

# Chapter 9

## Historical Trends in Mortality Risk over 100-Year Period in China with Recent Data: An Innovative Application of Age-Period-Cohort Modeling



Xinguang Chen

**Abstract** History is the best teacher. It is challenging to learn from history to address contemporary problems in the field of global health and epidemiology. A first challenge is the lack of data to examine medical and health problems in the past. Such challenge is more obvious in developing countries, such as China where no data were collected due to limited resources, wars, plaques, and natural disasters. Theoretical analysis and empirical data from age-period-cohort modeling indicate that recent data by age of a population contains information about the past. For example, mortality rate for people aged 90 in 1990 contains information about mortality risk in 1900 when they were born. Therefore, information contained in mortality by age functions like digital fossil; and the age-period-cohort modeling provides a tool to extract the information from the fossil. In this study we examined the mega-trends in mortality risk for China since 1901 when the 2000-year long feudalism was throughout, to 1949 when the independence was established, and up to 1980s when rapid economic growth emerged. We achieved the goal by analyzing data collected in recent years from 1990 to 2010 with the age-period-cohort modeling method and the intrinsic estimator. Findings of the study suggest the existence of four Sunny Periods and three Cloudy Periods during 1901–2010. These Sunny and Clouding Periods were in close coincident with significant social, cultural, political, economic events in the history of China. Findings of the study revealed that the highest mortality risk was associated with foreign invasion, the second highest risk was associated with civil wars, the third highest risk was associated with economic reform; the lower mortality risk was associated with the post-war period and the establishment of new China, and the longest period with reduced risk of mortality was associated with the Cultural Revolution Period. In

---

X. Chen (✉)

Department of Epidemiology, College of Public Health and Health Professions, College of Medicine, University of Florida, Gainesville, FL, USA

Global Health Institute, Wuhan University, Wuhan, China

e-mail: [jimax.chen@ufl.edu](mailto:jimax.chen@ufl.edu)

© Springer Nature Switzerland AG 2020

X. Chen, (Din) D.-G. Chen (eds.), *Statistical Methods for Global Health and Epidemiology*, ICSA Book Series in Statistics,

[https://doi.org/10.1007/978-3-030-35260-8\\_9](https://doi.org/10.1007/978-3-030-35260-8_9)

223

conclusion, age-period-cohort modeling provide a powerful tool for researchers to examine medical and health issues in the past with more recent data to advance epidemiology and global health. We highly recommend the use of this method in research in both developed and developing countries.

**Keywords** Historical epidemiology · APC modeling · Mortality in China · Social economic factors · Global health

## 9.1 Introduction

The development of global health and epidemiology requires us to investigate significant medical and health issues in the past with a historical perspective. However, it is rather challenging, if not impossible to examine things happened long time ago in the past when no research data were collected. In this chapter, we introduce the age-period-cohort (APC) modeling method, a classic epidemiological modeling tool; demonstrate its use in analyzing more recent data to investigate the time trends in the risk of mortality in China more than 100 years ago since 1906.

### 9.1.1 *Learn from History*

History can provide unique information supporting global health and epidemiology to better understand current medical and health problems we are facing. Significant events recorded in history, such as wars, famines, political, socioeconomic changes, often can exert broad and substantial impact on population health. While the majority of published studies use more recent data in research to understand current status of health and diseases; findings from a historical epidemiologic research studies will add new data from a different angle to strengthen current findings and to help clarify controversial issues observed from current studies. In addition to the advancement of our understanding of challenge medical and health issues with a global and temporal perspective, historical data provide information at the macro-level, guiding studies with newly collected data.

To examine historical events in medicine and public health, we often consult qualitative method by reviewing and summarizing events occurred in history and link them to changes in health status. For example, based on mortality recoded in several countries and regions, the epidemiological transition model posts that the life expectancy was rather low for people living in the developed countries in the 1800s when infectious diseases were predominant; and we attribute the rapid increases in life expectancy in modern society to industrialization, economic growth and advancement in medical technology and medicine (McKeown, 2009; Omran, 2005; Zhou et al., 2019). Also, we have no doubt that wars, famine also contributed to high mortality. However, few quantitative and systematic studies in the literature

that have associated historical events with changes in health status of a population. Although findings from qualitative studies provide useful information, adding more systematic data will further enhance our mastery of historical events on current health status and better inform decision-making to deal with significant medical and health challenges we are facing today and in the future.

### ***9.1.2 Challenges for Quantitative Historical Research***

One reason for the lack of quantitative and systematic investigation of historical events in epidemiology and global health is due to the lack of recorded data to quantify health status of a population in the past. Data are lacking on diseases and deaths during the primitive time for all countries in the world except fossils and grave records from archeology, genealogy records, and church records. Recorded population mortality data are not widely available before World War II because of frequent wars, population migration, and large-scale plagues. More data have gradually become available since the 1950s after the establishment of the World Health Organization (WHO). However, such data are often available primarily for industrialized and developed countries in the Europe and North America.

Collecting health and disease data could not be on agenda for poor countries during the period when people are struggling for survival. Although, changes in the health status of a population during difficult times may contain unique information for researchers to understand how natural (i.e., famine, flooding), social (i.e., social movements), political (revolution that moves a country from one system to another), economic, and technical factors may affect health; unfortunately such data are not available. Fortunately, following the developed countries and assisted by WHO and developed countries, many developing countries started to collect data on population health and diseases since the 1980s. China is among one of these countries. If data collected in more recent times contains the information about health status in the past, it may provide an opportunity to quantitatively examine the relationship between significant historical events and the corresponding changes in population health.

In this chapter, we take China as an example to demonstrate how to conduct an historical epidemiologic research with more recent data. We first list the major historical events since 1900s after China abolished the 2000-year long feudalistic society and entered into modern society through long-term frequently social, political, economic, technical and cultural changes. We then demonstrate that the health data collected since 1990 contain the information to measure health status from 1900s to 1990, corresponding to the significant historical events.

## 9.2 Timeline of Significant Events in China Since 1900

The recent history in China since 1900 presents a good case for epidemiology and global health to examine social, political, cultural and economic factors and population health using APC modeling method with a historical perspective. People in China have experienced a series of historical events with largest scope and highest impact on human kind from the 1900s till today. Relating significant historical events with the risk of population mortality will provide evidence unique for researchers, decision-makers and the general public to objectively assess the history of health policy and changes in population health, to think of the health and disease in the future based on knowledge from the past, and to make evidence-based strategic plans for disease prevention and health promotion.

In this study we focused on the time period since 1900 in China, starting with the period of overthrow of the 2000-year feudalistic society in 1911, the independence of China in 1949, and the open policy and rapid economic reform since 1978. In the following, we list the major historical events that are known to be influential to population health.

### 9.2.1 *Overthrow of the Feudalistic Society*

1899–1901: Yehetuan Movement or Boxer Rebellion: Happened in Northcentral China, and described as an anti-foreign invasion, anti-colonial, and anti-western culture, triggered by the power deteriorating of the Qing Dynasty, companied by a growing number of invasions from the earlier industrialized foreign countries.

1901–1911: Abdication of Qing Dynasty, the last feudalism society and the establishment of the Republic of China, the first modern society in Chinese history. The government, also the first time in Chinese history, established the Department of Sanitary in 1905 as the central government agency in charge of health.

