

The Concentration and Persistence of Medical Expenses among Middle-Aged and Elderly Adults: Findings from the China Health and Retirement Longitudinal Survey

Luyan Jiang

Nanjing Medical University

Qianqian Qiu

Nanjing Medical University

Lin Zhu

Nanjing Medical University

Zhonghua Wang (✉ wzh04@njmu.edu.cn)

Nanjing Medical University

Research Article

Keywords: Medical expenses, concentration, persistence, middle-aged and elderly, China

Posted Date: April 27th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1588484/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

1 **The Concentration and Persistence of Medical Expenses among Middle-Aged**
2 **and Elderly Adults: Findings from the China Health and Retirement**
3 **Longitudinal Survey**

4 **Authors and affiliations:**

5 Luyan Jiang¹, Qianqian Qiu¹, Lin Zhu¹, Zhonghua Wang^{1,2,3*}

6 ¹ School of Health Policy & Management, Nanjing Medical University, 211166 Nanjing, China

7 ² Public Health Policy and Management Innovation Research Group, Nanjing Medical University,
8 211166 Nanjing, China

9 ³ Center for Global Health, Nanjing Medical University, 211166 Nanjing, China

10 **Email address of each author:**

11 Luyan Jiang: jiangluyan_njmu@163.com

12 Qianqian Qiu: m15720800939@163.com

13 Lin Zhu: zhulin980@163.com

14 Zhonghua Wang: wzh04@njmu.edu.cn

15 Correspondence to Zhonghua Wang (email: wzh04@njmu.edu.cn)

16 *** Corresponding author:** Zhonghua Wang, School of Health Policy & Management, Nanjing
17 Medical University, 101Longmian Avenue, Jiangning District, Nanjing 211166, P.R. China. Email:
18 wzh04@njmu.edu.cn. Telephone and fax number: +81 025 86868258

19

20

1 **Abstract**

2 **Background:** Medical expenses, especially among middle-aged and elderly people, have increased
3 in China over recent decades. While the concentration and persistence of medical expenses have
4 attracted great attention, few studies have focused on Chinese residents or vulnerable groups. This
5 study sought to identify characteristics associated with the concentration of medical expenses and
6 estimate their persistence among Chinese middle-aged and elderly adults.

7 **Methods:** Data was obtained from the China Health and Retirement Longitudinal Study, a
8 nationally representative household survey of middle-aged and elderly adults that was conducted by
9 the National School of Development at Peking University in 2013, 2015, and 2018. Concentration
10 was measured using the cumulative percentages of ranked annual medical expenses and descriptive
11 statistics were used to define the characteristics of individuals with high medical expenses. The
12 persistence of medical expenses and associated factors were estimated using transfer rate
13 calculations and Heckman selection modeling.

14 **Results:** Total medical expenses were concentrated among a few middle-aged and elderly adults
15 and the concentration increased over time. Compared with middle-aged and elderly adults in the
16 lower 90% medical expense group, those in the top 10% high medical expense group were more
17 likely to be older, live in urban areas, be less wealthy, have chronic diseases, and attend higher-
18 ranking medical institutions. High medical expenses demonstrated strong persistence. Lagged
19 medical expenses had a persistent positive effect on the incidence and volume of current medical
20 expenses and the effect of a one-period lag was strongest. Individuals with chronic diseases during
21 the lagged period had a higher likelihood of experiencing persistent medical expenses. Factors such
22 as age ≥ 75 years, income, health insurance, comorbidities, and receiving a higher number of

1 medical services were significant positive contributors to higher medical expenses.

2 **Conclusion:** Policy efforts should focus on preventive management and optimizing the allocation
3 of health resources to reduce high medical expenses and inequitable access to health services among
4 middle-aged and elderly individuals in China. The government should establish a more efficient
5 care system and adjust social health insurance policies to alleviate the long-term financial healthcare
6 burden.

7 **Keywords:** Medical expenses, concentration, persistence, middle-aged and elderly, China.

8 **Background**

9 The concentration and persistence of medical expenses are well documented [1-3]. Health care
10 spending is overwhelmingly concentrated within a very small proportion of the population defined
11 as high-cost users (HCUs) [4]. The persistence of medical expenditures generally refers to ‘long-
12 term medical spending patterns’ [3]. However, the concentration and persistence of medical costs
13 have primarily been confirmed in the United States. Monheit et al. [5] found that 1% of patients
14 account for 27% of annual medical spending in the US and Riley et al. [6] reported that 5% accounts
15 for 34.4% of total health expenses. Kohn et al. [7] revealed that medical costs have significant
16 persistence while Newhouse et al. [8] reported that 40% of current medical expenses are associated
17 with past medical expenditures. Persistent high medical expenses have also been reported in many
18 developed high-income countries, such as Germany, Canada, the Netherlands, and Australia [9-12].
19 Studies in developed Asian regions such as Singapore and Japan have reported that identifying the
20 characteristics of high medical cost beneficiaries and long-term medical spending trends helps to
21 reduce medical expenses when governments are faced with tighter medical budgets [13,14].

22 Studies have explored the persistence of medical expenses since the 1980s and have found that

1 they can be anticipated by extended longitudinal profiles [8,15]. Other studies have shown that a
2 relatively small number of spenders with persistent high medical costs account for about 20% of the
3 medical expenses [2,5,12]. Researches have also focused on the impact of factors such as chronic
4 diseases, demographic indicators, medical service utilization and insurance, on the persistence and
5 concentration of medical expenses [16,17]. Age, poor health status, and comorbidities are strong
6 predictors of medical expense persistence [14,18], and researchers have highlighted the importance
7 of early intervention to curb high costs [19,20].

8 China is the largest developing country in the world. With improved medical technology and the
9 changing spectrum of disease, medical expenses have risen significantly in recent years. Since the
10 new round of healthcare-system reform in 2009, China's social medical insurance system has
11 gradually improved with urban employee and urban and rural resident medical insurance now
12 covering almost all citizens. The number of individuals with social medical insurance has reached
13 1.36 billion, the participation rate is more than 95% and universal coverage has basically been
14 realized in China. The financing and reimbursement of social medical insurance have also improved,
15 however, gaps in the level of medical security between urban and rural areas or other geographic
16 regions remain. Thus, the wide coverage and increasing compensation of social medical insurance
17 has resulted in higher medical expenses and increased healthcare utilization and cost inequalities.
18 Peng et al. [21] indicated that nearly 50% of the medical expenses in eastern China were paid by
19 10% of patients, and Ku L-JE et al. [22] demonstrated that being male, older, divorced, or having
20 chronic diseases was associated with high medical spending. Xu et al. [23] also found that prior
21 medical expenses had a positive persistent effect on current medical expenditures.

22 While many studies have examined the concentration and persistence of medical expenditures,

1 most focus on general populations, and few have specifically assessed China or vulnerable groups.
2 Middle-aged and elderly people are more likely to suffer from chronic diseases and account for a
3 larger share of total medical expenditures. Thus, it is particularly important to study the
4 concentration and persistence of medical expenses among this at-risk group. To our knowledge, no
5 study has analyzed the concentration or persistence of medical expenses among middle-aged and
6 elderly adults in China. The current study aims to identify the characteristics of this population with
7 high medical spending and establish dynamic regression models for medical expenses to evaluate
8 the persistence and associated factors of medical costs. Findings could provide information about
9 the concentration and persistence of medical expenses and help to alleviate social medical insurance
10 spending.

11 **Methods**

12 **Data Source**

13 Data were obtained from the China Health and Retirement Longitudinal Survey (CHARLS), a
14 nationally representative survey of elderly household residents, conducted in 2013, 2015, and 2018
15 (<https://charls.charlsdata.com/pages/Data/2013-charls-wave2/zh-cn>;
16 <https://charls.charlsdata.com/pages/Data/2015-charls-wave4/zh-cn>;
17 <https://charls.charlsdata.com/pages/Data/2018-charls-wave4/zh-cn>). This survey, which covered
18 150 counties in 28 provinces, was used to construct a high-quality, nationally representative sample
19 of Chinese household-dwelling adults aged ≥ 45 years for scientific research.

20 The survey used multi-stage stratified random sampling involving four steps. First, 150 county-
21 level units covering 28 of the 30 provinces in mainland China, excluding Tibet, were sampled based
22 on population size and stratified by GDP and district (urban or rural) using probability proportionate

1 to size (PPS) sampling. Second, using the latest population data, village and community units within
2 counties were chosen by referencing the National Bureau of Statistics. The administrative villages
3 in rural areas and communities in urban areas were used as primary sampling units (PSUs). Three
4 PSUs were selected within each county, and 450 PSUs were selected using the PPS. Third,
5 approximately 25 household units were selected in each PSU based on sampling frames constructed
6 using CHARLS GIS maps, considering the pre-investigation refusal rate. Fourth, one or two age-
7 eligible respondents in each household were administered the survey and followed up every 2–3
8 years.

