Utilizing Electronic Medical Records to Discover Changing Trends of Medical Behaviors Over Time*

Liangying Yin[1]; Zhengxing Huang[1,3]; Wei Dong[2]; Chunhua He[3]; Huilong Duan[1,3]

[1]College of Biomedical Engineering and Instrument Science, Zhejiang University, Hangzhou, China;
[2]Department of Cardiology, Chinese PLA General Hospital, Beijing, China;
[3]College of Medical Engineering Technology, Xinjiang Medical University, Urumqi, China

Summary

Objectives: Medical behaviors are playing significant roles in the delivery of high quality and cost-effective health services. Timely discovery of changing frequencies of medical behaviors is beneficial for the improvement of health services. The main objective of this work is to discover the changing trends of medical behaviors over time.

Methods: This study proposes a two-steps approach to detect essential changing patterns of medical behaviors from Electronic Medical Records (EMRs). In detail, a probabilistic topic model, i.e., Latent Dirichlet allocation (LDA), is firstly applied to disclose yearly treatment patterns in regard to the risk stratification of patients from a large volume of EMRs. After that, the changing trends by comparing essential/critical medical behaviors in a specific time period are detected and analyzed, including changes of significant patient features with their values, and changes of critical treatment interventions with their occurring time stamps.

Results: We verify the effectiveness of the proposed approach on a clinical dataset containing 12,152 patient cases with a time range of 10 years. Totally, 135 patients features and 234 treatment interventions in three treatment patterns were selected to detect their changing trends. In particular, evolving trends of yearly occurring probabilities of the selected medical behaviors were categorized into six content changing patterns (i.e, 112 growing, 123 declining, 43 updown, 16 down-up, 35 steady, and 40 jumping), using the proposed approach. Besides, changing trends of execution time of treatment interventions were classified into three occurring time changing patterns (i.e., 175 early-implemented, 50 steady-implemented and 9 delay-implemented).

Conclusions: Experimental results show that our approach has an ability to utilize EMRs to discover essential evolving trends of medical behaviors, and thus provide significant potential to be further explored for health services redesign and improvement.

Correspondence to:
Zhengxing Huang
College of Biomedical Engineering and
Instrument Science
Zhejiang University
Zhou Yiqin building 510
Zheda road 38#
Hangzhou 310008
China
E-mail: zhengxinghuang@zju.edu.cn

* Supplementary material published on our website https://doi.org/10.3414/ME16-01-0047

1. Introduction

Medical behaviors, as key components in the care of patients with specific clinical problems, greatly influence the clinical efficiency and medical costs of delivered therapies for patients [1, 2, 3, 4]. Medical behaviors are always patient-linked and time-specific [4]. ▶Figure 1 provides an example of how medical behaviors are arranged in a patient's treatment process. Notably, medical behaviors keep improving with the continuous development of medical technologies and regulation of medical strategies. As well, patient status keeps changing over time, which will end up with changes in medical behaviors in a patient's treatment process. Collectively, all these factors lead to changes in medical behaviors, which definitely influence the clinical outcome and medical quality of delivered health services. Undoubtedly, discovering the changing trends of medical behaviors in patients' treatment processes could open new opportunities for developing more efficient health services. Acknowledged by hospital managers and clinical experts, healthcare organizations can develop more time-efficient and cost-effective health services if they can detect changes of medical behaviors over time, e.g., they can re-adjust essential medical behaviors and rearrange their occurring time stamps in a specific disease's treatment process through utilizing detected changes.

A lot of work has been proposed for medical behavior analysis [2, 3, 4, 5, 6]. However, most of them are concerned with tracking typical treatment behaviors in treatment processes to obtain evidences for health services optimization or verify the efficiency of delivered health services for improving clinical outcomes, e.g., pathway

Figure 1
A detailed example of medical behaviors in a patient's treatment process.

modelling [2], variation prediction [6], latent treatment pattern mining [3, 4, 5], and effects of treatment processes on actual clinical practice [14, 15, 16], etc. To the best of our knowledge, detecting the changing trends of medical behaviors over time, as determining factors for developing/optimizing efficient treatment processes, has been significantly neglected. In our previous work [13], we have made efforts to detect differences between frequent medical behavior patterns generated respectively from two different time-periods. While this work is helpful to discover significant changes raised from critical treatment patterns in specific treatment phases (i.e., admission, before operation, operation, after operation and discharge), it lacks the ability to detect changing trends of essential/critical medical behaviors over time elapsing.

To capture nontrival changing trends of medical behaviors over time in a refine manner, we should first figure out what are typically patient features associated with a particular treatment process and what are conditioned treatment interventions given specific patient features during treatment processes of patients. Both patient features and treatment interventions compose a typical treatment pattern that forms the backbone of a disease's treatment process protocol. Note that a treatment pattern is one found in most executed patient traces of the same treatment procedure for patients presenting with specific symptoms, and is characterized by a set of representative patient features and treatment interventions. In this sense, medical behaviors are always changing over time with respect to both patient features and treatment interventions contained in treatment patterns.

Recently, with the growing availability of large electronic medical record (EMR) databases, clinical researchers are increasingly interested in the secondary use of these observational data [17]. Essentially, EMRs reflect the temporal nature of treatment processes, and thus may unlock the ability to analyze change frequencies of medical behaviors and identify critical changing patterns [22]. We believe that such patterns may provide important insights into how medical behaviors change over time and the effects of implemented interventions.

Motivated by the demand of capturing the evolving trends of medical behavior over time and the availability of observational data in EMRs, this work adopts a novel approach to detect evolving trends of medical behaviors in treatment processes. The proposed approach consists of two steps: 1) we employ a probabilistic topic model to discover essential/critical treatment patterns from a large volume of electronic medical records (EMRs) within different time periods; 2) we detect meaningful changing trends by comparing discovered medical behaviors in a specific time period, including changes of significant patient features with their values (e.g., patient symptoms, vital signs, lab test results, etc.), and changes of critical treatment interventions with their occurring time stamps (e.g., clinical examinations, medical orders, and surgeries, etc.). Contrary to most of existing methods that based on experiences and knowledge of clinical experts, the proposed method can automatically detect changing trends of medical behaviors from EMRs rather than manually interpret large amount of patient cases piece by piece, which is time effective and effort-saving. Note that a treatment pattern is assumed to correspond to a specific severity of patient conditions. This assumption has been confirmed by our clinical collaborators. In particular, they claim that a treatment pattern reflects specific clinical conditions of patients within the same clinical risk level. In this sense, we
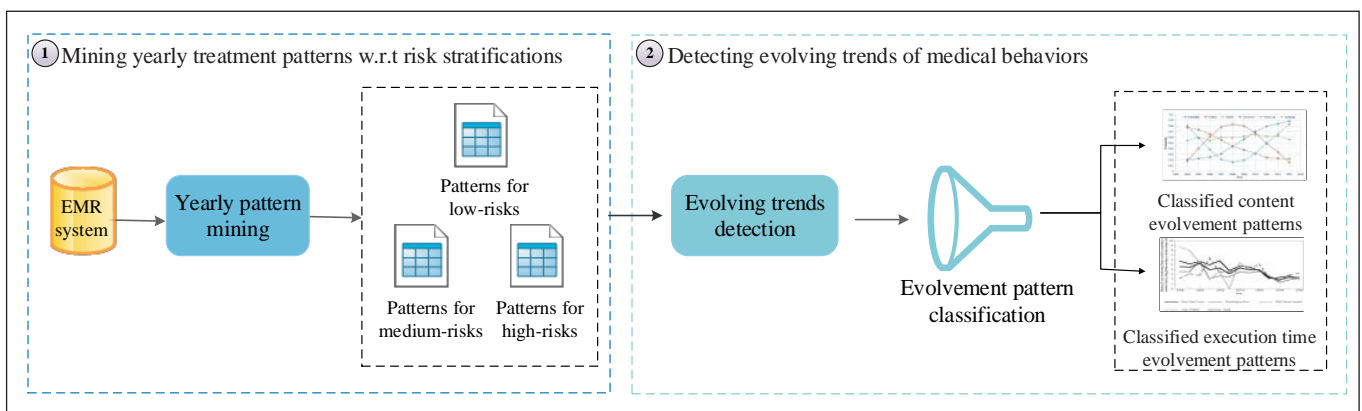
Figure 2   Detailed steps of the proposed approach.

argue that our approach can cluster patients into different risk classes based on their clinical conditions and then detect changing trends of medical behaviors with regard to different risk tiers. The proposed approach has been evaluated on a real clinical dataset pertaining to the unstable angina collected from Chinese PLA General hospital and from 2004 to 2013.

The reminder of this paper is organized as follows. Section 2 describes detailed steps for discovering evolving trends of medical behaviors over time. Experimental setting, results and assessment results are carefully summarized in Section 3. Finally, main discoveries and clinical implications of the proposed approach are given as well as the limitations of the proposed work in Section 4.

2. Method

The purpose of this study is to automatically discover essential evolving trends of medical behaviors over time by exploring a large volume of EMRs. In detail, the proposed approach consists of two steps, i.e., treatment pattern discovery, and changing pattern detection of medical behaviors, as shown in ▶ Figure 2.

2.1 Treatment Pattern Mining with Respect to Risk Stratifications of Patients

Generally speaking, patients following a particular treatment process protocol may have complex conditions, urging the treatment plans to be complied with their clinical status (i.e., risk level). Besides, a patient's clinical status may be very time-dynamic during his/her treatment process. For example, with the delivery of effective therapy procedures during the hospitalization, a medium-risk patient may turn into a low-risk one. All these factors in turn lead to a patient's EMR to be represented by a mixture of risk-specific treatment patterns, and each treatment pattern is characterized by a set of patient features and their values as well as treatment interventions and their occurring time stamps, corresponding to a specific risk tier. Latent Dirichlet Allocation (LDA) models have
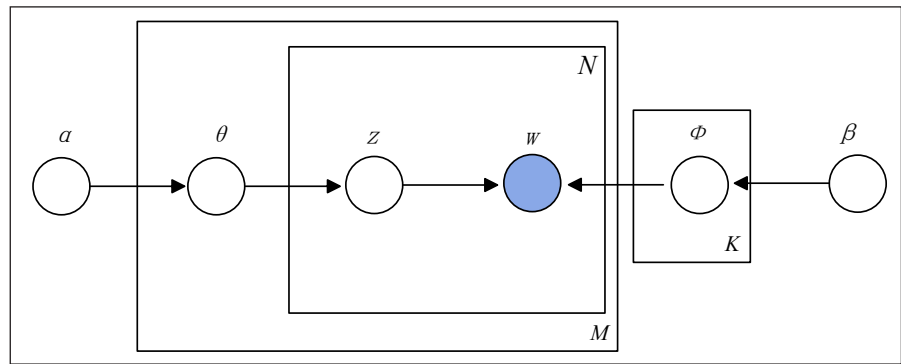

Figure 3 Graphical model reprentation of LDA [7].

proven effective in revealing the mixture risk-specific patterns of patients and stratify them into different risk tiers from temporal EMRs [10]. In this regard, we propose employing LDA to discover the latent treatment patterns jointly with clinical risk stratification on individual patient samples.

In general, it takes 1–2 years for hospitals to updates their treatment process protocols. Given this, the data collected from EMRs need to be split into a series of yearly segments, and each segment corresponds to the data recorded in one particular year. Then, we can employ a probabilistic topic model to infer essential treatment patterns from yearly datasets. As we mention above, a treatment pattern, describing the essential medical behaviors of patients within the same risk level during their hospitalizations [21], is a mixture distribution of multiple medical behavior items. Note that medical behaviors consist of both patient features and treatment interventions, and are regularly recorded in EMRs. Thus, the corresponding risk level of an extracted treatment pattern can be indicated implicitly by the mixture distributions of medical behaviors (i.e., corresponding risk factors) contained in that pattern.

