Scientific Research Publishing

Journal of Biosciences and Medicines, 2019, 7, 119-135
https://www.scirp.org/journal/jbm
ISSN Online: 2327-509X
ISSN Print: 2327-5081

Study on the Grouping of Patients with Chronic Infectious Diseases Based on Data Mining

Min Li[1,2]

[1]Shanghai Public Health Clinical Center, Fudan University, Shanghai, China
[2]College of Economics and Management, Nanjing University of Aeronautics and Astronautics, Nanjing, China
Email: minliji@aliyun.com

How to cite this paper: Li, M. (2019) Study on the Grouping of Patients with Chronic Infectious Diseases Based on Data Mining. *Journal of Biosciences and Medicines*, **7**, 119-135.
https://doi.org/10.4236/jbm.2019.711011

Received: July 4, 2019
Accepted: November 10, 2019
Published: November 13, 2019

Abstract

Objective: According to RFM model theory of customer relationship management, data mining technology was used to group the chronic infectious disease patients to explore the effect of customer segmentation on the management of patients with different characteristics. **Methods:** 170,246 outpatient data was extracted from the hospital management information system (HIS) during January 2016 to July 2016, 43,448 data was formed after the data cleaning. K-Means clustering algorithm was used to classify patients with chronic infectious diseases, and then C5.0 decision tree algorithm was used to predict the situation of patients with chronic infectious diseases. **Results:** Male patients accounted for 58.7%, patients living in Shanghai accounted for 85.6%. The average age of patients is 45.88 years old, the high incidence age is 25 to 65 years old. Patients was gathered into three categories: 1) Clusters 1—Important patients (4786 people, 11.72%, R = 2.89, F = 11.72, M = 84,302.95); 2) Clustering 2—Major patients (23,103, 53.2%, R = 5.22, F = 3.45, M = 9146.39); 3) Cluster 3—Potential patients (15,559 people, 35.8%, R = 19.77, F = 1.55, M = 1739.09). C5.0 decision tree algorithm was used to predict the treatment situation of patients with chronic infectious diseases, the final treatment time (weeks) is an important predictor, the accuracy rate is 99.94% verified by the confusion model. **Conclusion:** Medical institutions should strengthen the adherence education for patients with chronic infectious diseases, establish the chronic infectious diseases and customer relationship management database, take the initiative to help them improve treatment adherence. Chinese governments at all levels should speed up the construction of hospital information, establish the chronic infectious disease database, strengthen the blocking of mother-to-child transmission, to effectively curb chronic infectious diseases, reduce disease burden and mortality.

Keywords

Data Mining, K-Means Clustering Algorithm, C5.0 Decision Tree Algorithm, Customer Relationship Management, Patients with Chronic Infectious Disease

1. Introduction

The concept of data mining was first proposed at the "11th International Symposium on Joint Artificial Intelligence" held in Detroit in the United States in March 1989, referring to the extraction or excavation of potential from a large amount of noisy messy data. The process of information and patterns useful for decision making is widely used in disease spectrum analysis, gene profiling and hospital management [1] [2] [3] [4]. Data mining technology in the hospital customer relationship management (HCRM) commonly used clustering algorithm, classification algorithm, association rule analysis [5] [6]. Hospital customer relationship management is the customer relationship management theory in the hospital management application [7]. Commonly used data mining technology in the hospital customer relationship management (HCRM) include clustering algorithm, classification algorithm, and association rule analysis [5] [6]. Hospital customer relationship management is the application of customer relationship management theory in hospital management [7].

Customer Relationship Management (CRM) is through the provision of high quality customer satisfaction services or products, and through effective management tools, with a view to establishing a stable customer relationship, the formation of a loyal customer base, to maximize the benefits [8] [9] [10] [11]. RMF model is the most classic model of CRM theory, R (Recency) refers to the recent shopping time, F (Frequency) refers to the shopping frequency, M (Monetary) refers to the cost of shopping [5] [12] [13] [14] [15] [16].

Tuberculosis, a chronic respiratory infections, has become a serious harm to people's health and social public health; WHO reports that about 2 million people had died of tuberculosis each year all over the world [17] [18], China's annual new tuberculosis patients is up to 1 million, ranking the world's second [19]; WHO list China as a high burden of tuberculosis, high-risk countries, second only to India, tuberculosis is one of the key government control disease [19]. There are about 2 billion people infected with Hepatitis B Virus (HBV) in the world. HBV infection rate is as high as 57.63% in China. The positive rate of hepatitis B surface antigen (HBsAg) is more than 120 million, accounting for about 1/3 of the global proportion; In China, chronic hepatitis B patients is currently the largest group of chronic infectious diseases, China is a high hepatitis B prevalence area, about 100 million new cases of hepatitis B patients was found in China each year. The conservative estimate of hepatitis B patients is about 20 million people [20] [21] [22]. According to the WHO and UNAIDS report, as of

the end of 2014, the global existing HIV infection is about 36.9 million. In 2014 the new infected by 2 million. As of March 2015, more than 15 million people worldwide are receiving antiretroviral therapy [23] [24]. At present, although HIV infection is still low in China, the number of people living with HIV in the late 1990s and the beginning of the 21st century has been increasing, and the number of people living with HIV/AIDS has been increasing [24] [25] [26]. In 2003, 4592 cases of AIDS reported in China, accounting for only 5.4% of the estimated number of cases of infection; After the implementation of real-name detection and network direct reporting system, the number of reported cases of survival is 577,000 by 2015, accounting for 67.9% of the estimated number of surviving [24] [25] [26] [27] [28]. Tuberculosis, viral hepatitis B, AIDS belong to chronic infectious disease. Chronic infectious disease management research has now become one of the hot issues [14] [15].

Since the opening of the hospital information system in 2004, research units have accumulated massive patient data. In this study, from the HIS system of research unit extraction from January to July 2016 outpatient data 170,246 pen data, the data after the formation of 43,448 pen patient data, the application of customer relationship management in the classic RFM model, in this study, R refers to the patient The frequency of treatment, F refers to the frequency of treatment, M refers to medical expenses [8], the K-Means clustering algorithm will 43,448 people with chronic infectious diseases divided into different groups, to identify different characteristics of patients with different characteristics of treatment, To explore the impact of segmentation of patients with chronic infectious diseases on patient treatment and hospital development and to make suggestions and countermeasures for medical staff, hospital managers, government reference and provide decision-making basis.

