BMJ Open

Cohort profile: The Xinjiang Multiethnic Cohort (XMC) study

Luo Tao,[1] Tian Tian ![ORCID],[2] Lirong Liu,[1] Zewen Zhang,[1] Qi Sun,[3] Gaofeng Sun,[4] Jianghong Dai,[5] Hong Yan[6]

[1]Department of Epidemiology and Biostatistics, School of Public Health, Xinjiang Medical University, Urumqi, Xinjiang, China
[2]Shenzhen Campus of Sun Yat-sen University, School of Public Health (Shenzhen), Sun Yat-Sen University, Guangzhou, Guangdong, China
[3]Medical Research Design and Data Analysis Center, Academy of Traditional Chinese Medicine in Xinjiang Uygur Autonomous Region, Urumqi, China
[4]Urumqi Municipal Center for Disease Control and Prevention, Urumqi, China
[5]School of Public Health, Xinjiang Medical University, Urumqi, Xinjiang, China
[6]School of Public Health, Xi'an Jiaotong University, Xi'an, Shaanxi, China

Correspondence to
Dr Jianghong Dai;
epidjh@163.com

ABSTRACT

Purpose To investigate the potential causal link between heredity, geographical environment, diet and other lifestyle factors with long-term health consequences, we established the Xinjiang Multiethnic Cohort Study (XMC), the first large-scale prospective cohort in Xinjiang, China.

Participants XMC commenced in 2018 and enrolled participants from three study sites (Urumqi, Hotan and Ili) in Xinjiang, China. Data collected include standard baseline questionnaire, physical measurement, biological specimen. In addition, about one-third of participants were assessed habitual diet by a more detailed semiquantitative food frequency questionnaire which included 127 foods items at baseline

Findings to date Finally, a total of 30 949 participants, with 32.37% from Urumqi, 41.75% from Hotan, and 25.88% from Ili were recruited in XMC. The average age of participants was 56.21 years for men, and 54.75 years for women. More than 60% of participants in all three survey sites reported an average consumption of fruit and vegetable three or more times per week. In Hotan and Ili, the staple food was wheaten food, whereas, in Urumqi, rice and wheaten food was the food staples. Consumption of white meat, such as fish and poultry, was lower in the three survey locations. Based on self-reported disease from study participants, the five most common chronic diseases among participants across all three survey locations were dyslipidaemia, hypertension, cholecystitis, diabetes, ischaemic heart disease.

Future plans First, we will collect all health-related records of the study participants in January each year for the previous year. Second, 10% of subjects were randomly selected for telephone follow-up in the final year of cohort building. Finally, as planned, we will revisit the study subjects on site every 2–3 years. Again, we will conduct face-to-face questionnaires and collect biological specimens such as blood and urine from the study subjects.

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ The Xinjiang multiethnic cohort is the first population cohort study established in Xinjiang, China, broadly representative of the demographics of residents in Xinjiang.
⇒ The study collected information through face-to-face questionnaires and biological specimens such as blood and urine from the study subjects.
⇒ We have established a high-level biospecimen bank to manage our baseline samples, which will benefit our subsequent studies.
⇒ We have a comprehensive follow-up schedule that includes annual medical record follow-up and face-to-face visits approximately once every 2–3 years.
⇒ The study data are not freely available, but specific proposals for future collaborations are welcome.

INTRODUCTION

Over the past several decades, China has established several large-scale prospective population cohort studies. These have included the China Kadoorie Biobank (CKB) study,[1] a population study of 500 000 people among 10 provinces in China, and the Shandong Multi-centre Longitudinal Cohort for Health Management,[2] which included a research cohort of 100 million people and covered a study period of 12 years. The Taizhou Longitudinal Study[3] and Tianjin Chronic Low-Grade Systemic Inflammation and Health Cohort Study[4] have also been among the longitudinal cohort studies established during the same period. These cohort studies have focused on analysing the interactions between health-related risk factors and environmental and genetic factors with population health in China.[5] The main research outputs of these cohort studies have had a substantial impact on public health promotion in China. However, no prospective cohort studies have been carried out in the Xinjiang Uygur Autonomous Region (Xinjiang) (figure 1).

Xinjiang is China's largest provincial administrative region, accounting for one-sixth of China's landmass (1.66 million km^2). The Altai Mountains in the north of Xinjiang, the Kunlun Mountains in the south and the Tianshan Mountains in the middle divide the region into two distinct geographical areas; the Tarim Basin in the South and the Junggar Basin in the north. The topography of Xinjiang is complex, with geographical features varying from mountains and basins to grasslands, deserts and oases. Xinjiang has a multiethnic population with approximately
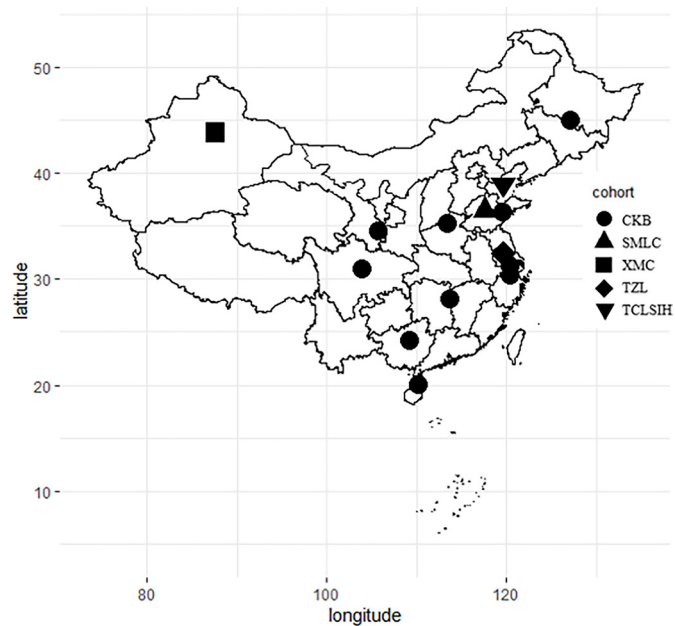
Figure 1 Locations of ongoing prospective cohorts in China. CKB, China Kadoorie Biobank; SMLC, Shandong multicentre Longitudinal Cohort; TCLSIH, Tianjin Chronic Low-Grade Systemic Inflammation and Health Cohort Study; TZL, Taizhou Longitudinal Study; XMC, Xinjiang Multiethnic Cohort.

