BMJ Open

Medical expenditure clustering and determinants of the annual medical expenditures of residents: a population-based retrospective study from rural China

Yan Zhang,[1,2] Shan Lu,[1,2] Yadong Niu,[1,2] Liang Zhang[1,2]

Check for updates

[1]School of Medicine and Health Management, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei, China
[2]Research Centre for Rural Health Service, Key Research Institute of Humanities & Social Sciences of Hubei Provincial Department of Education, Wuhan, Hubei, China

Correspondence to
Dr Liang Zhang;
zhangliang@mails.tjmu.edu.cn

ABSTRACT

Objective To identify the characteristics of high-cost (HC) patients and the determinants of the annual medical expenditures of Chinese rural residents.

Methods Medical expenditure clustering was performed by Lorentz curve and Gini index. T and X^2 tests were performed to identify the characteristics of the respondents, and a multilevel regression model examined the determinants of their annual medical expenditures.

Design A cluster sampling study was performed to identify those residents who availed healthcare services and to assign them to HC (top 5%), moderate-cost (top 30%) and low-cost (others) groups based on their annual medical expenditures.

Setting The annual healthcare utilisation was calculated by using data from the population-based database of the 2014 New Rural Cooperative Medical System.

Participants A total of 478 051 residents who availed healthcare services were recruited for the retrospective study in 2014. The annual medical expenditures of these residents were used as the research object.

Results The total medical expenditures of Macheng city residents for the year 2014 have a Gini index of 0.81 and around 68.01% of these expenditures can be attributed to HC patients. Female residents (51.5%) and persons aged over 60 years (34.48%) who are suffering from diseases that are difficult to diagnose have a high tendency to accumulate high medical costs. The annual medical expenditures of people living in the same village or town tend to be approximated. Age, disease category, inpatient status, healthcare utilisation and utilisation level are identified as the determinants of annual medical expenditures.

Conclusions The medical expenditures of rural residents are clustered at a remarkably high level, and HC patients are suffering from high economic burden. Therefore, policy-makers must guide these patients in seeking appropriate healthcare services and improve their management of healthcare quality to reduce the unnecessary healthcare utilisation of these patients.

Trial registration number ChiCTR-OOR-14005563.

Strengths and limitations of this study

► This study is the first to introduce the medical expenditure clustering technique, the findings of which can supplement the results of previous research on high-cost patients.
► The annual medical expenditures of residents are seldom reported at the population level. This study was conducted in Macheng, a city in Hubei province with 889 160 residents according to the New Rural Cooperative Medical System database, which also stores the inpatient and outpatient records of these residents.
► The Lorentz curve and Gini index were used to cluster the annual medical expenditure data, and a three-level linear regression model was used to aggregate these data at the residential and town levels.
► The age, gender, hospitalisation, geographical and disease data of the sample were included in the regression model. However, some individual factors were not included in the model.

BACKGROUND

The rapid increase in health expenditures greatly impedes the development of the New Rural Cooperative Medical System (NRCMS), the largest basic social health insurance system in rural China that covers 603.46 million rural residents. Specifically, the health expenditures per capita in China have increased from ¥513.8 (US$83.6) in 2012 to ¥1279.2 (208.2) in 2017 with an annual growth rate of 25.6%, which is much higher than the annual growth in fundraising per capita (16.02%).[1] Medical expenditure clustering is considered an important factor that motivates such rapid increase in health expenditures.[2]

Studies on the distribution of diseases within a population have defined clustering as the uneven distribution of disease morbidity in time or space.[3 4] In line with

this definition, medical expenditure clustering indicates the uneven distribution of the medical expenditures of a given population. In recent years, researchers have shown great interest in medical expenditure clustering and have specifically focused on high-cost (HC) patients, who are known for accumulating high annual medical expenditures[5] and comprise the top 5% biggest spenders in healthcare.[2] Previous studies have revealed that the medical expenditures of HC patients exceed those of the entire population by 50%.[6] For instance, in 2014, HC patients account for 52.3% of the total medical expenditures in the USA.[5] These patients and their healthcare quality management have attracted much research attention because of their high healthcare utilisation rate and inappropriate utilisation of healthcare services.[7] Improving the healthcare quality of these patients can discourage such inappropriate utilisation, reduce their health expenditures, conserve social health insurance funds and promote horizontal equity.

Medical expenditure clustering has also become a major concern in low-income and middle-income countries. In rural China, the rapid development of NRCMS significantly promoted the healthcare utilisation rates of residents. For instance, the annual hospitalisation rate in rural China increased from 8.7% in 2008 to 14.9% in 2017.[1] An empirical analysis of seven counties in China revealed that 78.6% of inpatient services in 2015 were distributed among one-third of all inpatients in the area. In addition, one patient in Qianjiang District used inpatient services 27 times in 2014. Another study involving 12 600 families in Jiangsu revealed that HC patients accounted for 44.9% of the total medical expenditures of the entire population.[8] Moreover, the medical expenditures clustered at the patient level are much higher than those clustered at the family level. We also suspect that China has a very high degree of medical expenditure clustering and that HC patients incur unnecessary medical expenditures because of the fragmented healthcare delivery system in rural China. Residents in rural China also seek healthcare services from a three-tier (village–town–county) healthcare delivery system where the higher tiers provide better services and charge higher medical costs.[9] Given that patients neither follow a specified order nor have limitations when seeking healthcare services and that general practitioners or consultants are unavailable in most parts of rural China, the residents of rural areas tend to make uninformed decisions when choosing among hospitals and various types of healthcare services, thereby leading to their inappropriate utilisation of healthcare services. For instance, some of these residents, especially HC patients, may be given inpatient services when they actually require outpatient services and may be unnecessarily admitted to higher-level hospitals, thereby incurring higher medical costs.[10] Yingchun revealed that the inappropriate hospital admission rate in five counties reached as high as 27.6% in 2014.[11]

