Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 119 (2017) 207–214

Procedia Computer Science

www.elsevier.com/locate/procedia

6th International Young Scientists Conference in HPC and Simulation, YSC 2017, 1-3 November 2017, Kotka, Finland

Towards a simulation-based framework for decision support in healthcare quality assessment

Ilia Kisliakovskii[a,*], Marina Balakhontceva[a], Sergey Kovalchuk[a], Nadezhda Zvartau[a,b], Alexandra Konradi[a,b]

[a]*ITMO University, St. Petersburg, Russian Federation*
[b]*Almazov National Medical Research Centre, St. Petersburg, Russian Federation*

Abstract

This article describes a simulation framework for healthcare quality assessment from the perspective of management and corresponding decision-makers. The proposed framework will allow simulating "what-if" scenarios and getting alternative outcomes in case of decision support systems. In our research, we are dealing with heterogeneous data sources, which combine within data-flow processes. The data flow of the presented framework conceptually combines several modeling methods: discrete-event simulation, agent-based modeling and also includes data analysis. The experiments were executed based on data from the Almazov National Medical Research Centre hospital information and access control systems.

Peer-review under responsibility of the scientific committee of the 6th International Young Scientist conference in HPC and Simulation

Keywords: value-based medicine; decision support system; simulation in healthcare; healthcare quality assessment; agent-based modeling; discrete-event simulation.

1. Introduction

Nowadays, healthcare is experiencing a paradigm shift. There is a transition from the volume-based to value-based approach, from maximum assistance to all patients to individual service and care quality improvement along

* Corresponding author. Tel.: +7-812-909-31-56.
E-mail address: kisliakovskiii@yandex.ru

1877-0509 © 2018 The Authors. Published by Elsevier B.V.
Peer-review under responsibility of the scientific committee of the 6th International Young Scientist conference in HPC and Simulation
10.1016/j.procs.2017.11.178

with efficient usage of resources. All healthcare processes become even more complex and consequently generate even more unstructured data. It makes management decisions more complicated. Existing HIS do not always allow proper decision-making support for medical professionals [1]. Therefore, in conjunction with extending of practical application and integration of HIS, it is important to develop new methods and technologies to support decision-makers within the paradigm of personalized medicine. The implementation of high-quality and well-timed decision-making support is only possible when using results of analysis of already existing data, and obtaining new knowledge based on the performed analysis.

At the same time, it's really necessary to improve the healthcare quality through, for example, an evaluation of patients' satisfaction. But in some cases, this type of evaluation is unavailable due to lack of feedbacks. That's why decision-makers can use some techniques from staff scheduling. These techniques allow optimizing workload but can't give any prediction in case of emergency or some unusual situations. For that reason, simulation and modeling had widespread application in healthcare.

2. Related works

Mankind experiences the growth of the life's length as well as the common population growth and it is the reason for pride. Meanwhile, the same achievements require strengthening of the economy, particularly, in the field of healthcare. There should be significant expenditures also because the world population is rapidly aging [2] and necessary technologies should be developed and deployed. Solutions of the costs' containment task vary in different countries. Nevertheless, the most useful were patient-oriented measures, e.g. more active treatment for people with multiple comorbidities and, in general, the increase in patient throughput in hospitals [3]. Also, staffing policies in healthcare are shown as ineffective since they consider volumes instead of the quality of care [4]. This marks the transition from volume-based to value-based approach and means the need to optimize the whole process of patient treatment depending on personal characteristics. In fact, the same idea has been voiced in the call of WHO to balance economy of means and quality improvement and appreciate patients' expectations [5].

One of the most important findings was that the nature of care and a warm attitude (primarily from nurses) determine the perception of the patient, not the technical side of care, as was previously thought [6]. This is also evidenced by patient-reported outcomes [7]. Reporting through surveys and questionnaires is the most widely used approach in assessing the quality of care [3,8–11]. It represents a fair feedback and allows giving an idea of overall healthcare statement at the hospital level. The nurse work environment is strongly associated with patient satisfaction [9] and it means the environment building is the mandatory step on the way to the main goal. Inattention to this leads to the burnout of nurses [3,8,11]. Even some studies state that nurse work overload has no effect on patient outcomes [8,10] it increases the number of errors of nurses the degrades the patient safety [12–14].

