

## ·大型队列研究·

## 肺癌社区高危人群队列研究进展

李鑫<sup>1,2</sup> 冯小双<sup>1,3</sup> 张愉涵<sup>1</sup> 崔宏<sup>1</sup> 李霓<sup>1</sup> 代敏<sup>1</sup>

<sup>1</sup>国家癌症中心/国家肿瘤临床医学研究中心/中国医学科学院北京协和医学院肿瘤医院  
癌症早诊早治办公室, 北京 100021; <sup>2</sup>河南省人民医院病案科/郑州大学人民医院, 郑州  
450003; <sup>3</sup>复旦大学附属肿瘤医院肿瘤预防部/复旦大学上海医学院肿瘤学系, 上海 200032  
通信作者: 代敏, Email: daimin2002@hotmail.com

**【摘要】** 肺癌社区高危人群队列是2017年启动的国家重点研发计划精准医学专项“肺癌专病队列研究”的一部分,其主要目标是在全国7个大区的7个城市,收集5万例以上社区人群肺癌危险因素暴露信息,识别肺癌高危人群,开展低剂量螺旋CT筛查,并进一步追踪肺癌确诊和死亡信息;同时采集生物样本,建立以社区人群为基础、涵盖危险因素暴露、人群危险评估、低剂量螺旋CT筛查、生物样本以及结局随访信息的肺癌社区高危人群队列,为后续开展肺癌精准医学研究提供数据和样本支持。

**【关键词】** 肿瘤,肺; 高危人群; 队列研究

**基金项目:** 国家重点研发计划(2017YFC0907900, 2017YFC0907901)

**Progress in cohort study of lung cancer in high-risk population in communities**Li Xin<sup>1,2</sup>, Feng Xiaoshuang<sup>1,3</sup>, Zhang Yuhan<sup>1</sup>, Cui Hong<sup>1</sup>, Li Ni<sup>1</sup>, Dai Min<sup>1</sup>

<sup>1</sup>Office of Cancer Screening, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100021, China; <sup>2</sup>Department of Medical Records, Henan Provincial People's Hospital, Zhengzhou University People's Hospital, Zhengzhou 450003, China; <sup>3</sup>Department of Cancer Prevention, Fudan University Shanghai Cancer Center, Department of Oncology, Shanghai Medical College, Fudan University, Shanghai 200032, China

Corresponding author: Dai Min, Email: daimin2002@hotmail.com

**【Abstract】** The cohort study of lung cancer in high-risk population in communities in China was a part of Lung Cancer Cohort Study initiated in 2017 and funded by Precision Medicine Research of National Key Research and Development Program. Around 50 000 participants from the communities were enrolled from 7 cities in 7 regions in China. Information about the risk factors for lung cancer were collected and the populations at high risk for lung cancer were identified. Then, low-dose CT (LDCT) screening of lung cancer was conducted in the populations at high risk, and further information about the diagnosis of lung cancer cases and death cases were collected. Therefore, a community population-based cohort was established for lung cancer risk factor exposure survey, high risk population evaluation, LDCT screening and lung cancer case and death follow up. Meanwhile, biological samples were collected from all the participants in the cohort to support the future precision medicine research of lung cancer.

**【Key words】** Neoplasm, lung; High-risk population; Cohort study

**Fund programs:** National Key Research and Development Program of China (2017YFC0907900, 2017YFC0907901)

## 一、研究背景

肺癌是严重威胁人类健康的恶性肿瘤之一,位

居世界及中国恶性肿瘤发病及死亡的首位<sup>[1]</sup>。据国家癌症中心肿瘤登记办公室最新数据显示,

DOI: 10.3760/cma.j.cn112338-20210118-00041

收稿日期 2021-01-18 本文编辑 万玉立

引用本文:李鑫,冯小双,张愉涵,等.肺癌社区高危人群队列研究进展[J].中华流行病学杂志,2021,42(7):1174-1178. DOI: 10.3760/cma.j.cn112338-20210118-00041.