To reduce the economic burden of rural residents in China, several policies and strategies, such as the Tiered Healthcare System and Serious Illness Medical Insurance (2016), have been proposed. However, the effectivity of these programmes has received mixed feedback and the focus of these policies remains unsubstantiated, which may be ascribed to the difficulty in identifying high-demand and HC patients.[12] Waxmonsky *et al* argued that those HC patients with multiple or complex conditions, behavioural disorders or socioeconomic problems are particularly difficult to monitor.[13] Moreover, only few studies have investigated the characteristics and determinants of HC patients by using population-level data because of the lack of necessary data for exploring their annual medical expenditures. Therefore, the critical values, average expenditures and inpatient service utilisation of HC patients in rural China remain unclear.

Previous studies reveal that HC patients often maintain a high level of medical expenditure for the following year. For instance, Robst found that 49.2% of patients under the Florida Medical Assistance Programme (Florida Medicaid) were continuously classified as HC patients from 2005 to 2010.[14] Meanwhile, Wodchis *et al* found that one-third of residents with public health insurance in Ontario, Canada continuously incurred high medical expenditures from 2009 to 2011.[15] Therefore, while HC patients are assumed to possess certain characteristics, the distribution of their medical expenditures remains unclear.

The growth of health expenditures must be controlled and the efficiency of insurance funds must be enhanced to improve the health insurance system of a specific area. Medical expenditure clustering can be used to monitor the cost efficiency of healthcare services. Based on the above findings, identifying HC patients is a necessary procedure and clustering the medical expenditures of a given population must be taken as the first step towards achieving such goal.

This research focuses on medical expenditure clustering in particular as well as on the distribution and characteristics of HC patients and the determinants of the annual medical expenditures of residents in general. This study also aims to guide policy-makers and health planners in predicting and planning the future needs of these patients.

METHODS
Study setting
We calculated the annual medical expenditures of residents based on the outpatient and inpatient services that they have availed within a calendar year. A population-based retrospective study was performed in Macheng, a typical rural area in Hubei and county-level city in central China (figure 1). Macheng has a total population of 889 160 and a GDP per capita of ¥22 758 (US$3704.83).

Macheng has 2 county hospitals, 22 township hospitals and 207 village clinics. The residents of this city are enrolled in NRCMS, which reimburses the medical expenditures of these residents for any type of healthcare service (eg, outpatient and inpatient services in various medical institutions, including tertiary hospitals in urban areas). The healthcare
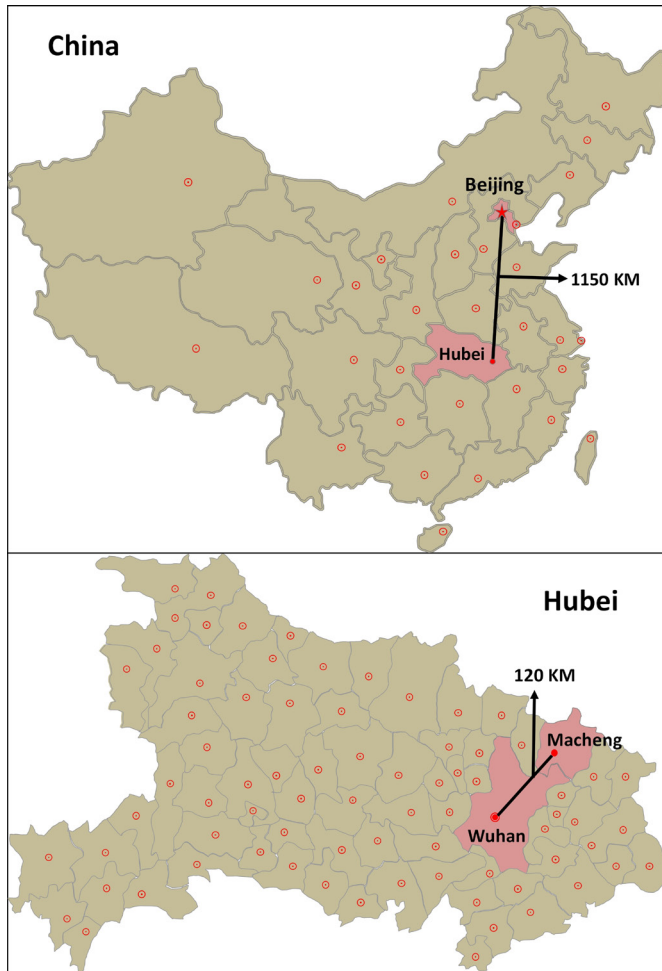
Figure 1 Map of Macheng city and geographical distribution of the residents.

utilisation information of these residents is recorded in the NRCMS database whenever they request for a reimbursement. Therefore, this study uses the 2014 Macheng city data from the NRCMS database.

