

“You’ve come a long way, baby”: The convergence in age patterns of lung cancer mortality by sex, United States, 1959–2013

Natalie A. Rivadeneira*
Andrew Noymer†

16 November 2016

Abstract

We analyze lung cancer mortality by age and sex in the United States, 1959–2013. It is already known that male lung cancer death rates exceed those of women and that tobacco use is the leading reason for the sex difference. We elaborate on this knowledge by showing that unlike most causes of death, lung cancer mortality patterns by age are a very good fit to a quadratic-Gompertz model, i.e., log mortality rates are quadratic by age, with the mode above age 70. With a little additional historical data on sex differences in tobacco use, the quadratic models help to paint a clear quantitative picture of behavior-led convergence in lung cancer mortality by sex. The shape of the sex-specific mortality curves have converged dramatically: since 1983, the sexes have had statistically-indistinguishable shapes of their quadratic-Gompertz mortality curves. Female lung cancer mortality patterns have shown a transformation from a non-smoking to a smoking pattern. The modal age of the quadratic-Gompertzian lung cancer death rates has potential application in countries at earlier stages of the global tobacco epidemic. Where the modal age is falling (year-on-year), we can infer that the local lung cancer epidemic is getting worse, and where it is rising, that it has begun to abate.

*Rollins School of Public Health, Emory University, Atlanta, Georgia.

†Department of Population Health and Disease Prevention, University of California, Irvine. To whom correspondence should be addressed. noymer@uci.edu

Introduction

The goal of this paper is quantitatively to characterize patterns of lung cancer mortality by age and sex in the United States, from 1959 to the present, the period of availability of detailed cause-specific mortality data. We then relate these patterns, qualitatively, to historical sex-specific changes in tobacco use. This results in two principal findings. First, lung cancer mortality is, very clearly, quadratic-Gompertz: when log mortality rates are plotted against age, the lung cancer pattern is a quadratic curve with a negative-valued squared term. This pattern holds for both sexes, and over time, although there are some differences, which we analyze. The modal age of the quadratic pattern plays an important role in our analysis. We suggest that this quantity, the modal age, can be used as an informative unidimensional summary of tobacco's influence on patterns of lung cancer mortality.

Our second principal finding pertains to sex differences. The data since 1959 show three distinct patterns, as follows. In the first pattern (roughly, 1959–64), male lung cancer mortality is characteristic of tobacco use, while the female pattern, which peaked much older, was more representative of the non-smoking, or background, lung cancer mortality rates. In the second period (roughly, 1965–82), the female pattern converges to the male pattern, undergoing a transition from background to smoking-related lung cancer mortality. Since 1983, the *pattern* (but not the level) of lung cancer mortality has converged between the sexes, with both males and females experiencing a smoking-like pattern of lung cancer mortality, with modal rates at age 80–89. However, both sexes are slowly transforming back to background lung cancer mortality patterns (with upward shifts in the modal age of the death rate) as smoking prevalence declines. Needless to say, the connection between smoking and lung cancer mortality is not novel (f.e., Hammond and Horn 1954). However, the present analysis of long-run changes in lung cancer mortality sex differences casts new light on sex differences in cause-specific mortality. This is the first analysis of which we are aware that makes the point that there has been a convergence in lung cancer age-mortality patterns between men and women (note, patterns not levels; see Pampel 2002, 2003 for important work on the convergence of levels).

It is difficult to analyze lung cancer mortality without considering tobacco use (f.e., Holford et al. 2014*b*, Siegel et al. 2015). Figure 1 presents age-standardized sex-specific data on smoking prevalence in the United States, 1933–2005 using data from Burns et al. (1997) and Holford et al. (2014*a*). A number of patterns are clear. In the United States, men smoke more than women, and always have — although the prevalence is converging. On an

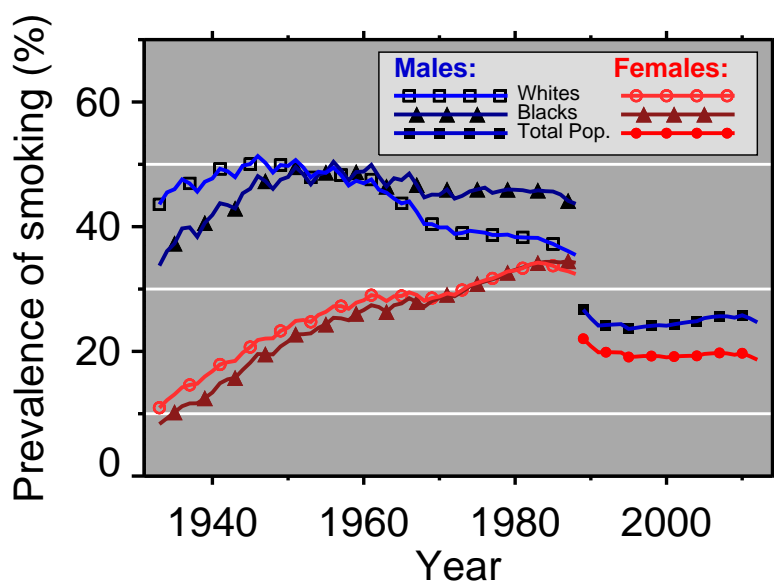


Figure 1: Age-standardized prevalence of tobacco smoking, USA, 1933–2012. Underlying data from Burns et al. (1997) (for 1933–88, by race and sex) and Holford et al. (2014a) (for 1989–2012, by sex). The two data series do not line-up perfectly but are reasonably congruent.

age-standardized basis, male smoking rates peaked in mid-century while female rates peaked later, in the early 1980s (albeit at a lower level). Women adopted smoking later in time, and female cigarette use was still growing at the time of the Surgeon General’s report (US Public Health Service 1964), a turning point in tobacco use among males in the United States. Moreover, in absolute terms, smoking declined more slowly among women than men (figure 1).

Data & Methods

Using mortality multiple cause of death (MCD) data from the US National Center for Health Statistics (2014), we extracted counts, by age and sex, of all lung cancer deaths (underlying cause; the specific ICD codes are given in Appendix I). The time span is 1959 to 2013, the full extent of mortality microdata availability. Death counts were converted to rates using person-years at risk (PYAR) from the Human Mortality Database (2015). Five-year groups (40–44, 45–49, ..., 95–99) were used to smooth heaping on preferential digits of age, as is common practice (f.e., Ho and Fenelon 2015). The exposure (i.e., PYAR) data from the Human Mortality Database are an appropriate match for the MCD data, since the MCD data, collapsed to all causes, is the same as the HMD mortality count data (the HMD does not, as of this writing, include cause-specific data). Hereinafter, rate always refers to sex-, age-, and cause-specific death rates.

We estimated quadratic-Gompertz models, separately by sex, by regressing logged death rates on age and age-squared, with a constant. The quadratic specification was chosen after inspection of the data; the canonical Gompertz model with only an intercept and slope clearly doesn’t fit the data. The regression equation is:

$$\log(M_x^{\text{LC}}) = \alpha + \beta_1 x + \beta_2 x^2 + \epsilon, \quad (1)$$

where M_x^{LC} are lung cancer sex- and age-specific death rates, x is age, α , β_1 , β_2 are parameters to be estimated, and ϵ is the error term. See also Yashin et al. (2012); equation 1 is a simplified version of the models therein. The age range for this analysis was 40 to 99. Below age 40 and above age 100, lung cancer deaths are very rare and thus subject to high variability, and are non-Gompertzian (even with the quadratic adjustment). The regressions are OLS weighted by the number of deaths. For example, in 2010, there were 14,094 lung cancer deaths among men age 70–74, and these were used as weights. Weighting has two advantages over unweighted regressions. Death rates at

the age limits (i.e., 40–44, 95–99) are typically the poorest-fitting points in a Gompertzian pattern. Removing these points does not really help, since the estimates then change, and 45–49 also becomes a poor fit, and so on. Weighting solves this problem because there are fewer deaths at the age limits, so these observations are down-weighted, reducing their leverage. The weighting by deaths also produces estimates that are closer to those that would be obtained by maximum likelihood (Abdullatif and Noymer 2016, pp. 206–7). Moreover, we do not see any disadvantage of using weights. All analyses were performed with Stata, version 13.1 (StataCorp LP, College Station, Texas).

