Demography (2024) 61(6):1715–1730
DOI 10.1215/00703370-11696463 Published online: 5 December 2024

Modeling Disability-Free Life Expectancy With Duration Dependence: A Research Note on the Bias in the Markov Assumption

Tianyu Shen and James O'Donnell

ABSTRACT Demographic studies on healthy life expectancy often rely on the Markov assumption, which fails to consider the duration of exposure to risk. To address this limitation, models like the duration-dependent multistate life table (DDMSLT) have been developed. However, these models cannot be directly applied to left-censored survey data, as they require knowledge of the time spent in the initial state, which is rarely known because of survey design. This research note presents a flexible approach for utilizing this type of survey data within the DDMSLT framework to estimate multistate life expectancies. The approach involves partially dropping left-censored observations and truncating the duration length after which duration dependence is assumed to be minimal. Utilizing the U.S. Health and Retirement Study, we apply this approach to compute disability-free/healthy life expectancy (HLE) among older adults in the United States and compare duration-dependent models to the typical multistate model with the Markov assumption. Findings suggest that while duration dependence is present in transition probabilities, its effect on HLE is averaged out. As a result, the bias in this case is minimal, and the Markov assumption provides a plausible and parsimonious estimate of HLE.

KEYWORDS Markov assumption • Duration dependence • Multistate model • Longitudinal survey data • Healthy life expectancy

Introduction

Multistate life tables (MSLTs) are a generalization of the life table to model processes involving multiple and recurrent types of events (Schoen 1988). They are most commonly used to capture "lifetime" expectancies of time spent in the states defined by the state space. Recent papers have used these methods in fields spanning sociology, demography, and epidemiology (e.g., Brown et al. 2021; Hosokawa et al. 2023; Neumann et al. 2022; Zaniotto et al. 2020). Almost all MSLT studies in this field rely on the Markov assumption. In the health context, the Markov assumption means that an individual's probability of transitioning between health states is a function of their current health status and perhaps a set of other characteristics, including age and sociodemographic characteristics, but *not* their past health histories.

ELECTRONIC SUPPLEMENTARY MATERIAL The online version of this article (https://doi.org/10.1215/00703370-11696463) contains supplementary material.

The fallibility of this assumption has been established across social, demographic, health, and economic contexts, particularly in demonstrating the importance of duration-specific transitions (Belanger 1989; Cai et al. 2008; Crowther and Lambert 2014; Maddox et al. 1994; O'Donnell 2021; van den Berg and van Ours 1996). In studies of health and disability, Maddox et al. (1994) and Cai et al. (2008) found that the risks of disability or impairment, along with the chances of recovery, are duration dependent, and specifically that prospective transition risks generally decrease the longer a person has been in the same health state. Although widely acknowledged as a potential limitation (Cha et al. 2021; Jia and Lubetkin 2020; Xu and Payne 2024), the Markov assumption is rarely tested or addressed, and popular multistate software, ImaCh, relies exclusively on the Markov assumption (Lièvre et al. 2003). In estimating multistate life expectancies, Markov models may average out the effects of duration dependence, preventing bias and any need for concern. However, bias is potentially introduced by the issue of right-censoring and the potential for duration dependence to impact on transition probabilities beyond survey observation windows.

Demographers and others have developed approaches for introducing duration dependence to multistate analyses (e.g., Steele et al. 2004; Wolf 1988). The MSLT with duration dependence (abbreviated to DDMSLT) has been discussed since the 1980s following the development of a discrete-time semi-Markov approach (e.g., Hennessey 1980; Littman and Mode 1977). Wolf (1988) proposed a generalized MSLT depending on the duration of risk exposure and showed that DDMSLT, though inherently a non-Markovian process, embeds the Markovian component. DDMSLT was developed and applied to study marital transitions, in recognition that the risk of divorce varies by the length of time spent married (Belanger 1989; Schoen 2021; Wolf 1988), but is applicable to a range of contexts. The approach assumes that no transitions occur between intervals. While this assumption may introduce potential bias in the results, Wolf and Gill (2009) demonstrated that existing models, such as the event-history model and embedded Markov chain model, perform comparably under this condition. Moreover, they noted that no model could perfectly reconstruct the unknown underlying truth.

Duration-dependent models partially relax the Markov assumption while maintaining computational efficiency. The popularity of the Markov assumption in health contexts stems from its link to biological aging (Crimmins et al. 2021; Hayflick 2007) and ability to simplify complex life histories. Markov models provide useful insights into health transitions and life expectancies based on chronological age. Including duration dependence adds complexity but may better capture aging by accounting for heterogeneity. While duration may not be a direct causal influence on health transitions, controlling for duration may help to control for heterogeneity in the aging process, on the premise, for example, that accelerated aging is related to extended durations of ill-health (Crimmins et al. 2021). Incorporating duration dependence may therefore help control for variations in biological aging rates not reflected by chronological age alone (Levine 2013; Rockwood and Mitnitski 2007).

Left-censoring in longitudinal surveys is a common problem limiting the application of duration-dependent models. One of the key inputs is information on how long people have been living in their present state. However, in left-censored surveys, the timing of events before the observation period is unknown and we do not know how long respondents have been in the state in which they were first observed (Payne and

Kobayashi 2022; Payne et al. 2013). Solutions to this problem have been proposed but are imperfect (Guo 1993). Discarding all left-censored observations (Allison 1984) is simple but involves a potentially large loss of information (Cai et al. 2006). As an alternative, Cai et al. (2006) suggested imputing the duration of left-censored observations. However, the imputation method is computationally heavy and relies on non-left-censored durations to impute the unobserved durations, ignoring any unmeasured differences between the censored and noncensored subsamples. These challenges contribute to gaps in understanding duration dependence in multistate outputs, including healthy life expectancy (HLE).

