An E-health Evaluation Case study : Evaluating The New Laboratory Information System

Güney Gürsel, Hüsamettin Gül
Department of Medical Informatics
Gülhane Military Medical Academy (GATA)
Ankara, TURKEY
ggursel@gata.edu.tr, hgul@gata.edu.tr

Kaya Kuru
School of Computing, Engineering and Physical Sciences
University of Central Lancashire
Southampton, UK
kkuru@uclan.ac.uk

***Abstract*—Healthcare is one of the technology-intensive areas. Almost all healthcare organizations use an information system, without it, managing daily works and providing the continuity of healthcare is impossible. Information technologies staff has to support, manage, and improve the information system. To do this, they are supposed to foretell the hardware and software requirements, improve the system they manage in the competitive environment, to survive, and further to pioneer. In this respect, evaluations are carried out to reveal the weak and strong sides of information systems in operation. In this sense, a case study is performed in this study, to evaluate a healthcare information system. Particularly, the recently deployed laboratory information system (LIS) is evaluated by means of questionnaires, applied to both patients and users of the laboratory information system. Laboratory information system is evaluated on the basis of Function sufficiency, Decreasing Work Load, Speed, Learning Ease, Improving Service Quality, Availability, Help Manuals, User Satisfaction, and Patient Satisfaction features. The features needing to be improved in terms of the effectiveness and efficiency of LIS are measured based on the threshold value. The results are presented in a variable table according to the threshold value selected by the evaluator. As the target threshold value increases, the number of features needing to be improved also increases.**

Keywords- evaluation; healthcare information system; laboratory information system; e-health evaluation

I. INTRODUCTION

Healthcare industry is growing and developing rapidly, not only in health services, but also in information technologies (IT) related to it. Particularly electronic health, E-Health, is the main IT related area to meet the immediate needs of this industry. E-Health is in the intersection of medical informatics, public health, and business; it can be defined as the use of information and communication technologies to improve healthcare [1]. From primary care institutions to big healthcare centers, every healthcare organization uses an information system, named as Healthcare Information System (HCIS). HCIS is the system composed of data, workflows, users, and technology; used to collect, store, process, and provide the needed information to support healthcare institutions and professionals [2]. HCISs are composed of several components such as hardware, software, data, database, workflows, business-driven intelligent approaches; to support healthcare institutions and professionals, in terms of collecting, storing, processing, and disseminating the required work-based information. Moreover, these systems are in the interest of many actors such as engineers, technicians, physicians, nurses, laboratory staff, administrative staff, managers, governmental and private social security/profession institutions, and patients. Hence, these systems include many different levels of actors; they should be open systems to help these actors interact with each other. The purpose of a HCIS is; to contribute to a high quality, efficient health care, for patients, consumers, and medical research [3]. HCISs are more complex when compared to other systems, because they incorporate into many sub-systems such as Radiology Information system (RIS), Laboratory Information System (LIS), Picture Archiving and Communication Systems (PACS), Hospital Information System (HIS). Thus, they should be supported by established intelligent mechanisms to manage this level of complexity.

With the growth in the industry, the need for qualified computer support of healthcare organizations grows proportionally. Almost all the hospitals have a HIS; all laboratories have a stand-alone or a HIS built-in LIS. Ironically, although health institutions invest huge amounts in Information Systems (IS), it is estimated that nearly 60-70% of IT implementation projects fail in healthcare [4]. IS projects in other fields share similar aftermath with the healthcare as well. They have bad reputation for exceeding budget and schedule, failure in realizing the expectations and having poor return of investment [5]. Literature tells, of 260.000 projects, 25% were cancelled before finish, 47% exceeded the budget [5]. These findings substantiate that, a huge amount of money is lost together with invaluable efforts and time. Loss of confidence to the systems is the worst of all.

Literature shows “Improving IT Quality” as one of the top five concerns that face IT staff [6]. To improve, first we should know the weak sides of our IS. Taking the current picture will reveal the points to improve, by measuring the level of success and failure. “You can’t manage it, if you

 ISBN: 978-1-61208-384-1

can't measure it" tells the importance of measuring the quality of your system [7].

To improve IS, in our context it is HCIS, it must/should be evaluated from the time being started to be developed, to the time taken out of operation, i.e. in the system's life cycle, iteratively [8-10]. These iterative evaluations help eliminate the reasons of bad reputations of HCISs given above, by means of early recognition of the problems. They also help eliminate the implementation problems by means of on-time interventions [11].