LDA is a three-level hierarchical Bayesian model originally designed for collections of text corpora [7]. A document can be represented as a sequence of words denoted by $d = (w_1, w_2, …, w_N)$ while a corpus is a collection of $|D|$ documents denoted by $D = \{d_1, d_2, … d_{|M|}\}$. It represents the documents as a mixture of latent topics, where each topic is characterized by a distribution over words. The dependency of all variables can be derived concisely with the

plate notation shown in ▶ Figure 3, where α is the parameter of the Dirichlet prior on the per-document topic distribution, β is the parameter of the Dirichlet prior on the per-topic word distribution, θ is the topic distribution for a specific document, Φ is the word distribution for K topics, z is a specific topic and w is an individual word. For more details about LDA, please be referred to [5, 7].

Shifting to clinical settings, each EMR corresponds to a particular patient trace of the hospitalization, and consists of a set of patient feature-value pairs {f<a,v>} (where a is the feature type and v is the feature value) and treatment intervention-occurring time stamp pairs {e<a,t>} (where a is the intervention type and t is the occurring time stamp of a). Patients with different risk levels may follow different treatment patterns, and these essential treatment patterns are exactly corresponding to treatment procedures for patients in different risk clusters. Hence, we can employ LDA to extract essential treatment patterns with respect to different risk levels. Formally, a yearly dataset is a collection of EMRs, and each EMR σ follows one or several specific treatment patterns in his/her treatment process. Based on the LDA model, a treatment pattern $z (z \in Z)$ is a mixture of medical behaviors W, which can be either a feature-value pair represented as $f = (a, v)$ or an intervention-time stamp pair represented as $e = (a, t)$. In this regard, each patient's EMR is a distribution over treatment patterns and can be represented by $\theta_\sigma (\theta\sigma \sim \text{Dir}(\alpha), Dir(\alpha)$ is the Dirichlet distribution of prior parameter α) and the occurring probability of a treatment pattern z

Table 1   Definitions of parameters.

Parameter	Definition
α	Prior parameter of Dirichlet distribution
β	Prior parameter of Dirichlet distribution
Z	Treatment pattern numbers
σ	EMR of a particular patient sample
z	A specific treatment pattern
θ_σ	Distribution of each σ, $\theta_\sigma \sim Dir(\alpha)$, $Dir(\alpha)$ is the Dirichlet distribution of prior parameter α
$\theta_{\sigma,z}$	Occurring probability of a treatment pattern z in a care journey σ
φ_z	Distribution of each treatment pattern z, $\varphi_z \sim Dir(\beta)$, $Dir(\beta)$ is the Dirichlet distribution of prior parameter β
$\varphi_{z,f(a,v)}$	Occurring probability of each feaure-value in a treatment pattern z
$\varphi_{z,e.a}$	Occurring probability of each treatment intervention in a treatment pattern z
$P(\varphi_{z,e(a,t)})$	Daily occurring probability of a treatment intervention in treatment pattern z
T	The largest length of stay of yearly collection of EMRs
M	Number of feaure-value pairs/intervention-occurring time pairs in treatment pattern z
α_*	Summing results of α over Z times
β_*	Summing results of β over M times

in a care journey σ can be represented as $\theta_{\sigma,z} = P(z|\sigma)$, each pattern is a distribution over medical behaviors and can be represented by φ_z ($\varphi_z \sim Dir(\beta)$, $Dir(\beta)$ is the Dirichlet distribution of prior parameter β). The occurring probability of each medical behavior in a treatment pattern can be represented as $\varphi_{z,w} = P(w|z)$ ($\varphi_{z,w} = \varphi_{z,f(a,v)}$ where w is a feature-value pair, $\varphi_{z,w} = \varphi_{z,e(a,t)}$ where w is an intervention-occurring time pair), which we are interested in to derive.

In this study, Gibbs sampling [19, 20] is employed to iteratively draw patient samples from the conditional distribution for each treatment pattern z in each piece of EMR. After marginalizing out the parameter, the conditional distribution, which represents the occurring probability of the ith medical behavior in σ assigned to treatment pattern z, can be represented as follows:

$$P(z_i^\sigma = z|\mathbf{z}_{-i,\sigma}, w, \alpha, \beta) \propto \frac{c_{\sigma,z}^{-i} + \alpha}{c_{\sigma,*}^{-i} + \alpha_*} \times \frac{n_{z,w}^{-i} + \beta}{n_{z,*}^{-i} + \beta_*} \quad (1)$$

Where $z_i^\sigma = z$ represented the assignments of the ith medical behavior in σ to treatment pattern z, $z_{-i,\sigma}$ represented all pattern assignment excluded the ith medical behavior, $C_{\sigma,z}^{-i}$ is the number of treatment pattern z assigned to σ not including

the current instance, $n_{z,w}^{-i}$ is the count of the ith medical behavior w assigned to treatment pattern z not including the current one. Furthermore, dot $*$ defines the summing operation at the corresponding index, e.g., $n_{z,*}$ is the summing of medical behaviors in pattern z.

Based on symmetric Dirichlet distribution, we can estimate $\theta_{\sigma,z}$ and $\varphi_{z,w}$ as follow:

$$\theta_{\sigma,z} = \frac{c_{\sigma,z} + \alpha}{c_{\sigma,*} + \alpha_*} \quad (2)$$

$$\varphi_{z,w} = \frac{n_{z,w} + \beta}{n_{z,*} + \beta_*} \quad (3)$$

Where $c_{\sigma,z}$ is the number of current treatment pattern z in EMR σ, $c_{\sigma,*}$ is the count of all treatment patterns occurred in EMR σ, α_* is the summing results of α over Z times; $n_{z,w}$ describes the number of medical behavior w occurred in treatment pattern z, $n_{z,*}$ is the number of all medical behaviors in treatment z, β_* is summing results of β over M (M is the number of medical behaviors in treatment pattern z) times. Notably, the occurring probabilities of feature-value pairs and intervention-time stamp pairs are separately calculated based on Eq. (3). Therefore, occurring probabilities of all patient feature-value/intervention-time

stamp pairs for patients in a specific risk level add up to 1. For more information about the inference process, please refer to [5, 7, 12, 19].

It also should be noted that a treatment intervention assigned to a specific treatment pattern may be performed in different time stamps. In this regard, the occurring probability of a particular treatment intervention in a specific treatment pattern $\varphi_{z,e.a}$ is the sum of probabilities of that treatment intervention executed in different time stamps. $\varphi_{z,e.a}$ can be calculated as follows:

$$\varphi_{z,e.a} = \sum_{t=1}^{T} \varphi_{z,e(a,t)} \quad (4)$$

Where T is the largest length of stay of the yearly collection of EMRs with respect to a specific CP, implicating that occurring probabilities of all treatment interventions in a treatment pattern corresponding to a specific risk level also add up to 1. Differently, the daily occurring probability of a treatment intervention in a specific treatment pattern $P(\varphi_{z,e(a,t)})$ can be computed as follows:

$$P(\varphi_{z,e(a,t)}) = \frac{\varphi_{z,e(a,t)}}{\sum_{\tau=1}^{T} \varphi_{z,e(a,t)}},$$

$$(t \in \{1, 2, \dots T\}) \quad (5)$$

Where t is the exactly executed time-stamp of a specific type of treatment intervention. Detailed definitions of used parameters in equations are addressed in ►Table 1.

In this study, yearly datasets of EMRs are considered to follow the same distributions. By employing the LDA-based approach, we can derive yearly essential treatment patterns from yearly dataset composing of both patient features and treatment interventions.

Once treatment patterns are discovered from yearly datasets, we need to align them into different risk tiers for the further purpose of detecting changing trends of medical behaviors embedding in them. Note that we assume that those highly correlated patterns extracted from different yearly datasets represent the same/similar medical behaviors and should be aligned into the same risk tier. To this end, we apply two state-of-the-art statistical methods (i.e., Pearson test, and student's t-test) to meas-

ure the correlational significance between a pair of patterns, and then align the highly correlated patterns extracted from different yearly datasets into the same risk tier. Notably, patient features from the yearly derived patterns can reflect the severity of patient conditions, and thus are important and direct indicators of patient risk tiers. And the discovered patterns that disclose the least, moderately, and most severe patient conditions can be separately aligned into low-, medium- and high-risk clusters, respectively. It must mention that the aligned patterns with their specific risk labels have been evaluated by our clinical collaborators after careful review of the aligned pattern groups. They indicate the aligned results are consistent with their judgments.

As we have indicated above, a patient's clinical status may be very complex and dynamic. In fact, with the delivery of effective treatment interventions during the treatment process execution, a medium-risk patient may turn into a low-risk one during his/her hospitalization. This in turn leads to a patient's journey to be represented by a mixture of risk-specific treatment patterns. To this end, we propose employing LDA to discover latent treatment patterns with their distributional probabilities on individual patient samples. In this sense, a patient can be categorized into specific risk tiers using the inferred distributional probabilities of treatment patterns on that patient sample.

