

А.А. Литвинов, А.Д. Мизюмская, В.В. Истомина
**ОСОБЕННОСТИ ПРОЦЕДУРЫ ОЦЕНКИ И ОБНАРУЖЕНИЯ
МЕДИЦИНСКИХ СТАНДАРТОВ
ДЛЯ ПЛАНИРОВАНИЯ РЕСУРСОВ**

Аннотация. В данной статье рассматривается алгоритм работы самообучающейся экспертной системы, для обнаружения новых и оценки существующих расширенных медико-экономических стандартов. Такой подход позволяет выявить объективную причину отклонения, обнаружить устойчивые статистические группы, усовершенствовать стандарты, которые создают основу системы планирования ресурсов медицинского учреждения.

Ключевые слова. Стандарт, экспертная система, обучение без учителя, алгоритм.

Введение. На текущий момент создание и внедрение медико-экономических стандартов играет важнейшую роль в учреждениях как государственного, так и частного типа. Основными целями при этом являются: клиническое и экономическое обоснование оказания медицинской помощи [1]; построение базы для достижения прогнозируемой, устойчивой работы клиники [2]; эффективное планирование активностей и ресурсов; прогнозирование и управление рисками. Таким образом, разработка эффективных подходов к построению стандартов, методы их выделения, контроля и оценки являются важной задачей при построении информационных систем, поддерживающих внедрение стандартизации. Стандарт медицинской помощи можно определить, как множество протоколов, которые отображают специфику услуг для клинического состояния [1, 3], учитывая при этом множество разных параметров, связанных с состоянием и образом пациента (напр.: возраст, пол, вес, место жительства, реакция на препараты и пр.) и влияют на необходимость выполнения тех или других лечебных / диагностических активностей.

Более формально протокол можно определить, как правило (if-then), описывающее реакцию системы (в данном случае клиники)

на определенное клиническое состояние пациента, с указанием ожидаемого результата. Данная реакция выражается в виде потока работ (workflow) состоящего из набора активностей, каждая из которых четко определена в терминах пространственных (ресурсы, материалы, исполнители) и временных (длительность пребывания, операции) ограничений.

$$s_i: d_j \rightarrow w_k, \quad s_i \in S, d_j \in D, w_k \in \mathbb{W}, \quad (1)$$

s_i – протокол стандарта, d_j – клиническое состояние пациента, w_k – поток работ.

$$w_k = a_1^p \ll a_1^s \ll \dots \ll a_n^r = \langle a_1^p, a_1^s, \dots, a_n^r \rangle, \quad (2)$$

$$a_1^p, a_1^s, \dots, a_n^r \in A; p, s, r \in A^t, \quad a_i = \langle t, \tau, e, c, r \rangle_i,$$

A^t – типы активностей, e – подмножество ролей исполнителей (количество и тип исполнителя), r – множество типизированных ресурсов, t – время начала выполнения активности, τ – продолжительность, c – затраты.

Важной составляющей протокола является прогнозирование ожидаемого результата, который выражается в таких экономических показателях как длительность пребывания пациента в стационаре bd , затраты на лечение c , время занятия операционной (в случае оперативного случая) ol .

$$\forall_{w_k \in \mathbb{W}}. (ol, bd, c) \quad (3)$$

На основании протоколов возможно осуществлять эффективное оперативное и долгосрочное (принимая во внимание структуру заявок на обслуживание) планирование работы стационара. Реальные ситуации могут требовать внесение поправок в набор активностей базового потока работ, что может оказывать влияние на результат, нарушая при этом работу механизма планирования. Для решения данной задачи следует предусмотреть процедуру переоценки существующих и обнаружения новых стандартов-протоколов. Вопрос нетривиального получения множества стандартов тесно связан с методами datamining, которые, как правило, базируются на создании обучающей выборки, выделении весомых характеристик, поиска функций-классификаторов (путем обучения системы), применении спец. продуктов. Вторым вопросом является обучение – кол-во классов заранее не известно, а при применении методов кластеризации без учителя в результате работы таких процедур возможно появление большого числа непонятных, плохо структурированных классов. Методы пере-

оценки таких классов также связаны с дополнительными затратами, не всегда поддаются автоматизации, требуют специалистов.

Постановка задачи. Целью данной работы является разработка алгоритма автоматизированной проверки существующих и поиска новых, расширенных стандартов, учитывающих весь комплекс параметров, влияющих на выбор лечебной тактики, применение ресурсов. При этом следует отметить, что мы имеем дело не с предопределенным набором классов-стандартов, по которым следует распределить данные, нашей целью является нахождения процедуры поиска новых стандартов на базе уже существующих, при этом следует отметить адаптивность и переобучение системы в зависимости от пополнения базы данных. Важными особенностями являются: полная автоматизация такого поиска с минимальными затратами со стороны ответственного за стандарты лица; минимизация ресурсов и обеспечение допустимого времени работы процедуры поиска. В работе [4] была рассмотрена базовая версия алгоритма, основная стратегия поиска. В данной работе рассмотрим уточненную версию.