Results and Discussion

Figure 2 presents lung cancer age-mortality profiles and fitted quadratic-Gompertz curves, for the first year of each decade (1960–2010), for both sexes (see Appendix II for all years, 1959–2013). Plotting symbol sizes are proportional to the absolute number of deaths. In 1960, men and women had distinct age-mortality profiles for lung cancer. Male death rates were considerably higher than females, at all ages. The fitted quadratic-Gompertz curves cross at age 95, implying convergence at that age, but the empirical data diverge from the fit at oldest ages, and males are always higher. The male death rates in 1960 have their mode at ages 70–74, whereas the female death rates are monotonic up to a mode at age 90–94 (this refers to the fitted pattern; the empirical rate falls a little in the 95–99 age group, but this is calculated from only 4 deaths).

Unlike all-cause mortality, or mortality from a number of specific causes (f.e., heart disease), lung cancer in the presence of tobacco smoking is non-Gompertzian (i.e., unless a quadratic term is used). Lung cancer mortality does not just keep going up with age, but reaches a mode and then declines (Horiuchi and Wilmoth 1997), as is the case with incidence (Dix et al., 1980). This refers to period data; Manton et al. (1986) and Manton et al. (2009) present some analysis of smoking-related mortality by age, holding cohort constant. The reason the mode lies inside the age range of human lifespan is thought to be frailty — i.e., that mortality selection in the 70s and 80s (of age) means that the heaviest smokers don't live into their 90s. In addition, cohort differences in tobacco use affect the observed patterns — for example, the male 1910 birth cohort were heavier smokers than men born in 1900 (Moolgavkar et al. 2012, Christopoulou and Lillard 2016). Moreover, smokers have higher non-lung cancer (Jacobs et al., 2015) and all-cause mortality

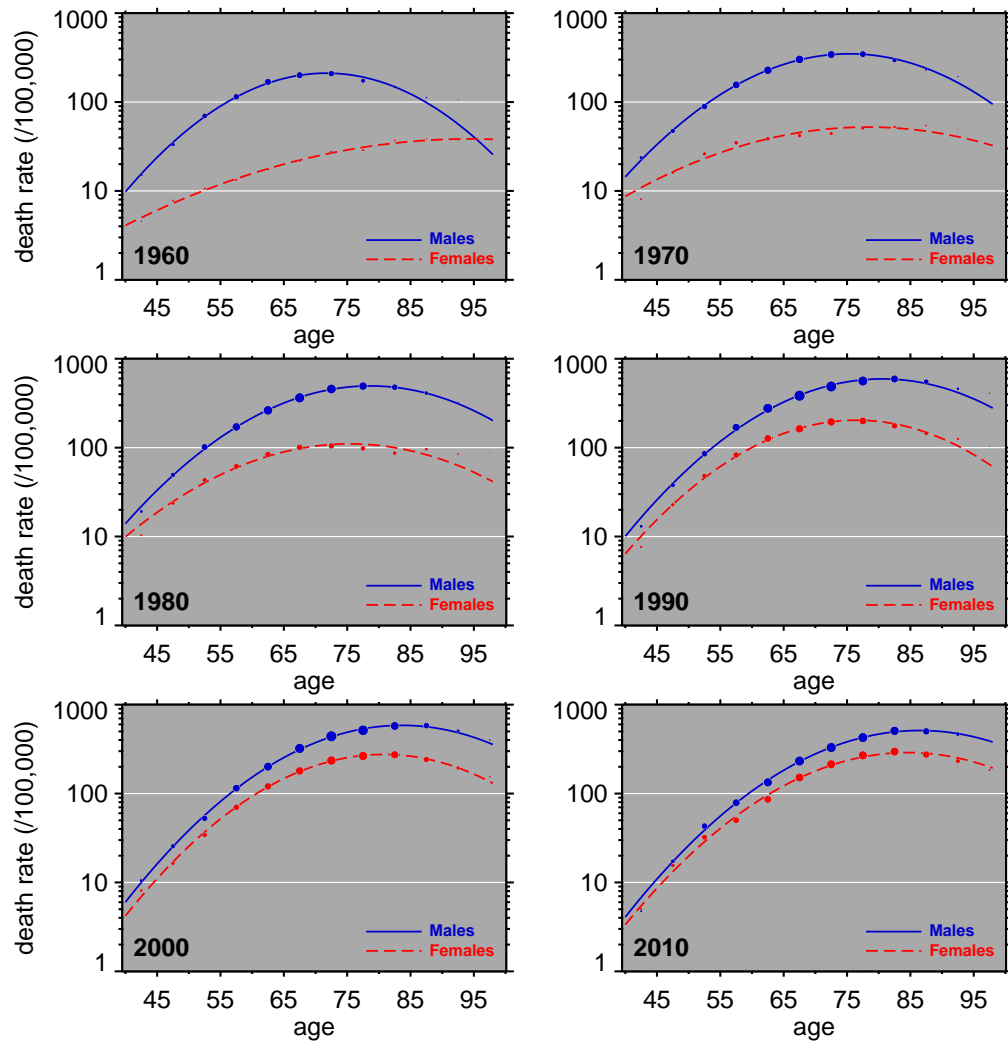


Figure 2: Age-mortality profiles, by sex. Lung cancer, USA, 1960–2010 (decennial). With quadratic-Gompertz fit.

(Rogers et al. 2005, Jha et al. 2013, Banks et al. 2015, Carter et al. 2015), so the forces of out-selection of smokers from the population at risk of lung cancer death are complex.

Given both the delayed effect of smoking on mortality, and the later adoption of smoking by women, the pattern of mortality seen in 1960 for women is much closer to a “background” mortality rate for lung cancer (Preston and Wang 2006). This is not only because the female rates are lower (the peak male lung cancer death rates were about *ten times* higher than those of women in 1960), but because they have their mode at an older age. Because female smoking was not zero, the female pattern cannot be said to be a perfect background mortality rate, but it is much closer to it than are males. Exposure to secondhand smoke among females is another reason it is not a true background rate, although the role of secondhand smoke in lung cancer is not clear (Boffetta et al. 1998, Jöckel et al. 1998).

Figures 3 (for 1960) and 4 (for 2005) show how unusual the quadratic-Gompertz pattern is, relative to other causes, across both sexes, and over time. All-cause mortality is Gompertzian, being linear on log scale. As are cardiovascular causes, and influenza and pneumonia (as a combined cause). This is true for both sexes, and in both time periods. Lung cancer, as discussed, shows a quadratic pattern. The data series for all cancers (excluding lung cancer) show a somewhat curvilinear pattern but nothing nearly as pronounced as the pattern for lung cancer. Moreover, in the 2005 data (figure 4) while the data series for all cancer curves down somewhat at older ages, the mode lies above age 99; i.e., the modal age is not realized in the data, unlike for lung cancer. In 1960, diabetes mellitus shows a curved pattern with a mode inside the data range, albeit with an older mode than lung cancer, especially for males (figure 3). However, by 2005, the pattern of diabetes has changed: while it still shows some curved departure from straight-line Gompertzian mortality, it increases monotonically up to age 99. The increases in the levels of diabetes mortality in the last 50 years (see, f.e., Gregg et al. 2012, Menke et al. 2015) have changed its pattern as well as its level — and not for the better: diabetes death rates no longer drop off at older ages as they once did, although competing risks and classification issues (discussed below) may also play a role. Figures 3 and 4 show that lung cancer’s quadratic-Gompertz is unusual among major causes of death, and merits the close scrutiny we accord it.

