International collaboration in cohort studies in China: opportunities and challenges

Zhengming Chen and colleagues outline the key value, opportunities, and challenges of international collaboration in cohort studies in China

Since the early 1950s, major developments in prospective cohort studies have greatly advanced our knowledge about the causes and prevention of chronic non-communicable diseases. China began conducting cohort studies in the 1970s, modelled initially on the British Doctor Study[1] and the Framingham Heart Study,[2] but later with increasing regional coverage, sample size, complexity, and international collaboration.[3-5]

China is ideally placed for undertaking large scale cohort studies, given its large population size; diversity in lifestyles, environments, and disease patterns; well established healthcare infrastructures; and cost effectiveness in study conduct. Recent advances in information technology and biomedical sciences have enabled China to enter a new phase of cohort development to address emerging health challenges and research gaps. Close international collaboration has been an important factor in the success of cohort studies in China and elsewhere and will continue to be so.

KEY MESSAGES

- Large cohort studies in China can capture unique characteristics in risk exposures, disease patterns, and genetic architecture that will lead to novel discoveries and improvement in disease prevention and treatment
- International collaboration on China based cohort studies, which may cover different stages of study lifecycle, can help to improve study quality, attract external funding, and maximise study potential and global impact
- Important challenges for international collaboration include cultural differences, data comparability, funding priorities, regulatory constraints, and geopolitical tensions
- Good governance framework, openness, transparency, cooperation, and data sharing are key elements to the success of collaborative cohort studies

We analysed the key opportunities and challenges for international collaboration involving cohort studies in China, and highlighted their potential to shape the future of population health and biomedical research, both nationally and globally. We focused on large population based cohort studies of adult populations since they share similar design principles and methodology and have pivotal roles in tackling the largest disease burden in society this century. However, the issues covered should also be relevant for cohort studies in other populations (eg, children) or settings (eg, hospitals).

Need for international collaboration in cohort studies

While the core design principles remain unchanged in the 21st century, contemporary cohort studies are often large (n>100 000), long term, and complex, with collection and long term storage of biological samples (eg, blood, urine, stool), comprehensive exposure assessment using conventional (eg, questionnaire) and novel approaches (eg, wearables, imaging, multiomics assays), and electronic linkages to different healthcare records. However, the complexity of the study procedures and data collected warrant carefully designed IT, data management, and quality control systems, as well as cross disciplinary collaboration throughout different phases of study. There is also an increasing need for cloud computing, AI, and data science to store, process, and analyse high dimensional big (eg, whole genome sequencing) data. As such, any single institute may not have all the essential infrastructure and expertise to effectively establish and manage a large cohort study and to fully utilise the resultant data. Collaboration at every stage of cohort development with researchers nationally and internationally is essential to maximise the scientific potential and global impact of a cohort.

Approaches of international collaboration

International collaboration on cohort studies in China typically involves contribution of scientific insights and analytical skills to specific papers; participation of Chinese cohorts in international consortiums; comparative investigations across Chinese and international cohorts; or long term partnerships between Chinese and overseas researchers, often covering the full study lifecycle. These approaches are not mutually exclusive, with some cohorts having multiple tiers of collaboration, all underpinned by openness and data sharing. Among existing cohorts involving extensive international collaboration and data sharing in China, the China Kadoorie Biobank (CKB) and UK Biobank, each with 0.5 million participants, are the leading examples, along with other large (n >50 000) Chinese cohorts (eg, Shanghai Women's Health Study (SWHS) and Shanghai Men's Health Studies (SMHS)) (table 1).

