MINI-REVIEW

Sources of Information on the Burden of Cancer in China

Ling Yang1,2, Donald Maxwell Parkin1, Liandi Li2, Yude Chen3

Abstract

Three data sources for the information on cancer in China are described in this paper: (i) mortality data from national retrospective surveys (1973-75 and 1990-92); (ii) mortality data obtained through special research projects (CHIS, DSP and mortality survey in 1986-89); and (iii) incidence and mortality data from cancer registries. Different combinations of mortality and incidence data can be used to estimate the pattern or burden of cancer in China. Registration of cancer incidence and mortality in China should be standardized and expanded, in order to enhance availability of accurate data for estimating cancer burden in China.

Key Words: Neoplasms - Information science - China

Introduction

Based on the most recently available international incidence and mortality data, it is estimated that there were 10.1 million new cases, 6.2 million deaths and 22 million people living with cancer (within 5 years of diagnosis) in 2000 (Parkin, 2001). But in 1990, the number was 8.1 million new cancer cases (Parkin et al., 1999) and 5.2 million deaths (Pisani et al., 1999). Driven principally by the world population growth and aging, cancer is becoming a major health problem for most countries. The estimation of global cancer burden and trends depends upon the availability of accurate incidence and mortality data. Furthermore, at the national level, accurate data are important for planning cancer control activities (such as cancer prevention, screening and treatment) and for monitoring the effects of such interventions. China is the largest country in the world with more than 1.3 billion citizens, comprising of one fifth of the world’s population. Since a large fraction of the global cancer burden is provided by China, inferences about global cancer trends and the overall effectiveness of cancer control worldwide are heavily weighted by this source, and accurate estimates of incidence and mortality are therefore very important in a global context, and for national policy setting. There are very big differences in social, economic and health status indicators within this huge country. The densely populated areas and those of higher socioeconomic status are mostly located in the eastern and southern parts, while the northern and western regions are less economically developed and more sparsely populated. It is therefore difficult to plan and implement nation wide cancer control strategies.

China has a long history of medical care, and data about cancer incidence and mortality have been collected for several decades, although until recently the resulting information has mainly been published in Chinese-language journals and reports (Tu 1985; Wang et al., 1995; Gao 1991; Chen 2000).

1Unit of Descriptive Epidemiology, International Agency for Research on Cancer (IARC), Lyon 69732, France, 2National Office on Cancer Prevention and Control (NOCPC), Beijing 100021, China, 3National Center for Health Information and Statistics of Ministry of Health, Beijing 100044, China.

Correspondence to Dr Ling Yang, Unit of Descriptive Epidemiology, International Agency for Research on Cancer, 150, cours Albert Thomas, 69572 Lyon Cedex 08, France Tel: (+33) (0)4 72 73 86 67 Fax: (+33) (0)4 72 73 86 96 E-mail: ling@iarc.fr

The national civil registration system currently in operation in China - the National Household Registration System - was established in 1954, implemented in 1955, and is run by the Public Security Department. This system collects information on vital events, through registration of births and deaths, on a national basis since 1958. China has regulations governing the registration of deaths, i.e. there is legal provision for death registration. The level of medical certification of causes of death, however, was low: only a quarter and perhaps as little as one-fifth of deaths were accompanied by a qualified report on the cause of death (WHO 1994). A nationwide check of registration data conducted prior to the census in 1982 revealed some underreporting, particularly of births. The government has since increased the quality of registration through various means, notably by increasing the number and qualifications of agents at all levels, stressing the independence of registration, encouraging uniformity of data and developing electronic data processing (WHO 1994).

The National Office for Cancer Prevention and Control (NOCPC) was established in Beijing in 1969. Its role is to promote cancer prevention and control in whole of China; update national norms for cancer control; monitor cancer rates and risk factors; improve cancer diagnosis and treatment; guide administration of cancer hospitals; publish the Bulletin of Chinese Cancer; organize task-forces for key projects etc. NOCPC has organized two national retrospective surveys of the cause of death in the country. These provided for the first time an indication of the sex- and age-specific mortality rates for the principal diseases affecting a fifth of the world's population. The results of these two national surveys provide a sound basis for establishing priorities for research in basic and clinical sciences in China.

