

建成环境暴露与急性心肌梗死 复发的关联识别及规划应对

(申请清华大学工学硕士学位论文)

培 养 单 位 ： 建筑学院
学 科 ： 城乡规划学
研 究 生 ： 刘 宁 睿
指 导 教 师 ： 龙 瀛 副 教 授

二〇二三年五月

建成环境暴露与急性心肌梗死复发的关联识别及规划应对

刘宁睿

Associations Between Urban Exposome and Recurrence of Acute Myocardial Infarction and Planning Responses

Thesis submitted to

Tsinghua University

in partial fulfillment of the requirement

for the degree of

Master of Science

in

Urban and Rural Planning

by

Liu Ningrui

Thesis Supervisor: Long Ying, Associate Professor

May, 2023

学位论文公开评阅人和答辩委员会名单

公开评阅人名单

党安荣	教授	清华大学
刘佳燕	副教授	清华大学

答辩委员会名单

主席	武廷海	教授	清华大学
委员	王崇烈	教授级高级工程师	北京市城市规划设计研究院
	王英	副教授	清华大学
	刘佳燕	副教授	清华大学
	龙瀛	副教授	清华大学
	梁思思	副教授	清华大学
	孙诗萌	助理研究员	清华大学
	郑筱津	教授级高级工程师	北京清华同衡规划设计研究院
秘书	耿丹		清华大学

关于学位论文使用授权的说明

本人完全了解清华大学有关保留、使用学位论文的规定，即：

清华大学拥有在著作权法规定范围内学位论文的使用权，其中包括：（1）已获学位的研究生必须按学校规定提交学位论文，学校可以采用影印、缩印或其他复制手段保存研究生上交的学位论文；（2）为教学和科研目的，学校可以将公开的学位论文作为资料在图书馆、资料室等场所供校内师生阅读，或在校园网上供校内师生浏览部分内容；（3）按照上级教育主管部门督导、抽查等要求，报送相应的学位论文。

本人保证遵守上述规定。

作者签名：

导师签名：

日 期：

日 期：

摘 要

健康城市近年来受到了国家和社会的广泛关注，一系列健康城市、健康社区相关的评价标准相继出台。现有数据表明，缺血性心脏病已经成为影响中国人群健康水平的罪魁祸首之一，其中，急性心肌梗死的预后复发问题又显得尤为突出。因此，本文提出的关键科学问题为：建成环境暴露与急性心肌梗死复发之间的定量关联规律是什么？如何将以上定量规律应用到规划实践当中，从而从建成环境的角度降低居民的急性心肌梗死复发风险？本研究围绕上述科学问题，以北京市为例，应用新兴的暴露组学（*exposomics*）分析方法开展了相关研究，主要内容和成果包括：

首先，提出了根据患者住址文本信息进行建成环境暴露系统评价的方法。本研究采用地址解析和人工审查结合的方法，实现了患者住址的空间落位，并根据暴露组全面性、动态性的特征，构建了六大类共 26 个小类的建成环境暴露指标体系，采用三种缓冲区的空间尺度与基线暴露、累积平均暴露两种时间尺度，对患者的建成环境暴露进行了全面测度和评价。

其次，揭示了建成环境暴露与急性心肌梗死复发风险之间的定量关联规律。本研究通过回顾性队列研究设计（*retrospective cohort study*），以 2013 年至 2019 年北京市心血管疾病监测系统内的急性心肌梗死患者为研究对象，以患者的首次急性心肌梗死复发为首要结局，采用生存分析中 Cox 模型和变量选择中 ENET 算法相结合的统计分析方法，探究建成环境暴露水平与急性心肌梗死复发之间的暴露反应关系。研究发现，在基线暴露水平下，500 米缓冲区内的城市功能混合度为显著保护因素，到地铁站的距离及 PM_{2.5} 是显著危险因素；在累积平均暴露水平下，除上述显著因素外，1000 米缓冲区内的蔬果店密度为显著保护因素，到公园的距离是显著危险因素。

最后，建立了建成环境的归因急性心肌梗死复发风险评估及规划应对的方法。本研究采用人口归因分数（*population attributable fraction, PAF*）的概念，基于北京市 2021 年上述五个显著指标的暴露水平，进行了建成环境的归因复发风险评估。本研究进一步对排名后 40% 的待提升小区进行分类，发现占比较高的待提升小区类别主要面临的是蔬果店缺乏、PM_{2.5} 浓度较高、到公园距离较远等问题。本研究分别给出了利用社区服务中心及腾退空间增设蔬果店、设置颗粒物滞纳型防护绿地、利用闲置小微空地及滨水空间营造多层级绿地结构等规划应对策略。

关键词：建成环境；急性心肌梗死；复发；暴露组；风险评估

Abstract

In recent years, healthy cities have been paid much attention to by our country and society. A series of evaluation standards about healthy cities and healthy communities have been put forward. The newest data show that ischemic heart disease has become one of the most serious health outcomes for Chinese people, and the prognosis and recurrence of acute myocardial infarction (AMI) is particularly prominent. Therefore, this study raised two key scientific questions: What are the quantitative associations between exposure to built environment and recurrence of AMI? How can the above quantitative associations be applied to planning practice from the perspective of built environment, in order to reduce the recurrence risk of AMI for residents? Centered on the above issues, this study took Beijing as an example and conducted relevant research using the emerging exposome approach. The main contents and achievements of this study include:

First, a method to systematically evaluate the urban exposures based on the text information of patients' home addresses was proposed. This study combined the method of automatic address geocoding and manual audit to accurately acquire the spatial coordinates of the patients' home addresses. Then, according to the comprehensive and dynamic characteristics of the exposome, this study constructed the index system of urban exposures, including six categories and 26 types of exposures. The urban exposures of patients were measured and evaluated using three different buffers zones and two different time scales (i.e., the baseline exposure and cumulative average exposure).

Second, the quantitative associations between urban exposure and the recurrence risk of AMI were revealed. Through a retrospective cohort study design, this study selected the AMI patients in the Beijing Cardiovascular Disease Monitoring System from 2013 to 2019, and took the first recurrence of AMI as the primary outcome. The Cox model in survival analysis was applied, combining the ENET algorithm for variable selection, to explore the exposure-response relationships between built environment and AMI recurrence. This study found that under the baseline exposure level, the urban function diversity within 500-m buffer was a significant protective

factor, and the distance to subway station and PM_{2.5} were significant risk factors. Under the cumulative average exposure level, besides the above significant factors, the density of vegetable and fruit shops within 1000-m buffer was a significant protective factor, while the distance to park was a significant risk factor.

