

# 基于“中国慢性病前瞻性研究”的遗传资源建设与应用

王文秀<sup>1</sup>, 黄涛<sup>1</sup>, 李立明<sup>1,2</sup>

1. 北京大学公共卫生学院流行病与卫生统计学系, 北京 100191

2. 北京大学公众健康与重大疫情防控战略研究中心, 北京 100191

**摘要:** 目前中国慢性病负担沉重, 加快推进慢性病早期预防的研究迫在眉睫。大型人群队列是慢性病病因学研究的首要设计之一, 有助于人们寻找到切实可行的慢性病干预措施。同时, 从大型人群队列的基因水平阐明危险因素与疾病的遗传关联, 也可开辟一条研究流行病学中因果关系的新途径。本文对由中国医学科学院、北京大学和英国牛津大学联合开展的“中国慢性病前瞻性研究”项目 China Kadoorie Biobank (CKB)进行了推介, 重点介绍了该项目遗传数据资源的建设进展以及近年来发表的遗传学研究成果, 以期为我国大型人群队列遗传资源的深入挖掘与利用提供参考。

**关键词:** 中国慢性病前瞻性研究; 遗传资源; 生物样本库; 大型人群队列

## Construction and application of human genetic resources in the China Kadoorie Biobank

Wenxiu Wang<sup>1</sup>, Tao Huang<sup>1</sup>, Liming Li<sup>1,2</sup>

1. Department of Epidemiology & Biostatistics, School of Public Health, Peking University, Beijing 100191, China

2. Peking University Center for Public Health and Epidemic Preparedness & Response, Beijing 100191, China

**Abstract:** China is facing a heavy burden of chronic diseases, so it is urgent to promote the relevant researches for early prevention of chronic diseases. Large population cohorts are one of the primary study designs for etiology evidence of chronic diseases, which are helpful to explore feasible intervention measures. Moreover, clarifying the genetic associations between risk factors and diseases from the genetic level of large population cohort can also open up a new way to the exploration of causality. This article aimed to introduce the study “China Kadoorie Biobank (CKB)” jointly carried out by the Chinese Academy of Medical Sciences, Peking University and University of Oxford, with emphasis on the construction progress of genetic resources and the recent published genetic studies, in order to provide reference for the in-depth mining and utilization of genetic resources in large population cohort in China.

**Keywords:** China Kadoorie Biobank; genetic resources; biobank; large population cohort

收稿日期: 2021-07-19; 修回日期: 2021-09-07

基金项目: 国家重点研发计划项目(编号 2020YFC2003401)资助[Supported by the National Key R&D Program of China (No. 2020YFC2003401)]

作者简介: 王文秀, 在读硕士研究生, 专业方向: 流行病与卫生统计学。E-mail: wwenxiu1996@126.com

通讯作者: 黄涛, 博士, 研究员, 研究方向: 慢性病流行病学。E-mail: huang.tao@pku.edu.cn

DOI: 10.16288/j.ycz.21-257

网络出版时间: 2021/9/14 16:26:34

URI: <https://kns.cnki.net/kcms/detail/11.1913.R.20210914.1433.002.html>

随着我国经济社会的飞速发展,国民生活方式和膳食营养结构发生了深刻的变化。加之人口老龄化及城镇化进程的加速,种种因素使得我国居民的疾病谱发生了很大的转变,慢性非传染性疾病的患者基数不断扩大,同时因慢性病死亡的比例也在不断增加。《中国居民营养与慢性病状况报告(2020年)》显示,我国 18 岁以上居民高血压患病率为 27.5%,糖尿病患病率为 11.9%,慢性阻塞性肺病患病率为 13.6%,因慢性病导致的死亡占总死亡率的 88.5%<sup>[1]</sup>。由此可见,中国慢性病负担沉重,预防、控制和治疗慢性病的工作迫在眉睫。解决这个问题应该从一级预防入手,综合防控慢性病危险因素。而目前已有的病因学研究多来自西方发达国家人群,并不完全适用于中国人群。因此,获得中国人群的慢性病病因学证据,寻找更加切实可行的干预措施,具有迫切的理论与实践意义。