The first cancer registry in China was set up in Shanghai in 1963, but cancer registration has developed slowly in the country. In 1990, NOCPC set up the Chinese Cancer Registry Coordination Committee, which has developed national norms and a manual for a cancer registration system for China. A Cancer Registry and Surveillance Association was established in 2001 in Beijing, to set up a national cancer registry network in collaboration with the International Agency for Research on Cancer (IARC). However, it is known that there are many registries established by local institutions in China not included in NOCPC’s registry network (NOCPC 2001).

Data Sources

The sources of information on cancer in China fall into three broad groups: (i) mortality data from national retrospective surveys, (ii) mortality data obtained through special research projects, and (iii) incidence and mortality data from cancer registries. Different combinations of these data sets have been used to estimate the pattern or burden of cancer in China (Parkin et al. 1999; Pisani et al. 1999; Ferlay et al. 2001; WHO 1997; World Bank 1993).

1. Mortality data from national retrospective surveys

(1) The first national retrospective survey of death in China was conducted in 1976 in 2392 counties/cities in 29 provinces, and covered about 850 million people, or 96.7% of the population of the country. Only 35 remote counties in Tibet and Sichuan Provinces were excluded. The cause of death was determined retrospectively for deaths in the preceding 3-year period, 1973-1975. The demographic information was based on local civil vital statistics departments or security departments’ records. Using a uniform questionnaire for cause of death, nearly one million health workers obtained information from various levels of the medical system, with corresponding level of accuracy. Interviews were carried out at community level, including with family members and community leaders. At that time, there were five grades of Chinese hospitals: village, commune, county, district in prefecture/city, and province. Four diagnostic categories of cause of death were used on the questionnaire: I – based on pathology, cytology and bone marrow examination; II - clinical investigations, including exploratory laparotomy (without biopsy), x-ray, ultrasound, isotope scans, biochemical and immunological laboratory tests; III - clinical case-history evaluation; IV - evaluation after death. Figure 1 shows the percentages of deaths recorded at each diagnostic level; 70.5% of deaths had an antemortem diagnosis made at level I or II.


With respect to deaths from cancer, 79.2% had been recorded from records at the county hospital level, implying a full range of medical services for determining cause. A further 13.3% of deaths were recorded from commune level hospitals in which there was likely to have been at least one physician present. Analysis of the data from this survey

![Figure 1. Percentage for Different Diagnosis Basis Levels for Cancer in 1973-75 Death Survey, China](image-url)
showed that the national average annual age-standardized cancer mortality rate (using the Chinese census data in 1964 as standard population) was 80.2 per 100,000 in males, (comprising 11.3% of all deaths) and 54.3 per 100,000 in females (8.6% of all deaths). During 1973-1975, cancer was the second leading cause of death (following respiratory disease) for males and ranked third for females (following respiratory, heart diseases other than coronary heart disease and cerebrovascular disease). But among the 2392 counties, there were 429 counties/cities where cancer ranked as the first cause of death for males and 161 counties/cities where cancer was the leading causes of death in females. For males, the most common cancer sites were stomach, liver, lung and colon/rectum; for females, they were stomach, cervix uteri, esophagus, liver and lung (NOCPC 1979). The results of data analysis show clearly that cancer risk varied substantially from region to region and each cancer has its particular geographic distribution. Based on the results of this study, the Atlas of Cancer Mortality in the People’s Republic of China was published in 1979 (Editorial Committee for the Atlas of Cancer Mortality 1979). Fig. 2 shows one of maps from that atlas. The maps and tables shown in this atlas provided a resource useful as a first step in the study of cancer causation. Geographic patterns and associations may suggest where further epidemiologic studies of risk factors should be concentrated.