Figure 2 Map of Xinjiang, China, showing the locations of three surveyed sites (red five-pointed StAR) in the XMC. XMC, Xinjiang Multiethnic Cohort.

20 million people. About 60% of the total population of Xinjiang belongs to Uyghur, Kazakh, Hui and other non-Han ethnic groups,[6] and a special diet may be correlated with health outcomes in the region.

The main risk factors associated with non-communicable chronic diseases are modifiable risk factors such as poor diet, lack of physical activity and alcohol and tobacco use. High-quality epidemiological data on diet, environmental and genetic determinants of non-communicable chronic diseases and long-term outcomes are essential for developing public health strategies to reduce the burden of non-communicable diseases. Results of statistical analyses carried out on surveillance data found that heart disease, malignant tumours and cardiovascular disease were the main contributors to the overall burden of chronic health conditions in 2015, and accounted for 77.67% of all deaths in Xinjiang.[7]

While extensive research on the aetiology of chronic diseases has been carried out across much of mainland China, the risk factors underpinning chronic disease may vary between populations. Substantial uncertainty exists as to how important these risk factors are in different settings. Because of the unique geographical environment, genetic background and population makeup of Xinjiang, the aetiologies of chronic disease deduced from other cohort studies carried out in China may not be fully applicable to the population of this region. The unique geographical environment, genetic background and population makeup also lead to the special eating habits of the residents in Xinjiang. The main objectives of this study were to identify diet risk factors and other

determinants of non-communicable diseases among the population of Xinjiang.

MATERIAL AND METHODS
Study design and organisation
The study was conducted in collaboration with the Xinjiang Medical University, the Academy of Traditional Chinese Medicine in Xinjiang Uygur Autonomous Region and the Urumqi Municipal Center for Disease Control and Prevention, and was completed under the guidance of Xi'an Jiaotong University.

The locations selected for inclusion in the cohort study, Urumqi, Hotan and Ili, were chosen by the geographical distribution of minority populations in the Xinjiang Uygur Autonomous Region (figure 2). Urumqi city is the capital of Xinjiang and is the cultural and political centre of the region. Surveys on factors such as health conditions, health-related behaviours and diet were carried out in 20 community health service centres in Urumqi to be representative of the urban population of Xinjiang. In Hotan, surveys were carried out in local villages where more than 95% of residents are of the Uyghur ethnicity. The highest level of educational attainment among this population is generally low and farming is the predominant occupation in this area. In addition to this, the population living in Hotan has distinctive dietary habits, the health outcomes of which may be possible to establish by comparing diet and health outcomes between this population and populations outside of this survey location. In Ili, surveys were conducted in townships where the distribution of Han, Uyghur, Kazak and Hui is relatively balanced. Most study participants recruited in this area were animal herders, the survey results, therefore, being reflective of health status and health-related risk factors of herdsmen in rural areas of Xinjiang. These diverse survey sites were selected to be reflective of the health status of both urban

and rural residents, workers or farmers and herders, and different ethnicities in Xinjiang.

Study participants

Our survey sites were set up in community health centres or village clinics with medical qualifications. For the selection of the study participants, we did not conduct a randomised sampling. We conducted extensive publicity campaigns, such as delivering and broadcasting introduction letters, to engage people in the study. And our baseline survey was conducted in tandem with the Xinjiang Universal Health Examination. Eligibility criteria for inclusion in the cohort study were adults aged 35–74 (born between 1943 and 1982), without any physical or communication disabilities (such as earning disabilities, language disorders and muscular dystrophy), with the ability to formally consent to participation in the study cohort, residing permanently (at least 1 year of residency) in study sites, and whose disease incidence data belonged to the local department of health. We excluded migratory populations and temporary residents, as these participants are more likely to be lost to follow-up. We also excluded armed forces and staff members stationed in the investigation site, as long-term follow-up is also challenging among this population. All participants gave written informed consent before the study began. First, the investigator introduced the participants to the information about Xinjiang Multiethnic Cohort (XMC). Second, the investigator asked the study participants if they would accept a questionnaire, a physical examination, a biospecimen collection and a search of their health information from a medical database. Finally, the subjects were given the option to withdraw at any time during the survey.

Patient and public involvement

Patients or the public were not involved in the design, recruitment, conduct, reporting or dissemination plans of our research. However, our researchers asked them whether they would be willing to participate in this study and they were told the approximate time required to participate, including length of survey response time and the number of potential years of involvement.