队列研究是流行病学研究最基本的研究设计之一,是一种重要的评价疾病危险因素的流行病学研究方法。小样本队列因样本量的限制常导致统计效率的不足,而大型人群队列因样本量大、随访时间长、覆盖广泛的人群与地区特征等优势,使其有机会更深入地探讨复杂的基因-环境相互作用、剂量-反应关系以及不同特征人群的亚组分析等,从而可以为病因学研究提供更为强有力的证据。我国人群队列建设具有约 50 年历史,相比于西方队列存在起步较晚、随访时间较短和规模较小等局限<sup>[2]</sup>。近 20 余年来,我国陆续建成了若干大型人群队列,其中重点关注慢性病的队列有:51.3 万样本量的中国慢性病前瞻性研究(China Kadoorie Biobank, CKB)、20 万样本量的泰州人群健康跟踪调查、7.5 万样本量的上海女性健康研究和 6.2 万样本量的上海男性健康研究等<sup>[3-6]</sup>。目前,以个体化医疗为特征的精准医学模式成为了研究焦点。“十三五”期间,“精准医学研究”被列为我国重点研究专项之一,启动了百万人以上的自然人群国家大型健康队列和重大疾病专病队列,旨在实施精准医学研究的协同攻关。目前已完成立项的研究包括京津冀自然人群队列研究、西南区域自然人群队列研究、呼吸系统疾病专病队列研究和乳腺癌专病队列研究等。

随着以测序技术与质谱技术为代表的高通量组学技术的发展,近年来的人群队列致力于采集、长

期保藏、利用研究对象的生物样本,建立了专门的队列生物样本库。流行病学家借助丰富的生物样本数据,结合传统流行病学的宏观性,通过建立大数据、多水平的计算模型,可以更为系统地阐明由因及果的致病机制,系统流行病学(systems epidemiology)的概念因此也应运而生<sup>[7]</sup>。系统流行病学通过生物样本库多水平、多层次的丰富数据,构建整合的模型,完善了传统流行病学的局限性,在探索病因和寻找可能的干预靶点中发挥着重要的作用。CKB 作为目前全球仅有的几项建立生物样本库的超大型前瞻性队列研究之一,集成了丰富的中国人群遗传资源与数据信息资源。因此,本文将以 CKB 为例,概述该项目的研究设计、遗传资源的采集与管理方法,总结其已发表的遗传学证据,为我国大型人群队列的建设提供参考。

## 1 CKB 研究项目概述

CKB 是 2002 年由卫生部批准立项,现由中国医学科学院、北京大学和英国牛津大学联合开展的一项大型慢性病国际合作研究项目。该项目旨在从遗传、环境和生活方式等多个层次和水平入手,深入研究危害中国人群健康的主要慢性疾病(如脑卒中、冠心病、糖尿病、癌症等)的致病因素、保护性因素、发病机制及流行规律和趋势,为有效制定重大慢性病防控策略和指南,开发新的治疗和干预手段,提供高质量的病因学证据。

本项目于 2004 年 6 月至 2008 年 8 月间在中国 10 个省/地区(黑龙江省哈尔滨市、山东省青岛市、江苏省苏州市、浙江省桐乡市、河南省辉县市、甘肃省天水市、四川省彭州市、广西壮族自治区柳州市、海南省海口市和湖南省浏阳市)开展,项目点的选择综合考虑了城市和农村的比例、各类慢性病的发病率和主要危险因素的分布、人口稳定性,以及死因和疾病登记报告制度的完整性等因素,并经过现场调研考核后确定。本项目采取整群抽样的方式,以城市区和农村县为单位进行募集和调查,共计招募了 512,891 名年龄在 30~79 岁、具有知情同意书及完整基线调查数据的研究对象,其中男性占 41%,农村人口占 56%,平均年龄为 52 岁<sup>[3,8]</sup>。项目于 2008 年、2014 年及 2020 年在各项目点随机抽取



图 1 CKB 项目基线与重复调查信息

Fig. 1 Baseline and resurvey information in CKB

约 5% 的研究对象各完成一次重复调查, 以了解队列人群中生活方式等因素的动态变化情况。基线及三次重复调查的开展时间、调查人数及调查内容见图 1。

在队列成员参加基线调查后 6 个月, 开展针对全部队列成员的结局随访工作。随访期间的死亡与发病信息通过多种途径获得, 包括当地卫生部门现有的死亡和常规疾病监测系统及社区工作人员主动定向监测。队列成员的医保覆盖率已经达到 98%, 医疗保险数据库也成为了补充死亡、常规发病信息的重要途径。为了保证结局随访工作的质量, 项目组还开展了对主要慢性病发病事件的病例复核工作, 确保报告和诊断的准确性。

