Cohort Profile: The China Multi-Ethnic Cohort (CMEC) Study

Xing Zhao,[1*] Feng Hong,[2] Jianzhong Yin,[3,4] Wenge Tang,[5] Gang Zhang,[6] Xian Liang,[7] Jingzhong Li,[8] Chaoying Cui,[9] Xiaosong Li[1] on behalf of the China Multi-Ethnic Cohort (CMEC) collaborative group

[1]West China School of Public Health and West China Fourth Hospital, Sichuan University, Chengdu, Sichuan, China

[2]School of Public Health, Guizhou Medical University, Guiyang, Guizhou, China

[3]Kunming Medical University, Kunming, Yunnan, China

[4]Baoshan College of Traditional Chinese Medicine, Baoshan, Yunnan, China

[5]Chongqing Centre for Disease Control and Prevention, Chongqing, China

[6]Sichuan Provincial Hospital for Women and Children Affiliated Women and Children's Hospital of Chengdu Medical College, Chengdu, Sichuan, China

[7]Chengdu Centre for Disease Control and Prevention, Chengdu, Sichuan, China

[8]Tibet Centre for Disease Control and Prevention, Lhasa, Tibet, China and [9]High Altitude Medical Research Centre, School of Medicine, Tibet University, Lhasa, Tibet, China

*Corresponding author: Xing Zhao

West China School of Public Health and West China Fourth Hospital, Sichuan University

No. 16, Section 3, Renmin South Road, Chengdu, Sichuan, CN 610041

E-mail: xingzhao@scu.edu.cn

Abstract

Cohort purpose: The China Multi-Ethnic Cohort (CMEC) is a community population-based prospective observational study aiming to address the urgent need for understanding NCD prevalence, risk factors and associated conditions in resource-constrained settings for ethnic minorities in China.

Cohort Basics: A total of 99 556 participants aged 30 to 79 years (Tibetan populations include those aged 18 to 30 years) from the Tibetan, Yi, Miao, Bai, Bouyei, and Dong ethnic groups in Southwest China were recruited between May 2018 and September 2019.

Follow-up and attrition: All surviving study participants will be invited for re-interviews every 3-5 years with concise questionnaires to review risk exposures and disease incidence. Furthermore, the vital status of study participants will be followed up through linkage with established electronic disease registries annually.

Design and Measures: The CMEC baseline survey collected data with an electronic questionnaire and face-to-face interviews, medical examinations and clinical laboratory tests. Furthermore, we collected biological specimens, including blood, saliva and stool, for long-term storage. In addition to the individual level data, we also collected regional level data for each investigation site.

Collaboration and data access: Collaborations are welcome. Please send specific ideas to corresponding author at: xingzhao@scu.edu.cn.

Key words: CMEC, Cohort Profile, Longitudinal Community Population-based Cohort, NCDs, Ethnic Variations

Why was the cohort set up?

Non-communicable diseases (NCDs), such as cardiovascular disease, cancer, chronic respiratory disease and diabetes, have become leading causes of death worldwide[1, 2]. It is estimated that 63% of all global deaths were due to NCDs[3], with an increasing contribution from low- and middle-income countries. The prevalence and incidence of NCDs are not evenly

distributed between and within countries because of variations in the environment, lifestyle, socioeconomic position, genetic factors, and access to health care services[4, 5].

China has been facing a growing burden of NCDs with strong regional variation[6]. Most primary data on the prevalence of and risk factors for NCDs come from Central and Eastern China[7, 8], with a mix of ethnicity, socioeconomic status, geography, and dietary, behavioural and lifestyle factors that is very different compared to Western China. Ethnicity in particular is increasingly seen as an important risk factor for NCDs throughout the world[9]. Ethnic minority groups are known to have different outcomes for cancer[10], diabetes[11], cardiovascular diseases[9] and respiratory diseases[12] compared to ethnic majority groups.

China, with its 55 different ethnic minority groups, represents the largest ethnic minority population in the world (114 million in 2010)[13]. However, remarkably little is known about ethnic variation in the prevalence, incidence and/or risk factors for NCDs in China. One national cross-sectional study suggested that Tibetans had a significantly lower prevalence of diabetes than the majority Han population[14]. Another national cross-sectional study found large ethnic differences in haemoglobin distribution and anaemia prevalence[15]. To our knowledge, these are the only studies examining national ethnic variation in NCDs in China.

China's Southwest region is home to the 56 ethnic groups living in China[13]. Ethnicity not only represents genetic diversity but also characterizes strong local identities for work and daily life, such as the Tibetan people in rural areas who usually live a nomadic life. Many ethnic minorities live in remote areas, speak different dialects, and share some unique dietary habits, for example, Tibetan people favouring butter tea[16] or the Miao people's diet rich in sour soup (fermented acidic liquid)[17]. Although substantial efforts have been made to improve the socioeconomic status of ethnic minorities[18], they still tend to be much poorer than their Han counterparts. Many ethnic minorities live in distinct geographical environments and climates. Tibetans, for example, live in the Tibetan Plateau with low barometric pressure, which could cause high-altitude hypoxia or decreased oxygen levels[19]. In contrast, more than 100 million Han people live in the low-altitude Sichuan Basin and are exposed to severe air pollution.

To address the urgent need for data on the prevalence of, risk factors for and associated conditions of NCDs across various ethnicities in resource-constrained settings, the China Multi-Ethnic Cohort (CMEC) study was launched. The study is based on a standardized survey approach and multiple stringent quality control (QC) measures. Such steps are essential considering not only the substantial research gap but also that the study of NCDs across multiple ethnicities is likely to reveal important and potentially novel avoidable causes, which may inform disease prevention programmes in other populations worldwide.

Who is in the sample?

The CMEC was established in five provinces of Southwest China (Figure 1), and the baseline survey was conducted between May 2018 and September 2019. Given a full consideration of China's ethnic characteristics, population size, and non-communicable disease patterns, we aimed to include participants from the Tibetan, Yi, Miao, Bai, Bouyei, and Dong ethnic groups. In addition, to compare those ethnic minorities to the majority of people in China, the Han ethnic population was also recruited into our cohort. A multistage, stratified cluster sampling method was used to obtain samples from community-based populations. In the first stage, one to two minority settlements for each ethnic group were selected as our study sites. To reach a better representativeness of variation with regard to geography and development state, settlements from high plateaus, basins, rural areas, and highly air-polluted regions were given special consideration. In the second stage, one to eight communities (depending on the size of communities) in each settlement were selected by the local Centres for Disease Control and Prevention (CDCs), taking into account migration status, local health conditions and, most importantly, ethnic structure. In the final stage, all participants who met our inclusion criteria were invited to participate in our studies in consideration of both sex ratio and age ratio.

The inclusion criteria we applied included (i) ages of 30-79 years on the day of the investigation (except for Tibetan populations, whose inclusion criteria for age was 18-79, given the shorter life expectancy in this population); (ii) permanent residents, capability of completing baseline surveys and availability to complete follow-up studies; and (iii) complete questionnaire

interviews, physical examinations and blood tests. The exclusion criteria were as follows: (i) inability to provide a unique national identification card; (ii) severe physical or mental diseases (*e.g.,* schizophrenia and bipolar disorder); and (iii) refusal to comply with the requirements of the study. All of the participants signed an informed consent form prior to data collection. Ethical approval was received from the Sichuan University Medical Ethical Review Board (K2016038).

-Figure 1 here-

How often will participants be followed up, and what is the rate of loss likely to be?