Describing the first survey of mortality, and the resulting cancer Atlas, Doll and Greenwald (Chen et al. 1990) wrote: “Few medical projects can ever have been successfully carried through that compare in scale with the national survey of the causes of death undertaken in China in 1976. That immense task, which involved identification of about 20 million deaths in the three previous years and retrospective diagnosis of their causes, provided for the first time in indication of the sex- and age-specific mortality rates for the principal diseases affecting a fifth of the world’s population. From the purely epidemiological point of view this result alone was of great value; but the value of the survey was immensely enhanced by the demonstration that the rates for every major cause of death varied widely from one part of the country to another, and varied, moreover, more widely than in any other country for which useful mortality data were available. This, it was realized, provided Chinese epidemiologists with a unique opportunity for testing hypotheses about the etiology of disease and possibly for obtaining some entirely new clues, if only the mortality rates could be married to reliable information about local environments and the ways of life of local populations.”

(2) A second retrospective death survey, also organized by NOCPC, was carried out in 1990-1992. The survey involved 263 units (74 cities & 189 counties) sampled from 2400 cities/counties in 27 provinces. It included 10% of the total Chinese population (335 million person years) were included in the sample. A two-stage randomized stratified cluster sampling method was used to select the 263 units, based on the distribution of mortality rates in the survey of
1973-1975. The basic unit was a county, or a district in large city. Fig.3 shows the sampling sites chosen. Using the data of the first national mortality survey in 1973-1975, the comparison between the sampling sites that were selected in 1990-1992 survey and the whole country was done (Li et al., 1996).

According to the result from that comparison, the mortality rate for all causes was 6.7 for the sampling units, 6.9 for the whole population; the mortality rate for cancer was 83.7 per 100,000, 74.6 per 100,000 respectively during 1973-1975. And the proportion for each malignant cancer and the distribution in sampling sites were also consistent as the country level distribution during that time. Thus the sampling sites used in 1990-1992 mortality survey were, at least as far as the results of the earlier (1973-1975) survey showed, highly representative of the country as a whole. Among the 362,900 deaths from cancer during 1990-1992, the cause for 91.0% was obtained from records from hospitals at county level or above, among them 50.9% from hospitals at province, prefecture or city levels. 31.8% of deaths had a cause of death based on pathology evaluation; when combined with autopsy, surgery, clinical investigations, clinical chemistry estimation, the proportion was about 90%.

According to this survey, cancer ranked second among all causes of death (after respiratory disease), the age-standardized mortality rate (using the Chinese census data in 1982 as standard population), was 94.4 per 100,000 (123.6 for males and 66.3 for females). The five most common cancers in males were stomach, liver, lung, esophagus and colon/rectum; in females they were stomach, esophagus, liver, lung and colon/rectum. The age-standard mortality rate for all cancers in rural areas was 96.5 per 100,000, higher than in urban areas (89.8 per 100,000). A similar geographic distribution of cancers in China was observed as in the 1973-1975 survey. However, the actual mortality rates and the ranking between city and rural areas were quite different. In general, cancer mortality rates had increased especially for lung, liver, stomach cancers and leukemia, but certain cancers, such as cervix uteri, nasopharynx and esophagus cancer, had declined. The increase in mortality was greater for rural populations than for cities, especially for cancer of the lung, liver and leukemia. The data is now kept in NOCPC in Beijing, some results have been published (Li et al. 1996; Li et al., 1997; Zhou et al., 1997).