## 2 队列遗传资源的采集与管理

CKB 作为一项多因素、多结局、多学科的大型慢性流行病学研究, 建立了世界上领先规模的成人生物样本库, 存储了项目 51.3 万人及重复调查的所有生物样本, 存储规模达到 58.8 万份<sup>[2]</sup>。与其他国内外队列相比, 其样本量庞大, 覆盖了我国东北、华东、华南、西北和西南具有不同文化背景、经济水平以及暴露谱和疾病谱的城市和农村地区; 随访时间长, 截至目前, 队列平均随访 13 年, 累计随访 600 余万人年。CKB 项目数据量大、类型多元、来源复杂。数据资源上既包括问卷调查、体格检查数据, 还包括了血生化检测数据、基因分型数据、代谢组学数据和结局事件数据等。因此, 数据的标准化采集、规范化管理成为队列建设的一项重要议题。基于此, CKB 项目以十余年来的成功建设经验为基础, 撰写并发布了一套完善的大型人群队列研究的规范化团体标准, 包括《大型人群队列现场调查管

理技术规范(T/CPMA 001-2019)》<sup>[9]</sup>、《大型人群队列研究数据安全技术规范(T/CPMA 002-2018)》<sup>[10]</sup>等, 以期为其他队列研究提供参考。

### 2.1 人类生物样本资源的采集与管理

CKB 项目建立了生物样本采集、分装、运输、存储等一系列的电子化标准操作程序和管理系统, 对所有样本实施全方位监管, 最大限度地降低人为错误的概率。所有生物样本的采集均获得了研究对象的知情同意, 由经过统一培训的专业技术人员根据标准操作程序完成采集和分装, 并统一运输至 CKB 项目样本库储存。项目样本库严格按照样本存储空间、电源、网络、温度、监控等各项标准进行管理。

CKB 项目的基因测序工作由武汉华大基因(Beijing Genomics institution, BGI)完成, 统一采用针对中国人群单倍体信息设计的芯片(Affymetrix Axiom myDesign)检测。项目组先后选取了两批研究对象, 采用两种不同版本的芯片进行测定。第一批选取了约 3.2 万名研究对象, 于 2015 年采用 700K 芯片进行基因测序, 其中包括 5000 名脑出血病例、5000 名缺血性中风病例、5000 名慢阻肺病例、1.5 万余名健康对照等。第二批采用更新后的 800K 芯片, 对队列中随机选取的约 5 万名研究对象, 以及第二次重复调查的 2.5 万人进行测序。后期参照千人基因组计划第三阶段数据进行填补, 最终填补 600 万~1000 万位点信息, 对全基因组的常见变异实现了大于 93% 的覆盖。

### 2.2 队列数据资源的采集与管理

队列遗传资源还包括与之相关联的健康数据来源信息, 主要来源于现场调查和长期随访监测, 包

括人口和社会经济学因素、生活方式、疾病史与用药史等详细的问卷数据, 身高、体重、腰臀围、肺功能、血压等体测数据以及随访结局事件数据等。

现场调查期间, 均由通过统一培训的工作人员进行数据资源采集。问卷调查人员熟悉当地语言、生活方式与习惯, 并采用计算机辅助面访及询问时录音等方式保证问卷质量。体检人员具有较强的医学背景, 具备熟练操作相关设备的能力, 考核通过方可开展体检, 并接受定期考核。不同项目点地区现场调查中用到的设备、物资等统一配备、规格一致, 定期维护。同时, 采用信息关联的方式获取结局信息, 与当地死亡和疾病监测、医保系统有机整合, 以及社区工作人员主动定向监测、项目组病例复核, 保证病例的真实性。

为了保证队列调查收集数据的准确性和完整性, 保障项目的质量控制、监测、长期随访, 项目组开发了一系列计算机系统, 可实现远程数据库管理、库存控制和设备校准、长期随访系统的链接等。现场调查产生的各种文件, 按保密级别及重要程度进行存档。涉及调查对象隐私数据的文件, 确保数据信息的安全, 设置相关责任人, 严格责任制。定期进行统计学监测, 及时了解数据质量及数据库动态, 识别可能存在的问题并及时反馈, 以保证数据质量。

### 3 队列遗传资源的利用

CKB 项目目前已经在遗传学研究中取得了重大阶段性成果, 相关研究论文在 *The Lancet*、*Nature Medicine*、*The Lancet Oncology*、*JAMA Cardiology* 等国际重要学术期刊发表, 产生了广泛的国际影响, 为探索我国人群慢性病病因的科学证据, 开发适合我国人群的遗传风险评分, 指导精准预防提供了重要参考。

#### 3.1 心脑血管疾病

酗酒是危害心血管健康的一大原因, 但是既往前瞻性研究发现适量摄入酒精对中风和心肌梗死具有一定的保护作用<sup>[11]</sup>。CKB 研究结果显示, 自报饮酒量与中风、急性心肌梗死的发病风险呈 U 型轨迹, 在调整了年龄、性别、教育程度、收入与吸烟状况后, 在男性中, 相比不饮酒以及酗酒者, 每周饮酒