The CKB was established in 2004-08 with a Sino-British collaboration covering the entire project lifecycle. In particular, the Oxford based team raised initial funding and developed detailed procedures and IT systems for study conduct, quality control, management, and monitoring; both teams worked closely in undertaking pilot studies and refining the study design, procedures, and fieldwork plans; and the China based team mobilised and managed immense human and logistical resources for implementation. Several other major cohorts adopted broadly similar approaches, with, for example, the SWHS and SMHS co-established as a China-US collaboration.[4 8]

While the CKB benefited from experience learnt in earlier cohort studies (eg, Mexico City Prospective Study, MCPS), it also informed the UK Biobank, which was established in 2006-10. As one of the most extensively used open medical data resources worldwide, the UK Biobank has also benefited China by enhancing capacity development and enabling rapidly growing high impact research among a large number of Chinese scientists (that is, 16% of total) (fig 1).

Recognising the unique value of biobank studies, during 2016-18 the

Table 1 | Main large population based prospective cohort studies in adults in China

Study name	Age range (years)	Sample size*	Year of baseline survey	Biosamples	Outcome follow-up	International collaborating institute†	Data sharing
Exemplar cohorts established before 2016							
Chinese Prospective Smoking Study (CPSS)[6 7]	40-79	225 721	1990-91	No	Passive (mortality)‡	University of Oxford (UK)	Collaboration
Shanghai Women's Health Study (SWHS)[4]	40-70	74 942	1996-2000	Yes	Active (survey) + passive (mortality and cancer)	Vanderbilt University (US)	Collaboration
Shanghai Men's Health Study (SMHS)[8]	40-74	61 480	2002-06	Yes	Active (survey) + passive (mortality and cancer)	Vanderbilt University (US)	Collaboration
China Kadoorie Biobank (CKB)[5]	30-79	512 714	2004-08	Yes	Passive (mortality and morbidity)	University of Oxford (UK)	Open access + collaboration
China Chronic Disease and Risk Factor Surveillance (CCDRFS) cohort[9]	15-69	732 472	2004-18 §	Yes[§]	Passive (mortality)	NA	Collaboration
Taizhou Biobank[10]	30-80	123 115	2007-16	Yes	Active (survey) + passive (mortality and morbidity)	Karolinska Institute (Sweden)	Collaboration
Seven cohorts of the Million Cohort Initiative established in 2016							
Shanghai Suburban Adult Cohort and Biobank (SSACB)[11]	20-74	69 116	2016-19	Yes	Active (survey) + passive (mortality and morbidity)¶	NA	Collaboration
Beijing, Tianjin, and Hebei (BTH) general population cohort[12]**	18-70+	82 138	2017-21	Yes	Active (survey)	NA	Internal
Regional Ethnic Cohort Study in North west China (RECS)[13]**	35-74	118 572	2018-19	Yes	Passive (mortality and morbidity) + active (survey in 10%)	NA	Collaboration
South China Cohort (SCC)[14]	25-89	116 520	2018-20	Yes	Passive (mortality and morbidity)		Collaboration
Central China Cohort Study of common chronic and non-communicable diseases[15 ††]	Mean 55.3	115 424	Not reported	Yes	Active (survey) + passive (mortality and morbidity)	NA	Collaboration
China South-west General Population Cohort (also known as China Multi-Ethnic Cohort (CMEC)[16]	18-79	99 556	2018-19	Yes	Passive (mortality and morbidity)	NA	Collaboration
North east China Biobank (NEC-Biobank)[17]**	17-79	101 773	2018-21	Yes	Active (survey) + passive (mortality and morbidity)	NA	Collaboration

*Sample size of adult participants. Some cohorts included children and the corresponding sample size was not indicated in this table.
†Primary international collaborating institutes involved since study conceptualisation.
‡Follow-up for mortality ended in 1999.
§Successions of periodic nationwide surveys conducted every 3 years as baseline; biosamples collected since 2010.
¶Passive follow-up restricted to township or street level healthcare records but not national systems, so follow-up may be incomplete.
**Multi-component cohort constituted of multiple sub-cohorts of distinctive population characteristics.
††Age range or baseline survey year was not clearly reported in the published material.