9 The CHARLS collects detailed information about each respondent and their spouse, including
10 basic demographics, chronic disease status, health status and behaviors, health care utilization and
11 insurance, income and retirement, and medical expenditures. Medical expenditure data includes
12 outpatient, hospitalization, and self-treatment expenses and total expenditures included the sum of
13 all three spending categories. Based on CHARLS responses, nine chronic diseases were chosen
14 based on their high prevalence or mortality rate in 2020 chronic disease status reports across China:
15 hypertension, diabetes (i.e., high blood sugar), cancer or malignant tumors (except minor skin
16 cancers), chronic lung disease (i.e., chronic bronchitis, emphysema), liver disease (except fatty liver,
17 tumors or cancer), heart disease (i.e., coronary heart disease, angina, congestive heart failure, and
18 other heart problems), kidney disease (except tumors or cancer), stomach and other digestive
19 diseases (except tumors or cancer) and arthritis or rheumatism. Cross-sectional data from 2013,
20 2015, and 2018 was matched to a three-period dynamic panel dataset using ID, restricting the sample
21 to individuals with data in each year. After dropping patients who had migrated or died, had negative
22 expenditures, or had incomplete information, 19,869 observations (6,623 samples) were included in

1 the final dataset.

2 **Data Analysis**

3 **Determining the concentration of medical expenses among middle-aged and elderly** 4 **participants**

5 Consistent with other surveys, most CHARLS survey respondents (47.2%) belonged to the “zero
6 medical expenses” group. Participants with positive medical expenses were ranked from highest to
7 lowest each year and divided into categories representing the top 1%, 10%, and 20%, and the bottom
8 50% of expenses. The medical expenses of each category were divided by the annual medical
9 expenditures to estimate the cumulative percentage of each expense category and used to determine
10 the concentration of total medical expenses as well as hospitalization, outpatient, and self-treatment-
11 specific expenses from 2013 to 2018.

12 **Determining the persistence of high medical expenses among middle-aged and elderly** 13 **participants**

14 The persistence of high medical expenditures was determined by calculating the transfer rate of high
15 medical expense samples. Rates were estimated by determining the number of spenders in the top
16 10%, 20% and 50% categories in 2013 that transferred into other categories or remained in the
17 original categories in 2015 and 2018 and then dividing these values by the number of spenders in
18 the top 10%, 20% and 50% categories in 2013. The top 1% category was excluded because its small
19 sample size made it difficult to obtain scientifically representative transfer rates.

20 **Descriptive Analysis**

21 Descriptive statistics were used to analyze the characteristics of middle-aged and elderly people
22 with high medical costs. While the top 1%, 10% and 50% categories were used to estimate the

1 concentration of medical expenses, individuals in the top 10% and 20% of annual total medical
 2 expenses were used for the descriptive analysis because the top 1% group was too small to calculate
 3 reliable rates for specific characteristics. Participants with the top 10% of total medical costs in 2013,
 4 2015 and 2018 were combined as the “Top 10% high medical expense group” while those with the
 5 bottom 90% of total medical expenses in 2013, 2015 to 2018 were combined as the “Bottom 90%
 6 medical expense group” and defined as the control group. Similarly, a “Top 20% high medical
 7 expense group” and “Bottom 80% medical expense group” were also created using combined data
 8 from each of the survey years.

9 **Regression analysis of medical expenses among middle-aged and elderly participants**

10 As a result of the large proportion of participants in the “zero medical expenses” group and the
 11 “long-tail distribution” of positive medical costs, the Heckman selection model was used to identify
 12 the contribution of related variables to medical spending incidence. Lagged medical expenditures
 13 were used as the independent variable to analyze the persistence of medical expenses. The analysis
 14 included a regression of medical expense incidence and a regression of total medical expenditures.

15 The first regression used the probit model for estimation as follows:

$$16 \quad I_{it} = \alpha + X_{it}\beta + \sum_k \gamma_{t-k} I_{i,t-k} + \sum_k \sum_j \lambda_{j,t-k} I_{i,t-k} D_{ij,t-k} + U_{it}(1)$$

17 In this model, the dependent variable, “incidence of medical costs” (I_{it}), is a binary variable
 18 defined as “whether an individual i had medical expenses in year t ” and dichotomized as “0”=“Non-
 19 incurrence of medical expenses” or “1”= “Incurrence of medical expenses.” $I_{i,t-k}$ denotes the lag
 20 term of “whether medical expenses occurred.” X_{it} are a series of control variables representing the
 21 demographic factors, age, sex, education level, marital status, retirement condition, and residence.
 22 $D_{ij,t-k}$ is a dummy indicator representing “whether individual i suffered from disease j in lagged

1 k period.” $I_{i,t-k}D_{ij,t-k}$ are cross-variables indicating “whether medical expenses incurred in the
 2 lag period” interacted with “the prevalence of diseases i in the lag k stage.” In the current paper,
 3 $k=1,2$, or k here does not refer to the specific value of the year. Instead, $t-1$ denotes lagged one
 4 period while $t-2$ denotes lagged two periods. U_{it} is the error term, with a normal distribution, and
 5 γ and λ are the parameters measured in this study. While γ assesses the persistence of incurring
 6 medical expenses, λ represents the effect of diseases in the lagged period on the incidence of
 7 current medical expenses.

8 The second model represented the total medical expenditures and OLS regression was used, as
 9 follows:

$$10 \quad \ln(Y_{it}|I_{it} > 0) = \delta + Z_{it}\phi + \sum_k \eta_{t-k} \ln(1 + Y_{i,t-k}) + \varepsilon_{it} \quad (2)$$

11 In this model, the dependent variable, Y_{it} , is the total medical expenditure of individual i in year
 12 t, and was made logarithmic to normalize the skewed medical expense data. The key independent
 13 variables, $Y_{i,t-k}$, are “the lagged term of medical expenses for each individual i over a certain
 14 number of years.” For example, if $Y_{i,t-k}$ includes $Y_{i,2015}$ and $Y_{i,2013}$ for 2018 then $k=2$.
 15 According to Peng et al. [21], $\ln(1 + Y_{i,t-k})$ is the logarithm of $Y_{i,t-k}$. If $Y_{i,t-k} = 0$, 1 should
 16 be added when transforming into a logarithm, and if $Y_{i,t-k}$ is not equal to 0, the log of the original
 17 lagged expenditure, k , should be used to define the equation (1). Z_{it} is a vector for covariates like
 18 socioeconomic and other endogenous factors, η_{t-k} determines how much the current medical
 19 expenses will increase when the lagged expenditures increase by 10%, and ε_{it} is a random error
 20 term. Statistical analyses were performed using Stata, version 16.0 (Stata Corp, Inc., Cary, TX,
 21 USA).

22 **Variable Selection**

1 **Dependent variables**

2 The dependent variables included total medical expenditures (Y) and annual incidence of medical
3 costs (I). Annual total medical costs were calculated by combining the outpatient, hospitalization
4 (fees paid to the hospital, including ward fees but not wages paid to a hired nurse, transportation
5 costs, and accommodation costs for oneself or family members), and self-treatment expenses
6 (medicines purchased by the patient but not including prescription medications). I is a dummy
7 binary variable set based on Y such that if the total medical expenses are not equal to zero, I is
8 defined as “1”, otherwise it is defined as “0.”

9 **Independent variables**

10 The independent indicators from three dimensions (predisposing factors, enabling factors, and need
11 factors) were selected using the Anderson health service utilization model [24], the mainstream
12 model used for health service utilization research. The model divides factors affecting health service
13 utilization into three categories: predisposing, enabling, and need factors. Predisposing factors
14 include demographic and social structure factors, such as age, sex, marital status, retirement
15 condition, education level, and residence. Enabling factors include financial and organizational
16 variables such as household income per capita, health insurance, inpatient visit times, type of
17 outpatient institution and visit times, and multi-type outpatient facility visits. Specifically, “multi-
18 type outpatient facility visits” is a virtual variable based on the multi-choice question “Which types
19 of medical facilities have you visited in the last 4 weeks for outpatient treatment?” The outpatient
20 institutions were divided into two categories based on the grade classification of the medical
21 facilities. While general, specialized and Chinese medicine hospitals = 0, community health centers,
22 town hospitals, and village clinics = 1. If the respondent received treatment from both high-ranking
23 hospitals and primary medical institutions, “type of outpatient institution” =1, and if the respondent

1 only received medical services from primary medical institutions, the “type of outpatient
2 institution”=0. Need factors included self-reported health status and chronic diseases. To investigate
3 the impact of several chronic diseases, a “comorbidity” dummy variable was developed. All
4 independent variables were set as categorical indicators. Detailed descriptions of the samples are
5 included in Table 1.