Follow-up survey

Beginning in September 2016, Xinjiang's government launched a comprehensive physical examination project within the region. All residents in Xinjiang can participate in this free annual physical examination, with the cost covered by the local government. The project will continue for decades. This free annual health examination for all residents in Xinjiang included but were not limited to physical examination, health-risk survey, biochemical blood examination, ultrasound and X-ray. During the past 4 years, Xinjiang has set up more than 2800 physical examination centres and invested a total of ¥3.90 billion in the Xinjiang Universal Health Examination. These universal health examinations are routinely attended by more than 80% of adults residing in the region. More than 50 million people have taken part in universal health examinations up to September 2018. Results of physical examination have formed a complete longitudinal database. After the cohort baseline survey, we can link with the database through individual identification (ID) numbers to follow our participants. We have matched the baseline survey data of XMC subjects with the personnel database of the national physical examination within Xinjiang in 2019. More than 70% of subjects in XMC participated in the national physical examination within Xinjiang in 2019.

Participants' follow-up will conduct by recording unique ID numbers associated with everyone across several common data systems. These data systems include medical record information system, medical insurance management system, maternal and child health information system, public security household registration management systems, civil affairs management system, chronic disease management information system, national central cancer registry of china and death registration information management system. Data from these different sources can be used concordantly for mutual review and for supplementing survey data.

In the final year of cohort building, 10% of subjects were randomly selected for telephone or face-to-face follow-up. Study participants given priority for this follow-up included those that had not had any measurable outcome during the study period. There are two main purposes for this strategy. First, we can reduce the loss of follow-up as much as possible. Second, we can check whether the medical database can cover all participants by comparing the results of telephone or face-to-face follow-up with the medical database matching.

As planned, we will revisit the study subjects on site every 2–3 years. Again, we will conduct face-to-face questionnaires and collect biological specimens such as blood and urine from the study subjects. We have completed a partial revisit of the study population in 2021.

Data and blood samples collection

The data collection methods applied for the XMC study included a questionnaire survey, physical examination, and collection of biological samples; measures which are essential for any longitudinal study. The questionnaire survey was conducted by medical students who had received consent training, and physical examinations were conducted by trained nurses or doctors.

The baseline questionnaire we used was mainly referred to as the baseline questionnaire of CKB[8] and has been slightly modified according to the opinion of experts from the medical colleges in Northwest China. The questionnaire collected information on socio-demographics, tea and coffee consumption, alcohol intake, tobacco use, dietary habits, passive smoking and indoor air pollution, personal and family medical history, physical activity, mental health and reproductive history (female study participants only; table 1). But our

Table 1 Questionnaire data collected in the Xinjiang Multiethnic Cohort study

Questionnaire survey	Description
Demographic data	Name, gender, ID card, medical insurance, date of birth, education level, marital status, occupation, financial income
Tea and coffee consumption	Frequency of tea drinking in the past year, previous tea drinking habits, types of tea drinking, frequency of coffee consumption
Alcohol intake	Frequency and amount of drinking in the past year, previous drinking habits (drinking refers to drinking in the past year under normal circumstances, not holidays or special periods, such as during marriage)
Tobacco use	Frequency and amount of smoking in the past year, previous smoking habits (current smoker defined as daily or almost daily smoking; former smoker defined as stopping smoking for at least 6 months; non-smoker defined as that the cumulative amount of smoking in lifetime does not exceed 100)
Dietary habits	Intake of 30 common foods, including staple foods, animal and plant foods, soy products, dairy products, vegetables and fruits; types of daily consumption oil; use of nutrient supplements; spicy food and vinegar consumption; household refrigerator use time
Passive smoking and indoor air pollution	Whether had history of live with current smoker; frequency of weekly exposure to passive smoking and cumulative exposure time; cooking, heating, household fuel use, and house decoration situation; occupational exposure history of air pollutants
Personal and family medical history	Self-health evaluation; chest and respiratory symptoms; history of 18 common diseases including diabetes, acute myocardial infarction, hypertension, asthma, and malignant tumours; history of blood transfusion and constipation; history of 5 common chronic diseases among family members
Physical activity	Work-related physical activity in the past year; the way of commuting and time spent on commuting; amateur physical exercise; housework activities; frequency and length of physical activity (ie, sweating, heartbeat, etc); weight loss in the past year
Mental health	Satisfaction with current living conditions; 10 major events that may have a serious impact on life in the past 2 years; sleep conditions (including insomnia, daily sleep time, whether to snoring while sleeping, etc); depression, anxiety disorders, and unexplained phobias; quality of life
Reproductive history (only for females)	Menstrual history of female (age at menopause and menarche), pregnancy, fertility, breast feeding, contraceptive use and surgery

collaborators conducted an additional dietary survey of the study participants in the Ili region, the data collected on dietary habits included specific information as to the types of food consumed, as well as the frequency and intake of specific food types. For example, information on the frequency and intake amount of fruit consumed, as well as specifics pertaining to fruit type, that is, number of apples, bananas and oranges consumed, was collected.

Data collected during the physical examination included height (using a medical height gauge with an accuracy of 0.1 cm), weight (using a medical electronic scale with an accuracy of 0.1 kg), waist circumference (using a soft measuring tape with an accuracy of 0.1 cm; the lower edge of the measuring tape was placed at the highest points on both sides of the hip bone and horizontal measurement of the waist circumference taken), heart rate and blood pressure (measured using a medical arm electronic blood pressure monitor with an accuracy of 2 mm Hg; two measurements per study participant were taken, following a rest period of 5–10 min), body fat composition (measured using a body composition analyzer, TANITA DC-430MA). Participants were asked to remove jackets, shoes, and hats while all physical measurements were being taken.

