Process Mining for Clinical Pathway

Literature Review and Future Directions

Wei Yang, Qiang Su
School of Economics and Management
Tongji University
Shanghai, P. R. China
yangvmail@163.com

***Abstract*—Clinical pathway (CP) is a tool to improve service quality and efficiency of the medical institutions. However, most of the clinical pathways designed with the traditional methods are static and non-adaptive. Recently, the process mining techniques are receiving increasing attentions. It can not only help the clinical pathways designers to discover the sequence of activities, but also provide execution information for analyzing variances and correcting design errors. In this article, first, a literature review is presented. In the analysis of 37 studies from the period 2004–2013, three research aspects (process discovery for clinical pathways design, variants analysis and control, continuous evaluation and improvement) are explored, and the weaknesses of the methods are analyzed. It is found that the mining algorithms developed are not efficient enough to deal with the unstructured processes, the models obtained cannot give a good explanation of the variants, and the lack of systemic thinking makes the improvement process of CP very tedious. Based on the analysis, finally, four key trends are identified: (1) further analysis of the variants, (2) integrated process management, (3) customization and (4) self-learning improvement of the clinical pathway.**

Keywords—clinical pathway; process mining; design and optimization

I. INTRODUCTION

Clinical pathway (CP) is a structured, multidisciplinary, patient care plan in which diagnostic and therapeutic interventions performed by physicians, nurses, and other staff for a particular diagnosis or procedure are sequenced on a timeline [1]. It is vital for improving the quality of healthcare delivery [2]. Clinical pathway is developed as a standard way to manage medical activities since the 1980's. It can have a positive impact on reducing the length of stay in hospital and costs of patient care. Other benefits identified include improved patient outcomes [3, 4], reduced complications [5], increased patient satisfaction with the service [6, 7], improved communication between doctors and nurses [8], increased participation of patient in patient treatment procedure [9]. Additionally, some research in information technology supported healthcare processes suggest that the IT supported CP can improve medical staff satisfaction [10] and reduce the time that health staff spent carrying out paperwork [11]. However, there are still many challenges lie ahead. Nowadays most of the clinical pathways designed with the traditional method are static and non-adaptive [12]. And more importantly, clinical pathways are defined by using big documents created in complex iterative processes, and those documents are affected by subjectivity of the creation members. Therefore, the methods using data mining and machine learning technologies to analyze clinical pathways based on associated event logs are receiving increasing attentions. These techniques, also known as process mining [13] can not only help the clinical pathways designers to discover the sequence of activities, but also provide execution information for analyzing variances and correcting design errors. They are more objective and effective techniques for achieving continuous clinical pathways evaluation [14, 15]. The application of process mining in healthcare is a relatively unexplored field. In this review, 37 studies were analyzed (shown in Table 1). They are surveyed in three aspects: process discovery, variants analysis and control, continuous evaluation and improvement [16].

TABLE I. STUDIES FROM THE PERIOD 2004-2013

	Authors	*Num.*
Process discovery	Hwang et al. (2004); Catley et al. (2008); Mans et al. (2008a); Blum et al. (2008); Lang et al. (2008); Mans et al. (2008b); Yang et al. (2009); Klimov et al. (2010); Mans et al. (2012); Yoo et al. (2012); Iwata et al. (2012); Álvarez et al. (2013); Jacob et al. (2013); Jindal & Roth (2013); Iwata (2012); Huang et al. (2012); Tsumoto et al. (2012); Fernandez-Llatas et al. (2013); Huang et al. (2013)	19
Variants analysis and control	Yang & Hwang (2006); Tsumoto et al. (2007); Gupta (2007); Quaglini (2009); Klundert (2010); Quaglini (2010); Poelmans et al. (2010); Díaz et al. (2012); Rebuge & Ferreira (2012); Du et al. (2012); Bouarfa & Dankelman (2012); Jacob & Ramani (2012); Eriksson et al. (2013)	13
Evaluation and improvement	Gunther et al.(2008); Garg et al.(2009); Zhou and Piramuthu (2009); Zhou & Piramuthu(2010); Fernandez-Llatas et al. (2011)	5

This research is funded by the Natural Science Foundation of China (NSFC) (Grant ID: 71090404 and 71072026) and sponsored by Research Center for Healthcare Management, SEM, Tsinghua University.