2. Materials and Methods

1) Define the purpose of the study and research field. The research field includes patient ID, gender, date of birth, health insurance card number, serial number, contact address, the main diagnosis, the cost of large items.

2) Data Extraction. 170,246 outpatient data was extracted from hospital management information system (HIS) of the research unit during January to July 2016.

3) Field expansion. The date of birth is expanded to age, and the extension number is expanded to the number of visits (frequency) and the last time (week). The contact address field is expanded to the province/city, city/area, residence (Shanghai, And other) three fields; cost items is expanded to four fields: the cost of outpatient expenses, medical expenses, inspection fees, and other costs.

4) Repeat checking and removing. First, make the patient ID, gender, birth date as a group of markers to find repeat patients; then make medical insurance card number, gender, date of birth as a group of markers to find repeat patients.

Each time check, need to always calculate the frequency of repeated patients, the cost of outpatient, medical expenses, medical expenses, inspection fees, other costs, health insurance costs, at their own expense, update the recent treatment time. At this point, 1 patient retained 1 data.

5) Field conversion. The cost of outpatient and treatment frequency is converted to outpatient cost. At this point, 43,448 outpatient data was final formed.

6) Variable Formation. After the field is expanded, the fields are converted into variables: sex, age, place of residence, frequency of treatment (time), last treatment time (weeks), outpatient total cost (yuan) and so on 7 fields.

7) Variable management. There is no missing value in this study. The outliers and extreme values were treated with mean ±3 standard deviations, mean ±5 standard deviation, respectively.

8) Data visualization by Tableau 10.0 and Excel 2016 analysis. Descriptive statistical analysis was done by SPSS 22.2. Data mining analysis was done by SPSS Modeler 18.0. In this study, two-step clustering algorithm model is found the optimal method by comparing clustering analysis K-Means, Kohonen, two-step clustering; According to two-step clustering algorithm, different types Of patients were classified, and important variables were extracted and identified using C5.0.

3. Results

3.1. Descriptive Statistical Analysis

In this study, male patients (58.7%) accounted for about 2/3. To patients (85.6%) living in Shanghai accounted for the vast majority, followed by the surrounding provinces such as Zhejiang, Jiangsu, Anhui and other places, see **Table 1**, **Table 2**, and **Figure 1**, **Figure 2**.

Table 1 shows the majority of patients are hepatitis in this study, it is consistent with the reported situation in China in the literature [20] [21] [22]. It is the same as suggesting that treatment and cure of all types of hepatitis will continue to be a long and arduous task in China. AIDS patients account for about 20% of the number of patients in this study, suggesting that China needs more people to know, and control the spread of AIDS. Tuberculosis had got some control.

Table 2 and **Figure 2** show that the average age of patients is 45.88 years old, the standard deviation is 17.30 years old, the minimum age is less than 1 year old baby, the maximum age is 106 years; 25 to 65 years old is high incidence, but infants and young children should also be pay attention to. In 7 months studied, the average patient's last visit was about 10 weeks ago, the average frequency is about 3.7 times, the average cost of outpatient is RMB 16,327 yuan. In China, patients with chronic diseases can only buy 2 weeks of medicine each time; The law of antiviral treatment for AIDS patients is 0 months, 0.5 months, 1 month, 2 months, 3 months, followed by regular visits every 3 months, suggesting that some patients with chronic infectious diseases in this study poor adherence. Outpatient cost is 1.1 - 13,473.4 yuan, the average of the cost is 2200.695 yuan;

Table 1. The sociological characteristics of patients.

		Frequency	Percent	disease	Frequency	Percent
Contact address	Shanghai	37,194	85.6	AIDS	7642	17.6
	Outland	6254	14.4	Hepatitis	22,893	52.7
Sex	Female	17,942	41.3	Tuberculosis	727	1.7
	Male	25,506	58.7	Others	12,186	28

Table 2. Treatment situation of patients.

	N	Minimum	Maximum	Mean	Std. Deviation
Age	43,448	0	106	45.88	17.3
Recency	43,448	1	30	10.17	8.358
Frequency	43,448	1	16	3.71	3.629
Monetary-mean	43,448	1.1	13,473.4	2200.695	2768.3964
Monetary	43,448	1	226,852	16,327.49	35,375.95

Figure 1. The map of the AIDS patient distributes.

the total outpatient cost is 1 - 226,852 yuan, the average of cost of patients is 16,327.49 yuan, the average medical cost is 2332 yuan/month, suggesting that the burden of direct medical expenses in Chinese patients with chronic infectious diseases is heavier.

3.2. Two-Step Clustering Algorithm Was Used to Classify 43,448 Patients with Different Types of Groups

After comparison, this study finally selected three important variables based on the customer relationship management RFM model: R Last time (weeks), F treatment frequency (times), M outpatient total cost (yuan), multiple clustering combined with clinical. This study suggests that outpatients with the final group of three are appropriate. Results of two-step clustering analysis are shown in **Table 3**.