100 g 的研究对象患病风险最低<sup>[12]</sup>。然而, 遗传研究证据则并不支持适度饮酒预防中风的保护性作用。在男性中, 基因型与地区决定的酒精摄入量的增加与脑卒中发病风险呈连续的正相关, 表现为每增加 280 g, 对应的缺血性与出血性中风的 *HR* 为 1.27 (95%*CI*: 1.13~1.43) 和 1.58 (1.36~1.84)。而女性由于饮酒率很低, 基因型不能反映实际的酒精摄入量, 与血压、中风和心肌梗死也不存在相关性。观察性研究与遗传证据结果的不一致可能是因为适量饮酒的保护性作用不一定来自于酒精本身, 研究对象本身健康状况不佳可能会影响酒精摄入, 从而导致其发病风险高于适量饮酒者。

血脂水平也与心脑血管疾病发病密切相关。既往观察性研究和随机对照试验已经证明, 低密度胆固醇(low-density lipoprotein, LDL)是缺血性心脏病和缺血性中风的重要危险因素之一<sup>[13,14]</sup>, 而 LDL 与脑出血风险之间的关系仍没有定论。此外, 中国人群与西方人群相比, 尽管有着较低的平均 LDL 浓度, 但是脑出血的发病率较高<sup>[15]</sup>, 并不能简单套用西方研究的结论。因此, CKB 项目在巢式病例对照研究中比较生化测定的 LDL 与出血性与缺血性中风的关联, 以及采用孟德尔随机化方法证实两种中风类型与 LDL 的因果关联<sup>[16]</sup>。研究显示, 血浆 LDL 每降低 1 mmol/L, 缺血性与出血性中风 *RR*=0.85 (0.80~0.91) 和 1.16(1.08~1.25)。基因预测的 LDL 每降低 1 mmol/L, 缺血性与出血性中风 *RR*=0.75 (0.60~0.95) 和 1.13 (0.91~1.40)。研究提示, 降低 LDL 对预防总体中风风险和主要冠状动脉事件依然具有净收益。此外, CKB 项目利用队列 15 万名研究对象的数据, 分析胆固醇酯转运蛋白(cholesterol ester transfer protein, CETP)基因中的变异, 进而评估高密度胆固醇(high-density lipoprotein, HDL)与心脑血管疾病的遗传性关联<sup>[17]</sup>。研究发现 CETP 变异与显著升高 HDL 水平, 但并未改变 LDL 水平, 与心脑血管疾病发病风险也无显著关联。这表明在不降低 LDL 水平的情况下, 单纯增加 HDL 并不能有效降低心脑血管疾病发病风险, 为药物研究路径提供了新的科学证据。

欧洲人群中的观察性研究显示, 血浆维生素 D (25[OH]D)低水平与较高的心脑血管疾病发病与死亡风险相关<sup>[18]</sup>。然而, 尚不清楚 25[OH]D 是导致心脑血管疾病风险降低的原因, 还是仅仅是健康状况



不佳的结果。CKB 项目与哥本哈根一般人群研究(Copenhagen General Population Study, CGPS)及哥本哈根心脏研究(Copenhagen City Heart Study, CCHS)合作,以证实维生素 D 是否与特定血管疾病风险存在因果关联<sup>[19]</sup>。孟德尔随机化研究结果显示,无论在中国还是欧洲成人中,均未发现血浆 25[OH]D 水平(通常用于评估维生素 D 水平)与中风、心肌梗死、总体心脑血管病发病或死亡之间存在因果关联。同样,25[OH]D 水平与血脂成分之间也不存在因果关联。CKB 结果表明,观察性研究中维生素 D 与心脑血管疾病之间的负相关可能是混杂的结果。

### 3.2 糖尿病

全基因组关联研究发现了许多 2 型糖尿病相关的遗传变异。然而,既往研究多在欧洲人群中进行,缺少中国人群的证据;通常基于高度选择的病例对照样本,缺少对人群水平效应大小的可靠估计。2016 年 CKB 项目在具有基因分型的 93,000 名研究对象中,检验了 56 个既往研究确定的位点与 2 型糖尿病的风险关联。其中,48 个位点与既往文献的效应方向一致,为欧洲人和东亚人之间 2 型糖尿病的共同遗传结构提供了进一步的证据<sup>[20]</sup>。另外,改进的位点效应大小估计可以构建更加准确的加权遗传风险评分,更好地应用于疾病预测。2020 年 CKB 的一项研究即采用以上验证的位点构建遗传风险评分,研究 2 型糖尿病遗传风险评分与健康生活方式(根据 BMI、饮酒、吸烟、体力活动和饮食定义)的交互作用与 2 型糖尿病发病风险的关联<sup>[21]</sup>。研究发现,高遗传风险组(最高五分位数)患 2 型糖尿病的风险比低遗传风险组(最低五分位数)高出 79% (60%, 100%)。在具有高遗传风险的人群中,采取健康生活方式的参与者患 2 型糖尿病的风险比不健康生活方式者低 63% (56%, 69%)。

