Quality & Quantity
https://doi.org/10.1007/s11135-023-01782-w

Analyzing the demographic coherence of selected US, Australian and Chinese biometric data sets used to price long-term care insurance and life care annuities

Carlos Vidal-Meliá[1,2] · **Manuel Ventura-Marco**[1] · **Anne M. Garvey**[3]

Accepted: 14 October 2023

Abstract
This paper examines the implicit healthy life expectancy used for actuarial calculations in some selected biometric data sets from the US, Australia and China. We are interested in checking the demographic/epidemiological coherence of these data sets because this health indicator is rarely presented when authors build their biometric data sets, nor when they are used to calculate long-term care insurance (LTCI) and life care annuity (LCAs) premiums, nor when they are employed in research articles to estimate the future demand for LTC services. We follow a methodology based on multistate life table methods that enables us to obtain a life expectancy matrix for individuals on the basis of their initial health state. We also present some additional indicators of longevity, mortality and morbidity, these being the median age at death, the interquartile range, the weighted modal age at death, the mortality ratio and the implicit LTC prevalence rates broken down by health state. We find several weaknesses that highlight the difficulty involved in building the biometric data sets needed to make an actuarially fair valuation of the premiums for LTCI and LCAs. We also verify the existence of the so-called "male–female health-survival paradox". From the perspective of a potential purchaser of this type of insurance products, disclosing and explaining the summary measures of health and longevity would make it easier for them to understand the need to protect themselves against the cost of possible LTC services and also make the computation of the premiums more transparent.

Keywords Activity limitations · Healthy life expectancy · Life care annuity · Long-term care insurance · States of dependence

JEL Classification G22 · G5 · I13 · J14 · J26

1 Introduction

The development of health measures that include detailed information on both the mortality and morbidity conditions of the population has been a focus of study since the 1960s (Molla et al. 2001), when the concept of health expectancy was first introduced. Given that life expectancy is not inextricably bound to health, researchers and policy makers were

Extended author information available on the last page of the article

searching for a population health indicator that would combine both mortality and morbidity (Saito et al. 2014) and could be used for assessing and monitoring population health.

Sullivan (1971) developed a method for combining mortality and morbidity rates in a single summary measure of population health known as disability-free life expectancy (DFLE), today also referred to as healthy life expectancy (HLE) or active life expectancy (ALE). In 1993 the DFLE measure became one of the health indicators used by the Organisation for Economic Co-operation and Development (OECD) (Di Lego 2021).

Healthy life expectancy (HLE) typically combines mortality and morbidity information to represent overall population health in a single indicator (Imai and Soneji 2007; Majer et al. 2013). It measures the number of remaining life years that a person of a certain age is likely to enjoy without activity limitations, assuming current disability and mortality conditions continue to apply, and is increasingly used to complement the conventional measure of life expectancy.

It is difficult to ignore how important and useful summary measures are for population health. Findings from HLE analyses are increasingly being used to shape policies and programmes, not only in the health sector but also in areas such as pension policy and sustainable development. Such measures are also important in the actuarial field.

Long-term care (LTC) involves a range of services including medical and nursing care, personal care services, assistance services and social services that help people live either independently or in residential settings when they can no longer carry out routine activities on their own. From the perspective of long-term care insurance (LTCI), the key indicators are HLE and life expectancy with activity limitation. LTCI is not life insurance, disability or health insurance and should not be confused with them, although they may sometimes be marketed together as hybrid products. Under an LTCI scheme, a healthy insured person (contributor) pays a premium (or contributes) up to a certain age (usually the statutory retirement age) or until disability occurs, at which point they may receive regular benefit payments to cover the costs associated with LTC. These regular payments can take the form of cash (cash-for-care benefits which might be indexed to inflation) or in-kind benefits involving the partial or total disbursement of actual recurring LTC costs.

A life care annuity (LCA) is a lifetime annuity in which the LTC benefit is defined in terms of an uplift with respect to the basic amount. In return for the payment of a premium (either in the form of a lump sum or an amount collected over time), the LCA provides a stream of fixed-income payments for the lifetime of the annuitant. It also provides an extra stream of payments if the annuitant requires LTC (Pla-Porcel et al. 2016; Vidal-Meliá et al. 2020).

In this paper we are very interested in investigating the biometric assumptions made by the insurers/sponsors of LTCI schemes. The inclusion of information showing the breakdown of life expectancy by health state is invaluable for illustrating the actuarial calculations used for pricing LTCI and LCAs.

We examine the implicit healthy life expectancy (HLE) used for actuarial calculations in some selected biometric data sets from the US (Robinson 1996), Australia (Hariyanto et al. 2014a, b) and China (Cui et al. 2022). It is very difficult to find this health indicator explicitly presented in the actuarial literature on LTC, and even more difficult to find it for LCAs. It is for this reason that we check the demographic/epidemiological coherence of the biometric data sets in two ways. First, we use each data set to compute the life expectancy values for both healthy and dependent individuals and then for dependent people in each of the different states of dependence. We then verify the existence of the so-called "male–female health-survival paradox", a phenomenon seen in developed countries where women have greater longevity but higher rates of disability and poorer health than men

Table 1 Main features of the biometric data sets. Source: Own

Data set	Country	Data years	Levels of dependence	Gender distinction	Recovery from illness state
Robinson (1996)	US	1982–89	6	Yes	Yes
Hariyanto et al. (2014a, b)	Australia	1998–2003	4	Yes	Yes
Cui et al. (2022)	China	2014–17	2	Yes	Yes

(Nielsen et al. 2021). And second, we present some additional longevity, mortality and morbidity indicators, namely the median age at death, the interquartile range, the weighted modal age at death, the mortality ratio and the implicit LTC prevalence rates curves broken down by health state.

It is worth noting that what is called the demographic coherence of the biometric data does not have a very precise definition. As data sets have more weaknesses (low number of dependence states, big differences in the amount of time likely to be spent in a dependent state when compared to official statistics and/or similar investigations, anomalies in life expectancy depending on the level of dependence and high recovery rates, to name just a few) they are arguably less coherent from a demographic point of view. It could be said that the concept of demographic coherence is a qualitative rather than a quantitative one, although quantitative elements are used to evaluate it.

The rest of the paper is structured as follows. Section 2 briefly describes the selected biometric data sets from the US, Australia and China, and the methodology used to compute the implicit healthy life expectancy (HLE) and check the demographic/epidemiological coherence of the data sets. Section 3 presents the main results. Section 4 discusses some issues arising from the results. The paper ends with our concluding comments, future research possibilities and a technical appendix, which explains the methodology used to compute the life expectancy matrix based on the individual's initial health state.

2 Data and methodology

2.1 Data

In this section we briefly describe the biometric data sets analysed. It is not our aim to include all the existing biometric data sets from all the countries we have selected. We could have included more, but for various reasons (difficulties in data collection, non-Markovian structure…) we decided not to. Table 1 summarizes the main features of the biometric data sets selected.

2.1.1 USA, data from Robinson's (1996) care transition model (CTM)

Based on data from the National Long-Term Care Survey (NLTCS) 1982–1989, the seven health states contained in Robinson's (1996) care transition model (CTM) differ according to three variables: (a) the number of instrumental activities of daily living

(IADLs) impaired, (b) the number of activities of daily living (ADLs) impaired, and (c) whether there is "cognitive impairment" (CI).

In state a, the individual has no impaired IADLs, no impaired ADLs and no cognitive impairment (i.e. they are able or healthy). In state d_1 the individual has one impaired IADL, no impaired ADLs and no CI. In state d_2, the individual has one impaired ADL and no CI. In state d_3 the individual has two impaired ADLs and no CI. In state d_4 the individual has three or more impaired ADLs and no CI. In state d_5 the individual has at most one impaired ADL but has CI. In state d_6 the individual has more than one impaired ADL and CI. In state 8 (f) the individual is dead.

2.1.2 Australia, data from Hariyanto et al. (2014a, b)

Hariyanto et al. (2014a, b) estimate transition probabilities between levels of disability as defined by the Australian Survey of Disability, Ageing and Carers, produced by the Australian Bureau of Statistics for the period 1998–2003. They assume that an individual in any state involving core activity limitation can only improve by one category over a one-year interval, if and only if they survive the year and do not deteriorate to a more severe state of dependence.

Their model provides five health states: one healthy and four dependent (with core activity limitation in their terminology). In state r, the individual is able or healthy with no activity limitation.

The four levels of core activity limitation are determined according to whether a person needs help with, has difficulty with or uses aids or equipment for any of the core activities (communication, mobility and self-care). A person's overall level of core activity limitation is determined by their highest level of limitation in these activities.

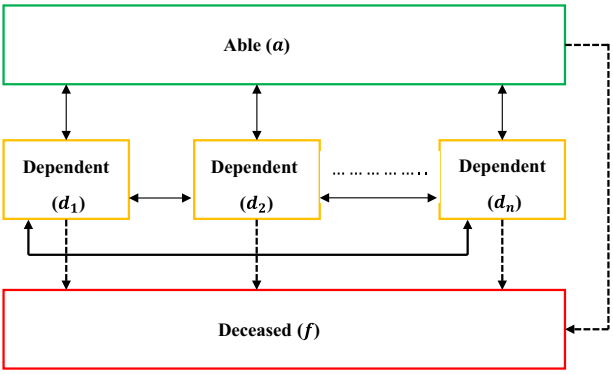
In state d_1, the individual has mild limitations. In state d_2, the individual has moderate limitations. They need no help but have difficulty with a core activity task.

