Annals of Epidemiology 24 (2014) 570–574

Contents lists available at ScienceDirect

Annals of Epidemiology

journal homepage: www.annalsofepidemiology.org

Original article

Mortality cohort effects from mid-19th to mid-20th century Britain: did they exist?

Yu-Kang Tu PhD [a,*], Katherine Keyes PhD [b], George Davey Smith MD, DSc [c]

[a] Department of Public Health, Institute of Epidemiology and Preventive Medicine, College of Public Health, National Taiwan University, Taipei, Taiwan
[b] Department of Epidemiology, Columbia University, New York, NY
[c] MRC Integrative Epidemiology Unit (IEU), University of Bristol, Bristol, UK

ARTICLE INFO

Article history:
Received 11 February 2014
Accepted 9 June 2014
Available online 14 June 2014

Keywords:
Age–period–cohort analysis
Mortality
Partial least squares
Cohort effects

ABSTRACT

Purpose: Identification is a central problem with age–period–cohort analysis. Because age + cohort = period, there is no unique solution to the linear effect using generalized linear modeling, but cohort effects have caused greater controversy than age and period effects. To illustrate the magnitude of cohort effects given the presence of collinearity, we reanalyze data from the seminal study by Kermack et al, with an update.
Methods: Relative mortality data in England and Wales between year 1845 and 1995 were analyzed using partial least squares regression. There were seven age groups ranging from 5 to 74 years old and 16 periods with 22 cohorts.
Results: Our reanalysis seemed to support the existence of cohort effects in the mortality trends. Period and cohort effects were generally consistent with changes in the social, economic, and environmental factors taking place in the last two centuries. Our analysis also showed a declining trend in period effects up to 1950s.
Conclusions: Partial least squares and related methods provide intuitive pointers toward the separation of linear age, period, and cohort effects. Because statistical algorithms cannot distinguish between relative and actual mortality rates, cohort effects may be underestimated because of contamination by negative age effects.

How to separate the distinct linear effects of age, period, and birth cohort on the changes in, for example, attitudes, behaviors, and health outcomes in the population has a long and controversial history in epidemiology and social sciences [1–9]. Age effects typically reflect changes in the distribution of an outcome across the life course, reflecting changes in the distributions of biological, developmental, and sociological risk factors independent of period, and generational experience [10]. Most approaches to assessing age, period, and cohort effects consider age effects as nearly universally present for most outcomes [1–4]. Period effects are usually considered to reflect environmental experiences common to a given period, which affects the whole population living at that time. People of different ages at a given period share the same experience. Birth cohort effects are collective environmental effects uniquely experienced by groups of individuals born around the same time [10]. They are generally considered to be the consequence of events occurring during particular sensitive periods of the life course, often early life, including the in-utero and early postnatal period. Although age, period, and birth cohort effects seem to be conceptually unique, they are nevertheless mathematically related because age + cohort = period. Consequently, when traditional regression models attempt to simultaneously consider age, period, and cohort effects as linear constructs, there is no unique solution, giving rise to the identification problem. Recently, several approaches have been proposed to deal with this collinearity; whereas some approaches make implicit constraints about the relations among the three effects to allow them to become estimable [11–14], others use smoothing functions to obtain age, period, and cohort effect patterns [15–17].

Despite having some resonance within epidemiology, particularly with regard to the fetal/developmental origins of health and disease hypothesis, cohort effects have probably caused greater controversy than age and period effects in the past [18–20]. Data used for most age–period–cohort analyses are aggregated data crosstabulated by age and period, in which cohort is represented by the diagonals. Early and recent cohorts therefore have fewer observations than those in the middle and are often more elusive to detect.

* Corresponding author. Institute of Epidemiology and Preventive Medicine, College of Public Health, National Taiwan University, Room 539, 17 Xu-Zhou Road, Taipei, Taiwan. Tel: +886 2 33668039.
E-mail address: yukangtu@ntu.edu.tw (Y.-K. Tu).

1047-2797/$ – see front matter
http://dx.doi.org/10.1016/j.annepidem.2014.06.002

Changes in generations' mortality in England and Wales

The pioneering study by Kermack et al. [21] was an important early work that apparently demonstrated cohort effects in mortality in England and Wales. Using the mortality at year 1845 as a reference, their study showed a clear decline in mortality for groups for age more than 5 years beginning around the year 1845 and seen in later periods for the same cohort. This finding has been interpreted in terms of the impact of social and environmental factors from the mid-19th century onward on population health. Although the study by Kermack et al. has been influential, cohort effects shown by their analysis are still being questioned [20]. In this article, we revisit the cohort effects in Kermack et al. using partial least squares regression to address the identification problem [22]. We will also discuss issues related to the use of relative mortality in identifying cohort effects.

Materials and methods

We used the relative mortality data in England and Wales between year 1845 and 1995, which was first presented by Kuh and Davey Smith [23,24] with updates obtained from the web site for Office for National Statistics, London [22,25]. There were seven age groups ranging from 5 to 74 years old and 16 periods with 22 cohorts. Following a similar approach by Roberson and Boyle [26], we first used a three-dimensional perspective plot to show the trends in the relative mortality rates with the use of the earliest period as the reference (Fig. 1). In Figure 1, as each diagonal from one of the age groups in different periods represents a cohort, a trend from the bottom left to the upper right across diagonals would suggest cohort effects; there seemed to be a decreasing trend in mortality after 1845 from the earliest cohort at the upper left corner to the most recent cohort at the lower right corner, which seems to suggest a cohort effect. However, as age, period, and cohort effects are not independent, a visual examination alone could be misleading.