Briefly, evaluation can be defined, by drawing from the literature, as "measuring the extent of meeting the specified criteria of a system, in a specified context" [12]. Evaluations can be made both by government and public sector organizations; fortunately, the number of evaluations is rapidly increasing [13].

Implementation of a new HCIS is not an easy process. There exist many problems and challenges [14]. Some of these problems and challenges may be technical (low speed system, frequent outages, etc.), and some of them may be organizational or user dependent (poor implementation planning, resistance to change etc.). With a rigorous early deployment evaluation, these problems and challenges can be determined early and improved before the problem deteriorates.

The structure of the manuscript is organized as follows: In section 2, "Materials and Methods", Materials used in the study and the methods used to get the study results will be described, in section 3, "Results", the results of the study will be presented without any comment, and these results will be further discussed in the section 4, "Discussion". The findings of the study and the proposed future work will be in the section 5, "Conclusion and Future Work" part.

A. Motivation

In a Hospital of 1700 HIS users, Biochemistry department outsourced its LIS and quit using the built-in LIS of the HIS. This new system takes the orders of hospital from biochemistry and needed information from the HIS, then disseminates these orders to the related auto analyzers. After the auto analyzers are through with the tests, it gives some facilities to the Biochemistry doctors (Such as delta checks). Finally, if the responsible doctor approves the test result, LIS sends the results to the HIS.

In the old system, the orders were seen in the work lists of the staff in built-in HIS module. An (only one) operator will make the "specimen received" action in the HIS and then the patient will attend a queue for giving specimen. Five nurses get the specimen simultaneously. All the auto analyzers were communicating with the HIS independently. The facilities provided with were limited.

In this study, evaluation of a newly implemented HCIS, namely LIS, is performed. The purpose of the study is to evaluate the LIS on the basis of Function sufficiency, Decreasing work load, Speed, Learning ease, Improving service quality, Availability, Help Manuals, User Satisfaction, Patient Satisfaction features; and get the early deployment evaluation results to determine the weak sides of the system to improve.

II. Materials and Methods

A. Questionnaires

Data are collected using the face-to-face questionnaire method. Two different questionnaires are prepared and applied to capture the evaluation results; one for the patients and one for the laboratory staff. Both staff and patients are asked to express their answers using 3-point Likert scale (Disagree, Partially Agree, Agree) ranging from 1 (Disagree) to 3 (Agree). 3 point Likert scale is used instead of 5 point scale, to prevent patients from hesitating between middle answers such as Moderately Agree, Moderately Disagree. In staff questionnaire, also 3 point Likert scale is used to keep the consistency with the patient questionnaire. Patient data are only used for evaluation for Patient Satisfaction whereas staff data are used for all features under evaluation.

B. Data

The questionnaires are applied to the patients visiting the laboratory in randomly selected days. 138 patients and 42 staff (all employees) have participated in the study voluntarily. Staff has biochemistry physicians, nurses, administrative staff, assistants, pharmacists, and biologists.

C. Statistical Analysis:

The internal consistencies of the answers to the questionnaires are measured by reliability coefficient (ρ) given in (1) to (5). Reliability is the degree of measurement being consistent and reproducible [15]. One important goal of a measurement study is to quantify the reliability of a measurement process. A reliability coefficient value of 0.7 or above is usually adequate, although higher reliability is always desirable. Measurements with reliabilities of 0.5 or less are rarely adequate for anything but preliminary research. In this study, ρ greater than 0.70, is considered reliable.

$$SS_{total} = \sum_{ij} X_{ij} - \frac{(\sum_{ij} X_{ij})^2}{n_i n_j} \quad (1)$$

$$SS_{objects} = \frac{\sum_{i=1}^{n_i} (\sum_{j=1}^{n_j} X_{ij})^2}{n_j} - \frac{(\sum_{ij} X_{ij})^2}{n_i n_j} \quad (2)$$

$$SS_{observations} = \frac{\sum_{j=1}^{n_i}(\sum_{i=1}^{n_j} X_{ij})^2}{n_i} - \frac{(\sum_{ij} X_{ij})^2}{n_i n_j} \quad (3)$$

$$SS_{error} = SS_{total} - SS_{objects} - SS_{observations} \quad (4)$$

(Reliability)

$$\rho = 1 - \left[\frac{SS_{error} / (n_i - 1)(n_j - 1)}{SS_{objects} / (n_i - 1)} \right] \quad (5)$$

Where

i is the number of users (or patients),
j is the number of questions,
X is the weight of the answer,
SS_{total} = Total sums of squares,
$SS_{objects}$ = Sums of squares for objects,
$SS_{objservations}$ = Sums of squares for answers,
SS_{error} = Sums of squares for error.