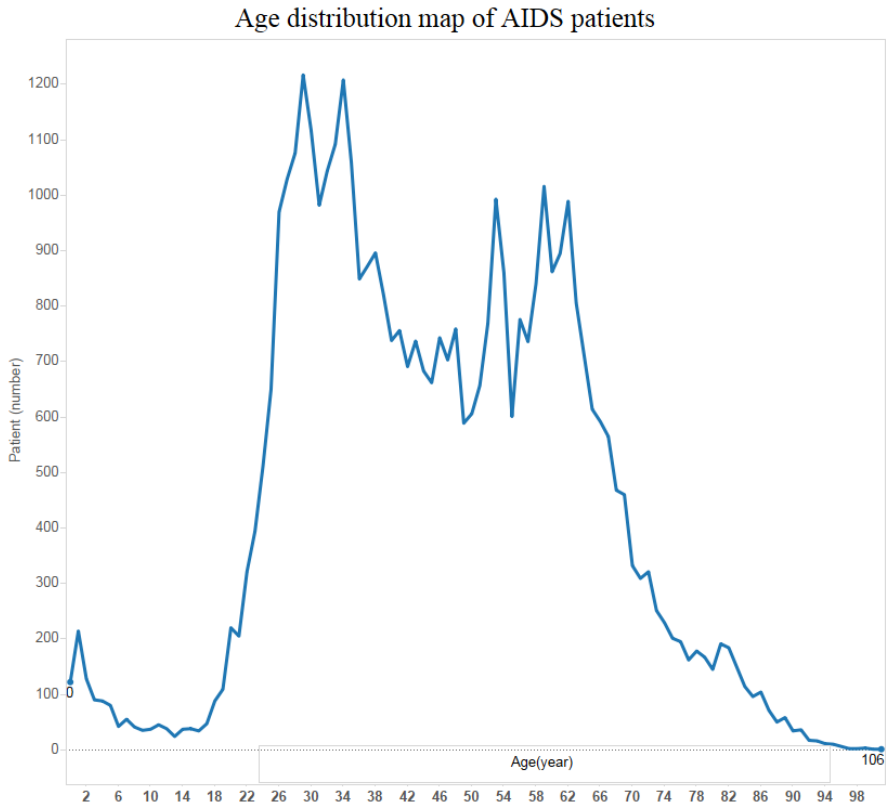

Figure 2. Age distribution map of AIDS patients.

Table 3. Map of two-step clustering model.

Clustering	Clustering-Two	Clustering-Three	Clustering-One
Description	M = 9146.39 R = 5.22 F = 3.45	F = 1.55 M = 1739.09 R = 19.77	F = 11.72 M = 84,302.95 R = 2.89
Proportion	53.2% (23,103)	35.8% (15,559)	11.0% (4786)
Enter the variable	Monetary 9146.39 Recency 5.22 Frequency 3.45	Frequency 1.55 Monetary 1739.09 Recency 19.77	Frequency 11.72 Monetary 84,302.95 Recency 2.89

From **Table 3** we can be found, the frequency of treatment, the recent treatment weeks, outpatient costs are important variables; the two-step clustering algorithm is of good quality and it can clustered patients into three categories: First, clustered 1—important patients (4786), accounting for 11.0% of the number of outpatient visits, they have visit for treatment about 11.72 times in 7 months, the final treatment time about 2.89 weeks ago, outpatient costs about 84,302.95 yuan. The final treatment time is about 2.89 weeks ago, their outpatient costs about 84,302.95 yuan. Second, clustering 2—Major patients (23,103), accounting for 53.2% of the number of outpatient visits, about 3.45 visits in 7

months, the final treatment time in 5.2 weeks ago, the outpatient cost of about 9146.39 yuan; Third, clustering 3—potential patients (15,559), accounting for 35.8% of the number of outpatient visits, about 1.55 visits in 7 months, the final treatment time in 19.77 weeks ago, outpatient costs 1739.09 yuan, this part of the patient may be physical examination, or patients who were treated outside Shanghai, who were cured, who were lost patient, or patients with poor adherence. Suggesting that the main patients in this study should be frequent patients, Potential patients can be maintained by improving the quality of service or become the annual physical health of the physical examination, the important patients, who may be outside Shanghai patients, and seriously ill patients, not only bring outpatient business, outpatient income, they may also be a violation of health insurance.

From **Table 4** we can be found that the number of potential patients is 15,563, 12,738 patients out of potential patients live in Shanghai, 2825 patients outside Shanghai, suggesting that some patients with chronic infectious diseases can't be properly treated for various reasons. From **Table 4**, **Table 5** we can be found, regardless of local or foreign AIDS patients, viral hepatitis patients, tuberculosis patients, pregnant women are likely to be lost patients.

3.3. Application of C5.0 Decision Tree Algorithm to Predict the Situation of Patients with Chronic Infectious Diseases

According to the two-step algorithm, 43448 chronic infectious diseases were divided into three groups. Application of 5.0 decision tree algorithm to predict the situation of patients with chronic infectious diseases, as shown in **Figure 3**.

It can be seen from **Figure 3** that patients with a final visit time (weeks) greater than 10 weeks and a visit frequency of 1 (see) can be predicted as cluster 3 (potential patient). Patients with a final visit time (weeks) greater than 10 weeks but the frequency of visits (times) is greater than or equal to 1 time, and Patients with a final treatment time (weeks) less than or equal to 10 weeks, outpatient total cost (yuan) less than or equal to 41,973.4 yuan, the treatment frequency is less than or equal to 8 times can be predicted for cluster 2 (primary patient). Patients with a final visit time (Weeks) lesser than or equal to 10 weeks,

Table 4. The cross-table of 3 groups of patients with gender and residence.

		Cluster 1 (Important patients)	Cluster 2 (Major patients)	Cluster 3 (Potential patients)	total
Contact address	Shanghai	4540 (12.20%)	19,916 (53.55%)	12,738 (34.25%)	37194
	Outland	242 (3.87%)	3187 (50.96%)	2825 (45.17%)	6254
	total	4782	23,103	15,563	43,448
Sex	Female	1970	8674	7298	17,942
	Male	2812	14,429	8265	25,506
	total	4782	23,103	15,563	43,448

Table 5. Cross-linked table of 3 groups of patients with patients having 7 different diseases.