This research note develops and tests options to address the gap in incorporating duration dependence into HLE estimates using survey data on disability transitions in the United States. It proposes parsimonious variations of semi-Markov multistate models that seek to incorporate duration dependence in ways that minimize data loss and bias, providing useful information to researchers in identifying and testing duration dependence and its impact on HLE estimates. The twofold objectives are to modify and improve the flexibility of the DDMSLT (Wolf 1988) to make it feasible on left-censored survey data and to enable examination of the impacts of duration dependence on estimates of HLE. Thus, we provide empirical evidence regarding whether the widely noted limitation of Markov models in ignoring duration dependence introduces serious bias in HLE estimates.

Methods

Models

The state space of a typical three-state Markov MSLT for estimating HLE is depicted in panel a of Figure 1. There are two transient states (healthy and unhealthy) and one absorbing state (death). If this model is by single year of age, the health state of the next age would depend only on the health state of the current age. A semi-Markov process (SMP) is very similar to the Markov model except that it assumes that the health state of the next age depends on the state in the current age and the duration in this current state (conceptually shown in panel b of Figure 1). Thus, the probabilities of transitioning from one state to the next in the SMP model are conditional on the length of time spent in the origin state.

Panel a of Figure 2 shows the state space and the pathways of transition in the DDMSLT model developed by Wolf (1988). DDMSLT treats duration as a categorical variable and transition risks as a piecewise constant. In our application of DDMSLT, duration-specific transition risks remain constant within one-year intervals, up to an open-ended duration category. For example, person-wave observations have a duration of zero years if they have been in the same state for less than one year (Dur 0). If an individual remains in the same health status for more than one year, that is indicated by Dur 1; if they remain in the same health state but advance to the next duration category, then Dur 2. After three years, individuals reach and remain in an open-ended duration category of 3+ years (Dur 3+). If their health status changes, they move to a new state and their duration is reset to Dur 0. An individual can die from any state, indicated aggregately by the blue arrows.

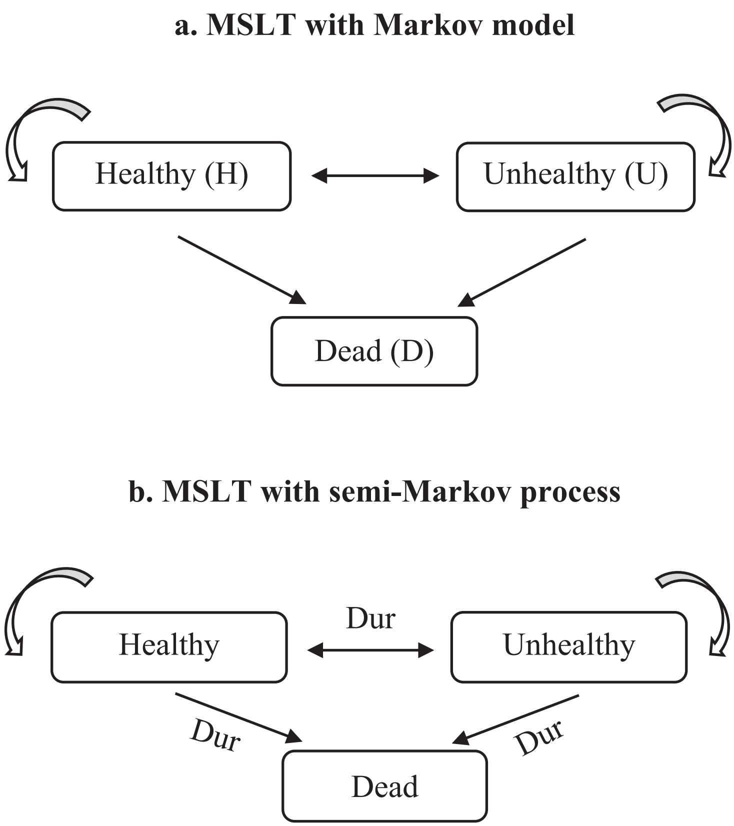

Fig. 1 State space of multistate life table (MSLT) models. "Dur" represents duration.

The model requires a known duration for each health state and so cannot handle left-censored observations without duration. We propose a modification—truncated DDMSLT or T-DDMSLT—to allow the model to utilize some of the observations without a known origin. Instead of discarding all left-censored observations (Allison 1984), we drop only observations with unknown origin up until a prespecified truncation point that aligns with the open-ended duration category. In panel b of Figure 2, this truncation point is set to three years. In other words, any observations in which the individual is observed in the same state for more than three years are included in Dur 3+. All other left-censored observations are dropped. In section 1 of the online appendix, we present some examples based on hypothetical observation.

The proposed modified model can be understood as a piecewise function—a semi-Markov process below a certain duration, and a Markov process independent of duration above that point. By incorporating duration in each health state up to the truncation, the approach retains more information than discarding all left-censored observations, at the cost of assuming duration dependence is no longer important after a certain time.

A further modification to the model utilizing this feature—truncated DDMSLT with health history, or T-DDMSLT-H—is shown in panel c of Figure 2. This modification includes an extra Markov state identifying individuals with no history of being

