# Forecasting Health Expectancy for Old Adults in the United States Using Cohort Smoking and Obesity History

**Bochen Cao**

**University of Pennsylvania**

## Introduction

Despite the continuing increase in life expectancy in the United States, whether the increase of life expectancy reflects a proportionally increase in life years spent without disability is still in debate (Nusselder 2003). Therefore, health expectancy (HE) becomes a good alternative measure to summarize overall population health using combined mortality and morbidity information (Robine et al. 2003). It measures the number of years that an individual at a certain age is expected to live without disability. The most common forms used for measuring health expectancy are disability-free life expectancy (DFLE) and life expectancy with disability (LwD).

Both mortality and morbidity can be affected by personal health behaviors. In the U.S., smoking and obesity in particular are the two most destructive factors to the health of population. Both behaviors are believed to be causes of a variety of chronic diseases . Smoking is found to be responsible for shortening both DFLE and LwD (Nusselder 2000; Reuser et al. 2009). Obesity, on the other hand, is associated with higher risk of disabling but non fatal conditions, such as osteoarthritis of weight-bearing joints and chronic back pain. Consequently, the extended life years spent with obesity-related disability may raise more concerns, especially for women who often experience increased survival and increased disability at the same time (Must et al. 1999).

The past trends have shown opposite directions for smoking and obesity. The prevalence of obesity has increased, and so has the mortality risk associated with obesity. In contrast, the prevalence of smoking has declined, and so has the mortality risk associated with smoking. In the next decades, the prevalence of obesity for both sexes is expected to increase and smoking is expected to continue declining (Preston et al. 2014).

This present study aims to produce forecasts of DFLE for the U.S. population above age 55, using a multistate life table approach (MSLT) based on the method proposed by Majer et al. (2013) to estimate transition rates from observed data and a modified Lee-Carter model to forecast future transition rates. Additionally, I include variables for cohort history of smoking and obesity to capture the trends for these two major risk factors to which a considerable proportion of deaths in the U.S. is attributable, and in order to avoid implausible long-term forecasts that may be produced by the Lee-Carter model (Wang and Preston 2009).

## Data and Methods

The population in this study is restricted to US elderly age 55 years and above. Population-level mortality data is drawn from Human Mortality Database (HMD). I use single-year of age-sex-specific mortality rates from age 55 to 85+ for every calendar year from 1970 to 2010. Individual-level data for disability, smoking and obesity are obtained from the National Health Interview Survey (NHIS), to calculate prevalence of disability, and cohort smoking and obesity history which are measured by the average number of years that members from a birth cohort smoked prior to age 40 and cohort prevalence of obesity (BMI>=30) at age 40.

**Estimating the transition rates** The age-period-specific transition rates among different states (healthy, disabled and dead) are estimated first. Due to unavailability of separate mortality rates for disabled and non-disabled population in the data sources, I calculate these rates using total population mortality, prevalence of disability and risk of disability on mortality, given the mortality rate for healthy and disabled can be expressed separately as:

$$m_{x,t}^{HD,g} = \frac{m_{x,t}^{g}}{HR_{x,t}^{g} \times p_{x,t}^{U,g} + (1 - p_{x,t}^{U,g})}$$

$$m_{x,t}^{UD,g} = m_{x,t}^{HD,g} HR_{x,t}^{g}$$

where $m_{x,t}^{g}$ , $m_{x,t}^{HD,g}$ , $m_{x,t}^{UD,g}$ , $p_{x,t}^{U,g}$ and $HR_{x,t}^{g}$ indicate the gender-specific (indicated by *g*) population mortality rates, estimated mortality rates of being healthy and disabled, prevalence of disability and hazard ratio of disability on mortality at age *x* and time *t*. The population mortality rates are extracted from HMD. The prevalence of disability is calculated from NHIS from 1970 to 2010 when respondents were asked whether they have any limitations of activity. Due to the changes in question wording across surveys, the variable that indicates disability is reconstructed as a dummy variable based on whether one reported any limitation in activities of daily living (ADLs) or limitation in instrumental activities of daily living (IADLs). The prevalence of disability is then estimated using NHIS sampling weights. The NHIS is linked to National Death Index (NDI) for years from 1986 to 2004. Using this record-linked survival dataset, the relative risk of being disability on mortality, $HR_{x,t}^{g}$ , can be estimated with Cox regression models, stratified by survey years. I found no age or year interactions, therefore constant hazard ratios of 2.05 and 1.95 are used for all age groups over time for men and women respectively.