A vacuum blood collection device with intravenous anticoagulant was used to collect a 20 mL blood sample for each participant. A 4 mL blood sample was used for biochemical examination and routine blood examination. These examinations were completed at the nearest township health service centre to the survey location. Whole blood samples (3 mL) were transferred to three cryopreservation tubes immediately after blood samples were collected. Blood samples used to separate plasma and white blood cells were centrifuged within 2 hours of blood sample collection (4°C at 3000 rpm for 10 min). Samples were stored at −196°C liquid nitrogen containers before and after bimonthly transportation to Urumqi. RNA protection solution was also added to some samples for subsequent RNA detection. All the blood and urine samples are placed in a special biological sample bank, which is managed by a specially assigned person using an electronic management system.

For a large study such as this to be practicable and economically feasible in a resource-poor setting, study procedures needed to be simple and streamlined. In this study, the full assessment carried out at recruitment, including obtaining informed consent, administering the questionnaire survey, carrying out the

physical examination and obtaining blood samples, took an average of 60 min to complete. Investigators from all three survey locations were trained by the same junior investigator and postgraduate of the Xinjiang Medical University. The survey equipment was also consistent across all three survey locations and equipment calibration was carried out every day.

Statistical analyses

For the baseline profile, descriptive statistics were calculated for baseline data regarding demographic data, tea and coffee consumption, alcohol intake, tobacco use, dietary habits, passive smoking and indoor air pollution, personal and family medical history, physical activity, mental health, reproductive history (only for female). All analyses were conducted using SAS V.9.4.

RESULTS

When a participant came to the assessment centre and met our eligibility criteria, they were first given a serial number, and a total of 31 778 participants were given a number, but some of them dropped out during the questionnaire or refused to collect a biological specimen. Finally, a total of 30 949 participants were recruited as part of the XMC study from Urumqi, Hotan and Ili in Xinjiang, China. The number of participants recruited from Urumqi, Hotan and Ili was 10 017, 12 921 and 8011, respectively (table 2). The average age of study participants was 56.21 years for men, and 54.75 years for women. In Urumqi, more than half of the subjects were aged over 65 years old, significantly higher than the other two regions. According to the Census data of Xinjiang,[9] Urumqi has the highest proportion of people aged 65 and over compared with the other two regions. Also, our baseline survey was conducted on a weekday, but in urban areas, younger people are more likely to be required to work on a weekday. The proportion of Han Chinese in the Urumqi study population was close to 90%, while the proportion of Uyghur in the Hotan study population was over 99%. The proportion of ethnic groups represented in the study population recruited in the Ili region, conversely, was relatively balanced, indicating that ethnic composition varies significantly from region to region. The average level of highest educational attainment and average income among the Urumqi population was significantly higher than that of the Hotan and Ili populations. The proportion of people who had been educated to primary or pre-primary level in Hotan and Yili were 87.65% and 71.14%, respectively.

Approximately 50% of participants surveyed in Urumqi had an average annual income of more than 50 000, whereas about 60% of subjects in Hotan had an average annual income of less than 10 000. Study participants in Urumqi were found to have the highest average height, while study participants in Ili were found to have the highest average weight, with study participants in Hotan having the highest average waist circumference. The average body mass index (BMI) of the subjects in the three regions exceeded the diagnostic threshold (BMI ≥25 kg/m^2)[10] for overweight in China, with study participants in the Ili region having the highest average BMI.

The five most common chronic diseases among participants across all three survey sites were dyslipidaemia, hypertension, cholecystitis, diabetes, ischaemic heart disease with a prevalence of 34.55%, 32.66%, 14.34%, 10.07% and 8.74%, respectively (table 3).

Disease prevalence varies substantially across the three survey locations, which may be attributed to differences in nationality, diet, education, socioeconomics and living habits between the three areas. After standardising for age and sex based on 2010 China census data, the prevalence of hypertension in Ili was 36.09%, higher than that in Urumqi (29.29%) and Hotan (32.64%). According to a survey carried out between 2012 to 2015 the average prevalence of hypertension in China was 23%.[11] The prevalence of hypertension in the XMC study sites, as determined by this study, was much higher than the national average. Study participants in Urumqi had a higher prevalence of diabetes (19.14%) compared with those in Ili and Hotan. The results of this study also found that the prevalence of diabetes in Ili and Hotan was slightly lower than the national average (10.9%).[12] Urumqi has a high prevalence of diabetes and hypertension, which may be associated with participants residing in urban communities and having a higher average income compared with rural dwelling populations in Hotan and Ili. As a result, Urumqi's people have better health resources so people with diabetes are more likely to be diagnosed correctly. In addition, Urumqi's people have a lower proportion of the physical labour workforce than the other two regions, so that they may be sedentary in work time, which may lead to a higher prevalence of diabetes in the Urumqi population. The prevalence of chronic kidney disease (CKD) among participants in Hotan (6.28%) was approximately 10 times higher than that of Urumqi (0.56%) and 6 times that of Ili (0.93%). The prevalence of cholecystitis in Hotan (21.28%) was approximately seven times that of the prevalence in Urumqi (3.12%). Results of this study also found that the prevalence of COPD in Xinjiang (0.62%) was much lower than the national level (3.84%).[13] Regarding Cholecystitis, the data on the national average were limited. A previous epidemiological survey showed that the incidence of gallstones in Xinjiang was 11.83%.[14] The prevalence of Cholecystitis in Ili (12.49%) and Hotan (21.28%) was much higher than previously recorded levels. This may be associated with Hotan's local dietary habits. The prevalence of ischaemic heart disease, stroke, chronic hepatitis and CKD in Hotan was also found to be higher than that of the other two sites.

