An Empirical Study of Hospital’s Outpatient Loyalty From a Medical Center in Taiwan

Shu-Hui Chao¹, Mu-Kuan Chen¹, and Hsin-Hung Wu²³⁴

Abstract
In a highly competitive medical industry, hospitals can continue to create medical values and competitive advantages using data mining technologies to identify patients’ needs and provide the medical services needed by various patients. This research focuses on the outpatients in a medical center in Taiwan and adopts recency, frequency, and monetary (RFM) model, self-organizing maps, and K-means method to construct a set of data exploration procedures so that the hospital can use the reference to deal with the related patient management issues, where R, F, and M measure the RFM spent for each outpatient in Year 2016. The results show that 321,908 outpatients can be classified into 12 groups and further categorized into loyal outpatients, new outpatients, and lost outpatients. The similarities and differences among groups can be further analyzed to allow hospital management to provide differentiation strategies to its patients. That is, with the model illustrated in this study, the hospital can establish a better and long-term relationship with its patients by increasing patient loyalty.

Keywords
patient loyalty, outpatient, RFM model, cluster analysis, self-organizing maps, K-means method

Introduction
Porter and Teisberg (2006) stated that the traditional competition in health care industry does not focus on increasing values for patients. Cambria et al. (2012) pointed out that everything we do that does not provide benefits to patients and their families is a waste. In contrast, health care organizations need to use a value-based competition by managing a patient’s personal health, the provision of a health plan based on value added, acting responsibility, and building a long-term relationship. That is, the competition among health care organizations needs to focus on value for patients instead of minimization of short-term costs (Porter and Teisberg, 2006). In fact, value in health care is related to the health outcome per dollar of cost. Dafny and Lee (2016) emphasized that health care organizations need to make the following changes in this competitive market, including put patients first, create choices, stop rewarding volume, standardize methods to pay for value, and make outcomes transparent.

MacStravic (1987) summarized that hospitals must put patient loyalty in a top priority in their marketing strategies because loyal patients are sources of repeat businesses, potential users of new services, and positive spokespersons in a word-of-mouth advertising. MacStravic (1994) also pointed out that patient loyalty can be observed through the intention to return to the same provider, resistance to changing to another provider, and the intention to recommend one’s provider to others. In addition, loyal patients are more likely to be both motivated and knowledgeable enough to provide some parts of their medical care themselves, thereby saving costs and promoting profitability, become less sensitive based on the price of medical care solely, and reduce patient anxiety when medical care is needed.

Patients in Taiwan have a wide variety of selections in health care organizations which might result in a fierce competition between health care service providers (Shieh et al., 2010). Rundle-Thiele and Russell-Bennett (2010) pointed out that the success of health care organizations does not result from having good technical skills solely but

¹Changhua Christian Hospital, Changhua, Taiwan
²National Changhua University of Education, Changhua, Taiwan
³Asia University, Taichung City, Taiwan
⁴State University of Malang, Indonesia

Corresponding Author:
Hsin-Hung Wu, Department of Business Administration, National Changhua University of Education, No. 2 Shida Road, Changhua City, Taiwan 500.
Email: hhwu@cc.ncue.edu.tw
also from meeting patients’ needs and encouraging them to return to the same health care organizations, that is, behavioral loyalty. That is, patient loyalty is beyond patient satisfaction when health care service providers are in an intense competition (Ravichandran, 2015). To measure patient loyalty effectively, actual behavioral or attitudinal loyalty such as consistency of use or return to the same health care organizations would be a better indicator than expressions of willingness or intention (MacStravic, 1994; Ravichandran, 2015).

Due to the complexity and the rise of data in health care, the deployment of artificial intelligence (AI) in health care is promising and very welcome (Bartoletti, 2019). Davenport and Kalakota (2019) pointed out that AI will be increasingly applied in some key categories of applications including patient engagement and adherence. For instance, natural language processing such as speech recognition, text analysis, and the like has been applied in health care industry, and the usefulness of natural language processing systems can analyze unstructured clinical notes on patients, prepare reports (on some medical examinations), transcribe patient interactions, and conduct conventional AI to classifications of clinical documentation and published researches (Davenport and Kalakota, 2019). On the contrary, Menger et al. (2019) stated that cluster analysis, a data-driven approach, can be used to stratify patient subgroups, which is an important process to support clinical cares in hospitals.

