

TECHNICAL ADVANCE

Open Access



New ways of estimating excess mortality of chronic diseases from aggregated data: insights from the illness-death model

Ralph Brinks^{1,2*}, Thaddäus Tönnies¹ and Annika Hoyer¹

Abstract

Background: Recently, we have shown that the age-specific prevalence of a disease can be related to the transition rates in the illness-death model via a partial differential equation (PDE). The transition rates are the incidence rate, the remission rate and mortality rates from the 'Healthy' and 'Ill' states. In case of a chronic disease, we now demonstrate that the PDE can be used to estimate the excess mortality from age-specific prevalence and incidence data. For the prevalence and incidence, aggregated data are sufficient - no individual subject data are needed, which allows application of the methods in contexts of strong data protection or where data from individual subjects is not accessible.

Methods: After developing novel estimators for the excess mortality derived from the PDE, we apply them to simulated data and compare the findings with the input values of the simulation aiming to evaluate the new approach. In a practical application to claims data from 35 million men insured by the German public health insurance funds, we estimate the population-wide excess mortality of men with diagnosed type 2 diabetes.

Results: In the simulation study, we find that the estimation of the excess mortality is feasible from prevalence and incidence data if the prevalence is given at two points in time. The accuracy of the method decreases as the temporal difference between these two points in time increases. In our setting, the relative error was 5% and below if the temporal difference was three years or less. Application of the new method to the claims data yields plausible findings for the excess mortality of type 2 diabetes in German men.

Conclusions: The described approach is useful to estimate the excess mortality of a chronic condition from aggregated age-specific incidence and prevalence data.

Trial registration: The article does not report the results of any health care intervention.

Keywords: Chronic disease epidemiology, Multi-state model, Prevalence, Incidence, Dementia, Diabetes, Partial differential equation, Bayes estimation

Background

Recently, we have shown that the age-specific prevalence of a health state or disease can be related to the transition rates in the illness-death model via a partial differential equation (PDE) [1, 2]. The transition rates are the incidence rate, the remission rate and mortality rates from the *Healthy* and *Ill* states (Fig. 1). In case of a chronic disease,

i.e. a disease with no remission, this relation can be used to estimate the incidence from a sequence of cross-sectional studies if information about mortality is available [3]. This might be an alternative way to estimate the incidence of a chronic condition in situations where follow-up studies are challenging to conduct or not feasible at all.

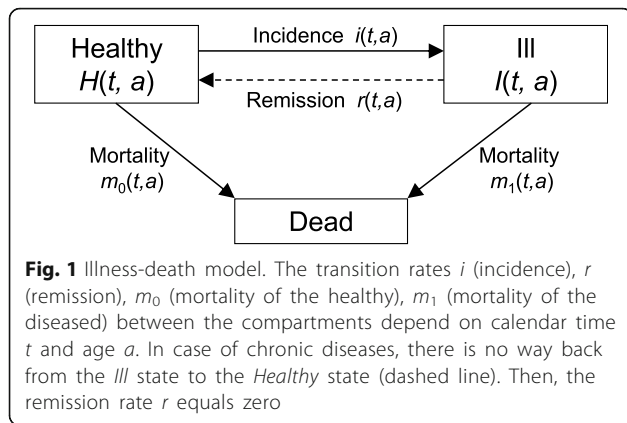
In this article, we demonstrate that it is also possible to estimate excess mortality from age-specific prevalence and incidence of a chronic disease. This can be useful for the analysis of data where it is difficult to observe mortality directly, for instance in disease registers [4] or health insurance claims data where cases of death might

* Correspondence: ralph.brinks@ddz.de

¹Institute for Biometry and Epidemiology, German Diabetes Center, Auf'm Hennekamp 65, 40225 Duesseldorf, Germany

²Department and Hiller Research Unit for Rheumatology, University Hospital Duesseldorf, Moorenstr. 5, 40225 Duesseldorf, Germany





be reported with a delay [5]. Another example where excess mortality of a chronic condition cannot be estimated directly is the US National Health Interview Survey (NHIS) from the National Center for Health Statistics [6]. NHIS is a yearly cross-sectional household interview survey with up to 90,000 participants each year. Usually, participants are followed up for mortality by linkage to the National Death Index. This implies that it is possible to check the vital status of a participant from a previous cross-sectional interview, but it is not possible to decide if a deceased participant who had been disease-free at the interview, has contracted the disease in the period between the cross-section and the date of death. With other words, for a subject disease-free at the interview, it is not possible to determine the disease status at death. Thus, in estimating the mortality it is uncertain to attribute this case to the mortality of the healthy or of the diseased subjects.