In state d_3, the individual has severe limitations. In state d_4, the individual has profound limitations. They are unable to do, or always need help with, a core activity task. In state 5 (f) the individual is dead.

2.1.3 China, data from Cui et al. (2022)

Two types of data are used in this study. The first comes from the sixth and seventh waves of the Chinese Longitudinal Healthy Longevity Survey (CLHLS) in 2014 and 2017. The second comes from the China Life Insurance Mortality Table, which is used to calculate the survival rate for all age groups.

To classify dependent people the authors, use three indexes—ADLs, IADLs and CI—to define four health states corresponding to four levels: health (state 1), health impairment (state 2), disability (state 3) and death (state 4). State 1 means no health obstacle in any of the three indexes, state 3 is defined as having three or more obstacles in daily activities or scoring less than 16 in cognitive function, and the complementary state is defined as state 2.



Source: Own

Fig. 1 A multistate model for valuing LCAs/LTCI with n levels of disability. Source: Own

2.2 Methodology

2.2.1 The life expectancy matrix based on the individual's initial health state

Healthy life expectancy (HLE) can be estimated using a variety of health attributes, and we should bear in mind that its values vary by definition, by measures of health and by methods of calculation. Commonly used terms for healthy years are disability-free life expectancy, active life expectancy, healthy life years and healthy life expectancy (Saito et al. 2014).

The most widely used measure is disability, from which we get disability-free life expectancy (DFLE), the most classic of all healthy life expectancy indicators. The measurement of disability can vary, however, and indeed does vary considerably from one study to another. This means that comparability between some estimates and others, even when dealing with the same indicator, cannot be assured.

Three commonly used methods for estimating DFLE (or "healthy life-years") are the Sullivan method, the multistate life table method (MLTM) and the double decrement method (DDM) (Imai and Soneji 2007; Majer et al. 2013; Di Lego 2021). However, they require different kinds of data and can yield different results.

The actuarial approach used to compute the life expectancy of active people broken down into healthy and unhealthy life years is clearly linked to the multistate life table method (MLTM). Each state represents a particular state for the policyholder.

A multiple state transition model applied to compute annuity rates for LTCI contracts describes a subject's movements within a set of various states: active (healthy) (a), disabled ($j \in \{1, 2, \ldots, n\}$-level dependent) ($d_j, j \in \{1, 2, \ldots, n\}$) and deceased ($f$).

This discrete-time Markov multistate model takes into account n levels of dependence where no more than one transition in a year is assumed, with level 1 being the least severe and level n the most. Although the Markov model (like any other) might be accused of oversimplification, for models that use variations in state space and intensities, the Markov set-up can deal with extremely complex phenomena (Norberg 2002).

Figure 1 shows the transitions into and out of the various states that will be considered in the model for computing life expectancies based on a multistate model for valuing

Table 2 Life expectancies at age x broken down into health states. Source: Own

Initial health state	States					Total
	a	d_1	d_2	…	d_n	
a	$\boldsymbol{e_x^{aa}}$	$e_x^{ad_1}$	$e_x^{ad_2}$	…	$e_x^{ad_n}$	e_x^a
d_1	$e_x^{d_1a}$	$\boldsymbol{e_x^{d_1d_1}}$	$e_x^{d_1d_2}$	…	$e_x^{d_1d_n}$	$e_x^{d_1}$
d_2	$e_x^{d_2a}$	$e_x^{d_2d_1}$	$\boldsymbol{e_x^{d_2d_2}}$	…	$e_x^{d_2d_n}$	$e_x^{d_2}$
…	…	…	…	…	…	…
d_n	$e_x^{d_na}$	$e_x^{d_nd_1}$	$e_x^{d_nd_2}$	…	$\boldsymbol{e_x^{d_nd_n}}$	$e_x^{d_n}$

LCAs/LTCI with n levels of disability (Pitacco 2014; Pla-Porcel et al. 2017; Baione and Levantesi 2018). The relationships denoted by the dotted lines in the diagram represent the different paths to the absorbent state.

In this type of Markov model the transition probabilities depend only on the current state of the process. This methodology (see the Technical Appendix) enables us to obtain the following life expectancy matrix based on the individual's initial health state (Table 2).

The diagonal line in Table 2 shows the life expectancy likely to be lived in that same (initial) health state. The last column shows the individual's total life expectancy linked to the initial health state, i.e. each cell in this column is the sum of the cells in its corresponding row, where e_x^a is the life expectancy for active people aged (x). This can be broken down into the health states (active or dependent) they can expect to experience. e_x^{aa} is the dependence-free life expectancy (or "healthy life years"). This indicates how many years of their total remaining life an active person aged (x) can expect to live free of activity limitation. $e_x^{ad_j}$ is the life expectancy for a $j \in \{1, 2, ..n\}$-level dependent person. This can be defined as the number of years an active person aged (x) can expect to spend with $(j) \in \{1, 2, ..n\}$ -level activity limitation. $e_x^{d_j}$ is the life expectancy for a $j \in \{1, 2, ..n\}$-level dependent person aged (x). This can be broken down into the health states (active or dependent) that they can expect to experience. $e_x^{d_id_i}$ indicates how many years of their total remaining life expectancy a person aged (x) in state (i) of dependence can expect to live in the same state of dependence. $e_x^{d_ia}$ is the life expectancy expected to be spent free of activity limitation for a person aged (x) in state (i) of dependence.

2.2.2 Other lifespan variation indexes: median age at death, the interquartile range (IQR) and the (weighted) adult modal age at death

There are several indexes for calculating lifespan variation and each has different underlying properties (Van Raalte and Caswell 2013). Median and modal ages at death are seldom proposed as measures for studying longevity. They complement one another with information on the "centre" of the distribution of deaths (Canudas-Romo 2010).

The median age at death, Md, is the age when half of the hypothetical cohort members have died, i.e. when the number of people surviving to the exact age of x (l_x) in any health state is equal to half the initial cohort aged x_e, $l_{Md} = \frac{l_{x_e}}{2}$. In our case the cohort is aged 65 or older and the health state is able (or with j-level activity limitation).

The interquartile range (IQR), also known as the middle 50% (Wilmoth and Horiuchi 1999), is a measure that equals the distance between the lower and upper quartiles of the age distribution of death in a life table. The measure decreases as age at death becomes

less variable. The IQR has a twofold appeal as a single measure of variability in the life table. Firstly, it is very simple to calculate because it equals the difference between the ages where the survivorship curve, S(x), crosses 0.25 and 0.75, and secondly, it is easy to interpret as it is the length of the span of ages containing the middle 50% of deaths.

The adult modal age at death, M, i.e. the age (beyond infancy) at which the largest single number of deaths occur, is used as an indicator to analyse mortality disparities at older ages. The mode is both a natural measure of the length of life and a good basis for measuring its dispersion (Kannisto 2001).

When the individuals that form part of a group can exit for reasons other than death (i.e. move from one group to another) as in our case, an improvement or worsening of their state of health is considered, and therefore the previous indicator needs to be modified. Given a starting age and an initial health state (able or disabled), there is a mode for each state ($M^{ij}_{x_e}$). By taking into account the percentages of deaths in the various states, we get the weighted adult modal age at death ($\overline{M}^{i}_{x_e}$). In our setting this new indicator is more consistent and informative than the unweighted adult modal age at death, especially in relation to life expectancy and the median.

2.2.3 Further indicators of mortality and morbidity

There are several other mortality and morbidity indicators that are very interesting from an LTC perspective when analysing biometric data sets. They can give us additional information about how realistic the data set is.

We show the implicit LTC prevalence rate, λ^{j}_{x+k}—which is the ratio between the number of dependent people in dependence level j and the number of individuals aged $x + k$—and the average LTC prevalence rate, $\overline{\lambda}^{j}_{x}$. The total prevalence rate is a key element in the public field, where there is a need to assess how the demand and cost burden of LTC will evolve over the coming years under a given set of biometric assumptions (Vanella et al. 2020).

We also explore the mortality ratio, δ^{j}_{x+k}, for dependent people aged $x + k$ in dependence level j. This is the ratio between the mortality rates for dependent people and active people and the average mortality ratio ($\overline{\delta}^{j}_{x}$).

3 Results

The procedure followed in this section is:

1. The results are presented by country/authors and summarized in tables or graphs.
2. The consistency of the results is studied via a series of questions: Is life expectancy in line with what would be expected according to official data and/or other similar studies? Is the percentage of time expected to be spent in a dependent health state reasonable? Is the mortality of dependents realistic? Is the life expectancy according to health state consistent? Is the male–female health-survival paradox fulfilled? Are the resulting age-specific prevalence rates consistent with the empirical evidence?
3. The main weaknesses are highlighted.

