Kadoorie Biobank Study of 0.5 million adults in China

Zhengming Chen

CTSU, Richard Doll Building

Nuffield Dept. Of Clinical Medicine

University of Oxford, UK

XVIII IEA World Congress of Epidemiology,
Porto Alegra, Brazil 20-24th Sep 2008

Comparison of premature death between UK and China, 2000

Age range	Probability of death	
	CHINA	UK
0-34	6%	3%
35-69	34%	25%
(70+)	(60%)	(72%)

Premature death in China

Of 17M births each year, 7M can expect to die before 70, at current death rates

1M age 0-34, all causes

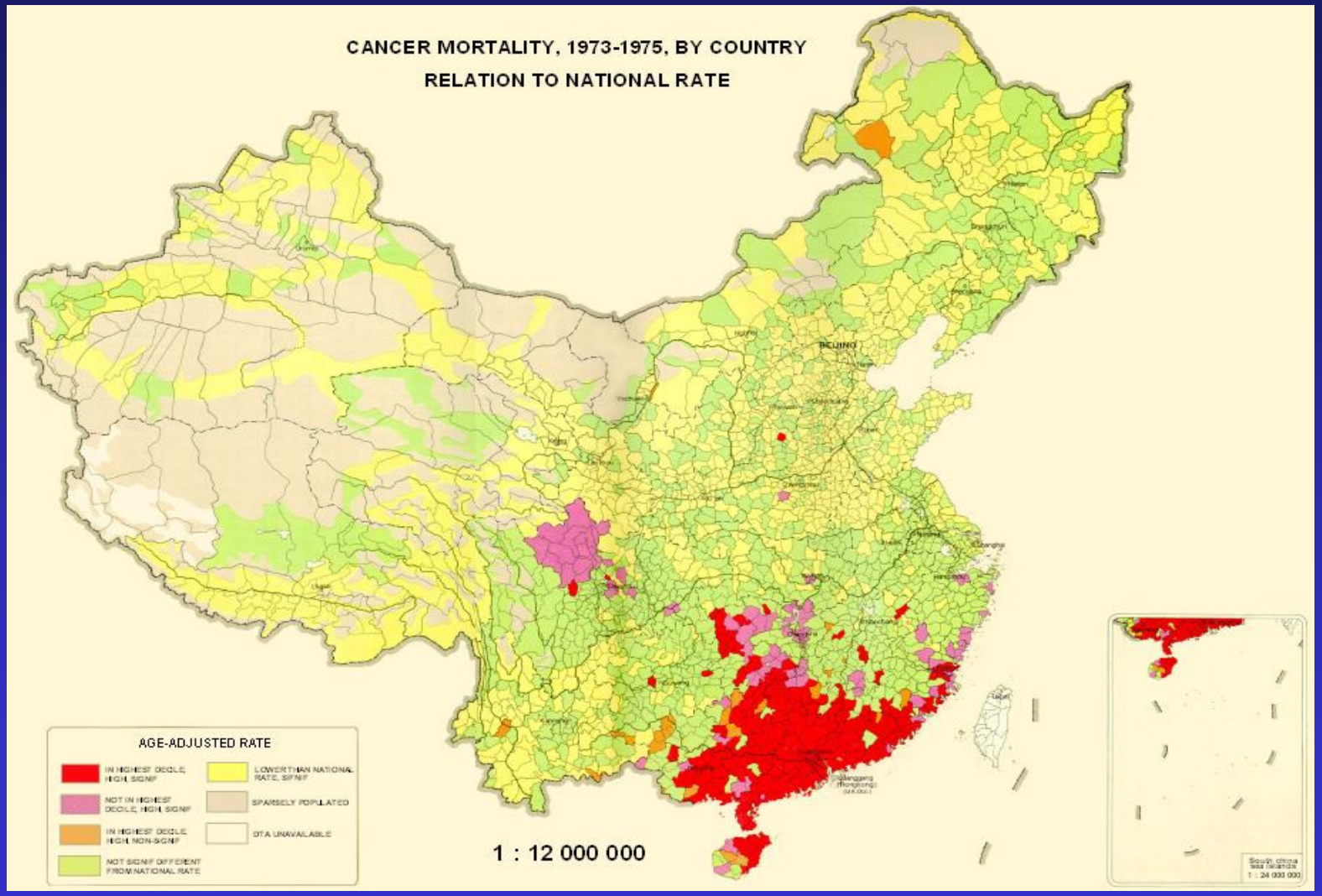
6M age 35-69:

- *2M CVD*
- *2M Cancer*
- *1M COPD*
- *1M Other causes, incl. suicide*

10M age 70+, all causes

Geographic variation of nasopharynx cancer mortality

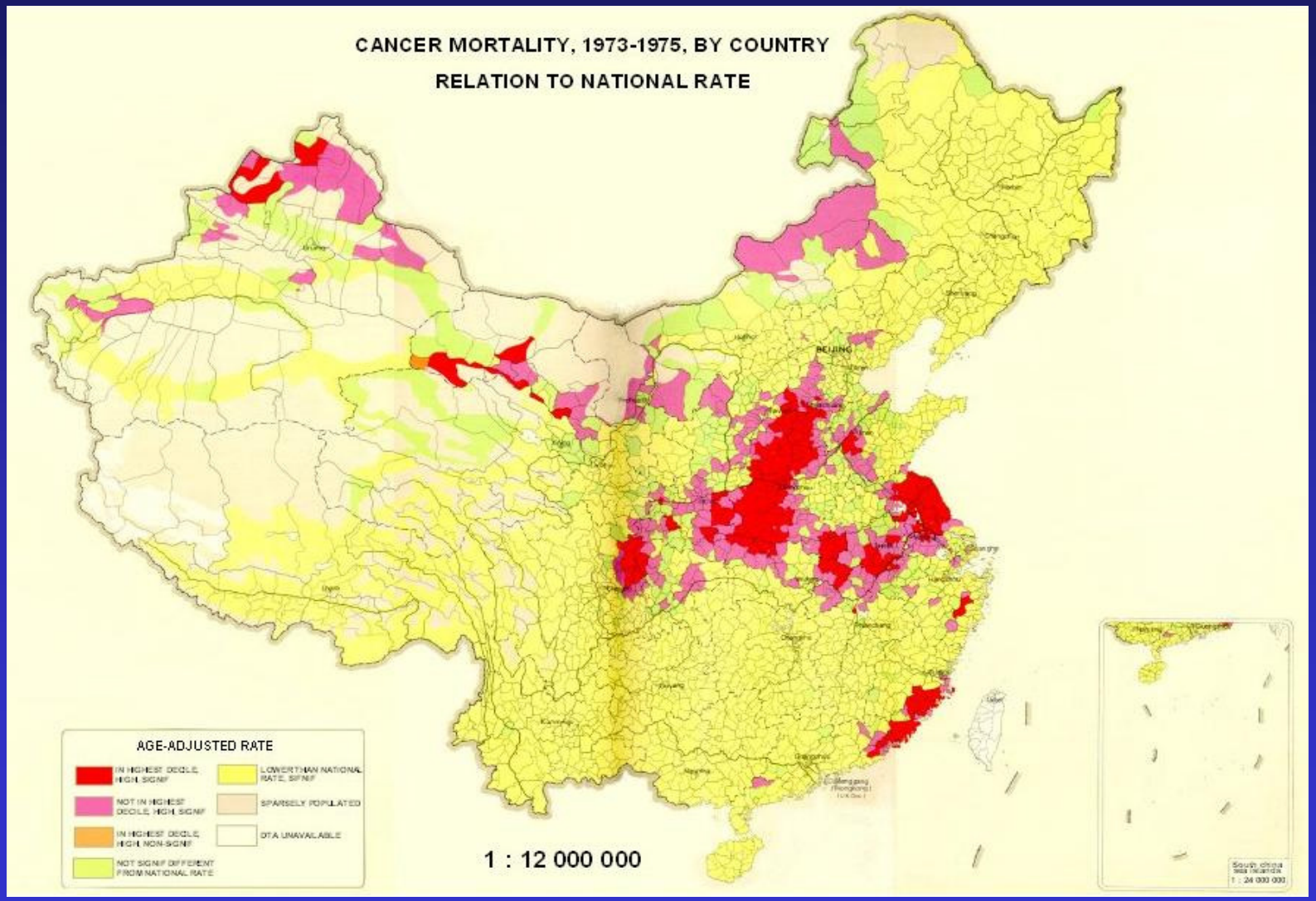
(The First National Mortality Survey in China, 1973-1975, Male)



Atlas of Cancer Mortality in the People's Republic of China
Published by China Map Press, 1976

