Theoretical Population Biology 114 (2017) 117–127

Contents lists available at ScienceDirect

Theoretical Population Biology

journal homepage: www.elsevier.com/locate/tpb

Theory of partitioning of disease prevalence and mortality in observational data

I. Akushevich [a,*], A.P. Yashkin [a], J. Kravchenko [b], F. Fang [a], K. Arbeev [a], F. Sloan [c], A.I. Yashin [a]

[a] Biodemography of Aging Research Unit, Center for Population Health and Aging, Duke University, Durham, NC, United States

[b] Department of Surgery, Duke University School of Medicine, Durham, NC, United States

[c] Department of Economics, Duke University, Durham, NC, United States

ARTICLE INFO

Article history:
Received 23 July 2016
Available online 24 January 2017

Keywords:
Time trend
Partitioning
Incidence
Prevalence
Mortality
Diabetes

ABSTRACT

In this study, we present a new theory of partitioning of disease prevalence and incidence-based mortality and demonstrate how this theory practically works for analyses of Medicare data. In the theory, the prevalence of a disease and incidence-based mortality are modeled in terms of disease incidence and survival after diagnosis supplemented by information on disease prevalence at the initial age and year available in a dataset. Partitioning of the trends of prevalence and mortality is calculated with minimal assumptions. The resulting expressions for the components of the trends are given by continuous functions of data. The estimator is consistent and stable. The developed methodology is applied for data on type 2 diabetes using individual records from a nationally representative 5% sample of Medicare beneficiaries age 65+. Numerical estimates show excellent concordance between empirical estimates and theoretical predictions. Evaluated partitioning model showed that both prevalence and mortality increase with time. The primary driving factors of the observed prevalence increase are improved survival and increased prevalence at age 65. The increase in diabetes-related mortality is driven by increased prevalence and unobserved trends in time-periods and age-groups outside of the range of the data used in the study. Finally, the properties of the new estimator, possible statistical and systematical uncertainties, and future practical applications of this methodology in epidemiology, demography, public health and health forecasting are discussed.

1. Introduction

Prevalence is an epidemiologic characteristic which is easily measured using survey data or medical records. Analyses of prevalence trends play an influential role in health policy planning and are widely used to assess the extent to which a given health problem affects the population. However, conclusions about the relative success or failure of a health policy change cannot be made directly from trends of disease prevalence because temporal changes in age-adjusted prevalence rates are the result of two simultaneously occurring competing processes: (i) changes in incidence and (ii) changes in survival. Health interventions and disease treatment guidelines are usually aimed at decreasing the incidence and increasing the survival rate for a disease. If successful, these measures will push the observed prevalence in different directions. A related quantity of interest is the mortality rate by cause or more generally, the mortality for individuals after the onset of a specific disease. This is also known as the incidence-based mortality rate (Chu et al., 1994). The time trend of incidence-based mortality (Mozaffarian et al., 2016; Smith et al., 2013; Thun et al., 2013) is defined by the same factors that define the time trends in the disease prevalence rate, as well as trends in mortality in the general population. In contrast to disease prevalence, improvements in incidence and survival push the observed incidence-based mortality for a specific disease in the same direction, because improved incidence reduces the total number of people with the disease and improved survival further reduces the number of deaths associated with the disease.

In this paper, we develop a new methodological approach for the decomposition of trends in disease prevalence and incidence-based mortality into their constituent components (such as trends in incidence, survival, and prevalence prior to observation) and for

* Correspondence to: Biodemography of Aging Research Unit, Social Science Research Institute, Duke University, 2024 W. Main Str., United States.
E-mail address: igor.akushevich@duke.edu (I. Akushevich).

http://dx.doi.org/10.1016/j.tpb.2017.01.003
0040-5809/

the evaluation of the strength and the direction of the contribution of each respective component. The methodology described in this study offers a number of distinct strengths: (i) computation of disease prevalence and incidence-based mortality as well as their partitioning through a set of exact formulas without making simplifying assumptions, (ii) evaluation of the individual contributions of each component to the total time trend by direct calculation using exact formulas applied to real data, and (iii) a set of natural generalizations including applications to medical costs, complications of a specific disease, the incorporation of disease risk factors and the use of the historical trends of each of the model components beyond the region directly measured in data.

The only previously existing methodological approach of this type was developed by Tunstall-Pedoe for the partitioning of mortality trends through the use of an approximate formula for the simple decomposition of the annual percent change (APC) for mortality as a sum of APC's of cardiovascular disease incidence and case fatality (percentage of 28-day fatalities) (Tunstall-Pedoe et al., 1999). This approximation is valid only for events (disease onset and death) occurring within a short time of each other and requires that the APC be small and the disease of interest be the primary cause of death. Other methods of decomposition used in demography and epidemiology (see Canudas-Romo, 2003; Horiuchi et al., 2008; Vaupel and Romo, 2003 for a comprehensive review) are not related to the decomposition of prevalence into its constituent components.

Although the primary focus of this paper is to introduce the methodology and describe the mathematics involved in its execution, an example involving type 2 diabetes mellitus is also considered. The application of the methodology to disease prevalence and mortality is intended to address an aspect of a current Public Health problem—with some notable exceptions such as cardiovascular disease (Will et al., 2014), the prevalence rate of many chronic diseases including diabetes has been increasing with time (Akinbami et al., 2012; Bauer et al., 2014; Coresh et al., 2007; Egan et al., 2010). Understanding the contribution each individual component makes to the overall effect on disease prevalence and mortality and how these contributions have changed over time in response to changes in health policy, population age-structure and epidemiologic characteristics could be of great use in identifying likely targets for pro-active policy interventions.

2. Theory

2.1. Mathematical formalism

Data collected in an observational study represent information on eligible individuals over given periods of age and time. In this study, we use a nationally representative 5% sample of the US Medicare population provided for research as restricted access public use files by the Centers for Medicare and Medicaid Services. This database provides individual heath related information on US Medicare beneficiaries after age 65 from 1991 to 2013. The long time period and level of detail provided by such data allow us to calculate disease prevalence and mortality at any point after a certain look-back period (12 months is used in this study) necessary to collect individual information for evaluation of disease presence. Fig. 1 presents the Lexis diagram in the plane over age (in years; denoted by x) and calendar time (in years; denoted by y). Each of the dashed lines in the Lexis diagram uniquely corresponds to a birth cohort with the birth time $y_b = y - x$ for any point (x, y) belonging to the cohort-specific dashed line. Therefore, epidemiologic characteristics at a given point of time are defined by the history of the cohort represented by a leftward move along the respective line in the Lexis diagram down to bounds of the available region. The bound is defined by an initial year (y_{00})

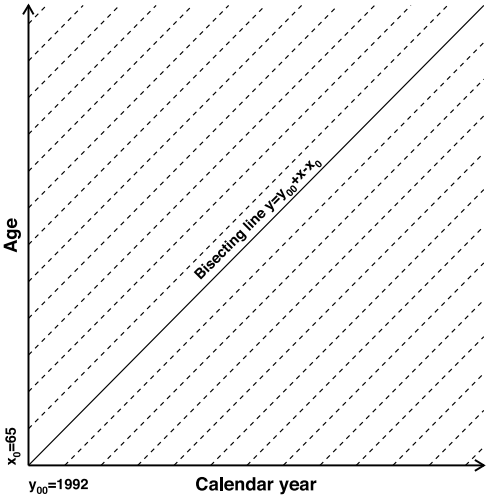

Fig. 1. The two dimensional diagram (Lexis diagram) to show the age–time area in which data are available and represent events (such as disease onset or deaths) that occur to individuals belonging to different cohorts. Calendar time is represented on the horizontal axis, while age is represented on the vertical axis. Dashed lines show time/age points for specific cohorts. Information about a cohort is available starting from bounding lines, i.e., either $y_{00} = 1992$ or $x_0 = 65$. Calculation of age-adjusted rates for a specific time requires integration over all ages starting from x_0, so regions both below and above bisecting line contribute to the integral for any $y > y_{00}$.

or minimal age (x_0) observed in the data. These two subareas are separated by the bisecting line defined as $y = y_{00} + x - x_0$. Above the bisecting line, the starting point is defined by the initial conditions $y = y_{00}$ with various ages while below the line the initial point is defined by boundary condition $x = x_0$ with various years. The cohort-specific bounding point is defined as $\bar{x}_0 = \max(x_0, y_{00} - y_b)$ and $\bar{y}_0 = y_b + \bar{x}_0$. Definitions of ages and times as well as functions of survival analyses used in the paper are collected in Table 1.

The idea for the representation of the formulas for prevalence is based on that the probability of being prevalent $P_c(x, y_b)$ at age x in cohort c with birth time y_b requires either

(i) being prevalent (represented by initial prevalence $P_c(\bar{x}_0, y_b)$) in the initial age $\bar{x}_0$ (and year $\bar{y}_0$) for the cohort and surviving to age x (represented by the survival probability $\bar{S}(x - \bar{x}_0, \bar{x}_0, \bar{y}_0)$ of a patient diagnosed no later than $\bar{x}_0$), or
(ii) being incident at an earlier age τ, $\bar{x}_0 < \tau \le x$ (represented by incidence density function $I_c(\tau, y_b)$) and having survival longer than $x - \tau$ (represented by survival probability $S(x - \tau, \tau, y_d)$ of a patient diagnosed at age τ and year y_d).