Answers to the questions are analyzed by nonparametric Kruskal Wallis test to determine if there is any difference between the branches of staff. $p < 0.05$ level is considered as statistically significant. Nonparametric test is used since data do not come from a normal distribution, regarding the normality test applied to the data.

D. Features Under Evaluation

Function sufficiency, Decreasing Work Load, Speed, Learning Ease, Improving Service Quality, Availability, Help Manuals, User Satisfaction, and Patient Satisfaction features of the LIS were evaluated.

TABLE I. RELIABILITIES

Staff	n	ρ
Physicians	8	0.81
Nurses	8	0.76
Operators	6	0.90
Laboratory assistants	9	0.79
Pharmacists	3	0.68
Biologists	8	0.96
Total	42	0.89

a. n = the number of staff
b. ρ = reliability

The final rating RF of the feature *j* is computed by

$$RF_j = \sum_{i=1}^{k} W_i R_i / n \quad (6)$$

where *k* is the number of Likert scales employed (3 for this study) , W is the weight (1 to 3) of the Likert scale *i*, R is the number of answers given as that Likert scale and n is the total number of answers.

RF can have values ranging from 1 to 3, where 1 is the worst and 3 is the best value. If RF of the feature is below the threshold value, the feature is considered as weak and needs to be improved.

III. RESULTS

A. Staff Data

As stated before, 42 staff have participated in the study. Although the study was on volunteered basis, whole staff participated.

In Table I, the reliability values which are calculated by (1) to (5) are given. Operators have the highest reliability with a value of 0.90, whereas Pharmacists have the lowest reliability with a value of 0.68. All the staff has satisfactory reliability values higher than 0.70, if we accept Pharmacist as 0.70 which is very near. Overall reliability is 0.89.

In Table II, RF values calculated by (6) are given. For RF values, Learning Ease is the first with a value of 2.86, whereas Availability is the last with a value of 2.36. Overall RF appeared to be 2.62.

To see the statistical significant difference among the branches of staff, statistical comparison is employed using Kruskal Wallis test. Only the difference in nurses appeared to be statistically significant in some features, and one in physicians and laboratory assistants. The descriptives about the statistically significant p values of the features under study of nurses are given in Table III. For the features Function Sufficiency, Decreasing Work Load, Availability and Help Manuals, nurses do not agree with the other groups.

In addition to nurses, physicians and laboratory assistants do not agree with other groups in Function Sufficiency as well ($p < 0.011$).

TABLE II. RF VALUES OF THE FEATURES UNDER STUDY

Feature	RF value
Function sufficiency	2.61
Decreasing Work Load	2.46
Speed	2.65
Learning Ease	2.86
Improving Service Quality	2.85
Availability	2.36
Help Manuals	2.46
User Satisfaction	2.82
Patient Satisfaction	2.59
Overall	**2.62**

a. RF = Final Rating

TABLE III. DESCRIPTIVES OF STAFF COMPARISON

Feature	p
Function sufficiency	0.011
Decreasing Work Load	0.009
Availability	0.019
Help Manuals	0.020

B. *Patient Data*

Reliability of the patients' questionnaire is 0.87. 38.40% of the patients stated that they applied for giving specimen for analysis and the rest were in the laboratory for taking results.

Majority of the patients (84.78%) stated that they had applied to the biochemistry department before, when the old system was in use.

The most important observation is the decrease in the duration of the processes. Patients who had applied before expressed that they have waited shorter than their previous application for transaction (86.96%).

The majority of the patients expressed that they were able to take the service more easily with fewer processes (94.20%).

IV. DISCUSSION

In this study, evaluation of a newly deployed HCIS, LIS, is performed using face-to-face questionnaires. Both the patients and the staff showed great interest in the study by a high rate of participation. The high participation shows that, stakeholders of the system (patients and staff in this study) consider the evaluations as an opportunity to express their feelings and problems faced, to whom in charge of developing and running the systems. The best way to send the message "your ideas are taken into account for improving the system" can be given by means of evaluations. That may certainly help increase user acceptance and attention to the information system. It won't be false if we disclose that the more user-centric the evaluation, the higher the participation is.