Основная часть. Предлагаемый подход базируется на предварительном описании некоторой группы стандартов по принципу «клинический диагноз - операция». После внедрения проводится мониторинг и оценка стандарта (с возможностью его корректировки, выделению новых протоколов). Для этого анализируются отклонения значений его оценочных параметров (время операции, кол-во койко-дней, затраты) от некоторой пороговой величины в ту или другую сторону. В случае если поток укладывается в заданные границы, связанные со временем, материальными и трудовыми затратами, стандарт становится устойчивым. В противном случае выясняются причины расхождения, и выявляется возможный дефект стандарта или определяется группа пациентов по совокупности дополнительных параметров (пол, вес, рост, возраст, реакция на препараты), которая будет всегда давать такое отклонение, т.е. на базе которой необходимо сформировать новый протокол.

Выявление причины расхождений базируется на сравнении отклонения с порогом-триггером, который включает функцию поиска причины отклонения перебирая группы по заданному порядку. Данный порог является гибким для каждого протокола, определяя последний порог срабатывания. Цель процедуры – обнаружить группу,

порог отклонения для которой превышает порог, отвечающий за выделение стандарта.

Алгоритм. Начальной точкой работы системы является формирование экспертом множества корневых элементов стандарта: пар «клинический диагноз – операция». На основании опыта в полуавтоматическом режиме (на базе существующей статистики) эксперт оценивает ресурс, который необходим для лечения пациентов с данным клиническим диагнозом. Основными характеристиками, определяющими ресурс являются: время операции, время пребывания больного в стационаре, обязательные активности и ресурсы, на базе которых определяются затраты на лечение. В дальнейшем происходит автоматическое уточнение системы стандартов на базе нового материала: комбинации «пациент-параметры» могли быть не выявлены по причине отсутствия необходимой информации в базе данных, отсутствия учета дополнительных параметров.

После определения факта отклонений целесообразно обнаружить причину отклонений, путем выделения устойчивой статистической группы пациентов. Данная группа создает прецедент появления отдельного протокола.

Для формального описания алгоритма поиска причины отклонения введем следующие обозначения:

$d_i \subseteq D$ – подмножество пациентов с клиническим диагнозом;

$o_j \subseteq O$ – подмножество пациентов с операцией;

$p_i \in P_i \subseteq P$ – параметр, множество значений которого является подмножеством общего множества параметров, определяющих образ и состояние пациента. Например, возрастные группы, пол, осложнение операции и пр.;

$\delta_i \in \Delta$ – отклонение по некоторому параметру, зафиксированному стандартом. Например, превышение послеоперационного пребывания в клинике на 2 дня или понижение времени операции на 10 минут.

На первом шаге определяется пару «клинический диагноз – операция», которая составляет стандарт в момент времени t – до начала процедуры поиска отклонений.

$$s_k^t \equiv d_i \cap o_j, s_k \in S^t, \quad k \leq |S^t| \quad (4)$$

θ – фиксированное пороговое значение (для всех стандартов) определяющее выделение нового стандарта: если доля пациентов с отклонением по тому или иному параметру превышает данный порог, происходит пересмотр данного стандарта или выделение нового. В нашем случае используется 0.75;

π_k^t – порог, определяющий вероятность наличия отклонений (скрытого стандарта) для k -го стандарта: порог для начала поиска причины отклонения и возможного выделения нового стандарта. Первоначально порог для начала поиска причины отклонения определяется соотношением $\pi_k^t = \frac{\theta}{2}$.

Для выбранной пары «диагноз-операция» на втором шаге анализируются параметры отклонений, основными из которых являются: отклонение по времени операции, отклонение по периоду лечения. Если доля пациентов с отклонениями превышает порог θ , необходимо провести корректировку стандарта.

$$\frac{|s_k^t \cap \delta_i|}{|s_k^t|} > \theta \rightarrow s_k^{t+1} = Edit(s_k^t, \delta_i) \quad (5)$$

Если отклонение присутствует (выше порога π_k^t) – необходимо провести исследование на предмет влияния дополнительных параметров на отклонение.

$$\theta > \frac{|s_k^t \cap \delta_i|}{|s_k^t|} > \pi_k^t \rightarrow Search(s_k^t, \delta_i, P) \quad (6)$$

Выразим начальное условие поиска – множество больных с отклонением от стандарта.