2.2 Detecting Changing Medical Behaviors

2.2.1 Detecting Changing Trends of Significant Patient Feature-value Pairs with Respect to Specific Risk Levels

Essential and critical patient features and their values disclose detailed patient conditions and thus reflect the severity of patient conditions. To figure out the changing trends of patient conditions with respect to a specific risk level, it is necessary to detect the changing frequencies of those patient feature-value pairs. The yearly occurring probabilities of all patient feature-value pairs can be derived from Eq. (3) where w is a feature-value pair and $\varphi_{z,w} = \varphi_{z,f(a,v)}$.
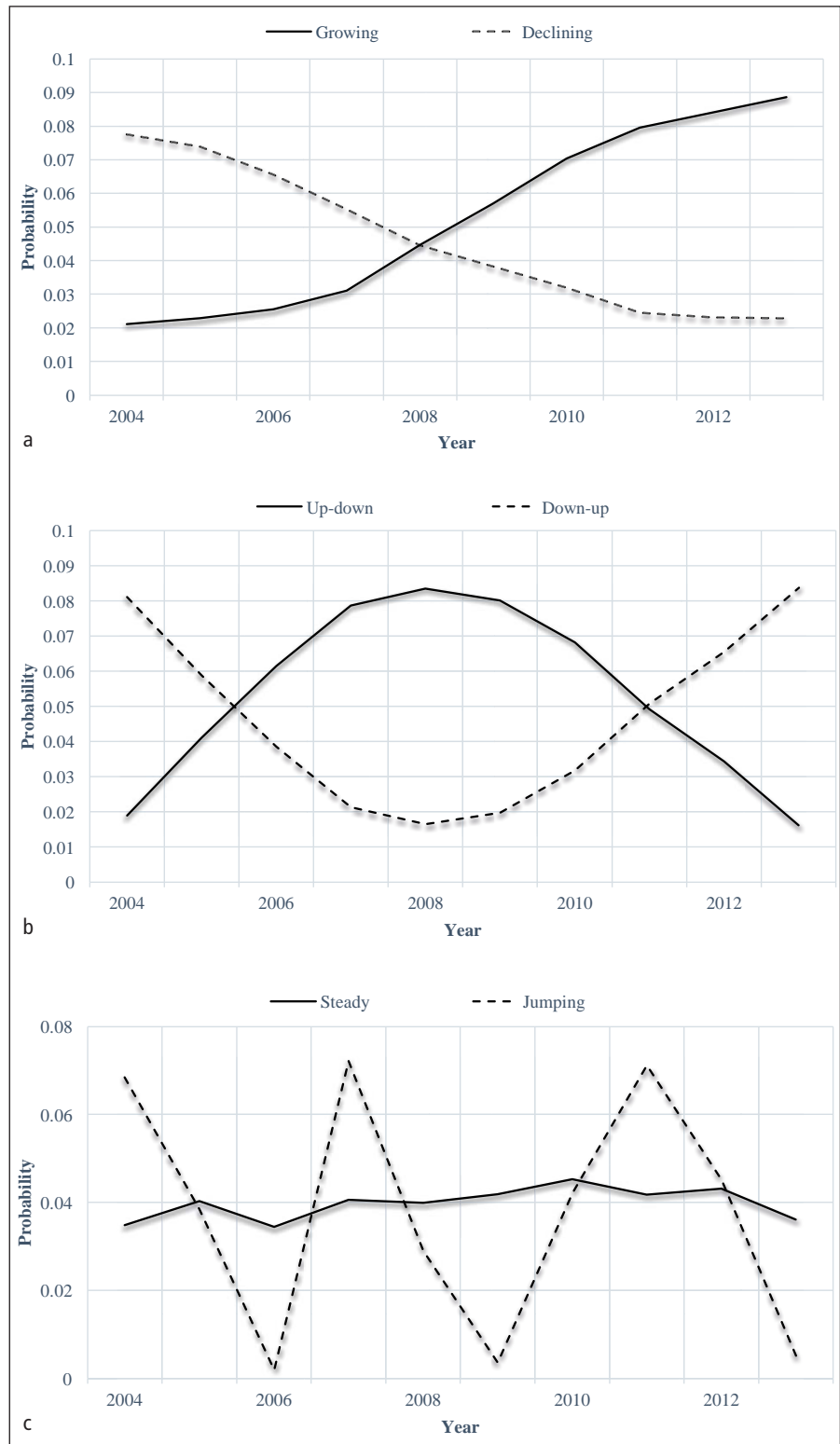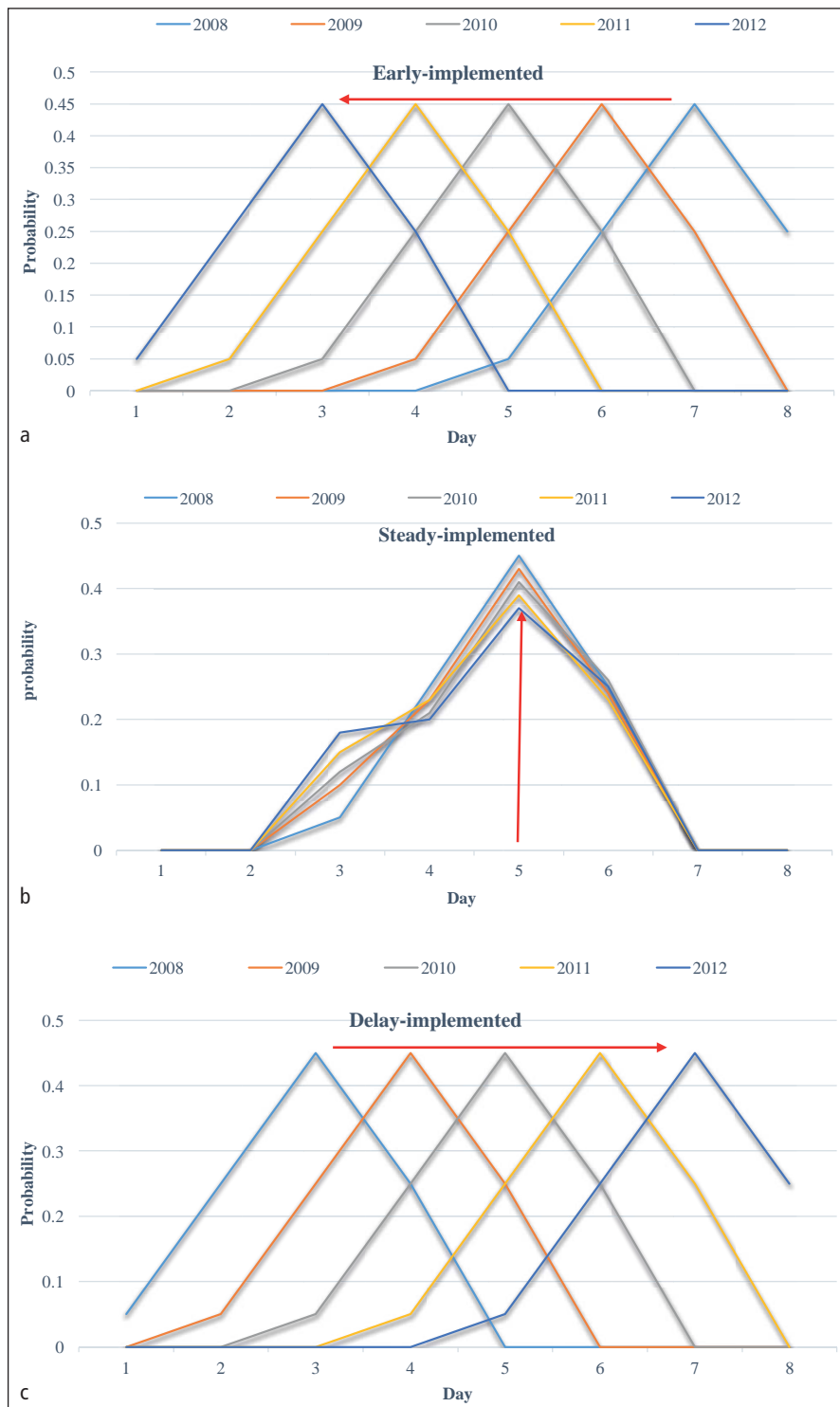

Figure 4 Three pairs of conversely changing patterns of feature-value pairs. (a) Growing and declining pattern, (b) Up-down and down-up pattern, and (c) Steady and jumping pattern.

Figure 5   Changing patterns of implementation time of treatment interventions. (a) Early-implemented, (b) Steady-implemented, and (c) Delay-implemented.

After graphically demonstrating yearly occurring probabilities of patient feature-value pairs, we found that changes of patient feature-value pairs follow certain changing patterns. So, these evolving trends of patient feature-value pairs can be clustered into different categories based on their changing trends. As shown in

▶ Figure 4, the changing trend of a feature-value pair can be classified as it follows:

- Growing changing patterns: the curve is increased with time elapsing (as shown in ▶ Figure 4(a))
- Declining changing pattern: the curve is decreased with time elapsing (as shown in ▶ Figure 4(a))
- Up-down changing pattern: the curve is increased with time elapsing at first, and then decreased with further elapsing of time (as shown in ▶ Figure 4(b))
- Down-up changing pattern: the curve is decreased with time elapsing at first, and then increased with further time elapsing (as shown in ▶ Figure 4(b))
- Steady changing pattern: the curve is remained stable regardless of time elapsing (as shown in ▶ Figure 4(c))
- Jumping changing pattern: the curve irregularly changes with obvious amplitude of variation as time elapsed (as shown in ▶ Figure 4(c))

Notably, these six changing patterns can be classified into three converse pairs, i.e., growing and declining changing patterns, up-down and down-up changing patterns, steady and jumping changing patterns. As shown in ▶ Figure 4, two changing patterns of medical behaviors in each pair are following totally different changing trends.

2.2.2 Detecting Changing Trends of Treatment Interventions with Respect to Specific Risk Levels

Obviously, conditioned treatment interventions are determinant factors for medical quality and medical cost. In this sense, detecting the changing frequencies of treatment interventions is beneficial to the improvement of medical service and control of medical cost. The yearly occurring probability of a pair of treatment intervention-occurring time stamp can be calculated from Eq. (3) where w is an intervention-occurring time stamp pair and $\varphi_{z,w} = \varphi_{z,e(a,t)}$. Then, by employing Eq. (4), we can derive the occurring probabilities of unique treatment interventions with regard to a specific risk level. Similarly with patient feature-value pairs, evolving trends of yearly occurring probabilities of treatment interventions can be categorized into either of six
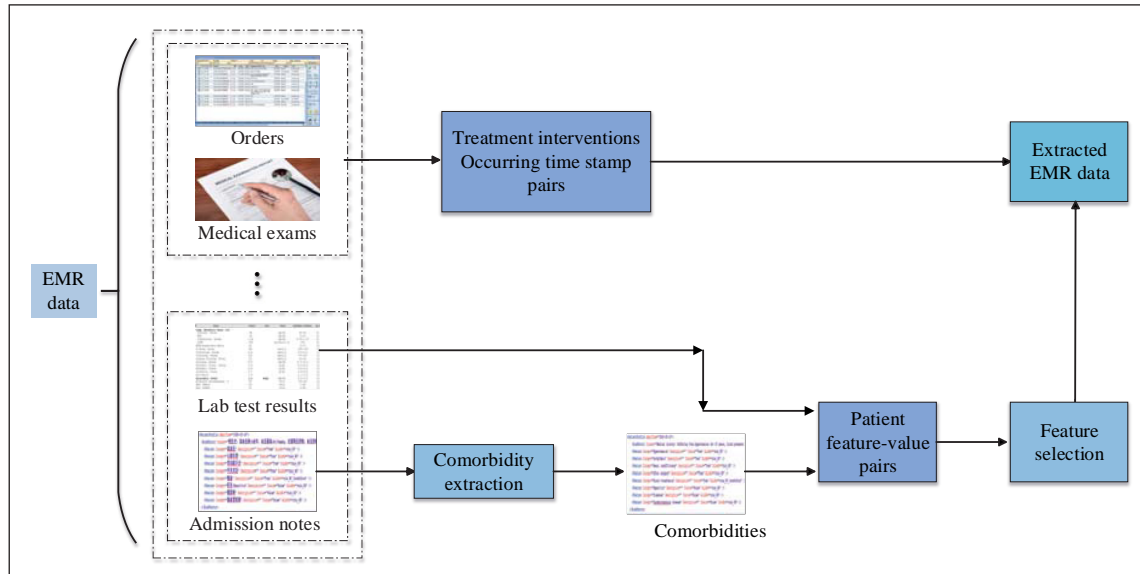
Figure 6
Patient data extraction from EMR system.

changing patterns, i.e.,growing, declining, steady, up-down, down-up and jumping changing pattern, as shown in ▶ Figure 4.

2.2.3 Detecting Changing Trends of Execution Time of Treatment Interventions

Treatment interventions defined in a treatment process protocol usually are sequences of specific execution time stamps. More specifically, they are scheduled on a daily basis. However, in actual clinical practice, execution time stamps of many treatment interventions are often inconsistent with the treatment process protocol for a specific disease and vary over time. For example, the treatment intervention "Discharge", originally performed in the 10th to 14th day after admission in 2006, takes place in advance to the 8th day after admission in 2013. Note that occurring time stamps of treatment interventions are intimately associated with the length of stay (LOS) of patients. Conclusively, daily duration is the proper granuity to sequence treatment interventions. To this end, it will be helpful to derive evolving trends of daily occurring probabilities of treatment interventions over time with respect to specific risk levels. The daily occurring probability of a treatment intervention in a treatment pattern can be calculated using Eq. (5). By analyzing changes of execution time stamps of treatment interventions, three

possible changing patterns are enumerated, i.e., early-implemented, steady-implemented, delay-implemented, as shown in ▶ Figure 5.

- Early-implemented: the curve of daily occurring probability moves forward as time goes, as indicated by the red dashed arrow in ▶ Figure 5(a).
- Steady-implemented: the curve of daily occurring probability remains stable over time, as indicated by the red dashed arrow in ▶ Figure 5(b).
- Delay-implemented: the curve of daily occurring probability moves afterward as time goes, as indicated by the red dashed arrow in ▶ Figure 5(c).

3. Experiments and Results

In this section, we present a case study of unstable angina to evaluate the feasibility of the proposed method on detecting significant evolving trends of treatment behaviors over time. Firstly, we introduce the collected clinical dataset of unstable angina in Section 3.1. And then we describe how to select the significant changing trends of medical behaviors from the experimental dataset in Section 3.2. Details of the selected changing trends are presented and analyzed in Section 3.3. Finally, we elaborate clinical assessment results of the selected changing trends and their clinical implications in Section 3.4.

