Constructing Life Tables from the Kaiser Permanente Smoking Study and Applying the Results to the Population of the United States

David Swanson

Center for Studies in Demography and Ecology

University of Washington, Seattle, WA 98195 USA

Email: dswanson@ucr.edu

Abstract

Following the path laid out in Abelin's seminal 1965 article, I construct life tables from cohort mortality data widely employed in efforts to examine smoking and health, which in this case is the Kaiser Permanente Smoking Study. The mortality data in this study have been used in terms of relative mortality and risk rates in regard to smoking behaviors. However, they have never been used to generate life tables. After describing life tables in general, I describe the KP smoking study data, then discuss the methods used to generate the life tables from them. Following these descriptions and discussion, I show the life tables developed from the KP smoking study. I then discuss the methods used to extend these life tables to the US population and create hazard rate and survivorship data that can be used as input to models designed to assess the population health impact of tobacco products.

Introduction

Abelin (1965) used a set of studies on smoking and health that relied on mortality rates and ratios as a jumping off point to illustrate the advantages of using survivorship and mortality measures based on life tables.[1] Specifically, Albelin (1965) converted age-specific mortality rates (deaths divided by number of person-years of exposure by age groups and smoking categories) found in three studies, Dorn (1959), Hammond (1964), and Hammond and Horn (1958) into the probabilities of dying using a standard actuarial formula:

$$ {}_nq_x = [({}_nm_x)(n)]/[1 + (0.5)(n)({}_nm_x)] \qquad [1] $$

where x = the beginning of an age group

n = the width of the age group in question

m = deaths per person-year

and q = the probability of dying between age x and age x+n

If n =1, then equation [1] becomes

$$ q_x = (m_x)/[1 + (0.5)(m_x)] \qquad [2] $$

As an example of using this formula to convert an age-specific mortality rate to the probability that those alive at the start of the age group (x) will die before reaching the end of it (n, where n =5 in this paper), suppose that deaths per-person year for those aged 60-64 is 0.015. Using equation [1] I find that the for those who reach their 60th birthday, the probability of dying before reaching the 65th birthday is 0.0723, where 0.0723 = (0.015*5)/[1 +

(.5*5*0.015)]. Once the probability of dying is estimated for a set of age groups, a complete life table can be generated (Keyfitz, 1970; Kintner, 2004), which provides a range of mortality measures, notably: (1) the probability of surviving from birthday to birthday as well as from age group to age group; and (2) expected years of life remaining at the start of any given age group. Life tables can be constructed for a wide range of variables – gender, race, ethnicity, and smoking status, for example.

Following in Abelin's (1965) pioneering footsteps, I construct life tables from cohort mortality data widely employed in efforts to examine smoking and health, namely the Kaiser Permanente Smoking Study (Friedman et al. 1997). The mortality data in this study (presented in tabular form) have been used in terms of relative mortality and risk rates as I shortly show. However, they have never been used to generate life tables.[2]

I believe that it is worthwhile to convert the mortality rates found in the KP Smoking Study data into life table data because the study has been fairly widely cited. Here are illustrative examples of how these data have been employed: (1) to examine the decline in excess risk of chronic obstructive pulmonary disease following quitting smoking (Lee, Fry, and Forey 2013); (2) in reviews of the effects of smoking on health (Ernster et al., 2000; Fry et al., 2013; Jacobs et al. 1999; Lee and Forey, 2013; Lee et al., 2012; Thun et al., 2000); (3) to assess the potential population health benefit of modified risk tobacco products (Bachand and Sulsky, 2013; Bachand et al. 2018, Muhammad-Kah et al., 2016); (4) as a basis for comparing study results (Bach et al. 2003; Bain et al., 2004; De Mattias et al., 2012;Blizzard and Dwyer, 2003; Thun et al., 2006); (5) as a reference point in a given study (Hunt et al., 2005; Lee et al., 2015; Levy et al., 2007; Poland and Teischinger, 2017 ; van Iersel et al., 2006)

Virtually all of these studies would have benefited from the use of mortality measures constructed from life tables, especially those aimed at assessing the population health benefit of modified risk tobacco products (Bachand and Sulsky, 2013; Bachand et al. 2018; Muhammad-Kah et al. 2016). I make this claim because a life table is, in essence, one form of combining mortality rates of a population at different ages into a single statistical model. A life table is specifically designed to measure the level of mortality of a population of interest (Kintner 2004). In the paper that follows, I first develop life tables from the KP smoking study mortality rates and then extend the life tables by first adjusting them to reflect the US population and then using models to create life tables for a much more comprehensive set of age groups than are found in the KP smoking study.

The remaining paper is organized as follows. After briefly discussing life tables in general, I describe the KP smoking study data, then discuss the method used to generate a set of preliminary life tables from them. Following these descriptions and discussion, I show the preliminary life tables developed from the KP smoking study. I then assess the preliminary life tables and identify inconsistencies. I then describe how I minimized these inconsistencies and go on to produce revised life tables that are displayed in terms of survivorship and life expectancy by age. I discuss the methods used to produce these revised life table and then discuss the fact that the KP population has higher life expectancy than its temporally equivalent US population. I then show how the KP life tables can be extended to their temporally equivalent and more recent US populations and include a discussion of the strengths and weaknesses of the KP life tables and the extension of them to US populations over time. The paper concludes with a discussion of the application of the paper's results to the assessment of the net population health.

Life Tables: General Comments and an Example of Research

Life tables represent an important component of demography. Not only do they serve as methodological and conceptual tools (Burch, 2018; Swanson and Tedrow 2012; and Yusuf, Martins and Swanson, 2014), but they support a wide range of both applied work (Abelin, 1965; Jiang et al., 2009; Kintner and Swanson, 1994; Owen et al., 2018; Preston, 1970a, 1970b; Retherford, 1972; Richards and Donaldson, 2010; Siegel, 2002; Smith, Tayman, and Swanson, 2013; Thomas and Bao, 2016); and academic work (Shkolnikov, Andreev and Begun, 2003; Swanson and Sanford, 2012; Trovato and Lalu, 2001; Villavicencio and Riffe, 2016; Wrycza, 2014; Yashin, Stallard and Land, 2016).

It should not be surprising that life tables have been widely studied and that many of their defining characteristics, such as their internal mathematical equalities, have been described (Kintner, 2004). As one of many possible examples, there is a line of research that has examined equalities and inequalities in the life table. In addition to the obvious equalities such as the crude birth rate being equal to the crude death rate, this research has revealed that: (1) mean years lived is equal to mean years remaining, which is known as Carey's Equality Theorem (Vaupel, 2009); and (2) the distribution of age composition is equal to the distribution of remaining lifetimes (Carey et al., 2008; Rao and Carey, 2014, Vaupel, 2009). To these equalities, the following can be added: (1) mean age is equal to mean years lived (Rao and Carey, 2014); and (2) mean age is equal to mean years remaining (Kim and Aron, 1989), which implies that mean years lived is equal to mean years remaining (Swanson and Tedrow, 2019). From this body of work, Swanson and Tedrow (2019) identified two hitherto unknown inequalities. The first is that at any given age, the sum of mean years lived and mean years remaining exceeds life expectancy at birth when age is greater than zero and less than the maximum lifespan. The second inequality applies to the entire population and shows that the sum of mean years lived and mean years remaining exceeds life expectancy at birth.

As an example of an application of this line of research, Vaupel (2009) used Carey's Equality Theorem (Carey et al., 2008, Rao and Carey, 2014) and a 2005 US life table to estimate that nearly half of the US population in 2009 will still be alive in 2050. Using the same US 2005 life table, Swanson and Tedrow (2018) found that on average the population lived 40.60 years and will live another 40.60 years on average, or 81.3 years in total, which is 3.67 years more than their life expectancy at birth of 77.63 years.

The KP Smoking Study Data

The mortality data from the KP smoking study I use as input for the initial life tables are found in tables 3, 4, 5, 6, 7, and 8 as shown by Friedman et al. (1997). By gender, these tables provide mortality data for all causes by selected age groups and cigarette smoking status, which are categorized as:

(1) Never

(2) Current

  (a) Quantity (< 20 cigarettes daily; 20+ daily

  (b) Duration smoking (< 20 years; 20-39 Years; 40+ years)

(3) Former

(a) Duration since quitting smoking (2-10 years, 11-20 years, 21+ years)

The KP smoking study provides overall mortality rates for all causes by selected racial groups, gender, and age in Table 1, but age groups are not provided, which precludes the construction of life tables from these data. In Table 2, mortality rates for all causes are provided by race, gender, and age group, but the small numbers preclude the construction of life tables from these data.

Methods

In constructing the preliminary life tables from the KP smoking study, I employ a different method than that employed by Abelin (1965) to convert age-specific mortality rates into the probabilities of dying using the formula shown as equation [1], which assumes that deaths occur evenly throughout a given age interval. Instead, I employ a conversion formula that assumes that deaths occur in increasing numbers within a given age interval, specifically in an exponential manner (Fergany, 1971):

$$ {}_nq_x = 1-e^{(-n*{}_nm_x)} \qquad [3] $$

where, as in equation [1],

x = the beginning of an age group

n = the width of the age group in question

m = deaths per person-year

q = the probability of dying between age x and age x + n and where $e \approx 2.71828$

As noted earlier, once the probability of dying is estimated for a set of age groups, then a complete life table can be generated (Kintner, 2004). I start this process with the number of Person-years lived during a given year by (lx) people alive at the start of that year, which is denoted by (L_x). This can be estimated as follows.

Let (l_x) be the number of survivors aged x at the beginning of a given age group, x. Of these (${}_nd_x$) will die before reaching age x+n , so that l_{x+n} would be alive, aged x + n at the beginning of the next age group x + n. Obviously, l_{x+n} lived for n whole years. To this, one must add the average period lived by the people who died aged ${}_nd_x$. To do this I divide the number of expected deaths, ${}_nd_x$ generated by equation [2] by the age-specific mortality rate, ${}_nm_x$
As defined in equation [1] (Ferangy, 1971; Kintner, 2004). Thus,

$${}_nL_x = {}_nd_x / {}_nm_x \qquad [4]$$

Because (T_x) is the total person-years lived beyond age x, and (l_x) is the number of persons alive at age x, the average number of person-years lived beyond age x (also referred to as the life expectancy at age x (e_x)) is calculated as $e_x = T_x/l_x$

Life tables not only reflect the mortality regimes of actual populations, they represent hypothetical populations that will evolve as a result of the following conditions affecting these same populations over long periods of time:

• the population will remain closed to migration

• every year it will be augmented by a constant number of births (l_0), which is equal to the number of deaths that deplete the population each year ($\sum d_x$)

• it will experience a constant schedule of mortality (${}_nq_x$) every year

• deaths will generally occur uniformly within each year.

Under these conditions, the characteristics of this “hypothetical” population (also known as a “complete” life table stationary population since its age structure and size remain constant over time) are

as follows:

- L_x is the number of persons at each age x
- T_0 is the total size of the population
- l_0 is the number of births (which is equal to the number of total deaths)
- d_x is the number of deaths at each age x
- l_0/T_0 = the crude birth rate = the crude death rate

where $T_0 = \sum L_x$,

$d_x = l_x q_x$

A complete life table uses single years from birth to the terminal age (beyond which no one lives). If the terminal age is 115, then the life table uses 0, 1, 2, 3, 4,…, 115.

An “abridged” life table uses age groups, where the width of the age group is provided by n. In this case, the characteristics of the hypothetical population are

- ${}_nL_x$ is the number of persons in the age group that is n years wide starting at age x,
- T_0 is the total size of the population
- l_0 is the number of births (which is equal to the number of total deaths)
- ${}_nd_x$ is the number of deaths in age group that is n years wide starting at age x
- l_0/T_0 = the crude birth rate = the crude death rate

Because I use Fergany's (1971) method to construct the preliminary life tables, there are several points that need discussion that are specific to this method. Fergany's (1971) method is advantageous because only the age-specific death rates are needed to construct an abridged life table. "In addition to its simplicity, it is, in contrast to other methods, self-contained in the sense that beyond making only the assumption of approximating the force of mortality by a step function (which is all we observe anyway) no further assumptions, approximations, or parameter estimates are required to compute all the life table functions." (Fergany 1971: 334). One disadvantage of this method in terms of the KP mortality data is that for the terminal open ended age group, where the hazard rate ($_{\omega}q_x$) is 1.00, an adjustment has to be made because the calculation of "Years lived ($_{\omega}L_x$) requires an age specific death rate for the terminal, open ended age group, which is not shown in the KP mortality data (Friedman et al. 1997).

With these conventions and characteristics in mind, I now turn to the construction of preliminary life tables from the KP Smoking Study. I describe the initial tables as preliminary because I expected – and encountered - some anomalies due to the small samples and conventions that characterize the KP Smoking Study data.

Constructing Preliminary Life Tables from the KP Smoking Study

Tables 1 through 12 show the preliminary life tables constructed from the "all-cause" mortality data reported in the KP Smoking study data by gender and smoking status (Friedman et al. 1997). The tables were generated using the age specific mortality rates found in: (1) table 3 (female never smokers and female current smokers by quantity smoked); (2) table 4 ((male never smokers and male current smokers by quantity smoked); (3) table 5 (female never smokers and female current smokers by duration smoked); (4) table 6 (male never smokers and male current smokers by duration smoked); (5) table 7 ((female never smokers and female former smokers by duration since quitting); (4) table 8 (male never smokers and male former smokers by duration since quitting). As an example of how to read the tables, Table 1 shows that for never smoking females in the KP study, life expectancy at age 35 is 45.8 years, while at 50 years of age, it is 33.27 years. By age 65, life expectancy is 19.21 years and at age 75, it is 10.26 years.

Table 1. Preliminary Life Table for Females, Never Smokers.

FEMALES			NUMBER OF DEATHS							LIFE TABLE			
age (x)	cigarette smoking status	quantity (cigarettes per day)	Person-years	All Causes	age-specific death rate ($_nm_x$)	width of age group*	$_nq_x$ (Fergany's method) Fergany (1971): $_nq_x = 1\text{-}e^{(-n*nmx)}$ where n is the width of the age interval.	l_x**	$_nd_x$	Number of years lived in interval (Fergany, 1971), where $_nL_x = {_nd_x}/{_nm_x}$ $_nL_x$	Total years lived to age x T_x	life expectancy at age x e_x	life expectancy of white females at age x *** e_x
35-49	Never	0	45,786	37	0.00081	15.0	0.01205	97,965	1,180	1,460,605	4,681,106	47.78	45.8
50-64	Never	0	49,744	118	0.00237	15.0	0.03496	96,785	3,383	1,426,245	3,220,502	33.27	31.6
65-74	Never	0	24,159	171	0.00708	10.0	0.06833	93,401	6,383	901,725	1,794,256	19.21	15.4
75-86	Never	0	12,285	299	0.02434	12.0	0.24964	87,019	21,723	892,531	892,531	10.26	12.0
87+	Never	0					1.00000	65,296	65,296				

* The KP population was not tracked to death over age 75; hence it is right truncated. To account for this, the final open-ended age interval was set to the expectation of life for a "white" female who reached the 75th birthday, which in 1990 was estimated to be 12.0 years (NCHS, 1994: Table 6-3.)