Finally, the method of risk assessment for the recurrence risk of AMI attributable to urban exposures and corresponding planning responses was established. This study adopted the concept of population attributable fraction (PAF), and estimated the attributable recurrence risk based on the exposure level of the above five urban exposures in Beijing in 2021. This study further classified the communities ranking the bottom 40% into ten types, and found that the majority of these communities are faced with the lack of vegetable and fruit shops, high PM_{2.5} concentration, and long distance to the park. This study proposed the corresponding planning responses to these problems, respectively, including using the community service center or the vacant space to add vegetable and fruit shops, increasing the protection greenbelt with trees, and taking full advantage of microscale vacant space and waterfront space to promote the multi-level green space system around the community.

Keywords: built environment; acute myocardial infarction; recurrence; exposome; risk assessment

目 录

摘 要	I
Abstract	II
目 录	IV
插图清单	VII
附表清单	X
符号和缩略语说明	XI
第 1 章 绪论	1
1.1 研究背景	1
1.1.1 中国城镇化面临的健康挑战与已有对策	1
1.1.2 中国居民心血管健康水平及预后问题应受重视	2
1.1.3 新方法与新数据环境为建成环境大规模精细化测度提供支持	2
1.2 研究目标与内容	4
1.2.1 研究目标与科学问题	4
1.2.2 研究内容	5
1.3 研究方法与技术路线	5
1.3.1 研究方法	5
1.3.2 技术路线	6
第 2 章 理论与研究与文献综述	8
2.1 建成环境的概念界定及主要暴露因素	8
2.2 建成环境暴露与心血管疾病之间的关联研究	9
2.2.1 绿地与心血管疾病关联的流行病学研究	9
2.2.2 其他建成环境因素与心血管疾病关联的流行病学研究	13
2.2.3 空气污染与心血管疾病关联的流行病学研究	17
2.3 暴露组学研究方法及其在建成环境领域的应用	18
2.3.1 暴露组学：概念与方法	18
2.3.2 暴露组学在建成环境领域的应用	19
2.4 健康视角下的规划应对研究	23
2.4.1 城市规划与公共健康的相互作用关系	23
2.4.2 规划应对策略与技术	25
2.5 文献评述与研究方向	26

第 3 章 研究对象与数据基础	28
3.1 研究范围与研究对象	28
3.1.1 研究范围	28
3.1.2 研究设计与研究对象	29
3.1.3 研究人群的随访结果与基线特征	29
3.2 建成环境数据类型与来源	30
3.2.1 基础城市空间数据	31
3.2.2 城市兴趣点 (POI) 数据	32
3.2.3 城市用地及建筑物数据	32
3.2.4 街景图像数据	34
3.2.5 高分遥感影像数据	35
3.2.6 其他建成环境数据	35
第 4 章 建成环境暴露评价	36
4.1 患者住址的空间化	36
4.1.1 患者住址空间化方法	36
4.1.2 患者住址空间化结果	39
4.2 建成环境暴露因素的选择	41
4.3 患者住址周边建成环境测度方法与结果	42
4.3.1 第 1 类建成环境暴露因素: 密度	42
4.3.2 第 2 类建成环境暴露因素: 多样性	47
4.3.3 第 3 类建成环境暴露因素: 设计	51
4.3.4 第 4 类建成环境暴露因素: 到公共交通的距离	59
4.3.5 第 5 类建成环境暴露因素: 目的地可达性	62
4.3.6 空气污染	69
4.4 患者建成环境暴露评价方法与结果	71
4.4.1 建成环境暴露评价方法	71
4.4.2 建成环境暴露评价结果	71
4.5 本章小结	77
第 5 章 建成环境暴露与 AMI 复发的关联分析	78
5.1 暴露组变量选择算法的比较	78
5.1.1 变量选择算法的原理	78
5.1.2 变量选择算法的性能比较方法	81
5.1.3 变量选择算法的性能比较结果	83

5.2 统计分析方法	87
5.2.1 暴露变量、协变量、结局变量的选择	87
5.2.2 生存分析的基本原理	88
5.2.3 统计分析方法与流程	89
5.3 单因素关联分析结果	90
5.4 多因素关联分析结果	94
5.4.1 所有 AMI 复发事件的关联分析结果	94
5.4.2 致死型、非致死型 AMI 复发事件的关联分析结果	95
5.4.3 北京市五环内子队列的关联分析结果	99
5.5 结果讨论	99
5.6 本章小结	101
第 6 章 北京市建成环境的归因 AMI 复发风险现状评估及规划应对	103
6.1 北京市建成环境归因 AMI 复发风险评估	103
6.1.1 建成环境归因 AMI 复发风险评估方法	103
6.1.2 乡镇街道尺度复发风险评估结果	106
6.1.3 居住小区尺度复发风险评估结果	111
6.2 北京市居住小区 AMI 预后的规划应对策略	115
6.2.1 相关标准与导则	115
6.2.2 待提升小区的识别与分类	118
6.2.3 各类待提升小区的规划应对策略	119
6.3 本章小结	129
第 7 章 研究结论与展望	131
7.1 研究结论	131
7.2 研究展望	132
参考文献	134
致 谢	144
声 明	145
个人简历、在学期间完成的相关学术成果	146
指导教师评语	149
答辩委员会决议书	150

插图清单

图 1.1 研究技术路线	7
图 2.1 建成环境暴露的健康传导机制	24
图 2.2 建成环境暴露对心血管疾病的影响机制	24
图 3.1 北京市全域及五环内区域	28
图 3.2 北京市基础城市空间数据	31
图 3.3 北京市 POI 数据	32
图 3.4 北京市用地数据	33
图 3.5 北京市五环内 2014 年建筑物数据	34
图 3.6 北京市五环内街景图像数据	35
图 4.1 患者住址空间化的基本流程	36
图 4.2 人工核查与修正工作指南部分页面	40
图 4.3 2013 年至 2019 年北京市急性心肌梗死患者住址空间分布	40
图 4.4 北京市 AMI 患者住址 1000 m 缓冲区 2019 年人口密度测度结果	43
图 4.5 北京市五环内 AMI 患者住址 1000 m 缓冲区 2018 年(a)建筑密度与(b)建筑容积率测度结果	45
图 4.6 北京市 AMI 患者住址 1000 m 缓冲区 2019 年城市功能密度测度结果	46
图 4.7 北京市 AMI 患者住址 1000 m 缓冲区 2014 年土地利用混合度测度结果	49
图 4.8 北京市 AMI 患者住址 1000 m 缓冲区 2019 年城市功能混合度测度结果	50
图 4.9 北京市 AMI 患者住址 1000 m 缓冲区 2019 年路网密度测度结果	52
图 4.10 北京市 AMI 患者住址 1000 m 缓冲区 2016 年到主干道距离测度结果	53
图 4.11 空间失序要素审计清单	54
图 4.12 北京市五环内 AMI 患者住址 1000 m 缓冲区 2019 年空间失序测度结果	55
图 4.13 北京市五环内 AMI 患者住址 1000 m 缓冲区 2014 年步行指数测度结果	56
图 4.14 北京市 AMI 患者住址 1000 m 缓冲区 2014 年绿地比例测度结果	57
图 4.15 北京市五环内 AMI 患者住址 1000 m 缓冲区 2017 年街道绿视率测度结果	58
图 4.16 北京市 2019 年 NDVI 测度结果 (a) NDVI 空间分布, (b) 五环内 AMI 患者住址 1000 m 缓冲区内测度结果	60
图 4.17 北京市五环内 AMI 患者住址 1000 m 缓冲区 2019 年地铁站测度结果	61
图 4.18 北京市五环内 AMI 患者住址 1000 m 缓冲区 2019 年公交站密度测度结果	