Disease	Place of residence	Cluster 1 (important patients)	Cluster 2 (Major patients)	Cluster 3 (potential patients)	total
AIDS	Shanghai	296	5226	934	6456
	No-Shanghai	63	765	358	1186
Hepatitis B	Shanghai	1142	3661	2046	6849
	No-Shanghai	43	618	469	1130
Chronic hepatitis B	Shanghai	1549	1530	254	3333
	No-Shanghai	47	138	39	224
Hepatitis B cirrhosis	Shanghai	225	216	87	528
	No-Shanghai	10	37	33	80
Other liver disease	Shanghai	197	434	303	934
	No-Shanghai	4	103	97	204
Tuberculosis	Shanghai	57	273	189	519
	No-Shanghai	11	118	79	208
Pregnancy with liver disease	Shanghai	18	578	646	1242
	No-Shanghai	2	101	61	164

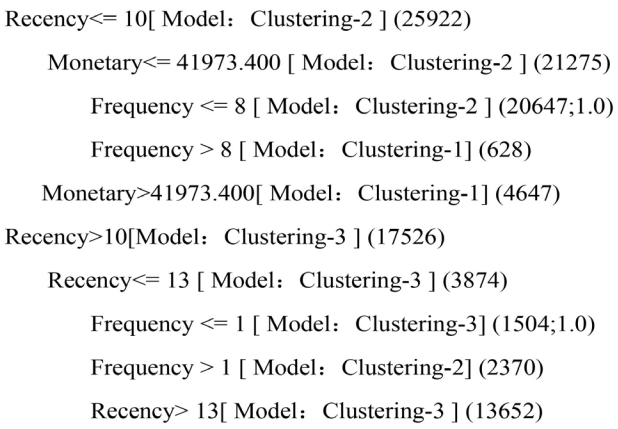

Figure 3. Model diagram of treatment situation for patients with chronic infectious diseases predicted by 5.0 Decision tree algorithm.

the total cost of the outpatient service (yuan) greater than 41,973.4 yuan, or patients with a final treatment time (weeks) lesser than or equal to 10 weeks, the total cost of outpatient (yuan) lesser than or equal to 41,973.4 yuan, the frequency of more than 8 times can be Predicted for cluster 1 (important patient). The most important variables to predict whether patients with chronic infectious diseases are on time to receive normal treatment is the final visit time (weeks), followed by the total cost of outpatient service (yuan) and the frequency of visits (times). The validation of the confusing model shows an accuracy rate of 99.94%, and the ROC curve is used to confirm the accuracy. As shown in **Table 6**.

Table 6. Validation table of variables predictive confusion model by C5.0 decision tree.

	N	Percent
correct	43,421	99.94%
error	27	0.06%
total	43,448	

4. Analysis and Discussion

4.1. Take the Initiative to Improve Patient Adherence Who Has Chronic Infectious Disease

China is a big country of hepatitis and tuberculosis. Therefore, it is necessary to concern about and supervise the treatment of chronic infectious diseases applying customer relationship management theory and data mining technology in China. The research unit is located in the suburbs of Shanghai, China. It is a class 3 first level specialist hospital for the treatment of statutory infectious diseases, with 660 beds approved and of nearly 900 medical staff.

Adherence refers to the extent to which a patient follows a treatment plan [14] [15]. Clustering 3, potential patients, is 15,559 people, accounting for 35.8% of the number of outpatient visits, 12,559 out of it live in Shanghai, 2825 people live out of Shanghai. Patients in this cluster are characteristic with the final treatment time (weeks) more than 10 weeks and the frequency of visits (times) is only 1. It includes persons who conduct medical examination, Non-Shanghai patients, cured, lost patients, and poor adherence patients. Poor adherence can't be properly treated possibly due to adverse drug reactions, long treatment time, self-feeling stable condition, doctors and patients are not in place, easy to forget, low level of education, medical inconvenience and small social support and other factors [29] [30] [31]. The research unit is located in the suburbs of Shanghai, China. It is the only class 3 first level specialist hospital admitted to the statutory infectious diseases in Shanghai. Hepatitis, tuberculosis, AIDS, are three kinds of most common disease, with 660 beds approved and of nearly 900 medical staff. In Shanghai City, there are Huashan Hospital, Zhongshan Hospital, Ruijin Hospital, Changzheng Hospital, Oriental Hepatobiliary Hospital, Renji Hospital, 6 well-known third-grade class-A general hospital are admitted to liver disease; Pulmonary hospital and Ruijin hospital admitted to tuberculosis; Obstetrics and Gynecology Hospital and the first maternal and child hospital admitted to pregnant women; AIDS is only in the research unit. The study found that both local and foreign AIDS patients, viral hepatitis patients, tuberculosis patients, pregnant patients may be lost patients, good medication adherence is the key to improve the cure rate of tuberculosis patients, reduce the recurrence rate [17] [18] [19]. The key to treatment of HBV is antiviral therapy, it needs long-term, standardized medication, and adherence is the basis of the effectiveness of drug treatment. The consequences of poor adherence are serious. Light consequences are adversely affected by the disease, leading to treatment failure, Severe cases lead to serious damage to liver function or even life-threatening [20] [21] [22].

Adherence has proven to be a key factor affecting HAART therapy [32], High adherence can reduce the number of drug-resistant strains and their further spread, can reduce opportunistic infections [33], reduce mortality and morbidity [34], reduce the cost of AIDS prevention and disease burden [32] [35]. Patients with chronic infectious diseases do not accept formal treatment is easy to lead to their own opportunistic infections are also easy to spread the disease, increased mortality and disease burden, easily lead to social instability. Therefore, the hospital should strengthen the adherence education for chronic infectious diseases, education for different chronic infectious diseases to carry out adherence education, make full use of mobile phone APP function, the establishment of hospital patient customer groups, the establishment of special disease patients customer base, regularly introduce adherence with different chronic diseases related content, active education to help patients to understand the importance of adherence, reduce the incidence of chronic infectious diseases, so that more patients with chronic infectious diseases to receive regular treatment.

The actually lost chronic infectious diseases patients is the focus of the hospital, government, national concern, They can't timely, on time take treatment is bound to aggravate the condition, the following-up treatment is bound to increase the patient's own and his family's medical burden, more importantly, it is to increase the social disease and economic burden. Therefore, we believe that if can find out the lost chronic infectious diseases patients and improve their compliance, it will improve the world's chronic infectious diseases patient's compliance, and reduce the mortality and medical burden.