** The radix (age 35) was set to the l_{35} value found for "white" females in Table 6.1 of the 1990 Life Tables (National Center for Health Statistics, 1994: Table 6.1)

*** The life expectancy values for "white" females are taken from Table 6-1 of the 1990 US Life tables (National Center for Health Statistics, 1994: Table 6.1)

Table 2. Preliminary Life Table for Females, Current Smokers, less than or equal to 19 Cigarettes Daily

FEMALES			NUMBER OF DEATHS							LIFE TABLE			
age (x)	cigarette smoking status	quantity (cigarettes per day)	Person-years	All Causes	age-specific death rate ($_nm_x$)	width of age group*	$_nq_x$ (Fergany's method) Fergany (1971): $_nq_x = 1-e^{(-n*nmx)}$ where n is the width of the age interval.	l_x**	$_nd_x$	Number of years lived in interval (Fergany, 1971), where $_nL_x = {_nd_x}/{_nm_x}$ $_nL_x$	Total years lived to age x T_x	life expectancy at age x e_x	life expectancy of all other females at age x *** e_x
35-49	Current	≤19	12,086	12	0.00099	15.0	0.01478	96,220	1,422	1,432,605	4,432,364	46.06	42.6
50-64	Current	≤19	10,205	40	0.00392	15.0	0.05710	94,798	5,413	1,380,969	2,999,758	31.64	29.2
65-74	Current	≤19	3,582	45	0.01256	10.0	0.11806	89,385	10,553	839,980	1,618,789	18.11	17.8
75+	Current	≤19	808	23	0.02847	11.5	0.28122	78,832	22,169	778,809	778,809	9.88	11.5

* The KP population was not tracked to death over age 75; hence it is right truncated. To account for this, the final open-ended age interval was set to the expectation of life for a "all other" female who reached the 75th birthday, which in 1990 was estimated to be 11.5 years (NCHS, 1994: Table 6-3.)

** The radix (age 35) was set to the l_{35} value found for "all other" females in Table 6.1 of the 1990 Life Tables (National Center for Health Statistics, 1994: Table 6.1)

*** The life expectancy values for "all other" females are taken from Table 6-1 of the 1990 US Life tables (National Center for Health Statistics, 1994: Table 6.1)

Table 3. Preliminary Life Table for Females, Current Smokers, 20 or more Cigarettes Daily

FEMALES			NUMBER OF DEATHS							LIFE TABLE			
age (x)	cigarette smoking status	quantity (cigarettes per day)	Person-years	All Causes	age-specific death rate ($_nm_x$)	width of age group*	$_nq_x$ (Fergany's method) Fergany (1971): $_nq_x = 1-e^{(-n*nmx)}$ where n is the width of the age interval.	l_x**	$_nd_x$	Number of years lived in interval (Fergany, 1971), where $_nL_x = {_nd_x}/{_nm_x}$ $_nL_x$	Total years lived to age x T_x	life expectancy at age x e_x	life expectancy of black females at age x *** e_x
35-49	Curent	20+	12,851	25	0.00195	15.0	0.02876	95,625	2,750	1,413,649	4,130,898	43.20	41.3
50-64	Curent	20+	10,950	69	0.00630	15.0	0.09019	92,875	8,376	1,329,311	2,717,248	29.26	28.2
65-74	Curent	20+	3,583	70	0.01954	10.0	0.17747	84,498	14,996	767,566	1,387,938	16.43	17.0
75+	Curent	20+	588	24	0.04082	11.2	0.36432	69,503	25,321	620,372	620,372	8.93	11.2

* The KP population was not tracked to death over age 75; hence it is right truncated. To account for this, the final open-ended age interval was set to the expectation of life for "black" females who reached the 75th birthday, which in 1990 was estimated to be 11.2 years (NCHS, 1994: Table 6-3.)

** The radix (age 35) was set to the l_{35} value found for "black" females in Table 6.1 of the 1990 Life Tables (National Center for Health Statistics, 1994: Table 6.1)

*** The life expectancy values for "black" females are taken from Table 6-1 of the 1990 US Life tables (National Center for Health Statistics, 1994: Table 6.1)

Table 4. Preliminary Life Table for Males, Never Smokers.

MALES		NUMBER OF DEATHS					LIFE TABLE					
age (x)	cigarette smoking status	quantity (cigarettes per day)	Person-years	All Causes	age-specific death rate ($_nm_x$)	width of age group*	$_nq_x$ (Fergany's method) Fergany (1971): $_nq_x = 1-e^{(-n*nmx)}$ where n is the width of the age interval.	l_x**	$_nd_x$	Number of years lived in interval (Fergany, 1971), where $_nL_x = {_nd_x}/{_nm_x}$ $_nL_x$	Total years lived to age x T_x	life expectancy at age x e_x	life expectancy of white males at age x *** e_x
35-49	Never	0	29,916	49	0.00164	15.0	0.02427	95,725	2,323	1,418,380	4,167,846	43.54	40.1
50-64	Never	0	24,020	97	0.00404	15.0	0.05878	93,402	5,490	1,359,438	2,749,467	29.44	26.7
65-74	Never	0	11,466	161	0.01404	10.0	0.13100	87,912	11,517	820,189	1,390,029	15.81	12.1
75-85.7	Never	0	4,486	203	0.04525	9.4	0.33754	76,395	25,786	569,840	569,840	7.46	9.4
85.8+							1.00000	50,609	50,609				

* The KP population was not tracked to death over age 75; hence it is right truncated. To account for this, the final open-ended age interval was set to the expectation of life for a "white" male who reached the 75th birthday, which in 1990 was estimated to be 9.4 years (NCHS, 1994: Table 6-3.)

** The radix (age 35) was set to the l_{35} value found for "white" males in Table 6.1 of the 1990 Life Tables (National Center for Health Statistics, 1994: Table 6.1)

*** The life expectancy values for "white" males are taken from Table 6-1 of the 1990 US Life tables (National Center for Health Statistics, 1994: Table 6.1)

Table 5. Preliminary Life Table for Males, Current Smokers, less than or equal to 19 Cigarettes Daily

MALES		NUMBER OF DEATHS					LIFE TABLE					
age (x)	cigarette smoking status	quantity (cigarettes per day)	Person-years	All Causes	age-specific death rate ($_nm_x$)	width of age group*	$_nq_x$ (Fergany's method) Fergany (1971): $_nq_x = 1-e^{(-n*nmx)}$ where n is the width of the age interval.	l_x**	$_nd_x$	Number of years lived in interval (Fergany, 1971), where $_nL_x = {_nd_x}/{_nm_x}$ $_nL_x$	Total years lived to age x T_x	life expectancy at age x e_x	life expectancy of all other males at age x *** e_x
35-49	Current	≤19	7,895	17	0.00215	15.0	0.03178	92,218	2,931	1,361,170	3,769,454	40.88	36.0
50-64	Current	≤19	5,575	49	0.00879	15.0	0.12352	89,287	11,029	1,254,775	2,408,285	26.97	23.9
65-74	Current	≤19	1,740	44	0.02529	10.0	0.22343	78,259	17,486	691,476	1,153,510	14.74	14.0
75+	Current	≤19	490	22	0.04490	9.1	0.34134	60,773	20,744	462,033	462,033	7.60	9.1

* The KP population was not tracked to death over age 75; hence it is right truncated. To account for this, the final open-ended age interval was set to the expectation of life for a "other races" male who reached the 75th birthday, which in 1990 was estimated to be 9.1 years (NCHS, 1994: Table 6-3.)

** The radix (age 35) was set to the l_{35} value found for "other races" males in Table 6.1 of the 1990 Life Tables (National Center for Health Statistics, 1994: Table 6.1)

*** The life expectancy values for "all other" males are taken from Table 6-1 of the 1990 US Life tables (National Center for Health Statistics, 1994: Table 6.1)

Table 6. Preliminary Life Table for Males, Current Smokers, 20 or more Cigarettes Daily

MALES			NUMBER OF DEATHS							LIFE TABLE			
age (x)	cigarette smoking status	quantity (cigarettes per day)	Person-years	All Causes	age-specific death rate ($_nm_x$)	width of age group*	$_nq_x$ (Fergany's method) Fergany (1971): $_nq_x = 1\text{-}e^{(-n*nmx)}$ where n is the width of the age interval.	l_x**	$_nd_x$	Number of years lived in interval (Fergany, 1971), where $_nL_x = {_nd_x}/{_nm_x}$ $_nL_x$	Total years lived to age x T_x	life expectancy at age x e_x	life expectancy of black males at age x *** e_x
35-49	Curent	20+	13,304	49	0.00368	15.0	0.05375	90,827	4,882	1,325,455	3,547,555	39.06	34.1
50-64	Curent	20+	10,838	116	0.01070	15.0	0.14832	85,945	12,748	1,191,015	2,222,100	25.85	22.5
65-74	Curent	20+	2,995	71	0.02371	10.0	0.21106	73,198	15,449	651,683	1,031,086	14.09	13.2
75+	Curent	20+	545	37	0.06789	8.6	0.44603	57,749	25,758	379,403	379,403	6.57	8.6

* The KP population was not tracked to death over age 75; hence it is right truncated. To account for this, the final open-ended age interval was set to the expectation of life for A "black" male who reached the 75th birthday, which in 1990 was estimated to be 8.6 years (NCHS, 1994: Table 6-3.)

** The radix (age 35) was set to the l_{35} valuefound for "black" males in Table 6.1 of the 1990 Life Tables (National Center for Health Statistics, 1994: Table 6.1)

*** The life expectancy values for "black" males are taken from Table 6-1 of the 1990 US Life tables (National Center for Health Statistics, 1994: Table 6.1)

Table 7. Preliminary Life Table for Females, Former Smokers, 2-10 Years since Quitting

FEMALES			NUMBER OF DEATHS							LIFE TABLE			
age (x)	cigarette smoking status	Duration (Years Since Smoked)	Person-years	All Causes	age-specific death rate ($_nm_x$)	width of age group*	$_nq_x$ (Fergany's method) Fergany (1971): $_nq_x = 1\text{-}e^{(-n*nmx)}$ where n is the width of the age interval.	l_x**	$_nd_x$	Number of years lived in interval (Fergany, 1971), where $_nL_x = {_nd_x}/{_nm_x}$ $_nL_x$	Total years lived to age x T_x	life expectancy at age x e_x	life expectancy of black females at age x *** e_x
35-49	Former	2 - 10	5,493	1	0.00018	15.0	0.00273	96,308	263	1,442,649	4,453,374	46.24	41.3
50-64	Former	2 - 10	3,750	15	0.00400	15.0	0.05824	96,045	5,593	1,398,312	3,010,724	31.35	28.2
65-74	Former	2 - 10	1,572	15	0.00954	10.0	0.09101	90,452	8,232	862,707	1,612,413	17.83	17.0
75-85.5	Former	2 - 10	394	15	0.03807	11.2	0.34714	82,220	28,542	749,706	749,706	9.12	11.2
85.6+	Former	2 - 10					1.00000	53,678	53,678				

THE RED FONT INDICATES THAT 1 DEATH WAS SUBSTITUTED FOR ZERO DEATHS SINCE THE LIFE TABLE FUNCTION Lx CANNOT BE CALCULATED USING FERGANY'S METHOD WHEN DEATHS = ZERO

* The KP population was not tracked to death over age 75; hence it is right truncated. To account for this, the final open-ended age interval was set to the expectation of life for "black" females who reached the 75th birthday, which in 1990 was estimated to be 11.2 years (NCHS, 1994: Table 6-3.)

** The radix (age 35) was set to the l_{35} value found for "black" females in Table 6.1 of the 1990 Life Tables (National Center for Health Statistics, 1994: Table 6.1)

*** The life expectancy values for "black" females are taken from Table 6-1 of the 1990 US Life tables (National Center for Health Statistics, 1994: Table 6.1)

Table 8. Preliminary Life Table for Females, Former Smokers, 11-20 Years since Quitting

FEMALES			NUMBER OF DEATHS							LIFE TABLE			
age (x)	cigarette smoking status	Duration (Years Since Smoked)	Person-years	All Causes	age-specific death rate ($_nm_x$)	width of age group*	$_nq_x$ (Fergany's method) Fergany (1971): $_nq_x = 1\text{-}e^{(-n*nmx)}$ where n is the width of the age interval.	l_x**	$_nd_x$	Number of years lived in interval (Fergany, 1971), where $_nL_x = {_nd_x}/{_nm_x}$ $_nL_x$	Total years lived to age x T_x	life expectancy at age x e_x	life expectancy of black females at age x *** e_x
35-49	Former	11-20	6,027	4	0.00066	15.0	0.00991	95,741	948	1,428,990	4,493,641	46.94	41.3
50-64	Former	11-20	5,467	16	0.00293	15.0	0.04295	94,793	4,071	1,391,131	3,064,651	32.33	28.2
65-74	Former	11-20	2,505	21	0.00838	10.0	0.08041	90,721	7,295	870,226	1,673,520	18.45	17.0
75 - 85.2	Former	11-20	722	23	0.03186	11.5	0.30674	83,426	25,590	803,294	803,294	9.63	11.5
85.6+	Former	11-20					1.00000	57,836	57,836				

* The KP population was not tracked to death over age 75; hence it is right truncated. To account for this, the final open-ended age interval was set to the expectation of life for a "all other" female who reached the 75th birthday, which in 1990 was estimated to be 11.5 years (NCHS, 1994: Table 6-3.)

** The radix (age 35) was set to the l_{35} value found for "all other" females in Table 6.1 of the 1990 Life Tables (National Center for Health Statistics, 1994: Table 6.1)

*** The life expectancy values for "all other" females are taken from Table 6-1 of the 1990 US Life tables (National Center for Health Statistics, 1994: Table 6.1)

Table 9. Preliminary Life Table for Females, Former Smokers, More than 20 Years since Quitting

FEMALES			NUMBER OF DEATHS							LIFE TABLE			
age (x)	cigarette smoking status	Duration (Years Since Smoked)	Person-years	All Causes	age-specific death rate ($_nm_x$)	width of age group*	$_nq_x$ (Fergany's method) Fergany (1971): $_nq_x = 1\text{-}e^{(-n*nmx)}$ where n is the width of the age interval.	l_x**	$_nd_x$	Number of years lived in interval (Fergany, 1971), where $_nL_x = {_nd_x}/{_nm_x}$ $_nL_x$	Total years lived to age x T_x	life expectancy at age x e_x	life expectancy of white females at age x *** e_x
35-49	Former	>20	1,279	2	0.00156	15.0	0.02318	97,807	2,267	1,450,033	4,460,359	45.60	45.8
50-64	Former	>20	4,405	7	0.00159	15.0	0.02355	90,502	2,132	1,341,478	3,010,327	33.26	31.6
65-74	Former	>20	2,641	20	0.00757	10.0	0.07293	88,370	6,445	851,071	1,668,848	18.88	15.4
75 - 85.2	Former	>20	852	27	0.03169	12.0	0.31633	81,925	25,915	817,778	817,778	9.98	12.0
86+	Former	>20					1.00000	56,010	56,010				

* The KP population was not tracked to death over age 75; hence it is right truncated. To account for this, the final open-ended age interval was set to the expectation of life for a "white" female who reached the 75th birthday, which in 1990 was estimated to be 12.0 years (NCHS, 1994: Table 6-3.)