The panels of figure 2 show that, over time, the female pattern progressively looks more and more like that of males. By 1980, female death rates very clearly have their mode in the 70s (of age), and are in the neighborhood of 100 per 100,000. By 1990, the female rates are much higher than

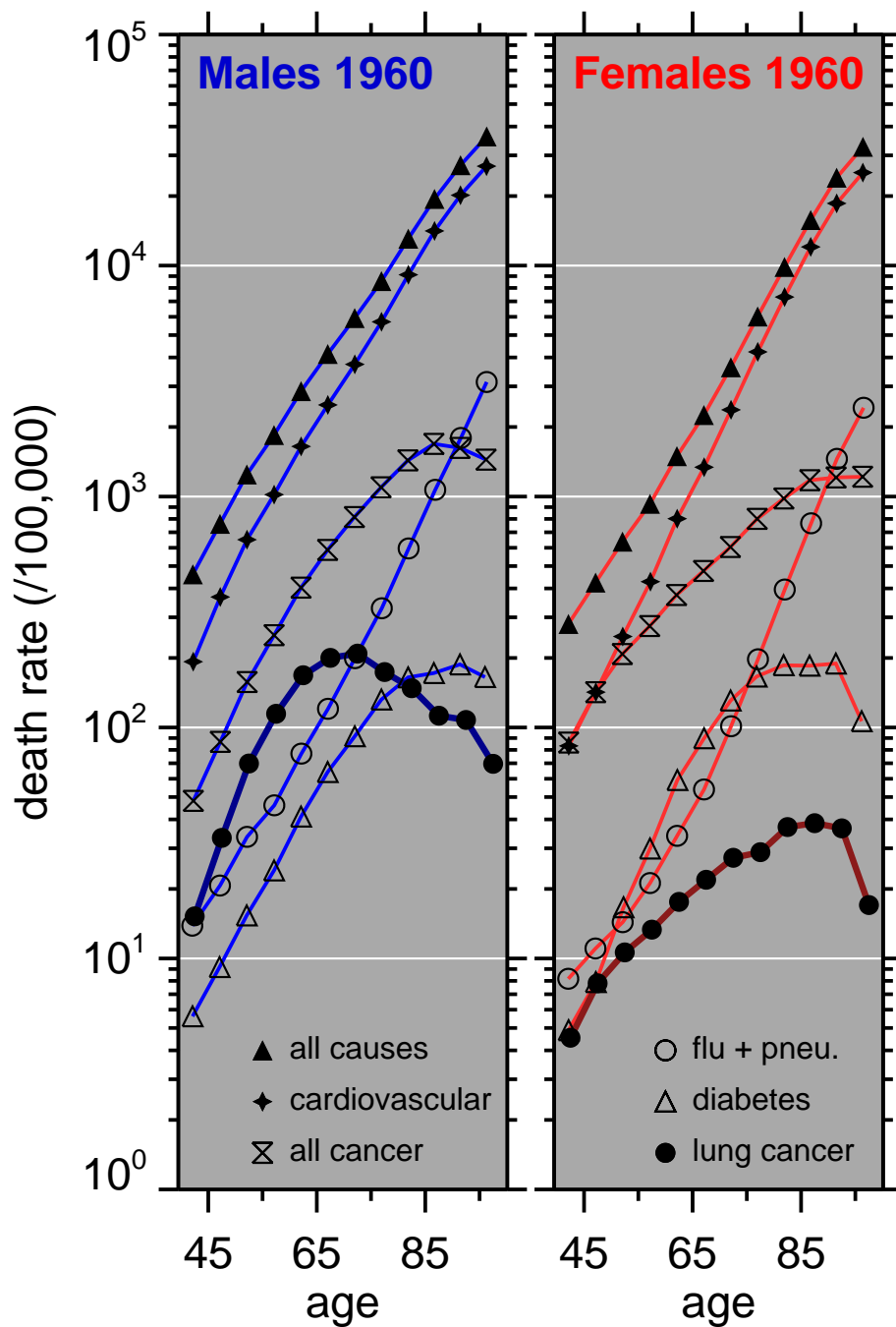


Figure 3: Death rates by age, from all causes, and selected causes, males and females, 1960. Note that “all cancer” denotes all cancer *excluding* lung cancer (which is shown separately).

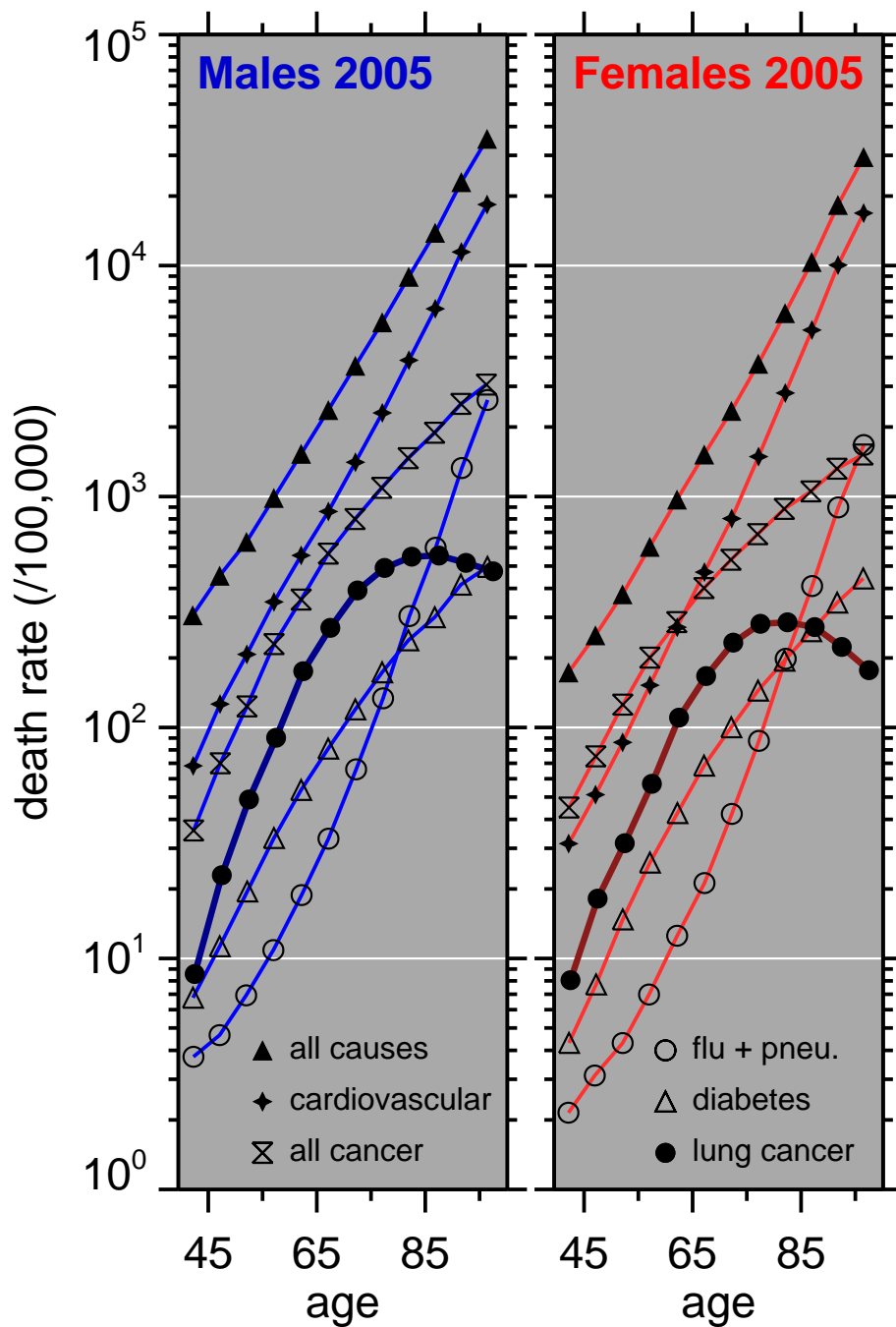


Figure 4: Death rates by age, from all causes, and selected causes, males and females, 2005. Note that “all cancer” denotes all cancer *excluding* lung cancer (which is shown separately).