.....	62
图 4.19 北京市五环内 AMI 患者住址 1000 m 缓冲区 2019 年公园密度测度结果	63
图 4.20 北京市五环内 AMI 患者住址 1000 m 缓冲区 2019 年到公园的距离测度结果	64
图 4.21 北京市五环内 AMI 患者住址 1000 m 缓冲区 2019 年到大型绿地的距离的测度结果	64
图 4.22 北京市五环内 AMI 患者住址 1000 m 缓冲区 2019 年烟酒店密度测度结果	65
图 4.23 北京市五环内 AMI 患者住址 1000 m 缓冲区 2019 年餐厅密度测度结果	65
图 4.24 北京市五环内 AMI 患者住址 1000 m 缓冲区 2019 年快餐店密度测度结果	66
图 4.25 北京市五环内 AMI 患者住址 1000 m 缓冲区 2019 年甜品/饮品/糕饼店密度测度结果	66
图 4.26 北京市五环内 AMI 患者住址 1000 m 缓冲区 2019 年蔬果店密度测度结果	67
图 4.27 北京市五环内 AMI 患者住址 1000 m 缓冲区 2019 年运动场馆密度测度结果	67
图 4.28 北京市五环内 AMI 患者住址 1000 m 缓冲区 2019 年综合医院测度结果 (a) 综合医院密度 (b) 到综合医院的距离	68
图 4.29 北京市五环内 AMI 患者住址 1000 m 缓冲区 2019 年药房密度测度结果	69
图 4.30 北京市 AMI 患者住址 1000 m 缓冲区内 2018 年 PM _{2.5} 浓度测度结果	70
图 4.31 建成环境暴露因素之间的 Spearman 相关系数矩阵	75
图 5.1 变量选择算法在主分析中 100 次计算的统计学性能指标均值	84
图 5.2 变量选择算法在情形 2 和 3 中 100 次计算的统计学性能指标均值	85
图 5.3 变量选择算法在情形 4 和 5 中 100 次计算的统计学性能指标均值	85
图 5.4 变量选择算法在情形 6 中 100 次计算的统计学性能指标均值	86
图 5.5 变量选择算法在情形 7 中 100 次计算的统计学性能指标均值	87
图 5.6 单因素分析结果的火山图: (a) 基线暴露, (b) 累积平均暴露	91
图 6.1 北京市乡镇街道尺度建成环境的总归因复发风险结果	107
图 6.2 北京市乡镇街道尺度城市功能混合度对 AMI 复发的 PAF 结果	108
图 6.3 北京市乡镇街道尺度到地铁站距离对 AMI 复发的 PAF 结果	108
图 6.4 北京市乡镇街道尺度到公园距离对 AMI 复发的 PAF 结果	109
图 6.5 北京市乡镇街道尺度蔬果店密度对 AMI 复发的 PAF 结果	109

图 6.6 北京市乡镇街道尺度 $PM_{2.5}$ 对 AMI 复发的 PAF 结果	110
图 6.7 北京市六环内小区尺度建成环境的总归因复发风险结果	112
图 6.8 北京市六环内小区尺度城市功能混合度对 AMI 复发的 PAF 结果	113
图 6.9 北京市六环内小区尺度到地铁站距离对 AMI 复发的 PAF 结果	113
图 6.10 北京市六环内小区尺度到公园距离对 AMI 复发的 PAF 结果	114
图 6.11 北京市六环内小区尺度蔬果店密度对 AMI 复发的 PAF 结果	114
图 6.12 北京市六环内小区尺度 $PM_{2.5}$ 对 AMI 复发的 PAF 结果	115
图 6.13 北京市六环内待提升小区的空间分布及分类结果	119
图 6.14 第 1 类待提升小区的空间分布	120
图 6.15 第 1 类待提升小区的规划应对模式图	120
图 6.16 居住区外围边界绿地配置方案	121
图 6.17 第 1 类待提升小区案例：大兴区义和庄东里	121
图 6.18 第 3 类待提升小区的空间分布	122
图 6.19 第 3 类待提升小区的规划应对模式图	123
图 6.20 第 3 类待提升小区案例：房山区碧桂园小区 A 区	123
图 6.21 第 4 类待提升小区的空间分布	124
图 6.22 第 4 类待提升小区的规划应对模式图	124
图 6.23 第 4 类待提升小区案例：朝阳区东坝家园 B 区	125
图 6.24 第 5 类待提升小区的空间分布	126
图 6.25 第 5 类待提升小区的规划应对模式图	126
图 6.26 第 5 类待提升小区案例：海淀区门头馨村北二区	127
图 6.27 第 7 类待提升小区的空间分布	128
图 6.28 第 7 类待提升小区的规划应对模式图	128
图 6.29 第 7 类待提升小区案例：昌平区二拨子新村	129

附表清单

表 2.1 绿地与心血管疾病之间关联的流行病学研究	11
表 2.2 其他建成环境因素与心血管疾病之间关联的流行病学研究	14
表 2.3 城市暴露组相关的流行病学研究	21
表 3.1 研究人群的基线特征水平	30
表 3.2 城市用地类型分类依据	33
表 4.1 不同来源的地址信息编号	38
表 4.2 地址信息来源选择示例	38
表 4.3 地址规范化示例	39
表 4.4 本研究选择的建成环境暴露因素	41
表 4.5 POI 分类依据	48
表 4.6 各功能的 POI 分类依据	62
表 4.7 北京市急性心肌梗死患者的建成环境暴露评价结果	72
表 5.1 变量选择算法在主分析中 100 次计算的统计学性能指标	84
表 5.2 单因素分析的 HR 值及 95%置信区间	92
表 5.3 建成环境基线暴露水平与所有、致死型、非致死型 AMI 复发事件之间的 HR 值及 95%置信区间	96
表 5.4 建成环境累积平均暴露水平与所有、致死型、非致死型 AMI 复发事件之 间的 HR 值及 95%置信区间	97
表 6.1 建成环境与 AMI 复发之间具有显著性的 HR 及其 95%置信区间	104
表 6.2 建成环境因素的 PAF 临界值	118
表 6.3 待提升小区分类	119

