Forecasting demand for long-term care based on multistate piecewise constant Markov process

Cui Xiaodong✉*, Zheng Yuhua

Nanjing Xiaozhuang University, Nanjing, China

Article Info

Abstract

This paper forecasts demand for long-term care based on multistate piecewise constant Markov process. Two types of data are mainly used in this study. The first type of data came from the Chinese Longitudinal Healthy Longevity Survey (CLHLS), 2008-2011. The second type of data came from the China's Demographic Yearbook of 2016, used as the number of population in the base period. It finds that the changes in health have a significant difference in gender and age. It also finds that under different health states, the distribution of duration of staying in each state in different gender and age groups shows a characteristic similar to the distribution of population size.

Introduction

Long-term care means *"providing care, including medical care, social care, home care, transportation or other support services for people with a chronic disease or dysfunction under disability within a long period"*[1]. Due to illness, disability, function decline or other reasons, the aged population becomes the main group with a demand for long-term care. According to the sixth national census data, the ratio of the elderly aged 60 and above who cannot care for themselves in China is nearly 3% (Bao, 2015). Moreover, with the increase of life expectancy and the decline of fertility level, China's aging population shows a high-speed and advanced-age trend. The aging rate is averagely increased by 0.34 every year from 2010 to 2025, and the ratio of the aged population is increased from 8.9% to 14%, which reflects a serious aging problem (Wei, Shanjun, & Chen, 2014). In the meantime, aging is associated with family miniaturization. According to the sixth national census data, the average number of members in each family is 3.10 in China, and the average number of members in each urban family is only 2.87. However, the number of members in each urban family was 3.50 in 1990 (Hongwei, 2015). Family miniaturization causes the weakening of traditional family-based elderly support function and the aging of population, and the combination of the two brings a severe challenge to China's pension system. As China's current social welfare and social security system are unable to solve the long-term care problem for most elderly people, the establishment of a long-term care security system is an inevitable choice for coping with aging. So far, China's long-term care security system is in the exploration stage. This study is aimed to research the actuarial basis of the long-term care security system, i.e., the evolvement rule and transition probability of health state, and to forecast such issues as the population size required for long-term care, and the long-term care time in China in the future so as to provide experience for exploring China's long-term care security system for the elderly.

The forecasting methods of the demand for long-term care are divided into macro scenario simulation forecasting and micro simulation data forecasting. Macro scenario simulation means constructing scenarios by setting different parameters of key factors, then speculating the demands for long-term care in different scenarios. There is a difference in the factor selection and setting of different documents. Overall, factor selection includes the factors reflecting macro socio-economic state and the factors at the micro individual level. Comas-Herrera et al. (2006) took the lead in applying a simulation method to forecast the demand for long-term care in four European countries, and found that the demand for long-term care is highly sensitive to

---

[1] From the definition given by the Health Insurance Association of America

*Corresponding author e-mail: cui_xiaodong@163.com

living standard, health state, nursing cost, as well as the availability of family or friend care; (Costa-Font et al., 2008) used the data provided by the actuarial department to forecast the demand for long-term care from 2005 to 2031 in his research on the UK, and found that the selection of such parameters as marital state, family members' dwelling state and degree of disability has a great influence on the forecasting. Costa-Font et al. (2008) constructed the PSSRU model, and forecasted the population with a demand for long-term care in the future by simulating the number of population under different circumstances and multiplying it by the ratio of the current population under long-term care to the total population; Minglai, (2009) and Yi, Huashuai, & Zhenglian (2012) speculated the number of the elderly needing long-term care based on the same method. Zeng Yi utilized the geometric growth rate of long-term care population on a period-on-period basis (Yi et al., 2012). The key factors of this method in simulation are birth rate and death rate. The results of macro simulation greatly rely on the selection and setting of parameters. Comas-Herrera et al. (2012) used the analytic hierarchy process to verify this conclusion by selecting different factors.

Micro simulation forecasting is based on micro survey data to follow up the evolution track of health state of the elderly. First, the transition probability matrix of different health state is constructed, and then the Markov process is utilized for forecasting. The key of this method is the construction of the transition probability matrix and the application of Markov process. Rickayzen & Walsh (2002) took the lead in applying Markov process to the forecasting of demand for long-term care, and divided the calculation of the transition probability matrix into two parts. First, the ratio of population under each state at the beginning of the period to the total population is used as the base period probability; then the base period probability is adjusted according to the variation trend of health state to obtain the probability of the next year so as to calculate the transition probabilities of the forecast period in sequence. Hare, Alimadad, Dodd, Ferguson, & Rutherford (2009) obtained the transition probability by fitting through the construction of a regression model based on the factors influencing the demand for long-term care in his research on the long-term care of the UK, and his research data came from medical departments and individual surveys. Based on the same regression method, Chahed, Demir, Chaussalet, Millard, & Toffa (2011) used the health state transition data of patients in London from 2008 to 2011, and Peng, Ling, & Qun (2010) utilized the health state survey data of China's aged population to construct the transition probability matrix of corresponding regions, and forecasted the elderly people's demand for long-term care under Markov time invariant hypothesis; In the research by Rong (2009) in China, the transition probability is borrowed from the national long-term care survey & research results of the USA; Feng & Chunjie (2012) and Hongwei (2015) respectively established a mortality probability model and a health state transition model through logit and ordinal logit regression models, and assumed that the health transition probability matrix is unchanged during the forecasting when applying the Markov process method.

In the comparison between the two forecasting methods, as macro scenario simulation is restricted by the selection and hypothesis of numerous key factors, the forecasting model based on micro data is relatively dominant in the case of data available. The "Chinese Longitudinal Healthy Longevity Survey" (CLHLS) program conducted by the Center for Healthy Aging and Development Studies, Peking University, has made some progress for the domestic research based on micro forecasting, but still has some room for improvement in application. First, in the aspect of the construction of the transition probability matrix, the direct borrowing of the overseas transition probability or the simple period-on-period method has low precision, which cannot meet the design requirement of the insurance system. The multiple regression model construction method is a more common method. But the complicated, diversified and multidimensional characteristics of changes in health state add diversity and subjectivity to the selection of explanatory variables in the regression model. Therefore, in the case of data available, only constructing the transition probability based on the tracking data can better capture the dynamic evolution of the health state of the elderly. Second, in the aspect of the time invariant hypothesis of Markov process, the nature of time invariant hypothesis is that the transition intensity is an invariant constant, and the transition probability is only related to time interval, but unrelated to its time point. This hypothesis is convenient for the expression of the transition probability function. However, the changes in health state are related to age. People of different ages have different changes in health, i.e., the transition probability is related to the time point. Thus, the time invariant hypothesis of Markov process is improper. Third, there are only few researches on the forecasting of time required for long-term care, and the expressions in existing literature are unclear, but the time required for long-term care and the population size thereof are also the key to forecasting the cost of long-term care.

In this regard, based on the micro simulation forecasting method, the changes in health state before and after the observation period of samples in the CLHLS survey program are tracked in this paper, and the

actuarial method is adopted to construct the transition probability matrix to avoid the subjectivity of variable selection in regression analysis; the Markov process method with transition intensities as piecewise constants and matrix multiplication by age is adopted to forecast the number of long-term care population, and to overcome the defect of Markov time invariant hypothesis not in conformity with the reality; on this basis, the expected time required for long-term care is forecasted based on the actuarial theory of life insurance. Considering that the micro simulation forecasting loses the reliability of medium and long-term forecast due to model refinement and forecast precision advantages, this paper is focused on forecasting the demand size of ten-year long-term care and the time duration required for long-term care.

Research Method

Data source and concept definition

Two types of data are mainly used in this study. The first type of data came from the CLHLS (2008-2011). This survey was carried out in 22 provinces of China, and related data were collected by means of questionnaire interviews in each household. Hundreds of scholars at home and abroad have registered for free use of CLHLS, the data quality of which is good and has been widely accepted by the academic circles. In this study, the elderly people aged 60 and above in the tracking survey samples were the objects of study. After the observed value lacking key information and the data unavailable for follow-up tracking were removed, the number of effective samples were 15,964, and the male to female ratio was 0.93:1. The CLHLS data were used to calculate the transition probability matrix and reckon the ratio of the population in each health state to the total population in the base period. The second type of data came from the China's Demographic Yearbook of 2016, used as the number of population in the base period.