Data processing

Those residents of Macheng who availed healthcare services in the past were identified through a retrospective study. After screening the data, 478 051 of the 889 160 resident records stored in the NRCMS database were included in the sample and processed by using MS Excel 2010. First, the outpatient and inpatient cases were input into a single Excel sheet. Second, those cases under the same patient identifier (ID number) were sorted chronologically. Third, the annual medical (inpatient and outpatient) expenditures of each patient, the number of outpatient and inpatient cases and other information were recorded. Fourth, the annual healthcare utilisation cases were input into a new database, where each case represents the annual healthcare utilisation of a resident. Finally, the residents were sorted in a descending order according to the annual medical expenditures stated in their records. Those residents who occupied the top 5% and 6%–30% of the sample were included in the HC and moderate-cost (MC) groups, respectively,

while the other patients were included in the low-cost (LC) group. These three groups represent the various degrees of medical expenditure clustering.

The main programming techniques employed in this paper include Excel functions (eg, COUNTIF, SUMPRODUCT, LOOKUP and IF) and case processing technologies (eg, split columns and removal of duplicates). The outstanding diseases of each resident were marked and the original International Classification of Diseases (ICD)-10 disease codes were adjusted to broader ones (eg, the disease code for chronic obstructive pulmonary disease (COPD) was adjusted from J44.900 to J44). Township hospitals were divided into four levels according to their scale and service capacity. The distance from and arrival time to county hospitals were individually captured by using Google Maps. The exchange rate of the US$ against the RMB in 2014 was 6.1428.

The land form of towns, the healthcare capacity of township hospitals and the sociological characteristics (eg, gender and age), arrival time to county hospitals, disease categories, healthcare utilisation (eg, annual length of stay (LOS)) and annual medical expenditures of the residents were collected to build the final database.[16]

Statistical analysis

First, the medical expenditures of the residents were clustered by using the Gini coefficient and Lorentz curve. The Gini coefficient is a digitised representation of medical expenditure clustering. A larger Gini coefficient corresponds to a higher degree of medical expenditure clustering. Then, the characteristics of the residents in the HC, MC and LC groups were compared by conducting t-test and χ^2 test in the IBM SPSS Statistics V.22.0 software. At last, the determinants of annual medical expenditure were then examined by conducting a linear logistic regression analysis.

Two key observations were obtained at this stage. First, the obtained data showed a hierarchical structure. Therefore, the determinants of annual medical expenditure were examined by conducting a multilevel linear logistic regression analysis using MLwiN V.2.30.[17] Second, the annual medical expenditure showed a skewed distribution as expected. Therefore, this variable was transformed to follow a normal distribution. The patient, village and town were assigned to levels 1, 2 and 3, respectively. After accepting the Log10 (x) translation, the following regression model was obtained:

$$
\begin{aligned}
Log10 \left(Annual\ medical\ expense \right) = \\
\beta_{0jk}\ cons + \beta_1 \text{Gender}_{ijk} + \beta_2\ Age_{ijk} + \beta 3\ Family_{ijk} \\
+ \beta_4\ Distance_{ijk} + \beta_5\ Time_{ijk} + \beta_6\ Capacity_{ijk} \\
+ \beta_7\ Disease_{ijk} + \beta_8\ LOS_{ijk} + \beta_9\ Coutinp_{ijk} + \beta_1 0 O\ utinp_{ijk} \\
+ \beta_1 1\ Coutinp_{ijk} + \beta_1 2\ Towinp_{ijk} beta_1 3\ Cououtp_{ijk} + \beta_1 4\ Towoutp_{ijk} \\
+ \beta_1 5\ Clioutp_{ijk} + More_{ijk} \\
\beta_{0j}k = \beta_0 + u_{0j} + w_{0jk}
\end{aligned}
$$

where β_{0jk} denotes the fixed-effects parameter while u_{oj} and w_{ojk} denote the random effects at the village and town levels, respectively.

Table 1 Medical expenditure clustering results for Macheng residents in 2014

Population ranking (%)	Critical value (¥)	Total medical expenditures (¥ 10T)	Cumulative percentage of medical expenditures (¥10T)	Cumulative percentage of medical expenditures (%)
1	19 123.63	22 634.57	22 634.57	38.73
1–5	4985.8	17 111.41	39 745.99	68.01
5–10	2021.45	7696.77	47 442.76	81.18
10–20	610.22	5179.36	52 622.12	90.04
20–30	347.29	2205.70	54 827.82	93.82
30–40	221.26	1329.19	56 157.01	96.09
40–50	147.37	867.84	57 024.85	97.58
50–60	99.54	582.81	57 607.67	98.57
60–70	63.35	384.42	57 992.09	99.23
70–80	39.95	244.01	58 236.10	99.65
80–90	20.19	141.11	58 377.20	99.89
90–100	3.20	64.17	58 441.37	100.00

Patient and public involvement

No patients or members of the public were involved in this research.

The patient information was anonymised and deidentified before the analysis.

RESULTS

Medical expenditure clustering in the sample area

Table 1 presents the clustering results for the medical expenditures of Macheng residents in 2014. Among these residents, the top 5% and 20% accounted for 68.01% and 90% of the total medical expenditures of the city, respectively. Figure 2 shows the Lorentz curve of the clustering results, which have a Gini coefficient of 0.814.

Table 2 shows the medical expenditure distribution of the HC, MC and LC groups. The HC group has an average annual expenditure per capita of over ¥15 000 (US$2441.6) and a minimum expenditure of ¥4985.80 (US$811.61) while the LC group has a maximum expenditure of ¥347.29 (US$56.54). Figure 3 presents the expenditure composition of these groups.

