Cohort Profile: The Guangzhou Biobank Cohort Study, a Guangzhou-Hong Kong-Birmingham collaboration

Chaoqiang Jiang[3], G Neil Thomas[1], Tai Hing Lam[1], C Mary Schooling[1], Weisen Zhang[3], Xiangqian Lao[1], Peymane Adab[2], Bin Liu[3], Gabriel M Leung[1], Kar Keung Cheng[2],

[1]Departments of Community Medicine, The University of Hong Kong, Hong Kong SAR, People's Republic of China

[2]Department of Public Health and Epidemiology, University of Birmingham, Birmingham B15 2TT, UK

[3]Guangzhou Occupational Disease Prevention and Treatment Centre, Guangzhou, People's Republic of China

All correspondence: Tai Hing Lam

Department of Community Medicine

5/F Academic and Administration Block

Faculty of Medicine Building

University of Hong Kong

21 Sassoon Road, Pokfulam,

Hong Kong

Fax: (852) 2855 9528

E-mail: conmed@hkucc.hku.hk

Words 2989

How did the study come about?

The Guangzhou Biobank Cohort Study (GCBS), initiated by TH Lam (Hong Kong), CQ Jiang (Guangzhou) and KK Cheng (Birmingham), has been in part modelled on the Kadoorie Study of Chronic Disease in China (KSCDC)[1] and the UK Biobank.[2] Initial seed funding came from the Universities of Hong Kong and Birmingham, Guangzhou 12th Hospital (Guangzhou Occupational Disease Prevention & Treatment Centre), Guangzhou Public Health Bureau and Guangzhou Science and Technology Bureau. The Clinical Trials Service Unit at Oxford University has provided technical advice, especially in the planning and setting up of the project.

Whereas the KSCDC recruits adults aged 35 to 74 years nationwide from five rural counties and five urban districts, the Guangzhou Biobank Cohort Study focuses on older people aged at least 50 years in a mega-city of approximately 10 million. Of these, 6.4 million are permanent residents with locally registered households and the remainder are mostly migrants from other parts of the country. Guangzhou, the provincial capital of Guangdong province in southern China, is one of the most economically developed regions of China (figure 1). Uniquely, the lifetime experience of older urban residents in Guangzhou has straddled two macro-environments. Their parents were born into a pre-industrial environment, where living standards had been essentially unchanged for millennia[3;4] and where male heights (an anthropometric marker of living standards) were similar to those in France at their pre-industrial nadir.[5-10] These circumstances largely continued until after the establishment of the People's Republic of China (PRC) in 1949, although there were some isolated pockets of transient, haphazard economic growth along the eastern seaboard and around the treaty ports of Shanghai and to a lesser extent Guangzhou during the first half of the 20th century.[11;12] However, with the more inclusive route to economic development taken by the PRC, and in

recent decades following the establishment of the special economic zones close to Hong Kong in 1978, Guangzhou has been transformed and its residents' lives with it. Inevitably, this very rapid economic progress has been neither smooth nor comprehensive. There have been periods of significant hardships, such as during the Japanese occupation in the 1940s, the Great Leap Forward in the late 1950s and the Cultural Revolution in the 1960s.[13] There are also inequalities, and one specifically relevant to this study is the current difficulty of access to medical care after the large-scale privatisation of health care facilities since the 1980s.[14] Nevertheless, despite rapid economic transformation, traditional aspects of Chinese culture and social norms have been retained, such as more communitarian values and a Confucian respect for authority, and specific dietary and lifestyle habits, such as low use of alcohol and low female smoking rates.[15]

In this setting, where the social and historical context can be clearly identified and which has a homogeneous population who share a Cantonese heritage with a common dialect, culture and cuisine, our main objective is to examine prospectively the determinants of health in a population in social transition over the life course. The older residents of Guangzhou provide a unique opportunity to gain an early insight into the effects of the epidemiological transition, where our findings will be relevant to other parts of China transitioning later or at a slower rate, such as the inland industrial cities or the rural areas in the KSCDC, and also to other parts of the world currently undergoing rapid economic development.

What does it cover?

We plan to examine genetic, lifestyle, occupational and environmental factors and life course causes of the common chronic diseases, which are emerging with economic development. In doing so we would take advantage of natural socio-economic experiments provided by

3

macro-environmental exposures unique to this cohort, some of which are outlined in the previous section. The geo-ethnic homogeneity of our cohort helps to reduce the risk of population stratification in nested case control studies on genetic association. In addition we intend to leverage the concept of Mendelian randomisation[16] provided by the distribution of common genetic polymorphisms in this population. For example, like other East Asian ethnic groups, Chinese have a relatively high frequency of ALDH2*2 alleles[17] which slows metabolism of potentially carcinogenic acetaldehyde, the first metabolite of alcohol.[18] People with these alleles have a toxic reaction, including hot flushing, increased heart rate and nausea on use of alcohol,[18] and the very limited evidence available suggests they may be at a greater risk of cancer from acetaldehyde exposure.[19;20] Genetic data from this cohort and the large proportion of never-drinkers will provide an opportunity to examine whether the observed protective effects of moderate alcohol use against cancer mortality in these populations[21-24] may be the artifactual results of genetic differences between users and non-users of alcohol. Similarly, the relatively high levels of chronic obstructive pulmonary disease (COPD) in women,[25] despite low female smoking rates provides an opportunity to investigate the aetiology of COPD, in the absence of confounding by smoking.

