ISSN:2790-1661 DOI: 10.56028/aemr.3.1.173

Statistics and Prediction: a Preliminary Study on the Birth Month and Diseases of The Elderly over 80 Years Old in China

Shangjie Ge-Zhang[1,3], Yating Liu[1,3], Qiaochu Wei [2], Lili Su[1,*]

[1]Northeast Forestry University, Harbin 150040, PR China.

[2] Xiamen University, Xiamen 361104, PR China.

[3]These authors contributed equally to this work.

*fangwen@bupt.edu.cn

Abstract. The month of birth is related to the risk difference of diseases, and this research has been widely verified internationally. However, in China, there is still no study covering the whole country's birth date-multiple diseases at present. This study initiated a nationwide study on the relationship between the birth date of the elderly and their susceptibility to diseases. SPSS 26.0 software is used to analyze the Chinese Longitudinal Healthy Longevity Survey (CLHLS) by descriptive statistics, chi-square test and multiple regression model. The experimental data covered 23 provinces and 11 common diseases, and this article found that 5 of them were related to birth date. It should be pointed out that this is only a preliminary conclusion, and further conclusions and data will be given in the follow-up study. Although the internal complex relationship has not been deeply discussed in this paper, the research results provide some basis for tracing the source of diseases, and help the elderly who have reached or are about to reach the age of 80 to prevent and understand their prone diseases.

Keywords: Birth month, Statistics, Incidence of diseases, Disease prevention.

1. Introduction

The month of birth is related to the congenital health condition and even the risk difference of diseases [1-5]. Tatonetti et al. [6] made logistic regression analysis on the data set of 1,749,400 patients in New York Presbyterian Church/Columbia University Medical Center (CUMC), and used chi-square test to prove the significance, pointing out that 55 diseases were significantly correlated with birth month。 Zhang et al. evaluated the association between birth month, birth season and total mortality rate and cerebrovascular disease mortality rate, and examined the role of family and socio-economic factors in these associations. They found that the mortality rate of cerebrovascular disease in women born in spring and summer, especially in March-July, increased, and family and socio-economic factors did not change these associations [7]. In Sweden, Ueda et al. [8,9] studied the relationship between birth month and specific causes of death based on the longitudinal cohort of population, and found that birth month is related to the mortality of cerebrovascular diseases and infections over 50-80 years old. Furthermore, birth month is related to the mortality risk of the age group over 50 years old, but the mortality results of 30-50 years old are uncertain. Yang et al. [10] adopted a time-stratified cross-case design, It was found that the exposure level of PM2.5 in the birth season reflected that it might affect lung development in the early stage, and had a potential impact on the burden of respiratory diseases related to environmental PM2.5 exposure in later life.

In China, this research mainly focuses on specific diseases and specific regions [11-16]. At present, there is no clear research that can cover the whole country's birth date and multiple diseases. In this paper, through SPSS 26.0, using descriptive statistics, chi-square test and multiple Logstic regression model, we analyzed a number of disease data sets in Chinese Longitudinal Healthy Longevity Survey (CLHLS), and got a wide range of relationships between birth date and susceptible diseases. Although we haven't deeply explored this inherent complex relationship, the research results provide some basis for the tracing of diseases, help the general public, especially the elderly, to know the relationship between diseases and time nodes, and provide reference for the disease prevention of the elderly.

ISSN:2790-1661 DOI: 10.56028/aemr.3.1.173

2. Sources and Methods

2.1 Objects

In this paper, the recorded data from the Chinese Longitudinal Healthy Longevity Survey (CLHLS) are used for analysis, and the subjects lacking key variables (demographics, disease history, birth month) are excluded, and 9,090 groups of valid data are extracted from them. Secondly, the same person tracked the survey data for many times, and selected the last valid data as the standard, thus avoiding repeated calculation.