2015 年中国肺癌发病率和死亡率分别为 57.26/10 万和 45.87/10 万,且呈逐年上升趋势<sup>[24]</sup>,疾病负担和社会负担日益严重。

大规模人群队列研究自 20 世纪 70 年代以后得到了快速发展,涉及肿瘤的大型队列研究主要有:欧洲的 EPIC 研究(The European Prospective Investigation into Cancer and Nutrition)<sup>[5]</sup>、美国的 NHS 研究(The Nurses' Health Study)<sup>[6]</sup>、日本的 JACC 研究(The Japan Collaborative Cohort Study)<sup>[7]</sup>等。我国的大型队列研究起步比较晚,主要有中国慢性病前瞻性队列研究<sup>[8]</sup>、泰州健康人群跟踪调查队列<sup>[9]</sup>、上海女性健康研究<sup>[10]</sup>和男性健康研究<sup>[11]</sup>等。随着这些研究的开展,大型队列研究的条件日益成熟。

国内的肺癌队列研究主要集中在高发现场和高危人群,如云南省个旧市和宣威市的肺癌高发现场研究等。2012 年启动的国家重大公共卫生服务项目“城市癌症早诊早治项目”<sup>[12]</sup>将肺癌的高危人群评估、低剂量螺旋 CT(LDCT)筛查以及随访作为重点,在全国 20 多个省份的数十个城市开展了我国目前覆盖人群最多的肺癌筛查项目,为后续开展前瞻性队列研究奠定了基础。

基于此,2017 年由 中国医学科学院肿瘤医院牵头、全国 20 家单位承担的国家重点研发计划精准医学专项“肺癌专病队列研究”正式启动。项目旨在系统整合不同特色社区人群和临床诊疗的 5 个队列资源,建立肺癌早期发现和临床诊治全过程的生物样本库及与之相匹配的临床诊疗和随访全链条数据库和共享平台。

肺癌社区高危人群队列作为肺癌专病队列研究的一部分,目前已在全国目标社区进行了研究对象的纳入,并开展了危险因素调查、危险因素评估、LDCT 筛查和随访,为精准医学研究奠定了很好的社区人群队列研究基础。

## 二、研究设计和目标

按照统一标准和规范,在全国选择有代表性的地区开展社区人群肺癌危险因素调查和高危人群评估,并开展 LDCT 筛查和随访,系统收集人群的流行病学资料、LDCT 筛查信息以及随访信息;同时,按照规范收集所有研究对象的生物学标本,建立由高危人群、阳性结节、癌症组成的肺癌筛查人群队列的数据库和生物样本库,为后续开展以人群筛查为基础的肺癌精准医学研究提供数据和样本支持。

## 三、研究内容和方法

1. 研究现场:在全国 7 个大区选择工作基础好、有代表性的 7 个省份(每个大区 1 个省份)开展该项目,分别为河北省(华北地区)、安徽省(华东地区)、广西壮族自治区(华南地区)、湖南省(华中地区)、黑龙江省(东北地区)、新疆维吾尔自治区(西北地区)、云南省(西南地区)。

### 2. 研究对象的选择:

(1)纳入和排除标准:纳入标准:年龄 40~74 岁,本地常住人口(在本地居住 3 年以上),无严重器官功能障碍或精神疾患,自愿参加并且能够完成问卷调查和其他临床检查者。排除标准:已经确诊为肺癌,或者有其他严重的疾病正在治疗者。

(2)选择方法和样本量:以社区为单位进行整群抽样,社区内所有符合纳入条件的人群都将被邀请参加本研究。共计纳入研究对象 5 万例以上,平均每个省份纳入 7 143 例。肺癌社区高危人群队列流程见图 1。