Table 3 Life expectancy matrix in years at age 65 and percentage of life expectancy likely to be spent in each state (males). Source: Own based on data from Robinson's (1996) care transition model (CTM)

States	a	d_1	d_2	d_3	d_4	d_5	d_6	Total
a	**12.15**	1.85	0.30	0.12	0.33	0.36	0.24	15.35
	79.14%	12.06%	1.98%	0.75%	2.12%	2.37%	1.57%	100%
d_1	4.72	**5.87**	0.72	0.30	0.93	0.64	0.61	9.88
	37.35%	**45.02%**	4.62%	1.79%	4.40%	4.16%	2.66%	100%
d_2	3.13	3.50	**1.12**	0.21	0.51	0.39	0.26	9.12
	34.26%	38.42%	**12.28%**	2.26%	5.64%	4.32%	2.82%	100%
d_3	2.11	2.62	0.64	**1.12**	0.83	0.33	0.32	7.95
	26.48%	32.89%	7.99%	**14.05%**	10.38%	4.17%	4.03%	100%
d_4	0.81	1.12	0.30	0.36	**2.34**	0.17	0.33	5.44
	14.94%	20.64%	5.60%	6.61%	**42.95%**	3.14%	6.12%	100%
d_5	2.15	2.63	0.55	0.28	0.80	**1.53**	0.44	7.95
	25.65%	31.45%	6.54%	3.36%	9.52%	**18.26%**	5.25%	100%
d_6	0.68	0.94	0.25	0.26	1.46	0.24	**1.30**	5.13
	13.29%	18.33%	4.93%	5.02%	28.43%	4.69%	**25.30%**	100%

Table 4 Life expectancy matrix in years at age 65 and percentage of life expectancy likely to be spent in each state (females). Source: Own based on data from Robinson's (1996) care transition model (CTM)

States	a	d_1	d_2	d_3	d_4	d_5	d_6	Total
a	**13.76**	2.69	0.51	0.21	0.72	0.58	0.59	19.02
	72.22%	14.14%	2.67%	1.08%	3.79%	3.02%	3.09%	100%
d_1	4.72	**5.87**	0.72	0.30	0.93	0.64	0.61	13.79
	34.22%	**42.57%**	5.19%	2.18%	6.77%	4.67%	4.40%	100%
d_2	4.12	4.90	**1.41**	0.34	1.05	0.63	0.60	13.04
	31.55%	37.58%	**10.84%**	2.59%	8.05%	4.79%	4.59%	100%
d_3	2.91	3.84	0.93	**1.33**	1.50	0.55	0.68	11.72
	24.80%	32.79%	7.90%	**11.30%**	12.78%	4.66%	5.77%	100%
d_4	1.30	1.96	0.53	0.56	**3.44**	0.33	0.69	8.83
	14.73%	22.25%	6.03%	6.39%	**38.97%**	3.78%	7.86%	100%
d_5	2.90	3.80	0.81	0.43	1.48	**1.82**	0.82	12.06
	24.01%	31.49%	6.76%	3.60%	12.23%	**15.08%**	6.83%	100%
d_6	1.12	1.71	0.47	0.44	2.50	0.41	**1.79**	8.44
	13.28%	20.29%	5.51%	5.25%	29.61%	4.90%	**21.16%**	100%

3.1 USA

Tables 3 and 4 give detailed information about life expectancy broken down into the various health states, while Tables 5 and 6 add further mortality and morbidity indicators. The expected number of years allocated to the possible range of health states depends on the starting state by age. The data shown for the life expectancy of active individuals (15.35 years and 19.02 years respectively for men and women) almost coincide with those

Table 5 Further indicators of mortality and morbidity (females). Source: Own based on data from Robinson's (1996) care transition model (CTM)

Indicators/states	a	d_1	d_2	d_3	d_4	d_5	d_6
$A^i\left(=65+e^i_{65}\right)$ (years)	84.05	78.79	78.04	76.72	73.83	77.06	73.44
Md^i (years)	84.67	77.76	76.75	74.77	70.99	74.97	70.83
Q^i_3 (years)	90.80	85.93	85.54	83.56	78.62	83.65	77.76
Q^i_1 (years)	77.78	70.81	69.90	68.94	67.68	69.74	67.61
IQR^i_{65} (years)	13.02	15.12	15.64	14.62	10.94	13.91	10.15
$\overline{M}^i_{65}$ (years)	84.73	72.76	72.05	68.48	67.70	68.78	67.91
$\overline{\lambda}^i_{65}\%$	71.55	14.51	2.73	1.11	3.87	3.09	3.13
$\overline{\delta}^d_{65}\%$	–	3.32	4.53	4.39	7.22	2.77	7.12

Table 6 Further indicators of mortality and morbidity (males). Source: Own based on data from Robinson's (1996) care transition model (CTM)

Indicators/states	a	d_1	d_2	d_3	d_4	d_5	d_6
A^i (years)	80.35	74.88	74.12	72.95	70.44	73.38	70.13
Md^i (years)	80.60	72.90	71.92	70.77	68.67	71.56	68.57
Q^i_3 (years)	86.62	80.53	78.97	76.87	72.67	77.60	71.91
Q^i_1 (years)	74.55	68.65	67.83	67.64	66.69	67.88	66.66
IQR^i_{65} (years)	12.07	11.88	11.14	9.23	5.98	9.72	5.25
$\overline{M}^i_{65}$ (years)	78.71	68.61	68.19	67.95	67.05	68.25	67.19
$\overline{\lambda}^i_{65}\%$	78.46	12.47	2.05	0.77	2.19	2.45	1.62
$\overline{\delta}^d_{65}\%$	–	3.38	4.56	4.48	7.32	2.87	7.20

calculated by Crimmins et al. (2016). However, there is a sizeable divergence in the time that active individuals are expected to spend in a state of dependence. In Robinson's model (1996), active life expectancy is 79.14 and 72.22% of total life expectancy for males and females respectively, whereas Crimmins et al. (2016) estimate 49.06 and 52.10% for males and females respectively.

The discrepancies are even greater but in the opposite direction when we look at the values provided by Manton and Land (2000), who use the same database and a similar period of time (the National Long-Term Care Surveys of older Americans for the years 1982 to 1996). These authors estimate 94.49 and 94.87% active life expectancy as a proportion of total life expectancy for males and females respectively.

In both tables our attention is drawn to state d_5 (those unable to perform one ADL and with cognitive impairment) because its associated total life expectancy is higher than that calculated for states d_2, d_3, d_4 and, logically, for state d_6 too. At first sight and from an epidemiological point of view, it might be said that this observation shows some kind of data anomaly because it is known that life expectancy decreases (or should decrease) when an individual's disability state worsens. However, there may not be an anomaly (Biessy 2017) if we consider the French experience, where two types of mortality were found in disabled people: the first due to severe illness causing entry into dependence

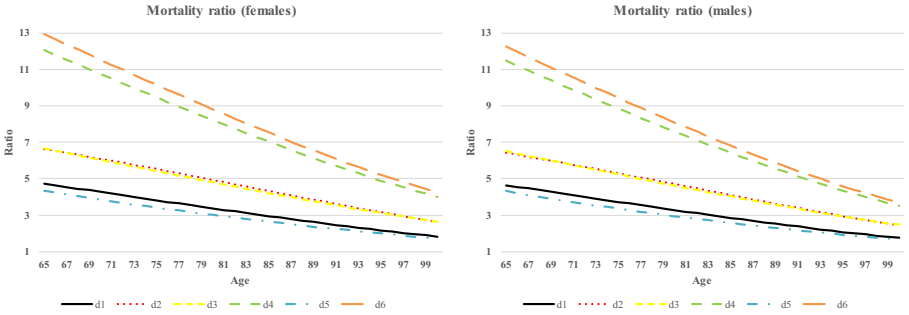


Fig. 2 Comparison of mortality ratios: dependent/active people. Source: Own based on data from Robinson (1996)

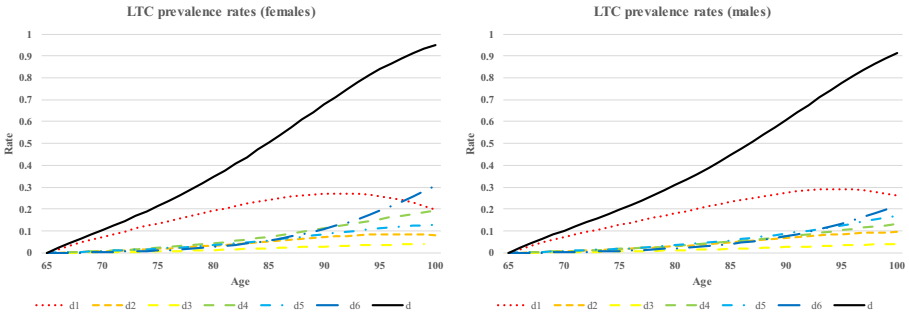


Fig. 3 Comparison of LTC prevalence rates. Source: Own based on data from Robinson (1996)

with a very short life expectancy, and the second caused by capacity erosion due to age and conditions such as dementia, which has a much longer mean duration.

The mortality ratio (dependent/healthy population) (Fig. 2) shows that dependent people have a higher (much higher) yearly probability of dying than healthy people, but the difference tends to decrease notably as they get older. The extra mortality for dependent individuals is very noticeable and the "anomaly" involving the group of people labelled d_5 is clear to see.