It should be noted that the perception of the quality of care for the hospital staff and the patient should be similar in order to avoid the loss of compliance [15] and managers should take this into account in corporate decision-making. Performance appraisals are required for effective management activities, i.e., promotions, training, coaching, etc. [16], which in turn leads to positive organizational and patient outcomes [10,11], including even patient mortality [17]. However, nurses are disappointed in the performance appraisals. There are identified subjectivity and lack of communication between managers and nurses as major points that must be eliminated in the performance appraisal system [18,19]. The system can be improved with the participation of all parties and improving the work environment while taking into account results of the assessment of the quality of care [11,20].

A fundamentally different approach to solving the problem of the quality assessment is the creation of data-driven models. Such models allow considering the interests of all parties objectively: to assess work environment from the staff point of view and the effectiveness of work from the management point. The key to objectivity is work with data from hospital information system (HIS) and electronic health records (EHR). Unfortunately, in most cases, the data does not conform to standardized terminology and has unusable (because of inaccuracy) timestamps [21]. A promising application of models is possible when assessing the workload of nurses. Stored data of admissions and discharges of patients allows obtaining accurate assessments [22]. The detection of overloads can have a positive effect on the quality of care as it was said earlier. To build a comprehensive assessment, it is necessary to consider nurse interventions based on HIS data, influence of doctors and managers.

3. Proposed framework

The proposed framework allows decision-makers to have a set of alternative scenarios in real-time. It consists of several models linked by data flow (Fig. 1). At the first stage, we need to model the patient inflow by using results of Electronic Health Record (EHR) data analysis. This model had been already developed in our research group and was presented in [23]. Combining patient population data with access control system (ACS) data and staffing tables we obtain the input data for the next stage. The next step in the framework is a discrete-event simulation of medical center activity. This step needs to evaluate possibilities and workload for each department separately and for the whole hospital as well. Such model can get departments' workload outputs and special key-values for departments which can help in the quality assessment. Although, this step can be based on machine learning techniques [24]. Analysis of the workload for different staff groups is a very important aspect in case of decision-makers' support. Further, it is necessary to optimize regular staff tables and schedules. Also, it helps to obtain schedules for each employee to determine peak hours. These data allows moving from discrete-event to agent-based modeling and receiving updated data for the emergency scenario. Agent-based modeling gives an opportunity to present internal hospital processes in continuous space. Simulation results of departments' dynamics are individual tracks in geometric space for each employee.