The follow-up stage will be conducted in both passive and active modes. In passive mode, we will conduct follow-up via linkage to the electronic records of disease and death registries, mainly relying on the National Electronic Disease Surveillance System (NEDSS) and the National Health Insurance (HI) claim databases. Through annual data linkage with these established electronic disease registries, various health outcomes of the study participants will be obtained. The diagnoses of these conditions are based on well-accepted international standards. In China, the underlying and contributing causes of death are coded according to the 10th version of the International Statistical Classification of Diseases (ICD-10)[20]. All deaths during follow-up will be checked regularly for cause-specific mortality through the death certificates reported to the regional CDCs and the death registries. In active mode, all surviving study participants will be invited to resurvey every 3-5 years. We will re-estimate the status of risk exposure and document the incidence of multiple diseases (*e.g*., cancer, stroke, heart attack, diabetes mellitus, hypertension) using the same measurements in the baseline survey (see more details in the following section). We aim to minimize the expected rate of loss to less than an estimated 8% in the next round of follow-up.

What has been measured?

Table 1 summarizes the measurements collected from the CMEC baseline survey, which consisted of an electronic questionnaire with face-to-face interviews, medical examinations and clinical laboratory tests. In addition, we collected biological specimens, including blood, saliva and stool, which were stored for further use. In addition to the individual level data for each participant, we also collected the regional level data for each investigation site. More details on these measurements are described below.

Collaboration and field workflow

The CMEC baseline survey was conducted through the collaborative efforts of multiple agencies, including academic institutions, CDCs, local clinical centres, Third-Party Medical Laboratory (TPML) companies and local governments. Each agency has a clear role in the process of the investigation. Typically, the local government began the CMEC propaganda or publicity campaigns a few weeks before the formal investigation. Meanwhile, the residents were fully informed of the benefits and requirements of participating in the CMEC study. Then, residents who were willing to participate made an appointment with the local clinical centre. The clinic visit for each participant typically took 90–120 min. The average daily recruitment rate was 70–80 participants per site.

-Table 1 here-

Electronic questionnaire

The main content of our questionnaire referred to the study of the China Kadoorie Biobank as the prototype[21]. Major modifications were made in sections regarding physical activity, habitual diet and psychological conditions to fully capture the unique features of the different ethnic minority groups. We used a tablet computer with a self-developed application (CMEC App) to collect the questionnaire information (see more details in Table 1). The information was collected by a face-to-face interview implemented by well-trained interviewers who were typically local college students with medical backgrounds. For a skilled interviewer, an average of 30 to 45 min was often required to complete a questionnaire.

The whole interview was audio recorded, and we documented the duration spent on each question. For data QC, on the same day of data collection, the data quality inspectors, who were chosen from excellent interviewers, drew random samples of questionnaires to assess their data quality by listening to the audio records. The sampling scheme was built on our computer system and could ensure that each interviewer was sampled at least once. The next day, the assessment report was fed back to the interviewers to help them improve their interviewing skills. Questionnaires that were classified as unqualified were excluded from the final analysis.

Medical examinations

We conducted medical examinations mainly using the resources and personnel at local clinical centres. We implemented standardized training for the doctors and nurses before the investigation.

For those testing devices that the local clinical centres were rarely equipped with, we provided unified devices for each site and trained the local staff to operate those devices, including bone mineral density densitometers (OSTEOKJ3000) and peak expiratory flowmeters (KOKA PEF-3). Particularly, for Tibet and high-altitude areas in Sichuan and Yunnan, qualified physicians used the chronic high-altitude disease scale, which consists of 7 questions or symptoms for assessing chronic high-altitude disease and its severity.

The CDC took the main responsibility of field QC guided by a handbook and serval forms to monitor key risk points. The main measures of QC included checking whether the measurement instruments were in good condition, inspecting whether the measurement was in accordance with SOPs, randomly sampling participants to re-test their results, *etc*.

Clinical laboratory tests and biobanking

All participants provided blood and urine samples on-site at the time of the baseline survey. Venous blood samples, collected after overnight fasting (at least 8 hours), were used for clinical laboratory testing, including routine blood tests, fasting blood glucose, lipid levels and hepatic function. (Table 1). A total of 14 mL blood was collected into two ethylenediamine

tetraacetic acid dipotassium (EDTA-K2) tubes (one 5 mL and one 2 mL), one EDTA-K2 tube with sodium fluoride (2 mL) and one vacuum tube without anticoagulants (5 mL). The blood samples were temporarily stored at 4°C before delivery for biobanking or testing by 2 Third-Party Medical Laboratories (TPMLs)/independent testing organizations via a cold chain. A total of 15~20 mL mid-stream urine was collected for routine urine testing. A total of 4 000 saliva and 1 000 stool samples were collected from Tibetans and Han individuals residing in Sichuan Province for multi-omics analysis, such as the microbiome and systemic metabolism.

All the samples were centralized in a biobank located at West China School of Public Health, Sichuan University and stored at -80°C in cryogenic refrigerators. Computer systems for sample entry, access and temperature monitoring were developed and have been applied.

QC was implemented throughout the processes of sample collection, shipment, testing at the TPMLs, and storage at the biobanking facility. QC lists were strictly followed, including critical items to ensure high-quality samples, such as checking the identity of the sample and the participant ID, standardizing the volume of the samples, and regulating the storage time and temperature conditions of samples on site. To avoid any sample thawing during the shipment from each site to the biobanking facility and the TPMLs, electronic thermometers were placed in each sample's dry ice box, and the temperature of samples was kept below -20°C before entry into the biobanking facility or testing by the TPMLs. The QC process for sample testing at the TPMLs was approved by the clinical testing centre of the China Health Committee and carried out every day during the study period. Blind panels of samples were periodically sent to the TPMLs for QC testing. Batches of controls were simultaneously put into the biobanking facility with samples and taken out for testing at regular intervals as a storage QC procedure.

Computer system and data cleaning

All of the above measurements were stored and managed in electronic form by a self-developed computer system, with the central server located at West China School of Public Health, Sichuan University. Different datasets were linked by the unique participant's ID. A visualized data organization schematic can be seen in Figure 2.

In addition to data storage and management, our computer system also had some built-in functions to help with data cleaning based on electronic data. The whole process can be divided into two parts: detection and verification. In the detection part, we checked for duplication, completeness, outliers and logic errors via predesigned algorithms. Once suspicious data errors were detected, we ran a further verification process by listening to the audio recordings, looking up examination reports or calling back the participants. Then, the data were revised, and the whole correction process was recorded by our computer system.

-Figure 2 here-

What has it found?

Table 2 shows the baseline characteristics of the participants. There were more women than men across the different ethnicities. The proportions of women ranged from 54.39% in the Han population in Basin (Chongqing, Chengdu) to 70.38% in the Bai population in the high plateau (Yunnan). While we observed a similar age distribution and marital statuses across participants of different ethnicities, significant disparities were found in terms of highest education completed. For instance, 71.40% of Tibetans in Aba reported no formal schooling, while this number was as low as 10.66% among the Han in Basin. The proportions of participants who never smoked (91.27%) or drank alcohol (93.09%) were relatively high among Tibetans in Aba compared to other populations in our study. Additionally, we noted variations regarding some lifestyle variables within an ethnic group, *e.g.,* the proportion that never consumed tea was 32.31% for Tibetans in Aba and 8.12% for Tibetans in Lhasa.