3.1 Experimental Dataset

With the approvement of the research IRB in Chineses PLA General hospital, EMRs of patients who followed the unstable angina CP from the Department of Cardiology were collected from the EMR system with a time range of ten years (i.e., from 2004 to 2013), to demonstrate the feasibility of the proposed approach in identifying changing trends of medical behaviors over time. Unstable angina, as a type of irregular angina pectoralis, is generally classified as a type of acute coronary syndromes [22]. As an extremely common clinical disease occurred in aged people, it's a disease whose treatment process is typically thought of evolving trends of treatment behaviors over time (e.g., the intended surgery day is affected by patients' status) [7]. Unstable angina is a complex disease with many comorbidities and complications, which may lead patients to different risk levels and make it suitable for us to study evolving trends of medical behaviors by considering the risk stratifications of unstable angina patients. Clinical risk level reflects the severity of patient conditions. And it is capable of determining the chance of experiencing unwanted outcome [19, 21]. And it is commonly used for unstable angina patients to identify their risk levels, which facilitates to provide specific and accurate treatment interventions for patients. Notably, laboratory test results, medical exams

Table 2   Demographic attributes of the experimental dataset.

Year	Traces	Male	Female	Average age	Average stay	Feature types	Interven-tion types
2004	657	517	140	64.9	15.87	106	280
2005	1022	775	247	63.1	11.6	123	303
2006	1268	925	343	65.3	11.34	135	317
2007	1445	1040	405	62.6	9.04	136	349
2008	1694	1214	480	62.2	9.25	141	350
2009	1541	1092	449	63.0	9.03	135	333
2010	1003	707	296	61.6	7.5	83	307
2011	1470	1056	414	61.8	6.53	97	344
2012	1670	1167	503	61.8	7.14	93	370
2013	382	252	130	60.9	7.12	66	231
Total	**12152**	**8745**	**3407**	**62.7**	**9.14**	**225**	**606**

Table 3   Changing patterns of medical behaviors detected from the experimental dataset.

Risk levels	Patient features			Treatment interventions		
	low	medium	high	low	medium	high
Growing	8	17	15	23	25	24
Declining	8	17	20	22	24	32
Steady	5	0	0	10	9	11
Up-down	3	8	8	3	14	7
Down-up	6	0	0	3	5	2
Jumping	2	8	10	3	7	10

as well as admission note are selected in each patient's EMR to exhibit those medical behaviors implemented in the treatment process (as shown in ▶Figure 6). In this study, all examinations, orders were chosen as treatment interventions of the unstable angina treatment process. Contrary to this, only those comorbiditites and lab tests that are comparatively related to unstable angina were chosen as patient features. As shown in ▶Figure 6, comorbidities of patients were extracted from admission notes recorded in the EMR system. ▶Table 2 demonstrates detailed demographic attributes of the experimental datasets. Typically, an unstable angina treatment process is arranged in a day-by-day format, enumerating the essential/critical medical behaviors (e.g. examination, lab test, surgery, medication, and care, etc) that spread the whole length of hospital stay.

3.2 Selection of Significant Changing Trends of Medical Behaviors

The described method in section 2 is directly employed for changing trends detection of unstable angina patients. Accordingly, the extracted dataset was divided into 10 sub-datasets to detect yearly treatment patterns. Note that based on the guidance document enacted by National Health and Family Planning Commission of China, it takes about 1–2 years for hospitals to updates their treatment process protocols of selected diseases. Taking this into consideration, it proper to select yearly duration as the time granularity for medical pattern learning.

In experiments, dirichlet prior parameters α and β are set to 0.1, pattern number Z is set to 3 for 10 sub-datassets. Based on

patient features, derived treatment patterns of unstable angina patients can be classified into different risk levels. Specifically, derived patterns reflect the least, moderately, and most severe patient conditions can be separately categorized into low-, medium-, and high-risk levels, respectively.

Note that we assume that treatment patterns extracted from different yearly datasets and with highest correlations should be aligned into the same risk tier. To this end, we employed three state-of-the-art statistical methods (i.e., Pearson test, Levene's test for equality of variances, and t-test for equality of means) to measure the correlational significance between a pair of patterns, and then aligned highly correlated patterns into the same risk level. The results of correlational analysis on between discovered treatment patterns are exhibited in Table 1 of ▶Online Appendix. In particular, we found that aligned patterns are significantly correlated with each other. Conversely, those patterns at different risk levels are weakly or moderately correlated than those aligned ones. Besides, the Levene's test and t-test were also applied to validate our assumption. Comparatively, there exists a significant difference in variances among patterns at the different risk tiers, as shown in Table 2 of ▶Online Appendix. Moreover, values of means among patterns at the different risk levels are much smaller than that of aligned patterns, as shown in Table 3 of ▶Online Appendix. Though some of the significance values of variances and means among patterns at different risk levels are greater than the frequently used p-value threshold (i.e., 0.05), it's quite obvious that significance values are much smaller than that of these aligned patterns in the same risk level, which are quite enough to discriminate these discovered patterns. By examining the statistical analysis results, it is quite evident for us to align the discovered treatment patterns with respect to the same risk tier.

Once yearly essential treatment patterns have been derived, yearly occurring probabilities of all feature-value pairs, intervention-occuring time pairs and treatment interventions can be figured out to assist the analysis of changing trends of medical behaviors over time. Note that it does not need extra efforts to analyze all types of

Table 4
Changing patient feature-value pairs detected from the experimental dataset with respect to different risk levels.

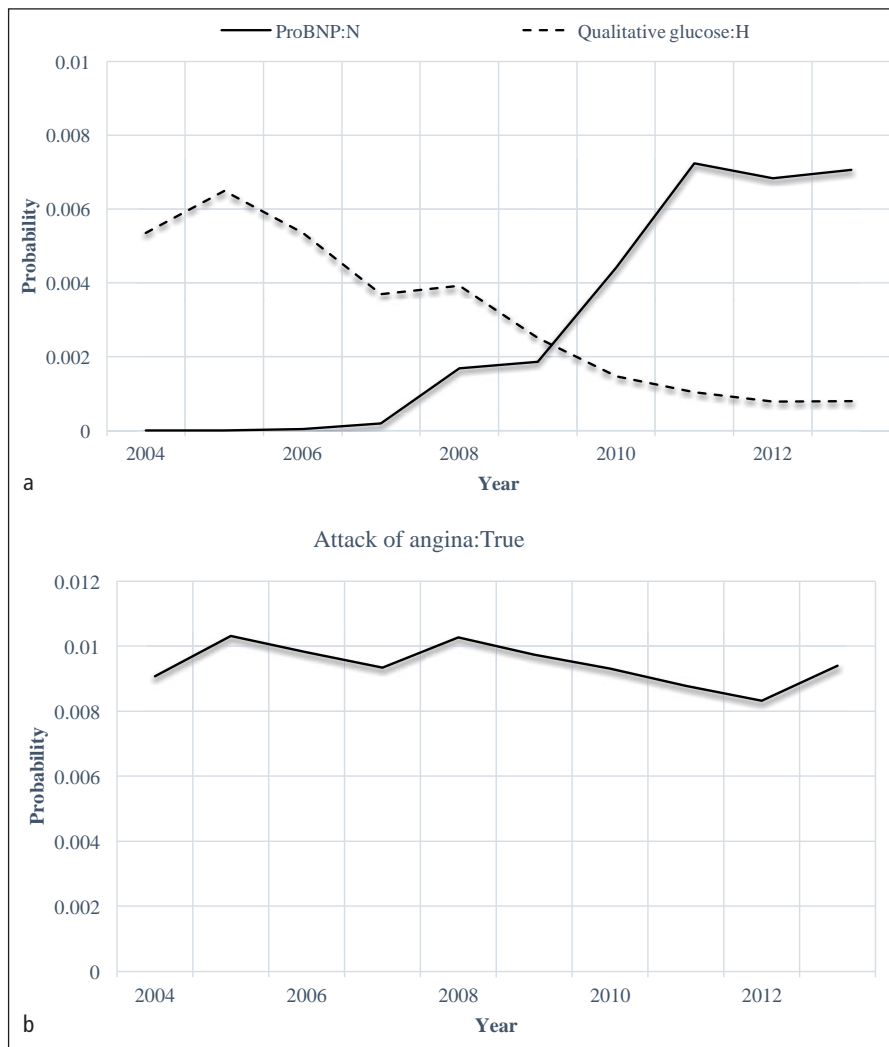| Low risk	
Growing	Total cholesterol :L, Atherosclerosis :True, Quantitative determination of Isoenzyme of creatine kinase :N, Glucose :L, Mean platelet volume :L, Creatine kinase :L, ProBNP :N, Triglycerides :L
Declining	Gender:Male, Creatine kinase:N, Triglycerides:H, Unsaturated iron binding capacity:L, Hypertension:True, Qualitative urine test:H, TIBC:L, Total cholesterol:H
Steady	Sodium:L, Creatinine:L, Age,Attack of angina:True, Lactate dehydrogenase:L
Up-down	High-density lipoprotein cholesterol:H, Low-density lipoprotein cholesterol:H, Isoenzyme of creatine kinase:N
Down-up	Hemoglobin:L, Potassium:H, Platelet:L, Troponin T:N, Potassium:L, Low-density lipoprotein cholesterol:N
Jumping	Mean corpuscular hemoglobin :L, Hemoglobin :L
Medium risk	
Growing	Attack of angina:True, After PCI:True, Cardiac insufficiency:True, Atherosclerosis:True, TEG-MA:L, Age, Acute coronary syndrome:True, Myocardial bridge:True, TEG-E:L, High-sensitivity C-reactive protein:N, TEG-K, Platelet hematocrit, Platelet volume distribution width, History of coronary heart disease:True, TEG-R:L, After CABG:True, U-mALb/Ucr:N
Declining	Hypertension:True, Diabetes:True, Prostate disease:True, Hyperlipidemia:True, Renal insufficiency:True, Tumor:True, Cerebral infarction:True, Artery stenosis:True, Hyperlipidemia:True, Pulmonary disease:True, Sick sinus syndrome:True, Ischemic cardiomyopathy:True, Insufficient blood supply:True, Anemia:True, Abnormal glucose tolerance:True, Bleeding:True, Renal cyst:True
Up-down	Isoenzyme of creatine kinase:N, Unsaturated iron binding capacity:L, Troponin T:N, Fecal transferrin: Negative Quantitative determination of Isoenzyme of creatine kinase:N, 2 hours glucose tolerance test, ProBNP:N, Serum ferritin:L
Jumping	Troponin T:H, Cardiac arrhythmia:True, Fasting serum insulin, Qualitative urine test:N, Urinary potassium:L, Platelet aggregation:L, Urine creatinine:L, Cardiac enlargement:True
High risk	
Growing	Troponin T:H, Attack of angina:True, After PCI:True, Atherosclerosis:True, Platelet volume distribution width, Platelet hematocrit, TEG-K, History of coronary heart disease:True, Acute coronary syndrome:True, MCHC:L, After CABG:True, 2 hours glucose tolerance test, Low-density lipoprotein cholesterol:N, Urine creatinine:L, Age
Declining	Hypertension:True, Cerebral infarction:True, Diabetes:True, Prostate disease:True, Renal insufficiency:True, Tumor:True, Thrombosis:True, Hyperlipidemia:True, Thrombocytopenia:True, Insufficient blood supply:True, Hyperuricemia:True, Ischemic cardiomyopathy:True, Valvular heart disease:True, Sick sinus syndrome:True, Hypokalemia:True, Anemia:True, Artery stenosis:True, Renal cyst:True, Abnormal glucose tolerance:True, Bleeding:True
Up-down	Quantitative determination of Isoenzyme of creatine kinase:N, Mean platelet volume:H, Creatine kinase:L, Mean platelet volume:L, Isoenzyme of creatine kinase:N, High-sensitive C-reactive protein:N ,Troponin T:N, Hematoma:True
Jumping	Cardiac insufficiency:True, Hyponatremia:True, 2-hour postprandial serum insulin, Cardiac arrhythmia:True, Fasting serum insulin, Plasma renin activity, High-sensitive C-reactive protein:H, Myocardial bridge:True, Serum ferritin:L ProBNP:N

medical behaviors regarding the treatment process of a specific disease, and only essential/critical medical behaviors in each treatment pattern should be selected as targets for changing pattern detection. To this end, the union of yearly 20 top-ranked patient feature-value pairs ($\varphi_{z,f(a,v)}$) and 50 top-ranked treatment interventions ($\varphi_{z,e}$) of each pattern were selected. These selected patient feature-value pairs and treatment interventions clearly embrace almost all critical/essential medical behaviors with

respect to specific risk levels of a disease. Accordingly, 32 patient feature-value pairs and 64 treatment interventions in low-risk pattern, 50 patient feature-value pairs and 84 treatment interventions of medium-risk pattern, 53 patient feature-value pairs and 86 treatment interventions in high-risk pattern were selected to detect their changing trends over time.