2.2 Follow-Up of Research

The survey covers 23 provinces in China, and the subjects are the elderly aged 65 and above (the elderly were still alive at the last follow-up) and adult children aged 35-64 (the elderly have passed away). The respondents are respectively applicable to the following two types of questionnaires: the questionnaire of living respondents and the questionnaire of family members of deceased elderly people. Following the baseline survey in 1998, the survey project was followed up in 2000, 2002, 2005, 2008-2009, 2011-2012, 2014 and 2017-2018. A total of 113,000 home visits were conducted, and detailed data on the health status, quality of life and medical care needs of the elderly were collected. The database provides data to scholars free of charge, and the procedures followed conform to ethical standards.

2.3 Determination of Experimental Data

The date of birth of the participants was registered in the questionnaire in detail, which was confirmed by the ID information and the Chinese lunar calendar date. First, translate the Chinese lunar calendar date dictated by the elderly into the Gregorian calendar date (Chinese elderly usually take the lunar calendar date as their birth date), and then verify it again according to the ID card information. The illness is provided by the interviewee or his family members, and the illness status is confirmed according to the autopsy report, death certificate, summary of medical records or information obtained from family members.

2.4 Statistical Methods

SPSS 26.0 (IBM Inc, Chicago, IL, USA) was used for analysis. Descriptive statistics clarify the general characteristics of the respondents, the number of cases of various diseases and their respective proportions. Chi-square test was used to explore the correlation between birth date and various diseases in China. For diseases with significant correlation with birth month, multiple regression models were established to further explore the influencing factors of different diseases. $\alpha = 0.05$ is the statistical test level, and $P < 0.05$ is considered to be statistically significant.

3. Results and Discussion

3.1 Statistical Description

Among the respondents, the average age was 93.761 ± 7.764 (Mean ± SD). Among all the participants, 5,453 were women, accounting for 59.99%, which is in line with the sex ratio of this age group [17]. The survey subjects and their diseases are shown in Table 1. In the later analysis of each disease, this study adopted the number of people corresponding to the "yes" part of the table and their birth month.

ISSN:2790-1661 DOI: 10.56028/aemr.3.1.173

Table 1 Data statistics of diseases suffered by respondents

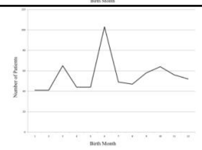
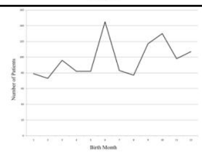
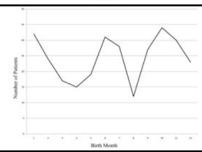
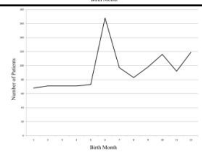
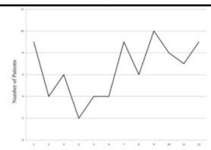
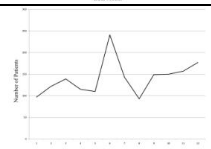
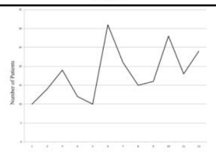
Type of Disease	Answer Classification	Frequency	Percent (%)
Hypertension	Yes	1169	12.9
	No	7212	79.3
	Don't Know	451	5.0
	Missing	258	2.8
Diabetes	Yes	75	0.8
	No	8277	91.1
	Don't Know	461	5.1
	Missing	277	3.0
Cardiopathy	Yes	664	7.3
	No	7721	84.9
	Don't Know	444	4.9
	Missing	261	2.9
Cerebrovascular Disease	Yes	292	3.2
	No	8129	89.4
	Don't Know	397	4.4
	Missing	272	3.0
Bronchitis, Emphysema, Pneumonia, Asthma	Yes	1133	12.5
	No	7338	80.7
	Don't Know	357	3.9
	Missing	262	2.9
Tuberculosis	Yes	78	0.9
	No	8310	91.4
	Don't Know	425	4.7
	Missing	277	3.0
Cataract	Yes	1693	18.6
	No	6725	74.0
	Don't Know	424	4.7
	Missing	248	2.7
Glaucoma	Yes	218	2.4
	No	8154	89.7
	Don't Know	442	4.9
	Missing	276	3.0
Cancer	Yes	48	0.5
	No	8284	91.1
	Don't Know	480	5.3
	Missing	278	3.1
Gastric Or Duodenal Ulcer	Yes	293	3.2
	No	8126	89.4
	Don't Know	403	4.4
	Missing	268	2.9
Parkinson's Disease	Yes	86	0.9
	No	8272	91.0
	Don't Know	453	5.0
	Missing	279	3.1