About baseline dietary intake (see figure 3), more than 60% of participants in all three survey locations reported an average consumption of fruit and vegetable three or more times per week. In Hotan and Ili, the staple food was wheaten food, whereas, in Urumqi, rice

Table 2 Demographic characteristics of the study participants by site

	Urumqi (n=10017)			Hotan (n=12921)			Ili (n=8011)		
	Men	Women	Total	Men	Women	Total	Men	Women	Total
Age (years)									
<45	231 (5.87)	330 (5.43)	561 (5.60)	1040 (21.17)	2337 (29.18)	3377 (26.14)	1120 (29.60)	1250 (29.57)	2370 (29.58)
45–54	570 (14.47)	1021 (16.80)	1591 (15.88)	1408 (28.66)	2871 (35.85)	4279 (33.12)	1248 (32.98)	1537 (36.36)	2785 (34.76)
55–64	824 (20.92)	1358 (22.34)	2182 (21.78)	1364 (27.76)	1848 (23.08)	3212 (24.86)	870 (22.99)	994 (23.52)	1864 (23.27)
65	2313 (58.74)	3370 (55.44)	5683 (56.73)	1101 (22.41)	952 (11.89)	2053 (15.89)	546 (14.43)	446 (10.55)	992 (12.38)
Ethnicity									
Han	3457 (87.79)	5432 (89.36)	8889 (88.74)	6 (0.12)	6 (0.07)	12 (0.09)	572 (15.12)	522 (12.35)	1094 (13.66)
Hui	310 (7.87)	392 (6.45)	702 (7.01)	4 (0.08)	15 (0.19)	19 (0.15)	1336 (35.31)	1333 (31.54)	2669 (33.32)
Uyghur	105 (2.67)	168 (2.76)	273 (2.73)	4897 (99.67)	7969 (99.51)	12866 (99.57)	908 (24.00)	1361 (32.20)	2269 (28.32)
Kazakh	41 (1.04)	53 (0.87)	94 (0.94)	6 (0.12)	13 (0.16)	19 (0.15)	883 (23.34)	951 (22.50)	1834 (22.89)
Other	25 (0.63)	34 (0.56)	59 (0.59)	NA	5 (0.06)	5 (0.04)	85 (2.25)	60 (1.42)	145 (1.81)
Education									
Primary/less	1064 (27.02)	2534 (41.68)	3598 (35.92)	4093 (83.31)	7232 (90.31)	11325 (87.65)	2582 (68.23)	3117 (73.74)	5699 (71.14)
Secondary	2320 (58.91)	3020 (49.68)	5340 (53.31)	774 (15.75)	734 (9.17)	1508 (11.67)	1166 (30.81)	1071 (25.34)	2237 (27.92)
Degree or above	553 (14.04)	517 (8.50)	1070 (10.68)	32 (0.65)	18 (0.22)	50 (0.39)	28 (0.74)	26 (0.62)	54 (0.67)
Refused/missing	1 (0.03)	8 (0.13)	9 (0.09)	14 (0.28)	24 (0.30)	38 (0.29)	8 (0.21)	13 (0.31)	21 (0.26)
Marriage									
Married	3696 (93.85)	5045 (82.99)	8741 (87.26)	4319 (87.91)	6165 (76.99)	10484 (81.14)	3500 (92.49)	3586 (84.84)	7086 (88.45)
Widowed/separated	197 (5.00)	1006 (16.55)	1203 (12.01)	538 (10.95)	1773 (22.14)	2311 (17.89)	227 (6.00)	614 (14.53)	841 (10.50)
Never married	42 (1.07)	19 (0.31)	61 (0.61)	19 (0.39)	6 (0.07)	25 (0.19)	43 (1.14)	6 (0.14)	49 (0.61)
Refused/missing	3 (0.08)	9 (0.15)	12 (0.12)	37 (0.75)	64 (0.80)	101 (0.78)	14 (0.37)	21 (0.50)	35 (0.44)
Income (RMB)*									
Under 10000	632 (16.05)	912 (15.00)	1544 (15.41)	2689 (54.73)	5139 (64.17)	7828 (60.58)	414 (10.94)	616 (14.57)	1030 (12.86)
10000–19999	412 (10.46)	585 (9.62)	997 (9.95)	1196 (24.34)	1649 (20.59)	2845 (22.02)	1311 (34.65)	1712 (40.50)	3023 (37.74)
20000–34999	411 (10.44)	737 (12.12)	1148 (11.46)	453 (9.22)	501 (6.26)	954 (7.38)	1157 (30.58)	1093 (25.86)	2250 (28.09)
35000–49999	504 (12.80)	885 (14.56)	1389 (13.87)	148 (3.01)	198 (2.47)	346 (2.68)	393 (10.39)	339 (8.02)	732 (9.14)
50000 or above	1978 (50.23)	2952 (48.56)	4930 (49.22)	416 (8.47)	500 (6.24)	916 (7.09)	503 (13.29)	459 (10.86)	962 (12.01)
Refused/missing	1 (0.03)	8 (0.13)	9 (0.09)	11 (0.22)	21 (0.26)	32 (0.25)	6 (0.16)	8 (0.19)	14 (0.17)
Anthropometric index									
Height	169.03±6.26	157.98±19.33	162.33±16.47	163.32±7.29	153.60±6.81	157.30±8.44	167.13±6.98	155.98±7.05	161.24±8.95
Weight	73.82±10.69	63.10±11.46	67.31±12.33	66.71±12.38	59.63±12.04	62.32±12.64	72.24±12.06	65.20±12.33	68.52±12.69

