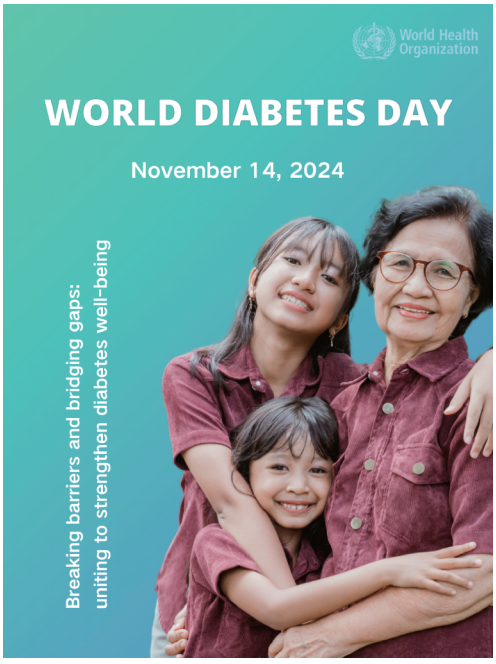
ISSN 2096-7071 (Print)
ISSN 2096-3101 (Online)
CN 10-1629/R1

CHINA CDC WEEKLY

Vol. 6 No. 45 Nov. 8, 2024

中国疾病预防控制中心周报



ISSN 2096-7071
9 772096 707248 45>

China CDC Weekly

Cover Image: adapted from the World Health Organization, https://www.who.int/campaigns/world-diabetes-day/2024.

Vital Surveillances

Trends in the Mortality Rate of Major Kidney Diseases — China, 2014–2021

Xinhui Yu[1]; Jinlei Qi[1]; Peng Yin[1]; Limin Wang[1]; Yunning Liu[1]; Maigeng Zhou[1]; Lijun Wang[1,#]

ABSTRACT

Introduction: Kidney disease represents a significant public health issue in China, yet there is a lack of comprehensive knowledge regarding national and regional trends in its mortality and causes. This study evaluated the mortality, causes, and regional distribution of major kidney diseases in China from 2014 to 2021.

Methods: Data pertaining to kidney disease were obtained from the National Death Cause Surveillance System. Estimates were made for both the mortality rates and age-standardized mortality rates (ASMR) for kidney cancer, glomerular disease, tubulointerstitial nephritis, and kidney failure. Additionally, the average annual percent change (AAPC) was calculated to illustrate trends by sex, urban/rural distinctions, and regional differences from 2014 to 2021.

Results: There was a significant reduction in the ASMR for all kidney diseases combined from 2014 to 2021. Analysis of age-specific mortality rates reveals a gradual increase beginning at age 35, with a sharp rise after age 60. However, the ASMR for glomerular disease and kidney failure in females, as well as for tubulointerstitial nephritis in males, displayed a notable decrease. Regionally, the ASMR for glomerular disease in the eastern region and kidney failure in the western region significantly decreased by AAPC of –4.6% and –2.3%, respectively. Conversely, the ASMR for kidney cancer in the central region rose significantly (AAPC=2.1%).

Conclusion: From 2014 to 2021, the ASMR for major kidney diseases remained high among men, urban residents, and individuals in western China. Future prevention and control initiatives should prioritize these disparities to mitigate the impact of kidney diseases.

Kidney diseases encompass a variety of progressive disorders that disrupt the structure and function of the kidneys because of diverse etiologies. Notable among these are kidney cancer, glomerular disease, tubulointerstitial nephritis, and kidney failure (*1*). These diseases significantly affect global health; notably, chronic kidney disease caused approximately 1.2 million deaths globally in 2017, reflecting a 41.5% increase in the global chronic kidney disease mortality rate from 1990 to 2017 (*2*). According to the 2019 GBD study, the mortality rate from kidney cancer in China was 2.79 per 100,000 persons in 2019, and the mortality rates of both kidney cancer and chronic kidney disease have been rising since 1990 (*3*). The aging population has further exacerbated the burden, making kidney diseases a progressively common cause of death. Remarkably, nearly one-fifth of all global chronic kidney disease patients are in China (*4*). Despite the significance of this issue, few studies have examined mortality due to major kidney diseases at a national level. Therefore, this study utilizes data from the National Cause of Death Surveillance System to assess the mortality trends associated with major kidney diseases among Chinese residents from 2014 to 2021, aiming to inform targeted prevention and control strategies.

METHODS

This study sourced its data from 605 surveillance sites within the National Death Cause Surveillance System, covering the period from 2014 to 2021. Surveillance sites reporting an all-cause mortality rate below 5 ‰ were systematically excluded annually. Major kidney diseases were classified according to the International Classification of Disease 10th Revision (ICD-10) and included kidney cancer (C64–C65), glomerular disease (N00–N08), tubulointerstitial nephritis (N10–N16), and kidney failure (N17–N19). Standard population demographics were derived from the Yearbook of the Seventh National Census of 2020 published by the National Bureau of Statistics.

The study examined various mortality indicators concerning major kidney diseases from 2014 to 2021.

These included the crude mortality rate, age-specific mortality rate, and age-standardized mortality rates (ASMRs) for kidney cancer, glomerular disease, tubulointerstitial nephritis, and kidney failure across different demographics (males and females), as well as urban versus rural settings and regional distinctions (Eastern, Central, and Western). These regions were further divided into 31 provincial-level administrative divisions (PLADs): 11 in the Eastern region (Beijing, Tianjin, Hebei, Liaoning, Shanghai, Jiangsu, Zhejiang, Fujian, Shandong, Guangdong, Hainan), 8 in the Central region (Shanxi, Jilin, Heilongjiang, Anhui, Jiangxi, Henan, Hubei, Hunan), and 12 in the Western region (Inner Mongolia, Guangxi, Chongqing, Sichuan, Guizhou, Yunnan, Xizang, Shaanxi, Gansu, Qinghai, Ningxia, Xinjiang). For analysis, SAS (version 9.4; SAS Institute Inc., Cary, USA) calculated the crude and ASMRs, excluding data from surveillance sites with an all-cause mortality rate below 5 ‰. Additionally, Joinpoint software (version 5.0.2, Applications Branch, National Cancer Institute, Bethesda, USA) was used to assess trends in ASMRs of major kidney diseases and to compute the average annual percent change (AAPC) with 95% confidence intervals (*CI*).

Personnel at each surveillance site undergo an annual training course on cause-of-death surveillance, which is organized by the China CDC. They must pass an examination to be registered. Additionally, the examination of cause-of-death registration and reporting is conducted regularly. Regular investigations into underreporting are also performed on a sample of reporting units to ensure data quality.

RESULTS

The crude mortality rates for major kidney diseases demonstrated an increasing trend in both males and females, with AAPCs of 1.8% (95% *CI*: 0.8%, 2.9%) in males and 0.7% (95% *CI*: 0%, 1.4%) in females, and these changes were statistically significant ($P<0.05$). Conversely, the ASMR for females declined from 6.9 per 100,000 to 5.7 per 100,000, with an AAPC of –3.0% (95% *CI*: –4.5%, –1.5%). However, this downward trend was not statistically significant in males ($P>0.05$).

The crude mortality rates for kidney cancer have risen in both males and females, with AAPCs of 4.9% (95% *CI*: 4.0%, 5.9%) and 5.1% (95% *CI*: 3.8%, 6.3%), respectively. In contrast, the ASMR for kidney cancer increased slightly in both genders, with AAPCs of 1.1% (95% *CI*: –0.5%, 2.8%) for males and 1.0% (95% *CI*: –0.6%, 2.6%) for females, though these changes were not statistically significant (both $P>0.05$). Significant declines in ASMR were noted for glomerular disease and kidney failure in females, with AAPCs of –2.9% (95% *CI*: –4.8%, –1.0%) and –3.6% (95% *CI*: –5.3%, –1.8%), respectively. Males displayed a significant reduction in the ASMR for tubulointerstitial nephritis, with an AAPC of –4.7% (95% *CI*: –7.0%, –2.5%) (Table 1).

The age-specific mortality rates for the four principal kidney diseases demonstrated consistent trends, with low rates observed under the age of 35, a gradual rise between ages 35 and 60, a sharp increase after age 60, and peaking in the 85 and older age group. Among these conditions, glomerular diseases had the highest crude mortality rate, followed by kidney failure, kidney cancer, and tubulointerstitial nephritis (Figure 1).

The ASMR for all types of kidney disease demonstrated a decreasing trend in urban areas, with statistically significant reductions ($P<0.05$). The most notable decrease was observed in tubulointerstitial nephritis, which exhibited an AAPC of –7.3% (95% *CI*: –10.7%, –3.9%). Conversely, in rural areas, only the ASMR for tubulointerstitial nephritis followed a similar declining trend, with an AAPC of –4.9% (95% *CI*: –7.3%, –2.3%). The ASMRs for other kidney disease types remained stable (Table 2).

The highest ASMR for kidney cancer was found in the eastern region, surpassing rates in both the central and western regions. Notably, the central region experienced the most significant increase in ASMR for kidney cancer, with an AAPC of 2.1% (95% *CI*: 0.5%, 3.7%). In contrast, ASMRs for other kidney-related diseases have generally decreased across different regions. The eastern region recorded the steepest declines in ASMR for tubulointerstitial nephritis and glomerular disease, with AAPCs of –7.9% (95% *CI*: –9.5%, –6.3%) and –4.6% (95% *CI*: –8.0%, –1.0%), respectively. The central region saw significant reductions in ASMR for tubulointerstitial nephritis and kidney failure, with AAPCs of –8.9% (95% *CI*: –14.9%, –2.5%) and –2.8% (95% *CI*: –5.1%, –0.6%). Additionally, the western region exhibited the most rapid decrease in ASMR for kidney failure, with an AAPC of –2.3% (95% *CI*: –4.3%, –0.2%) (Table 3).

DISCUSSION

This study investigated the trends in mortality rates for major kidney diseases in China from 2014 to 2021.

TABLE 1. Mortality rate and age-standardized mortality rate (per 100,000) for major kidney diseases in China, 2014–2021.

Year	Male										Female									
	Kidney cancer		Glomerular disease		Tubulointerstitial nephritis		Kidney failure		Total		Kidney cancer		Glomerular disease		Tubulointerstitial nephritis		Kidney failure		Total	
	CMR	ASMR	CMR	ASMR	CMR	ASMR	CMR	ASMR	CMR	ASMR	CMR	ASMR	CMR	ASMR	CMR	ASMR	CMR	ASMR	CMR	ASMR
2014	1.1	1.4	4.4	5.5	0.4	0.6	2.4	3.0	8.2	10.4	0.6	0.7	3.5	3.8	0.3	0.4	1.8	2.0	6.2	6.9
2015	1.2	1.5	4.7	6.0	0.4	0.5	2.3	2.9	8.6	10.8	0.7	0.8	3.6	4.0	0.3	0.4	1.7	1.9	6.3	7.0
2016	1.3	1.5	5.0	6.0	0.4	0.5	2.2	2.6	8.8	10.6	0.7	0.8	3.7	3.9	0.4	0.4	1.6	1.7	6.4	6.8
2017	1.3	1.6	5.3	6.3	0.4	0.5	2.2	2.7	9.1	11.0	0.7	0.8	3.8	4.1	0.3	0.3	1.6	1.7	6.4	6.8
2018	1.4	1.6	5.1	5.9	0.4	0.5	2.3	2.7	9.1	10.6	0.8	0.8	3.8	3.9	0.3	0.3	1.7	1.7	6.5	6.7
2019	1.5	1.7	5.0	5.7	0.4	0.4	2.2	2.5	9.0	10.3	0.8	0.8	3.7	3.7	0.3	0.3	1.6	1.6	6.5	6.5
2020	1.5	1.6	4.8	5.3	0.4	0.4	2.3	2.5	9.0	9.8	0.8	0.8	3.6	3.3	0.3	0.3	1.6	1.5	6.3	5.9
2021	1.6	1.6	5.3	5.3	0.4	0.4	2.6	2.6	9.8	9.8	0.9	0.8	3.7	3.1	0.3	0.2	1.9	1.6	6.7	5.7
AAPC (%)	4.9	1.1	2.1	−1.1	−1.6	−4.7	1.6	−2.2	1.8	−1.1	5.1	1.0	0.6	−2.9	−2.4	−7.6	0.0	−3.6	0.7	−3.0
95% *CI* lower	4.0	−0.5	−1.7	−3.2	−3.4	−7.0	−1.5	−5.1	0.8	−2.7	3.8	−0.6	−0.6	−4.8	−6.9	−15.4	−2.5	−5.3	0.0	−4.5
95% *CI* upper	5.9	2.8	6.1	1.1	0.3	−2.5	4.9	0.8	2.9	0.5	6.3	2.6	1.8	−1.0	2.3	0.9	2.6	−1.8	1.4	−1.5
P	0.001	0.175	0.286	0.341	0.078	0.002	0.318	0.141	0.005	0.191	0.001	0.168	0.332	0.003	0.310	0.078	0.967	0.003	0.049	0.001

Abbreviation: CMR=crude mortality rate; ASMR=age-standardized mortality rate; AAPC=average annual percent change; *CI*=confidence intervals.

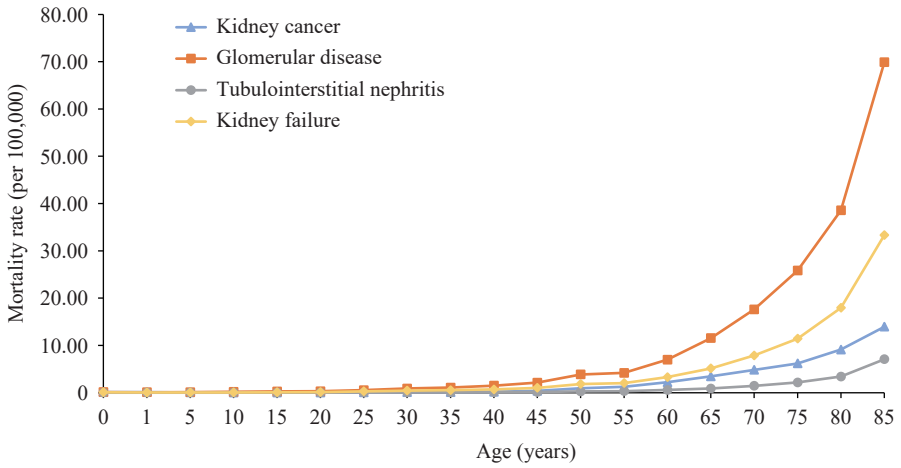


FIGURE 1. Age-specific mortality rates for major kidney diseases in China, 2014–2021.