Assuming all transitions occur in the middle of the age interval, transition probabilities can be calculated based on the values of transition rates. For simplicity, I apply the method of Majer et al. (2013) and assume there is no recovery and hence transition rate from disabled to healthy $m_{x,t}^{UH,g}$ is 0. Therefore, the transition rate from healthy to disabled , $m_{x,t}^{HU,g}$ , can be considered as the net incidence rate, which corresponds to the number of transitions from healthy to disabled minus the number of transitions from disabled to healthy, divided by the person-years lived in healthy state. Then given the prevalence of disabled $p_{x,t}^{U,g}$ of age *x* at time *t* and $p_{x+1,t+1}^{U,g}$ of age *x+1* at time *t+1*, the mortality rates of non-disabled $m_{x,t}^{HD,g}$ and disabled $m_{x,t}^{UD,g}$, the corresponding transition probability from health to disabled can be calculated as these quantities are interrelated and can define each other mutually, using the following relationship (Majer et al. 2013; Guillot and Yu 2009)

$$p_{x+1,t+1}^{U,g} = \frac{p_{x,t}^{U,g} - p_{x,t}^{U,g} q_{x,t}^{UD,g} + (1 - p_{x,t}^{U,g}) q_{x.t}^{HU,g}}{1 - (1 - p_{x,t}^{U,g}) q_{x,t}^{HD,g} - p_{x,t}^{U,g} q_{x,t}^{UD,g}}$$

where $q_{x,t}^{HD,g}$, $q_{x,t}^{UD,g}$ and $q_{x.t}^{HU,g}$ are the probability of death of healthy and disabled, and transition probability from health to disabled, respectively, as $q_{x.t}^{HU,g}$ is the only unknown here.

**Forecasting the transition rates** I use Lee-Carter model that incorporates cohort smoking and obesity history to fit and forecast the transition probabilities. Given the two leading risk factors, smoking and obesity, of mortality are adjusted for, a common temporal trend in both mortality and morbidity is assumed for males and females (Wang and Preston 2009). The model can be expressed as:

$$\ln m_{x,t}^{g,i} = \alpha_x^{g,i} + \beta_x^{g,i} \kappa_t^i + \theta_x^{g,i} S_{x,t}^g + \lambda_x^{g,i} O_{x,t}^g + \varepsilon_{x,t}^{g,i}$$

where *i* specifies the three types of transition: healthy to unhealthy/disabled (HU), healthy to death (HD), unhealthy/disabled to death (UD). The parameter $\alpha$ is the average of the log transition rate at age *x* over time, $\kappa$ is the time-dependent latent process that quantifies the development of transition rates over time and is assumed to be the same for males and females when smoking and obesity are adjusted. $\beta$ is the changes in transition rates at age *x* in response to changes in $\kappa$ over time. *S* and *O* are variables for cohort history of smoking and obesity, and $\theta$ and $\lambda$ are corresponding coefficients that measure the effect of smoking and obesity on the specific transition rates respectively. The cohort smoking history is measured by the average number of years that member from a birth cohort smoked prior to age 40. The cohort obesity history is measured by the prevalence of obesity (BMI>=30) at age 40 for a birth cohort. Age 40 is chosen because the mortality is relatively low at age 40, thus minimizing the problem of reverse causation. Birth cohorts are constructed by 5-year groups.