1914–1926: The first and second civil wars among major warlords extended from the last administration of Qing Dynasty, after establishment of the Republic of China. The well-known May fourth Movement occurred in 1919 led by students and participated by people from all walks of life against Japanese and warlords.

1927–1934: Ending of the civil war through Northern Anti-Warlord Military Campaign and starting a period of peace and country's reconstruction, particularly the pioneer work of village health workers as part of the Rural Reconstruction Movement for rural uplift. A 3-year famine (1928–1930) occurred during this period, killing millions of people.

1937–1945: The War Anti-Japanese invasion in a full scale as part of the World War II until the Japanese surrendered in 1945. There was also a famine during 1942–1943, killing millions of people.

1945–1949: Start and end of the last Civil Ware and the inauguration of the People's Republic of China.

## 9.2.2 *Early Period After Independence*

1949–1957: Reconstruction of the country, including the high-impact Patriotic Health and Hygiene Movement to change the unhealthy lifestyles, to improve environmental health, and to control infectious diseases.

1957–1966: Domestic political movements, 1959–1962 famine, and the Great Leap Forward started and the Campaign of Removing Four Harm Pests (mosquitoes, flies, rats and sparrows) in 1958.

1966–1976: The 10-Year Cultural Revolution with destruction of traditional cultures and cultivation of new culture, national policies and campaigns of “Prevention First”, “Shifted the Priority of Healthcare from Urban to Rural”, “Barefoot Doctors”, “Combination of Traditional Chinese Medicine with Western Medicine”, and establishment of the 3-Tier Primary Care System.

## 9.2.3 *The Period of Open Policy and Economic Reform*

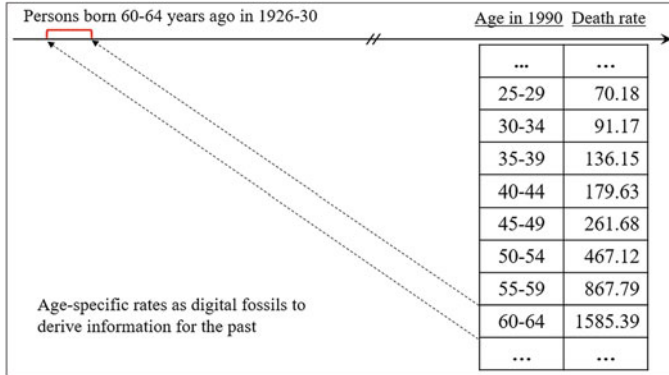
1978–present: Reform and open policy with unprecedented rural-to-urban population movement, rapid industrialization and economic growth the influence western culture and the 1989 Students Movement, shift from prevention first and free healthcare to market-oriented healthcare with reduced attention to prevention, and shift back in full scope in 2005 to re-emphasis of prevention, the free healthcare for urban residents and collaborative healthcare for rural residents.

## 9.3 **Age-Specific Data and APC Modeling Analysis**

### 9.3.1 *Age-Specific Data as Digital Fossils*

Epidemiologists are very familiar with mortality data tabulated by age of a population. Figure 9.1 shows how can we think of mortality data tabulated by age as a digital fossil containing information about risk of death in the past. In the figure, it contains a truncated part of population mortality rate per 100,000 by age groups from 25–29 to 60–64 in 1990 in China. For example, the mortality rate for all persons aged 60–64 was 1585.36/100,000 and this rate was the ratio of all persons died in 1990 over the total persons who were born 60–65 years ago in 1925–1930.

This figure tells the mortality rate for persons in an age or age group contains two important pieces of information: mortality risk in the year when the data were collected (the year 1990 as shown in the example), and the mortality risk related to the year of birth, the period 60–65 years ago. Although no data on mortality were collected for China 60–65 years ago during the 1925–1930 period (indicated by the red bracket) when the country were in the wars for domestic division and Japanese



**Fig. 9.1** Schematic illustration of age-specific mortality in a current year as digital fossil that contains information about the mortality risk in the past in addition to informing current status. **Note:** Mortality rate for persons aged 60–64 in 1990 contains information related to mortality risk in 1926–1930 when these people were born 60–64 years ago in 1926–1930; likewise, mortality rate for persons aged 80–84 in 1990 contains information of mortality risk in 1906–1910

invasion, it is likely to obtain data to describe the mortality risk during the period using the information contained in the death rate for persons aged 60–65 in 1990 when systematic data are available for use. It is in this regard, we conclude that *age-specific data can be considered as digital fossils* to examine the impact of historical events on mortality and potentially other health outcomes.

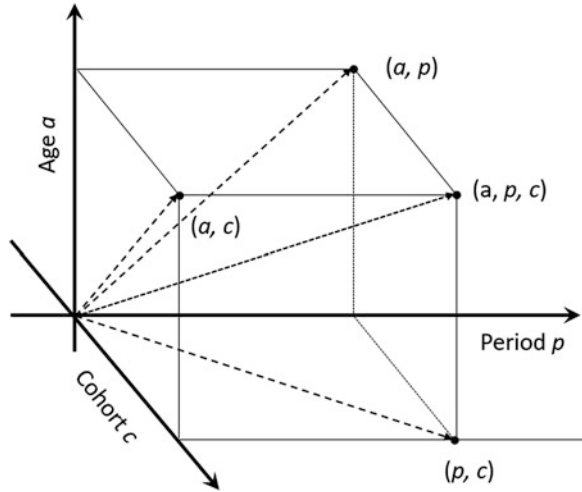
Likewise, the higher the age groups for data in more recently years, and further we can describe the past using more recent data. For example, mortality rates for persons aged 80–84 in 1990 contains the mortality risk for the period 80–84 years back during 1906–1910 when these persons were born. Although no data were available for the period long time ago, mortality data from recent times do contain the information to assess the health status during the period. The challenge is how to extract the information from the digital fossil—the age-specific mortality data.

### 9.3.2 APC Model to Extract the Historical Information

APC modeling provides one method to extract information from digital fossils—age-specific mortality data to examine the mortality risk of a population in the past when no systematic data were available for analysis (Chen et al., 2019; Chen & Wang, 2014a, 2014b). The basic principle of APC modeling analysis is to decompose the observed mortality by age of a year into three independent components. Let  $M_{ij}$  = the mortality rate for persons in age group  $i$  and period  $j$ ,  $M_{ij}$  can be described using a general regression model:

$$f(M_{ij}) = \mu + \alpha_i + \beta_j + \gamma_k + \epsilon, \tag{9.1}$$

**Fig. 9.2** Decomposition of mortality rate into three independent components as age, period and cohort effects. **Note:** In the endless time axis (period), people are continuous born to the world (cohort). At a point in time, one cohort was born. As time goes, this cohort ages. When reaching age a, this these people are corresponding to one point (a, p, c) in a 3D coordinate system. Other three meaningful points are (a, p), (a, c) and (p, c) respectively



where,  $\mu$  represents the grant mean,  $\alpha_i$  represents the impact of age  $i$  of the participants or “age effect”,  $\beta_j$  represents the impact of the time period  $j$  or “period effect”, and  $\gamma_k$  represents the impact of birth year  $k$  or “cohort effect”, and  $\epsilon$  are residuals.  $f()$  is a function that links a mortality measure with the three effects,  $f()$  can be identity for linear model, Poison for counting data and logistic for binary data.