Characteristics of HC patients at various clustering levels

Table 3 shows the demographic characteristics of the patients. The three groups showed significant differences (p<0.001) for nine demographic items. Specifically, females accounted for 51.5% and 47.42% of the HC and LC groups, respectively. Most of the residents aged over 60 years were assigned to the HC group. The residents
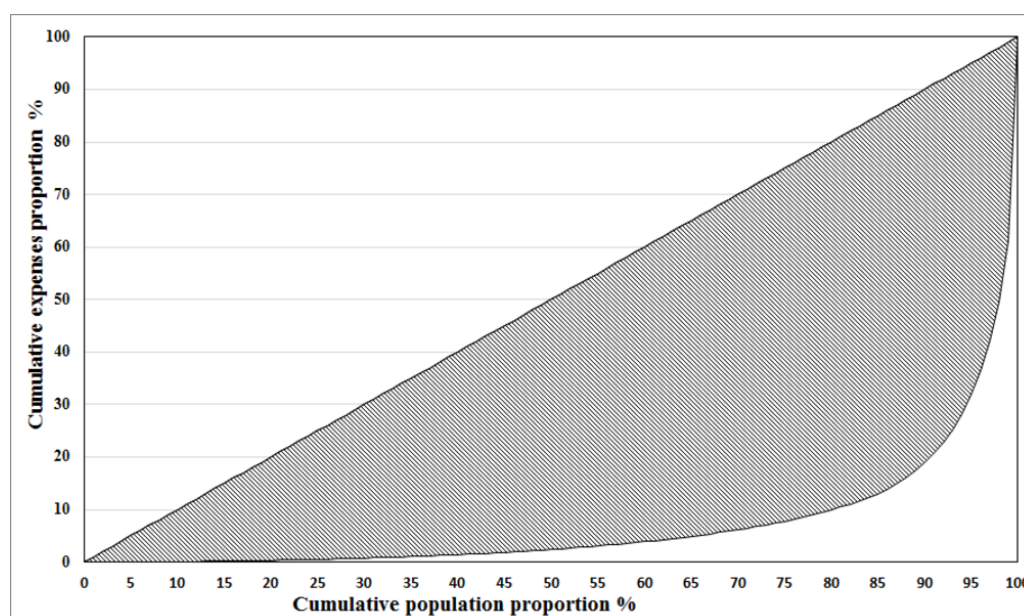

Figure 2 Lorentz curve of the medical expenditure clustering for Macheng in 2014.

Table 2 Expenditure distribution of the HC, MC and LC groups (¥)

Variable	All	Cost groups		
		HC	MC	LC
Annual expenditure per capita M (SD)	1222.49 (6253.01)	16 618.78 (22 817.95)	1261.36 (1112.90)	107.99 (89.26)
Minimum expenditure	424 962.10	424 962.10	4985.60	347.29
Maximum expenditure	3.20	4985.80	347.30	3.20
Inpatient expenditure per capita M (SD)	976.09 (6165.55)	15 866.07 (22 745.43)	728.3 (1182.59)	0.15 (6.39)
Outpatient expenditure per capita M (SD)	246.4 (559.70)	752.71 (2083.85)	533.06 (417.68)	107.84 (89.17)

HC, high cost; LC, low cost; MC, moderate cost.

in the HC and LC groups generally had small (4.01) and large (4.23) family sizes, respectively. Family size, distance from county hospitals and arrival time to county hospitals all showed the same change trends. Most of the residents living near a high-capacity township hospital were assigned to the HC group (15.19%). By contrast, 30.76% of the members in the LC group were living near a low-capacity township hospital. These three groups also showed similar distributions across varying geographical and traffic conditions. Specifically, those members of the LC group who were living in poor areas (mountains and county roads) showed high accident rates (12.54% and 32.36%, respectively). In addition, the members of the HC group were highly likely to develop cancer, circulatory, digestive and urinary diseases as well as haematological disorders while the members of the LC group often developed respiratory diseases.

Table 4 shows the distribution of the healthcare utilisation of all Macheng residents in 2014. The residents in all three groups showed a decreasing trend in their average number of inpatient cases (2.36, 0.45 and 0.02 for the HC, MC and LC groups, respectively), average number of inpatient cases and annual LOS (25.69, 3.28 and 0.1).

However, the opposite trend was observed in the annual LOS (0.07, 62.94 and 99.95) and average number of outpatient cases (8.28, 13.57 and 5.05) of those residents without prior hospitalisation experience in the HC, MC and LC groups. The healthcare utilisation of these residents also showed an inverted V-shaped distribution at the town and clinic levels. Specifically, the residents in the MC group had a high average number of outpatient cases at the town (3.69) and clinic (8.61) levels, but the residents in all three groups showed a decreasing trend in their average number of outpatient cases (1.76, 1.27 and 0.21 for the HC, MC and LC groups, respectively).