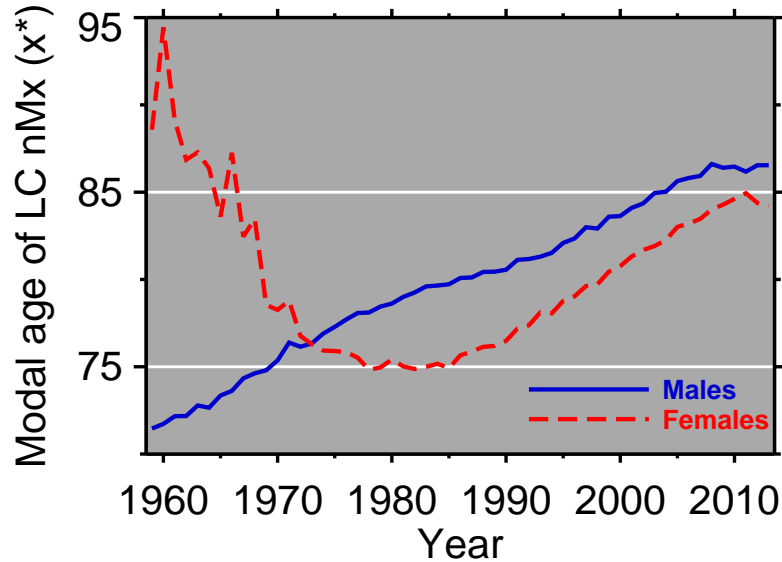


Figure 5: Modal age of lung cancer mortality rates (x^*), based on derivative of quadratic-Gompertz model.

100 per 100,000. The female pattern in 1990 looks somewhat parallel to that of males, at a lower level; it looks a lot like the male pattern in 1960. By 2010, the female lung cancer mortality rates are nested neatly just below those of males.

The *modal age* of the death rates plays an important role in this analysis of sex differences in lung cancer mortality; it provides a unidimensional summary of the effect of past tobacco use and its effect on mortality. The older the modal age, the closer the lung cancer pattern is to background mortality, and the younger the modal age, the closer it is to a tobacco-influenced pattern. To quantify the quadratic-Gompertzian modal age, x^* , we solved the following equation:

$$\frac{d}{dx} \log(M_x^{\text{LC}}) = \beta_1 + 2\beta_2 x^* = 0, \quad (2)$$

(the logarithm is a monotone transformation, so x^* is the same for logged or untransformed death rates). Thus, $x^* = -0.5\beta_1/\beta_2$, which must be positive since $\beta_1 > 0$ and $\beta_2 < 0$ in the lung cancer patterns. Referring to x^* as the modal age is consistent with other applications (Horiuchi 2003, Steinsaltz and Wachter 2006), but note that we are looking at the modal death rate, not

the modal age of death (in which the death rate is weighted by exposure). In the present application, our interest is the mortality rates, per se.

Figure 5 shows the pattern of x^* for males and females. Consider males first. The mean age of male lung cancer mortality has been rising, almost monotonically, for over half a century. An obvious question is, if increasing mean age of mortality indicates a return to a background rates, why does it occur during a period when *modal* rates (i.e., M_x^{LC}) are increasing and then decreasing (cf. figure 2)? The answer is that smoking partly is a cohort phenomenon and as smoking prevalence declines, it does so on a cohort basis.

The modal age is increasing essentially because smoking is going down and the pattern is (slowly) returning (approximately) to the pattern of “background” (i.e., tobacco-free) lung cancer mortality, of which 1960 females is the best exemplar here. However, this pattern should be interpreted cautiously, because modal age is also affected by competing risks. As mortality for other causes (most notably, cardiovascular disease) has declined since the 1960s (Doll 1987, Tate et al. 2016), more people are surviving longer, and must die of something else at older ages. Multiple cause mortality is a complex phenomenon since cardiovascular deaths averted at younger ages may simply be delayed to later ages as opposed to necessarily transferring to another cause (see Preston 1987 and Honoré and Lleras-Muney 2006 for detailed models, and Pearl 1929 for an early example of similar thinking). Moreover, the declining use of heart-related conditions as a “garbage code”, especially at advanced ages (Preston 1976), may mean that some of the lung cancer deaths were there before, so to say, but only became coded as malignant neoplasm when cause of death classification became more refined. Also, the effects of tobacco on cardiovascular health have been noted almost as long its effects on lung cancer (Russek et al. 1955, Bronte-Stewart 1961, Wald et al. 1973).

Figure 6 shows the goodness-of-fit (R^2) of the annual quadratic-Gompertz models. The fit is excellent, with all models above 93%. The goodness-of-fit was calculated as:

$$R^2 = 1 - \left[\sum_x w_x (\log(M_x^D) - \log(M_x^M))^2 / \sum_x w_x (\log(M_x^D) - M^N)^2 \right], \quad (3)$$

where $\log(M_x^D)$ are the logs of the rates in the data, $\log(M_x^M)$ are the modeled log rates, w_x are the weights (viz., the deaths in each cell), and M^N is the null model (DeGroot and Schervish 2002, p. 658). In this case, $M^N = \sum_x w_x \log(M_x) / \sum_x w_x$, or the weighted mean of the log rates; $\exp(M^N)$ is the weighted geometric mean death rate, which is a useful summary statis-

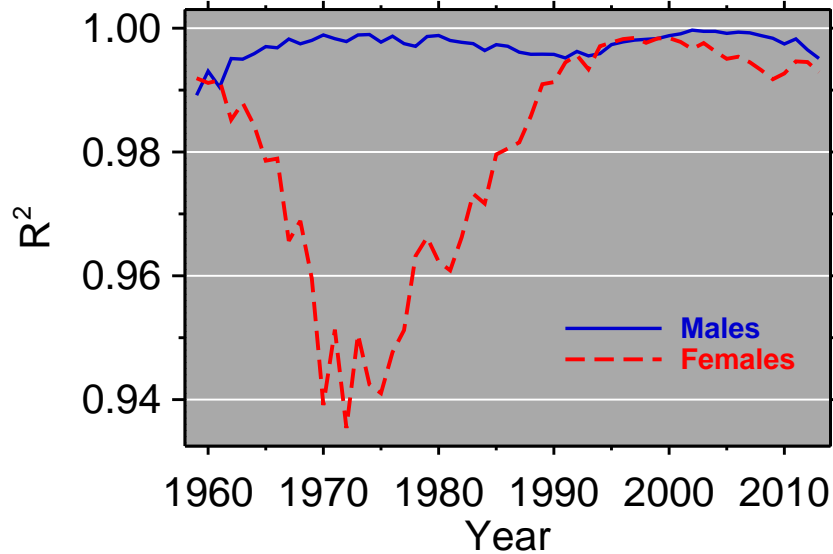


Figure 6: Goodness-of-fit (R^2) of quadratic-Gompertz models.

tic in demography (Schoen, 1970). Kvålseth (1985) notes the importance of making explicit the formula used for goodness of fit, since there are multiple (and non-equivalent) routes to its calculation; cf. Willett and Singer (1988) for discussion of the (weighted) R^2 as in eqn. (3).

The evolution of R^2 depicted in figure 6 is especially interesting for females, which begins with a truly excellent fit (99%), then dips, and then rises again. During the early 1960s, the model is a good fit because it is mostly cohorts of women who weren't heavy smokers. The model fit declines as the heavier-smoking female cohorts move through the ages, and the fit returns to good later on when a new equilibrium of lower-smoking cohorts is in place. During the transition period to a mortality pattern that fully reflects smoking, the female models don't fit the data quite as well, hence the decline and then increase in R^2 . We miss this transition in the male series because the detail mortality data in the US (viz., machine-readable cause of death data) begins in 1959.