Continued

Table 2 Continued

	Urumqi (n=10017)			Hotan (n=12921)			Ili (n=8011)		
	Men	Women	Total	Men	Women	Total	Men	Women	Total
Waist circumference	92.40±17.00	87.09±15.47	89.18±16.30	91.88±11.50	91.44±12.29	91.61±12.00	90.56±11.02	88.49±11.80	89.46±11.48
Body mass index	25.81±3.31	25.35±4.42	25.53±4.02	24.96±4.04	25.24±4.67	25.14±4.44	25.82±3.75	26.76±4.51	26.31±4.20

N (%) for categorical variable; mean±SD for continuous variables.
*RMB1=US$0.14524 and RMB1=€0.1277 in December 2018.
NA, not available.

and wheaten food was the food staples. Dietary habits in Hotan, which has a predominantly Uyghur population, consisted predominantly of mutton, with more than 70% consuming mutton three or more times per week. The proportion of the population who consumed pork, mutton, or beef three or more times per week in Urumqi by contrast, where the population is predominantly Han, was 28.26%, 20.24 and 24.28%, respectively. Consumption of white meat, such as fish and poultry, was lower in the three survey locations. Fresh meat consumed by Xinjiang residents consisted mainly of red meat such as beef and mutton. Some studies have suggested that too high an intake of red meat might increase the risk of metabolic syndrome.[15 16] The study also showed that consumption of fresh milk was associated with a lower risk of Metabolic Syndrome.[16] In this study, the proportion of people drinking milk three or more times per week was lower in Ili and particularly in Hotan (<10% of participants surveyed). More than 60% of people in Ili reported drinking milk tea ≥3 times a week, however, with few of them drinking fresh milk directly.

A diet high in salt and fat is associated with an increased risk of high blood pressure and ischaemic heart disease.[17] The consumption of coarse grains, eggs, milk and yoghurt three or more times per week among people surveyed in Urumqi was higher than among people surveyed in Ili and Hotan. In addition to this, consumption of fowl, fish or seafood, and beans three or more times per week was low in all three survey locations (<10% of participants surveyed). The results of the survey on dietary habits reported here are generally consistent with a recent survey,[18] which found that consumption of cereals, meat, and oil in Xinjiang exceeded maximum recommended allowances, while consumption of eggs, seafood products, dairy products, nuts and legumes were below the minimum recommended allowances. Notably, this study reported higher consumption of fruit and vegetable in all three regions than previous studies on fruit and vegetable consumption in Xinjiang.[16 19]

DISCUSSION

This is the largest prospective study that has been undertaken to date in Xinjiang, China. The main strength of the XMC study is the inclusion of over 20 000 survey participants from diverse ethnic backgrounds such as Uyghur, Kazakh and Hui, which allows prevalence of chronic conditions, as well as dietary habits, various health determinants and sociodemographics between distinct populations to be compared. It is already well established that the dietary habits of Uyghur, Kazakh and Hui differ from those of Han Chinese. We have investigated the dietary habits of all the study participants as part of this cohort study, in conjunction with prevalence of non-communicable diseases to compare health determinants and prevalence of chronic conditions between these three survey locations.

Table 3 Self-reported disease prevalence among subjects in different sites (%)

	Urumqi		Hotan		Ili		Total	
Diseases	Crude	Adjusted	Crude	Adjusted	Crude	Adjusted	Crude	Adjusted
Cardiovascular diseases								
Hypertension*	47.19	29.29	37.00	32.64	38.61	36.09	40.74	32.66
Hypertension	32.64	17.87	27.35	23.76	16.71	15.23	26.31	19.80
IHD	8.18	3.28	13.54	12.15	8.44	7.72	10.48	8.74
Other heart disease	0.57	0.32	2.63	2.34	0.69	0.62	1.46	1.27
Stroke	3.43	1.62	4.91	4.31	1.61	1.47	3.58	2.82
Metabolic disease and kidney diseases								
Diabetes*	23.93	19.14	7.26	6.74	9.80	9.19	13.11	10.07
Diabetes	14.73	8.81	4.25	3.87	4.69	4.36	7.76	5.54
Dyslipidaemia*	35.00	35.08	33.53	34.05	34.83	35.08	34.32	34.55
CKD	0.68	0.56	6.20	6.28	0.99	0.93	3.06	3.17
Anaemia*	3.31	5.30	5.89	5.55	5.30	5.26	4.98	5.44
Chronic respiratory diseases								
Chronic bronchitis	4.48	2.16	15.89	14.17	3.31	3.12	8.94	7.87
Emphysema	0.29	0.15	3.00	2.57	0.51	0.49	1.48	1.25
COPD	0.40	0.17	1.22	1.14	0.20	0.18	0.69	0.62
Asthma	1.11	0.65	2.86	2.45	0.22	0.20	1.61	1.36
Tuberculosis	0.69	0.57	2.84	2.55	0.54	0.50	1.55	1.37
Digestive diseases								
Chronic hepatitis	0.49	0.53	4.85	4.91	1.34	1.33	2.53	2.82
Cholecystitis	6.50	3.12	23.28	21.28	13.66	12.49	15.36	14.34
Peptic ulcer	0.91	0.74	6.61	6.33	2.11	2.01	3.60	3.61
Skeletal disorders								
Osteoporosis	2.15	0.83	7.48	6.50	3.74	3.52	4.79	4.21
Fracture	1.93	1.24	5.82	6.10	5.75	5.87	4.54	4.96
Cancer	0.63	0.32	0.67	0.54	0.46	0.43	0.60	0.48

Adjusted: age and gender adjusted based on 2010 China census data.
Other heart disease, included rheumatic heart disease and cor pulmonale.
*Self-reported combined with biological detection.
CKD, chronic kidney disease; COPD, Chronic obstructive pulmonary disease; IHD, ischaemic heart disease.