In this study, the definition of nursing state is subject to the ADLs (activities of daily living) independence criterion widely used in the academic circles. According to the criterion, in case that a person needs help from others in one or more of the six items reflecting the activities of daily living of the elderly (bathing, dressing, indoor mobility, getting to the toilet, self-feeding, urine & feces control), it is defined as self-care disability; if a person has three or more daily activity disorders, such person is deemed to need long-term care. Besides, considering that one of the objectives of this study is providing a basis for the long-term care security system, cognitive impairment is also defined as self-care disability so as to enable the definition of long-term care state to be consistent with the long-term care security system.

In CLHLS, the respondents' health state is investigated from multiple dimensions. In addition to the six items of ADLs, the eight indexes reflecting the instrumental activities of daily living (IADLs) of the elderly are collected as well, including dropping in on neighbors, 5kg weightlifting, bathing, cooking, and so on. If a respondent is unable to complete a daily activity without the sustained help from others, such respondent is deemed to lose the activity of daily living. Moreover, in the survey, the Chinese version of simple cognitive function scale conforming to China's cultural traditions and socio-economic state is used to measure the cognitive function of the elderly. The simple scale includes such aspects as direction positioning ability, reaction ability, attention, calculative ability, recalling ability, as well as language, understanding and self-coordination ability, and covers 24 questions and 30 points in total. The scoring method is used in CLHLS, wherein, the ability to complete an activity or a correct answer will add 1 point for the respondent.

To sum up, health state is divided into four types in this study, respectively, 1 health, 2 health impairment, 3 dysfunction and 4 death, among which, 1, 2 and 3 belong to transition states, i.e., the states are interchangeable; 4 is the absorption state, i.e., 1, 2 and 3 can be transferred to 4, but 4 cannot be transferred to others. The definition of health state is based on three indexes, namely, ADLs, IADLs and cognitive ability. If a person has no disorder in none of the three, such person is deemed to in a health state 1. If a person has three or more than three daily activity disorders, i.e., the ADLs score is larger or equal to 3 points or the cognitive function score is below 16 points (30 points in total), the person is deemed to be in the state of dysfunction 3, namely, the state of needing long-term care. The complementary state is deemed as health impairment state, i.e., a certain disorder exists, but does not reach the state 2 of needing long-term care.

Mechanism analysis of Markov process

With the help of the multi-state model in long-term health insurance actuarial science, a four-state probability transition model is adopted pursuant to the above definition in this paper. Defining random process $\{S(x). x \geq 0\}$: the value of $S(x + t)$ represents the health state when the variable is $x + t$, and the health state space is $\{1, 2, 3, 4\}$, wherein, 4 is the absorption state.

Defining transition probability: $_t p_x^{ij} = p\{S(x+t) = j \mid S(x) = i\}; j = 1,2,3,4; i = 1,2,3$, it represents that the state is i when at time x; it is transformed into j probability at time $x + t$; it is deemed as one-step transition probability $_1 p_x^{ij}$ when $t = 1$. The treatment method of the multi-state transition probability model is the Markov process.

Markov process

A Markov process refers to a set of random variables $\{X(t), t \in T\}$ relying on the variable parameter t, the set T of all possible values of the variable parameter t is known as a parameter space, and the value S of $X(t)$ constitutes the state space of a stochastic process. If it is known that the time t system is under a state condition, when the state of the time $\tau(\tau > t)$ is unrelated to the previous state of the time t, this process is a Markov process.

In this paper, the four-state probability transition model can be regarded as a Markov process with T discrete and S discrete. T is a set of ages, and S represents four health states. In other words, the transition probability of health states is only related to the state before transition, which can be expressed as the following mathematical symbol:

$$p\{S(x+t) = h_t \mid S(x+t-1) = h_{t-1}, S(x+t-2) = h_{n-2}, \text{L}, S(x+1) = h_1\}$$
$$= p\{S(x+t) = h_t \mid S(x+t-1) = h_{t-1}\} \tag{1}$$

For all $t \in T, h_t \in S$ is satisfied.

Transition probability matrix and transition intensity matrix

States 1, 2 and 3 are transferable states while state 4 is the absorption state. The transition probability meets the Chapman-Kolmogorov equation:

$$_2 p_x^{ij} = \sum_k {_1 p_x^{ik}} {_1 p_{x+1}^{kj}}, \, _3 p_x^{ij} = \sum_k {_2 p_x^{ik}} {_1 p_{x+2}^{kj}}, \text{L}$$
$$_t p_x^{ij} = \sum_k {_{t-1} p_x^{ik}} {_1 p_{x+t-1}^{kj}} = \sum_k {_1 p_x^{ik}} {_1 p_{x+1}^{ik}} {_1 p_{x+2}^{ik}} \text{L} {_1 p_{x+t-1}^{kj}} \tag{2}$$
$$i = 1,2,3; k = 1,2,3; j = 1,2,3,4$$

For different x, i.e., different ages in this paper, with different transition probabilities, the transition probabilities of different ages are expressed as matrixes. The one-step transition probability is denoted as matrix $P(1)$; the t-step transition probability matrix is $P(t)$. According to Equation (2),

$$P(t) = P^{(m)}(1) . P^{(m+2)}(1) . \dots P^{(m+t-1)}(1). \tag{3}$$

Where, $P^{(m+t-1)}(1)$ represents the one-step transition matrix at time $m + t - 1$.

When the transition probability is only related to the initial state i and the arrival state j, but unrelated to the initial time x, the Marko chain has a stationary transition probability, also known as time-homogeneous transition probability.

Transition intensity represents the instantaneous transition of state, which is denoted as:

$$\sim^{ij}(t) = \lim_{\Delta t \to 0} \frac{p^{ij}(t + \Delta t) - p^{ij}(t)}{p^{ij}(t) \cdot \Delta t} = \frac{(p^{ij}(t))'}{p^{ij}(t)} \tag{4}$$

From this, the relation between transition intensity and transition probability can be inferred:

$$_t p_x^{ij} = e^{\int_x^{x+t} \sim^{ij}(s)ds} \tag{5}$$

When the transition intensity is the constant $\sim$: $_t p_x^{ij} = e^{-t}$. Under the time invariant hypothesis, transition probability and transition intensity meet Kolmogorov forward differential equation:

$$\frac{d}{dt} {_t p_x^{ij}} = \sum_{l=1}^{s} {_t p_x^{ih}} u^{hj}(x+t), \text{ see the appendix for the specific equation.}$$

Markov process method with transition intensity as piecewise constant

The nature of the Markov process time invariant hypothesis is that the transition intensity is an invariant constant, i.e., the time interval of state transition is subject to exponential distribution, and the state transition probability is only related to the time interval, but unrelated to the time point. This hypothesis is convenient for the expression of a transition probability function, but is improper in many applications. For example, if we assume that the health state transition probability is time-homogeneous, it means that no matter 60, 70 or 80, as long as the time intervals of transition observed are the same, the health state transition probabilities are the same. In other words, the people aged 60, 70 or 80 have the same changes in health state within the following equal time, which is apparently not in conformity with the reality. Therefore, based on the non-homogeneous characteristic of transition intensity varying with age, the forecasting method with transition intensity as a piecewise constant is adopted so as to maintain the easy controllability of constant transition intensity in this paper. That is to say, a given forecast interval (s,t) is divided into m short intervals. We assume that the transition intensity of each short interval is a constant, but the transition intensities of the m short intervals are different. In the process of calculating the transition probabilities of (s,t) the transition intensity of each short interval is calculated first, then the transition intensity of this short interval, finally the transition probability matrix of the entire interval.