The very recent history of the epidemiological transition in this cohort potentially provides a unique window into some of the unexplained period effects in the epidemics of chronic diseases associated with economic development (obesity, ischaemic heart disease (IHD), diabetes, hypertension and some cancers), which may yield aetiological clues to their environmental determinants. For example, the social patterning of IHD, hypertension and obesity appear to vary with epidemic stage;[26-28] male disadvantage in IHD risk appears to increase with economic development,[29;30] whilst a female disadvantage in obesity fades.[31] Specifically, we intend to test whether empirically-driven hypotheses mainly developed in

response to observations in economically developed western societies apply in a population with a very different history, at a different epidemiological stage, without universal access to health care, and potentially with a social patterning of disease more consistent with the early stages of the epidemiological transition. Many of these analyses require prospective events to accumulate and thus are medium to longer-term aims. Our shorter term objectives are to demonstrate the validity of GBCS by confirming the usual expected relationships between biological variables in the baseline data; to exploit this study as a retrospective historical cohort for the examination of the impact of occupational and lifetime environmental factors on baseline physical and mental health, while paying special attention to the possibility that people with certain combinations of characteristics may have been excluded from the study through differential survivorship, access or otherwise restricted from participation in the study; and to verify our hypothesis that the relation between socio-economic variables and cardiovascular risk may be more typical of a society at an earlier stage of the epidemiological transition. Clearly a final principal advantage of a large cohort study such as GBCS is the open-ended nature of enquiry made possible by the prospective collection of a wide range of epidemiologic data and biomaterials thus allowing subsequent prosecution to serve the broadest spectrum of research aims, in the tradition of the Framingham,[32] Nurses Health[33;34] and British Doctors[35;36] cohorts.

Who is in the sample?

In developed countries, patient registration lists with defined primary care providers are often used as sampling frames for studies similar to the present one. In China, the infrastructure to facilitate such studies is not available, so a community social and welfare association "Guangzhou Zunlao Kangle Xiehui" or "The Guangzhou Health and Happiness Association for the Respectable Elders" (GHHARE) aligned with the municipal government, was chosen

as a sampling frame. This is a large unofficial organisation with branches throughout Guangzhou and membership open to anyone 50 years of age or older for a nominal, monthly fee of 4 yuan (1 USD = 8 yuan). Figure 2 shows the sample selection strategy. Figure 3 shows the overall study plan. Subjects are included if they are permanent Guangzhou residents capable of consenting and have agreed to participate, ambulatory, and not receiving treatment modalities which if omitted may result in immediate life threatening risk, such as chemotherapy or radiotherapy for cancer, or dialysis for renal failure. The study has ethics approval from the Guangzhou Medical Ethics Committee of the Chinese Medical Association, Guangzhou, China.

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

All subjects will be followed up indefinitely for health services use and cause-specific mortality. Currently, health services use and mortality is being tracked continuously through GHHARE and cross checked with death records from the Guangzhou Public Health Bureau. To ensure completeness and to correct for regression dilution of long-term usual levels of certain risks, subjects will be re-contacted every 3-5 years, for interview, examination and retrieval of additional information from patient-held discharge summaries issued routinely by hospitals, as most health care is fee-for-service and record keeping is an individual responsibility. Where possible, morbidity data will be verified against hospital records. Some subjects with specific diseases, such as diabetes and COPD, will also be selected for more intensive follow-up and monitoring of the disease progress and development of complications.

A pilot follow-up study 18 months after initial recruitment of the first 1,100 subjects achieved re-attendance of 72.2% and telephone interview with a further 22.4%. Information on the remaining subjects was sought from family members, close friends and GHHARE, after

which only 0.4% remained uncontacted. As a validity check, mortality rates in GBCS will be compared with age, sex and cause specific rates in Guangzhou overall.

What has been measured?

To maximise participation, safety and standardisation, after an overnight fast, all subjects come to a newly designed centre at the Guangzhou 12th Hospital, where full time trained nurses and technicians carry out examinations and interviews. After the taking of fasting bloods, the subjects are given free breakfast, and then circulate through pre-set stations for interviews, and for the rest of the physical examination covering measurement of blood pressure, anthropometric indices, 12-lead electrocardiography, pulmonary function testing, and chest radiograph (Table 1). Fasting blood samples are drawn using a vacutainer tube. Lipids, glucose, markers of liver, renal and cardiac function, rheumatoid factor, and high sensitivity C-reactive protein are determined automatically in the hospital laboratory. A complete blood count is performed. Aliquots of plasma are stored at -80°C, as well as two aliquots of bar-coded plasma in liquid nitrogen. White cells extracted using Ficoll-Paque extraction are stored in two bar-coded tubes in liquid nitrogen for subsequent immortalisation. Urinary cotinine levels are measured as an index of smoking intensity or second hand smoke exposure, and to validate the smoking questionnaire.[37] Weight, standing height, sitting height, waist circumference and hip circumference are measured with light indoor clothing and without shoes. Using a face-to-face interview in the appropriate regional dialect, detailed information is entered directly by the interviewer onto a computer based standardised structured questionnaire, including occupational exposures, socio-economic status, family and personal disease histories (in layman terms, as appropriate), cognitive function, smoking history, alcohol use, detailed diet (semi-quantitative food frequency)[38] and physical activity (International Physical Activity Questionnaire short-form) (Table 2).