** The radix (age 35) was set to the l_{35} value found for "white" females in Table 6.1 of the 1990 Life Tables (National Center for Health Statistics, 1994: Table 6.1)

*** The life expectancy values for "white" females are taken from Table 6-1 of the 1990 US Life tables (National Center for Health Statistics, 1994: Table 6.1)

Table 10. Preliminary Life Table for Males, Former Smokers, 2-10 Years since Quitting

MALES			NUMBER OF DEATHS							LIFE TABLE			
age (x)	cigarette smoking status	Duration (Years Since Smoked)	Person-years	All Causes	age-specific death rate ($_nm_x$)	width of age group*	$_nq_x$ (Fergany's method) Fergany (1971): $_nq_x = 1-e^{(-n*nmx)}$ where n is the width of the age interval.	l_x**	$_nd_x$	Number of years lived in interval (Fergany, 1971), where $_nL_x = {_nd_x}/{_nm_x}$ $_nL_x$	Total years lived to age x T_x	life expectancy at age x e_x	life expectancy of black males at age x *** e_x
35-49	Former	2 - 10	5,571	12	0.00215	15.0	0.03179	92,632	2,945	1,367,273	3,748,205	40.46	34.1
50-64	Former	2 - 10	3,625	26	0.00717	15.0	0.10200	89,687	9,148	1,275,462	2,380,932	26.55	22.5
65-74	Former	2 - 10	977	14	0.01433	10.0	0.13350	80,539	10,752	805,387	1,105,470	13.73	13.2
75-83.1	Former	2 - 10	253	16	0.06324	8.6	0.41950	69,787	29,276	300,083	300,083	4.30	8.6
83.2+	Former	2 - 10					1.00000	40,511	40,511				

* The KP population was not tracked to death over age 75; hence it is right truncated. To account for this, the final open-ended age interval was set to the expectation of life for A "black" male who reached the 75th birthday, which in 1990 was estimated to be 8.6 years (NCHS, 1994: Table 6-3.)

** The radix (age 35) was set to the l_{35} valuefound for "black" males in Table 6.1 of the 1990 Life Tables (National Center for Health Statistics, 1994: Table 6.1)

*** The life expectancy values for "black" males are taken from Table 6-1 of the 1990 US Life tables (National Center for Health Statistics, 1994: Table 6.1)

Table 11. Preliminary Life Table for Males, Former Smokers, 11-20 Years since Quitting

MALES			NUMBER OF DEATHS							LIFE TABLE			
age (x)	cigarette smoking status	Duration (Years Since Smoked)	Person-years	All Causes	age-specific death rate ($_nm_x$)	width of age group*	$_nq_x$ (Fergany's method) Fergany (1971): $_nq_x = 1-e^{(-n*nmx)}$ where n is the width of the age interval.	l_x**	$_nd_x$	Number of years lived in interval (Fergany, 1971), where $_nL_x = {_nd_x}/{_nm_x}$ $_nL_x$	Total years lived to age x T_x	life expectancy at age x e_x	life expectancy of all other males at age x *** e_x
35-49	Former	11-20	6,210	5	0.00081	15.0	0.01200	91,596	1,100	1,365,677	3,920,859	42.81	36.0
50-64	Former	11-20	6,107	29	0.00475	15.0	0.06875	90,496	6,222	1,310,229	2,555,182	28.24	23.9
65-74	Former	11-20	2,548	52	0.02041	10.0	0.18460	84,275	15,557	762,315	1,244,953	14.77	14.0
75-82.6	Former	11-20	671	40	0.05961	9.1	0.41869	68,717	28,771	482,639	482,639	7.02	9.1
82.7+	Former	11-20					1.00000	39,946	39,946				

* The KP population was not tracked to death over age 75; hence it is right truncated. To account for this, the final open-ended age interval was set to the expectation of life for a "other races" male who reached the 75th birthday, which in 1990 was estimated to be 9.1 years (NCHS, 1994: Table 6-3.)

** The radix (age 35) was set to the l_{35} value found for "other races" males in Table 6.1 of the 1990 Life Tables (National Center for Health Statistics, 1994: Table 6.1)

*** The life expectancy values for "all other" males are taken from Table 6-1 of the 1990 US Life tables (National Center for Health Statistics, 1994: Table 6.1)

Table 12. Preliminary Life Table for Males, Former Smokers, More than 20 Years since Quitting

MALES			NUMBER OF DEATHS							LIFE TABLE			
age (x)	cigarette smoking status	Duration (Years Since Smoked)	Person-years	All Causes	age-specific death rate ($_n m_x$)	width of age group*	$_n q_x$ (Fergany's method) Fergany (1971): $_n q_x = 1-e^{(-n*nmx)}$ where n is the width of the age interval.	l_x**	$_n d_x$	Number of years lived in interval (Fergany, 1971), where $_n L_x = {_n d_x}/{_n m_x}$ $_n L_x$	Total years lived to age x T_x	life expectancy at age x e_x	life expectancy of white males at age x *** e_x
35-49	Former	>20	1,149	3	0.00261	15.0	0.03841	97,807	3,757	1,438,747	3,845,300	39.32	40.1
50-64	Former	>20	4,670	19	0.00407	15.0	0.05920	80,909	4,790	1,177,344	2,406,553	29.74	26.7
65-74	Former	>20	3,507	43	0.01226	10.0	0.11539	76,119	8,784	716,374	1,229,209	16.15	12.1
75-82.6	Former	>20	1,442	67	0.04646	9.4	0.35387	67,335	23,828	512,835	512,835	7.62	9.4
82.7+	Former	>20					1.00000	43,507	43,507				

* The KP population was not tracked to death over age 75; hence it is right truncated. To account for this, the final open-ended age interval was set to the expectation of life for a "white" male who reached the 75th birthday, which in 1990 was estimated to be 9.4 years (NCHS, 1994: Table 6-3.)

** The radix (age 35) was set to the l_{35} value found for "white" males in Table 6.1 of the 1990 Life Tables (National Center for Health Statistics, 1994: Table 6.1)

*** The life expectancy values for "white" males are taken from Table 6-1 of the 1990 US Life tables (National Center for Health Statistics, 1994: Table 6.1)

Discussion of the Preliminary Life Tables

As a starting point for discussing the preliminary life tables, I have a set of a priori assumptions:

1) Those who are younger have a longer life expectancy than those who are older, all else equal;
2) Women will generally have longer life expectancies than men, all else equal;
3) Current smokers will have a shorter life expectancy than non-smokers, all else equal; and
4) Among prior smokers - at any given age, those who stopped smoking more recently will have a *lower* life expectancy than those who stopped smoking in the more distant past, all else equal;

In the course of constructing the preliminary life tables, results comport with these assumptions – with the exception of life expectancy among prior smokers, which can be seen in Tables 13 and 14.

Table 13 – Female Life Expectancy by Age: Total, Never & Prior Smokers by Duration Since Quitting

Smoking Status	**Age 35***	**Age 50***	**Age 65***	**Age 75***
Never Smoker	47.92 (37,45786)	33.41 (118, 49,744)	19.35 (171, 24,159)	10.41 (299, 12,285)
Duration since Quitting (years)		,		
2-10	46.24 (1, 5,493)	31.35 (15, 3,750)	17.83 (15, 1,572)	9.12 (15, 394)
11-19	46.94 (4, 6,027)	32.33 (16, 5,467)	18.45 (21, 2,505)	9.63 (23, 723)
20+	45.60 (2, 1,279)	33.26 (7, 4,405)	18.88 (20, 2,641)	9.98 (27, 852)

*Life expectancy references expected years of life remaining at an exact age. The numbers in parentheses shown within each cell represent, respectively, all-cause deaths and person years.

Table 14 – Male Life Expectancy by Age: Total, Never Prior Smokers by Duration Since Quitting

Smoking Status	**Age 35***	**Age 50***	**Age 65***	**Age 75***
Never Smoker	43.70 (49, 29,916)	29.60 (97, 24,020)	15.98 (161, 11,466)	7.66 (203, 4,486)
Duration since Quitting (years)				
2-10	40.65 (12, 5,571)	26.55 (26, 3,625)	13.73 (14, 977)	4.30 (16, 253)
11-19	42.81 (5, 6,210)	28.24 (29, 6,107)	14.77 (52, 2,548)	7.02 (40, 671)
20+	39.32 (3, 1,499)	29.74 (19, 4,670)	16.15 (43, 3,507)	7.62 (67, 1,442)

*Life expectancy references expected years of life remaining at an exact age. The numbers in parentheses shown within each cell represent, respectively, all-cause deaths and person years.

For females, an anomaly, contrary to my a priori assumption, is found at age 35, where, the life expectancy of those who quit smoking more than 20 years ago (45.60), is less than both those who quit 2-10 years ago (46.29) and those who quit 11-20 years ago (46.94). For males, anomalies contrary to my a priori assumptions are found at age 35, and at both age 50 and age 65, as follows:

(1) At age 35, male life expectancy of those who quit smoking more than 20 years ago (39.32), is less than both those who quit 2-10 years ago (40.46) and those who quit 11-20 years ago (42.81);

(2) At age 50, male life expectancy of those who quit more than 20 years ago (29.74) is, as expected, higher than both those who quit 2-10 years ago (26.55) and those who quit 11-20 years ago (28.24). However, it is slightly above the life expectancy of those males who never smoked (29.60); and

(3) Similarly, at age 65, male life expectancy of those who quit more than 20 years ago (16.15) is, as expected, higher both those who quit 2-10 years ago (13.73) and those who quit 11-20 years ago (14.77). However, it is slightly above the life expectancy of those males who never smoked (15.90).

In the original KP smoking study publication (Friedman et al., 1997), there are acknowledgements to the anomalies found in the preliminary life tables. The first of these acknowledgments is for women, “In the youngest group, 35- to 49-year-olds, all-cause mortality was the highest among those who had quit smoking for more than twenty years, but this was based on only two deaths.” (Friedman et al., 1997: 490). The second acknowledgment is in terms of men, “All-cause deaths among men showed decreasing risks with increased duration of quitting only in the 50- to 64 and 75+ year age groups…” (Friedman et al., 1997: 490). A third acknowledgement generalizes the anomalies, "An inverse relationship of risk with duration of quitting was often but not consistently seen." (Friedman et al., 1997: 493).

I agree with Friedman et al. (1997:490) that small numbers play a role in these anomalies in a specific instance they note where an anomaly is "…based on only two deaths." I also believe that small numbers play a role in all of the anomalies they note, which are more explicitly stated by us. The issue of small numbers is important in the KP smoking data because they can reflect either the stochastic uncertainty associated with a small population, sampling uncertainty, or a combination of both.[2]

As an example of these uncertainties, I can use the coefficient of variation (CV), which is designed to make comparisons of relative uncertainty, whether stochastic or sample-based. First consider the case of males aged 35-49 who have never smoked. The mean is 49 (deaths), where 49 = 29,916*0.00164 and .00164 = 49/29,916. The variance is 48.98, where 48.98 = 29,916*0.00164*(1- 0.00164) and the standard deviation is 6.99 = $(48.98)^{0.5}$. Thus, the coefficient of variation is 0.1427 =6.99/49. Next, consider the case of males aged 35-49 who quit smoking 20 or more years ago. The mean is 3, where 3 = 1,149*0.00261 and 0.00261 = 3/1,149. The variance is 2.99, where 2.99 = 1,149*0.00261*(1-0.00261) and the standard deviation is 1.729 = $(2.99)^{0.5}$. Thus, the coefficient of variation is 0.576 = 1.729/3. The CV for males aged 35-49 who quit smoking 20 or more years ago is four times higher than the CV for males aged 35-49 who have never smoked (0.576/0.1427 = 4.04), which indicates that the uncertainty for the latter is four times greater than the former. A similar situation exists in terms of the CVs for females aged 35-49 who never smoked and females aged 35-49 who quit smoking 20 or more year ago.

Because of the widespread use of the KP Smoking Study data, I believe that it is worth the effort to resolve the anomalies identified here.[3] To this end, I first interpolate the hazard rates (${}_nq_x$ values) found in the preliminary life tables so that I have a set of hazard rates for age groups of five –year widths, starting at age 35-40 and ending at age 80-85. I then use these interpolated hazard rates as input to Gompertz-type regression models, which are used to generate a “smoothed” set of hazard rates specific to each group associated with the 12 preliminary life tables that encompass a wider range of five year age groups, where feasible (e.g., for never smokers, these estimated hazard rates start at age 20-24 and end at age 90-94; however, for current smokers who have smoked for more than 40 years, the estimated hazard rates start at age 55-59 and end at age 90-94). These estimated hazard rates were examined in terms of the a priori assumptions. At this point, an anomaly remained for males who reported that they were current smokers, but had smoked for less than twenty years. The adjustment consisted of replacing zero deaths with three deaths in each of two oldest age groups, recalculating the hazard rates and using these revised hazard rates as the input to the Gompertz-type model.

With the estimated hazard rates in hand, an adjustment was used to convert them so that they would apply to the US population in 1990. To this end, a 1990 US life table was used (details here) as a “standard table” (Kintner, 2004; United Nations, 1982) and a gender-specific ratio, 1/(US e35/KPe35), was formed for all smoking groups employed in the KP Smoking Study. These adjusted hazard rates were then graphed and examined for anomalies. Three anomalies were found. The first was that females who reported being former smokers who had quit more than 20 years ago generally had lower hazard rates than females who reported never smoking.

The second was that males who reported being former smokers who had quit more than 20 years ago generally had lower hazard rates than males who reported never smoking. The third anomaly was that males who reported being former smokers who had quit between 2 and 10 years ago generally had lower hazard rates than those who quit 11-20 years ago. The first and second anomalies were resolved using simple averages at each group between the hazard rates for former smokers who had quit 20+ and 11-20 years ago, respectively. The third anomaly was resolved by using simple averages at each group between the hazard rates for former male smoker who had quit between 11-20 years ago and 2-10 years ago, respectively. There is a sound justification for using this approach to resolve the each of the three anomalies.[4] Recall that age-specific death rates (${}_nm_x$), life-table death rates, also known as hazard rates, (${}_nq_x$), and survival ratios (${}_nS_x$), though differently derived, are closely related to each other. If one of these functions is known, reference to a system for constructing life tables makes it possible to estimate immediately the approximate levels of the other two functions. Fergany's method (1971), for example, converts ${}_nmx$ into ${}_nq_x$, and ${}_nS_x$ is simply found by subtracting ${}_nq_x$ from 1. Because ${}_nq_x$ directly generates l_x and ${}_nd_x$, and in combination with ${}_nm_x$ generates ${}_nL_x$, and, hence, T_x, it is considered to be the fundamental life table function.

The expectation of life at a given age, e_x, is in a different category than ${}_nq_x$. It is both the result of the cumulative addition of specific values (T_x) and a ratio because $e_x = T_x/l_x$. It is powerful in that it represents the one synthetic measure by which the "general" level of mortality can be summarized in a single figure (United Nations, 1982: 25). This is evident in tables 12 and 13, where the life expectancy values clearly show the inconsistencies I have noted. However, life expectancy (e_x) cannot be used to construct ${}_nq_x$ because, T_x, the numerator used to created e_x, is the result of the cumulative addition of ${}_nL_x$ while lx, the denominator used to create e_x, is the result of the cumulative subtraction of ${}_nd_x$ from l_x. What these relationships suggest is that inconsistencies in the KP life tables need to be dealt with by revising the underlying ${}_nq_x$ values (or equivalently, the underlying ${}_nS_x$ values). Thus, In terms of resolving inconsistency that females who quit smoking 20+ years ago have a lower life expectancy at age 35 than females who quit smoking 2-10 years ago and females who quit smoking 11-20 years ago, I can take the average ${}_nq_x$ at each age between females who never smoked and those who quit smoking 20+ years ago.