To overcome these problems, we examine mathematical relations of the illness-death model and associated PDEs to develop reliable estimators for excess mortality.

Methods

Illness-death model

We consider the illness-death model as shown in Fig. 1. Each subject of the population is in one of the relevant disease states, *Healthy* (with respect to the considered chronic disease) *Ill* or *Dead*. Let the number of people aged a at calendar time t in the *Healthy* and *Ill* states be denoted by $H(t, a)$ and $I(t, a)$, respectively. Subjects can transit from both states into the (absorbing) state *Dead*. The transition rates between the three states are the incidence rate (i), the remission rate (r), the mortality rate of the healthy (m_0) and the mortality rate of the diseased (m_1). These rates usually depend on calendar time t and on age a . Henceforth, we consider only chronic, i.e., irreversible diseases, which is equivalent to a remission rate of zero ($r = 0$).

To develop estimators for the excess mortality $\Delta m = m_1 - m_0$, we use mathematical relations between the incidence, prevalence $p(t, a) = I(t, a) / \{I(t, a) + H(t, a)\}$ and the mortality rates in the illness-death model.

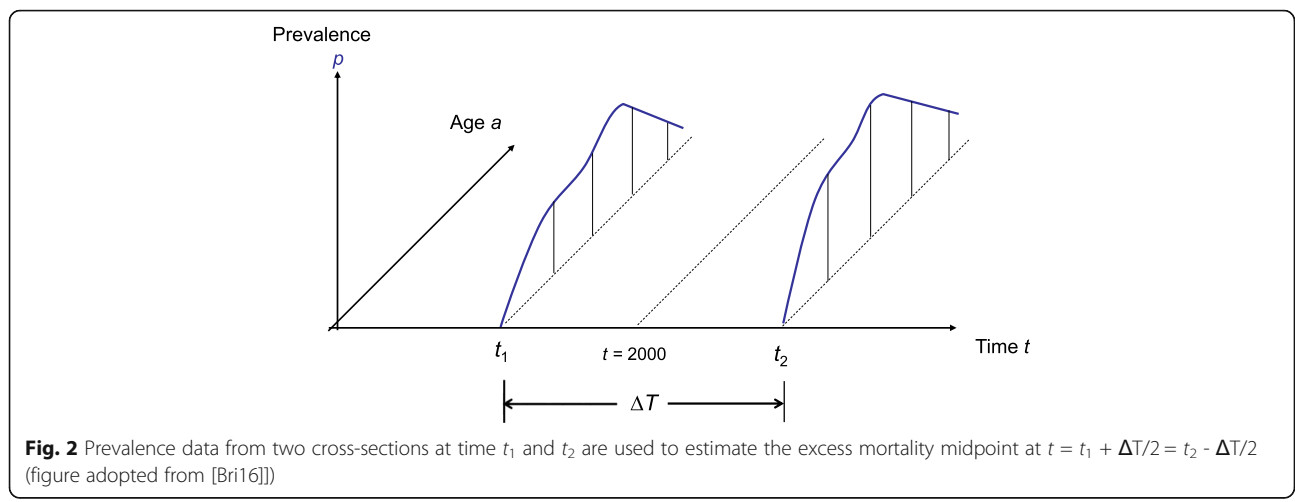
An alternative epidemiological measure to Δm for assessing discrepancies between the mortality rates m_0 and m_1 , is the mortality rate ratio $R = m_1/m_0$ which is of potential interest for practitioners. The mortality rate ratio R expresses the mortality rate of the diseased people *relative* to the non-diseased at the same age. Due to this plain interpretation, R is more often used than the (absolute) excess mortality Δm . Both measures, Δm and R , are related by $R = 1 + \Delta m/m_0$.

Direct estimation in simulated data about dementia

To illustrate how measures of excess mortality in a chronic disease can directly be estimated from incidence and prevalence data, we conduct a simulation study. We mimic a sequence of two cross-sectional studies for a chronic disease in two different years t_1 and t_2 centered at the year $t = 2000$. Let $\Delta T = t_2 - t_1$ denote the difference between t_1 and t_2 , i.e. $t_1 = 2000 - \Delta T/2$ and $t_2 = 2000 + \Delta T/2$. In each of the cross-sectional studies at t_1 and t_2 , the age-specific prevalence p is surveyed (Fig. 2).

The aim is to estimate the excess mortality at year $t = 2000$ from the cross-sectional prevalence data at t_1 and t_2 and the incidence. To assess the impact of the temporal difference between the cross-sectional studies, we vary ΔT from 0.1 to 10 (years). Together with the age-specific incidence rate i at $t = 2000$, the prevalence data in the two years t_1 and t_2 serve as input values to estimate the excess-mortality in the year $t = 2000$. The estimated excess mortality is then compared with the rates used to set up the simulation study in terms of absolute and relative bias.