RFM (recency, frequency, and monetary) model is based on the practice/foundation that recency (R), frequency (F), and monetary (M) are three important variables to summarize a customer’s purchase history (Zhang et al., 2015). Hughes (1996) and Wei et al. (2010) also emphasized that RFM model is a behavior-based model to analyze the behavior of a customer based on his or her past actual behaviors in the database. That is, these three variables can be used to observe customers’ attitudes toward the product, brand, benefit, or even loyalty from the database.

In the past, Lee (2012) used RFM model to analyze 14,072 inpatients for patient loyalty. Hosseini and Mohammadzadeh (2016) also applied RFM-based model to analyze 234,690 patients from emergency clinics from September 2008 to October 2012 for customer value analysis. That is, RFM model is practical to be used in medical industry for patient loyalty. However, no studies have been found to analyze outpatients’ patient loyalty based on RFM model along with cluster analysis. There is a need to establish an RFM-based model to analyze patient loyalty to provide better health care services and, therefore, to establish long-term relationships with outpatients.

No studies have been found to use RFM model to analyze outpatients’ value from a medical center particularly with more than 320,000 outpatients together with more than 1,000,000 visits in 2016. In reality, each hospital has its limited resources to take great care of patients and to meet individual’s needs. Thus, the philosophy of triage in the emergency room can be applied. That is, the patients who are critically ill or in poor health care conditions should be in a top priority for treatment and monitored relentlessly when a patient-centered medical care is applied.

Methods

This study uses a database of outpatients’ records from a medical center in Taiwan from January 1 to December 31, 2016. Each record from an outpatient includes the patient’s gender, date of visit, age, department (clinic) of visit, cost of visit, and zip code. The clinical trial approval certificate (new protocol) was approved by the Institutional Review Board of Changhua Christian Hospital in Changhua County, Taiwan with the protocol number of CCH IRB 1711103 with the waiver of documentation of informed consent. In order to use RFM model, four general guidelines are applied for data cleaning. First, exclude the patient’s records due to nonillness to the hospital such as preventive care or vaccination. Second, exclude the patients from other designated hospitals to the case hospital for an inspection. Third, if a patient visits the case hospital for more than once or for two or more departments on the same date, this study considers the patient only seeks the medical treatment once on that particular date. The medical expenses for more than one visit on the same date are added to become a total amount of medical expenditure on that date. Fourth, if a patient was born in 2016, his or her age is less than 1 year old. The age of the patient is assigned to be zero.

Following the above procedures for data cleaning, there were 321,908 outpatients with 1,683,065 visits in 2016. The definitions of R, F, and M variables in this study are as follows. Recency is defined as the number of days since the last visit from January 1, 2016, to December 31, 2016. Unlike the traditional definition that a smaller number of recency indicates a more recent visit, Wei et al. (2013) used a different philosophy to define recency. Based on the definition by Wei et al. (2013), January 1 is given by a number of 1, January 2 is assigned by a number of 2, and December 31 has the number of 366. That is, a higher recency value indicates a patient visits the case hospital more recent. Frequency is to count the number of visits for a patient in 2016. Finally, monetary is to accumulate the medical expenditures in 2016 for each patient.

To perform cluster analysis, both self-organizing maps (SOMs) and K-means methods are very popular methods, where SOM can also be viewed as a constrained version of K-means method because cluster centers in SOM tend to lie in a low-dimensional manifold in the feature or attribute space (Han & Kamber, 2006). To achieve a better cluster analysis performance, Kuo et al. (2002) proposed a
two-stage method by combining SOM and $K$-means method. Their results showed that the combination of SOM and $K$-means method was superior than the traditional approaches in cluster analysis. Tang et al. (2018) also reported that a combination of SOM and $K$-means method was a better approach to perform cluster analysis. Their study pointed out that the result of the final clusters employed by $K$-means method depended on the selection of the initial centroids. Thus, the preliminary results obtained by SOM was later used by $K$-means method for polymers classifications.

