Constructing Life Tables from the Kaiser Permanente Smoking Study and Applying the Results to Models Designed to assess the Population Health Impact of Reduced Risk Tobacco Products.

David Swanson

[2] **University of California Riverside (Professor Emeritus), Riverside, CA USA**
and Center for Studies in Demography and Ecology (Faculty Affiliate),
University of Washington, Seattle, WA USA
email: dswanson@ucr.edu

OUTLINE

1. INTRODUCTION
2. THE KAISER PEREMAMENTE (KP) SMOKING STUDY DATA
3. OVERVIEW OF PROCESS USED TO GENERATE PRELIMINARY KP LIFE TABLES
4. PRELIMINARY KP LIFE TABLES
5. DISCUSSION OF THE PRELIMINARY KP LIFE TABLES
6. OVERVIEW OF PROCESS OF CREATING FINAL KP LIFE TABLES
7. FINAL KP LIFE TABLES
8. APPLYING KP LIFE TABLE RESULTS TO THE GENERAL US POPULATION

INTRODUCTION

Following the path laid out in Abelin's seminal 1965 article, we construct life tables from cohort mortality data widely employed in efforts to examine smoking and health, which in this case is the Kaiser Permanente (KP) 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. We describe the KP smoking study data, then provide an overview of the methods used to generate the life tables from them. Following the description and overview, we show an illustration of the life tables developed from the KP smoking study using two females (1) never smokers; and (2) current smokers, 20+ cigarettes daily.

We 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 reduced risk tobacco products.

KAISER PERMANENTE (KP) SMOKING STUDY DATA

The mortality data from the KP smoking study we use as input for the initial life tables are provided by Freidman 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, but age groups are not provided, which precludes the construction of life tables from these data. Where mortality rates for all causes are provided by race, gender, and age group, the small numbers preclude the construction of life tables from these data.

OVERVIEW OF PROCESS USED TO GENERATE PRELIMINARY KP LIFE TABLES

We employ a conversion formula that assumes that deaths occur in increasing numbers within a given age interval, specifically in an exponential manner (Fergany, 1971):

$${}_{n}q_{x} = 1\text{-}e\text{^}(-n * {}_{n}m_{x})$$

where,

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

and where e ≈ 2.71828

OVERVIEW OF PROCESS USED TO GENERATE PRELIMINARY KP LIFE TABLES

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 any-way) 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 available for the KP mortality data (Freidman et al. 1997).

Preliminary KP 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): $_nq_x = 1-e^{(-n*nmx)}$ 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
35-49	Never	0	45,786	37	0.00081	15.0	0.01205	97,807	1,178	1,458,249	4,673,556	47.78
50-64	Never	0	49,744	118	0.00237	15.0	0.03496	96,629	3,378	1,423,945	3,215,307	33.27
65-74	Never	0	24,159	171	0.00708	10.0	0.06833	93,251	6,372	900,271	1,791,363	19.21
75+	Never	0	12,285	299	0.02434	11.8	0.24964	86,879	21,688	891,092	891,092	10.26

Preliminary KP 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
35-49	Curent	20+	12,851	25	0.00195	15.0	0.02876	95,741	2,753	1,415,364	4,135,909	43.20
50-64	Curent	20+	10,950	69	0.00630	15.0	0.09019	92,988	8,387	1,330,923	2,720,545	29.26
65-74	Curent	20+	3,583	70	0.01954	10.0	0.17747	84,601	15,014	768,497	1,389,621	16.43
75+	Curent	20+	588	24	0.04082	11.1	0.36432	69,587	25,352	621,124	621,124	8.93

DISCUSSION OF PRELIMINARY KP LIFE TABLES

As a starting point for discussing the preliminary life tables, we 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.

DISCUSSION OF PRELIMINARY KP LIFE TABLES

For females, an anomaly, contrary to our 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).

DISCUSSION OF PRELIMINARY KP LIFE TABLES

For males, anomalies contrary to our 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).

DISCUSSION OF PRELIMINARY KP LIFE TABLES

In the original KP smoking study publication (Friedman et al., 1997), there are acknowledgements to the anomalies found in our 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).

OVERVIEW OF PROCESS USED TO GENERATE FINAL KP LIFE TABLES

Because of the widespread use of the KP Smoking Study data, we believe that it is worth the effort to resolve the anomalies identified here. To this end, we first interpolate the hazard rates (${}_nq_x$ values) found in the preliminary life tables so that we have a set of hazard rates for age groups of five –year widths, starting at age 35-40 and ending at age 80-85.

OVERVIEW OF PROCESS USED TO GENERATE FINAL KP LIFE TABLES

We 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 our 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.

OVERVIEW OF PROCESS USED TO GENERATE FINAL KP LIFE TABLES

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 e_{35}/KP e_{35}), was formed for all smoking groups employed in the KP Smoking Study. These adjusted hazard rates were then graphed and examined for anomalies.

OVERVIEW OF PROCESS USED TO GENERATE FINAL KP LIFE TABLES

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.

OVERVIEW OF PROCESS USED TO GENERATE FINAL KP LIFE TABLES

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. 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.

OVERVIEW OF PROCESS USED TO GENERATE FINAL KP LIFE TABLES

Because nqx directly generates l_x and ${}_nd_x$, and in combination with nmx 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 from the inconsistencies we noted using life expectancy values.

OVERVIEW OF PROCESS USED TO GENERATE FINAL KP LIFE TABLES

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 l_x, the denominator used to create e_x, is the result of the cumulative subtraction of ${}_nd_x$ values from preceding l_x values. 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, where ${}_nS_x = 1 - {}_nq_x$).

Thus, In terms of resolving the 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, we can take the average ${}_nq_x$ at each age between females who never smoked and those who quit smoking 20+ years ago.

OVERVIEW OF PROCESS USED TO GENERATE FINAL KP LIFE TABLES

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 nqx values via their reciprocals, nSx values. We can do this because ${}_nS_x = 1 - {}_nq_x$. We 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:

$$e_0 = S_0 + (S_0 * S_1) + (S_0 * S_1 * S_2) +,\dots,+(S_0 * S_1 * S_2,\dots,*S_x)$$

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

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

$$e_0 = {}_nS_0 + ({}_nS_0 * {}_nS_1) + ({}_nS_0 * {}_nS_1 * {}_nS_2) +,\dots,+({}_nS_0 * {}_nS_1 * {}_nS_2,\dots,* {}_nS_x)$$

Final Life Table for KP 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

Final Life Table for KP Females, Current Smokers, 20+ 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

Final 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 FINAL KP LIFE TABLES

The preceding table 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.

DISCUSSION OF FINAL KP LIFE TABLES

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.

Given the two remaining anomalies for males, we 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 we 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. 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.

APPLYING KP LIFE TABLE RESULTS TO THE GENERAL US POPULATION

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			

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 ($_nq_x$)	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		

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

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

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 $\ln({}_{n}q_{x})$ =-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 multicollinarity is present and affects the significance tests

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

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