Overall, the ASMR was higher in males than females and higher in urban areas than in rural areas. Except for kidney cancer, the ASMR was higher in western regions than in the central and eastern regions. The crude mortality rate of all types of major kidney diseases gradually increased with age. In addition to an increase in the ASMR of kidney cancer in the central region, that of major kidney diseases decreased in most areas and populations.

The crude mortality rates for both male and female patients with kidney cancer and major kidney diseases have shown an upward trend. However, there is no significant increase in the ASMR for kidney cancer in either gender, likely reflective of the gradual aging of the population. This study highlights that, excluding kidney cancer, the ASMR for glomerular disease, tubulointerstitial nephritis, and kidney failure generally displays a decreasing trend across most regions and populations. Particularly in economically prosperous areas, notably urban and eastern regions, this decline can be attributed to advances in medical technology, enhancements in healthcare services, and greater access to health education. These factors contribute to more timely and effective diagnosis and treatment of kidney diseases, thereby improving the prognosis for those diagnosed.

TABLE 2. Differences in age−standardized mortality rates (per 100,000) for major kidney diseases between urban and rural areas in China, 2014–2021.

Year	Urban					Rural				
	Kidney cancer	Glomerular disease	Tubulointerstitial nephritis	Kidney failure	Total	Kidney cancer	Glomerular disease	Tubulointerstitial nephritis	Kidney failure	Total
2014	1.7	4.1	0.4	2.8	9.0	0.7	4.8	0.5	2.3	8.3
2015	1.7	4.4	0.4	2.5	9.0	0.9	5.2	0.4	2.3	8.7
2016	1.7	4.4	0.4	2.5	8.9	0.9	5.2	0.5	2.0	8.5
2017	1.7	4.4	0.4	2.6	9.1	0.9	5.5	0.4	1.9	8.7
2018	1.6	4.3	0.4	2.5	8.7	0.9	5.1	0.4	2.0	8.4
2019	1.6	4.0	0.3	2.3	8.2	1.0	5.0	0.4	1.9	8.3
2020	1.4	3.5	0.3	2.1	7.3	1.0	4.7	0.4	1.9	7.9
2021	1.4	3.3	0.2	2.1	7.1	1.0	4.5	0.3	2.1	7.9
AAPC (%)	−2.7	−3.7	−7.3	−3.6	−3.7	4.1	−1.3	−4.9	−2.4	−1.1
95% *CI*	(−5.1, −0.3)	(−5.9, −1.4)	(−10.7, −3.9)	(−5.5, −1.7)	(−5.4, −2.0)	(−1.0, 9.4)	(−2.9, 0.4)	(−7.3, −2.3)	(−6.0, 1.5)	(−2.8, 0.6)
P	0.028	0.002	0.002	0.004	<0.001	0.119	0.132	0.004	0.222	0.214

Abbreviation: AAPC=average annual percent change; *CI*=confidence intervals.

TABLE 3. Differences among regions in age−standardized mortality rate (per 100,000) for major kidney diseases in China, 2014–2021.

Year	Eastern region					Central region					Western region				
	Kidney cancer	Glomerular disease	Tubulointerstitial nephritis	Kidney failure	Total	Kidney cancer	Glomerular disease	Tubulointerstitial nephritis	Kidney failure	Total	Kidney cancer	Glomerular disease	Tubulointerstitial nephritis	Kidney failure	Total
2014	1.3	3.5	0.3	2.3	7.5	0.9	5.5	0.5	2.1	8.9	0.6	5.5	0.6	3.2	10.0
2015	1.5	3.4	0.3	2.3	7.5	1.0	5.9	0.5	1.8	9.2	0.8	6.2	0.6	3.0	10.5
2016	1.5	3.6	0.3	2.0	7.4	1.0	5.5	0.5	1.8	8.8	0.8	6.3	0.6	2.8	10.4
2017	1.5	3.6	0.3	2.2	7.5	1.0	6.0	0.5	1.8	9.2	0.8	6.7	0.5	2.7	10.7
2018	1.5	3.3	0.2	2.2	7.2	1.0	5.6	0.4	1.8	8.8	0.8	6.6	0.6	2.7	10.7
2019	1.5	3.1	0.2	2.2	7.0	1.1	5.7	0.4	1.6	8.8	0.8	6.0	0.6	2.6	9.9
2020	1.5	2.6	0.2	1.8	6.0	1.1	5.4	0.3	1.6	8.4	0.7	5.6	0.6	2.7	9.6
2021	1.5	2.6	0.2	2.0	6.2	1.0	5.2	0.3	1.7	8.2	0.7	5.4	0.5	2.8	9.4
AAPC (%)	1.2	−4.6	−7.9	−2.8	−3.0	2.1	−0.8	−8.9	−2.8	−1.2	0.7	−0.7	−1.4	−2.3	−1.1
95% *CI* lower	−0.6	−8.0	−9.5	−5.1	−5.9	0.5	−4.0	−14.9	−5.1	−2.8	−4.7	−2.9	−4.0	−4.3	−2.6
95% *CI* upper	3.0	−1.0	−6.3	−0.5	0.1	3.7	2.5	−2.5	−0.6	0.5	6.3	1.6	1.3	−0.2	0.5
P	0.193	0.013	<0.001	0.025	0.054	0.018	0.635	0.007	0.023	0.154	0.810	0.550	0.254	0.035	0.174

Abbreviation: AAPC=average annual percent change; *CI*=confidence intervals.

A study encompassing 195 countries reported that higher Socio-demographic Index (SDI) levels correlate with an increased incidence of kidney cancer burden. Notably, the ASMR for this cancer is higher in affluent areas, including Latin America and Eastern Europe, with significant rises observed in Eastern and Southern Asia (*5*). Further studies in the Middle East and North Africa have confirmed these findings, establishing a positive correlation between SDI and kidney cancer ASMR (*6*). These results partially explain the observed disparities in kidney cancer ASMR worldwide. In China, data shows an upward trend in kidney cancer ASMR within the central region from 2014 to 2021. Studies using Global Burden of Disease 2019 data have shown that this pattern is similar evident in Japan, The Republic of Korea, Singapore, and other countries (*7*). Many studies have linked risk factors such as smoking, obesity, hypertension, and elevated fasting blood glucose levels to higher rates of deaths associated with kidney cancer (*8*–*9*). With socioeconomic growth and improved living standards in China, the prevalence of these risk factors has risen annually, potentially contributing to the continual increase in kidney cancer ASMR (*10*). In alignment with the Healthy China 2030 initiative, it is crucial to strengthen tobacco control, raise awareness about the risks of tobacco use, and promote physical activity to effectively mitigate kidney cancer risk factors and reduce its burden.

The ASMR for non-cancerous kidney diseases was observed to be higher in rural and western areas. Chronic conditions such as cardiovascular disease, diabetes, and hypertension, along with prolonged use of anti-inflammatory medications, are identified as risk factors for mortality due to chronic kidney disease (*11*). Regular physical examinations to detect and manage these risk factors could markedly improve outcomes for these vulnerable populations. Additionally, the availability of medical care influences kidney disease mortality rates (*12*). Since a significant portion of individuals in underdeveloped areas often seek medical treatment across different regions (*13*), it is crucial to enhance primary healthcare services, optimize the allocation of resources related to kidney disease, and ensure better access to medical services for residents in rural and western regions.

This study was subject to some limitations. First, the data were sourced from the national cause-of-death surveillance system, which may suffer from underreporting, consequently leading to a possible underestimation of kidney disease mortality rates. To mitigate the impact of this underreporting and improve data reliability, this study excluded data from regions where the all-cause mortality rate was under 5‰ during the data compilation process. Additionally, there is an inherent delay in death surveillance data, with the most current data available only up to 2021. Despite these limitations, the data presented can still lay the groundwork for developing preventive strategies for kidney diseases. Future research should incorporate data from various sources to more accurately analyze the mortality rates and causes of death due to kidney diseases.

In conclusion, the trends identified in this study suggest that future preventive strategies should be region- and population-specific. Enhanced focus is essential on the health of elderly populations, as well as on improving health education for males, rural communities, and western regions. Furthermore, there is a need for strengthened tobacco control, efficient allocation of health resources, and measures to decrease the societal burden of kidney diseases.

Conflicts of interest: No conflicts of interest.

Funding: Supported by the National Key Research and Development Program of the Ministry of Science and Technology, China (2021YFC2500201).

doi: 10.46234/ccdcw2024.236

[#] Corresponding author: Lijun Wang, wanglijun@ncncd.chinacdc.cn.

[1] National Center for Chronic and Noncommunicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing, China.

Submitted: May 20, 2024; Accepted: July 01, 2024

REFERENCES

1. Luo PF, Yu H, Han RQ, Zhou JY, Wu M. Mortality and years of life lost caused by major kidney diseases in Jiangsu province, 2010-2018. Dis Surveill 2020;35(6):489 – 94. https://doi.org/10.3784/j.issn.1003-9961.2020.06.008.
2. GBD Chronic Kidney Disease Collaboration. Global, regional, and national burden of chronic kidney disease, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2020;395 (10225):709 – 33. https://doi.org/10.1016/S0140-6736(20)30045-3.
3. IHME. Global burden of disease study 2019 (GBD2019) data resources. https://vizhub.healthdata.org/gbd-results/.[2024-3-23].
4. Li Y, Ning YC, Shen B, Shi YQ, Song NN, Fang Y, et al. Temporal trends in prevalence and mortality for chronic kidney disease in China from 1990 to 2019: an analysis of the Global Burden of Disease Study 2019. Clin Kidney J 2023;16(2):312 – 21. https://doi.org/10.1093/ckj/sfac218.
5. Safiri S, Kolahi AA, Mansournia MA, Almasi-Hashiani A, Ashrafi-Asgarabad A, Sullman MJM, et al. The burden of kidney cancer and its attributable risk factors in 195 countries and territories, 1990-2017. Sci Rep 2020;10(1):13862. https://doi.org/10.1038/s41598-020-70840-2.
6. Safiri S, Hassanzadeh K, Ghaffari Jolfayi A, Mousavi SE, Motlagh Asghari K, Nejadghaderi SA, et al. Kidney cancer in the Middle East and North Africa region: a 30-year analysis (1990-2019). Sci Rep

2024;14(1):13710. https://doi.org/10.1038/s41598-024-64521-7.
7. Wang HX, Fan WL, Yang XY, Chen DY, Huang Q, Pan SY, et al. Trend analysis and forecast of kidney cancer burden in five Asian countries and regions. South China J Prev Med 2023;49(1):1 – 4,9. https://doi.org/10.12183/j.scjpm.2023.0001.
8. Zi H, He SH, Leng XY, Xu XF, Huang Q, Weng H, et al. Global, regional, and national burden of kidney, bladder, and prostate cancers and their attributable risk factors,1990-2019. Mil Med Res 2021;8(1): 60. https://doi.org/10.1186/s40779-021-00354-z.
9. Chow WH, Dong LM, Devesa SS. Epidemiology and risk factors for kidney cancer. Nat Rev Urol 2010;7(5):245 – 57. https://doi.org/10.1038/nrurol.2010.46.
10. Qin JM. Epidemic trends, problems and countermeasures of chronic diseases and related risk factors in China. Chin J Public Health 2014;30 (1):1-4. http://dx.doi.org/10.11847/zgggws2014-30-01-01. (In Chinese)
11. Ammirati AL. Chronic kidney disease. Rev Assoc Med Bras (1992) 2020;66Suppl 1(Suppl 1):s03-9. http://dx.doi.org/10.1590/1806-9282.66.S1.3.
12. Mokdad AH, Dwyer-Lindgren L, Bertozzi-Villa A, Stubbs RW, Morozoff C, Shirude S, et al. Trends and patterns of disparities in diabetes and chronic kidney disease mortality among US counties, 1980-2014. Popul Health Metr 2022;20(1):9. https://doi.org/10.1186/s12963-022-00285-4.
13. Wang FL, Yang C, Li PF, Wen FY, Zhang LX. Individual factors contributing to the intercity patient mobility for chronic kidney diseases in China. Chin J Blood Purif 2023;22(5):330 – 4. https://doi.org/10.3969/j.issn.1671-4091.2023.05.003.

Preplanned Studies

Emerging Sialylated Class R Lipooligosaccharides in *Campylobacter jejuni* from Seagulls Has the Potential to Trigger Guillain-Barré Syndrome — Yunnan Province, China, 2018–2023

Xiaoli Chen[1]; Yixin Gu[1]; Hairui Wang[1]; Guilan Zhou[1]; Xin Zhang[1]; Chang Liu[1]; Zhujun Shao[1]; Jianzhong Zhang[1]; Maojun Zhang[1,#]

Summary

What is already known about this topic?

Lipooligosaccharides from *Campylobacter jejuni* (*C. jejuni*) have a mimicry antigen structure with gangliosides, which explains the mechanism by which *C. jejuni* caused Guillain-Barré syndrome (GBS).

What is added by this report?

All 12 *C. jejuni* strains with class R LOSs and specific serotypes were isolated from seagulls in south China. These emerging *C. jejuni* strains had ganglioside-mimicry antigen structures and possessed a high potential for triggering GBS.

What are the implications for public health practice?

Sialylated lipooligosaccharides (LOS) class R with GBS-associated serotypes isolated from seagulls highlight the risk of induced GBS around coastal or lakeside areas.

Campylobacter jejuni lipooligosaccharides (LOS) possess a mimicry antigen structure with gangliosides, explaining the mechanism by which *C. jejuni* causes Guillain-Barré syndrome (GBS) (*1*). This study investigated 25 *C. jejuni* strains isolated from seagulls in southern China, 12 of which possessed class R LOSs. Notably, all 12 strains belonged to specific serotypes (HS2, HS4c, HS19, HS41) previously shown to be strongly associated with GBS (*2*). Further genetic dissection revealed that these 12 strains could be classified into 2 groups based on LOS sialylation: sialylated (8 strains) and non-sialylated (4 strains). Antigenicity analysis using immunized rabbit serum and serum from GBS patients demonstrated that antigen-antibody reactions occurred only in strains with sialylated LOSs, termed here as emerging *C. jejuni*. These results indicate that emerging *C. jejuni* strains, possessing both GBS-associated serotypes and sialylated class R LOSs, exhibit ganglioside mimicry antigen structures and a high potential for triggering GBS.

