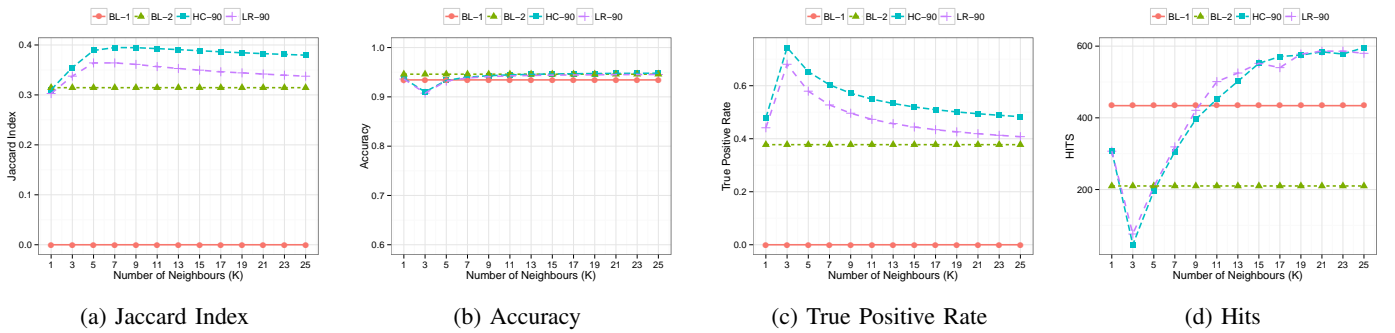


Fig. 4: Effectiveness of summarization module; x-axis varies  $k$  and y-axis captures the average of a respective quality measure.

accuracy of the recommendations (ACCY); 4) True positive rate (TPR) or Recall<sup>4</sup>. For each pair of a set of recommended procedures and a set of observed procedures, we define a true positive ( $tp$ ) case, if a recommended procedure appears in the observed procedure set. We define a true negative ( $tn$ ) case, if a non-recommended procedure does not appear in the observed procedure set. False positive ( $fp$ ) occurs when the recommended procedure does not appear in the observed set. False negative ( $fn$ ) occurs when the non-recommended procedure actually appears in the observed procedure set.

2) *Effectiveness of Summarization*: In this set of experiments, we vary  $k$  (number of nearest neighbor) to create summary and observe the effectiveness of the generated recommendations based on the 4 quality measures described above. We also implement two baseline algorithms, referred to as BL-1 and BL-2. For brevity, we compare them with only one structure learning algorithm (HC) with 90 attributes and the LR baseline. Since, our data suggests that most of the time no procedure is recommended to a patient, BL-1 simply suggests 0 procedure. On the other hand, BL-2 does not use learning, but consider the entire patient population and suggests the

<sup>4</sup>Note that Precision is not relevant in our settings, as it captures the actually correct recommendations out of all recommendation.

top-3 most frequent procedures always. Figure 4 contain the results that exhibits that the baseline algorithms are clearly inappropriate for TPR. Unsurprisingly, with increasing  $k$ , the quality improves only upto certain extent, that suggests that a reasonably small number of similar rules (small  $k$ ) is adequate for effective intervention recommendation.

3) *Recommendation Effectiveness & Discussion*: The experimental results are presented in Figure 3 for  $k = 7$ . Understandably the HIT values are in the lower side for all the algorithms, while the other three measures (especially Accuracy) are reasonable and demonstrates the effectiveness of our proposed methods. Figure 3 and Table V show our experiment results. We perform paired t-test to further understand the statistical significance of the obtained results. The significance level is set to  $p$ -value  $< 0.05$ . The results indicates that, for JAC and ACCY, HC significantly outperform GS, HY as well as LR (for the same number of attributes). The results also demonstrate that JAC and ACCY of HY with 30 attributes outperform the other two variants – HY with 60 attributes and HY with 90 attributes. On the other hand, HC with 90 attributes achieved significantly better results for JAC and ACCY compared to other two structure learning algorithms (GS and HY) and to the settings of using 30 and 60 attributes. In terms of TPR, all the Bayesian Network based algorithms