We assessed the reproducibility of the questionnaire responses in 200 subjects from phase 1 re-interviewed after one month. Examples of the kappa values for a range of categorical variables include: smoking (0.96 and 0.88 for the two questions on smoking status), drinking (0.60), physical activity (0.58), education (0.90), occupation (0.80), household income (0.60), and use of contraceptive pills (0.67), and the intraclass correlation coefficient for continuous variables for example are: age at menarche (0.92) and age at first pregnancy (0.50). These findings support the reproducibility of the study data collected. The detailed food frequency data and self-reports of physical activity will be validated in a subset of subjects at different times of the year.

What has it found?

Following the initial pilot, recruitment of subjects in the first phase took place during September 2003 to November 2004, when 10,413 subjects were examined. The sample mean age was 64 years, ranging from 50 to 93 years, with 70.6% women. Most of the subjects were married and the most common longest held occupation was as a factory worker for 49.5% and 37.9% of females and males, respectively. A higher proportion of men than women were ever smokers or ever-users of alcohol (57.8 and 71.9% in men and 5.3 and 36.6% in women). More women (51.1%) than men (22.5%) were centrally obese (waist circumference $\geq$90cm for men, or $\geq$80cm for women), although there was less difference in general obesity (body mass index ≥ 25 kg/m^2), the prevalence of which was 35.5% of women and 31.1% of men.

As per our stated short-term objectives, we have confirmed that internal relationships at baseline between biological risk factors are similar to results previously demonstrated elsewhere. For example, increasing glycaemia, even in the normal range, is associated with

increasing prevalence of vascular risk factors[39] (Figure 4), and snoring is associated with a history of vascular disease.[40] Investigation of inter-relationships between inflammatory markers and vascular risk factors is ongoing. Taking a life course perspective, increased parity was found associated with increased risk of the metabolic syndrome in women but number of children was not associated with such risk in men,[41] suggesting a specific biological effect of pregnancy rather than a general effect of child rearing (figure 5). Turning to setting GBCS within its historical period and epidemiological time, there are different associations between socio-economic status and disease from those often seen[42], for example diabetes was unrelated to socio-economic status (as determined by educational level and longest held occupation) in women, and possibly associated with higher socio-economic status in men (figure 6).

What are the main strengths and weaknesses?

This large, carefully designed cohort study with detailed baseline characterisation of the subjects has been conceived to investigate a wide range of chronic diseases in a southern Chinese population that has undergone rapid economic transition. GBCS will provide important evidence in the determination of risk factors of chronic disease that emerge with economic development.

Study participants are members of a well-organised association, permanent residents of Guangzhou recruited at one centre closely connected with a well-equipped general hospital, so recruitment is efficient and follow up is facilitated. Using the elderly association is a practicable way to recruit and track large numbers of subjects at low cost without relying on networks of social control such as street and village committees, and the response rate is high. Standardised and mostly automated collection of detailed baseline and follow up information

in a single centre by a small well-trained team with the same high quality, calibrated equipment maximise reliability, validity and data access. Recruitment of subjects steadily throughout the year, coupled with post hoc adjustment as appropriate, removes the possibility of biases caused by potentially large seasonal variation in baseline factors, such as dietary pattern, systolic blood pressure[43] or C-reactive protein.[44]

The older ages of cohort participants means that sufficient numbers of incident outcome events can be expected within a relatively short time period. Results from the prospective and nested case-control studies of relevance to the rest of China transitioning at a slower rate can be expected within a short time frame.

The collection of fasting blood samples and live white cells for later immortalisation are unique among large cohort studies in China. Tests conducted after 3 months of storage found the rate of viability of the white cells was more than 95%. Cell cultures can be used to provide unlimited genomic DNA and are an invaluable source for cytogenetic studies and testing for drug treatment, proteins for proteomic studies, and mRNA. Additional storage of blood at $-80^{o}C$ allows DNA extraction without retrieving the live white cells.

Several challenges bear mention. Our subjects are unlikely to be completely representative of the older population of Guangzhou although they are an unselected sample of GHHARE. Specifically women are over-represented, which we are urgently addressing in the ongoing rounds (phases 2 and 3) of recruitment. Nevertheless within sex and age-group, the phase 1 subjects had fairly similar levels of diabetes and hypertension to nationally representative samples of urban Chinese[45;46] (Figure 7). Notwithstanding these similarities and the lack of easy access to health care, we cannot rule out the possibility of "healthy volunteer bias"