Therefore

$$P_c(x, y_b) = P_c(\bar{x}_0, y_b)\bar{S}(x - \bar{x}_0, \bar{x}_0, \bar{y}_0) + \int_{\bar{x}_0}^{x} I_c(\tau, y_b)S(x - \tau, \tau, y_d)d\tau \quad (1)$$

where we integrate over all possible ages at diagnosis. Similarly, for mortality (we consider incidence-based mortality, i.e., mortality after disease onset) the probability of dying in the age interval $(x, x + dx)$ requires having death in the interval $(x, x + dx)$ and either being prevalent at the boundary point $(\bar{x}_0, \bar{y}_0)$ for this cohort or being incident at an earlier age $x - \tau$. Death is represented by a respective density function $M_c(x, y_b)$ such that

$$M_c(x, y_b) = P_c(\bar{x}_0, y_b)\bar{f}_c(x - \bar{x}_0, \bar{x}_0, \bar{y}_0) + \int_{\bar{x}_0}^{x} I_c(\tau, y_b)f_c(x - \tau, \tau, y_d)d\tau. \quad (2)$$

The densities $\bar{f}_c(x - \bar{x}_0, \bar{x}_0, \bar{y}_0)$ and $f_c(x - \tau, \tau, y_d)$ in (2) are related to respective survival functions in (1): $\bar{f}_c() = -\bar{S}'_c()$ and $f_c() =$

$-S_c'()$, where derivatives are taken in respect to the first argument. Details of derivation of Eqs. (1) and (2) and some properties of the contributed functions are given in Appendix A.

The exact definition of $P_c(x, y_b)$ is the fraction of individuals born in year y_b and living with the disease at age x of the total number of individuals born in year y_b. Similarly, $I_c(\tau, y_b)$ and $M_c(x, y_b)$ are the cohort incidence and mortality densities defined through the number of new incident and death cases per cohort size (i.e., the number of individuals born in year y_b). However, the cohort size for the studied population is not usually known with sufficient accuracy. What is known (or can be estimated) is the current population at risk, i.e., the population currently living in the same age and calendar year (denoted as $y = y_b+x$) or calendar year of diagnosis (denoted as $y_d = y_b + \tau$). Therefore, we avoid dealing with cohort prevalence and incidence/mortality densities and use their standard definitions involving the population at risk rather than birth cohort size. Within these definitions, the cohort prevalence and incidence/mortality densities are expressed through accepted definitions of prevalence and hazard functions of incidence and mortality: $P_c(x, y_b) = P(x, y)S_t(x, 0, y_b)$, $M_c(x, y_b) = M(x, y_d)S_t(x, 0, y_b)$, and $I_c(\tau, y_b) = I(\tau, y_d)S_t(\tau, 0, y_b)$, where $S_t(x, 0, y_b)$ is the survival function of the cohort born during year y_b. Using these expressions in Eq. (1) results in occurrence of three survival functions on the right hand side which can be combined in the relative survival functions (i.e., the ratios of survival probabilities for individuals with the disease and the general population):

$$\begin{aligned}\bar{S}^r(x-\bar{x}_0,\bar{x}_0,\bar{y}_0) &= \frac{\bar{S}(x-\bar{x}_0,\bar{x}_0,\bar{y}_0)}{S_t(x-\bar{x}_0,\bar{x}_0,\bar{y}_0)}\\ &= \frac{S_t(\bar{x}_0,0,y_b)}{S_t(x,0,y_b)}\bar{S}(x-\bar{x}_0,\bar{x}_0,\bar{y}_0),\\ S^r(x-\tau,\tau,y_d) &= \frac{S(x-\tau,\tau,y_d)}{S_t(x-\tau,\tau,y_d)}\\ &= \frac{S_t(\tau,0,y_b)}{S_t(x,0,y_b)}S(x-\tau,\tau,y_d).\end{aligned} \tag{3}$$

The resulting expression for age-specific prevalence is:

$$P(x,y) = P(\bar{x}_0,\bar{y}_0)\bar{S}^r(x-\bar{x}_0,\bar{x}_0,\bar{y}_0) + \int_{\bar{x}_0}^{x} I(\tau,y_d)S^r(x-\tau,\tau,y_d)d\tau \tag{4}$$

where $\bar{y}_0 = y_b + \bar{x}_0 = y - x + \bar{x}_0$ and $y_d = y - x + \tau$.

Combining survival probabilities for mortality results in the ratio of density f_c (or $\bar{f}_c$) to the survival probability in the general population, that can be further transformed as

$$\begin{aligned}\frac{f(x-\tau,\tau,y_d)}{S_t(x-\tau,\tau,y_d)} &= -\frac{S'(x-\tau,\tau,y_d)}{S_t(x-\tau,\tau,y_d)}\\ &= -\frac{(S_t(x-\tau,\tau,y_d)S^r(x-\tau,\tau,y_d))'}{S_t(x-\tau,\tau,y_d)}\\ &= f^r(x-\tau,\tau,y_d) + S^r(x-\tau,\tau,y_d)\mu(x,y).\end{aligned} \tag{5}$$

The function $\mu(x, y)$ is the mortality hazard function in the general population. Note because $S_t(x-\tau,\tau,y_d) = \exp\left(-\int_\tau^x \mu(v, y_d + v-\tau)dv\right)$, we have $S_t'(x-\tau,\tau,y_d) = -\mu(x, y_d+x-\tau)S_t(x-\tau, \tau, y_d)$. We see that all terms at $\mu(x, y)$ give the prevalence and finally we obtain the incidence-based mortality in terms of $\mu(x, y)$ and functions previously derived.

$$M(x,y) = P(x,y)\mu(x,y) + P(\bar{x}_0,\bar{y}_0)\bar{f}^r(x-\bar{x}_0,\bar{x}_0,\bar{y}_0) + \int_{\bar{x}_0}^{x} I(\tau,y_d)f^r(x-\tau,\tau,y_d)d\tau. \tag{6}$$

Explicit representation (i.e., avoiding a function maximum used in the definition of $\bar{x}_0$) of prevalence and mortality as functions of x and y that allows for expressing prevalence and mortality in terms of x, y, and constants, requires considering two regions below and above the bisecting line, i.e., the regions defined by inequalities $y \geq y_{00} + x - x_0$ and $y < y_{00} + x - x_0$. Thus

$$\begin{aligned}P(x,y) &= P(x_{00},y_{00})\bar{S}^r(y-y_{00},x_{00},y_{00}) + \int_{x_{00}}^{x} I(\tau,y_d)S^r(x-\tau,\tau,y_d)d\tau,\\ M(x,y) &= P(x,y)\mu(x,y) + P(x_{00},y_{00})\bar{f}^r(y-y_{00},x_{00},y_{00}) + \int_{x_{00}}^{x} I(\tau,y_d)f^r(x-\tau,\tau,y_d)d\tau\end{aligned} \tag{7}$$

for $y < y_{00} + x - x_0$ and

$$\begin{aligned}P(x,y) &= P(x_0,y_0)\bar{S}^r(x-x_0,x_0,y_0) + \int_{x_0}^{x} I(\tau,y_d)S^r(x-\tau,\tau,y_d)d\tau,\\ M(x,y) &= P(x,y)\mu(x,y) + P(x_0,y_0)\bar{f}^r(x-x_0,x_0,y_0) + \int_{x_0}^{x} I(\tau,y_d)f^r(x-\tau,\tau,y_d)d\tau\end{aligned} \tag{8}$$

for $y \geq y_{00} + x - x_0$. The formulas (7) and (8) coincide for $y = y_{00} + x - x_0$. In these formulas we denote the age at y_{00} as $x_{00} = y_{00} - y + x$ and year at x_0 as $y_0 = y - x + x_0$.

The quantity of interest is the time trend of age adjusted prevalence (over the age region $(x_0, x_{\max})$) and mortality as well as their partitioning. Age-adjusted prevalence and incidence-based mortality based on (7) and (8) are:

$$\begin{aligned}P(y) = \int_{x_0}^{\infty} \Big(&P(x_0,y_0)\bar{S}^r(x-x_0,x_0,y_0)I(x \leq y-y_{00}+x_0)\\ &+P(x_{00},y_{00})\bar{S}^r(y-y_{00},x_{00},y_{00})I(x > y-y_{00}+x_0)\\ &+\int_{\max(x_0,x_{00})}^{x} I(\tau,y_d)S_d^r(x-\tau,\tau,y_d)d\tau\Big)p(x)dx\end{aligned} \tag{9}$$

and

$$\begin{aligned}M(y) = \int_{x_0}^{\infty} \Big(&P(x,y)\mu(x,y) + P(x_0,y_0)\bar{f}^r(x-x_0,x_0,y_0)\\ &\times I(x \leq y-y_{00}+x_0)\\ &+P(x_{00},y_{00})\bar{f}^r(y-y_{00},x_{00},y_{00})I(x > y-y_{00}+x_0)\\ &+\int_{\max(x_0,x_{00})}^{x} I(\tau,y_d)S^r(x-\tau,\tau,y_d)d\tau\Big)p(x)dx\end{aligned} \tag{10}$$

where $I()$ is the indicator function and $p(x)$ is the density of age distribution in a standard year. Recall, $x_{00} = y_{00}-y+x$, $y_d = y-x+\tau$, and $y_0 = y-x+x_0$ are functions of y and the integration variables x and τ. Age-adjusted prevalence and mortality are functions of three and four contributing factors, respectively:

$$\begin{aligned}P(y) &= P_0(y) + P_{00}(y) + P_{is}(y),\\ M(y) &= M_{P\mu}(y) + M_0(y) + M_{00}(y) + M_{is}(y).\end{aligned} \tag{11}$$

The derivative of $P(y)$ with respect to y represents the time trend of age-adjusted prevalence and are determined by trends in the respective components including initial prevalence (i.e., prevalence at x_0 or y_{00}), incidence rates, relative survival after disease onset and in patients with the disease at initial point of observation. Explicit differentiation results in seven terms (note that $\max(x_0, x_{00})$ depends on y because of x_{00}). Thus,

$$\begin{aligned}P_y'(y) = &T_{p0}(y) + T_{p00}(y) + T_{\bar{S}}(y) + T_{\bar{S}00}(y)\\ &+T_{X00}(y) + T_{inc}(y) + T_S(y)\end{aligned} \tag{12}$$

Table 1
Summary of mathematical functions.