The input data for the simulation are motivated from survey data about dementia in the female population of Europe [7]. Dementia is a major health problem in many countries with potentially increasing prevalence in the future [8]. The age-specific prevalence p for each of the two years t_1 and t_2 is calculated analytically with the incidence rate i from [7]. The age-specific mortality rate m_0 of the dementia-free population is chosen to be $m_0(t, a) = \exp(-10.7 + 0.1a + t \ln(0.99))$ aiming to approximate the mortality of the European population based on the Gompertz-Makeham law of mortality [9]. In addition, we assume that the mortality m_1 of the diseased people can be written as a product of m_0 and R : $m_1(t, a) = R(t, a) \times m_0(t, a)$ with $\log R(t, a) = \log(3) + [\log(1.5) - \log(3)] (a - 60)/(90 - 60)$. The rationale for choosing this R is based on the idea that m_1 also follows a Gompertz-Makeham law. Then, the logarithm of the quotient m_1/m_0 is a straight line as given



here. The specific numerical values in the definition of R are chosen to mimic the age-dependency as reported in [10], where R was found to be about 3 and 1.5 at 60 and 90 years of age, respectively. Note, however, that in this simulation we want to demonstrate feasibility of the method in a realistic range of parameters. We do not aim for the best obtainable agreement between our input data and the observed data.

Bayes estimation and application to claims data

After describing the direct estimation, we present an estimation method in the framework of Bayesian inference. Bayes methods are increasingly used in applied statistics because they provide a flexible framework for the analysis of scientific problems and quantifying uncertainty in their solution [11]. As an application of the Bayesian approach, we estimate the excess mortality of type 2 diabetes in the year 2012 from claims data comprising 35 million German men. Goffrier and colleagues [12] reported the age-specific prevalence of diabetes among German men in the years $t_1 = 2009$ and $t_2 = 2015$ as shown in Fig. 3. In the same work, the age-specific incidence rate i in 2012 has been surveyed. The data for this analysis is publicly available and can be found in [12].

Our aim in the diabetes example is to estimate the age-specific mortality rate ratio R in the range 50 to 90 years of age. Recently, for a smaller age range the mortality rate ratio has been estimated in Tönnies et al. [13]. Compared to [13] we extend the age range by the novel Bayesian approach.

The idea for the Bayes method is that for given age-specific prevalence p , incidence rate i and general mortality m , an estimate of the excess mortality in terms of the mortality rate ratio R is desired. According to the Theorem of Bayes [11] we obtain.

$$f(R|p) \propto f(p|R) \times f(R) \tag{1}$$

where $f(R|p)$ is the posteriori distribution of R , $f(p|R)$ is the probability density function of p given R and $f(R)$ denotes the priori distribution of R . For clarity, we assume that i and m are known. Motivated by empirical findings from the Danish Diabetes Register [14], we assume that the logarithm of the age-specific mortality rate ratio R approximately is a straight line in the age range 50 to 90 years:

$$\log(R(a)) = \log(R(50)) + [\log(R(90)) - \log(R(50))] \frac{(a - 50)}{(90 - 50)} \tag{2}$$

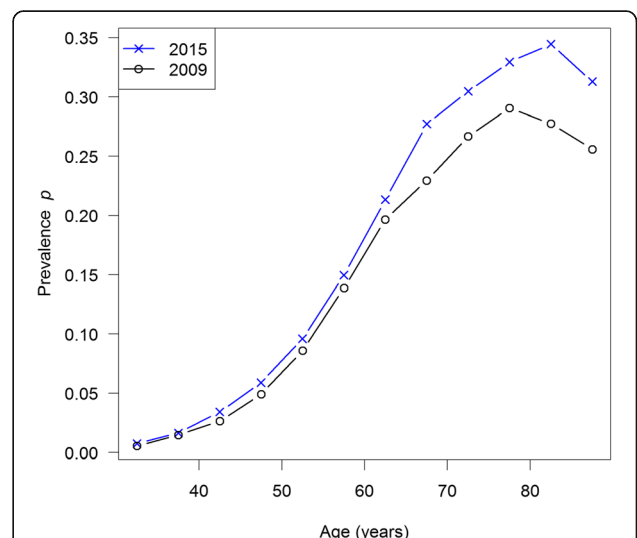


Fig. 3 Surveyed age-specific prevalence p of type 2 diabetes in German men in 2009 (black line with circles) and 2015 (blue with crosses)

For estimation of $R(50)$ and $R(90)$ in Eq. (2), we use weakly informative prior distributions $R(50) \sim U(2; 9)$ and $R(90) \sim U(1; 2)$; again inspired by the Danish diabetes register. $U(\nu; w)$ means the continuous uniform distribution with minimum and maximum value ν and w , respectively. In Bayesian terminology, our aim is to estimate the joint a-posteriori distribution for $R(50)$ and $R(90)$.