Constructing Revised Life Tables for the Kaiser Permanente Population

It should be clear from the preceding discussion that the complexities found in the life table lead us to a method that allows us to directly assess the ${}_nq_x$ values via their reciprocals, ${}_nS_x$ values. I can do this because ${}_nS_x = 1 - {}_nq_x$. I can do this using a method described by Swanson and Tedrow (2012). In this approach, note that when the radix of a life table is equal to 1 (l_0 = 1.00) life expectancy at birth can be computed directly from the expression:

$$e_0 = S_0 + (S_0*S_1) + (S_0*S_1*S_2) +,...,+(S_0*S_1*S_2,...,*S_x) \qquad [5]$$

where

e_0 = life expectancy at birth

S_0 = survivorship from t=0 (e.g., birth) to t=1(e.g., age 1)

S_1 = survivorship from t=1 (e.g., age 1) to t=2(e.g., age 2)

and so on through S_x

and $S_x = {}_1L_x / {}_1L_{(x-n)}$

Equation [5] is set up for single year age groups. However, I can generalize it to other age groups: ${}_nS_x = {}_nL_x / {}_nL_{(x-n)}$, so that

$$e_0 = {}_nS_0 + ({}_nS_0 * {}_nS_1) + ({}_nS_0 * {}_nS_1 * {}_nS_2) +,...,+({}_nS_0 * {}_nS_1 * {}_nS_2,...,* {}_nS_x) \quad [6]$$

As equation [5] and equation [6] both imply, the fundamental life table function is inherent in this method. That is via the ${}_nS_x$ values, I have ${}_nq_x$ values. Swanson and Tedrow (2012) show a derivation of the relationship between survivorship rates and life expectancy as shown in equation [5] and generalized to equation [6]. Using this approach, I have generated "revised" life tables from the preliminary life tables (to include the processing I did to resolve the anomalies noted earlier). The results are found in tables 14 through 31.

Table 14. Revised Life Table for Kaiser Permanente Females, Never Smokers: Survivorship and Life Expectancy by Age

SMOKING STATUS'	AGE	$_5S_x$	female e_x
NEVER	**20**	0.99989	66.25
NEVER	**25**	0.99968	61.25
NEVER	**30**	0.99921	56.27
NEVER	**35**	0.99830	51.32
NEVER	**40**	0.99671	46.41
NEVER	**45**	0.99411	41.56
NEVER	**50**	0.99008	36.81
NEVER	**55**	0.98411	32.17
NEVER	**60**	0.97557	27.69
NEVER	**65**	0.96372	23.39
NEVER	**70**	0.94767	19.27
NEVER	**75**	0.92640	15.33
NEVER	**80**	0.89875	11.55
NEVER	**85**	0.86338	7.85
NEVER	**90**	0.81877	4.09

Table 15. Revised Life Table for Kaiser Permanente Males, Never Smokers: Survivorship and Life Expectancy by Age

SMOKING STATUS'	AGE	$_5S_x$	male e_x
NEVER	**20**	0.99972	62.05
NEVER	**25**	0.99922	57.07
NEVER	**30**	0.99822	52.12
NEVER	**35**	0.99642	47.21
NEVER	**40**	0.99344	42.38
NEVER	**45**	0.98880	37.66
NEVER	**50**	0.98194	33.08
NEVER	**55**	0.97215	28.69
NEVER	**60**	0.95866	24.52
NEVER	**65**	0.94055	20.57
NEVER	**70**	0.91677	16.87
NEVER	**75**	0.88616	13.40
NEVER	**80**	0.84740	10.13
NEVER	**85**	0.79905	6.95
NEVER	**90**	0.73951	3.70

Table 16. Revised Life Table for Kaiser Permanente Females, Former Smokers who quit Between 2 and 10 Years Ago: Survivorship and Life Expectancy by Age.

SMOKING STATUS'	AGE	$_5S_x$	female e_x
FORMER, 2-10	**20**	0.99999	62.31
FORMER, 2-10	**25**	0.99997	57.31
FORMER, 2-10	**30**	0.99987	52.31
FORMER, 2-10	**35**	0.99958	47.32
FORMER, 2-10	**40**	0.99883	42.34
FORMER, 2-10	**45**	0.99715	37.39
FORMER, 2-10	**50**	0.99368	32.49
FORMER, 2-10	**55**	0.98698	27.70
FORMER, 2-10	**60**	0.97483	23.06
FORMER, 2-10	**65**	0.95385	18.66
FORMER, 2-10	**70**	0.91910	14.56
FORMER, 2-10	**75**	0.86357	10.84
FORMER, 2-10	**80**	0.77755	7.56
FORMER, 2-10	**85**	0.64789	4.72
FORMER, 2-10	**90**	0.45710	2.29

Table 17. Revised Life Table for Kaiser Permanente Males, Former Smokers who quit Between 2 and 10 Years ago: Survivorship and Life Expectancy by Age,

SMOKING STATUS'	AGE	$_5S_x$	male e_x
FORMER, 2-10	**20**	0.99955	59.09
FORMER, 2-10	**25**	0.99880	54.12
FORMER, 2-10	**30**	0.99733	49.18
FORMER, 2-10	**35**	0.99477	44.32
FORMER, 2-10	**40**	0.99062	39.55
FORMER, 2-10	**45**	0.98430	34.92
FORMER, 2-10	**50**	0.97512	30.48
FORMER, 2-10	**55**	0.96226	26.26
FORMER, 2-10	**60**	0.94480	22.29
FORMER, 2-10	**65**	0.92168	18.59
FORMER, 2-10	**70**	0.89171	15.17
FORMER, 2-10	**75**	0.85360	12.01
FORMER, 2-10	**80**	0.80589	9.07
FORMER, 2-10	**85**	0.74699	6.26
FORMER, 2-10	**90**	0.67518	3.38

Table 18. Revised Life Table for Kaiser Permanente Females, Former Smokers who quit Between 11 and 20 Years ago: Survivorship and Life Expectancy by Age

SMOKING STATUS'	AGE	${}_5S_x$	female e_x
FORMER, 11-20	**20**	0.99994	64.70
FORMER, 11-20	**25**	0.99978	59.70
FORMER, 11-20	**30**	0.99940	54.72
FORMER, 11-20	**35**	0.99859	49.75
FORMER, 11-20	**40**	0.99705	44.82
FORMER, 11-20	**45**	0.99436	39.95
FORMER, 11-20	**50**	0.98992	35.18
FORMER, 11-20	**55**	0.98295	30.54
FORMER, 11-20	**60**	0.97247	26.07
FORMER, 11-20	**65**	0.95721	21.81
FORMER, 11-20	**70**	0.93564	17.78
FORMER, 11-20	**75**	0.90588	14.00
FORMER, 11-20	**80**	0.86570	10.46
FORMER, 11-20	**85**	0.81245	7.08
FORMER, 11-20	**90**	0.74304	3.72

Table 19. Revised Life Table for Kaiser Permanente Males, Former Smokers who quit Between 11 and 20 Years ago: Survivorship and Life Expectancy by Age

SMOKING STATUS'	AGE	${}_5S_x$	male e_x
FORMER, 11-20	**20**	0.99974	59.40
FORMER, 11-20	**25**	0.99928	54.41
FORMER, 11-20	**30**	0.99832	49.45
FORMER, 11-20	**35**	0.99651	44.54
FORMER, 11-20	**40**	0.99337	39.69
FORMER, 11-20	**45**	0.98823	34.96
FORMER, 11-20	**50**	0.98020	30.37
FORMER, 11-20	**55**	0.96813	25.99
FORMER, 11-20	**60**	0.95055	21.84
FORMER, 11-20	**65**	0.92562	17.98
FORMER, 11-20	**70**	0.89105	14.42
FORMER, 11-20	**75**	0.84408	11.19
FORMER, 11-20	**80**	0.78137	8.25
FORMER, 11-20	**85**	0.69893	5.56
FORMER, 11-20	**90**	0.59208	2.96

Table 20. Revised Life Table for Kaiser Permanente Females, Former Smokers who quit more than 20 Years ago: Survivorship and Life Expectancy by Age

SMOKING STATUS'	AGE	$_5S_x$	female e_x
FORMER, 20+	**20**	0.99985	65.39
FORMER, 20+	**25**	0.99957	60.40
FORMER, 20+	**30**	0.99901	55.43
FORMER, 20+	**35**	0.99796	50.48
FORMER, 20+	**40**	0.99617	45.59
FORMER, 20+	**45**	0.99328	40.76
FORMER, 20+	**50**	0.98884	36.04
FORMER, 20+	**55**	0.98228	31.44
FORMER, 20+	**60**	0.97290	27.01
FORMER, 20+	**65**	0.95984	22.76
FORMER, 20+	**70**	0.94207	18.72
FORMER, 20+	**75**	0.91837	14.87
FORMER, 20+	**80**	0.88730	11.19
FORMER, 20+	**85**	0.84721	7.61
FORMER, 20+	**90**	0.79619	3.98

Table 21. Revised Life Table for Kaiser Permanente Males, Former Smokers who quit more than 20 Years ago: Survivorship and Life Expectancy by Age

SMOKING STATUS'	AGE	$_5S_x$	male e_x
FORMER, 20+	**20**	0.99964	60.67
FORMER, 20+	**25**	0.99911	55.69
FORMER, 20+	**30**	0.99810	50.74
FORMER, 20+	**35**	0.99632	45.83
FORMER, 20+	**40**	0.99338	41.00
FORMER, 20+	**45**	0.98872	36.28
FORMER, 20+	**50**	0.98162	31.69
FORMER, 20+	**55**	0.97112	27.28
FORMER, 20+	**60**	0.95597	23.09
FORMER, 20+	**65**	0.93461	19.16
FORMER, 20+	**70**	0.90510	15.50
FORMER, 20+	**75**	0.86502	12.12
FORMER, 20+	**80**	0.81147	9.02
FORMER, 20+	**85**	0.74094	6.11
FORMER, 20+	**90**	0.64926	3.25

Table 22. Revised Life Table for Kaiser Permanente Females, Current Smokers, Less than 20 Cigarettes Daily: Survivorship and Life Expectancy by Age

SMOKING STATUS'	AGE	$_5S_x$	female e_x
CURRENT ≤ 19 CIGS DAILY	**20**	0.99986	63.25
CURRENT ≤ 19 CIGS DAILY	**25**	0.99958	58.26
CURRENT ≤ 19 CIGS DAILY	**30**	0.99894	53.28
CURRENT ≤ 19 CIGS DAILY	**35**	0.99771	48.34
CURRENT ≤ 19 CIGS DAILY	**40**	0.99551	43.45
CURRENT ≤ 19 CIGS DAILY	**45**	0.99187	38.65
CURRENT ≤ 19 CIGS DAILY	**50**	0.98618	33.97
CURRENT ≤ 19 CIGS DAILY	**55**	0.97766	29.44
CURRENT ≤ 19 CIGS DAILY	**60**	0.96538	25.11
CURRENT ≤ 19 CIGS DAILY	**65**	0.94820	21.01
CURRENT ≤ 19 CIGS DAILY	**70**	0.92477	17.16
CURRENT ≤ 19 CIGS DAILY	**75**	0.89352	13.56
CURRENT ≤ 19 CIGS DAILY	**80**	0.85263	10.17
CURRENT ≤ 19 CIGS DAILY	**85**	0.80003	6.93
CURRENT ≤ 19 CIGS DAILY	**90**	0.73334	3.67

Table 23. Revised Life Table for Kaiser Permanente Males, Current Smokers, Less than 20 Cigarettes Daily: Survivorship and Life Expectancy by Age

SMOKING STATUS'	AGE	$_5S_x$	male e_x
CURRENT ≤ 19 CIGS DAILY	**20**	0.99951	57.73
CURRENT ≤ 19 CIGS DAILY	**25**	0.99870	52.76
CURRENT ≤ 19 CIGS DAILY	**30**	0.99709	47.83
CURRENT ≤ 19 CIGS DAILY	**35**	0.99426	42.97
CURRENT ≤ 19 CIGS DAILY	**40**	0.98965	38.22
CURRENT ≤ 19 CIGS DAILY	**45**	0.98261	33.62
CURRENT ≤ 19 CIGS DAILY	**50**	0.97232	29.21
CURRENT ≤ 19 CIGS DAILY	**55**	0.95786	25.04
CURRENT ≤ 19 CIGS DAILY	**60**	0.93816	21.15
CURRENT ≤ 19 CIGS DAILY	**65**	0.91199	17.54
CURRENT ≤ 19 CIGS DAILY	**70**	0.87797	14.23
CURRENT ≤ 19 CIGS DAILY	**75**	0.83458	11.21
CURRENT ≤ 19 CIGS DAILY	**80**	0.78012	8.43
CURRENT ≤ 19 CIGS DAILY	**85**	0.71274	5.81
CURRENT ≤ 19 CIGS DAILY	**90**	0.63040	3.15

Table 24. Revised Life Table for Kaiser Permanente Females, Current Smokers, 20 or more Cigarettes Daily: Survivorship and Life Expectancy by Age

SMOKING STATUS'	AGE	$_5S_x$	female e_x
CURRENT 20+ CIGS DAILY	**20**	0.99962	59.39
CURRENT 20+ CIGS DAILY	**25**	0.99897	54.41
CURRENT 20+ CIGS DAILY	**30**	0.99766	49.47
CURRENT 20+ CIGS DAILY	**35**	0.99532	44.58
CURRENT 20+ CIGS DAILY	**40**	0.99145	39.79
CURRENT 20+ CIGS DAILY	**45**	0.98545	35.13
CURRENT 20+ CIGS DAILY	**50**	0.97660	30.65
CURRENT 20+ CIGS DAILY	**55**	0.96404	26.39
CURRENT 20+ CIGS DAILY	**60**	0.94676	22.37
CURRENT 20+ CIGS DAILY	**65**	0.92363	18.63
CURRENT 20+ CIGS DAILY	**70**	0.89332	15.17
CURRENT 20+ CIGS DAILY	**75**	0.85439	11.98
CURRENT 20+ CIGS DAILY	**80**	0.80521	9.02
CURRENT 20+ CIGS DAILY	**85**	0.74397	6.21
CURRENT 20+ CIGS DAILY	**90**	0.66870	3.34

Table 25. Revised Life Table for Kaiser Permanente Males, Current Smokers, 20 or more Cigarettes Daily: Survivorship and Life Expectancy by Age