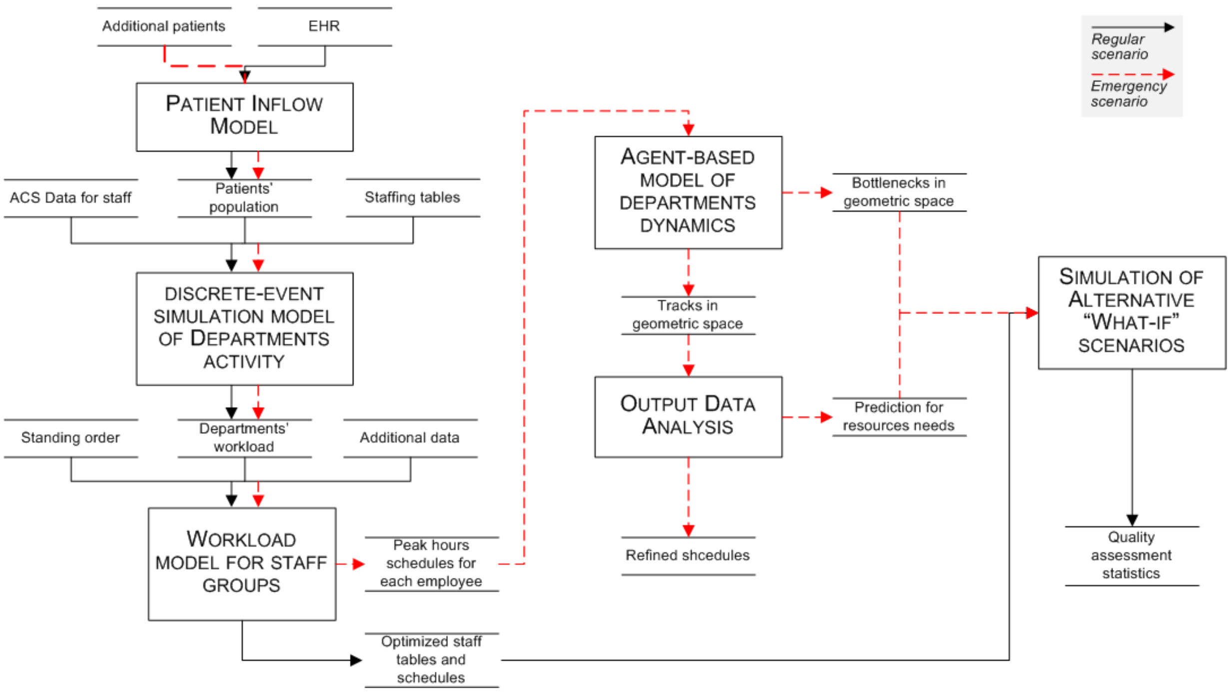

Fig. 1 Simulation framework data flow diagram.

This framework can support decision makers through "what-if" scenarios and alternative outcomes. It is really necessary for decision-making under condition of high risks, incompleteness, and unreliability of information. Two scenarios were chosen for the framework evaluation. They are characterized by different external environmental data: a regular medical center workload and an emergency workload. The first scenario is regular; it means that there are no unexpected additional patients in the hospital. This scenario can provide management decisions about staff schedules and other resources. The emergency scenario can help in unpredictable situations; in a case with additional patients.

The framework can be adapted to use additional datasets, e.g., feedbacks from patients, results of surveys. After data output analysis step, first, we have the set of refined individual schedules and can rerun agent-based model;

second, we have predictions for resource needs for a quality assessment and can obtain alternative scenarios for decision-makers.

4. Methods and case study

For the test scenario, we used data from HIS and ACS of the Almazov National Medical Research Centre. First, we focused on one group of specialists. Daily nurses (ones who works twenty-four-hour shifts) were chosen because this staff group already have the most complex and volatile workload at the cardiology department of Almazov Centre. The cardiology department was chosen because of diverse patient population. It allows developing regular and emergency scenarios.

The dataset consists of ~2’000 EHR and four-month access control data for each daily nurse (~10’000 records). The first task was to evaluate measures for workload assessment for each employee. But we are faced with the following difficulties: it is responsibilities of this staff group are not strictly regulated; it is impossible to obtain numerical criteria for work quality assessments.

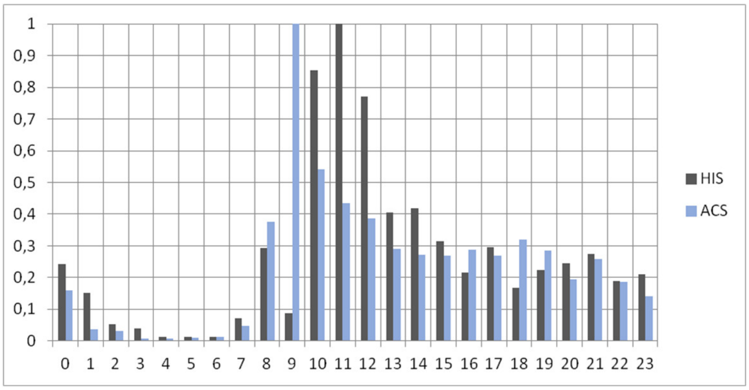

Fig. 2 Analysis of events from ACS and HIS.