Table 3 presents the characteristics of the participants' medical examinations. The mean height was between 154.42 cm in Miao peoples and 161.00 cm in Tibetans from Lhasa. The mean weight was between 56.93 kg in the Han population in the high plateau (Yunnan) and 65.95 kg in Tibetans from Aba. The proportion of subjects who had systolic blood pressure (SBP) higher than 140 was low (13.03%) in Lhasa Tibetans, and the proportion who had diastolic blood pressure (DBP) higher than 90 was low (15.52%) in Basin Han individuals. In

terms of bone density, osteoporosis was prevalent in Aba Tibetans (14.67%) compared with other populations.

Table 4 indicates the biochemical characteristics of the participants measured at baseline. The mean fasting blood glucose (FBG) was significantly low in Tibetans from Lhasa (4.53 mmol/L) and Tibetans from Aba (4.62 mmol/L). The mean HbA1C was between 5.69% and 6.01% across different ethnicities. With regard to triglycerides (TG), the highest TG (1.97 mmol/L) was measured among the Han ethnic group in the high plateau (Yunnan), while the lowest TG (1.00 mmol/L) was found among Tibetans in the high plateau of Aba.

Figure 3 presents the prevalence of self-reported diseases, which demonstrates the diverse disease patterns among different ethnic groups. The highest prevalence of hypertension was observed in the Bai (20.50%) and Yi (20.17%) ethnic groups in the high plateau, while the lowest was reported in Tibetans from Aba (11.41%). Diabetes was mostly prevalent among the Han ethnic group in Basin (6.52%). Regarding coronary heart disease, substantial disparity was observed. The most prevalent group was Tibetans from Lhasa (3.69%), while the least prevalent group was their rural counterpart, Tibetans from Aba (0.16%). Overall, cancer was not frequently reported, with the highest observation in the Bouyei ethnic group in Guizhou (1.37%).

-Table 2 here-

-Table 3 here-

-Table 4 here-

-Figure 3 here-

What are the main strengths and weaknesses?

This study has some limitations worth mentioning. First, we relied on self-reported data for some variables, such as lifestyle factors, which may incur recall bias. Second, we excluded populations aged outside of 30-79 (except for the Tibetan populations in which 18-30-year-

olds were also included), which may have ignored information on early life exposures. Despite the above limitations, our study is unique. First, to our knowledge, this is the first large-scale cohort study focusing on ethnic minority groups in China. Due to the substantial disparity in terms of social and economic development among the ethnic minority groups, as well as their distinct and diverse exposures and outcomes, the findings based on our cohort study have the potential to not only add evidence to the extant body of literature on the prevalence and risk factors of certain NCDs but also reveal novel knowledge that may serve as a reference to other global populations. Second, our study involved multi-ethnic individuals, involving mixed groups of urban and rural residents, plateau and basin inhabitants, populations living in highly concentrated and air polluted areas and those in remote mountainous areas. This would allow one to conduct comparative studies of the effects of various environmental exposures on people's health outcomes from different angles. For instance, Tibetan populations in rural areas (Aba prefecture) live a nomadic life, which is not shared by their urban counterparts (those in Lhasa). One thus will be able to investigate significant disparities of lifestyle even within the same ethnic group. Third, our baseline survey collected detailed and comprehensive information. For instance, in addition to demographics, socioeconomics, lifestyles, and health-related histories, we also collected biological specimens and performed medical examinations and clinical laboratory tests for every participant. Fourth, our study is funded by the National Key Research and Development Program. As the highest level research program in China, this study is strongly supported by the different levels of government, CDCs and local clinic centres. Therefore, related resources, field surveys, public cooperation and electronic medical records checks can be guaranteed and thus may improve the data quality and minimize the loss of participants at follow-up. Fifth, follow-up information will be obtained through annual data linkage to electronic records of disease and death registries, as well as re-interviews every 3-5 years. This follow-up strategy enables accurate and reliable tracking of both exposures and outcomes. Finally, various methods have been rigorously implemented at every step to ensure the quality of the data.

Can I get hold of the data? Where can I find out more?

To maximize the use of the CMEC data, we welcome collaboration from all over the world. Currently, the database is not accessible to be downloaded publicly because of some sensitive information. However, researchers interested in our study could contact the corresponding author at xingzhao@scu.edu.cn by providing specific ideas for more information.

Profile in a nut shell

- The China Multi-Ethnic Cohort (CMEC) is a community population-based prospective observational study aiming to address the urgent need for understanding NCD prevalence, risk factors and associated conditions in resource-constrained settings for ethnic minorities in China.

- A total of 99 556 participants aged 30 to 79 years (Tibetan populations include those aged 18 to 30 years) from the Tibetan, Yi, Miao, Bai, Bouyei, and Dong ethnic groups in Southwest China were recruited between May 2018 and September 2019.

- All surviving study participants will be invited for re-interviews every 3-5 years with concise questionnaires to review risk exposures and disease incidence. Furthermore, the vital status of study participants will be followed up through linkage with established electronic disease registries annually.

- The CMEC baseline survey collected data with an electronic questionnaire and face-to-face interviews, medical examinations and clinical laboratory tests. Furthermore, we collected biological specimens, including blood, saliva and stool, for long-term storage. In addition to the individual level data, we also collected regional level data for each investigation site.

- Collaborations are welcome. Please send specific ideas to corresponding author at: xingzhao@scu.edu.cn.

Funding

Acknowledgements

Conflict of interest: None declared.

References:

1.   Gowshall M, Taylor-Robinson SD. The increasing prevalence of non-communicable diseases in low-middle income countries: the view from Malawi. *Int J Gen Med* 2018; **11**: 255-64.

2.   collaborators NC. NCD Countdown 2030: worldwide trends in non-communicable disease mortality and progress towards Sustainable Development Goal target 3.4. *Lancet* 2018; **392**: 1072-88.

3.   Wang Yf, Lim H, Wu Y. Growing global burden of chronic noncommunicable diseases and an alarming situation in China. *Beijing Da Xue Xue Bao Yi Xue Ban* 2012; **44**: 688-93.

4.   Rembert N, He K, Judd SE, McClure LA. The geographic distribution of trace elements in the environment: the REGARDS study. *Environ Monit Assess* 2017; **189**: 84.

5.   Wu F, Guo Y, Chatterji S, et al. Common risk factors for chronic non-communicable diseases among older adults in China, Ghana, Mexico, India, Russia and South Africa: the study on global AGEing and adult health (SAGE) wave 1. *BMC Public Health* 2015; **15**: 88.

13

6.   Zhou M, Wang H, Zeng X, et al. Mortality, morbidity, and risk factors in China and its provinces, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* 2019; **394**: 1145-58.

7.   Liu X, Mao Z, Li Y, et al. Cohort Profile: The Henan Rural Cohort: a prospective study of chronic non-communicable diseases. *Int J Epidemiol* 2019; **48**: 1756-j.

8.   Wang X, Lu M, Qian J, et al. Rationales, design and recruitment of the Taizhou Longitudinal Study. *BMC public health* 2009; **9**: 223.

9.   Perini W, Snijder MB, Peters RJG, Kunst AE. Ethnic disparities in estimated cardiovascular disease risk in Amsterdam, the Netherlands : The HELIUS study. *Neth Heart J* 2018; **26**: 252-62.

10. Fang CY, Tseng M. Ethnic density and cancer: A review of the evidence. *Cancer* 2018; **124**: 1877-903.

11. Zhu Y, Sidell MA, Arterburn D, et al. Racial/Ethnic Disparities in the Prevalence of Diabetes and Prediabetes by BMI: Patient Outcomes Research To Advance Learning (PORTAL) Multisite Cohort of Adults in the U.S. *Diabetes Care* 2019; **42**: 2211-9.