For forecasting purpose, random walk model with drift is used to produce future values of *k* for year 2011 to 2040. I use cohort-specific rather than period-specific measures of smoking and obesity, as making projections using the latter will require extrapolating the measures of smoking and obesity at all ages for all future periods, while only few cohort-specific measures need to be extrapolated since most cohort behaviors have already been observed for the rather old population I study here. Specifically, people from the cohort born in 1970 reach age 40 by 2010, the last year for which data are available, and age 55 by 2025, while people from the cohort born in 1985 reach age 55 by 2040, the last year of the forecasting period. Consequently, smoking and obesity variable need to be extrapolate for only three cohorts (1975, 1980 and 1985). The extrapolation is performed by regressing the values of the smoking and obesity variables at age 40 on their corresponding values at age 35, age 30 and age 25. Regressions are estimated on data for the cohorts for which we have complete data up to age 40. Then the future values of $k$, as well as corresponding cohort smoking and obesity variables are used to estimate the future transition rates from 2011 to 2040, which are eventually translated into disability-free life expectancy (DFLE) and life expectancy with disability (LwD).

## Preliminary Results

Figure 1 shows the cohort trends of smoking and obesity. While the average number of years a cohort smoked before age 40 declines, the obesity prevalence at age 40 continues to increase for cohorts born after 1930 for both sexes. This implies that mortality and morbidity that are attributable to smoking will decline, but morbidity attributable to obesity will likely to increase. The results shown in Table 1 confirm this hypothesis. The improvement of life expectancy at age 55 for men accelerates after 2010, as a result of the extinction of the heaviest smoking cohorts. Although men are expected to spend more years disable-free, reflecting the increases in total life expectancy, the proportion of DFLE relative to total life expectancy will decline after 2020, as the relatively more obese cohorts reach their old ages. In contrast, because of women's delayed smoking pattern and the heaviest smoking cohorts start to replace the lighter ones after 2015, the rate of increase in life expectancy for women shrinks. Accordingly, women are expected to experience virtually no gain in life expectancy until around 2030, when their gain in life expectancy starts to accelerate again, as the heaviest smoking cohorts disappear. Due to the relatively higher mortality attributable to smoking, many women may not have survived long enough to become disabled; the increase in DFLE is likely a result of disproportionally deaths of severely sick people. Therefore not only the absolute increase in DFLE is slow, but the proportion of DFLE relative to total LE also tend to increase until 2030.

## Discussion and Future Direction

The results from this study suggest that increase in both DFLE and LwD can be expected in the next decades for both sexes. However, due to different patterns of smoking history and the consequently different mortality, unhealthy men and women tend to have different distributions conditional on their smoking history. Therefore, over the next decades, men are expected to gain DFLE in absolute term, but lose in relative term; while women are likely to gain both absolute and relative DFLE only slowly before the proportion of DFLE starts to decline. The proportionally decline in DFLE is likely a consequence of increasing obesity. The next step would be decomposing the changes in DFLE to those attributable to smoking and those attributable to obesity. Also, the MSLT method used here could be modified to include transition from disabled to healthy.