If we analyze the number of events in the data on the peak day of the week and consider the upper value for the maximum, then it turns out that in essence, nurses should not have any overloads. But there is overload; this is evidenced by the nurses themselves: there is a shortage of personnel and it is impossible to accurately calculate how many hours a nurse needs to spend on each patient for satisfactory quality of care. Fig. 2 presents a simple data analysis that tells us that the peak hours for HIS and ACS are different. It seems like data is entered into the system when nurses have a break, but not at the time of the event's occurrence. Therefore, at this stage, we choose data from the ACS to assess the quality of the indirect characteristics as the most reliable reflecting the real dynamics.

4.1. Clustering

The data from the access control system represents the sequence of events, each of which is a pass through the door using a security badge. Therefore, the event consists of the nurse’s identifier, the date and time of the pass, the door’s identifier (includes the floor’s number). The source dataset contains 3931 lines (events). The advantage of this sort of data is that it is collected automatically and that’s why it cannot be arbitrarily changed or shifted in time by a nurse. Furthermore, reasons, directions, and amount of displacements undoubtedly related to the workload.

First of all, we grouped sequences (or “chains”) of the events to strings where each symbol is a floor to track the movement of nurses within the hospital. One sequence corresponds to the movement of one nurse during one twenty-four-hour shift. Contrary to expectations, collected sequences did not look alike, though there must be some patterns among usual work shifts. It should be noted that a sequence includes symbols of floors, but not concrete door (e.g. in the sequence “A00011” each digit denotes the floor of the hospital building, A-letter denotes checkpoints at the entrance to the hospital territory). This approach has been chosen in order to first estimate nurses' movements in more general scale. To reveal some features of nurses' routes it was decided to cluster the sequences.

In the context of machine learning, we can assume that any sequence of characters is a word. Therefore, to cluster the sequences we can use word-specific algorithms. One of the main methods of calculating the distance between words is the Damerau-Levenshtein method. Calculating the distance, it takes into account substitutions and transpositions of characters and also insertions and deletions which in turn helps to solve the problem of the alignment of vectors (different lengths of sequences). Another way to calculate distances between the sequences comes from the counting the number of occurrences of each pair of characters in one sequence (i.e. the count of "00" pair in the sequence is the first component of the vector, the count of "01" pair – second, etc.), which to some extent is similar to the formation of "clickstreams" [25]. As a result, we obtained plain numeric vectors to which classical methods of distance calculation can be applied. As an experiment, we reproduced the formation of "clickstreams" by collecting sequences from pairs of symbols where the first symbol of pair was still the number of the floor, but the second one was the time gap (in hours) between current and next event (e.g. "3M" means the third floor and 22-hour gap before the next event). Unfortunately, use of such sequences led to poor results of clustering in comparison with other implemented approaches.

There were selected two different algorithms for clustering: Agglomerative Clustering was selected among hierarchical algorithms and Affinity Propagation – among nonhierarchical algorithms. The idea was to observe separability of sequences and estimate the potential number of clusters during agglomerative clustering, then compare the obtained number to another computed by second algorithm. One of the main features of Affinity Propagation is the selection of the number of clusters by the algorithm itself.

The main point of agglomerative clustering is that all vectors, initially being separate clusters, gradually merge into larger clusters and it is possible to track all the history of merging. Before performing clustering, you need to choose a method for calculating the distances between clusters and a distance function for pairs of vectors. In our case, Farthest Point method used together with the above-mentioned Damerau-Levenshtein function for vectors-strings and also Ward's method – with the Euclidean function for numeric vectors. The second combination allowed to obtain more relevant results with more clear division into clusters when the first one has shown extremely uneven distribution between clusters. Truncated by 12 last mergers dendrogram representing the results of hierarchical clustering are shown in Fig. 3.