To use Eq. (1) for the estimation of R given p , we apply three steps: 1) values for $R(50)$ and $R(90)$ are drawn from the uniform prior distributions, 2) solving the PDE with initial condition $p(2009; a)$ as given in [12] and 3) comparing the calculated solution p in 2015 with the surveyed values.

For solving the PDE, we use the Method of Characteristics [15] to first convert the PDE into an ordinary differential equation (ODE) and then, second, solve the ODE by the Runge-Kutta Method of fourth order [16]. Next, the calculated prevalence in 2015, $p(2015; a)$, is compared with the observed prevalence in 2015 given by [12]. The age-specific prevalences p in the years 2009 and 2015 are shown as black and blue lines in Fig. 3, respectively. As conditional distribution $f(p|R)$, we chose the multivariate normal distribution.

$$f(p|R) \propto \exp.(-(\mathbf{p}_{\text{mod}} - \mathbf{p}_{\text{obs}})^t \Sigma^{-1} (\mathbf{p}_{\text{mod}} - \mathbf{p}_{\text{obs}})/2)$$

where $\mathbf{p}_{\text{mod}} = \mathbf{p}_{\text{mod}}(R)$ is the solution of the PDE for a given R . The conditional distribution $f(p|R)$ assesses the differences between the modeled \mathbf{p}_{mod} and observed prevalences \mathbf{p}_{obs} . The covariance matrix Σ is estimated by following diagonal matrix:

$$\Sigma = \text{diag}(p_j (1 - p_j)/n_j)$$

with age-specific prevalences p_j and the corresponding number of people n_j in the age group j . Choosing the covariance matrix as a diagonal matrix makes the implicit assumption that the prevalences p_j are stochastically independent. A justification for this assumption is the fact that people belonging to one age group are different from the people in another age group.

In a sensitivity analysis, we released the assumption of weakly informative priors ($R(50) \sim U(2; 9)$, $R(90) \sim U(1; 2)$) and examined the impact on the estimation of $R(50)$ and $R(90)$. For this, we choose $R(50)$ and $R(90)$ from a bivariate normal distribution with mean (5.5, 1.5), standard deviation of 1 and 0.1 in $R(50)$ and $R(90)$, respectively, and a correlation coefficient of 0.9 between $R(50)$ and $R(90)$. These assumptions lead to the following covariance matrix for the joint distribution of $R(50)$ and $R(90)$:

$$\begin{pmatrix} 1^2 & 0.9 \times 1 \times 0.1 \\ 0.9 \times 1 \times 0.1 & 0.1^2 \end{pmatrix}$$

Results

Illness-death model

The age-specific prevalence $p(t, a) = I(t, a) / \{H(t, a) + I(t, a)\}$ i.e., the percentage of people aged a at time t who are chronically ill, is the solution of the following partial differential equation (PDE):

$$(\partial_t + \partial_a)p = (1-p)\{i - p(m_1 - m_0)\} \tag{3}$$

In Eq. (3), ∂_t and ∂_a denote the partial derivatives with respect to t and a , respectively. The mathematical proof for Eq. (1) can be obtained from examining the change rates of the number of healthy and ill people in the illness-death model (H and I in Fig. 1) [17] or by using the theory of stochastic processes [2].

Eq. (3) implies that the excess mortality $\Delta m = m_1 - m_0$ can directly be estimated from the incidence rate i , prevalence p and the temporal change of the prevalence $((\partial_t + \partial_a)p)$:

$$\Delta m = \left[i - \frac{(\partial_t + \partial_a)p}{1-p} \right] / p \tag{4}$$

Note that for direct estimation of the excess mortality Δm by Eq. (2) only the incidence rate i and the prevalence based figures p and $(\partial_t + \partial_a)p$ are necessary. No additional data are needed.