To test the sex differences in the shapes of the quadratic-Gompertz lung cancer models, we pooled the male and female data and estimated models with full-interaction (sex, sex \times age, sex \times age-squared):

$$\log(M_{x,s}^{LC}) = \dot{\alpha} + \dot{\beta}_1 x + \dot{\beta}_2 x^2 + \gamma s + \delta(s \times x) + \zeta(s \times x^2) + \varepsilon, \quad (4)$$

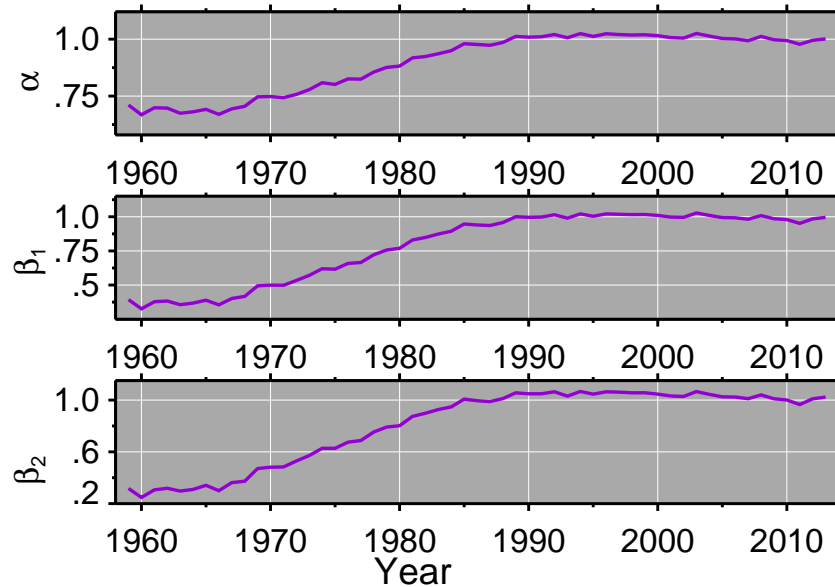


Figure 7: Female/male ratio, quadratic Gompertz coefficients over time.

where s is sex (coded 1 for males and 0 for females) and is x age. These models (one per year) recapitulate the coefficients in the single-sex models (eqn. 1), with $\hat{\alpha}$, $\hat{\beta}_1$, and $\hat{\beta}_2$ being the same coefficients as in the female-only models. For the men, the previous intercept is equal to $\hat{\alpha} + \gamma$, the previous age term is $\hat{\beta}_1 + \gamma$, and the previous quadratic term is $\hat{\beta}_2 + \zeta$. The advantage of the pooled-sex equation above is that it permits testing the sex differences among the suite of coefficients: intercept (α), slope (β_1), quadratic term (β_2) in each year. The test of the sex difference is the tests of significance of γ , δ , and ζ , respectively. This speaks to the shape of the quadratic curves, not whether men have higher overall lung cancer death rates — with thousands more male lung cancer deaths, the overall male excess is always significant.

The question here is not, are male lung cancer death rates higher (they are), but, are the male and female *patterns* distinguishable? Table 1 gives the results, as p -values on γ , δ , and ζ . The evolution is clear: from 1959–80 males and females have statistically-distinguishable differences in the shape of their quadratic-Gompertz lung cancer mortality patterns. This is followed by two years of transition, and from 1983 to the end of the data set, males and females have no distinguishable shape differences in the lung cancer mortality patterns.

| year | α [γ] | β_1 [δ] | β_2 [ζ] | year | α [γ] | β_1 [δ] | β_2 [ζ] |
|------|-----------------------|------------------------|-----------------------|------|-----------------------|------------------------|-----------------------|
| 1959 | 0.000 | 0.000 | 0.000 | 1987 | 0.534 | 0.356 | 0.885 |
| 1960 | 0.000 | 0.000 | 0.000 | 1988 | 0.729 | 0.523 | 0.882 |
| 1961 | 0.000 | 0.000 | 0.000 | 1989 | 0.749 | 0.994 | 0.430 |
| 1962 | 0.000 | 0.000 | 0.000 | 1990 | 0.813 | 0.943 | 0.490 |
| 1963 | 0.000 | 0.000 | 0.000 | 1991 | 0.771 | 0.974 | 0.477 |
| 1964 | 0.000 | 0.000 | 0.000 | 1992 | 0.532 | 0.776 | 0.296 |
| 1965 | 0.000 | 0.000 | 0.000 | 1993 | 0.868 | 0.862 | 0.660 |
| 1966 | 0.000 | 0.000 | 0.000 | 1994 | 0.452 | 0.676 | 0.272 |
| 1967 | 0.000 | 0.000 | 0.000 | 1995 | 0.653 | 0.933 | 0.364 |
| 1968 | 0.000 | 0.000 | 0.000 | 1996 | 0.338 | 0.587 | 0.176 |
| 1969 | 0.000 | 0.000 | 0.000 | 1997 | 0.381 | 0.618 | 0.172 |
| 1970 | 0.000 | 0.000 | 0.000 | 1998 | 0.454 | 0.686 | 0.231 |
| 1971 | 0.000 | 0.000 | 0.000 | 1999 | 0.394 | 0.648 | 0.205 |
| 1972 | 0.000 | 0.000 | 0.000 | 2000 | 0.453 | 0.758 | 0.252 |
| 1973 | 0.000 | 0.000 | 0.000 | 2001 | 0.735 | 0.945 | 0.437 |
| 1974 | 0.000 | 0.000 | 0.000 | 2002 | 0.825 | 0.898 | 0.509 |
| 1975 | 0.000 | 0.000 | 0.001 | 2003 | 0.223 | 0.403 | 0.113 |
| 1976 | 0.001 | 0.000 | 0.001 | 2004 | 0.573 | 0.798 | 0.363 |
| 1977 | 0.001 | 0.000 | 0.002 | 2005 | 0.909 | 0.898 | 0.660 |
| 1978 | 0.003 | 0.001 | 0.008 | 2006 | 0.959 | 0.867 | 0.669 |
| 1979 | 0.005 | 0.001 | 0.012 | 2007 | 0.823 | 0.714 | 0.855 |
| 1980 | 0.011 | 0.003 | 0.021 | 2008 | 0.728 | 0.879 | 0.576 |
| 1981 | 0.088 | 0.033 | 0.163 | 2009 | 0.950 | 0.821 | 0.897 |
| 1982 | 0.101 | 0.047 | 0.237 | 2010 | 0.885 | 0.750 | 0.995 |
| 1983 | 0.145 | 0.077 | 0.372 | 2011 | 0.507 | 0.380 | 0.604 |
| 1984 | 0.288 | 0.168 | 0.552 | 2012 | 0.899 | 0.807 | 0.899 |
| 1985 | 0.629 | 0.420 | 0.921 | 2013 | 0.975 | 0.953 | 0.795 |
| 1986 | 0.578 | 0.370 | 0.949 | | | | |

Table 1: Annual sex difference test. These are p -values (not coefficients) for sex differences of each coefficient. The test of the sex difference in α from equation 1 is γ from equation 4; the test of β_1 is δ and the test of β_2 is ζ .

A more direct approach is shown in figure 7. Rather than calculate p -values in an interactive model, it just plots the female to male ratio of the coefficient values. It is clear that there has been enormous convergence, with the ratios being very close to 1.0 since the mid 1980s. The very rapid convergence between 1980–83 implied by table 1 is an artifact of having a fixed level of significance. Figure 7 shows clearly that the convergence was a slow and steady phenomenon which began around 1966 and was complete by the mid 1980s. It was in the early 1980s that this pattern crosses the arbitrary statistical significance threshold of 0.05. While it is worth knowing that the patterns are indistinguishable at some conventional level of significance (and in which years), the convergence period in terms of statistical significance will necessarily be much shorter than that of the overall process.

Conclusion

In the United States, there is no gendered *pattern* of lung cancer mortality, though there are sex differences in *level*. Male lung cancer death rates exceed those of women, but the shapes of the age-mortality profiles have been statistically-indistinguishable since 1983. Mortality very often reflects behavior, never more so than with lung cancer and cigarette smoking. The shape of the lung cancer age-mortality profile reflects cohort histories of cigarette use, and these are now similar enough for both sexes that the pattern (but not the level) of lung cancer mortality is the same for both sexes. Prior research has already established sex differences in level as well as the tobacco nexus (see, f.e., Retherford 1972, Preston and Wang 2006).

What we have shown more originally is that period lung cancer mortality by age, especially in tobacco-smoking populations, is quadratic-Gompertz in pattern. Although this has been partially documented previously (Horiuchi and Wilmoth, 1997), the present study is the most comprehensive treatment of which we are aware. This framework allows us to demonstrate the convergence in patterns of lung cancer mortality between the sexes. Historically, men started smoking earlier — in fact, in some jurisdictions within the United States, smoking was once even illegal for women (Brandt, 2007). On a same-year basis, women have always had lower smoking rates, but the sex differences in smoking rates have been converging since the 1960s.

One attractive aspect of the quadratic Gompertz model is that it has a modal age, proportional to the ratio of the linear to the quadratic coefficients. This may be exploited as a unidimensional summary measure. The convergence story that is demonstrated by the interaction models is mir-

rored by the evolution of the modal ages of the quadratic pattern (seen in figure 5), and even the goodness-of-fit of the quadratic models follows a pattern in-line with tobacco use patterns (figure 6).