ISSN:2790-1661 DOI: 10.56028/aemr.3.1.173

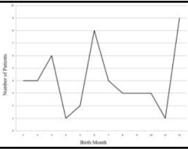
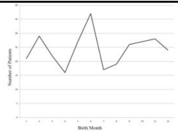
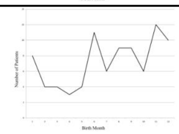
3.2 Chi-Square Test

Chi-square test was used to explore the correlation between the subjects born in different months and their diseases. The specific results are shown in Table 2 below. Hypertension ($\chi 2$=54.499, $P<0.05$), diabetes ($\chi 2$=56.066, $P<0.05$), cerebrovascular disease ($\chi 2$=59.633, $P<0.05$), tuberculosis ($\chi 2$=51.296, $P<0.05$), cataract ($\chi 2$=49.108， $P<0.05$) were related to the birth month of the investigated person, and the test results were statistically significant. Previous studies have found that rising temperatures and heat waves have a direct impact on health [18]. In this study, the birth months with high incidence of hypertension, diabetes, cerebrovascular disease, tuberculosis and cataract are also concentrated in summer, that is, June to September. Cerebrovascular diseases are more common in people born in autumn (October), and the lowest incidence corresponds to the birth month in August.

Table 2 The results of chi-square test

Type of Disease	Birth Month-Illness Relationship Diagram	χ2Value	P Value	Risk of Birth Month	
				Highest	Lowest
Hypertension		54.499	0.011	June	February
Diabetes		56.066	0.007	September	May & July
Cardiopathy		45.358	0.074	June	January & February
Cerebrovascular Disease		59.633	0.003	October	August
Bronchitis, Emphysema, Pneumonia, Asthma		41.244	0.154	June	January
Tuberculosis		51.296	0.022	September	April
Cataract		49.108	0.035	June	August
Glaucoma		40.481	0.174	June	January & May

ISSN:2790-1661 DOI: 10.56028/aemr.3.1.173

Cancer		34.824	0.381	December	April & November
Gastric Or Duodenal Ulcer		41.188	0.155	June	April
Parkinson's Disease		35.831	0.337	November	April

3.3 Multivariate Logstic Regression Model

The multivariate Logstic regression model was established for five statistically significant diseases detected by Chi-square. According to the Omnibus test results of model coefficients, the significance of Hypertension, Diabetes, Cerebrovascular disease and Cataract are all less than 0.05, and the variables included in the model are statistically significant. However, the significance of Tuberculosis is 0.294, which is greater than 0.05, and the variables included in the model have no statistical significance. In addition, all the models have passed the Hosmer-Lemeshow test, and the goodness of fit of the models is high. The data are suitable for establishing binary logstic regression models.

In this study, gender had been proved to influence cerebrovascular diseases and cataracts. Women are more likely to suffer from cerebrovascular diseases than men, while cataracts show the opposite trend. In addition, hypertension, diabetes, cerebrovascular disease and cataract all showed a increasing prevalence rate with the increase of age, which are the diseases that the elderly should pay more attention to in prevention. Women are more likely to suffer from cerebrovascular diseases than men, which may be due to the fact that, relatively speaking, most women are physically weaker than men and are more susceptible to heat [19]. Secondly, many studies show that there is a connection between environment, nutrition and human health in early life [20-23]. In China, at the beginning of the 20th century, due to the lack of advanced planting and cultivation techniques, food supply largely followed the objective seasonal law. In summer or at the turn of summer and autumn, it is very likely that there will be food shortage, which can be expressed by the idiom "when the new crop is still in the blade and the old one is all consumed", that is, a temporary food shortage period when old food is finished and new food is not harvested. The prevalence rate may also be related to the lack of food at that time and the inability to supplement vitamins in time [24]. The sex of cataract patients is mostly male, and we guess that this may be related to acquired life and work. Smokers have a higher risk of cataract [25-27]. In addition, men's heavy industry work in the last century may also be an important reason for the high prevalence rate, because heavy industry will produce a lot of gases and particles harmful to eyes [28-30].