2 Mortality data from special research projects

(1) In 1973, the Ministry of Health in China set up a National Mortality Statistics system derived from a 10% sample of the population (between 100 and 120 million persons) yielding at least 600,000 deaths annually. The Center for Health Information and Statistics (CHIS) of the Ministry of Health is now responsible for this system (WHO 1994). The geographic area covered by the sample consists essentially of easily accessible prefectures/cities and counties, on the eastern seaboard, and corresponds to the most densely populated areas of China. In many cases, sites are included on the basis of their voluntary participation in the data collection exercise (NOCPC 2001). According to the Annual Report published in 1999 (National Center of Health Information and Statistics 1999), the sample population comprised 15 big cities, 21 middle-sized or small
The sites participating include:

**Big Prefectures/cities:** Beijing, Tianjin, Haerbin, Changchun, Shenyang, Dalian, Anshan, Shanghai, Nanjing, Wuhan, Guangzhou, Chengdu, Chongqing, Xi’an, Zhengzhou.

**Middle or small cities:** Suzhou, Xuzhou, Anqing, Hefei, Ma’anshan, Tongling, Fuzhou, Xiamen, Bengpu, Yichang, Huangshi, Fuoshan, Zigong, Guiulin, Chishan, Xiantan, Heyang, Changde, Wulumuqi, Pingdingshan, Xinyang.

**Counties:** 85 counties in 15 provinces/municipalities, that include Beijing, Tianjin, and Shanghai, Chongqin, Jiangsu, Zhejiang, Fujian, Hubei, Hunan, Guangdong, Anhui, Sichuan, Henan, Guizhou, Gansu.

The sample sites are not representative of the national population as a whole. According the research on the classification of health status in China based on principal factor and cluster analysis, Chinese counties are divided into four different classes: from class 1 to 4, the health status declined (Rao et al. 1989). Among the 85 counties which involved in the National Mortality Statistics system, the counties in Beijing, Tianjin, Shanghai, Jiangsu, Zhejiang belong to the first class; counties in Fujian, Hubei, Hunan, Guangdong, Sichuan, Henan belong to the second class and counties in Anhui, Guizhou, Gansu belong to the second class. There are no counties participating in NMS system from the fourth class. The annual report includes data on causes of death, incidence and mortality for infectious disease and so on. These data have been reported to World Health Organization (WHO) by CHIS since 1987. Several years of data have already been published in the WHO for urban and rural sample sites separately, as well as for both combined. The data for the years 1987 – 1999 are on the IARC – DEP mortality database, at: http://www-dep.iarc.fr/who/mortality.htm

(2) Disease Surveillance Points System (DSP): From 1980, under the direction of the Anti-Epidemic Department of the Ministry of Health and the Chinese Academy of Preventive Medicine (CAPM), Anti-Epidemic Stations in 29 provinces/municipalities were recruited to practice a health and vital statistics reporting system called ‘Disease Surveillance Points’ (DSP). At first, the DSP focused mainly on acute infectious disease, and included 71 points in both city and rural areas, covering a population of about 10 million people (one percent of the Chinese population). Although it provided a lot of information about vital events, it was hardly representative on a national scale, since all these surveillance points were volunteer participants in the system, and most were located in the big cities or counties with high socioeconomic status. Thus in 1989, a second stage of the national disease surveillance points was introduced, using a multistage stratified sampling method, based on data from the third national census, to get a more representative sample for disease information at a national level (NOCP 2001; Yang et al. 1992). Geographic, administrative and various health indicators of the areas used as surveillance units were taken into account during the sample. First, 13 geographic areas were selected, based on geographic strata, then according to the health status in each area, a total of 71 strata were selected. Finally, 145 surveillance points were selected, each of them covering between 30,000 and 100,000 people. During sampling, 15% of the points were replaced by other points in the same strata to guarantee the quality of data collection. The information on mortality is based on the death certificate card from hospitals, with the ICD-9 code added during data input to computer (Yang et al., 1992).

The surveillance points were distributed between the different provinces as follows: Beijing (2), Tianjin (3), Hebei (6), Shanxi (5), Inner-Mongolia (3), Liaoning (5), Jilin (4), Heilongjiang (4), Shanghai (2), Jiangsu (5), Zhejiang (6), Anhui (6), Fujian (5), Jiangxi (5), Shandong (11), Henan (6), Hubei (5), Hunan (6), Guangdong (5), Guangxi (6), Sichuan (7), Guizhou (6), Yunnan (4), Tibet (2), Shaanxi (6), Gansu (6), Qinghai (3), Ningxia (4), Xinjiang (5), Hainan (2).