6 *<Insert Table 1 near here>*

7

8 **Results**

9 **The concentration of medical expenses among middle-aged and elderly participants**

10 The cumulative percentages of medical expenditures spent by middle-aged and elderly people who
11 were in the top 1%, 10%, and 50% annual medical expense categories in 2013, 2015, and 2018 are
12 shown in Figures 1 and 4. Four types of expenses were reported: total medical expenses, and
13 inpatient, outpatient, and self-treatment medical expenses.

14 In 2013, 2015, and 2018, middle-aged and elderly people in the top 1% category of total medical
15 expenditures spent 22.0%, 27.9% and 35.2% of the total annual medical expenses, respectively
16 (Figure 1). Medical expenses for middle-aged and elderly people in the top 10% category accounted
17 for 64.2%, 68.6%, and 76.1% of the total medical costs, respectively, while those with expenses in
18 the top 50% accounted for 95.7%, 96.1%, and 97.2% of the total medical costs, respectively. The
19 proportion of total medical expenses spent by the top 1%, 10% and 50% of middle-aged and elderly
20 spenders increased over time.

21 During 2013, 2015, and 2018, middle-aged and elderly people with the top 1% of hospitalization
22 expenses spent 6.9%, 6.0%, and 7.1% of annual hospitalization costs, respectively (Figure 2). Those

1 with the top 10% of expenditures spent 41.8%, 46.6%, and 41.0% of the annual inpatient expenses
2 in 2013, 2015, and 2018, respectively, while those with the top 50% spent 83.9%, 89.7%, and 88.8%,
3 respectively. The proportion of individuals with the top 1%, 10%, and 50% of expenditures
4 fluctuated slightly from 2013 to 2018.

5 During 2013, 2015, and 2018, middle-aged and elderly people with the top 1% of outpatient
6 expenses spent 19.2%, 17.5%, and 22.9% of the annual costs, respectively (Figure 3). Those with
7 the top 10% spent 51.7%, 62.5%, and 61.7% of the annual outpatient expenses in 2013, 2015, and
8 2018, respectively, while those with the top 50% spent 90.7%, 94.0%, and 93.2%, respectively. The
9 proportion of individuals with the top 1%, 10%, and 50% outpatient expenses demonstrated an
10 increasing trend from 2013 to 2018.

11 In 2013, 2015, and 2018, middle-aged and elderly people with the top 1% of self-treatment
12 expenses spent 21.6%, 20.4% and 24.4% of the annual costs, respectively (Figure 4). Those with
13 the top 10% spent 59.2%, 61.4%, and 59.2% of the annual self-treatment medical costs in 2013,
14 2015, and 2018, respectively, and those with the top 50% spent 94.5%, 94.0%, and 92.3%,
15 respectively. The proportion of individuals with the top 1%, 10% and 50% annual self-treatment
16 medical expenses fluctuated slightly from 2013 to 2018.

17 *<Insert Fig 1 to 4 near here>*

18 **The characteristics of middle-aged and elderly participants with high medical expenses**

19 There were significant differences between the characteristics of the middle-aged and elderly
20 participants with the top 10% and 20% medical expenditures and those with the bottom 80% and
21 90% medical expenditures (Tables 2 and 3). Those in the top 10% medical expense group were more
22 likely to be female (56.84% vs 50.86%; $p < 0.000$), older (age ≥ 75 years: 9.56% vs 7.35%; $p < 0.000$),

1 retired (17.34% vs 9.66%; $p < 0.000$), more educated (tertiary education: 1.07% vs 0.44%; $p < 0.000$)
2 and to live in an urban area (28.78% vs 22.50%; $p < 0.000$) than those in the bottom 90% medical
3 expense group. Middle-aged and elderly adults with the top 10% medical expenses were more likely
4 to have a lower income level (household income per capita \leq ¥8,000: 83.20% vs 76.41%; $p < 0.000$),
5 receive outpatient treatment in higher-ranking hospitals (58.18% vs 19.71%; $p < 0.000$) or have
6 multi-types of outpatient visits (8.67% vs 0.51%; $p < 0.000$). In addition, middle-aged and elderly
7 participants with the top 10% medical expenses tended to have more hospitalizations (number of
8 hospitalizations >1 : 28.87% vs 0.90%; $p < 0.000$), outpatient treatments (number of outpatient visits
9 ≥ 36 : 17.78% vs 1.17%; $p < 0.000$) and multiple chronic diseases (comorbidity: 65.95% vs 36.71%).

10 The top 10% of medical spenders demonstrated a significantly higher prevalence of chronic
11 diseases (hypertension: 36.73% vs 24.52%, diabetes: 12.24% vs 5.65%, cancer: 2.86% vs 0.97%,
12 chronic lung diseases: 19.48% vs 9.00%, liver diseases: 10.01% vs 4.03%, heart diseases: 29.31%
13 vs 11.35%, kidney diseases: 12.96% vs 5.67%, stomach or digestive diseases: 36.37% vs 22.29%,
14 arthritis or rheumatism: 48.70% vs 33.77%) than the bottom 90% of medical spenders (Table 3).

15 *<Insert Table 2 and 3 near here>*

16 **Persistence of high medical expenses**

17 No middle-aged and elderly people with the top 10% medical expenses in 2013 remained in the top
18 10% in 2015 or 2018 (Table 4). Similarly, only 30.83% and 26.46% of those with the top 20%
19 medical expenses in 2013 still had the top 20% expenses in 2015 and 2018, respectively. In contrast,
20 46.42% and 50% of those with the top 50% medical expenses in 2013 still had the top 50% expenses
21 in 2015 and 2018.

22 Most (70%) of those with the top 10% medical expenses in 2013 transferred to the top 20% in

1 2015 and 2018, and 80% and 70% of those with the top 10% high medical expenses transferred to
2 the top 50% in 2015 and 2018, respectively. In contrast, 20% and 30% of participants with the top
3 10% of medical expenses in 2013 transferred to the bottom 50% in 2015 and 2018, and 55.03% and
4 55.29% of those with the top 20% of medical expenses in 2013 transferred to the top 50% of medical
5 expenses in 2015 and 2018. While 44.97% and 44.71% of those with the top 20% of medical
6 expenses in 2013 transferred to the bottom 50% in 2015 and 2018, respectively, 21.48% and 20.64%
7 with the top 50% medical expenses in 2013 transferred to the top 20% in 2015 and 2018, respectively,
8 and 53.58% and 50% of those with the top 50% medical expenses in 2013 transferred to the bottom
9 50% in 2015 and 2018.

10 *<Insert Table 4 near here>*

11 **Heckman Selection Model Regression**

12 **Incidence of medical expenses**

13 Results of the two regression models, including the lagged one-period and two-period variables, are
14 shown in Table 5. Once medical expenses occurred, the probability of incurring persistent medical
15 expenditures in the next period increased significantly by 16.5% and 15.8%, respectively. Incurring
16 medical expenses in lagged two periods increased the incidence of current medical expenses by
17 11.7%. Females had a significantly higher likelihood of incurring medical expenses than males
18 ($dy/dx=0.033$, 0.027), and chronic diseases had a persistent effect on the incidence of medical
19 expenses. The probability of incurring medical expenses in the current period increased significantly
20 by 29.3% once middle-aged and elderly adults with cancer had medical expenses in lagged one
21 period ($dy/dx=0.293$). Meanwhile, those suffering from hypertension ($dy/dx=0.102$, 0.069),
22 diabetes ($dy/dx=0.098$), chronic lung diseases ($dy/dx=0.078$, 0.081), liver diseases ($dy/dx=0.049$),
23 heart diseases ($dy/dx=0.074$), kidney diseases ($dy/dx=0.089$, 0.115), digestive diseases

1 (dy/dx=0.055) and arthritis or rheumatism (dy/dx=0.060) in the lagged one-period significantly
2 increased the probability of incurring medical expenses. In addition, suffering from hypertension,
3 liver disease, arthritis, or rheumatism in the lagged two-period also increased the incidence of
4 current medical costs (dy/dx=0.103, 0.111, and 0.036, respectively).

5 *<Insert Table 5 near here>*

6 **Total medical expenses**

7 Factors associated with total medical expenses among middle-aged and elderly participants in the
8 second part of Heckman selection model are shown in Table 6. The medical expenses were
9 logarithmically transformed in the regression model. When other variables were controlled, the
10 current level of medical expenses was strongly affected by past medical expenditures. For each 10%
11 increase in medical costs during the lagged one-period, the current medical expenses increased by
12 2.25% and 1.65%, respectively, and for each 10% increase in medical expenses during the lagged
13 two-period, current medical expenses increased by 1.13%.