4.2. Establishment of Customer Relationship Management Database for Patients with Chronic Infectious Diseases

This study found that 1) the most important indicator of the treatment of chronic infectious diseases was the final treatment time (weeks); 2) the actual number of patients was 43,448 people out of the treatment of 170,246 people, only 23,103 (53.2%) patients were clear diagnosed with stable condition from January 2016 to July 2016 in the research unit, including AIDS patients. Therefore, how to use the last visit time (week), the important predictive index, to good service and manage this 23,103 patients with chronic infectious diseases? At present, the research unit has not established the hospital customer relationship management, but with reference to the short-term chemotherapy (DOTS) program under full supervision recommended by the World Health Organization (WHO) [18], the research unit has set up a management records management records Model for AIDS patients, equipped with four case managers. There are currently information about 10,000 copies, including the information about the death of patients, foreign patients but converted to other hospital units for treatment, foreign resident patients, local patients, patients are scheduled to visit the next time and information was print and given to each patient at the end of each visit to consult the, and the necessary telephone follow-up to urge AIDS patients to visit doctor in a timely manner. At the same time, the number of AIDS

patients in the research unit is about 1/6 of the total number of 43,448 people, but the number of visits is more than 1/5 of the total treatment of 170,246 people.

Therefore, this paper believes that China's infectious diseases hospital should establish a customer relationship management database of patients with chronic infectious diseases, to understand the situation of these patients, and the following work can be carried out according to the final treatment time (week): 1) telephone follow-up; 2) The last (as of April) medical information book will be placed in the next (as of April) treatment column, this can allow hospital manager to confirm whether patients with chronic infectious diseases have got treatment by phone, e-mail, SMS, QQ SMS, WeChat, or mobile phone APP 1 to 3 days in advance, or found overdue patients and allow hospital manager to contacted and managed these patients timely by telephone, QQ SMS, microphones, mobile APP [36] [37] [38] [39] [40]. Only effective management of chronic infectious diseases can effectively reduce the mortality and disease burden. Chinese governments at all levels should speeded up the construction of hospital information, centralized management of patient data, well done the job of customer relationship management for patients with chronic disease, chronic infectious diseases, sent drugs for basic chronic diseases and chronic infectious diseases to community hospitals, both to achieve a two-way referral of medical care, but also easy to do community hospital chronic diseases, chronic infectious diseases management work in order to help curb chronic infectious diseases in order to help reduce mortality and disease burden.

4.3. Increase the Proportion of Medical Investment, Improve the Proportion of Medical Reimbursement

In this study, 4786 people accounted for 11.0% of the number of outpatient were treated about 11.72 times in 7 months. The final treatment time is about 2.89 weeks ago. The outpatient total cost is about 84,302.95 yuan, the average of the cost is 12,043 yuan/month. But the average medical cost of patients with stable condition is about 1300 yuan/month. This data show that changes in condition or complication has resulted in increased medical burden. Although China has medical insurance for urban workers, medical insurance for urban residents, new rural cooperative medical insurance, but the proportion of personal cash health expenditure to total health expenditure is too high, the growth rate of government health expenditure is lower than the growth rate of fiscal expenditure, the gap between urban and rural health care services is growing, but the direct health care costs are significantly higher than non-medical costs, and the severity of the disease is positively correlated with the economic burden [41] [42]. China's development is uneven between eastern and western region, there is a large gap in medical investment in the western region, it exacerbated back to poverty due to illness [43]; People in rural or ethnic minority areas are often treated by selling their own assets or borrowing from family members. Support for free or subsidized treatment and medical insurance forms is considered es-

sential, but not enough to reduce the patient's medical burden [44].

Therefore, we should increase the proportion of medical investment, investment intensity is not lower than the previous year's local economic GDP gains; Speed up the realization of universal health insurance, and improve the proportion of medical Reimbursement，reduce the proportion of individual pays for chronic infectious diseases, guide residents to use drugs rationally, and improve the efficiency of health care fund use [45] [46].

4.4. It Is a Long Way to Curb Chronic Infectious Diseases for China

In this study, the ratio of male patient to female patient was 1.4:1. The ratio of patients living in Shanghai to foreign patients was 5.9:1. Patients distribute throughout China provinces. Patients age distribute from 1 to 106 years old, with 25 to 65 years old patients to be the high incidence people. This group of patients is the main labor force, when dinner, they should as much as possible to use public chopsticks, public spoon, and accept regular treatment, to reduce viral hepatitis and other infectious diseases [47]; When sexual life, appropriate protective measures should be taken to avoid being infected. Some patients are younger than 10 years old. In fact, viral hepatitis, AIDS mother-to-child blockade has been achieved in the research unit. There are still unblocking mother to child spread of viral hepatitis and AIDS in China is due to following two cause: First, some pregnant women do not know that they have been infected with the virus or have been sick; Second, part of the maternal is known to be infected or sick, but they do not know the existence of medical technology to block mother-to-child transmission. Therefore, the Chinese government should expand and increase surveillance of sentinel sites [48]. Free or small payment of maternal screening for infectious diseases, the detection and treatment coverage should be further expanded with a view to early detection and comprehensive implementation of integrated interventions for infected pregnant women and their children [49] [50].

5. Conclusion

Out of the 43,448 patients, 4786 (11.0%) were important patients, they have high frequency of visits and high medical costs. They may be patients with conditional changes or complex condition, it should increase the proportion of medical investment, improve the proportion of medical claims reimbursement. 23,103 people (53.2%) were major patients, they may be patients with clear diagnosis and stable condition. A good job of medical services should be done for major patients. Potential patients 15,559 (35.8%), their last treatment time is long and the frequency is low. They may be patients with annual medical examinations, patients who are treated outside Shanghai, patients who have been cured, who are losing patients, or who are poorly adherence. According to the forecast, the most important indicator of the treatment of chronic infectious diseases is the final treatment time (weeks). Medical institutions can strengthen the adherence

education to patients with chronic infectious diseases, establish the customer relationship management database for patients with chronic infectious diseases, take initiative to help them improve treatment adherence by telephone, e-mail, QQ SMS, WeChat, mobile phone APP to reduce mortality and disease burden. Chinese governments at all levels should speed up the construction of hospital information construction, establish the patient database, sent drugs for basic chronic diseases and chronic infectious diseases to community hospitals, both to achieve a two-way referral of medical care, it will be conducive to curb chronic infectious diseases, reduce mortality and disease burden. In addition, Chinese governments at all levels should also strengthen the screening of maternal infectious diseases, with a view to early detection and effective implementation of integrated interventions to pregnant women with infectious disease.