A total of 25 *C. jejuni* strains isolated from the feces of seagulls in Kunming City, Yunnan Province from 2018 to 2023, and 4 laboratory-stocked GBS-associated strains coded ICDCCJ07001 (HS:41, LOS class A, GD3-mimic), ICDCCJ07002 (HS:41, LOS class A, ganglioside mimic as GT1a, GM1, GD1a, GD1b, GT1b), HB-CJGB-QYT (HS:2, LOS class A, ganglioside mimic as GT1a, GM1, GD1a, GD1b, GT1b), and HB-CJGB-ZB (HS:19, LOS class A, ganglioside mimic as GD1a, GM1) were included in this study (Table 1). *Campylobacter* spp. isolation was carried out using the Campylobacter Isolation Kit and the membrane filter method (ZC-CAMPY-002, Qingdao Sinova Biotechnology Co., Ltd., Qingdao, China). Suspected colonies were picked and identified by Gram stain and biochemical tests. Polymerase chain reaction (Real-time PCR) tests were used to confirm and identify the species according to a previous study (*3*). The potential for ganglioside mimicry of these four strains was previously determined in Dr. Gilbert's lab via mass spectrometry analysis (data not shown). All studied strains were serotyped using a set of 25 commercial antisera for the serotyping of heat-stable (Penner) antigens of *C. jejuni* by the passive hemagglutination method. LOSs for immunization experiments were purified from strains ICDCCJ07001, ICDCCJ07002, HB-CJGB-QYT, and HB-CJGB-ZB by the hot aqueous-phenol method. New Zealand White rabbits (2.0 to 2.5 kg) were immunized with 400 μg of LOS in complete Freund's adjuvant. Blood was collected from the immunized rabbits until the serum titer reached 1∶100,000. Seven serum samples from GBS patients in a GBS outbreak in Jilin (northern China) were stocked in the laboratory (*4*). Ganglioside antibodies were

TABLE 1. *Campylobacter jejuni* strains used in this study.

Strain	LOS locus class	CPS genotypes	Penner serotype
ICDCCJ07001	A	HS41	HS41
ICDCCJ07002	A	HS41	HS41
HB-CJGB-QYT	A	HS2	HS2
HB-CJGB-ZB	A	HS19	HS19
KW_CHFB_005	R1	HS2	HS2
KW_CHFB_014	R1	HS2	HS2
KW_CHFB_028	R1	HS2	HS2
KW_CHFB_017	R2	HS4c	HS4c
KW_CHFB_022	R2	HS4c	HS4c
KW_CHFB_024	R2	HS4c	HS4c
KW_CHFB_029	R2	HS4c	HS4c
KW_DBFB_008	R1	HS19	HS19
KW_CHFB_006	R1	HS41	HS41
KW_CHFB_009	R1	HS41	HS41
KW_CHFB_011	R1	HS41	HS41
KW_CHFB_020	R1	HS41	HS41
KW_CHFB_008	W	HS21	HS21
KW_CHFB_026	W	HS23_36	HS23_36_53
KW_CHFB_027	C	HS4c	HS4c
KW_DBFB_007	F	HS44	HS44
KW_CHFB_023	K	HS44	HS44
KW_DBFB_009	D	HS5_31	HS5
KW_DBFB_004	M	HS40	UT
KW_CHFB_001	CDC13	HS8_17	HS8
KW_CHFB_007	CDC14	UT	UT
KW_CHFB_010	CDC16	UT	UT
KW_DBFB_005	CDC17	HS18	HS18
KW_CHFB_021	CDC18	HS8_17	HS8
KW_DBFB_002	CDC19	HS60	UT

Abbreviation: UT=untypeable; LOS=lipooligosaccharides; CPS=capsular polysaccharides.

previously measured, and antibodies against gangliosides GT1a and GM1 were positive. The antigenicity of the LOSs was analyzed using Western blotting.

A total of 13 LOS classes and 12 capsular genotypes were identified in 25 *C. jejuni* strains (*5*) (Table 1). GBS-associated capsular serotypes (HS2, HS4c, HS19, HS41, HS1_44, HS23_36) accounted for 64% (16/25) of strains. Besides, 12 of 25 (48%) strains displayed class R LOS. All 12 class R LOS strains possessed GBS-associated capsular genotypes, including HS2 (3/12, 25%), HS4 (4/12, 33.3%), HS19 (1/12, 8.3%), and HS41 (4/12, 33.3%). The phenotypic serotypes were consistent with the CPS genotypes listed in Table 1. The class R locus contains a large portion of the LOS class A locus; the first 9,700 nt showed 97% sequence identity to LOS class A, including *orf3r* (encoding a two-domain glucosyltransferase, Cj1135), *orf4r* (Cj1136), *orf5r* (CgtA), *orf6r* (CgtB), *orf7r* (CstII), *orf8r* (NeuB), *orf9r* (NeuC), and *orf10r* (NeuA). However, *orf14r* is specific to class R and distinguishes class A from class R (Figure 1A). The *orf5a1* of strain ATCC43446 encodes CgtA, whose acceptor is a sialylated galactosyl residue, while the CgtA of strain ATCC43438 utilizes a non-sialylated acceptor. These alternative alleles are also present in class R. Based on gene content and organization of the LOS biosynthesis loci from *C.*

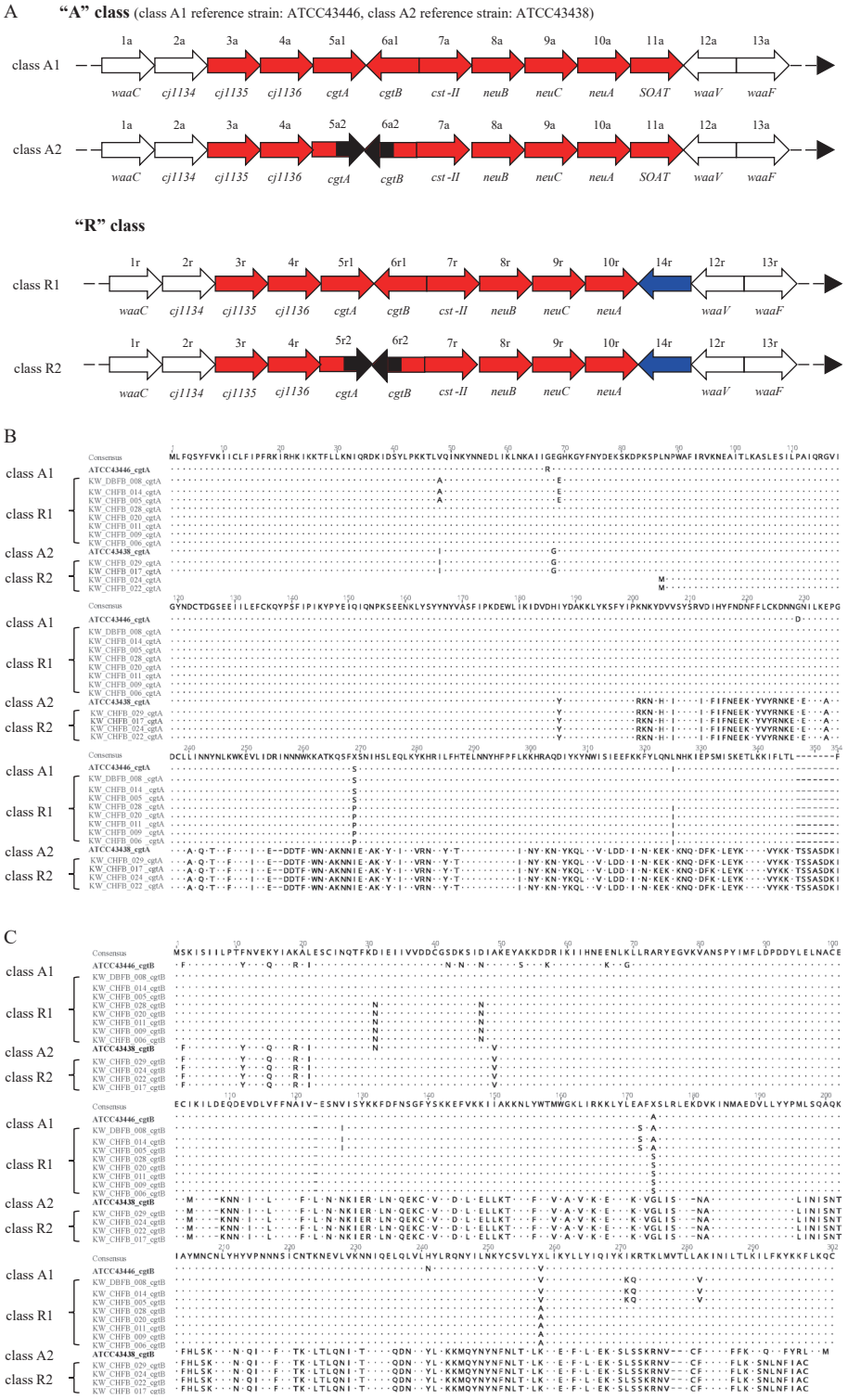


FIGURE 1. LOS class R related to class A. (A) Genetic organization of LOS biosynthesis locus of class A and R strains. (B) Alignment of the *cgtA* and (C) *cgtB* educed amino acid sequences.

Note: Four genes (*orf* 1, *orf* 2, *orf* 12, *orf* 13) were previously recognized as responsible for LOS inner core biosynthesis. Eight genes (*orf* 3 to *orf* 10, shown in red) have a high identity to LOS class A. There is a gene (*orf* 14r, shown in blue), unique to class R. Alignment of the *cgtA* (*orf* 5) (1B) and *cgtB* (*orf* 6) (1C) deduced amino acid sequences showed that *cgtA* (*orf* 5) and *cgtB* (*orf* 6) of LOS class R strains use either sialylated (99% identity with class A1 strain ATCC43446) or non-sialylated (99% identity with class A2 strain ATCC43438) acceptors, and class R strains were classified into class R1 and class R2.

Abbreviation: LOS=lipooligosaccharides.

jejuni class R strains, the loci were grouped into two major groups: group R1 strains with sialylated LOS (8 strains) and group R2 with non-sialylated LOS (4 strains). Amino acid sequences of CgtA and CgtB are nearly identical (98% identity) between ATCC 43446 and class R1 strains and 99% identical between ATCC 43438 and class R2 strains (Figure 1B and 1C).

Based on the genetic characteristics and previous investigations, the mimic antigen structures of GBS-associated *C. jejuni* strains ICDCCJ07002, HB-CJGB-QYT, HB-CJGB-ZB, and ICDCCJ07001 were demonstrated in Figure 2A. LOSs from six strains from two groups of class R (group R1 and group R2) were purified and verified with silver staining for PAGE (Figure 2B). The reactions between different class R LOS antigens from different strains and the serum of rabbits immunized with GBS-associated strains were presented in Figure 2C. Serum immunized with LOS of strain ICDCCJ07001 did not react with any of the purified class R LOSs (Figure 2C). LOSs from three class R1 strains showed a strong reaction with the serum from seven antiganglioside-positive GBS patients. None of the R2 LOSs reacted with the patients' serum. Sera of healthy controls did not react with any of the class R LOSs tested.

DISCUSSION

Molecular mimicry between *C. jejuni* LOS and host gangliosides leads to the production of cross-reactive antibodies directed against the peripheral nerves of the host to trigger GBS (*6*–*7*). Genes involved in the synthesis of sialylated LOS and ganglioside mimics are present in the LOS classes A, B, C, M, and R (*8*). Numerous reports indicate that *C. jejuni* strains with classes A, B, and C LOS are frequently isolated from the stools of patients with GBS, and classes A and B are associated with GBS and MFS, respectively. Penner serotyping showed that the combination of LOS loci that contain genes involved in LOS sialylation with the capsular types HS1/44c, HS2, HS4c, HS19, HS23/36c, and HS41 are associated with the development of GBS and MFS and are therefore markers for these neurological diseases (*2*). In the present study, 16 (64%, 16/25) of these seagulls isolates had GBS-associated serotypes, and 75% (12/16) of them displayed class R LOS. These results raise significant concern about the risk of these strains inducing GBS around coastal or lakeside areas.

Further pathogenicity analysis revealed that class R LOS strains could be classified into two groups: strains with sialylated LOS (group R1) and strains with non-sialylated LOS (group R2). Group R2 strains showed no reactivity with GBS patient sera and failed to induce anti-LOS antibody responses in immune serum from animals immunized with LOS from ganglioside-mimic strains. These results confirmed this study's hypothesis that emerging *C. jejuni* strains isolated from seagulls — which possess specific serotypes (HS2, HS4c, HS19, HS41) and sialylated class R LOS — have ganglioside-mimicking antigen structures and a high potential risk for triggering GBS.

C. jejuni requires specific gene combinations to express ganglioside mimics. An important feature in ganglioside mimicry is the presence of sialic acid (N-acetylneuraminic acid, NANA) in both LOS and gangliosides. Sialic acid is a rare constituent of bacterial surface polysaccharides, and sialylation of bacterial proteins has thus far been found only in *Campylobacter* species (*9*). Sialylated LOS is also involved in other aspects of *C. jejuni* pathogenesis. *C. jejuni* strains expressing sialylated LOS invade human intestinal epithelial cells significantly more frequently than strains expressing non-sialylated LOS (*10*). Sera from seven GBS patients in this study were previously identified as being infected with *C. jejuni* strains bearing a GT1a-like LOS (*7*). Moreover, antigenicity reactions only occurred in class R1 LOS strains with four immunized sera, which were raised against LOS from the GBS-associated *C. jejuni* strains. This indicates the presence of additional epitopes in class R1 LOS that mimic the ganglioside structures in GT1a or GM1, GD1a, GD1b, and GT1b.

In this study, only the pathogenic characteristics of the *C. jejuni* strains were investigated. There was no data about the GBS patients. In the future, enhanced surveillance for both the *C. jejuni* infection and the GBS patients in this region will be needed for the prevention and control of GBS caused by *C. jejuni* infection.

In conclusion, this study indicated that these emerging *C. jejuni* strains, which had both GBS-associated serotypes and sialylated class R LOSs, had gangliosides mimicry antigen structures and possessed high potential for triggering GBS. The results highlight the risk of *C. jejuni*-induced GBS around coastal or lakeside areas which reminds us that we should strengthen the surveillance and prevention of *C. jejuni* infection in these areas.

Conflicts of interest: No conflicts of interest.