978-1-4799-3134-7/14/$31.00 ©2014 IEEE

II. Process Discovery For CP Design

Process discovery is to discover the process model by inferring the ordering relations between various tasks in the event log.

Mans et al. [17] applied process mining to discover how stroke patients are treated in different hospitals. The ProM framework was used along with the Heuristic Miner to gain insights about the control-flow perspective of the process. Different practices that are used to treat similar patients were discovered, together with unexpected behavior as well. The discovered process model was converted to a Petri net. The performance of the process was then analyzed by projecting performance indicators onto the Petri net. In further work [18], process mining is used to analyze the careflow of gynecological oncology patients. The control-flow, the organizational, and the performance perspectives of the process were analyzed. To discover the control-flow perspective, the Heuristic Miner was used first, which resulted in a spaghetti model that was not useful for analysis. The authors explain this difficulty based on the complex and unstructured nature of healthcare processes. Trace Clustering and the Fuzzy Miner were then used to separate the regular behavior from the infrequent one, and more understandable process models were discovered for each cluster. Nevertheless the traditional process mining approaches have problems dealing with unstructured processes that can be found in a medical environment. It is necessary to develop new techniques and extend existing techniques. Lang et al. [15] reached the same conclusions in their research. They did a detailed hands-on evaluation and analysis of 7 process-mining approaches, and concluded that none of the discussed approaches is able to meet all major challenges of mining clinical processes (e.g. noise, incompleteness, multiple occurrences of activities, or richness of process types and variants). Nevertheless, the concept of process mining carries great potential in helping to understand clinical workflows and their variations. The innovation of the mining approaches is important.

Huang et al. [19] developed a polynomial-time algorithm to build a concise and comprehensive summary that describes the entire structure of a clinical pathway, while revealing essential/critical medical behaviors. Given an input event log, the approach summarizes the pathway by segmenting the observed time period of the pathway into continuous and overlapping time intervals, and discovering frequent medical behavior patterns in each specific time interval from the log. But this approach only gives suggestions for CP deployment. It cannot help the designers to create a clinical pathway completely. Iwata et al. [20] proposed a temporal data mining method to create clinical pathway. This method extracted histories of nursing events from hospital information system, and the events are classified into several groups by using clustering and multidimensional scaling method. Feature selection is applied to the data and important features for classification are extracted. They obtained two major groups in the nursing events: ones were indispensable to the treatment, and the others were specific to the status of patients. Huang et al. [21] proposed a novel approach using closed clinical pathway pattern's sequence mining algorithm. This approach mines the clinical activity sequences at first, and then mines chronicles on the sequences to generate closed clinical pathway patterns, finally finds the closed clinical pathway patterns. As compared to the general sequence pattern mining algorithm, this approach consumes less processing time, generates quite a smaller number of clinical pathway patterns, and has a linear scalability in terms of execution time against the increasing size of data sets. Even though this approach is efficient for mining precise and complete set of regular medical behaviors in clinical pathways, there are still a number of infrequent behaviors that are missing in the discovered patterns. Note that these infrequent behaviors may represent interesting variants in clinical pathways, and thus need to be discovered and analyzed. It is also a challenge to develop an effective algorithm to identify and detect these variants. Moreover, the approaches above are usually based on the Event-based approach. They only pay attention to the information of the action name and the starting time, but forget the action results. Clinical pathways usually base their decisions on the results of activities. So the traditional process mining approaches should be improved when applied to CP design and optimization. Fernandez-Llatas et al. [22] combined a pattern recognition algorithm with the classical process mining framework, developed a Parallel Activity-based Log Inference Algorithm (PALIA). This algorithm require samples with the actions, the starting and the finish time of those actions as well as the actions' indicators, thereby achieving inference of medical process according to the activities' results. But the processing time is increased, because it is required the correct inference of not only the pattern flow, but also the indicator that fires that transition.

Therefore, it is a novel application of process mining in healthcare; however the diversity of medical behaviors and the complexity of chronicle information among medical behaviors in clinical pathways are far higher than that of common business processes. Traditional process mining techniques have many problems and challenges when used for mining clinical pathway patterns. Furthermore, the high dynamic and uncertain medical environment makes it much more difficult to design the clinical pathways. The clinical pathways designed using most recent approaches cannot make adjustment effectively when the external environment changes. In other words, it is non-adaptive.