Chinese government invested 400 million Chinese Yuan (£41m; $56m; €48m) to establish more than 25 population and disease specific cohorts, including seven community based cohorts as part of the Million Cohort Initiative, for which the CKB was selected as a demonstration project (table 1).[18] Moreover, as medical record linkages become increasingly feasible, many nationwide representative cross sectional surveys (eg, the China Chronic Disease and Risk Factor Surveillance (CCDRFS)) have been converted into prospective cohort studies.[9] In 2017, the China Cohort Consortium (CCC; https://chinacohort.bjmu.edu.cn/) was established to facilitate collaboration, involving 154 cohorts of over 100 million individuals to date.[19] Chinese cohorts are increasingly sharing data nationally and internationally, but most require direct collaboration on specific projects. Although the CKB has been able to share individual participant data worldwide since 2015, it only involves non-genetic data owing to policy constraints in China and the number of open access projects was much smaller compared with the UK Biobank. Altogether, cohort studies in China offer uniquely rich opportunities for international collaboration, especially if the genetic data can be shared more widely, and the post-2016 cohorts have laid promising ground for future collaboration on rapidly evolving health issues in younger generations.

Value of international collaboration

Cohort studies in China can capture a different range of risk exposures, genetic architecture, disease patterns, and social context from those undertaken in other populations. These greatly diversify and enrich global population health research, shedding new light on longstanding evidence gaps that could not be properly addressed by cohorts in western high income populations alone. For example, cohort studies in China could reliably examine the associations of blood cholesterol with ischaemic heart disease at levels well below those typically seen in western populations,[20 21] informing the design of statin trials in the west, which went on to demonstrate the benefits of lipid lowering at all levels.[22] Using East Asian specific "flushing" variants (absent in white people), a Mendelian randomisation study in China reliably refuted a longstanding belief in the benefits of moderate alcohol drinking on the risk of cardiovascular disease.[23]

Studies in China also enable investigation of specific diseases that are now relatively rare (and thus difficult to study) in western and other high income East Asia (eg, Japan) populations, as exemplified by the contrasting causal associations of low density lipoprotein cholesterol with ischaemic stroke (positive)