Geographic variation of oesophagus cancer mortality

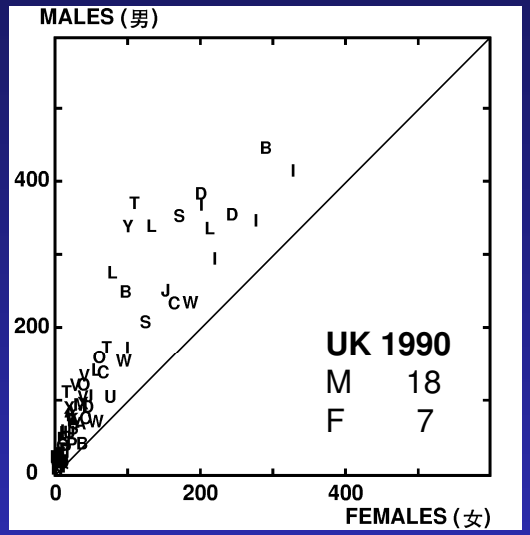
(The First National Mortality Survey in China, 1973-1975, Male)



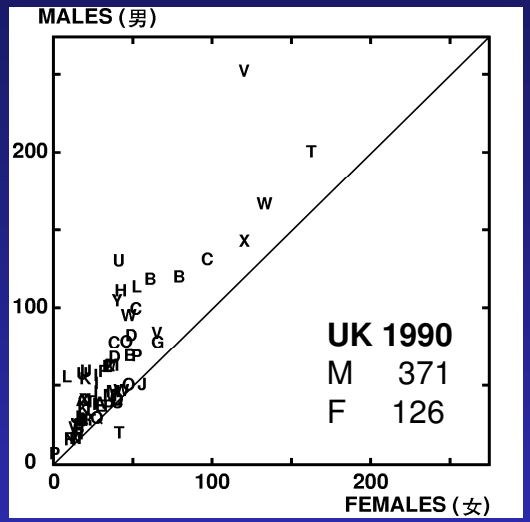
Atlas of Cancer Mortality in the People's Republic of China
Published by China Map Press, 1976

Large unexplained variations in mortality rates from major chronic diseases in China, aged 35-69, 1986-88

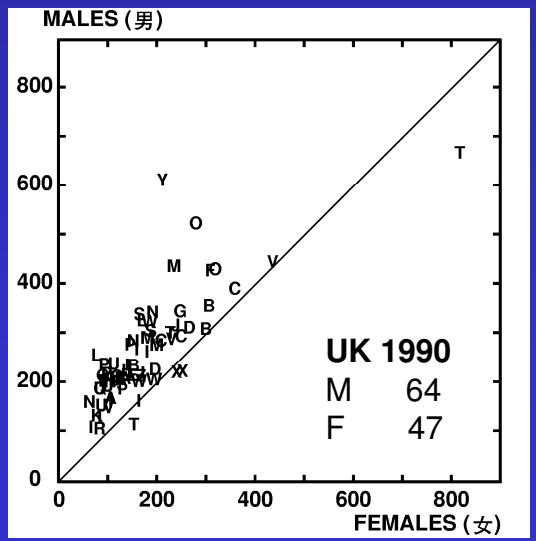
Oesophagus cancer



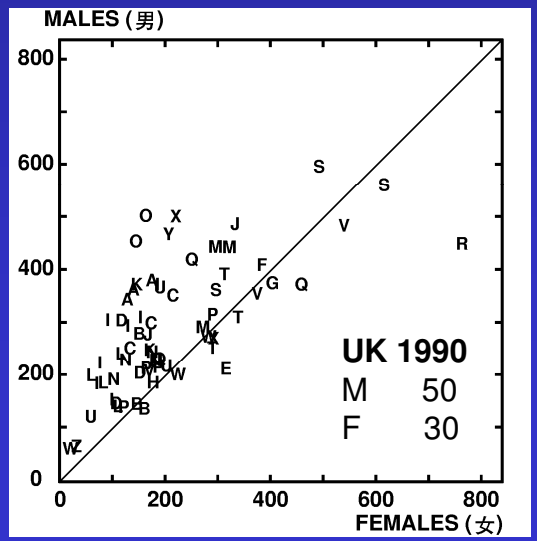
IHD



Stroke



COPD



Kadoorie Biobank Study

An open-ended study with broad research aims

To study the separate and combined effects of lifestyle, environmental and genetic factors on morbidity & mortality of common chronic conditions in a Chinese population

A large collaborative project between Oxford CTSU and China Centre for Disease Control, funded by KCF

KSCDC
中英慢病合作项目

Kadoorie Biobank Study: design

- 0.5M adults, aged 35-74 years, in 10 areas
- Baseline survey: questionnaire, physical measurements & collection of blood sample
- Morbidity & mortality follow up for decades
- Periodic resurvey of a sample of participants

KSCDC
中英慢病合作项目

The location of 10 survey sites in China

(with different risk exposure and disease patterns)

Urban

Qindao
Harbin
Liuzhou
Suzhou
Haikou

Rural

Sichuan
Zhejiang
Hunan
Gansu
Henan

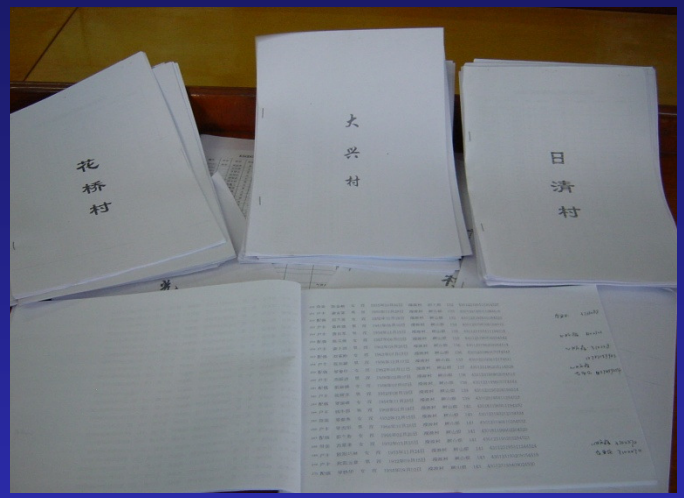
KSCDC
中英慢病合作项目

Survey team at RCC travelling out to clinics on a daily basis

About 160 staff in China working on the project

Standard procedures in assessment centre

Identifying participants by record



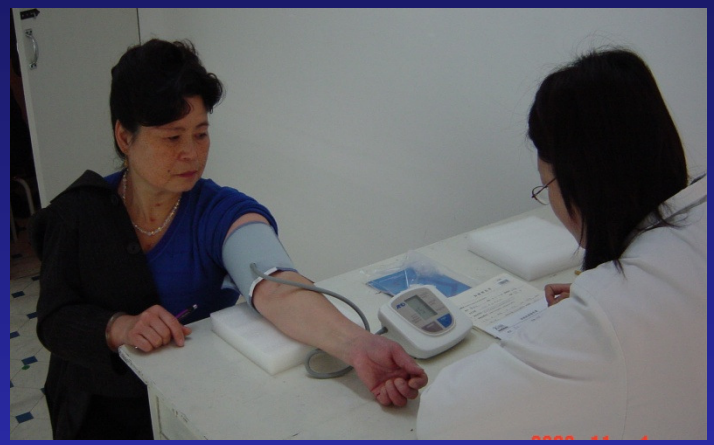
Attending clinics by invitation

Obtaining written consent

Measurements then interview

KSCDC 中英慢病合作项目

A range of physical measurements



- Standing & sitting heights
- Weight & bio-impedance
- Hip and waist circumference
- Blood pressure & heart rate
- Lung function (FEV1, FVC)
- CO level
- Random blood glucose
- HBsAg

A 10 ml blood sample will be collected for long-term storage

KSCDC 中英慢病合作項目

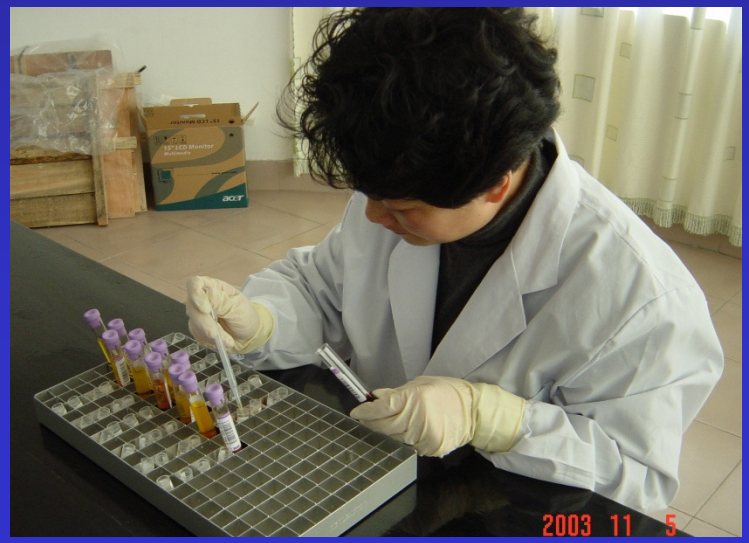
Interview using laptop-based questionnaire