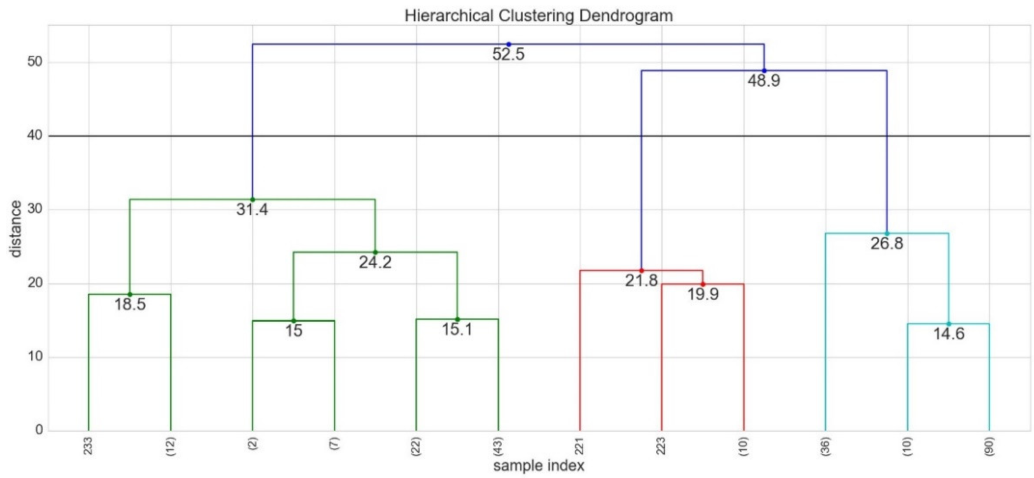

Fig. 3 Dendrogram of agglomerative clustering (numbers in brackets mean counts of samples in the cluster).

The affinity propagation method is based on the metaphor of elections. The points (vectors) can pass messages to each other, elect a leader-point and merge around it into a cluster [26]. In this case, Damerau-Levenshtein, Euclidean (Squared-Euclidean) and Manhattan (to check outliers' impact) pairwise distance functions were used. The best results (Silhouette coefficient 0.57) were obtained using Euclidean function. The total number of clusters was 3 what is equal to the optimum number obtained from results of hierarchical clustering, as seen in Figure 1. "Optimum" was named the number, transition to which require a minimal increase in distance allowed for merging clusters (y-axis). These clusters are situated under the black line in Fig. 3.

Both clustering methods give approximately the same distribution of sequences between clusters. The results of clustering indicate that a sequence from a cluster cannot refer to a specific group of nurses or a concrete weekday. In other words, it means that a sequence of events does not depend on a nurse personality or a weekday, but presumably depends on emergency cases or workload, in general.

4.2. Frequency analysis

Performed clustering did not reveal specific patterns in sequences of events. The fact of the independence of the sequences from nurses themselves and success in the processing of sequences by means of division into pairs of symbols (bigrams) led us to the idea to build common, or average, sequences through the frequency analysis. The most frequent combination of door codes per each half an hour presented in Fig. 4. Each symbol means one event from ACS.

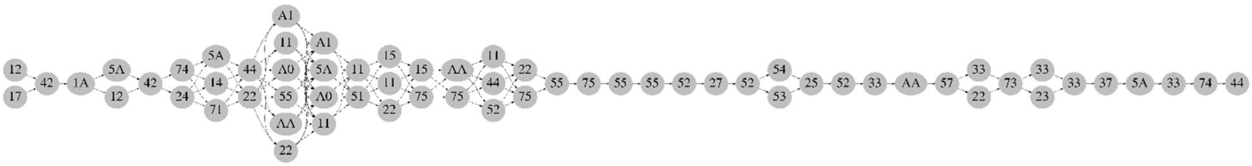

Fig. 4 Graph of the most frequent combinations of events

The cardiology department, considered for this case study, is located on the 7th floor, which corresponds to code '7'. The remaining codes also correspond to the floor number on which the door is located, the code 'A' means passing through the checkpoint (entrance/exit control). Thus, we managed to get an average daily scenario of the nurse's behavior. In the future, this graph must be expanded with events from the HIS. At the current stage, it is impossible to use data from the HIS, because according to observations, nurses contribute data to the system spontaneously.