Instead of using Eq. (3) for a relation between the incidence, prevalence and mortality, an alternative way is possible by considering the prevalence-odds $\theta(t, a) = I(t, a) / H(t, a)$. For the prevalence-odds θ we find the following PDE, which is equivalent to Eq. (3):

$$(\partial_t + \partial_a)\theta = i - \theta(m_1 - m_0 - i) \tag{5}$$

Equation (5) was first published by Brunet and Struchiner [18]. The derivation is given in an additional file [Additional file 1]. Compared to Eq. (3) the PDE (5) has the advantage of being linear. Solving PDEs like Eq. (3) and (5) is usually accomplished by transformation into an equivalent ordinary differential equation by the Method of Characteristics [15]. In case of Eq. (3), the resulting ordinary differential equation is of Ricatti type [19], which in general can only be solved numerically because an explicit representation of the general solution does not exist [20]. In case of the equivalent Eq. (5), however, an explicit representation of the solution indeed is possible. As detailed in the additional file [see Additional file 1] it holds:

$$\theta(t, a) = \int_0^a i(t-s, a-s) \exp(-\phi_{t,a}(s)) ds \tag{6}$$

For brevity, in Eq. (6) it was set $\phi_{t,a}(x) = \int_0^x [m_1 - m_0 - i](t-x + \tau, a-x + \tau) d\tau$

The explicit representation of the solution θ in Eq. (6) allows to calculate θ with any prescribed accuracy, e.g. by Romberg integration [16], which we will use in the examples below. Applying the back-transformation $p = \theta/(1 + \theta)$ yields the prevalence p .

For later use, we note that Eq. (3) can also be expressed in terms of the mortality rate ratio R and the general mortality $m = p m_1 + (1 - p) m_0$:

$$(\partial_t + \partial_a)p = (1-p) \left\{ i - m \frac{p(R-1)}{1 + p(R-1)} \right\} \tag{7}$$

Direct estimation: dementia in the female population of Europe

After calculating the prevalence-odds θ in years t_1 and t_2 by Eq. (6), the associated prevalences $p = \theta/(1 + \theta)$ are calculated. Figure 4 shows the age-specific prevalences for the years $t_1 = 1990$ (dashed line) and $t_2 = 2010$ (solid line). To demonstrate that our simulated prevalence has a reasonable range, we additionally plotted the surveyed values for European women reported in [8]. The proposed method to estimate the excess mortality Δm in the year $t = 2000$ is the direct application of Eq. (4). The partial derivative $(\partial_t + \partial_a)p$ in Eq. (4) is approximated by a finite difference:

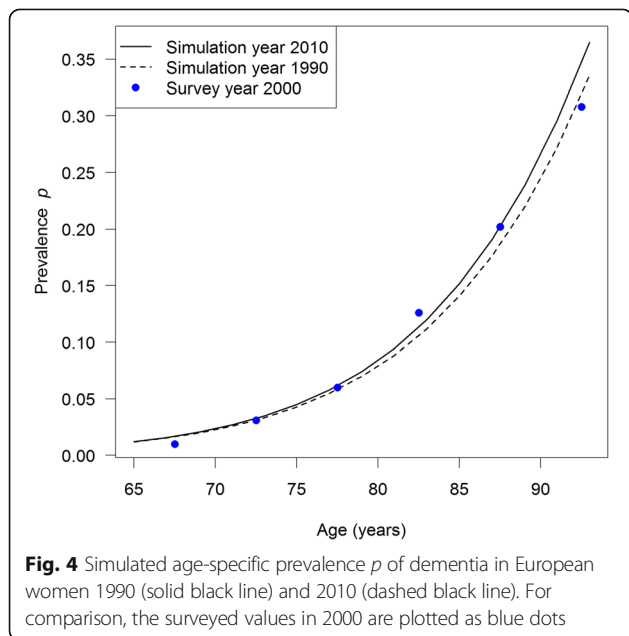


Fig. 4 Simulated age-specific prevalence p of dementia in European women 1990 (solid black line) and 2010 (dashed black line). For comparison, the surveyed values in 2000 are plotted as blue dots

$$(\partial_t + \partial_a)p(t, a) \approx p(t + \Delta T/2, a + \Delta T/2) - p(t - \Delta T/2, a - \Delta T/2) / \Delta T$$

Then, the excess mortality Δm can be estimated by plugging these numbers into Eq. (4). In case the mortality rate m_0 of the non-diseased is known, the age-specific mortality rate ratio can be calculated by $R = 1 + \Delta m/m_0$. Table 1 shows the true and estimated values for R at different ages and various choices of ΔT .

From Table 1 we can see that the absolute relative Error increases as the temporal difference ΔT between the cross-sections increases and that absolute relative error increases as the age decreases. In the extreme case (age 60, $\Delta T = 10$), the absolute relative error reaches nearly 30%. This indicates that in case two cross-sectional studies are separated by more three years (i.e., $\Delta T > 3$) the method yields feasible results only in the higher age groups.