12. Aarab R, Vijverberg SJH, Prins M, et al. Prevalence of and factors associated with adult-onset asthma in different ethnic groups: The HELIUS study. *Respir Med* 2019; **150**: 113-9.

13. *National Bureau of Statistics of China. Tabulation on the 2010 population census of the People's Republic of China.* [cited; Available from: http://www.stats.gov.cn/tjsj/tjgb/rkpcgb/2012.7.23

14. Wang L, Gao P, Zhang M, et al. Prevalence and Ethnic Pattern of Diabetes and Prediabetes in China in 2013. *JAMA* 2017; **317**: 2515-23.

15. Zhang X, He Y, Xie X, Ji M, Ma X, Yu Z. Distribution of hemoglobin and prevalence of anemia in 10 ethnic minorities in China. *Medicine* 2017; **96**.

16. Xu S, Jiayong Z, Li B, et al. Prevalence and Clustering of Cardiovascular Disease Risk Factors among Tibetan Adults in China: A Population-Based Study. *PLoS One* 2015; **10**: e0129966.

17. Qingli T. Study on characteristic components of fermented acidic liquid from miao-nationality in Guizhou provice [J]. *Science and Technology of Food Industry* 2005; **9**.

18. Bao Y, Sirijimodege. Suggestions on Promoting the Economic Development of Chinese Minorities in the New Era. *Journal of Financial and Economic Theory* 2019; **5**: 9.

19. West JB. High-altitude medicine. *Am J Respir Crit Care Med* 2012; **186**: 1229-37.

20. Organization WH. ICD-10: international statistical classification of diseases and related health problems: tenth revision. 2004.

21. Chen Z, Chen J, Collins R, et al. China Kadoorie Biobank of 0.5 million people: survey methods, baseline characteristics and long-term follow-up. *Int J Epidemiol* 2011; **40**: 1652-66.

Table 1. Summary of measurements at the baseline survey in the CMEC study

Measurements	No. of variables	Variables
Electronic questionnaire		
Demographics	21	National ID number, date of birth, sex, marital status, ethnic group, education, occupation, household income, household expenses, composition of household, insurance coverage, contact information, *etc*.
Smoking and indoor air pollution	20	Smoking status, smoking cessation, passive smoking exposure, pesticide exposure, indoor air pollution, *etc*.
Alcohol use	11	Frequency of drinking, types of drink, alcohol consumption, post-drinking symptoms, *etc*.
Tea and other beverages	12	Frequency of drinking tea, types of tea, frequency of drinking other beverages, types of beverage, *etc*.
Personal and family health	137	Self-rated health status (EQ-5D-5L), disease history, family disease history, defecation situation, gum situation, *etc*.
Physical activity	24	Job-related physical activity, transportation physical activity, leisure-time physical activity, housework, weight change during last 12 months, weight in early adulthood, *etc*.
Reproductive history (for women)	19	Age of menarche, menopause status, experience of pregnancy, history of breast feeding, history of contraceptive pills use, history of hysterectomy and of ovary/breast surgery, *etc*.
Habitual diet	139	Flavouring, food frequency questionnaire (FFQ) [a], nutritional supplements, history of severe food

Measurements	No. of variables	Variables
		shortage, history of refrigerator usage, spicy food intake, Sichuan pepper intake, *etc*.
Life events, social support and psychological conditions	28	Life satisfaction, traumatic events, sleep situation, napping, probable depression (PHQ-2), probable anxiety (GAD-2), social support and social capital, *etc*.
Medical examination		
Fasting condition	1	Fasting time
Medication history	2	Antihypertensive drugs, antiarrhythmic drugs
Anthropometric measurement	4	Weight, height, hip circumference, wrist circumference.
Physical examination	3	Fasting blood pressure (measured three times, OMRON HEM-8711), fasting heart rate (measured three times, OMRON HEM-8711), visual inspection (international standard visual acuity chart)
Lung function	1	Peak expiratory flow rate (KOKA PEF-3)
Osteopathic examination	1	Mineral density of anklebone (OSTEOKJ3000)
Electrocardiogram (ECG)	1	12-lead ECG
Abdominal ultrasonography	5 for male	Liver, gallbladder, spleen, pancreas, kidney
	8 for female	Uterus, ovaries and fallopian tubes (additional for female participants)

Measurements	No. of variables	Variables
Chest X-ray	2	Cardiovascular, lungs
Chronic high-altitude disease [b]	7	Chronic high-altitude disease scale (7 questions or symptoms)
Clinical laboratory test		
Routine blood	24	White blood cells (WBCs), red blood cells (RBCs), haemoglobin (HGB), platelets (PLTs), lymphocytes, monocytes, neutrophils, eosinophils, *etc*.
Fasting blood glucose	1	Fasting blood glucose (FBG)
Lipid levels	4	Triglycerides (TGs), cholesterol (CHOL), high-density lipoprotein cholesterol (HDL-CH) and low-density lipoprotein cholesterol (LDL-CH)
Hepatic function	5	Total protein (TP), albumin, globulin, alanine transaminase (ALT), aspartate amino transferase (AST), and gamma-glutamyl transpeptidase (GGT)
Kidney function	3	Creatinine, urea, and uric acid
Routine urine	12	Specific gravity, urobilinogen, urinary bilirubin, urinary ketone body, urine glucose, urinary protein, pH, urine nitrite, *etc*.
Alkaline phosphatase	1	Alkaline phosphatase (ALP)
Creatine kinase-MB	1	Creatine kinase-MB (CK-MB)

Measurements	No. of variables	Variables
Total bilirubin	1	Total bilirubin (TBIL)
Glycosylated haemoglobin	1	Glycated haemoglobin (HbA1C)
Biological specimens		
Blood sample	6	2 aliquots of serum (0.5 mL × 2), 2 aliquots of plasma (0.5 mL × 2), 1 aliquot of buffy coat (0.5 mL) and 1 aliquot of blood cells (0.5 mL)
Saliva	2	2 aliquots of saliva (1 mL × 2)
Stool	1	1 aliquot of stool (1~2.5 g × 2)
Regional level data		
Hourly air pollution data	6	Sulfur dioxide, nitrogen dioxide, nitric oxide, fine particulate matter with particle size below 10 microns (PM10), fine particulate matter with particle size below 2.5 microns (PM2.5), ozone (China National Environmental Monitoring Centre).
Daily meteorological data	9	Air pressure, air temperature, precipitation, evaporation, relative humidity, wind direction and wind speed, sunshine hours and 0 cm ground temperature elements (National Meteorological Information Centre).
Socioeconomic data	-	Specific variables selected according to research purposes (Yearbook data of national and provincial statistical departments)

Measurements	No. of variables	Variables
Health related data	-	Specific variables selected according to research purposes (Yearbook data of national and provincial health departments)
Geographic data	-	Specific variables selected according to research purposes (National Earth System Science Data Sharing Platform)

a. Includes common food items and ethnic group-specific food items.

b. Only available for high-altitude areas.