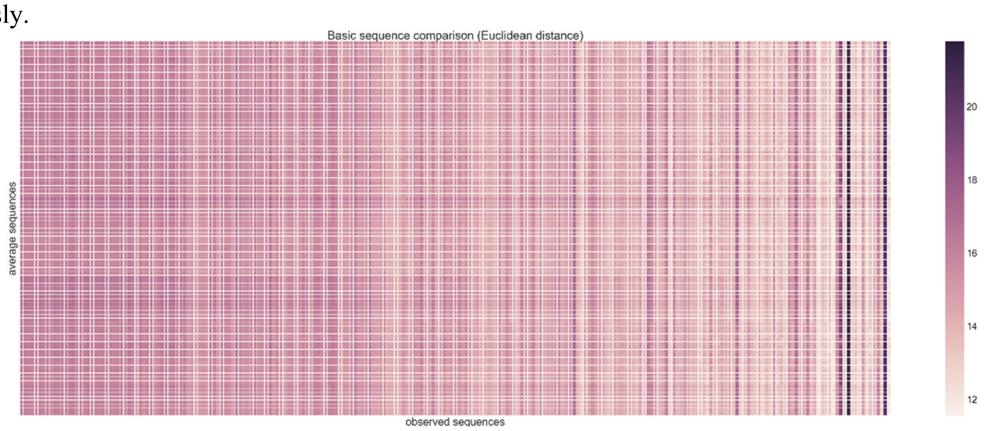

Fig. 5 Heatmap of distances between average and observed sequences.

According to the graph (Fig. 4) all possible “average” sequences were built. There were 96 sequences what is enough to try to reveal a common pattern, but it does not cover completely the space of possible sequences of events. The number of sequences depends on degrees of the graph’s vertices, but the majority of degrees is equal to one, which in turn indicates a significant lack of data. In order to discover how well average sequences are close to observed ones, we compared them accordingly as numeric vectors by the algorithm described above. The results of the comparison are shown in Fig. 5. While a row of the heatmap is one of average sequences, a column is one of observed sequences, their intersection is the Euclidean distance between these sequences. As can be seen on the map, average sequences slightly differ from each other, observed sequences are removed from them by approximately the same distance, and outliers, that mean unique work shifts, (dark purple columns on the map) are rarely detected.

4.3. Agent-based modeling

Problems of optimization of healthcare processes require more complex solutions. In particular, the computerized medical systems usually operate in a distributed environment, whose components continuously exchange data. Flexibility, reliability, and agility of multi-agent systems allow them to work effectively in such conditions. The distributed architecture enables multi-agent systems to solve performance issues and resource constraints. Each agent can be described separately. In the interaction of realistically modelled entities, complex forms of behavior arise naturally. This makes it possible to effectively use agent models to study the processes and build forecasts under the real dynamics reproduction.

For the agent-based modelling, we used PULSE framework [27], which allows presenting each agent separately and configuring common behaviour for each group of staff. In Fig. 6 (a) you can see box-plots which show the relative nurse workload in cardiology department based on the number of patients. The maximum allowable nurse workload, according to the expert assessment, corresponds to a value of 1.0. It gives us the opportunity to estimate peak hours. Fig. 6 (b) shows track proximity metrics for daily nurses. These metrics received through agent-based modeling results. Obtained tracks of nurses do not give us the right to conclude and should be specified by additional data of events at the department. In this regard, at this stage, it is impossible to go to the stage of developing alternative scenarios for decision-makers.