Another important strength of this study was the survey questionnaire, which collected details on major risk factors such as alcohol intake, tobacco use (we investigated current as well as past usage quantity, frequency, type), mental health and quality of life. The three survey locations selected for this study were broadly representative of the different populations residing in Xinjiang, including animal herders, and urban and rural residents. The results of the survey conducted here can be used to compare differences in prevalence and type of non-communicable chronic diseases among these three populations.

There were also several limitations with the XMC study. First, we did not use random sampling to select study participants, but rather a convenience sampling strategy whereby eligible study participants from predetermined survey locations were recruited until an adequate sample size had been met. We excluded those who had lived in the local area for less than 1 year, which may reduce how representative the study population is for Urumqi regions. However, for the Ili and Hotan regions, the population is more stable and may not have an impact on the representation of the population. As this study is the first large-scale cohort study to have been carried out in Xinjiang, and data collected as part of previous cross-sectional health surveys of Xinjiang residents have not yet been published, we cannot compare differences in the distribution of diseases and related factors between our cohort study population and the wider Xinjiang population. Another limitation is that diseases and exposures were self-reported. Although a detailed and comprehensive questionnaire was carried out, some results are subject to reporting and recall bias.
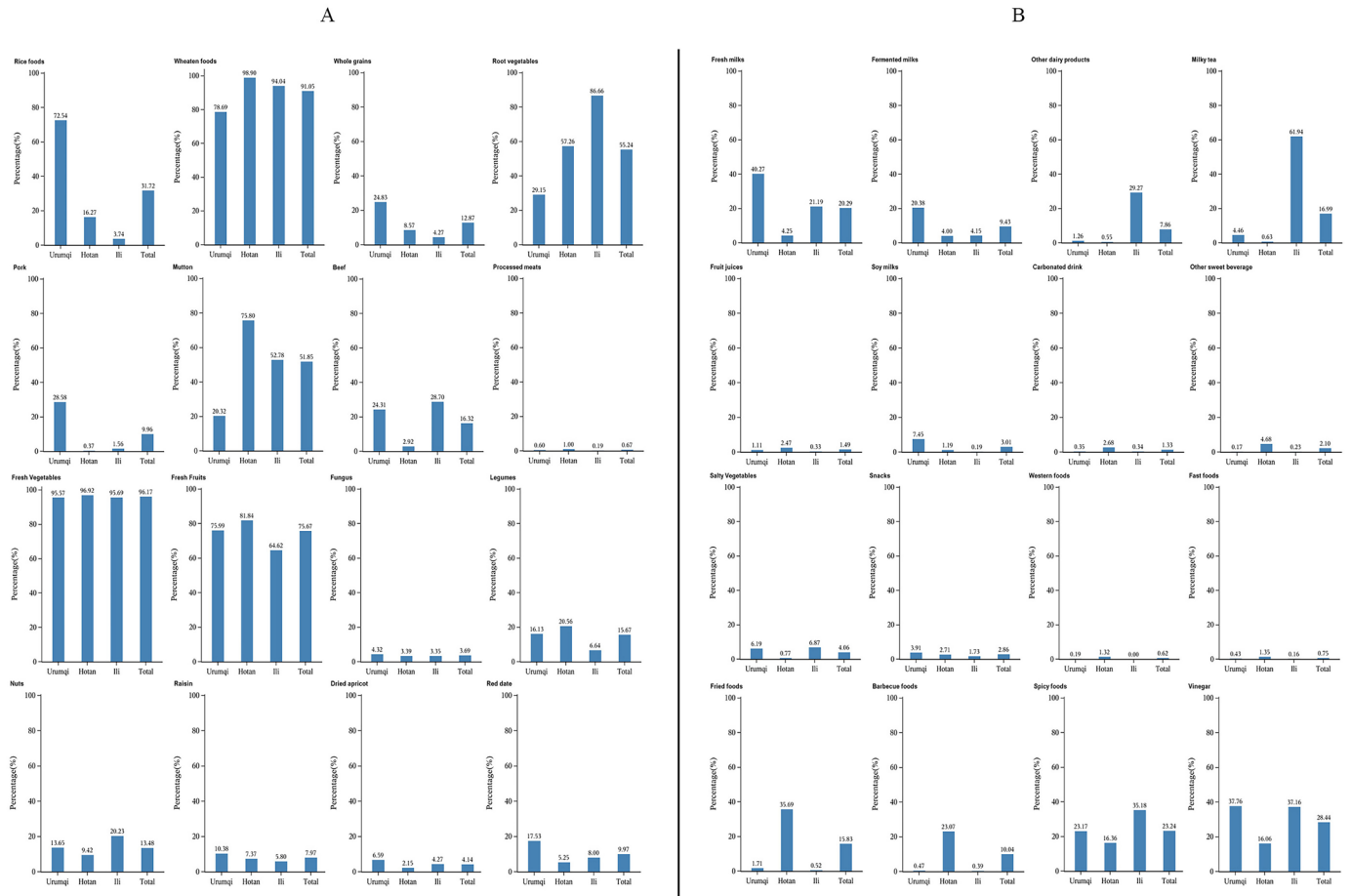
Figure 3 Percentage of common food intake in different sites in the XMC. XMC, Xinjiang Multiethnic Cohort.

