Data mining of hospital characteristics in online publication of medical quality information*

Victor B. Kreng, Shao-Wei Yang#

Department of Industrial and Information Management, National Cheng Kung University, Tainan, Taiwan; #Corresponding Author: <u>r3893112@mail.ncku.edu.tw</u>

Received 7 February 2013; revised 30 March 2013; accepted 1 May 2013

Copyright © 2013 Victor B. Kreng, Shao-Wei Yang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT

Information disclosure can reduce information asymmetry between health care providers and patients, thus improving both patient safety and medical quality. The National Bureau of Health Insurance (NBHI) in Taiwan currently publishes health-related information online in order to enhance service efficiency and enable the public to monitor the country's medical system. A data mining technique, classification and regression tree (CART), is used in this work to investigate online public quality information to compare the characteristics of hospital. The hospital quality indicators and characteristics data are available on the websites of the NBHI

(http://www.nhi.gov.tw/AmountInfoWeb/Index.as
px) and the Department of Health

(http://www.doh.gov.tw/). The full classification and regression tree presented in this work, grown using the hospitals' quality medical indicators and characteristic values, classifies all hospitals into seven groups. The rate of stays longer than 30 days, which is the dependent variable in this study, is most influenced by the number of medical staff. This reflects the fact that the fewer medical staffs that are employed, the smaller the hospital is, and patients who are likely to have longer stays tend to go to the medium or large hospitals. Policy makers should work to decrease or eliminate persistent healthcare disparities among different socioeconomic groups and offer more online healthrelated services to reduce information asymmetry between

health care providers and patients.

Keywords: Hospital Characteristics; Data Mining; Classification and Regression Tree; Information Disclosure

1. INTRODUCTION

Advances in information, telecommunication, and network technologies have led to the emergence of a revolutionary new paradigm for health care that some refer to as e-health. By reducing information asymmetry between health care providers and patients, information disclosure can improve both patient safety and medical quality. With more disclosure, patients are capable of searching for more appropriate health-related knowledge, while health care providers are encouraged to provide higher quality services in order to attract patients [1-3].

To balance information asymmetry is pressured to health care providers in hope that public awareness and an informed consumer/patient will indirectly lead to improvements in quality of care [4,5]. For example, if information asymmetry exists with regard to prices, then this may cause increased prices for health services, because information holders, *i.e.* the health care providers, can charge monopoly prices [6]. In order to reduce information asymmetry, the information disclosed should also include unobservable quality measures in health care services, such as hospital mortality rates or the extent to which appropriate care is provided.

In the Health Maintenance Organization (HMO) market in the US, the quality scores of health care service providers have been disclosed since 1996 as part of the Health Plan Employer Data and Information Set (HEDIS). The demand for quality information came from HMOs themselves, who wished to demonstrate

^{*}Conflict of interest: The authors declare that they have no conflict of interest.

Funding source: The study was self-funded by the authors and their institution.

their quality improvement efforts [7]. Since HMOs with low quality scores are less likely to attract customers, such companies choose not to disclosure such information, especially if nondisclosure carries little stigma. This implies that voluntary disclosure of quality data, the national mechanism for HMO quality oversight in the US is failing to meet its stated goals to improve consumer decision making, provided incentives to raise quality, and increase public accountability [8]. However, Jung (2010) finds positive effects of disclosure on HMO quality, supporting the view that the public release of quality information may lead to improvements in quality [9]. High-quality care may be offered in markets with consumers who have greater willingness to pay for quality, and high-quality plans, which benefit from data disclosure, tend voluntarily to release quality information to the public.

In 1995, the National Health Insurance (NHI) program was introduced in Taiwan, offering a comprehensive, unified, and universal health care system. The NHI program is a single-payer one that is managed by the Bureau of National Health Insurance (BNHI), a government agency, to offer comprehensive benefit coverage. To raise the quality of medical care services. mandatory publication of service quality information was adopted in 2007. The Department of Health (DOH) is representatives for assisting in the medical care quality indicators and choices of relative information, as well as establishing cooperative planning information with BNHI. All the relevant medical care quality indicators were selected by the NHI Committee on Quality of Medical Care Services (CMCQ), composed of clinical medicine, medical management, information, law, and health education experts, as well as representatives from consumers, patients, the media, and various associations.