Ages and calendar times used as arguments in the survival functions			
x	Current age in years	y	Current (calendar) time in years
x_0	Minimal age observed in data (see Fig. 1)	y_{00}	Initial time in data (see Fig. 1)
$\bar{x}_0 = \max(x_0, y_{00} - y_b)$	The cohort-specific bounding (minimal) age	$\bar{y}_0 = y_b + \bar{x}_0$	The cohort-specific bounding (minimal) time
$\tau,\ \bar{x}_0 < \tau \le x$	Age at diagnosis	$y_d = y - x + \tau$	Time of diagnosis
$x_{00} = y_{00} - y + x$	The cohort specific age at y_{00}	$y_0 = y - x + x_0$	The cohort-specific time of reaching age x_0
		$y_b = y - x$	Time of birth for a cohort
Survival analysis functions for age-specific prevalence and mortality			
$P_c(x, y_b)$	Probability of being prevalent at age x in cohort with birth time y_b		
$I_c(\tau, y_b)$	Incidence density function for birth cohort y_b. The normalization rule for the density is $\int_{\bar{x}_0}^{\infty} I_c(\tau, y_b)d\tau = 1$		
$M_c(x, y_b)$	Mortality density function for birth cohort y_b. The normalization rule for the density is $\int_{\bar{x}_0}^{\infty} M_c(x, y_b)dx = 1$		
$\bar{S}(x - \bar{x}_0, \bar{x}_0, \bar{y}_0)$	Survival function and respective density function of a patient group formed at age $\bar{x}_0$ and time $\bar{y}_0$, diagnosed before $\bar{x}_0$ and survived to x (i.e., living $x - \bar{x}_0$ years after cohort forming)		
$\bar{f}_c(x - \bar{x}_0, \bar{x}_0, \bar{y}_0) = -\bar{S}'_x(x - \bar{x}_0, \bar{x}_0, \bar{y}_0)$			
$S(x - \tau, \tau, y_d)$	Survival function and respective density function of a patient group diagnosed at age τ and time y_d, and survived to x (i.e., living $x - \tau$ years after cohort forming)		
$f_c(x - \tau, \tau, y_d) = -S'_x(x - \tau, \tau, y_d)$			
$S_t(x - \tau, \tau, y_d)$ and $\mu(x, y)$	Survival function and respective mortality hazard function in the general population for the cohort formed at age τ and y_d, i.e., $S_t(0, \tau, y_d) = 1$.		
Survival analysis functions for age-adjusted prevalence and mortality			
$P(x, y)$	Prevalence at age x and time y		
$I(x, y)$	Incidence hazard function at age x and time y		
$M(x, y)$	Incidence based mortality hazard function at age x and time y		
$P(y)$	Age-adjusted prevalence at time y		
$M(y)$	Age-adjusted incidence-based mortality hazard function at time y		
$S^r(x - \tau, \tau, y_d)$	Relative survival and respective density function of a patient group diagnosed at age τ and year y_d, and reached age x (i.e., living $x - \tau$ years after diagnosis and cohort forming)		
$f^r(x - \tau, \tau, y_d) = -S^{r\prime}_x(x - \tau, \tau, y_d)$			
$\bar{S}^r(x - \bar{x}_0, \bar{x}_0, \bar{y}_0)$	Relative survival and respective density function of a patient group formed at age $\bar{x}_0$ and time $\bar{y}_0$, diagnosed before $\bar{x}_0$, and survived to age x (i.e., living $x - \bar{x}_0$ years after cohort forming)		
$\bar{f}^r(x - \bar{x}_0, \bar{x}_0, \bar{y}_0) = -\bar{S}^{r\prime}_x(x - \bar{x}_0, \bar{x}_0, \bar{y}_0)$			
$p(x)$	The density of age distribution in a standard year		
$T_{\dots}(y)$ and $\hat{T}_{\dots}(y)$	Specific contributions to the time trends of age-adjusted disease prevalence and incidence-based mortality (discussed in detail in Interpretation of partitioning components)		

with explicit expressions for terms

$$
\begin{aligned}
T_{p0}(y) &= \int_{x_0}^{x_0+y-y_{00}} P'_y(x_0, y_0)\bar{S}^r(x - x_0, x_0, y_0)p(x)\,dx,\\
T_{p00}(y) &= \int_{x_0+y-y_{00}}^{\infty} P'_y(y_{00} - y + x, y_{00})\bar{S}^r(y - y_{00}, y_{00} - y + x, y_{00})p(x)\,dx,\\
T_{\bar{S}}(y) &= \int_{x_0}^{x_0+y-y_{00}} P(x_0, y_0)\bar{S}^{r\prime}_y(x - x_0, x_0, y_0)p(x)\,dx,\\
T_{\bar{S}00}(y) &= \int_{x_0+y-y_{00}}^{\infty} P(x_{00}, y_{00})\bar{S}^{r\prime}_y(y - y_{00}, y_{00} - y + x, y_{00})p(x)\,dx,\\
T_{X00}(y) &= \int_{x_0+y-y_{00}}^{\infty} I(x_{00}, y_{00})S^r(y - y_{00}, x_{00}, y_{00})p(x)\,dx,\\
T_{inc}(y) &= \int_{x_0}^{\infty}\int_{\max(x_0,x_{00})}^{x} I'_y(\tau, y_d)S^r(x - \tau, \tau, y_d) \times p(x)\,d\tau dx,\\
T_S(y) &= \int_{x_0}^{\infty}\int_{\max(x_0,x_{00})}^{x} I(\tau, y_d)S^{r\prime}_y(x - \tau, \tau, y_d) \times p(x)\,d\tau dx.
\end{aligned}
\tag{13}
$$

Derivative of mortality includes nine terms

$$
\begin{aligned}
M'_y(y) &= \hat{T}_\mu(y) + \hat{T}_P(y) + \hat{T}_{p0}(y) + \hat{T}_{p00}(y) + \hat{T}_{\bar{S}}(y)\\
&\quad + \hat{T}_{\bar{S}00}(y) + \hat{T}_{X00}(y) + \hat{T}_{inc}(y) + \hat{T}_S(y)
\end{aligned}
\tag{14}
$$

with

$$
\begin{aligned}
\hat{T}_\mu(y) &= \int_{x_0}^{\infty} \mu'_y(x, y)P(x, y)p(x)\,dx,\\
\hat{T}_P(y) &= \int_{x_0}^{\infty} \mu(x, y)P'_y(x, y)p(x)\,dx,\\
\hat{T}_{p0}(y) &= \int_{x_0}^{x_0+y-y_{00}} P'_y(x_0, y_0)\bar{f}^r(x - x_0, x_0, y_0)p(x)\,dx,\\
\hat{T}_{p00}(y) &= \int_{x_0+y-y_{00}}^{\infty} P'_y(y_{00} - y + x, y_{00})\bar{f}^r(y - y_{00}, x_{00}, y_{00})p(x)\,dx,\\
\hat{T}_{\bar{S}}(y) &= \int_{x_0}^{x_0+y-y_{00}} P(x_0, y_0)\bar{f}^{r\prime}_y(x - x_0, x_0, y_0)p(x)\,dx,\\
\hat{T}_{\bar{S}00}(y) &= \int_{x_0+y-y_{00}}^{\infty} P(x_{00}, y_{00})\bar{f}^{r\prime}_y(y - y_{00}, x_{00}, y_{00})p(x)\,dx,\\
\hat{T}_{X00}(y) &= \int_{x_0+y-y_{00}}^{\infty} I(x_{00}, y_{00})f^r(y - y_{00}, x_{00}, y_{00})p(x)\,dx,\\
\hat{T}_{inc}(y) &= \int_{x_0}^{\infty}\int_{\max(x_0,x_{00})}^{x} I'_y(\tau, y_d)f^r(x - \tau, \tau, y_d)p(x)\,d\tau dx,\\
\hat{T}_S(y) &= \int_{x_0}^{\infty}\int_{\max(x_0,x_{00})}^{x} I(\tau, y_d)f^{r\prime}_y(x - \tau, \tau, y_d)p(x)\,d\tau dx.
\end{aligned}
\tag{15}
$$

Non-trivial technical aspects of derivation of the derivatives (13) and (15) are discussed in Appendix B.

2.2. Interpretation of partitioning components

The three terms contributing to disease prevalence in Eq. (11) correspond to the contributions of individuals with disease onset (i) before $x_0 = 65$ (the age of eligibility for Medicare coverage

for the majority of the general population) for the cohorts with $y \geq y_{00} + x - x_0$ (i.e., cohorts below the bisecting line the Lexis diagram in Fig. 1), (ii) before $y_{00} = 1992$ for the cohorts $y < y_{00} + x - x_0$ (i.e., cohorts above the bisecting line in Fig. 1), and (iii) after $x_0 = 65$ and after $y_{00} = 1992$ (i.e., in the shaded area in the Lexis diagram in Fig. 1). Mortality in Eq. (11) has four terms, $M_{P\mu}(y)$ and three others: $M_0(y)$, $M_{00}(y)$, and $M_{is}(y)$ which have the same meaning as the three equivalent terms in prevalence. These three terms represent the mortality rates of individuals with disease onset before x_0, before y_{00}, and after both x_0 and y_{00} respectively. These terms are expressed in terms of relative survival and therefore represent the mortality of individuals with the disease relative to the mortality in the general population. The additional term in Eqs. (11), $M_{P\mu}(y)$, represents mortality for the prevalent population with the mortality rate as in the general population. In sum, Eq. (11) models two components of mortality: (i) the effect of prevailing trends in the general population and (ii) the effect of relative mortality in individuals with the disease.