III. Variants Analysis And Control

To analyze the variants of the clinical pathway, the researchers mainly use the conformance of process mining to check if observed behavior in the event log conforms to a given model. Recently, the related studies have achieved some results.

Yang and Hwang [14] proposed a data-mining framework to facilitate automatic and systematic construction of detection model. This framework need identify the characteristics of some specific variants, and then find out the CPs that these variants exist in. However this detection model is constructed for some specific variants; it is strongly affected by noise on the data. Gupta [23] evaluated the capabilities of the Heuristic Miner and also the DWS Algorithm to analyze the careflows.

The Heuristic Miner produced inaccurate and confusing models, and it was unable to identify process variants. The clustering approach of the DWS Algorithm was able to discover some behavioral patterns; however, the discriminants rules were hard to understand. None of them was considered to be useful to gain insight about exceptional medical cases or about variants of careflows. To handle this problem, the author introduced the Association Rule Miner (ARM) plug-in, which aims at discovering association rules and frequent item sets in the event log, thereby mining the process variants. But the process model discovered by this approach also contained some unclear AND/XOR join/splits and missed activities. Du et al. [24] developed a rule extraction approach based on improved hybrid particle swarm optimization algorithm (PSO) to discovery the previously unknown and potentially complicated nonlinear relationship between key parameters and variances handling measures of CP. But the classification accuracy and the efficiency of the approach should be further improved. Poelmans et al. [25] proposed a new approach based on Hidden Markov Models and Formal Concept Analysis (FCA) to discover process inefficiencies, exceptions and variations immediately and to search for the root causes of inefficiencies. But mishandling of the noise data will significantly affect the accuracy of the model. Rebuge and Ferreira [26] proposed a methodology for the application of process mining techniques that leads to the identification of regular behavior, process variants, and exceptional medical cases. But this method didn't provide an indication for number of clusters and a quality measurement for the results. Bouarfa and Dankelman [27] proposed to use tree-guided multiple sequence alignment to derive consensus workflow from multiple surgical activity logs, and use a global pair-wise sequence alignment algorithm to detect variants. However they cannot give a good explanation for these variants.

There are still many technical challenges for the identification and analysis of the CP variants. According to the different research purposes, the variants can be divided into many types. Recently, the identification and analysis methods always lack pertinence, pay scant attention to the sources and influences of the variants. It will lead to the low detection efficiency. The variants should be analyzed from a systemic perspective, and the influences on the adjustment of the clinical pathways should be considered.

IV. Continuous Evaluation And Improvement

Incremental mining techniques are always used for business process reengineering and continuous improvement, and related studies have achieved some results [28, 29]. Incremental mining is to analyze and use the new data that continuously added to the database to extend or redesign the recent process model, thereby guaranteeing practicality of the model. Along with the unceasing development and improvement of process mining techniques, many researchers focus on the process discovery and variants analysis, but very few studies have noted the continuous evaluation and improvement of the clinical pathway.

Garg et al. [30] presented a non-homogeneous Markov model to efficiently extract the frequent patterns. This method can also identify pathways having high expected cost and high readmission rate; these pathways require more attention to ensure optimum resource allocation. It can help the medical institution to evaluate the effectiveness of the improvement policy. However the information it can provide is limited, the causes of high expected cost and readmission rate need to be further analyzed. And the deficiency of variants identification ability has an effect on the accuracy of the evaluation result. Fernandez-Llatas et al. [31] developed a tool able to compare the designed Clinical Pathways with the real implantation cases in order to detect their differences. This tool can be used for continuous evaluation of the clinical pathways. However, it cannot explain the causes of the variants in detail; therefore it is not a self-learning system. The medical staffs have to handle these variants themselves subjectively. Zhou and Piramuthu [32] proposed an adaptive learning framework based on an RFID instance-level tracing/tracking system to analyze and evaluate the medical processes. It generates appreciable improvement both in terms of efficiency and service quality. However the authors didn't consider the trigger condition of the self-adaptive adjustment. And the complexity of computing makes it difficult to analyze the improvement policy.

In the recent studies, the evaluation and improvement are regarded as the redesign of the clinical pathways. However the root causes that triggered the improvement are always ignored, leading to too much work in the improvement process. It will severely affect the efficiency and effectiveness of the continuous improvement. The continuous evaluation and improvement should not be an isolated process. It needs the support of variants analysis that can identify the trigger causes to ensure the effectiveness and pertinence of improvement.