A ten-year transition probability matrix is forecasted in this paper. It is divided into ten intervals on a yearly basis. We assume that the transition intensity every year is a constant[2], but the transition intensities in different years are different. Then according to Equation (3), the future ten-year transition probability matrix can be expressed as:

$$P(0,10) = P^0 P^1 \dots P^9 \qquad (6)$$

It should be noted that unlike the time invariant hypothesis, $P^0, P^1 \dots$ here, i.e., the transition probabilities of all intervals, are unequal to one another. Theoretically, we might track the health state transition of samples every year to obtain ten section matrixes for multiplication, but it is actually infeasible and unnecessary for short and medium-term forecasting[3]. For the section matrix of health transition of each age group obtained by the utilization of samples, dislocation multiplication is applied during forecasting. For the forecasting of the future 10-year health transition probability of people aged (x), we assume that each year is a short interval, then we can apply the transition probability of age (x) in the section data to the first interval, the transition probability of age $(x+1)$ in the section data to the second interval, the transition probability of age $(x+2)$ in the section data to the third interval. It should be also noted that P^0, P^1 etc. are one-year transition probabilities, and the three-year transition probabilities are obtained according to CLHLS survey data from 2008 to 2011. Therefore, the one-year transition intensities and one-year transition probabilities need to be obtained pursuant to Equation (5), and then substituted into Equation (6).

Expected time of state transition and duration

State duration means the time $T(x)$ of state i transitioning to and remaining in state. The expected time is expressed as $E(T(x))$. For example, if it is forecasted that a person aged 60 is in a health state, then the expected time of the state transitioning to and remaining in disability state from age 60 to 70 is $E(T(60))$. According to the actuarial theory[4], the value can be given by the following equation (for more specific derivation, see the annex):

$$
\begin{aligned}
E(T(60)) &= \int_0^{10} t \times {}_t p^{ij}_{60} \times \sim^{ij}_{60+t} dt \\
&= \int_0^{10} {}_t p_{60}{}^{ij} dt = \sum_{x=60}^{69} \int_x^{x+1} {}_t p_{60}{}^{ij} dt \\
&= \sum_{x=60}^{69} \int_x^{x+1} \sum_{h=1}^{k} ({}_{x-60} p_{60}{}^{ih} {}_{t-x} p_x{}^{hi}) dt \\
&= \sum_{x=60}^{69} \sum_{h=1}^{k} {}_{x-60} p_{60}{}^{ih} \int_x^{x+1} {}_{t-x} p_x{}^{hi} dt
\end{aligned}
\qquad (7)
$$

Where, x is a nonnegative integer, $i = 1,2,3; k = 1,2,3; j = 1,2,3,4,$

---

[2] As long-term care often results from chronic and long-term diseases, it is reasonable to set the health transition probability within one year as a constant (the explanation comes from *Health Insurance*, exam book required by the China Association of Actuaries).
[3] In the short to medium-term, the macro-environment has an insignificant impact on health (Zeng Yi).
[4] See *Actuarial Principles*

To simplify the above equation, the aforesaid hypothesis is followed, i.e., the transition intensity or transition probability is a constant within one year, and initial and final states are only considered in state changes within one year, irrespective of the mutual transformation state in the middle period. For instance, if the initial state is 1, the mid-year state is changed to 2 and the year-end state is changed to 3, then we can denote the state changes as 1 to 3[5]. Hence, in the following equation:

$$_{x-60}p_{60}^{ih} = \prod_{t=0}^{x-60-1} {}_1p_{60+t}^{ih} \, , \quad \int_{x}^{x+1} {}_{t-x}p_{x}^{hi}dt = {}_1p_{x}^{hi} \tag{8}$$

When Equation (8) is substituted into Equation (7), Equation (7) can be simplified to:

$$T = \int_{60}^{70} {}_tP_{60}^{ij}dt = \sum_{x=60}^{69}(\prod_{t=0}^{x-60} p_{60+t}^{ij}) \tag{9}$$

Equation (9) is applicable to the estimation of average time of any age, health state and forecast interval.

Results and Discussion

Three-year health state transition probability matrix

As the sample data are three-year tracking survey data, the three-year health state transition probability matrix is calculated first. Table 1 is the three-year health state transition probability matrix of different ages and genders completed with SPSS software based on the CLHLS survey data from 2008 to 2011. In the process of sorting data, considering that age and gender are important influencing factors of health state, first, the samples are classified by age and gender. Every five years is an age group (if every age is a group, more accurate results will be obtained, but there will be too many data. Moreover, when there is only a small age difference, the transition probability difference will be insignificant). Each individual of each category is classified as different health state according to the definition standard of health state, then, the health states of individuals in different states at the end of the period are tracked in each category so as to calculate the ratio of the number of people in each state at the end of the period to that at the beginning of the period, which can be counted as the corresponding three-year transition probability. As shown in Table 1, the samples are classified into different groups by age and gender, and each group is divided into different state. The data at the intersection represents the corresponding state transition probability. For instance, among the elderly men aged 60-64 who are in State 1 (i.e., health) at the beginning of the period, the ratio of the elderly men remaining in State 1 at the end of the period is 0.7867; the ratio of the elderly men in State 2 is 0.14; the ratio of the elderly men in State 3 is 0.0067; the ratio of the elderly men in State 4 (i.e., death) is 0.0666. Similarly, among the elderly men aged 60-64 who are in State 2 at the beginning of the period, the ratio of the elderly men in State 1 at the end of the period is 0.6444; the ratio of the elderly men in State 2 is 0.3011; the ratio of the elderly men in State 3 is 0.0222; the ratio of the elderly men in State 4 (i.e., death) is 0.0323. The remaining data have the same meaning.

In order to observe the impact of age and gender on health state in a more intuitive way and support the grouping standard in this paper, Fig.1 and Fig.2 are drawn according to the data in Table 1. Seen from the trend of each figure, no matter how the initial state is, disability probability and death probability both progressively increase with age[6]. The decline in people aged above 90 in Fig. 1b and the people aged above 85 in Fig. 1c is due to the sharp increase of the mortality risk in the period. This figure re-illustrates that it is inconsistent with the reality to set the transition probability as a constant in the entire forecast period. Seen from all figures, there is a significant difference in the transition probabilities between men and women. In Fig.1, the disability transition probability of men in the same initial state and the same age group is smaller than that of women, especially the men aged 75-79 in Fig.1b and the men aged 70-84 in Fig.1c, who have a significant health advantage. As shown in Fig.2, the death probability of men is larger than that of women on the whole, especially the people aged 85-89 in Fig. 2b and the people aged 70-79 in Fig. 2c, who have a significant difference. This phenomenon forms a complementary explanation with Fig.1, with a large disability probability and a small death probability; a small disability probability and a large death probability respectively. Certainly, the two are not linearly related or equal to 1. But, no matter in State 2 or 3, the people

---

[5] From *Health Insurance*, exam book required by the China Association of Actuaries
[6] It can be used to explain why the Markov process does not meet the time invariant hypothesis as abovementioned.

in the age group have a fairly low possibility of transition to health. Thus, it can be concluded: the intensified aging, increased aged population, prolonged life expectancy and increased ratio of aged population in the future will certainly bring an increase of the demand for long-term care; compared to men, women have a high disability probability and a low death rate, resulting in a severer situation of women's demand for long-term care. The future demand size for long-term care is our next content. Due to the restriction on forecasting hypothesis, we only forecast the demand for long-term care in the next 10 years.

Table 1 Three-Year Health State Transition Matrix

Age	State	men				women			
		1	2	3	4	1	2	3	4
[60 64]	1	0.7867	0.1400	0.0067	0.0666	0.7500	0.2065	0.0326	0.0109
	2	0.6444	0.3011	0.0222	0.0323	0.4837	0.4140	0.0943	0.0080
	3	0.3902	0.3698	0.1800	0.0600	0.3710	0.3720	0.2500	0.0070
[65 69]	1	0.6754	0.2403	0.0237	0.0606	0.5333	0.3767	0.0500	0.0400
	2	0.3869	0.4201	0.1003	0.0927	0.3768	0.4589	0.1063	0.0580
	3	0.2010	0.4000	0.2490	0.1500	0.1923	0.4231	0.2692	0.1154
[70 74]	1	0.5279	0.2817	0.0711	0.1193	0.5233	0.2746	0.1451	0.0570
	2	0.3948	0.3542	0.1107	0.1403	0.2073	0.5793	0.1189	0.0945
	3	0.2364	0.2545	0.1455	0.3636	0.1837	0.1429	0.4286	0.2448
[75 79]	1	0.3202	0.3900	0.0969	0.1929	0.3006	0.4247	0.1561	0.1186
	2	0.1968	0.4124	0.1860	0.2049	0.1654	0.3780	0.2703	0.1863
	3	0.0625	0.2813	0.2188	0.4375	0.0645	0.2903	0.3710	0.2742
[80 84]	1	0.3196	0.3144	0.1443	0.2217	0.1900	0.4364	0.2216	0.1520
	2	0.1119	0.3776	0.2005	0.3100	0.0502	0.3480	0.3542	0.2476
	3	0.0518	0.2798	0.2642	0.4042	0.0190	0.2531	0.3742	0.3537
[85 89]	1	0.0755	0.4854	0.1946	0.2445	0.0456	0.4154	0.3051	0.2339
	2	0.0363	0.3189	0.2104	0.4344	0.0297	0.3051	0.3729	0.2923
	3	0.0178	0.1684	0.2579	0.5559	0.0117	0.1202	0.3806	0.4875
[90 94]	1	0.0339	0.3266	0.2358	0.4037	0.0385	0.1923	0.3077	0.4615
	2	0.0149	0.1983	0.2593	0.5275	0.0045	0.2123	0.4234	0.3598
	3	0.0101	0.0911	0.2376	0.6612	0.0058	0.0764	0.3299	0.5880
95+	1	0.0833	0.2500	0.2500	0.4167	0.0156	0.0940	0.4005	0.4899
	2	0.0185	0.1806	0.2454	0.5556	0.0107	0.1584	0.3083	0.5226
	3	0.0033	0.0542	0.1921	0.7504	0.0011	0.0272	0.2484	0.7233