Table 2. Baseline characteristics of the participants in the CMEC study

Characteristics	Han ethnic group in Basin (n=44900)	Han ethnic group in Yunnan (n=10543)	Yi ethnic group in Yunnan (n=6283)	Bai ethnic group in Yunnan (n=6310)	Tibetans in Aba (n=4993)	Tibetans in Lhasa (n=7737)	Dong ethnic group in Guizhou (n=7239)	Bouyei ethnic group in Guizhou (n=5992)	Miao ethnic group in Guizhou (n=5559)
Gender, n(%)									
male	20478(45.61)	3696(35.06)	2088(33.23)	1869(29.62)	1927(38.59)	3222(41.64)	2594(35.83)	1852(30.91)	2067(37.18)
female	24422(54.39)	6847(64.94)	4195(66.77)	4441(70.38)	3066(61.41)	4515(58.36)	4645(64.17)	4140(69.09)	3492(62.82)
Ages(years), n(%) [a]									
<30	-	-	-	-	-	925(11.96)	-	-	-
30-39	8382(18.67)	1037(9.84)	784(12.48)	442(7)	1268(25.4)	1748(22.59)	850(11.74)	773(12.9)	869(15.63)
40-49	13564(30.21)	3112(29.52)	1805(28.73)	1759(27.88)	1614(32.33)	1870(24.17)	1977(27.31)	1783(29.76)	1627(29.27)
50-59	10413(23.19)	3605(34.19)	2148(34.19)	2124(33.66)	1216(24.35)	1934(25)	2248(31.05)	1897(31.66)	1572(28.28)
60-69	8499(18.93)	2110(20.01)	1157(18.41)	1440(22.82)	655(13.12)	912(11.79)	1484(20.5)	1091(18.21)	981(17.65)
≥70	4042(9)	679(6.44)	389(6.19)	545(8.64)	240(4.81)	348(4.5)	680(9.39)	448(7.48)	510(9.17)
Marital status, n(%)									

Married or cohabiting	39893(88.85)	9501(90.12)	5516(87.75)	5735(90.9)	4358(87.28)	6484(83.81)	6377(88.09)	5222(87.13)	4817(86.64)
Widowed	2466(5.49)	190(1.8)	127(2.02)	106(1.68)	224(4.49)	154(1.99)	180(2.49)	214(3.57)	200(3.6)
Separated or divorced	2019(4.5)	799(7.58)	553(8.8)	461(7.31)	235(4.71)	319(4.12)	636(8.79)	503(8.39)	476(8.56)
Never married	519(1.16)	53(0.5)	90(1.43)	7(0.11)	176(3.52)	780(10.08)	46(0.64)	54(0.9)	67(1.21)

Highest education completed, n(%)

No formal school	4787(10.66)	2876(27.28)	1569(24.96)	2300(36.45)	3565(71.4)	3797(49.08)	3603(49.77)	2210(36.88)	2454(44.14)
Primary school	9709(21.63)	3870(36.71)	2625(41.76)	2519(39.92)	1058(21.19)	2299(29.71)	1021(14.1)	1458(24.33)	761(13.69)
Middle school	14499(32.29)	2740(25.99)	1671(26.58)	1227(19.45)	192(3.85)	790(10.21)	1211(16.73)	1514(25.26)	1083(19.48)
High school	8073(17.98)	653(6.19)	279(4.44)	199(3.15)	53(1.06)	405(5.23)	596(8.23)	462(7.71)	584(10.5)
College or university	7829(17.44)	403(3.82)	142(2.26)	65(1.03)	125(2.5)	446(5.76)	808(11.16)	349(5.82)	678(12.19)

Annual income, whole family (Yuan/year), n(%)

<12000	4586(10.23)	2261(21.46)	1579(25.14)	979(15.56)	877(17.57)	1771(22.9)	2252(31.17)	1968(32.85)	1438(25.89)
12000-19999	5435(12.12)	2262(21.47)	1798(28.63)	1344(21.36)	1693(33.91)	2191(28.33)	1515(20.97)	1384(23.1)	925(16.65)
20000-59999	16080(35.87)	4570(43.38)	2378(37.87)	2652(42.16)	2038(40.83)	2627(33.97)	1944(26.91)	1784(29.78)	1797(32.35)
60000-99999	9480(21.14)	805(7.64)	345(5.49)	726(11.54)	246(4.93)	630(8.15)	794(10.99)	519(8.66)	824(14.83)

≥100000	9253(20.64)	636(6.04)	180(2.87)	590(9.38)	138(2.76)	514(6.65)	720(9.97)	336(5.61)	571(10.28)
Smoking history, n(%)									
No smoking	31650(70.51)	7839(74.35)	4941(78.62)	4827(76.51)	4556(91.27)	5724(73.99)	5679(78.46)	4933(82.34)	4322(77.73)
Smoking [b]	10273(22.89)	2388(22.65)	1169(18.6)	1258(19.94)	242(4.85)	1651(21.34)	1269(17.53)	863(14.4)	1030(18.53)
Quit smoking [c]	2964(6.6)	316(3)	175(2.78)	224(3.55)	194(3.89)	361(4.67)	290(4.01)	195(3.25)	208(3.74)
Drinking of alcohol, n(%)									
Never	21161(47.15)	7382(70.02)	4442(70.69)	5234(83.11)	4647(93.09)	5269(68.11)	4125(57)	2940(49.07)	2505(45.05)
Occasional [d]	18373(40.94)	1925(18.26)	1270(20.21)	669(10.62)	314(6.29)	2078(26.86)	2469(34.12)	2474(41.3)	2635(47.39)
Regular [e]	5347(11.91)	1236(11.72)	572(9.1)	395(6.27)	31(0.62)	389(5.03)	643(8.88)	577(9.63)	420(7.55)
Drinking of tea, n(%)									
Never	31092(69.28)	5804(55.06)	5308(84.5)	4406(69.96)	1613(32.31)	628(8.12)	6228(86.05)	5239(87.45)	4815(86.6)
Occasional [f]	2286(5.09)	483(4.58)	191(3.04)	117(1.86)	195(3.91)	406(5.25)	196(2.71)	138(2.3)	188(3.38)
Regular [g]	11503(25.63)	4255(40.36)	783(12.46)	1775(28.18)	3184(63.78)	6700(86.63)	814(11.25)	614(10.25)	557(10.02)
Total physical activity (MET hours/day), n(%) [h]									
<10	8043(17.99)	877(8.37)	435(6.94)	839(13.4)	1511(30.49)	2157(27.98)	993(13.8)	909(15.31)	851(15.45)