Acknowledgements

This study was supported by Three-year Action Program of Shanghai Municipality for Strengthening the Construction of Public Health System (No.15GWZK0103), the Program to Support Medical Science of the Shanghai Science and Technology Committee (No.16411953800), the Program to Support Medical Science of Hospital Development Center (No.SHDC2015636), and the Hospital Management Research Fund of Shanghai Hospital Association (No.1601036 and No.201701005). The authors have no conflicts of interest to disclose.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

[1] Sung, S.F., Hsieh, C.Y., Kao Yang, Y.H., Lin, H.J., Chen, C.H., Chen, Y.W. and Hu, Y.H. (2015) Developing a Stroke Severity Index Based on Administrative Data Was Feasible Using Data Mining Techniques. *J Clin Epidemiol*, **68**, 1292-1300. https://doi.org/10.1016/j.jclinepi.2015.01.009

[2] Leary, A., Cook, R., Jones, S., Smith, J., Gough, M., Maxwell, E., Punshon, G. and Radford, M. (2016) Mining Routinely Collected Acute Data to Reveal Non-Linear Relationships between Nurse Staffing Levels and Outcomes. *BMJ Open*, **6**, e011177. https://doi.org/10.1136/bmjopen-2016-011177

[3] Teufel, A. (2015) Bioinformatics and Database Resources in Hepatology. *J Hepatol,* **62**, 712-719. https://doi.org/10.1016/j.jhep.2014.10.036

[4] Cao, B., Kong, X., Kettering, C., Yu, P. and Ragin, A. (2015) Determinants of HIV-Induced Brain Changes in Three Different Periods of the Early Clinical Course: A Data Mining Analysis. *Neuroimage Clin*, **9**, 75-82. https://doi.org/10.1016/j.nicl.2015.07.012

[5] Zare Hosseini, Z. and 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. *Iran J Pharm Res*, **15**, 355-367.

[6] Lee, E.W. (2012) Data Mining Application in Customer Relationship Management for Hospital Inpatients. *Healthc Inform Res*, **18**, 178-185. https://doi.org/10.4258/hir.2012.18.3.178

[7] Kondasani, R.K. and Panda, R.K. (2015) Customer Perceived Service Quality, Satisfaction and Loyalty in Indian Private Health Care. *Int J Health Care Qual Assur*, **28**, 452-467. https://doi.org/10.1108/IJHCQA-01-2015-0008

[8] Poku, M.K., Behkami, N.A. and Bates, D.W. (2017) Patient Relationship Management: What the U.S. Healthcare System Can Learn from Other Industries. *J Gen Intern Med.*, **32**, 101-104. https://doi.org/10.1007/s11606-016-3836-6

[9] Walker, D.D., van Jaarsveld, D.D. and Skarlicki, D.P. (2017) Sticks and Stones Can Break My Bones but Words Can Also Hurt Me: The Relationship between Customer Verbal Aggression and Employee Incivility. *J Appl Psychol*, **102**, 163-179. https://doi.org/10.1037/apl0000170

[10] Orenga-Roglá, S. and Chalmeta, R. (2016) Springer plus Social Customer Relationship Management: Taking Advantage of Web 2.0 and Big Data Technologies. *Springerplus*, **5**, 1462-1478. https://doi.org/10.1186/s40064-016-3128-y

[11] Chen, Y.S., Cheng, C.H., Lai, C.J., Hsu, C.Y. and Syu, H.J. (2012) Identifying Patients in Target Customer Segments Using a Two-Stage Clustering-Classification Approach: A Hospital-Based Assessment. *Comput Biol Med.*, **42**, 213-221. https://doi.org/10.1016/j.compbiomed.2011.11.010

[12] Cubillas, J.J., Ramos, M.I., Feito, F.R., González, J.M., Gersol, R. and Ramos, M.B. (2015) Importance of Health CRM in Pandemics and Health Alerts. *Aten Primaria*, **47**, 267-272. https://doi.org/10.1016/j.aprim.2014.05.013

[13] Quero, M., Ramos, M.B., López, W., Cubillas, J.J., González, J.M. and Castillo, J.L. (2016) Use of Customer Relationship Management to Improve Healthcare for Citizens. The 24h Andalusian Health Service: Health line. *Gac Sanit*, **30**, 397-400. https://doi.org/10.1016/j.gaceta.2016.01.001

[14] Kardas, P., Lewek, P. and Matyjaszczyk, M. (2013) Determinants of Patient Adherence: A Review of Systematic Reviews. *Front Pharmacol*, **25**, 91-106. https://doi.org/10.3389/fphar.2013.00091

[15] Bright, C.E. (2017) Measuring Medication Adherence in Patients With Schizophrenia: An Integrative Review. *Arch Psychiatr Nurs*, **31**, 99-110. https://doi.org/10.1016/j.apnu.2016.09.003

[16] Lo-Ciganic, W.H., Donohue, J.M., Jones, B.L., Perera, S., Thorpe, J.M., Thorpe, C.T., Marcum, Z.A. and Gellad, W.F. (2016) Trajectories of Diabetes Medication Adherence and Hospitalization Risk: A Retrospective Cohort Study in a Large State Medicaid Program. *J Gen Intern Med*, **31**, 1052-1060. https://doi.org/10.1007/s11606-016-3747-6

[17] Lönnroth, K., Castro, G.K., Chakaya, J.M., Chauhan, L.S., Floy, K., Glaziou, P. and Raviglione, M.C. (2010) Tuberculosis Control and Elimination 2010-50: Cure, Care, and Social Development. *Lancet*, **375**, 1755-1757. https://doi.org/10.1016/S0140-6736(10)60483-7

[18] Cremers, A.L., Gerrets, R., Kapata, N., Kabika, A., Birnie, E., Klipstein-Grobusch, K. and Grobusch, M.P. (2016) Tuberculosis Patients' Pre-Hospital Delay and Non-Adherence with a Longstanding DOT Programme: A Mixed Methods Study in Urban Zambia. *BMC Public Health*, **16**, 1130. https://doi.org/10.1186/s12889-016-3771-9