$$c_0^k \equiv s_k^t \cap \delta_i, \quad b_0^k \equiv s_k^t \quad (7)$$

В случае превышения порога происходит выделение нового стандарта

$$\forall p_i \in P, p_i \in b_0^k \frac{|c_0^k \cap p_i|}{|b_0^k \cap p_i|} \geq \theta \rightarrow NewStandard(c_0^k \cap p_i) \quad (8)$$

Процедура выделения нового стандарта может быть записана следующим образом:

$$s_k^{t+1} = s_k^{t+1}, s_k^{t+1}, s_{k+1}^{t+1} \in S^{t+1}, n = |S^t|, \quad k \leq n \quad (9)$$

Если ни один из параметров не дал оснований для выделения стандарта – для каждого из параметров, которые могут влиять на отклонение, вычисляется расстояние до порога выделения стандарта δ_i .

$$\forall p_i \in P, p_i \in S_i^k, \frac{|c_0 \cap p_i|}{|S_i^k \cap p_i|} < \theta \Rightarrow b_i = \frac{|c_0 \cap p_i|}{|S_i^k \cap p_i|} - \theta \quad (10)$$

Результатом такой процедуры будет упорядоченный вектор расстояний, в котором минимальное расстояние является первым элементом. Исходя из этого, сортируется и набор параметров.

$$b_1^k \leq b_2^k \dots \leq b_n^k \Rightarrow p_1^k \preccurlyeq p_2^k \preccurlyeq \dots \preccurlyeq p_n^k \quad (11)$$

Далее производится формирование условия для выполнения следующей итерации по следующим правилам.

Если доля больных по условию c_i^k больше порогового значения θ , следует выделить новый стандарт по процедуре описанной формулой (9).

$$\frac{|c_i^k|}{|b_i^k|} \geq \theta \Rightarrow \text{halt}; \text{NewStandard}(c_i^k); \pi_k = b_i. \quad (12)$$

Если доля больных с отклонениями от стандарта превышает долю, полученную на предыдущей итерации, но меньше порога выделения стандарта – параметр, рассматриваемый в момент текущей i -ой итерации, учитывается в последующей выборке.

$$\frac{|c_i^k|}{|b_i^k|} - \theta < b_{i-1} \Rightarrow c_{i+1}^k = c_i^k \cap p^{i+1}; b_{i+1}^k = b_i^k \cap p^{i+1}; b_i = \frac{|c_i^k|}{|b_i^k|} - \theta. \quad (13)$$

Здесь запись p^{i+1} означает параметр, соответствующий рассматриваемому расстоянию b_{i+1} .

$$\frac{|c_i^k|}{|b_i^k|} - \theta \geq b_{i-1} \Rightarrow c_{i+1}^k = c_{i-1}^k \cap p_{i+1}; b_{i+1}^k = b_{i-1}^k \cap p_{i+1}; b_i = b_{i-1}. \quad (14)$$

Работа данного алгоритма может выполняться в параллельном режиме, разными потоками. Процедура может запускаться в фоновом режиме, формируя предложения по формированию новых стандартов, которые после соответствующего подтверждения вступают в силу.

Вывод. В данной работе описан алгоритм самообучающейся экспертной системы типа «без учителя» суть которого заключается в исследовании существующих стандартов с поэтапным включением в анализ дополнительных параметров. Такой подход позволяет выявить объективную причину отклонения, обнаружить устойчивые статистические группы, усовершенствовать стандарты, которые создают основу для планирования ресурсов медицинского учреждения. Данный алгоритм ориентирован на автоматическое накопление и формирование знаний с использованием процедур анализа и обобщения данных.

ЛИТЕРАТУРА

1. Дзяк Г.В., Березницкий Я.С., Филиппов Ю.А. и др. Библиотека практического врача. Унифицированные клинико–статистические классификации болезней органов пищеварения (ведомственная инструкция). – Киев, 2004. – 93 с.
2. Mary Beth Chrissis. CMMI® for Development Guidelines for Process Integration and Product Improvement, Addison-Wesley Professional; 3 edition (March 20, 2011). – 688 p.
3. Литвинов О. А. Формалізація клінічного діагнозу і модель електронної класифікації діагнозів для Медичних Інформаційних Систем України. // Системные технологии. Рег. межвуз. сб. научн. работ. – Выпуск 2(49). – Днепропетровск, 2007. – С.46-56.
4. А.А Литвинов, Г.В. Полухин, В.В. Истомина. Оценка и обнаружение стандартов клиники хирургического профиля. Системные технологии. Регион. меж вуз. сборник научных работ.– Выпуск 1(90) –Днепропетровск, 2014.-С.48-53