Bayesian estimation of excess mortality in male diabetics from Germany

The log-likelihood of the a-posteriori distribution $f(R|p) \propto f(p|R) \times f(R)$ is shown in Fig. 5. The black cross indicates the maximum aposteriori (MAP) estimator for these data, which is given by $R_{MAP}(50) = 4.47$ and $R_{MAP}(90) = 1.39$. We obtain the estimates for $R(50)$ and $R(90)$ including 95% credibility intervals as shown in Table 2.

These values agree well with the empirical findings from the Danish Diabetes Register [14], where values slightly below 4 and slightly above 1.5 have been found for ages 50 and 90 years, respectively.

In the sensitivity analysis with bivariate normal prior distributions, the MAP estimator changed only slightly to $R_{MAP}(50) = 4.54$ and $R_{MAP}(90) = 1.38$.

Discussion

In this work, we have described how the illness-death model can be used to obtain information about excess mortality in case prevalence and incidence are given. It turns out that the excess mortality can be calculated by the incidence rate, the prevalence and the temporal change of the prevalence (see Eq. (4)). In data where these figures are estimable, insights into the excess mortality of people with chronic diseases compared to the people without the disease can be gained.

As applications, simulated data about dementia and claims data about diabetes have been analyzed. For the dementia example we estimated the excess mortality directly and for the diabetes data we formulated a Bayesian approach. Both methods were based on aggregated data only (age-specific prevalence and incidence rate)

Table 1 True and estimated mortality rate ratios

Age	True R	$\Delta T = 0.1$		$\Delta T = 0.5$		$\Delta T = 1$		$\Delta T = 3$		$\Delta T = 5$		$\Delta T = 10$	
		Est. R	rel. Err. (%)	Est. R	rel. Err. (%)	Est. R	rel. Err. (%)	Est. R	rel. Err. (%)	Est. R	rel. Err. (%)	Est. R	rel. Err. (%)
65	2.717	2.717	0.01	2.716	-0.06	2.708	-0.32	2.635	-3.0	2.495	-8.2	1.942	-29
70	2.461	2.461	0.02	2.460	-0.04	2.455	-0.23	2.406	-2.2	2.312	-6.0	1.938	-21
75	2.229	2.229	0.00	2.228	-0.03	2.225	-0.18	2.191	-1.7	2.128	-4.6	1.868	-16
80	2.019	2.019	0.00	2.018	-0.02	2.016	-0.14	1.993	-1.3	1.948	-3.5	1.765	-12
85	1.829	1.829	0.00	1.828	-0.01	1.827	-0.11	1.810	-1.0	1.779	-2.7	1.647	-9.9
90	1.656	1.656	0.02	1.666	0	1.655	-0.07	1.643	-0.76	1.621	-2.1	1.528	-7.7
95	1.500	1.501	0.10	1.502	0.1	1.503	0.18	1.505	0.36	1.511	0.76	1.492	-0.54

Table legend: The estimated mortality rate ratios (Est. R) at different temporal distances between the two cross-sectional studies (ΔT) and different ages (first column) are compared to the true mortality rate ratios (True R , second column). The difference between the estimated and the true mortality rate ratio are given in terms of the relative error (rel. Err., in %)

and do not require data from individual subjects. Aggregated data can be found frequently in the literature, which makes the proposed method suitable for many applications, especially when the research question is aimed at population-wide measures. Here, we have chosen aggregated data about diabetes from the statutory health insurance in Germany based on about 35 million men. Based on the age-specific prevalence in 2009, we used non-informative priors for mortality rate ratio R and the PDE (7) to estimate the a posteriori likelihood of R given the age-specific prevalence in 2015. In this way, the PDE can therefore be seen as the data generating process underlying the prevalence data. In a sensitivity analysis, we used more informative prior distributions (bivariate normal) and found that the estimated values for the mortality rate ratios changed only slightly. Main reason for this robustness is due to the large number of people in the prevalence data.

Our approach has two limitations. The first limitation stems from the fact that Eqs. (3) and (5) are only valid if migration into and or from the considered population does not take place or if the prevalence of the chronic condition in migrants is similar to the prevalence in the resident population [21]. If migration happens on a considerable magnitude and if the prevalence in the migrants is substantially different from the residents, adoptions to Eq. (3) are possible [21]. The second limitation of our novel approach becomes visible in the simulation study about dementia: The two (or more) cross-sectional surveys for estimating the change of the prevalence should not be separated too much. In our simulation, the surveys should be conducted within a period of three years (or less) (i.e., $\Delta T \leq 3$) to have a relative error below 5%. If the two cross-sections are separated by ten years ($\Delta T = 10$), the relative error has reached up to 30%. In the diabetes example, the two cross-sections were separated by six years ($\Delta T = 6$). Based on this, we expect the relative errors of our estimates $R(50)$ and $R(90)$ to be about 10%. For comparison, the width of the credibility intervals for our estimates $R(50)$ and $R(90)$ have a similar magnitude. Thus, we would conclude a relative error of 10% in the mortality rate ratio is a rough estimate of the magnitude of accuracy that can be obtained from our method applied to these data.