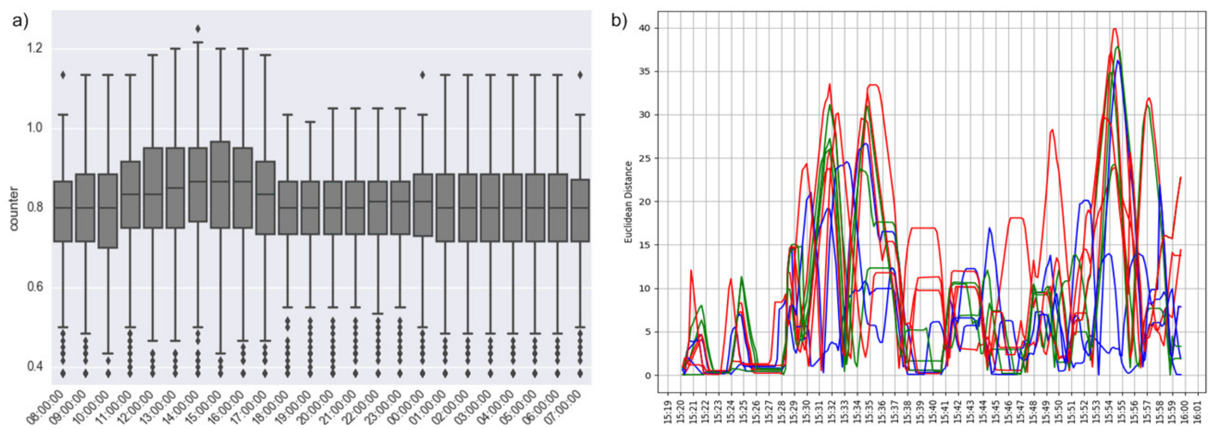

Fig. 6 (a) number of patients based cardiological department workload; (b) number of patients based cardiological department workload.

5. Future works

In the future, it is planned to develop a mechanism for updating individual schedules of nurses by analyzing text data from EHR using text-mining methods. It is also necessary to obtain additional survey data for a more accurate assessment of the work of nurses. Finally, it is possible to extend the data set using RFID sensors. Also, we plan to use the annual dynamics data analysis from ACS and include some external data in our framework.

It is worth noting that in the construction of a complex assessment of the quality of medical care, in the future is planned to use the System Dynamics approach. This will allow us to find implicit behavioral patterns of staff groups, as well as reflect the dynamics of an ever-changing system, for example, in case of new laws are adopted in healthcare.

Acknowledgements

This research is financially supported by The Russian Scientific Foundation, Agreement #17-71-10259.

References

[1] Morquin D, Ologeanu-Taddei. Professional Facing Coercive Work Formalization: Vicious Circle of the Electronic Medical Record (EMR) Implementation and Appropriation. Procedia Comput Sci 2016;100:652–7. doi:http://dx.doi.org/10.1016/j.procs.2016.09.207.

[2] World Health Organiztion. Ageing and health. WHO 2015. http://www.who.int/mediacentre/factsheets/fs404/en/ (accessed July 11, 2017).

[3] Aiken LH, Sermeus W, den Heede K, Sloane DM, Busse R, McKee M, et al. Patient safety, satisfaction, and quality of hospital care: cross sectional surveys of nurses and patients in 12 countries in Europe and the United States. Bmj 2012;344:e1717.

[4] Sermeus W, Aiken LH, den Heede K, Rafferty AM, Griffiths P, Moreno-Casbas MT, et al. Nurse forecasting in Europe (RN4CAST): rationale, design and methodology. BMC Nurs 2011;10:6.

[5] World Health Organiztion. The world health report 2000: health systems: improving performance. WHO; 2000.

[6] Attree M. Patients' and relatives' experiences and perspectives of "Good" and "Not so Good" quality care. J Adv Nurs 2001;33:456–66.

[7] Leonova MV. Patient-reported outcomes in clinical studies. Качественная Клиническая Практика 2016.

[8] You L, Aiken LH, Sloane DM, Liu K, He G, Hu Y, et al. Hospital nursing, care quality, and patient satisfaction: cross-sectional surveys of nurses and patients in hospitals in China and Europe. Int J Nurs Stud 2013;50:154–61.