符号和缩略语说明

主要符号

A_i	患者住址缓冲区内第 i 类用地的归一化面积, 无量纲
E	环境暴露变量的暴露水平
E_0	环境暴露变量的单位水平
$E_{\max}$	队列研究中该因素暴露水平的 99%分位数
$h(t)$	时间为 t 时的风险函数, 无量纲
$h_0(t)$	时间为 t 时的基准风险函数, 无量纲
HR_0	每增加一个单位水平暴露变量下的危险比, 无量纲
M	患者住址缓冲区内土地利用或城市功能混合度, 无量纲
N_i	患者住址缓冲区内第 i 类 POI 的归一化个数, 无量纲
N_{11}	被算法选择出的真实变量个数, 个
N_{1+}	设定的真实变量个数, 个
N_{+1}	被算法选择出的所有变量个数, 个
N_{21}	被算法选择出的错误变量个数, 个
N_{22}	没有被算法选择出的错误变量个数, 个
N_{2+}	设定的错误变量个数, 个
n_B	算法选择出的变量个数, 个
n_{total}	待选变量总个数, 个
NIR	近红外波段反射值, 无量纲
p	环境暴露因素在某暴露水平下的概率密度函数, 无量纲
p_i	患者住址缓冲区内第 i 类 POI 个数或用地面积的归一化占比, 无量纲
R	红光波段反射值, 无量纲
r_s	环境暴露变量间的 spearman 相关系数, 无量纲
$RMSE$	模型的均方误差
$S(t)$	时间为 t 时的生存函数, 无量纲
X	环境暴露变量观测值
Y	健康结局观测值
Z	控制变量观测值

希腊字母

β	回归模型中环境暴露因素的回归系数
γ	回归模型中控制变量的回归系数
σ	回归模型正态随机误差的标准差
ε	回归模型的随机误差项
λ	ENET 算法中的惩罚系数
Σ	多元正态的暴露数据集的协方差矩阵

缩略语

AMI	acute myocardial infarction, 急性心肌梗死
BH	Benjamini and Hochberg, 一种多重比较校正方法
DSA	deletion-substitution-addition algorithm, 删除-替代-添加算法
ENET	elastic net, 弹性网模型
EWAS	environment-wide association studies, 全环境关联研究
EWAS-MLR	environment-wide association studies – multiple linear regression, 全环境关联研究-多元线性回归法
FDP	false discovery proportion, 错误发现率
GBD	Global Burden of Disease, 全球疾病负担项目
GIS	Geographic Information System, 地理信息系统
GUESS	graphical unit evolutionary stochastic search, 图形单元进化随机搜索
GWAS	genome-wide association studies, 全基因组关联研究
HELIX	Human Early-Life Exposome, 人类早期暴露组项目
HIA	Health Impact Assessment, 健康影响评估
HR	hazard ratio, 危险比
IER	integrated exposure-response relationship, 综合暴露反应关系
IQR	interquartile range, 四分位间距
LASSO	least absolute shrinkage and selection operator, 套索回归
Lowess	locally weighted scatterplot smoothing, 局部加权回归平滑曲线
MAE	mean absolute error, 平均绝对误差
MAE TP	mean absolute error of true predictors, 真实变量的平均绝对误差
NDVI	normalized difference vegetation index, 归一化植被指数
OR	odds ratio, 比值比
PAF	population attributable fraction, 人口归因分数
PCA	principle component analysis, 主成分分析

PCR	principle component regression, 主成分回归
PH	proportional hazard, 比例风险
PLS	partial least squares, 偏最小二乘回归
POI	point of interest, 兴趣点
RR	relative risk, 相对危险度
sPLS	sparse partial least squares regression, 稀疏偏最小二乘回归
TMREL	theoretical minimum risk exposure level, 理论最小风险暴露水平

第 1 章 绪论

1.1 研究背景

1.1.1 中国城镇化面临的健康挑战与已有对策

中国在过去二十年正经历着前所未有的快速城镇化进程，截至 2022 年，中国城镇化率达到 65.2%，超过半数的中国人居住在城市环境当中。快速城镇化的背景，除了显著提升了人民生活水平及城市经济发展水平之外，也带来了一系列的挑战——城市工业化带来的环境破坏问题、建成时间较早的老城区的空间品质下降问题、居民的生活方式发生显著转变等，这些挑战最终会对居民健康产生不良影响。

在这样的背景之下，中共中央国务院在 2016 年颁布了《健康中国 2030 规划纲要》。纲要指出，要以健康优先、改革创新、科学发展、公平公正为指导思想，以全民健康作为建设健康中国的根本目的。纲要第五篇为“建设健康环境”，其中既涉及大气、水、土壤等城市自然环境的污染防治，也涉及工业污染源排放等城市人工环境控制，并关注道路交通安全、食品药品安全、生产过程职业健康等方面。纲要第六篇“发展健康产业”当中，也对城市中的医疗服务、健身休闲运动等健康服务产业提出了要求。从中可以看出，纲要已经从城市自然环境、城市交通、城市健康服务基础设施等多个维度出发，旨在促进城市居民的健康水平。

以《健康中国 2030 规划纲要》作为引领者，国家相关部门以及建筑与规划行业相继出台了一系列标准，来推动健康城市建设。国家卫生健康委员会下辖的国家疾病预防控制中心（后文简称为国家卫健委疾控局）基于 2016 年发布的《关于开展健康城市健康村镇建设的指导意见》，于 2018 年颁布了《全国健康城市评价指标体系》，指标体系中包括健康环境、健康社会、健康服务、健康人群、健康文化 5 个一级指标，下设 20 个二级指标与 42 个三级指标。国家卫健委疾控局每年对所有国家卫生城市依据此指标开展评价工作。同时，在建筑与规划行业，2016 年首先颁布了《健康建筑评价标准》^[1]，2020 年相继颁布了《健康社区评价标准》^[2]与《健康小镇评价标准》^[3]，三个标准均从空气、水、舒适、健身、人文、服务六个建成环境维度，以控制项和评分项两个等级的形式，对健康人居环境的营造提出了可参考的规范要求。

目前，我国的城市发展正从增量开发逐步转向存量更新，继《关于进一步加强城市规划建设管理工作的若干意见》这一文件出台之后，城市更新、城市体检、城市高质量发展与城市双修（生态修复、城市修补）等发展战略成为城市规划领

域的潮流。但是，目前健康城市的相关发展战略，仍需要大量的建成环境与居民各种健康结局之间的定量关联规律作为科学支撑，而这方面的研究成果及数据基础还相对薄弱，需要更多研究来补充完善。