The time trend of disease prevalence, represented by the first derivative of the age-adjusted prevalence, has seven terms. The main contributions are $T_{inc}(y)$ and $T_S(y)$ that reflect effects of trends in disease incidence and survival after the disease onset. Occurrence of five other terms reflects the fact that we observe individual follow-up not from their birth date. They can be combined in two terms reflecting the effects on two bounds $x = x_0$ and $y = y_{00}$: $T_0(y) = T_{p0}(y) + T_{\bar{S}}(y)$ and $T_{00}(y) = T_{p00}(y) + T_{\bar{S}00}(y) + T_{X00}(y)$, respectively. The terms $T_{p0}(y)$ and $T_{\bar{S}}(y)$ reflect the effects of time trends in initial prevalence (i.e., prevalence at $x_0 = 65$) and trends in survival of these individuals. The contributions of these terms can be considered separately if the respective hypotheses are of interest. The three terms contributing to $T_{00}(y)$ (i.e., $T_{p00}(y)$, $T_{\bar{S}00}(y)$, and $T_{X00}(y)$) are the only terms contributing to disease prevalence that survive in the limit $y \to y_{00}$. They are responsible for the reconstruction of the correct derivative in the region of $y \sim y_{00}$. Specifically, the first and second terms characterize the time trend in initial prevalence and survival for $y = y_{00}$, respectively. The last term equals age-adjusted incidence rate in the limit $y \to y_{00}$. Its occurrence reflects the lack of information about incidence before y_{00}. With time the fraction of unknown information about incidence goes down and the contribution from this term to the total time trends decreases.

Similarly, seven of the nine terms in the decomposition of mortality can be combined into four terms: $\hat{T}_{inc}(y)$ and $\hat{T}_S(y)$ represent the effects of incidence and survival for individuals diagnosed after x_0 and y_{00}, while $\hat{T}_0(y) = \hat{T}_{p0}(y) + \hat{T}_{\bar{S}}(y)$ and $\hat{T}_{00}(y) = \hat{T}_{p00}(y) + \hat{T}_{\bar{S}00}(y) + \hat{T}_{X00}(y)$ reflect the effects on two bounds $x = x_0$ and $y = y_{00}$. Two additional terms occurring in the formula for mortality are $\hat{T}_\mu(y)$ and $\hat{T}_P(y)$. They reflect the effects of trends in mortality in the general population and in prevalence of the given disease.

3. Statistical estimation of model parameters from observational data

The quantities of interest (i.e., $T_{\dots}(y)$ and $\hat{T}_{\dots}(y)$) are expressed in terms of derivatives of survival analysis functions in respect of time. In our approach, we use explicit analytic parameterization for all functions for which derivatives are needed. An alternative approach based on numerical differentiation would require us to deal with numerical instabilities typical for numerical evaluation of derivatives. Since integration is performed numerically, the integrand must be calculated with maximal accuracy—this condition is satisfied by our approach involving analytic differentiation of the parametric models of these functions. Specifically we need to develop and estimate three disease-specific models for a specific disease which are involved in the expressions for prevalence and mortality as well as their derivatives: (i) models for prevalence at x_0 (i.e., prevalence at the starting age of observation) and all years $y \geq y_{00}$, and prevalence at y_{00} (i.e., at the beginning year of observation) and all ages $x \geq x_0$, (ii) the model for the incidence rate for all $x \geq x_0$ and $y \geq y_{00}$; and (iii) models for relative survival of individuals prevalent at x_0, prevalent at y_{00}, and incident at $x > x_0$ and $y > y_{00}$. Furthermore, for the modeling of incidence-based mortality, models for mortality in the general population need to be developed.

We use individual medical records from a nationally representative 5% sample of Medicare beneficiaries age 65+ to estimate the model parameters for the models enumerated above. Medicare data provide individual records for individuals above age 65 (i.e., $x_0 = 65$) and starting in 1992 (i.e., $y_{00} = 1992$). Collection of all records with the disease-specific ICD-9-codes for an individual allows us to reconstruct individual disease-specific trajectories and then create the following datasets for further analyses using the methods from Akushevich et al. (2012):

- D1: Prevalence for boundaries of the region (Fig. 1), i.e., one-year-specific prevalence rates (i) for x_0 and all years $y \geq y_{00}$ and (ii) for y_{00} and all ages $x \geq x_0$.
- D2: Incidence rates in one-year groups over age and year.
- D3: Individual survival times. The dataset contains individual records including age and year at the first record interpreted as incident or prevalent cases and time to death/censoring. For prevalent cases $x = x_0$ or $y = y_{00}$.
- D4: Prevalence and mortality in one-year groups over age and year. This dataset will be only used for comparison to the results of modeling of these measures.

Estimation strategy of model parameters involves B-splines in order to evaluate y- and x_{00}-dependences occurring in the expressions for prevalence, mortality, and their derivatives. An important feature of B-splines necessary for our study is that they allow for the calculation of derivatives explicitly and without additional simplifying assumptions. Other dependencies such as age-dependencies of incidence, survival, and mortality in the general population as well as survival time dependence are modeled by appropriate (known or empirically based) models adopted for them, such as the linear model of disease incidence, the Gompertz model for age patterns for mortality in the general population, and the Weibull model for survival time distribution.

3.1. Model for prevalence at boundaries

First, the y_0-dependence of initial prevalence (dataset D1) are modeled using B-splines as $P(x_0, y_0) = \sum_i \alpha_i B_{i,n}(y_0)$, where n is the degree of B-splines ($n = 3$ in our analysis) and i runs over all B-splines the number of which is defined by the number of used knots. The functions $B_{i,n}(y_0)$ are polynomial functions completely known when the sets of knots are fixed and parameters α_i are subject for estimation. The first derivative of $P(x_0, y_0)$ is then explicitly calculated because $B'_{i,n}(y_0)$ is represented in terms of B-splines of a lower degree for that we also have explicit representation. Note also that the approach gives the derivative of $P(x_0, y_0)$ with respect to y_0, however since $y_0 = y - x + x_0$ it is equal to the derivative with respect to y: $dP(x_0, y_0)/dy_0 = dP(x_0, y_0)/dy$. Similarly, B-splines provide the fit for x_{00} (where $x_{00} = y_{00} - y + x$) dependence of $P(x_{00}, y_{00})$ together with the first derivative in respect of x_{00} thus providing the derivative in respect of y: $dP(x_{00}, y_{00})/dx_{00} = -dP(x_{00}, y_{00})/dy$. Empirical estimates and the B-spline models for both y_0-dependence of $P(x_0, y_0)$ and x_{00}-dependence of $P(x_{00}, y_{00})$ are shown in Fig. 2.

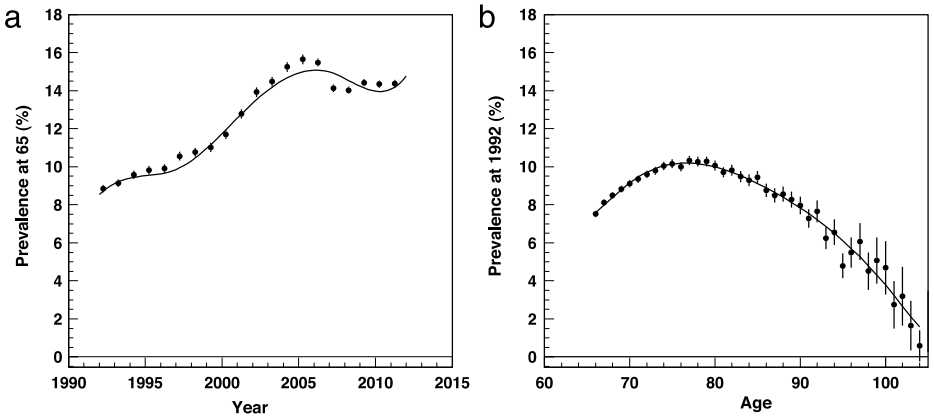

Fig. 2. Prevalence at 65, $P(x_0, y_0)$ vs. y_0, (left panel) and at 1992, $P(x_{00}, y_{00})$ vs. x_{00}, (right panel) of diabetes: empiric estimates (dots) and B-spline model (solid line).

3.2. Model for incidence

Assume that for each y_d the age-dependence of incidence rates $I(\tau, y_d)$ from dataset D2 is explicitly parameterized through the sets of model parameters $\beta_{inc} = \{\beta_i^{inc}(y_d)\}$ dependent on y_d (e.g., linearly $I(\tau, y_d) = \tau\beta_1^{inc}(y_d) + \beta_2^{inc}(y_d)$), and y_d-dependence of each parameter $\beta_i^{inc}(y_d)$ is fitted by B-splines providing the first derivative $d\beta_i^{inc}(y_d)/dy_d = d\beta_i^{inc}(y_d)/dy$. Thus

$$\frac{dI(\tau, y_d)}{dy} = \sum_i \frac{\partial I(\tau, y_d)}{\partial \beta_i^{inc}} \frac{d\beta_i^{inc}(y_d)}{dy_d}.$$

Fig. 3 presents the age-dependence of age-specific rates for two selected years (left panel) and y_d-dependence of age-adjusted incidence rates (right panel) together with the B-spline models fitting y_d-dependences of age-specific rates with subsequent age-adjustment for the second case. The results presented in Fig. 3 justify the choice of the linear model for the age-specific rates of diabetes (another model can be chosen for another disease). Note that the age adjusted rates can be represented by a linear model only approximately and the spline approximation that provides partial smoothing of this effect could be an alternative.