Fig.1a. from State 1 to 3   Fig.1b. from State 2 to 3   Fig.1c. from State 3 to 3

Figure 1. Three-Year Transition Probabilities from Different States to Disability State

Fig.2a. from State 1 to 4   Fig.2a. from State 2 to 4   Fig.2a. from State 3 to 4
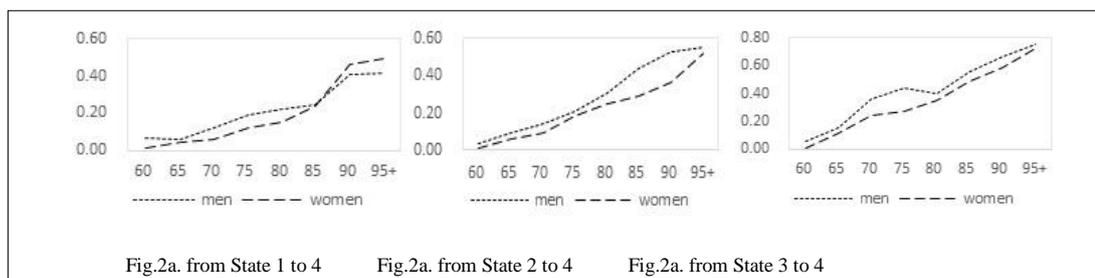
Figure 2. Three-Year Transition Probabilities from Different States to Death State

One-year transition probability matrix

The three-year transition probability matrix $P(3)$ is obtained in Table 1. To calculate one-year transition probabilities, we may conduct a staged time invariant hypothesis, i.e., assuming that the transition probability intensity of each category within the three years is a constant to reversely calculate $P(1)$ according to Equation (3). In this paper, mathematic programming is utilized to calculate the transition intensity, then to calculate the one-year transition probability matrix. Due to limited space, the matrix is not listed in the text. Please refer to Table 1 in the Appendix.

Ten-year transition probability matrix of piecewise constant transition intensities

In accordance with the one-year transition probability matrix, the transition probability in each category and different state after any period can be calculated via Equation (5). However, due to the complexity of health state transition, only ten-year forecasting is carried out in this study, and each five years is still used as an age unit in each category, e.g., the forecasting of probabilities of healthy men (State 1) aged 60-64 in States 1, 2, 3 and 4 ten years later. As each age group is on a five-year basis, with a forecast period of ten years, the forecast period is divided into two parts. The transition probability of the age group in Table 2 is applied in the first 5 years, and the transition probability of the next age group is applied in the last 5 years. For instance, for the forecasting of the ten-year transition probability matrix of people aged 60-64, the transition probability every year in the first five years is the one-year transition probability of the age group of 60-64; then 5 years later, the people aged 60-64 enter the age group of 65-69, so, the one-year transition probability matrix of the age group of 65-69 is adopted in the last 5 years[7]. In the meantime, it is assumed that the transition probability matrix every year in the five years meets time homogeneity. $P_{60\sim64}(1)$ represents the one-year transition probability of the age group of 60-64; $P_{65\sim69}(1)$ represents the one-year transition probability of the age group of 65-69. Then, the ten-year transition probability of the age group of 60-64 is $P_{60\sim64}(10) = (P_{60\sim64}(1))^5 \cdot (P_{65\sim69}(1))^5$. The calculation methods of the remaining groups are the same. The people aged 95+ have a fairly low survival rate, so it is assumed that they are in State 4. In this way, the ten-year transition probability matrix of each group is obtained by means of programming calculation, as shown in Table 2. Seem from the table, the ratio of healthy men in the age group of [60 64] who remain healthy ten years later is 0.4980; the ratio of health impairment is 0.2514; the probability of dysfunction is 0.0441; the death probability is 0.2065. The remaining groups have the same meaning.

Table 2. Ten-Year Health State Transition Probability Matrix

Age	State	men				women			
		1	2	3	4	1	2	3	4
[60 64]	1	0.4980	0.2514	0.0441	0.2065	0.4262	0.3925	0.0830	0.0983
	2	0.5096	0.2656	0.0488	0.1761	0.4166	0.3950	0.0883	0.1000
	3	0.4729	0.2667	0.0551	0.2054	0.4077	0.3958	0.0930	0.1036
[65 69]	1	0.3710	0.2475	0.0745	0.3070	0.2935	0.3382	0.1546	0.2137
	2	0.3428	0.2338	0.0718	0.3516	0.2777	0.3295	0.1533	0.2395
	3	0.3049	0.2124	0.0666	0.4161	0.2434	0.2995	0.1483	0.3088
[70 74]	1	0.1484	0.2588	0.1072	0.4856	0.1388	0.2882	0.2064	0.3666
	2	0.1405	0.2476	0.1038	0.5082	0.1253	0.2691	0.2003	0.4053
	3	0.0995	0.1772	0.0751	0.6482	0.0911	0.2039	0.1598	0.5452
[75 79]	1	0.0806	0.1816	0.1096	0.6282	0.0361	0.2072	0.2221	0.5347
	2	0.0730	0.1717	0.1051	0.6503	0.0300	0.1818	0.1995	0.5887
	3	0.0463	0.1148	0.0714	0.7675	0.0239	0.1537	0.1729	0.6495
[80 84]	1	0.0151	0.1274	0.1004	0.7570	0.0109	0.1104	0.1965	0.6822
	2	0.0119	0.1015	0.0832	0.8033	0.0086	0.0878	0.1633	0.7402
	3	0.0096	0.0821	0.0690	0.8393	0.0071	0.0722	0.1366	0.7842
[85 89]	1	0.0039	0.0431	0.0734	0.8796	0.0022	0.0436	0.1321	0.8221
	2	0.0027	0.0305	0.0525	0.9142	0.0019	0.0377	0.1156	0.8448
	3	0.0020	0.0219	0.0384	0.9377	0.0013	0.0242	0.0759	0.8986
[90 94]	1	0.0019	0.0177	0.0359	0.9444	0.0006	0.0095	0.0445	0.9455
	2	0.0014	0.0128	0.0264	0.9595	0.0006	0.0106	0.0507	0.9381
	3	0.0009	0.0084	0.0176	0.9731	0.0004	0.0060	0.0301	0.9636

---

[7]If 1 age unit is a group, this hypothesis will not exist.