Determinants of the annual medical expenditures of residents

A three-level linear regression was performed where the patient, village and town were assigned to levels 1, 2 and 3, respectively. Table 5 displays the results for the explanatory variables that are used to fit the three-variance component models. Age, disease category and healthcare utilisation were identified as the major determinants of the annual medical expenditures of residents. No significant relationship was observed among gender, family size, distance to county hospitals, arrival time to county
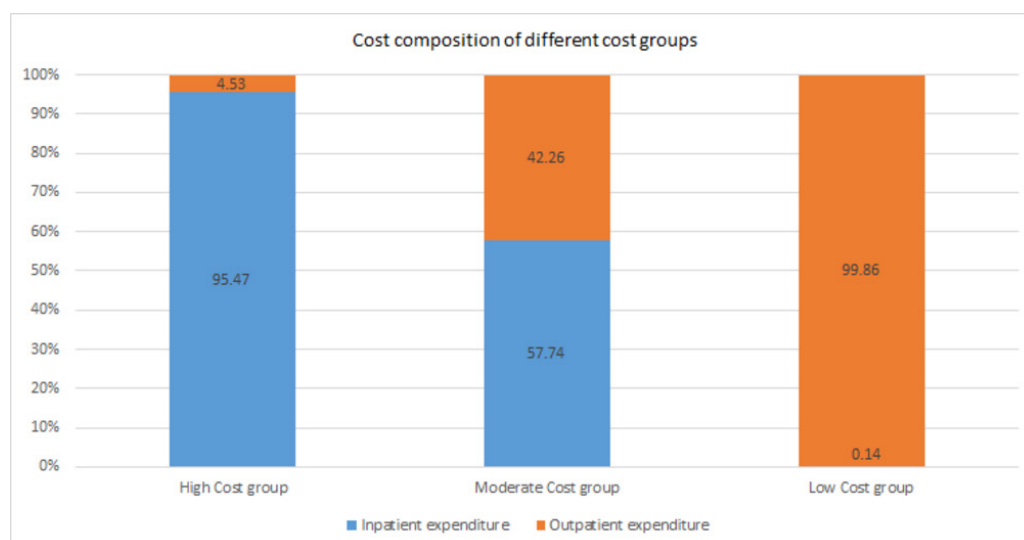

Figure 3 Cost composition of the three groups of rural residents in Macheng in 2014.

Table 3 Distribution of the demographic characteristics of residents in the three cost groups (n=478 051)

Variable	All n (%)	Cost groups n (%)			T/F values
		HC	MC	LC	
All	478 051	23 922 (5.00)	119 492 (25.00)	334 637 (70.00)	
Gender					
Male	242 244 (50.67)	11 602 (48.5)	54 704 (45.78)	175 938 (52.58)	1673.95*
Female	235 804 (49.33)	12 320 (51.5)	64 786 (54.22)	158 698 (47.42)	
Age (years)					
Mean (SD)	41.97 (20.11)	49.45 (18.47)	45.56 (20.14)	40.14 (19.92)	4739.48†
Less than 20	77 706 (16.4)	1374 (5.75)	15 218 (12.79)	61 114 (18.47)	10 792.85*
20–39	114 619 (24.19)	5142 (21.51)	23 166 (19.47)	86 311 (26.08)	
40–59	181 426 (38.29)	9147 (38.27)	48 561 (40.81)	123 718 (37.39)	
60–79	93 338 (19.7)	7753 (32.44)	29 926 (25.15)	55 659 (16.82)	
More than 79	6706 (1.42)	487 (2.04)	2118 (1.78)	4101 (1.24)	
Family size					
Mean (SD)	4.2 (1.54)	4.01 (1.62)	4.15 (1.59)	4.23 (1.51)	350.63†
Distance from CH (km)					
Mean (SD)	30.7 (15.15)	29.34 (15.48)	30.16 (15.44)	31 (15.01)	234.87†
Arrival time to CH (min)					
Mean (SD)	46.5 (18.03)	44.88 (18.6)	45.95 (18.64)	46.81 (17.76)	200.56†
Capacity of TH					
First level (strong)	56 866 (11.9)	3633 (15.19)	15 868 (13.28)	37 365 (11.17)	690.05*
Second level (general)	276 701 (57.88)	13 304 (55.61)	69 057 (57.79)	194 340 (58.07)	
Third level (weak)	144 484 (30.22)	6985 (29.2)	34 567 (28.93)	102 932 (30.76)	
Geography					
Plain	175 767 (36.77)	9753 (40.77)	47 877 (40.07)	118 137 (35.3)	2701.41*
Hilly	117 197 (24.52)	5792 (24.21)	31 915 (26.71)	79 490 (23.75)	
Hilly and mountainous	126 341 (26.43)	5641 (23.58)	25 665 (21.48)	95 035 (28.4)	
Mountainous	58 746 (12.29)	2736 (11.44)	14 035 (11.75)	41 975 (12.54)	
Traffic condition					
National standard roads	214 571 (44.88)	10 499 (43.89)	53 544 (44.81)	150 528 (44.98)	885.97*
Provincial standard roads	113 392 (23.72)	6561 (27.43)	31 008 (25.95)	75 823 (22.66)	
County roads	150 088 (31.4)	6862 (28.68)	34 940 (29.24)	108 286 (32.36)	
Disease category					
Cancer	4665 (0.98)	2686 (11.23)	1452 (1.22)	527 (0.16)	87 301.73*
ENT disease	41 772 (8.74)	2475 (10.35)	7114 (5.95)	32 183 (9.62)	
Respiratory disease	181 929 (38.06)	3312 (13.84)	30 142 (25.23)	148 475 (44.37)	
Circulatory disease	24 721 (5.17)	3781 (15.81)	9701 (8.12)	11 239 (3.36)	
Digestive disease	45 723 (9.56)	2838 (11.86)	9539 (7.98)	33 346 (9.96)	
Urinary disease	8048 (1.68)	1360 (5.69)	3364 (2.82)	3324 (0.99)	
Endocrinology disease	5785 (1.21)	133 (0.56)	792 (0.66)	4860 (1.45)	
Haematological disorders	2995 (0.63)	427 (1.78)	739 (0.62)	1829 (0.55)	
Bones and muscles	41 675 (8.72)	2988 (12.49)	11 993 (10.04)	26 694 (7.98)	
Obstetrics and gynaecology	15 844 (3.31)	2223 (9.29)	6226 (5.21)	7395 (2.21)	
Others	104 894 (21.95)	1699 (7.1)	38 430 (32.16)	64 765 (19.36)	

*Pearson's χ^2 test.
†ANOVA.
ANOVA, analysis of variance; CH, county hospital; HC, high cost; LC, low cost; MC, moderate cost; TH, township hospital.