From Eq. 9.1 it can be seen that the **cohort effect**  $\gamma_k$  is what we need—*information extracted from more recent data to quantify the mortality risk in the past when no data were available.*

Figure 9.2 visually depicts the age, period and cohort effect from Eq. 9.1 in a 3D format. In the chart, the three effects of age  $\alpha_i$ , period  $\beta_j$  and cohort  $\gamma_k$  are mutually independent from an APC model as indicated by the three coordinators at the point (a, p, c).

Figure 9.2 also shows other three models routinely used in APC modeling analysis, and they are (1) the age-period model with the age and period effects showing by the two coordinators at the point (a, p); (2) the age-cohort model with the age and cohort effects showing by another two coordinators at the point (a, c); and the period-cohort model with the period and cohort effects showing by two more coordinators at the point (p, c).

### 9.3.3 Challenges to APC Modeling

Among many challenges to APC modeling, two are critical. The first is that APC modeling method cannot model individual level data and the second is that an APC model is mathematically not identifiable, therefore no valid solution is possible.

Methods are now available for analyzing individual-level data, and more details can be found in the book by Yang and Land (2013). Since modeling individual-level data is beyond the scope of this chapter, we provide some description about the non-identifiability issue with APC modeling and progress in methodology research to deal with the issue.

Although the APC model described in the previous section is geometrically beautiful and very attractive to many researchers, the defined model is mathematically not identifiable because the three predictor variable age, period and cohort are not independent from each other. Given any two variable, the third one is determined, and can be derived with the other two. For example, if we know the age of a person and the year he/she participated in survey, the year of birth = survey year – age. Namely,

$$\text{Year of birth (c)} = \text{survey year (p)} - \text{age (a)} \quad (9.2)$$

A number of methodologists in demography, public health, and epidemiology have devoted their efforts attempting to find a solution, including the dedicated work by Mason and Winsboro (1973), Fienberg and Mason (1979), Holford (1983, 1991), Hobcraft, Menken, and Preston (1982), Yang and colleagues (Yang, Fu, & Land, 2004; Yang, Schulhofer-Wohl, Fu, & Land, 2008). A more detailed description of the methodology research can be found in O'Brien (2013); Yang and Land (2013), and Chen & Chen in Chap. 10 of this book. Despite much progress, none of the reported methods is very appealing to researchers till the report of the intrinsic estimator (IE) method established by a group of methodologists through a series of mathematical analyses, simulation studies and empirical application as detailed and summarized in the book by Yang and Land (2013).

The invention of the IE method makes it possible the first time to obtain a set of unique solution to an APC model with a minimum number of constraints that do not affect/altering the estimated age, period and cohort effects. After invention, the IE method has been widely used in reported studies on population mortality (Chen & Wang, 2014a, 2014b), chronic disease mortality (Yang, 2008); suicide death (Chen et al., 2019). In this study, we used the IE method for parameter estimation.

### ***9.3.4 New Data Selection Method to Correctly Estimate Cohort Effect***

In APC modeling analysis, researchers almost always use the following format for data selection: Age-specific data matched and aggregated with the years of data collection (O'Brien, 2013; Yang & Land, 2013). I also used this approach when I conducted my first APC modeling analysis (Chen, Li, Unger, Liu, & Johnson, 2003). Table 9.1 presents an example with hypothetical data to demonstrate the method.



**Table 9.1** Hypothetic mortality data (1/100,000) for APC modeling analysis

Age/period	1990–1994	1995–1999	2010–2004	2005–2009	2010–2014
...					
10–14	53	58	54	48	46
15–19	70	73	64	57	51
20–24	88	86	70	64	64
25–29	123	123	90	80	86
30–34	183	185	135	115	119
35–39	275	282	221	186	166
...					

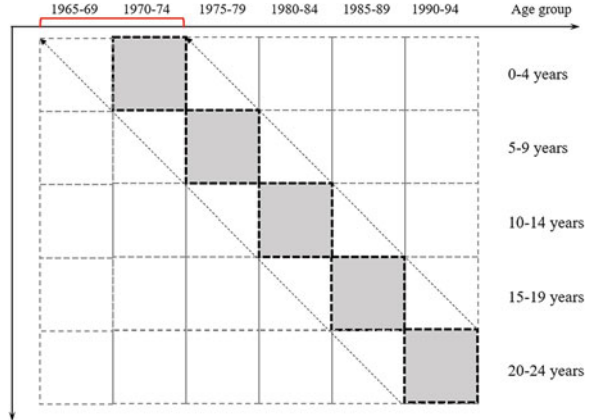
**Fig. 9.3** Conventional approach to determine cohort. **Note:** Persons aged 20–24 in 1990–1994 are coded as born in 1970–1975, following the shaded boxes (a synthetic cohort) as in most reported studies using APC modeling, but birth cohort defined using this method is subject to large error



When mortality data are tabulated with 5 years as the age-group interval, the mortality rate must also be tabulated for 5-year period. Even if mortality data by single year are available, such data must be aggregated by summarizing data from five consecutive years into one group. For example, the mortality rate for 1990–1994 for each age group can be calculated by adding all deaths in an age group during this period and divided by summation of total population of the same age group of all 5 years.

In several of my recent studies, I was stopped by a problem using the aggregate-data in the Table 9.1 format in estimating the birth cohort (Chen et al., 2019; Chen & Wang, 2014a, 2014b). Figure 9.3 presents a hypothetical example. Following the conventional method, to calculate the cohort (the years of birth) for people aged 20–24 years in 1990–1994, we simply trace back of this group of persons from the bottom right to the top left (the shaded boxes) as if they were a cohort or synthetic cohort. By tracking, we simple code the years of birth of the persons aged 20–24 in 1990–1994 as if they were born during 1970–1974 for analysis.

**Fig. 9.4** Miss calculation of birth cohort using conventional approach. **Note:** Persons tabulated by 5-year age group and 5-year period were born not in 5 years but 10 years



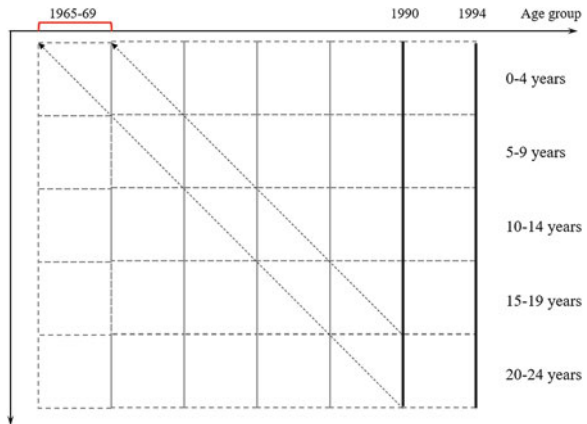
From Fig. 9.3, at first no one can find anything wrong or want to challenge the method. However, a careful examination of this method, we can see that not all persons aged 20–24 in 1990–1994 were born in 1970–1974. This can be illustrated with Fig. 9.4. Persons aged 20–24 in 1990–1994 including those who were 20–24 years old in 1990, 21–34 in 1991, 22–24 in 1992 and 23–24 in 1993 who were born in 1965–1969 as the second dashed line indicated in the figure. With the conventional methods, these persons are treated as if they were all in the year 1994 rather than the period of 1990–1994.