According to Table 3, we can conclude that compared to women, men still have a health advantage; regardless of the initial state, men's probability of transitioning to health ten years later is larger than women's probability; however, after the age 70, men's health advantage is gradually weakened, and there is no significant difference between men and women after age 80 (Fig.3). As it is ten-year forecasting, the people at the age of 70 now will be at the age of 80 ten years later, and the men at the age of 80 above have no significant health advantage compared to women, which conforms to the previous results. Next, we will focus on the probability of transition to disability state in a ten-year period. Fig.4 provides a comparison among the ten-year transition probabilities from different states to disability state. According to the figure, we can see that under three states, women's probability of transition to disability 3 (any age group) is larger than men's probability. In the age group of 70-74, the difference reaches the maximum value. In other words, the ratio of women aged from 75 to 79 needing long-term care is significantly larger than that of men. The inverted u-shaped difference is because that men blow the age of 70-74 has a health advantage (as shown in Fig.3), and the death rate of men above the age of 75 is higher than that of women, as shown in Fig. 4.
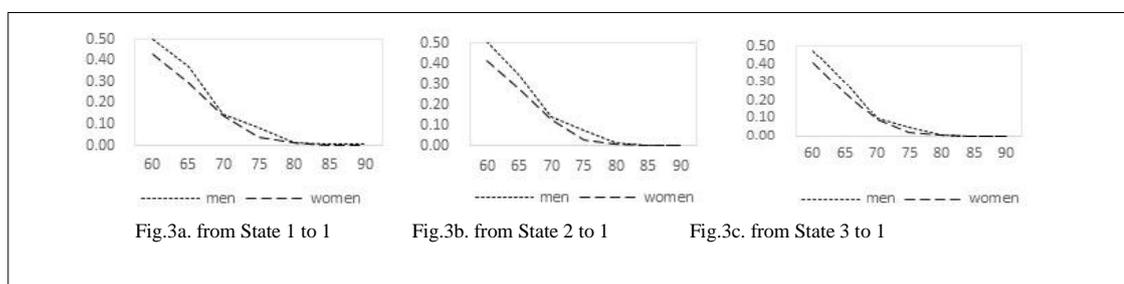

Fig.3a. from State 1 to 1   Fig.3b. from State 2 to 1   Fig.3c. from State 3 to 1

Figure 3. Ten-Year Transition Probability from Different States to Health State

Fig.3a. from State 1 to 3   Fig.3b. from State 2 to 3   Fig.3c. from State 3 to 3
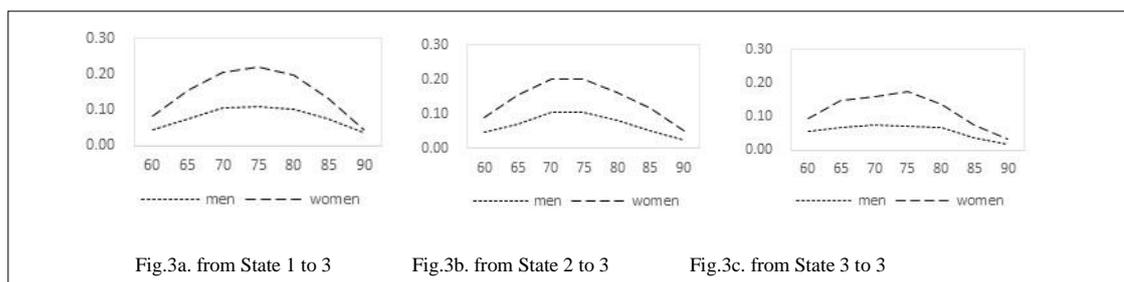
Figure 4. Ten-Year Transition Probability from Different States to Disability State

Forecasting of population size with a demand for long-term care

1.   Forecasting of population size in the same health state

Based on the demographic structure data of China's Demographic Yearbook of 2016, the ratio of people in each age group and each health state is calculated according to the CLHLS survey data from 2008 to 2011. The number of population in the base period multiplied by corresponding ratio vector is the heath state population vector of different age groups in the base period. The ratio of men at the age of 60-64 to healthy people in samples multiplied by the total number of men at the age of 60-64 is the number of men aged 60-64 in the health state (i.e., State 1) in the base period. The rest are similar. According to the definitions of health state and standard in this paper, the ratio (6.6%) of disabled population at and above the age of 60 is calculated[8]. Due to limited space, the number of population of each group in the base period is shown in Table 2 of the Appendix.

The health state population vector of each age group in the base period multiplied by the corresponding ten-year transition probability matrix is the number of population in each state ten years later, as shown in the table below. The datum 302672 in the table shows that among the men at the age of 60-64,

---

[8] The ratio is higher than the result 6.25% forecasted by the Research Team of the China Research Center on Aging (2011). It may be because of different definitions of the disability state. The definition of disability state in this paper includes the cognitive impairment with a score below 18.

the number of healthy population is 302,672 ten years later; the number of population with health impairment is 154,248; the number of population with a dysfunction (with a demand for long-term care) is 27,471; the number of deaths is 120,534. The rest are similar. The original age group of 95+ enters the age group of 105+ ten years later. It is presumed that they are all in the death state, so they are not included in the table.

Table 3. Number of Population in Different Health State after the Ten-Year Period (one thousand people)

Age	men				women			
	1	2	3	4	1	2	3	4
[60 64]	302672	154248	27471	120534	255950	238704	51749	60118
[65 69]	153640	103189	31241	136966	121572	141976	65526	96963
[70 74]	39185	68672	28616	139877	36383	77365	56975	116042
[75 79]	14666	33875	20616	126460	6628	39805	43519	125691
[80 84]	1357	11512	9377	89480	1186	12059	22368	105921
[85 89]	118	1309	2259	41160	107	2082	6431	58291
90+	14	128	266	11431	10	164	804	20959

The data in the table show that the number of healthy old men is more than that of women in the same age group. Moreover, the number of female deaths is more than that of male deaths before the age of 80-84, because the population size of old men in the later period is reduced, which indicates that men have a heath advantage and a survival disadvantage. Seen from the age distribution of disabled population size, the number of old men and women shows a phenomenon of increasing first and decreasing later, and the number of the disabled elderly in the age group of 65-69 is the largest.

2. Comparison of disabled population size before and after the ten-year period

Fig.5 and Fig.6 respectively provide the population size comparison diagram of disabled men before and after the ten-year period, as well as the population size comparison diagram of disabled women before and after the ten-year period. The two figures show the changes in the number of disabled population within ten years. A comparison of the same age before and after a ten-year period is made, e.g., the comparison between the number of disabled population at the age of 70-74 at present and that at the age of 70-74 ten years later. In the process of data processing, as the first age group of samples is 60-64, corresponding to 70-74 ten years later, the forecasted number of disabled population in the age group of 70-74 is compared with the current number of disabled population in the age group of 70-74. The rest can be compared in a similar fashion. The three age groups of 85-89, 90-94 and 95+ will enter the 95+ age group ten years later, so the sum of the three group data is compared with the data of the 95+ age group.

Seen from the figures, compared to the number of disabled population in the base period, the number of both men and women in all age groups is increased to different extent, wherein, the number of disabled population in the age group of 75-80 is increased significantly; the number of male disabled population is 1.5 times the number in the base period; the number of female disabled population is 1.57 times the number in the base period; the total number of disabled population is 1.53 times the number in the base period[9]. Seen from the age characteristics of the disabled population changes, the disabled population from the age group of 70-75 to the age group of 75-80 shows a progressive increase, reaches the maximum change in the age group of 80-85, and begins showing a downtrend at the age of 85+. Fig.7 depicts a gender contrast of the number of disabled population. Seen from the figure, the number of disabled old women in all age groups is more than that of disabled old men, and the total number of disabled women is 2.06 times that of disabled men.