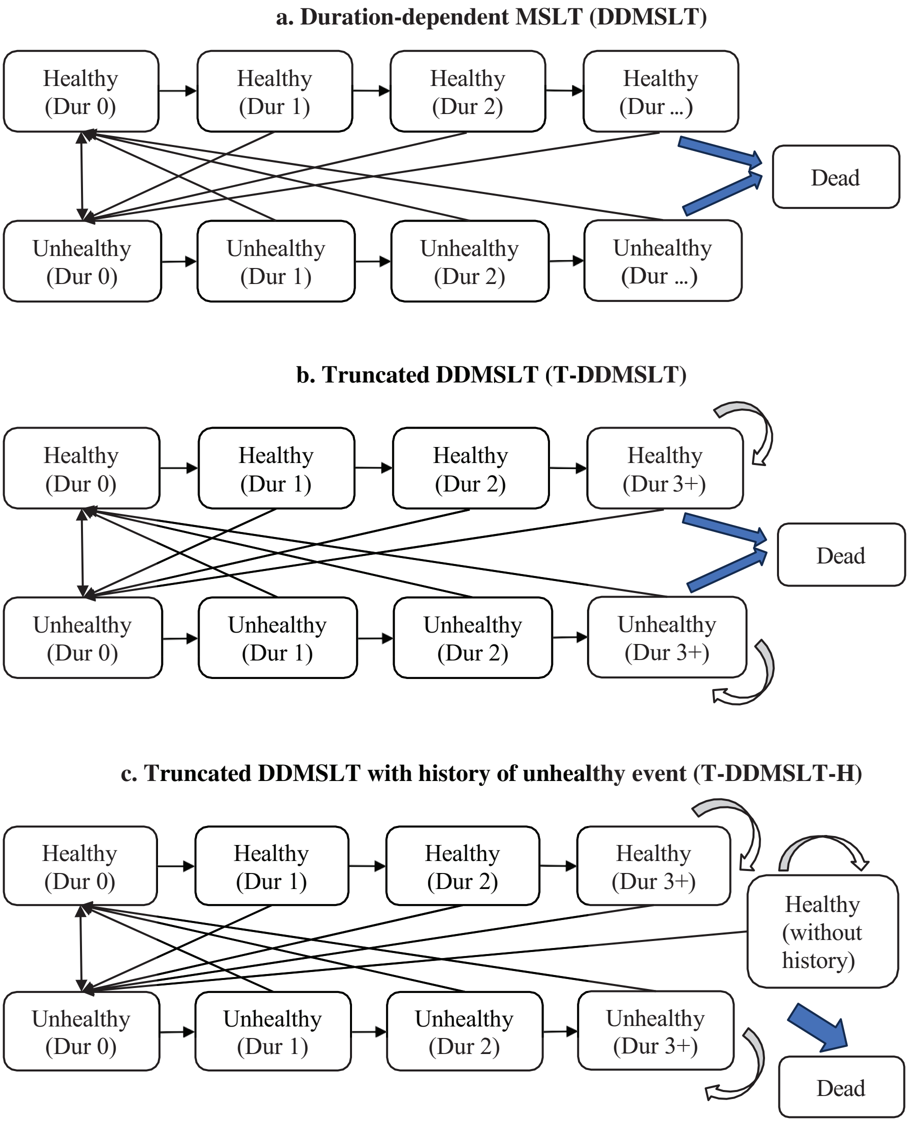

Fig. 2 State space and transition pathway of different models. "Dur" represents duration. In panel a, "Dur . . ." aggregates other states with longer duration. For example, "Dur 3+" in panels b and c includes the states with 3 and above. "3" can be changed to any other truncation of duration. The blue arrows represent transitions to death from all states.

unhealthy during the survey observation period (*cf.* Bardenheier et al. 2016). This state records left-censored observations after a truncation point (e.g., the third year) as healthy with no history, assuming no prior unhealthy events. The other "healthy" state groups individuals with a history of being unhealthy (see section 1 of the online appendix). As in the previous model, their transition probabilities are duration dependent up to the truncation point. This model better differentiates between long-term healthy individuals and those making transitions. By employing these different model designs in Figures 1 and 2, transition probabilities are estimated from empirical data and life expectancy estimates are compared.

Data

We use data from the U.S. Health and Retirement Survey ([HRS] 2023), a biannual, national, longitudinal survey, from Wave 5 (2000) to Wave 15 (2020). We select a birth cohort, 1936–1945, with an average age of about 60 in 2000, to estimate the cohort disability-free/healthy life expectancy.

The HRS is conducted biannually,[1] so to better model duration we start by converting the data to a single-year scale. This conversion has two stages. The first stage is based on the standard approach in the literature, in which the interwave state is imputed using two consecutive preceding and succeeding waves by random assignment (Lang and Little 2018; Payne 2022; Raymo et al. 2019). In this approach, individuals with the same health status at two consecutive waves would remain in that state within the two-year interval. The second stage seeks to relax this assumption and impute unobserved interwave transitions.

Using the first imputation, we estimate the transition probabilities by age and sex and simulate 500,000 individual trajectories from ages 50 to 100. We impute the interwave health status by using one of the simulated trajectories best matched to the full observed history of an individual with a two-year interval and filling the interwave period with the simulated trajectories as that sex and age.[2] Unobserved health transitions are imputed through these simulated trajectories and added to individual health histories. The transition probabilities used to simulate the trajectories contain Markovian properties, however, the simulated trajectories are calculated and constrained by the observed health status at the start and end of each interwave period. Nevertheless, we acknowledge that the approach may still understate interwave transitions.

In this study, we focus on disability to derive estimates of "healthy" and "unhealthy" life expectancies, while acknowledging that health is a broader and more multidimensional aspect of life. Disability is measured by reports of difficulty in any of the five basic activities of daily living; individuals are classified as "disabled" or "unhealthy" if they report any difficulty (Freedman et al. 2002) and as "healthy" or "disability-free" otherwise. Mortality is captured through linkage to the U.S. death registry. Importantly, these models can be replicated with a wide range of indicators, including, for example, self-rated health and disease incidence.

[1] Observation intervals are typically two years, based on age differences between interviews. Two-year or three-year intervals are converted to annual transitions, while leaving the one-year interval unchanged. Intervals exceeding three years are not imputed. While not exact, this method approximates annual changes.

[2] Typically, one individual can be matched to several simulated trajectories, which we would randomly select from. In rare cases, typically in older age, where no simulated trajectory can fully match an individual's observed health history, we would use the trajectory with the most matches. Observed health statuses remain unchanged during imputation.

Estimation Procedures

We estimate and compare four models: MSLT, T-DDMSLT, T-DDMSLT-H, and a general SMP model in which all left-censored observations are discarded (referred to as T-SMP). Two sets of inputs are required to calculate the HLE: the baseline disability distribution and age- and duration-specific transition probabilities. The 1936–1945 cohort was aged 55–64 in the year 2000. Because there is no duration information in the first wave, we discard the first five years of observation and construct the baseline on the basis of respondents aged 60–69. The starting point for each model is therefore a cohort in their 60s, with an average age of 65 and a health state corresponding to the observed distribution among respondents five years into the survey. A similar test is conducted with the older 1926–1935 cohort at an average initial age of 75 (see section 2 of the online appendix).