3.3 Experimental Results

3.3.1 Results of Changing Patient Feature-value Pairs

►Table 6 displays classification results of yearly change trends of selected patient feature-value pairs with respect to each risk level of unstable angina patients. It should be noted that feature values of some lab tests are categorized into 3 different classes: L (Low), N (Normal) and H (High). As shown in ►Table 3, both growing and de-

Figure 7   Changing trends of example feature-value pairs. (a) "ProBNP:N" and "Qualtitive glucose:H" for low risk patients, and (b) "Attack of angina: True" for low risk patients.

ing patient feature-value pairs caused by the replacement of new types of lab tests due to the development of medical technology. For example, "*Creatine kinase*" is gradually substituted by "*Troponin T*" in clinical practice. Regarding steady patient features, they tend to be associated with obligatory lab test results or comorbidities of unstable angina patients, e.g., "*Attack of angina:True*" (▶Figure 7 (b)) is an important feature of low-risk unstable angina patients. (3) Differently, feature-value pairs following jumping changing pattern are hard to confirm whether they are meaningful for the clinical practice. (4) As well, it's difficult to determine whether up-down, and down-up feature-value pairs are comparatively relevant features of patients or not.

Normally, "*Age*" is a significant demographic attribute of patients with a specific disease. The average age distribution of patients in different risk clusters is demonstrated in ▶Figure 8. It is clear that the yearly average age of patients (i.e. ▶Figure 8 (a), ▶Figure 8(b) and ▶Figure 8 (c)) in corresponding risk levels keep decreasing over time, indicating that patients are inclined to suffer from unstable angina in early age, regardless of their risk stratifications. All changes of patient feature-value pairs need to be further assessed by clinicians to confirm whether they have clinical significance.

3.3.2 Results of Changing Treatment Interventions

Classification results of yearly evolving trends of selected treatment interventions are shown in ▶Table 5, as well. It's obvious that growing, declining and steady changing patterns are dominant change patterns of treatment interventions in all three clusters. Comparatively, there are only a few treatment interventions whose changing trends follow down-up, up-down, and jumping change patterns over time in each risk cluster. Moreover, we can notice that the distributions of all six types of changing patterns are basically similar to each other, regardless of risk levels of unstable angina patients. The detailed detection results of categorized change trends of selected treatment interventions in three risk clusters are exhibited in ▶Table 8.

clining changing patterns are two dominant ones in all three clusters while there are less patient feature-value pairs that follow up-down, down-up and jumping changing patterns. Regarding steady patient features, they only appear in the low-risk cluster, while there are more up-down, down-up and jumping patient feature-value pairs in the medium- and high-risk clusters than that in the low-risk cluster. ▶Table 4 demonstrates detailed examples of classified changing trends of selected patients feature-value pairs.

By analyzing these significant detected changing trends, we have the following discoveries: (1) growing patient feature-value pairs might be associated with obligatory

lab test results or emerging comorbidities of unstable angina, e.g., "*ProBNP:N*" (▶Figure 7(a)) as a growing feature-value pair, is growing into an important comorbidity of low-risk unstable angina patients, implicating that low-risk patients are less likely to suffer from cerebral infarction. (2) Conversely, declining patient feature-value pairs might be associated with fading symptoms of unstable angina patients, e.g., "*Qualtitive glucose:H*" (▶Figure 7(a)) discovered as a declining feature-value pair in low-risk treatment pattern, gradually becomes an insignificant comorbidity of low-risk patients, indicating that patients with low risk are less likely to be diagnosed with diabetes as time elapsing. However, there are declin-

By analyzing detected changes, we can better understand the clinical value of these discovered changing trends. Finally, we have the following findings from the experimental results: (1) Growing treatment interventions are caused by the increase of execution frequency. It means that they might be essential for the treatment, such as "*Stent implantation*" (▶Figure 9(a)) as a fairly effective treatment behavior for unstable angina, is gradually becoming essential for the therapy and treatment of unstable angina patients. (2) Contrary to growing treatment interventions, declining ones may be caused by the decrease of execution frequency. Note that a declining treatment intervention might indicate that it is not used in the treatment process protocol anymore and could be eliminated ultimately. For example, "*Angiotensin-converting enzyme inhibitors (ACEI)*" (▶Figure 9(a)), as a significant declining treatment behavior, is gradually eliminated for the unstable angina treatment process, implicating that this kind of drugs are unnecessary for the treatment of patients. (3) Some interventions may be replaced by other ones over time, consequently there are correlations between some declining and growing treatment interventions, i.e. a declining treatment intervention may be replaced by a growing one. For example, the declining intervention "*Thrombosis examination*" (▶Figure 9(b)) is eventually substituted by the growing intervention "*Blood coagulation examination*" (▶Figure 9(b)). (4) For steady treatment interventions, they are essential medical behaviors regardless of changes of occurring time stamps, e.g., "*Anti-coagulation drugs*" (▶Figure 9(c)), as one of the most essential kind of drugs for low-risk patients, remains stable as time elapsed. However, it's difficult to determine whether these jumping treatment interventions are essential to the treatment process, e.g. "*X-ray*" (▶Figure 9(c)) as a jumping treatment intervention, is not essential for low-risk unstable angina patients. (5) With respect to up-down, down-up and jumping treatment interventions, they all have observable changing trends and may be induced by different clinical factors. For example, "*Glucose injection*" is a down-up intervention. And it may be caused by drug abuse, or inaccurate data recording in the
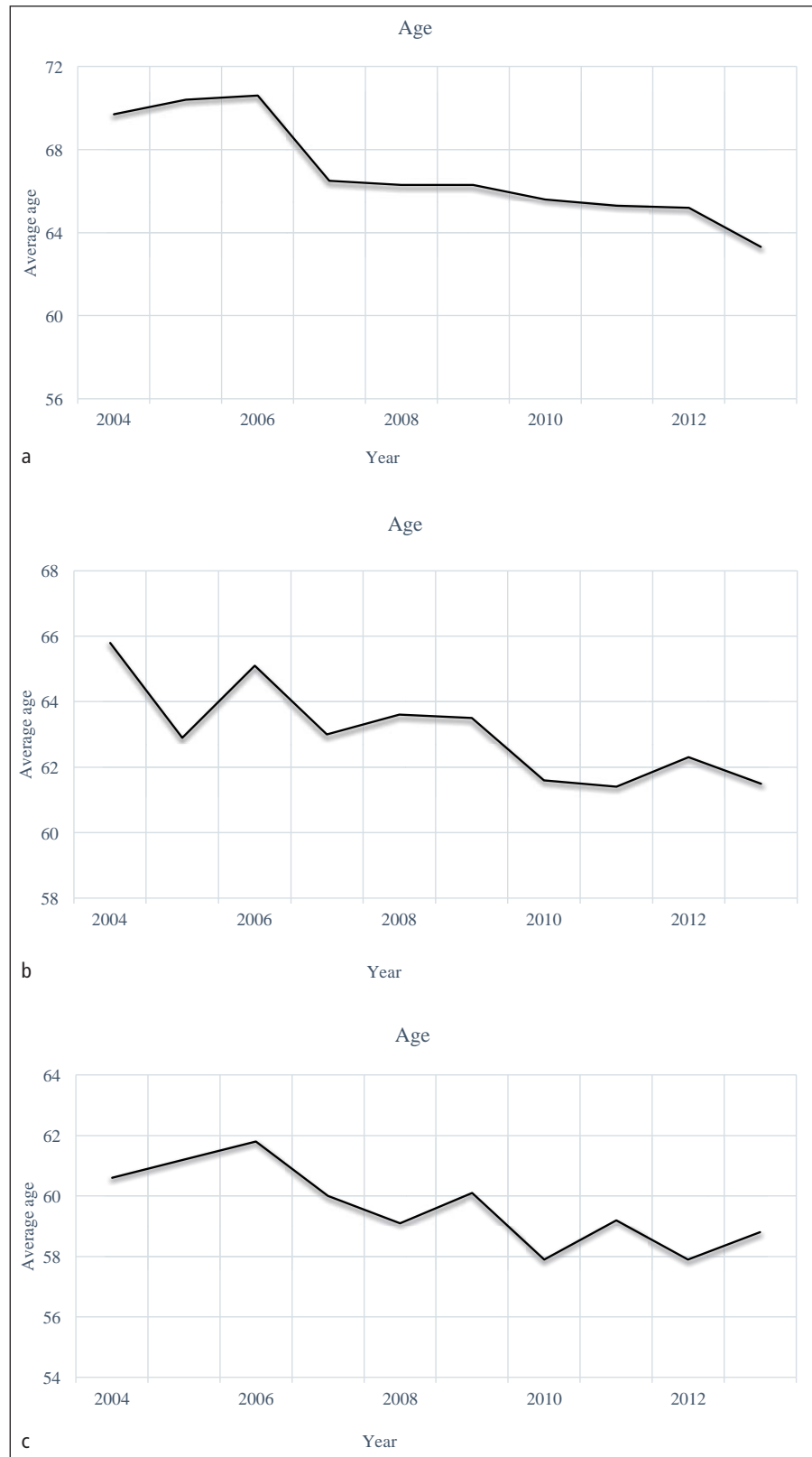
Figure 8   Average age of patients in the corresponding risk level. (a) Average age of patients in low risk level, (b) Average age of patients in medium risk level, and (c) Average age of patients in high risk level.

Table 5   Changing treatment interventions detected from the experimental dataset with respect to different risk levels.

Low risk	
Growing	Antiplatelet drugs, Ultrasound, Lipid regulators, 0.9% sodium chloride injection, Proton pump inhibitors, Coagulation test, Coronary angiography, PTCA, Consultation, Drug exchange, Blood pressure measurement, Glucose regulators, Blood collection, Occult blood test, Radionuclide examination, Peripheral vasodilators, CT, Thyroid function, Puncture, Monitoring of ECG and blood oxygen, Angiotensin receptor blocker, Anesthesia, Diabetes inspection
Declining	Antianginal drugs, Blood routine, Biochemical routine, β adrenergic receptor blockers, First-grade nursing, Thrombosis Inspection, Diuretics, Specific diet, Urine routine, Ganmao Qingre, Calcium channel blocker, Maren Runchang, Xinqingning, Blood typing, Angiotensin converting enzyme inhibitors, Oxygen Inhalation, Mucosal protective agents, Transferred, Sedative and antianxiety drugs, Multifunctional monitors, Troponin T, Stool routine
Steady	Drugs for cardiovascular diseases, Routine nursing, Skin preparation, Anti-diabetes drugs, Injection methods, Anticoagulant drugs, Discharge, Second-grade nursing, Serum test, Electrolyte regulators
Up-down	Local anesthetics, Analysis of blood ion, Parenteral nutrition drugs
Down-up	Glucose injection, Basic care, Vasodilators
Jumping	ECG, Tumor marker test, X-Ray
Medium risk	
Growing	0.9% sodium chloride injection, Anticoagulant drug, Troponin T, Drug exchange, Urine routine, Consultation, Injection methods, Anesthesia, Antiplatelet drug, Blood collection, Puncture, Blood pressure measurement, Discharge, Intake and output, Basic care, Ganmao Qingre, Coronary angiography, Proton pump inhibitors, Monitoring of ECG and blood oxygen, Renal function regulators, Adjuvant drugs for liver diseases, PTCA, Angiotensin receptor blocker, TEG, Gastrointestinal prokinetic agent,
Declining	Antianginal drugs, Thrombosis Inspection, Plasma and plasma substitutes, Electrolyte regulators, Diuretics, Skin test, Specific diet, Sedative and antianxietic drugs, β adrenergic receptor blockers, Digitalis glycosides, Glucose injection, Anti-diabetes drugs, Angiotensin converting enzyme inhibitors, Second-grade nursing, Routine nursing, Transferred, PH balance regulators, Multifunctional monitors, Tuberculosis examination, Vasoactive drugs, β-lactamase inhibitors, Coagulation test, Narcotic analgesics, Vitamins
Steady	Biochemical routine, Glucose regulators, Oxygen Inhalation, Blood routine, Skin preparation, Cephalosporins, Lipid regulators, Enema, Stool routine