SMOKING STATUS'	AGE	$_5S_x$	male e_x
CURRENT 20+ CIGS DAILY	**20**	0.99889	55.37
CURRENT 20+ CIGS DAILY	**25**	0.99738	50.43
CURRENT 20+ CIGS DAILY	**30**	0.99471	45.57
CURRENT 20+ CIGS DAILY	**35**	0.99041	40.81
CURRENT 20+ CIGS DAILY	**40**	0.98396	36.20
CURRENT 20+ CIGS DAILY	**45**	0.97475	31.79
CURRENT 20+ CIGS DAILY	**50**	0.96211	27.62
CURRENT 20+ CIGS DAILY	**55**	0.94530	23.70
CURRENT 20+ CIGS DAILY	**60**	0.92352	20.08
CURRENT 20+ CIGS DAILY	**65**	0.89589	16.74
CURRENT 20+ CIGS DAILY	**70**	0.86149	13.68
CURRENT 20+ CIGS DAILY	**75**	0.81931	10.88
CURRENT 20+ CIGS DAILY	**80**	0.76830	8.28
CURRENT 20+ CIGS DAILY	**85**	0.70734	5.78
CURRENT 20+ CIGS DAILY	**90**	0.63524	3.18

Table 26. Revised Life Table for Kaiser Permanente Females, Current Smokers, less than 20 years of smoking: Survivorship and Life Expectancy by Age

SMOKING STATUS'	AGE	$_5S_x$	female e_x
CURRENT < 20	**20**	0.99991	64.83
CURRENT < 20	**25**	0.99972	59.84
CURRENT < 20	**30**	0.99926	54.86
CURRENT < 20	**35**	0.99834	49.90
CURRENT < 20	**40**	0.99665	44.98
CURRENT < 20	**45**	0.99379	40.13
CURRENT < 20	**50**	0.98920	35.38
CURRENT < 20	**55**	0.98219	30.77
CURRENT < 20	**60**	0.97189	26.33
CURRENT < 20	**65**	0.95722	22.09
CURRENT < 20	**70**	0.93689	18.08
CURRENT < 20	**75**	0.90937	14.29
CURRENT < 20	**80**	0.87285	10.72
CURRENT < 20	**85**	0.82524	7.28
CURRENT < 20	**90**	0.76413	3.82

Table 27. Revised Life Table for Kaiser Permanente Males, Current Smokers, less than 20 years of smoking: Survivorship and Life Expectancy by Age

SMOKING STATUS'	AGE	$_5S_x$	male e_x
CURRENT < 20	**20**	0.99906	61.20
CURRENT < 20	**25**	0.99792	56.26
CURRENT < 20	**30**	0.99600	51.38
CURRENT < 20	**35**	0.99307	46.58
CURRENT < 20	**40**	0.98883	41.91
CURRENT < 20	**45**	0.98298	37.38
CURRENT < 20	**50**	0.97520	33.03
CURRENT < 20	**55**	0.96514	28.87
CURRENT < 20	**60**	0.95243	24.91
CURRENT < 20	**65**	0.93668	21.16
CURRENT < 20	**70**	0.91748	17.59
CURRENT < 20	**75**	0.89441	14.17
CURRENT < 20	**80**	0.86703	10.84
CURRENT < 20	**85**	0.83487	7.50
CURRENT < 20	**90**	0.79746	3.99

Table 28. Revised Life Table for Kaiser Permanente Females, Current Smokers, 20 -39 Years of smoking: Survivorship and Life Expectancy by Age

SMOKING STATUS'	AGE	${}_5S_x$	female e_x
CURRENT, 20-39	**20**	0.99961	62.22
CURRENT, 20-39	**25**	0.99900	57.24
CURRENT, 20-39	**30**	0.99783	52.30
CURRENT, 20-39	**35**	0.99581	47.41
CURRENT, 20-39	**40**	0.99259	42.61
CURRENT, 20-39	**45**	0.98777	37.93
CURRENT, 20-39	**50**	0.98084	33.40
CURRENT, 20-39	**55**	0.97125	29.05
CURRENT, 20-39	**60**	0.95835	24.91
CURRENT, 20-39	**65**	0.94143	20.99
CURRENT, 20-39	**70**	0.91969	17.30
CURRENT, 20-39	**75**	0.89226	13.81
CURRENT, 20-39	**80**	0.85817	10.48
CURRENT, 20-39	**85**	0.81638	7.21
CURRENT, 20-39	**90**	0.76576	3.83

Table 29. Revised Life Table for Kaiser Permanente Males, Current Smokers, 20 – 39 years of smoking: Survivorship and Life Expectancy by Age

SMOKING STATUS'	AGE	${}_5S_x$	male e_x
CURRENT, 20-39	**20**	0.99929	55.97
CURRENT, 20-39	**25**	0.99818	51.01
CURRENT, 20-39	**30**	0.99607	46.10
CURRENT, 20-39	**35**	0.99250	41.28
CURRENT, 20-39	**40**	0.98685	36.59
CURRENT, 20-39	**45**	0.97843	32.08
CURRENT, 20-39	**50**	0.96642	27.79
CURRENT, 20-39	**55**	0.94988	23.75
CURRENT, 20-39	**60**	0.92776	20.01
CURRENT, 20-39	**65**	0.89888	16.56
CURRENT, 20-39	**70**	0.86194	13.43
CURRENT, 20-39	**75**	0.81551	10.58
CURRENT, 20-39	**80**	0.75805	7.97
CURRENT, 20-39	**85**	0.68785	5.51
CURRENT, 20-39	**90**	0.60312	3.02

Table 30. Revised Life Table for Kaiser Permanente Females, Current Smokers, 40+ Years of smoking: Survivorship and Life Expectancy by Age

SMOKING STATUS'	AGE	$_5S_x$	female e_x
CURRENT, 40+	**50**	0.97018	31.72
CURRENT, 40+	**55**	0.95700	26.27
CURRENT, 40+	**60**	0.93994	22.45
CURRENT, 40+	**65**	0.91833	18.88
CURRENT, 40+	**70**	0.89144	15.56
CURRENT, 40+	**75**	0.85852	12.46
CURRENT, 40+	**80**	0.81873	9.51
CURRENT, 40+	**85**	0.77122	6.61
CURRENT, 40+	**90**	0.71507	3.58

Table 31. Revised Life Table for Kaiser Permanente Males, Current Smokers, 40+ years of smoking: Survivorship and Life Expectancy by Age

SMOKING STATUS'	AGE	$_5S_x$	male e_x
CURRENT, 40+	**50**	0.95364	26.50
CURRENT, 40+	**55**	0.93412	22.06
CURRENT, 40+	**60**	0.90921	18.62
CURRENT, 40+	**65**	0.87805	15.47
CURRENT, 40+	**70**	0.83975	12.62
CURRENT, 40+	**75**	0.79335	10.03
CURRENT, 40+	**80**	0.73784	7.65
CURRENT, 40+	**85**	0.67220	5.36
CURRENT, 40+	**90**	0.59532	2.98

Discussion

Table 32.a provides a summary of the e_{35} results for never and former smokers by gender. It shows that the results both within and across gender by smoking status are consistent. Life expectancy at age 35 is higher for females at each smoking status and within gender is highest for never smokers, lowest

for those who quit smoking 2-10 years ago and increases monotonically for former smokers as the years increase since quitting.

Table 32a. Revised Life Table Summary Results by Smoking Status for Kaiser Permanente Members

	Life Expectancy at Age 35	
smoking status	Female	Male
Never	51.32	47.21
Former, 2-10 Yrs Since Quitting	47.32	44.32
Former, 11-20 Yrs Since Quitting	49.75	44.54
Former, 20+ Yrs Since Quitting	50.48	45.83

Table 32.b provides a summary of the e_{35} results for never and current smokers by gender and number of cigarettes smoked daily. It shows that the results both within and across gender by smoking status are consistent. Life expectancy at age 35 is highest for female never smokers and lowest for males who smoke 20 or more cigarettes daily.

Table 32b. Revised Life Table Summary Results by Smoking Status for Kaiser Permanente Members

	Life Expectancy at Age 35	
smoking status	Female	Male
Never	51.32	47.21
Current, < 20 Cigarettes Daily	48.34	42.97
Current, ≥20 Cigarettes Daily	44.58	40.81

Table 32.b provides a summary of the e_{55} results by gender for never smokers, duration since quitting for former smokers, and duration smoked for current smokers. It shows that the results both within and across gender by smoking status are consistent in terms of: (1) never smokers v. former and current smokers; and (2) never smokers v. current smokers. It also shows that females who quit smoking 20 or more years ago have higher e_{55} values than either those who quit more recently or current smokers. In addition, life expectancy at age 55 is highest for female never smokers and lowest for males who have smoked for 40 or more years. However, for males, there are two inconsistencies: (1) the highest e_{55} value among former and current smokers is found for males who are current smokers but have smoked less than 20 years; and (2) e_{55} for males 2-10 years since quitting is higher than e_{55} for males who quit smoking 11-20 years ago.

Table 32c. Revised Life Table Summary Results by Smoking Status for Kaiser Permanente Members

	Life Expectancy at Age 55	
smoking status	Female	Male
Never	32.17	28.69
Former, 20+ Yrs Since Quitting	31.44	27.28
Former, 11-20 Yrs Since Quitting	30.54	25.99
Former, 2-10 Yrs Since Quitting	27.70	26.26
Current, < 20 Yrs Duration	30.77	28.87
Current, 20-39 Yrs Duration	29.05	23.75
Current, 40+ Yrs Duration	26.27	22.06

Discussion of Revised Life Tables for Kaiser Permanente Population

Given the two remaining anomalies for males, I nonetheless find the results encouraging in that the life tables by smoking status are otherwise consistent, especially considering the small sample size as represented by the KP study population and other limitations, namely, that I do not know: (1) how many cigarettes were smoked daily by duration for current smokers; and (2) how long former smokers smoked and how many cigarettes they smoked daily. These factors would clearly cause differences in mortality and are likely to be underlying this specific anomaly and others that ae not apparent to us (See also Endnote 3). This situation is known as “hidden heterogeneity” among demographers (Vaupel and Missov 2014). Unfortunately, what is hidden to us in the KP study is likely to remain hidden.

Discussion of Revised KP Life Tables applied to the US Population

The development of adequate life tables and survivorship information from the KP Smoking Study is of interest historically but to be relevant, the information needs to be applied to a more current and broader range of people, such as found in the work of Bachand and Sulsky (2013) and Bachand, Sulsky, and Curtin (2018). To effect this application, the starting point is the recognition that the Kaiser Permanente population as represented in the smoking study (Friedman et al. 1997) has higher life expectancy values than does the comparable US population from the same time period.

I came to this working conclusion after weighting the revised life expectancy values by age and smoking statuses to produce a mean life expectancy by selected age and smoking status and comparing the results to US life tables produced by the Human Mortality Database project (Wilmoth et al. 2017) for the period 1985-89, which corresponds to the time period of the KP smoking study (Friedman et al. 1997).[5] The weighted results take into account proportions of the KP study population in the smoking status categories (never current by time, former by years since quitting, and current by years smoked).[6] For example, e_{55} for US females 1985-89 is shown by the Human Mortality Database (HMD) is 26.69 while e_{55} for the weighted KP population (never, former by years since quitting, and current by years smoked) is 30.85; for US males 1985-89, the HMD life expectancy at age 55 is e_{55} =

21.75 while the weighted KP population (never, former by years since quitting, and current by years smoked) is 27.05. These differences suggest that the KP life tables need to be adjusted in order to better approximate the US population. This can be done by applying scalar values to the $_nS_x$ values found in the revised KP life tables (tables 14-31), a process I describe later.

There are several reasons that the KP subject population has a higher life expectancy than the comparable US population. First, the KP subject population is from the San Francisco Bay area, and at the time of the study, life expectancy in California was higher than the US average. Springborn (2006), for example, shows life expectancy at birth for the total population of California in 1979-81 and 1990, respectively, as 74.8 and 76.2 (Table 1) while for the total population of the US in the same years, it is 73.9 and 75.4, respectively (Table 2). Second, KP members who joined as individuals did not have serious pre-existing conditions (Friedman, 2019). In general, it is likely safe to attribute the longer life expectancy of KP members relative to the general US population to differences in healthcare access, income, race, ethnicity, and smoking status.

The authors of the KP smoking study did not report ethnic and income groups (likely because the study population's sub-sets would have become too small to allow meaningful analysis). Similarly, they will become too small if I attempted to weight it not only by age and gender, but also race. Thus, I weight the e_x values by age and gender by smoking status proportions. Among females, CDC found that US females age 17 and over, 54.7 percent were never smokers, 20.9 percent were former smokers, and 24.4 percent were current smokers. Among US males age 17 and over, CDC found that 27 percent were never smokers, 35.1 percent were former smokers, and 37.9 percent were current smokers. Using the KP "person years" I find that the estimated proportions are very different: Among females 70.71 percent were never smokers and the remaining 29.29 percent were either current or former smokers; among males, 61.7 percent were never smokers and the remaining 38.3 percent were either current or former smokers (Friedman et al. 1997). Thus, the weighting of the KP population is done by applying information of the percent of the US population in each of three smoking status categories, never, former, and current by the CDC (US Center for Disease Control 1990) as described in endnote 6. Because the “current” and “former” smoking status categories were not broken down by the CDC in terms of years smoked or years since quitting (or by the number of cigarettes smoked by daily by those who were listed as current smokers), I distributed the proportions in the current and former categories to the same categories by age as found in the KP smoking study. For example, among females, CDC found that US females age 17 and over, 54.7 percent were never smokers, 20.9 percent were former smokers, and 24.4 percent were current smokers. Consequently, I weighted the KP e_x values for female never smokers by .547 and the female former smokers e_x values of those who quit 2-10 years ago by 0.209/3 = 0.0697, the female former smoker e_x values of those who quit 11-19 years ago by 0.209/3 = 0.0697, and the female former smoker e_x values of those who quit 20+ years ago by 0.209/3 = 0.0697. I weighted the e_x values of current KP female smokers who smoked for less than 20 years by 0.244/3 = 0.081333, the e_x values of current KP female smokers who smoked between 20 and 39 years by 0.244/3 = 0.081333, and the e_x values of current KP female smokers who smoked for 40 or more years by 0.244/3 – 0.081333 in each smoking status by age and sex (Friedman et al. 1997).

One of the reasons I selected the approach described by Swanson and Tedrow (2012) to develop the revised tables is that it allows us to use a simple adjustment such that the adjustment will modify them and their two functions (survivorship and life expectancy) in order to represent other populations, namely US populations . I use the term “scalar” adjustment for this process because I need apply only a single number to effect the desired changes so that the life tables approximate the Human Mortality database life tables for the US population for the any of the following relevant periods, 1985-89, 1990-94, 1995-1999, 2000-2004, 2005-2009, 2010-2014, 2015-2016. These periods are relevant because the first corresponds with the time period over which the KP smoking study data were gathered and the latter (especially 2000-2004, which corresponds to the 2000 US decennial census correspond to potential US populations that can be used in the population health modeling process, which is the final outcome desired for this project.

As an example of the scalar process, consider the adjustment of the life table functions (survivorship and life expectancy) for KP females by smoking status to US life table for females, 2000-2004. The approximation was accomplished by an iteration process that started with a scalar equal to 1.000 (no adjustment) and used successive scalar values to bracket the target life expectancy in the HMD life table for US females for 2000-2004 by applying the scalar to the KP ${}_nS_x$ values for females . The iterative process yielded a final scalar value of 0.9775, which was found by applying the successive scalar values to the ${}_nS_x$ values by smoking status and calculating the resulting mean life expectancy for all smoking status groups by weighting the e_x values using the proportion in each smoking status shown in exhibits 1 and 2.