14 Middle-aged and elderly people who were ≥ 75 years of age generally had more medical expenses
15 than those < 55 years of age (Coef. =0.447). Females had higher medical expenses than males (Coef.
16 =0.328,0.297) and married respondents had higher medical expenses than the unmarried (Coef.
17 =0.258). Respondents living in urban areas generally spent more on medical expenses than those
18 living in rural areas (Coef. =0.221).

19 Middle-aged and elderly adults with more outpatient visits had higher medical costs (Coef. =24–
20 36 and ≥ 36 outpatient visits=0.331, 0.827) as did those with more hospitalizations (Coef. =1.067,
21 0.911). In addition, respondents receiving medical services in primary medical institutions had
22 significantly lower medical costs than those seen in general or specialized hospitals (Coef. =-1.167,

1 -1.151). Medical expenses were significantly higher for middle-aged and elderly individuals with a
2 household per capita income of ¥15,600–30,000 and \geq ¥30,000 than those with an income \leq ¥8,000
3 (Coef. of ¥15,600–30,000 and \geq ¥30,000=0.284, 0.307). In addition, respondents with health
4 insurance generally spent more on medical expenses (Coef. =2.734, 2.012). A significantly positive
5 association was found between those with chronic diseases and higher medical expenses.
6 Respondents with hypertension (Coef. =0.383, 0.315), diabetes (Coef. =0.408, 0.328), chronic lung
7 diseases (Coef. =0.220), heart disease (Coef. =0.346, 0.308) and arthritis or rheumatism (Coef.
8 =0.197) generally had higher medical expense. The total medical expenses for those with
9 comorbidities were significantly higher (Coef. =0.274, 0.654). In addition, middle-aged and elderly
10 adults with poor self-reported health status had higher medical costs (Coef. =1.388,1.811).

11 *<Insert Table 6 near here>*

12 **Discussion**

13 The medical expenses of middle-aged and elderly people account for a large proportion of all
14 medical expenses because of the higher incidence of age-related chronic disease in this population
15 [21-23]. Thus, it is important to study the concentration and persistence of medical expenses in this
16 group. To the best of our knowledge, this study of a nationally representative longitudinal Chinese
17 household survey population is the first to analyze the extent and characteristics of the concentration
18 and persistence of medical expenses among Chinese middle-aged and elderly adults.

19 **The concentration of medical expenses among middle-aged and elderly participants**

20 This study found that 22.0%–35.2% of total annual medical expenses were spent by the top 1% of
21 middle-aged and elderly medical spenders, while 64.2%–76.1% were spent by the top 10% and
22 95.7%–97.2% were spent by the top 50%. These findings demonstrate that total medical expenses

1 are concentrated in a small portion of middle-aged and elderly adults, supporting findings from
2 previous studies [6,10]. These results indicate that there is an inequitable utilization of health
3 resources among middle-aged and elderly adults. As a result, policy efforts should focus on
4 optimizing the allocation of health service utilization by different groups to enhance social welfare
5 [25].

6 The concentration of total medical expenses increased from 2013 to 2018, which contrasts with
7 some studies [4,6]. This suggests that the total medical expenses of Chinese middle-aged and elderly
8 people were more concentrated among a small number of individuals in recent years, which may be
9 the result of advancements in medical treatment technologies and universal social health insurance
10 coverage in China. Thus, differentiated measures of social health insurance for high medical
11 expense spenders among these demographics should be a priority concern for public policy.
12 Meanwhile, the concentration of outpatient medical expenses also increased from 2013 to 2018.
13 While, inpatient medical expenses in the top 10% decreased slightly from 2015 to 2018, this may
14 be the result of recent reforms to social medical insurance and standardization of the clinical care
15 process in China. The concentration of self-treatment medical expenses for participants in the top
16 1% increased from 2015 to 2018, likely as a result of worsening inequities in self-treatment expenses.
17 Chinese health policy will require more focus on improving the utilization fairness of over-the-
18 counter medications.

19 **Persistence of high medical expenses among middle-aged and elderly participants**

20 Some middle-aged and elderly participants with high medical expenses in 2013 remained in the
21 same high medical expense categories in 2015 and 2018, supporting prior studies illustrating the
22 persistence of high medical costs [7,22,26]. This may be linked to the transitioning disease spectrum

1 and high prevalence of chronic diseases among high medical expense groups. Policy efforts should
2 focus on preventive management for individuals at risk of incurring high medical expenses [2,6,20].

3 Most middle-aged and elderly participants with the top 10% of medical expenses in 2013
4 transferred to the top 20% or 50% categories in 2015 and 2018. This is consistent with previous
5 surveys [27,28] and indicates that high medical expenses may have been brought under control with
6 reform to the medical system and the advancement of medical technology in China [12,15].
7 However, it should be noted that almost 20% of middle-aged and elderly adults in the top 50%
8 spending category in 2013 transferred to the top 20% spending category in 2015 and 2018. Thus,
9 policy efforts will need to better define the characteristics of this group to effectively control rising
10 medical expenses.

11 **Characteristics of middle-aged and elderly participants with high medical expenses**

12 Middle-aged and elderly participants with high medical expenses were more likely to be female and
13 older, which is in agreement with previous studies [5,29]. This may be the result of the increasing
14 incidence of menopause and severe disease in older women [30,31]. A high education level was also
15 associated with higher medical expenses, possibly resulting from better health awareness and a
16 stronger motivation to seek medical services [32]. In addition, participants with high medical
17 expenses were more likely to live in urban than rural areas. This may be attributed to socioeconomic
18 differences between urban and rural areas in China [33]. Urban residents tend to have higher
19 incomes and are thus able to incur higher medical expenses and advanced medical technologies and
20 resources tend to be more concentrated in urban areas. Thus, initiatives that optimize the allocation
21 of health resources in urban and rural areas should be considered to narrow the gap between these
22 regions. Participants with a lower income were also more likely to be in the higher medical expenses

1 group, which is consistent with some studies [11, 34]. This may be because less-wealthy groups
2 have lower health management awareness and more risky lifestyle behaviors that increase the
3 prevalence of severe diseases [35]. Thus, health sectors should strengthen health management,
4 disease prevention, and financial assistance for poorer individuals.

5 In this study, participants receiving medical services in general or specialized hospitals had high
6 medical expenses, potentially as a result of excessive medical treatment. This should be addressed
7 by reforming hierarchical health services and promoting rational utilization. Middle-aged and
8 elderly participants who visited multi-type medical institutions also demonstrated high medical
9 expenses. Social health insurance policies should focus on reducing the medical expenses of these
10 at-risk groups.

11 Middle-aged and elderly people with high medical expenses demonstrated a significantly higher
12 prevalence of hypertension, diabetes, cancer, chronic lung diseases, liver diseases, heart diseases,
13 kidney diseases, digestive diseases, and arthritis or rheumatism, as well as an increased prevalence
14 of comorbidities, which supports previous survey findings [17,36]. Prevention management and
15 early screening and treatment for chronic diseases should help to reduce medical expenses in this
16 population.

17 **Factors associated with the persistence of total medical expenses**

18 This study found that the incidence of medical expenses for middle-aged and elderly people
19 increased by 15.8%–16.5% from 2013 to 2015 and by 11.7% from 2013 to 2018. Current medical
20 expenses increased by 1.65%–2.25% and 1.13% for every 10% increase in the lagged one and
21 lagged two-period medical expenses, respectively, which is consistent with previous survey results
22 [5,37]. These findings demonstrate that medical expenses are significantly persistent. While

1 previous medical expenses had a persistently positive effect on current medical expenses, the effect
2 of the lagged one period was the strongest, which is also supported by other studies [15,18]. This
3 effect may be attributed to the long-term treatment of chronic diseases. It is worth noting that
4 middle-aged and elderly people with diabetes, cancer, chronic lung diseases, heart diseases, kidney
5 diseases, and digestive diseases had a high likelihood of having persistent medical expenses for two
6 periods, while those with hypertension, liver diseases, arthritis or rheumatism were more likely to
7 have persistent medical expenses for three periods. In addition, those with hypertension, diabetes,
8 chronic lung diseases, heart diseases, arthritis, rheumatism, or comorbidities generally had higher
9 medical expenses, supporting prior studies [16,38]. To control persistent medical expenses and
10 higher medical expenses, preventive measures should focus on reducing the incidence of chronic
11 diseases. Health sectors should consider joining primary medical institutions to establish persistent
12 integrated care systems for middle-aged and elderly people with chronic diseases to reduce the
13 overuse of medical resources during long-term treatment regimens. In addition, the reimbursement
14 of medical insurance should be improved for those with chronic diseases to alleviate persistent
15 medical expense burdens.