Funding: Supported by National Key Research and Development Program of China (2021YFC2301000),

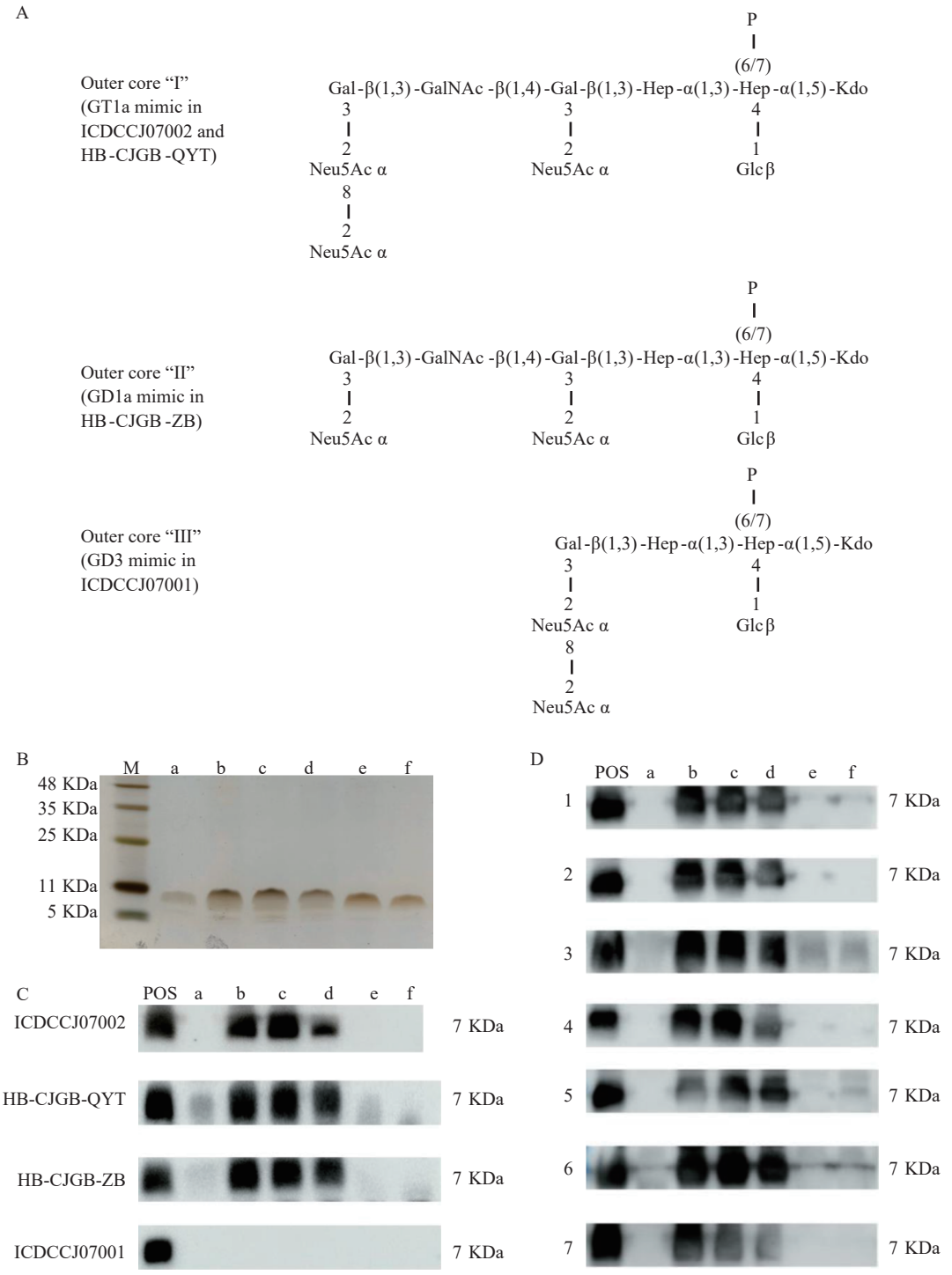


FIGURE 2. Reaction of class R LOS with sera of either immunized rabbit or GBS patient. (A) Structures of various *C. jejuni* LOS used for immunization of rabbits and corresponding ganglioside mimics. (B) Silver staining of the purified class R LOS (three strains from group R1 LOS and group R2 LOS were selected in this figure, respectively). (C) Western blot analysis of class R LOS with sera from LOS-immunized rabbits. (D) Western blot analysis of class R LOS with GBS patient serum.
Note: M, marker. POS, positive control, purified LOS from different strains to immunize rabbits. a LOS from strain KW_CHFB_024 (group R2); b LOS from strain KW_CHFB_009 (group R1); c LOS from strain KW_CHFB_011 (group R1); d LOS from strain KW_CHFB_014 (group R1); e LOS from strain KW_CHFB_017 (group R2); f LOS from strain KW_CHFB_022 (group R2).
Abbreviation: LOS=lipooligosaccharides.

Research on detection and tracing of important pathogen combinations, identification and prevention of drug resistance hazards (SZSM201803081), and Project for Gonorrhea vaccine development (2023110024000074).

doi: 10.46234/ccdcw2024.237

[#] Corresponding author: Maojun Zhang, zhangmaojun@icdc.cn.

[1] National Key Laboratory of Intelligent Tracking and Forecasting for Infectious Diseases, National Institute for Communicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing, China.

Submitted: March 24, 2024; Accepted: July 11, 2024

REFERENCES

1. Hao YL, Wang WF, Jacobs BC, Qiao BJ, Chen MS, Liu DQ, et al. Antecedent infections in Guillain-Barré syndrome: a single-center, prospective study. Ann Clin Transl Neurol 2019;6(12):2510 – 7. https://doi.org/10.1002/acn3.50946.
2. Heikema AP, Islam Z, Horst-Kreft D, Huizinga R, Jacobs BC, Wagenaar JA, et al. *Campylobacter jejuni* capsular genotypes are related to Guillain-Barrésyndrome. Clin Microbiol Infect 2015;21(9):852.e1 – 9. https://doi.org/10.1016/j.cmi.2015.05.031.
3. Liang H, Wen ZY, Li Y, Duan YX, Gu YX, Zhang MJ. Comparison of the filtration culture and multiple real-time PCR examination for *Campylobacter* spp. From stool specimens in diarrheal patients. Front Microbiol 2018;9:2995. https://doi.org/10.3389/fmicb.2018.02995.
4. Zhang MJ, Li Q, He LH, Meng FL, Gu YX, Zheng MH, et al. Association study between an outbreak of Guillain-Barre syndrome in Jilin, China, and preceding *Campylobacter jejuni* infection. Foodborne Pathog Dis 2010;7(8):913 – 9. https://doi.org/10.1089/fpd.2009.0493.
5. Wang JQ, Chen XL, Zhou GL, Wang HR, Gu YX, Zhang JZ, et al. Genetic characteristics of lipooligosaccharide and capsular polysaccharide of *Campylobacter jejuni* from different sources in China. Biomed Environ Sci 2022;35(12):1106 – 14. https://doi.org/10.3967/bes2022.140.
6. Rodríguez Y, Rojas M, Pacheco Y, Acosta-Ampudia Y, Ramírez-Santana C, Monsalve DM, et al. Guillain-Barré syndrome, transverse myelitis and infectious diseases. Cell Mol Immunol 2018;15(6):547 – 62. https://doi.org/10.1038/cmi.2017.142.
7. Zhang MJ, Gilbert M, Yuki N, Cao FF, Li JJ, Liu HY, et al. Association of anti-GT1a antibodies with an outbreak of Guillain-Barré syndrome and analysis of ganglioside mimicry in an associated *Campylobacter jejuni* strain. PLoS One 2015;10(7):e0131730. https://doi.org/10.1371/journal.pone.0131730.
8. Parker CT, Gilbert M, Yuki N, Endtz HP, Mandrell RE. Characterization of lipooligosaccharide-biosynthetic loci of *Campylobacter jejuni* reveals new lipooligosaccharide classes: evidence of mosaic organizations. J Bacteriol 2008;190(16):5681 – 9. https://doi.org/10.1128/JB.00254-08.
9. Heikema AP, Bergman MP, Richards H, Crocker PR, Gilbert M, Samsom JN, et al. Characterization of the specific interaction between sialoadhesin and sialylated *Campylobacter jejuni* lipooligosaccharides. Infect Immun 2010;78(7):3237 – 46. https://doi.org/10.1128/IAI.01273-09.
10. Kuijf ML, Samsom JN, van Rijs W, Bax M, Huizinga R, Heikema AP, et al. TLR4-mediated sensing of *Campylobacter jejuni* by dendritic cells is determined by sialylation. J Immunol 2010;185(1):748 – 55. https://doi.org/10.4049/jimmunol.0903014.

Preplanned Studies

Longitudinal Study on Puberty Onset and Adolescent Obesity — Suzhou City, Jiangsu Province, China, 2012–2022

Huiling Lyu[1]; Ziyao Ding[2]; Bowen Chen[3]; Di Han[2]; Xi Wang[1]; Xiao Jiang[2]; Jiale Hu[1]; Hui Shen[1,2,4,#]; Jia Hu[1,2,4,#]

Summary

What is already known about this topic?

The early onset of puberty increases the risk of cardiovascular diseases. Limited research has focused on the association between the onset of puberty, based on sudden increases in height, and obesity in late adolescence.

What is added by this report?

This report assessed the age at take-off and age at peak height velocity for children and adolescents in China from 2012 to 2020. The results indicated that age at take-off and age at peak height velocity were negatively associated with the risk of late adolescence obesity in boys, independent of childhood body mass index; however, no similar association was found in girls.

What are the implications for public health practice?

Puberty timing influences adiposity. Effective monitoring and management of pubertal development are necessary, and attention should be paid to sex differences.

The onset of puberty is an important biological event in human growth and development. In recent years, a trend toward earlier pubertal onset has been observed among children and adolescents in China (*1*). Early pubertal timing increases the risk of cardiovascular diseases, such as obesity and hypertension in adulthood; however, research focusing on late adolescence is relatively limited (*2*). Meanwhile, the evaluation indicators for pubertal timing are mostly based on secondary sexual characteristics, with limited reports concerning age at take-off (ATO) and age at peak height velocity (APHV), which reflect characteristics of sudden increases in height (*1*). Therefore, using data from the Health Promotion Program for Children and Adolescents (HPPCA) in Suzhou, China (*3*), we conducted further analyses using ATO and APHV to evaluate the association between pubertal onset and adolescent obesity. Ultimately, 9,269 adolescents aged 15–17 years from 20 schools were enrolled, including 67,383 historical physical examination data from 2012 to 2020. In boys, early pubertal timing, as determined by ATO and APHV, was negatively associated with endpoint body mass index (BMI) and risk of obesity in late adolescence, independent of childhood BMI. APHV was inversely associated with endpoint BMI among girls. Our findings indicate that ATO and APHV can serve as early risk indicators of adiposity in adolescents. Thus, effective monitoring and management of pubertal development are essential, and preventive interventions should prioritize the reduction of childhood adiposity.

This study was a retrospective cohort analysis based on a large-scale, ongoing, school-based monitoring program from 2012 to 2020 in Suzhou City, Jiangsu Province, China. Detailed information about HPPCA has been previously published (*3*). In 2020, 19,824 students aged 15–17 years were selected from 20 schools. Height and weight were measured, and BMI was calculated by dividing weight (kg) by the square of height (m^2). Additionally, 9,760 participants were excluded for the following reasons: fewer than four height measurements between 2012 and 2019; no height measurement at 6–8 years old; or fewer than three height measurements at 9–14 years old. Another 795 participants were excluded because of invalid values due to logical inconsistencies in fitting ATO and APHV. Ultimately, 9,269 adolescents, contributing 67,383 physical examination data points, were included in the study.

The Preece & Baines growth model, a set of nonlinear regression models proposed by Preece and Baines in 1978 to fit individual growth curves, was used to fit the ATO and APHV for each subject (*4*). By analyzing velocity (first derivative) and acceleration (second derivative), the model estimates the onset rate of sudden height increase, PHV, ATO, and APHV of individuals (*4*). Puberty timing onset was determined as early (<P25) or non-early (≥P25) according to the 25th percentile (P25) of ATO (boys: 8.19 years, girls: 6.49 years) and APHV (boys: 11.85 years, girls: 9.83

years), respectively. Obesity was defined by the age- and sex-specific BMI cutoffs [Screening for Overweight and Obesity Among School-Age Children and Adolescents (WS/T 586-2018)].

Categorical variables are presented as numbers and percentages, while continuous variables are reported as mean±standard deviation. In this study, *t*-tests and chi-square tests were used to compare differences between groups. We established initial participation in the HPPCA at 6–8 years of age as the baseline during 2012–2014 and designated the final measurement at 15–17 years of age (late adolescence) in 2020 as the endpoint. Linear and logistic regression models were used to analyze the associations of ATO and APHV with BMI levels and risks of obesity in late adolescence. We also analyzed the risks of elevated BMI levels and obesity in late adolescence for participants in the early group compared to the non-early group. Coefficients (β), odds ratios (*OR*s), and 95% confidence intervals (*CI*s) for the unadjusted model (Model 1) and the adjusted model (Model 2, adjusted for sex, area, age, and baseline BMI) were reported. The fitting of ATO and APHV was performed in Stata Statistical Software (version 16, Stata Corp., TX, USA) using the "pbreg" package. SPSS Statistics (version 22, IBM SPSS Inc., Chicago, IL, USA) was used for other analyses. A *P*-value of less than 0.05 was considered statistically significant.

The final analysis included 9,269 adolescents aged 15 to 17 years, comprising 4,857 boys (52.40%) and 4,412 girls (47.60%), with 5,212 urban (56.23%) and 4,057 rural (43.77%) participants. The ages at baseline and endpoint were 8.16±0.51 years and 15.96±0.50 years, respectively. The overall prevalence of obesity at baseline and endpoint was 14.79% and 10.18%, respectively, with a higher prevalence of obesity found in boys than in girls ($P<0.05$). The average number of follow-ups was 7.27±0.78. The ATO for girls was 7.93±2.19 years, occurring 1.24 years earlier than for boys ($P<0.05$). The APHV for girls was 10.75±1.22 years, which was 1.67 years earlier than for boys ($P<0.05$). Detailed general characteristics of the included children and adolescents are presented in Table 1. Additionally, the main characteristics of the included and excluded participants are outlined in Supplementary Table S1 (available at https://weekly.chinacdc.cn/).

Table 2 illustrates the associations between the onset of puberty timing and BMI and obesity in late adolescence. After adjusting for potential confounders, both ATO and APHV exhibited negative associations with endpoint BMI in late adolescence (β=−0.04, 95% *CI*: −0.07, −0.02; β=−0.17, 95% *CI*: −0.22, −0.12). ATO was negatively associated with obesity risk in late adolescence (*OR*=0.96, 95% *CI*: 0.93, 1.00). When stratified by gender, the association between ATO and endpoint BMI was only significant among boys. Similarly, ATO and APHV showed significant associations with obesity risks among boys but not girls.