4. Conclusion

On the basis of the existing literature, this paper studied the data sample set of the Chinese Longitudinal Healthy Longevity Survey. The main advantage of this research lies in the wide database range and long follow-up time, so this research has the ability to analyze the potential relationship between birth month of the elderly and diseases. In addition, this is the first nationwide study on the relationship between birth date and various diseases in China. However, this research also has some limitations. First of all, this research is only based on statistics in the form of data, but it still lacks comprehensiveness of conditions. We didn't investigate and analyze the climate, region and living habits of the interviewees, nor did we have a more detailed understanding of their economic status, parents' educational level and mothers' nutritional status during pregnancy [19].

ISSN:2790-1661 DOI: 10.56028/aemr.3.1.173
Secondly, there is no detailed qualitative research on morbidity statistics of people, such as physique and genetic inheritance, whether such factors are also closely related to diseases is unknown. Third, social, economic and spiritual factors in old age may also be related to diseases. Finally, although there were 9,090 groups in the database, the number of patients with some diseases was only a few hundred, which would cause errors, so it is necessary to expand the database. A larger amount of data is needed to make the analysis result more accurate and get an effective and reliable forecast. Although this study needs more information to determine the deeper complex relationship, it has already revealed the potential relationship between birth month and the existence of five diseases, which can bring some hints for disease prevention of the elderly. However, it should be noted that this study cannot be considered suitable for people of all ages, because some serious diseases may cause few people to live over 80 years old, which will result in no sample records in the database of this study. This research is suitable for people who have reached or will reach the age of 80, and is used to predict and prevent potential diseases.

References

[1] Lippi G, Salvagno G L, Montagnana M, et al. Birth season predicts the values of red blood cell distribution width (RDW) in adulthood[J]. Clinical Chemistry and Laboratory Medicine (CCLM), 2016, 54(4): 667-671.

[2] Li L, Boland M R, Miotto R, et al. Replicating cardiovascular condition-birth month associations[J]. Scientific reports, 2016, 6(1): 1-7.

[3] Vassallo M F, Banerji A, Rudders S A, et al. Season of birth and food allergy in children[J]. Annals of Allergy, Asthma & Immunology, 2010, 104(4): 307-313.

[4] Auger N, Fraser W D, Sauve R, et al. Risk of congenital heart defects after ambient heat exposure early in pregnancy[J]. Environmental Health Perspectives, 2017, 125(1): 8-14.

[5] Zhang X, Wang Y, Chen X, et al. Associations between prenatal sunshine exposure and birth outcomes in China[J]. Science of The Total Environment, 2020, 713: 136472.

[6] Boland M R, Shahn Z, Madigan D, et al. Birth month affects lifetime disease risk: a phenome-wide method[J]. Journal of the American Medical Informatics Association, 2015, 22(5): 1042-1053.

[7] Zhang Y, Devore E, Strohmaiers, et al. Birth month, birth season, and overall and cardiovascular disease mortality in US women: prospective cohort study[J]. bmj, 2019, 367.

[8] Ueda P, Edstedt Bonamy A K, Granath F, et al. Month of birth and cause-specific mortality between 50 and 80 years: a population-based longitudinal cohort study in Sweden[J] . European journal of epidemiology, 2014, 29(2): 89-94.

[9] Ueda P, Edstedt Bonamy A K, Granath F, Et al. Month of Birth and Mortality in Sweden: A Nation-wide Population-based Cohort Study [J]. Plos One, 2013, 8 (2): E56425.

[10] Yang T, He T, Huang J, et al. Impact of birth season on the years of life lost from respiratory diseases in the elderly related to ambient PM2. 5 exposure in Ningbo, China[J]. Environmental health and preventive medicine, 2021, 26(1): 1-9.