Exhibit 1. Example of an adjustment of KP life tables for females by smoking status to serve as a life table for US females, 2000-2004

	ORIGINAL		ADJUSTED
II. ADJUSTMENT USING PROPORTION SMOKING IN US 1986	FEMALE PERSON YRS		FEMALE
WEIGHTED KP STUDY e55	**30.85**		**27.55**
smoking status	PROPORTION		PROPORTION
Never	0.547		0.547
Former, 20+ Yrs Since Quiting	0.069666667		0.069666667
Former, 11-20 Yrs Since Quitting	0.069666667		0.069666667
Former, 2-10 Yrs Since Quitting	0.069666667		0.069666667
Current, < 20 Yrs Duration	0.081333333		0.081333333
Current, 20-39 Yrs Duration	0.081333333		0.081333333
Current, 40+ Yrs Duration	0.081333333		0.081333333
	1		1
WEIGHTED KP STUDY e55 (1985-89)	30.85		27.55
HMD US 1985-89 e55	26.69		26.69
HMD US 1990-94 e55	27.17		27.17
HMD US 1995-99 e55	27.55		27.55
HMD US 2000-04 e55	27.55		27.55
HMD US 2005-09 e55	28.42		28.42
HMD US 2010-14 e55	29.04		29.04
HMD US 2015-16 e55	29.2		29.2
References			
Centers for Disease Control (1990). Smoking and Health: A National Status Report,			
2nd Edition: A Report to Congress. Public Health Service. USDHHS Publication no. 87-8369. Rockville, MD			
S Department of Health and Human Service			
Human Mortality Database. US Female Life Tables 5x5, 1933-2015			
Human Mortality Database. US Male Life Tables 5x5, 1933-2015			

Exhibit 2. Example of an adjustment of KP life tables for males by smoking status to serve as a life table for US males, 2000-2004

	ORIGINAL		ADJUSTED
II. ADJUSTMENT USING PROPORTION SMOKING IN US 1986	MALE PERSON YRS		mALE
WEIGHTED KP STUDY e55	**27.05**		**24.01**
smoking status	PROPORTION		PROPORTION
Never	0.27		0.27
Former, 20+ Yrs Since Quiting	0.117		0.117
Former, 11-20 Yrs Since Quitting	0.117		0.117
Former, 2-10 Yrs Since Quitting	0.117		0.117
Current, < 20 Yrs Duration	0.126333333		0.126333333
Current, 20-39 Yrs Duration	0.126333333		0.126333333
Current, 40+ Yrs Duration	0.126333333		0.126333333
	1		1
WEIGHTED KP STUDY e55 (1985-89)	27.05		24.01
HMD US 1985-89 e55	21.75		21.75
HMD US 1990-94 e55	22.57		22.57
HMD US 1995-99 e55	23.19		23.19
HMD US 2000-04 e55	24.02		24.02
HMD US 2005-09 e55	25.06		25.06
HMD US 2010-14 e55	25.71		25.71
HMD US 2015-16 e55	25.85		25.85
References			
Centers for Disease Control (1990). Smoking and Health: A National Status Report,			
2nd Edition: A Report to Congress. Public Health Service. USDHHS Publication no. 87-8369. Rockville, MD			
S Department of Health and Human Service			
Human Mortality Database. US Female Life Tables 5x5, 1933-2015			
Human Mortality Database. US Male Life Tables 5x5, 1933-2015			

Application of Results to Assess Population Health Impacts of Tobacco Products

Once the revised KP life table functions are adjusted to reflect the desired US population life table (i.e., the mortality of the 2000 US general population as found in the HMDB life US life tables for 2000-2004). The next step is to develop a model so that the needed survivorship rates can be fed into a population health model for the US (e.g., an agent based model). The process is grounded in the

Gompertz model, which views the human mortality rate increasing exponentially as a function of age (Greenwood, 1928). In its general form, the model is expressed as:

$$h(age)=B*c^{age} \quad [7]$$

where h(age) is the mortality hazard rate for the given age, and B and c are constants.

If I let $B=e^{\alpha}$ and $c= e^{\beta}$then the model translates into:

$$h(age)= e^{(\alpha+\beta*age)} \quad [8]$$

where α= intercept and β= the age coefficient

While mortality rates and doubling periods differ across various populations, the Gompertz model has been empirically validated over two centuries across a wide range of countries (Lee, Boscardin, Kirby, & Covinsky, 2014; Olshansky & Carnes, 1997). The model is widely used in various fields, including insurance and actuarial science (Bowers, Gerber, Hickman, Jones, & Nesbitt, 1997; Dickson, Hardy, & Waters, 2009) demography (Impagliazo, 1985; Smith & Keyfitz, 1977) and biology (Greenwood, 1928).

Previously, the Gompertz model was thought to be appropriate only thru around age 80, as mortality appeared to decelerate at higher ages. However recent research suggest otherwise, showing the Gompertz model applies up to low to mid 100's. Gavrilov et al. (2014) find that the mortality trajectory at advanced ages follows the Gompertz model up to the ages 102–105 years without a noticeable deceleration. The Gompertz model may be viewed as a sub-model under a broader range of risk models known as proportional hazards model, first proposed by David Cox (1972). The general model is given by:

$$h_i(t)=\lambda_0(t)*e^{(\beta_1*x_{i1}+\cdots+\beta_k*x_{ik})} \qquad [9]$$

For individual i at time t, the hazard is the product of: (1) the function $\lambda_0(t)$, which is left unspecified with the only condition that its value cannot be negative, and (2) the linear function $\beta_1 * x_{i1} + \ldots + \beta_k * x_{ik}$ comprised of k covariates that are then exponentiated. If for example one wanted to model US general population mortality starting in the year 2000, the revised KP life tables are not appropriate "as is" because they are not representative of the US year 2000 population. This is due to issues discussed earlier: (1) the KP study observed mortality was collected during the 1980's (2) the KP study population has different characteristics, particularly access to health care in that KP is both the insurer and the provider. Thus, the KP revised life tables reflect lower mortality rates than found in the general US population and would not be representative of that of ALCS ABM's target US population. The Human Mortality Database mortality tables provide the data for the required 2000 US population mortality experience. Using this approach, I created models applicable to US females and males in 2000. The input data for females are shown in Exhibit 1.a and the model in Exhibit 1.b. For the model applicable to US males, the input data are shown in Exhibit 2.a and the model in Exhibit 2.b.

Exhibit 1a. DATA INPUT FOR THE FEMALE COMPREHENSIVE GOMPERTZ MODEL FOR NQX USING 4 COVARIATES, AGE, SMOKING STATUS & AGE, US FEMALE POPULATION, 2000 (2000-04)

FEMALES								
ADJUSTED KP nSx	nqx (nqx = 1-nSx)	LN ($_{N}q_{x}$)	AGE	SMOKING STATUS	YRS SMOKED (OR SINCE LAST SMOKED)	AGE x YRS SMOKED	1	CURRENT SMOKER
0.9773953	0.0226047	-3.78960	20	0	0	0	-1	FORMER SMOKER
0.9771847	0.0228153	-3.78032	25	0	0	0	0	NEVER SMOKER
0.9767236	0.0232764	-3.76031	30	0	0	0		
0.9758366	0.0241634	-3.72292	35	0	0	0		
0.9742816	0.0257184	-3.66055	40	0	0	0		
0.9717395	0.0282605	-3.56629	45	0	0	0		
0.9678033	0.0321967	-3.43589	50	0	0	0		
0.9619685	0.0380315	-3.26934	55	0	0	0		
0.9536226	0.0463774	-3.07094	60	0	0	0		
0.9420347	0.0579653	-2.84791	65	0	0	0		
0.9263459	0.0736541	-2.60838	70	0	0	0		
0.9055589	0.0944411	-2.35978	75	0	0	0		
0.8785280	0.1214720	-2.10807	80	0	0	0		
0.9774940	0.0225060	-3.79397	20	-1	6	-6		
0.9774676	0.0225324	-3.79280	25	-1	6	-6		
0.9773710	0.0226290	-3.78852	30	-1	6	-6		
0.9770853	0.0229147	-3.77598	35	-1	6	-6		
0.9763597	0.0236403	-3.74480	40	-1	6	-6		
0.9747171	0.0252829	-3.67763	45	-1	6	-6		
0.9713181	0.0286819	-3.55149	50	-1	6	-6		
0.9647745	0.0352255	-3.34599	55	-1	6	-6		
0.9529005	0.0470995	-3.05549	60	-1	6	-6		
0.9323921	0.0676079	-2.69403	65	-1	6	-6		
0.8984221	0.1015779	-2.28693	70	-1	6	-6		
0.8441387	0.1558613	-1.85879	75	-1	6	-6		
0.7600545	0.2399455	-1.42734	80	-1	6	-6		
0.9774367	0.0225633	-3.79143	20	-1	15.5	-15.5		
0.9772835	0.0227165	-3.78466	25	-1	15.5	-15.5		
0.9769089	0.0230911	-3.76831	30	-1	15.5	-15.5		
0.9761183	0.0238817	-3.73464	35	-1	15.5	-15.5		
0.9746167	0.0253833	-3.67366	40	-1	15.5	-15.5		
0.9719833	0.0280167	-3.57496	45	-1	15.5	-15.5		
0.9676430	0.0323570	-3.43092	50	-1	15.5	-15.5		
0.9608364	0.0391636	-3.24001	55	-1	15.5	-15.5		
0.9505886	0.0494114	-3.00757	60	-1	15.5	-15.5		
0.9356753	0.0643247	-2.74381	65	-1	15.5	-15.5		
0.9145884	0.0854116	-2.46027	70	-1	15.5	-15.5		
0.8854993	0.1145007	-2.16717	75	-1	15.5	-15.5		
0.8462212	0.1537788	-1.87224	80	-1	15.5	-15.5		
0.9773489	0.0226511	-3.78755	20	-1	20	-20		
0.9770820	0.0229180	-3.77583	25	-1	20	-20		
0.9765296	0.0234704	-3.75202	30	-1	20	-20		
0.9755055	0.0244945	-3.70931	35	-1	20	-20		
0.9737524	0.0262476	-3.64018	40	-1	20	-20		
0.9709282	0.0290718	-3.53799	45	-1	20	-20		
0.9665907	0.0334093	-3.39892	50	-1	20	-20		
0.9601824	0.0398176	-3.22345	55	-1	20	-20		
0.9510139	0.0489861	-3.01622	60	-1	20	-20		
0.9382462	0.0617538	-2.78460	65	-1	20	-20		
0.9208733	0.0791267	-2.53670	70	-1	20	-20		
0.8977027	0.1022973	-2.27987	75	-1	20	-20		
0.8673366	0.1326634	-2.01994	80	-1	20	-20		
0.9774137	0.0225863	-3.79041	20	1	10.5	10.5		
0.9772219	0.0227781	-3.78195	25	1	10.5	10.5		
0.9767761	0.0232239	-3.76258	30	1	10.5	10.5		
0.9758749	0.0241251	-3.72450	35	1	10.5	10.5		
0.9742257	0.0257743	-3.65838	40	1	10.5	10.5		
0.9714258	0.0285742	-3.55525	45	1	10.5	10.5		
0.9669428	0.0330572	-3.40952	50	1	10.5	10.5		
0.9600936	0.0399064	-3.22122	55	1	10.5	10.5		
0.9500235	0.0499765	-2.99620	60	1	10.5	10.5		
0.9356846	0.0643154	-2.74396	65	1	10.5	10.5		
0.9158138	0.0841862	-2.47472	70	1	10.5	10.5		
0.8889098	0.1110902	-2.19741	75	1	10.5	10.5		
0.8532106	0.1467894	-1.91876	80	1	10.5	10.5		
0.9771221	0.0228779	-3.77758	20	1	29.5	29.5		
0.9765223	0.0234777	-3.75170	25	1	29.5	29.5		
0.9753745	0.0246255	-3.70397	30	1	29.5	29.5		
0.9734014	0.0265986	-3.62690	35	1	29.5	29.5		
0.9702615	0.0297385	-3.51531	40	1	29.5	29.5		
0.9655455	0.0344545	-3.36811	45	1	29.5	29.5		
0.9587743	0.0412257	-3.18869	50	1	29.5	29.5		
0.9493972	0.0506028	-2.98375	55	1	29.5	29.5		
0.9367891	0.0632109	-2.76128	60	1	29.5	29.5		
0.9202494	0.0797506	-2.52885	65	1	29.5	29.5		
0.8989998	0.1010002	-2.29263	70	1	29.5	29.5		
0.8721828	0.1278172	-2.05715	75	1	29.5	29.5		
0.8388603	0.1611397	-1.82548	80	1	29.5	29.5		
0.9483469	0.0516531	-2.96320	50	1	40	40		
0.9354641	0.0645359	-2.74053	55	1	40	40		
0.9187893	0.0812107	-2.51071	60	1	40	40		
0.8976650	0.1023350	-2.27950	65	1	40	40		
0.8713859	0.1286141	-2.05094	70	1	40	40		
0.8391997	0.1608003	-1.82759	75	1	40	40		
0.8003076	0.1996924	-1.61098	80	1	40	40		

Exhibit 1b. FEMALE COMPREHENSIVE GOMPERTZ MODEL FOR NQX USING 4 COVARIATES, AGE, SMOKING STATUS & AGE, US FEMALE POPULATION, 2000 (2000-04)

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.948948392
R Square	0.900503051
Adjusted R Square	0.895528204
Standard Error	0.220639936
Observations	85

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	4	35.24793331	8.811983327	181.0111893	3.02572E-39
Residual	80	3.894558502	0.048681981		
Total	84	39.14249181			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-4.875850538	0.077135785	-63.21126466	4.66553E-70	-5.029355643	-4.722345433	-5.029355643	-4.722345433
X Variable 1	0.033065319	0.001302638	25.3833487	5.01548E-40	0.030472986	0.035657652	0.030472986	0.035657652
X Variable 2	-0.099276455	0.063382377	-1.566310065	0.121223304	-0.225411404	0.026858495	-0.225411404	0.026858495
X Variable 3	0.004328614	0.002668883	1.62188189	0.108763444	-0.000982634	0.009639861	-0.000982634	0.009639861
X Variable 4	0.007554462	0.003378203	2.236236939	0.028121348	0.000831624	0.014277301	0.000831624	0.014277301

Variable

1 = age

2 = smoking status

3 = years smoked (+)/ Years since last smoked (-)

4 Age*Years smoked (+)/Years since last smoked (-)

The estimated equation is $\ln({}_nq_x)$ =-4.87585015861879 + 0.0330653163522129 * age - 0.0992771811245758 * smokingstatus + 0.00432862862681633 * years + 0.00755450916062156 * ageXyears

The model for females appears to be adequate, with the exception that multi-collinarity is present and affects the significance tests. It has a high coefficient of variation (R^2= 0.9005) and the adjusted value is R^2=0.8955, which suggests that there is no distortion of the goodness of fit due to a small sample size (n=85) relative to the number of independent variables (k=4). Two of the four coefficients are statistically significant at the .05 level, with years smoked (p= 0.1088) and smoking status (p = 0.1212) not statistically significant. One would expect both of these variable to be statistically significant and it is likely the case that they are not because each of them is highly correlated with age x years smoked: r= 0.90 for smoking status and age x years smoked; and r = 0.57 for years smoked and age x years. High correlations like these tends to inflate the standard errors estimated for their coefficients, which can lead to a determination that any or all are not statistically significant.