This study has limitations. We do not have individual data on smokers and non-smokers. The goal of this study is to do a descriptive analysis of lung cancer age-mortality patterns in the United States, net of changing composition of smokers and nonsmokers over time. This is not to deny, however, that microdata would permit a richer analysis (see, f.e., Ho and Fenelon 2015). Data by cohort would provide additional insights, subject, itself, to the limitation of much incomplete cohort data. Nonetheless, policy questions relevant to demographic data are often framed in period, not cohort, terms (Ní Bhrolcháin, 1992). As such, we believe that the male and female convergence in lung cancer mortality pattern is a noteworthy finding, while acknowledging that many of the underlying mechanisms align with cohorts (tobacco use, for example).

Our results suggest possibilities for replication in other countries. Standard international databases provide lung cancer death rates in 10-year age groups and with the open-ended interval beginning at age 75 (Pampel, 2003), which is not the ideal level of granularity for reproduction of this work. However, at least in principle, the foregoing analysis can be replicated in other settings. It is also possible that the modal age analysis can be applied profitably to understanding lung cancer epidemiology in regions whose current tobacco use patterns resemble earlier stages in the United States, such as Asia (Choe et al. 2004, TFI 2015). Specifically, in countries where the modal age of the quadratic-Gompertzian lung cancer death rates is falling (year-on-year), we can infer that their lung cancer epidemic is getting worse, and where the modal age is rising, that it has begun to abate.

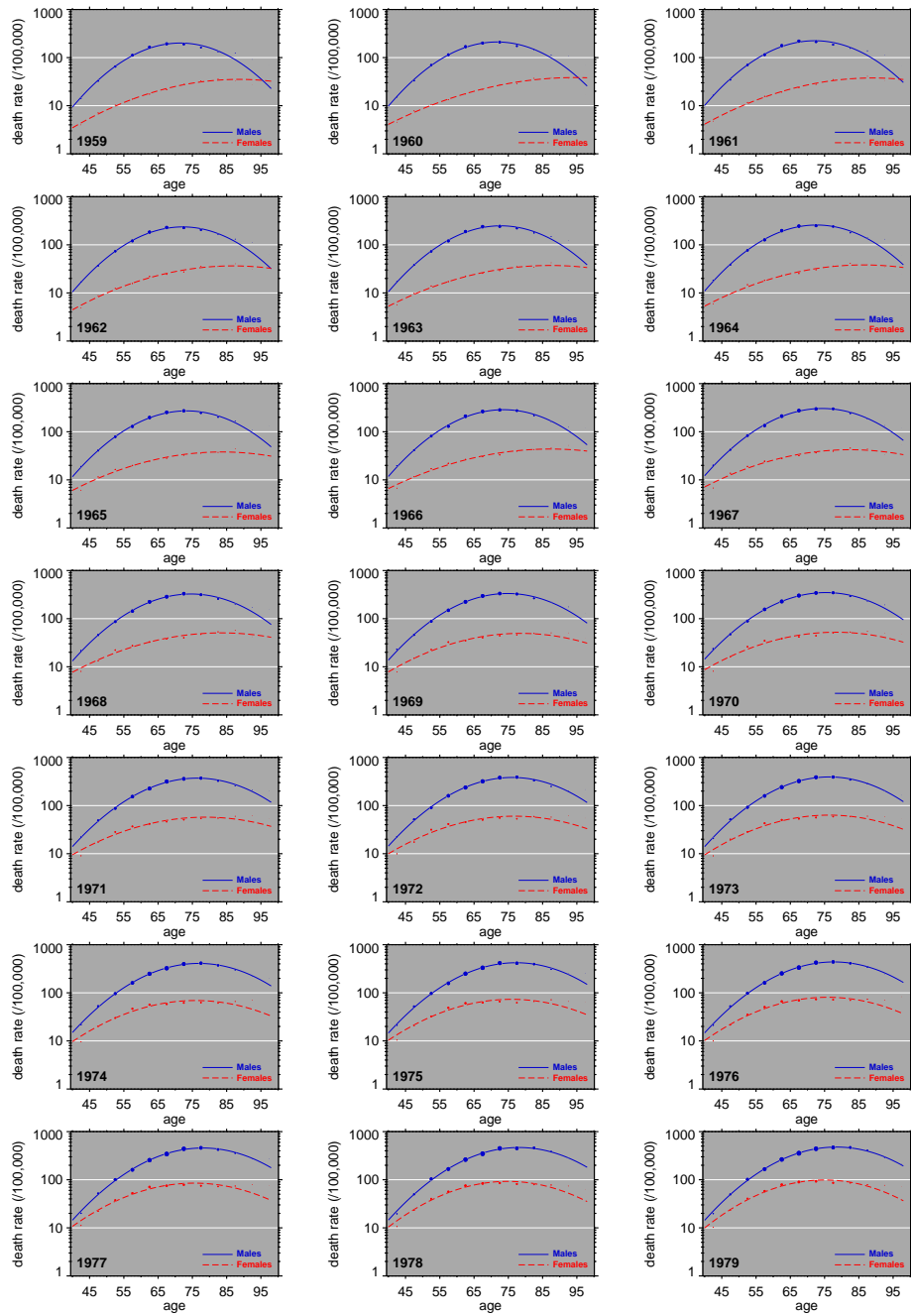
Acknowledgments

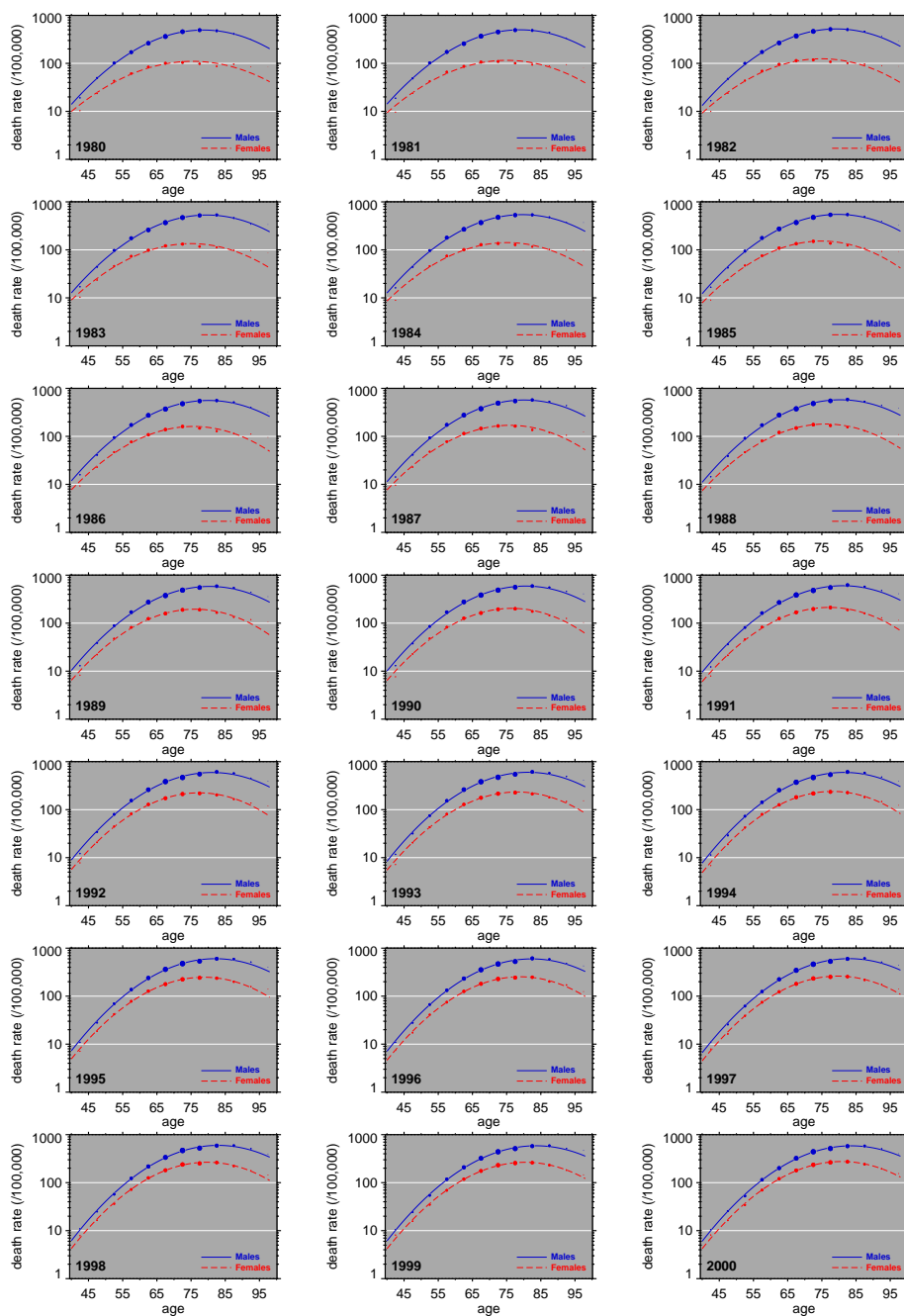
We thank the anonymous referees for constructive criticism. This work was presented at the 2016 Population Association of America meeting and the 2016 NBER Cohort Studies Meeting; we thank participants for feedback. We thank Olivia A. Quan for research assistance.