---

[9] According to Huang Feng's research results, if it is assumed that the growth rate is unchanged, the number of disabled women is 2.0 times the number in the base period after the ten-year period (i.e., 2026 in this study); the number of disabled men is basically stable, and the total number of disabled population is 1.57 times the number in the base period. Lin Bao measured and calculated that the disabled population shows an annual growth rate of 3% before 2032, so the estimated number of disabled population in 2026 is 1.35 times the current number according to the ratio.
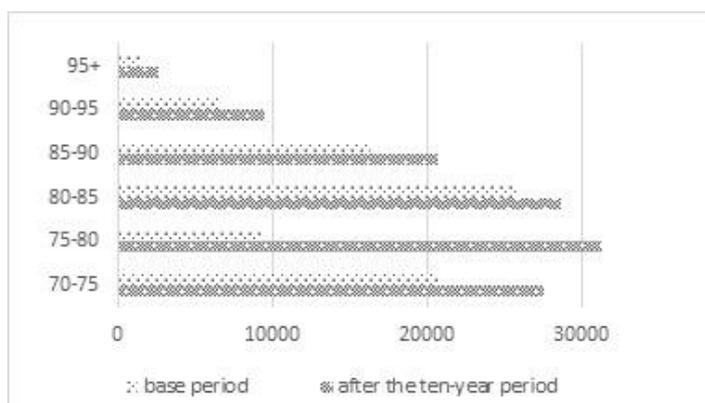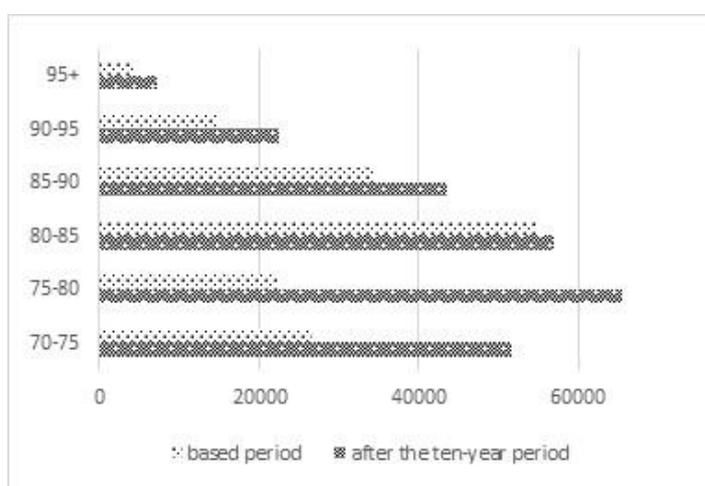
Figure 5. Comparison of the Number of Disabled Men

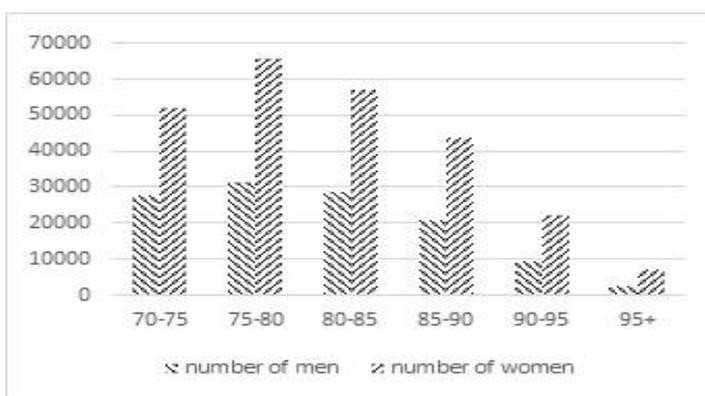**Figure 6.** Comparison of the Number of Disabled Women

Figure 7. Comparison of Number of Disabled Men and Women

Estimation of Expected Time Required for Long-Term Care

For the study of long-term care, in addition to estimating the probability and population size of disability, we also need to estimate the duration of disability (if any). Thus, the average time of each group transferred to and remaining in the target state in the future ten years is forecasted. Likewise, as the original 95+ age group enters the 105+ age group ten years later, it is presumed that they are all in the death state. The data of the group are not included in the table. Based on Equation 8, the duration of each state is obtained as shown in

Table 4. Seen from the data in the table, the duration of healthy (State 1) men in the age group of 60-65 who remain healthy (State 1) in the future ten years is 6.8257 years; the average duration of staying in State 2 is 1.8225 years; the average duration of staying in State 3 is 0.1960 year; the average duration of staying in State 4 is 1.1558 years. The rest are similar.

Table 4. Duration of Staying in Each State in Future Ten Years (Unit: Year)

Age	State	men				women			
		1	2	3	4	1	2	3	4
[60 64]	1	6.8257	1.8225	0.1960	1.1558	6.2592	2.8542	0.5066	0.3800
	2	5.7986	3.0656	0.3017	0.8341	4.5081	4.2656	0.8573	0.3691
	3	4.3174	3.0740	1.5048	1.1038	3.8873	3.6994	2.0363	0.3770
[65 69]	1	5.6410	2.4418	0.4785	1.4386	4.5385	3.5616	0.9217	0.9782
	2	3.6918	3.6316	0.8567	1.8199	3.3108	4.2868	1.2131	1.1893
	3	2.5648	3.0706	1.9361	2.4285	2.1845	3.6246	2.3854	1.8055
[70 74]	1	3.8653	2.7964	0.8400	2.4984	3.7555	2.9287	1.6737	1.6421
	2	2.7980	3.4355	1.0514	2.7153	1.7653	4.6372	1.5769	2.0205
	3	1.7430	2.2485	1.4236	4.5850	1.4081	1.8199	3.2668	3.5052
[75 79]	1	2.4220	2.9842	1.0441	3.5499	2.0782	3.3512	1.8755	2.6931
	2	1.4296	3.3319	1.5254	3.7143	1.0476	3.2266	2.4058	3.3200
	3	0.5530	2.0893	1.8167	5.5422	0.4658	2.3416	3.1019	4.0906
[80 84]	1	2.0208	2.4834	1.2837	4.2121	1.3483	2.9682	2.1660	3.5175
	2	0.6277	2.9304	1.4972	4.9448	0.3069	2.6244	2.7452	4.3235
	3	0.3161	1.9507	2.0890	5.6442	0.1347	1.7321	3.0144	5.1188
[85 89]	1	0.6765	2.9163	1.3918	5.0156	0.4396	2.6495	2.2327	4.6784
	2	0.2078	2.1454	1.4318	6.2150	0.1734	2.1109	2.6287	5.0870
	3	0.0991	0.9970	1.9195	6.9844	0.0652	0.7607	2.7597	6.4146
90+	1	0.4374	2.0371	1.2344	6.2911	0.5094	1.1069	1.8160	6.5678
	2	0.0891	1.4207	1.5058	6.9844	0.0259	1.4991	2.5009	5.9743
	3	0.0600	0.5104	1.6733	7.7563	0.0332	0.4249	2.2359	7.3070

Note: Due to rounding, the sum of duration is not fully equal to 10.

According to the data in the table, the distribution of old men remaining healthy in the same age group and the same state in the future is completely different from that of women. In each age group and each initial state, the duration of old men in health state 1 is longer than that of women, which shows that men have a health advantage; besides, the duration of men in the death state is longer than that of women (except the women aged 90 above whose initial health state is State 1), which shows that women have a survival advantage. For further observation, the duration of staying in the disability state of different ages, different genders and different states is shown in Fig.8.
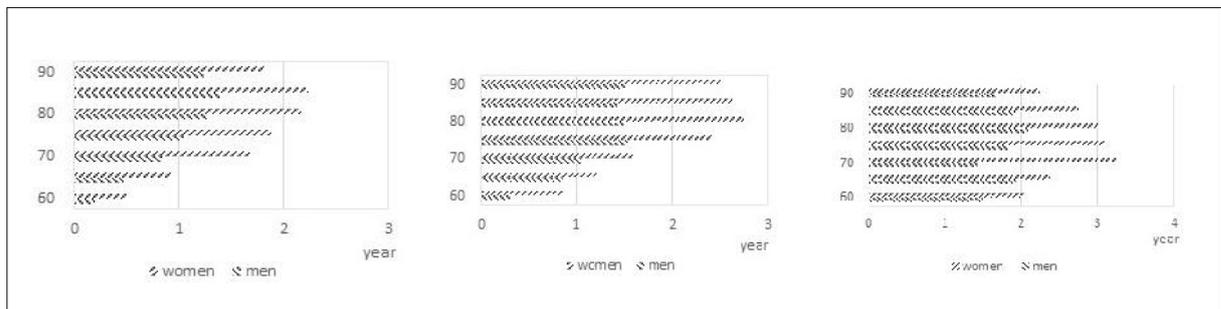

Figure 8. Duration of Disability from Different States in the Future

Fig.8 shows that regardless of the initial state and genders, the duration of disability has no linear correlation with age. On the whole, it shows an increase first and a slow decrease later. In the age group of 70-85, the average time required for long-term care is relatively longer. In a comparison among Fig. 8a, Fig. 8b and Fig. 8c, Fig. 8c shows the longest duration of disability at the same age, and Fig.8a shows a smaller duration, so the elderly who are in poor health initially have an urgent demand for long-term care; for the comparison of duration of disability between different genders, the average time of old women needing long-

term care is longer than that of old men, regardless of their states and age groups, and the difference is significant in the age groups of 70-75 above, until the age group of 90-95 as shown in Fig. 8b. Therefore, on the whole, the women around the ages of 70-85 who are in poor health initially constitute the major group with a demand for long-term care in the future.