10~	15790(35.31)	1770(16.89)	841(13.42)	1581(25.25)	1288(25.99)	2096(27.19)	1643(22.83)	1406(23.68)	1268(23.02)
20~	10515(23.51)	2164(20.65)	1315(20.99)	1339(21.39)	755(15.23)	1661(21.55)	1587(22.05)	1318(22.2)	1162(21.09)
30~	5344(11.95)	2148(20.49)	1447(23.09)	932(14.89)	602(12.15)	872(11.31)	1259(17.49)	947(15.95)	892(16.19)
40~	5027(11.24)	3522(33.6)	2228(35.56)	1570(25.08)	800(16.14)	923(11.97)	1716(23.84)	1358(22.87)	1336(24.25)
Frequency of physical exercise, n(%)									
Never	16642(37.09)	7267(68.98)	4739(75.44)	3516(55.99)	2626(52.82)	4288(55.46)	5143(71.07)	4105(68.59)	4016(72.26)
Occasional [i]	8381(18.68)	716(6.8)	513(8.17)	927(14.76)	1271(25.56)	1710(22.12)	609(8.42)	498(8.32)	466(8.38)
Regular [j]	19852(44.24)	2552(24.22)	1030(16.4)	1837(29.25)	1075(21.62)	1734(22.43)	1485(20.52)	1382(23.09)	1076(19.36)
Reproductive history in women, n(%)									
Age at menarche <13 years	4218(17.28)	880(12.89)	233(5.55)	493(11.15)	159(5.21)	141(3.12)	380(8.19)	331(8.01)	312(8.95)
Age at first live birth <20 years	740(3.13)	325(4.85)	610(14.71)	163(3.72)	892(30.93)	390(9.88)	277(6.05)	202(4.96)	300(8.77)
Three or more live births	2338(9.73)	891(13.16)	637(15.33)	1151(26.28)	1708(59.04)	975(24.65)	1923(41.96)	1478(36.21)	1212(35.42)
Age at menopause <50	6070(52.92)	1841(51.6)	1075(50.76)	1351(55.87)	609(54.91)	838(57.32)	1648(62.85)	1188(56.22)	900(52.94)
Ever used contraceptive pill	5265(21.56)	464(6.79)	603(14.37)	173(3.91)	120(3.93)	621(13.77)	797(17.17)	631(15.28)	721(20.66)
High fat diet, n(%) [k]	31409(70.99)	8717(82.94)	4185(66.73)	5206(84.31)	3502(70.51)	4255(55.27)	4582(64.07)	3924(67.08)	3281(60.03)

High salt diet, n(%) [l]	12227(27.63)	4390(41.77)	1826(29.11)	1269(20.55)	1256(25.29)	3317(43.09)	3662(51.2)	2614(44.68)	2785(50.95)
Regular consumption of certain foodstuffs, n(%) [m]									
Rice	44603(99.42)	10501(99.82)	6267(99.76)	6177(99.87)	4606(92.64)	6950(90.05)	7037(97.25)	5949(99.73)	5411(97.44)
Pasta	25483(56.8)	5652(53.73)	1174(18.69)	3195(51.66)	2064(41.51)	5137(66.56)	2036(28.14)	3531(59.2)	2303(41.47)
Coarse Cereals	8686(19.36)	1100(10.46)	89(1.42)	322(5.21)	29(0.58)	480(6.22)	115(1.59)	131(2.2)	161(2.9)
Tubers	12706(28.32)	5442(51.73)	1780(28.33)	4822(77.96)	1537(30.91)	4458(57.76)	305(4.22)	330(5.53)	473(8.52)
Meat/poultry	40094(89.37)	7419(70.52)	3919(62.38)	4396(71.06)	4109(82.64)	5566(72.12)	6552(90.56)	5345(89.59)	4890(88.08)
Fish/sea food	1918(4.28)	380(3.61)	64(1.02)	321(5.19)	NA	36(0.47)	238(3.29)	97(1.63)	94(1.69)
Eggs	25207(56.19)	4425(42.06)	971(15.46)	2316(37.44)	551(11.08)	1499(19.42)	1321(18.26)	2149(36.01)	1326(23.88)
Fresh vegetables	44484(99.16)	10291(97.82)	6212(98.89)	6117(98.88)	4582(92.16)	7232(93.7)	7166(99.05)	5914(99.11)	5498(99.03)
Soybeans	7468(16.65)	1445(13.74)	519(8.26)	1066(17.23)	61(1.23)	482(6.25)	535(7.39)	1189(19.93)	596(10.73)
Fresh fruits	31865(71.03)	8669(82.4)	3423(54.49)	4860(78.56)	2698(54.26)	3563(46.16)	3894(53.82)	3455(57.91)	2453(44.18)
Dairy products	17012(37.92)	1354(12.87)	630(10.03)	1070(17.3)	2250(45.25)	1845(23.91)	391(5.4)	610(10.22)	470(8.47)
Spicy foods	28429(63.37)	8837(84)	3233(51.46)	4160(67.25)	1850(37.21)	2861(37.07)	6738(93.14)	5396(90.46)	4899(88.25)
Black pepper-seasoned foods	25582(57.02)	7677(72.99)	2430(38.68)	3306(53.44)	1257(25.28)	3205(41.53)	1909(26.39)	2612(43.79)	2261(40.73)

Sleeping status, n(%)

Difficulty initiating or maintaining sleep [n]	17300(38.57)	3057(29.08)	2225(35.42)	1606(26.25)	2236(44.97)	1807(23.41)	2694(37.24)	2194(36.79)	2107(37.96)
Early morning awakening [o]	13689(30.52)	2747(26.13)	2202(35.06)	1440(23.54)	2003(40.29)	1511(19.58)	2096(28.97)	1752(29.38)	1643(29.6)
Taking sleeping pills [p]	1048(2.34)	280(2.66)	211(3.36)	190(3.11)	1305(26.25)	164(2.12)	75(1.04)	115(1.93)	104(1.87)
Daytime dysfunction [q]	7430(16.56)	1636(15.56)	1250(19.9)	489(7.99)	1667(33.53)	717(9.29)	1024(14.16)	979(16.42)	966(17.4)
Noon break	27322(60.91)	3574(34)	1225(19.5)	737(12.05)	1329(26.73)	2074(26.87)	3525(48.73)	3089(51.79)	2378(42.84)

a. The inclusion criteria for age was 30-79 years on the day of investigation (except for the Tibetan population in Lhasa whose inclusion criteria for age was 18-79)

b. Smoking is defined as a total of more than 100 cigarettes smoked to date.

c. Quit smoking is defined as quitting for more than half a year.

d. Occasional drinking of alcohol is defined as drinking an average of 1-2 days per week or less in the past year.

e. Regular drinking of alcohol is defined as drinking an average of 3-5 days per week or more in the past year.

f. Occasional drinking of tea is defined as drinking an average of 1-2 days per week or less in the past year.

g. Regular drinking of tea is defined as drinking an average of 3-5 days per week or more in the past year.

h. Total physical activity includes job-related physical activity, transportation physical activity, leisure-time physical activity and housework.

i. Occasional physical exercise is defined as 1-2 times per week or less in the past year.

j. Regular physical exercise is defined as 3-5 times per week or more in the past year.

k. A high-fat diet is defined as an average daily oil intake above 25 g.

l. A high salt diet is defined as an average daily salt intake above 6 g.

m. Regular consumption of certain foodstuffs is defined as an average of 3 times per week or more in the past year.

n. Difficulty initiating or maintaining sleep is defined as at least 3 days per week in the past month.

o. Early morning awakening is defined as at least 3 days per week in the past month.

p. Taking sleeping pills is defined as at least 1 day per week in the past month.

q. Daytime dysfunction is defined as at least 3 days per week in the past month.