All surveillance points are divided into 7 regional-economic strata. There are 3 urban strata (C1 - big city with more than 1 million population; C2 - middle city with population between 0.5 and 1 million; C3 - small city with less than 0.5 million people). Rural areas are divided into 4 strata, based on socioeconomic status (L1 – high level, L2 – middle level, L3 – low-level, L4 – lowest level). In the second stage of DSP, the new points were recruited to all 7 strata, and population weightings are said to be available for the whole country. Seven indices were used to test the representativity of the surveillance points: GNP, illiterate rate, birth rate, infant death rate, crude death rate, proportion of the population in age-group 0-14, proportion of the population in the over 65 age-group. Based on the census data in 1982, there were no significant differences between the distribution of these indices in the survey sites and the general population; the 145 sampling sites are therefore presumably representative of the country as far as health status in concerned (Yang et al., 1992; Yang et al. 1996).

However, the DSP population is certainly much younger than the national population, and this should be taken into account when analyzing the data. This system collects a wide range of health indicators, including occurrence of infectious disease, chronic disease, birth and death (including cause of death). In 1993, a set of indices for data quality control was produced and a program to investigate under-reporting, based on capture-recapture methods, was prepared. This quantitative estimate is used to calculate an under-reporting ratio, at three-year intervals. The estimated under-reporting is used to adjust the reported mortality, using formula:

\[
\text{adjusted rate} = \frac{\text{reported rate}}{1 - \text{underreporting ratio}}
\]

(Yang et al., 1996; Chinese Academy of Preventive Medicine 1997).

The mortality data from the DSP system have been used to review the health situation in China in World Bank report (World Bank 1993). The data are held in CAPM and an
annual report is prepared including data quality evaluation, demographic surveillance report, causes of death and surveillance for infectious disease (Chinese Academy of Preventive Medicine 1997).

Based on the DSP system, CAPM and the Clinical Trial Service Unit (CTSU) in UK carried out a nationally representative, prospective survey in 1990-1991 to monitor the evolving epidemic of mortality from tobacco in China, following the large increase in male cigarette smoking in recent decades (Niu et al., 1998). All men aged more than 40 in two or three units (urban street committees or rural villages) in 45 representative disease surveillance points were sought, and about 80% (224,500) were interviewed on smoking, drinking and medical history. Mortality was monitored prospectively through local residential records. Causes of death recorded on the official death certificate were supplemented by medical notes, and were coded to ICD-9 by trained staff, blinded to baseline information. By January 1996, 13,412 deaths had occurred. The results show that the overall mortality of smokers was significantly higher than that of never smokers (RR=1.2, P<0.00001). The excess mortality among smokers chiefly involved neoplasms (1.3; P<0.0001), respiratory diseases (1.4, P<0.0001) and cardiovascular diseases (1.1, P<0.01). Long-term follow-up is continuing, with periodic re-surveys of all middle-aged adults living in the study areas, to monitor the evolution of the epidemic of tobacco related death.

(3) During 1989-1991, CAMS, CAPM and CTSU organized the world’s largest retrospective study of smoking and death (Liu et al., 1998). In 98 areas of China (24 cities (non-randomly) and 74 rural counties (random sampling based on the 1970s national death survey)), five hundred researchers interviewed the surviving family members (or sometimes, in rural areas, other informants) of the one million people who died during 1986-1988. When the interviewee was the surviving spouse (300,000 cases), their smoking habits were also recorded. This cause-specific mortality data is now available on the Internet at: http://www.ctsu.ox.ac.uk/projects/cecology1989/.