### 3.3 恶性肿瘤

肺癌是中国发病率与死亡率最高的癌症类型<sup>[22]</sup>,其中非小细胞肺癌(non-small cell lung cancer, NSCLC)约占肺癌总数的 85%<sup>[23]</sup>。然而,肺癌的主要遗传位点尚未完全确定,尤其是在中国人群中。2019 年发表的纳入 27,120 例病例与 27,355 例对照的全基因组关联研究,发现了 19 个与 NSCLC 相关的遗传位

点<sup>[24]</sup>。利用这些位点,在应用于 CKB 研究的 95,408 名研究对象时,位点的多基因遗传评分与肺癌的发病率存在剂量反应关联( $P_{趋势}=2.02\times 10^{-9}$ ),成功验证了其在肺癌风险预测中的有效性,并有可能应用于肺癌的个体化预防。

超过 40% 的胃癌新发与死亡病例发生在中国<sup>[22]</sup>。2020 年发表的全基因组关联荟萃分析纳入了 10,254 例病例及 10,914 例对照,并利用发现的 112 个遗传位点构建多基因风险得分。在 CKB 项目的 100,220 名研究对象中,研究基因与健康生活方式因素(定义为不吸烟、从不饮酒、较少摄入腌制食品以及经常摄入新鲜水果)的交互作用与胃癌发病风险的关联<sup>[25]</sup>。研究发现,多基因风险得分与胃癌发病风险显著相关( $P_{趋势}<0.0001$ );与同样具有高遗传风险但采取不健康生活方式的对象相比,采取健康生活方式的研究对象胃癌发病风险较低,HR 值为 0.53(0.29~0.99),其绝对风险降低 1.12%(0.62~1.56)。

### 3.4 其他慢性病

肺功能受损是反映慢性阻塞性肺病的重要指标。然而,既往肺功能的全基因组关联研究都基于欧洲人群,缺乏亚洲人群的关联证据。CKB 项目利用 100,285 名研究对象的遗传资源数据,开展了目前为止最大的亚裔肺功能全基因组关联研究,识别了 18 个既往未报道过的肺功能相关基因位点,并发现肺功能和肥胖指标之间存在共同的遗传基础<sup>[26]</sup>。该研究提供了中国人群肺功能易感基因的证据,从遗传的角度揭示了肺功能与肥胖的关联,这些发现可能为未来开发个体化干预治疗以预防肺功能下降开辟新的途径。

东亚人群中 BMI 与肝胆疾病的关联证据有限。2020 年发表的一篇孟德尔随机化研究发现,基因预测的 BMI 与肝胆疾病风险增加有关,在疾病亚型之间也观察到一致的关联<sup>[27]</sup>。研究表明,肥胖是中国人群患有肝胆疾病的一项重要危险因素。因此,研究支持采用生活方式干预的手段,包括减重和饮食干预,以达到肝胆疾病的一级预防。

## 4 结语与展望

目前,CKB 项目已经获得了一系列适用于中国

人群的慢性病病因学遗传证据, 成为了我国高质量病因学证据的重要来源, 对制定慢性病防控的策略和措施和指导我国慢性病精准防治具有重要意义。由于大样本队列研究是一项长期系统的工程, CKB 项目已对 10 万余名研究对象进行了基因分型, 但后续样本的基因测序仍需要投入大量的经费与人力资源。此外, 项目组对表观遗传组学、蛋白组学、代谢组学等多组学数据的检测与利用较少, 有待进一步开展多时点多组学检测, 构建精细队列。基于 CKB 高质量的人群队列生物样本数据, 如果可以得到足够的经费支持、长期维持队列, 必将带来大量的研究机遇, 产出更多的研究成果, 为我国慢性病流行病学领域的发展做出突出贡献。