16 This study found that females, age ≥ 75 years, and living in an urban area were all positive factors
17 for higher medical expenses, which is consistent with other surveys results [21,39]. In addition,
18 married individuals had higher medical costs than unmarried, possibly because married people
19 utilize more health services as a result of supervision from their spouses. Middle-aged and elderly
20 people with high per capita income and health insurance generally incurred higher medical expenses
21 indicating that those with high income could afford higher medical expenses. Moreover, social
22 health insurance may promote an increase in medical expenses to some extent. Participants who

1 received more outpatient visits and hospitalizations or who visited higher-ranking hospitals
2 generally had higher medical expenses, which is consistent with a previous study [40]. Thus,
3 promoting the primary medical treatment of hierarchical health service in China may help to reduce
4 total medical expenses.

5 There are some limitations to the current study. Because the CHARLS does not include
6 information on chronic disease diagnoses, the chronic disease prevalence was obtained by asking
7 adults whether they had been diagnosed with a chronic disease. Thus, individuals who might have
8 had a chronic disease but were not or did not recall being diagnosed may have been excluded in
9 which case the true prevalence of a particular condition may be underestimated. In addition, data on
10 missing samples were removed from the study, potentially resulting in an undercount of individuals
11 with persistent medical expenses. Finally, medical expenses are likely to be influenced by additional
12 factors which are not included in the claims data.

13 **Conclusion**

14 This study found that total medical expenses were concentrated among a few middle-aged and
15 elderly individuals and the concentration increased over time. Some health and socioeconomic
16 characteristics such as chronic disease, age, education level, residence, employment status, income
17 level, and medical service utilization condition were significantly associated with high medical
18 expenses. Thus, policy effects should address these factors to optimize the allocation of health
19 resources to lower medical costs and ensure the equitable utilization of health services. High medical
20 expenses also demonstrated strong persistence. Lagged medical expenses had a significantly
21 persistent positive effect on current medical expenses and the effect of lagged one period was the
22 strongest. Those having chronic diseases in the lagged periods were more likely to have persistent

1 medical expenses. To address this, the government should establish a more efficient care system and
2 adjust social health insurance policy to alleviate long-term medical financial burden.

3

4 **List of abbreviations**

5 HCUs, high-cost users; CHARLS, the China Health and Retirement Longitudinal Study; PPS,
6 probability proportionate to size; PSUs, primary sampling units. GDP, gross domestic product; GIS,
7 geographic information system.

8 **Declarations**

9 **Ethics approval and consent to participate**

10 This study was approved by the Academic Research Ethics Committee of Nanjing Medical
11 University; reference number: 2022460. All procedures were in accordance with the ethical
12 standards of the Helsinki Declaration. Participants provided informed consent prior to data
13 collection.

14 **Consent for publication**

15 Not applicable.

16 **Availability of data and materials**

17 The datasets used in the current study are not publicly available due to the confidential policy but
18 are available from the corresponding author on reasonable request.

19 **Competing interests**

20 The authors declare that they have no competing interests.

21 **Funding**

22 This study is funded by the major research project of philosophy and social sciences in colleges
23 and universities in Jiangsu Province (no.2021SJZDA148). The funding bodies were not involved in

1 the design of the study, or data collection, analysis, and interpretation or in writing the manuscript.

2 **Authors' contributions**

3 LJ led the analysis of the data and wrote the first draft of the manuscript; ZW contributed to the
4 study design, interpretation of the data and helped in the writing of the final draft of the manuscript;
5 LZ and QQ helped in data analysis and contributed in writing. All authors read and approved the
6 final manuscript.

7 **Acknowledgements**

8 The authors are grateful to many people and groups who made this study possible. We especially
9 thank the head of the National School of Development, Peking University. We would also like to
10 thank all study participants for their time to be interviewed.

11 **Reference**

- 12 [1] Jones AM, Lomas J, Rice N. Healthcare Cost Regressions: Going Beyond the Mean to Estimate
13 the Full Distribution. *Health Economics*. 2015;24(9):1192-1212.
- 14 [2] Wodchis WP, Austin PC, Henry DA. A 3-year study of high-cost users of health care. *Canadian
15 Medical Association Journal*. 2016;188(3):182-188.
- 16 [3] Hirth RA, Gibson TB, Levy HG, Smith JA, Calonico S, Das A. New Evidence on the Persistence
17 of Health Spending. *Medical Care Research and Review*. 2015;72(3):277-297.
- 18 [4] Ebrahimoghli R, Sadeghi-Bazargani H, Janati A, Hamishehkar H, Khalili-Azimi A. A 4-Year
19 Investigation of Ambulatory Health Care Expenditure Concentration and High-Cost Patients: An
20 Experience From a Developing Country. *The Journal of ambulatory care management*.
21 2020;43(2):169-178.
- 22 [5] Monheit AC. Persistence in health expenditures in the short run: Prevalence and consequences.

- 1 Medical Care. 2003;41(7):53-64.
- 2 [6] Riley GF. Long-term trends in the concentration of Medicare spending. Health Affairs.
3 2007;26(3):808-816.
- 4 [7] Kohn JL, Liu JS. THE DYNAMICS OF MEDICAL CARE USE IN THE BRITISH
5 HOUSEHOLD PANEL SURVEY. Health Economics. 2013;22(6):687-710.
- 6 [8] Newhouse JP, Manning WG, Keeler EB, Sloss EM. Adjusting capitation rates using objective
7 health measures and prior utilization. Health care financing review. 1989;10(3):41-54.
- 8 [9] Wammes JJG, van der Wees PJ, Tanke MAC, Westert GP, Jeurissen PPT. Systematic review of
9 high-cost patients' characteristics and healthcare utilisation. Bmj Open. 2018;8(9).
- 10 [10] Lange L, Pimperl A, Schulte T, Groene O, Tanke M. High-cost patients in Germany: General
11 description of utilization and costs. Zeitschrift fur Evidenz, Fortbildung und Qualitat im
12 Gesundheitswesen. 2020;153-154:76-83.
- 13 [11] Lemstra M, Mackenbach J, Neudorf C, Nannapaneni U. High health care utilization and costs
14 associated with lower socio-economic status: results from a linked dataset. Canadian journal of
15 public health = Revue canadienne de sante publique. 2009;100(3):180-183.
- 16 [12] Longden T, Wong CY, Haywood P, Hall J, van Gool K. The prevalence of persistence and
17 related health status: An analysis of persistently high healthcare costs in the short term and medium
18 term. Social Science & Medicine. 2018;211:147-156.
- 19 [13] Ng SH-X, Rahman N, Ang IYH, et al. Characterization of high healthcare utilizer groups using
20 administrative data from an electronic medical record database. BMC Health Services Research.
21 2019;19.
- 22 [14] Teraoka E, Kunisawa S, Imanaka Y. Trajectories of end-of-life medical and long-term care

1 expenditures for older adults in Japan: retrospective longitudinal study using a large-scale linked
2 database of medical and long-term care claims. *Bmc Geriatrics*. 2021;21(1).

3 [15] Cohen SB, Ezzati-Rice T, Yu W. The utility of extended longitudinal profiles in predicting
4 future health care expenditures. *Medical Care*. 2006;44(5):45-53.

5 [16] Robst J. Developing Models to Predict Persistent High-Cost Cases in Florida Medicaid.
6 *Population Health Management*. 2015;18(6):467-476.

7 [17] Liptak GS, Shone LP, Auinger P, Dick AW, Ryan SA, Szilagyi PG. Short-term persistence of
8 high health care costs in a nationally representative sample of children. *Pediatrics*.
9 2006;118(4):E1001-E1009.

10 [18] Lin P-J, Biddle AK, Ganguly R, Kaufer DI, Maciejewski ML. The Concentration and
11 Persistence of Health Care Expenditures and Prescription Drug Expenditures in Medicare
12 Beneficiaries With Alzheimer Disease and Related Dementias. *Medical Care*. 2009;47(11):1174-
13 1179.

14 [19] Berk ML, Fang Z. Most Americans Have Good Health, Little Unmet Need, And Few Health
15 Care Expenses. *Health Affairs*. 2017;36(4):742-746.

16 [20] Krause JS, Murday D, Corley EH, DiPiro ND. Concentration of Costs Among High Utilizers
17 of Health Care Services Over the First 10 Years After Spinal Cord Injury Rehabilitation: A
18 Population-based Study. *Archives of Physical Medicine and Rehabilitation*. 2019;100(5):938-944.

19 [21] Peng X, Du C. Research on the concentration and persistence of medical expenditure : micro
20 empirical evidence from China. *World Economy*. 2019;42(12):51-76.

21 [22] Ku L-JE, Chiou M-J, Liu L-F. Variations in the persistence of health expenditures and the
22 implications for the design of capitation payments in Taiwan. *Journal of health services research &*

1 policy. 2015;20(3):146-153.