Table 4 Distribution of the healthcare utilisation characteristics of the three groups (n=478 051)

Variable	All n (%)	Cost groups n (%)/mean (SD)			F values
		HC	MC	LC	
Outpatient no					
Mean (SD)	7.34 (8.26)	8.28 (9.1)	13.57 (11.63)	5.05 (4.89)	58 639.94*
Outpatient level (time)					
County	0.55 (1.51)	1.76 (2.85)	1.27 (2.26)	0.21 (0.66)	34 430.07*
Township	1.96 (4.14)	2.16 (4.5)	3.69 (6.48)	1.32 (2.57)	15 398.27*
Clinic	4.83 (7.71)	4.36 (7.72)	8.61 (11.78)	3.52 (4.94)	20 854.66*
Inpatient no					
Mean (SD)	0.23 (1.21)	2.36 (4.33)	0.45 (1)	0 (0.02)	55 409.19*
0	409 682 (85.70)	17 (0.07)	75 207 (62.94)	334 458 (99.95)	274 894.92†
1–3	64 955 (13.59)	20 840 (87.12)	43 936 (36.77)	179 (0.05)	
4–7	2736 (0.57)	2447 (10.23)	289 (0.24)	0 (0)	
More than 8	678 (0.14)	618 (2.58)	60 (0.05)	0 (0)	
Inpatient level (time)					
Outside the county	0.04 (0.38)	0.58 (1.53)	0.04 (0.22)	0.01 (0.01)	28 920.75*
County	0.14 (1.07)	1.49 (4.09)	0.25 (0.90)	0.01 (0.01)	24 788.76*
Township	0.06 (0.32)	0.29 (0.96)	0.16 (0.45)	0.01 (0.02)	18 940.21*
Annual LOS					
Mean (SD)	2.09 (10.66)	25.46 (36.69)	3.28 (7.91)	0.01 (0.1)	88 485.86*

*Pearson's χ^2 test.
†ANOVA.
ANOVA, analysis of variance; HC, high cost; LC, low cost; LOS, length of stay; MC, moderate cost.

hospitals, geography, traffic condition and capacity of township hospitals. The medical expenditures of these residents increased along with age, number of outpatient cases, LOS and healthcare utilisation level while the other factors were kept constant. Those patients with cancer, circulatory, digestive, urinary, obstetric and gynaecological diseases or haematological disorders had a higher probability to incur high annual medical expenditures compared with those patients with respiratory, ENT, endocrinal, skeletal and muscular diseases.

DISCUSSION
Clustering of the medical expenditures of residents in rural China

The medical expenditures of rural residents showed an extremely uneven distribution. A Gini coefficient of 0.814 was recorded in 2014, which indicated that the medical expenditures of this population was clustered at the minority level. The HC group accounted for 68.01% of the total medical expenditures, which was much higher than that recorded in the USA (52.3%).[5] Meanwhile, the LC group, which included 70% of the residents in the sample, accounted for 2.42% of the total medical expenditures. The annual medical expenditure per capita of the entire population was ¥1222.49 (US$199.01), which was nearly similar to that

of the MC group (¥1261.36, US$205.18). However, this value was only 7.35% of the annual medical expenditure per capita of the HC group. The maximum annual medical expenditure of this population was ¥424 962.1 (US$69 180.52), which is more than four times larger than the annual reimbursements provided by NRCMS (¥100 000, US$16 279.22).

The three groups showed obvious differences in their expenditure structures. Hospitalisation expenditures accounted for over 95% of the medical expenditures of the HC group, thereby supporting the findings of Driessen et al.[18] Outpatient expenditures accounted for most of the medical expenditures of the LC group while outpatient and hospitalisation expenditures equally accounted for the medical expenditures of the MC group. The medical expenditures of the HC group showed a dispersed distribution and an SD that was much higher than the mean (¥22 817.95 vs ¥16 618.78, US$3714.58 vs US$2705.41). By contrast, the medical expenditures of the MC and LC groups showed a relatively concentrated distribution.

The HC group faces a very high economic burden. In fact, those residents with medical expenditures of over ¥5000 (US$813.96) were included in this group. This group had an annual medical expenditure per capita of over ¥16 000 (US$2604.68), of which only

Table 5 Three-level linear regression model analysis of annual medical expenditures (n=478 051)

	Parameter estimate	SE	χ^2	P values
Fixed factors				
Constant	1.842	0.063	862.479	<0.001
Gender (baseline:male)				
Female	0.003	0.002	5.051	0.025
Age (years)	0.003	0.001	4690.838	<0.001
Family size (person)	0.001	0.001	1.536	0.215
Distance to CH (km)	−0.001	0.001	0.276	0.599
Time to CH (min)	0.001	0.001	0.006	0.938
Capacity of TH	−0.073	0.030	5.797	0.016
Geography (baseline:plains)	0.013	0.020	0.423	0.515
Traffic condition (baseline:national)	0.004	0.022	0.036	0.849
Disease category (baseline:respiratory)				
Cancer	0.759	0.008	8286.833	<0.001
ENT disease	−0.095	0.003	1084.129	<0.001
Circulatory disease	0.303	0.004	4948.426	<0.001
Digestive disease	0.024	0.004	11.12	0.001
Urinary disease	0.304	0.006	2197.804	<0.001
Endocrinology disease	−0.070	0.007	6.756	0.009
Haematological disorders	0.113	0.010	127.488	<0.001
Bones and muscles	0.030	0.008	10.168	0.001
Obstetrics and gynaecology	0.390	0.005	6056.819	<0.001
Else	0.126	0.003	0.294	0.588
Annual length of stay (days)	0.015	0.001	18 287.060	<0.001
Inpatient level (times)				
Outside the county	0.275	0.002	15 149.683	<0.001
County	0.064	0.001	5024.366	<0.001
Township	0.260	0.003	8241.838	<0.001
Outpatient level (times)				
County	0.134	0.001	62 795.795	<0.001
Township	0.049	0.001	56 609.642	<0.001
Clinic	0.031	0.001	82 166.996	<0.001
Random factors				
Town variance	0.005	0.002	9.114	0.003
Village variance	0.007	0.000	298.486	<0.001
Patient scale parameter	1	0.00	–	–