The synthetic LTC prevalence rates by age and gender that would result from combining mortality and incidence rates are shown in Fig. 3. The rates for the most severe states of dependence—d_5 and d_6—would be particularly high for the very elderly (especially females). The highest average prevalence rate ($\overline{\lambda}_{65}^{d_1}$) (see Tables 7 and 8) would be for the least severe state of dependence, 12.47% for males and 14.51% for females, whereas the lowest would be for dependent people in state d_3 ($\overline{\lambda}_{65}^{d_3}$), barely reaching 0.77% for males and 1.11% for females. The total average LTC prevalence rate ($\overline{\lambda}_{65}^{d_j}$) would be 21.54 and 28.45% for males and females respectively.

According to the above prevalence rates by age, the total number of dependent people would exceed that of the active population from age 87 in males and age 85 in females. For males, the least severe level of dependence d_1 would predominate over the entire range of ages considered, while for females it would be the most important up to age 97, when level d_6 would become most predominant.

Table 7 Life expectancy matrix in years at age 65 and percentage of life expectancy likely to be spent in each state (males). Source: Own based on data from Hariyanto et al. (2014a, b)

States	a	d_1	d_2	d_3	d_4	Total
a	**10.32**	2.99	1.38	0.98	1.95	17.62
	58.55%	16.96%	7.84%	5.56%	11.09%	100%
d_1	6.28	**6.53**	1.52	1.06	2.04	17.43
	36%	**37.47%**	8.71%	6.09%	11.73%	100%
d_2	3.26	4.54	**5.66**	1.25	2.26	16.97
	19.19%	26.76%	**33.34%**	7.38%	13.33%	100%
d_3	1.15	1.94	2.94	**5.62**	2.13	13.78
	8.35%	14.07%	21.3%	**40.82%**	15.46%	100%
d_4	0.35	0.69	1.13	2.7	**5.17**	10.04
	3.53%%	6.84%	11.26%	26.86%	**51.5%**	100%

Table 8 Life expectancy matrix in years at age 65 and percentage of life expectancy likely to be spent in each state (females). Source: Own based on data from Hariyanto et al. (2014a, b)

States	a	d_1	d_2	d_3	d_4	Total
a	**11.39**	2.32	1.71	1.52	4.07	21.01
	54.25%	11.03%	8.13%	7.22%	19.37%	100%
d_1	6.8	**6.13**	1.93	1.66	4.24	20.76
	32.76%	**29.54%**	9.28%	7.99%	20.43%	100%
d_2	3.44	4.27	**5.98**	1.96	4.58	20.23
	16.99%	21.12%	**29.56%**	9.69%	22.64%	100%
d_3	1.24	1.9	3.18	**6.32**	4.55	17.19
	7.21%	11.05%	18.49%	**36.76%**	26.47%	100%
d_4	0.39	0.71	1.29	3.4	**8.22**	14.01
	2.77%	5.04%	9.23%	24.27%	**58.69%**	100%

Another aspect that stands out in these data is the different range of values observed for $IQR_{65}^{a/d}$ for males and females. In the case of males, the increase in the level of activity limitation—apart from that mentioned in the case of d_5—implies a decrease in the average age of death for people still alive at 65. The same applies to the median, mode and IQR_{65}^{i} values. For females, however, this is not the case, despite the fact that the increase in the level of activity limitation also implies a decrease in the average age of death for people still alive at 65, the median, the mode, and the interquartile range value increases, which would seem to indicate greater heterogeneity within each group. Finally, it can be seen from Table 6 that for both groups in the able state, the average age of death for people still alive at 65 and the median and the mode have very close values. However, when states of dependence are taken into account, the weighted average mode is very close to Q_1 (and well below the mean and median values) and the mean begins to differ from the median, which is higher in all cases.

To end this subsection, it is worth mentioning that Friedberg et al. (2014) revised and updated Robinson's CTM by adding a linear time trend. They used data from the 1982–2004 National Long-Term Care Survey (NLTCS) and the 1998–2010 Health and Retirement Study (HRS). On the basis of their LTC transition rate parameter estimate (Table 5 in their paper), a new set of health transition matrices were built. This biometric data set presents greater problems of coherence than Robinson's. The analysis of this biometric data set is available upon request from the authors.

Table 9 Further indicators of mortality and morbidity. Source: Own based on data from Hariyanto et al. (2014a, b)

Indicators/states	Males					Females				
	a	d_1	d_2	d_3	d_4	a	d_1	d_2	d_3	d_4
A^i (years)	82.62	82.44	81.97	78.78	75.04	86	85.74	85.22	82.19	79
Md^i (years)	82.77	82.66	81.87	78.53	72.9	86.77	86.65	85.87	82.88	77.98
Q_3^i (years)	88.73	88.65	87.9	85.57	80.77	91.87	91.8	91.61	89.74	86.87
Q_1^i (years)	76.78	76.65	75.88	71.69	67.97	80.88	80.67	79.78	74.74	70.63
IQR_{65}^i (years)	11.95	12	12.02	13.88	12.8	10.99	11.13	11.83	15	16.24
$\overline{M}_{65}^i$ (years)	82.10	79.54	79.91	75.04	68.61	87.56	85.70	85.65	80.90	68.31
$\overline{\lambda}_{65}^i$ %	57.34	17.45	8.07	5.73	11.42	53.14	11.29	8.33	7.40	19.84
$\overline{\delta}_{65}^d$ %	–	1.00	1.12	2.45	3.10	–	1.00	1.05	2.29	2.97

3.2 Australia

The structure of the information for Australia is identical to that of the US, but the biometric data include only four states of dependence, as can be seen in Tables 7, 8 and 9.

These data show that active people in Australia have a higher life expectancy than those in the US and are likely to spend a smaller percentage of their life expectancy in the same health state. To some extent it is understandable that the life expectancies deriving from the data for Australia are appreciably higher than those for the US because they are based on much more recent observations.[1]

The data for the life expectancy of active people (17.62 and 21.01 years respectively for males and females) are surprisingly similar to those provided by the Australian Institute of Health and Welfare (2014)—to be found in Appendix A, Table A2 for 2003 (the reference year for the data in Hariyanto et al. 2014a, b)—which are 17.6 years for men and 21.00 for women. However, there is some divergence regarding the time that active people are likely to spend in a state of dependence. In the Hariyanto et al. (2014a, b) model, active life expectancy is 54.55 and 54.25% of total life expectancy for males and females respectively, whereas AIHW (2014) report values of 43.18% and 41.90% for males and females respectively.

Tables 7 and 8 clearly show that the so-called "male–female health-survival paradox" is verified for this biometric data set, i.e. women have greater longevity than men but are likely to spend more time in any state of dependence, irrespective of the initial health state.

From a demographic/epidemiological point of view, the data from Australia are coherent because life expectancy decreases when an individual's disability state worsens. However, the mortality ratios ($\overline{\delta}_{65}^d$) are lower than those shown for the case of the US and their time structure seems well graduated and decreases with age.

The synthetic LTC prevalence rates by age and gender are shown in Fig. 4.

[1] The cross-country comparisons are made for the sake of illustration. A more accurate comparison would require all data sets to be based on the same period.

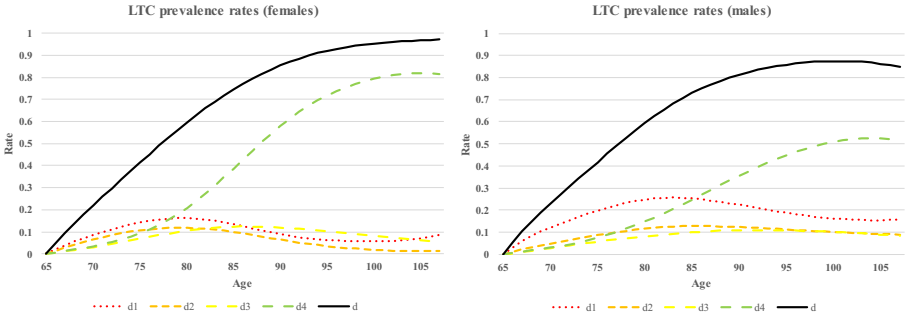


Fig. 4 Comparison of LTC prevalence rates. Source: Own based on data from Hariyanto et al. (2014a, b)

As in the case of the US, the rate for the most severe dependence state d_4 would be highest for the very elderly (especially females). However, unlike for the US, the highest average prevalence rate for males ($\overline{\lambda}^{d_4}_{65}$) would be in the most severe state of dependence at 19.87%, whereas for females it would be in the least severe state of dependence ($\overline{\lambda}^{d_1}_{65}$) at 17.45%. For both males and females, the lowest rate would be for dependent people in state d_3 ($\overline{\lambda}^{d_3}_{65}$). The total average LTC prevalence rate ($\overline{\lambda}^{d_j}_{65}$) would be 42.66% and 46.86% for males and females respectively, which would be far higher than those calculated for the US.

In the case of men, the least severe level of dependence d_1 would be predominant up to age 85, but from then on the highest prevalence rates would be for people at a more severe level of dependence (d_4). In the case of females, d_1 would be the predominant level of dependence up to age 78, while from then on the most severe level d_4 would predominate.