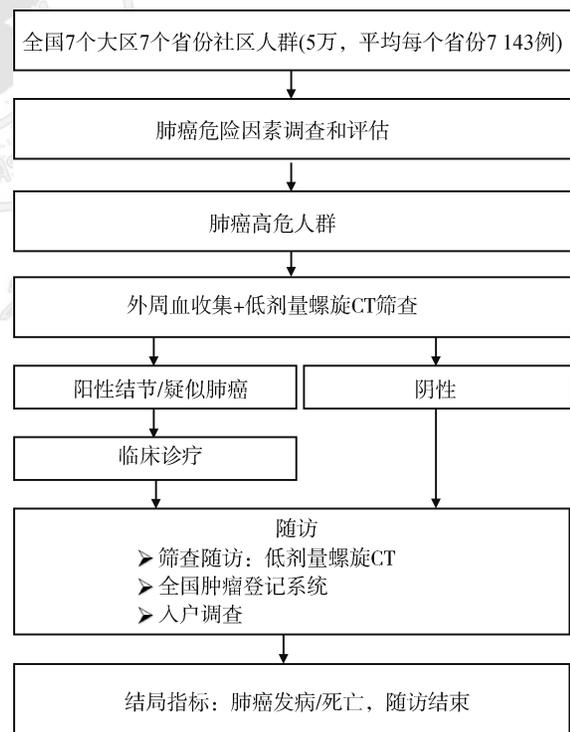


图1 肺癌社区高危人群队列流程

### 3. 基线调查:

(1)问卷调查:包括研究对象基本信息、饮食习惯、生活环境、生活方式和习惯、心理和情绪、疾病既往史、癌症家族史等。所有研究对象在研究前均签署知情同意书,然后在 CDC/社区专业技术人员指导下填写危险因素调查问卷或者由经过培训的

调查员询问调查对象后填写问卷。质控后由工作人员录入数据库,应用已开发的高危人群评估模型及其后台软件(由中国医学科学院肿瘤医院提供),计算得出其肺癌风险度,高危人群将被邀请参加 LDCT 筛查。

(2)LDCT 筛查:肺癌高危人群在约定的时间内被邀请前往指定医院接受 LDCT 筛查。筛查流程:预约、抽血、LDCT 筛查、筛查结果的反馈及进一步诊疗的指导。各医院按照项目组统一的筛查技术说明进行操作。

CT 扫描技术规范:选用多排(64 排)螺旋 CT 扫描仪(如参加单位条件不能达到者最少符合 CT 扫描仪 $\geq 16$  排)。 $\textcircled{1}$ 扫描参数:管电压 120 kVp,管电流 $\leq 30$  mAs;扫描层厚 5 mm,层间距 5 mm;重建层厚 1.00~1.25 mm 连续(层间隔为 0); $\textcircled{2}$ 扫描范围:从肺尖到肋膈角(包括全部肺),受检者吸气末一次屏气完成扫描(检查时应有专人训练受检者屏气)。

图像观察:胸部专业放射科医师在 CT 工作站或图像储存与传输系统(picture archiving and communication systems, PACS)专用监视器观察图像,采用标准肺窗(窗宽/窗位:1 600~2 000 Hu/-600~-700 Hu)、纵隔窗(软组织窗:350~380 Hu/10~15 Hu)及骨窗(2 000/400 Hu)观察。

结节测量:用电子测量尺(工作站或 PACS 系统内自带)通过结节最大截面测量长径及宽径(长径指结节最大截面的最大径;宽径指与长径垂直的最大径)。

阳性肺结节的定义: $\textcircled{1}$ 影像学特征倾向或者不完全除外恶性的病灶; $\textcircled{2}$ 直径 $\geq 15$  mm 的所有类型(包括实性、部分实性及非实性)结节; $\textcircled{3}$ 直径 5~14 mm 的实性/部分实性结节; $\textcircled{4}$ 直径 8~14 mm 的非实性结节。

对基线 CT 筛查的非钙化性结节进行进一步处理的推荐方案:对于 $\geq 15$  mm 结节(包括实性、部分实性及非实性)可选择以下对策: $\textcircled{1}$ 由副高级或以上职称放射科医师判断是否进入临床干预; $\textcircled{2}$ 抗炎治疗 5~7 d 后与基线 CT 间隔 1 个月复查;若结节部分吸收,则与基线 CT 间隔 3 个月时进行 LDCT 复查,如果结节增大或无变化,由副高级或以上职称放射科医师判断是否进入临床干预。

对于 5~14 mm 的实性/部分实性结节及 8~14 mm 非实性结节,应建议与基线 CT 时间间隔 3 个月时进行 LDCT 复查,如果结节增大,由副高级或以上职称放射科医师判断是否进入临床干预。

(3)生物样本的收集以及生物样本库的建立:LDCT 筛查前,每位研究对象分别抽取 2 管全血(各 5 ml),1 管为抗凝血、1 管为非抗凝血。按照标准的血样分离方法,将抗凝血分离成血浆、白细胞和红细胞,非抗凝血分离成血清和血细胞。血清、血浆、白细胞分离后,分装至冻存管,并粘贴对应冻存标签。每位参加者共收集 11 份样本,贴好对应标签后分别存至不同冻存盒的同一位置(血清 5 盒、血浆 5 盒、白细胞 1 盒)。