The quality assurance programs agreed with the medical sectors establishing health care quality indicators and posting quality information on the BNHI website as a reference for medical institutions to help them continue improving the quality of their care. The BNHI is committed to making health-related information more open and transparent to improve service efficiency and enable the public to monitor the country's medical system, with 73 quality indicators being posted on the agency's website

(http://www.nhi.gov.tw/AmountInfoWeb/Index.aspx) having received 3,497,611 hits as of the end of March 2012.

In practical terms, the timely availability of information is required for effective public health policy making and decision making [10]. However, online publication of quality information only reports the

reference statistics, and not the meanings that they contain. Data mining is the analysis of data sets to find hidden relationships and extract useful patterns and rules. Data mining in medicine is most often used for building classification models for diagnosis, prognosis or treatment planning [11-14]. Healthcare process are typically data rich, and thus many patterns can be discovered by the use of different algorithms [15-18], following techniques derived from machine learning, artificial intelligence, and statistics. The effectiveness of data mining has been proven in improving marketing campaigns, detecting fraud, and predicting diseases based on medical records [17,19,20].

As noted above, technological advances in information and communication technologies, the widespread use of the Internet has a number of implications for medical practice [21-24]. The use of patient population characteristics as surrogates for the characteristics of a particular patient have been widely reported in the medical literature [20,25-27]. However, there are few studies exploring the providers' characteristics when health care information is disclosed [28,29]. This study thus uses data mining to investigate the differences in quality information in order to compare hospitals. The aim of this study is thus to identify the conditions that facilitate voluntary disclosure of quality information, and to further develop policies to encourage this, so that the characteristics of a hospital, such as its human resources, can be considered by patients.

2. METHODOLOGY AND DATA

2.1. Data Mining

Over the last decade, there has been widespread use of medical information systems and an explosive growth of medical databases. Data mining is defined as the nontrivial extraction of implicit, previously unknown and potentially useful information from data [30]. To aid healthcare management, data mining applications can be developed to better identify and track chronic disease states and high-risk patients, design appropriate interventions, and reduce the number of hospital admissions and claims.

In general, data mining in medicine is most often used for building classification models, these being used for diagnosis, prognosis or treatment planning. In various types of data mining methods, the classification and regression tree (CART) technique, which uses a top-down greedy approach to tree construction, is used to produce accurate predictions or classifications based on a few logical if-then conditions. The main difference between decision trees and regression trees

is that decision tree construction involves classification into a finite set of discrete classes, whereas in regression tree learning the decision variable is continuous and the leaves of the tree either consist of a prediction with a numeric value or a linear combination of variables.

2.2. CART

CART is a nonparametric statistical procedure that identifies mutually exclusive and exhaustive subgroups of a population whose members share common characteristics that influence the dependent variable of interest [11]. Either continuous or categorical variables can be taken as input in CART, and no distributional hypothesis is required for these variables. CART has been applied to the problem of mining a diabetic data warehouse composed of a complex relational database with time series and sequencing information [17,31-33].

Based on the fact that tree methods are nonparametric and nonlinear, the simplicity of the results that CART provides is useful not only for the rapid classification of new observations, but can also be used to explain why observations are classified or predicted in a particular manner. The final results of using a binary tree structure for classification can be summarized in a series of logical if-then conditions. Therefore, tree methods can often reveal simple relationships between just a few variables that could have easily gone unnoticed using other analytic techniques.

In general, CART analysis consists of four basic steps, described as follows:

1) Tree building. A tree is built using recursive splitting of nodes. Each resulting node is assigned a predicted class, based on the distribution of classes in the learning dataset which would occur in that node and the decision cost matrix. The assignment of a predicted class to each node occurs whether or not that node is subsequently split into child nodes.

- 2) Stopping the tree building. At this point a "maximal" tree has been produced, which probably greatly overfits the information contained within the learning dataset.
- 3) Pruning a tree. To create the sequence of simpler and simpler trees, through the cutting off of increasingly important nodes.
- 4) To select an optimal tree. The tree, which is selected from among the sequence of pruned trees, fits the information in the learning dataset, but does not overfit the information.