Given that the input data are not from a random sample, the fact that one or more of the coefficients is not statistically significant is not a huge concern. However, this correlation may have led to some level of inaccuracy in the estimation of the values of the coefficients for either or both of these variables, and, as such, degrade the predictability of the model. This suggests that it may be wise to omit the variable, age x years. In total, the diagnostic evaluation suggests that with the exception of multicollinearity, the model does not substantially violate the underlying assumptions of OLS regression models and is adequately specified. When, however, the variable age x years is removed, the indications of multicollinearity disappear without a noticeable decline in the coefficient of variation (R^2 =0.8953), which supports the use of this revised model:

$Ln({}_nq_x)$ =

-4.9068227324992 + 0.0332947010597161 * age + 0.0275251439138526 * smokingstatus + 0.0075854124754397 * years

Exhibit 2a. DATA INPUT FOR THE MALE COMPREHENSIVE GOMPERTZ MODEL FOR NQX USING 4 COVARIATES, AGE, SMOKING STATUS & AGE, US MALE POPULATION, 2000 (2000-04)

MALES								
ADJUSTED KP nSx	nqx (nqx = 1-nSx)	LN ($_{N}q_{x}$)	AGE	SMOKING STATUS	YRS SMOKED (OR SINCE LAST SMOKED)	AGE x YRS SMOKED	1	CURRENT SMOKER
0.9822230	0.0177770	-4.02985	20	0	0	0	-1	FORMER SMOKER
0.9817371	0.0182629	-4.00288	25	0	0	0	0	NEVER SMOKER
0.9807544	0.0192456	-3.95047	30	0	0	0		
0.9789852	0.0210148	-3.86253	35	0	0	0		
0.9760556	0.0239444	-3.73202	40	0	0	0		
0.9714993	0.0285007	-3.55783	45	0	0	0		
0.9647515	0.0352485	-3.34533	50	0	0	0		
0.9551418	0.0448582	-3.10425	55	0	0	0		
0.9418883	0.0581117	-2.84539	60	0	0	0		
0.9240918	0.0759082	-2.57823	65	0	0	0		
0.9007298	0.0992702	-2.30991	70	0	0	0		
0.8706514	0.1293486	-2.04524	75	0	0	0		
0.8325719	0.1674281	-1.78720	80	0	0	0		
0.9770568	0.0229432	-3.77473	20	-1	6	-6		
0.9763246	0.0236754	-3.74332	25	-1	6	-6		
0.9748921	0.0251079	-3.68457	30	-1	6	-6		
0.9723843	0.0276157	-3.58937	35	-1	6	-6		
0.9683297	0.0316703	-3.45237	40	-1	6	-6		
0.9621549	0.0378451	-3.27425	45	-1	6	-6		
0.9531794	0.0468206	-3.06143	50	-1	6	-6		
0.9406108	0.0593892	-2.82364	55	-1	6	-6		
0.9235400	0.0764600	-2.57099	60	-1	6	-6		
0.9009376	0.0990624	-2.31201	65	-1	6	-6		
0.8716495	0.1283505	-2.05299	70	-1	6	-6		
0.8343930	0.1656070	-1.79814	75	-1	6	-6		
0.7877538	0.2122462	-1.55001	80	-1	6	-6		
0.9822467	0.0177533	-4.03119	20	-1	15.5	-15.5		
0.9817937	0.0182063	-4.00599	25	-1	15.5	-15.5		
0.9808462	0.0191538	-3.95525	30	-1	15.5	-15.5		
0.9790681	0.0209319	-3.86648	35	-1	15.5	-15.5		
0.9759814	0.0240186	-3.72893	40	-1	15.5	-15.5		
0.9709313	0.0290687	-3.53809	45	-1	15.5	-15.5		
0.9630455	0.0369545	-3.29807	50	-1	15.5	-15.5		
0.9511897	0.0488103	-3.01981	55	-1	15.5	-15.5		
0.9339181	0.0660819	-2.71686	60	-1	15.5	-15.5		
0.9094191	0.0905809	-2.40151	65	-1	15.5	-15.5		
0.8754569	0.1245431	-2.08310	70	-1	15.5	-15.5		
0.8293077	0.1706923	-1.76789	75	-1	15.5	-15.5		
0.7676920	0.2323080	-1.45969	80	-1	15.5	-15.5		
0.9821484	0.0178516	-4.02566	20	-1	20	-20		
0.9816281	0.0183719	-3.99693	25	-1	20	-20		
0.9806338	0.0193662	-3.94423	30	-1	20	-20		
0.9788865	0.0211135	-3.85784	35	-1	20	-20		
0.9759942	0.0240058	-3.72946	40	-1	20	-20		
0.9714195	0.0285805	-3.55503	45	-1	20	-20		
0.9644429	0.0355571	-3.33662	50	-1	20	-20		
0.9541206	0.0458794	-3.08174	55	-1	20	-20		
0.9392374	0.0607626	-2.80078	60	-1	20	-20		
0.9182552	0.0817448	-2.50415	65	-1	20	-20		
0.8892564	0.1107436	-2.20054	70	-1	20	-20		
0.8498828	0.1501172	-1.89634	75	-1	20	-20		
0.7972689	0.2027311	-1.59587	80	-1	20	-20		
0.9815774	0.0184226	-3.99418	20	1	10.5	10.5		
0.9804523	0.0195477	-3.93490	25	1	10.5	10.5		
0.9785721	0.0214279	-3.84306	30	1	10.5	10.5		
0.9756868	0.0243132	-3.71673	35	1	10.5	10.5		
0.9715214	0.0284786	-3.55860	40	1	10.5	10.5		
0.9657772	0.0342228	-3.37486	45	1	10.5	10.5		
0.9581336	0.0418664	-3.17327	50	1	10.5	10.5		
0.9482483	0.0517517	-2.96130	55	1	10.5	10.5		
0.9357594	0.0642406	-2.74512	60	1	10.5	10.5		
0.9202853	0.0797147	-2.52930	65	1	10.5	10.5		
0.9014256	0.0985744	-2.31694	70	1	10.5	10.5		
0.8787620	0.1212380	-2.11000	75	1	10.5	10.5		
0.8518589	0.1481411	-1.90959	80	1	10.5	10.5		
0.9817980	0.0182020	-4.00622	20	1	29.5	29.5		
0.9807071	0.0192929	-3.94802	25	1	29.5	29.5		
0.9786430	0.0213570	-3.84637	30	1	29.5	29.5		
0.9751286	0.0248714	-3.69404	35	1	29.5	29.5		
0.9695814	0.0304186	-3.49270	40	1	29.5	29.5		
0.9613094	0.0386906	-3.25216	45	1	29.5	29.5		
0.9495083	0.0504917	-2.98595	50	1	29.5	29.5		
0.9332591	0.0667409	-2.70694	55	1	29.5	29.5		
0.9115253	0.0884747	-2.42504	60	1	29.5	29.5		
0.8831507	0.1168493	-2.14687	65	1	29.5	29.5		
0.8468573	0.1531427	-1.87639	70	1	29.5	29.5		
0.8012434	0.1987566	-1.61567	75	1	29.5	29.5		
0.7447814	0.2552186	-1.36563	80	1	29.5	29.5		
0.9369493	0.0630507	-2.76382	50	1	40	40		
0.9177752	0.0822248	-2.49830	55	1	40	40		
0.8933006	0.1066994	-2.23774	60	1	40	40		
0.8626890	0.1373110	-1.98551	65	1	40	40		
0.8250541	0.1749459	-1.74328	70	1	40	40		
0.7794619	0.2205381	-1.51168	75	1	40	40		
0.7249314	0.2750686	-1.29073	80	1	40	40		

Exhibit 2B. MALE COMPREHENSIVE GOMPERTZ MODEL FOR NQX USING 4 COVARIATES, AGE, SMOKING STATUS, YRS, & AGE BY YEARS, US MALE POPULATION, 2000 (2000-04)

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.979717457
R Square	0.959846296
Adjusted R Square	0.95783861
Standard Error	0.171856908
Observations	85

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	4	56.48069777	14.12017444	478.0860504	5.55205E-55
Residual	80	2.362783759	0.029534797		
Total	84	58.84348153			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-5.194625011	0.060081225	-86.46003847	8.80513E-81	-5.314190459	-5.075059563	-5.314190459	-5.075059563
X Variable 1	0.041879635	0.001014628	41.27586765	1.07031E-55	0.039860462	0.043898808	0.039860462	0.043898808
X Variable 2	-0.221450075	0.049368666	-4.485640231	2.40472E-05	-0.319696851	-0.123203298	-0.319696851	-0.123203298
X Variable 3	0.002517455	0.002078799	1.211014189	0.229455356	-0.001619487	0.006654397	-0.001619487	0.006654397
X Variable 4	0.014044315	0.002631289	5.337426912	8.58834E-07	0.008807882	0.019280748	0.008807882	0.019280748

Variable

1 = age

2 = smoking status

3 = years smoked (+)/ Years since last smoked (-)

4 Age*Years smoked (+)/Years since last smoked (-)

The estimated equation is $\ln({}_nq_x)$ = -5.19462495486153 + 0.0418796505958745 * age - 0.221450058632551 * smokingstatus + 0.00251742357642365 * years + 0.0140443131891837 * ageXyears

As was the case for the female model, the model for males appears to be adequate except for the presence of multi-collinearity, which affects the significance tests. It has a high coefficient of variation (R^2=0.9598) and the adjusted value is R^2=0.9578, which suggests that there is no distortion of the goodness of fit due to a small sample size relative to the number of independent variables. Three of the four coefficients are statistically significant at the .05 level, with years smoked being the only one that is not ($p= 0.2295$). One would expect this variable to be statistically significant and it is likely not because this variable is highly correlated with age x years smoked. ($r= 0.89$). High correlation like this between any two of the independent variables tends to inflate the standard errors estimated for their coefficients, which can lead to a determination that either or both are not statistically significant.

Given that the input data are not from a random sample, the fact that one or more of the coefficients is not statistically significant is not a huge concern. However, the presence of multicollinearity may also have led to some level of inaccuracy in the estimation of the values of the coefficients for either or both of these variables, and, as such, degrade the predictability of the model. This suggests that it may be wise to omit the variable, age x years. In sum, the diagnostic evaluation suggests that with the exception of multicollinearity, the model does not substantially violate the underlying assumptions of OLS regression models and is adequately specified. As was the case with the female Gompertz model when the variable age x years is removed, the indications of multicollinearity disappear without a noticeable decline in the coefficient of variation ($R^2 =0.9455$), which supports the use of this revised model for males:

$Ln({}_nq_x) =$

-5.2522049403514 + 0.0423060913618448 * age + 0.0142835314779996 * smokingstatus + 0.00857199196558667 * years

Conclusion

Following the path laid out in Abelin's seminal 1965 article, I have provided life tables from cohort mortality data widely employed in efforts to examine smoking and health, which in this case is the Kaiser Permanente Smoking Study. The mortality data in this study have been used in terms of relative mortality and risk rates in regard to smoking behaviors. However, they have never been used to generate life tables. After describing life tables in general, I described the KP smoking study data, then discussed the methods and procedures used to generate the life tables from them. I showed the life tables developed from the KP smoking study. I then discussed the methods used to extend these life tables to the US population and created hazard rate and survivorship data that can be used as input to models designed to assess the population health impact of tobacco products.

Acknowledgements

The author thanks Gary Friedman, Lai Wei, Thad Hannel, Raheema Muhammad-Kah, Thomas Bryan, and Simeon Chow for their comments and assistance.

Endnotes

1. To my knowledge, adequate life tables have not been constructed from any of the major cohort studies other than the work in this paper, which uses data from the KP Smoking Study. The other major cohort studies include Hammond-Horn (1958), CPS I, CPS II and CPS-3 (American Cancer Society, no date), the 16 cohorts in the seven country study (https://www.sevencountriesstudy.com/home/). There have, however, been studies that attempt to construct approximate life tables from other data sources, including Cancer Research UK Cancer Survival Group (2009), Ederer Axtell and Cutler (1961) and Ellis Coleman, and Rachet (2014). Blakeley and Wilson (2005) point out shortcomings in some of these approximations.

2. Friedman et al. (1997) provided a data set that addresses the mortality of current cigarette smokers, former cigarette smokers, and never smokers at various ages. These data were obtained from a KP Medical Care Program Cohort (KP) study. The study obtained baseline information on more than 60,000 subjects between 1979 and 1986, age 35 years and older and followed the cohort for mortality through 1987. The data set contains information about how long a person in the study has been a cigarette smoker and whether the person had quit smoking and if so how long the person has been a former cigarette smoker. The study subjects logged over 372,000 person-years, with 54% categorized as never smokers, 19% as former smokers and 26% as current smokers.

3. I employ here the definition that stochastic uncertainty is the manifestation of a process representing numerical values of some system randomly changing over time (Doob, 1953); and the definition of sampling uncertainty as an estimate attached to a test result that characterizes the range of values within which the true value is asserted to lie (Ramsey and Ellison, 2007). Stochastic uncertainty is far more likely to yield noticeable changes in a small population than in a large population (Swanson, forthcoming).

   I also believe that there is an unacknowledged confounding factor that may play a role in the inconsistencies by duration since quitting. This unacknowledged factor is length of time smoked (and or quantity smoked) before quitting. As an extreme example of this potentially confounding effect, consider those aged 35 who reported that they quit smoking more than 20 years ago, which implies that they would have started smoking at age 14 or younger. How long could somebody age 14 years or younger been smoking before they quit in order to state they had not smoked for 20 or more years? This potentially confounding factor may not play as large a role as small numbers, but it may contribute to the inconsistencies.

4. In addition to following the United Nations (1982) guidelines for constructing life tables from incomplete data, I also use principles developed by Green and Armstrong (2015) who compared the efficacy of complex forecasting models to simple ones, and found justification for using simple methods in that they suggest that clients who prefer accuracy should use forecasts based on simple evidence-based procedures. I believe that the adjustment methods I employ in this paper is consistent with their principles and advice.
5. The Human Mortality Database (HMD) is a collaborative project sponsored by the University of California at Berkeley and the Max Planck Institute for Demographic Research in Rostock, Germany. The purpose of the database is to provide researchers around the world with easy access to detailed and comparable national mortality data via the Internet. The database contains original life tables for almost 40 countries or areas, as well as all raw data used in constructing the tables. By design, populations in the HMD are restricted to those with data (both vital statistics and census information) that cover the entire population and that are very nearly complete. As such, the HMD covers almost all of Europe, plus Australia, Canada, Japan,

New Zealand, Chile, Israel, Taiwan and the United States. Outside this set, there are very few countries that possess the kind of data required for the HMD. Details are available at either of the following addresses: www.mortality.org and www.humanmortality.de (Wilmoth et al. 2017).