- Social-economic status
- Tobacco use
- Alcohol & green tea
- Diet & food supplement
- Indoor air-pollution
- Physical activity
- Sleep & psychological status
- Medical history & treatment
- Reproductive history (for women)

The total clinic visit time: ~60 minutes

KSCDC 中英慢病合作项目

Blood samples processed within 24 h



KSCDC
中英慢病合作项目

Short- & long-term blood storage facilities

-40° C freezers at RCC

Equipped with alarm & power generator

Nitrogen tank facility at Oxford

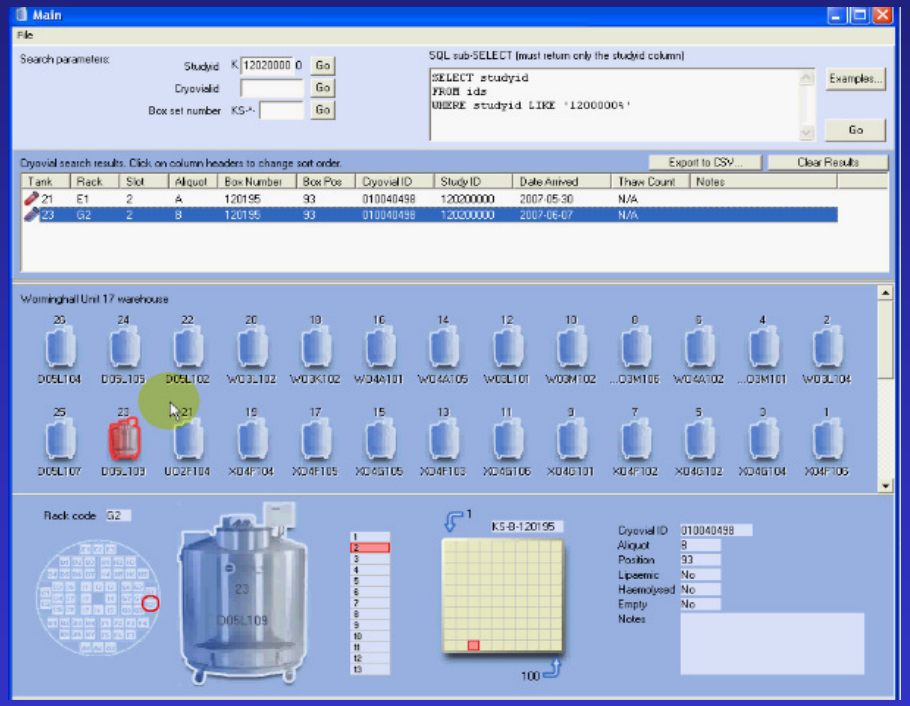
Each tank holds 70,200 aliquots

KSCDC
中英慢病合作项目

Computerised sample monitoring, tracking, shipment and storage management system

Monitoring & tracking blood samples

Central sample storage management

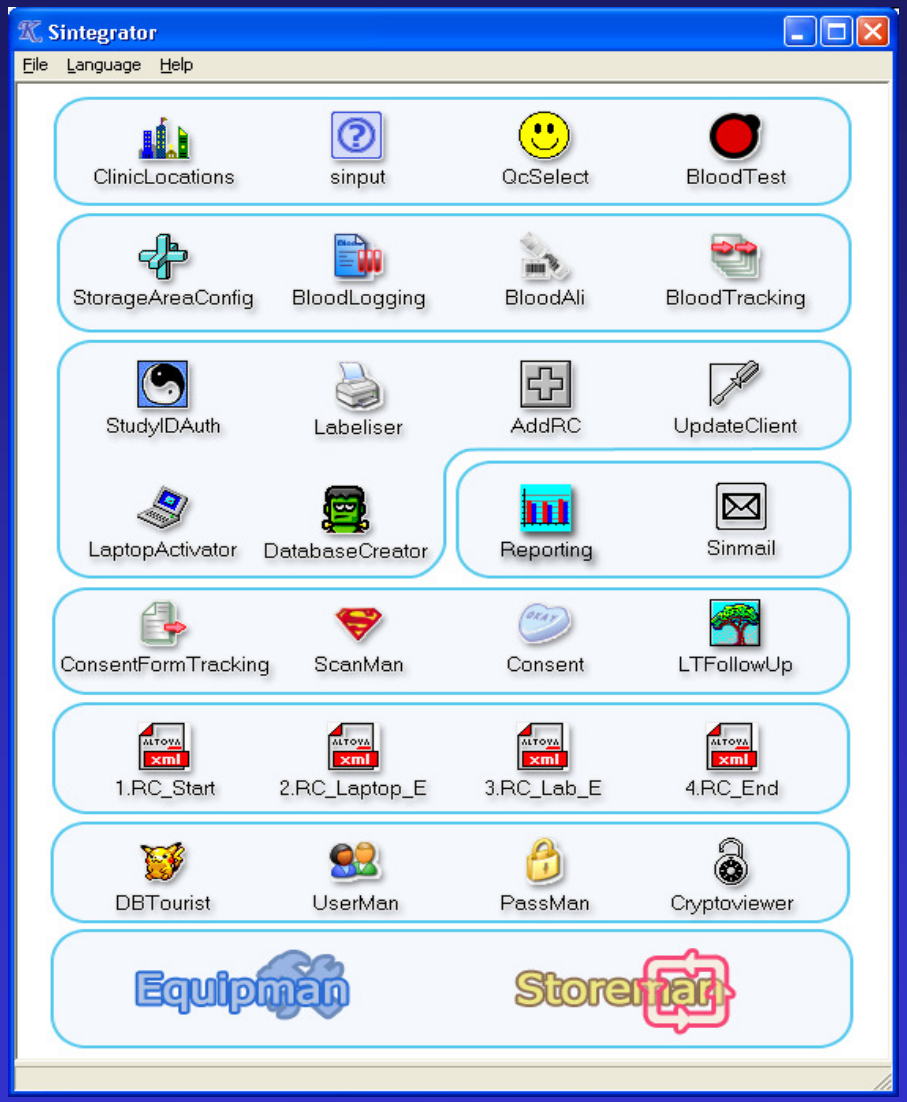


KSCDC 中英慢病合作项目

Fully computerised systems for the study

Main IT functionalities

- Survey data collection
- Blood processing, tracking and shipment control
- Remote database & system management
- QC monitoring & reporting
- Web-based progress report & internal email system
- Long-term follow up linkage
- Stock control & equipment calibration monitoring



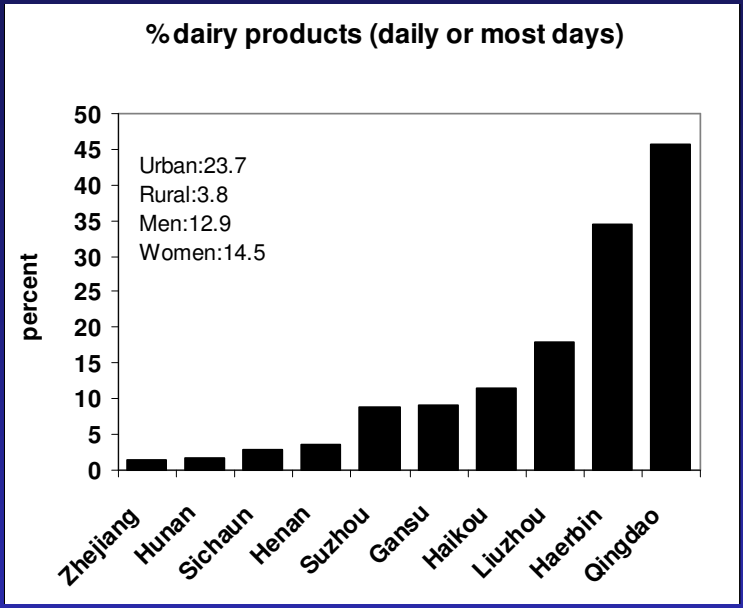
KSCDC 中英慢病合作项目

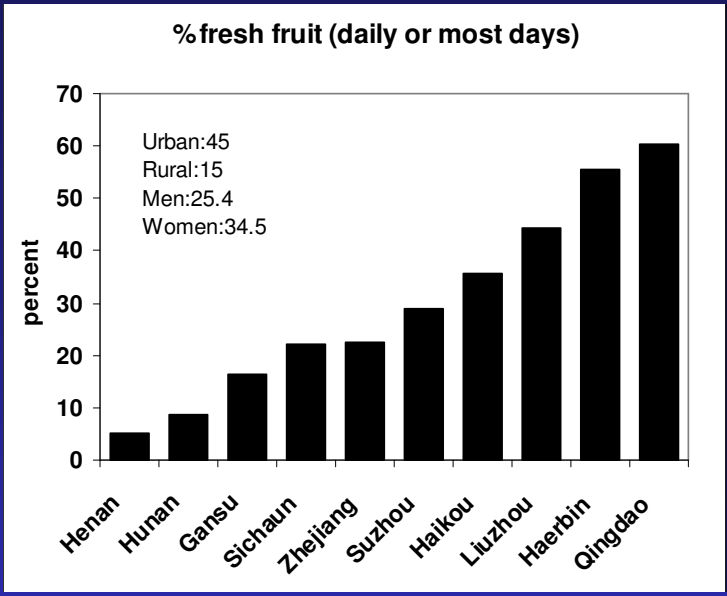
515,000 participants recruited between 2004-8