1.1.2 中国居民心血管健康水平及预后问题应受重视

根据国家统计局数据，2022 年中国人口死亡率为 7.37‰^[4]；全球疾病负担项目（Global Burden of Disease, GBD）在著名医学期刊《柳叶刀》最新发文显示，2019 年中国每 10 万人的疾病负担（burden of disease）为 26,871 人年^[5]。目前，中国人群的疾病构成中，慢性非传染性疾病占据主导。2019 年中国心血管疾病导致的死亡人数达到每 10 万人 322 人，占到总死亡人数的 43%，其中，脑卒中贡献了 21%，缺血性心脏病贡献了 18%。从疾病负担角度，中国心血管疾病同样贡献最大，达到每 10 万人 6,463 人年，占到总疾病负担的 24%，其中脑卒中 12%、缺血性心脏病 9%^[5]。因此，缺血性心脏病、脑卒中这两类疾病是影响中国人群健康水平的罪魁祸首。

缺血性心脏病当中，超过 90% 的病例属于大家熟知的冠心病，其中，最为严重的一种冠心病就是急性心肌梗死（AMI, acute myocardial infarction）。根据《中国心血管报告 2021》，2019 年中国城市与农村的急性心肌梗死死亡率分别达到了每 10 万人 60 人和 78 人，在整个心血管疾病（城市与农村分别为每 10 万人 278 人和 323 人）当中占了将近 30%^[6]。除此之外，在医疗发达的国家或地区，急性心肌梗死的病死率较低，比如北京市 2012 年急性心肌梗死的病死率仅为 9.0%^[7]。另一项全国性研究也指出，中国 ST 段抬高型心肌梗死的入院病死率仅为 10.2%^[8]。上述研究结果表明，大量患上急性心肌梗死的居民会康复，从而转变为冠心病的长期慢性患者，随时面临再次复发的风险。从以上的大量数据中可见，急性心肌梗死的预后工作显得尤为重要，有效的预后工作可以降低急性心肌梗死复发的风险，保障患者的健康水平。

然而目前，在健康城市的大背景之下，城市的各种建成环境暴露因素与急性心肌梗死复发之间的关联规律尚缺乏足够的定量研究，使得如何从建成环境维度助力解决急性心肌梗死预后问题尚不明晰。

1.1.3 新方法与新数据环境为建成环境大规模精细化测度提供支持

要揭示建成环境暴露与急性心肌梗死复发之间的定量关联规律，离不开对急性心肌梗死患者的建成环境暴露水平的大规模、精细化评价。传统的城市研究当中，受到开源数据的样本量、时空间精度，以及对于图片等多源数据的分析技术

的限制,难以对居民建成环境多个维度的暴露水平开展大规模、精细化测度。随着新数据环境、新技术方法的不断涌现,该难题得到了实质改善。现有的建成环境新数据主要包括以下几种类型:

(1) 在线地图矢量数据:主要包括路网数据(包含道路位置、几何形状、道路名称、道路等级等信息)、建筑物数据(包含建筑轮廓、建筑层数、建筑面积等信息)、兴趣点数据(points of interest, POI, 比如餐厅、地铁站、公园、医院等,包含名称、位置、分类等属性)等。这些城市矢量数据可以通过百度地图、高德地图、Open Street Map (OSM)等在线地图开源数据获得,并通过二维和三维空间分析的方法,对城市中建成环境的相应物理要素进行暴露评价,比如路网密度、建筑容积率、各类兴趣点密度等^[9, 10]。

(2) 遥感影像数据:主要包括城市自然环境与人工环境的遥感影像数据。前者通过光谱分析等手段,识别地表的水体等自然要素边界、植被覆盖水平(例如归一化植被指数(normalized difference vegetation index, NDVI))^[11]、城市热岛效应^[12]、近地表大气污染状况(例如PM_{2.5}浓度、二氧化氮浓度分析)^[13, 14]等。后者主要侧重观测城市中的建成区,可用于识别建筑物分布等人工地表信息,夜光遥感影像数据亦属于此类数据^[15]。该类数据可通过Landsat系列遥感、MODIS系列遥感、珞珈一号夜光遥感、NASA Earthdata 数据开放平台等渠道获得。

(3) 街景图片数据:指运用街景图片采集车等设备收集的真实街道照片,能反映街道周边的建成环境。该类数据可通过百度地图、谷歌地图等在线地图的街景图片数据获取,并可以进一步通过人工审计、计算机深度学习等技术方法,对图像内容特征进行学习,实现图像内要素的识别、分类、检测,比如城市建筑风貌、街道绿视率、街道视野开阔度等指标的测度^[16-18]。该类数据往往还具备历史数据查询功能,从而实现多年回溯的暴露评价。

(4) 人口空间分布数据:指基于一定区域的人口统计数据,结合其他城市新数据(如到达主干道的距离、地表高程、夜光遥感等)作为协变量,将统计人口分配到高精度栅格层面的人口分布数据集。该类公开数据目前主要包括南安普敦大学开发的WorldPop数据集、美国橡树岭国家实验室开发的LandScan数据集、欧盟委员会联合研究中心的GHS-POP数据集等。

(5) 人群活动数据:主要包括手机信令数据、公交卡刷卡数据、公共交通轨迹数据、社交媒体数据等。手机信令数据是指手机用户与通讯公司固定发射基站之间的通信记录,带有准确的时间、位置属性和通话、流量等数据信息,可用于研究就业时空分布、城市人群流动、出行OD分析、道路拥堵分析等。公交卡刷卡数据包含公交车上下车站位置信息、刷卡时间、持卡类型等数据,可用于城市

职住平衡、公交线路优化等方面的研究。公共交通轨迹数据主要包括公交车、网约车、共享单车在运营过程中的 GPS 数据，可与路网数据、交通出行调查数据、交通分析小区边界等结合，用于道路拥堵分析、主要商圈客源吸引力分析等研究。社交媒体数据种类繁多，如微博数据，其包含微博发文信息、发文位置等，再比如大众点评的用户评论数据，该类数据可用于城市活力分析等方面。

上述基于新数据和新技术的建成环境暴露评价方法具备效率高、成本低的优势，能够实现空间、时间两个维度上的客观评价，使得建成环境的健康效应研究成为可能。

1.2 研究目标与内容

1.2.1 研究目标与科学问题

基于以上研究背景，本研究重点关注多种建成环境暴露对急性心肌梗死复发的关联识别及其规划应对，并提出以下两个科学问题：

(1) 建成环境暴露与急性心肌梗死复发之间的定量关联规律是什么？

(2) 如何将以上定量规律应用到规划实践当中，从而降低居民的急性心肌梗死复发风险？

具体而言，对于第一个科学问题，本研究首先需要识别众多建成环境暴露指标当中哪些与急性心肌梗死复发之间存在关联，其次回答这些建成环境指标与急性心肌梗死之间的关联效应有多强，具有怎样的暴露反应关系（*exposure-response relationship*）。研究难点在于，急性心肌梗死患者同时暴露于多种建成环境因素当中，且这些因素之间往往存在较强的多重共线性问题，如何利用统计分析工具进行暴露变量的有效选择是需要重点考虑的问题。此外，本研究将采用急性心肌梗死患者队列数据而非横断面数据来进行分析，以往的线性回归、逻辑回归等分析方法无法体现队列数据的时间尺度，如何利用生存分析等统计分析工具来处理复杂的多变量生存数据也是一大难点。