among people who have chosen to join GHHARE, and by extension the potential for a type of "reverse Berkson's bias[*]" particularly in subsequent nested case-control sub-studies where the prevalence of exposure to a health risk would expectedly be artifactually lower in our sample than the general population thus possibly leading to an overestimation of its effect size. Specifically, the exclusion criteria as stated might result in a lower prevalence of general ill health, cancer and renal failure. As such, GBCS is unsuited to estimate disease prevalences, and in common with many non-representative cohort studies the internal validity of relationships is open to bias, if for example subjects with specifically different relationships from those observed are missing. We are planning to investigate, through the use of simulations, what effects such missing data could have. Additionally, in common with all cohorts in developing countries, our subjects are with age increasingly strongly selected survivors. In particular, our subjects survived childhood infections in a pre-antibiotic environment, lived through periodic social turmoil up until the 1970s, and did not move to Hong Kong where until the 1980s immigration from Guangdong province had been commonplace. If survivorship were an issue, we would expect different relationships in the older compared with the younger subjects, which we will check for routinely. The present sample size is not sufficient for studying diseases of lower incidence, such as specific types of cancer. The amount of plasma stored is limited (two cryovials of 1.8ml each per subject), as we can only collect about 18ml of blood, some of which is used for baseline biochemical assays.

Where can I find out more?

---

[*] Whereas Berkson's bias as originally conceived points to the fact that individuals with two diseases may seek treatment for either one, thus increasing the probability that they will present to the health care system and be diagnosed as having the other condition. Thus the prevalence of a second condition that is estimated from a treated (for the first condition) sample is likely to over-represent both the second condition and the co-occurrence of the two conditions in the general population. By reverse Berkson's bias, therefore, we mean that the prevalence estimates of a second condition from our healthy volunteer sample are likely to underestimate the

11

We aim to complete recruitment of the final target of 30,000 subjects by the end of 2007. Although we have no immediate plans to make the data freely available in the public domain, specific proposals for collaboration are welcomed. We are developing a website for the study (www.hku.hk/gbcs) which should be available by December 2006 and further details may be obtained from the site or from commed@hkucc.hku.hk.

---

true burden of disease in the underlying population.

Table 1: Summary of clinical measures collected at baseline in the GBCS

Variables	Number of measure-ments	Equipment used	Method
Standing height	Once	Manufactured equipment	
Sitting height	Once		
Hip circumference	Once	Standard tape measure	Around the maximal girth of the hips
Waist circumference	Once		Horizontally around the smallest circumference between ribs and iliac crest, or at the navel, if no natural waistline
Weight	Once	RGZ-120-RT, Wuxi weighing apparatus factory(无锡市衡器厂)	
Resting blood pressure	Three	Omron 705CP	Seated blood pressure, one minute apart, after a 3-minute rest.
Forced expiratory volume in 1 sec (FEV1)	Best three of up to eight	COSMED microQuark	As per American Thoracic Society / European Respiratory Society (ATS/ERS) task force recommendations, but without reversibility testing
Forced vital capacity (FVC)			
12-lead electrocardiography	Once	Marquette CE0366	
Chest X-ray	Once	Toshiba 500mA	
Fasting plasma glucose	Once	Shimadzu CL-8000 Clinical Chemistry Analyzer	
Lipid profile[a]	Once		
Markers of liver[b], renal[c] and cardiac[d] function	Once		
Rheumatoid factor	Once		
C-reactive protein	Once		
Complete blood count[e]	Once	SYSMAX K-21	
Urinary cotinine	Once	Agilent 6890 series GC System, Agilent 5973 Network Mass Selective Detector	

[a] total, LDL-, and HDL-cholesterol, triglycerides, apolipoprotein A1 and B
[b] alanine aminotransferase, aspartate transaminase
[c] total protein, albumin, globulin, total bilirubin, urea, creatinine, uric acid, phosphate, total calcium, magnesium
[d] creatine kinase, lactate dehydrogenase, hydroxybutyrate dehydrogenase
[e] white (and subtypes) and red blood cell counts, haemoglobin levels, and platelet counts
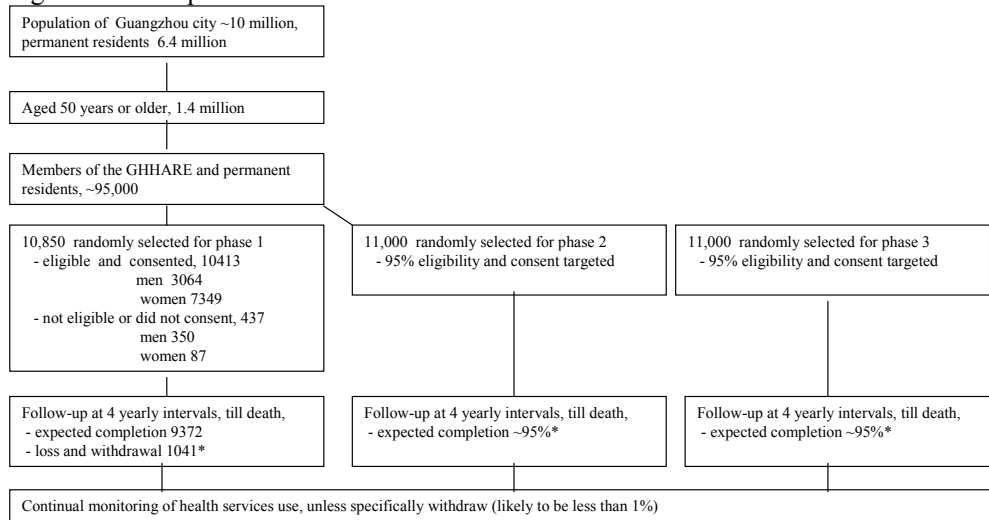
Table 2: Summary of questionnaire data collected at baseline in the GBCS