2.3. Data

Under health care system in Taiwan, hospitals must register details of their beds, medical staff and related operating item with the Department of Health (DOH). The data set, during the first quarter of 2011, in this study consists of hospital characteristics data and quality indicators of licensed health care facilities in the NBHI website

(http://www.nhi.gov.tw/AmountInfoWeb/Index.aspx). Data related to the number of beds, physicians, man-

power data and hospital characteristics is collected from DOH (http://www.doh.gov.tw/). The database was merged into a data set using hospital registered identification codes, and the descriptive statistics are listed in detail in **Table 1**.

A quality indicator, the rate of hospital stays longer than 30 days, is used as the dependent variable in the regression tree. There are seven independent variables: acute beds, chronic beds, physicians, nursing staff, pharmacists, medical staff and hospital type. The detail definitions are as follows:

Table 1. Descriptive statistics of different type hospitals.

Variable -	Specialist Hospital (n1 = 7)		Chronic Hospital (n2 = 7)		General Hospital (n3 = 146)		Hospital $(n4 = 227)$		Total (N = 387)		F	
	Mean	Standard error	Mean	Standard error	Mean	Standard error	Mean	Standard error	Mean	Standard error	Г	
Acute bed	41.0	27.6225	6.4	17.0084	402.3	433.7922	45.1	50.3228	179.1	320.1349	53.5865	***
Chronic bed	0.0	0.0000	79.9	43.0984	15.4	48.2646	3.5	22.6263	9.3	36.4545	13.2846	***
Physician	0.9	2.2678	1.7	4.5356	44.2	49.4724	2.6	6.3749	18.3	36.7874	32.8743	***
Pharmacist	16.3	13.9608	9.3	8.9762	141.3	205.4430	9.1	11.1568	59.1	141.5436	48.2323	***
Nursing staff	6.0	5.5377	3.9	5.3984	35.0	39.1121	4.2	4.5245	15.8	28.4525	49.3468	***
Medical staff	0.4	0.7868	2.3	1.9760	8.6	13.8683	0.8	1.1475	3.8	9.3311	57.3382	***
The rate of hospital stays longer than 30 days	0.0009	0.0019	0.0483	0.0568	0.0209	0.0651	0.0223	0.0892	0.0218	0.0795	0.4285	

Note: ***indicates p < 0.01.

Acute beds: the number of acute beds served by a hospital.

Chronic beds: the number of chronic beds served by a hospital.

Physician: the number of full-time physicians employed by a hospital.

Nursing staff: the number of full-time nursing staffs employed by a hospital.

Pharmacist: the number of full-time pharmacists employed by a hospital.

Medical staff: the number of full-time medical staff employed by a hospital.

Hospital type: different sizes of hospitals in Taiwan are classified into hospitals, specialist hospitals, chronic hospitals and general hospitals.

According to the related attributes and tasks, hospitals in Taiwan are classified into hospitals, specialist hospitals, chronic hospitals and general hospitals. The primary difference of general hospitals and hospitals is that the beds and the departments they offered. General hospitals must have over 100 beds and serve health care services including six departments at least, such as the departments of Medicine, Surgery, Obstetrics and Gynecology, Pediatric, Anesthesiology, and Radiology. The hospitals usually have less than 100 beds and serve one department or several specialist departments. Only hospital type is categorical, while the other variables are continuous.

3. RESULTS

The disclosure of quality information related to individual hospitals in the Taiwan NHI system is intended to identify the quality differences among hospitals and thus enable people to make informed choices to best meet their needs for health care services. Disclosure should not simply be focused on the perspective that patients should only receive disclosures and providers should only give them in an effort to promote patient safety.

To focus on the relationship between hospital characteristics and medical quality in the Taiwan NHI system, CART is used to classify different clusters with minimum internal variance, but with maximum variance between clusters, using a tree structure. The full tree, grown using the hospitals' quality indicators and characteristic values, contains seven the predictor variables and seven terminal nodes. The full tree is shown in **Figure 1**. This tree successfully classified 387 cases into seven groups, with significant differences among these (as shown in **Table 2**). Simple resubstitution classification rates can suffer considerable bias, and a more realistic assessment of the performance of this tree is to apply it to data other than that used in its construction.