3.3. Models for survival

The models describing age- and survival-time-dependences for the three specific relative survival functions $\bar{S}^r(x - x_0, x_0, y_0)$, $\bar{S}^r(y - y_{00}, x_{00}, y_{00})$, and $S_d^r(x - \tau, \tau, y_d)$ have to be specified, parameterized, and estimated. We use the approach based on maximizing the likelihood for individual survival data (Dickman et al., 2004), which can be outlined in general terms as follows and specified for the three relative survival functions below. An individual i in dataset D3 is characterized by (i) the age of diabetes diagnosis or initial age of follow-up (x_{0i}), (ii) final age of follow-up (x_i), and (iii) the death/censoring indicator d_i at age x_i. Denoting survival function for an individual i as $S(x_i, x_{0i})$ and using the standard likelihood for total survival $L = \prod_i (h(x_i))^{d_i} S(x_i, x_{0i})$ and the definitions of relative survival, $S(x_i, x_{0i}) = S_t(x_i, x_{0i})S^r(x_i, x_{0i}; \beta)$, and respective hazard functions $h(x) = h_t(x) + h^r(x, \beta)$, we construct the log likelihood as

$$l(\beta) = -\sum_i \int_{x_{0i}}^{x_i} h_t(u)du - \sum_i \int_{x_{0i}}^{x_i} h^r(u, \beta)du + \sum_i d_i \log(h_t(x_i) + h^r(x_i, \beta)). \quad (16)$$

Here β is the set of parameters for the relative survival and respective hazard. The first term does not depend on β and therefore can be omitted. The only item that we need to know about the general population is the population hazards at the age of death for all individuals in the datasets. This information is obtained from the Human Mortality Database.

Specific parameterization is required to describe the age- and survival-time-dependences of relative survival functions. We assume that the Weibull model is flexible enough (Carroll, 2003; Zhu et al., 2011) and can be applied for the three relative survival functions involved in (7) and (8).

For $\bar{S}^r(x-x_0, x_0, y_0)$ we use $\bar{S}^r(x-x_0, x_0, y_0) = \exp(-\exp(\sigma^{-1}(\log(x-x_0)-\mu)))$ in which parameters $\mu = \mu(y_0)$ and $\sigma = \sigma(y_0)$ are estimated for each y_0 using maximizing the likelihood (16), and then y_0-dependences of μ and σ are fitted by B-splines providing derivatives $d\mu/dy_0 = d\mu/dy$ and $d\sigma/dy_0 = d\sigma/dy$. Thus,

$$\frac{d\bar{S}^r(x - x_0, x_0, y_0)}{dy} = \frac{\partial \bar{S}^r(x - x_0, x_0, y_0)}{\partial \mu} \frac{d\mu}{dy_0} + \frac{\partial \bar{S}^r(x - x_0, x_0, y_0)}{\partial \sigma} \frac{d\sigma}{dy_0}.$$

The partial derivatives are calculated explicitly and derivatives of μ and σ in respect to y_0 are provided by B-splines.

Similarly, we use $\bar{S}^r(y-y_{00}, x_{00}, y_{00}) = \exp\big(-\exp(\sigma^{-1}(\log(y-y_{00}) - \mu))\big)$, and parameters $\mu = \mu(x_{00})$ and $\sigma = \sigma(x_{00})$ depend on x_{00} and are estimated using (16). The dependences are given by B-splines and $d\mu/dx_{00} = -d\mu/dy$ and $d\mu/dx_{00} = -d\mu/dy$. Now the relative survival function depends on y explicitly, therefore

$$\frac{d\bar{S}^r(y - y_{00}, x_{00}, y_{00})}{dy} = \frac{\partial \bar{S}^r(y - y_{00}, x_{00}, y_{00})}{\partial y} - \frac{\partial \bar{S}^r(y - y_{00}, x_{00}, y_{00})}{\partial \mu} \frac{d\mu}{dx_{00}} - \frac{\partial \bar{S}^r(y - y_{00}, x_{00}, y_{00})}{\partial \sigma} \frac{d\sigma}{dx_{00}}.$$

Finally, we use $S^r(x-\tau, \tau, y_d) = \exp(-\exp(\sigma^{-1}(\log(x-\tau)-\mu)))$ where the dependence of μ and σ on τ are explicitly represented through the quadratic function of τ with the sets of parameters $\beta_\mu = \{\beta_\mu^i\}$ and $\beta_\sigma = \{\beta_\sigma^i\}$ and each parameter $\beta_{\mu,\sigma}^i$ is y_d-specific (i.e., estimated for each year y_d). Then y_d dependence of each parameter $\beta_{\mu,\sigma}^i$ is fitted by B-splines providing respective derivatives $d\beta_{\mu,\sigma}^i/dy_d = d\beta_{\mu,\sigma}^i/dy$. Therefore,

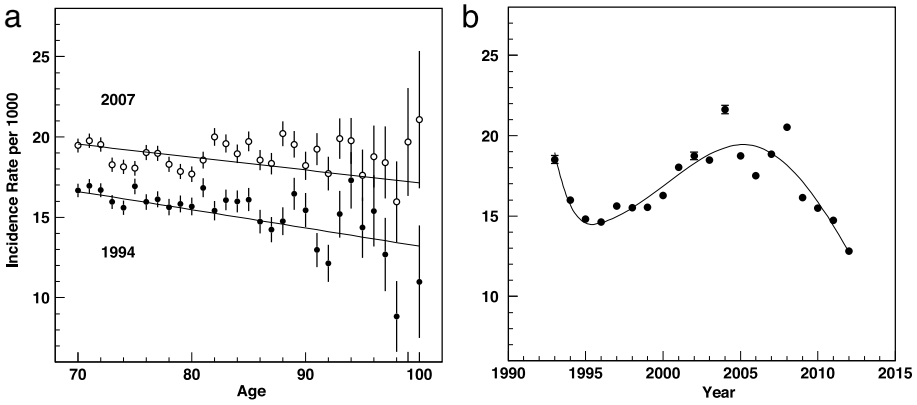

Fig. 3. Incidence rate of type 2 diabetes $I(\tau, y_d)$: age pattern for selected years of diagnosis (left plot) and year-at-diagnosis patterns for age-adjusted rate (right panel). Dots show empiric estimates and curves show the models: linear model for age-patterns (left plot) and B-spline model for age-adjusted rates (right plot).

$$\frac{dS^r_d(x-\tau,\tau,y_d)}{dy} = \sum_i \frac{\partial S^r_d(x-\tau,\tau,y_d)}{\partial \beta^i_\mu}\frac{d\beta^i_\mu}{dy_d} + \sum_i \frac{\partial IS^r_d(x-\tau,\tau,y_d)}{\partial \beta^i_\sigma}\frac{d\beta^i_\sigma}{dy_d}.$$

Fig. 4 presents the results for the relative survival functions. Projections of time survival for selected ages and years are shown for all three survival functions involving in (7) and (8). Note that empirical estimates of relative survival are not necessary for our modeling so they are not presented in Fig. 4.

4. Partitioning for diabetes prevalence and mortality and their time trends

Application of estimated models to 5%-Medicare data resulted in predicted age-adjusted prevalence, $P(y)$, and incidence-based mortality, $M(y)$, according Eqs. (9) and (10). Their patterns are presented by thick lines in the upper panels of Fig. 5; the actual empirical patterns (dots) are provided for comparison. The three terms ($P_0(y)$, $P_{00}(y)$, and $P_{is}(y)$) contributing to prevalence and four terms ($M_{P\mu}(y)$, $M_0(y)$, $M_{00}(y)$, and $M_{is}(y)$) contributing to the mortality rate according to Eq. (11) are also shown. The curves in these plots are marked by labels corresponding to subscripts of the respective contributions from Eq. (11). Excellent agreement between the theoretical predictions and the empirical estimates is detected for both prevalence and the incidence-based mortality of diabetes.

The term with the double integral in Eq. (9), $P_{is}(y)$, that contains the product of incidence and relative survival gives the most essential contribution to diabetes prevalence. The contribution of prevalence at 1992, $P_{00}(y)$, decreases with time because of two reasons: (i) mortality of individuals prevalent at 1992 is larger than mortality in the general population and (ii) the relative contribution of the region above the bisecting line to the integral with respect to x (see Fig. 1) decreases with time. In contrast the contribution of prevalence at age 65, $P_0(y)$, goes up because of increased prevalence at 65 with time (Fig. 2) and the increased contribution of the region below the bisecting line to the integral with respect to x in (9). The major contribution to the incidence-based mortality is the term containing the product of prevalence and mortality in the general population (i.e., $M_{P\mu}(y)$). This term would be the only contribution if the mortality rates of individuals with diabetes and of the general population are the same. The gap between this term and mortality, given by the thick line (i.e., $M(x, y)$), is due to three remaining contribution terms showing the effects of the individuals prevalent at the boundaries ($M_0(y)$ and $M_{00}(y)$ represented by the curves '0' and '00') and the individuals diagnosed during follow-up ($M_{is}(y)$ and the curve 'is'). The patterns of these curves reproduce those observed for prevalence and are explained by the same reasoning.