Exposure category	Variable / exposure
Demographics and socio-economics	Housing type
	Household size
	Occupation
	Income
	Education
	Health insurance coverage
Personal health behaviour	Smoking
	Alcohol
	Tea
	Spicy food
	Afternoon nap
	Physical activity (at work and leisure)
	Diet (300-item food frequency questionnaire)
Environment	Occupational exposures
	Cooking and heating fuels
	Passive smoking exposure
Family history	Family history of coronary heart disease, diabetes, cancers, other chronic diseases
	Parental age, age of death and cause of death
Specific health questions/ disability/impairment	Rose angina screening questionnaire
	MRC respiratory health questionnaire
	Self-reported general health
	Arthritis screening / activities of daily living (ADL)
	Dementia screening (delayed word recall test)
	Eyesight
	Hearing
	Dental health
Past medical history	Diagnosed medical conditions (specifically itemised)
	Use of health services
	Use of Chinese medicines
	Western and other medication use
Reproductive history	Age of menarche
	Parity and breastfeeding history
	Menopause status and associated timing
	Contraceptive history
	History of hormone-replacement therapy use
	Hysterectomy or other gynaecological interventions
Psychological status	Psychological morbidity scales / stress
	Anger / hostility
	Optimism
	Social isolation

Figure 1: Location of major urban areas in China, with populations of more than about 10 million

Beijing
Tianjin
Chengdu
Chongqing
Shanghai
Guangzhou

Guangdong Province
Guangzhou
Hong Kong

Figure 2: Participant recruitment

Population of Guangzhou city ~10 million, permanent residents 6.4 million

Aged 50 years or older, 1.4 million

Members of the GHHARE and permanent residents, ~95,000

10,850 randomly selected for phase 1
- eligible and consented, 10413
       men 3064
       women 7349
- not eligible or did not consent, 437
       men 350
       women 87

11,000 randomly selected for phase 2
- 95% eligibility and consent targeted

11,000 randomly selected for phase 3
- 95% eligibility and consent targeted

Follow-up at 4 yearly intervals, till death,
- expected completion 9372
- loss and withdrawal 1041*

Follow-up at 4 yearly intervals, till death,
- expected completion ~95%*

Follow-up at 4 yearly intervals, till death,
- expected completion ~95%*

Continual monitoring of health services use, unless specifically withdraw (likely to be less than 1%)

* Lost to follow-up may be tracked in local mortality registers
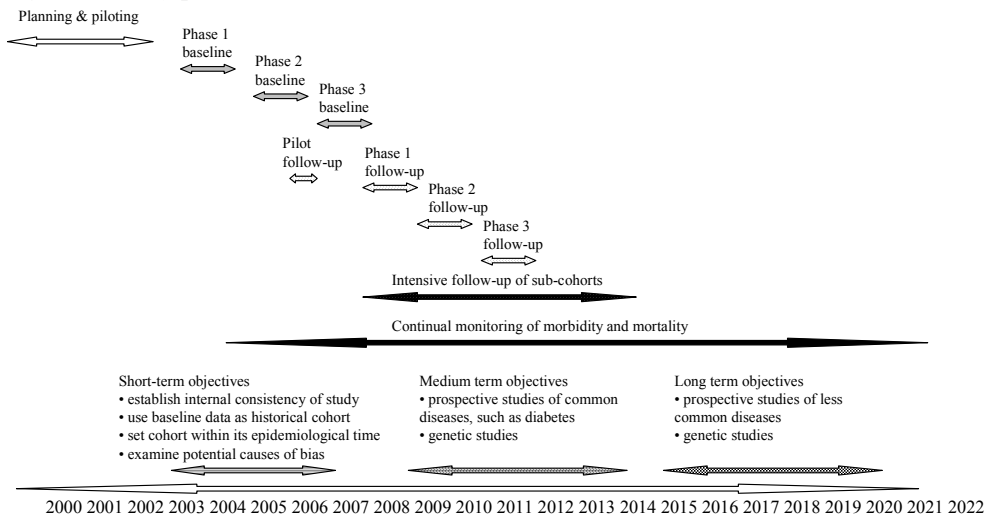
15