Obviously, cohort or year of birth determined using the conventional methods have two problems. First, the cohort determined is incorrect using the conventional method because approximately 1/3 (the unshaded upper triangle in the top left of Fig. 9.4) of the persons with the coded year of births not equal the true year of birth. According to Fig. 9.4, the correct cohort for persons aged 20–24 in 1990–1994 should be 1965–1974, rather than 1970–1974 as show in this figure and Fig. 9.3. The second problem because of the first problem of misclassification, is the reduction in time resolution for cohort as seen in reported studies (Wang, Hu, Sang, Luo, & Yu, 2017). With the conventional method, the time resolution for cohort effect is reduced 50% from 5 to 10 years. Consequently significant change in cohort effects over time could be smoothed out when the conventional method is used to code year of birth with tabulated data as shown in a few studies.

### 9.3.5 Using Single Year of Data 5 Years Apart as a Solution

To solve the problem discussed in the previous section, we invented a new method in our previous study—using single year of data that are 5 years apart if the mortality data are aggregated using 5 years as the age interval (Chen et al., 2019; Chen & Wang, 2014a, 2014b). As shown in Fig. 9.5, instead of aggregating data from all 5

**Fig. 9.5** New data selection method to correctly determine birth cohort



years from 1990 to 1994, data in 1990 were used. Clearly, persons aged 20–24 in 1990 were born in 5 years during 1965–1969.

With this approach we solved the two problems at the same time—the cohort is now corrected measured, and the time resolution of the cohort has doubled. In addition, this new approach will increase work efficiency since there is no need to aggregate 5-year data; and there is also no need to have data year by year.

## 9.4 Materials and Methods

### 9.4.1 Source of Data

Age-specific mortality rates of urban Chinese population during 1990–2010 were derived from the *China Health Statistical Yearbook*. Mortality data were collected through the Vital Registration System and the Ministry of Health and Hygiene of China and population data were from the National Census of China (Yang et al., 2005). Mortality rates (1/100,000) were computed as the number of deaths over the population, overall and by 5-year age group.

We limited the age-group range from 20–24 to 80–84 years with a total of 13 age groups. Data for people younger than 20 years of age were excluded because the mortality during this period showed a declining trend, different from those 20 years of age and older. The last age group 85+ was excluded because APC model cannot handle open-ended age group (Holford, 1991; Mason & Winsboro, 1973). With data for participants aged 80–84 in 1990, we could assess the risk of mortality as early as 1903–1907 (1990–84 = 1905 and 1990–80 = 1910).

In conventional APC modeling analysis, 5-year average rates are used (Holford, 1991; Mason & Winsboro, 1973; Yang, 2008) (Riggs, McGraw, & Keefover, 1996; Skegg & Cox, 1991; Stack, 2000; Wang et al., 2016). As we discussed early in this

**Table 9.2** Age-specific mortality rate (1/100,000) of people living in urban China, 1990–2010

Age group	1990	1995	2000	2005	2010
20–24	64.25	49.08	35.1	42.22	39.27
25–29	70.18	61.08	51.145	51.64	42.46
30–34	91.17	87.37	73.995	84.72	61.91
35–39	136.15	121.33	118.15	138.12	101.79
40–44	179.63	209.75	197.915	209.68	166.06
45–49	261.68	306.02	310.135	306.01	259.72
50–54	467.12	400.97	398.63	502.66	375.04
55–59	867.79	685.52	559.18	738.03	563.96
60–64	1585.39	1434.75	1096.18	1199.14	956.08
65–69	2661.59	2464.04	2145.73	1967.18	1544.13
70–77	4614.63	4311.49	3787.095	3752.49	2766.82
75–79	7455.59	6751.83	6120.66	5938.69	4889.94
80–84	12641.64	11663.7	10043.995	10092.88	8754.1

**Data sources and definition:** The *Health Statistical Yearbook* of China with mortality rates calculated by dividing the number of deaths with the total population. Data were collected through the Vital Registration System and the Ministry of Health and Hygiene of China and population data were from the National Census of China

chapter, with this method, the cohort could not be correctly defined. Also, persons in a 5-year age group within 5-year period were born in a 10-year period, reducing the time resolution to link years of birth with changes in the risk of mortality. To overcome the problem, we used single-year data 5-year apart in 1990, 1995, 2000, 2005, and 2010. With this approach, data for persons in a 5-year age group in 1 year were all born within 5 years—birth cohort (Chen & Wang, 2014a, 2014b). The mortality data were presented in Table 9.2.

### 9.4.2 APC Modeling Analysis

Let  $m_{ij}$  and  $n_{ij}$  be the mortality rate and number of persons in age group  $i$  and time period  $j$ . By definition, the expected number of deaths in age group  $i$  and time period  $j$  would be  $E_{ij} = m_{ij} \times n_{ij}$ . Assuming a Poisson distribution of the number of deaths, the following log linear model was used to extract cohort effect stretching back to 1905 with data during 1990–2010:

$$\log(E_{ij}) = \mu + \alpha_i + \beta_j + \gamma_k + \epsilon$$

where  $E_{ij}$  = expected number of deaths in age group  $i$  and year  $j$ ;  $\mu$  denotes the intercept or grand mean;  $\alpha_i$  denotes the effect of age-group  $i$  ( $i = 20\text{--}24, 25\text{--}29, \dots, 80\text{--}84$ );  $\beta_j$  denotes the effect of period  $j$  ( $j = 1990, 1995, \dots, 2010$ ); and  $\gamma_k$

denotes the effect of birth cohort  $k$  ( $k = 1906\text{--}2010, 1910\text{--}1914, \dots, 1986\text{--}1990$ ); and  $\epsilon \sim N(0, \sigma)$ .

With this model specification, a total of 42 parameters were estimated (one grand mean  $\mu$ , 13 age effects, 5 period effects, and 17 cohort effects). The 5 period effects plus the 25 cohort effects provided data regarding the risk of mortality among urban Chinese over a period of more than 100 years from 1906–1910 to 2010, with the cohort effects fill in the period from 1906–1989 with no data collection. All the parameters were estimated using the IE method (Fu, 2000; Yang, 2008; Yang et al., 2004). The data-model fit was evaluated using the fit deviance, AIC and BIC. The APC modeling analysis was conducted using the STATA-based software package `apc_ie` with Poisson as the distribution option and `logit` as the link function.

Before the full-scale APC modeling, we fit the data with three 2-component models, including AP (age and period) model, AC (age and cohort) model, and PC (period and cohort) model (see Fig. 9.2). These three models are identifiable, but each of them omitted one predictor. Results from these two-component models were compared with those from the full-scale APC model to show the differences in the estimated effects using different methods and to deepen our understanding of the APC modeling approach.

To effectively describe the historical dynamics of mortality risk change over time through visualization, numerical differentiation was performed over the estimated cohort effect (Chen & Wang, 2014a, 2014b). The numerical differentiations provide an approximate of the changing speed in mortality risk over time. In plotting the changes, y-axis was reversed for the change speed such that any upward trend as a *Sunny Cohort* during which suicide risk was reducing and any downward trend as a *Cloudy Cohort* during which suicide risk was increasing.