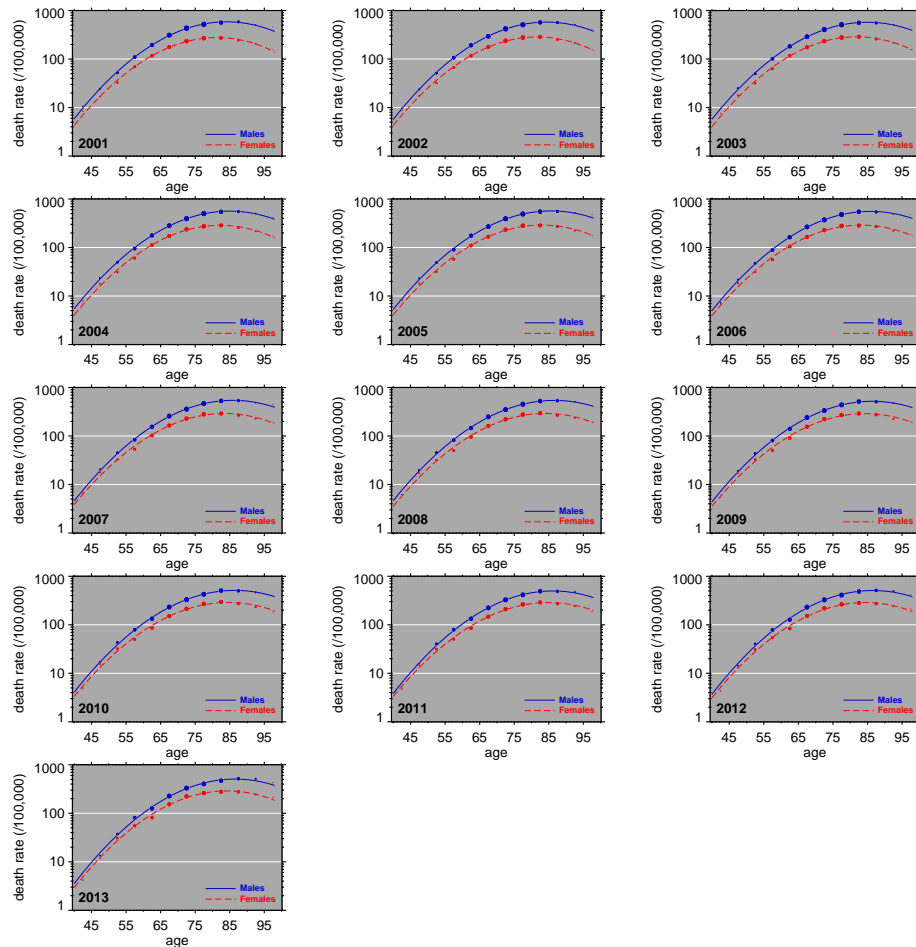
Appendix I: ICD codes for lung cancer mortality

| Years | Cause of death | |
|--------------------|----------------|--|
| | ICD code | Description (m.n., malignant neoplasm) |
| 1959–1967 (ICD 7) | 162 | m.n. of bronchus and trachea, and of lung specified as primary |
| | 163 | m.n. of lung, unspecified as to whether primary or secondary |
| 1968–1978 (ICD 8) | 162 | m.n. of trachea, bronchus and lung |
| 1979–1998 (ICD 9) | 162 | <i>ibid.</i> |
| 1999–2013 (ICD 10) | C33 | m.n. of trachea |
| | C34 | m.n. of bronchus and lung |

Appendix II: Graphs for all years







Works Cited

- Abdullatif, Viytta N. and Andrew Noymer. 2016. "Clostridium difficile infection: An emerging cause of death in the twenty-first century." *Biodemography and Social Biology* 62(2):198–207.
- Banks, Emily, Grace Joshy, Marianne F. Weber, Bette Liu, Robert Grenfell, Sam Egger, Ellie Paige, Alan D. Lopez, Freddy Sitas, and Valerie Beral. 2015. "Tobacco smoking and all-cause mortality in a large Australian cohort study: Findings from a mature epidemic with current low smoking prevalence." *BMC Medicine* 13(38):10pp.

- Boffetta, Paolo, Antonio Agudo, Wolfgang Ahrens, Ellen Benhamou, Simone Benhamou, Sarah C. Darby, Gilles Ferro, Cristina Fortes, Carlos A. Gonzalez, Karl-Heinz Jöckel, Martin Krauss, Lothar Kreienbrock, Michaela Kreuzer, Anabela Mendes, Franco Merletti, Fredrik Nyberg, Göran Pershagen, Hermann Pohlabein, Elio Riboli, Giovanni Schmid, Lorenzo Simonato, Jean Trédaniel, Elise Whitley, Heinz-Erich Wichmann, Carlos Winck, Paola Zambon, and Rodolfo Saracci. 1998. "Multicenter case-control study of exposure to environmental tobacco smoke and lung cancer in Europe." *Journal of the National Cancer Institute* 90(19):1440–1450.
- Brandt, Allan M. 2007. *The cigarette century: The rise, fall, and deadly persistence of the product that defined America*. Basic Books, New York.
- Bronte-Stewart, B. 1961. "Cigarette smoking and ischaemic heart disease." *British Medical Journal* 1(5223):379–384.
- Burns, David M., Lora Lee, Larry Z. Shen, Elizabeth Gilpin, H. Dennis Tolley, Jerry Vaughn, and Thomas G. Shanks. 1997. "Cigarette smoking behavior in the United States." In David M. Burns, Lawrence Garfinkel, and Jonathan M. Samet (eds.), *Changes in cigarette-related disease risks and their implications for prevention and control, Tobacco Control Monograph Series*, vol. 8, chap. 2, pp. 13–112. National Cancer Institute, Bethesda, MD. http://cancercontrol.cancer.gov/brp/tcrb/monographs/8/m8_2.pdf.
- Carter, Brian D., Christian C. Abnet, Diane Feskanich, Neal D. Freedman, Patricia Hartge, Cora E. Lewis, Judith K. Ockene, Ross L. Prentice, Frank E. Speizer, Michael J. Thun, and Eric J. Jacobs. 2015. "Smoking and mortality: Beyond established causes." *New England Journal of Medicine* 372(7):631–640.
- Choe, Minja Kim, Shyam Thapa, Chai Podhisita, Corazon Raymundo, Hui sheng Lin, and Sulistina Achmad. 2004. "The teen tobacco epidemic in Asia: Indonesia, Nepal, Philippines, Taiwan, and Thailand." *Journal of Youth Studies* 7(1):73–87.
- Christopoulou, Rebekka and Dean R. Lillard. 2016. "A novel indicator of life-course smoking prevalence in the United States combining popularity, duration, quantity, and quality of smoking." *American Journal of Public Health* 106(7):1329–1335.
- DeGroot, Morris H. and Mark J. Schervish. 2002. *Probability and statistics*. Addison-Wesley, Reading, MA, third ed.