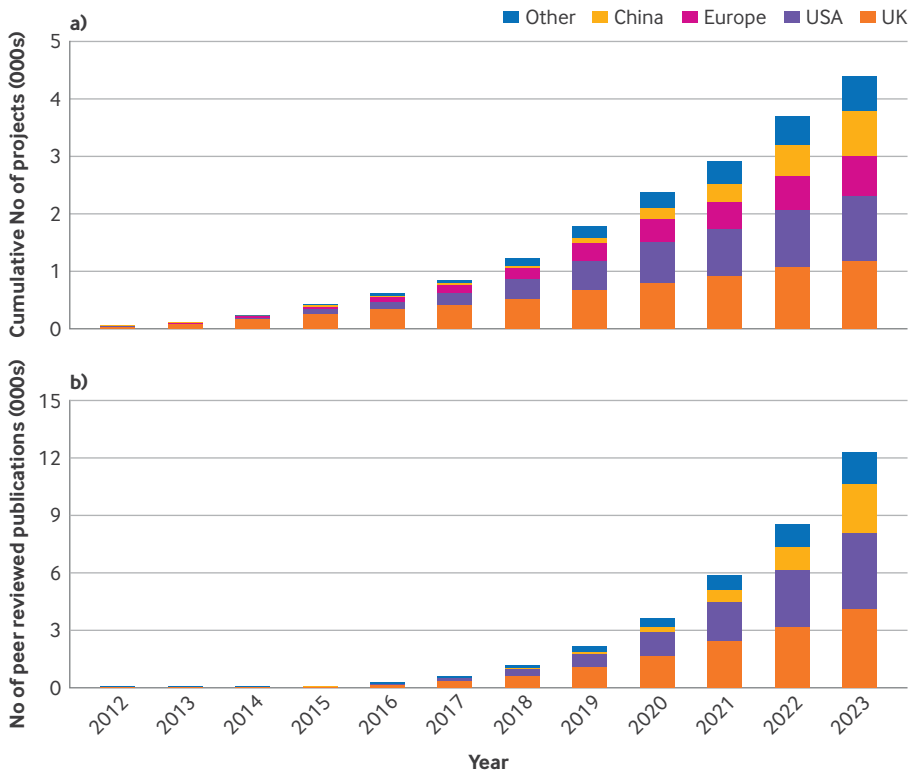

Fig 1 | UK Biobank research statistics on cumulative number of active research projects by region in 2012-23 and number of peer reviewed publications by region in 2012-23

and intracerebral haemorrhage (inverse) in Chinese adults, which informed the safe use of statins in populations with high rates of intracerebral haemorrhage.[21 22] Moreover, the complex interplay of genes and environment in influencing disease risks may operate differently across populations. This was demonstrated by findings of the importance of higher homocysteine levels for increased risk of stroke, but not ischaemic heart disease, in a low folate intake Chinese population,[24] in contrast to previous null findings in populations with high folate intake. Furthermore, successive large Chinese cohorts have documented the growing tobacco related mortality among men (in contrast to the declining trends among Chinese women and in high income populations), informing China's national tobacco control policy[6 25] and long term monitoring of smoking (and other exposures) in national representative cohorts (eg, CCDRFS).

International collaboration can also facilitate knowledge exchange and capacity development, improve study design and ethical standards, attract funding from multiple sources, introduce best practices, enhance data quality, and increase study visibility and global impact.

At the conceptualisation stage, international collaboration helps to identify research priorities of global importance, while ensuring the study design, main objectives, and operation are tailored to the local context. At the planning and implementation stages, there could be sharing of expertise and identification of best practices in study conduct (eg, data security and management). In particular, as ethics and data management regulations often differ across countries, international collaboration would naturally require the study to be conducted to the highest standards across all countries.

For study conduct, traditional paper based data collection poses major logistical and quality control challenges to large cohort studies. The CKB was the first cohort in China to be paper-free, enabled by a laptop based electronic questionnaire with built-in logic and error checks as part of a comprehensive bespoke IT system for data collection and study management.[26]

The size and complexity of contemporary cohort studies require substantial investment not only to enrol participants but also to maintain the study long term (eg, outcome follow-up, periodic re-survey, storage of biosamples) and to turn samples into data using new technologies (eg, whole genome sequencing, multiomics). International collaboration can effectively pool funding from multiple sources. For example, the CKB was initially funded by a philanthropic seed grant from the Hong Kong-based Kadoorie Charitable Foundation, and the successful establishment of the cohort has subsequently attracted further public and charitable funding from various sources in China (eg, National Natural Science Foundation, Ministry of Science and Technology) and the UK (eg, Medical Research Council, Wellcome Trust) and industry (eg, Merck, GSK, BGI).

Similarly, as a global open access resource, the UK Biobank has attracted substantial funding from industry for large scale assays (eg, whole genome sequencing, proteomics, metabolomics),[27 28] which would not be feasible with traditional national government or charitable funding. Collaboration with industry also helps to improve assay technology and reduce costs. The data generated through industry funding are also released to the wider research community after a short period of exclusive use, benefiting a large number of users globally. By charging a modest access fee (with discount for users from low income countries) for cost recovery purpose, it also helps to improve the long term sustainability of the UK Biobank leveraged by numerous research projects using the data. This "open access" model is increasingly common in large scale epidemiological research, as funders, researchers, and scientific journals recognise the data generated as common goods to humanity. Data sharing and international collaboration can also help improve the quality of research by enabling reproducible science with cross validation across and within cohorts, as well as reducing publication bias.