4. DISCUSSION

The aim of public disclosure of medial quality data is not only to reveal information to patients, but also to stimulate quality improvement efforts in hospitals [3]. For example, the public release of performance data has been proposed as a mechanism to improve quality of care [34] by providing more transparency and greater accountability of health care providers [35]. In the US, voluntary national efforts to publicly report on hospital quality include pilot projects that have tested the use of a standardized instrument, the Hospital CAHPS Survey, to measure patient perspectives on hospital care [36].

Different from most previous studies concerning data mining application in health care services, which classifies patients' characteristics, this study focuses on the relationship between hospital characteristics and medical quality. In this study, CART is adopted to discover the hidden connections between hospital

Table 2. Descriptive statistics	f end nodes classified	by CART.
----------------------------------------	------------------------	----------

Variable	Group I (n = 84)		Group II (n = 179)		Group III (n = 2)		Group IV (n = 10)		Group V (n = 4)		Group VI (n = 6)		Group VII (n = 102)	
	Mean	Standard error	Mean	Standard error	Mean	Standard error	Mean	Standard error	Mean	Standard error	Mean	Standard error	Mean	Standard error
Acute bed	25.93	10.017	350.81	408.310	79.50	14.849	10.30	20.298	31.00	13.342	68.00	32.168	34.74	15.808
Chronic bed	0.60	5.455	16.65	49.895	0.00	0.000	55.60	47.376	0.00	0.000	0.00	0.000	0.20	1.393
Physician	4.57	2.191	120.93	190.492	11.50	0.707	6.50	2.173	3.75	1.500	12.00	0.000	6.58	2.463
Pharmacist	2.46	1.046	30.78	36.548	4.00	1.414	2.50	1.354	1.00	0.000	5.50	1.643	3.36	1.296
Nursing staff	20.60	12.606	433.41	537.187	70.00	2.828	29.00	13.157	21.00	12.780	69.50	20.226	32.58	16.541
Medical staff	2.54	1.312	76.26	85.715	56.00	18.385	9.10	3.381	6.50	1.000	10.83	3.545	9.08	3.894
The rate of hospital stays longer than 30 days	0.0039	0.0182	0.0162	0.0210	0.6947	0.1317	0.1017	0.1907	0.1500	0.3000	0.1158	0.2670	0.0150	0.0484

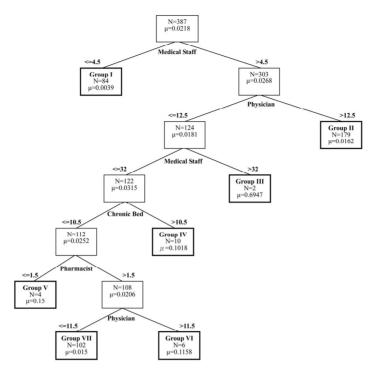


Figure 1. Classification and regression tree. Dependent variable: the rate of hospital stays longer than 30 days.

characteristics and medical quality indicators. The most important factor that influences the length of a hospital stay is the number of medical staff, such as medical laboratory scientists, physical therapists and occupational therapists. With regard to the services offered by a hospital, the physicians provide the direct health care services needed to improve patients' health, with medical staff providing supportive health care services. The number of medical staff employed is based on the organizational structure of a specific hospital, and reflects both the operating scale and services provided. For example, a small hospital with less than five physicians may only provide primary acute health care services, without a laboratory scientist or physical therapist. Therefore, patients with an expected length of stay over 30 days will prefer to seek health care services at medium or large hospitals, as these have more medical staff and better equipment to serve more health care services.

The hospital characteristics are classified as shown in **Table 3**, based on the association rules derived from CART. The results show that the number of medical staff is the most important factor to classify groups. Almost 22% of the hospitals, with less than or equal to four medical staff, are classified into group I. Such hospitals offer very limited services, and are common local district hospitals, offering primary health care services in rural areas. Middle and large-sized hospitals with more than 12 physicians, like Na-

Table 3. Association rules of hospital characteristics classified by CART.