If we start from patients' results, having 84.78% patients that took service in both systems gives us a healthy comparison chance. Of them, 94.20% state they get the same services in fewer steps. That means; the new system has shortened the workflow and eliminated some outmoded steps. This is good for a new system; actually one of the most expected virtue of the ISs is to make renovations in the business. The majority of the patients experiencing both systems states that the service time is shorter (86.96%), which fortifies the renovation of the new system. Drawing from the findings of patients data, we can say that the new LIS satisfied the patients.

The results of reliability measures given in Table I substantiate that the study has a high reliability. It shows the internal consistency of the answers, which leads us to the true and unbiased results.

TABLE IV. WEAK FEATURES ACCORDING TO THRESHOLD VALUES

Feature	RF	t = 2.40	t = 2.50	t= 2.75
Function sufficiency	2.61			✓
Decreasing Work Load	2.46		✓	✓
Speed	2.65			✓
Learning Ease	2.86			
Improving Service Quality	2.85			
Availability	2.36	✓	✓	✓
Help Manuals	2.46			✓
User Satisfaction	2.82			
Patient Satisfaction	2.59			✓

RF = Final Rating
t = threshold
✓ = Improvement needed

Almost all the RF values are near or above 2.5. That is also good from the staff point of view. User Satisfaction, Improving Service Quality and Learning Ease have the highest values. Having a high value in Improving Service Quality feature is compatible with the patients' results. With these findings, it can be definitely said that, there is an increase in service quality with this system change. As in patients, the system appeared to have satisfied the staff as well. It seems, it is an easy to learn system.

Speed, Availability and Help manuals are the least ranked features, when compared to the others. In this study, for the new system, we can say that, these features are the main improvement needed areas, although they are not so bad.

The areas of improvement can be determined by a context and management dependent threshold value. We do not propose a threshold in this study. This threshold is relative to the context and situation. For a newly adopted system it can be 2.30 or something, while 2.70 or higher for a high standard-like management. In Table IV, the change in improvement needed areas is given according to the three different thresholds. It is one when the threshold is 2.4, it becomes three when the threshold is increased to 2.5 and becomes six out of nine when the threshold is 2.75.

According to the statistical analysis, there are problems in nurses, physicians and laboratory assistants. Different from other groups, they don't think the functions of the LIS are sufficient. When we think that these groups are the core staff of the laboratory, this finding should be seriously taken into consideration.

Evaluations give some extra messages as well. If the findings are lower than expected, there can be a lack of communication or training in the target users. They may not know some important features or they may not know how to use the system efficiently. Because of these reasons, the results may be lower than expected. The best solution would be eradicating the reasons and then reevaluate the system to see the difference before and after.

Unrealistic expectations are another point of bias in evaluations. In other words, the evaluation results become lower if the expectations of the users are unrealistically high. Some methodologies should be employed to keep the expectations in a realistic level. In this respect, Nevo and Chan find in their study that managers are able to generate

 ISBN: 978-1-61208-384-1

realistic expectations [16]. As the Ryker et al. put forth, if these groups' expectations from HCIS can be found unrealistic (very relative issue, so the management must be very careful to make this decision), the management can organize some committees and arrange interviews with these users to set realistic expectations [17].

As we have stated above, if there is a problem with the communication (if a variable is expected to give higher values but the result is low) then IS staff should organize on-site trainings and improve the communication channels with users.

V. CONCLUSION AND FUTURE WORK

Many international institutes, both governmental and nongovernmental, regulate many standards about the HCIS; they employ classification and nomenclature systems, security and privacy measures, and many other great effort products and mechanisms. Despite these huge great efforts, they fail. Consecutively some questions arise: How can HCISs be measured, to determine if they are good enough? Do we evaluate them properly? Do they really meet the needs of the owner institutes? Do we really need them? The answers to these questions are mostly overlooked, and eventually HCISs fail.

Evaluations should be done iteratively, both to get user acceptance, and improve the system. Users should be in the center, because, it is the users that makes a system better, it is the users that makes a wonderful system useless. It is the managements' ability that makes the users use the system properly. Evaluating the system in a user centric manner is an option to accomplish.

For a future work, this study can be deepened in staff group basis, to customize the system. A user group (like nurses in our study), can be unhappy with the system, while others are happy. To eradicate the problems of this group, a deeper study is required.