TABLE V: Recommendation evaluation for the alternative whitelist. \* denotes the value that is significantly higher

Alg		HITS	JAC	ACCY	TPR
GS	orig	284	0.3067*	0.9355*	0.4796
	alt	217	0.3044	0.9307	0.4876*
HY	orig	291	0.3082*	0.9355*	0.4834
	alt	217	0.3031	0.9341	0.4840

show a significant statistically improvement over the LRs and especially the results of HY are the statistical best. On the other hand for GS and HY, the structure learning algorithms with 90 attributes outperforms the ones with 30 and 60 attributes in TPR while 30 is the best for HC. These results corroborate that based on the underlying algorithm and the input diagnoses, the effectiveness of different algorithm varies for the task of intervention recommendation.

4) *Effects of Prior Knowledge:* In order to investigate the effect of prior clinical knowledge encoded in the whitelist, we perform an experiment using a different set of arcs in the whitelist for the setting of 30 diagnosis attributes with GS and HY algorithms. Naturally, LR algorithm is not applicable here anymore. Our alternative whitelist replaces two procedures that are highly associated with readmission with two less correlated procedures. We apply paired t-test with significance level at 0.05 to examine whether differences exist between the original and alternative whitelist.

Table III-D4 compares the averaged results between the original and the alternative whitelist. It shows that using the original whitelist produces significantly better results in JAC and ACCY for both GS and HY structure learning algorithms. However, in the case of GS, we observe that the alternative whitelist has small but significant higher TPR.

#### IV. CASE STUDIES : RECOMMENDATION EFFECTIVENESS

For clinical validation, we involve a cardiologist to clinically validate the effectiveness of the recommended interventions for 5 real patient cases randomly chosen from the data. The cases are summarized in Table VI, using different combinations of demographics and diagnosis as the inputs to evaluate the procedure recommendations. All three network learning algorithms are used in this process and the generated rules are summarized using  $k = 7$ . Table VII shows the results. We ask two questions specifically - a) How likely this procedures would be used for this patient? b) How likely the recommended procedures could reduce the readmission risk? These two questions are referred to as Ques 1 and Ques 2 respectively in Table VIII. The cardiologist rates in a scale of 1–5 (least likely to most likely). These results in Table VIII demonstrate that our proposed rules are reasonably effective.

#### V. RELATED WORK

To the best of our knowledge, no prior work has investigated the intervention recommendation problem for heart failure. Several effective solutions to predict the risk of readmission for congestive heart failure [16], [14], [17] is known that do not have any easy extension for intervention recommendation.

TABLE VIII: Ratings provided by the cardiologist describing the effectiveness of the generated recommendations in Table VII. The results indicate that our automated recommendation generation process is quite effective.

Case	Ques 1 Score - HC, GS, HY	Ques 2 Score - HC, GS, HY
1	5,5,5	4,4,4
2	4,4,4	4,4,4
3	3,3,3	3,3,3
4	4,4,4	4,4,4
5	4,4,4	4,4,4

Existing research proposes different modeling to enable decision support for generating treatment plan for other diseases such as coronary diseases [18], ulcers [19], sepsis [20], and depression [21]. Unlike us, none of these work deal with the problem of high dimensionality, scale, or multiple layers. Unlike us, they do not generate recommendation rules, nor do they perform large scale validation.

A recent work [22] has leveraged “big” healthcare claims data for the knowledge discovery process. Although our studied problem is fundamentally different; nevertheless, our proposed framework could benefit from the proposed large scale data analysis solutions. A recent research [23] studies the problem of identifying risk signals of potential adverse drug reactions (ADRs) through Bayesian network. It is easy to observe that our effort is orthogonal to these related work. Nevertheless, we wish to explore more advanced large scale data mining techniques in the future.