2 [23] Xu L, Gong G, Wang H. Research on Medical Expenditure under Basic Medical Insurance for
3 Urban Employees—Unnegligible Persistence Characteristics. *Journal of Finance and economics*.
4 2011;37(05):25-35.

5 [24] de Boer AG, Wijker W, de Haes HC. Predictors of health care utilization in the chronically ill:
6 a review of the literature. *Health policy (Amsterdam, Netherlands)*. 1997;42(2):101-115.

7 [25] Berk ML, Monheit AC. The concentration of health care expenditures, revisited. *Health Affairs*.
8 2001;20(2):9-18.

9 [26] Figueroa JF, Lyon Z, Zhou X, Grabowski DC, Jha AK. Persistence and Drivers of High-Cost
10 Status Among Dual-Eligible Medicare and Medicaid Beneficiaries: An Observational Study. *Ann*
11 *Intern Med*. 2018;169(8):528-534.

12 [27] Coughlin TA, Long SK. Health Care Spending and Service Use among High-Cost Medicaid
13 Beneficiaries, 2002-2004. *Inquiry-the Journal of Health Care Organization Provision and Financing*.
14 2009;46(4):405-417.

15 [28] Liptak GS, Shone LP, Auinger P, Dick AW, Ryan SA, Szilagyi PG. Short-term persistence of
16 high health care costs in a nationally representative sample of children. *Pediatrics*.
17 2006;118(4):e1001-1009.

18 [29] Tanke MAC, Feyman Y, Bernal-Delgado E, et al. A challenge to all. A primer on inter-country
19 differences of high-need, high-cost patients. *Plos One*. 2019;14(6).

20 [30] Perkins AJ, Kroenke K, Unutzer J, et al. Common comorbidity scales were similar in their
21 ability to predict health care costs and mortality. *Journal of Clinical Epidemiology*.
22 2004;57(10):1040-1048.

- 1 [31] Longman JM, Rolfe MI, Passey MD, et al. Frequent hospital admission of older people with
2 chronic disease: a cross-sectional survey with telephone follow-up and data linkage. *Bmc Health*
3 *Services Research*. 2012;12.
- 4 [32] Rosella LC, Fitzpatrick T, Wodchis WP, Calzavara A, Manson H, Goel V. High-cost health care
5 users in Ontario, Canada: demographic, socio-economic, and health status characteristics. *Bmc*
6 *Health Services Research*. 2014;14.
- 7 [33] Kunna R, San Sebastian M, Stewart Williams J. Measurement and decomposition of
8 socioeconomic inequality in single and multimorbidity in older adults in China and Ghana: results
9 from the WHO study on global AGEing and adult health (SAGE). *International journal for equity*
10 *in health*. 2017;16(1):79-79.
- 11 [34] Selden TM. The Within-Year Concentration of Medical Care: Implications for Family Out-of-
12 Pocket Expenditure Burdens. *Health Services Research*. 2009;44(3):1029-1051.
- 13 [35] Fitzpatrick T, Rosella LC, Calzavara A, et al. Looking Beyond Income and Education
14 Socioeconomic Status Gradients Among Future High-Cost Users of Health Care. *American Journal*
15 *of Preventive Medicine*. 2015;49(2):161-171.
- 16 [36] Anderson GF. Medicare and chronic conditions. *New England Journal of Medicine*.
17 2005;353(3):305-309.
- 18 [37] Rettenmaier AJ, Wang ZJ. Persistence in Medicare reimbursements and personal medical
19 accounts. *Journal of Health Economics*. 2006;25(1):39-57.
- 20 [38] Chang H-Y. The impact of morbidity trajectories on identifying high-cost cases: using Taiwan's
21 National Health Insurance as an example. *Journal of Public Health*. 2014;36(2):300-307.
- 22 [39] Wouterse B, Meijboom BR, Polder JJ. THE RELATIONSHIP BETWEEN BASELINE

1 HEALTH AND LONGITUDINAL COSTS OF HOSPITAL USE. Health Economics.
2 2011;20(8):985-1008.
3 [40] Qureshi D, Isenberg S, Tanuseputro P, et al. Describing the characteristics and healthcare use
4 of high-cost acute care users at the end of life: a pan-Canadian population-based study. BMC Health
5 Services Research. 31 2020;20(1)

1 **Table1** Description of the dependent and independent variables

Variable	Category	Indicator/survey question
Dependent variables		
Total medical expenses	¥	Question: What was the total medical expenses during the past year?
Incidence of medical expenses	=0, non-incurrence of medical expenses; =1, incurrence of medical expenses	Whether the total medical expenses were equal to zero.
Independent variables		
Predisposing factors		
Sex	=0, male; =1, female	
Age (years)	=0,45–55; =1,55–65; =2,65–75; =3, if≥75	
Marital status	=0, not married; =1, married	Question: What is your marital status? (Separation, divorce, widowhood, and cohabitation belong to “not married”.)
Retirement status	=0, not retired; =1, retired	Question: What is your retirement status?
Education level	=0, less than lower secondary; =1, upper secondary, vocational training; =2, tertiary	Question: What is the highest level of education you have attained?
Area of residence	=0, rural; =1, urban	Type of residence.
Enabling factors		
Income level	=0, ≤ ¥8,000; =1, ¥8,000–15,600; =2, ¥15,600–30,000; =3, ≥ ¥30,000	Yearly household income divided by the number of household members.
Social health insurance	=0, have no insurance; =1, have insurance	Question: Do you have any social health insurance?
Type of outpatient medical facilities	=0, general hospital, specialized hospital, Chinese medicine hospital; =1, community healthcare center, township hospital, village clinic	Question: Which types of medical facilities have you visited in the last 4 weeks for outpatient treatment?
Multi-type outpatient facility visits	=0, no multi-type outpatient facility visits; =1, muti-type outpatient facility visits	Whether attended multiple types of outpatient facilities.

1 **Table 1** Description of dependent and independent variables (Continued)

	Categories	Indicators/survey questions
Number of outpatient visits	=0, ≤24; =1, 24 - 36; =2, ≥36	Question: How many times did you visit/ been visited by during the last month?
Number of hospitalizations	=0, ≤1; =1, >1	Question: How many times have you received inpatient care during the past year?
Need factors		
Self-reported health status	=0, very good; =1, good; =2, fair; =3, poor	
Comorbidity	=1, ≤1; =2, >1	Number of chronic diseases.
Hypertension	=0, no; =1, yes	
Diabetes	=0, no; =1, yes	
Cancer	=0, no; =1, yes	Excluding minor skin cancers
Chronic lung diseases	=0, no; =1, yes	Excluding tumors or cancer
Liver diseases	=0, no; =1, yes	Excluding fatty liver, tumors, cancer
Heart diseases	=0, no; =1, yes	
Kidney diseases	=0, no; =1, yes	Excluding tumors or cancer
Stomach or other digestive diseases	=0, no; =1, yes	Excluding tumors or cancer
Arthritis or rheumatism	=0, no; =1, yes	

2

3 **Table 2** Characteristics of middle-aged and elderly participants with the top 10% and 20% of
4 medical expenses

	Top10% n=1119 (5.63%)	Bottom 90% n=18750 (94.37%)	P	Top 20% n=2180 (10.97%)	Bottom 80% n=17689 (89.03%)	P
Predisposing factors						
Sex			0.000			0.000
Male	43.16%	49.14%		42.89%	49.53%	
Female	56.84%	50.86%		57.11%	50.47%	
Age (years)			0.000			0.000
45–55	23.77%	28.79%		25.28%	28.91%	
55–65	38.25%	40.12%		38.53%	40.20%	
65–75	28.42%	23.73%		26.51%	23.69%	
≥75	9.56%	7.35%		9.68%	7.20%	
Marital status			0.817			0.374
Not married	11.80%	11.57%		12.16%	11.51%	
Married	88.20%	88.43%		87.84%	88.49%	
Retirement status			0.000			0.000
Not retired	82.66%	90.34%		83.62%	90.68%	

Retired	17.34%	9.66%		16.38%	9.32%	
Education level			0.000			0.002
Less than lower secondary	85.08%	88.82%		86.79%	88.83%	
Upper secondary, vocational training	13.85%	10.74%		12.39%	10.74%	
Tertiary	1.07%	0.44%		0.82%	0.43%	
Area of residence			0.000			0.000
Rural	71.22%	77.50%		71.74%	77.81%	
Urban	28.78%	22.50%		28.26%	22.19%	
Enabling factors						
Income level			0.000			0.000
≤ ¥8,000	83.20%	76.41%		81.19%	76.25%	
¥8,000–15,600	4.29%	7.66%		5.55%	7.71%	
¥15,600–30,000	5.36%	7.76%		5.55%	7.88%	
≥ ¥30,000	7.15%	8.18%		7.71%	8.17%	
Social health insurance			0.456			0.455
Have no insurance	1.97%	2.31%		2.06%	2.32%	
Have insurance	98.03%	97.69%		97.94%	97.68%	