In the current analysis, no attempt has been taken to examine the effect of smaller population sizes, i.e., how sampling uncertainty in the age-specific prevalence and incidence affects the estimates of the excess mortality. Furthermore, we have not analyzed the robustness of the estimation methods against misclassification error (i.e.,

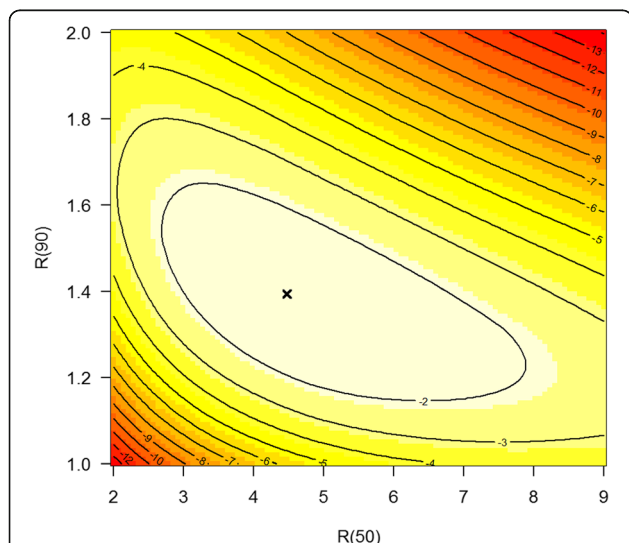


Fig. 5 Contour plot of the posteriori likelihood of the mortality rate ratio R at ages 50 (abscissa) and 90 (ordinate). The maximum a posteriori (MAP) estimator is indicated as a black cross

Table 2 Estimated mortality rate ratios for the diabetes data

Age	Mortality rate ratio R	95% credibility interval	
50	4.47	4.17	4.78
90	1.39	1.33	1.46

Table legend: The estimated mortality rate ratios (R) at ages 50 and 90 with 95% credibility intervals

false positive and false negative rates in input prevalence and incidence data). Questions about sample sizes and misclassification are currently analyzed and will be subject to a future paper providing more technical details.

Conclusion

The described approach is useful to estimate the excess mortality of a chronic condition from aggregated incidence and prevalence data. The feasibility has been demonstrated in a simulation study about dementia and in claims data about diabetes in German men.

Additional files

Additional file 1: Microsoft Word file (.doc) providing details of the mathematical background of Eqs. (5) and (6). (DOC 34 kb)

Additional file 2: Script (plain text file, accessible via any text editor, e.g., Notepad, GNU Emacs etc) for the dementia simulation study, intended to use with the statistical software R (The R Foundation of Statistical Software). (R 4 kb)

Abbreviations

MAP: Maximum a posteriori; NHIS: National Health Interview Survey; PDE: Partial differential equation

Acknowledgements

The authors wish to thank the Zentralinstitut für Kassenärztliche Versorgung, Berlin, for making the claims data available.

Authors' contributions

RB had the initial idea for this work, developed the source code and drafted the manuscript. TT and AH critically discussed the ideas and revised the manuscript. All authors gave substantial intellectual contributions, read and approved the final manuscript.

Funding

This research did not receive any funding.

Availability of data and materials

The source code for generating the data about dementia in Europe is available as an electronic supplement to this published article [Additional file 2]. The data for the diabetes example were taken from a published source [12], which has been cited in the text.

Ethics approval and consent to participate

This study does not involve data from human participants (simulation about dementia) or solely relies on publicly available secondary data (aggregated claims data [12]). Therefore, consent to participate is not required. The Ethics Board of the University Hospital Duesseldorf has confirmed that in case of published data, no review of the Ethics Board is necessary.

Consent for publication

Not necessary because this manuscript does not contain data from any individual person.

Competing interests

The authors declare that they have no competing interests.