“健康中国 2030”战略的推进需要具有更广泛人群和地域特征的中国人队列研究。包括 CKB 研究在内的大型人群队列以及已逐步建立完善的“精准医学研究”专项队列, 正在形成一个多层次的生物医学大数据平台和遗传多样化的数据资源体系。未来, 通过应用更先进的大数据网络分析处理方法并整合多层次多组学的数据, 这些队列将在分子、细胞、组织、社会与生态环境等多水平上进行深入地挖掘与探索, 进一步解释和明确疾病发生发展过程, 为疾病的预防、控制和治疗提供重要支持; 依靠智能手机、可穿戴移动设备等移动终端收集更为详尽、准确的个体化暴露组学信息; 加强数据共享平台的建设, 优化数据清理、开发和有机整合, 同时密切关注生物信息安全, 将会是大型人群队列遗传数据资源的发展方向。

## 参考文献(References):

- [1] National Health Commission of the People's Republic of China. Report on the nutrition and chronic diseases status of Chinese residents 2020. Beijing: Peking University Medical Press, 2020.  
国家卫生健康委员会. 中国居民营养与慢性病状况报告(2020 年). 北京: 北京大学医学出版社, 2020. [DOI]
- [2] Li LM, Lv J. Large prospective cohort studies: a review and update. *Chin J Epidemiol*, 2015, 36(11): 1187–1189.  
李立明, 吕筠. 大型前瞻性人群队列研究进展. 中华流行病学杂志, 2015, 36(11): 1187–1189. [DOI]
- [3] Chen ZM, Chen JS, Collins R, Guo Y, Peto R, Wu F, Li LM. China Kadoorie Biobank of 0.5 million people: survey methods, baseline characteristics and long-term follow-up. *Int J Epidemiol*, 2011, 40(6): 1652–1666. [DOI]
- [4] Zheng W, Chow WH, Yang G, Jin F, Rothman N, Blair A, Li HL, Wen WQ, Ji BT, Li Q, Shu XO, Gao YT. The Shanghai Women's Health Study: rationale, study design, and baseline characteristics. *Am J Epidemiol*, 2005, 162(11): 1123–1131. [DOI]
- [5] Shu XO, Li HL, Yang G, Gao J, Cai H, Takata Y, Zheng W, Xiang YB. Cohort profile: the Shanghai Men's Health Study. *Int J Epidemiol*, 2015, 44(3): 810–818. [DOI]
- [6] Wang XF, Lu M, Qian J, Yang YJ, Li SL, Lu D, Yu SZ, Meng W, Ye WM, Jin L. Rationales, design and recruitment of the Taizhou Longitudinal Study. *BMC Public Health*, 2009, 9: 223. [DOI]
- [7] Huang T, Li LM. Systems epidemiology. *Chin J Epidemiol*, 2018, 39(05): 694–699.  
黄涛, 李立明. 系统流行病学. 中华流行病学杂志, 2018, 39(05): 694–699. [DOI]
- [8] Chen ZM, Lee L, Chen JS, Collins R, Wu F, Guo Y, Linksted P, Peto R. Cohort profile: the Kadoorie Study of Chronic Disease in China (KSCDC). *Int J Epidemiol*, 2005, 34(6): 1243–1249. [DOI]
- [9] Chinese Preventive Medicine Association. Technical specification of management for field survey in large population-based cohort study (T/CPMA 001-2019). *Chin J Epidemiol*, 2019, 40(7): 739–747.  
中华预防医学会. 大型人群队列现场调查管理技术规范(T/CPMA 001-2019). 中华流行病学杂志, 2019, 40(7): 739–747. [DOI]
- [10] Chinese Preventive Medicine Association. Technical specification of data security for large population-based cohort study (T/CPMA 002-2018). *Chin J Epidemiol*, 2019, 40(1): 12–16.  
中华预防医学会. 大型人群队列研究数据安全技术规范(T/CPMA 002-2018). 中华流行病学杂志, 2019, 40(1): 12–16. [DOI]
- [11] Wood AM, Kaptoge S, Butterworth AS, Willeit P, Warnakula S, Bolton T, Paige E, Paul DS, Sweeting M, Burgess S, Bell S, Astle W, Stevens D, Koulman A, Selmer RM, Verschuren WMM, Sato S, Njølstad I, Woodward M, Salomaa V, Nordestgaard BG, Yeap BB, Fletcher A, Melander O, Kuller LH, Balkau B, Marmot M, Koenig W, Casiglia E, Cooper C, Arndt V, Franco OH, Wennberg P, Gallacher J, De La Cámara AG, Völzke H, Dahm CC, Dale CE, Bergmann MM, Crespo CJ, Van Der