6. The US Centers for Disease Control (1990) estimated that at that time among: (1) US females age 17 and over, 54.7 percent were never smokers, 20.9 percent were former smokers, and 24.4 percent were current smokers; and (2) US males age 17 and over, 27 percent were never smokers, 35.1 percent were former smokers, and 37.9 percent were current smokers. Using the KP "person years," I find that the estimated proportions are very different: Among females 70.71 percent were never smokers and the remaining 29.29 percent were either current or former smokers; among males, 61.7 percent were never smokers and the remaining 38.3 percent were either current or former smokers (Friedman et al. 1997).

References

Ablin, T. (1965). Application of Life Table Methods to Results of Epidemiologic Follow-Up Studies on Smoking and Mortality. *American Journal of Epidemiology* 81 (2): 254-269.

American Cancer Society (no date), History of the cancer preventions studies (available online at https://www.cancer.org/research/we-conduct-cancer-research/epidemiology/history-cancer-prevention-study.html).

Bach, P., M. Kattan, M. Thornquist, M. Kris, R. Tate, M. Barnett, L. Hsieh, and C. Begg (2003). Variations in lung cancer risk among smokers. *Journal of the National Cancer Institute* 95 (6): 470-478.

Bachand, A. and S. Sulsky (2013). A dynamic population model for estimating all-cause mortality due to lifetime exposure history. *Regulatory Toxicology and Pharmacology* 67: 246-251.

Bachand, A., S. Sulsky, and G. Curtin (2018). Assessing the likelihood and magnitude of a population health benefit following the market introduction of a modified-risk tobacco product: Enhancements to the dynamic population modeler, DPM (+1). *Risk Analysis* 38 (1): 151-162.

Bain, C., D. Feskanich, F. Speizer, M. Thun, E. Hertzmark, B. Rosner, G. Colditz (2004). Lung cancer rates in men and women with comparable histories of smoking. *Journal of the American Cancer Institute* 96 (11): 826-834.

Bell, F. and M. Miller (2005). *Life tables for the United States Social Security Area, 1900-2100*. Actuarial Study no. 120. SSA Publication no. 11-11536. Office of the Chief Actuary, Social Security Administration. Washington, D.C.. U.S. Social Security Administration.

Blakely, T. and N. Wilson (2005). The contribution of smoking to inequalities in mortality by education varies over time and by sex: two national cohort studies, 1981–84 and 1996–99. *International Journal of Epidemiology* 34:1054–1062.

Blizzard, L. and T. Dwyer (2003). Case-control study of lung cancer during 1994-1997 in the birth cohort in Tasmania, Australia, with an excess of female cases during 1983-1992. *Cancer Causes and Control* 14: 123-129.

Bonnie, R., K. D. Stratton, and L Kwan (Eds). (2015). *Public Health Implications of raising the minimum age of legal access to tobacco products.* Institute of Medicine, Washington, D.C. The National Academies Press.

Bowers, N., H. Gerbe, J. Hickman, D. Jones, D., & C. Nesbitt. (1997). Actuarial Mathematics: Society of Actuaries.

Burch, T. (2018). *Model-Based Demography: Essays on integrating data, technique, and theory*. Springer. Dordrecht, The Netherlands.

Cancer Research, UK Cancer Survival Group (2009). *Life tables for England and Wales by sex, calendar period, region and deprivation.* London School of Hygiene & Tropical Medicine. http://www.lshtm.ac.uk/eph/ncde/cancersurvival/tools/index.html

Carey, J.R., R., Papadopoulos, H-G Müller, B. Katsoyannos, B., N. Kouloussis, J-L Wang, K. Wachter, W. Yu, and P. Liedo. (2008). Age structure and extraordinary life span in wild medfly populations. *Aging Cell* 7: 426-437.

Cox, D. R. (1972). Regression models and life tables (with discussion). *Journal of the Royal Statistical Society*. *Series B: Statistics in Society* 34(2), 33.

De Matteis, S., D. Consonni, A. Pesatori, A. Beren, P. Bertazzi, N. Caporaso, J. Lubin, S. Wacholder, and M. Landi (2012). Are women who smoke at higher risk for lung cancer than men who smoke? *American Journal of Epidemiology* 177 (7): 601-612.

Dickson, D., M. Hardy, & H. Waters. (2009). *Actuarial Mathematics for Life Contingent Risks*. Cambridge: Cambridge University Press

Doob, J. (1953). *Stochastic Processes*. New York, NY. John Wiley & Sons.

Dorn, H. (1959). Tobacco consumption and mortality from cancer and other diseases. *Public Health Reports* 74: 581-593.

Ederer F., L. Axtell, and S. Cutler. (1961). *The relative survival rate: a statistical methodology*. National Cancer Institute Monograph (6):101–121.

Ellis, L. M. Coleman, and B. Rachet. (2014). The impact of life tables adjusted for smoking on socio-economic difference in net survival for laryngeal and lung cancer. *British Journal of Cancer* 111(1): 195-202.

Impagliazo, J. (1985). *Deterministic Aspects of Mathematical Demography: An Investigation of the Stable Theory of Population Including an Analysis of the Population Statistics of Denmark*: Springer-Verlag.

Ernster, V., N. Kaufman, M. Nichter, J. Samet, and S. Yoon (2000). Women and Tobacco: Moving from policy to action. *Bulletin of the World Health Organization* 78 (7): 891-901.

Fergany, N. (1971). On the human survivorship function and life table construction. *Demography* 8 (3): 331-334.

Friedman, G. (2019). Personal communication.

Friedman G., I. Tekawa, M. Sadler, and S. Sidney (1997). Smoking and mortality: The Kaiser Permanente Experience. pp. 477–499 in D. Shopl and D. Burns, L. Garfinkel, J. Samet (eds). *Changes in Cigarette-Related Disease Risks and Their Implication for Prevention and Control.* Rockville, MD: U.S. Department of Health and Human Services, Public Health Service, National Institutes of Health, National Cancer Institute.

Fry, j., P. Lee, B. Forey, and K. Coombs (2013). Dose-response relationship of lung cancer to amount smoked, duration, and age starting. *World Journal of Meta-Analysis* 1 (2): 57-77.

Gavrilov, L. N. Gavrilova, C. Stone, and A. Zissu. (2014). New Findings on older people's life expectancies confirm Gompertz Law: The impact on the value of securitized life settlements. *Journal of Structured Finance* 20(2) 66-73.

Green, K. and J. S. Armstrong (2015). Simple versus complex forecasting: The evidence. *Journal of Business Research* (68): 1678-1685.

Greenwood, M. (1928). "Laws" of Mortality from the Biological point of view. *Journal of Hygiene* 28(3), 267-294.

Hammond, E. (1964). Smoking in relation to mortality and morbidity. Findings in the first 34 months of follow-up in a prospective study started in 1959. *Journal of the National Cancer Institute* 32: 1161-1188.

Hammond, E. and D. Horn (1958) Smoking and death rates—report on forty-four months of follow-up of 187,783 men. I. Total mortality. II. Death rates by cause. *Journal of the American Medical Association 166:* 1159-1172; 1294-1308.

Hunt, D., T. Blakely, A. Woodward, and N. Wilson (2005). The smoking-mortality association varies over time and by ethnicity in New Zealand. *International Journal of Epidemiology* 24: 1020-1028.

Jacobs, D., H. Adachi, I. Mulder, D. Kromhout, A. Menotti, A. Nissinen, H. Blackburn (1999). Cigarette smoking and mortality risk: Twenty-five-year follow-up of the seven countries study. *Archives of Internal Medicine* 159: 733-740.

Jiang, J., B. Liu, F. Sitas, J. Li, X. Zeng, W. Han, X. Zou, Y. Wu, and P. Zhao (2010). Smoking-attributable deaths and potential years of life lost from a large, representative study in China. *Tobacco Control*: 19: 7-12. doi:10.1136/tc.2009.031245

Keyfitz, N. 1970. Finding Probabilities from Observed Rates or How to Make a Life Table. *The American Statistician* 24 (1): 28-33.

Kim, Y. and J. Aron. (1989). On the equality of average age and average expectation of remaining life in a stationary population. *SIAM Review* 31 (1): 110-113.

Kintner, H. (2004). The Life Table. pp. 301-340 in J. Siegel and D. Swanson (Eds.) *The Methods and Materials of Demography, 2nd Edition.* San Diego, CA: Elsevier Academic Press.

Lee, P., and B. Forey (2013). Indirectly estimated absolute lung cancer mortality rates by smoking status and histology type based on a systematic review. *BMC Cancer* 13: 1-90.

Lee, P., B. Forey, and K. Coombs (2012). Systematic review with meta-analysis of the epidemiological evidence in the 1900s relating smoking to lung cancer. *BMC Cancer* 12: 385

Lee, P., J. Hamling, J. Fry, and B. Forey (2105). Using the negative exponential model to describe changes in risk of smoking-related diseases following changes in exposure to tobacco. *Advances in Epidemiology* (ID 487876).

Lee, S. J., W. Boscardin, K. Kirby, & K. Covinsky. (2014). Individualizing life expectancy estimates for older adults using the Gompertz Law of Human Mortality. *PLoS ONE* 9(9), e108540. doi:10.1371/journal.pone.0108540

Levy, D., E. Mumford, and D. Gerlowski (2007). Examining trends in quantity smoked. *Nicotine & Tobacco Research* 9 (11): 1287-1296.

Moylan, S., K. Gustavson, E. Karevold, S. Øverland, F. Jacka, J. Pasco, and M. Berk. (2013). The impact of smoking in adolescence on early adult anxiety symptoms and the relationship between infant vulnerability factors for anxiety and early adult anxiety symptoms: The TOPP study. *PLOS one*. https://doi.org/10.1371/journal.pone.0063252.

Muhammad-Kah, R., Y. Pithawalla, M. Gogova, L. Wei, and E. Boone (2016). An Agent-Based Modeling Approach for Tobacco Product Risk Assessments. Paper presented at the Joint Statistical Meeting, Chicago, Illinois.

Olshansky, S. J., & B. Carnes. (1997). Ever since Gompertz. *Demography* 34(1), 1-15.

Owen, A., S. Maulida, E. Zomer, and D. Liew (2018). Productivity burden of smoking in Australia: a life table modelling study. *Tobacco Control*: 1-8., doi:10.1136/tobaccocontrol-2018-054263

Poland, B., and F. Teischinger (2017). Population modeling of modified risk tobacco products accounting for smoking reduction and gradual transitions of relative risk. *Nicotine & Tobacco Research* 19 (11): 1277-1283

Preston, S. (1970a). An international comparison of excessive adult mortality. *Population Studies* 24:5-20. .

Preston, S. 1970b. *Older Male Mortality and Cigarette Smoking*. Berkeley: University of California Press.

Ramsey, M. and S. Ellison (Eds.) (2007). *Eurachem/EUROLAB/CITAC/Nordest/AMC Guide: Measurement uncertainty arising from sampling, a guide to methods and approaches*. Tornio, Italy. Eurachem Secretariat.

Rao, A., and J. Carey. (2014). Generalization of Carey's equality and a Theorem on stationary population." *Journal of Mathematical Biology* DOI 10.1007/s00285-014-0831-6. http://entomology.ucdavis.edu/files/203430.pdf

Retherford, R. (1972). Tobacco smoking and the sex mortality differential. *Demography* 9 (2): 203-216.

Richards, H. and M. Donaldson (2010). *Life and worklife expectancies*. Lawyers and Judges Publishing. Tucson, AZ.

Siegel, J, (2002).*Applied Demography*. Academic Press. San Diego, CA.

Smith, D., &. Keyfitz. (1977). *Mathematical demography: selected papers*. Springer-Verlag.

Smith, S. J. Tayman, and D. Swanson. (2013). *A practitioner's guide to state and local population projections*. Springer. Dordrecht, The Netherlands.

Springborn, R. (2006). *Abridged Life Tables for California, 2004*. California Center for Health Statistics (https://www.dhcs.ca.gov/services/ltc/documents/lifetables2004.pdf).

Swanson, D. A. (forthcoming). "Estimating the underlying infant mortality rates for small populations, even those reporting zero infant deaths: A case study of 66 local health areas in British Columbia." *Canadian Studies in Population.*

Swanson, D. and A. Sanford (2012). Socio-Economic Status and Life Expectancy in the United States, 1990-2010: Are We Reaching the Limits of Human Longevity? *Population Review* 51 (2): 16-41.

Swanson D., and L. Tedrow (2012). Using cohort change ratios to estimate life expectancy in populations with negligible migration: A new approach." *Canadian Studies in Population* 39: 83-90.

Swanson, D., and L. Tedrow (2019). On Mathematical Equalities and Inequalities in the Life Table: Something Old and Something New. *Canadian Studies in Population* (forthcoming).

Thomas, J. and L. Bao (2016). Modeling the dynamics of an HIV epidemic. pp. 91- 114 in R. Schoen (Ed.) *Dynamic Demographic Analysis.* Springer. Dordrecht, The Netherlands.

Thun, M., E. Calle, C. Rodriguez, and P. Wingo (2000). Epidemiological research at the American Cancer Society. *Cancer Epidemiology, Biomarkers and Prevention* 9: 861-868.

Thun, M., S. Henley, D. Burns, A. Jemal, T. Shanks, and E. Calle (2006). Lung cancer rates in lifetime nonsmokers. *Journal of the National Cancer Institute* 98 (10): 691-699.

Trovato, F. and N. M. Lalu (2001). Narrowing sex differences in life expectancy: Regional variations, 1971-1991. *Canadian Studies in Population* 28 (1): 89-110.

United Nations (1982). *Model Life Tables for Developing Countries*. Department of International Economic and Social Affairs, Population Studies no. 77. United Nations. New York, New York.

United States Centers for Disease Control (1990). *Smoking and Health: A National Status Report 2nd Edition: A Report to Congress.* Public Health Service. USDHHS Publication no. 87-8369. Rockville, M Department of Health and Human Service.

Van Iersel, C., H de Koning, G. Draisma, W. Mali, E. Scholtem, K. Nackaerts, M. Prokop, J. Habbema, M. Oudkerk, and R. van Klaveren (2006). Risk-based selection from the general population in a screening trial: Selection criteria, recruitment and power for the Dutch-Belgian randomized lung cancer multi-slice CT screening trial (NELSON). *International Journal of Cancer* 120: 868-874.

Vaupel, J. (2009). Life lived and left: Carey's equality. (2009). *Demographic Research* 20: 7-10.

Vaupel, J. and T. Missov (2014). Unobserved population heterogeneity: A review of formal relationships *Demographic Research* 31 (22): 659-686.

Villavicencio, F., and T. Riffe. (2016). Symmetries between life lived and left in finite stationary populations. *Demographic Research* 35: 381-398.

Wilmoth, J., K. Andreev, D. Jdanov, D.A. Glei and T. Riffe with the assistance of C. Boe, M. Bubenheim, D. Philipov, V. Shkolnikov, P. Vachon, C. Winant, M. Barbieri (2017). Methods Protocol for the Human Mortality Database (Version 6). Available online at www.mortality.org

Wrycza, T. (2014). Variance in age at death equals average squared remaining life expectancy at death. *Demographic Research* 30 (50): 1405-1412

Yashin, A., E. Stallard, and K. Land (2016). *Biodemography of Aging: Determinants of Healthy Life Span and Longevity*. Springer. Dordrecht, The Netherlands.

Yusuf, F., J. Martins, and D. Swanson (2014). *Methods of Demographic Analysis*. Springer. Dordrecht, The Netherlands.