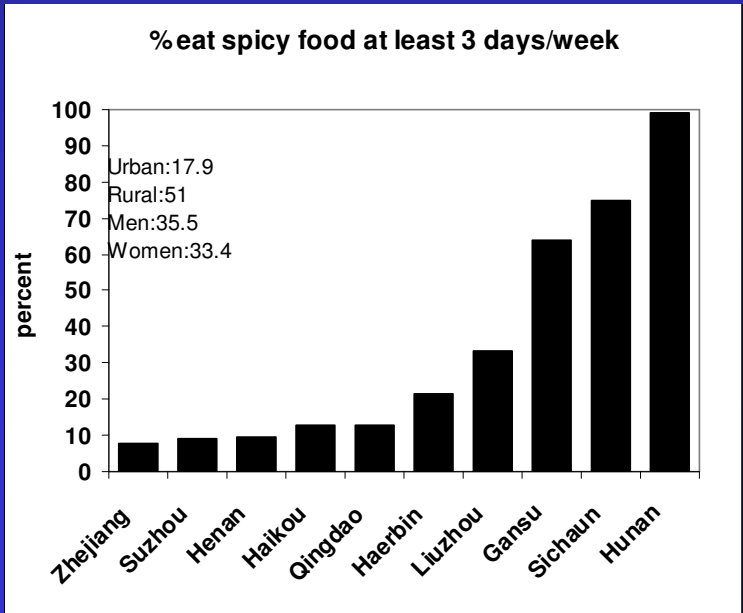
(41% men; mean age: 52 yrs; mean BMI: 23.6 kg/m^2)

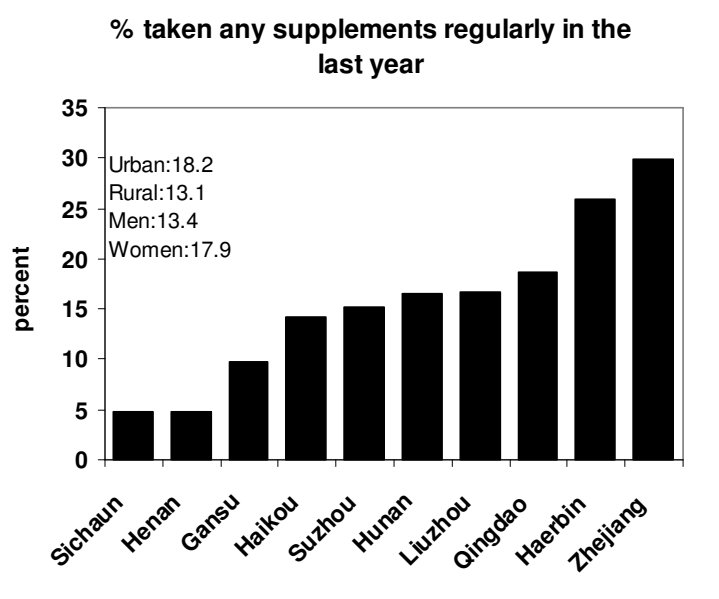
~100% blood collection & 97% processed <24 h,
& all have their national ID No. recorded for follow-up

Dietary patterns between different areas



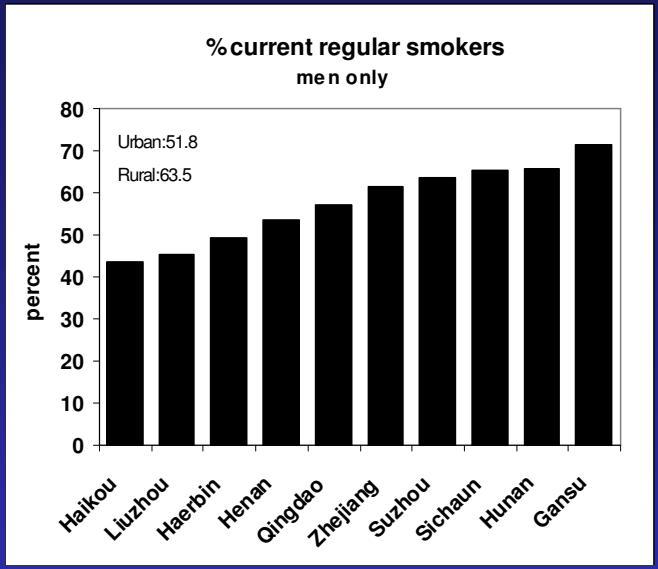


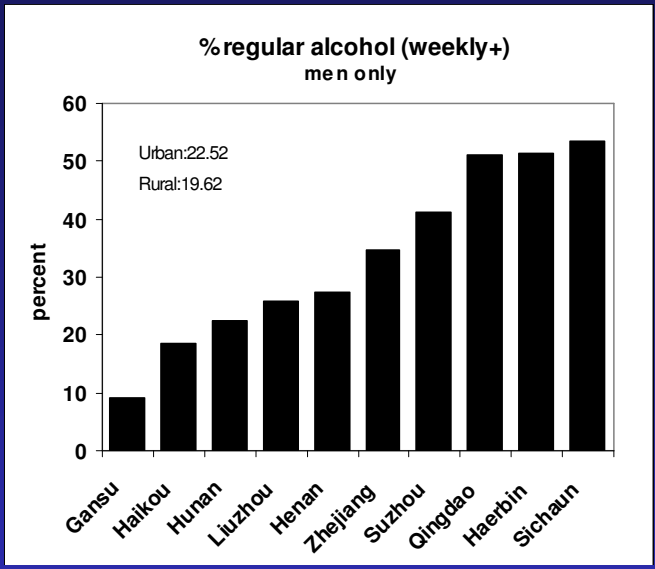


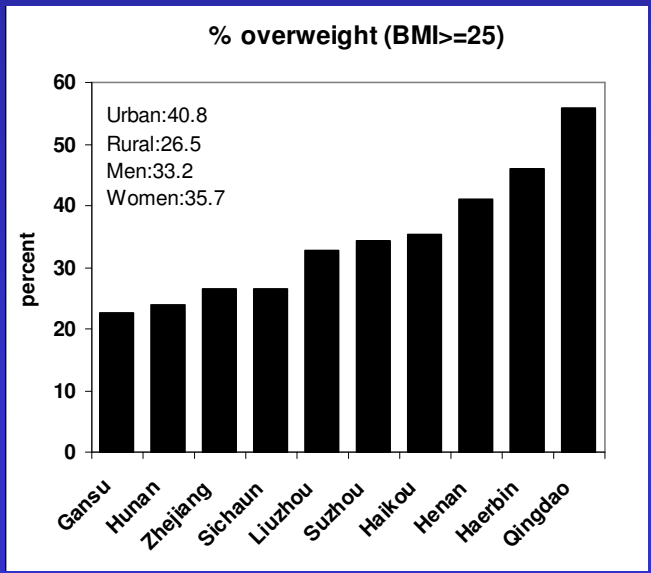


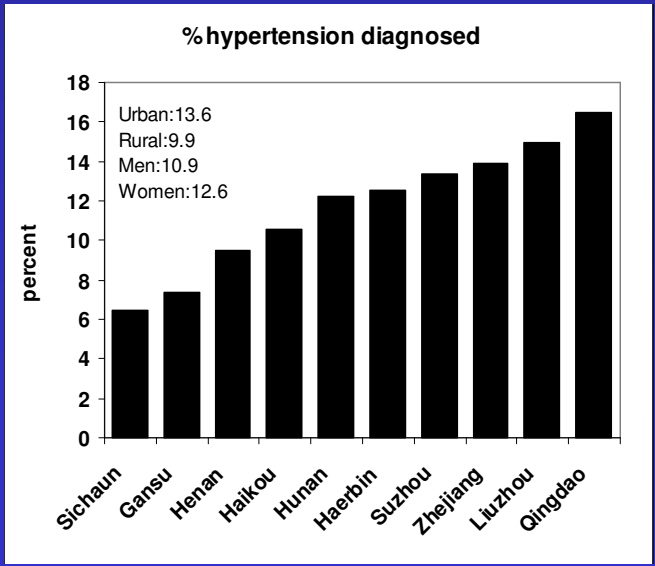
KSCDC 中英慢病合作项目

Regional variations in prevalence of smoking, alcohol drinking, overweight and hypertension