V. Future Directions

Process mining carries great potential in medical process management. It can be an objective way of analyzing clinical pathways as it is not biased by perceptions or normative behaviors. Note that the medical process is more complex than the common business process. The co-ordination of the hospital departments is very important to the implementation of the clinical pathways. Accordingly, besides the innovation of the process mining techniques, the systemic management of the clinical pathway should be an important direction of the future research. To this end, the key topics and problems may be summarized as follows:

A. Variants Identification and Analysis

Although there are a lot of studies in this field, few of them conducted from a systemic perspective considering the relationship between the variants and the handling measures. In a medical process, there exist some key variants that may significantly affect the design and implementation of the clinical pathway. Some of them should be avoided, and some may be used to improve the processes. Process mining techniques can help us to identify and classify them according to their different influences on the clinical pathway. The classification of the variants is also the basis for adaptive adjustment and redesign of the clinical pathway (as shown in Fig.1). The crucial problems involve analyzing the sources and key characteristics of the variants, determining the influence of the variants, identifying the variants that may trigger the adjustment or improvement of the clinical pathway.

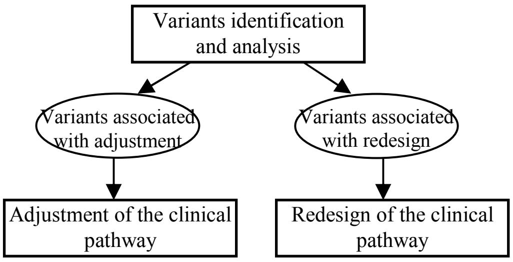

Fig. 1. The role of variants identification and analysis

B. Customization of Clinical Pathway

The change of external environment will affect the execution of the clinical pathway (e.g. the change of patient condition, equipment condition, and medical staff condition). Current application of clinical pathway is predefined, static and non-customized. However a particular patient should follow a special clinical pathway instead of a standard one, depending on his concrete health situation and hospital's resource. Consequently, the adaptability of the clinical pathway should be improved. In other words, the clinical pathway should adjust to changing situations. There will be some essential problems to solve, such as identifying the trigger conditions of the CP adjustment, determining the relationship between the adjustments and the performances of the clinical pathway, developing criteria for the adjustment evaluation.

C. Self-learning Improvement of Clinical Pathway

Along with the emergence of new medicine and treatment techniques, the current clinical pathway should be improved. It is an ongoing process. Note that the improvement of the clinical pathway is not just a redesign process. It needs to achieve the self-learning improvement through establishing automatic adjustment mechanism. The mechanism should be obtained from the historical data, and it is always related to some typical variants. Some interesting problems are raised. Which kind of the variants will trigger the improvement process? What are the characteristics of these variants? How to improve according to the variants analysis? How to verify the effectiveness of the improvement? The evaluation of the current pathways is also important. An appropriate evaluation system should be established to guide the improvement. Data mining techniques may provide a way to solve these problems.

D. Integrated Medical Process Management

Clinical pathway is an important tool to improve the service quality and process efficiency. However, the local optimization is not enough. Integrated medical process management should be considered. It pays attention to the whole life cycle of the disease, including activities of prevention, pre-hospital care, hospital treatment and rehabilitation (shown in Fig.2). Under this perspective, clinical pathway is a part of the whole process. Although the process discovery and variants analysis of the clinical pathway are also applicable to the whole process, there are some problems should be noted. The medical process of a specific disease may contain several similar clinical pathways. The rules for clustering of the similar clinical pathways should be determined.