Both prevalence and mortality increase with time so their derivative over calendar time is positive. Explicit calculations allowed us to evaluate time patterns of the components responsible for these trends of prevalence (12), (13) and mortality (14), (15). The total derivative of prevalence ($P'_y(y)$ and thick curve) largely reproduces the shape of the curve marked by 'inc', that is the total prevalence trend is defined primarily by the dynamics of diabetes incidence. The term containing the derivative of incidence, $T_{inc}(y)$, is negative, i.e., incidence is decreasing over time driving the prevalence downwards. The effect of survival ($T_S(y)$ and the curve marked by "S") pushes the prevalence upwards reflecting increased life-span of patients with diabetes. The curve marked by "0" contains two contributions, $T_0(y) = T_{p0}(y) + T_{\bar{S}}(y)$, the first representing prevalence at 65 is dominant and increasing driving the total prevalence up. Another contribution $T_{\bar{S}}(y)$ reflecting survival of patients prevalent at 65 is positive (similarly to $T_S(y)$) and small. The remaining contribution marked by "00" comprises all effects related to the boundary at y_{00}, i.e., $T_{00}(y) = T_{p00}(y)+T_{\bar{S}00}(y)+T_{X00}(y)$. This contribution is largely technical because it reflects the fraction of the effects coming from incidence and survival trends before 1992. As expected this fraction decreases with time. In sum, the total prevalence increases over time as the three contributions pulling the prevalence up overpower the downward effect of incidence.

The presentation of the partitioning of mortality trends (Fig. 5(d)) largely reflects the picture obtained for prevalence in Fig. 5(c). The most important contribution to the mortality trend is the term containing the derivative of prevalence $\hat{T}_P(y)$ (marked by "P", thick dashed curve). Its shape reproduces the shape of total prevalence change (thick line in Fig. 5(c)) and deviates only because of the factor containing the mortality rate in the general population which is time dependent. Other curves on Fig. 5(d) reflect the effects of relative survival and respective mortality. Although the size of the effects is not large (as follows from Fig. 5(b)), they can have significant contributions to the mortality time trend (Fig. 5(d)). Their signs, sizes, and time trends reflect what we observed for prevalence with the exception of the terms reflecting survival that change sign. For example, incidence and survival result in decreasing mortality. However, mortality still increases. This is a consequence of negative tendencies in past (i.e., before 1992) that are represented by the term "00", i.e., $\hat{T}_{00}(y) = \hat{T}_{p00}(y) + \hat{T}_{\bar{S}00}(y) + \hat{T}_{X00}(y)$. An additional term, $\hat{T}_\mu(y)$, is marked by μ (dashed curve), this reflects the effect of the general change in mortality in the general population, that is the mortality of patients with diabetes not related to diabetes itself (i.e., death due to other causes) is going down just as in the general population.

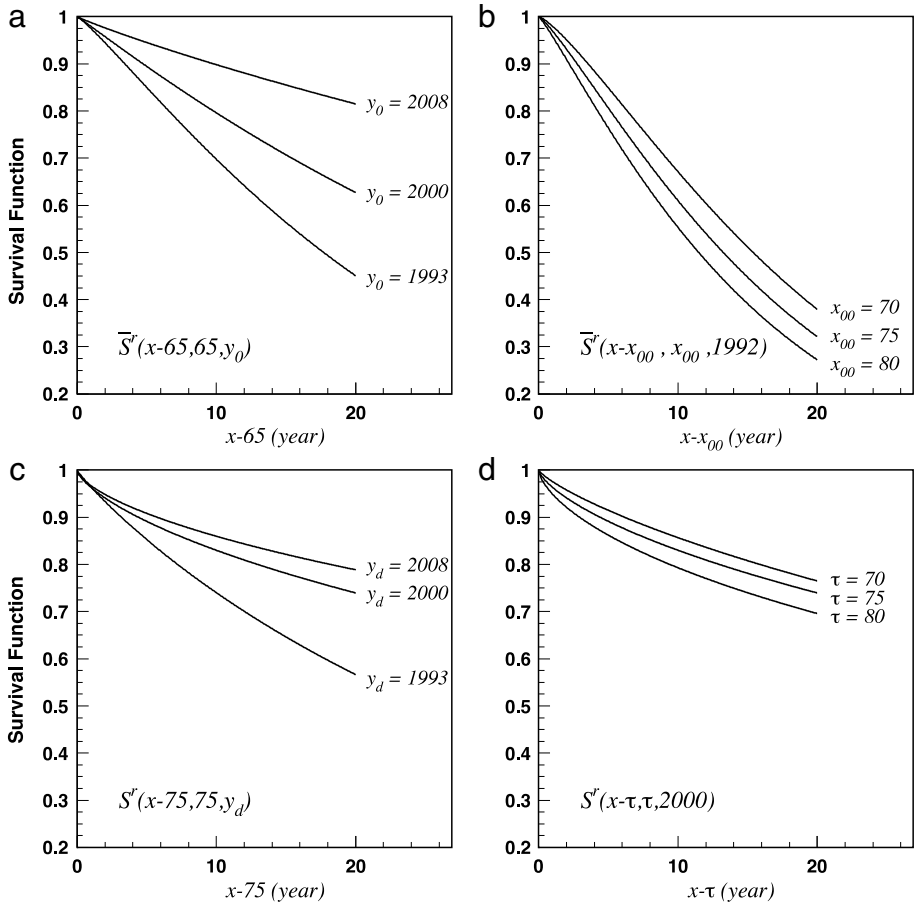

Fig. 4. Relative survival functions for selected ages and years at diagnoses. Specifically, (a) relative survival vs. time after 65 for individuals prevalent at 65 for three y_0 (years at $x_0 = 65$), (b) relative survival vs. time after x_{00} for individuals prevalent at 65 for three x_{00} (age at $y_{00} = 1992$), (c) relative survival vs. time after diagnosis for individuals diagnosed at age 75 in three years of diagnosis y_d, and (d) relative survival vs. time after diagnosis for individuals diagnosed in 2000 in three ages of diagnosis τ.

5. Discussion and conclusion

In this study, we developed an approach for the modeling of disease prevalence and incidence-based mortality (i.e., mortality for individuals who had a diagnosis earlier in life). The model provides analytical expressions for these epidemiologic characteristics which allows for the analysis of the relative contributions of incidence and survival as important components of total prevalence and incidence-based mortality.

The approach provides expressions for the partitioning of the time trends of these quantities. All of the components that are responsible for trends in disease prevalence and mortality are evaluated through explicit expressions. These components include disease incidence and survival as well as effects at the boundaries of the region available for analysis (in the case of 5%-Medicare we observe the effects after 1992 for individuals aged 65+). For mortality, additional information about the mortality rate in the general population is required.

The results of the partitioning analyses are presented in Fig. 5 and described in detail in "Partitioning for Diabetes Prevalence and Mortality and Their Time Trends" earlier in this text. Based on this analysis we can conclude that (i) the theory describes empirical estimates for prevalence and mortality with good accuracy, (ii) among the possible contributions $T_{...}(y)$ and $\hat{T}_{...}(y)$ in Eqs. (12) and (14), the contributions of incidence and survival after age 65 have the greatest effect on diabetes prevalence and diabetes-related mortality, and (iii) the dynamics of diabetes prevalence and mortality are generated by causes consistent with improvements in population health: decreased incidence and improved survival.

Use of our methodology offers new opportunities in public health. Researchers obtain the opportunity to clearly identify the sources of observed processes at the level of disease prevalence and mortality. The methodology presented in this paper provides a formal method for the decomposition of an observed trend in prevalence and incidence-based mortality into their constituent parts. Practically, this can be used as a public health planning tool, to identify areas of concern which, either due to the size of the effect, the direction of the trend, or the observed rate of change, require targeted attention from health agencies. Furthermore, over time improvements in diagnostic technology and the body of knowledge on the pathological characteristics of a disease lead to improved ascertainment (i.e. the ability to identify the presence of a disease) and more tightly defined guidelines for making a valid diagnosis. Improved ascertainment is likely to lead to an increase in the incidence of a disease as individuals who were previously left undiagnosed are identified. The effect of changes in diagnostic guidelines is more ambiguous as depending on whether elements were added or removed from the definition of a valid diagnosis incidence and by extent prevalence could be pushed in either direction. A standardized method for the partitioning of an existing prevalence trend into the time-trends of its components and, more importantly, the relative strength of the contribution of each component over time, will aid in both correctly assessing the relative success of a health intervention and identifying time-periods of special interest for more in depth analysis (e.g. a sharp spike in the relative strength of the contribution associated with incidence could indicate either an area of public health interest, or an improvement in ascertainment).