As a convention, before statistical modeling analysis, we also plotted the mortality data using several methods to show the potential age, period and cohort effect.

## 9.5 Main Study Findings

### 9.5.1 Visual Presentation of the Mortality Data

Figure 9.6 depicts the age-specific mortality by year. Data from the figure indicate that during the 30-year period from 1990 to 2010, there was a progressive decline in the mortality for people from all age groups. Furthermore, the mortality declining was more pronounced for people 60 years of age and older. A careful review of the five curves, we also noted that there were no observable declines in the mortality from 2000 to 2005 for people in all age groups, including the older residents.

We plot the same data from Fig. 9.6 using the age-group and birth-cohort method (Fig. 9.7). Different from Fig. 9.6, a logarithm scale was used for y-axis to better separate the mortality curves for different age groups. In the figure, each line

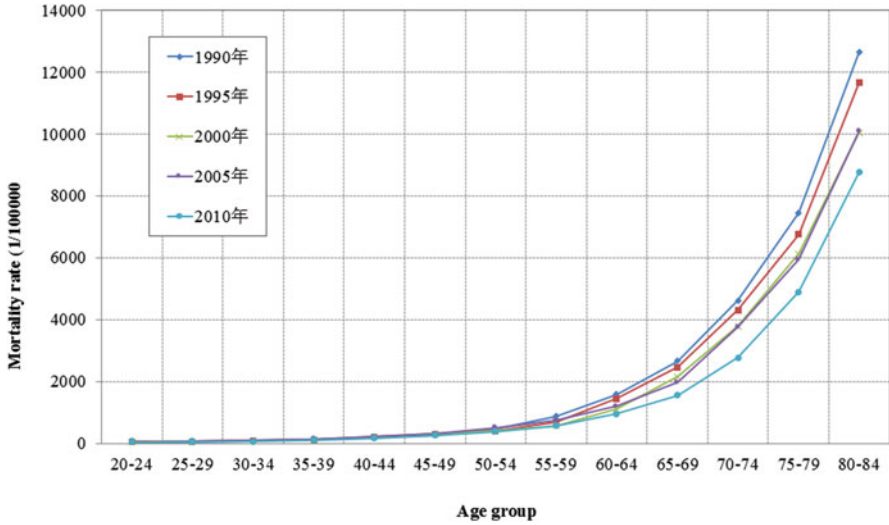
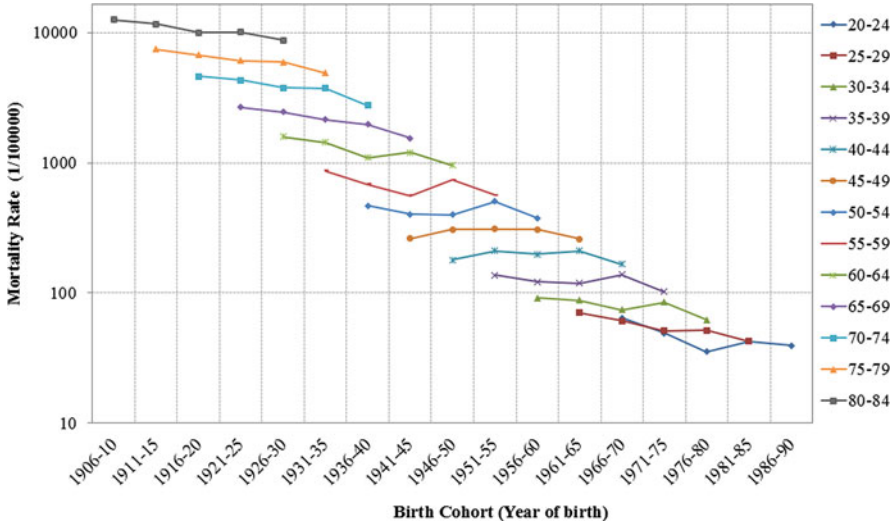


Fig. 9.6 Age-specific mortality of urban Chinese in 1990, 1995, 2000, 2005 and 2010

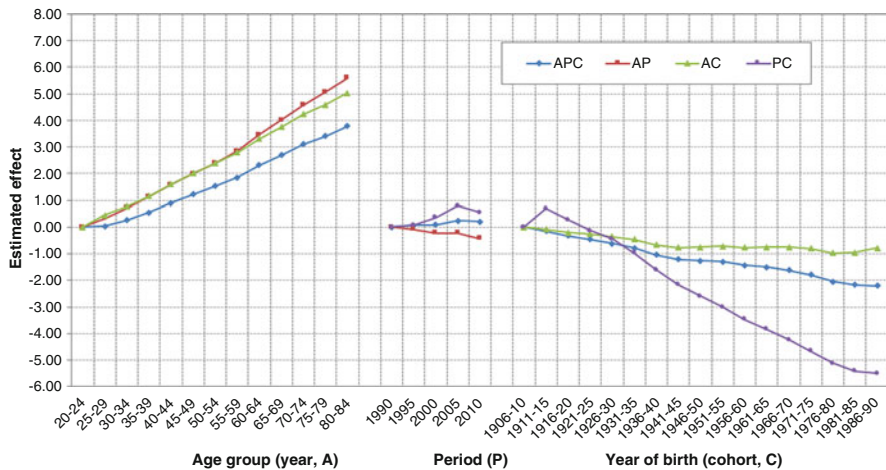
represents an age group with mortality rates in 1990, 1995, 2000, 2005 and 2010. Results from the figure suggest a progressive decline in mortality associated with year of birth from 1906 to 1990. In addition, the four lines located on the top of the map suggest that the mortality decline was smoother and more consistent for those aged 65 years and older, compared to those in all younger age groups. In fact, the mortality in most of the younger working age groups increased from 2000 to 2005 before it declined.

### 9.5.2 Comparison of Results from Four Different APC Models

Results in Fig. 9.8 shows that the age, period and cohort effects estimated with the AP, AC and PC models all differed from those estimated with the APC model. This result suggests that the mortality was related to all three predictors. Specifically, with reference to the result from the APC model, the age effect was over estimated with either the AP or the AC model; the PC model resulted in an excaudate increasing trend in period effect while the AP model suppressed the increasing trend. The differences were more dramatic for the cohort effect. Cohort effect obtained from the PC model increased for the first two cohort and then progressively decline to way below the effect from the APC model while the effect from AC model was progressively higher than that from the APC model.