around 50% were reimbursed by NRCMS. Therefore, HC patients had an average out-of-pocket expenditures over ¥8000 (US$1302.34), which was nearly 80% of the total consumer spending per capita of rural residents (¥10 129.8, US$1649.05).[1] This ratio can easily lead to a catastrophic health spending.

Aggregating annual medical expenditures at the village and town levels

The multilevel linear regression revealed a hierarchical structure (town–village–residents) in the collected data.

The annual medical expenditures of residents were clustered at the town and village levels while the annual medical expenditures of residents living in the same village and town tend to be approximated. The distribution of annual medical expenditures as shown in table 3 significantly differs from that shown in table 5 in terms of geography, traffic condition and capacity of township hospitals. Such differences were mostly observed at the town level because the residents were living in towns with the same geography, traffic conditions and capacity of township hospitals. Meanwhile, 30 towns in the sample

showed differences in their social customs, geographical locations and capacity of township hospitals. These findings altogether show that those residents living in areas with favourable geographical conditions (eg, plains and national standard roads) and high-capacity hospitals tend to accumulate high medical expenditures. Similarly, the effects of distance from and arrival time to county hospitals were mostly observed at the village level. Given their convenient locations, those residents who are living near county hospitals have a higher probability to avail healthcare services in these institutions and accumulate higher medical expenditures compared with those who are living far from these hospitals.

Determinants of annual medical expenditures

Annual medical expenditures are directly determined by healthcare utilisation. Table 5 shows that a higher utilisation of healthcare services corresponds to higher medical expenditures. The village, town and county levels had regression coefficients of 0.031, 0.049 and 0.134 for outpatient cases, which were lower than the corresponding coefficients for inpatient cases. Similar to hospital utilisation, an increase in LOS corresponds to an increase in medical expenditures.

In addition, a higher age corresponds to a higher expenditure. This case is particularly true for those residents aged above 60 years, who account for 34.48% of HC group, in other words, 8.5% of the elderly population was defined as HC, while only 5% of total population was defined as HC. Such high expenditures can be attributed to their poor physical condition and high tendency to develop one or multiple diseases. A WHO report revealed that an ageing population would increase the healthcare expenditures of a country, but the extent of such increase is less than expected. Although health-related needs often increase along with age, the relationship between healthcare utilisation and health expenditure varies as a person grows older.[19] In fact, the medical expenditures of people from high-income countries gradually decline after they reach the age of 75. Other studies even show that those people aged 80 years and above have a much lower share in the consumption of medical resources compared with the total population. In terms of gender, the average medical expenditures of female residents do not differ from those of male residents. The relationship between gender and medical expenditures also warrants further study.[20]

Those patients with respiratory, urinary, endocrinology, skeletal and muscular diseases have the same degree of medical expenditure clustering. By contrast, those patients with cancer, circulatory, digestive, obstetrical and gynaecological diseases or haematological disorders have significantly high annual medical expenditures, with patients with cancer having the highest regression coefficient of 0.759. Developing diseases that are fatal and difficult to assess can easily lead to high medical expenditures.[21] These findings can be attributed to the limited ability of healthcare professionals at townships, most of whom lack specialisation in treating urinary and cardiovascular diseases or haematological disorders; therefore, those residents suffering from such diseases tend to be admitted to county hospitals or hospitals outside their respective counties.[22]

Redesigning the healthcare delivery system for the HC group

The intense concentration of medical expenditures reveals a great imbalance in the healthcare utilisation of Macheng city residents. Therefore, the rural healthcare delivery system as well as the current efforts in reducing the costs and promoting the cost efficiency of healthcare services should focus on the HC patients.

First, a monitoring mechanism for HC patients must be established and the NRCMS database can be used as a source of information that can facilitate the monitoring of these patients. Although HC patients are identified based on their medical expenditure or health claims, a patient can be predicted as HC high probability in advance based on several risk factors. Robst and Wodchis *et al* found that a patient identified as HC in a year is more than 40% likely to be identified as an HC patient in the following year given that these patients often maintain a high level of medical expenditure for the following year.[14 15] Therefore, those residents with a remarkably high healthcare utilisation are exposed to many risk factors and have been identified as HC patients in the previous year warrant special attention. Moreover, HC patients need guidance when choosing healthcare services and must be given priority access to primary healthcare.