In the case of women, the highest dependence rate (when all levels are included) would be reached at age 107, around 97.18%, while for men it would be at age 100, with a total rate of dependence of 87.47%.

With the above rates of prevalence by age, the total dependent population would exceed the active population from age 78 in both men and women.

In contrast to the case of the US, the range of values for IQR^i_{65} is very similar for both men and women. In the case of males, the increase in the level of activity limitation implies a decrease in the average age of death for those still alive at 65, while a decrease is also observed in the median and the adult weighted average mode of death. However, an increase in the value of IQR^i_{65} is also observed, which means a significant increase in heterogeneity as far as the death of individuals is concerned. This also occurs in the case of females but is even more pronounced than for males, showing a change in value from 10.99 (a) to 16.24 d_4.

Finally, it can be seen in Table 9 that for women in the first three states considered—a, d_1, d_2—the values are very similar for the average age of death for people still alive at 65, and for the median and the adult weighted average mode age of death. However, when the most severe states of dependence are considered, the mode moves away from these values and, even in the most severe state of dependence, is below Q_1 (and well below the mean and median values). For men this behaviour is different, with the mode being close to the other two indicators only in the able state. From then on the divergence is more accentuated, although it does not fall below Q_1 even when the mean and median values separate from each other in this particular state.

Table 10 Life expectancy matrix in years at age 65 and percentage of life expectancy likely to be spent in each state. Source: Own based on data from Cui et al. (2022)

States	Males				Females			
	a	d_1	d_2	Total	a	d_1	d_2	Total
a	**7.44**	4.99	1.81	14.24	**5.97**	6.44	3.79	16.2
	52.26%	35.03%	12.71%	100%	**36.86%**	39.76%	23.38%	100%
d_1	4.88	**6.51**	2.13	13.52	4.18	**7.55**	4.01	15.74
	36.13%	**48.18%**	15.69%	100%	26.58%	**48%**	25.42%	100%
d_2	3.6	5.2	**3.59**	12.39	2.93	6.12	**5.46**	14.51
	29.07%	41.99%	**28.94%**	100%	20.18%	42.18%	**37.64%**	100%

Table 11 Further indicators of mortality and morbidity. Source: Own based on data from Cui et al. (2022)

Indicators/states	Males			Females		
	a	d_1	d_2	a	d_1	d_2
A^i (years)	79.24	78.52	77.39	81.19	80.74	79.5
Md^i (years)	78.82	77.91	76.82	80.98	80.8	79.71
Q_3^i (years)	84.8	83.98	83.59	86.83	86.74	85.88
Q_1^i (years)	73.73	72.66	70.86	75.76	75.46	72.93
IQR_{65}^i (years)	11.07	11.32	12.73	11.07	11.28	12.95
$\overline{M}_{65}^i$ (years)	79.00	73.55	72.50	80.79	81.72	69.10
$\overline{\lambda}_{65}^i$%	50.53	36.30	13.17	34.85	41.02	24.13
$\overline{\delta}_{65}^d$%	–	1.26	3.50	–	2.07	4.69

3.3 China

The structure of this information is identical to the cases of the US and Australia, although in the biometric data there are only two states of dependence, as can be seen in Tables 10 and 11.

The data shown for active life expectancy (14.24 years and 16.2 years respectively for men and women) are slightly lower than those presented for the 1940 cohort by Jiao (2019), according to whom life expectancy at age 65 would fluctuate between 16.45 and 16.65 years for women and 14.74 and 15.06 years for men. For 2015 Zhen et al. (2022) estimate a life expectancy of 14.36 and 17.81 years for men and women respectively.

There is more divergence in the time that active people are likely to spend in a state of dependence. In the Cui et al. (2022) model, active life expectancy is 52.26 and 36.86% of total life expectancy for males and females respectively, whereas Jiao (2019) and Zhen et al. (2022) report a range of values between 86.60 and 94.60% for males and 89.44 and 96.53% for females as a proportion of total life expectancy.

As we saw with the biometric data sets for other countries, Table 12 reflects again the so-called "male–female health-survival paradox". The paradox is verified for this data set, although in this case the large proportion of time that both men and women in dependence level d_2 would likely spend in better health states is striking. In the case of men, 71.06% of

Table 12 Main weaknesses of the biometric data sets selected. Source: Own elaboration

Data set	Weaknesses
Robinson (1996)	Different levels of heterogeneity at the same level of dependence for males and females
	Big differences in the amount of time likely to be spent in a dependent state when compared to the official statistics and/or similar investigations
Hariyanto et al. (2014a, b)	People in any state of core activity limitation can only improve by one category over a one-year interval
	Very high prevalence rates
Cui et al. (2022)	Only two states of dependence
	Unhealthy life expectancy very high in comparison with other recent studies
	High recovery rates
	The mortality ratio presents sizeable irregularities for sections that are very difficult to understand and justify. We see inexplicable jumps for both males and females
	Lack of graduation in incidence and mortality rates
	Very high prevalence rates

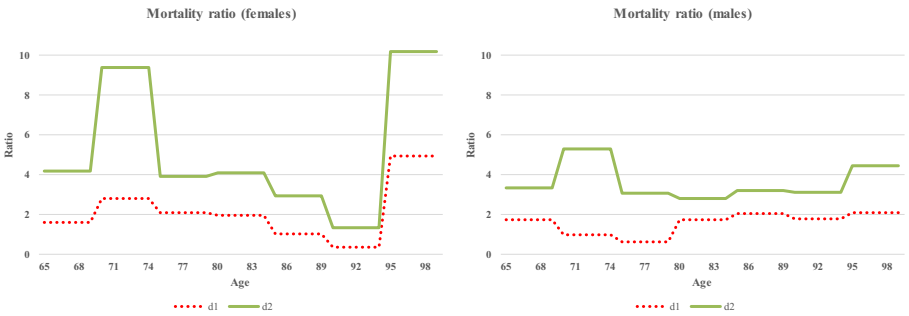


Fig. 5 Comparison of mortality ratios: dependent/active people. Source: Own based on data from Cui et al. (2022)

life expectancy would be in a better level of activity limitation, while for women it would be 62.36%.

From a demographic/epidemiological standpoint, the data for China are apparently coherent because life expectancy decreases when an individual's disability state worsens. However, the mortality ratios ($\overline{\delta}_{65}^{d}$) are lower than those shown for the US and higher than those for Australia, although the comparison is not entirely reliable because there are only two levels of dependence as opposed to 6 and 4 in the US and Australia respectively. The mortality ratio by age (Fig. 5, Graph 1 for females and Graph 2 for males) shows large irregularities for some sections that are very difficult to understand and justify. Inexplicable jumps occur for both males and females.

The synthetic LTC prevalence rates by age and gender are shown in Fig. 6. The highest average rate for males and females ($\overline{\lambda}_{65}^{d_1}$) (see Table 12) would be for the least severe state of dependence at 36.60 and 41.42% for males and females respectively. The total average rate ($\overline{\lambda}_{65}^{d_j}$) would be 49.47 and 65.15% for males and females respectively, which is far higher than the figures calculated in the case of Australia.

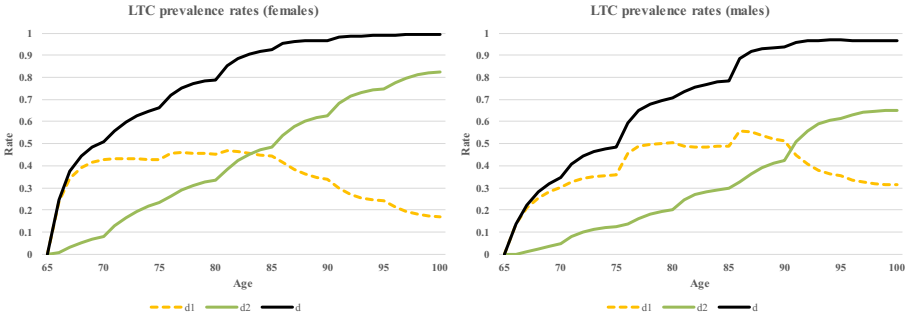


Fig. 6 Comparison of LTC prevalence rates. Source: Own based on data from Cui et al. (2022)

In the case of men, the least severe level of dependence d_1 would be predominant up to age 90, while for women it would be the most prevalent level up to age 82, with d_2 predominating from then on.

In the case of women, those aged 83 and over would already have an LTC prevalence rate of over 90%, while for men this level would be reached at age 87. The highest dependence rate (for all levels) would be attained at age 100, being 96.18% and 99.44% for males and females respectively.

According to the previous rates of LTC prevalence by age, the total number of dependent people would be greater than the active population after age 76 for men and age 70 for women.

Like we saw in the case of Australia, the range of values for the IQR^i_{65} is very similar for both men and women (Table 11). The increase in the level of activity limitation means a decrease in the average age of death for people still alive at age 65, and also for the median and the mode. However, an increase in the IQR^i_{65} value is observed, which means a significant increase in heterogeneity as regards individuals' deaths. We can see that the main difference between males and females is the value of the mode in the active state. The average age of death for people still alive at 65 and the median and the mode have similar values. However, although the value of the mode for women remains close to the other two indicators in d_1, in d_2 it decreases sharply and reaches a value lower than Q_1. The divergence in the case of men can already be observed in d_1, but in d_2 the mode stays clearly above the level of Q_1. The explanation for the different mode values by gender lies in the average number of deaths in each state. It should be remembered that the value is calculated as an average of the mode levels reached in each of the states separately.