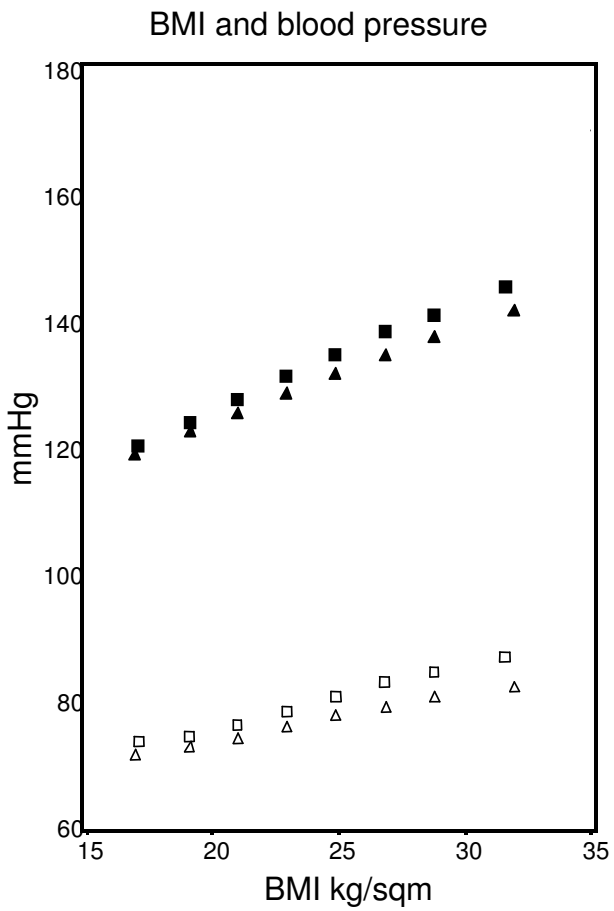


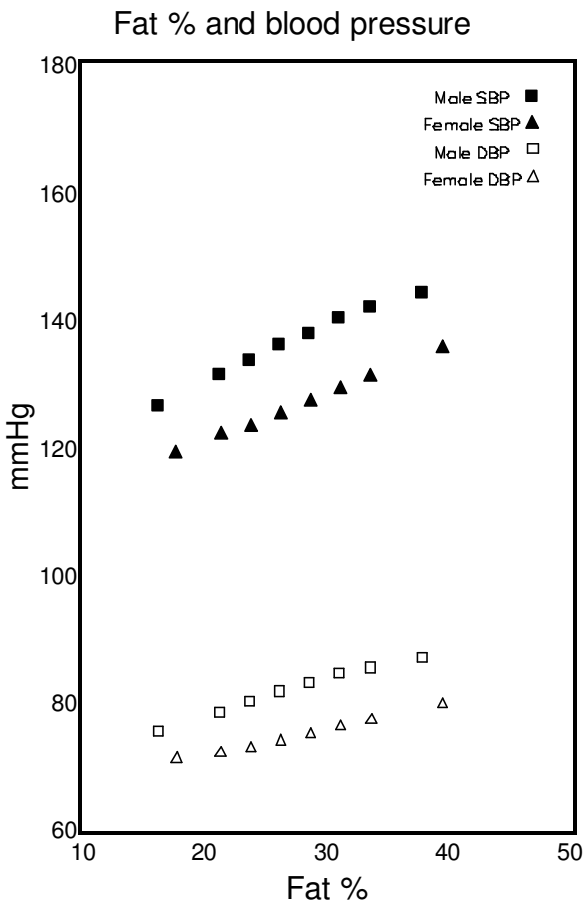


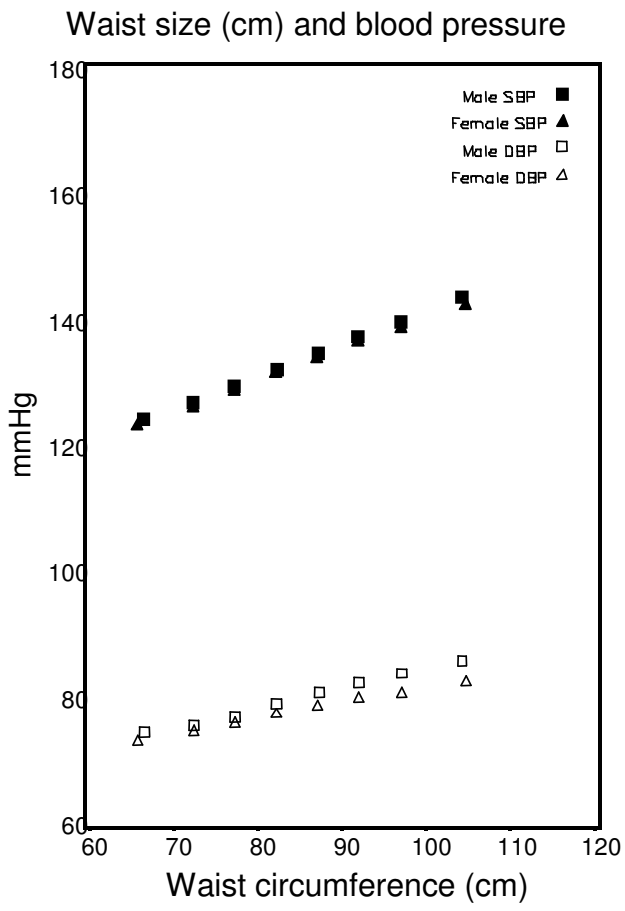


KSCDC
中英慢病合作项目

Associations of blood pressure with different measures of adiposity in the Kadoorie Biobank







Long-term follow-up of morbidity & mortality through electronic linkage by national ID

Restricted access for searching and linkage using name, hospital, NID

Electronic linkage to health insurance claim systems by unique national ID

(eg, Ischaemic stroke confirmed by MRI at county hospital)

桐乡市城乡居民合作医疗

信息管理系统

机构名称: 桐乡市第一人民医院　合作医疗卡号:　全部　未完成　已完成　查询(Q)　审核(A)

姓名:　日期: 2007- 7- 1 至 2007- 7-20　类别: 住院　重置(R)　关闭(C)

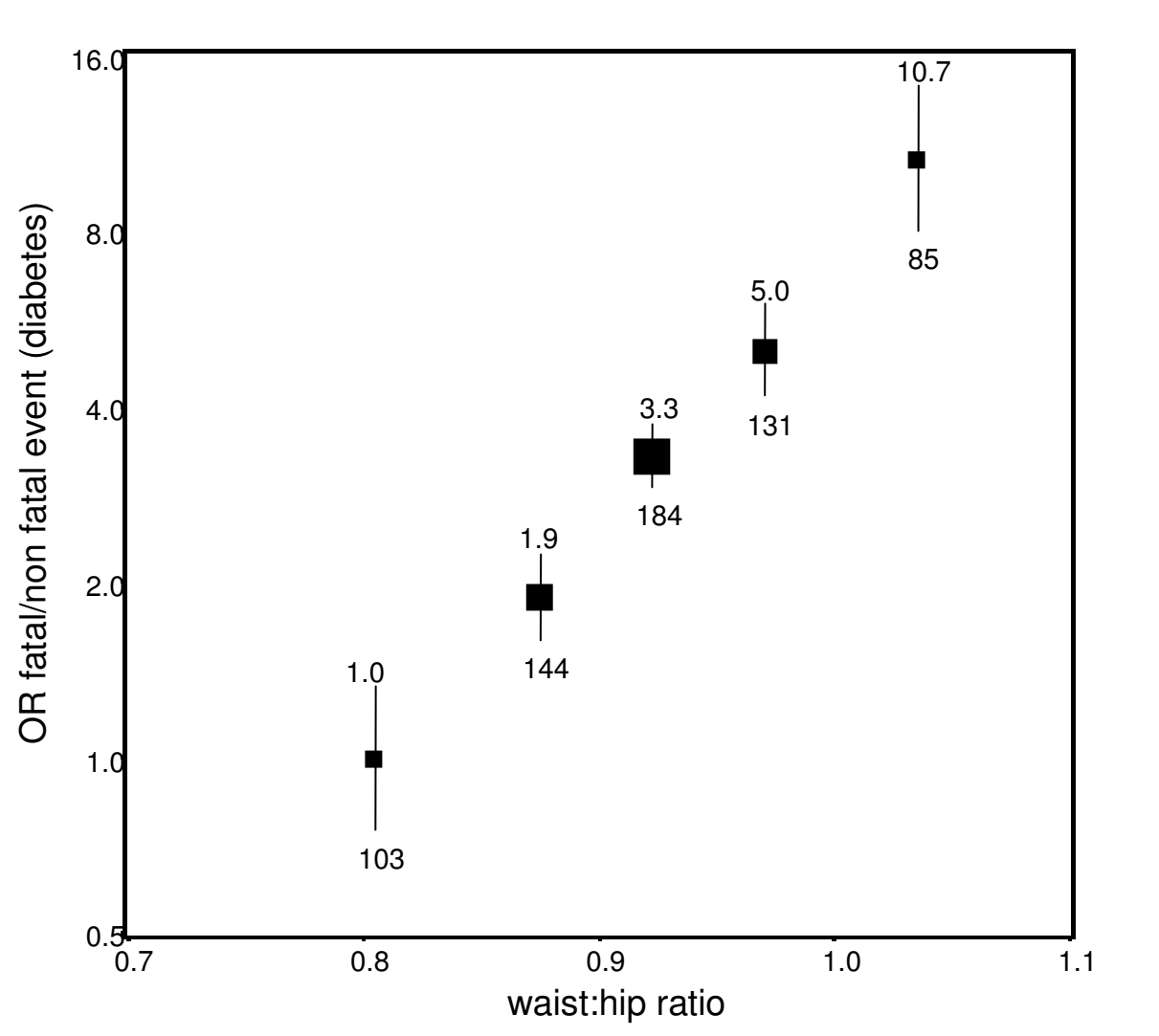
姓名	性别	人员类型	结算单号	结算次数	合作医疗卡号	就诊时间	结算日期	冲票标记	冲票日期	疾病名称
钱永泉	男	普通	7893	1	33048312210498060001	2007-07-07至2007-07-16	2007-07-16 10:03:11	正常		脑梗塞
李汉清	男	普通	7555	1	33048311220411010001	2007-06-28至2007-07-01	2007-07-01 09:31:26	正常		脑梗塞
吕杏娥	女	普通	7868	1	33048308210081010001	2007-07-05至2007-07-14	2007-07-14 09:24:27	正常		脑梗塞
张卫平	女	城镇低保	7488	1	33048301071165010001	2007-06-25至2007-07-01	2007-07-01 15:46:25	正常		脑梗塞
郁明和	女	普通	8297	1	33048311200808020001	2007-07-15至2007-07-16	2007-07-16 11:19:06	正常		脑出血
张金洪	男	普通	7163	1	33048301160951050001	2007-06-17至2007-07-13	2007-07-13 14:08:56	正常		脑出血
王阿大	女	普通	7698	1	33048306160807010001	2007-07-01至2007-07-14	2007-07-14 10:01:29	正常		脑出血
邢国华	男	普通	7684	1	33048304010532010001	2007-06-29至2007-07-04	2007-07-04 07:58:16	正常		心肌炎