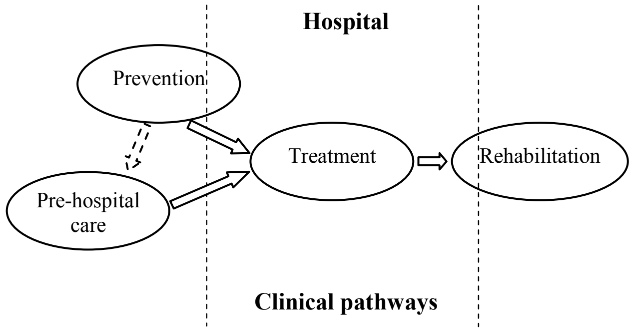

Fig. 2. Integrated medical process management

VI. Conclusion

In this paper, the studies on process mining techniques for clinical pathways were reviewed. Three research aspects were explored, respectively are process discovery for clinical pathways design, variants analysis and control, continuous evaluation and improvement. Although the process mining techniques had successfully used in the field of business process management, it was found that, in a medical environment, there still many challenges lie ahead. The medical processes are always dynamic and unstructured. The efficiency of the mining algorithm is strongly affected by the noise, incompleteness, multiple occurrences of activities, and richness of process types and variants. And most of the methods proposed in the studies only pay attention to the information of the action name and the starting time, but not the action results. It always makes the CP non-adaptive. Nevertheless, the concept of process mining carries great potential in helping to understand clinical workflows and their variations. To meet the challenges of the medical environment, the future research should consider more about the dynamic and unstructured nature of the process associated with the conditions of the patients and the variants of the CP. To this end, four key trends are identified. The further analysis of the variants is the basis for adaptive adjustment and improvement. The customization of the CP is an important way to solve the problem of dynamic conditions. And the integrated process management can help the medical institutions to improve the service quality and process efficiency more comprehensively.

Acknowledgment

This research was supported by NSFC projects (71090404, 71072026) and sponsored by Research Center for Healthcare Management, SEM, Tsinghua University.

References

[1] Ireson C. L., "Critical pathways: effectiveness in achieving patient outcomes", J. Nurs. Adm, Vol. 27, No. 7, pp. 16–23,1997.

[2] Berg A. O., Atkins D., Tierney W., “Clinical practice guidelines in practice and education”, Journal of General Internal Medicine, Vol. 12, No. 2, pp. S25-S33, 1997.

[3] Ogilvie-Harris D. J.,Botsford D.J., R. W. Hawker R. W., “Elderly patients with hip fractures: Improved outcome with the use of care maps with high-quality medical and nursing protocols”, J. Orthop. Trauma, Vol. 7, No. 5, pp. 428–437, 1993.

[4] Panella M., Marchisio S., Barbieri A., Di Stanislao F., “A cluster randomized trial to assess the impact of clinical pathways for patients with stroke: rationale and design of the Clinical Pathways for Effective and Appropriate Care Study”, BMC Health Service Research, Vol. 8, 2008.

[5] Preston S. R., Markar S. R., Baker C. R., Soon Y., Singh S., Low D. E., “Impact of a multidisciplinary standardized clinical pathway on perioperative outcomes in patients with oesophageal cancer”, British Journal of Surgery, Vol. 100, No. 1, pp. 105-112,2013.

[6] Stead L., Arthur C., Cleary A., “Do multidisciplinary pathways of care affect patient satisfaction?” Health Care Risk Rep, pp. 13–15,1995.

[7] van Dam P. A., Verheyden G., Sugihara A. et al., “A dynamic clinical pathway for the treatment of patients with early breast cancer is a tool for better cancer care: implementation and prospective analysis between 2002-2010”, World Journal of Surgical Oncology, Vol. 11, 2013.

[8] Mosher C., Cronk P., Kidd A., et al., “Upgrading practice with critical pathways”, Amer. J. Nurs., Vol. 92, No. 1, pp. 41–44,1992.

[9] Williams J. G., Roberts R., Rigby M. J., “Integrated patient records: Another move towards quality for patients?” Qual. Health Care, Vol. 2, No. 2, pp. 73–74, 1993.

[10] Schuld J., Schafer T., Nickel S., Jacob P., Schilling M. K., Richter S., “Impact of IT-supported clinical pathways on medical staff satisfaction. A prospective longitudinal cohort study”, International Journal of Medical Informatics, Vol. 80, No. 3, pp. 151-156, 2011.

[11] Dimitrios S., Pantelis M., and Kalliopi S., “Comparing the effects of eccentric training with eccentric training and static stretching exercises in the treatment of patellar tendinopathy. A controlled clinical trial”, Clinical Rehabilitation, Vol. 26, No. 5, pp. 423-430, 2012.

[12] Yang H. Q., Li S. Z., Zhao J. P., “Study of clinical pathway based on organizational semiotics methods”, Computer Engineering and Design, Vol. 30, No. 13, pp. 3189-3192, 2009.

[13] van der Aalst. W. M. P., Deseland J. A Oberwcis., Business process management: model, techniques, and empirical studies. Volume 1806 of Lecture Notes in Computer Science. Berlin: Springer-Verlag, 2000.