4 Summary

The analysis carried out in the previous subsection showed that the biometric data sets analysed have several weaknesses that we summarize in Table 12.

The above table shows how difficult it is to build the biometric data sets needed to carry out an actuarially fair valuation of the prices for LTCI and LCA contracts.

The construction of biometric data sets is a difficult task, particularly when more than one state of dependence is considered along with the recovery assumption (Guibert and Planchet 2019). Bearing in mind that insurers have so far dealt with only a low volume of

cases, that the nature of the situation is constantly evolving, that the mortality law for LTC claimants consists of a mixture of pathologies, and that the estimation of crude death rates is very volatile (Planchet and Tomas 2016), it is not surprising that we found a number of flaws in the data sets analysed.

5 Discussion

From an LTC insurance perspective, HLE and life expectancy with activity limitation are key indicators and we are therefore very interested in the biometric assumptions made by the LTC insurance scheme's insurer or sponsor. Including information by breaking down life expectancy by health state is invaluable for illustrating the actuarial calculations used to price LTCI in general and LCAs in particular.

According to the equivalence principle used in insurance, the expected present values of premiums and benefits should be equal. Roughly speaking, therefore, premiums and benefits will balance on average (Norberg 2002). The main assumptions to be made when pricing LTCI are (Pitacco 2014; SOA 2016): longevity; morbidity/incidence rates; recovery rates; lapse rates; the interest rate; the loading for expenses on acquisition, collection and administration; inflation protection; and indexation of benefits in payment.

Why, then, are we so interested in LCAs? The uncertainty surrounding the need for LTC services presents a significant challenge in retirement planning. Preparing for the possible costs of future impairment and LTC is a task that everyone has to confront as they get older. The probability of losing physical functioning increases dramatically with age and is therefore a highly relevant consideration for the population aged 65 and above. According to Robinson's (1996) model that was used in the American insurance industry, for a healthy 65-year-old man, the probabilities of needing LTC services when he reaches the ages of 70, 80 and 90 are 8.39, 30.91 and 60.75% respectively, whereas for a healthy woman at the same age the probabilities are 10.19, 31.32 and 64.37%.

LCAs have several advantages. First, combining LTCI and life annuities would reduce adverse selection for the annuity portion, resulting in lower premiums and allowing the relaxation of underwriting standards for the LTCI portions. Second, LCAs would allow people to delay the purchase of LTCI until closer to retirement, which would have the further benefit of allowing insurers to better account for trends in disability, longevity, investment returns and other sources of risk in designing and pricing policies. The cash-for-care benefit in a combined annuity has the potential to simplify choices and reduce uncertainty about future benefits, which are two frequently cited reasons for explaining the low market penetration of conventional LTCI. Finally, if the individual has an existing health issue, they might find it easier to be approved for an annuity with an LTC rider instead of a standalone LTCI. This is a very important point, given that Cornell et al. (2016) show that in the target age range (50–71 years), approximately 30% of those whose wealth meets minimum industry standards for suitability for LTCI would have their applications rejected for this type of insurance at the underwriting stage.

In those countries for which the data sets have been analysed, the private LTCI market either does not exist or is very small. Private LTCI is not currently available in Australia (Sherris 2022), and in China the LTC insurance market is still at an early stage of development. Policy and regulation need to be improved before the market can enter a phase of high growth (Swiss Re 2022).

The private LTCI market in the US had a promising start at the beginning of the 80 s, but is now in a downturn. The LTCI premium rates developed by insurers from at least 10 years ago have generally turned out to be underpriced, and therefore many premium rate increases have been filed in this area (AAA 2018). By 2014 the average policy premiums (for the same benefits) had increased by 215% of the premiums charged in 2000 (AAA 2016). The variability of insurance premiums means that some of the spending risk of unknown LTC expenditure is transferred back to the policyholder. This means they participate in the risk of increased LTC costs or in claims that exceed those anticipated in the underwriting process. Indeed, in the US the issue of increased LTC premiums is a nationwide trend.

Most insurance companies underestimated the cost of paying claims and overestimated the number of people cancelling policies. When original LTCI policy forms were issued in the 1980s and 1990s, morbidity assumptions were often based on statistics for the general population, while lapse and mortality assumptions were based on experience with non-LTC insurance products. Inadequate medical underwriting was another source of deviation (AAA 2018).

Insurance companies are forced to adjust premiums to compensate for inaccurate pricing assumptions to ensure there are enough reserves to pay LTC claims under each plan. The result of all this along with other reasons that go beyond the scope of this paper is that the number of individual LTCI policies sold in the US declined to 49,000 in 2020. This was after they increased from 380,000 in 1990 to a peak of 754,000 in 2002. According to the National Association of Insurance Commissioners (NAIC), the number of insurers offering LTCI coverage has fallen from just over 100 in 2004 to around a dozen in 2021. Between 2012 and 2021 the number of individuals with an active or "in-force" stand-alone LTCI policy declined from 7.2 million to 5.3 million (AALTCI 2022). The continuous reduction in the number of LTCI policies sold and the declining size of the market could to some extent be a result of the very high prevalence of bad experiences undergone by policyholders due to premium increases, which in turn has led to a lessening of trust in insurance companies (Courbage and Nicolas 2021).

To conclude this section, we should ask ourselves whether it should be compulsory to disclose information about the HLE and the other longevity risk and morbidity indicators used in the technical bases for computing LTCI and/or LCA premiums?

The actuarial black box is an old term (Gutterman 1996). It can be defined as a situation in which the actuarial analysis has not been adequately explained to its users. This might be due to the use of a complicated actuarial model, a lack of understanding of what the actuary does, or simply poor communication. Individuals or couples who are interested in purchasing LTCI or an LCA in particular might feel they are up against something similar. For them, a black box could be the actuarial calculation done to compute the premium for an LCA.

The market for individual LTC products would greatly benefit from the elimination of as many of these black boxes as possible. We could start to eliminate them by disclosing more information about the biometric assumptions made to compute the premiums and by spending more time with possible purchasers to explain the meaning of concepts such us life expectancy, healthy life expectancy, and longevity and morbidity risks. Fuino et al. (2022) find that factors relating to the awareness and understanding of LTC are extremely relevant when purchasing LTCI. Boyer et al. (2020) suggest that to spark an interest in buying LTCI policies it is essential to make the population aware of the frequency and severity of the risk insured, and also to provide a clear explanation of the benefits that come from having an LTCI policy.

6 Conclusions and future research

The methodology used to obtain the life expectancy matrix based on an individual's initial health state plus the analysis carried out has enabled us to show various health and demographic indicators to help assess the coherence and quality of the different data sets analysed. These indicators are rarely presented when authors build their biometric data sets nor when they are used to calculate LTCI or LCA premiums, nor when they are used in research articles to estimate the future demand for LTC services in a particular country.

The data sets analysed confirm the existence of the so-called "male–female health-survival paradox".

Despite the fact that all the biometric data sets we have analysed have their own problems, we would say that the most coherent from a demographic/epidemiological standpoint are those in Hariyanto et al. (2014a, b) and Robinson (1996). On the basis of all the information we have analysed, the data set with the most consistency problems would be Cui et al. (2022).

Given that the construction of biometric data sets is a difficult task, particularly when more than one state of dependence is considered along with the recovery assumption, it is no surprise that insurance companies are becoming less willing to provide LTCI and LCA contracts, since there is a big problem with the biometric aspect when it comes to calculating actuarially fair premiums with any degree of certainty. In those countries (except the US) whose data sets we have analysed, the market for this type of product is either non-existent or underdeveloped.

From the point of view of potential purchasers of this type of insurance product, the disclosure of the summary health/longevity measures used in the technical bases to compute LTCI and LCA premiums would help them to understand the need to be protected against the cost of requiring LTC services and also make the computation of the actuarial factors transparent (Ventura-Marco et al. 2023). This information could also be embedded in the retirement calculators and other tools used by financial advisors, given that many people are simply unaware of longevity (and morbidity) risks (Hurwitz et al. 2022). This would be in line with the proposal to increase the transparency of complex products and strategies during the decumulation phase (Bär and Gatzert 2022).

Finally, it could be said that LTCI is one of the least secure insurance products for the insured person since the periodic increases in premiums clearly distort the traditional concept of insurance. Seen from this perspective, LCAs are certainly better for the annuitant because the premium is generally paid as a lump-sum when the individual is close to retirement.

Future research could extend the analysis carried out in this paper to other biometric data sets that we initially excluded for various reasons, such as the difficulty of collecting all the necessary data or because they did not use a Markovian structure.