The models we test require us to discard varying numbers of observations. At one end of the spectrum, the standard Markov MSLT keeps all observations. At the other end of the spectrum, the T-SMP requires us to discard all left-censored (unknown duration) observations. In between, the proposed T-DDMSLT models require us to drop all left-censored observations up to prespecified truncation points. We present results with two truncation points: three years and five years. Four data subsets were used: (1) full data; (2) data dropping unknown duration of three years (Truncated 3); (3) data dropping five years (Truncated 5); and (4) data without unknown origins (no unknown).

The first significant contribution of our proposed approach is in substantially reducing data loss and retaining the original sample's representativeness. Panel A of Table 1 presents the characteristics of the baseline of the full data for MSLT and the subsamples dropping unknown durations of three and five years for T-DDMSLT and all the unknown durations for T-SMP. Panel B of Table 1 presents the characteristics of the transitions in each subsample. Discarding all unknown durations severely reduces the sample size and observed number of transitions, while skewing the demographic profile of the sample. The sample for T-SMP records substantially fewer transitions (23%) than the full data and is particularly poor in capturing transitions from the disability-free state (15% of the full data), reflecting the fact that most survey participants were disability-free on first entering the study. As a result, the T-SMP sample is substantially skewed toward participants with a history of disability and who are older, less likely to be White, and less likely to have high school or postschool qualifications. By contrast, the samples for T-DDMSLT preserve 70–81% of transitions and retain the same demographic and socioeconomic profile as the full data.

Following common practice in the literature (e.g., Cai et al. 2010; Cai et al. 2006; Shen and Payne 2023), we apply multinomial logistic regression to estimate the transition probabilities. In the MSLT model, apart from disability status time t, the other covariates include age at time t, age squared, sex, and interactions between age and sex. For the T-SMP model, the duration is included as a continuous independent variable, and the age is recorded at the start of the duration with the rest of the covariates the same. The T-DDMSLT models are similar except that duration is treated as a categorical variable. The resulting age- and duration-specific transition probabilities are combined with the respective baseline in a multistate life table to estimate disability-free life expectancy. Bootstrap resampling from the

Table 1 Characteristics (N and %) of the sample for birth cohort 1936–1945

	Full Data	Truncated 3	Truncated 5	No Unknown
A. Baseline at Age 65				
N	6,897	6,420	6,039	1,804
Disability-free (all)	6,042	5,520	5,009	556
0		320	393	445
1		52	52	52
2		31	31	31
3/3+		5,117	19	19
4			9	9
5/5+			4,505	0
Disabled (all)	855	900	1,030	1,248
0		460	702	1,159
1		56	56	56
2		19	19	19
3/3+		365	9	9
4			4	4
5/5+			240	1
Sex (%)				
Male	44.6	43.7	43.0	40.5
Female	55.4	56.3	57.0	59.5
Race and ethnicity (%)				
White	71.6	72.1	72.8	65.2
Black	16.2	15.8	15.4	19.7
Hispanic	10.0	9.9	9.6	12.5
Other	2.2	2.2	2.2	2.6
Educational attainment (%)				
Below high school	22.8	22.4	22.0	30.7
High school	36.9	37.0	37.3	36.0
Above high school	40.2	40.6	40.7	33.3
B. Transitions Between 65 and 80				
Number of transitions (all durations)	104,502	84,269	73,158	24,290
HH	84,126	66,788	56,798	11,759
HU	3,870	3,265	2,959	1,448
HD	1,219	1,054	956	302
UH	2,738	2,319	2,228	2,146
UU	11,462	9,852	9,258	7,804
UD	1,087	991	959	831
Age (average)	68.9	70.5	71.4	71.5
Sex (%)				
Male	42.7	42.2	41.9	39.5
Female	57.3	57.8	58.1	60.5
Race and ethnicity (%)				
White	72.6	72.9	72.8	65.5
Black	15.2	15.0	14.9	18.4
Hispanic	9.9	9.9	10.0	13.5
Other	2.2	2.2	2.2	2.6
Educational attainment (%)				
Below high school	21.1	20.9	20.9	29.7
High school	36.8	36.8	36.8	35.6
Above high school	42.0	42.3	42.3	34.6

Note: "H" represents healthy/disability-free, "U" unhealthy/disabled, and "D" dead.

Source: Authors' calculations based on Health and Retirement Study, 2000–2020 (2023).

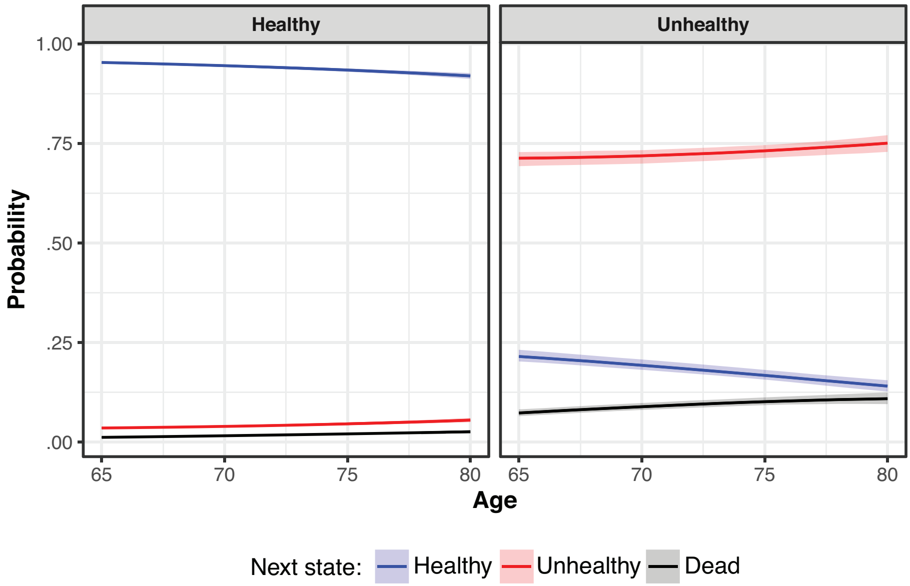

Fig. 3 Age-specific transition probabilities for the male cohort 1936–1945 with MSLT (full data). Shading represents 95% confidence intervals. *Source:* See Table 1.

original dataset is used to generate confidence intervals. All calculations are done in R 4.3.0 (R core team 2023) and can be replicated (see https://github.com/tyaSHEN/HLEmarkov).