		4 < 1				
Physician	Medical Staff	Chronic	Bed < 11	Chronic	Medical Staff >32	
·	≤4	Pharmacist = 1	harmacist Pharmacist Bed≥11			
≤11	Group I (n = 84)	Group V	Group VII (n = 102)	Group IV	Group III	
=12		(n=4)	Group VI (n = 6)	(n = 10)	(n=2)	
≥13						

tional Taiwan University Hospital, Taipei Veterans General Hospital, Tzu-Chi General Hospital, Chang-Gung Memorial Hospital and Mackay Memorial Hospital, are classified into group II, and these are able to offer more comprehensive health care services. Focusing on groups I and II, it is reasonable to assume that the similar length of stay data found in these hospitals is due to their size. In general, small hospitals with around 20 beds have only a single department, such as internal or family medicine, and fewer patients go to such hospitals for health care services. The results for group IV, like Tainan State Hospital Sin-Hua Branch and Wan-Hwa Hospital, reflect the fact that more chronic beds offer patients more opportunities to increase the length of their hospital stays. Besides, some small or chronic hospitals, such as group IV,

have a shortage of pharmacists. Most hospitals in group VII are small and middle-sized ones, only Chiu Hospital, You-Chang United Hospital, Yang-Ming Hospital (Taoyuan County) and En-Hua Hospital have more than 70 beds and the others don't.

Health policy makers have given considerable attention to the effect of information disclosure on medical quality, such as improving the performance of hospitals and physicians. However, from the results of this study, the difference of medical quality indicator results from the different scale of hospitals, including equipment and medical human source. The small hospitals have better performance in medical quality indicator than great ones, because the patients with serious illness usually go to big hospitals for adequate health care services. According to the online publication of medical quality information, the patients might misunderstand that small hospitals have better performance on quality indicator than big ones and ignore that the great hospitals have the ability to care for the patients with serious, chronic, and terminal illness. Hence, the selection of quality indicators for disclosure should reveal different clinical information, which is easy to read and understand for people.

The online publication of information on the quality of medical care services offers patients a chance to reduce information asymmetries existed between patient and physician and to seek services. However, access to outside medical information is linked to a patient's socioeconomic status [25-27,37,38], in which can create a digital divide. Policy makers should thus target educational efforts to decrease or eliminate the persistent healthcare disparities among different socioeconomic groups. Educating and empowering ehealth care consumers through online information enables them to become active participants in their own health care, thus potentially resulting in higher satisfaction.

5. CONCLUSIONS

Unlike traditional studies, hospital characteristics, instead of patients, have been explored in this study to reveal the hidden issues behind quality care of health service providers. Similar hospital characteristics with nearly quality indicator have been classified into a group with a tree structure through data mining. Medical staff plays an important role to reflect different size of hospital, and reveal the department or services that hospital provided. Hence, the fewer medical staffs are employed, the smaller size hospital is, which presents that patients seek for health care service towards medium or large hospital with higher rate of stay length over 30 days.

Nowadays, online publication of medical quality information in NHI system of Taiwan help people identify the quality differences among individual hospitals and to make an informed choice to seek for health care service, which is helpful to reduce the information asymmetry existed in health care delivery and to improve health care quality. Health policy maker should enhance people among different socioeconomic groups to access medical information through internet.