全部
- 西药费
- 中成药
- 中草药
- 检查费
- 手术费
- 接生费
- 其他费
- 空调费
- 陪客费
- 伙食费
- 婴儿费
- 化验费
- 治疗费
- 材料费
- 放疗费
- 特治费

	医院项目名称	中心项目名称	单位	项目单价	项目限价	项目数量	项目金额	自负比例	自负金额	扣减数量	扣减原因	扣减金额
31	尿有形成份定量	尿有形成份(定量)	次	8.0000	0.00	1	8.000	0.00%	0.000	0		0.00
32	磁共振平扫	MRI平扫(一部位)	部位	350.0000	0.00	1	350.000	15.00%	52.500	0		0.00
33	磁共振血管成像	MRI各类成像	人次	50.0000	0.00	1	50.000	15.00%	7.500	0		0.00
34	磁共振平扫加部位	MRI平扫同时增强加收	部位	50.0000	0.00	1	50.000	15.00%	7.500	0		0.00
35	激光胶片(富士干式)14*17进口	自费材料费	张	26.0000	0.00	3	78.000	100.00%	78.000	0		0.00
36	粪便检查(粪便常规)	粪便常规	次	3.0000	0.00	1	3.000	0.00%	0.000	0		0.00
37	粪寄生虫镜检(寄生虫、原虫、虫卵镜检)	粪寄生虫镜检	次	3.0000	0.00	1	3.000	0.00%	0.000	0		0.00
38	隐血试验	粪便隐血试验(OB)	项	3.0000	0.00	1	3.000	0.00%	0.000	0		0.00
							3613.46		244.01			0.01

Waist-hip ratio and incidence of diabetes

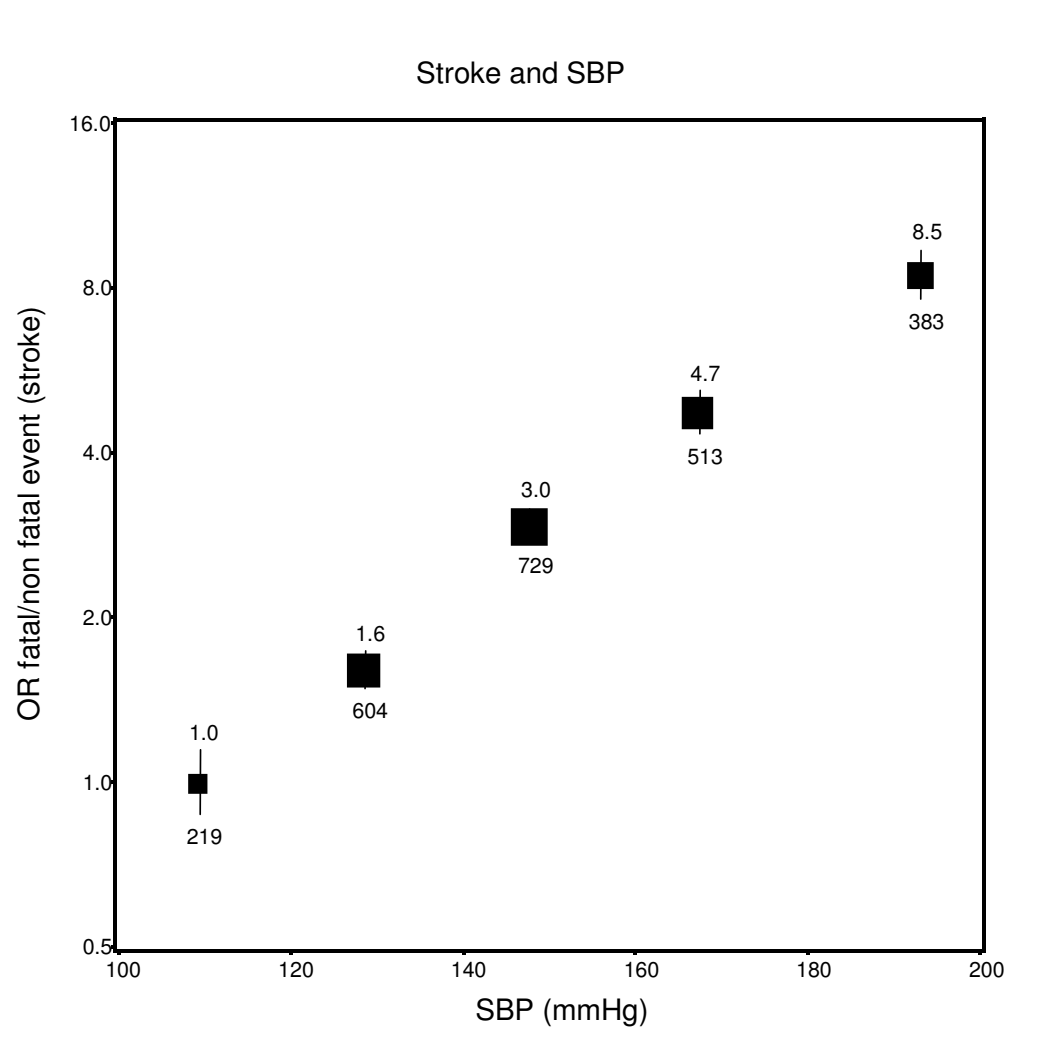
(Excluding diagnosed diabetes or abnormal blood glucose at baseline)



KSCDC
中英慢病合作项目

Systolic BP and incidence of stroke in the Kadoorie Biobank Study

(2400 new cases by July 2008, ~75% confirmed by CT/MRI)



Estimated number of incident cases for main diseases in the Kadoorie Biobank

Types	By 2013	By 2018	By 2023
Cancer	10,000	20,000	35,000
IHD	4,000	8,000	12,000
Stroke	12,000	30,000	50,000
Diabetes	2,500	6,000	10,000
COPD*	4,000	10,000	20,000
Any death	25,000	60,000	110,000

** Death only*

By 2013, there will be ~125,000 hospitalised events

KSCDC 中英慢病合作項目

Nested case-control approach for blood analyses of particular conditions

Unbiased comparison between cases and matched survivors (controls) on:

- *genome-wide association analysis*
- *lipids / lipoproteins / particle size*
- *antibodies / inflammatory factors*
- *hormones / growth factors*
- *micronutrients*
- *plus, many more (eg, proteomics)*