**Figure 1. Smoking and Obesity Trends by Cohort**

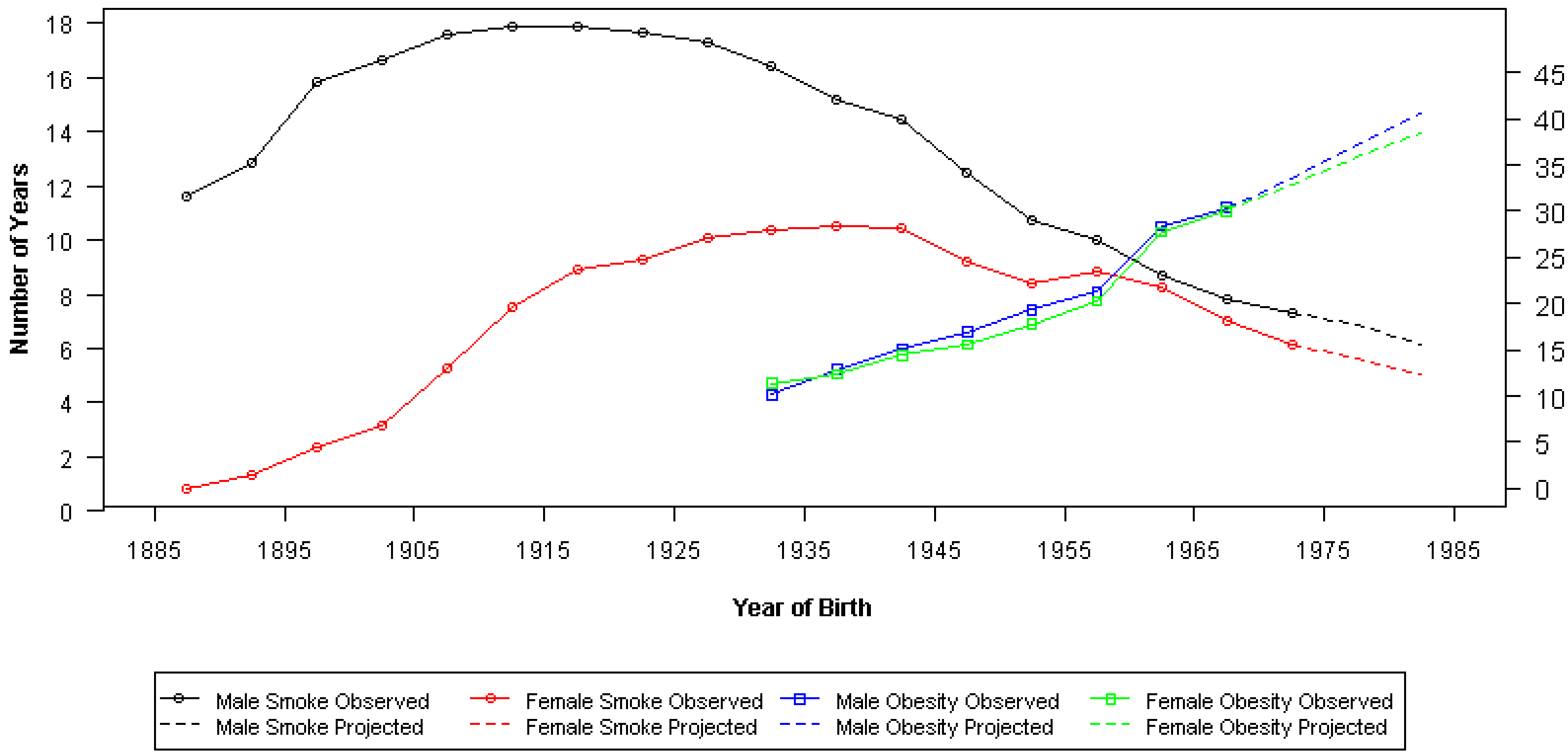

**Table 1: Life Expectancy at Age 55 by Disable Status**

| Year | Males | | | | Females | | | |
|---|---|---|---|---|---|---|---|---|
| | Life Expectancy at 55 | DFLE at 55 | LwD at 55 | DFLE/LE (%) | Life Expectancy at 55 | DFLE at 55 | LwD at 55 | DFLE/LE (%) |
| 1970 | 19.44 | 13.06 | 6.38 | 67.19 | 24.92 | 17.65 | 7.27 | 70.81 |
| 1980 | 21.02 | 13.25 | 7.77 | 63.04 | 26.67 | 17.88 | 8.79 | 67.05 |
| 1990 | 22.42 | 16.47 | 5.95 | 73.47 | 27.54 | 19.34 | 8.20 | 70.21 |
| 2000 | 23.82 | 18.26 | 5.56 | 76.66 | 27.60 | 20.20 | 7.40 | 73.19 |
| 2010 | 25.99 | 20.17 | 5.82 | 77.60 | 29.54 | 21.68 | 7.86 | 73.38 |
| 2020 | 26.58 | 20.77 | 5.81 | 78.15 | 29.67 | 21.85 | 7.82 | 73.65 |
| 2030 | 27.41 | 21.11 | 6.30 | 77.02 | 29.72 | 22.15 | 7.57 | 74.53 |
| 2040 | 28.16 | 21.46 | 6.70 | 76.19 | 30.68 | 22.48 | 8.20 | 73.26 |

## References


Brown, D. C., Hayward, M., Montez, J. K., Hummer, R. A., Chiu, C.-T., Hidajat, M. M. (2012). The significance of education for mortality compression in the United States. *Demography, 49,* 819–840.