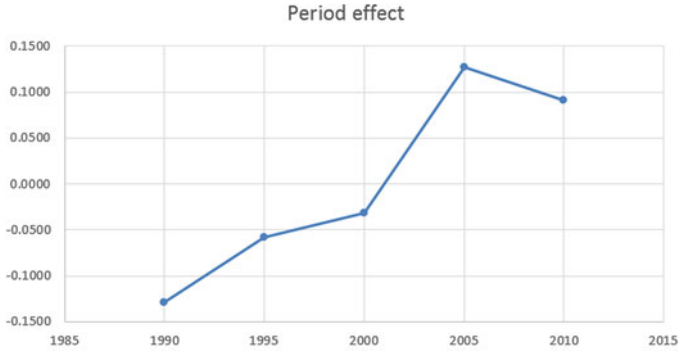
**Fig. 9.7** Age-group and birth-cohort plotting of the mortality rate of the urban Chinese in 1990, 1995, 2000, 2005 and 2010



**Fig. 9.8** Estimated age, period and cohort effects from different APC models. **Data source:** Mortality data for urban Chinese collected in 1990, 1995, 2000, 2005, and 2010

### 9.5.3 Period Effect for Mortality Risk Change over 1990–2010

The period effect estimated from the APC modeling over 1990–2010 is presented in Fig. 9.9. Results in the figure indicate a progressive increase in the risk of death from 1990 to 2005 before it declined. This effect was independent from changes in the age composition of the urban population in China and the year of births



**Fig. 9.9** Period effect  $\beta$  assessing the risk of mortality for urban residents of China, 1990–2010

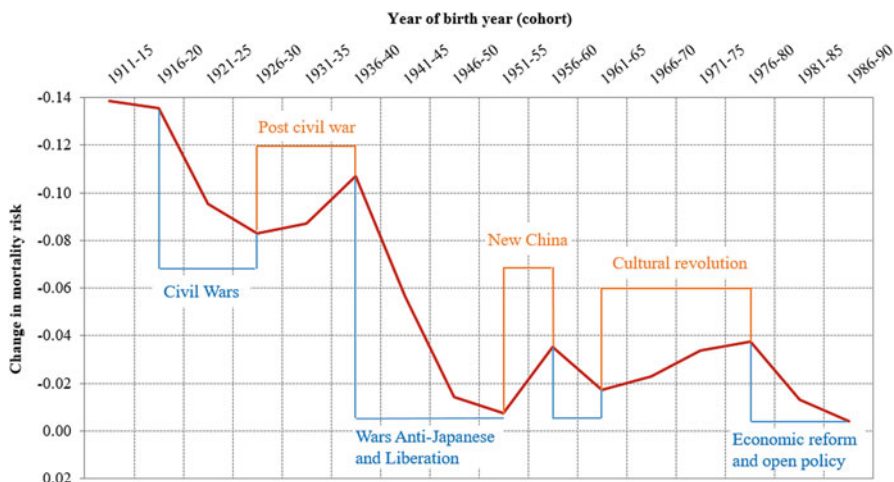
of all urban residents during this period. The period effect extracted through APC modeling reflected the trends in the risk of death all urban residents experienced during the period of 1990–2010.

#### 9.5.4 Changes in Cohort Effect Through Numerical Differentiation

Cohort effect is the information this study intended to extract from the mortality data collected during 1990–2010 to describe changes in the mortality risk among urban Chinese since the early 1900s. The estimated cohort effect estimated using the full-scale APC model for the 1905–1990 was presented in Fig. 9.8 (blue line). Despite an overall declining trends with small ups and downs, the estimated cohort effect did not present any more information to associated with significant historical events in China. To better use the historical mortality risk information derived from the APC modeling analysis, Fig. 9.10 presents the estimated cohort effect through numerical differentiation. For example, the number of the 1911–1915 is the differences in the mortality risk (cohort effect) of this cohort relative to the previous cohort of 1905–1910.

For efficient presentation, in Fig. 9.10 we purposefully reversed the y-axes such that an increase in the curve representing a period of *deceleration* in mortality risk (termed as Sunny Period) and a decline in the curve representing a period of *acceleration* in mortality risk (termed as Cloudy Period).





**Fig. 9.10** Historical changes in mortality risk among urban Chinese, 1911–1990: Numerical differentiation of the cohort effect extracted using an APC model. Note: The Y-axis is reversed such that an upward trend indicating deceleration in mortality risk and a downward trend indicating acceleration in mortality risk

### 9.5.5 Sunny Periods in Historical China

Results in Fig. 9.10 reveal three obvious Sunny Periods: 1926–1936, 1951–1955, and 1961–1975 during which deceleration in the risk of death was observed. The first Sunny Period (1926–1930) was corresponding to the post-civil wars after the success of the Northern Anti-Warlord Military Campaign. The country returned to peace and started the home reconstruction. The slowing down of the mortality risk reduction during the latter half of the period was associated with the historically recorded famine that lasted for 3 years during 1928–1930.

The second Sunny Period (1951–1955) was corresponding to the time after the liberation of the country and the establishment of the People’s Republic of China was established in 1949. Although short, the reduction of mortality risk was reasonable. With a newly liberated and united country after a long period of anti-foreign invasion and domestic war, people started to enjoy peace, reconstruction of the war-broken country and their homes and reconstruction of new lives. Furthermore, the national Patriotic Health and Hygiene Movement was launched during this period for infectious disease control.

The third Sunny Period (1961–1975). This was the longest “healthy period” and overlapped with most part of the 10-year Cultural Revolution from 1966 to 1976. This period is also one of the most controversial periods in contemporary China. Despite the long-standing political and cultural movement, destruction of the traditional Chinese culture, punishment and harsh time for many political leaders and intellectuals, the prolonged declines in the risk of mortality for the whole urban

population could be due to the emphasis of prevention and the construction of the three-tier primary healthcare network system and the promotion of the combination of the Traditional Chinese Medicine with Western Medicine. In addition to health and medicine, researchers, politicians, policymakers and historians can learn more from this listen in China.

### **9.5.6 *Cloudy Period in Historical China***

Results in Fig. 9.10 also showed Four Cloudy Periods: 1916–1925; 1936–1945, 1956–1960, and 1976–1990. The first Cloudy Period (1916–1925) was corresponded with the first and second civil wars (1914–1926) in China after the establishment of the Republic of China in 1911. This was only a part of the long transitions from the 2000-year long feudalistic society to the modern society for a country with the largest population in the world.

The second Cloudy Period (1936–1950). This period was directly related to the Anti-Japanese War as part of the World War II from 1936 to 1945 and the War for Liberation (1946–1949). Among the four cloudy periods, the impact of this one is the largest. China as country dominated with agriculture was invaded by Japan, an industrialized country that attempt to dominate Asia by occupying the largest neighbor country and massacring millions of people with war machines. A recorded famine in 1942–1943 further worsened the survival of the population in China. The cohort effect revealed through this APC modeling provides the first time systematic data showing the increased mortality by Japanese invasion.

The third Cloudy Period (1956–1960). This period, fortunately short, was associated with a number of natural, social and political events, including the Great Leap Forward, the Anti-Righteous Agents Movement, and the sudden destruction of the diplomatic relationship with the former Soviet Union, and another recorded famine in 1959 in the history of China.

The last Cloudy Period (1976–1990). This period was corresponding to the most well-known time in recent China, the Economic Reform and the Open Policy with unprecedented rural-to-urban migration, rapid economic growth, quick industrialization, speedy technological advancement, and dramatic cultural change and diversification. Unfortunately, the estimated cohort effect indicate that risk of mortality accelerated for the urban Chinese during this period, consistent with the recorded mortality rates in the data and the estimated period effect from 1990 to 2005.

## **9.6 Discussion and Conclusions**

In this study, we demonstrated on how to analyze data collected in recent years using APC modeling approach to address historical epidemiological problems. With

data collected in more recent years from 1990 to 2010 among urban population, systematic information was extracted from the data to accurately describe the changes in mortality risk since the early 1900s when the country had been torn by civil wars, foreign invasion, natural disasters, political and cultural revolutions, industrialization, and urbanization. The type of data we used in this study are widely available (see Chap. 1 in this book) and the APC modeling is easy to implement with the newly invented IE method (Y. Yang & Land, 2013). The revision to use single year data as well as the introduction to use numerical differentiation further sharpen the APC modeling method for researchers to gain more detailed insight into historical changes with recent data (Chen et al., 2019; Chen & Wang, 2014a, 2014b). In addition to the total mortality, the same analysis can be used to assess individual causes of deaths as well as health behaviors as reported in the literature (Chen et al., 2003).