REFERENCES

[1] G. Eysenbach, "What is e-health?" Journal of medical Internet research, vol.3. issue 2, e-20, Jun. 2001, doi: 10.2196/jmir.3.2.e20.

[2] K. A. Wager, F. W. Lee, and J. P. Glaser, "Managing Healthcare Information Systems: A Practical Approach for Healthcare Executives", Wiley, San Francisco, 2005, pp. 92.

[3] R. Haux, "Health information systems- past, present, future", International Journal of Medical Informatics, vol. 75, Mar-Apr 2006, pp. 268-281, doi:10.1016/j.ijmedinf.2005.08.002

[4] E. Ammenwerth, C. Iller, and C. Mahler, "IT-adoption and the interaction of task, technology and individuals: a fit framework and a case study", BMC Med Inform Decis Mak, vol.6, Jan. 2005, pp.1-13, doi : 10.1186/1472- 6947-6-3.

[5] R. Stanley and L. Uden, "Why Projects Fail, from the Perspective of Service Science". The Seventh International Conference on Knowledge Management in Organizations: Service and Cloud Computing, Springer Berlin Heidelberg, January 2013, pp. 421-429, ISBN 978-3-642-30866-6.

[6] Z. S. Yang, L. Hao, and Y. H. Zhang, "A New Method for Information System with Good Performance". Advanced Materials Research, vol. 753, Aug. 2013, pp. 3014-3017, doi: 10.4028/www.scientific.net/AMR.753-755.3014, retrieved: Dec., 2014.

[7] N. Gorla, T. M. Somers, and B. Wong, "Organizational impact of system quality, information quality, and service quality". The Journal of Strategic Information Systems, vol. 19, issue 3, Sep. 2010, pp.207-228, doi: 10.1016/j.jsis.2010.05.001.

[8] D. A. Protti, "Proposal to use a balanced scorecard to evaluate Information for Health: an information strategy for the modern NHS (1998–2005)", Comput Biol Med, vol.32, , May. 2002, pp. 221-236, doi: 10.1016/S0010-4825(02)00017-3

[9] E. Ammenwerth, S. Gräber, G. Herrmann, T. Bürkle, and J. König, "Evaluation of health information systems—problems and challenges", International Journal of Medical Informatics vol. 71, Sep. 2003, pp.125-135, doi: 10.1016/S1386-5056(03)00131-X.

[10] H. Al-Yaseen, S. Al-Jaghoub, M. Al-Shorbaji, and M. Salim, "Post-Implementation Evaluation of HealthCare Information Systems in Developing Countries", The Electronic Journal Information Systems Evaluation, vol. 13, issue 1, 2010, pp. 9-16.

[11] A. W. Kushniruk and V. Patel, "Cognitive and usability engineering methods for the evaluation of clinical information systems", Journal of Biomedical Informatics, vol. 37, Feb. 2004, pp.56–76, doi : 10.1016/j.jbi.2004.01.003.

[12] G. Gürsel, N. Zayim, K. H. Gülkesen, A. Arifoğlu, and O. Saka, "A new approach in the evaluation of hospital information systems," Turkish Journal of Electrical Engineering & Computer Sciences, vol. 22 issue 1, Jan. 2014, pp.214-222, doi: 10.3906/elk-1110-15

[13] J. E. Furubo, R. C. Rist, and R. Sandahl, "International Atlas of Evaluation". Transaction Publishers. London, 2002, pp. 1–26.

[14] M. Berg, "Implementing information systems in health care organizations: myths and challenges", Int J Med Inf, vol. 64, Dec.2001, pp. 143–156, doi:10.1016/S1386-5056(01)00200-3.

[15] C. P. Friedman and J. C. Wyatt, "Evaluation Methods in Medical Informatics", Springer, 1997, pp. 89-117.

[16] D. Nevo and Y. E. Chan, "A temporal approach to expectations and desires from knowledge management systems", Decision Support Systems, vol. 44, Nov. 2007, pp. 298-312, doi : 10.1016/j.dss.2007.04.003.

[17] R. Ryker, R. Nath, and J. Henson, "Determinants Of Computer User Expectations And Their Relationships With User Satisfaction: An Empirical Study", Information Processing & Management, vol. 33, issue 4, Jul. 1997, pp. 529-537, doi : 10.1016/S0306-4573(97)00016-2.