Figure 3: Study plan

Figure 4: Relationship between increasing fasting plasma glucose in the normal range and selected cardiovascular risk factors (LOWESS plot)
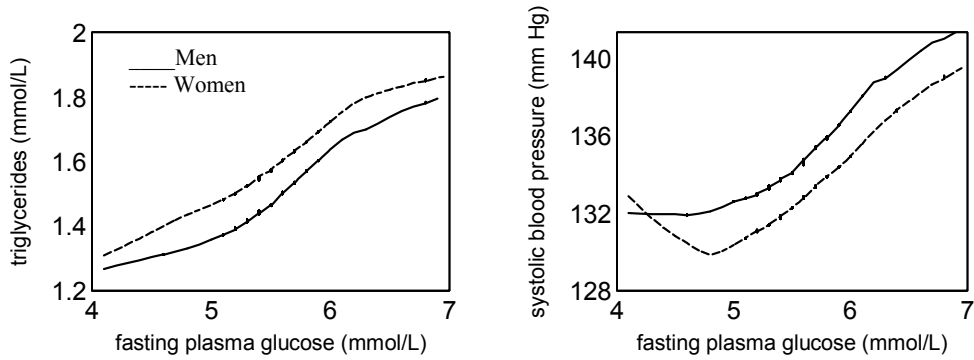
Figure 5: Relationship between number of children and prevalence of the metabolic
syndrome (MS) in women and men, using the International Diabetes Federation definition[47],
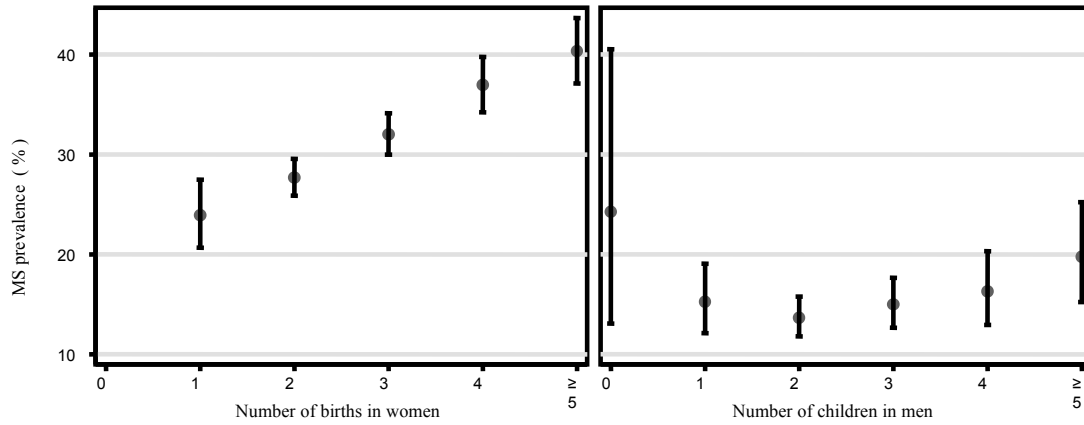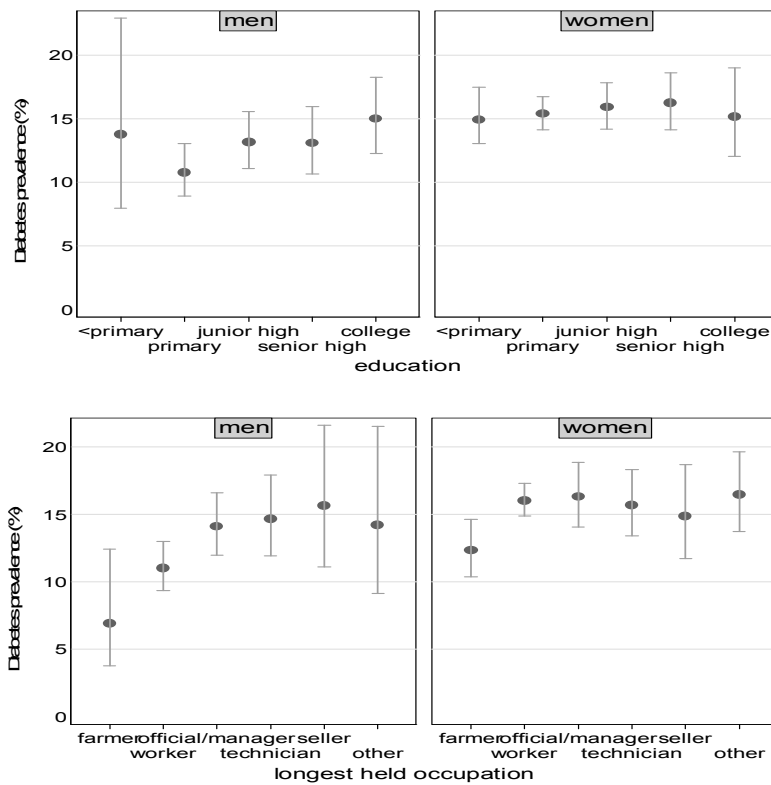adjusted for age, socio-economic status, use of alcohol, smoking and physical activity.

Figure 6: Relationship between prevalence of diabetes* and educational level and longest
held occupation by sex, adjusted for age, smoking, use of alcohol and physical activity

* diabetes defined as fasting plasma glucose >=7mmol/L or use of anti-diabetic medication