Based on the findings from Kuo et al. (2002) and Tang et al. (2018), this study uses the combination of SOM and $K$-means method for cluster analysis in this case study. To perform cluster analysis, IBM SPSS Modeler 14.1 is the software, and the “Kohonen node” (namely SOM) with default values is used. The input variables are R, F, and M. The Kohonen mode is set to “simple” for cluster analysis to generate the appropriate number of clusters. Later, $K$-means method is employed to cluster 321,908 outpatients based on the determined number of clusters by SOM.

## Results

Among 321,908 outpatients, the numbers of male and female outpatients are 138,595 and 183,313, representing 43% and 57% of total outpatients, respectively. The youngest and oldest outpatients for both male and female are 0 and 104, while the average ages of male and female outpatients are 42 and 45, respectively. The distributions of age groups are as follows: there are 56,080 outpatients whose ages are 17 years old or less, 80,097 outpatients from 18 to 39 years old, 118,622 outpatients from 40 to 64 years old, and 67,109 outpatients with 65 years old or more, consisting of 17%, 25%, 37%, and 21%, respectively. Nearly 97% of outpatients are from Changhua County with 250,205 (77.73%), Taichung City with 23,182 (7.20%), Nantou County with 20,160 (6.26%), and Yunlin County with 17,499 (5.44%). Changhua County is the location of this case medical center, while Taichung City, Nantou County, and Yunlin County are the adjacent counties/cities. Therefore, the outpatients’ visits are geographically related.

To transform 1,683,065 visits into R, F, and M variables, Table 1 summarizes the descriptive statistics of these variables. Specifically, the minimum, average, and maximum values of recency are 1, 239, and 366, respectively. The minimum, average, and maximum values of frequency are 1, 5, and 220, respectively. The minimum, average, and maximum total medical expenditures in 2016 are $18, $19,690, and $35,416,990 in terms of New Taiwan dollars, respectively.

Before the use of SOM, the values of R, F, and M are normalized such that each individual value ranges from zero to one. In contrast, $K$-means method uses Euclidean distance for cluster analysis such that data normalization is not necessarily required. When R, F, and M variables are the input variables, SOM suggests that the number of clusters among 321,908 outpatients is 12. Table 2 depicts the descriptive statistics of 12 clusters in terms of sample size, average numbers of R, F, and M, and the symbol(s) of R, F, and M greater than the averages of R, F, and M when $K$-means method is performed. Most of outpatients are in Clusters 1, 2, 3, 4, 5, 7, and 8. Ha and Park (1998) defined four types of customers in terms of R, F, and M variables. Vulnerable customers include $R \downarrow F \uparrow M \downarrow$ and $R \downarrow F \uparrow M \uparrow$. New customers have the characteristic of $R \uparrow F \downarrow M \downarrow$. Loyal customers are $R \uparrow F \uparrow M \uparrow$, and prospect customers are $R \uparrow F \downarrow M \uparrow$. Based on the philosophy of Ha and Park (1998), 12 clusters as shown in Table 2 can be further classified into three types of outpatients in terms of loyal, new, and lost outpatients depicted in Table 3.

The patients with the feature of $R \uparrow F \uparrow M \uparrow$ can be viewed as loyal outpatients with more recent visit, high frequency, and high medical expenditure as shown in Table 4. Patients with high medical expenses typically indicate that they are critically ill patients that need long-term treatment such as division of hematology—oncology (cancer outpatient chemotherapy) and division of nephrology (hemodialysis), and so on. The major characteristics of patients are as follows. Eighty percent of outpatients are 40 years old or more. More than 60% of the patients live in Changhua County. The medical center should put more resources to provide regular and active care to monitor each patient’s condition before and after each treatment and provide relevant medical knowledge consultation and services through continuous care to closely maintain the patient’s loyalty and long-term relationship with the hospital. In practice, patients can be included in the disease case management system to allow the case manager to actively track each patient’s condition, care for the patient, and carry out the health education program or invite the patients to join a group of patients in their disease category through the patient exchange activities held by the case medical center. The patient exchange activities give patients psychological comfort and professional medical knowledge for treatment. In doing so, the case medical center can effectively establish a long-term close relationship with patients.