3 Incidence and mortality data from Cancer Registries:

The first cancer registry in China was established in Shanghai urban area in 1963. In the 1970s, cancer registries were created in Beijing, Tianjin, Nanjin, Zhongshan County, Qidong County, Fusui County and Linxian County. Since the 1980s, more and more cities/counties have set up cancer registries, such as Shenyang, Dalian, Jinzhou, Anshan and Haerbin. In 1995, ‘Chinese cancer incidence, mortality and risk factor surveillance programme’ was organized in eleven cities/counties by NOCPC and CHIS. These experimental units included 5 big cities (Beijing, Tianjin, Shanghai, Wuhan and Haerbin) and 6 high-risk areas for certain cancers (Cixian, Linzhou, Changle, Qidong, Jiashan, Fusui). The cancer registries are all population-based, and follow the methodology and format described in ‘Cancer Incidence in Five Continents (CIFC)’; the ICD-9 was used for coding cancer site (Li, et al., 2000). The results from the programme have been published as ‘Cancer Incidence and Mortality Rates in Experimental Cities/Counties in China (1988-1992)’ (WHO 1994; Li et al., 2000). Furthermore, NOCPC will continue to publish ‘Cancer Incidence and Mortality Rates in Experimental Cities/Counties in China (1993-1997)’ and ‘Cancer Risk Factor Research in Experimental Cities/Counties in China’ in future. Although these experimental units are not representative of the national population, the program provides experience and a standard for other cancer research.

Table 1. Cancer Registries Included in Different Volume of Cancer Incidence in Five Continents (CIFC)

<table>
<thead>
<tr>
<th>Cancer Registry</th>
<th>Cancer Incidence in Five Continents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tianjin</td>
<td></td>
</tr>
<tr>
<td>Qidong</td>
<td></td>
</tr>
<tr>
<td>Beijing</td>
<td></td>
</tr>
<tr>
<td>Jiashan</td>
<td></td>
</tr>
<tr>
<td>Wuhan</td>
<td></td>
</tr>
<tr>
<td>Changle</td>
<td></td>
</tr>
<tr>
<td>Cixian</td>
<td></td>
</tr>
</tbody>
</table>

: for all cancer sites;  : for certain cancer sites only
Data for Cancer Burden in China

registries being established in China. In Cancer Incidence in Five Continents Volume VIII (CIFC-VIII), data from eight Chinese Cancer Registries in the mainland will be included. These registries are: Beijing, Shanghai, Tianjin, Wuhan, Qidong County, Jiashan County, Changle County, Cixian County. The last two will be included as special cancer registries, for certain cancer sites only. Table 1 shows the cancer registry data from China mainland that been published in successive volumes of Cancer Incidence in Five Continents (Waterhouse et al., 1982; Muir et al., 1987; Parkin et al., 1992; Parkin et al., 1997). Fig. 4 shows the map of these cancer registries in China.

Apart from this national cancer register network, other cancer registries, sometimes part of research projects or monitoring cancer in special cancer high-risk areas are known to exist. Such registries are not part of any formal network, so receive no central technical or financial support, and the quality of the data collected has not been evaluated.

Conclusions

Cancer is becoming a major health problem in China. It has increased a lot during last thirty years and will continue to do so in the next few decades, along with the aging of the population. The huge difference of cancer patterns and social-economic status between urban and rural areas make cancer prevention and control work difficult. But from the purely epidemiological point of view, the presence of high-risk areas for certain cancers provides good resources and opportunities for cancer research and study, for testing hypotheses about the etiology of disease and possibly for obtaining some entirely new clues. Cancer register data will enhance this work. Thus, registration of cancer incidence and mortality in China should be standardized and expanded, in order to enhance availability of accurate data for estimating the pattern and burden of cancer in China.

Acknowledgement

This work was supported by International Agency of Research on Cancer. The authors appreciate the great help on data providing from Dr. Chalapati Rao (World Health Organization) and Mr. Siwei Zhang (National Office on Cancer Prevention and Control, China).

References