样本保存条件为 $-80$  °C 冰箱,各中心收集后暂时自行保存,最终转运至中国医学科学院肿瘤医院生物样本管理中心进行集中保存。

4. 随访:根据首次 LDCT 筛查的结果,对研究对象开展有针对性的随诊和随访,详细记录其确诊为肺癌或其他疾病的情况以及进一步筛查的结果。从入组后第二年起,每年都对纳入队列中的研究对象进行随访,监测其疾病发生情况。

随访采用主动随访与被动随访相结合的形式进行。其中,主动随访包括入户随访、电话随访和筛查随访等,被动随访包括链接国家癌症中心全国肿瘤登记信息、CDC 死因监测信息和职工医保数据(河北省唐山市)等。

随访结局指标:经临床或病理诊断确诊为肺癌;死亡;随访结束。

5. 信息平台的建立:建立互联互通生物样本和大数据共享平台,平台系统具有问卷信息采集、肺癌风险评估、LDCT 筛查、随访管理以及生物样本管理等功能。

#### 6. 质量控制:

(1)制定队列实施标准和规范: $\textcircled{1}$ 纳入人群基线信息收集标准规范:项目立项之初,即以原有队列人群基线信息收集标准化用表为基础,在肺癌社区高危人群队列代表性现场(河北省和云南省)深入调研信息收集标准的可行性,进一步完善队列构建所需的关键基线信息,制定了标准化调查问卷,已应用于所有队列人群的招募; $\textcircled{2}$ 危险因素调查标准规范:使用电子化问卷进行面对面调查,执行调查的工作人员在开始调查前均接受统一规范化培训; $\textcircled{3}$ 肺癌高危人群界定标准:采用经全国 20 个省份前期验证的可有效筛选肺癌高危人群的城市社区人群肺癌风险评估标准,通过综合多项流行病学信息(包括吸烟与被动吸烟情况、饮食情况、生活环境、方式和习惯、慢性呼吸系统疾病史、肺癌家族史等因素),进行模型运算,确定肺癌高危人群危险因

素和相关赋值,筛选出肺癌高危人群;④随访标准规范:已制定标准化肺癌患者随访量表和随访指南,并已使用该量表完成部分患者的首次随访;⑤生物样本收集及生物样本库的标准规范:已制定规范化外周血收集、分装、储存操作流程,并建立相应的生物样本库和信息库系统;⑥数据编码、录入及上传的标准规范:已完成对基线信息、危险因素和随访量表的统一变量编码模式,开发电子问卷系统,并已启用,可进行规范的数据录入和上传。

(2)队列质量控制:本项目实施之前已制定肺癌社区高危人群队列现场操作方案,各单位按照统一的标准和流程纳入研究对象、收集相关信息和生物样本,按照统一的样本编码规则进行编码和保存样本;所有资料由各参与单位组织人员双录入,经质控员检查合格后将纸质版和电子版按期逐级进行上报。同时指定专业人员录入生物样本信息,生物样本入库、出库均详细记录,并及时录入到生物样本信息化平台;承担 LDCT 筛查任务的医疗单位应为具有癌症专业化诊断和治疗能力的三级以上肿瘤专科医院或具有肿瘤专科的综合性医院,同时筛查医院选派具有一定资历的诊断科和临床科室的人员承担筛查任务。

7. 统计学分析:以经审核无误的数据为基础,计算 2017 年 9 月 1 日至 2020 年 11 月 30 日期间肺癌高危社区项目高危人数和参与率。高危率(%)=肺癌评估为高危例数/参与评估总例数 $\times$ 100,筛查率(%)=评估为高危且参与临床筛查例数/评估为高危例数 $\times$ 100。

#### 四、研究进展

##### 1. 队列建设情况:

(1)风险评估和高危人群人数:截至 2020 年 11 月 30 日,已完成 39.08 万例社区人群的危险因素调查。其中 8.55 万例被评估为肺癌高危人群,高危率为 21.88%。在 8.55 万例高危人群中,男性 4.88 万例(57.08%),女性 3.67 万例(42.92%);在年龄分布上,50~54 岁组高危人数最多,为 15 903 例,占全部肺癌高危人群的 18.60%,70~74 岁组高危人数最少,为 6 071 例,占 7.10%。见图 2。

(2)筛查人数及筛查参与率:在 8.55 万例高危人群中,4.17 万例高危人群完成了 LDCT 筛查,筛查参与率为 48.77%;不同年龄段筛查参与率不同,其中 55~59 岁、60~64 岁组筛查参与率较高,而 70~74 岁组筛查参与率较低。见图 3。剩余高危人群筛查在继续进行中。