17

Figure 7: Prevalence of Diabetes* and Hypertension† in GBCS compared with a nationally representative urban sample[45;46], with 95% confidence intervals
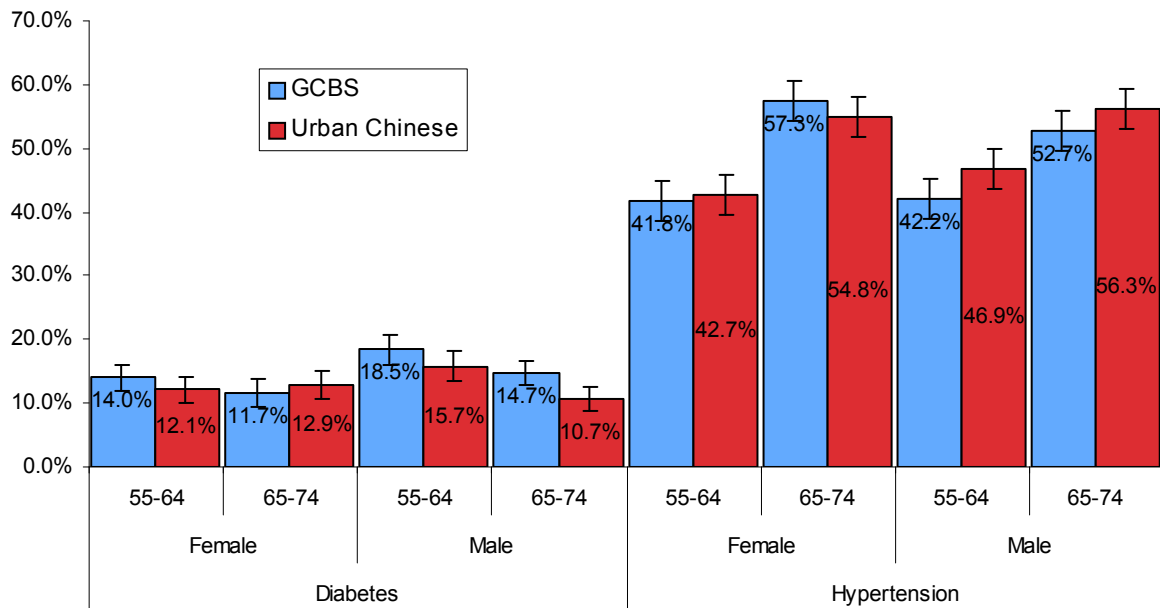
* defined as fasting plasma glucose ≥7.0 mmol/L or use of anti-diabetic medication

† defined as blood pressure≥ 140/90mm Hg or use of anti-hypertensive medication

---Page Break---

References

1.   Chen Z, Lee L, Chen J *et al*. Cohort profile: the Kadoorie Study of Chronic Disease in China (KSCDC). *Int.J Epidemiol.* 2005; 34:1243-9.

2.   UK Biobank. http://www.ukbiobank.ac.uk/ . 2006.

3.   Maddison A. Chinese economic performance in the long run. Paris: OECD , 1998.

4.   Eckstein A. China's Economic Development. Ann Arbor: University of Michigan, 1975.

5.   Komlos J. An anthropometric history of early-modern France. *European Review of Economic History* 2003; 7:159-89.

6.   Shirokogorov SM. Anthropology of eastern China and Kwangtung Province. Shanghai: Royal Asiatic Societyof Great Britian and Ireland. North China Branch, 1925.

7.   Cadbury WW. Height, weight and chest measurements of Mongolian Peoples with especial reference to Southern Chinese. *The Phillipine Journal of Science* 1924; 25:733-59.

8.   Hutcheson AC. Report on the height, weight and chest measurements of healthy Chinese. *China Medical Journal* 1920; 34:13-6.

9.   Whyte GD. The height, weight and chest measurements of Healthy Chinese. *The China Medical Journal* 1918; 32:210-6.

10. Stevenson PH. Collected Anthropomteric data on the Chinese. *The China Medical Journal* 1925; XXXIX:855-98.

11. Rawski T. Economic Growth in Pre-war China. Berkeley: University of California Press, 1990.

12. Brandt L. Reflections on China's Late 19th and Early 20th-Century Economy. *The China Quarterly* 1997; 150:282-308.

13. MacFarquhar R. *Origins of the Cultural Revolution: Great Leap Forward 1958-60*. New York: Columbia University Press, 1983.

14. Leung GM, Wagstaff A, Lindelow M, Lu JR. The health systems of China, Hong Kong and Taiwan. In Heggenhougen HK, ed. *The encyclopedia of public health.*, p in press. Oxford: Elsevier, 2006.

15. He J, Gu D, Wu X *et al*. Major causes of death among men and women in China. *N.Engl.J Med.* 2005; 353:1124-34.

16. Smith GD,.Ebrahim S. Mendelian randomization: prospects, potentials, and limitations. *Int.J Epidemiol.* 2004; 33:30-42.

17. Wall TL, Peterson CM, Peterson KP *et al*. Alcohol metabolism in Asian-American men with genetic polymorphisms of aldehyde dehydrogenase. *Ann.Intern.Med.* 1997; 127:376-9.

18. Boffetta P,.Hashibe M. Alcohol and cancer. *Lancet Oncol.* 2006; 7:149-56.

19. Yokoyama A,.Omori T. Genetic polymorphisms of alcohol and aldehyde dehydrogenases and risk for esophageal and head and neck cancers. *Jpn.J.Clin.Oncol.* 2003; 33:111-21.

20. Lewis SJ,.Smith GD. Alcohol, ALDH2, and esophageal cancer: a meta-analysis which illustrates the potentials and limitations of a Mendelian randomization approach. *Cancer Epidemiol.Biomarkers Prev.* 2005; 14:1967-71.

21. Lin Y, Kikuchi S, Tamakoshi A *et al*. Alcohol consumption and mortality among middle-aged and elderly Japanese men and women. *Ann.Epidemiol.* 2005; 15:590-7.