Table 3. Physical examinations of the participants in the CMEC study

Characteristics	Han ethnic group in Basin	Han ethnic group in Yunnan	Yi ethnic group in Yunnan	Bai ethnic group in Yunnan	Tibetans in Aba	Tibetans in Lhasa	Dong ethnic group in Guizhou	Bouyei ethnic group in Guizhou	Miao ethnic group in Guizhou
Height (cm)	158.47±8.28	158.25±7.54	157.09±7.7	159.27±7.7	158.62±8.29	161±8.57	155.35±7.77	155.14±7.45	154.42±7.94
Weight (kg)	61.67±10.47	56.93±9.49	58.26±10.64	58.32±9.89	65.95±13.44	65.82±10.78	57.81±10.3	58.08±10.08	59.65±10.11
BMI (kg/m2)									
< 18.5	832(1.89)	777(7.67)	351(5.71)	437(7.1)	60(1.25)	116(1.86)	310(4.66)	218(3.92)	110(2.15)
18.5~	19615(44.5)	6117(60.37)	3262(53.1)	3545(57.6)	1646(34.17)	1699(27.21)	3230(48.53)	2641(47.5)	1953(38.2)
24~	17535(39.78)	2684(26.49)	1846(30.05)	1753(28.48)	1625(33.73)	3400(54.44)	2339(35.15)	2006(36.08)	2114(41.35)
28~	6096(13.83)	555(5.48)	684(11.13)	420(6.82)	1486(30.85)	1030(16.49)	776(11.66)	695(12.5)	935(18.29)
Waist circumference (cm)	82.67±9.51	77.29±9.33	80.23±10.22	82.8±9.69	86.92±12.21	91.63±11.86	82.95±10.12	82.59±9.5	83.55±9.85
Hipline (cm)	93.75±6.69	89.35±6.07	91.88±6.65	92.97±6.33	97.58±9.14	100.07±8.86	90.96±6.34	92.55±6.32	92.23±6.39
Heart rate (/min) [a]	77.21±10.62	75.53±11.72	73.12±11.65	74.83±11.38	71.89±11.28	74.29±10.91	74.68±10.68	75.09±10.66	73.81±11.09
SBP (mmHg) [b]									

< 120	16628(37.71)	5208(51.38)	2126(34.6)	2290(37.16)	1889(39.31)	3720(59.71)	3009(45.25)	2538(45.78)	2198(43.19)
120-139	16938(38.41)	3308(32.64)	2274(37.01)	2298(37.29)	2045(42.55)	1698(27.26)	2282(34.32)	1778(32.07)	1763(34.64)
≥140	10533(23.88)	1620(15.98)	1745(28.4)	1574(25.54)	872(18.14)	812(13.03)	1358(20.42)	1228(22.15)	1128(22.17)
DBP (mmHg) [c]									
< 80	24697(56.01)	5671(55.94)	3149(51.25)	3414(55.4)	2726(56.8)	3464(55.65)	3476(52.25)	2579(46.45)	2450(48.11)
80-89	12555(28.47)	2771(27.34)	1719(27.98)	1721(27.93)	1277(26.61)	1637(26.3)	1894(28.47)	1720(30.98)	1497(29.39)
≥90	6843(15.52)	1695(16.72)	1276(20.77)	1027(16.67)	796(16.59)	1124(18.06)	1283(19.28)	1253(22.57)	1146(22.5)
Pulmonary function (L/min)	355.76±125.1	349.52±100.58	328.88±104.29	286.62±89.07	292.35±132.59	200.63±94.81	360.96±94.25	348.94±84.75	366.42±97.03
Bone density									
Normal	23976(54.38)	6421(63.32)	4330(70.45)	3331(54.06)	2922(60.74)	4043(64.84)	3995(59.98)	3503(62.85)	3035(59.3)
Osteopenia	17082(38.74)	3479(34.31)	1683(27.38)	2490(40.41)	1183(24.59)	1686(27.04)	2435(36.56)	1844(33.08)	1893(36.99)
Osteoporosis	3034(6.88)	240(2.37)	133(2.16)	341(5.53)	706(14.67)	506(8.12)	231(3.47)	227(4.07)	190(3.71)

a, b, c. Heart rate and systolic and diastolic blood pressure (SBP and DBP) were measured three times for each participant, and the statistics are based on the average of the three measurements.

Table 4. Biochemical characteristics of the participants in the CMEC study

Biochemical indexes	Han ethnic group in Basin	Han ethnic group in Yunnan	Yi ethnic group in Yunnan	Bai ethnic group in Yunnan	Tibetans in Aba	Tibetans in Lhasa	Dong ethnic group in Guizhou	Bouyei ethnic group in Guizhou	Miao ethnic group in Guizhou
FBG (mmol/l)	5.51±1.49	5.15±1.24	5.56±1.37	5.46±1.2	4.62±1.1	4.53±1.31	5.55±1.36	5.42±1.37	5.5±1.35
ALP (U/l)	86.66±29.39	84.08±27.73	83.99±29.01	90.24±28.48	114.71±43.52	113.24±37.92	92.56±32.68	74.27±23.77	81.19±26.88
TBIL (μmol/l)	12.65±7.1	11.45±5.72	10.17±5.05	12.03±6.33	12.73±8.5	13.86±6.9	9.92±5.11	10.07±4.56	10.27±5.36
HBA1C (%)	5.71±0.94	5.83±0.84	5.74±0.76	5.76±0.77	5.93±0.83	6.01±1.04	5.74±0.97	5.79±0.97	5.69±0.95
WBC (10^9/l)	5.87±1.67	6.64±2.04	6.41±1.75	6.43±1.68	5.73±1.58	6.18±1.72	6.56±1.77	6.15±1.65	6.49±1.8
RBC (10^12/l)	4.78±0.55	4.97±0.48	4.8±0.47	5.02±0.49	5.32±0.59	5.08±0.81	4.73±0.59	4.7±0.59	4.8±0.56
HGB (g/l)	144.19±16.44	150.63±15.12	147.8±14.72	151.55±15.46	160.29±21.22	154.45±28.46	140.75±16.45	137.35±16.21	143.61±16.07
PLT (10^9/l)	192.88±67.78	217.14±59.4	232.58±63.08	203.41±62.05	240.14±64.6	233.27±88.28	217.59±71.26	227±70.9	239.17±65.94
TG (mmol/l)	1.64±1.42	1.97±1.74	1.89±2.09	1.56±1.26	1±0.58	1.27±0.83	1.95±1.79	1.76±1.6	1.9±1.86
CHOL (mmol/l)	4.98±1	5.02±0.98	5.51±1.14	5.11±0.97	4.82±0.97	4.65±0.99	4.94±0.99	4.96±0.96	5.02±0.98
HDL-CH (mmol/l)	1.45±0.39	1.48±0.4	1.57±0.44	1.69±0.44	1.32±0.28	1.29±0.28	1.49±0.39	1.53±0.29	1.48±0.34
LDL-CH (mmol/l)	2.88±0.78	2.98±0.83	3.52±1	3.04±0.84	3.13±0.81	2.8±0.74	2.98±0.88	2.51±0.7	2.88±0.85