Compared with participants in the non-early group, those in the early group with earlier ATO and APHV exhibited higher BMI levels, with a β of 0.13 (95% *CI*: 0.00, 0.25) for ATO and 0.28 (95% *CI*: 0.16, 0.41) for APHV. For boys, these associations appeared stronger, with a β of 0.22 (95% *CI*: 0.03, 0.40) for ATO and 0.48 (95% *CI*: 0.29, 0.67) for APHV; however, no analogous patterns were observed among girls ($P>0.05$). Boys, but not girls, in the early-onset group identified by ATO and APHV had increased obesity risks than their counterparts in the non-early group, with an *OR* of 1.25 (95% *CI*: 1.00, 1.56) for ATO and 1.26 (95% *CI*: 1.02, 1.55) for APHV. Figure 1 includes more detailed information.

DISCUSSION

This study used the Preece & Baines growth model to fit height data from 9,269 children and adolescents collected from 2012 to 2020. This model facilitated the construction of a height development curve and calculation of ATO and APHV. The results showed that the average ATO for boys and girls in Suzhou was 9.17 years and 7.93 years, respectively, while the average APHV was 12.42 years and 10.75 years, respectively. Notably, sex differences were observed in the associations between early pubertal timing and late adolescence obesity. In boys, both ATO and APHV were negatively correlated with the risk of obesity in late adolescence, independent of childhood BMI. Conversely, no similar associations were found in girls.

A study including 28,251 adolescents in Changzhou City, Jiangsu Province, China, reported that ATO was 9.3 years for boys and 8.0 years for girls, while APHV was 12.6 years for boys and 10.6 years for girls (*5*). Chen et al. analyzed 13,143 children in a longitudinal cohort from 2006 to 2016 in Zhongshan City, Guangdong Province, and found that ATO was 7.9 and 9.1 years in girls and boys, respectively, and APHV was 10.8 years in girls, 1.9 years earlier than in boys (*6*). The findings of this study align with other regional data within China. Nevertheless, they show

TABLE 1. General characteristics of enrolled children and adolescents [*n* (%)].

Variable	Total (*n*=9,269)	Boys (*n*=4,857)	Girls (*n*=4,412)	t/χ^2-value	*P*
Endpoint					
Age (years)*	15.96±0.50	15.95±0.50	15.96±0.50	−0.707	0.480
Height (cm)*	169.19±8.15	174.80±5.91	163.02±5.31	101.085	<0.001
Weight (kg)*	61.98±13.33	67.85±13.86	55.52±9.08	51.098	<0.001
BMI (kg/m^2)*	21.56±3.79	22.17±4.18	20.88±3.19	16.773	<0.001
Area				0.292	0.589
Uran	5,212 (56.23)	2,744 (56.50)	2,468 (55.94)		
Rural	4,057 (43.77)	2,113 (43.50)	1,944 (44.06)		
BMI status				389.973	<0.001
Thinness	595 (6.42)	346 (7.12)	249 (5.65)		
Normal	6,182 (66.70)	2,816 (57.98)	3,366 (76.29)		
Overweight	1,548 (16.70)	996 (20.51)	552 (12.51)		
Obesity	944 (10.18)	699 (14.39)	245 (5.55)		
Numbers of follow−up*	7.27±0.78	7.27±0.79	7.26±0.78	0.559	0.576
ATO (years)*	8.58±2.24	9.17±2.13	7.93±2.19	27.582	<0.001
APHV (years)*	11.62±1.42	12.42±1.08	10.75±1.22	69.477	<0.001
Baseline					
Age (years)*	8.16±0.51	8.16±0.51	8.16±0.50	−0.137	0.891
Height (cm)*	130.32±5.98	130.98±5.95	129.59±5.94	11.261	<0.001
Weight (kg)*	28.84±5.96	29.99±6.33	27.58±5.24	20.019	<0.001
BMI (kg/m^2)*	16.87±2.55	17.36±2.71	16.33±2.25	19.924	<0.001

Abbreviation: BMI=body mass index; ATO=age at take-off; APHV=age at peak height velocity.
* Continuous variables were presented as mean ± standard deviation.

TABLE 2. Associations of the onset of pubertal timing with BMI and obesity in late adolescence.

Group	BMI [β (95% *CI*)]		Obesity [*OR* (95% *CI*)]	
	Model 1	Model 2	Model 1	Model 2
ATO (age)				
Total	0.02 (−0.02, 0.05)	−0.04 (−0.07, −0.02)	1.02 (0.99, 1.05)	0.96 (0.93, 1.00)
Boys	−0.17 (−0.23, −0.12)	−0.08 (−0.12, −0.04)	0.93 (0.90, 0.96)	0.95 (0.90, 0.99)
Girls	0.04 (0.00, 0.08)	0.00 (−0.03, 0.03)	1.03 (0.97, 1.10)	0.99 (0.92, 1.06)
APHV (age)				
Total	−0.15 (−0.20, −0.09)	−0.17 (−0.22, −0.12)	1.00 (0.95, 1.05)	0.94 (0.87, 1.01)
Boys	−0.97 (−1.08, −0.87)	−0.29 (−0.37, −0.21)	0.69 (0.64, 0.74)	0.90 (0.82, 0.99)
Girls	−0.33 (−0.41, −0.26)	−0.07 (−0.12, −0.01)	0.83 (0.75, 0.92)	0.99 (0.88, 1.11)

Note: Model 1: unadjusted, Model 2: adjusted for gender, area, age, and BMI at baseline.
Abbreviation: BMI=body mass index; β=coefficients; *OR*=odds ratio; *CI*=confidence interval; ATO=age at take-off; APHV=age at peak height velocity.

some discrepancies compared with data from other countries. A cohort study of 7,495 US children investigated early-life growth patterns and the age at pubertal onset, revealing APHV values of 12.9 years for boys and 10.8 years for girls (*7*). Additionally, a study aimed at estimating the height growth curve for Mexican children, involving 7,097 boys and 6,167 girls, found that ATO and APHV were 8.6 and 12.4 years for boys, respectively, and 7.0 and 9.9 years for girls, respectively (*8*). The observed variations are potentially due to differences in the timing of selection, ethnicity, environmental factors, and socioeconomic status of the participants.

Research on the associations of pubertal timing with

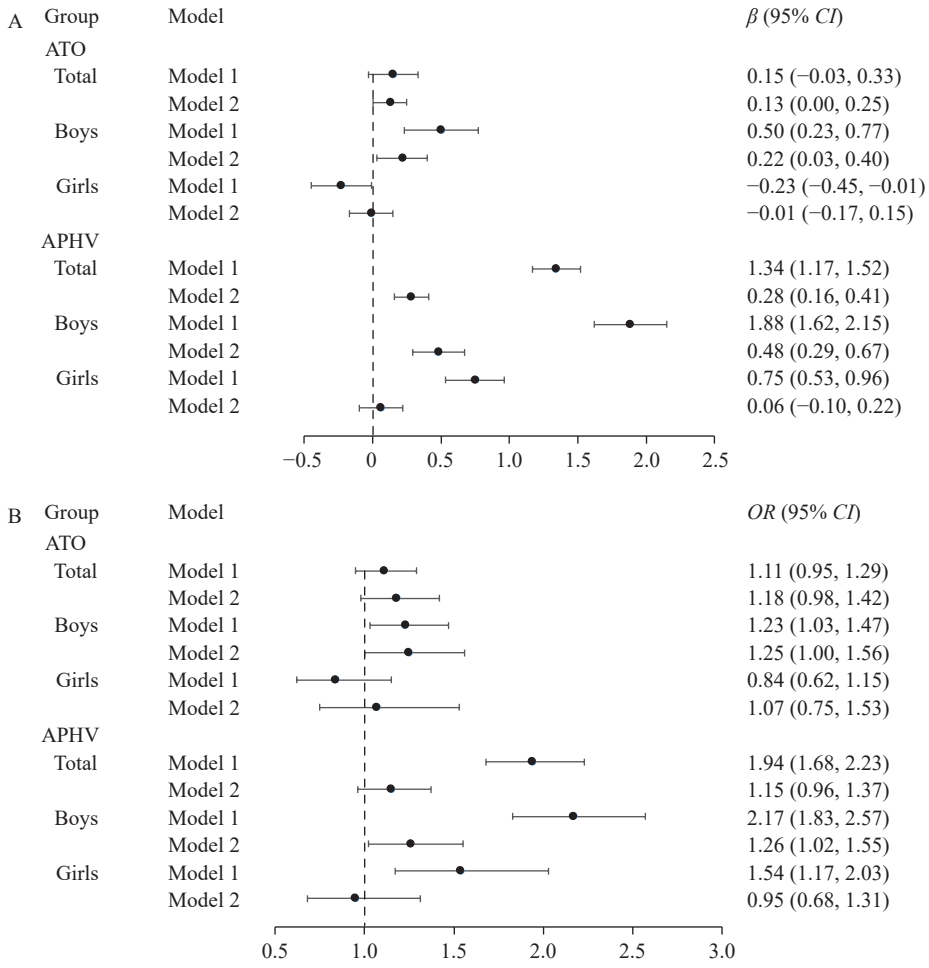


FIGURE 1. The risks of elevated BMI levels and obesity for participants in the early onset of pubertal timing group compared to those in the non-early group. (A) BMI levels. (B) Obesity.
Note: Model 1: unadjusted. Model 2: adjusted for gender, area, age, and BMI at baseline.
Abbreviation: β=coefficients; *OR*=odds ratio; *CI*=confidence interval; ATO=age at take-off; APHV=age at peak height velocity; BMI=body mass index.

obesity, based on ATO and APHV, is limited, especially in late adolescence. Limited research indicates that advanced APHV is associated with increased BMI levels in late puberty and adulthood. A cohort study by Chen et al. (*6*) showed that early APHV (*RR*=1.16, 95% *CI*: 1.03, 1.30) was associated with an increased risk of overweight and obesity in late adolescence. While this study found a negative correlation between girls' APHV and BMI in late adolescence, there was no statistically significant correlation of ATO and APHV with the risk of obesity. This finding aligns with the results of O'Keeffe et al. (*9*); however, Chen et al. (*6*) reported an increased risk of obesity associated with APHV in girls, indicating inconsistencies across studies. These discrepancies may stem from various factors, including different growth modeling methods, such as the SITAR model, Preece-Baines growth model, and growth curves. Even when employing the same model, variations in sample size, number of measurements, data balance, and measurement errors can lead to fitting deviations. Furthermore, the continuous nature of nutritional status throughout different life stages and the link between childhood obesity and the premature onset of puberty must be considered. Failing to account for pre-pubertal nutritional status could lead to misinterpretations of the associations between ATO, APHV, and obesity during late puberty. Thus, further research is necessary to elucidate the links between ATO, APHV, and obesity in late adolescence and adulthood. Additionally, this study found gender-specific associations between pubertal onset and adolescent obesity. Gender differences in this association may be attributed to sexual dimorphism in fat distribution and metabolism, coupled with a robust synergistic relationship between estrogen and leptin in

the regulation of reproductive and energy homeostasis (*6*).

Based on the Chinese National Survey on Students Constitution and Health, the prevalence of obesity in 2019 was 9.6% among children and adolescents aged 7–18 years, with 11.9% in children aged 7–8 years and 6.7% in adolescents aged 15–17 years in China. The prevalence of obesity among participants in this study was higher than the national average, consistent with trends in eastern China (*10*). This emphasizes that obesity is a major public health concern, particularly in coastal and economically developed areas. Given these findings, when boys present with early rapid height growth, healthcare providers should prioritize their nutritional status and implement health education and lifestyle interventions, including dietary modifications, exercise, and behavioral changes.

This study has several limitations. First, because participants were exclusively from Suzhou, China, caution is warranted when extrapolating these findings to other populations. Differences in ethnicity, living environments, and socioeconomic levels may influence the results. Second, the research design did not account for potentially significant confounding factors, including genetic influences; pregnancy and childbirth information; early-life exposures; and lifestyle factors such as diet, exercise behaviors, and sleep patterns. These omissions could affect the study's conclusions regarding the associations between the timing of puberty onset and obesity.

Conflicts of interest: No conflicts of interest.

Acknowledgments: All the children, parents, and workers who were involved in this study.

Funding: Supported by Gusu Health Talents Program Training Project in Suzhou, China (GSWS2020100) and the National Natural Science Foundation of China (82204070).

doi: 10.46234/ccdcw2024.239

[#] Corresponding authors: Hui Shen, shenhuicdc123@163.com; Jia Hu, hujia200606@163.com.

[1] School of Public Health, Nanjing Medical University, Nanjing City, Jiangsu Province, China; [2] Department of School Health, Suzhou Center for Disease Prevention and Control, Suzhou City, Jiangsu Province, China; [3] Department of Computer and Information Science, University of Pennsylvania, Philadelphia, PA, USA; [4] Suzhou Institute of Advanced Study in Public Health, Gusu School, Nanjing Medical University, Suzhou City, Jiangsu Province, China.