Results

In Figures 3 and 4, we present estimated transition probabilities among males utilizing various models and data samples. Figure 3 delineates age-specific transition probabilities from 65 to 80 estimated using the MSLT model with full data, stratified by the current health state for the cohort born between 1936 and 1945. The left panel illustrates disability-free states, while the right panel portrays disability states. Colored lines denote the probability of transitioning to the subsequent state. For instance, the probability of remaining disability-free (indicated by the red line in the left panel) stands at approximately 95% at age 65.

Figure 4 illustrates the probabilities of remaining healthy (panel a) and unhealthy (panel b) for males at age 70 for each sample and model. Markov MSLT probabilities are calculated on each subsample and displayed in each panel to show the effect of data loss in each model. Notably, these probabilities are similar between the full and truncated samples, but significantly lower in the sample with all left-censored observations discarded. The disparity in probabilities between T-SMP and T-DDMSLT arises partly from T-SMP treating duration as a continuous function. In T-DDMSLT, retention probabilities exhibit moderate fluctuations over initial durations before gradually increasing, while T-SMP shows a smoother increase with expanding confidence intervals. Retention probabilities in both T-SMP and T-DDMSLT models are

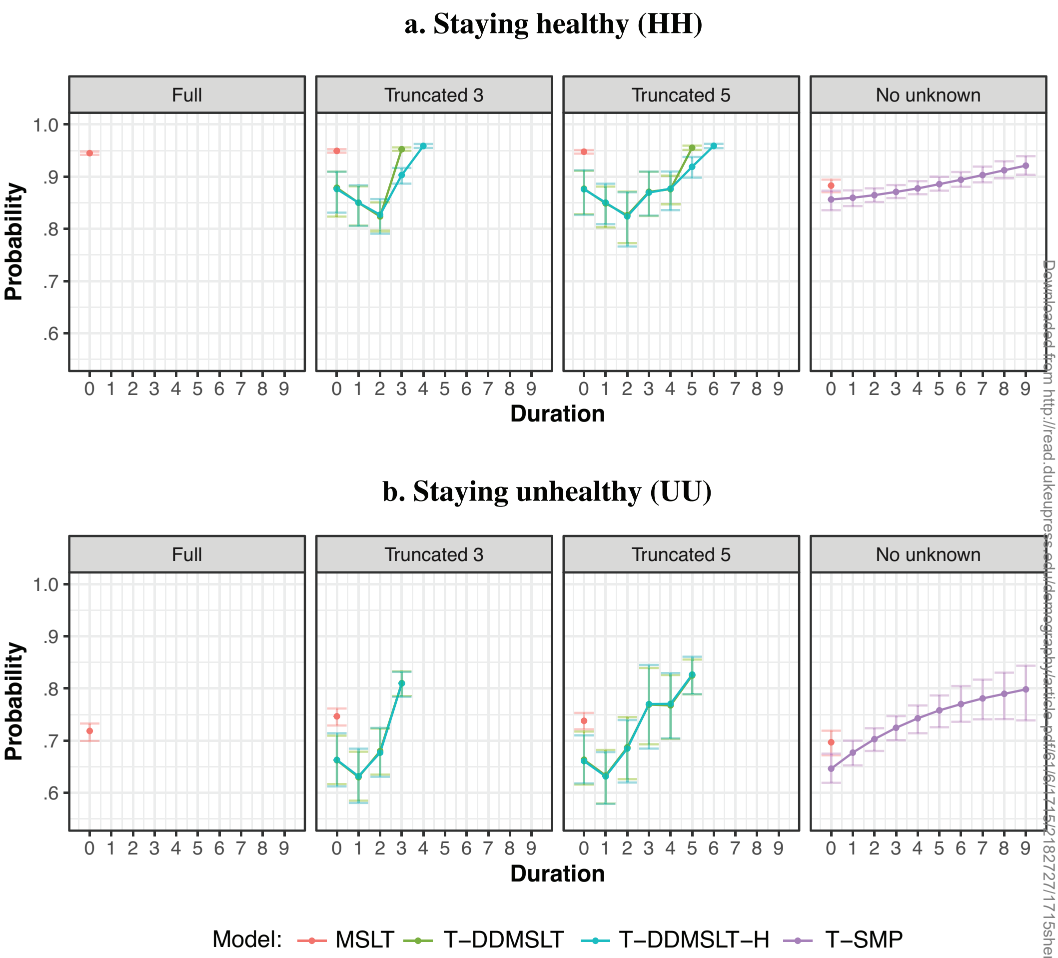

Fig. 4 Transition probabilities for different models at age 70 for the male cohort 1936–1945. In panel a, last duration in the model T-DDMSLT-H refers to the special state: healthy without history of unhealthy events recorded within the truncation of years. Error bars represent 95% confidence intervals. The T-SMP model is used to estimate the subsample with no unknown duration. *Source:* See Table 1.

notably lower for shorter durations compared with MSLT, increasing with duration. This demonstrates duration dependence in the transition risk, with individuals most likely to transition in the first 1–2 years.