The patients with the feature of $R \uparrow F \downarrow M \downarrow$ belong to new outpatients with more recent visit but low frequency and low medical expenditure as shown in Table 5. Most of the

| Variable | Minimum | Average | Maximum | SD |
|----------|---------|---------|---------|----|
| R        | 1       | 239     | 366     | 111|
| F        | 1       | 5       | 220     | 7  |
| M        | 18      | 19,690  | 35,416,990 | 129,961 |
outpatient visits include the followings: department of ophthalmology, department of obstetrics and gynecology, department of emergency medicine, and division of gastroenterology and hepatology. Fifty-five percent of outpatients are 40 years old or more. More than 95% of the outpatients live in Changhua County. The hospital should encourage patients and their families to use the mobile medical app and increase the postmedication satisfaction survey, return reminder, medication reminder, and other functions so as to track outpatients’ conditions as well as postmedication satisfaction to better understand the patients’ needs. In addition, the mobile medical app can provide appropriate medical advices and medical counseling services to improve patient satisfaction and establish long-term relationships with patients.

The patients with the feature of $R^{↓}F^{↓}M^{↓}$ are defined as lost outpatients due to far from the recent visit, low frequency, and low medical expenditure as shown in Table 6. The case hospital can further analyze why the outpatients are unwilling to come as time goes by. For instance, the patients might come to this medical center for the second medical opinion. The case hospital does not meet the patients’ needs such that the patients are reluctant to go back for treatment. Therefore, the hospital needs to figure out the reasons in detail so as to improve the patient-related procedures or increase the medical services to meet the needs of patients.

### Conclusions

Health care organizations need to focus on value for patients instead of minimization of short-term costs such as managing a patient’s personal health, the provision of a health plan, acting responsibly, and building a long-term relationship. Patient loyalty is essential for hospitals because patient loyalty is beyond patient satisfaction when health care service providers are in an intense competition (Ravichandran, 2015). RFM model commonly used to measure the behavioral or attitudinal loyalty based on the past purchase history is applied to measure patient loyalty. Moreover, RFM model with a combination of other approaches such as cluster analysis has been applied to analyze patients’ value or loyalty such as Lee (2012), Chen et al. (2012), and Hosseini and Mohammadzadeh (2016). Therefore, it is a practical approach to adopt RFM model to combine with cluster analysis approaches to evaluate patient loyalty of a hospital.

A case study of analyzing outpatients’ value from the database of a medical center in Taiwan is illustrated with 321,908 outpatients together with 1,683,065 visits in 2016. A combination of SOMs and $K$-means method is employed to cluster outpatients into 12 groups when $R$, $F$, and $M$ are input variables. Each cluster has its average $R$, $F$, and $M$ values and can be further compared with the grand averages of $R$, $F$, and $M$ values. That is, three major types of patients are found,
Table 4. Descriptive Statistics of Loyal Outpatients.

| Type of RFM model | R↑F↑M↑ |
|-------------------|--------|
| Sample of size    | 48,585 |
| Average R value   | 356    |
| Average F value   | 17     |
| Average monetary value | 78,888 |
| Expenditure from loyal patients as a percentage of total medical expenditure in 2016 | 60% |
| The number of visits from loyal patients as a percentage of the total number of visits in 2016 | 50% |
| Gender            |        |
| Female            | 52%    |
| Male              | 48%    |
| Age               |        |
| 17 years old or less | 8%    |
| 18–39 years old   | 12%    |
| 40–64 years old   | 40%    |
| 65 years old or more | 40%  |
| Residence         |        |
| Changhua County   | 60%    |

Note. RFM = recency, frequency, and monetary.