Submitted: May 06, 2024; Accepted: September 03, 2024

REFERENCES

1. Wang SS, Sun Y, Tao FB. Advances in evaluation methods of pubertal timing in children. Chin J Sch Health 2023;44(6):951 – 5,960. https://doi.org/10.16835/j.cnki.1000-9817.2023.06.035.
2. Reinehr T, Roth CL. Is there a causal relationship between obesity and puberty? Lancet Child Adolesc Health 2019;3(1):44-54. http://dx.doi.org/10.1016/S2352-4642(18)30306-7.
3. Ge WX, Han D, Ding ZY, Yi LP, Yang ZQ, Wang XN, et al. Pediatric body mass index trajectories and the risk of hypertension among adolescents in China: a retrospective cohort study. World J Pediatr 2023;19(1):76 – 86. https://doi.org/10.1007/s12519-022-00626-1.
4. Preece MA, Baines MJ. A new family of mathematical models describing the human growth curve. Ann Hum Biol 1978;5(1):1 – 24. https://doi.org/10.1080/03014467800002601.
5. Mao SH, Li HB, Jiang J, Sun X, Cheng JCY, Qiu Y. An updated analysis of pubertal linear growth characteristics and age at menarche in Ethnic Chinese. Am J Hum Biol 2011;23(1):132 – 7. https://doi.org/10.1002/ajhb.21116.
6. Chen L, Su BB, Zhang Y, Ma T, Liu JY, Yang ZG, et al. Association between height growth patterns in puberty and stature in late adolescence: a longitudinal analysis in Chinese Children and adolescents from 2006 to 2016. Front Endocrinol (Lausanne) 2022;13:882840. https://doi.org/10.3389/fendo.2022.882840.
7. Aris IM, Perng W, Dabelea D, Ganiban JM, Liu C, Marceau K, et al. Analysis of early-life growth and age at pubertal onset in US children. JAMA Netw Open 2022;5(2):e2146873. https://doi.org/10.1001/jamanetworkopen.2021.46873.
8. Flores LA, Datta Banik S, Cameron N, Fragoso IJ. Growth in height and its association with overweight and obesity in Mexican children: an evaluation based on a nationally representative sample (ENSANUT 2018). Front Public Health 2024;12:1339195. https://doi.org/10.3389/fpubh.2024.1339195.
9. O'Keeffe LM, Frysz M, Bell JA, Howe LD, Fraser A. Puberty timing and adiposity change across childhood and adolescence: disentangling cause and consequence. Hum Reprod 2020;35(12):2784 – 92. https://doi.org/10.1093/humrep/deaa213.
10. Dong YH, Chen L, Liu JY, Ma T, Zhang Y, Chen MM, et al. Epidemiology and prediction of overweight and obesity among children and adolescents aged 7-18 years in China from 1985 to 2019. Chin J Prev Med 2023;57(4):461 – 9. https://doi.org/10.3760/cma.j.cn112150-20220906-00881.

SUPPLEMENTARY MATERIAL

SUPPLEMENTARY TABLE S1. Main characteristics of included and excluded participants at endpoint [*n* (%)].

Variable	Total (*n*=19,824)	Included (*n*=9,269)	Excluded (*n*=10,555)	t/χ^2-value	*P*
Age (years)*	16.16±0.58	15.96±0.50	16.34±0.58	−50.161	<0.001
Height (cm)*	169.05±8.20	169.19±8.15	168.92±8.24	2.300	0.021
Weight (kg)*	62.12±13.46	61.98±13.33	62.23±13.57	−1.283	0.199
BMI (kg/m^2)*	21.64±3.83	21.56±3.79	21.71±3.86	−2.742	0.006
Gender				15.203	<0.001
Boys	10,095 (50.92)	4,857 (52.40)	5,238 (49.63)		
Girls	9,729 (49.08)	4,412 (47.60)	5,317 (50.37)		
Area				0.001	0.969
Uran	11,150 (56.24)	5,212 (56.23)	5,938 (56.26)		
Rural	8,674 (43.76)	4,057 (43.77)	4,617 (43.74)		
BMI status				4.407	0.221
Thinness	1,320 (6.66)	595 (6.42)	725 (6.87)		
Normal	13,212 (66.65)	6,182 (66.70)	7,030 (66.61)		
Overweight	3,348 (16.89)	1,548 (16.70)	1,800 (17.05)		
Obesity	1,944 (9.80)	944 (10.18)	1,000 (9.47)		

Abbreviation: BMI=Body mass index.

* Continuous variables were presented as mean±standard deviation.

Methods and Applications

AI-Based Hematological Age Predictors and the Association Between Biological Age Acceleration and Type 2 Diabetes Mellitus — Chongqing Municipality, China, 2015–2021

Zhe Yin[1,2,&]; Yingnan Song[3,&]; Junhui Zhang[4,&]; Qiaoyun Dai[1,2]; Xinyuan Zhang[1,2]; Xueying Yang[1,2]; Na Nie[4]; Cuixia Chen[1,2]; Zongfu Cao[1,#]; Xu Ma[2,3,#]

ABSTRACT

Introduction: Biological age (BA) can represent the actual state of human aging more accurately than chronological age (CA).

Methods: Using hematological data from 112,925 participants in southwestern China, collected between 2015 and 2021, this study constructed BA predictors using 7 machine learning (ML) methods (tailored separately for male and female populations). This study then analyzed the association between BA acceleration and type 2 diabetes mellitus (T2DM) within this data using logistic regression. Additionally, it examined the impact of glycemic control on BA in individuals with diabetes.

Results: Among all ML models, deep neural networks (DNN) delivered the best performance in male [mean absolute error (MAE)=6.89, r=0.75] and female subsets (MAE=6.86, r=0.74). BA acceleration showed positive correlations with T2DM in both male [odds ratio (*OR*): 2.22, 95% confidence interval (*CI*): 1.77–2.77] and female subsets (*OR*: 3.10, 95% *CI*: 2.16–4.46), while BA deceleration showed negative correlations in both male (*OR*: 0.32, 95% *CI*: 0.27–0.39) and female subsets (*OR*: 0.42, 95% *CI*: 0.33–0.53). Individuals with diabetes with normal fasting glucose had significantly lower BAs than those with impaired fasting glucose in all CA groups except for patients older than 80.

Discussion: Artificial intelligence (AI)-based hematological BA predictors show promise as advanced tools for assessing aging in epidemiological studies. Implementing AI-based BA predictors in public health initiatives could facilitate proactive aging management and disease prevention.

Aging is a process characterized by the gradual decline of various molecular functions and the progressive accumulation of senescent cells, increasing the risk of diseases such as neurodegenerative diseases (*1*), cerebral ischemia (*2*), cancer (*3*), and vasculitis (*4*). Research has evolved the definition of aging from a simple increase in chronological age to a systematic assessment of overall physiological function. Different bodily systems and organs may exhibit varying aging rates, suggesting the presence of multiple "clocks" (*5*). Hematological features have been employed to develop biological age (BA) prediction models in populations from the USA (*6*), Singapore (*7*), the Republic of Korea (*8*,*9*), Canada (*9*), and eastern Europe (*9*). BA serves as a valuable tool for identifying biomarkers and exploring risk and protective factors associated with aging. Machine learning (ML) techniques, particularly deep learning, are playing an increasingly important role in accurately predicting BA. This advancement not only enhances our understanding of healthy aging but also provides the public health sector with a powerful tool. Deep learning, an important ML method used to predict BA since 2016 (*6*), leverages deep neural networks (DNN) to improve model interpretability through enhanced nonlinear fitting, learning, and generalization abilities (*10*).

This study established and evaluated hematological BA prediction models in a large Chinese population. 7 ML methods based on 20 blood test features were used. Further analysis showed an association between BA acceleration and T2DM. The analysis also showed that controlling fasting blood glucose within normal levels may decrease BA. The study further revealed the association between BA and chronological age (CA), indicating that even among individuals with the same CA, there can be significant differences in BA, which often reflect variations in health status among individuals.

METHODS

Data Collection and Preprocessing

29 blood test features from 145,645 individuals were collected from the physical examination data center of the First Affiliated Hospital of Chongqing Medical University between 2015 and 2021. Supplementary Table S1 (available at https://weekly.chinacdc.cn/) provides a comprehensive overview of the participants' demographic profiles. This dataset includes five categories: blood routine examination, cardiovascular efficiency, diabetes mellitus, liver function, and renal function (Supplementary Table S2, available at https://weekly.chinacdc.cn/). Outliers for each variable in the dataset were removed using the quartile method (*11*). After removing features with multicollinearity, the remaining features were normalized to a range of 0 to 1. Multicollinearity can impact the individual effects of each explanatory variable, as well as model stability and generalization error. To ascertain multicollinearity within the dataset of 29 features, the variance inflation factor (VIF) was calculated. A feature was considered to exhibit multicollinearity if its VIF exceeded a threshold of 10.

BA Predictor Models

The datasets for male and female participants were randomly split into training, validation, and testing datasets at a ratio of 7∶2∶1. Seven ML methods, namely, DNN, support vector regression (SVR), stochastic gradient descent (SGD), kernel ridge regression (KRR), Least Absolute Shrinkage and Selection Operator (LASSO), K-nearest neighbors (KNN), and gradient boosting for regression (GBR), were used to build BA prediction models, with each model being optimized through parameter adjustment. To assess the performance of each model, Pearson's correlation coefficient (r) and the mean absolute error (MAE) were calculated. Additionally, permutation feature importance (PFI) was used to measure the contribution of features in DNN models. This process involved 100 permutations for each feature, and the average decrease in the coefficient of determination (R^2) was calculated before and after the permutation. In summary, the study comprehensively analyzed features of blood tests to develop and evaluate BA prediction models by applying various ML methods.

Association Analysis between BA Acceleration and T2DM

Participants with a difference between BA and CA greater than 7 years were classified as the BA acceleration group, while those with a difference less than −7 years comprised the BA deceleration group. The remaining participants constituted the control group. The T2DM status of each participant was self-reported based on previous medical history. Logistic regression, with BMI and CA as covariates, was used to analyze the association between BA acceleration and T2DM. Participants with T2DM were further divided into two groups based on glycemic control: those with impaired fasting glucose (IFG) and those with normal fasting glucose (NFG) after treatment. All individuals were categorized into six CA groups. The Kruskal-Wallis H test was used to analyze differences in BA-CA across the three groups (non-diabetic, IFG, and NFG) for each CA group in both male and female participants.

Statistical Analysis

SPSS (version 25.0; IBM Corporation, Armonk, NY, USA), R software (version 3.6.1; R Foundation, Vienna, Austria), Python (version 3.8.3; maintained by the Python Software Foundation, Chicago, IL, USA), and TensorFlow software (version 2.7.0; developed by Google Brain Team, Mountain View, CA, USA) were used to perform the analyses.

RESULTS

Dataset Preprocessing and Subsets Description

In this study, CA ranged from 11.44 to 99.82, with a detailed distribution shown in Supplementary Figure S1 (available at https://weekly.chinacdc.cn/). After outlier detection using the quartile method, 32,720 of 145,645 individuals were removed. Due to multicollinearity and strong correlations with other features, nine features were excluded (Supplementary Figure S2 and Supplementary Table S2, available at https://weekly.chinacdc.cn/). Figure 1 presents the process flowchart and sample size for each subset.

Hematological BA Models & Evaluation

Seven optimal BA prediction models were constructed using ML methods. Among these, the DNN model showed the highest coefficient and lowest mean absolute error (MAE) across the training, validation, and testing datasets (Supplementary Table S3, available at https://weekly.chinacdc.cn/). Subsequently, the DNN BA predictors were used to

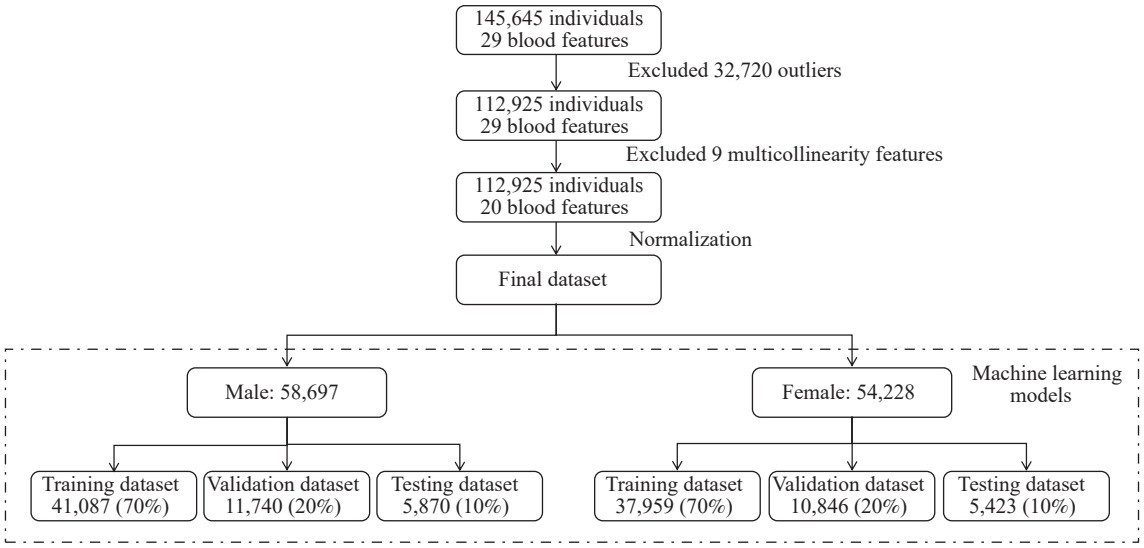


FIGURE 1. Optimized workflow for data processing and management.

predict hematological BA for both male and female subsets. Notably, strong correlations between hematological BA and CA were observed in the male and female training, validation, and testing datasets (Figure 2).

Feature Importance Analysis

The average decrease in model performance was calculated by randomly permuting features. The differences in the coefficients of determination (R^2) of features before and after random permutation were calculated as the feature importance values (Supplementary Figure S3, available at https://weekly.chinacdc.cn/). The top five important features in the BA prediction model in male subjects were albumin, fasting blood glucose, platelet hematocrit, mean corpuscular volume, and serum total cholesterol (Supplementary Figure S3A), whereas the top five features in the BA prediction model in female subjects were albumin, fasting blood glucose, serum total cholesterol, triglycerides, and urea (Supplementary Figure S3B). In both models, albumin and fasting blood glucose ranked first and second in terms of importance. Moreover, albumin was twice as important in the male model as in the female model, with almost no difference in fasting blood glucose.

Association between BA and T2DM

A total of 1,362 individuals with diabetes and 100,818 individuals without diabetes were included in this study to investigate the risk of BA acceleration or deceleration on diabetes incidence. Using logistic regression analysis, BA acceleration showed positive correlations with T2DM in both male [odds ratio (*OR*): 2.22, 95% confidence interval (*CI*): 1.77–2.77] and female subjects (*OR*: 3.10, 95% *CI*: 2.16–4.46), while BA deceleration showed negative correlations with T2DM in both male (*OR*: 0.32, 95% *CI*: 0.27–0.39) and female subjects (*OR*: 0.42, 95% *CI*: 0.33–0.53) (Table 1).

To observe whether glycemic control affected hematological BA in six CA groups, 1,362 diabetic subjects were divided into IFG and NFG groups (Supplementary Table S4, available at https://weekly.chinacdc.cn/). Using the Kruskal–Wallis H test, NFG diabetic subjects had significantly lower BAs than IFG diabetic subjects in all CA groups except those older than 80 years. Concurrently, there were no significant BA differences between NFG diabetic and nondiabetic subjects, except for individuals younger than 40 years (Figure 3).