#### VI. CONCLUSION

We investigate the problem of recommending interventions to minimize 30-day risk of readmission for heart failure patients. Our proposed solution relies on learning the structure and parameters of a hierarchical Bayesian network from the given data. After that, we propose algorithms to generate rules and summarize them that could be used to recommend interventions. Our implementation addresses the scalability and high dimensionality issues using implementation on Windows Azure. Our experimental results as well as case studies demonstrate the effectiveness of the proposed framework.

#### VII. ACKNOWLEDGMENTS

We acknowledge the Microsoft Azure for Research team for their support in this research. Additionally, we are thankful to cardiologists Dr. Jane Wilcox at Northwestern Medical School for her participation in the case study.

#### REFERENCES

- [1] M. D. Naylor *et al.*, “Comprehensive discharge planning and home follow-up of hospitalized elders: A randomized clinical trial,” *JAMA*, 1999.
- [2] M. W. Rich *et al.*, “A multidisciplinary intervention to prevent the readmission of elderly patients with congestive heart failure,” *New England Journal of Medicine*, 1995.
- [3] J. Schneider *et al.*, “A medication discharge planning program measuring the effect on readmissions,” *Clinical Nursing Research*, 1993.

TABLE IV: Statistical significance of quality results of Figure 3 is further explained using paired-t-test. \* denotes that the result is significantly better than the others among the three variations of attribute numbers. † denotes whether the result is significantly better than the others among the three structure learning algorithms and logistic regressions.

	HC			GS			HY			LR		
	30	60	90	30	60	90	30	60	90	30	60	90
HIT	284	229	306	298	295	306	306	151	151	11	311	319
JAC	0.3925†	0.3834†	0.3948*†	0.3865	0.3907	0.3920*	0.3894*	0.3791	0.3795	0.2797	0.3632	0.3641*
ACCY	0.9394	0.9372	0.9400*†	0.9370	0.9380	0.9383*†	0.9444*†	0.9344	0.9344	0.9152	0.9380	0.9384*
TPR	0.596	0.5983	0.5905*	0.5979	0.5995	0.6002*	0.6020†	0.6145†	0.6164*†	0.5149	0.5284*	0.5274

TABLE VI: Case study scenarios: LOS is length of stay, CHRON is # chronic conditions

Case	gender, age	LOS, CHRON	Diagnoses
1	Male,40-49	1,2	Diabetes, hypertension, Peripheral vascular disorders, Congestive Heart Failure, Coronary Atherosclerosis Of Autologous Vein Bypass Graft
2	Male,40-49	3,3	Chronic pulmonary disease, hypertension, Subendocardial infarction, Atherosclerosis Of Autologous Vein Bypass Graft
3	Male,30-39	2,2	hypertension, Subendocardial infarction, Ventricular Fibrillation
4	Male,50-59	1,2	Hypothyroidism,Chronic Systolic Heart Failure, Other Specified Forms Of Chronic Ischemic Heart Disease
5	Female,50-59	2,1	Hypertension, Acute Systolic Heart Failure, Subendocardial infarction.

TABLE VII: Recommended utilities and procedures for the case studies in Table VI; unsurprisingly, the different belief network may still reach the same conclusions for a given patient in some cases. These procedures includes Major operating room procedure (A), Cardiac Catheterization Lab (B),Electrocardiogram (C),Other Implants (D),CCU (E),Chest X-Ray (F),Echocardiology (G),ICU (H),Emergency Room (I),Pacemaker (J),INSERT DRUGELUTING CRNRY AR (PR-3607),Percutaneous Transluminal Coronary Angioplasty (PR-0066), INSERT 1 VASCULAR STENT (PR-0045), LEFT HEART CARDIAC CATH (PR-3722),Lt Heart Angiocardigram (PR-8853)