Burns D, et al.*(1998) in Smoking and Tobacco Control Monograph No. 8, eds Burns D, Garfinkel L, Samet J (Natl Inst Health, Bethesda), pp 13–112.*

Cheung, S. L., & Robine, J. M. (2007). Increase in common longevity and the compression of mortality: The case of Japan. *Population Studies, 61,* 85–97.

Edwards, R.D.and S. Tuljapurkar. 2005. "Inequality in life spans and a new perspective on mortality convergence across industrialized countries." *Population and Development Review* 31(4):645-674.

Fries, J.F. 1980. "Aging, natural death, and the compression of morbidity." *N Engl J Med* 303(3):130-135.

Guillot, M., Yu, Y. Estimating health expectancies from two cross-sectional surveys: The intercensal method. Demographic Research 2009;21: 503-534.

Lee, R. D., & Carter, L. R. (1992). Modeling and forecasting U.S. mortality. Journal of the American Statistical Association, 87, 659–671.

MAJER, I., R. STEVENS, W. NUSSELDER, J. MACKENBACH, AND P. BAAL (2013): “Modeling and Forecasting Health Expectancy: Theoretical Framework and Application,” Demography, pp. 1–25.

Manton, K.G. 1982. "Changing concepts of morbidity and mortality in the elderly population." *Milbank Mem Fund Q Health Soc* 60(2):183-244.

Manton, K. G., Gu, X., & Lamb, V. L. (2006). Long-term trends in life expectancy and active life expectancy in the United States. Population and Development Review, 32, 81–105.

Mokdad, A. H., J. S. Marks, D. F. Stroup and J. L. Gerberding. 2004. "Actual Causes of Death in the United States, 2000." *Journal of the American Medical Association* 291:1238-45.

Mokdad, A. H., J. S. Marks, D. F. Stroup and J. L. Gerberding, 2005. “Correction: Actual Causes of Death in the United States”. *JAMA* 293(3):293-4.

Must A, Spadano J, Coakley EH *et al*. The disease burden associated with overweight and obesity. *JAMA* 1999;282:1523–1529.

Nusselder, W.J. 2003. *Compression of Morbidity*: John Wiley & Sons, Ltd.

Nusselder WJ, Looman CW, Marang-van de Mheen PJ, van de, Mheen H, Mackenbach JP. Smoking and the compression of morbidity. *J Epidemiol Community Health* 2000;54:566–574.

OECD. 2009. *Pensions at a Glance 2009: Retirement-Income systems in OECD countries*.Paris: OECD.

Olshansky, S.J., M.A. Rudberg, B.A. Carnes, C.K. Cassel, and J.A. Brody. 1991. "Trading Off Longer Life for Worsening Health." Pp. 194-216.

Preston, Samuel, Andrew Stokes, Neil Mehta and Bochen Cao. (2014). "Projecting the effect of changes in smoking and obesity on future life expectancy in the United States." Demography Pp.27-49

Robine, J.M., C. Jagger, C.D. Mathers, E.M. Crimmins, and R.M. Suzman. 2003. *Determining health expectancies*. Chichester, England: Wiley.

Reuser M, Bonneux LG, Willekens FJ. Smoking kills, obesity disables: a multistate approach of the US Health and Retirement Survey. Obesity 2009; 17: 783–789.

Sullivan, D.F. 1971. "A single index of mortality and morbidity." *HSMHA Health Rep* 86(4):347-354.

Wang, Haidong and Samuel Preston. 2009. “Forecasting United States Mortality Using Cohort Smoking Histories” *Proceedings of the National Academy of Sciences* Vol. 106(2): 393-98.