Received: 5 April 2019 Accepted: 19 June 2019

Published online: 28 June 2019

References

1. Brinks R, Landwehr S. Change rates and prevalence of a dichotomous variable: simulations and applications. *PLoS One*. 2015;10(3). <https://doi.org/10.1371/journal.pone.0118955>.
2. Brinks R, Hoyer A. Illness-death model: statistical perspective and differential equations. *Lifetime Data Anal*. 2018;24(4):743–54. <https://doi.org/10.1007/s10985-018-9419-6>.
3. Brinks R, Hoyer A, Landwehr S. Surveillance of the incidence of non-communicable diseases (NCDs) with sparse resources: a simulation study using data from a National Diabetes Registry, Denmark, 1995–2004. *PLoS One*. 2016;11(3):e0152046. <https://doi.org/10.1371/journal.pone.0152046>.
4. Egeberg A, Kristensen LE. Impact of age and sex on the incidence and prevalence of psoriatic arthritis. *Ann Rheum Dis*. 2018;77:e19. <https://doi.org/10.1136/annrheumdis-2017-211980>.
5. Tamayo T, Brinks R, Hoyer A, Kuß OS, Rathmann W. The prevalence and incidence of diabetes in Germany. *Dtsch Arztebl Int*. 2016;113(11):177–82. <https://doi.org/10.3238/arztebl.2016.0177>.
6. National Center for Health Statistics of the Centers for Disease Control and Prevention (CDC) About the National Health Interview Survey https://www.cdc.gov/nchs/nhis/about_nhis.htm. Accessed on 5 Apr 2019.
7. Fratiglioni L, Launer LJ, Andersen K, Breteler MM, Copeland JR, Dartigues JF, Lobo A, Martinez-Lage J, Soininen H, Hofman A. Incidence of dementia and major subtypes in Europe: a collaborative study of population-based cohorts. *Neurologic diseases in the elderly research group*. *Neurology*. 2000;54(11 Suppl 5):S10–5.
8. Lobo A, Launer LJ, Fratiglioni L, Andersen K, Di Carlo A, Breteler MM, Copeland JR, Dartigues JF, Jagger C, Martinez-Lage J, Soininen H, Hofman A. Prevalence of dementia and major subtypes in Europe: a collaborative study of population-based cohorts. *Neurologic Diseases in the Elderly Research Group*. *Neurology*. 2000;54(11 Suppl 5):S4–9.
9. Missov TI, Lenart A. Gompertz–Makeham life expectancies: expressions and applications. *Theor Pop Bio*. 2013;90:29–35.
10. Rait G, Walters K, Bottomley C, Petersen I, Iliffe S, Nazareth I. Survival of people with clinical diagnosis of dementia in primary care: cohort study. *BMJ*. 2010;341:c3584. <https://doi.org/10.1136/bmj.c3584>.
11. Gelman A, Stern HS, Carlin JB, Dunson DB, Vehtari A, Rubin DB. *Bayesian data analysis*. London: Chapman and Hall/CRC; 2013.
12. Goffrier B, Schulz M, Bätzing-Feigenbaum J. Administrative Prävalenzen und Inzidenzen des diabetes mellitus von 2009 bis 2015. *Versorgungsatlas*. 2017. <https://doi.org/10.20364/VA-17.03>.
13. Tönnies T, Hoyer A, Brinks R. Excess mortality for people diagnosed with type 2 diabetes in 2012 - estimates based on claims data from 70 million Germans. *Nutr Metab Cardiovasc Dis*. 2018;28(9):887–91. <https://doi.org/10.1016/j.numecd.2018.05.008>.
14. Carstensen B, Kristensen JK, Ottosen P, Borch-Johnsen K. Steering Group of the National Diabetes Register. The Danish National Diabetes Register: trends in incidence, prevalence and mortality. *Diabetologia*. 2008;51(12):2187–96. <https://doi.org/10.1007/s00125-008-1156-z>.
15. Polyanin AD, Zaitsev VF, Moussiaux A. *Handbook of first-order partial differential equations*: CRC Press; 2001.
16. Dahlquist G, Björck A. *Numerical methods*. Englewood Cliffs: Prentice-Hall; 1974.
17. Brinks R, Landwehr S. A new relation between prevalence and incidence of a chronic disease. *Mathematical Medicine and Biology*. 2015;32(4):425–35. <https://doi.org/10.1093/imammb/dqu024>.
18. Brunet RC, Struchiner CJ. A non-parametric method for the reconstruction of age- and time-dependent incidence from the prevalence data of irreversible diseases with differential mortality. *Theor Pop Bio*. 1999;56(1):76–90.
19. Brinks R. Illness-death model in chronic disease epidemiology: characteristics of a related differential equation and an inverse problem. *Comp Math Meth Med*. 2018. <https://doi.org/10.1155/2018/5091096>.
20. Kamke E. *Differentialgleichungen Lösungsmethoden und Lösungen*. Leipzig: Teubner Verlag; 1983.
21. Brinks R, Landwehr S. Age- and time-dependent model of the prevalence of non-communicable diseases and application to dementia in Germany. *Theor Popul Biol*. 2014;92:62–8. <https://doi.org/10.1016/j.tpb.2013.11.006>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.