Technical appendix

The life expectancy matrix depending on an individual's initial health state

The discrete multistate model (Fig. 1, Sect. 2.2.1) can be expressed as an $(n+2)$-state-age non-homogeneous discrete-time Markov chain. The yearly transition probabilities forming the stochastic one-year transition matrix (M_{x+k}) are shown in Table 13:

p_{x+k}^{aa} is the probability that an active person aged $x+k$, $k \in \{0, 1, \cdots, w-(x)\}$ will reach age $x+k+1$ in the same state. $p_{x+k}^{ad_j}$ is the probability that an active person aged $x+k$ will reach age $x+k+1$ in state of dependence d_j, $j \in \{1, 2, \ldots n\}$. p_{x+k}^{af} is the probability that

Table 13 M_{x+k}. Source: Own

Starting status, i	Ending status, j					
	a	d_1	d_2	…	d_n	f
a	p_{x+k}^{aa}	$p_{x+k}^{ad_1}$	$p_{x+k}^{ad_2}$	⋯	$p_{x+k}^{ad_n}$	p_{x+k}^{af}
d_1	$p_{x+k}^{d_1a}$	$p_{x+k}^{d_1d_1}$	$p_{x+k}^{d_1d_2}$	⋯	$p_{x+k}^{d_1d_n}$	$p_{x+k}^{d_1f}$
d_2	$p_{x+k}^{d_2a}$	$p_{x+k}^{d_2d_1}$	$p_{x+k}^{d_2d_2}$	⋯	$p_{x+k}^{d_2d_n}$	$p_{x+k}^{d_2f}$
⋯	…	…	…	…	….	….
d_n	$p_{x+k}^{d_na}$	$p_{x+k}^{d_nd_1}$	$p_{x+k}^{d_nd_2}$	⋯	$p_{x+k}^{d_nd_n}$	$p_{x+k}^{d_nf}$
f	0	0	0	0	0	1

an active person aged $x+k$ will die during the year. $p_{x+k}^{d_id_j}$ is the probability that a person ($i \in \{1,2,\dots n\}$-level dependent) aged $x+k$ will reach age $x+k+1$ in state of dependence $d_j, j \in \{1,2,\dots n\}$.

The rows and columns of the matrix embedded in Table 13 relate to the starting and ending health status respectively of each individual in the cohort.

As we are working with Markov processes, we apply classic recurrent Chapman-Kolmogorov equations to obtain the corresponding multiyear transition probabilities.

A key element when it comes to computing the value of LCAs is the probability that an active person aged x will reach age $x+k$ in any state of $j \in \{1,2,\dots,n\}$ dependence $({}_kp_{x_e+A}^{rd_j})$.

This probability can be obtained from the vector (or matrix row) of the multiyear transition probabilities $({}_kM_x)$, given the initial state a at age x. This vector is the product of the annual probabilities of transition among the different states, which can also be expressed as an annual transition matrix for each age attained. At time t, for the group of healthy people aged x, the resulting probabilities are:

$$\begin{aligned}\left({}_kp_x^{aa}, {}_kp_x^{ad_1}, {}_kp_x^{ad_2}, \dots, {}_kp_x^{ad_n}, {}_kp_x^{af}\right) &= {}_kM_x \\ &= \overbrace{u_1 \cdot M_x \cdot M_{x+1} \cdot \dots \cdot M_{x+k-2}}^{{}_{k-1}M_x} \cdot M_{x+k-1}\end{aligned} \tag{1}$$

where vector row u_1, with value one in the first position and zero in the rest of the positions, sets the true initial state a at age x. The individual could have become dependent in any year within the corresponding age period $(x, x+k]$.

${}_kp_x^{ad_j}$ is the probability that a healthy person aged x will reach age $x+k$ in level $j \in \{1,2,\dots,n\}$ of dependence:

$${}_{k-1}p_a^{aa} \cdot p_{x+k-1}^{ad_j} + \sum_{i=1}^{n} {}_{k-1}p_x^{ad_i} \cdot p_{x+k-1}^{d_id_j} \tag{2}$$

${}_kp_x^{aa}$ is the probability that a healthy individual aged x will reach age $x+k$ in the same state:

$${}_{k-1}p_x^{aa} \cdot p_{x+k-1}^{aa} + \sum_{j=1}^{n} {}_{k-1}p_x^{ad_j} \cdot p_{x+k-1}^{d_ja} \tag{3}$$

and finally, ${}_kp_x^{af}$ is the probability that an active person aged x will not reach age $x+k$. The individual could have died (active or dependent) in any year within the period:

$$ {}_{k-1}p_x^{aa} \cdot p_{x+k-1}^{af} + \sum_{j=1}^{n} {}_{k-1}p_x^{ad_j} \cdot p_{x+k-1}^{d_j f} \tag{4} $$

Under the framework described above, the life expectancy of active people disaggregated into healthy and unhealthy life years can be computed as:

$$ \begin{aligned} e_x^a &= \sum_{k=0}^{w-x} {}_kp_x^{aa} + \sum_{k=1}^{w-x} {}_kp_x^{ad_1} + \cdots + \sum_{k=1}^{w-x} \cdot {}_kp_x^{ad_{n-1}} + \sum_{k=1}^{w-x} {}_kp_x^{ad_n} \\ &= \overbrace{e_x^{aa}}^{\text{healthy life years}} + \underbrace{\sum_{j=1}^{n} e_x^{ad_j}}_{\substack{\text{unhealthy life years} \\ \text{with } j\text{-level LTC}}} \end{aligned} \tag{5} $$

If the individual aged (x) is in state (i), the vector row in this case with value one in the position corresponding to state (i) and zero in the rest of the positions, i.e. their health state is not considered able, then life expectancy can be expressed as:

$$ e_x^{d_i} = \overbrace{\sum_{k=0}^{w-x} {}_kp_x^{d_i d_i}}^{\text{In the same state}} + \overbrace{\sum_{k=1}^{w-x} {}_kp_x^{d_i a}}^{\text{Total recovery}} + \underbrace{\sum_{k=1}^{w-x} \sum_{\substack{j=1 \\ j \neq i}}^{n} {}_kp_x^{d_i d_j}}_{\text{Transitions to other states}} = \overbrace{e_x^{d_i d_i}}^{\text{In the same state}} + \underbrace{e_x^{d_i a}}_{\substack{\text{Free of} \\ \text{activity} \\ \text{limitation}}} + \overbrace{\sum_{j=1}^{i-1} e_x^{d_i d_j}}^{\substack{\text{In better} \\ \text{states of} \\ \text{dependence}}} + \underbrace{\sum_{j=i+1}^{n} e_x^{d_i d_j}}_{\substack{\text{In worse} \\ \text{states of} \\ \text{dependence}}} \tag{6} $$

where ${}_kp_x^{d_i d_i}$ is the probability that a person with level d_i dependence aged (x) will reach age $(x+k)$ in the same state of dependence. The probability that a person will remain continuously in the same state for (k) length of time, occupancy probability term ${}_kp_x^{\overline{d_i d_i}}$, is included here:

$$ \begin{aligned} {}_kp_x^{d_i d_i} &= {}_kp_x^{\overline{d_i d_i}} + \sum_{s=1}^{k} {}_{s-1}p_x^{d_i a} \cdot p_{x+s-1}^{ad_i} \cdot {}_{k-(x+s)}p_{x+s}^{\overline{d_i d_i}} \\ &+ \sum_{\substack{j=1 \\ j \neq i}}^{n} \sum_{s=1}^{k} {}_{s-1}p_x^{d_i d_j} \cdot p_{x+s-1}^{d_j d_i} \cdot {}_{k-(x+s)}p_{x+s}^{\overline{d_i d_i}} \end{aligned} \tag{7} $$

${}_kp_x^{d_i d_j}$ is the probability that a person with level d_i dependence aged (x) will reach age $(x+k)$ in level j of dependence. $\sum_{j=1}^{i-1} e_x^{d_i d_j}$ indicates how many years of their total remaining life expectancy the individual can expect to live in better states of dependence. $\sum_{j=i+1}^{n} e_x^{d_i d_j}$ indicates how many years of their total remaining life expectancy the individual can expect to live in worse states of dependence.

Acknowledgements We would like to thank the seminar participants at the University of Salamanca (Asepuma 2023), the University of Alcalá, Spain, the University of Evora, Portugal (Asepelt 2023), and Peter Hall for his help with the English text. The comments and suggestions made by two anonymous referees were extremely helpful in improving the paper. Any errors are entirely due to the authors. A previous version (working paper) was published by the Instituto Complutense de Análisis Económico (ICAE): ICAE Working Paper, ISSN-e 2341-2356, N°. 2, 2023, págs. 1-49 https://eprints.ucm.es/id/eprint/77044/1/2302.pdf.

Author contributions CVM conceived the article and was in charge of overall direction and planning. MVM and CVM designed the model. CVM, AMG and MVM verified the analytical methods. MVM and CVM contributed to sample preparation. MVM designed and performed the computational framework. CVM, AMG and MVM contributed to the interpretation of the results. AMG and CVM wrote the paper with input from all authors. All authors have accepted responsibility for the entire content of this manuscript and approved its submission.

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature. This work was supported by the Spanish Ministry of Science, Innovation and Universities MCIN/AEI/, I+D+i PID2020-114563GB-I00, and by the Basque Government "BIRTE" IT1461-22.