Second, enhancing the healthcare quality management of patients plays a vital role in discouraging their unnecessary utilisation of healthcare services, such as inappropriate admission to hospitals and excessive utilisation of inpatient services.[23]

Third, exploring new mechanisms, such as the comprehensive management of patients with chronic diseases, setting a global budget for multilevel institutions, integrated management programmes and integrated delivery of medicine and nursing services to aged residents, can motivate doctors to deliver continued care to HC patients. In doing so, healthcare professionals can avoid unnecessary service duplications, improve their healthcare efficacy and help their patients take advantage of primary healthcare services.[24]

CONCLUSION

From the city-level population perspective, the medical expenditures of rural residents in China have an intense clustering level. Apart from demographic characteristics (eg, age and disease), healthcare utilisation was identified as a primary determinant of medical expenditure clustering. Therefore, policy-makers must guide HC patients in choosing healthcare services and improve their healthcare quality management to discourage their unnecessary utilisation of healthcare services. Doctors must also

be motivated to deliver continued care to this group of patients.

LIMITATION

This study has two limitations. First, hospitalisation information, geographical factors, referral status and diseases were included in the regression model while other individual factors, such as economic status, education and preference, were ignored. Second, several studies show that HC residents with a high tendency to develop multiple chronic diseases tend to use multidisciplinary and multi-institutional services.[25] However, this work only considered the main diseases of these patients as captured in the NRCMS database. These limitations may affect the stability of the findings and should be examined in further studies.

REFERENCES

1. NHaFPCoC. *China statistics yearbook of health and family planning*. Beijing: China union medical university press, 2017.
2. Coughlin TA, Long SK. Health care spending and service use among high-cost Medicaid beneficiaries, 2002-2004. *Inquiry* 2010;46:405–17.
3. Murakami Y, Okamura T, Nakamura K, *et al*. The clustering of cardiovascular disease risk factors and their impacts on annual medical expenditure in Japan: community-based cost analysis using gamma regression models. *BMJ Open* 2013;3:e002234.
4. Leech RM, McNaughton SA, Timperio A. The clustering of diet, physical activity and sedentary behavior in children and adolescents: a review. *Int J Behav Nutr Phys Act* 2014;11:4.
5. Zulman DM, Pal Chee C, Wagner TH, *et al*. Multimorbidity and healthcare utilisation among high-cost patients in the US veterans affairs health care system. *BMJ Open* 2015;5:e007771.
6. Zulman DM, Ezeji-Okoye SC, Shaw JG, *et al*. Partnered research in healthcare delivery redesign for high-need, high-cost patients: development and feasibility of an intensive management patient-aligned care team (ImPACT). *J Gen Intern Med* 2014;29:861–9.
7. Hong CS, Siegel AL, Ferris TG. Caring for high-need, high-cost patients: what makes for a successful care management program? *Issue Brief* 2014;19:1–19.
8. Miao Y, Qian D, Sandeep S, *et al*. Exploring the characteristics of the high-cost population from the family perspective: a cross-sectional study in Jiangsu Province, China. *BMJ Open* 2017;7:e17185.
9. Li X, Lu J, Hu S, *et al*. The primary health-care system in China. *Lancet* 2017;390:2584–94.
10. Yip W, Hsiao W. Harnessing the privatisation of China's fragmented health-care delivery. *Lancet* 2014;384:805–18.
11. Zhang Y, Chen Y, Zhang X, *et al*. Current level and determinants of inappropriate admissions to township hospitals under the new rural cooperative medical system in China: a cross-sectional study. *BMC Health Serv Res* 2014;14:649.
12. Bleich SN. Systematic review of programs treating high-need and high-cost people with multiple chronic diseases or disabilities in the United States, 2008-2014 (vol 12, E197, 2015). *Preventing Chronic Disease* 2016;13.
13. Waxmonsky JA, Giese AA, McGinnis GF, *et al*. Colorado access' enhanced care management for high-cost, high-need medicaid members: preliminary outcomes and lessons learned. *J Ambul Care Manage* 2011;34:183–91.
14. Robst J. Developing models to predict persistent high-cost cases in Florida medicaid. *Popul Health Manag* 2015;18:467–76.
15. Wodchis WP, Austin PC, Henry DA. A 3-year study of high-cost users of health care. *CMAJ* 2016;188:182–8.
16. Herrin J, St Andre J, Kenward K, *et al*. Community factors and hospital readmission rates. *Health Serv Res* 2015;50:20–39.
17. Longford NT. Letter to the editors: improved approximations for multilevel models with binary responses. *J R Stat Soc Ser A Stat Soc* 1997;160:593.
18. Driessen RJ, Bisschops LA, Adang EM, *et al*. The economic impact of high-need psoriasis in daily clinical practice before and after the introduction of biologics. *Br J Dermatol* 2010;162:1324–9.
19. Beard JR, Officer AM, Cassels AK. The World Report on Ageing and Health. *Gerontologist* 2016;56:S163–S166.
20. Owens GM. Gender differences in health care expenditures, resource utilization, and quality of care. *J Manag Care Pharm* 2008;14:2–6.
21. Ross JS, Chen J, Lin Z, *et al*. Recent national trends in readmission rates after heart failure hospitalization. *Circ Heart Fail* 2010;3:97–103.
22. Hartl S, Lopez-Campos JL, Pozo-Rodriguez F, *et al*. Risk of death and readmission of hospital-admitted COPD exacerbations: European COPD Audit. *Eur Respir J* 2016;47:113–21.
23. Feder JL. Predictive modeling and team care for high-need patients at healthcare partners. *Health Aff* 2011;30:416–8.
24. Ye T, Sun X, Tang W, *et al*. Effect of continuity of care on health-related quality of life in adult patients with hypertension: a cohort study in China. *BMC Health Serv Res* 2016;16.
25. Zulman D, Yoon J, Cohen DM, *et al*. Multimorbidity and health care utilization among high-cost patients: implications for care coordination. *J Gen Intern Med* 2013;281:S123–S124.