对于第二个科学问题，本研究首先需要基于第一个科学问题给出的暴露反应关系，给出北京市现状乡镇街道及居住小区尺度 AMI 复发总风险的评估方法，其次需要基于复发总风险评估结果，识别出待提升小区，并从城市规划领域出发提出空间提升的应对策略。研究难点在于，如何处理从个体尺度推演到区域尺度（乡镇街道或居住小区）的尺度转变，实现复发风险评估；如何将定量的复发风险评估结果，转译为定性的城市规划应对策略，从而为健康城市提供科学依据。

1.2.2 研究内容

本研究将针对以上两个科学问题，分别展开如下研究内容：

针对第一个科学问题，本研究将采用暴露组学的研究方法，基于 2013-2019 年北京市医疗系统内所有急性心肌梗死的患者队列数据，对患者住址周边尽可能多的建成环境暴露因素进行不同空间、时间尺度的大规模测度。其次，采用多种统计回归模型，对多因素的暴露水平与急性心肌梗死的复发概率之间的关系进行统计分析，从而筛选出显著的建成环境暴露变量，并计算出暴露因素每增加一个单位水平带来急性心肌梗死复发风险提升或降低的百分比及其置信区间，从而得到这些因素与 AMI 复发之间的暴露反应关系。

针对第二个科学问题，本研究先对 2021 年北京市乡镇街道及居住小区的上述显著建成环境暴露指标进行了测度，再利用前一步得到的暴露反应关系，利用人口归因分数的概念，对北京市乡镇街道及居住小区的 AMI 复发总风险进行评估与可视化。本研究将评估结果与基准水平（北京市各小区风险值的 60%分位数）进行比较，识别出北京市范围内复发风险较高的待提升小区（即后 40%）。然后将筛选出的建成环境暴露指标分为优、差两档，并将这些待提升小区根据暴露指标的优劣情况进行分类，对每一类小区展开规划应对策略的讨论，为各个建成小区的空间品质健康提升提供理论依据。

1.3 研究方法与技术路线

1.3.1 研究方法

（1）文献综述

本研究首先通过文献综述方法，收集目前已有的建成环境暴露与心血管疾病之间的流行病学研究，为本研究的备选建成环境暴露指标提供参考依据，也能为本研究最终筛选出的显著的建成环境暴露指标提供证据支持。其次，本研究通过对暴露组学研究方法的文献综述，确定了将暴露组学方法应用于建成环境领域研究的基本方法，并帮助梳理了统计模型的性能比选及分析方法。

（2）开源数据获取

本研究除了使用少量传统统计数据之外，侧重采用多类开源新数据来进行建成环境暴露评价，包括在线地图数据、街景图片数据、遥感影像数据等。其中，在线地图矢量数据（如 POI）和街景图片的多年历史数据主要基于多年在线地图 API 接口进行爬取，遥感影像数据则通过 Landsat 卫星影像开源数据库直接获取。

(3) GIS 空间分析

本研究利用地理信息系统软件 ArcGIS 和 QGIS 对空间数据进行空间落位、空间计算、统计分析,包括缓冲区分析、叠加分析(相交、融合、裁剪等)、距离计算、按照字段属性进行融合数据或统计分析等方法,从而实现点、线、面多个尺度下对建成环境各指标的密度、多样性、距离、比例等多维度的空间分析,并能提供相应的可视化效果。

(4) 统计建模分析

本研究首先对常见的几类暴露组学变量选择技术进行了统计性能的比较,遴选出准确度、稳健性最优的变量选择算法供后续统计模型嵌套使用。然后,针对队列研究设计及生存资料的特征,采用生存分析中最常用的 Cox 比例风险模型,结合变量选择算法,对各个建成环境暴露指标的危险比(hazard ratio, HR)进行评估,并进行比例风险假设和线性性假设的检验与修正,得到最终的暴露反应关系。

(5) 归因疾病负担评价方法

本研究最终利用获得的暴露反应关系,采用人口归因分数(population attributable fraction, PAF)的基本概念与相关方法,来表征建成环境因素为急性心肌梗死复发带来的归因疾病负担水平,并在乡镇街道和居住小区两个空间尺度上开展评价。

1.3.2 技术路线

围绕研究目标,借助上述研究方法,本研究的技术路线如图 1.1 所示,整项研究分为 3 个步骤进行:第一,对建成环境暴露水平进行空间、时间两个维度的大规模、精细化测度;第二,对建成环境暴露与急性心肌梗死复发间的暴露反应关系开展变量选择与生存分析;第三,对北京市现阶段乡镇街道及居住小区的急性心肌梗死复发风险进行评估,并识别出待提升小区,分类后提供针对性的规划应对策略。

因此,本论文的结构如下:第 1 章介绍研究背景、研究目标及内容、研究方法和技术路线;第 2 章针对建成环境与心血管疾病之间的已有研究结果、常见研究方法以及城市规划应对等理论开展文献综述,并结合文献的局限之处提出本研究的改进方向;第 3 章介绍本研究的研究对象(即患者队列信息)和可使用的数据基础;第 4 至 6 章详细展开上述技术路线的三个步骤,第 4 章介绍建成环境暴露水平的评价方法及结果,第 5 章介绍建成环境暴露与急性心肌梗死复发之间关联的流行病学研究方法,第 6 章基于第 5 章得到的暴露反应关系,对北京市现状进行风险评估,并给出规划应对策略;第 7 章对本研究进行总结与展望。

