УДК 614.2+574/578+004.38

A.A. Litvinov, A.D. Mizumskaya, V.V. Istomina

SPECIFIC OF MEDICAL STANDARDS ESTIMATION AND DISCOWERY PROCEDURE FOR RESOURCE PLANNING

Annotation. The paper is devoted to the algorithm of self-learning expert system to detect high-quality medical standards estimation and discovery. This approach allows to discover the reason of deviation, stable statistical groups, to polish and improve the standards which can be regarded as a foundation of resources planning system.

Keywords. Standard, expert systems, learning without a teacher, the algorithm.

Introduction. Today the creation and implementation of medical and economic standards plays a crucial role in the institutions, both public and private type. The main objectives in this case are: clinical and economic substantiation of health care [1]; building a database to achieve predictable, stable operation of the clinic [2]; effective planning of activities and resources; forecasting and risk management. Thus, the development of effective approaches to building standards, methods of selection, monitoring and evaluation is an important task in the construction of information systems that support the introduction of standardization. Standard of care can be defined as a set of protocols that reflect the specificity of services for the clinical condition [1, 3], taking into account many different parameters associated with the patient's condition and properly (age, sex, weight, reaction to drugs etc.) and affect the need to perform some or other medical / diagnostic activities.

More formally, the protocol can be defined as if–then rules describing the reaction of the system (in this case the clinic) for a specific clinical condition of the patient, indicating the expected result. This reaction is expressed as a workflow consists of a set of activities, each of which is clearly defined in terms of space (resources, materials, performers) and time (duration of stay, surgery) restrictions.

$$s_i : d_j \to w_k,\ s_i \in S,\ d_j \in D,\ w_k \in W, \tag{1}$$

s_i –protocol standard, d_j – the clinical condition of the patient, w_k – workflow.

$$w_k = a_i^p \square a_l^s \square \ldots \square a_n^r = \langle a_i^p, a_l^s, \ldots, a_n^r \rangle, \tag{2}$$

$$a_i^p, a_l^s, \ldots, a_n^r \in A;\ p, s, r \in A^t,\ \begin{matrix} \langle t, \tau, e, c, r \rangle_i \\ a_i = \end{matrix}$$

ISSN 1562-9945

A^t – the types of activities, e – subset of the roles of performers (the number and type performer), r – set of typed resources, t – time start of the activity, τ – duration, c – costs.

An important part of the protocol is the prediction the expected result, which is expressed in economic indicators such as the duration of hospital stay *bd*, treatment costs *c*, time in the operating room (in the case of surgical cases) *ol*.

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

On the basis of the protocols is possible to perform efficient operational and long-term (taking into account the structure of the service requests) planning of the hospital. Real-life situations may require the amendment of the set of active base workflow that can influence the result, disrupting the operation of the mechanism of planning. Solve of this problem should be envisaged revaluation of existing and new standards of detection protocols. Question nontrivial obtain multiple standards is closely related to the methods of data mining, which is usually based on the creation of a training sample, the allocation of tangible characteristics, search functions, classifiers (by training the system), use of special products. The second issue is education - number of classes is not known beforehand, and the application of clustering methods without a teacher as a result of such procedures may cause a large number of obscure, poorly structured classes. Methods of revaluation of these classes are also associated with additional costs are not always amenable to automation, require specialists.

Task definition. The goal of the paper is providing a method of automated check and improving of existed standards and search of new, detailed standards, taking into account all the specter of parameters influencing the treatment strategy, the use of resources. It should be noted that we are not dealing with a predefined set of classes standards by which to distribute the data, our goal is to find the search for new standards based on existing ones, it should be noted adaptability and retraining system, depending on the replenishment of the database. The important features are: complete automation of such a search with minimal cost to the person responsible for the standards; minimizing resource and provide a reasonable time of the procedure search. In [4] considered the basic version of the algorithm, the main search strategy. In this paper, we consider an amended version.

Main part. The suggested solution is based on a preliminary description of a groups of standards a “clinical state – surgical operation”. Then, after its implementation into the operational environment, an automated research and estimation of the raw standard could be made, using the results of monitoring. The research is based on exploring the deviations of the selected parameters (the duration

of surgical operation, the number of hospitalization days, resources and costs). The exploration basically depends on a defined threshold value (it might be used an interval) used to detect the deviation in either direction. Whether the deviations are absent, the standard is mature and robust and the clinical situation defined by the parameters becomes predictable. The fact of deviation causes the automated exploration procedure searches of the reasons of its appearance. The goal of such exploration is to get a sample of patients united by the parameters (sex, age, weight, reaction to drugs), which always makes such deviation and thus causes the creation of a new protocol.

Identifying reasons for discrepancies is based on a comparison deviation with threshold - trigger that includes a search function causes deviations fingering group from a given order. This threshold is flexible for each protocol, defining the latter threshold. Purpose of procedure - to detect the group to which the threshold deviation exceeds a threshold, is responsible for the selection standard.