1

2 **Table 2** Characteristics of middle-aged and elderly participants with the top 10% and 20% of
3 medical expenses (Continued)

	Top10% n=1119 (5.63%)	Bottom 90% n=18750 (94.37%)	P	Top 20% n=2180 (10.97%)	Bottom 80% n=17689 (89.03%)	P
Type of outpatient medical facilities			0.000			0.000
General hospital, specialized hospital, Chinese medicine hospital	58.18%	19.71%		48.46%	13.87%	
Community healthcare center, township hospital, village clinic	41.82%	80.29%		51.54%	86.13%	
Multi-type outpatient facility visits			0.000			0.000
No multi-type outpatient facility visits	91.33%	99.49%		93.39%	99.72%	
Multi-type outpatient facility visits	8.67%	0.51%		6.61%	0.28%	
Number of outpatient visits			0.000			0.000
≤ 24	72.39%	97.60%		77.20%	98.52%	

24–36	9.83%	1.23%		8.76%	0.84%	
≥ 36	17.78%	1.17%		14.04%	0.64%	
Number of hospitalizations			0.000			0.000
≤ 1	71.13%	99.10%		80.73%	99.60%	
> 1	28.87%	0.90%		19.27%	0.40%	
Need factors						
Self-reported health status			0.000			0.000
Very good	3.84%	14.31%		3.72%	14.96%	
Good	6.70%	15.17%		7.11%	15.63%	
Fair	39.59%	53.00%		44.54%	53.20%	
Poor	49.87%	17.51%		44.63%	16.21%	
Comorbidity			0.000			0.000
Without comorbidity	34.04%	63.29%		36.19%	64.78%	
With comorbidity	65.95%	36.71%		63.81%	35.22%	

1 Note: Consolidated observations in the top 10% and top 20% medical expense groups from 2013 to 2018 were
2 defined as the “Top 10%” and “Top 20%” medical spending groups, respectively.

3

4 **Table 3** Chronic diseases in middle-aged and elderly participants with the top 10% and 20% of
5 medical expenses

	Top10% n=1119 (5.63%)	Bottom 90% n=18750 (94.37%)	P	Top 20% n=2180 (10.97%)	Bottom 80% n=17689 (89.03%)	P
Hypertension			0.000			0.000
No	63.27%	75.48%		63.35%	76.21%	
Yes	36.73%	24.52%		36.65%	23.79%	
Diabetes			0.000			0.000
No	87.76%	94.35%		88.67%	94.63%	
Yes	12.24%	5.65%		11.33%	5.37%	
Cancer			0.000			0.000
No	97.14%	99.03%		97.71%	99.08%	
Yes	2.86%	0.97%		2.29%	0.92%	
Chronic lung diseases			0.000			0.000
No	80.52%	91.00%		82.25%	91.41%	
Yes	19.48%	9.00%		17.75%	8.59%	
Liver diseases			0.000			0.000
No	89.99%	95.97%		91.10%	96.20%	
Yes	10.01%	4.03%		8.90%	3.80%	
Heart diseases			0.000			0.000
No	70.69%	88.65%		74.17%	90.43%	
Yes	29.31%	11.35%		25.83%	9.57%	
Kidney diseases			0.000			0.000

No	87.04%	94.33%	87.57%	94.70%	
Yes	12.96%	5.67%	12.43%	5.30%	
Stomach or other digestive diseases			0.000		0.000
No	63.63%	77.71%	64.77%	78.41%	
Yes	36.37%	22.29%	35.23%	21.59%	
Arthritis or rheumatism			0.000		0.000
No	51.30%	66.23%	50.37%	67.25%	
Yes	48.70%	33.77%	49.63%	32.75%	

1 Note: Consolidated observations in the top 10% and top 20% medical expense groups from 2013 to 2018 were
2 defined as the “Top 10%” and “Top 20%” medical spending groups, respectively.

3

4 **Table 4** Persistence of total medical expenses during 2013, 2015, and 2018

Expense Ranking in 2013	Expense ranking in 2015			Expense ranking in 2018		
	Top10%	Top20%	Top50%	Top10%	Top20%	Top50%
Top 10%	0	70.00%	80.00%	0	70.00%	70.00%
Top 20%	0.26%	30.82%	55.03%	0.40%	26.46%	55.29%
Top 50%	0.28%	21.48%	46.42%	0.22%	20.64%	50.00%

5

6 **Table 5** Heckman probit coefficients: Regression on incidence of medical expenses

	(1)		(2)	
	dy/dx	Std. Err.	dy/dx	Std. Err.
Lagging item of with or without expenses (ref. without)				
L1	0.165***	0.010	0.158***	0.015
L2			0.117***	0.015
Demographic variables				
Age (ref. 45–55)				
55–65	-0.009	0.011	-0.013	0.016
65–75	0.007	0.012	0.011	0.017
≥75	0.019	0.017	0.015	0.023
Gender (ref. male)	0.033***	0.008	0.027**	0.012
Education (ref. less than lower secondary)				
Upper secondary, vocational training	-0.012	0.014	-0.022	0.020
Tertiary	-0.030	0.063	-0.016	0.096
Marriage (ref. not)	-0.001	0.013	-0.003	0.018

Employ (ref. not)	0.006	0.014	-0.021	0.019
Residence (ref. rural)	0.008	0.012	-0.017	0.018
Lagging item of with or without expenses*disease (ref. without)				
L1*Hypertension	0.102***	0.014	0.069***	0.025
L1*Diabetes	0.098***	0.030	0.071	0.053
L1*Cancer	0.056	0.064	0.293***	0.127
L1*Chronic lung diseases	0.078***	0.021	0.081**	0.035
L1*Liver diseases	0.049*	0.030	-0.036	0.050
L1*Heart diseases	0.074***	0.020	-0.009	0.036
L1*Kidney diseases	0.089***	0.026	0.115***	0.044
L1*Stomach or other digestive diseases	0.055***	0.014	0.020	0.023
L1*Arthritis or rheumatism	0.060***	0.012	0.016	0.020
L2*Hypertension			0.103***	0.025
L2*Diabetes			0.013	0.055
L2*Cancer			-0.037	0.108
L2*Chronic lung diseases			-0.018	0.037
L2*Liver diseases			0.111**	0.055
L2*Heart diseases			0.020	0.038
L2*Kidney diseases			0.014	0.047
L2*Stomach or other digestive diseases			0.035	0.024
L2*Arthritis or rheumatism			0.036*	0.021

1 Note: *** indicates p value <0.01, ** indicates p value <0.05, * indicates p value <0.1; (1) and (2) indicate lag 1 and
2 lag 2, respectively.

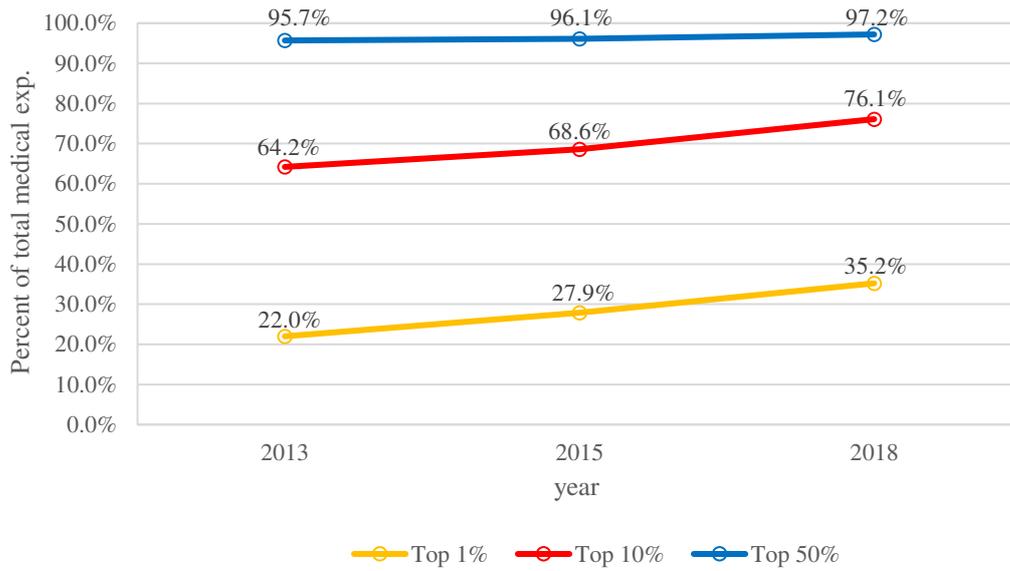
3

4 **Table 6** Heckman regression coefficients: Regression on total medical expenses