Challenges of international collaboration

Clearly, international collaboration entails important challenges. Firstly, cultural and language barriers between different stakeholders are inevitable. Experiences of successful Chinese cohorts have shown clear advantages of having key personnel with diverse cultural understanding and research experience. For example, both the CKB and SWHS have benefited significantly from having at least one lead investigator of Chinese origin with research training both within and outside

Box 1: Key opportunities and challenges of international collaboration on Chinese cohorts

Opportunities
- Capture diverse risk exposures, disease patterns, genetic architecture, and social context of global relevance
- Contribution to international monitoring efforts on exposure and disease patterns
- Sharing of expertise
- Identification of best practices
- Effective pooling of funding from diverse sources

Challenges
- Language and culture barriers, including differences in research assessment culture
- Different ethical standards
- Geopolitical tension
- Data security and confidentiality concerns
- Nationalism and funding priorities
- Comparability of variables and laboratory results

China, who can facilitate collaboration and resolve conflicts at all stages of study conduct. During the long term collaboration, mutual academic visits and training of junior researchers are important to strengthen relationships and ensure sustainability from the bottom up.

Ethics is one key aspect that requires in-depth mutual understanding of research practices and culture, especially given the ever evolving and distinctive ethical standards and regulations internationally. The concept of informed consent may differ by population, and the wordings used in information sheets may not always be directly translatable across cultures. It is therefore essential to conduct iterative internal review, pilot testing, translation, and back translation of different study materials.

Another major challenge lies in the research assessment culture. Researchers in China are under immense pressure from their highly competitive assessment system, which, for example, does not fully recognise shared first or senior authorships—something that is common and often inevitable in large scale collaborations. This issue also exists outside China, but there has been promising development to promote recognition beyond mere ranking of authorship by focusing on individuals' specific contributions.[29] A successful collaboration also requires a clear governance framework and agreement on data and intellectual property ownership, which may differ depending on the nature of collaboration and levels of investment from each party.

Major challenges also arise from rapidly evolving geopolitical tensions and regulatory requirements internationally. In 2019, the Chinese government passed a new law regulating the collection, use, and sharing, both cross border and within China, of human genetic materials and information.[30] While the law serves to ensure the appropriate use of genetic data for research and clinical translation, the initial lack of detailed implementation regulations has largely stalled many international collaborations. China has also introduced cyberspace security regulation to govern cross border transfer of both commercial and research related data.

Similarly, with emerging geopolitical tensions, some western countries have imposed restrictions on collaboration with research institutions in China. Again, the lack of clear guidelines has resulted in confusion and widely varying interpretation, with some UK institutions, for example, now requiring application of specific data export licences for collaboration with certain Chinese institutions and enterprises. Unlike UK Biobank, the US All of Us cohort does not permit data sharing with researchers from China. These measures may have been introduced to tackle legitimate concerns about highly sensitive national security related technologies, but they also pose major obstacles to other politically neutral scientific collaboration that would benefit population worldwide.

There are also legitimate concerns about sharing personal data. The conventional approach to deliver datasets to users, even with anonymisation, still entails the risk of untraceable secondary sharing in breach of data sharing agreements, especially in places with weak contractual culture. There is also growing concern about confidential data leakage or participant re-identification with the use of large language models in processing and analysing cohort data. However, increasingly, these concerns can be minimised by using secure cloud based analysis platforms (eg, UK Biobank Research Analysis Platform (RAP)) that can avoid direct data transfer and document a clear audit trail of data utilisation. Federated analysis is also increasingly common in multicohort collaborations, with individual cohorts analysing their data locally and integrating results using meta-analysis, thereby avoiding data transfer.[31]

The emerging major cut in medical research and foreign aid funding is also threatening international collaboration.[32] Behind this is a worrying global trend of nationalism that prioritises domestic interest over international collaboration, posing considerable challenges to the funding and sustainability of major cohorts founded on international collaboration. In China, for example, there is a growing trend to discount publications based on overseas databases (eg, UK Biobank) when researchers are evaluated for funding applications—something yet to be seen elsewhere. In the UK, the Wellcome Trust no longer funds research projects in China owing to the lack of a representative office in China, as required by the recently established Chinese law governing overseas non-governmental organisation operations.