### ***9.6.1 Findings and Implications for China***

Findings about the historical changes in mortality risk among urban population in China since the early 1900s has significant implications.

Good health policies as the best health protection measure.

The first and the most striking finding is the continuous improvement in mortality risk during the 10-year Cultural Revolution. Although substantial damages from Cultural Revolution are reported, the mortality risk persistently declined for the longest period in modern China, but why? A careful review of the related literature indicated that during the Cultural Revolution period, a series of important health policies were established and implemented, including the policy of “Prevention First”, “Emphasis of Primary Healthcare”, “Establishment of the Three-Tiered Primary Care System”, and “Combination of Traditional Chinese Medicine with Modern Medicine”. Despite the large-scale and long-standing social and political turmoil with very low levels of economy and medical technology in China during that period, no other factor can convincingly explain the continuous improvement in people’s health at the national level except these significant and well-known health policies and programs.

Peace is the best and most efficient guard for population health.

The second important findings of this study is the provision of systematic data document the impact of foreign invasion and domestic wars on increased mortality risk. Although such risk is obvious, but no systematical and valid data has been reported in the literature. In the history of more than 100 years in China since the 1900s, mortality risk was the highest corresponding to the time of Japanese invasion and the anti-Japanese War during the World War II. In addition, the impact of domestic or civil wars was also substantial; and the least was the social political turmoil and natural disasters, which showed the smallest impact in magnitude and lasted short in time.

### 9.6.2 *Economic and Technic Advancement Not Equal to Good Health*

Finding from the historical analysis of the data in China indicate that many advancements in recent China, including economy, technology, and culture since the 1970s did not translated into population health by reducing mortality. China has made substantial advancement in economy, technology and culture after implementing the reform and open policy since the 1980. However, evidence from our analysis indicate an increase in risk of mortality for urban Chinese during this period, including the cohort effect derived from age-specific mortality during 1976–1990, and the net period effect (after controlling for age and cohort) during the 1990–2010. This increased mortality occurred in a period when the per capita GDP in China increased from \$156 in 1978 to \$3838 in 2010 (World Bank); health expenditure increased at 11.32% annually from 1978 to 2010 (China Embassy in the United, 2012).

Fortunately, findings of this study revealed a small decline in the population mortality risk in China since 2005 as the economy of the country continued to grow. This decline was exactly in consistency with the health policy change in China (Gao, Tang, Tolhurst, & Rao, 2001; Wang, Gusmano, & Cao, 2011) from a market-oriented healthcare back to re-establishing the free health care system, re-emphasizing the prevention first policy and re-building the Three-Tiered Primary Care System that once benefited people's health at the time before the Reform and Open Policy started in China (Hsiao, 1995).

### 9.6.3 *APC Modeling for Historical Epidemiology*

By conducting APC modeling analysis and teaching the method to graduate students, I proposed the term historical epidemiology with APC modeling approach (Chen et al., 2019; Chen & Wang, 2014a, 2014b). The idea of historical epidemiology started with the concept that *data by age are digital fossils*, and these fossils contain information regarding the past of an issue we would like to know. Although contained in the age-specific data, the information is indexed by the year of birth—cohort effect. Furthermore, such information can be extracted with APC modeling method, a readily available tool to extract the indexed information. Findings from this study add new evidence to my previous research, supporting the historical epidemiology research with APC modeling method and aggregated data.

Several issues are to be considered when conducting APC modeling for historical epidemiological research. First, your data have not to be in 5-year age group, shorter age groups are even better to describe the risk in the past. For example, if single-year data are available, cohort effect can thus be estimated by single year of birth, more accurate than 5-year interval. Cohort effect estimated by single year will be better than by multi-year to associate the risk with historical events. Second, the older the

age in data, the longer the research can trace back to the history. If data are available up to 100 years of age, such data can be used to trace back literally to 100 years back in the past. Third, data for only 1 year cannot be analyzed using APC modeling method; you must have data for multiple years. Mathematically, data for a minimum of 3 years are needed to obtain valid effect estimates.

Application of the APC modeling in research has been hindered for long time because of the non-identifiability issue inherited with the method. The invention of the IE method has promoted the use of APC modeling in research (Chen et al., 2019; Chung, Yip, Chan, & Wong, 2016; Li, Wang, Gao, Xu, & Chen, 2016; Wang et al., 2017; Yang & Land, 2013). Despite many strengths, commercial software is needed to implement the IE method. Furthermore, the IE method reads complex even for researchers with training in mathematics and statistics, which may reduce the confidence of a researcher to select the for APC modeling. In recent methodology study, we have developed another method to deal with the identifiability issue inherited with the APC model—Solving APC model using generalized inverse matrix method. This method is based on the well-established Moore and Penrose generalized inverse matrix theory (Moore & Barnard, 1935; Penrose, 1955) and can be executed using R, and software free of charge. Chapter 10 that follows next provides more detailed description about this method.

#### ***9.6.4 Implication for Research in Other Countries***

One purpose of this chapter is to introduce this method for anyone to conduct research studies in countries all over the world. With research findings from more countries, they will inform local health policy and decision-making with both a historical and a global perspective. The success of this study indicate that similar studies can be conducted for many countries in the world as long as data on population health are available in recent one or two decades. For example, if the earliest data are available for a health measure for people in a country up to 80 years of age in 1980, such data can be used to assess health status since 1900. Such data will help us much to investigate social, political, economic and technological development in affecting population health. As we will see, such information is unique and important for researchers, policy-makers, and the general public to know the health history of their country, to make decisions and to from relevant policies and strategies to handle challenge medical and health issues by adding a historical perspective. When data from many countries are reported in the literature, we can form a historical and global perspective to understand human health and to deal with global and local health challenges with a broader spatiotemporal perspective.

### 9.6.5 *Limitations and Conclusion Remarks*

There are several limitations. First, no mortality data by single year of age was available from the data source. This prevented us from using single but 5-year interval to assessing the timing of an observed change in the mortality risk over time. Second, we did not quantify the relationship between the historical events and the mortality risk. Conducting such studies need to have solid methods to code the historical events, which is technically challenging. More methodological research is needed to code these events for quantitative (i.e., correlation and regression) analyses.

Despite the limitations described above, this study is the first one to demonstrate the potentials to examine historical events on population health with APC modeling methods and recent data. It sets an example for researchers in different countries in the world to conduct historical epidemiological research, to obtain new data supporting global health and modern epidemiology.