While both panels exhibit fluctuations in initial durations, the likelihood of remaining disabled (Figure 4, panel b) increases slightly faster than the probability of remaining disability-free (panel a). For instance, in the T-SMP, the probability of staying disabled at duration 5 is significantly higher than the probability at duration 0. Additionally, the transition probability at the truncation duration surpasses the MSLT estimate, suggesting higher stickiness for long-term states but greater fluidity for short-term states compared with the Markov MSLT model. In the T-DDMSLT-H model, which records past periods of disability, duration 4 (or 6) in panel a represents a special state: disability-free without a history of disability. By separating individuals

Table 2 Partial healthy life expectancy (number of years) of cohort 1936–1945 aged 65–80, by sex

		Markov (MSLT)			Semi-Markov (T-DDMSLT/T-SMP)			Semi-Markov With History (T-DDMSLT-H)		
Sex	Model/Data	Healthy (1)	Unhealthy (2)	Total (3)	Healthy (4)	Unhealthy (5)	Total (6)	Healthy (7)	Unhealthy (8)	Total (9)
Male	Full data	11.5	1.8	13.3	—	—	—	—	—	—
		(11.3, 11.7)	(1.6, 1.9)	(13.0, 13.4)						
	Truncated 3	11.5	1.8	13.3	11.4	1.8	13.2	11.5	1.8	13.3
		(11.2, 11.7)	(1.6, 1.9)	(13.1, 13.5)	(11.2, 11.7)	(1.7, 1.9)	(13.0, 13.4)	(11.3, 11.7)	(1.6, 1.9)	(13.1, 13.5)
	Truncated 5	11.3	1.9	13.2	11.3	1.9	13.2	11.3	1.9	13.2
		(11.0, 11.5)	(1.7, 2.0)	(13.0, 13.4)	(11.0, 11.6)	(1.7, 2.0)	(13.0, 13.4)	(11.1, 11.6)	(1.7, 2.0)	(13.0, 13.4)
	No unknown	7.3	4.0	11.3	7.2	4.0	11.2	—	—	—
		(6.8, 7.8)	(3.7, 4.4)	(10.9, 11.9)	(6.6, 7.7)	(3.7, 4.4)	(10.6, 11.7)			
Female	Full data	11.6	2.3	13.9	—	—	—	—	—	—
		(11.4, 11.8)	(2.2, 2.5)	(13.8, 14.1)						
	Truncated 3	11.5	2.4	13.9	11.5	2.4	13.9	11.6	2.3	13.9
		(11.4, 11.7)	(2.2, 2.5)	(13.8, 14.1)	(11.3, 11.7)	(2.2, 2.5)	(13.7, 14.1)	(11.4, 11.8)	(2.2, 2.5)	(13.8, 14.1)
	Truncated 5	11.3	2.5	13.8	11.4	2.4	13.8	11.4	2.4	13.8
		(11.1, 11.6)	(2.3, 2.6)	(13.6, 14.0)	(11.1, 11.6)	(2.3, 2.6)	(13.6, 14.0)	(11.2, 11.6)	(2.3, 2.6)	(13.6, 14.0)
	No unknown	7.5	4.8	12.3	7.3	4.8	12.1	—	—	—
		(7.0, 7.9)	(4.5, 5.2)	(11.9, 12.6)	(6.9, 7.8)	(4.5, 5.2)	(11.7, 12.5)			

Notes: 95% confidence intervals are shown in parentheses. MLST = multistate life table. T-DDMLST = truncated duration-dependent multistate life table. T-SMP = truncated semi-Markov process. H = with history of unhealthy event.

Source: See Table 1.

with and without any history, the probability of remaining disability-free for duration 3 (or 5) and above is much lower than in the T-DDMSLT model, indicating that individuals with a history are likely frailer and more prone to transitioning out of the disability-free state. Therefore, separating groups with and without a history of disability could lead to more robust and realistic estimations.

Applying transition probabilities to the different models yields estimates of healthy and unhealthy life expectancies (HLE and ULE, respectively). Table 2 shows the 15-year partial life expectancy from 65 to 80 by sex, while Table A2 in the online appendix presents results for older cohorts aged 75–90. Columns display partial life expectancy estimates using the different models and performed on each of the datasets required to run the Markov and duration-dependent models. The methods are comparable within each column, so the differences in life expectancy across rows derive from the data samples. Thus, differences in results using the Markov MSLT (columns 1–3) relative to the first row of results measure errors introduced by discarding the observations necessary to run the duration-dependent models. Taking the males in Table 2 as an example, dropping all left-censored observations (required for T-SMP) significantly affects the Markov MSLT estimates: HLE decreases by 37% (from 11.5 to 7.3 years), ULE more than doubles (from 1.8 to 4.0 years), and the proportion of disability-free life declines from 87% to 65%. The T-SMP model (columns 4–6) estimates remain nearly identical to the Markov estimates (columns 1–3) on the reduced dataset, so do not compensate for the data loss.

The samples for the proposed truncated models substantially reduce the difference in life expectancy compared with the T-SMP model. Markov estimates of life expectancy are nearly identical between those performed on the full data and those on Truncated 3 (columns 1–3). Truncated 5 produces slightly smaller estimates of HLE (11.3 vs. 11.5 years) and higher estimates of ULE (1.9 vs. 1.8 years) for males in Table 2. These gaps between Truncated 5 and full data are slightly more noticeable when estimating HLE in older cohorts (see section 2 of the online appendix 2). It is important to acknowledge that they are likely a reflection of a real bias, introduced by discarding a disproportionately larger number of survey respondents who were healthy on entering the study.

The semi-Markov models on the same truncated datasets produce near identical results as the Markov model in columns 1–3, indicating that model choice is less of an issue and that the effects of duration on the transition probabilities average out over duration when measuring expectancies.

Discussion and Conclusion

Observable differences in multistate transition probabilities arise when duration dependence is considered. In our case, the probabilities of remaining in the same healthy or unhealthy state increase over duration. These trends are consistent with findings from previous studies, including those by Cai et al. (2006) and Cai et al. (2008), and warrant the past and continued scholarly attention to how duration dependence can be understood and measured. Duration itself may not be the underlying cause of these patterns, though accounting for duration may still help control for unobserved drivers. Regardless, duration dependence did not translate into differences in HLE in this study. Importantly, the features of the study and sample have allowed Markov probabilities to average out the effects of duration dependence. The differences we observe are primarily driven by the treatment of left-censored observations. Given the cost of duration-dependent models in terms of data loss, the Markov MSLT is perhaps the best choice for estimating life expectancies in this study, even in the presence of duration dependence. Yet, duration-dependent models may produce more realistic results if the research focus is on outputs other than life expectancy, such as the number and patterns of transitions.