Case	HC	GS	HY
1	includes procedures A,B,C,D, PR-3607 .PR-0066 .PR-0045	Same as HC plus procedure E	Same as GS
2	includes procedures A,B,C,D,F,G,H,PR-3722,PR-8853,PR-0066	Same as HC	same as HC
3	includes procedures A,B,C,F,H,I,PR-3607,PR-0066,PR-8853,PR-3722,PR-0045	includes procedures A,B,C,D,E,I, PR-3607, PPR-0066,PR-8853, PR-0045	includes procedures A,B,C,D,F, I,H,PR-3607,PR-0066,PR-8853,PR-3722,PR-0045
4	includes procedures A,C,D,F,H,J	Same as HC plus PR-0066	Same as HC
5	includes procedures A,B,C,D,F,H,PR-3607,PR-0066,PR-8853,PR-3722, PR-0045	Same as HC	same as HC

- [4] T. M. Koelling *et al.*, “Discharge education improves clinical outcomes in patients with chronic heart failure,” *Circulation*, vol. 111, no. 2, pp. 179–185, Jan. 2005. [Online]. Available: <http://circ.ahajournals.org/content/111/2/179>
- [5] R. Amarasingham *et al.*, “An automated model to identify heart failure patients at risk for 30-day readmission or death using electronic medical record data,” *Medical Care*, 2010.
- [6] G. E. Box and G. C. Tiao, *Bayesian inference in statistical analysis*. John Wiley & Sons, 2011, vol. 40.
- [7] R. E. Neapolitan, *Learning bayesian networks*. Pearson Prentice Hall Upper Saddle River, 2004.
- [8] I. Tsamardinos *et al.*, “The max-min hill-climbing bayesian network structure learning algorithm,” *Machine learning*, 2006.
- [9] P. Spirtes and C. Meek, “Learning bayesian networks with discrete variables from data,” in *KDD*, 1995.
- [10] F. Bromberg *et al.*, “Efficient markov network structure discovery using independence tests,” in *SDM*, 2006.
- [11] F. V. Jensen and T. D. Nielsen, *Bayesian networks and decision graphs*. Springer, 2007.
- [12] W. Li *et al.*, “Marker selection by akaike information criterion and bayesian information criterion.” *Genetic Epidemiology*, 2000.
- [13] S. C. Kramer and H. W. Sorenson, “Bayesian parameter estimation,” *Automatic Control, IEEE Transactions on*, 1988.
- [14] N. Meadem *et al.*, “Exploring preprocessing techniques for prediction of risk of readmission for congestive heart failure patients,” in *DMH-KDD*, 2013.
- [15] J. Han and M. Kamber, *Data mining: concepts and techniques*, 2006.
- [16] K. Zolfaghar *et al.*, “Risk-o-meter: an intelligent clinical risk calculator,” in *KDD*, 2013.
- [17] —, “Big data solutions for predicting risk-of-readmission for congestive heart failure patients,” in *IEEE Big Data*, 2013.
- [18] J. Bittl *et al.*, “Bayesian methods affirm the use of percutaneous coronary intervention to improve survival in patients with unprotected left main coronary artery disease,” *Circulation*, 2013.
- [19] I. Cho *et al.*, “Using ehr data to predict hospital-acquired pressure ulcers: A prospective study of a bayesian network model,” *medical informatics*, 2013.
- [20] E. Gultepe *et al.*, “From vital signs to clinical outcomes for patients with sepsis: a machine learning basis for a clinical decision support system,” *JAMA*, 2013.
- [21] K. et.al, “Estimating mental states of a depressed person with bayesian networks,” in *Contemporary Challenges and Solutions in Applied Artificial Intelligence*. Springer, 2013.
- [22] V. Chandola *et al.*, “Knowledge discovery from massive health-care claims data,” in *SIG-KDD*, 2013.
- [23] R. Harpaz *et al.*, “Empirical bayes model to combine signals of adverse drug reactions,” in *SIG-KDD*, 2013.