Algorithm. The first step is creating a set of pairs «clinical diagnosis – operation». It could be done in semi-automated mode by an expert who is responsible for resource estimation on the base of experience and statistics. The main features of resource are as follows: time of operation, time of hospitalization, required activities and resources caused the expenses on the treatment. Later there is an automatic update system standards based on the new material: a combination of "patient-parameters" could not be identified due to lack of necessary information in the database, the lack of consideration of additional parameters.

After determining deviations advisable to discover the cause of deviations, through the provision of sustainable statistical group of patients.

This group sets a precedent occurrence of a particular protocol.

Let us take the following notation for formal description of the algorithm:

$d_i \subseteq D$ - subset of patients with the same clinical diagnosis;

$o_j \subseteq O$ - subset of operated patients with the same surgery;

$p_i \in P_l \subseteq P$ - parameter defined by the set of values, which is a subset of the total set of parameters that define the image and the state of the patient. For example, age group, gender, and complication of operation, etc.

$\delta_i \in \Delta$ - deviation on the parameter defined in the standard. For example, the fact of 2 days divergence for the post-operation period provided by the standard.

The first step is creating a set of pairs «clinical diagnosis – operation», which amounts the standard at time *t* before the search procedure start working.

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

θ - fixed threshold value (for all standards) determines the allocation of the new standard, if the proportion of patients with a deviation of a particular parameter exceeds this threshold, there is a revision of this standard or release a new one. In our case we use 0.75.

π_k^t – threshold that determines the probability that deviations (hidden standard) for the ***k***-th standard: the threshold to begin searching the reasons for rejection and possible release of the new standard. Initially, the threshold to begin searching the reason for rejection is given by $\pi_k^t = \frac{\theta}{2}$.

For a selected pair of "diagnosis-operation" analyze the deviations: operation time deviation, the period of treatment deviation. If the value increases the threshold θ , it causes the correction of protocol.

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

If the deviation exceeds the π_k^t , it causes additional parameters influence research:

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

Express the initial condition of the search - the set of patients with deviation from the standard.

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

In case of exceeding the threshold of the new standard is released:

$$\forall_{p_i \in P, p_i \Rightarrow b_0^k} \cdot \frac{|c_0^k \cap p_i|}{|b_0^k \cap p_i|} \ge \theta \Rightarrow NewStandard(c_0^k \cap p_i) \quad (8)$$

The procedure for selection of the new standard can be written as follows:

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

If none of the options did not give reasons for the selection of the standard - for each of the parameters that can influence the deviation, calculated the distance to the threshold selection standard d_i .

$$\forall_{p_i \in P, p_i \Rightarrow s_k^t} \cdot \frac{|c_0 \cap p_i|}{|s_k^t \cap p_i|} < \theta \Rightarrow d_i = \frac{|c_0 \cap p_i|}{|s_k^t \cap p_i|} - \theta \quad (10)$$

The result of this procedure is an ordered vector of distances in which the minimum distance is the first element. On this basis, sorted and set parameters.

$$d_1^i \le d_2^j \ldots . \le d_n^k \Rightarrow p_i^1 \Rightarrow p_j^2 \Rightarrow \ldots \Rightarrow p_k^n \quad (11)$$

Next is the formation of the conditions for the next iteration of the following rules.

If the proportion of patients on the condition c_i^k is greater than the threshold θ - the new standard should be allocated according to the procedure described by the formula (9).

$$\left|c_i^k\right| / \left|b_i^k\right| \geq \theta \square \quad halt\,;\, NewStandard\left(c_i^k\right);\, \pi_k = d_i. \tag{12}$$

If the proportion of patients with deviations from the standard is higher than the proportion obtained in the previous iteration, but less than the threshold standard selection a parameter considered in the current moment of the ***i***-th iteration, taken into account in the next sample.

$$\left|b_i^k\right| - \theta < d_{i-1} \square \quad c_{i+1}^k = c_i^k \cap p^{i+1};\, b_{i+1}^k = b_i^k \cap p^{i+1};\, d_i = ¿ \left|c_i^k\right| / \left|b_i^k\right| - \theta. \tag{13}$$

$$\left|c_i^k\right| / ¿$$

Here record p^{i+1} is a parameter corresponding to the distance considered d_{i+1}.

$$\left|c_i^k\right| / \left|b_i^k\right| - \theta \geq d_{i-1} \square \quad c_{i+1}^k = c_{i-1}^k \cap p_{i+1};\, b_{i+1}^k = b_{i-1}^k \cap p_{i+1};\, d_i = d_{i-1}. \tag{14}$$

It is notable, that procedure can be optimized running concurrent threads. The procedure can be run in background mode, making the propositions of new standards, which can be then confirmed and started using by the system.

Conclusion. The algorithm of self-learning expert system without teacher of estimation and discovery of surgical clinic medical standards with step-by-step exploration of additional parameters was described. This approach allows to discover the objective reason of «standard – fact» deviation, get stable statistic groups, improve standards, build the strong foundation for surgical clinic resources planning. This algorithm is destined for automatic accumulation and creation of knowledge using the procedures of analysis and generalization of data.

REFERENCES

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