We will take advantage of the existing monitoring systems or database of Xinjiang, China to obtain information on health outcomes (such as morbidity and death events) for all participants of the cohort study. Repeated cross-sectional surveys of study participants will be conducted on a biannual basis with surveys focusing on research outcome being the primary concern. Data will be obtained by administering questionnaires used for baseline research, as well as additional information on health determinants. Although the study population is relatively large, the availability of repeated measures, extensive biobank blood and urine samples, and health information systems linked by an ID number will allow this study to assess within-person and between-person variability over time in major risk factors among the population of people living in Xinjiang, China.

Map disclaimer The inclusion of any map (including the depiction of any boundaries therein), or of any geographic or locational reference, does not imply the expression of any opinion whatsoever on the part of BMJ concerning the legal status of any country, territory, jurisdiction or area or of its authorities. Any such expression remains solely that of the relevant source and is not endorsed by BMJ. Maps are provided without any warranty of any kind, either express or implied.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Consent obtained directly from patient(s)

Ethics approval This study involves human participants and was approved by Ethics Committee of Academy of Traditional Chinese Medicine in Xinjiang Uygur Autonomous Region (2018XE0108).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request. The study data are not freely available, but specific proposals for future collaborations are welcome. Address to the research leader of Xinjiang Multiethnic Cohort study.

ORCID iD
Tian Tian http://orcid.org/0000-0002-6103-1666

REFERENCES

1 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.

2 Zhang T, Zhang C, Zhang Y, *et al*. Metabolic syndrome and its components as predictors of nonalcoholic fatty liver disease in a northern urban Han Chinese population: a prospective cohort study. *Atherosclerosis* 2015;240:144–8.

3 Wang X, Lu M, Qian J, *et al*. Rationales, design and recruitment of the Taizhou longitudinal study. *BMC Public Health* 2009;9:223.

4 Song K, Du H, Zhang Q, *et al*. Serum immunoglobulin M concentration is positively related to metabolic syndrome in an adult population: Tianjin chronic low-grade systemic inflammation and health (TCLSIH) cohort study. *PLoS One* 2014;9:e88701.

5 Li L, Lyu J. [Large prospective cohort studies: a review and update]. *Zhonghua Liu Xing Bing Xue Za Zhi* 2015;36:1187–9. doi:10.3760/cma.j.issn.0254-6450.2015.11.001

6 He J, Guo S, Liu J, *et al*. Ethnic differences in prevalence of general obesity and abdominal obesity among low-income rural Kazakh and Uyghur adults in far Western China and implications in preventive public health. *PLoS One* 2014;9:e106723.

7 ZHE W, PUERHATI W, LIU L. Analysis of the death causes of residents in 15 surveillance sites in Xinjiang in 2015. *Modern Preventive Medicine* 2017;44:4186–90.

8 Li L-ming, Lv J, Guo Y, *et al*. [The China Kadoorie Biobank: related methodology and baseline characteristics of the participants]. *Zhonghua Liu Xing Bing Xue Za Zhi* 2012;33:249–55.

9 Office of the Leading Group of the Xinjiang Uygur Autonomous Region for the Seventh National Population Census, 2021Statistic Bureau of Xinjiang Uygur Autonomous Region. Available: Communiqué of the seventh national population census of the Xinjiang Uygur autonomous region

10 Reynolds K, Gu D, Whelton PK, *et al*. Prevalence and risk factors of overweight and obesity in China*. *Obesity* 2007;15:10–18.

11 Wang Z, Chen Z, Zhang L, *et al*. Status of hypertension in China: results from the China hypertension survey, 2012-2015. *Circulation* 2018;137:2344–56.

12 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.

13 Chan KY, Li X, Chen W, *et al*. Prevalence of chronic obstructive pulmonary disease (COPD) in China in 1990 and 2010. *J Glob Health* 2017;7:020704.

14 Zhu L, Aili A, Zhang C, *et al*. Prevalence of and risk factors for gallstones in Uighur and Han Chinese. *World J Gastroenterol* 2014;20:14942–9.

15 de Oliveira Otto MC, Alonso A, Lee D-H, *et al*. Dietary intakes of zinc and heme iron from red meat, but not from other sources, are associated with greater risk of metabolic syndrome and cardiovascular disease. *J Nutr* 2012;142:526–33.

16 Guo H, Gao X, Ma R, *et al*. Prevalence of metabolic syndrome and its associated factors among multi-ethnic adults in rural areas in Xinjiang, China. *Sci Rep* 2017;7:17643.

17 Zhai FY, Du SF, Wang ZH, *et al*. Dynamics of the Chinese diet and the role of urbanicity, 1991-2011. *Obes Rev* 2014;15 Suppl 1:16–26.

18 Yin J, Yang D, Zhang X, *et al*. Diet shift: considering environment, health and food culture. *Sci Total Environ* 2020;719:137484.

19 Zhai F, He Y, Wang Z, *et al*. [Status and characteristic of dietary intake of 12 minority nationalities in China]. *Wei Sheng Yan Jiu* 2007;36:539–41.