We then used partial least squares regression, to estimate the age, period, and cohort effects [27–30]. As details of partial least squares regression have been described elsewhere, we only gave a brief explanation here. For the analysis of crosstabulated data with seven age groups and 16 periods, the linear model for the natural log transformation of relative mortality rate (MR) with all the dummy variables for age, period, and cohort groups is written as follows:

$$\log(\mathrm{MR}) = a_0 + \sum_{i=1}^{7}\alpha_i \mathrm{age}_i + \sum_{j=1}^{16}\beta_j \mathrm{period}_j + \sum_{k=1}^{22}\gamma_k \mathrm{cohort}_k + \varepsilon \quad (1)$$

where a_0 is the intercept, ε the residual error term, and α_i, β_j, and γ_k are the regression coefficients for the dummy variables. As first shown by Kupper et al. [31], after all the variables are centered, there are four identification problems in Equation 1.

$$\sum_{i=1}^{7}\mathrm{age}_i = \sum_{j=1}^{16}\mathrm{period}_j = \sum_{k=1}^{22}\mathrm{cohort}_k = 0 \quad (2)$$

and

$$\sum_{i=1}^{a}\left(i-\frac{(a+1)}{2}\right)\mathrm{age}_i - \sum_{j=1}^{p}\left(i-\frac{(p+1)}{2}\right)\mathrm{period}_j + \sum_{k=1}^{a+p-1}\left(i-\frac{(a+p)}{2}\right)\mathrm{cohort}_k = 0 \quad (3)$$

where a and p are the number of groups of age and period, that is, seven and 16, respectively. As shown in our previous study [30],

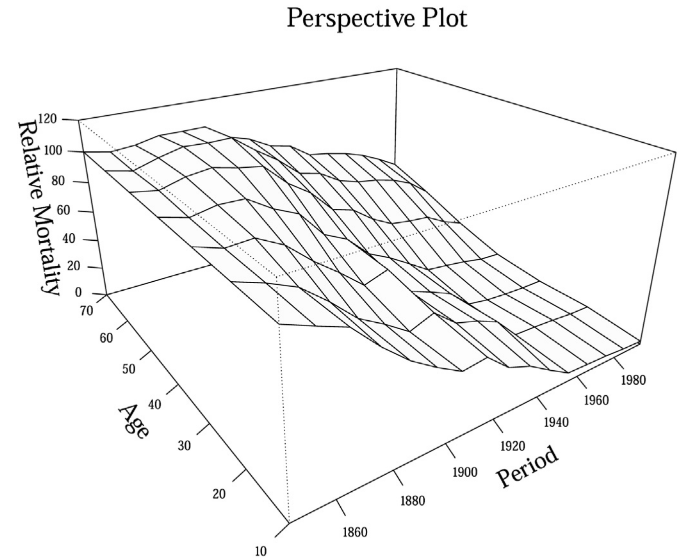

Fig. 1. Perspective plot for relative mortalities in England and Wales between 1845 and 1985.

partial least squares regression implicitly applied the following constraints on the parameters to make Equation 1 estimable.

$$\sum_{i=1}^{7}\alpha_i = \sum_{j=1}^{16}\beta_j = \sum_{k=1}^{22}\gamma_k = 0 \quad (4)$$

and

$$\sum_{i=1}^{a}\left(i-\frac{(a+1)}{2}\right)\alpha_i - \sum_{j=1}^{p}\left(i-\frac{(p+1)}{2}\right)\beta_j + \sum_{k=1}^{a+p-1}\left(i-\frac{(a+p)}{2}\right)\gamma_k = 0 \quad (5)$$

Fu [16] and later Yang et al. [11,12] proposed the use of the Equations 4 and 5 as a solution to the identification problem, and their proposed solution is widely known as the intrinsic estimator. From a mathematical viewpoint, imposing Equations 4 and 5 in the estimation is equivalent to using the Moore–Penrose generalized inverse method to obtain a unique inverse matrix for the perfectly collinear design matrix that consist of all the age, period, and cohort variables [7,11,30].

Partial least squares analysis is a data-reduction technique and successively extracts orthogonal components, which are weighted combinations of explanatory variables [30]. The first component has the largest covariance with the outcome variable and the second component has the second largest covariance, etc. Usually, the first few components can explain most of covariance with the outcome. For seven age groups, 16 period groups, and 22 cohort groups, the total number of dummy variables is 45, but the rank of the design matrix is only 41 owing to four collinearity constraints (Equations 2 and 3). The maximum number of components can be extracted is therefore 41. As shown in our previous studies [29,30], the weights in any partial least squares component always satisfy the constraints in Equations (4) and (5), and consequently, the partial least squares regression coefficients will satisfy these constraints irrespective of the number of components. We have shown elsewhere that when the maximum number of components is extracted, results from partial least squares regression is equivalent to that from the Moore–Penrose generalized inverse method. Consequently, when the maximum number of partial least squares components is extracted, the intrinsic estimator and partial least