Whether duration dependence impacts life expectancy in other contexts may depend on the circumstances. In this study, the MSLT averages out the impacts of duration dependence, resulting in comparable life expectancy estimates with and without explicitly modeling duration dependence. However, duration dependence could have a greater bearing when specific subpopulations are studied and compared, as well as in research contexts in which duration dependence is a more important influence, including, for example, contraceptive use and marital status (Schoen 2021; Steele et al. 2004). Future research should therefore continue to explore the presence and influence of duration dependence on the measurement of health and societal outcomes.

The proposed truncated duration-dependent model is a pragmatic option for exploring and accounting for duration dependence. Discarding all left-censored observations heavily biases samples away from relatively healthy populations, as seen in our example. The Cai et al. (2006) imputation approach for unknown durations is a methodological advance but may be impacted by systematic differences between left- and non-left-censored observations. Such differences place doubt on the implicit assumption that left-censored observation are missing at random (Dempster et al. 1977). While not claiming superiority over their imputation approach, our truncated DDMSLT model is viable, easy to implement, and a useful addition to the multistate toolkit. Future research could explore combining truncation and imputation approaches by imputing observations below the truncation point rather than dropping them, preventing data loss while strengthening the missing at random assumption validity by reducing differences between observations with and without missing durations.

A key limitation of the proposed model is the assumption that duration becomes irrelevant after a specific truncation point. These points represent a crucial trade-off between data loss and reimposing the Markov assumption. Higher truncations risk increased data loss and bias, while lower points nullify duration dependence modeling. Although we have not specified criteria, data-driven approaches for selecting truncation points are possible but require extended observation periods to identify if and when duration dependence disappears. A further limitation is that our data and approach incorporate no historical health information prior to the current state and potentially fail to fully capture health episodes and trajectories between survey waves. Arguably, these issues are intrinsic to the complexity of individual health trajectories and the limitations of data-gathering processes (Wolf and Gill 2009). Future research could nevertheless explore duration-dependent models that allow for transitions between intervals, such as the embedded Markov chain model proposed by Laditka and Wolf (1998).

In conclusion, we find duration dependence in the different models tested. However, in estimating healthy life expectancy in this study, any bias induced by not considering duration dependence is not so serious when the sample for the estimation is the same. As discussed by Guo (1993), the left-censored data are practically an intractable issue. There is no gold standard model without specific assumptions and limitations. The standard MSLT can produce comparable estimates to the different duration-dependent models and at the least cost in terms of data loss. Where that is not the case, the truncated DDMSLT proposed in this study is a viable, pragmatic, and parsimonious compromise that explicitly allows researchers to find the appropriate trade-off between data loss and duration dependence. ■

Acknowledgments The authors thank Collin F. Payne and Christian Dudel for their valuable comments. The views expressed herein are those of the authors and all errors are our own. This work was supported by the Australian National University Futures Scheme at the School of Demography, Australian National University. J.O. also acknowledges support from an Australian Research Council Discovery Early Career Researcher Award (DE240100232).

References

Allison, P. D. (1984). *Quantitative applications in the social sciences: Vol. 46. Event history analysis: Regression for longitudinal event data*. Newbury Park, CA: SAGE Publications. https://doi.org/10.4135/9781412984195

Bardenheier, B. H., Lin, J., Zhuo, X., Ali, M. K., Thompson, T. J., Cheng, Y. J., & Gregg, E. W. (2016). Compression of disability between two birth cohorts of U.S. adults with diabetes, 1992–2012: A prospective longitudinal analysis. *Lancet Diabetes & Endocrinology, 4,* 686–694.

Belanger, A. (1989). Multistate life table with duration dependence: An application to Hungarian female marital history. *European Journal of Population / Revue Européenne de Démographie, 5,* 347–372.

Brown, R., Diaz-Ramirez, L. G., Boscardin, W. J., Cappola, A. R., Lee, S. J., & Steinman, M. A. (2021). Changes in the hierarchy of functional impairment from middle age to older age. *Journals of Gerontology, Series A: Biological Sciences and Medical Sciences, 77,* 1577–1584.

Cai, L., Hayward, M. D., Saito, Y., Lubitz, J., Hagedorn, A., & Crimmins, E. (2010). Estimation of multistate life table functions and their variability from complex survey data using the SPACE program. *Demographic Research, 22,* 129–158. https://doi.org/10.4054/demres.2010.22.6

Cai, L., Schenker, N., & Lubitz, J. (2006). Analysis of functional status transitions by using a semi-Markov process model in the presence of left-censored spells. *Journal of the Royal Statistical Society, Series C: Applied Statistics, 55,* 477–491.

Cai, L., Schenker, N., Lubitz, J., Diehr, P., Arnold, A., & Fried, L. P. (2008). Evaluation of a method for fitting a semi-Markov process model in the presence of left-censored spells using the Cardiovascular Health Study. *Statistics in Medicine, 27,* 5509–5524.

Cha, H., Farina, M. P., & Hayward, M. D. (2021). Socioeconomic status across the life course and dementia-status life expectancy among older Americans. *SSM–Population Health, 15,* 100921. https://doi.org/10.1016/j.ssmph.2021.100921

Crimmins, E. M., Thyagarajan, B., Kim, J. K., Weir, D., & Faul, J. (2021). Quest for a summary measure of biological age: The Health and Retirement Study. *GeroScience, 43,* 395–408.

Crowther, M. J., & Lambert, P. C. (2014). A general framework for parametric survival analysis. *Statistics in Medicine, 33,* 5280–5297.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B: Methodological, 37,* 1–22.

Freedman, V. A., Martin, L. G., & Schoeni, R. F. (2002). Recent trends in disability and functioning among older adults in the United States. *JAMA, 288,* 3137–3146.

Guo, G. (1993). Event-history analysis for left-truncated data. *Sociological Methodology, 23,* 217–243.

Hayflick, L. (2007). Biological aging is no longer an unsolved problem. *Annals of the New York Academy of Sciences, 1100,* 1–13.