Table 5. Descriptive Statistics of New Outpatients.

| Type of RFM model | R↑F↓M↓ |
|-------------------|--------|
| Sample of size    | 147,925|
| Average R value   | 305    |
| Average F value   | 4      |
| Average monetary value | 12,469 |
| Expenditure from new patients as a percentage of total medical expenditure in 2016 | 29% |
| The number of visits from new patients as a percentage of the total number of visits in 2016 | 36% |
| Gender            |        |
| Female            | 56%    |
| Male              | 44%    |
| Age               |        |
| 17 years old or less | 20%  |
| 18–39 years old   | 26%    |
| 40–64 years old   | 36%    |
| 65 years old or more | 19%  |
| Residence         |        |
| Changhua County   | 95%    |

Note. RFM = recency, frequency, and monetary.

that is, loyal outpatients with R↑F↑M↑, new outpatients with R↑F↓M↓, and lost outpatient with R↓F↓M↓.

Hospital management can further summarize each type of outpatients by its average R, F, and M values, gender, age, residence, and clinic visit, and the like so as to provide different health care needs to different types of patients. Besides, with the detailed outpatients’ demographic variables available from the database, this hospital can provide tailor-made medical services to meet individual’s needs. The purpose of using RFM model in outpatients from a medical center is not to make profits but to focus on what outpatients’ health conditions are in terms of recency, frequency, and total medical expenditure and what might be the optimal medical care for them for treatment.

In practice, loyal outpatients might need much more attention because they visit more recent with high frequency and pay a larger amount of medical expenditure even under National Health Insurance program in Taiwan. That is, they might have poor health conditions and need extra health care. To provide value-based medical care for patients, the medical center should monitor their health conditions relentlessly and provide relevant medical knowledge and health education program for disease recovery. The major contribution of this study is to combine RFM model with cluster analysis (SOM and K-means method) to analyze outpatients’ value from a medical center in Taiwan. With the model illustrated in this study, the hospital can establish a better and long-term relationship with its patients by increasing patient loyalty.

This study has its limitations. First, using RFM model and cluster analysis approaches can be viewed as a basic understanding of patients’ profiles by grouping the similarities of patients in their behaviors. When patients are clustered, classification methods such as decision trees can
be used for each cluster to analyze and identify critical factors to influence patient loyalty (Lee, 2012) or to find the specific needs for patients in terms of departments or divisions (Chen et al., 2012). Second, to establish a patient-centric hospital, sentient computing can be further implemented to deal with structured and unstructured data from patients (Cambria et al., 2012). That is, hospital management can use sentient computing to further analyze patients’ needs based on the results of cluster analysis to provide better and a wide variety of health care services for its patients.

**Declaration of Conflicting Interests**

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Funding**

The author(s) received no financial support for the research, authorship, and/or publication of this article.

**ORCID iD**

Hsin-Hung Wu https://orcid.org/0000-0002-3702-276X

**References**

Bartoletti, I. (2019). AI in healthcare ethical and privacy challenges. In D. Riano, S. Wilk, & A. ten Teije (Eds.), Lecture notes in artificial intelligence: Vol. 11526. Artificial intelligence in medicine (pp. 7–10). Springer Nature.

Cambria, E., Benson, T., Eckl, C., & Hussain, A. (2012). Sentic PROMs: Application of sentient computing to the development of a novel unified framework for measuring health-care quality. Expert Systems with Applications, 39, 10533–10543.

Chen, Y.-S., Cheng, C.-H., Lai, C.-J., Hsu, C.-Y., & Syu, H.-J. (2012). Identifying patients in target customer segments using a two-stage clustering-classification approach: A hospital-based assessment. Computers in Biology and Medicine, 42, 213–221.

Dafny, L. S., & Lee, T. H. (2016). Health care needs real competition. Harvard Business Review, 94(12), 76–87.

Davenport, T., & Kalakota, R. (2019). The potential for artificial intelligence in healthcare. Future Healthcare Journal, 6(2), 94–98.