Declarations

Conflict of interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

American Academy of Actuaries (AAA): Understanding Premium Rate Increases on Private Long-term Care Insurance. Issue Brief https://www.actuary.org/content/understanding-premium-rate-increases-private-long-term-care-insurance (2016). Accessed 26 Dec 2022

American Academy of Actuaries (AAA): Long-Term Care Insurance: Considerations for Treatment of Past Losses in Rate Increase Requests. Issue Brief https://www.actuary.org/content/long-term-care-insurance-considerations-treatment-past-losses-rateincrease-requests (2018). Accessed 14 Dec 2022

American Association for Long-Term Care Insurance (AALTCI): Long-Term Care Insurance Facts-Data-Statistics-2022 Reports. www.aaltci.org (2022). Accessed 11 Dec 2022

Australian Institute of Health and Welfare (AIHW): Healthy life expectancy in Australia: patterns and trends 1998 to 2012. Bulletin 126. https://www.aihw.gov.au/reports/life-expectancy-death/healthy-life-expectancy-in-australia-patterns/formats (2014). Accessed 7 Oct 2022

Baione, F., Levantesi, S.: Pricing critical illness insurance from prevalence rates: Gompertz versus Weibull. N. Am. Actuar. J. **22**(2), 270–288 (2018)

Bär, M., Gatzert, N.: Products and strategies for the decumulation of wealth during retirement: insights from the literature. N. Am. Actuar. J. (2022). https://doi.org/10.1080/10920277.2022.2078374

Biessy, G.: Continuous-time semi-Markov inference of biometric laws associated with a long-term care insurance portfolio. ASTIN Bull. **47**(2), 527–561 (2017)

Boyer, M., Donder, P.D., Fluet, C., Leroux, M.L., Michaud, P.C.: Long-term care insurance: information frictions and selection. Am. Econ. J. Econ. Policy **12**(3), 134–169 (2020)

Canudas-Romo, V.: Three measures of longevity: time trends and record values. Demography **47**(2), 299–312 (2010)

Cornell, P.Y., Grabowski, D.C., Cohen, M., Shi, X., Stevenson, D.G.: Medical underwriting in long-term care insurance: market conditions limit options for higher-risk consumers. Health Aff. **35**(8), 1494–1503 (2016)
Courbage, C., Nicolas, C.: Trust in insurance: the importance of experiences. J. Risk Insur. **88**(2), 263–291 (2021)
Crimmins, E.M., Zhang, Y., Saito, Y.: Trends over 4 decades in disability-free life expectancy in the United States. Am. J. Public Health **106**(7), 1287–1293 (2016)
Cui, X., Duan, X., Chang, C.H.T., Jiang, S.: Health transition probability and long-term care cost estimation. Math. Probl. Eng. (2022). https://doi.org/10.1155/2022/7980111
Di Lego, V.: Health expectancy indicators: what do they measure? Cad. Saude Colet. **29**, 115–129 (2021)
Friedberg, L., Hou, W., Sun, W., Webb, A., Li, Z.: New Evidence on the Risk of Requiring Long-Term Care. WP 2014-12. Center for Retirement Research at Boston College, Chestnut Hill (2014)
Fuino, M., Ugarte-Montero, A., Wagner, J.: On the drivers of potential customers' interest in long-term care insurance: evidence from Switzerland. Risk Manag. Insur. Rev. **25**, 271–302 (2022)
Guibert, Q., Planchet, F.: Measuring long-term insurance contract biometric risks. In: Dupourqué, E., Planchet, F., Sator, N. (eds.) Actuarial Aspects of Long Term Care. Springer Actuarial. Springer, Cham (2019)
Gutterman, S.: The actuarial black box. The Actuary, p. 3 (1996).
Hariyanto, E., Dickson, D., Pitt, G.W.: Estimation of disability transition probabilities in Australia I: preliminary. Ann. Actuar. Sci. **8**(1), 131–155 (2014a)
Hariyanto, E., Dickson, D., Pitt, G.W.: Estimation of disability transition probabilities in Australia II: implementation. Ann. Actuar. Sci. **8**(1), 156–175 (2014b)
Hurwitz, A., Mitchell, O.S., Sade, O.: Testing methods to enhance longevity awareness. J. Econ. Behav. Organ. **204**, 466–475 (2022)
Imai, K., Soneji, S.: On the estimation of disability-free life expectancy. J. Am. Stat. Assoc. **102**(480), 1199–1211 (2007)
Jiao, K.: Inequality of healthy life expectancy for the Chinese elderly and its trend. J. Chin. Sociol. **6**, 22 (2019)
Kannisto, V.: Mode and dispersion of the length of life. Popul. Engl. Sel. **13**(1), 159–171 (2001)
Majer, I.M., Stevens, R., Nusselder, W.J., van Mackbenbach, J.P., Baal, P.H.M.: Modeling and forecasting health expectancy: theoretical framework and applications. Demography **50**, 673–697 (2013)
Manton, K.G., Land, K.C.: Active life expectancy estimates for the U.S. elderly population: a multidimensional continuous-mixture model of functional change applied to completed cohorts, 1982–1996. Demography **37**(3), 253–265 (2000)
Molla, M.T., Wagener, D.K., Madans J.H.: Summary measures of population health: methods for calculating healthy life expectancy. Healthy People 2010 Stat Notes, pp. 1–11 (2001)
Nielsen, C., Juel, L., Jeune, A.B., Christensen, K., Lindahl-Jacobsen, R.: Healthy life expectancy by frailty state in Europe from 2004 to 2015: findings from SHARE. Eur. J. Public Health **31**(3), 554–560 (2021)
Norberg, R.: Basic life insurance mathematics. https://web.math.ku.dk/~mogens/lifebook.pdf (2002) Accessed 11 Nov 2022
Pitacco, E.: Health Insurance. Basic actuarial Models. EAA Series. Springer (2014)
Planchet, F., Tomas, J.: Uncertainty on survival probabilities and solvency capital requirement: application to long-term care insurance. Scand. Actuar. J. **4**, 279–292 (2016)
Pla-Porcel, J., Ventura-Marco, M., Vidal-Meliá, C.: Converting retirement benefit into a life care annuity with graded benefits. Scand. Actuar. J. **10**, 829–853 (2017)
Pla-Porcel, J., Ventura-Marco, M., Vidal-Meliá, C.: Life care annuities (LCA) embedded in a notional defined contribution (NDC) framework. ASTIN Bull. **46**(2), 331–363 (2016)
Robinson, J.: A long-term care status transition model. In: Proceedings of the Old-Age Crisis—Actuarial Opportunities: The 1996 Bowles Symposium, pp. 72–79, Georgia State University, Atlanta (1996)
Saito, Y., Robine, J.M., Crimmins, E.M.: The methods and materials of health expectancy. Stat. J. IAOS **30**(3), 209–223 (2014)
Sherris, M.: Interviewed in Australia looks toward long-term care insurance to provide aged car. https://www.ltcnews.com/articles/australia-looks-toward-long-term-care-insurance-to-provide-aged-care (2022). Accessed 19 Nov 2022
Society of Actuaries (SOA): Long-Term Care Insurance: The SOA Pricing Project. https://www.soa.org/globalassets/assets/files/static-pages/sections/long-term-care/ltc-pricing-project.pdf (2016). Accessed 25 September 2022.
Sullivan, D.F.: A single index of mortality and morbidity. HSMHA Health Rep. **86**(4), 57–74 (1971)

Swiss Re: Private long-term care insurance opportunities in China. https://www.swissre.com/institute/research/topics-and-risk-dialogues/china/long-term-care-insurance-china.html (2022). Accessed 25 Sept 2022

Van Raalte, A., Caswell, H.: Perturbation analysis of indices of lifespan variability. Demography **50**(5), 1615–1640 (2013)

Vanella, P., Heß, M., Wilke, C.B.: A probabilistic projection of beneficiaries of long-term care insurance in Germany by severity of disability. Qual. Quant. **54**, 943–974 (2020)

Ventura-Marco, M., Vidal-Meliá, C., Pérez-Salamero González, J.M.: Joint life care annuities to help retired couples to finance the cost of long-term care. IInsur. Math. Econ. **113**, 122–139 (2023). https://doi.org/10.1016/j.insmatheco.2023.08.002

Vidal-Meliá, C., Ventura-Marco, M., Pla-Porcel, J.: An NDC approach to helping pensioners cope with the cost of long-term care. J. Pension Econ. Finance **19**(1), 80–108 (2020)

Wilmoth, J.R., Horiuchi, S.: Rectangularization revisited: variability of age at death within human populations. Demography **36**(4), 475–495 (1999)

Zhen, Z., Junhan, D., Chenyuan, Z., Qiang, L.: Trends of healthy life expectancy of the elderly in China in 1994–2015: revisiting from the perspective of morbidity transition. Front. Public Health **9**, 774205 (2022)

Authors and Affiliations

Carlos Vidal-Meliá[1,2] · Manuel Ventura-Marco[1] · Anne M. Garvey[3]

✉ Carlos Vidal-Meliá
Carlos.Vidal@uv.es

Manuel Ventura-Marco
Manuel.Ventura@uv.es

Anne M. Garvey
Anne.Garvey@uah.es

[1] Department of Financial Economics and Actuarial Science, University of Valencia, Valencia, Spain

[2] Instituto Complutense de Análisis Económico (ICAE), Complutense University of Madrid (Research Affiliate), Madrid, Spain

[3] Department of Economics and Management Sciences, University of Alcalá, Madrid, Spain