## References

- Chen, X., Li, G., Unger, J. B., Liu, X., & Johnson, C. A. (2003). Secular trends in adolescent never smoking from 1990 to 1999 in California: An age-period-cohort analysis. *American Journal of Public Health, 93*(12), 2099–2104.
- Chen, X., Sun, Y., Li, Z., Yu, B., Gao, G., & Wang, P. (2019). Historical trends in suicide risk for the residents of mainland China: APC modeling of the archived national suicide mortality rates during 1987-2012. *Social Psychiatry and Psychiatric Epidemiology, 54*(1), 99–110. <https://doi.org/10.1007/s00127-018-1593-z>
- Chen, X., & Wang, P. G. (2014a). Social change and national health dynamics in China. *Chinese Journal of Population Science (in Chinese), 2*, 63–73.
- Chen, X., & Wang, P. G. (2014b). Social change and the dynamic change in national health status in China (in Chinese). *Chinese Journal of Population Science (in Chinese), 26*, 63–73.
- Chung, R. Y., Yip, B. H., Chan, S. S., & Wong, S. Y. (2016). Cohort effects of suicide mortality are sex specific in the rapidly developed Hong Kong Chinese population, 1976-2010. *Depression and Anxiety, 33*(6), 558–566. <https://doi.org/10.1002/da.22431>
- Fienberg, S. E., & Mason, W. M. (1979). Identification and estimation of age-period-cohort models in the analysis of discrete archival data. *Sociological Methodology, 10*(1979), 1–67.
- Fu, W. J. J. (2000). Ridge estimator in singular design with application to age-period-cohort analysis of disease rates. *Communications in Statistics-Theory and Methods, 29*(2), 263–278. <https://doi.org/10.1080/03610920008832483>
- Gao, J., Tang, S., Tolhurst, R., & Rao, K. (2001). Changing access to health services in urban China: Implications for equity. *Health Policy and Planning, 16*(3), 302–312. <https://doi.org/10.1093/heapol/16.3.302>
- Hobcraft, J., Menken, J., & Preston, S. (1982). Age, period, and cohort effects in demography - A review. *Population Index, 48*(1), 4–43. <https://doi.org/10.2307/2736356>
- Holford, T. R. (1983). The estimation of age, period and cohort effects for vital rates. *Biometrics, 39*(2), 311–324.
- Holford, T. R. (1991). Understanding the effects of age, period, and cohort on incidence and mortality rates. *Annual Review of Public Health, 12*, 425–457. <https://doi.org/10.1146/annurev.pu.12.050191.002233>

- Hsiao, W. C. (1995). The Chinese health care system: Lessons for other nations. *Social Science & Medicine*, 41(8), 1047–1055.
- Li, Z., Wang, P. G., Gao, G., Xu, C. L., & Chen, X. G. (2016). Age-period-cohort analysis of infectious disease mortality in urban-rural China, 1990–2010. *International Journal for Equity in Health*, 15, 55. <https://doi.org/10.1186/s12939-016-0343-7>
- Mason, K. O., & Winsboro, H. (1973). Some methodological issues in cohort analysis of archival data. *American Sociological Review*, 38(2), 242–258. <https://doi.org/10.2307/2094398>
- McKeown, R. E. (2009). The epidemiologic transition: Changing patterns of mortality and population dynamics. *American Journal of Lifestyle Medicine*, 3(1 Suppl), 19S–26S. <https://doi.org/10.1177/1559827609335350>
- Moore, E. H., & Barnard, R. W. (1935). *General analysis part I*. Philadelphia, PA: The American Philosophical Society.
- O'Brien, R. M. (2013). Comment of Liying Luo's article, "assessing validity and application scope of the intrinsic estimator approach to the age-period-cohort problem". *Demography*, 50(6), 1973–1975. <https://doi.org/10.1007/s13524-013-0250-0>
- Omran, A. R. (2005). The epidemiologic transition: A theory of the epidemiology of population change. *Milbank Quarterly*, 83(4), 731–757. <https://doi.org/10.1111/j.1468-0009.2005.00398.x>. (Original work published, 1971).
- Penrose, R. A. (1955). Generalized inverse for matrices. *Proceedings of the Cambridge Philosophical Society*, 51, 406–413.
- Riggs, J. E., McGraw, R. L., & Keefover, R. W. (1996). Suicide in the United States, 1951–1988: Constant age-period-cohort rates in 40- to 44-year-old men. *Comprehensive Psychiatry*, 37(3), 222–225. [https://doi.org/10.1016/S0010-440x\(96\)90039-5](https://doi.org/10.1016/S0010-440x(96)90039-5)
- Skegg, K., & Cox, B. (1991). Suicide in New-Zealand 1957–1986 - The influence of age, period and birth-cohort. *Australian and New Zealand Journal of Psychiatry*, 25(2), 181–190. <https://doi.org/10.1080/00048679109077733>
- Stack, S. (2000). Suicide: A 15-year review of the sociological literature part II: Modernization and social integration perspectives. *Suicide and Life-threatening Behavior*, 30(2), 163–176.
- States, China Embassy in the United. (2012). *Medical and Health Services in China (government white paper)*. Embassy of the People's Republic of China in the United States of America.
- Wang, H., Gusmano, M. K., & Cao, Q. (2011). An evaluation of the policy on community health organizations in China: Will the priority of new healthcare reform in China be a success? *Health Policy*, 99(1), 37–43. <https://doi.org/10.1016/j.healthpol.2010.07.003>
- Wang, Z. K., Hu, S. B., Sang, S. P., Luo, L. S., & Yu, C. H. (2017). Age-period-cohort analysis of stroke mortality in China: Data from the Global Burden of Disease Study 2013. *Stroke*, 48(2), 271–275. <https://doi.org/10.1161/Strokeaha.116.015031>
- Wang, Z. K., Wang, J. Y., Bao, J. Z., Gao, X. D., Yu, C. H., & Xiang, H. Y. (2016). Temporal trends of suicide mortality in mainland China: Results from the age-period-cohort framework. *International Journal of Environmental Research and Public Health*, 13(8), 784. <https://doi.org/10.3390/Ijerp13080784>
- Yang, G., Hu, J., Rao, K. Q., Ma, J., Rao, C., & Lopez, A. D. (2005). Mortality registration and surveillance in China: History, current situation and challenges. *Population Health Metrics*, 3(1), 3. <https://doi.org/10.1186/1478-7954-3-3>
- Yang, Y. (2008). Trends in US adult chronic disease mortality, 1960–1999: Age, period, and cohort variations. *Demography*, 45(2), 387–416. <https://doi.org/10.1353/Dem.0.0000>
- Yang, Y., Fu, W. J. J., & Land, K. C. (2004). A methodological comparison of age-period-cohort models: The intrinsic estimator and conventional generalized linear models. *Sociological Methodology*, 34, 75–110. <https://doi.org/10.1111/j.0081-1750.2004.00148.x>
- Yang, Y., & Land, K. C. (2013). *Age-period-cohort analysis: New models, methods, and empirical applications*. Boca Raton, FL: Chapman & Hall/CRC.

- Yang, Y., Schulhofer-Wohl, S., Fu, W. J. J., & Land, K. C. (2008). The intrinsic estimator for age-period-cohort analysis: What it is and how to use it. *American Journal of Sociology*, *113*(6), 1697–1736. <https://doi.org/10.1086/587154>
- Zhou, M., Wang, H., Zeng, X., Yin, P., Zhu, J., Chen, W., . . . Liang, X. (2019). Mortality, morbidity, and risk factors in China and its provinces, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet*. [https://doi.org/10.1016/S0140-6736\(19\)30427-1](https://doi.org/10.1016/S0140-6736(19)30427-1)