- Dix, Douglas, Patricia Cohen, and John Flannery. 1980. "On the role of aging in cancer incidence." *Journal of Theoretical Biology* 83(1):163–173.
- Doll, Richard. 1987. "Major epidemics of the 20th century: From coronary thrombosis to AIDS." *Journal of the Royal Statistical Society, Series A* 150(4):373–395.
- Gregg, Edward W., Yiling J. Cheng, Sharon Saydah, Catherine Cowie, Sanford Garfield, Linda Geiss, and Lawrence Barker. 2012. "Trends in death rates among U.S. adults with and without diabetes between 1997 and 2006." *Diabetes Care* 35(6):1252–1257.
- Hammond, E. Cuyler and Daniel Horn. 1954. "The relationship between human smoking habits and death rates: A follow-up study of 187,766 men." *Journal of the American Medical Association* 155(15):1316–1328.
- Ho, Jessica Y. and Andrew Fenelon. 2015. "The contribution of smoking to educational gradients in U.S. life expectancy." *Journal of Health and Social Behavior* 56(3):307–322.
- Holford, Theodore R., David T. Levy, Lisa A. McKay, Lauren Clarke, Ben Racine, Rafael Meza, Stephanie Land, Jihyoun Jeon, and Eric J. Feuer. 2014a. "Patterns of birth cohort-specific smoking histories, 1965–2009." *American Journal of Preventive Medicine* 46(2):e31–e37.
- Holford, Theodore R., Rafael Meza, Kenneth E. Warner, Clare Meernik, Jihyoun Jeon, Suresh H. Moolgavkar, and David T. Levy. 2014b. "Tobacco control and the reduction in smoking-related premature deaths in the United States, 1964–2012." *JAMA* 311(2):164–171.
- Honoré, Bo E. and Adriana Lleras-Muney. 2006. "Bounds in competing risks models and the war on cancer." *Econometrica* 74(6):1675–1698.
- Horiuchi, Shiro. 2003. "Interspecies differences in the life span distribution: Humans versus invertebrates." *Population and Development Review* 29(Suppl):127–151.
- Horiuchi, Shiro and John R. Wilmoth. 1997. "Age patterns of the life table aging rate for major causes of death in Japan, 1951–1990." *Journal of Gerontology: Biological Sciences* 52A(1):B67–B77.
- Human Mortality Database. Accessed 15 January 2015. *U.S.A. Complete Data Series: Exposure-to-risk*. <http://www.mortality.org/>.

- Jacobs, Eric J., Christina C. Newton, Brian D. Carter, Diane Feskanich, Neal D. Freedman, Ross L. Prentice, and W. Dana Flanders. 2015. "What proportion of cancer deaths in the contemporary United States is attributable to cigarette smoking?" *Annals of Epidemiology* 25(3):179–182.e1.
- Jha, Prabhat, Chinthanie Ramasundarahettige, Victoria Landsman, Brian Rostron, Michael Thun, Robert N. Anderson, Tim McAfee, and Richard Peto. 2013. "21st-Century hazards of smoking and benefits of cessation in the United States." *New England Journal of Medicine* 368(4):341–350.
- Jöckel, Karl-Heinz, Hermann Pohlmann, Wolfgang Ahrens, and Martin Krauss. 1998. "Environmental tobacco smoke and lung cancer." *Epidemiology* 9(6):672–675.
- Kvålseth, Tarald O. 1985. "Cautionary note about R^2 ." *American Statistician* 39(4):279–285.
- Manton, Kenneth G., Igor Akushevich, and Julia Kravchenko. 2009. *Cancer mortality and morbidity patterns in the U.S. population: An interdisciplinary approach*. Springer, New York.
- Manton, Kenneth G., Eric Stallard, and James W. Vaupel. 1986. "Alternative models for the heterogeneity of mortality risks among the aged." *Journal of the American Statistical Association* 81(395):635–644.
- Menke, Andy, Sarah Casagrande, Linda Geiss, and Catherine C. Cowie. 2015. "Prevalence of and trends in diabetes among adults in the United States, 1988–2012." *JAMA* 314(10):1021–1029.
- Moolgavkar, Suresh H., Theodore R. Holford, David T. Levy, Chung Yin Kong, Millenia Foy, Lauren Clarke, Jihyoun Jeon, William D. Hazelton, Rafael Meza, Frank Schultz, William McCarthy, Robert Boer, Olga Gorlova, G. Scott Gazelle, Marek Kimmel, Pamela M. McMahon, Harry J. de Koning, and Eric J. Feuer. 2012. "Impact of reduced tobacco smoking on lung cancer mortality in the United States during 1975–2000." *Journal of the National Cancer Institute* 104(7):541–548.
- National Center for Health Statistics. 2014. *Mortality multiple cause-of-death data files*, http://www.cdc.gov/nchs/nvss/mortality_public_use_data.htm. National Center for Health Statistics. Accessed 15 January 2015.
- Ní Bhrolcháin, Máire. 1992. "Period paramount? A critique of the cohort approach to fertility." *Population and Development Review* 18(4):599–629.

- Pampel, Fred C. 2002. "Cigarette use and the narrowing sex differential in mortality." *Population and Development Review* 28(1):77–104.
- . 2003. "Declining sex differences in mortality from lung cancer in high-income nations." *Demography* 40(1):45–65.
- Pearl, Raymond. 1929. "A Note on the association of diseases." *Science* 70(1808):191–192.
- Preston, Samuel H. 1976. *Mortality patterns in national populations: With special reference to recorded causes of death*. Academic Press, New York.
- . 1987. "Relations among standard epidemiologic measures in a population." *American Journal of Epidemiology* 126(2):336–345.
- Preston, Samuel H. and Haidong Wang. 2006. "Sex mortality differentials in the United States: The role of cohort smoking patterns." *Demography* 43(4):631–646.
- Retherford, Robert D. 1972. "Tobacco smoking and the sex mortality differential." *Demography* 9(2):203–216.
- Rogers, Richard G., Robert A. Hummer, Patrick M. Krueger, and Fred C. Pampel. 2005. "Mortality attributable to cigarette smoking in the United States." *Population and Development Review* 31(2):259–292.
- Russek, Henry I., Burton L. Zohman, and Virgil J. Dorset. 1955. "Effects of tobacco and whiskey on the cardiovascular system." *Journal of the American Medical Association* 157(7):563–568.
- Schoen, Robert. 1970. "The geometric mean of the age-specific death rates as a summary index of mortality." *Demography* 7(3):317–324.
- Siegel, Rebecca L., Eric J. Jacobs, Christina C. Newton, Diane Feskanich, Neal D. Freedman, Ross L. Prentice, and Ahmedin Jemal. 2015. "Deaths due to cigarette smoking for 12 smoking-related cancers in the United States." *JAMA Internal Medicine* 175(9):1574–1576.
- Steinsaltz, David R. and Kenneth W. Wachter. 2006. "Understanding mortality rate deceleration and heterogeneity." *Mathematical Population Studies* 13(1):19–37.
- Tate, Steven, Jamie J. Namkung, and Andrew Noymer. 2016. "Did the 1918 influenza cause the twentieth century cardiovascular mortality epidemic in the United States?" *PeerJ* 4:e2531.

- TFI (Tobacco Free Initiative). 2015. "Global trends and projections for tobacco smoking, 2000–2025." <http://www.who.int/entity/tobacco/publications/surveillance/prevalencetrendsresults2010to2025.xlsx>. World Health Organization. Accessed 26 June 2015.
- US Public Health Service. 1964. "Smoking and health: Report of the advisory committee to the Surgeon General of the Public Health Service." Public Health Service Publication 1103, US Department of Health, Education, and Welfare, Washington DC.
- Wald, Nicholas, Susannah Howard, P. G. Smith, and Knud Kjeldsen. 1973. "Association between atherosclerotic diseases and carboxyhaemoglobin levels in tobacco smokers." *British Medical Journal* 1(5856):761–765.
- Willett, John B. and Judith D. Singer. 1988. "Another cautionary note about R^2 : Its use in weighted least-squares regression analysis." *American Statistician* 42(3):236–238.
- Yashin, A. I., K. G. Arbeev, I. Akushevich, A. Kulminski, S. V. Ukraintseva, E. Stallard, and K. C. Land. 2012. "The quadratic hazard model for analyzing longitudinal data on aging, health, and the life span." *Physics of Life Reviews* 9(2):177–188.