REFERENCES

- [1] Gallagher, T.H. and Levinson, W. (2005) Harmful medical errors to patients—A time for professional action. *Archives of Internal Medicine*, **165**, 1819-1824. doi:10.1001/archinte.165.16.1819
- [2] Kitamura, T. (2005) Stress-reductive effects of information disclosure to medical and psychiatric patients. *Psychiatry and Clinical Neurosciences*, 59, 627-633. doi:10.1111/j.1440-1819.2005.01428.x
- [3] Marshall, M.N., Romano, P.S. and Davies, H.T. (2004) How do we maximize the impact of the public reporting of quality of care? *International Journal for Quality Health Care*, **16**, 57-63. doi:10.1093/intqhc/mzh013
- [4] Berwick, D.M. (2002) Public performance reports and the will for change. *JAMA*, 288, 1523-1524. doi:10.1001/jama.288.12.1523
- [5] Chassin, M.R. (2002) Achieving and sustaining improved quality: Lessons from New York State and cardiac surgery. Health Affair, 21, 40-51. doi:10.1377/hlthaff.21.4.40
- [6] De Fraja, G. (2000) Contracts for health care and asymmetric information. *Journal of Health Economics*, 19, 663-677. doi:10.1016/S0167-6296(00)00037-0
- [7] Jin, G.Z. (2005) Competition and disclosure incentives: An empirical study of HMOs. *Rand Journal of Economics*, 36, 93-112.
- [8] McCormick, D., Woolhandler, S., Wolfe, S.M. and Bor, D.H. (2002) Relationship between low quality-of-care scores and HMOs' subsequent public disclosure of quality-of-care scores. *The Journal of the American Medical Association*, 288, 1484-1490. doi:10.1001/jama.288.12.1484
- [9] Jung, K. (2010) Incentives for voluntary disclosure of quality information in HMO markets. *Journal of Risk and Insurance*, 77, 183-210. doi:10.1111/j.1539-6975.2009.01339.x
- [10] Joseph, T. (2005) E-health care information systems: An introduction for students and professionals. Jossey Bass, San Francisco.
- [11] Breault, J.L., Goodall, C.R. and Fos, P.J. (2002) Data mining a diabetic data warehouse. *Artificial Intelligence* in Medicine, 26, 37-54. doi:10.1016/S0933-3657(02)00051-9
- [12] Kaur, H. and Wasan, S.K. (2006) Empirical study on applications of data mining techniques in healthcare. *Journal of Computer Science*, 2, 194-200.

- doi:10.3844/jcssp.2006.194.200
- [13] Kononenko, I. (2001) Machine learning for medical diagnosis: History, state of the art and perspective. Artificial Intelligence in Medicine, 23, 89-109. doi:10.1016/S0933-3657(01)00077-X
- [14] Obenshain, M.K. (2004) Application of data mining techniques to healthcare data. *Infection Control and Hospital Epidemiology*, 25, 690-695. doi:10.1086/502460
- [15] Cho, S.B. and Won, H.H. (2007) Cancer classification using ensemble of neural networks with multiple significant gene subsets. *Applied Intelligence*, 26, 243-250. doi:10.1007/s10489-006-0020-4
- [16] Delen, D., Walker, G. and Kadam, A. (2005) Predicting breast cancer survivability: A comparison of three data mining methods. *Artificial Intelligence in Medicine*, 34, 113-127. doi:10.1016/j.artmed.2004.07.002
- [17] Garzotto, M., Beer, T.M., Hudson, R.G., Peters, L., Hsieh, Y.C., Barrera, E., Klein, T. and Mori, M. (2005) Improved detection of prostate cancer using classification and regression tree analysis. *Journal of Clinical Oncology*, 23, 4322-4329. doi:10.1200/JCO.2005.11.136
- [18] Harper, P.R. (2005) A review and comparison of classification algorithms for medical decision making. *Health Policy*, 71, 315-331. doi:10.1016/j.healthpol.2004.05.002
- [19] Bellazzi, R. and Zupan, B. (2008) Predictive data mining in clinical medicine: Current issues and guidelines. *International Journal of Medical Informatics*, 77, 81-97. doi:10.1016/j.ijmedinf.2006.11.006
- [20] Phillips-Wren, G., Sharkey, P. and Dy, S.M. (2008) Mining lung cancer patient data to assess healthcare resource utilization. *Expert Systems with Applications*, 35, 1611-1619. doi:10.1016/j.eswa.2007.08.076
- [21] Cline, R.J. and Haynes, K.M. (2001) Consumer health information seeking on the internet: The state of the art. Health Education Research, 16, 671-692. doi:10.1093/her/16.6.671
- [22] Dedding, C., van Doorn, R., Winkler, L. and Reis, R. (2011) How will e-health affect patient participation in the clinic? A review of e-health studies and the current evidence for changes in the relationship between medical professionals and patients. *Social Science & Medicine*, 72, 49-53. doi:10.1016/j.socscimed.2010.10.017
- [23] Detmer, W.M. and Shortliffe, E.H. (1997) Using the internet to improve knowledge diffusion in medicine. *Communications of the ACM*, 40, 101-108. doi:10.1145/257874.257897
- [24] Powell, J.A., Darvell, M. and Gray, J.A. (2003) The doctor, the patient and the world-wide web: How the internet is changing healthcare. *Journal of the Royal Society of Medicine*, 96, 74-76. doi:10.1258/jrsm.96.2.74
- [25] Bernheim, S.M., Ross, J.S., Krumholz, H.M. and Bradley, E.H. (2008) Influence of patients' socioeconomic status on clinical management decisions: A qualitative study. *The Annals of Family Medicine*, 6, 53-59. doi:10.1370/afm.749
- [26] Hansen, R.P., Olesen, F., Sorensen, H.T., Sokolowski, I. and Sondergaard, J. (2008) Socioeconomic patient characteristics predict delay in cancer diagnosis: A Danish co-