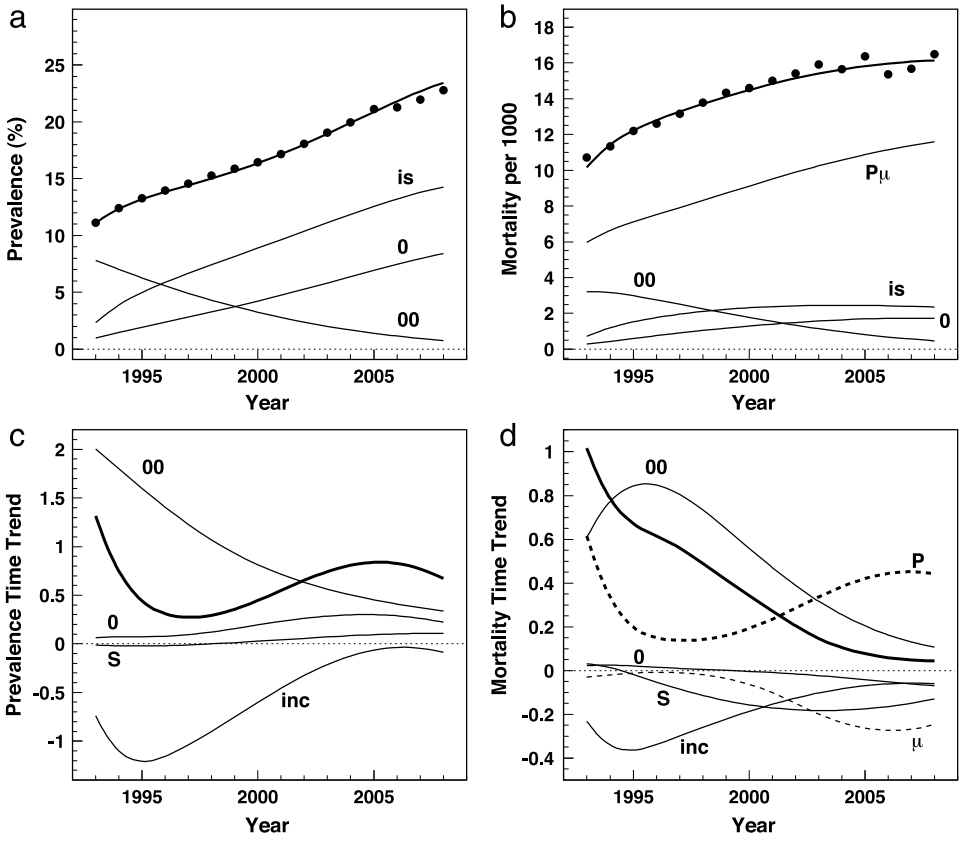

Fig. 5. Prevalence and incidence-based mortality, their time trends as well as partitioning of these measures: (a) type 2 diabetes prevalence, (b) incidence-based mortality, (c) time trend of prevalence and (d) time trend of the incidence-based mortality. Dots in upper plots show empiric estimates for prevalence and mortality. Thick curves show theoretical predictions given by Eqs. (9), (10), (12)–(13), and (14)–(15), respectively. Thin curves represent components of the theoretical predictions (Eqs. (11)–(15)). Label for each thin curve exactly corresponds to the subscript of respective term in Eqs. (11) for upper plots and (13), (15) for lower plots.

Our approach lends itself to multiple natural generalizations allowing for the estimation and partitioning of quantities such as (i) the effects of disease-specific medical costs (respective costs are added into integrand) (Akushevich et al., 2011, 2016), (ii) the effects of recovery and/or long-term remission (respective survival functions have to be used as additional factors) (Akushevich et al., 2013b), and (iii) the effects of complications for patients with a specific disease (specific patient selections and respective changes in mathematical formalism have to be included) (Akushevich et al., 2013a; Yashkin et al., 2015). The expressions for prevalence and mortality can also be used in improving the accuracy of future projections in health forecasting.

The components contributing to time trends are obtained not using a fitting procedure and/or maximum likelihood, but direct calculation using expressions representing each component as continuous functions of available data. Statistical estimates of parameters characterizing the time patterns of prevalence at the bounds, incidence and survival are obtained using B-splines. Since B-splines provide consistent estimates of model parameters (Strawderman and Tsiatis, 1996) and the trend components are continuous functions of the B-spline parameters, the estimates of the trend components are consistent. The estimates are largely model independent; therefore, the risk of model misspecification is minimal. The only model we use is the Weibull model for survival time. The model is quite flexible and its choice is not critical for the estimation procedure: two-dimensional splines for survival can be used instead.

The level of detail our approach provides is highly dependent on the length and scope present in the data. When Medicare data or a dataset of a similar size is used, statistical uncertainties are not expected to be large. In the general case the statistical uncertainty has to be estimated using a bootstrapping approach or through analytic estimates of error propagation. However, systematic uncertainties (biases) could be noticeable. Possible sources for the systematical uncertainties include (i) the possibility of non-precise separation of incident and prevalent cases, (ii) the effect of time trends in the fraction of individuals covered by Medicare Advantage (a private alternative to traditional Medicare which does not contribute data to Medicare datasets), and/or (iii) changes in the structure of the population of Medicare beneficiaries due to specific events such as initiation of Medicare coverage of Part D in 2006. Separate research to evaluate the contributions of these factors to the total systematical error is required. In sum, the estimation of the time trend components is consistent and stable, however further investigation of systematic uncertainties would improve overall accuracy of the estimates.

In summary, notable strengths of our approach include: (i) modeling of all components used in our models using explicit expressions; (ii) lack of simplifying assumptions; (iii) stability and consistency of the resulting estimates; and (iv) wide availability of large administrative health data like that used in the study. The application of this approach to the case of diabetes found that both its prevalence and incidence-based mortality increase with time. The primary driving factor of the observed prevalence increase is improved survival and increased prevalence at age 65. The increase in diabetes-related mortality is driven by increased prevalence and unobserved trends beyond the region observed in the data.

Acknowledgment

This study has been supported by the National Institute on Aging (grants R01-AG017473, R01-AG046860, P01-AG043352).

Appendix A. Derivation of expressions for prevalence (1) and mortality (2)

The formulae (1) and (2) can be understood in terms of the numbers of individuals. Let N_0 be the size of a birth cohort and $N_I(\tau)\,\Delta\tau$ be the number of individuals with disease onset within the age period $\Delta\tau$. The total number of sick (and alive) individuals at age x ($N_d(x)$) is the sum of all individuals who survived to age x after diagnoses at τ over all age periods, i.e.,

$$N_d(x) = \sum_n N_I(\tau_n)(\Delta\tau)_n S(x-\tau_n, \tau_n)$$

where $S(x-\tau_n, \tau_n)$ is the survival function of individuals diagnosed at age period $[\tau_n, \tau_n + (\Delta\tau)_n]$ who survived to age x. Considering infinitely small age periods (i.e., $\Delta\tau \to 0$) and defining $P_c(x, y_b) = N_d(x)/N_0$ and $I_c(\tau, y_b) = N_I(\tau)/N_0$, we obtain the formula $P_c(x, y_b) = \int_0^x I_c(\tau, y_b)S(x-\tau, \tau, y_d)d\tau$. Then the formula (1) is obtained when we split the integration region in two parts (from 0 to $\bar{x}_0$ and from $\bar{x}_0$ to x) and use respective notation for the first part: $\int_0^{\bar{x}_0} I_c(\tau, y_b)S(x-\tau, \tau, y_d)d\tau = P_c(\bar{x}_0, y_b)\bar{S}(x-\bar{x}_0, \bar{x}_0, \bar{y}_0)$. Exactly the definition of $\bar{S}(x-\bar{x}_0, \bar{x}_0, \bar{y}_0)$ is:

$$\bar{S}(x-\bar{x}_0, \bar{x}_0, \bar{y}_0) = \frac{\int_0^{\bar{x}_0} I_c(\tau, y_b)S(x-\tau, \tau, y_d)d\tau}{\int_0^{\bar{x}_0} I_c(\tau, y_b)S(\bar{x}_0-\tau, \tau, y_d)d\tau}.$$

The survival function in the numerator can split as $S(x-\tau, \tau, y_d) = S(x-\bar{x}_0, \bar{x}_0, \bar{y}_0|\tau)S(\bar{x}_0-\tau, \tau, y_d)$. The function $S(x-\bar{x}_0, \bar{x}_0, \bar{y}_0|\tau)$ is the survival function for a cohort of patients formed at age $\bar{x}_0$ and time $\bar{y}_0$ and diagnosed at age τ, $\tau < \bar{x}_0$. If the dependence on τ is weak for a disease, we can put $S(x-\bar{x}_0, \bar{x}_0, \bar{y}_0|\tau) \approx S(x-\bar{x}_0, \bar{x}_0, \bar{y}_0)$, and therefore obtain $\bar{S}(x-\bar{x}_0, \bar{x}_0, \bar{y}_0) \approx S(x-\bar{x}_0, \bar{x}_0, \bar{y}_0)$. Thus, the difference between empiric estimates or estimated models for the two survival functions reflects the force of dependence of $S(x-\bar{x}_0, \bar{x}_0, \bar{y}_0|\tau)$ on τ.

Let the number of individuals diagnosed at the age interval $[\tau, \tau+\Delta\tau]$ and then died during age interval $[x, x+\Delta x]$ be $N_{IM}(\tau, x)$. The total number of individuals died in the age interval $[x, x+\Delta x]$ is defined through the density of the incidence-based mortality $M(x, y_b)$ and equals:

$$M(x, y_b)\Delta x = \sum_n N_{IM}(\tau_n, x)(\Delta\tau)_n\Delta x = \sum_n N_I(\tau_n)\frac{N_{IM}(\tau_n, x)}{N_I(\tau_n)}(\Delta\tau)_n\Delta x.$$

Considering infinitely small age periods (i.e., $\Delta\tau \to 0$ and $\Delta x \to 0$) and defining $f_c(x-\tau_n, \tau_n, y_d) = \frac{N_{IM}(\tau_n, x)}{N_I(\tau_n)}$ we obtain the formula $M_c(x, y_b) = \int_0^x I_c(\tau, y_b)f_c(x-\tau, \tau, y_d)d\tau$. Splitting the integration region and using definitions of $\bar{f}_c(x-\bar{x}_0, \bar{x}_0, \bar{y}_0)$ similarly as in the case of disease prevalence considered above we obtain Eq. (2).