[19] Lei, X., Huang, K., Liu, Q., Jie, Y.F. and Tang, S.L. (2016) Are Tuberculosis Patients

Adherent to Prescribed Treatments in China? Results of a Prospective Cohort Study. *Infect Dis Poverty*, **5**, 38. https://doi.org/10.1186/s40249-016-0134-9

[20] Song, Z.L., Cui, Y.J., Zheng, W.P., Teng, D.H. and Zheng, H. (2015) Application of Nucleoside Analogues to Liver Transplant Recipients with Hepatitis B. *World J Gastroenterol*, **21**, 12091-12100. https://doi.org/10.3748/wjg.v21.i42.12091

[21] Hegde, S.K., Fathima, F.N., Agrawal, T. and Misquith, D. (2016) Adherence to Prescribed Medications for Chronic Illnesses among Older Adults in a Rural Community, Karnataka, India. *Geriatr Gerontol Int*, **16**, 1339-1345. https://doi.org/10.1111/ggi.12637

[22] Wójcik, K., Piekarska, A. and Jabłonowska, E. (2016) Adherence to Antiviral Therapy in HIV or HBV-Infected Patients. *Przegl Epidemiol*, **70**, 27-32, 115-118.

[23] Petersen, T., Townsend, K., Gordon, L.A., Sidharthan, S., Silk, R., Nelson, A., Gross, C., Calderón, M., Proschan, M., Osinusi, A., Polis, M.A., Masur, H., Kottilil, S. and Kohli, A. (2016) High Adherence to All-Oral Directly Acting Antiviral HCV Therapy among an Inner-City Patient Population in a Phase 2a Study. *Hepatol Int*, **10**, 310-319. https://doi.org/10.1007/s12072-015-9680-7

[24] Wang, W., Liu, W., Chen, T., Liu, N.P., Zheng, Y.J., Ye, S.D., Zhang, Y., Wang, X.M., Wang, G.Z. and Zhang, H.B. (2017) Factors Influencing Antiretroviral Therapy Adherence among HIV-Infected People on Antiretroviral Therapy in Ili Kazakh Autonomous Prefecture. *Zhonghua Yu Fang Yi XueZa Zhi*, **51**, 160-164.

[25] Unger, N.R., Worley, M.V., Kisgen, J.J., Sherman, E.M. and Childs-Kean, L.M. (2016) Elvitegravir for the Treatment of HIV. *Expert Opin Pharmacother*, **17**, 2359-2370. https://doi.org/10.1080/14656566.2016.1250885

[26] Qin, Q., Guo, W., Tang, W., Mahapatra, T., Wang, L., Zhang, N., Ding, Z., Cai, C., Cui, Y. and Sun, J. (2017) Spatial Analysis of the Human Immunodeficiency Virus Epidemic among Men Who Have Sex with Men in China, 2006-2015. *Clin Infect Dis*, **64**, 956-963. https://doi.org/10.1093/cid/cix031

[27] Yan, J., Liu, J., Su, B., Pan, X., Wang, Z., Wu, J., Zhang, J., Ruan, Y., His, J., Liao, L., Shao, Y. and Xing, H. (2016) Lamivudine Concentration in Hair and Prediction of Virologic Failure and Drug Resistance among HIV Patients Receiving Free ART in China. *PLoS One*, **11**, e0154421. https://doi.org/10.1371/journal.pone.0154421

[28] Xie, N., Xu, J., Yao, Z., Zhou, W. and Wang, X. (2016) Mode of Antiretroviral Treatment Based on the Medical Needs of People Living with HIV/AIDS. *Zhonghua Liu Xing Bing Xue Za Zhi*, **37**, 539-542.

[29] Egelund, E.F., Dupree, L., Huesgen, E. and Peloquin, C.A. (2017) The Pharmacological Challenges of Treating Tuberculosis and HIV Coinfections. *Expert Rev Clin Pharmacol*, **10**, 213-223. https://doi.org/10.1080/17512433.2017.1259066

[30] Bandera, A., Colella, E., Rizzardini, G., Gori, A. and Clerici, M. (2017) Strategies to Limit Immune-Activation in HIV Patients. *Expert Rev Anti Infect Ther*, **15**, 43-54. https://doi.org/10.1080/14787210.2017.1250624

[31] Palma, A.M., Rabkin, M., Nuwagaba-Biribonwoha, H., Bongomin, P., Lukhele, N., Dlamini, X., Kidane, A. and El-Sadr, W.M. (2016) Can the Success of HIV Scale-Up Advance the Global Chronic NCD Agenda? *Glob Heart*, **11**, 403-408. https://doi.org/10.1016/j.gheart.2016.10.012

[32] Escolano, A., Dosenovic, P. and Nussenzweig, M.C. (2017) Progress toward Active or Passive HIV-1 Vaccination. *J Exp Med*, **214**, 3-16. https://doi.org/10.1084/jem.20161765

[33] Chaisson, L.H., Saraceni, V., Cohn, S., Seabrook, D., Cavalcante, S., Chaisson, R.E., Golub, J. and Durovni, B. (2019) CD4 Count Stratification to Guide Tuberculosis Preventive Therapy for People Living with HIV. *AIDS*. https://doi.org/10.1097/QAD.0000000000002398

[34] Kagujje, M., Mubiana, M.L., Mwamba, E. and Muyoyeta, M. (2019) Implementation of Isoniazid Preventive Therapy in People Living with HIV in Zambia: Challenges and Lessons. *BMC Public Health*, **19**, 1329. https://doi.org/10.1186/s12889-019-7652-x

[35] Shah, M., Risher, K., Berry, S.A. and Dowdy, D.W. (2016) The Epidemiologic and Economic Impact of Improving HIV Testing, Linkage, and Retention in Care in the United States. *Clin Infect Dis*, **62**, 220-229. https://doi.org/10.1093/cid/civ801

[36] Fang, Q., Hao, Y.H., Li, L., Zhang, Z.H., Wang, J.Z., Chen, M.F., Kang, J. and Yang, D.L. (2013) Chronic HBV Carrier's Acceptance of Regular Surveilling Program in China. *JHuazhong UnivSci Technolog Med Sci*, **33**, 288-292. https://doi.org/10.1007/s11596-013-1113-z