- Schouw YT, Kaaks R, Simons LA, Lagiou P, Schoufour JD, Boer JMA, Key TJ, Rodriguez B, Moreno-Iribas C, Davidson KW, Taylor JO, Sacerdote C, Wallace RB, Quiros JR, Tumino R, Blazer 2nd DG, Linneberg A, Daimon M, Panico S, Howard B, Skeie G, Strandberg T, Weiderpass E, Nietert PJ, Psaty BM, Kromhout D, Salamanca-Fernandez E, Kiechl S, Krumholz HM, Grioni S, Palli D, Huerta JM, Price J, Sundstrom J, Arriola L, Arima H, Travis RC, Panagiotakos DB, Karakatsani A, Trichopoulou A, Kuhn T, Grobbee DE, Barrett-Connor E, Van Schoor N, Boeing H, Overvad K, Kauhanen J, Wareham N, Langenberg C, Forouhi N, Wennberg M, Despres JP, Cushman M, Cooper JA, Rodriguez CJ, Sakurai M, Shaw JE, Knuiman M, Voortman T, Meisinger C, Tjonneland A, Brenner H, Palmieri L, Dallongeville J, Brunner EJ, Assmann G, Trevisan M, Gillum RF, Ford I, Sattar N, Lazo M, Thompson SG, Ferrari P, Leon DA, Smith GD, Peto R, Jackson R, Banks E, Di Angelantonio E, Danesh J. Risk thresholds for alcohol consumption: combined analysis of individual-participant data for 599 912 current drinkers in 83 prospective studies. *Lancet*, 2018, 391(10129): 1513–1523. [DOI]
- [12] Millwood IY, Walters RG, Mei XW, Guo Y, Yang L, Bian Z, Bennett DA, Chen YP, Dong CX, Hu RY, Zhou G, Yu B, Jia WF, Parish S, Clarke R, Davey Smith G, Collins R, Holmes MV, Li LM, Peto R, Chen ZM. Conventional and genetic evidence on alcohol and vascular disease aetiology: a prospective study of 500 000 men and women in China. *Lancet*, 2019, 393(10183): 1831–1842. [DOI]
- [13] Baigent C, Blackwell L, Emberson J, Holland LE, Reith C, Bhala N, Peto R, Barnes EH, Keech A, Simes J, Collins R. Efficacy and safety of more intensive lowering of LDL cholesterol: a meta-analysis of data from 170,000 participants in 26 randomised trials. *Lancet*, 2010, 376(9753): 1670–1681. [DOI]
- [14] Di Angelantonio E, Sarwar N, Perry P, Kaptoge S, Ray KK, Thompson A, Wood AM, Lewington S, Sattar N, Packard CJ, Collins R, Thompson SG, Danesh J. Major lipids, apolipoproteins, and risk of vascular disease. *JAMA*, 2009, 302(18): 1993–2000. [DOI]
- [15] Zhou MM, Wang HD, Zhu J, Chen WQ, Wang LH, Liu SW, Li YC, Wang LJ, Liu YN, Yin P, Liu JM, Yu SC, Tan F, Barber RM, Coates MM, Dicker D, Fraser M, González-Medina D, Hamavid H, Hao YT, Hu GH, Jiang GH, Kan HD, Lopez AD, Phillips MR, She J, Vos T, Wan X, Xu GL, Yan LLJ, Yu CH, Zhao Y, Zheng YF, Zou XN, Naghavi M, Wang Y, Murray CJL, Yang GH, Liang XF. Cause-specific mortality for 240 causes in China during 1990-2013: a systematic subnational analysis for the Global Burden of Disease Study 2013. *Lancet*, 2016, 387(10015): 251–272. [DOI]
- [16] Sun LL, Clarke R, Bennett D, Guo Y, Walters RG, Hill M, Parish S, Millwood IY, Bian Z, Chen YP, Yu CQ, Lv J, Collins R, Chen JS, Peto R, Li LM, Chen ZM. Causal associations of blood lipids with risk of ischemic stroke and intracerebral hemorrhage in Chinese adults. *Nat Med*, 2019, 25(4): 569–574. [DOI]
- [17] Millwood IY, Bennett DA, Holmes MV, Boxall R, Guo Y, Bian Z, Yang L, Sansome S, Chen YP, Du HD, Yu CQ, Hacker A, Reilly DF, Tan YL, Hill MR, Chen JS, Peto R, Shen HB, Collins R, Clarke R, Li LM, Walters RG, Chen ZM. Association of CETP gene variants with risk for vascular and nonvascular diseases among Chinese adults. *JAMA Cardiol*, 2018, 3(1): 34–43. [DOI]
- [18] Al Mheid I, Quyyumi AA. Vitamin D and cardiovascular disease: controversy unresolved. *J Am Coll Cardiol*, 2017, 70(1): 89–100. [DOI]
- [19] Huang T, Afzal S, Yu CQ, Guo Y, Bian Z, Yang L, Millwood IY, Walters RG, Chen YP, Chen NY, Gao RQ, Chen JS, Clarke R, Chen ZM, Ellervik C, Nordestgaard BG, Lv J, Li LM. Vitamin D and cause-specific vascular disease and mortality: a Mendelian randomisation study involving 99,012 Chinese and 106,911 European adults. *BMC Med*, 2019, 17(1): 160. [DOI]
- [20] Gan W, Walters RG, Holmes MV, Bragg F, Millwood IY, Banasik K, Chen YP, Du HD, Iona A, Mahajan A, Yang L, Bian Z, Guo Y, Clarke RJ, Li LM, McCarthy MI, Chen ZM. Evaluation of type 2 diabetes genetic risk variants in Chinese adults: findings from 93,000 individuals from the China Kadoorie Biobank. *Diabetologia*, 2016, 59(7): 1446–1457. [DOI]
- [21] Li HX, Khor CC, Fan JN, Lv J, Yu CQ, Guo Y, Bian Z, Yang L, Millwood IY, Walters RG, Chen YP, Yuan JM, Yang Y, Hu C, Chen JS, Chen ZM, Koh WP, Huang T, Li LM. Genetic risk, adherence to a healthy lifestyle, and type 2 diabetes risk among 550,000 Chinese adults: results from 2 independent Asian cohorts. *Am J Clin Nutr*, 2020, 111(3): 698–707. [DOI]
- [22] Chen WQ, Zheng RS, Baade PD, Zhang SW, Zeng HM, Bray F, Jemal A, Yu XQ, He J. Cancer statistics in China, 2015. *CA Cancer J Clin*, 2016, 66(2): 115–132. [DOI]
- [23] Alberg AJ, Brock MV, Ford JG, Samet JM, Spivack SD. Epidemiology of lung cancer: Diagnosis and management of lung cancer, 3rd ed: American College of Chest