Up-down	Analysis of blood ion, Peripheral vasodilators, Calcium regulators, Tumor marker test, CT, Local anesthetics, Parenteral nutrition drugs, Radionuclide examination, Antiarrhythmic drugs, α-adrenergic receptor blockers, Expectorants, Xinqingning, Occult blood test, Glycyrrhizae Mixture
Down-up	First-grade nursing, Drugs for cardiovascular diseases, ECG, Ultrasound, Seriously ill
Jumping	Calcium channel blocker, Drugs for bronchitis, Adrenocorticotropic hormone, Drugs for prostate diseases, Sterilized distilled water, X-Ray, Maren Runchang
High risk	
Growing	Biochemical routine, Blood collection, Radionuclide examination, Drug exchange, Analysis of blood ion, CT, Blood pressure measurement, Xinhuang, Gastrointestinal prokinetic agent, Adjuvant drugs for liver diseases, Digitalis glycosides, Diuretics, Proton pump inhibitors, Coagulation test, Antiarrhythmic drugs, Transferred, Puncture, Expectorants, Troponin T, Intake and output, Local anesthetics, Body temperature lowering, X-Ray, Enema
Declining	Antianginal drugs, Calcium channel blocker, Glucose injection, Vasodilators, Sedative and antianxietic drugs, β adrenergic receptor blockers, Specific diet, ECG, Drugs for cardiovascular diseases, Thrombosis Inspection, Angiotensin converting enzyme inhibitors, Immunity test, Skin preparation, Second-grade nursing, Skin test, First-grade nursing, Vasoactive drugs, Narcotic analgesics, Multifunctional monitors, PH balance regulators, Adrenocorticotropic hormone, Nitroglycerin membrane, Electrophysiological examination, Maren Runchang, Antiepileptic and anticonvulsant drugs, Antiplatelet drug, Tuberculosis examination, Furacin, Urine routine, Gastrointestinal antispasmodic drugs, Plasma and plasma substitutes, Diabetes inspection
Steady	0.9% sodium chloride injection, Blood routine, Routine nursing, Glucose regulators, Sterilized distilled water, Oxygen Inhalation, Basic care, Injection methods, Stool routine, Occult blood test, Discharge
Up-down	Electrolyte regulators, Cerebral circulation regulators, Monitoring of ECG and blood oxygen, Ultrasound, β-lactamase inhibitors, Penicillins, Intravenous anesthetics
Down-up	Anticoagulant drug, Glycyrrhizae Mixture
Jumping	Cephalosporins, Consultation, Lipid regulators, Calcium regulators, Oral care, Antipyretic and analgesic drugs, Vitamins, Anti-diabetes drugs, Blood transfusion, Parenteral nutrition drugs

Table 6   Changing trends of execution time for treatment interventions.

Risk levels	Steady-implemented	Early-implemented	Delay-implemented
low	48	13	3
medium	2	78	4
high	0	84	2

Table 7   Changing trends of execution time for treatment interventions.

Low risk	
Early-implemented	Discharge, Skin preparation, Drug exchange, Injection methods, Coronary angiography, Anesthesia, Transferred, Electrolyte regulators, Ganmao Qingre, PTCA, Local anesthetics, Puncture, Consultation
Delay-implemented	Multifunctional monitors, CT, Monitoring of ECG and blood oxygen
Steady-implemented	Antianginal drugs, Specific diet, Drugs for cardiovascular diseases, Routine nursing, Blood routine, Ultrasound, Biochemical routine,β adrenergic receptor blockers, Thrombosis Inspection, Urine routine, Antiplatelet drugs, Stool routine, ECG, Anticoagulant drugs, Calcium channel blocker, Second-grade nursing, First-grade nursing, Serum test, Lipid regulators, Oxygen Inhalation, 0.9% sodium chloride injection, Blood collection, Basic care, Angiotensin converting enzyme inhibitors, Blood typing, Sedative and antianxiety drugs, Maren Runchang, Glucose regulators, Anti-diabetes drugs, Vasodilators, Tumor marker test, Blood pressure measurement, X-Ray, Mucosal protective agents, Coagulation test, Occult blood test, Analysis of blood ion, Peripheral vasodilators, Radionuclide examination, Thyroid function, Xinqingning, Diuretics, Troponin T, Angiotensin receptor blocker, Proton pump inhibitors, Parenteral nutrition drugs, Diabetes inspection, Glucose injection

| Medium risk	
Early-implemented	Antianginal drugs, Blood routine, 0.9% sodium chloride injection, Diuretics, Biochemical routine, Basic care, Specific diet, Anticoagulant drug, Sedative and antianxietic drugs, Glucose injection, Consultation, Vitamins, Cephalosporins, Expectorants, ECG, Tuberculosis examination, Injection methods, Anesthesia, Urine routine, Glucose regulators, Xinqingning, Drug exchange, Skin preparation, β adrenergic receptor blockers, PTCA, Calcium channel blocker, Antiplatelet drug, Transferred, Blood collection, Electrolyte regulators, Troponin T, First-grade nursing, Blood pressure measurement, Antiarrhythmic drugs, Oxygen Inhalation, Lipid regulators, Second-grade nursing, Multifunctional monitors, Routine nursing, Discharge, Plasma and plasma substitutes, Angiotensin converting enzyme inhibitors, Coronary angiography, Thrombosis Inspection, Maren Runchang, Analysis of blood ion, Narcotic analgesics, Sterilized distilled water, Proton pump inhibitors, Seriously ill, Ganmao Qingre, Anti-diabetes drugs, Vasoactive drugs, Radionuclide examination, Local anesthetics, Enema, PH balance regulators, Digitalis glycosides, β-lactamase

Early-implemented	inhibitors, Drugs for bronchitis, Coagulation test, Parenteral nutrition drugs, Adrenocorticotropic hormone, Glycyrrhizae Mixture, Drugs for prostate diseases, Puncture, Peripheral vasodilators, Angiotensin receptor blocker, Drugs for cardiovascular diseases, Skin test, Calcium regulators, Tumor marker test, Gastrointestinal prokinetic agent, Adjuvant drugs for liver diseases, Occult blood test, CT, TEG, Stool routine
Delay-implemented	α-adrenergic receptor blockers, Monitoring of ECG and blood oxygen, Ultrasound, X-Ray
Steady-implemented	Renal function regulators, Intake and output

High risk	
Early-implemented	0.9% sodium chloride injection, Antianginal drugs, Expectorants, Biochemical routine, Specific diet, Discharge, Blood routine, First-grade nursing, Tuberculosis examination, β adrenergic receptor blockers, Stool routine, Drug exchange, Electrolyte regulators, Sedative and antianxietic drugs, Second-grade nursing, Vasodilators, Drugs for cardiovascular diseases, Glucose regulators, Cephalosporins, Skin preparation, Blood collection, Thrombosis Inspection, Calcium channel blocker, Angiotensin converting enzyme inhibitors, Routine nursing, Maren Runchang, Anticoagulant drug, Transferred, Oxygen Inhalation, Urine routine, Sterilized distilled water, Plasma and plasma substitutes, Injection methods, Blood pressure measurement, Parenteral nutrition drugs, Multifunctional monitors, Vitamins, Analysis of blood ion, Vasoactive drugs, Coagulation test, Penicillins, Gastrointestinal antispasmodic drugs, Radionuclide examination, Adrenocorticotropic hormone, Diuretics, Enema, Troponin T, Narcotic analgesics, Lipid regulators, Local anesthetics, Blood transfusion, Consultation, Gastrointestinal prokinetic agent, Xinhuang, β-lactamase inhibitors, Glycyrrhizae Mixture, Diabetes inspection, Basic care, Ultrasound, Digitalis glycosides, Nitroglycerin membrane, Furacin,Calcium regulators, Puncture, PH balance regulators, Intake and output, Body temperature lowering, Intravenous anesthetics, Oral care, ECG, Antiepileptic and anticonvulsant drugs, Cerebral circulation regulators, Monitoring of ECG and blood oxygen, Antiplatelet drug, Anti-diabetes drugs, Proton pump inhibitors, Antiarrhythmic drugs, Occult blood test, X-Ray, Antipyretic and analgesic drugs, CT, Adjuvant drugs for liver diseases, Skin test
Delay-implemented	Immunity test, Electrophysiological examination

EMR system. Therefore, it is difficult to determine whether these changes of treatment interventions are essential to the treatment process of unstable angina. In this regard, these treatment interventions need to be further analyzed by clinicians to investigate the clinical significance of these changes for the treatment process of unstable angina.

3.3.3 Results of Changing Trends of Occurring Time Stamps of Treatment Interventions

Regarding these selected treatment interventions, changing trends of their occurring time stamps can be derived from the experimental dataset as well, as shown in