There are also many operational and data harmonisation challenges. For example, data collected under different sociocultural contexts may not be directly comparable (eg, Townsend Deprivation Index in UK Biobank versus education in the CKB, as best measures of socioeconomic status). Moreover, the quality and standards of disease diagnoses may differ across cohorts. In China, while most of the stroke diagnoses were based on brain imaging (about 95% in CKB), only about 10% of cases of chronic obstructive pulmonary disease were diagnosed by spirometry, even though a large proportion of hospital inpatient cases were found to be valid based on clinical assessment.[33 34] Furthermore, laboratory results may not always be comparable, owing to differences in assay technologies used (eg, antibody versus mass spectrometry based proteomics assays). These pose major challenges for data harmonisation and combined analyses. To tackle these issues, major efforts are now underway

to standardise and harmonise core data (eg, variable names, definitions, and values used) collected across major cohorts as a part of new initiatives by IHCC (https://globalgenomics.org/ihcc), which could serve as a template for future new cohorts.

In summary, large population based cohort studies are now at the forefront of population health and biomedical research, with great potential to transform the development of precision health. Despite the uncertainties and potential challenges, there are ample opportunities and value in international collaboration including Chinese cohorts and researchers (box 1). In future, the development of cloud based research analysis platform (as already commenced in the CKB to facilitate sharing of recently completed cohort-wide whole genome sequence data) represents a way forward, which will further maximise the scientific potential of the data generated while protecting the integrity, confidentiality, and security in accordance with national and international regulations. Meanwhile, we must continue to facilitate international collaboration by enhancing knowledge transfer, attracting inward investment, improving study quality, reducing geographical barriers, and maximising the value of study resources.

Contributors and sources: KHC has rich experience in leading multi-national collaboration on lifestyle and environmental epidemiological research based on large cohort studies; BL has studied and reported widely on population health metrics, health inequalities, and the enhancement of large-scale cohort studies, particularly the UK Biobank; JL, SZ, LL, and ZC have decades of experience in establishing and leading large-scale international collaboration on population health research involving cohorts in China, with SZ having established the China Cohort Consortium and LL and ZC co-establishing and leading CKB. In preparing this article, the authors reviewed both internal and published materials on the conduct and development of large international collaboration involving a wide range of cohort studies in China. KHC, BL, and ZC conceptualised the study, KHC and ZC wrote the first draft, and all authors reviewed and edited the manuscript. ZC is the guarantor.

Competing interests: We have read and understood BMJ policy on declaration of interests and declare no conflict of interest.

Provenance and peer review: Commissioned; externally peer reviewed.

This article is part of a collection developed in partnership with West China Hospital and the First Hospital of Jilin University. Article open access fees were funded by authors and their institutions. *The BMJ* commissioned, peer reviewed, edited, and made the decisions to publish the articles. John Ji and Jocalyn Clark were the lead editors for *The BMJ*.

Ka Hung Chan, senior research fellow[1]
Ben Lacey, associate professor[1,2]
Jun Lv, professor[3,4,5]
Siyan Zhan, professor[3,4]
Liming Li, professor[3,4,5]
Zhengming Chen, professor[1]

[1]Clinical Trial Service Unit and Epidemiological Studies Unit, Nuffield Department of Population Health, University of Oxford, UK
[2]UK Biobank, Stockport, UK
[3]Department of Epidemiology and Biostatistics, School of Public Health, Peking University, Beijing, China
[4]Key Laboratory of Epidemiology of Major Diseases (Peking University), Ministry of Education, Beijing, China
[5]Peking University Center for Public Health and Epidemic Preparedness and Response, Beijing, China