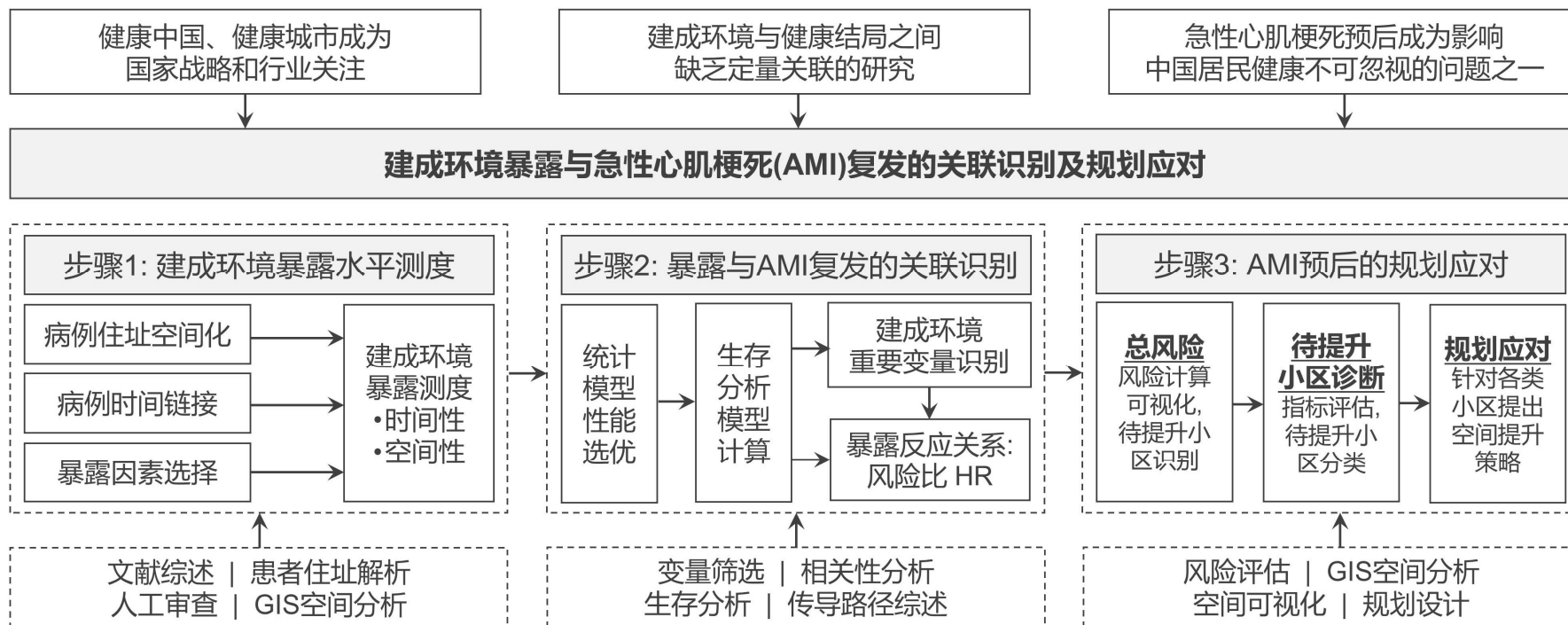


图 1.1 研究技术路线

第2章 理论与文献综述

2.1 建成环境的概念界定及主要暴露因素

在城市规划领域中,物质环境(physical environment)泛指包括土地、植被、水体、建筑、交通和市政基础设施以及所有的自然资源等为人类生存和社会经济发展提供支持的客观实体存在的集合^[19]。在物质环境当中,建成环境(built environment)指的是人类生产、生活活动而形成的人居环境状态,范围上从聚落整体到具体建筑物,同时也包括各种支持性基础设施^[19]。建成环境从土地利用的视角来看,可进一步分为建筑占据和未占据两种类型的空间,前者主要针对建筑功能所赋予的基础服务设施及其相应的室内环境,而后者可被认为是城市中广义的开敞空间,包括绿地、广场、公园、道路、水面等^[20]。本论文中的建成环境将涉及城市中广义的开敞空间环境及建筑本身所提供的基础服务或设施,而不包含对室内环境(如室内热环境、室内光环境、室内空气污染物等)的探讨。但同时,人暴露于建成环境的同时,还会暴露于自然环境当中,其中既包含空气污染等物质环境因素,也包含噪声、气温、辐射等非物质的物理因素。因此,考虑建成环境对人群行为和健康的影响时,通常还需要同时考虑自然环境的影响。

建成环境暴露因素众多,所以研究建成环境相关问题时需要能够提供潜在的建成环境暴露因素清单。目前应用最广泛的建成环境因素分类框架之一,是由 Reid Ewing 在研究城市交通问题时提出的 5D 理论,将建成环境各类因素划分为密度(density)、混合度(diversity)、设计(design)、到交通设施的距离(distance to transport)、目的地可达性(destination accessibility)这五个维度^[21]。邻里环境(neighborhood environment)作为建成环境中人活动时间最长、频率最高的空间,也备受学者关注。Zhang 等人^[22]通过“通往公平和健康城市的道路”国际工作组的长期研讨,总结得到了 33 个可能影响公共健康的邻里风险因素,包括物理环境因素、服务与商业设施环境因素、污染因素、社会环境因素、安全与伤害因素这五大类别。其中,与物质环境相关的主要是前三大类的 18 个风险因素,包括绿色空间、可步行性、建筑密度、蔬果商店可达性、社区健康设施可达性、空气污染等。在多源数据新环境下,大规模、精细化测度这些建成环境暴露因素已经成为现实,而这些建成环境暴露因素的分类与清单,可为筛选和识别影响公共健康的建成环境因素提供理论基础。

2.2 建成环境暴露与心血管疾病之间的关联研究

本节首先对建成环境与心血管疾病之间关联的流行病学研究进行文献综述,从而对本研究筛选影响急性心肌梗死复发的建成环境因素提供理论依据。总体上,有关绿地的心血管流行病学证据较为充分,而关于其他建成环境要素,如建筑密度、功能密度及混合度、交通可达性、步行指数、公共交通可达性、各类目的地可达性等因素,流行病学研究仍较为缺乏,且多为横断面研究。除建成环境之外,心血管健康的相关研究还常纳入空气污染这一典型自然环境因素,本节最后对空气污染与心血管疾病的流行病学研究进行了文献综述。

2.2.1 绿地与心血管疾病关联的流行病学研究

绿地是大部分流行病学研究集中关注的建成环境要素,如表 2.1 所示。绿地的暴露水平在文献当中有多种测度方式,主要包括绿地面积占比、到达绿地的最近距离、NDVI 这三种。

绿地面积占比这一类指标,往往采用选定区域内或一定距离缓冲区内绿地面积所占比例的方式来测度绿地的暴露水平。Maas 等人^[23]在荷兰开展了一项 34.5 万人的横断面研究,发现 1 km 缓冲区内的绿地占比对高血压、心脏病、冠心病、中风的发病风险均有显著保护作用,具体而言,每升高 10% 的绿地占比四种健康结局的 OR 值分别为 0.99 (95% CI: 0.98 ~ 1.00)、0.98 (0.97 ~ 0.99)、0.97 (0.95 ~ 0.99)、0.98 (0.95 ~ 1.00)。Mitchell 等人^[24]对英国 65 岁以下成年人的一项横断面研究表明,绿地占比位于最低四分位区间的人群,其循环系统疾病发病危险比 (IRR = 2.19) 显著高于绿地占比最高的四分位人群 (IRR = 1.54)。Richardson 等人^[25]在一项英国的横断面研究中还探讨了绿地对心血管疾病死亡风险影响的性别差异,研究结果显示,绿地对于男性的心血管疾病死亡风险有显著的保护作用 (IRR = 0.95, 95% CI: 0.91 ~ 0.98),但对女性的保护作用并不显著 (IRR = 1.00, 95% CI: 0.95 ~ 1.06)。此外,还有多篇横断面和生态学研究聚焦于绿地与心血管疾病死亡或发病风险的影响,均得到了保护效应结果,只是部分研究效应不显著,这可能与样本量太小以及研究设计有关^[26-29]。