Conclusion

In this paper, the survey data of CLHLS from 2008 to 2011 are utilized to investigate the dynamic evolution process of the health state of the elderly. According to the activities of daily living, instrumental activities of daily living and cognitive ability of the elderly, the health state of the elderly is divided into three types, namely, health, health impairment and dysfunction, and the people with a dysfunction are the major group needing long-term care as defined in this paper. Based on the Markov theory, tracking survey data are used to construct probability transition matrixes in this paper. Considering the age characteristics of changes in health states, the piecewise constant Markov process method is adopted to forecast the population size of different states in the future, and the time of transitioning to and remaining in each state is forecasted as well based on the actuarial method. According to the research results: (1) the changes in health have a significant difference in gender and age. Compared to men, old women have a significant survival advantage; compared to women, old men have a health advantage. The superposition of the two enables the number of disabled women of different ages to be around 2 times the number of disabled men in 2026, and the difference reaches the maximum value in the age group of 70-80. The mortality risk and the health risk are increased with the growth of age and the worsening of the initial health state. As a result, the people who are in poor health initially have a bigger possibility of needing long-term care. (2) Under different health states, the distribution of duration of staying in each state in different gender and age groups shows a characteristic similar to the distribution of population size. In the event that other conditions are the same, the duration of men staying healthy is longer than that of women, with a significant health advantage; the duration of men in the death state is longer than that women, which shows women have a survival advantage; the duration of men staying in disability is shorter than that of women; disability duration shows a nonlinear change with age. On the whole, it shows a trend of increasing first and progressively decreasing later. The duration of disability in the age groups of 70-85 is relatively long. As the disability duration is related to the initial health state, the people who are in poor health initially have a longer duration of disability.

   Population aging is a global yet irreversible phenomenon. As China has the largest number of aged population in the world, accompanied by miniaturized and hollow family size, the long-term care for the aged population not only adds a financial burden to families, but also forms a huge recessive aging debt of a country in the future. Boosting the long-term security system for the elderly is not only the huge progress in aging and social security undertakings made by a country, but also an inevitable strategic choice for a country to cope with population aging. This study contributes to the understanding and recognition of the natural law of body function changes of the elderly, helps policymakers to plan and prepare medical and nursing services required by the elderly, and is beneficial to the accumulation of basic information for the construction of China's long-term care security system for the elderly, thereby, providing a decision basis for the government to develop an aging strategy.

References

Bao, L. (2015). An analysis on the status and trend of old population without self-care ability in China. *Population & Economics*, *4*, 77–84.

Chahed, S., Demir, E., Chaussalet, T. J., Millard, P. H., & Toffa, S. (2011). Measuring and modelling occupancy time in NHS continuing healthcare. *BMC Health Services Research*, *11*, 163–178. https://doi.org/10.1186/1472-6963-11-155

Comas-Herrera, A., Northey, S., Wittenberg, R., Knapp, M., Bhattacharyya, S., & Burns, A. (2012). Future costs of dementia-related long-term care: Exploring future scenarios. *Int Psychogeriatric*, *23*(1), 20–30. https://doi.org/10.1017/S1041610210000025

Comas-Herrera, A., Wittenberg, R., Costa--Font, J., Gori, C., Di Maio, A., Patxo, C., & Rothgang, H. (2006). Future long-term care expenditure in Germany, Spain, Italy and the United Kingdom. *Ageing & Society*, *26*(2), 285–302. https://doi.org/10.1017/S0144686X05004289

Costa-Font, J., Wittenberg, Raphael Patxot, Concepció Comas-Herrera, A., Gori, C., Di Maio, A., Pickard, L., Pozzi, A., & Rothgang, H. (2008). Projecting long-term care expenditure in four European Union member states: The influence of demographic scenarios. *Social Indicators Research*, *86*(2), 303–321. https://doi.org/10.1007/s11205-007-9140-4

Feng, H., & Chunjie, W. (2012). A study of long-term-care demand of the elderly in China: Based on multistatus transition model. *Economic Research Journal*, *2*, 119–130.

Hare, W., Alimadad, A., Dodd, H., Ferguson, R., & Rutherford, A. (2009). A deterministic model of home and community care client counts in British Columbia. *Health Care Manag Sci*, *12*(1), 80–98.

Hongwei, H. (2015). Estimation and prediction of demand of Chinese elderly long-term care service. *Chinese Journal of Population Science*, *3*, 79–89.

Minglai, Z. (2009). The Analysis of Demand for Long-term Care and its Insurance System Constructing in China. *Chinese Journal of Health Policy*, *2*(7), 32–38.

Peng, R., Ling, L., & Qun, H. (2010). Self-rated health status transition and long-term care need of the oldest Chinese. *Health Policy*, *97*(2), 259–266.

Rickayzen, B., & Walsh, D. (2002). A multi-state model of disability for the UK: implications for need for long term care for the elderly. *Br Actuar J*, *8*(2), 341–393.

Rong, P. (2009). Analysis of senior population nursing needs based on Markov model. *Statistics & Information Forum*, *3*, 77–79.

Wei, Z., Shanjun, L., & Chen, K. (2014). Characteristics and trend of population aging in China and its potential impact on economic growth. *The Journal of Quantitative & Technical Economics*, (8), 3–38.

Yi, Z., Huashuai, C., & Zhenglian, W. (2012). Analysis on trends of future home-based care needs and costs for elderly in China. *Economic Research Journal*, *10*, 134–149.

Appendix

1. Transition Probability and Transition Intensity.

Transition probability: $_t p_x^{ij} = p\{S(x+t) = j \mid S(x) = i\}$; $j = 1,2,3,4; i = 1,2,3$, it represents that the state is i when at time x; it is transformed into j probability at time $x+t$; States 1, 2 and 3 are transferable states while 4 is the absorption state.

Transition intensity represents the instantaneous transition of state, which is denoted as:

$$\sim^{ij}(t) = \lim_{\Delta t \to 0} \frac{p^{ij}(t+\Delta t) - p^{ij}(t)}{\Delta t}$$

Under the time invariant hypothesis, transition probability and transition intensity meet Kolmogorov forward differential equations:

$$\frac{d}{dt}(_t p_x^{11}) = -_t p_x^{11}(\sim_{x+t}^{12} + \sim_{x+t}^{13} + \sim_{x+t}^{14}) + _t p_x^{12} \sim_{x+t}^{21} + _t p_x^{13} \sim_{x+t}^{31}$$

$$\frac{d}{dt}(_t p_x^{12}) = -_t p_x^{12}(\sim_{x+t}^{21} + \sim_{x+t}^{23} + \sim_{x+t}^{24}) + _t p_x^{11} \sim_{x+t}^{12} + _t p_x^{13} \sim_{x+t}^{32}$$

$$\frac{d}{dt}(_t p_x^{13}) = -_t p_x^{13}(\sim_{x+t}^{31} + \sim_{x+t}^{32} + \sim_{x+t}^{34}) + _t p_x^{11} \sim_{x+t}^{13} + _t p_x^{12} \sim_{x+t}^{23}$$

$$\frac{d}{dt}(_t p_x^{14}) = _t p_x^{11} \sim_{x+t}^{14} + _t p_x^{12} \sim_{x+t}^{24} + _t p_x^{13} \sim_{x+t}^{34}$$

$$\frac{d}{dt}(_t p_x^{21}) = -_t p_x^{21}(\sim_{x+t}^{12} + \sim_{x+t}^{13} + \sim_{x+t}^{14}) + _t p_x^{22} \sim_{x+t}^{21} + _t p_x^{23} \sim_{x+t}^{31}$$

$$\frac{d}{dt}(_t p_x^{22}) = -_t p_x^{22}(\sim_{x+t}^{21} + \sim_{x+t}^{23} + \sim_{x+t}^{24}) + _t p_x^{21} \sim_{x+t}^{12} + _t p_x^{23} \sim_{x+t}^{32}$$