Correspondence to: Z Chen
zhengming.chen@ndph.ox.ac.uk

1 Doll R, Peto R, Boreham J, Sutherland I. Mortality in relation to smoking: 50 years' observations on male British doctors. *BMJ* 2004;328:1519. doi:10.1136/bmj.38142.554479.AE
2 Tsao CW, Vasan RS. Cohort Profile: The Framingham Heart Study (FHS): overview of milestones in cardiovascular epidemiology. *Int J Epidemiol* 2015;44:1800-13. doi:10.1093/ije/dyv337
3 Niu SR, Yang GH, Chen ZM, et al. Emerging tobacco hazards in China: 2. Early mortality results from a prospective study. *BMJ* 1998;317:1423-4. doi:10.1136/bmj.317.7170.1423
4 Zheng W, Chow WH, Yang G, et al. The Shanghai Women's Health Study: rationale, study design, and baseline characteristics. *Am J Epidemiol* 2005;162:1123-31. doi:10.1093/aje/kwi322
5 Chen Z, Chen J, Collins R, et alChina Kadoorie Biobank of 0.5 million people: survey methods, baseline characteristics and long-term follow-up. *Int J Epidemiol* 2011;40:1652-66. doi:10.1093/ije/dyr120
6 Chen Z, Peto R, Zhou M, et al. Contrasting male and female trends in tobacco-attributed mortality in China: evidence from successive nationwide prospective cohort studies. *Lancet* 2015;386:1447-56. doi:10.1016/S0140-6736(15)00340-2
7 Niu SR, Yang GH, Chen ZM, et al. Emerging tobacco hazards in China: 2. Early mortality results from a prospective study. *BMJ* 1998;317:1423-4. doi:10.1136/bmj.317.7170.1423
8 Shu XO, Li H, Yang G, et al. Cohort profile: The Shanghai Men's Health Study. *Int J Epidemiol* 2015;44:810-8. doi:10.1093/ije/dyv013
9 Zhang M, Wang L, Wu J, et al. Data resource profile: China Chronic Disease and Risk Factor Surveillance (CCDRFS). *Int J Epidemiol* 2022;51:e1-8. doi:10.1093/ije/dyab255
10 Wang X, Lu M, Qian J, et al. Rationales, design and recruitment of the Taizhou Longitudinal Study. *BMC Public Health* 2009;9:223. doi:10.1186/1471-2458-9-223
11 Zhao Q, Liu X, Jiang Y, et al. [Establishment of Shanghai suburban adult cohort and biobank] [Chinese]. *Chin J Epidemiol* 2023;44:28-33.
12 He H, Pan L, Hu Y, et al. The diverse life-course cohort (DLCC): protocol of a large-scale prospective study in China[Online First: 20220719]. *Eur J Epidemiol* 2022;37:871-80. doi:10.1007/s10654-022-00894-1
13 Li C, Liu Y, Shi G, et al. Cohort profile: Regional Ethnic Cohort Study in Northwest China. *Int J Epidemiol* 2022;51:e18-26. doi:10.1093/ije/dyab212
14 Yang J, Luo S, Liu Y, et al. Cohort profile: South China Cohort. *Int J Epidemiol* 2024;53:dyae028. doi:10.1093/ije/dyae028
15 Zhang H, Wang C, Liu X, et al. [Design and performance of a prospective cohort study of common chronic and non-communicable diseases in central China] [Chinese] . *Chin J Epidemiol* 2023;44:34-9.
16 Zhao X, Hong F, Yin J, et al. Cohort profile: the China Multi-Ethnic Cohort (CMEC) study. *Int J Epidemiol* 2021;50:721-721l. doi:10.1093/ije/dyaa185
17 Chang Q, Wu Q, Xia Y, et al. Cohort profile: The Northeast China Biobank (NEC-Biobank). *Int J Epidemiol* 2023;52:e125-36. doi:10.1093/ije/dyac172
18 Yang J, Huang W, Huang P. [Established and on-going cohort studies in China: a literature study] [Chinese]. *Chin J Publ Health* 2019;35:1393-9.