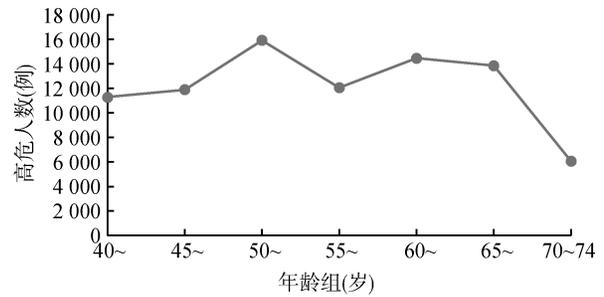


图2 不同年龄组高危人群数量分布

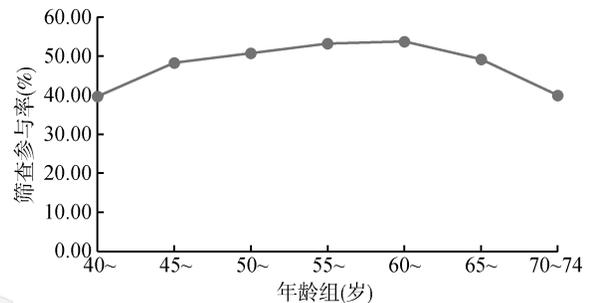


图3 不同年龄组筛查参与率分布

(3)生物样本收集及队列随访:收集生物样本 3.55 万例,匹配率为 85.13%;目前已完成随访 43 709 例,失访率为 3.36%。

2. 信息化平台建设情况:平台由项目牵头单位中国医学科学院肿瘤医院项目团队统筹与管理,可实现队列间一键切换,实施质量控制与监察。建立以受试者为中心的全流程自动化管理,同时通过数据质控确保数据完整性和准确性,实现数据的实时高质量收集,利用大数据技术,对受试者数据信息进行全方位挖掘和分析,实现基于肺癌的多维度检索等。截至目前,已完成 8 万余例高危人群数据收集,将为后续精准医学研究提供数据支撑。

#### 五、讨论

1. 队列招募对象规模较大,数据海量且繁杂:队列采用统一的标准和规范采集每个环节的数据和信息,以保证数据的整齐和准确;同时,采用规范的信息系统平台,将海量数据进行规范整理,提高数据质量;而且,增加现场数据的管理及审核环节,保障数据质量和数据安全,且及时录入数据平台。

2. 队列生物样本收集及完整保存难度较大:队列采用相同的生物样本收集和分装原则,保障样本质量和安全;同时中国医学科学院肿瘤医院专门建立生物样本库用于项目中生物样本的长期保存。

3. 队列随访周期长、结局要求严格,且需保证随访率:本研究采用主动随访和被动随访相结合的方式,同时收集研究对象的发病和死亡信息,保证

随访率和随访质量,尽可能降低失访率。

4. 针对不同年龄段人群筛查参与率的不同,项目组综合考虑筛查居民及筛查工作人员两个方面因素:首先,项目组积极组织宣传和动员活动,提升居民肺癌风险认知能力;同时针对不同年龄组,制定不同的筛查方案,提高筛查工作人员工作效率,最终提升各个年龄组筛查参与率。

#### 六、本研究的特色和前景

肺癌社区高危人群队列在精准医学队列建设工作中有如下创新性:①信息和样本收集与管理规范、精准:队列数据库和样本库的建立,是在标准化筛查方案和规范化诊疗方案基础上的数据收集和样本采集,保证了基线、筛查等信息和样本的统一及规范;在主动随访的基础上,充分利用全国肿瘤登记网络和肿瘤大数据网络平台,使用身份证识别信息进行二次随访,保证高随访率以及发病、死亡等结局信息的精准;②强调全国人群代表性、多学科联合攻关:队列均来自全国多中心,覆盖全国7个大区的7个城市,人群代表性好,结果产出和工作机制均可为全国以及不同地域的人群开展肺癌相关研究提供借鉴。工作团队覆盖现场人员、临床诊疗人员、实验室人员等,多学科合作再次体现了精准医学的多层次要求。