[14] Yang W. S., Hwang S. Y., “A process-mining framework for the detection of healthcare fraud and abuse”, Expert System With Applications, Vol. 31, No. 1, pp. 56-68, 2006.

[15] Lang M., Burkle T., Laumann S. Prokosch H. U., Process mining for clinical workflows: challenges and current limitations. 21st International Congress of the European Federation for Medical Informatic (MIE2008)

[16] Mans R. S., van der Aalst W. M. P., Vanwersch R. J. B., Moleman A. J., Process mining in healthcare: data challenges when answering frequently posed questions. ProHealth 2012/KR4HC 2012, Tallinn, Estonia, September 2012.

[17] Mans R. S., Schonenberg M. H., Song A., van der Aalst. W. M. P., Bakker P. J. M., “Application of process mining in healthcare - A case study in a Dutch hospital”, Biomedical Engineering Systems and Technologies, Vol. 25, pp. 425-438, 2008.

[18] Mans R. S., Schonenberg M. H., Leonardi G. et al. Process mining techniques: an application to stroke care. 21st International Congress of the European-Federation-for-Medical-Informatic (MIE2008)

[19] Huang Z. X., Lu X. D., Duan H. L., Fan W., “Summarizing clinical pathways from event logs”, Journal of Biomedical Informatics, Vol. 46, No. 1, pp. 111-127, 2013.

[20] Iwata H., Hirano S. and Tsumoto S., Maintenance of domain knowledge for nursing care using data in hospital information system. The 20th International Symposium on Methodologies for Intelligent Systems - ISMIS 2012.

[21] Huang Z. X., Lu X. D. and Duan H. L., “On mining clinical pathway patterns from medical behaviors”, Artificial Intelligence in Medicine, Vol. 56, No. 1, pp. 35-50, 2012.

[22] Fernandez-Llatas C., Meneu T., Benedi J. M., Traver V., Activity-based process mining for clinical pathways computer aided design. 32nd Annual International Conference of the IEEE EMBS 2010.

[23] Gupta S. Workflow and process mining in healthcare. Master's Thesis, Technische Universiteit Eindhoven, 2007.

[24] Du G., Jiang Z. B., Diao X. D., Yao Y., “Knowledge extraction algorithm for variances handling of CP using integrated hybrid genetic double multi-group cooperative PSO and DPSO”, Journal of Medical Systems, Vol. 36, No. 2, pp. 979-994, 2012.

[25] Poelmans J., Dedene G., Verheyden G. et al. Combining business process and data discovery techniques for analyzing and improving integrated care pathways. 10th Industrial Conference on Data Mining, Berlin, Germany.

[26] Rebuge A., Ferreira D. R., “Business process analysis in healthcare environments: A methodology based on process mining”, Information Systems, Vol. 37, No. 2, pp. 99-116, 2012.

[27] Bouarfa L., Dankelman J., “Workflow mining and outlier detection from clinical activity logs”, Journal of Biomedical Informatics, Vol. 45, No. 6, pp. 1185-1190, 2012.

[28] Gaaloul W., Gaaloul K., Bhiri S., Haller A., Hauswirth M., “Log-based transactional workflow mining”, Distributed and Parallel Databases, Vol. 25, No. 2, pp. 193-240, 2009.

[29] Rozinat A., de Jong I S M., Gunther C. W., van der Aalst W. M. P., “Process mining applied to the test process of wafer scanners in ASML”, IEEE Transactions on Systems Man and Cybernetics Part C – Applications and Reviews, Vol.39, No. 4, pp. 474-479, 2009.

[30] Garg L., McClean S., Meenan B., “Non-homogeneous Markov models for sequential pattern mining of healthcare data”, IMA Journal of Management Mathematics, Vol. 20, No. 4, pp. 327-344, 2009.

[31] Fernandez-Llatas C., Meneu T., Benedi J. M., Traver V., Continuous clinical pathways evaluation by using automatic learning algorithms. Proceedings of the International Conference on Health Informatics, 4th International Conference on Health Informatics (HEALTHINF 2011), Rome, Italy.

[32] Zhou W. and Piramuthu S., “Framework, strategy and evaluation of health care processes with RFID”, Decision Support Systems, Vol. 50, No.1, 2010, pp. 222-233, 2010.