Health and Retirement Study. (2023). *RAND HRS Longitudinal File 2020* (V1) [Public use dataset]. Ann Arbor: University of Michigan. Available from https://www.rand.org/well-being/social-and-behavioral-policy/portfolios/aging-longevity/dataprod.html

Hennessey, J. C. (1980). An age dependent, absorbing semi-Markov model of work histories of the disabled. *Mathematical Biosciences, 51,* 283–304.

Hosokawa, R., Ojima, T., Myojin, T., Kondo, K., & Kondo, N. (2023). Geriatric symptoms associated with healthy life expectancy in older people in Japan. *Environmental Health and Preventive Medicine, 28,* 44. https://doi.org/10.1265/ehpm.22-00300

Jia, H., & Lubetkin, E. I. (2020). Life expectancy and active life expectancy by marital status among older U.S. adults: Results from the U.S. Medicare Health Outcome Survey (HOS). *SSM–Population Health, 12,* 100642. https://doi.org/10.1016/j.ssmph.2020.100642

Laditka, S. B., & Wolf, D. A. (1998). New methods for analyzing active life expectancy. *Journal of Aging and Health, 10,* 214–241.

Lang, K. M., & Little, T. D. (2018). Principled missing data treatments. *Prevention Science, 19,* 284–294.

Levine, M. E. (2013). Modeling the rate of senescence: Can estimated biological age predict mortality more accurately than chronological age? *Journals of Gerontology, Series A: Biological Sciences and Medical Sciences, 68,* 667–674.

Lièvre, A., Brouard, N., & Heathcote, C. (2003). The estimation of health expectancies from cross-longitudinal surveys. *Mathematical Population Studies, 10,* 211–248.

Littman, G. S., & Mode, C. J. (1977). A non-Markovian stochastic model for the Taichung medical IUD experiment. *Mathematical Biosciences, 34,* 279–302.

Maddox, G. L., Clark, D. O., & Steinhauser, K. (1994). Dynamics of functional impairment in late adulthood. *Social Science & Medicine, 38,* 925–936.

Neumann, J. T., Thao, L. T. P, Callander, E., Carr, P. R., Qaderi, V., Nelson, M. R., . . . McNeil, J. J. (2022). A multistate model of health transitions in older people: A secondary analysis of ASPREE clinical trial data. *Lancet Healthy Longevity, 3,* e89–e97. https://doi.org/10.1016/S2666-7568(21)00308-1

O'Donnell, J. (2021). Does social housing reduce homelessness? A multistate analysis of housing and homelessness pathways. *Housing Studies, 36,* 1702–1728.

Payne, C. F. (2022). Expansion, compression, neither, both? Divergent patterns in healthy, disability-free, and morbidity-free life expectancy across U.S. birth cohorts, 1998–2016. *Demography, 59,* 949–973. https://doi.org/10.1215/00703370-9938662

Payne, C. F., & Kobayashi, L. C. (2022). Changes in life expectancy and disability-free life expectancy in successive birth cohorts of older cancer survivors: A longitudinal modeling analysis of the U.S. Health and Retirement Study. *American Journal of Epidemiology, 191,* 104–114.

Payne, C. F., Mkandawire, J., & Kohler, H.-P. (2013). Disability transitions and health expectancies among adults 45 years and older in Malawi: A cohort-based model. *PLoS Medicine, 10,* e1001435. https://doi.org/10.1371/journal.pmed.1001435

R Core Team. (2023). *R: A language and environment for statistical computing* (Version 4.3.0) [Computer software]. Available from http://www.r-project.org/

Raymo, J. M., Pike, I., & Liang, J. (2019). A new look at the living arrangements of older Americans using multistate life tables. *Journals of Gerontology, Series B: Psychological Sciences and Social Sciences, 74,* e84–e96. https://doi.org/10.1093/geronb/gby099

Rockwood, K., & Mitnitski, A. (2007). Frailty in relation to the accumulation of deficits. *Journals of Gerontology, Series A: Biological Sciences and Medical Sciences, 62,* 722–727.

Schoen, R. (1988). *Modeling multigroup populations.* New York, NY: Springer Science+Business Media. https://doi.org/10.1007/978-1-4899-2055-3

Schoen, R. (2021). Recognizing duration effects in multistate population models. *Genus, 77,* 32. https://doi.org/10.1186/s41118-021-00120-y

Shen, T., & Payne, C. F. (2023). Disability and morbidity among U.S. birth cohorts, 1998–2018: A multidimensional test of dynamic equilibrium theory. *SSM–Population Health, 24,* 101528. https://doi.org/10.1016/j.ssmph.2023.101528

Steele, F., Goldstein, H., & Browne, W. (2004). A general multilevel multistate competing risks model for event history data, with an application to a study of contraceptive use dynamics. *Statistical Modelling, 4,* 145–159.

van den Berg, G. J., & van Ours, J. C. (1996). Unemployment dynamics and duration dependence. *Journal of Labor Economics, 14,* 100–125.

Wolf, D. A. (1988). The multistate life table with duration-dependence. *Mathematical Population Studies, 1,* 217–245.

Wolf, D. A., & Gill, T. M. (2009). Modeling transition rates using panel current-status data: How serious is the bias? *Demography, 46,* 371–386.

Xu, K. Q., & Payne, C. F. (2024). A growing divide: Trends in social inequalities in healthy longevity in Australia, 2001–20. *Population Studies, 78,* 231–250.

Zaniotto, P., Head, J., & Steptoe, A. (2020). Behavioural risk factors and health life expectancy: Evidence from two longitudinal studies of ageing in England and the U.S. *Scientific Reports, 10,* 6955. https://doi.org/10.1038/s41598-020-63843-6

Tianyu Shen (corresponding author)
Tianyu.Shen@anu.edu.au

Shen • School of Demography, Research School of Social Sciences, College of Arts and Social Sciences, Australian National University, Acton, Australian Capital Territory, Australia; https://orcid.org/0000-0002-9954-1709

O'Donnell • School of Demography, Research School of Social Sciences, College of Arts and Social Sciences, Australian National University, Acton, Australian Capital Territory, Australia; https://orcid.org/0000-0002-3093-4141