22. Tsugane S, Fahey MT, Sasaki S, Baba S. Alcohol consumption and all-cause and cancer mortality among middle-aged Japanese men: seven-year follow-up of the JPHC study Cohort I. Japan Public Health Center. *Am.J.Epidemiol.* 1999; 150:1201-7.

23. Yuan JM, Ross RK, Gao YT, Henderson BE, Yu MC. Follow up study of moderate alcohol intake and mortality among middle aged men in Shanghai, China. *BMJ* 1997; 314:18-23.

24. Inoue M,.Tsugane S. Impact of alcohol drinking on total cancer risk: data from a large-scale population-based cohort study in Japan. *Br.J.Cancer.* 2005; 92:182-7.

25. Xu F, Yin X, Zhang M, Shen H, Lu L, Xu Y. Prevalence of physician-diagnosed COPD and its association with smoking among urban and rural residents in regional mainland China. *Chest.* 2005; 128:2818-23.

26. Gonzalez MA, Rodriguez AF, Calero JR. Relationship between socioeconomic status and ischaemic heart disease in cohort and case-control studies: 1960-1993. *Int J Epidemiol.* 1998; 27:350-8.

27.	Colhoun HM, Hemingway H, Poulter NR. Socio-economic status and blood pressure: an overview analysis. *J Hum Hypertens.* 1998; 12:91-110.

28.	Monteiro CA, Moura EC, Conde WL, Popkin BM. Socioeconomic status and obesity in adult populations of developing countries: a review. *Bull.World Health Organ* 2004; 82:940-6.

29.	Lawlor DA, Ebrahim S, Davey SG. Sex matters: secular and geographical trends in sex differences in coronary heart disease mortality. *BMJ.* 2001; 323:541-5.

30.	Nikiforov SV,.Mamaev VB. The development of sex differences in cardiovascular disease mortality: a historical perspective. *Am.J.Public Health.* 1998; 88:1348-53.

31.	Seidell JC. Epidemiology of obesity. *Semin.Vasc.Med.* 2005; 5:3-14.

32.	Oppenheimer GM. Becoming the Framingham Study 1947-1950. *Am J Public Health.* 2005; 95:602-10.

33.	Colditz GA, Manson JE, Hankinson SE. The Nurses' Health Study: 20-year contribution to the understanding of health among women. *J Womens Health.* 1997; 6:49-62.

34.	Colditz GA,.Hankinson SE. The Nurses' Health Study: lifestyle and health among women. *Nat.Rev.Cancer.* 2005; 5:388-96.

35.	Doll R, Peto R, Boreham J, Sutherland I. Mortality in relation to smoking: 50 years' observations on male British doctors. *BMJ.* 2004; 328:1519.

36.	Doll R, Peto R, Boreham J, Sutherland I. Mortality in relation to alcohol consumption: a prospective study among male British doctors. *Int.J Epidemiol.* 2005; 34:199-204.

37.	Cao M, Zhang Y, Jiang CQ, Lam TH, Cheng KK, Zhang WS. Methodology of measuring urine cotinine. *Chinese Journal of Industrial Hygiene and Occupational Diseases [Chinese]* 2006; A24:G39-G41.

38.	Woo J, Leung SS, Ho SC, Sham A, Lam TH, Janus ED. Dietary practices and lipid intake in relation to plasma lipid profile in Hong Kong Chinese. *Eur J Clin Nutr* 1997; 51:467-71.

39.	Thomas GN, Jiang CQ, McGhee SM *et al*. Association of vascular risk factors with increasing glycaemia even in normoglycaemic subjects in an older Chinese population: the Guangzhou Biobank Cohort Study. *Metabolism* 2006; in press.

40.	Thomas GN, Jiang CQ, Lao XQ *et al*. Snoring and vascular risk factors and disease in a low-risk Chinese population: the Guangzhou Biobank Cohort Study. *Sleep* 2006; in press.

41.	Lao X, Thomas GN, Jiang CQ *et al*. Parity and the metabolic syndrome in older Chinese women. *Clinical Endocrinology* 2006; in press.

42.	Robbins JM, Vaccarino V, Zhang H, Kasl SV. Socioeconomic status and diagnosed diabetes incidence. *Diabetes Res.Clin.Pract.* 2005; 68:230-6.

43.	Sega R, Cesana G, Bombelli M *et al*. Seasonal variations in home and ambulatory blood pressure in the PAMELA population. Pressione Arteriose Monitorate E Loro Associazioni. *J Hypertens.* 1998; 16:1585-92.

44.	Sung KC. Seasonal variation of C-reactive protein in apparently healthy Koreans. *Int.J.Cardiol.* 2006; 107:338-42.

45.   Gu D, Reynolds K, Duan X *et al* . Prevalence of diabetes and impaired fasting glucose in the Chinese adult population: International Collaborative Study of Cardiovascular Disease in Asia (InterASIA). *Diabetologia* 2003; 46:1190-8.

46.   Reynolds K, Gu D, Muntner P *et al*. Geographic variations in the prevalence, awareness, treatment and control of hypertension in China. *J Hypertens.* 2003; 21:1273-81.

47.   Alberti KG, Zimmet P, Shaw J. The metabolic syndrome--a new worldwide definition. *Lancet* 2005; 366:1059-62.