- Physicians evidence-based clinical practice guidelines. *Chest*, 2013, 143(Suppl 5): e1S–e29S. [DOI]
- [24] Dai JC, Lv J, Zhu M, Wang YZ, Qin N, Ma HX, He YQ, Zhang RX, Tan W, Fan JY, Wang TP, Zheng H, Sun Q, Wang LJ, Huang MT, Ge ZJ, Yu CQ, Guo Y, Wang TM, Wang J, Xu L, Wu WB, Chen L, Bian Z, Walters R, Millwood IY, Li XZ, Wang X, Hung RJ, Christiani DC, Chen HQ, Wang MY, Wang C, Jiang Y, Chen KX, Chen ZM, Jin GF, Wu TC, Lin DX, Hu ZB, Amos CI, Wu C, Wei QY, Jia WH, Li LM, Shen HB. Identification of risk loci and a polygenic risk score for lung cancer: a large-scale prospective cohort study in Chinese populations. *Lancet Respir Med*, 2019, 7(10): 881–891. [DOI]
- [25] Jin GF, Lv J, Yang M, Wang MY, Zhu M, Wang TP, Yan CW, Yu CQ, Ding YB, Li G, Ren CL, Ni J, Zhang RX, Guo Y, Bian Z, Zheng Y, Zhang NS, Jiang Y, Chen JP, Wang YN, Xu DZ, Zheng H, Yang L, Chen YP, Walters R, Millwood IY, Dai JC, Ma HX, Chen KX, Chen ZM, Hu ZB, Wei QY, Shen HB, Li LM. Genetic risk, incident gastric cancer, and healthy lifestyle: a meta-analysis of genome-wide association studies and prospective cohort study. *Lancet Oncol*, 2020, 21(10): 1378–1386. [DOI]
- [26] Zhu ZZ, Li JC, Si JH, Ma BS, Shi HWB, Lv J, Cao WH, Guo Y, Millwood IY, Walters RG, Lin K, Yang L, Chen YP, Du HD, Yu B, Hasegawa K, Camargo Jr CA, Moffatt MF, Cookson WOC, Chen JS, Chen ZM, Li LM, Yu CQ, Liang LM. A large-scale genome-wide association analysis of lung function in the Chinese population identifies novel loci and highlights shared genetic etiology with obesity. *Eur Respir J*, 2021. [DOI]
- [27] Pang YJ, Kartsonaki C, Lv J, Millwood IY, Yu CQ, Guo Y, Chen YP, Bian Z, Yang L, Chen JS, Clarke R, Walters R, Wu SK, Li HM, Holmes MV, Li LM, Chen ZM. Observational and genetic associations of body mass index and hepatobiliary diseases in a relatively lean Chinese population. *JAMA Netw Open*, 2020, 3(10): e2018721. [DOI]

(责任编辑: 方向东)