19 Sun Y, Pei Z, Zhao H, et al. Data resource profile: China Cohort Consortium (CCC). *Int J Epidemiol* 2020;49:1436-1436m. doi:10.1093/ije/dyaa102
20 Chen Z, Peto R, Collins R, MacMahon S, Lu J, Li W. Serum cholesterol concentration and coronary heart disease in population with low cholesterol concentrations. *BMJ* 1991;303:276-82. doi:10.1136/bmj.303.6797.276
21 Sun L, Clarke R, Bennett D, et al. Causal associations of blood lipids with risk of ischemic stroke and intracerebral hemorrhage in Chinese adults. *Nat Med* 2019;25:569-74. doi:10.1038/s41591-019-0366-x
22 Heart Protection Study Collaborative Group. MRC/BHF Heart Protection Study of cholesterol lowering with simvastatin in 20,536 high-risk individuals: a randomised placebo-controlled trial. *Lancet* 2002;360:7-22. doi:10.1016/S0140-6736(02)09327-3
23 Millwood IY, Walters RG, Mei XW, et al. Conventional and genetic evidence on alcohol and vascular disease aetiology: a prospective study of 500 000 men and women in China. *Lancet* 2019;393:1831-42. doi:10.1016/S0140-6736(18)31772-0
24 Bennett DA, Parish S, Millwood IY, et al. MTHFR and risk of stroke and heart disease in a low-folate population: a prospective study of 156000 Chinese adults. *Int J Epidemiol* 2023;52:1862-9. doi:10.1093/ije/dyad147
25 Chan KH, Xiao D, Zhou M, Peto R, Chen Z. Tobacco control in China. *Lancet Public Health* 2023;8:e1006-15. doi:10.1016/S2468-2667(23)00242-6
26 Chen Z. *Population biobank studies: a practical guide.* Springer Singapore, 2020.doi:10.1007/978-981-15-7666-9 .
27 Sun BB, Chiou J, Traylor M, et al. Plasma proteomic associations with genetics and health in the UK Biobank. *Nature* 2023;622:329-38. doi:10.1038/s41586-023-06592-6
28 Li S, Carss KJ, Halldorsson BV, et al. Whole-genome sequencing of half-a-million UK Biobank participants. *medRxiv* 2023:2023.12.06.23299426. [Preprint] doi:10.1101/2023.12.06.23299426.
29 McNutt MK, Bradford M, Drazen JM, et al. Transparency in authors' contributions and responsibilities to promote integrity in scientific publication. *Proc Natl Acad Sci U S A* 2018;115:2557-60. doi:10.1073/pnas.1715374115
30 The State Council of The People's Republic of China. Regulations on management of human

genetic resources. 2019. http://english.www.gov.cn/policies/latest_releases/2019/06/10/content_281476708945462.htm
31 Sun YX, Pei ZC, Zhan SY. [Data harmonization and sharing in study cohorts of respiratory diseases] [Chinese]. *Zhonghua Liu Xing Bing Xue Za Zhi* 2018;39:233-9.
32 Tanne JH. Trump's NIH research cuts are temporarily blocked as 22 US states sue. *BMJ* 2025;388:r303. doi:10.1136/bmj.r303
33 Wang C, Xu J, Yang L, et al. Prevalence and risk factors of chronic obstructive pulmonary disease in China (the China Pulmonary Health [CPH] study): a national cross-sectional study. *Lancet* 2018;391:1706-17. doi:10.1016/S0140-6736(18)30841-9
34 Chan KH, Kurmi OP, Bennett DA, et al. Solid fuel use and risks of respiratory diseases. A cohort study of 280,000 Chinese never-smokers. *Am J Respir Crit Care Med* 2019;199:352-61. doi:10.1164/rccm.201803-0432OC