Ha, S. H., & Park, S. C. (1998). Application of data mining tools to hotel data mart on the internet for database marketing. Expert Systems with Applications, 15(1), 1–31.

Han, J., & Kamber, M. (2006). Data mining concepts and techniques (2nd ed.). Morgan Kaufmann.

Hosseini, Z. Z., & Mohammadzadeh, M. (2016). Knowledge discovery from patients’ behavior via clustering-classification algorithms based on weighted eRFM and CLV model: An empirical study in public health care services. Iranian Journal of Pharmaceutical Research, 15(1), 355–367.

Hughes, A. M. (1996). Boosting response with RFM. Marketing Tools, 3(3), 4–10.

Kuo, R. J., Ho, L. M., & Hu, C. M. (2002). Integration of self-organizing feature map and K-means algorithm for market segmentation. Computers & Operations Research, 29, 1475–1493.

Lee, E. W. (2012). Data mining application in customer relationship management for hospital inpatients. Healthcare Informatics Research, 18(3), 178–185.

MacStravic, R. S. (1987). Loyalty of hospital patients: A vital marketing objective. Health Care Management Review, 12(2), 23–30.

MacStravic, R. S. (1994). Patient loyalty to physicians. Journal of Healthcare Marketing, 14(4), 53–58.

Menger, V., Spruit, M., van der Klift, W., & Scheepers, F. (2019). Using cluster ensembles to identify psychiatric patient subgroups. In D. Riano, S. Wilk, & A. ten Teije (Eds.), Lecture notes in artificial intelligence: Vol. 11526. Artificial intelligence in medicine (pp. 252–262). Springer Nature.

Porter, M., & Teisberg, E. O. (2006). Redefining health care: Creating value-based competition on results. Harvard Business School Press.

**Table 6. Descriptive Statistics of Lost Outpatients.**

| Type of RFM model | RFM*ML |
|------------------|--------|
| Sample of size   | 125,398|
| Average R value  | 115    |
| Average F value  | 2      |
| Average monetary value | 5,271 |
| Expenditure from lost patients as a percentage of total medical expenditure in 2016 | 10%    |
| The number of visits from lost patients as a percentage of the total number of visits in 2016 | 15%    |
| Gender           |        |
| Female           | 60%    |
| Male             | 40%    |
| Age              |        |
| 17 years old or less | 18%   |
| 18–39 years old  | 29%    |
| 40–64 years old  | 37%    |
| 65 years old or more | 16%  |
| Residence        |        |
| Changhua County  | 42%    |

Note. RFM = recency, frequency, and monetary.
Ravichandran, N. (2015). Dimensions of patient loyalty and switching intention: Relational outcomes for benchmarking approach. *International Journal of Healthcare Management, 8*(4), 209–216.

Rundle-Thiele, S., & Russell-Bennett, R. (2010). Patient Influences on Satisfaction and Loyalty for GP Services. *Health Marketing Quarterly, 27*(2), 195–214.

Shieh, J.-I., Wu, H.-H., & Huang, K.-K. (2010). A DEMATEL method in identifying key success factors of hospital service quality. *Knowledge-based Systems, 23*(3), 277–282.

Tang, Y., Guo, Y., Sun, Q., Tang, S., Li, J., Guo, L., & Duan, J. (2018). Industrial polymers classification using laser-induced breakdown spectroscopy combined with self-organizing maps and K-means algorithm. *Optik, 165*, 179–185.

Wei, J. T., Lee, M. C., Chen, H. K., & Wu, H. H. (2013). Customer relationship management in the hairdressing industry: An application of data mining techniques. *Expert Systems with Applications, 40*(18), 7513–7518.

Wei, J. T., Lin, S. Y., & Wu, H. H. (2010). A review of the application of RFM model. *African Journal of Business Management, 4*(19), 4199–4206.

Zhang, Y., Bradlow, E. T., & Small, D. S. (2015). Predicting customer value using clumpiness: From RFM to RFMC. *Marketing Science, 34*(2), 195–208.