[9] Kutney-Lee A, McHugh MD, Sloane DM, Cimiotti JP, Flynn L, Neff DF, et al. Nursing: a key to patient satisfaction. Health Aff 2009;28:w669--w677.

[10] Teng C-I, Hsiao F-J, Chou T-A. Nurse-perceived time pressure and patient-perceived care quality. J Nurs Manag 2010;18:275–84.

[11] Aiken LH, Sloane DM, Clarke S, Poghosyan L, Cho E, You L, et al. Importance of work environments on hospital outcomes in nine countries. Int J Qual Heal Care 2011;23:357–64.

[12] Kane RL, Shamliyan T, Mueller C, Duval S, Wilt TJ. Nurse staffing and quality of patient care. Evid Rep Technol Assess (Full Rep) 2007;151:1–115.

[13] Liu L-F, Lee S, Chia P-F, Chi S-C, Yin Y-C. Exploring the association between nurse workload and nurse-sensitive patient safety outcome indicators. J Nurs Res 2012;20:300–9.

[14] Gabbay U, Bukchin M. Does daily nurse staffing match ward workload variability? Three hospitals' experiences. Int J Health Care Qual Assur 2009;22:625–41. doi:10.1108/09526860910986885.

[15] Janse AJ, Gemke R, Uiterwaal C, Van Der Tweel I, Kimpen JLL, Sinnema G. Quality of life: patients and doctors don't always agree: a meta-analysis. J Clin Epidemiol 2004;57:653–61.

[16] Cardy RL, Korodi C. Nurse appraisal systems: characteristics and effectiveness. Soc Sci Med 1991;32:553–8.

[17] West MA, Borrill C, Dawson J, Scully J, Carter M, Anelay S, et al. The link between the management of employees and patient mortality in acute hospitals. Int J Hum Resour Manag 2002;13:1299–310.

[18] Spence DG, Wood EE. Registered nurse participation in performance appraisal interviews. J Prof Nurs 2007;23:55–9.

[19] Nikpeyma N, Abed_Saeedi Z, Azargashb E, Alavi_Majd H. Problems of Clinical Nurse Performance Appraisal System: A Qualitative Study. Asian Nurs Res (Korean Soc Nurs Sci) 2014;8:15–22. doi:10.1016/j.anr.2013.11.003.

[20] West M. How can good performance among doctors be maintained? 2002.

[21] De Cordova PB, Lucero RJ, Hyun S, Quinlan P, Price K, Stone PW. Using the nursing interventions classification as a potential measure of nurse workload. J Nurs Care Qual 2010;25:39.

[22] Beswick S, Hill PD, Anderson MA. Comparison of nurse workload approaches. J Nurs Manag 2010;18:592–8.

[23] Abuhay TM, Krikunov A V, Bolgova E V, Ratova LG, Kovalchuk S V. Simulation of Patient Flow and Load of Departments in a Specialized Medical Center. Procedia Comput Sci 2016;101:143–51. doi:http://dx.doi.org/10.1016/j.procs.2016.11.018.

[24] Funkner A, Kovalchuk S, Bochenina K. Preoperational Time Prediction for Percutaneous Coronary Intervention Using Machine Learning Techniques. Procedia Comput Sci 2016;101:172–6. doi:http://dx.doi.org/10.1016/j.procs.2016.11.021.

[25] Wang G, Zhang X, Tang S, Zheng H, Zhao BY. Unsupervised Clickstream Clustering for User Behavior Analysis. Proc. 2016 CHI Conf. Hum. Factors Comput. Syst. - CHI '16, 2016, p. 225–36. doi:10.1145/2858036.2858107.

[26] Dueck D, Frey BJ. Clustering by Passing Messages Between Data Points. Science (80-) 2007;315:972–6. doi:10.1126/science.1136800.

[27] Karbovskii V, Voloshin D, Karsakov A, Bezgodov A, Gershenson C. Multimodel agent-based simulation environment for mass-gatherings and pedestrian dynamics. Futur Gener Comput Syst 2016. doi:http://dx.doi.org/10.1016/j.future.2016.10.002.