	(1)		(2)	
	Coefficient	Std. Err.	Coefficient	Std. Err.
Lagging item of medical expenses				
L1 of medical exp.(ln)	0.225***	0.018	0.165***	0.022
L2 of medical exp.(ln)			0.113***	0.021
Predisposing factors				
Age (ref. 45–55)				
55–65	0.023	0.101	0.100	0.179
65–75	-0.002	0.113	0.229	0.187
≥75	0.088	0.160	0.447*	0.249
Gender (ref. male)	0.328***	0.084	0.297**	0.132

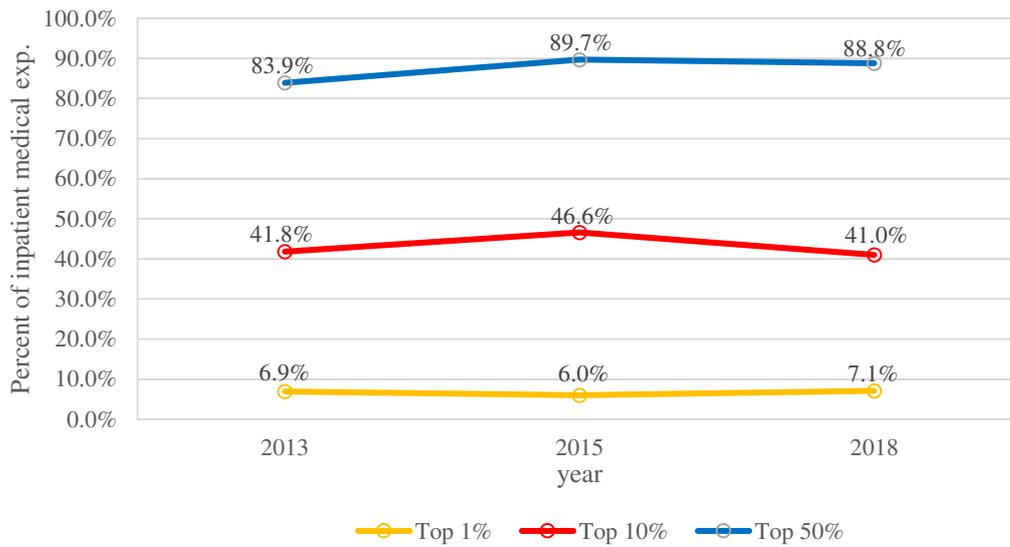
Education (ref. less than lower secondary)				
Upper secondary, vocational training	0.089	0.131	-0.043	0.198
Tertiary	0.637	0.622	0.904	1.289
Marriage (ref. not)	0.258**	0.123	0.252	0.189
Employ (ref. not)	-0.045	0.148	-0.164	0.208
Residence (ref. rural)	0.221*	0.124	0.155	0.206
Enabling factors				
Number of outpatient visits (ref. ≤24)				
24–36	0.331***	0.097	0.203	0.152
≥36	0.827***	0.097	0.684***	0.150
Number of hospitalizations(ref. ≤1)	1.067***	0.142	0.911***	0.224
Income (ref. ≤ ¥8,000)				
¥8,000–15,600	0.042	0.157	0.009	0.254
¥15,600–30,000	0.284*	0.166	0.391*	0.227
≥ ¥30,000	0.307*	0.167	0.330	0.239
Health insurance (ref. no)	2.734***	0.294	2.012***	0.344
Type of outpatient medical facilities (ref. general, specialized, Chinese medicine hospital)				
Community healthcare center, township hospital, village clinic	-1.167***	0.088	-1.151***	0.144
Multi-type outpatient facility visits (ref. no)	0.050	0.126	0.317	0.194
Need factors				
Hypertension	0.383***	0.103	0.315**	0.154
Diabetes	0.408***	0.137	0.328*	0.173
Cancer	0.088	0.320	0.535	0.489
Chronic Lung diseases	0.220**	0.107	-0.003	0.151
Liver diseases	0.013	0.153	-0.091	0.208
Heart diseases	0.346***	0.108	0.308**	0.150
Kidney diseases	0.204	0.134	0.024	0.181
Stomach or other digestive diseases	0.057	0.095	0.029	0.144
Arthritis or rheumatism	0.197**	0.095	0.222	0.146
Comorbidity (ref. ≤1)	0.274**	0.131	0.654***	0.197
Self-reported health status (ref. very good)				
Good	1.334***	0.259	1.821***	0.418
Fair	1.149***	0.202	1.784***	0.353
Poor	1.388***	0.209	1.811***	0.368

1 Note: *** indicates p value <0.01, ** indicates p value <0.05, * indicates p value <0.1 ; (1) and (2) indicate lag 1
2 and lag 2, respectively.



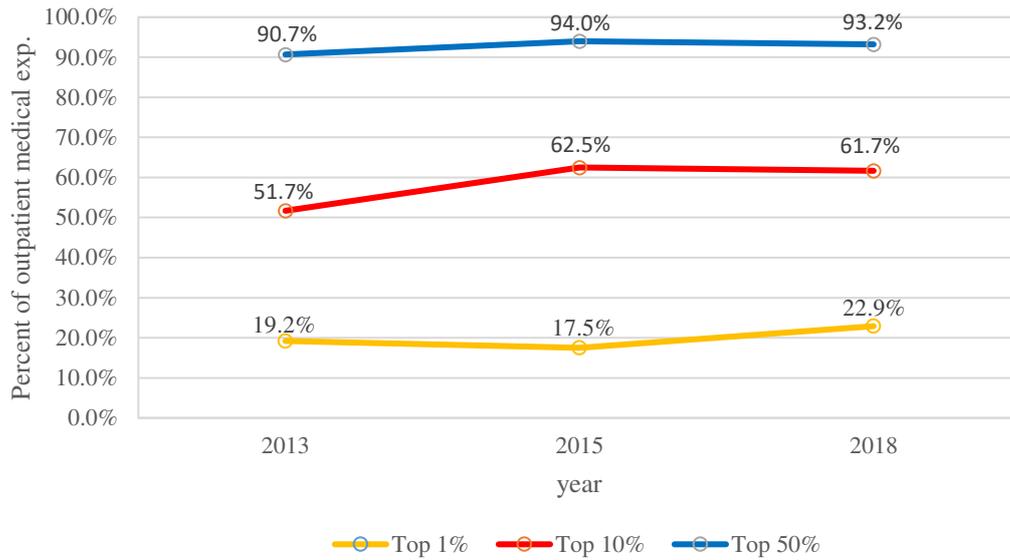
1
2
3

Fig. 1 The concentration of total medical expenses by year



4
5
6

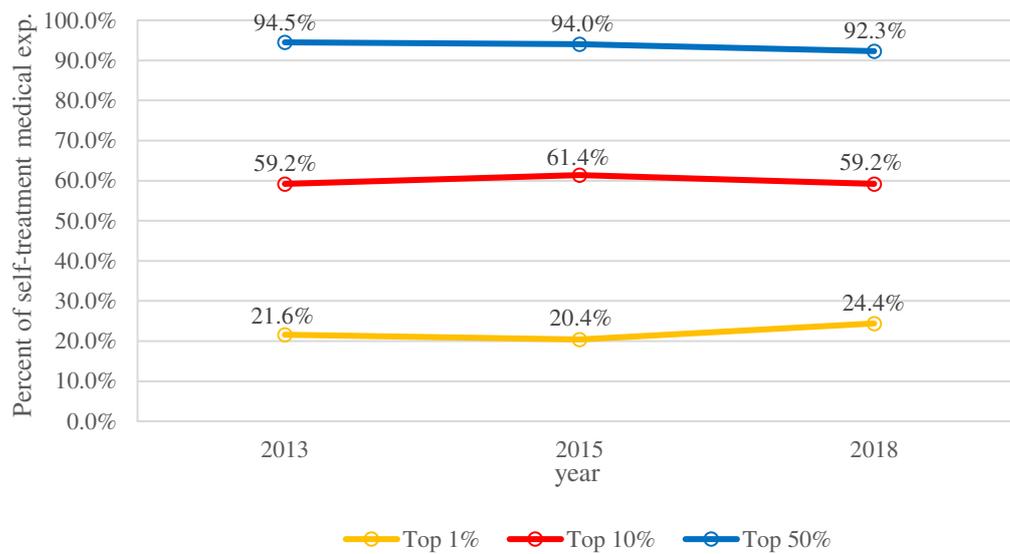
Fig. 2 The concentration of inpatient medical expenses by year



1

2 **Fig. 3** The concentration of outpatient medical expenses by year

3



4

5 **Fig. 4** The concentration of self-treatment medical expenses by year