DISCUSSION

In this study, sex-specific hematological BA prediction models constructed using the DNN algorithm showed good performance in a southwestern Chinese population. DNN outperformed six other ML methods, yielding the smallest MAE and largest correlation coefficient. This study also showed that hematological BA acceleration may be a strong risk factor for T2DM in males and females, while BA deceleration may be protective. Controlling fasting blood glucose within normal levels in those with T2DM may decrease hematological BA. This

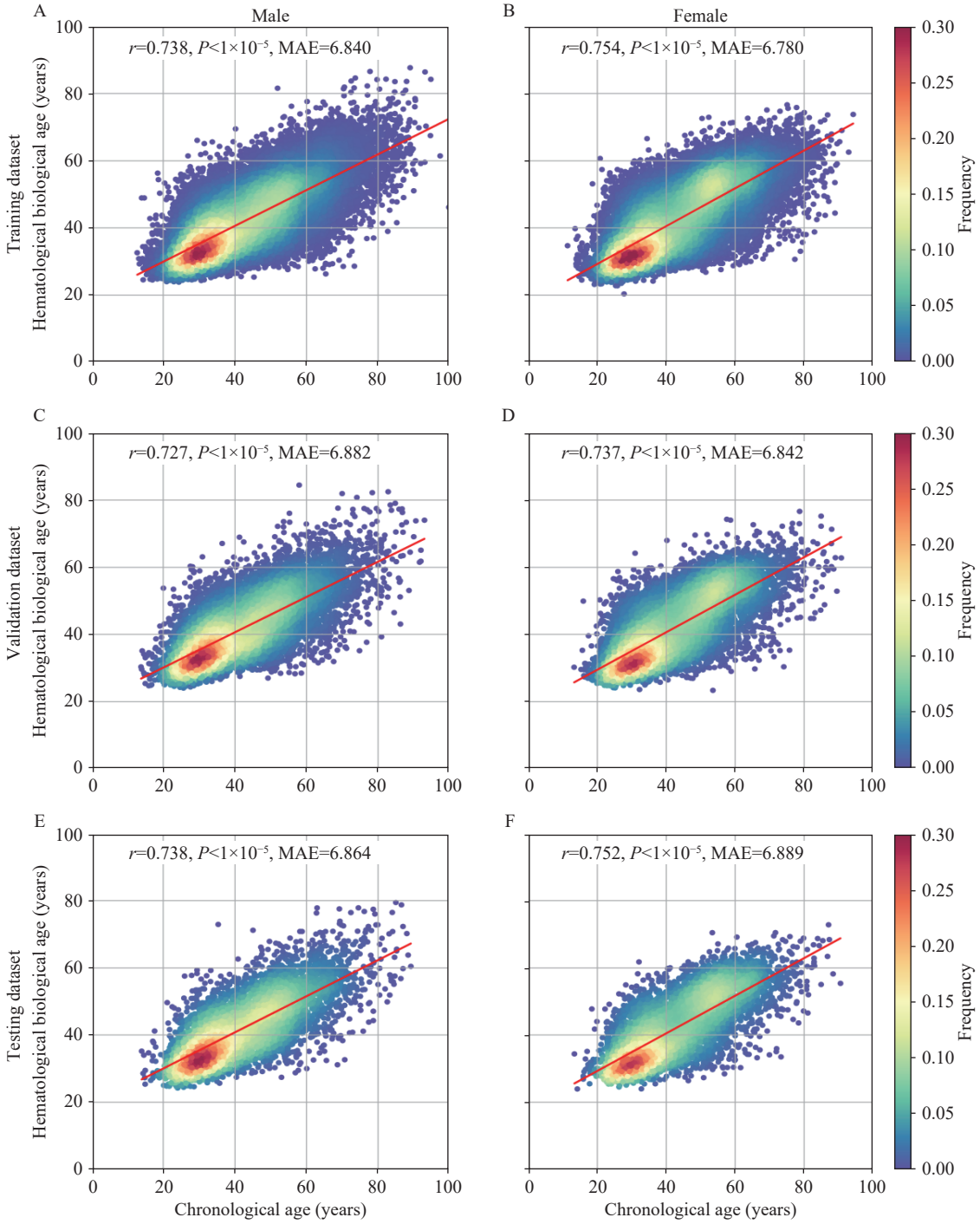


FIGURE 2. Correlations between hematological biological age and chronological age based on deep neural network models in the training, validation, and testing datasets. (A) Training dataset — Male; (B) Training dataset — Female; (C) Validation dataset — Male; (D) Validation dataset — Female; (E) Testing dataset — Male; (F) Testing dataset — Female.
Note: The colors represent the Gaussian kernel density estimation value.

observation highlights the importance of leveraging BA prediction tools to monitor the biological aging process in middle-aged and older adults (*12*), thereby informing personalized health management strategies for those at heightened risk for T2DM. Identifying individuals with a high risk of T2DM enables prompt actions to advocate for lifestyle changes, including a balanced diet, increased physical activity, and stress management. These proactive measures not only prevent the onset of T2DM but also improve the overall health of at-risk populations, halting disease progression before stages requiring intensive medical

TABLE 1. Gender differences in the impact of biological age acceleration and deceleration, body mass index, and chronological age on type 2 diabetes mellitus risk: estimation of odds ratios.

Factors	Male			Female		
	Coefficients	*ORs (95% CI)*	*P*	Coefficients	*ORs (95% CI)*	*P*
Body mass index	0.05	1.05 (1.03–1.06)	<0.001	0.03	1.03 (1.01–1.04)	0.006
Chronological age	0.11	1.12 (1.11–1.12)	<0.001	0.14	1.14 (1.13–1.16)	<0.001
Accelerate	0.80*	2.22 (1.77–2.77)*	<0.001	1.13*	3.10 (2.16–4.46)*	<0.001
Decelerate	−1.13*	0.32 (0.27–0.39)*	<0.001	−0.87*	0.42 (0.33–0.53)*	<0.001

Abbreviation: *OR*=odds ratio. *CI*=confidence interval.
* indicates that the acceleration of biological age is positively correlated with type 2 diabetes mellitus in both male and female subjects, while the deceleration of biological age is negatively correlated with type 2 diabetes mellitus.

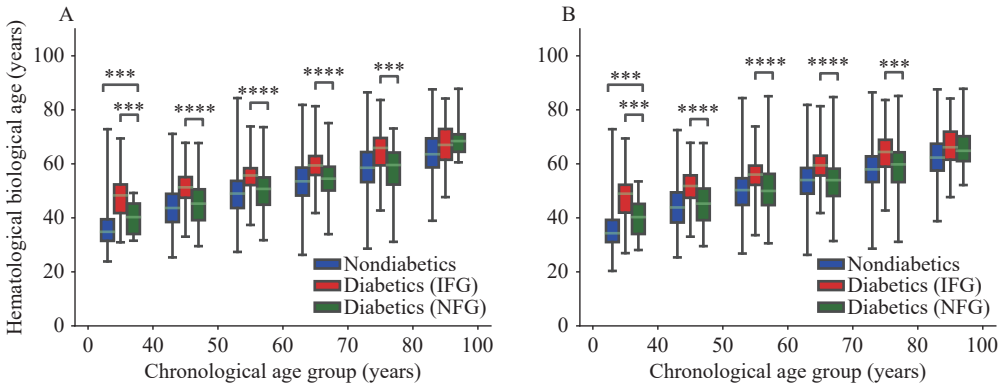


FIGURE 3. Boxplot of biological age in different age groups by gender for diabetic and non-diabetic individuals (A) Male; (B) Female.
Note: The significance levels were Bonferroni corrected.
Abbreviation: IFG=impaired fasting glucose; NFG=normal fasting glucose.
*** means $P \leq 0.001$;
**** means $P \leq 0.0001$.

care and diminishing healthcare costs.

Because information on the included features is widely available and easily obtained from routine hematological examinations, the BA prediction models may be widely applicable for assessing BA acceleration status, especially in retrospective cohort studies. Therefore, large, multicenter cohorts based on clinical or physical examination data could be established to facilitate extensive studies on the causes and adverse outcomes of accelerated aging. For example, Lu Chen et al. demonstrated an association between BA acceleration and the risk of cardiovascular events and all-cause mortality; specifically, integrating BA acceleration into mortality prediction increased Harrell's C-index from 0.813 to 0.821 (*13*).

This study's BA prediction models demonstrate significant potential for large-scale population health research, reshaping public health strategies by identifying individuals at risk of accelerated aging and elevated disease susceptibility. This precision fosters personalized preventive measures and targeted therapies. It has been reported that a majority of hematological markers are associated with both gender and age (*14*). Tracking hematological markers facilitates the mapping of aging patterns across diverse populations, informing strategic resource allocation and the design of tailored health preservation initiatives. However, this study was cross-sectional, and future cohort studies will assess the associations between hematological BA acceleration and other diseases.

In conclusion, AI-based hematological BA predictors using DNNs were established and demonstrated good performance in a large Chinese population, suggesting that they may be promising methods for assessing accelerated BA status in epidemiological studies of aging.

Conflicts of interest: No conflicts of interest.

Acknowledgements: Professor Wang Hong of Peking University for her suggestions in the analysis.

Funding: Supported by the National Human Genetic Resources Sharing Service Platform (Grant No. 2005DKA21300).

doi: 10.46234/ccdcw2024.240

[#] Corresponding authors: Zongfu Cao, caozongfu@egene.org.cn; Xu Ma, maxu_ky@nrifp.org.cn.

[1] National Human Genetic Resources Center, National Research Institute for Family Planning, Beijing, China; [2] National Human Genetic Resources Sharing Service Platform, National Research Institute for Family Planning, Beijing, China; [3] Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China; [4] The First Affiliated Hospital of Chongqing Medical University, Chongqing, China.
[&] Joint first authors.

Submitted: July 01, 2024; Accepted: November 03, 2024

REFERENCES

1. Hou YJ, Dan XL, Babbar M, Wei Y, Hasselbalch SG, Croteau DL, et al. Ageing as a risk factor for neurodegenerative disease. Nat Rev Neurol 2019;15(10):565 – 81. https://doi.org/10.1038/s41582-019-0244-7.
2. Popa-Wagner A, Petcu EB, Capitanescu B, Hermann DM, Radu E, Gresita A. Ageing as a risk factor for cerebral ischemia: Underlying mechanisms and therapy in animal models and in the clinic. Mech Ageing Dev 2020;190:111312. https://doi.org/10.1016/j.mad.2020.111312.
3. Smetana Jr K, Lacina L, Szabo P, Dvořánková B, Brož P, Šedo A. Ageing as an important risk factor for cancer. Anticancer Res 2016;36(10):5009 – 17. https://doi.org/10.21873/anticanres.11069.
4. Gloor AD, Berry GJ, Goronzy JJ, Weyand CM. Age as a risk factor in vasculitis. Semin Immunopathol 2022;44:281 – 301. https://doi.org/10.1007/s00281-022-00911-1.
5. Nie C, Li Y, Li R, Yan YZ, Zhang DT, Li T, et al. Distinct biological ages of organs and systems identified from a multi-omics study. Cell Rep 2022;38(10):110459. https://doi.org/10.1016/j.celrep.2022.110459.
6. Putin E, Mamoshina P, Aliper A, Korzinkin M, Moskalev A, Kolosov A, et al. Deep biomarkers of human aging: Application of deep neural networks to biomarker development. Aging (Albany NY) 2016;8(5):1021 – 33. https://doi.org/10.18632/aging.100968.
7. Zhong X, Lu YX, Gao Q, Nyunt SZ, Fulop T, Monterola CP, et al. Estimating biological age in the Singapore longitudinal aging study. J Gerontol A Biol Sci Med Sci 2020;75(10):1913 – 20. https://doi.org/10.1093/gerona/glz146.
8. An S, Ahn C, Moon S, Sim EJ, Park SK. Individualized biological age as a predictor of disease: Korean Genome and Epidemiology Study (KoGES) Cohort. J Pers Med 2022;12(3):505. https://doi.org/10.3390/jpm12030505.
9. Mamoshina P, Kochetov K, Putin E, Cortese F, Aliper A, Lee WS, et al. Population Specific Biomarkers of Human Aging: A Big Data Study Using South Korean, Canadian, and Eastern European Patient Populations. J Gerontol A Biol Sci Med Sci. 2018;73(11):1482–1490. http://dx.doi.org/10.1093/gerona/gly005.
10. Zhavoronkov A, Li R, Ma CDC, Mamoshina P. Deep biomarkers of aging and longevity: from research to applications. Aging (Albany NY) 2019;11(22):10771 – 80. https://doi.org/10.18632/aging.102475.
11. Manoj K, Senthamarai Kannan K. Comparison of methods for detecting outliers. Int J Sci Eng Res 2013;4(9):709-14. https://www.ijser.org/researchpaper/Comparison-of-methods-for-detecting-outliers.pdf.
12. Cao XQ, Yang GL, Jin XR, He L, Li XQ, Zheng ZT, et al. A machine learning-based aging measure among middle-aged and older Chinese adults: the China Health and Retirement Longitudinal Study. Front Med (Lausanne) 2021;8:698851. https://doi.org/10.3389/fmed.2021.698851.
13. Chen L, Zhang YQ, Yu CQ, Guo Y, Sun DJY, Pang YJ, et al. Modeling biological age using blood biomarkers and physical measurements in Chinese adults. eBioMedicine 2023;89:104458. https://doi.org/10.1016/j.ebiom.2023.104458.
14. Sebastiani P, Thyagarajan B, Sun FG, Honig LS, Schupf N, Cosentino S, et al. Age and sex distributions of age-related biomarker values in healthy older adults from the long life family study. J Am Geriatr Soc 2016;64(11):e189 – 94. https://doi.org/10.1111/jgs.14522.

SUPPLEMENTARY MATERIAL

SUPPLEMENTARY TABLE S1. Demographic characteristics of study participants.

Year	2016	2017	2018	2019	2020	2021
Sample size	15,689	16,937	21,521	26,066	16,021	16,691
Age, median (SD)	40.75 (12.85)	38.87 (12.73)	39.32 (13.36)	34.56 (13.07)	40.19 (13.65)	37.90 (13.60)
Age, range	14.42–99.82	13.42–94.78	12.76–90.00	13.38–93.15	13.34–93.42	11.44–97.37
Sex						
Female (%)	6,644 (42.35)	7,812 (46.12)	10,092 (46.89)	12,492 (47.92)	8,187 (51.10)	9,001 (53.93)
Male (%)	9,045 (57.65)	9,125 (53.88)	11,429 (53.11)	13,574 (52.08)	7,834 (48.90)	7,690 (46.07)
BMI (SD)	23.14 (5.19)	24.26 (4.00)	24.00 (6.51)	22.83 (3.37)	22.83 (3.61)	22.94 (3.66)

Note: The number of individuals undergoing health examinations in 2015 was 475, and due to the smaller sample size, it was combined with that of 2016.
Abbreviation: SD=standard deviation; BMI=body mass index.