►Table 6. It is obvious that most of selected treatment interventions for low-risk patients are stable while most of those with medium-risk level and high-risk level are early-occurred. ►Table 7 demonstrates detailed classification results of detected changing trends.
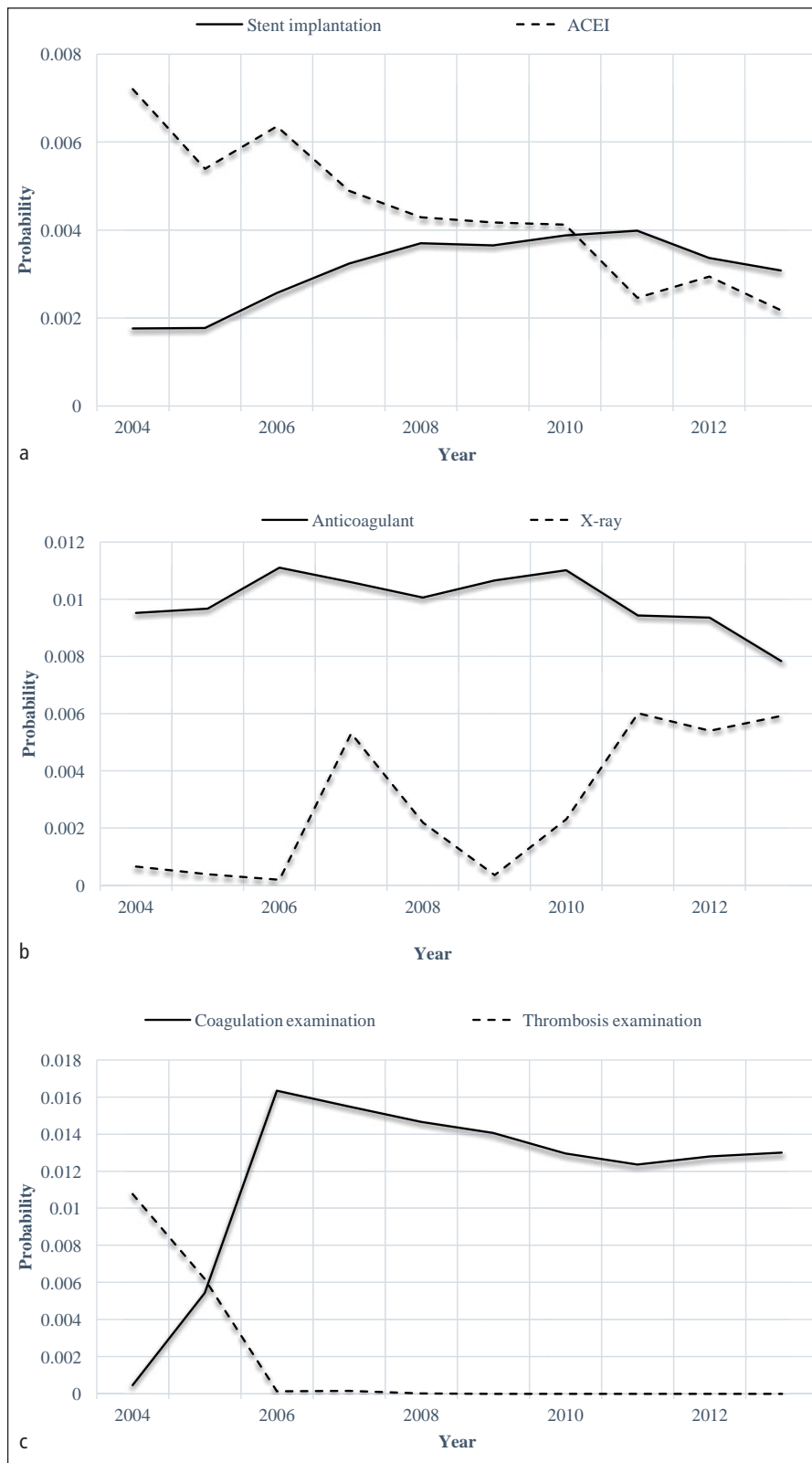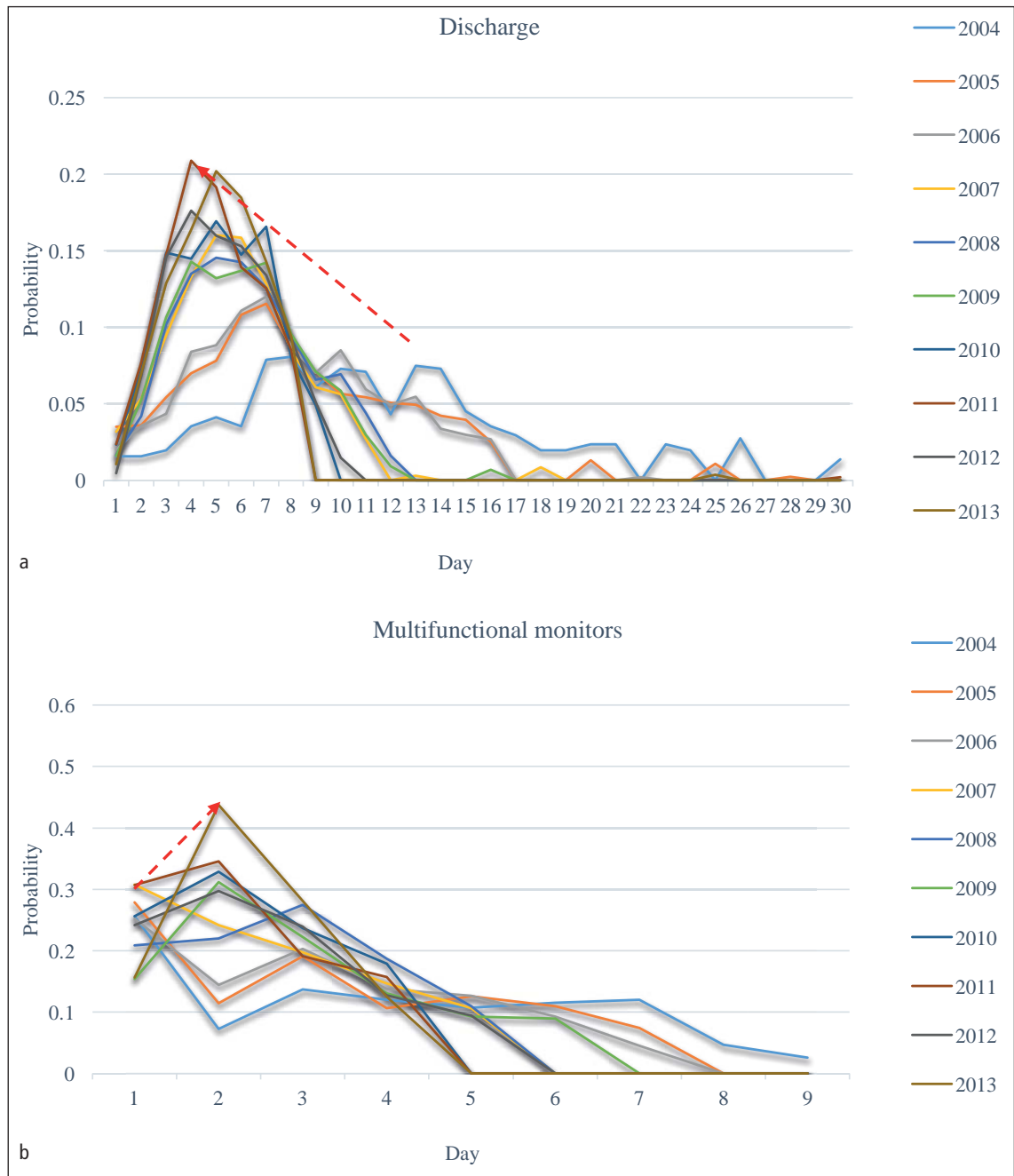
Figure 9   Changing trends of example treatment behaviors. (a) "stent implantation" and "ACEI" for low-risk patients, (b) "Coagulation examination" and "Thrombosis examination" for low-risk patients, and (c) "Anticoagulant" and "X-ray" for low-risk patients.

Table 8   Unperceived changing trends of medical behaviors beyond clinicians' expectations.

Patient feature-value pairs	
Low risk	Hemoglobin:L(down-up), Platelet:L(down-up), Mean corpuscular hemoglobin:L(jumping)
Medium risk	Bleeding:True(declining), Quantitative determination of Isoenzyme of creatine kinase:N(up-down)
High risk	High-sensitive C-reactive protein:H(up-down), Hyperlipidemia:True(declining), Renal insufficiency:True(declining)
Treatment interventions	
Low risk	Analysis of blood ion(up-down), Thyroid function(growing), Tumor marker test(jumping)
Medium risk	Drugs for prostate diseases(jumping), Gastrointestinal prokinetic agent(growing), Ganmao Qingre(growing)
High risk	Gastrointestinal prokinetic agent(growing), Xinhuang(growing), Antiplatelet drugs(declining), Skin preparation(declining), Lipid regulators(jumping)

Since that treatment interventions may not be performed within the desired timeframe scheduled in a treatment process protocol in the actual clinical practice, we can better understand the evolution of these treatment interventions with respect to each risk level, by analyzing changing trends of their execution time stamps. Specifically, occurring time stamps of essential treatment interventions for low-risk patients are mostly steady-implemented, implicating that their execution time stamps are arranged appropriately for patients with low-risk level. However, there are 13 early-implemented treatment interventions (e.g., "*Discharge*" (as shown in ►Figure 10(a)), as a important indicator of clinical outcome, is a significant early-implemented treatment intervention, implicating that patients in low-risk tend to be discharged in a shorter length of stay as time elapsed.) detected from the experimental dataset, which indicates that their execution time stamps may not be scheduled in proper time instants and should be rearranged to earlier days in the pathway. In addition,
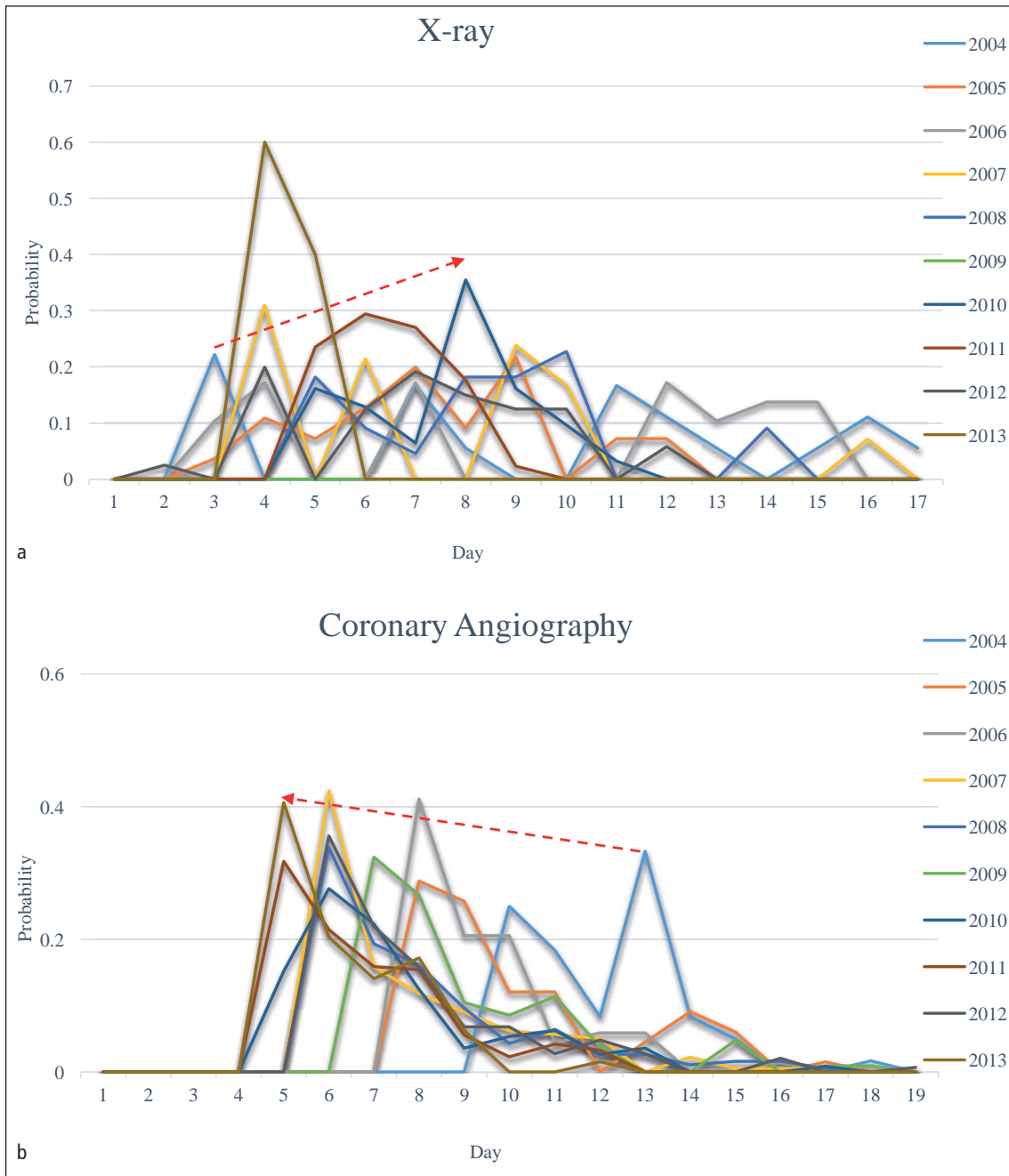
Figure 10
Examples of the changing trends of execution time for treatment interventions. (a) "Discharge" for low-risk patients, and (b) "Multifunctional monitors" for low-risk patients.

there are 3 delay-implemented treatment interventions, i.e., "*Multifunctional monitors*" (as shown in ▶Figure 10(b)), "*CT*" and "*Monitoring of ECG and blood oxygen*", in the treatment pattern for low-risk patients, suggesting that occurring time stamps of these treatment interventions could be rescheduled to later days in the unstable angina treatment process.

In contrast with low-risk patients, most of treatment interventions in treatment patterns for medium-risk and high-risk patients are early-implemented, demonstrating that most of them are scheduled improperly and can be rearranged to earlier time stamps. These findings contribute to redesigning/optimizing treatment process protocols to deliver efficient health services and reduce LOS of patients. Meanwhile, there are only 4 and 2 delay-implemented treatment interventions, respectively, for patients with medium-risk and high-risk

levels. Similar with delay-implemented treatment interventions for patients with low-risk level, their execution time stamps may need to be rescheduled to later days. These detected evolving trends provide convinced evidences for the rearrangement of efficient treatment interventions for unstable angina patients.