|									
TP (g/l)	76.26±4.55	75.56±4.64	76.22±4.75	73.98±4.34	77.58±5.07	76.54±5.31	79.22±4.62	77.51±4.53	78.58±5.36
ALT (U/l)	24.24±19.03	20.51±16.49	21.41±15.99	23.29±15.78	30.76±47.71	32.3±26.18	24.9±21.56	24.52±20.41	23.9±19.35
AST (U/l)	26.13±13.54	25.6±12.46	25.38±14.58	28±13.16	28.44±24.51	29.35±18.18	28.58±18.45	27.97±17.66	27.68±15.56
GGT (U/l)	34.21±49.62	39.7±63.11	47.2±92.03	38.78±54.15	42.28±46.62	52.38±90.48	43.85±80.96	38.75±53.1	41.07±68.46
Cr (µmol/l)	67.92±18.44	78.3±20.37	76.56±20.77	72.07±18.9	65.81±16.73	66.57±15.04	70.42±29.67	64.31±24.32	66.75±24.75
UREA (mmol/l)	5.46±1.51	5.06±1.48	5.26±1.57	5.16±1.5	5.2±1.5	4.7±1.38	5.18±1.67	5.29±1.64	5.08±1.66
UA (µmol/l)	322.25±86.65	318.55±87.74	299.82±87.69	286.09±83.08	316.72±89.8	334.31±89.25	337.6±96.57	315.7±91.51	336±95.32
CK-MB (U/l)	16.66±13.75	-	-	-	16.69±9.05	-	-	-	-
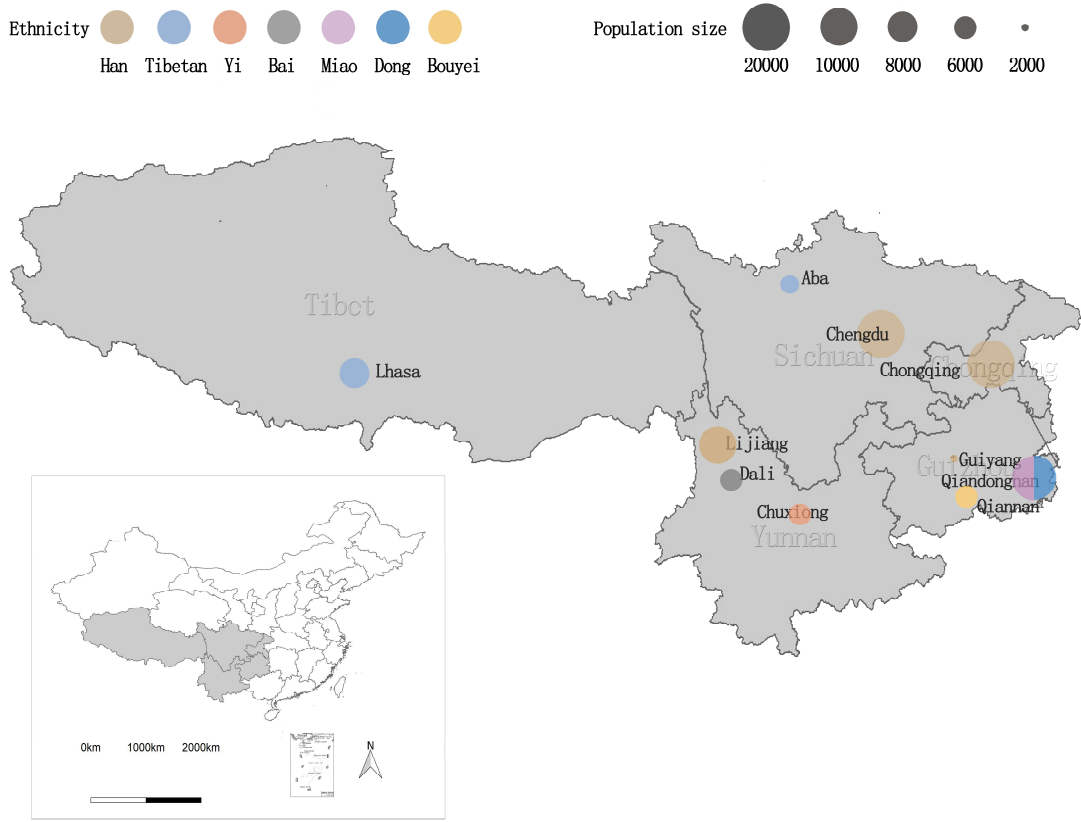
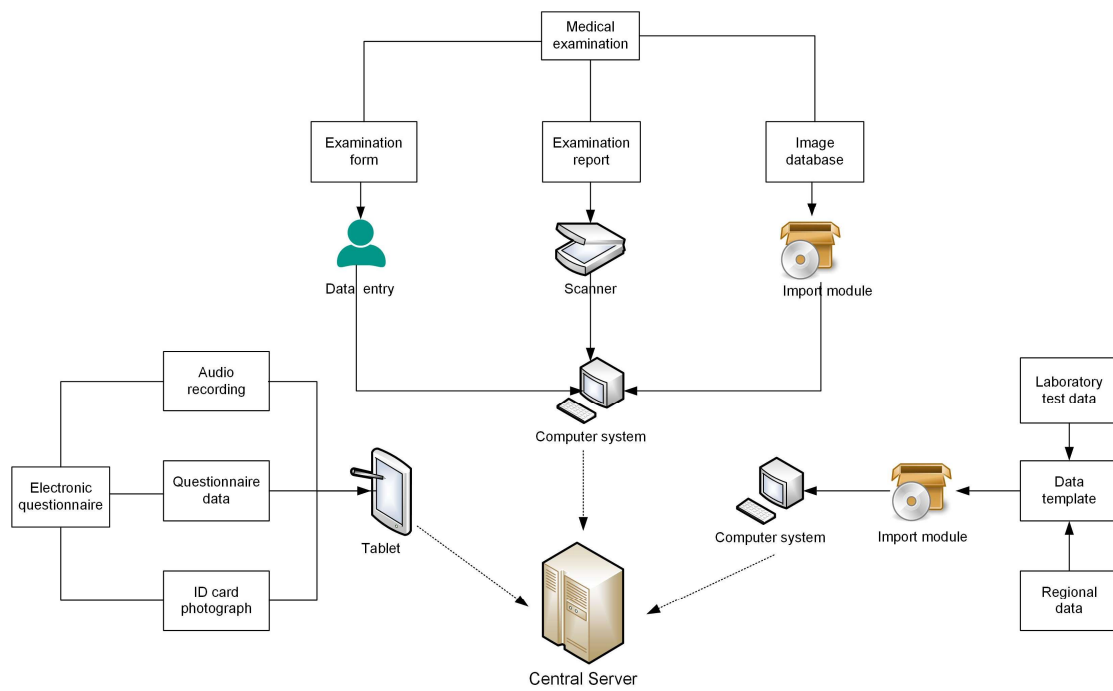**Figure 1.** Spatial distribution of baseline populations

Figure 2. Schematic graph of the data collection and computer system
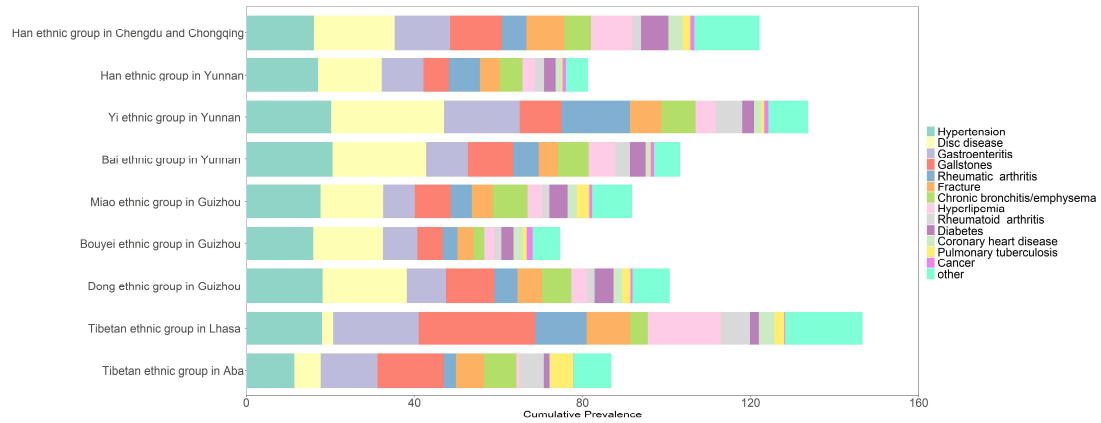
Figure 3. Prevalence of some selected self-reported diseases among the participants in different ethnic groups.