The sample access will be semi-public

KSCDC 中英慢病合作项目

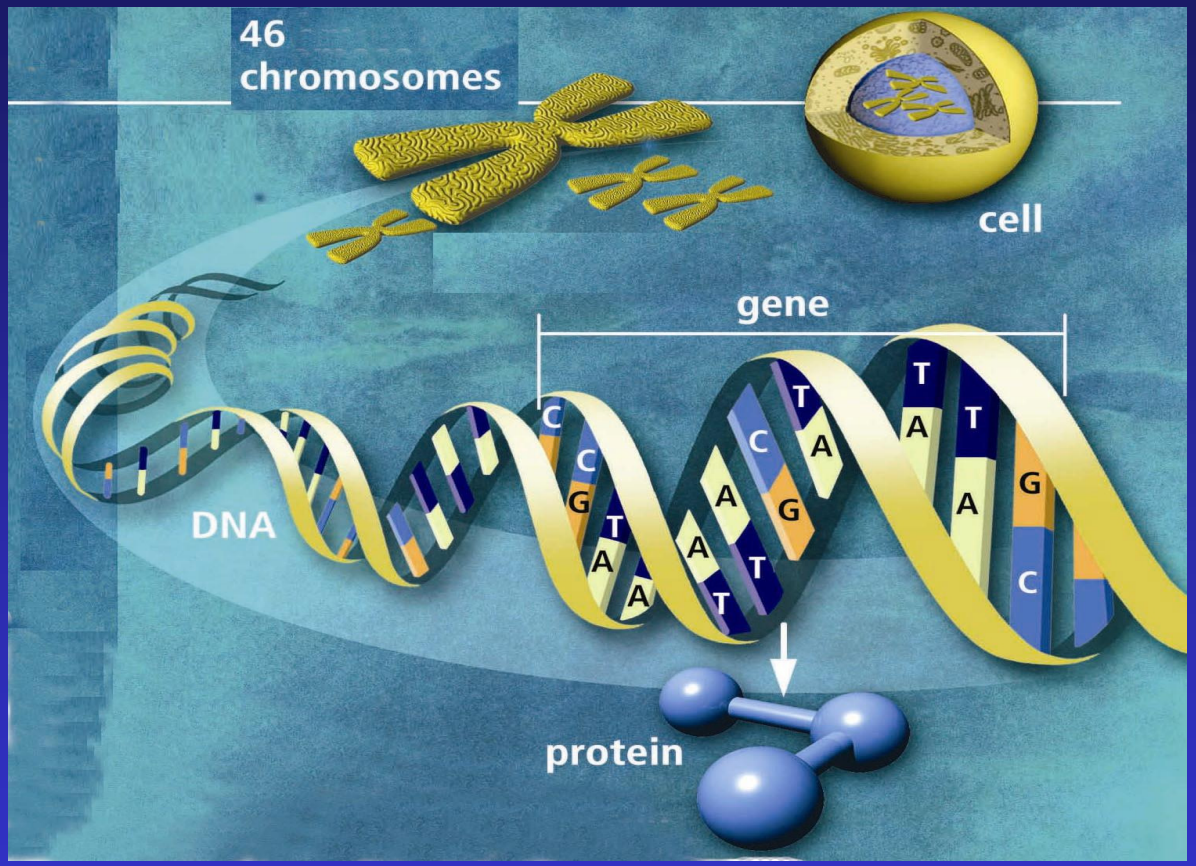
Genome-wide association study

Human genome

3000 million basepairs

10-20 million SNPs

~20 K genes



Typical genome-wide screen 0.3 -1 million SNPs

Recent success with genome-wide association studies in common chronic diseases

A Genome-Wide Association Study Identifies *IL23R* as an Inflammatory Bowel Disease Gene

Richard H. Duerr,[1,2] Kent D. Taylor,[3,4] Steven R. Brant,[5,6] John D. Rioux,[7,8] Mark S. Silverberg,[9] Mark J. Daly,[8,10] A. Hillary Steinhart,[9] Clara Abraham,[11] Miguel Regueiro,[1] Anne Griffiths,[12] Themistocles Dassopoulos,[5] Alain Bitton,[13] Huiying Yang,[3,4] Stephan Targan,[4,14] Lisa Wu Datta,[5] Emily O. Kistner,[15] L. Philip Schumm,[15] Annette T. Lee,[16] Peter K. Gregersen,[16] M. Michael Barmada,[2] Jerome I. Rotter,[3,4] Dan L. Nicolae,[11,17] Judy H. Cho[18]*

www.sciencemag.org SCIENCE VOL 314 1 DECEMBER 2006

doi:10.1038/nature05616 nature

ARTICLES

A genome-wide association study identifies novel risk loci for type 2 diabetes

Robert Sladek[1,2,4], Ghislain Rocheleau[1]*, Johan Rung[4]*, Christian Dina[5]*, Lishuang Shen[1], David Serre[1], Philippe Boutin[5], Daniel Vincent[4], Alexandre Belisle[4], Samy Hadjadj[6], Beverley Balkau[7], Barbara Heude[7], Guillaume Charpentier[8], Thomas J. Hudson[4,9], Alexandre Montpetit[4], Alexey V. Pshezhetsky[10], Marc Prentki[10,11], Barry I. Posner[2,12], David J. Balding[13], David Meyre[5], Constantin Polychronakos[1,3] & Philippe Froguel[5,14]

©2007 Nature Publishing Group

LETTERS

nature genetics

A genome-wide association scan of nonsynonymous SNPs identifies a susceptibility variant for Crohn disease in *ATG16L1*

Jochen Hampe[1,2,10], Andre Franke[1,10], Philip Rosenstiel[1,9], Andreas Till[1], Markus Teuber[1], Klaus Huse[3], Mario Albrecht[4], Gabriele Mayr[4], Francisco M De La Vega[5], Jason Briggs[5], Simone Günther[5], Natalie J Prescott[6], Clive M Onnie[6], Robert Häsler[1], Bence Sipos[7], Ulrich R Fölsch[2], Thomas Lengauer[4], Matthias Platzer[3], Christopher G Mathew[6], Michael Krawczak[8] & Stefan Schreiber[1,2]

://www.nature.com/naturegenetics

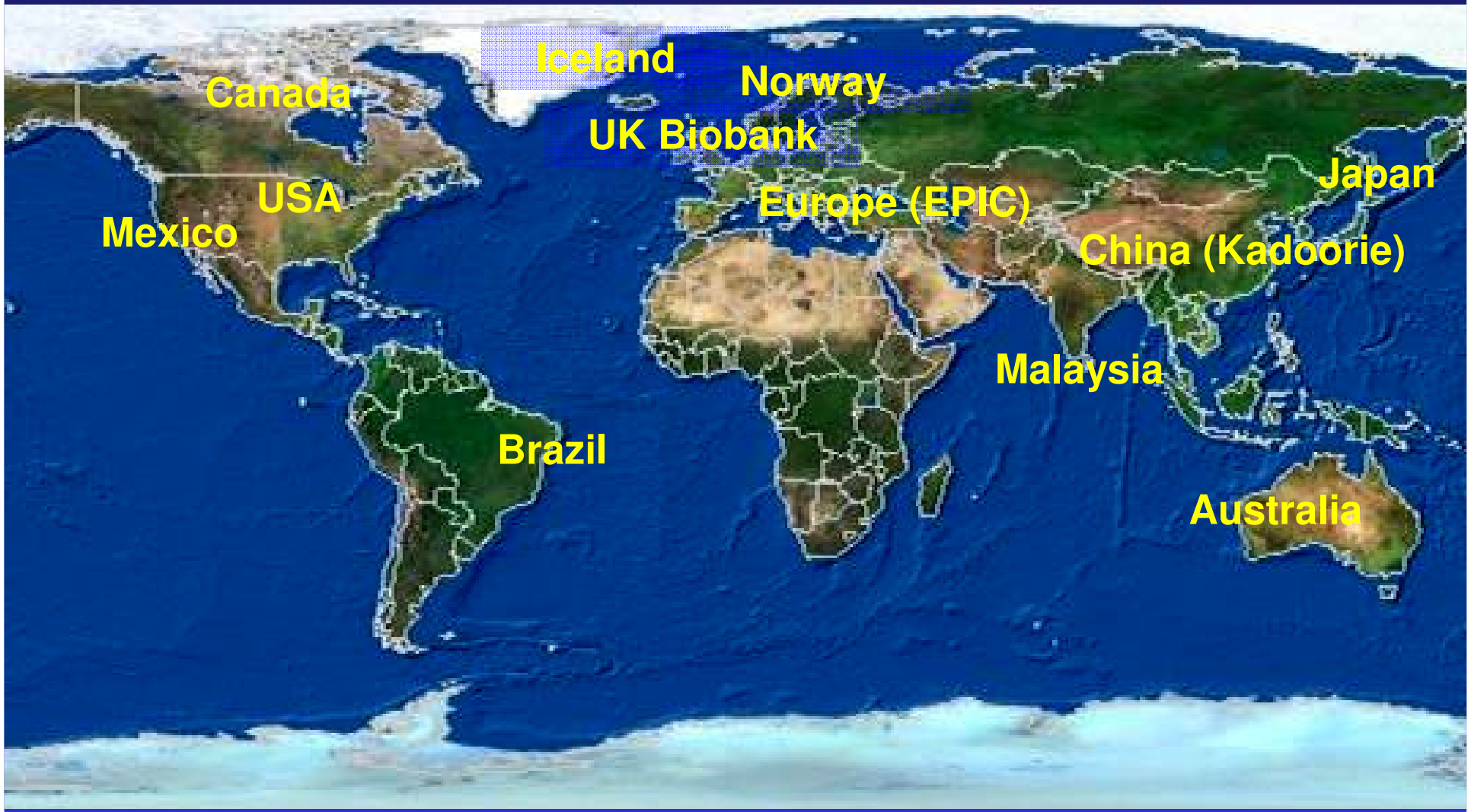
NATURE GENETICS VOLUME 39 | NUMBER 2 | FEBRUARY 2007 207