Appendix B. Derivation of the derivatives of prevalence and mortality

In this Appendix technical aspects of derivation of the derivatives (13) and (15) are discussed. Rewrite Eq. (9) by rewriting integration limits explicitly:

$$\begin{aligned}P(y) = &\int_{x_0}^{y-y_{00}+x_0} P(x_0, y_0)\bar{S}^r(x-x_0, x_0, y_0)p(x)dx\\ &+ \int_{y-y_{00}+x_0}^{\infty} P(x_{00}, y_{00})\bar{S}^r(y-y_{00}, x_{00}, y_{00})p(x)dx\\ &+ \int_{x_0}^{y-y_{00}+x_0}\left(\int_{x_0}^{x} I(\tau, y_d)S_d^r(x-\tau, \tau, y_d)d\tau\right)p(x)\,dx\\ &+ \int_{y-y_{00}+x_0}^{\infty}\left(\int_{y_{00}-y+x}^{x} I(\tau, y_d)S_d^r(x-\tau, \tau, y_d)d\tau\right)p(x)\,dx. \quad (17)\end{aligned}$$

Derivation of the right hand side is based on the Barrow's Fundamental Theorem of Calculus which can be adopted for our case as: $\left(\int_a^x f(x, y)dy\right)'_x = f(x, x) + \int_a^x f'_x(x, y)dy$, i.e., when we need to differentiate a function over an argument that is both in the integration limits and in integrand we have the two terms with and without integration. Differentiation of the first term in $P(y)$ results:

$$P(x_0, y_{00})\bar{S}^r(y-y_{00}, x_0, y_{00})p(y-y_{00}+x_0) + T_{p0}(y) + T_{\bar{S}}(y).$$

The second term of $P(y)$ gives

$$-P(x_0, y_{00})\bar{S}^r(y-y_{00}, x_0, y_{00})p(y-y_{00}+x_0) + T_{p00}(y) + T_{\bar{S}00}(y).$$

As we see first terms of these two contributions cancel in the sum resulting in $T_{p0}(y) + T_{\bar{S}}(y) + T_{p00}(y) + T_{\bar{S}00}(y)$.

Differentiation of the third term in Eq. (17) results:

$$\begin{aligned}&p(y-y_{00}+x_0)\int_{x_0}^{y-y_{00}+x_0} I(\tau, y_{00}-x_0+\tau)\\ &\quad\times S_d^r(y-y_{00}+x_0-\tau, \tau, y_{00}-x_0+\tau)d\tau\\ &\quad+\int_{x_0}^{y-y_{00}+x_0}\left(\int_{x_0}^{x} I'_y(\tau, y_d)S^r(x-\tau, \tau, y_d)d\tau\right)p(x)\,dx\\ &\quad+\int_{x_0}^{y-y_{00}+x_0}\left(\int_{x_0}^{x} I(\tau, y_d)S_y^{r\prime}(x-\tau, \tau, y_d)d\tau\right)p(x)\,dx. \quad (18)\end{aligned}$$

The last term in the expression for $P(y)$ contains y in integration limits for both first and second integrals, therefore the Barrow's theorem has to be applied twice resulting:

$$\begin{aligned}&-p(y-y_{00}+x_0)\int_{x_0}^{y-y_{00}+x_0} I(\tau, y_{00}-x_0+\tau)\\ &\quad\times S_d^r(y-y_{00}+x_0-\tau, \tau, y_{00}-x_0+\tau)d\tau\\ &\quad+\int_{y-y_{00}+x_0}^{\infty}\left(\int_{y_{00}-y+x}^{x} I(\tau, y_d)S_d^r(x-\tau, \tau, y_d)d\tau\right)'_y\\ &\quad\times p(x)\,dx\\ &=-p(y-y_{00}+x_0)\int_{x_0}^{y-y_{00}+x_0} I(\tau, y_{00}-x_0+\tau)\\ &\quad\times S_d^r(y-y_{00}+x_0-\tau, \tau, y_{00}-x_0+\tau)d\tau\\ &\quad+\int_{y-y_{00}+x_0}^{\infty} I(y_{00}-y+x, y_{00})\\ &\quad\times S_d^r(y-y_{00}, y_{00}-y+x, y_{00})p(x)dx\\ &\quad+\int_{y-y_{00}+x_0}^{\infty}\left(\int_{y_{00}-y+x}^{x} I'_y(\tau, y_d)S^r(x-\tau, \tau, y_d)d\tau\right)\\ &\quad\times p(x)\,dx\\ &\quad+\int_{y-y_{00}+x_0}^{\infty}\left(\int_{y_{00}-y+x}^{x} I(\tau, y_d)S_y^{r\prime}(x-\tau, \tau, y_d)d\tau\right)\\ &\quad\times p(x)\,dx. \quad (19)\end{aligned}$$

First terms of (18) and (19) cancel, second term of (19) is $T_{X00}(y)$, and sum of remaining terms (last two terms of (18) and (19)) gives $T_{inc}(y) + T_S(y)$. The sum of surviving terms gives finally the right hand side of Eq. (12).

The calculation for derivative of mortality is similar.

References

Akinbami, L.J., Moorman, J.E., Bailey, C., Zahran, H.S., King, M., Johnson, C.A., Liu, X., 2012. Trends in asthma prevalence, health care use, and mortality in the United States, 2001–2010. NCHS Data Brief 94 (94), 1–8.

Akushevich, I., Kravchenko, J., Akushevich, L., Ukraintseva, S., Arbeev, K., Yashin, A.I., 2011. Medical cost trajectories and onsets of cancer and noncancer diseases in US elderly population. Comput. Math. Methods Med. 2011, 857892.

Akushevich, I., Kravchenko, J., Arbeev, K.G., Ukraintseva, S.V., Land, K.C., Yashin, A.I., 2016. Medical cost trajectories and onset of age-associated diseases. In: Biodemography of Aging. Springer, pp. 143–162.

Akushevich, I., Kravchenko, J., Ukraintseva, S., Arbeev, K., Kulminski, A., Yashin, A.I., 2013a. Morbidity risks among older adults with pre-existing age-related diseases. Exp. Gerontol. 48 (12), 1395–1401.

Akushevich, I., Kravchenko, J., Ukraintseva, S., Arbeev, K., Yashin, A.I., 2012. Age patterns of incidence of geriatric disease in the US elderly population: Medicare-based analysis. J. Am. Geriatr. Soc. 60 (2), 323–327.

Akushevich, I., Kravchenko, J., Ukraintseva, S., Arbeev, K., Yashin, A.I., 2013b. Recovery and survival from ageing-associated diseases. Exp. Gerontol. 48 (8), 824–830.

Bauer, U.E., Briss, P.A., Goodman, R.A., Bowman, B.A., 2014. Prevention of chronic disease in the 21st century: elimination of the leading preventable causes of premature death and disability in the USA. Lancet 384 (9937), 45–52.

Canudas-Romo, V., 2003. Decomposition Methods in Demography. Rozenberg Publishers.

Carroll, K.J., 2003. On the use and utility of the Weibull model in the analysis of survival data. Control. Clin. Trials 24 (6), 682–701.

Chu, K.C., Miller, B.A., Feuer, E.J., Hankey, B.F., 1994. A method for partitioning cancer mortality trends by factors associated with diagnosis: An application to female breat cancer. J. Clin. Epidemiol. 47 (12), 1451–1461.

Coresh, J., Selvin, E., Stevens, L.A., Manzi, J., Kusek, J.W., Eggers, P., Van Lente, F., Levey, A.S., 2007. Prevalence of chronic kidney disease in the United States. Jama 298 (17), 2038–2047.

Dickman, P.W., Sloggett, A., Hills, M., Hakulinen, T., 2004. Regression models for relative survival. Stat. Med. 23 (1), 51–64.

Egan, B.M., Zhao, Y., Axon, R.N., 2010. US trends in prevalence, awareness, treatment, and control of hypertension, 1988–2008. Jama 303 (20), 2043–2050.

Horiuchi, S., Wilmoth, J.R., Pletcher, S.D., 2008. A decomposition method based on a model of continuous change. Demography 45 (4), 785–801.

Mozaffarian, D., Benjamin, E.J., Go, A.S., Arnett, D.K., Blaha, M.J., Cushman, M., Das, S.R., de Ferranti, S., Després, J.-P., Fullerton, H.J., 2016. Executive summary: Heart disease and stroke statistics-2016 update: A report from the American heart association. Circulation 133 (4), 447.

Smith, R.A., Brooks, D., Cokkinides, V., Saslow, D., Brawley, O.W., 2013. Cancer screening in the United States, 2013. CA: Cancer J. Clin. 63 (2), 87–105.

Strawderman, R.L., Tsiatis, A.A., 1996. On consistency in parameter spaces of expanding dimension: an application of the inverse function theorem. Statist. Sinica 917–923.

Thun, M.J., Carter, B.D., Feskanich, D., Freedman, N.D., Prentice, R., Lopez, A.D., Hartge, P., Gapstur, S.M., 2013. 50-year trends in smoking-related mortality in the United States. N. Engl. J. Med. 368 (4), 351–364.

Tunstall-Pedoe, H., Kuulasmaa, K., Mähönen, M., Tolonen, H., Ruokokoski, E., Amouyel, P., 1999. Contribution of trends in survival and coronar y-event rates to changes in coronary heart disease mortality: 10-year results from 37 WHO MONICA Project populations. Lancet 353 (9164), 1547–1557.

Vaupel, J.W., Romo, V.C., 2003. Decomposing change in life expectancy: A bouquet of formulas in honor of Nathan Keyfitz's 90th birthday. Demography 40 (2), 201–216.

Will, J.C., Yuan, K., Ford, E., 2014. National trends in the prevalence and medical history of angina: 1988 to 2012. Circ.: Cardiovasc. Qual. Outcomes 7 (3), 407–413.

Yashkin, A.P., Picone, G., Sloan, F., 2015. Causes of the change in the rates of mortality and severe complications of diabetes mellitus: 1992–2012. Med. Care 53 (3), 268.

Zhu, H.P., Xia, X., Chuan, H.Y., Adnan, A., Liu, S.F., Du, Y.K., 2011. Application of Weibull model for survival of patients with gastric cancer. BMC Gastroenterol. 11 (1), 1.