$$\frac{d}{dt}(_t p_x^{23}) = -_t p_x^{23}(\sim_{x+t}^{31} + \sim_{x+t}^{32} + \sim_{x+t}^{34}) + _t p_x^{21} \sim_{x+t}^{13} + _t p_x^{22} \sim_{x+t}^{23}$$

$$\frac{d}{dt}(_t p_x^{24}) = _t p_x^{21} \sim_{x+t}^{14} + _t p_x^{22} \sim_{x+t}^{24} + _t p_x^{23} \sim_{x+t}^{34}$$

$$\frac{d}{dt}(_t p_x^{31}) = -_t p_x^{31}(\sim_{x+t}^{12} + \sim_{x+t}^{13} + \sim_{x+t}^{14}) + _t p_x^{32} \sim_{x+t}^{21} + _t p_x^{33} \sim_{x+t}^{31}$$

$$\frac{d}{dt}(_t p_x^{32}) = -_t p_x^{32}(\sim_{x+t}^{21} + \sim_{x+t}^{23} + \sim_{x+t}^{24}) + _t p_x^{31} \sim_{x+t}^{12} + _t p_x^{33} \sim_{x+t}^{32}$$

$$\frac{d}{dt}(_t p_x^{33}) = -_t p_x^{33}(\sim_{x+t}^{31} + \sim_{x+t}^{32} + \sim_{x+t}^{34}) + _t p_x^{31} \sim_{x+t}^{13} + _t p_x^{32} \sim_{x+t}^{23}$$

$$\frac{d}{dt}(_t p_x^{34}) = _t p_x^{31} \sim_{x+t}^{14} + _t p_x^{32} \sim_{x+t}^{24} + _t p_x^{33} \sim_{x+t}^{34}$$

If the transition intensity is a constant, the above equations are a set of correlative differential equations which are available to a simultaneous solution. If it is not a constant, nonlinear differential equations should be used for a solution.

2. Expected Time of State Continuance

(x) represents the people at the age of x. First, the remaining life of (x) is derived, i.e., the time of maintaining a survival state, which is denoted as $T(x)$. $T(x)$ represents a continuous random variable, and its distribution function $G(t)$ is: $G(t) = P(T(x) \le t), t \ge 0$,

In fact, $G(t)$ is a conditional probability, i.e., the probability of people at the age of x who will be dead in the next t years, which is expressed as symbol $_t q_x$. Besides, $s(x)$ represents a survival function, then:

$$_t q_x = \frac{s(x) - s(x+t)}{s(x)},$$

$g(t)$ represents the probability density of $T(x)$, then:

$$g(t) = G'(t) = \frac{d}{dt} {}_t q_x$$

$$= \frac{d}{dt}\left(\frac{s(x) - s(x+t)}{s(x)}\right)$$

$$= -\frac{1}{s(x)} \times s'(x+t)$$

$$= \frac{s(x+t)}{s(x)} \times \left(-\frac{s'(x+t)}{s(x+t)}\right)$$

$$= {}_t p_x \times \tilde{\ }_{x+t}$$

Where, $_t p_x$ represents the survival probability of people from the age x to the age $x+t$, $_t p_x = 1 - {}_t q_x$, and $\tilde{\ }_{x+t}$ represents the intensity at time $x+t$.

On this basis, the expected value of $T(x)$ can be derived:

$$E(T(x)) = \int_0^\infty t \times {}_t p_x \times \tilde{\ }_{x+t} dt = \int_0^\infty {}_t p_x dt$$

The above indicates the calculation of the expected time of staying in the survival state. In case of calculating the expected time of staying in any other state, we just need to change the probability to the corresponding transition probability, and use the same calculation form.

3. One-Year Health Transition Matrix.

Table 1. One-Year Health State Transition Matrix

Age	State	men				women			
		1	2	3	4	1	2	3	4
[60 64]	1	0.8949	0.0788	0.0022	0.0241	0.8799	0.1066	0.0097	0.0038
	2	0.3650	0.6109	0.0196	0.0045	0.2458	0.6816	0.0702	0.0024
	3	0.0828	0.3393	0.5516	0.0262	0.1347	0.2686	0.5947	0.0020
[65 69]	1	0.8473	0.1331	0.0008	0.0188	0.7434	0.2367	0.0084	0.0115
	2	0.2111	0.6766	0.0798	0.0325	0.2374	0.6638	0.0794	0.0194
	3	0.0388	0.3105	0.5874	0.0633	0.0279	0.3182	0.6039	0.0500
[70 74]	1	0.7465	0.1867	0.0301	0.0366	0.7792	0.1373	0.0726	0.0109
	2	0.2606	0.5988	0.1043	0.0363	0.1011	0.8127	0.0556	0.0306
	3	0.0234	0.2368	0.5435	0.1963	0.0265	0.0611	0.8093	0.1031
[75 79]	1	0.6086	0.3178	0.0008	0.0728	0.5877	0.3775	0.0054	0.0294
	2	0.1639	0.6175	0.1728	0.0458	0.1473	0.5739	0.2183	0.0605
	3	0.0041	0.2687	0.5044	0.2228	0.0014	0.2348	0.6500	0.1137
[80 84]	1	0.6558	0.2146	0.0662	0.0634	0.5384	0.4217	0.0033	0.0366
	2	0.0809	0.6482	0.1606	0.1103	0.0481	0.5870	0.2930	0.0720
	3	0.0148	0.2313	0.5786	0.1763	0.0012	0.2088	0.6383	0.1518
[85 89]	1	0.3321	0.5722	0.0154	0.0803	0.2321	0.6830	0.0115	0.0734
	2	0.0447	0.6097	0.1831	0.1625	0.0432	0.5992	0.2851	0.0725
	3	0.0117	0.1370	0.5953	0.2560	0.0081	0.0856	0.6913	0.2149
[90 94]	1	0.2660	0.6301	0.0005	0.1034	0.3295	0.2795	0.1867	0.2042
	2	0.0193	0.5228	0.2727	0.1852	0.0003	0.5518	0.3776	0.0702
	3	0.0132	0.0839	0.5816	0.3213	0.0074	0.0653	0.6549	0.2724
95+	1	0.4128	0.3466	0.1530	0.0875	0.2244	0.1611	0.5791	0.0354
	2	0.0264	0.5202	0.2713	0.1821	0.0235	0.5189	0.2957	0.1619
	3	0.0012	0.0607	0.5483	0.3898	0.0004	0.0277	0.6157	0.3563

Note: Due to rounding, the sum of duration is not fully equal to 10

For the advanced age groups of [90 94] and 95+, as there are only a few individual sample data, probability transition has certain instability. However, it has little influence on the ten-year forecasting, because ten years later, people in the two age groups are all above the age of 100. It is presumed that they are in the death state.

4.   Ratio and Number of Population in Different Health State in the Base Period.

Table 2. Ratio and Number of Population in Different Health State in the Base Period

Age	Item	men			women		
		1	2	3	1	2	3
[60 64]	Ratio (%)	0.7462	0.2386	0.0152	0.3460	0.0474	0.7462
	Number of population (one thousand people)	451402	144357	9166	367930	209835	28745
[65 69]	Ratio (%)	0.7114	0.2516	0.0370	0.3677	0.0448	0.7114
	Number of population (one thousand people)	302356	106938	15742	249903	156428	19077
[70 74]	Ratio (%)	0.5539	0.3709	0.0752	0.5535	0.0933	0.5539
	Number of population (one thousand people)	153056	102499	20777	101289	158729	26747
[75 79]	Ratio (%)	0.4218	0.5310	0.0471	0.6728	0.1035	0.4218
	Number of population (one thousand people)	82510	103865	9222	48249	145075	22319
[80 84]	Ratio (%)	0.2313	0.5384	0.2303	0.4604	0.3869	0.2313
	Number of population (one thousand people)	25846	60151	25729	21612	65161	54761
[85 89]	Ratio (%)	0.1434	0.4931	0.3635	0.4303	0.5141	0.1434
	Number of population (one thousand people)	6433	22113	16301	3723	28789	34398
[90 94]	Ratio (%)	0.0802	0.3831	0.5367	0.3024	0.6754	0.0802
	Number of population (one thousand people)	949	4535	6353	487	6634	14814
95+	Ratio (%)	0.0307	0.2467	0.7225	0.1346	0.8605	0.0307
	Number of population (one thousand people)	60	480	1407	24	666	4257