肺癌社区高危人群队列有以下预期经济、社会效益:①队列针对LDCT筛查发现的阳性结节开展连续的随访能够有效发现早期肺癌患者,显著提高患者生存情况并减少医疗开支;②同时本研究将利用数据抽取技术和网络信息技术,探索共建、共享管理机制,实现样本库和数据库互联互通,有效推进信息化平台建设,实现多源医疗数据整合和信息共享平台的建立,使数据开发使用最大化,更好服务于全国乃至全球肺癌防治的科研和决策。

利益冲突 所有作者均声明不存在利益冲突

#### 参 考 文 献

- [1] Chen WQ, Zheng RS, Baade PD, et al. Cancer statistics in China, 2015[J]. CA Cancer J Clin, 2016, 66(2): 115-132. DOI:10.3322/caac.21338.
- [2] 孙可欣,郑荣寿,张思维,等. 2015年中国分地区恶性肿瘤发病和死亡分析[J]. 中国肿瘤, 2019, 28(1):1-11. DOI: 10.11735/j.issn.1004-0242.2019.01.A001.  
Sun KX, Zheng RS, Zhang SW, et al. Report of cancer incidence and mortality in different areas of China, 2015[J]. China Cancer, 2019, 28(1):1-11. DOI:10.11735/j.issn.1004-0242.2019.01.A001.
- [3] 陈万青,孙可欣,郑荣寿,等. 2014年中国分地区恶性肿瘤发病和死亡分析[J]. 中国肿瘤, 2018, 27(1):1-14. DOI: 10.11735/j.issn.1004-0242.2018.01.A001.  
Chen WQ, Sun KX, Zheng RS, et al. Report of cancer incidence and mortality in different areas of China, 2014[J]. China Cancer, 2018, 27(1):1-14. DOI:10.11735/j.issn.1004-0242.2018.01.A001.
- [4] 陈万青,郑荣寿,张思维,等. 2013年中国恶性肿瘤发病和死亡分析[J]. 中国肿瘤, 2017, 26(1):1-7. DOI:10.11735/j.issn.1004-0242.2017.01.A001.  
Chen WQ, Zheng RS, Zhang SW, et al. Report of cancer incidence and mortality in China, 2013[J]. China Cancer, 2017, 26(1):1-7. DOI:10.11735/j.issn.1004-0242.2017.01.A001.
- [5] Riboli E, Hunt KJ, Slimani N, et al. European Prospective Investigation into Cancer and Nutrition (EPIC): study populations and data collection[J]. Public Health Nutr, 2002, 5(6B):1113-1124. DOI:10.1079/PHN2002394.
- [6] Colditz GA, Manson JE, Hankinson SE. The Nurses' Health Study: 20-year contribution to the understanding of health among women[J]. J Women's Health, 1997, 6(1): 49-62. DOI:10.1089/jwh.1997.6.49.
- [7] Tamakoshi A, Yoshimura T, Inaba Y, et al. Profile of the JACC study[J]. J Epidemiol, 2005, 15 Suppl 1: S4-8. DOI: 10.2188/jea.15.s4.
- [8] Chen ZM, Chen JS, Collins R, et al. China Kadoorie Biobank of 0.5 million people: survey methods, baseline characteristics and long-term follow-up[J]. Int J Epidemiol, 2011, 40(6):1652-1666. DOI:10.1093/ije/dyr120.
- [9] Wang XF, Lu M, Qian J, et al. Rationales, design and recruitment of the Taizhou Longitudinal Study[J]. BMC Public Health, 2009, 9:223. DOI:10.1186/1471-2458-9-223.
- [10] Zheng W, Chow WH, Yang G, et al. The Shanghai Women's Health Study: rationale, study design, and baseline characteristics[J]. Am J Epidemiol, 2005, 162(11): 1123-1131. DOI:10.1093/aje/kwi322.
- [11] Shu XO, Li HL, Yang G, et al. Cohort profile: the Shanghai Men's Health Study[J]. Int J Epidemiol, 2015, 44(3): 810-818. DOI:10.1093/ije/dyv013.
- [12] 代敏,石菊芳,李霓. 中国城市癌症早诊早治项目设计及预期目标[J]. 中华预防医学杂志, 2013, 47(2):179-182. DOI: 10.3760/cma.j.issn.0253-9624.2013.02.018.  
Dai M, Shi JF, Li N. Project design and expected goals of early diagnosis and treatment of cancer in urban Chinese [J]. Chin J Prev Med, 2013, 47(2): 179-182. DOI:10.3760/cma.j.issn.0253-9624.2013.02.018.