[37] Sun, W.-C., Hsu, P.-I., Yu, H.-C., , Lin, K.-H. Tsay, F.-W., Wang, H.-M., Tsai, T.-J., Chen, W.-C., Lai, K.-H. and Cheng, J.-S. (2015) The Adherence of Doctors with Viral Hepatitis B Screening and Antiviral Prophylaxis in Cancer Patients Receiving Cytotoxic Chemotherapy Using a Hospital-Based Screening Reminder System. *PLoS One*, **10**, 1-16. https://doi.org/10.1371/journal.pone.0116978

[38] Ford, N., Matteelli, A., Shubber, Z., Hermans, S., Meintjes, G., Grinsztejn, B., Waldrop, G., Kranzer, K., Doherty, M. and Getahun, H. (2016) TB as a Cause of Hospitalization and In-Hospital Mortality among People Living with HIV Worldwide: A Systematic Review and Meta-Analysis. *J Int AIDS Soc*, **19**, 20714. https://doi.org/10.7448/IAS.19.1.20714

[39] Klein, C.G., Cicinnati, V., Schmidt, H., Ganten, T., Scherer, M.N., Braun, F., Zeuzem, S., Wartenberg-Demand, A., Niemann, G., Schmeidl, R. and Beckebaum, S. (2013) Adherence and Tolerability of Subcutaneous Hepatitis B Immunoglobulin Self-Administration in Liver Transplant Patients: A Prospective, Observational, Multicenter Study. *Ann Transplant*, **13**, 677-684. https://doi.org/10.12659/AOT.889269

[40] Terasaki, J., Singh, G., Zhang, W., Wagner, P. and Sharma, G. (2015) Using EMR to Improve Adherence with Clinical Practice Guidelines for Management of Stable COPD. *Respir Med*, **109**, 1423-1429. https://doi.org/10.1016/j.rmed.2015.10.003

[41] Zhang, S., Ma, Q., Liang, S., Xiao, H., Zhuang, G., Zou, Y., Tan, H., Liu, J., Zhang, Y., Zhang, L., Feng, X., Xue, L., Hu, D., Cui, F. and Liang, X. (2016) Annual Economic Burden of Hepatitis B Virus-Related Diseases among Hospitalized Patients in Twelve Cities in China. *J Viral Hepat*, **23**, 202-210. https://doi.org/10.1111/jvh.12482

[42] Xiao, J., Lin, H., Liu, T., Zeng, W., Li, X., Shao, X., Tan, Q., Xu, Y., Xu, X., Zheng, H. and Ma, W. (2015) Disease Burden from Hepatitis B Virus Infection in Guangdong Province, China. *Int J Environ Res Public Health*, **12**, 14055-14067. https://doi.org/10.3390/ijerph121114055

[43] Che, Y.H., Chongsuvivatwong, V., Li, L., Sriplung, H., Wang, Y.Y., You, J., Ma, S.J., Yan, Y., Zhang, R.Y., Shen, T., Chen, H.M., Rao, S.F. and Zhang, X.L. (2016) Financial Burden on the Families of Patients with Hepatitis B Virus-Related Liver Diseases and the Role of Public Health Insurance in Yunnan Province of China. *Public Health*, **130**, 13-20. https://doi.org/10.1016/j.puhe.2015.03.015

[44] Hutchison, C., Khan, M.S., Yoong, J., Lin, X. and Coker, R.J. (2017) Financial Bar-

riers and Coping Strategies: A Qualitative Study of Accessing Multidrug-Resistant Tuberculosis and Tuberculosis Care in Yunnan, China. *BMC Public Health*, **17**, 221. https://doi.org/10.1186/s12889-017-4089-y

[45] Are, C., Meyer, B., Stack, A., Ahmad, H., Smith, L., Qian, B., Song, T. and Chowdhury, S. (2017) Global Trends in the Burden of Liver Cancer. *J Surg Oncol*, **115**, 591. https://doi.org/10.1002/jso.24518

[46] Ding, P., Li, X., Jia, Z. and Lu, Z. (2017) Multidrug-Resistant Tuberculosis (MDR-TB) Disease Burden in China: A Systematic Review and Spatio-Temporal Analysis. *BMC Infect Dis*, **17**, 57. https://doi.org/10.1186/s12879-016-2151-5

[47] Gish, R.G., Yuen, M.F., Chan, H.L., Given, B.D., Lai, C.L., Locarnini, S.A., Lau, J.Y., Wooddell, C.I., Schluep, T. and Lewis, D.L. (2015) Synthetic RNAi Triggers and Their Use in Chronic Hepatitis B Therapies with Curative Intent. *Antiviral Res*, **121**, 97-108. https://doi.org/10.1016/j.antiviral.2015.06.019

[48] Joshi, R.K. and Mehendale, S.M. (2016) Can We Replace HIV Sentinel Surveillance Platform with Prevention of Parent-to-Child Transmission (PPTCT) Program Data to Assess HIV Burden and Trends in India? *Trans R Soc Trop Med Hyg*, **110**, 393-399. https://doi.org/10.1093/trstmh/trw045

[49] Yotebieng, M., Thirumurthy, H., Moracco, K.E, Kawende, B., Chalachala, J.L., Wenzi, L.K., Ravelomanana, N.L., Edmonds, A., Thompson, D., Okitolonda, E.W. and Behets, F. (2016) Conditional Cash Transfers and Uptake of and Retention in Prevention of Mother-to-Child HIV Transmission Care: A Randomised Controlled Trial. *Lancet HIV*, **3**, e85-93. https://doi.org/10.1016/S2352-3018(15)00247-7

[50] Ilboudo, D., Simpore, J., Ouermi, D., Bisseye, C., Sagna, T., Odolini, S., Buelli, F., Pietra, V., Pignatelli, S., Gnoula, C., Nikiema, J.B. and Musumeci, S. (2010) Towards the Complete Eradication of Mother-to-Child HIV/HBV Coinfection at Saint Camille Medical Centre in Burkina Faso, Africa. *Braz J Infect Dis*, **14**, 219-224. https://doi.org/10.1590/S1413-86702010000300004