- hort study. *BMC Health Services Research*, **8**, 49. doi:10.1186/1472-6963-8-49
- [27] Willems, S., De Maesschalck, S., Deveugele, M., Derese, A. and De Maeseneer, J. (2005) Socio-economic status of the patient and doctor-patient communication: Does it make a difference? *Patient Education & Counseling*, 56, 139-146. doi:10.1016/j.pec.2004.02.011
- [28] Keeler, E.B., Rubenstein, L.V., Kahn, K.L., Draper, D., Harrison, E.R., Meginty, M.J., Rogers, W.H. and Brook, R.H. (1992) Hospital characteristics and quality of care. *Journal of American Medical Association*, 268, 1709-1714. doi:10.1001/jama.1992.03490130097037
- [29] Lehrman, W.G., Elliott, M.N., Goldstein, E., Beckett, M.K., Klein, D.J. and Giordano, L.A. (2010) Characteristics of hospitals demonstrating superior performance in patient experience and clinical process measures of care. *Medical Care Research and Review*, 67, 38-55. doi:10.1177/1077558709341323
- [30] Frawley, W.J., Piatetskyshapiro, G. and Matheus, C.J. (1992) Knowledge discovery in databases—An overview. *Ai Magazine*, 13, 57-70.
- [31] Fonarow, G.C., Adams Jr., K.F., Abraham, W.T., Yancy, C.W. and Boscardin, W.J. (2005) Adhere scientific advisory committee SG, investigators. Risk stratification for in-hospital mortality in acutely decompensated heart failure: Classification and regression tree analysis. *JAMA*, 293, 572-580. doi:10.1001/jama.293.5.572
- [32] Rovlias, A. and Kotsou, S. (2004) Classification and regression tree for prediction of outcome after severe head injury using simple clinical and laboratory variables. *Journal of Neurotrauma*, 21, 886-893. doi:10.1089/0897715041526249
- [33] Zlobec, I., Steele, R., Nigam, N. and Compton, C.C. (2005) A predictive model of rectal tumor response to preoperative radiotherapy using classification and regression tree methods. *Clinical Cancer Research*, 11, 5440-5443. doi:10.1158/1078-0432.CCR-04-2587
- [34] Berwick, D.M., James, B. and Coye, M.J. (2003) Connections between quality measurement and improvement. *Medical Care*, 41, 130-138. doi:10.1097/00005650-200301001-00004
- [35] Lansky, D. (2002) Improving quality through public disclosure of performance information. *Health Affairs*, 21, 52-62. doi:10.1377/hlthaff.21.4.52
- [36] Barr, J.K., Giannotti, T.E., Sofaer, S., Duquette, C.E., Waters, W.J. and Petrillo, M.K. (2006) Using public reports of patient satisfaction for hospital quality improvement. Health Services Research, 41, 663-682. doi:10.1111/j.1475-6773.2006.00508.x
- [37] Colledge, A., Car, J., Donnelly, A. and Majeed, A. (2008) Health information for patients: Time to look beyond patient information leaf lets. *Journal of the Royal Society of Medicine*, 101, 447-453. doi:10.1258/jrsm.2008.080149
- [38] Lopez, L., Green, A.R., Tan-McGrory, A., King, R. and Betancourt, J.R. (2011) Bridging the digital divide in health care: The role of health information technology in addressing racial and ethnic disparities. *Joint Commission Journal on Quality and Patient Safety*, 37, 437-445.