3.4 Assessment of Detected Changing Trends of Medical Behaviors

The experimental results, reflecting the realities in EMRs, have been evaluated by clinical collaborators to assess and reveal their clinical meanings. They have pointed out that most of detected changing trends are complied with their knowledge and experiences in the changing frequencies of medical behaviors for the unstable angina treatment process. More specifically, our clinical collaborators indicate that growing or steady feature-value pairs/treatment interventions are obligatory/essential to the unstable angina treatment process, while those medical behaviors following a declining changing pattern are not essential any more over time, and thus can be removed from the treatment process protocol. For example, the patient feature-value pair "*after PCI (Percutaneous coronary intervention): True*" is progressively growing into an important symptom for medium-risk patients, indicating that most of high-risk patients have already been performed the PCI surgery in their previous hospitalizations. The treatment intervention "*ACEI (Angiotensin converting enzyme inhibitors)*", which follows significant declining trends in all risk clusters, is gradually becoming nonob-
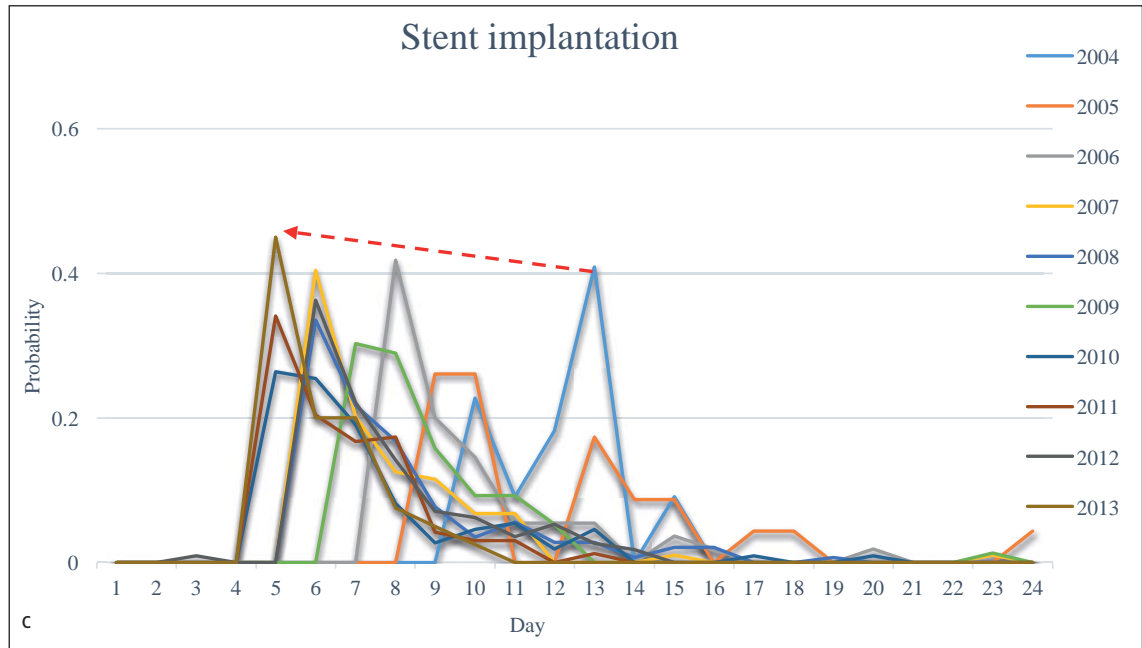
Figure 11
continued

ligatory for unstable angina patients in all risk clusters.

While most of detected changing trends are consistent with clinicians' perception of treatment changing frequencies, there are still a few that are beyond their expectations, as shown in ▶ Table 8. Our clinical collaborators have carefully assessed these evolving trends and confirmed that these unperceived evolving trends indeed occur in clinical practice. In this sense, the proposed approach can be an objective way of analyzing changing trends of essential medical behaviors over time as it is not biased by perceptions or normative behaviors, and thus can provide valuable references for the treatment process protocol redesign and optimization.

In addition, clinical collaborators have stated that changing trends of execution time stamps of some treatment interventions are meaningful and convinced evidences for rearranging them in the treatment process protocols. In particular, patients are gradually inclined to be discharged from hospital with shorter LOSs (as shown in ▶ Figure 10(a)), as many essential treatment interventions in the unstable angina treatment process are standing a better chance to be performed in early days after admission. It should be noted that treatment interventions following early-occurred or delay-occured change patterns provide important references for the redesign/optimizion of treatment process protocols. I.e., they are considered to be rescheduled to the proper time range during the treatment process execution, e.g., coronary angiogram, stent implantation, x-ray, etc, as illustrated in ▶ Figure 11.

4. Conclusion

In this study, we propose a novel approach to utilize EMRs for discovering changing trends of essential/critical medical behaviors over time. The feasibility of our approach has been evaluated on a real clinical dataset consisting of 12,512 unstable angina patient cases in cooperation with the Cardiology department of Chinese PLA General hospital. Experimental results implicate that the proposed approach is feasible in discovering essential changing trends in medical behaviors over time. Even though there are huge variations among different diseases, the proposed method is not limited to unstable angina, because domain knowledge for a particular disease is not incorporated into the proposed method. In fact, the proposed approach is a generalized methodology and can be applied to any types of diseases'

treatment processes to discover non-trivial evolving trends in medical behaviors from EMRs. Note that the possibilities to utilize a large volume of EMRs are currently weak. Our study suggests that there is significant potential to improve this utilization. This could help clinical experts further explore the detected changing trends in medical behaviors based on their domain knowledge. Since detected changing trends reflect the realities of patient-care processes other than conceptualized treatment process protocols, they can be potentially served as convincing evidences for the improvement of delivered health services.

Compared with most of existing methods that rely on the experiences and knowledge of clinical experts, the proposed approach can automatically discover changing trends of medical behaviors as well as their execution time stamps from EMRs rather than elaborate patient cases piece by piece. In this sense, our approach is time-effective and effort-saving. In addition, our approach can cluster the treatment processes of patients into different treatment patterns reflecting the severity of patient conditions. Therefore, our approach can category patients into different risk levels based on their clinical conditions and then detect the changing trends of medical behaviors regarding different risk

tiers. To our best knowledge, existing approaches lack this ability.

It has to be noted that there are some limitations of our work. First of all, we only selected these top-ranked (i.e., essential and frequent) behaviors in treatment processes to detect changing trends, however, there are a number of infrequent behaviors that are missing in discovered treatment patterns. Note that many of these infrequent behaviors are correlated with the treatment of comorbidities of patients, and thus need to be analyzed with respect to their changes over time [25]. In addition, we assume that occurring probabilities of medical behaviors are independent to each other. Somehow, this assumption neglects the correlations between medical behaviors. In clinical practice, occurrences of certain medical behaviors may have influences to each other. Hence, correlations of medical behaviors should be investigated to have a fine-grained analysis on changing trends of medical behaviors. Furthermore, in this study, we focus on detecting changing trends of medical behaviors over time, while derivations of behaviors/ treatments from the treatment process protocols is not considered. Moreover, data segments are predefined based on legal years for the treatment process protocol updates, and it could be valuable to define segments automatically independent from legal years. We plan to address these issues in our future work.

Acknowledgment

References

1. Rotter T, Kinsman L, James E, et al. Clinical pathway: effects on professional practice, patient outcomes, length of stay and hospital costs. Cochrane Database Syst Rev. 2010; (3): CD006632.
2. Adeyemi S, Demir E, Chaussalet T. Towards an evidence-based decision making healthcare system management: Modelling patient pathways to improve clinical outcomes. Decision Support System. 2013; 55(1): 117–125.
3. Zhang Y, Padman R, Wasserman L, Patel N, Teredesai P, Xie Q. On clinical pahtway discovery from Electornic Health Record data. IEEE Intelligent Systems. 2015; 30(1): 70–75.
4. Peleg M, Mulyar N, van Der Aalst WM. Pattern-based analysis of computer-interpretable guidelines: Don't forget the context. Artificial Intelligence in Medicine. 2012; 54(1): 73–74.
5. Huang Z, Lu X, Duan H. Latent treatment pattern discovery for clinical processes. Journal of Medical Systems. 2013; 37(2): 1–10.
6. Klundert J, Gorissen P, Zeemering S. Measuring clinical pathway adherence. Journal of Biomedical Informatics. 2010; 43(6): 861–872.
7. Blei D, A. Ng A, Jordan M. Latent Dirichlet allocation. Journal of Machine Learning Research. 2003; 3: 993–1022.
8. Catherwood E, O'Rouke DJ. Critical pathway management of unstable angina. Progress in Cardiovascular Diseases. 1994; 37(3): 121–148.
9. Dong G, Li J. Efficient mining of emerging patterns: Discovering trends and differences. In: Proceedings of the fifth International Conference on Knowledge Discovery and Data Mining; 1999. p. 43–52.
10. Huang Z, Dong W, Duan H. A probabilistic topic model for clinical risk stratification from electronic health records. Journal of Biomedical Informatics. 2015; 58: 28–36.
11. Wang C, Blei D, Heckerman D. Continuous time dynamic topic models. In: Proceedings of the Twenty-Fourth Conference on Uncertainty in Artificial Intelligence; 2008. p. 579–586.
12. Huang Z, Dong W, Peter B, Ji L, Duan H. On mining latent medical behavior patterns from electronic medical records. Data Mining and Knowledge Discovery. 2015; 29(4): 914–949.
13. Hall D, Jurafsky D, Manning C. Studying the history of ideas using topic models. In: Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing; 2008. p. 363–371.
14. Huang Z, Gan C, Lu X, Duan H. Mining the changes of medical behaviors in clinical pathway. In: Proceedings of the 2013 Conference on MedInfo; 2013. p. 117–121.
15. Basse L, Jakonbsen D, Billesbolle P, Werner M, Kehlet H. A clinical pathway to accelerate recovery after colonic resection. Annals of Surgery. 2000; 232(1): 51.
16. Stephen A, Berger D. Shortened length of stay and hospital cost reduction with implementation of accelerated clinical care pathway after elective colon resection. Surgery. 2003; 133(3): 277–282.
17. Porter G, Pisters P, Mansyur C, Bisenz A, et al. Cost and utilization impact of a clinical pathway for patients undergoing pancereaticoduodenectomy. Annals of Surgical Oncology. 2000; 7(7): 484–489ß.
18. Ainsworth J, Buchan I. Combining health data uses to ignite health system learning. Methods of Information in Medicine. 2015; 54: 479–487, 2015.
19. Porteous I, Newman D, Ihler A, Asuncion A, Smyth P. Fast collapsed gibbs sampling for latent dirichlet allocation. In: Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2008. p. 569–577.
20. Huang Z, Dong W, Ji L, et al. Discovery of clinical pathway patterns from event logs using probabilistic topic models. Journal of Biomedical Informatics. 2014; 47: 39–57.
21. Miller C, Reardon M, Safi H. Risk stratification: a practical guide for clinicians. Cambridge: Cambridge University Press; 2001.
22. Engelhardt B, König J, Blettner M, et al. Combining Cross-sectional Data on Prevalence with Risk Estimates from a Prediction Model. Methods of Information in Medicine. 2014; 53(5): 371–379.
23. de Lusignan S. An education intervention, involving feedback of collected computer data, to improve cardiovascular cisease canagement in UK primary care. Methods of Information in Medicine. 2007; 46(1): 57–62.
24. Paoin W. Lessons learned from data mining of WHO mortality database. Methods of Information in Medicine. 2011; 50(4): 380–385.
25. Huang Z, Dong W, Ji L, Duan H. Predictive monitoring of clinical pathways. Expert Systems with Applications. 2016; 56: 227–241.