SUPPLEMENTARY TABLE S2. The information of twenty-nine features and correlation between features and chronological age.

Abbreviation	Full name	Class	Pearson correlation	*P*	VIF
ALB	Albumin	Liver function	−0.37	0	21.7
FBG	Fasting blood glucose	Diabetes mellitus	0.34	0	1.13
TC	Serum total cholesterol	Cardiovascular efficiency	0.29	0	17.0
LDL	Low density lipoprotein	Cardiovascular efficiency	0.26	0	13.7
Urea	Urea	Renal function	0.25	0	1.15
MCV	Mean corpuscular volume	Blood routine examination	0.24	0	642
PCT	Platelet hematocrit	Blood routine examination	−0.22	0	27.1
PLT	Platelet count	Blood routine examination	−0.22	0	39.7
TP	Total protein	Liver function	−0.19	0	19.6
RBC	Red blood cell count	Blood routine examination	−0.18	0	425
AG	ALB/GLB	Liver function	−0.17	0	22.8
TG	Triglyceride	Cardiovascular efficiency	0.17	0	2.26
MCH	Mean corpuscular hemoglobin	Blood routine examination	0.14	0	845
MCHC	Mean corpuscular hemoglobin concentration	Blood routine examination	−0.13	0	229
HGB	Hemoglobin	Diabetes mellitus	−0.11	5.55×10^{-294}	1,094
HCT	Red blood cell specific volume	Blood routine examination	−0.08	1.01×10^{-176}	1,206
DBIL	Direct bilirubin	Liver function	−0.08	3.46×10^{-165}	4.20
MPV	Mean platelet volume	Blood routine examination	0.08	2.95×10^{-154}	115
P-LCR	Platelet-larger cell ratio	Blood routine examination	0.08	8.62×10^{-146}	114
WBC	White blood cell count	Blood routine examination	−0.07	1.24×10^{-128}	1.45
LY%	Percentage of lymphocyte	Blood routine examination	−0.06	6.90×10^{-102}	13.4
GLB	Globulin	Liver function	0.06	4.53×10^{-84}	.
MONO%	Percentage of monocytes	Blood routine examination	0.05	5.62×10^{-61}	1.54
HDL	High-density lipoprotein cholesterol	Cardiovascular efficiency	0.05	3.86×10^{-54}	3.68
TBIL	Total bilirubin	Liver function	0.03	2.73×10^{-27}	3.76
NEU%	Percentage of granulocyte	Blood routine examination	0.03	9.93×10^{-27}	13.8
Cr	Creatinine	Renal function	0.03	1.74×10^{-17}	1.79
PDW	Platelet distribution width	Blood routine examination	0.02	7.83×10^{-16}	1.17
ALT	Alanine aminotransferase	Liver function	−0.01	0.003	1.37

Note: "." note the VIF value is excessively large, the tool can't be able to display it properly. After performing multicollinearity diagnosis, HCT, RBC, LDL, AG, PCT, P_LCR, MCH, NEU_PCT and GLB were excluded.
Abbreviation: VIF=variance inflation factor.

SUPPLEMENTARY TABLE S3. Performance comparison of seven machine learning methods by gender in each dataset based on MAE and *r*.

Model	Training dataset				Validation dataset				Testing dataset			
	Male		Female		Male		Female		Male		Female	
	MAE	*r*	MAE	*r*	MAE	*r*	MAE	*r*	MAE	*r*	MAE	*r*
DNN*	6.84	0.74	6.78	0.75	6.88	0.73	6.84	0.74	6.89	0.75	6.86	0.74
SVR	6.92	0.73	6.94	0.74	6.99	0.72	7.00	0.73	7.11	0.73	6.98	0.73
SGD	7.50	0.69	7.55	0.70	7.45	0.69	7.50	0.69	7.59	0.70	7.49	0.69
KRR	7.86	0.65	7.64	0.69	7.82	0.65	7.60	0.68	7.68	0.69	7.88	0.65
LASSO	7.85	0.67	7.92	0.69	7.79	0.67	7.80	0.68	7.97	0.69	7.82	0.68
KNN	7.62	0.71	7.52	0.72	7.73	0.69	7.55	0.70	7.72	0.70	7.70	0.70
GBR	7.10	0.72	7.12	0.74	7.20	0.71	7.19	0.72	7.26	0.73	7.18	0.72

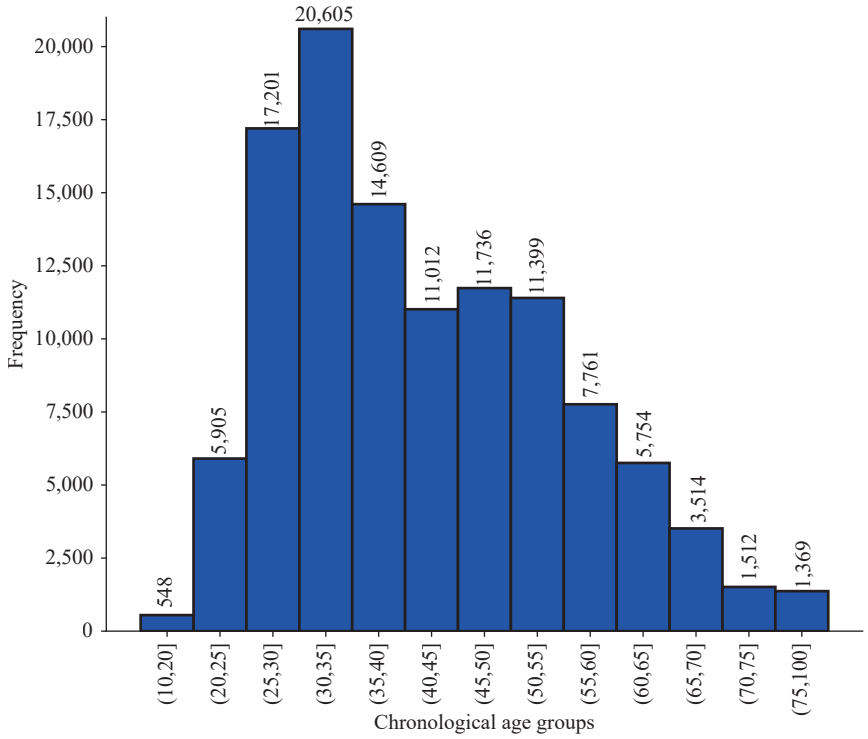
Abbreviation: MAE=mean absolute error; *r*=pearson's correlation coefficient; DNN=deep neural networks; SVR=support vector regression; SGD=stochastic gradient descent; KRR=kernel ridge regression; LASSO=Least Absolute Shrinkage and Selection Operator; KNN=K-nearest neighbors; GBR=gradient boosting for regression.

* highlights the superior performance of the DNN model in comparison to other models across the training, validation, and testing datasets.

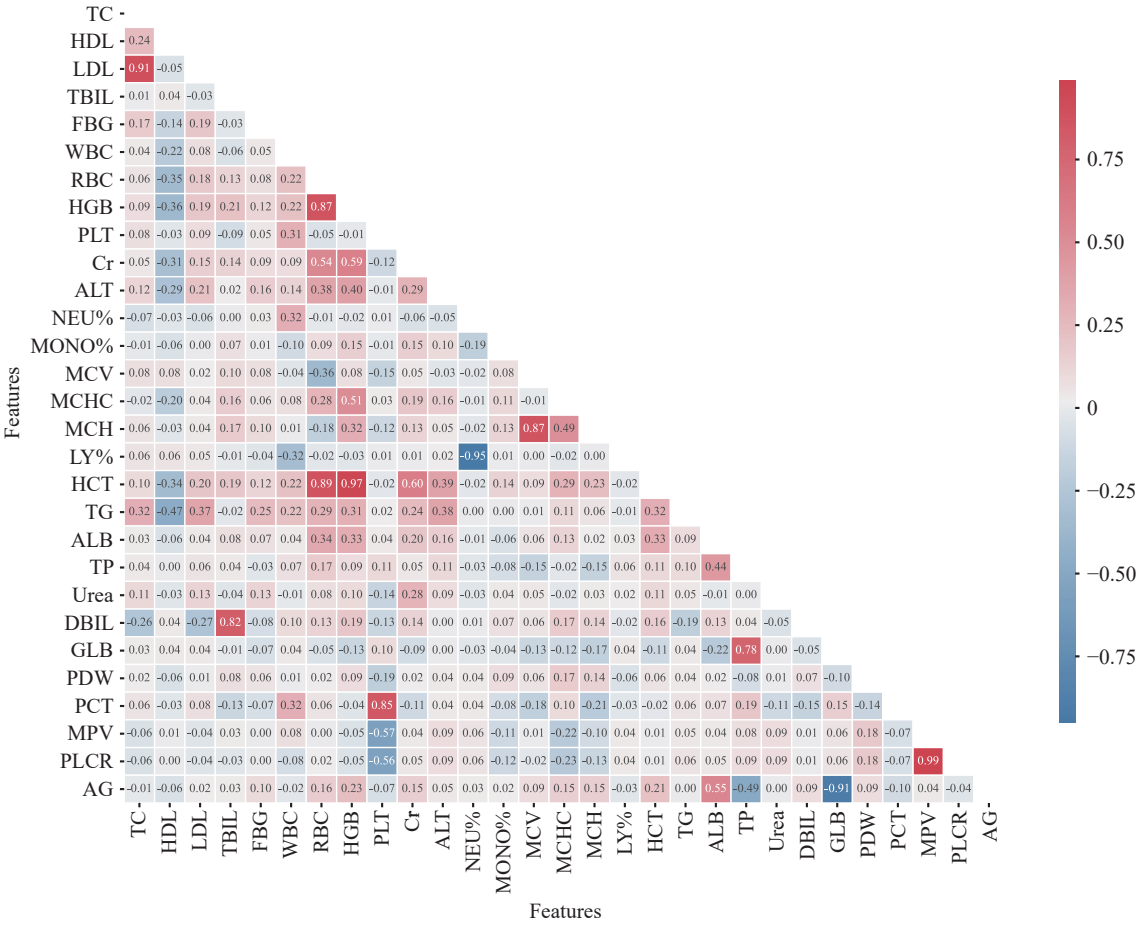
SUPPLEMENTARY TABLE S4. Statistics on the number of male and female individuals in different age ranges for each group.

Groups	Nondiabetics		Diabetics IFG		Diabetics NFG	
	Male	Female	Male	Female	Male	Female
(10,40]	28,668	24,094	29	5	14	12
(40,50]	11,347	9,005	123	17	40	6
(50,60]	8,400	8,680	215	92	92	51
(60,70]	3,941	4,262	167	117	62	62
(70,80]	1,013	966	68	54	31	38
(80,100]	279	163	30	13	12	12
Sum	53,648	47,170	632	298	251	181

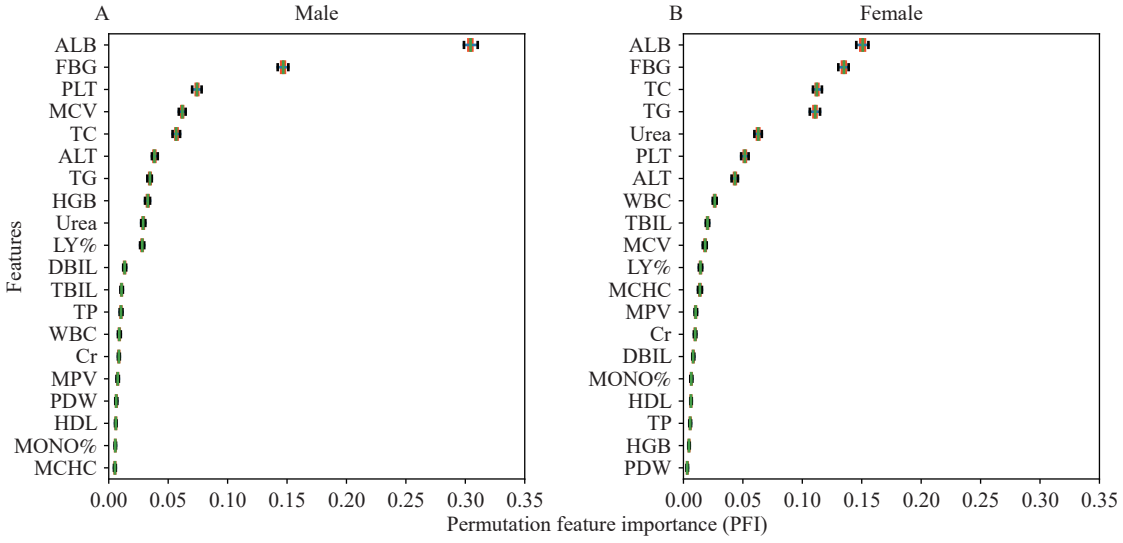
Abbreviation: IFG=Impaired fasting glucose; NFG=Normal fasting glucose.



SUPPLEMENTARY FIGURE S1. Histogram illustrating the distribution of 112,925 individuals across different chronological age groups.



SUPPLEMENTARY FIGURE S2. Heatmap of Pearson correlation between twenty-nine features.
Note: The intensity of color represents the magnitude of the Pearson correlation coefficient.



SUPPLEMENTARY FIGURE S3. Deep Neural Networks age-predictor model's feature importance for (A) male and (B) female subset.

Youth Editorial Board

Indexed by Science Citation Index Expanded (SCIE), Social Sciences Citation Index (SSCI), PubMed Central (PMC), Scopus, Chinese Scientific and Technical Papers and Citations, and Chinese Science Citation Database (CSCD)

The inauguration of *China CDC Weekly* is in part supported by Project for Enhancing International Impact of China STM Journals Category D (PIIJ2-D-04-(2018)) of China Association for Science and Technology (CAST).

CCDC weekly

Vol. 6 No. 45 Nov. 8, 2024

Responsible Authority
National Disease Control and Prevention Administration

Sponsor
Chinese Center for Disease Control and Prevention

Editing and Publishing
China CDC Weekly Editorial Office
No.155 Changbai Road, Changping District, Beijing, China
Tel: 86-10-63150501, 63150701
Email: weekly@chinacdc.cn

CSSN
ISSN 2096-7071 (Print)
ISSN 2096-3101 (Online)
CN 10-1629/R1