A Common Variant in the *FTO* Gene Is Associated with Body Mass Index and Predisposes to Childhood and Adult Obesity

Timothy M. Frayling,[1,2]* Nicholas J. Timpson,[3,4]* Michael N. Weedon,[1,2]* Eleftheria Zeggini,[3,5]* Rachel M. Freathy,[1,2] Cecilia M. Lindgren,[3,5] John R. B. Perry,[1,2] Katherine S. Elliott,[3] Hana Lango,[1,2] Nigel W. Rayner,[3,5] Beverley Shields,[2] Lorna W. Harries,[2] Jeffrey C. Barrett,[3] Sian Ellard,[2,6] Christopher J. Groves,[5] Bridget Knight,[2] Ann-Marie Patch,[2,6] Andrew R. Ness,[7] Shah Ebrahim,[8] Debbie A. Lawlor,[9] Susan M. Ring,[9] Yoav Ben-Shlomo,[9] Marjo-Riitta Jarvelin,[10,11] Ulla Sovio,[10,11] Amanda J. Bennett,[5] David Melzer,[1,12] Luigi Ferrucci,[13] Ruth J. F. Loos,[14] Inês Barroso,[15] Nicholas J. Wareham,[14] Fredrik Karpe,[5] Katharine R. Owen,[5] Lon R. Cardon,[3] Mark Walker,[16] Graham A. Hitman,[17] Colin N. A. Palmer,[18] Alex S. F. Doney,[19] Andrew D. Morris,[19] George Davey Smith,[4] The Wellcome Trust Case Control Consortium,† Andrew T. Hattersley,[1,2]‡§ Mark I. McCarthy[3,5]‡

www.sciencemag.org SCIENCE VOL 316 11 MAY 2007 889

Large biobank studies around the world
(Planned or existing)



Main Oxford-China KSCDC collaborators

Modern development of cohort studies

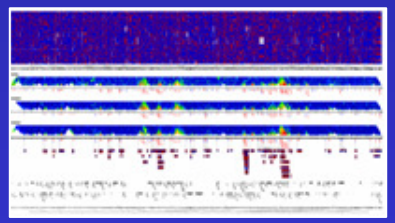
(Kadoorie Biobank, UK Biobank, EPIC, etc)

- Really large, some up to 0.5M individuals
- Not necessarily hypothesis driven, but with open-ended research objectives
- Long-term storage of biological samples
- Extensive use of IT for data collection, quality control, monitoring and management
- Ability to assess a range of health outcomes through electronic linkage

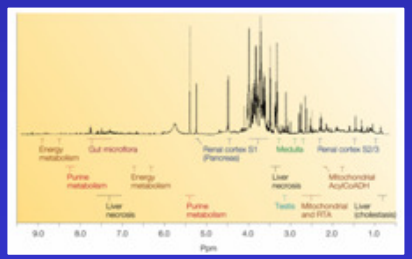
Advance in technology will revolutionise future medical research

Genomic Epidemiology

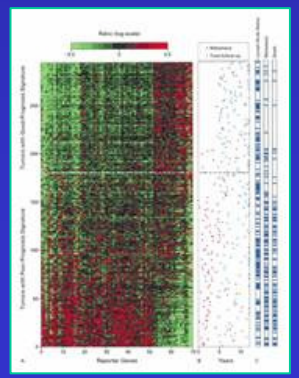
- Genes
  - Large scale genetic association



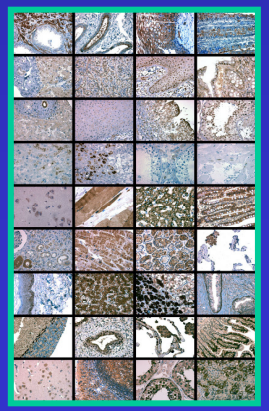
- Small molecules
  - Metabonomics



- mRNA
  - Transcription analysis



- Proteomics
  - Arrays MS



Requirements for the appropriately reliable assessment of risk factors for disease

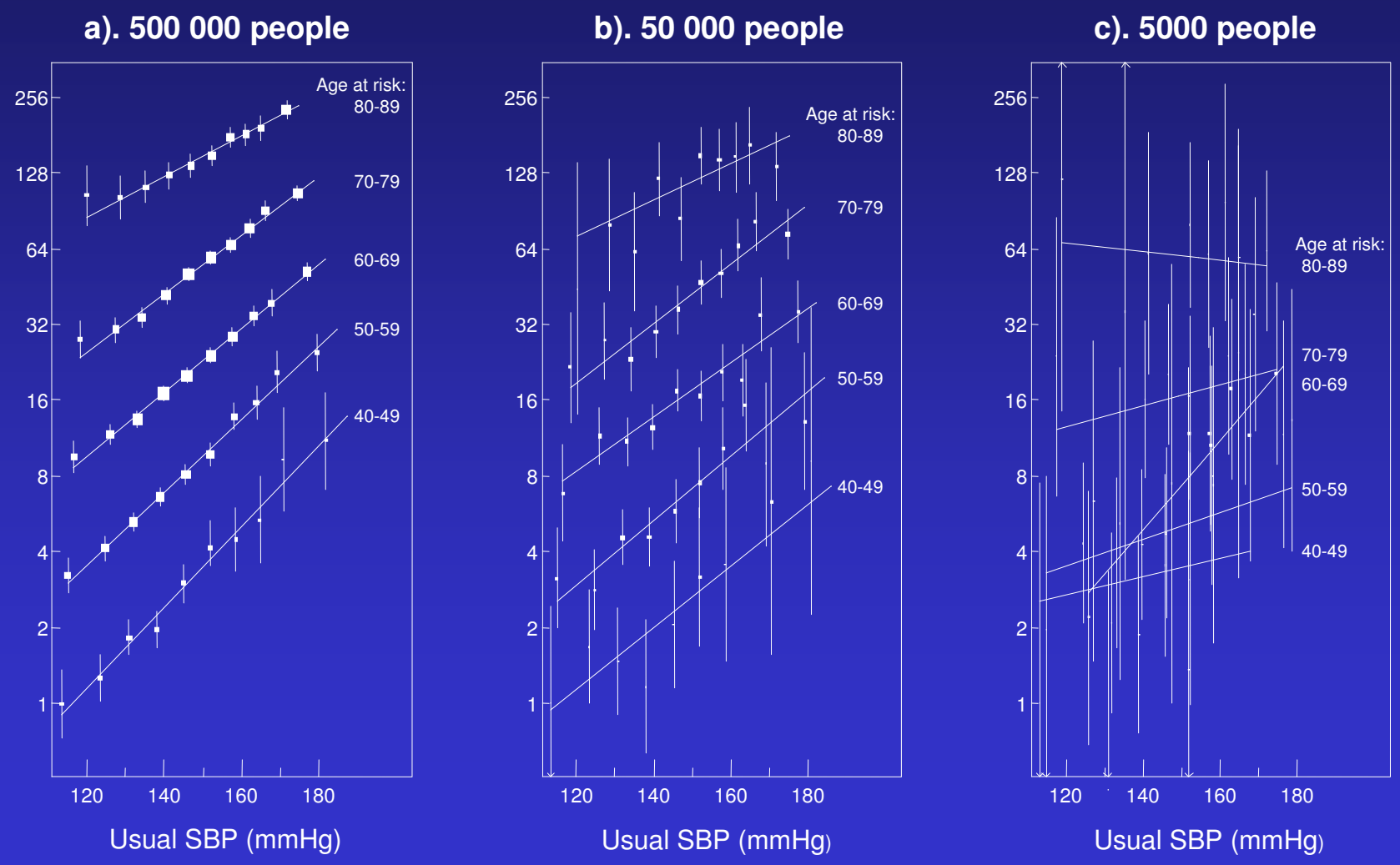
Study many more “cases” of the disease of interest in a wider range of circumstances than is currently customary

Pay greater attention to important statistical details in analyses (e.g. potential biases, confounders, and selective emphasis)

Gather information not just on death but also on a range of other outcomes (more powerful, better diagnosis)

Otherwise, risk factor associations with disease may be seriously under- and over-estimated

Size matters: Informativeness of 0.5M *vs* 50K *vs* 5K people on relation of SBP and IHD in PSC



Importance of studying the weak associations in epidemiology

Weak associations showing relative risk of, saying, less than 2 are often due to chance, bias, or confounding; but they are the best that we can hope for in most of current epidemiological research. They could be more important than strong associations with rare exposure if these associations can be established reliably to be causal and relate to common diseases and the risk exposures are prevalence in the population.

- Richard Doll