[11] Guo J, Xu L, Wang J, et al. The month of birth has a seasonal effect in Chinese patients with narcolepsy and cataplexy[J]. Journal of Clinical Sleep Medicine, 2022, 18(2): 461-467.

[12] Zhao L, Chen L, Yang T, et al. Birth prevalence of congenital heart disease in China, 1980–2019: a systematic review and meta-analysis of 617 studies[J]. European journal of epidemiology, 2020, 35(7): 631-642.

[13] Bai Y, Shang G, Wang L, et al. The relationship between birth season and early childhood development: Evidence from northwest rural China[J]. PLoS One, 2018, 13(10): e0205281.

[14] Ma Q, Xu W, Zhou X, Et. The relationship of season of birth with reflexive error in very young children in eastern China [J]. PLOS ONE, 2014, 9 (6): E100472.

[15] Xie D, Fang J, Liu Z, et al. Epidemiology and major subtypes of congenital heart defects in Hunan Province, China[J]. Medicine, 2018, 97(31).

ISSN:2790-1661 DOI: 10.56028/aemr.3.1.173
[16] Yu M, Ping Z, Zhang S, et al. The survey of birth defects rate based on birth registration system[J]. Chinese medical journal, 2015, 128(01): 7-14.

[17] Cheng Z, Ma J. Evolution of China's aging population and countermeasures (in Chinese) [J]. Academic Exchange, 2018 (12): 101-109.

[18] Arbuthnott K G, Hajat S. The health effects of hotter summers and heat waves in the population of the United Kingdom: a review of the evidence[J]. Environmental health, 2017, 16(1): 1-13.

[19] Gao Z, Liu S, Dai Y, et al. Born in winter or spring more susceptible to all-cause and cardiovascular disease death in rural areas of China: results from a 11.9-year follow-up study[J]. Journal of human hypertension, 2021, 35(12): 1170-1179.

[20] Maccini s, Yang D. Under the weather: health, scholling, and economic consequences of early-life rainfall[J]. American Economic Review, 2009, 99(3): 1006-26.

[21] Rolland-Cachera M F, Maillot M, Deheeger M, et al. Association of nutrition in early life with body fat and serum leptin at adult age[J]. International journal of obesity, 2013, 37(8): 1116-1122.

[22] Elo I T, Preston S H. Effects of early-life conditions on adult mortality: a review[J]. Population index, 1992: 186-212.

[23] Joseph K S, Kramer M S. Review of the evidence on fetal and early childhood antecedents of adult chronic disease[J]. Epidemiologic reviews, 1996, 18(2): 158-174.

[24] Pasco J A, Wark J D, Carlin J B, et al. Maternal vitamin D in pregnancy may influence not only offspring bone mass but other aspects of musculoskeletal health and adiposity[J]. Medical hypotheses, 2008, 71(2): 266-269.

[25] Chowdhury A, Choudhury A, Chakraborty S, et al. p-Benzoquinone-induced aggregation and perturbation of structure and chaperone function of α-crystallin is a causative factor of cigarette smoke-related cataractogenesis[J]. Toxicology, 2018, 394: 11-18.

[26] Lee H J, Kim C H, Lee J S, et al. Association between cataract and cotinine-verified smoking status in 11 435 Korean adults using Korea National Health and Nutrition Examination Survey data from 2008 to 2016[J]. Journal of Cataract & Refractive Surgery, 2020, 46(1): 45-54.

[27] Sánchez A G, Colligris B, Teruel Y Á, et al. Effects of cigarette smoke extract on the zebrafish visual system[J]. Investigative Ophthalmology & Visual Science, 2021, 62(8): 3163-3163.

[28] Lin C C, Chiu C C, Lee P Y, et al. The Adverse Effects of Air Pollution on the Eye: A Review[J]. International Journal of Environmental Research and Public Health, 2022, 19(3): 1186.

[29] Amegah A K. Global health risk factors: Air pollution[J]. Handbook of Global Health, 2020: 1-19.

[30] Mosley S. Environmental history of air pollution and protection[M]//The basic environmental history. Springer, Cham, 2014: 143-169.