第二种测度方式是用到达绿地的最近距离来表征居民对绿色空间的暴露水平, Hu 等人^[30]进行了一项生态学研究,借助贝叶斯层次 Poisson 回归模型,发现到达各类绿色空间的距离越近,对脑卒中有显著保护作用 (回归系数 $\beta = -0.161$, 95% CI: -0.289 ~ -0.031)。Tamosiunas 等人^[31]在立陶宛进行了一项 5112 人的队列研究,发现到达最近的面积大于 1 公顷的绿地的距离越长,居民心血管疾病死亡风险越高,距离最高三分位数的居民比最低三分位数的居民风险高出 36% (HR =

1.36, 95% CI: 1.03 ~ 1.80)。

此外, 较多文献采用的是第三种绿地暴露测度方式——利用遥感影像获得 NDVI。NDVI 是通过遥感影像中红光波段和近红外波段的反射值来计算的, 可以反应地表的植被覆盖程度。由于遥感影像数据具有多年度易获取、空间分辨率较高的特征, 因此, NDVI 这种测度方式使得队列研究成为可能, 而不像前两种绿地暴露测度方式均以横断面研究、生态学研究为主。Villeneuve 等人^[32]在加拿大开展的一项 57 万余人的队列研究发现, 居民住址周边 500 米缓冲区内每升高 1 个四分位距 (interquartile range, IQR) 的 NDVI 水平, 缺血性心脏病和中风的死亡风险均可以显著降低 4%。Wilker 等人^[33]一项美国的队列研究表明, 住址周边 250 米缓冲区内 NDVI 对缺血性中风死亡具有显著的保护作用, NDVI 最高与最低四分位数之间的 HR 为 0.78 (95% CI: 0.63 ~ 0.97)。Chen 等人^[34]的一项持续 14 年、共计 136 万人的加拿大队列研究也发现了 250 米缓冲区内 NDVI 对心血管疾病死亡、急性心肌梗死发病和心脏衰竭发病都具有显著的保护作用。其余研究多为横断面研究, 但也能够揭示绿地暴露对各类心血管疾病有较强的保护作用, 尽管部分研究结果显示该保护效应并不显著^[35-37]。

此外, 部分学者对绿地暴露的心血管健康影响效应开展的系统综述与荟萃分析 (meta analysis) 研究。Twohig-Bennett 等^[38]对 103 篇观察性研究和 40 篇干预性研究进行了荟萃分析, 结果表明, 增加的绿色空间暴露能够显著降低心血管疾病死亡风险 (OR = 0.84, 95% CI: 0.76 ~ 0.93), 不过, 该篇研究未区分绿色空间暴露的测度方式, 暴露评价过程描述得比较笼统。Gascon 等的荟萃分析以及 James 等的文献综述研究^[39, 40]也同样得到了绿地对心血管健康显著保护的结论。总体而言, 绿色空间对心血管疾病的保护作用有较多的流行病学研究支撑, 且不乏因果推断较强的队列研究以及系统综述类研究, 属于建成环境暴露研究中较为成熟的部分。

表 2.1 绿地与心血管疾病之间关联的流行病学研究

作者	研究年份	研究设计	研究人群	暴露	健康结局	统计方法	健康效应
Bixby 等 ^[26]	2002-2009	生态学研究	英国最大的 50 个城市的 15-64 岁人群	城市的绿地覆盖率	心血管疾病死亡	Poisson 回归	最高与最低五分位数比较： 男性 RR = 0.95 (0.86 ~ 1.23) 女性 RR = 0.94 (0.83 ~ 1.07)
Lachowycz 等 ^[27]	2007-2008	横断面研究	英国 165,424 位成人	普查区域内、5 km/10 km 缓冲区内绿地占比	循环系统疾病死亡	负二项回归	绿地占比越高，循环系统疾病死亡风险越低，以贫困人群最为突出（仅提供图，未提供具体数值）
Maas 等 ^[23]	2001	横断面研究	荷兰 345,143 人	1 km、3 km 缓冲区内绿地占比	高血压、心脏病、冠心病、中风患病	多层次 logistic 回归	1 km 缓冲区每升高 10%绿地占比： 高血压 OR = 0.99 (0.98 ~ 1.00) 心脏病 OR = 0.98 (0.97 ~ 0.99) 冠心病 OR = 0.97 (0.95 ~ 0.99) 中风 OR = 0.98 (0.95 ~ 1.00)
Mitchell 等 ^[24]	2001-2005	横断面研究	英国 65 岁内的 40,813,216 人	普查区域内的绿地占比	循环系统疾病患病	负二项回归	最低五分位数：IRR = 2.19 (2.04 ~ 2.34) 最高五分位数：IRR = 1.54 (1.38 ~ 1.73)
Richardson 等 ^[25]	2001	横断面研究	英国 16 至 64 岁的 2860 万人口	普查区域内的绿地占比	心血管疾病死亡	负二项回归	最高与最低四分位数比较： 男性 IRR = 0.95 (0.91 ~ 0.98) 女性 IRR = 1.00 (0.95 ~ 1.06)
Richardson 等 ^[28]	1996-2005	生态学研究	新西兰 15 至 64 岁的 1,546,405 人	普查区域内的绿地占比	心血管疾病死亡	负二项回归	最高与最低四分位数比较： 所有绿地 IRR = 1.01 (0.91 ~ 1.11) 可用绿地 IRR = 0.96 (0.90 ~ 1.03)
Richardson 等 ^[29]	2004	生态学研究	美国最大的 49 座城市的 4300 万人	普查区域外加 1 km 缓冲区内绿地占比	心脏病死亡	线性回归	每升高 1%绿地占比： 男性 β = -0.51 (-3.19 ~ 2.16) 女性 β = -0.49 (-2.17 ~ 1.19)

续表 2.1 绿地与心血管疾病之间关联的流行病学研究

作者	研究年份	研究设计	研究人群	暴露	健康结局	统计方法	健康效应
Hu 等 ^[30]	1998-2002	生态学研究	美国 17 个州的人口	住址到达绿色空间的最短距离	中风死亡	贝叶斯层次 Poisson 回归	$\beta = -0.161 (-0.289 \sim -0.031)$
Tamosiunas 等 ^[31]	2006-2011	队列研究	立陶宛 45 至 72 岁的 5,112 人	住址到绿地的最短距离	心血管发病	Cox 比例风险模型	最高与最低三分位数比较: HR = 1.36 (1.03 ~ 1.80)
James 等 ^[35]	2000-2008	队列研究	美国 108,630 位女性	250 m 缓冲区内 NDVI	冠心病、中风死亡	Cox 比例风险模型	每升高 0.1 NDVI: 冠心病 HR = 1.02 (0.87 ~ 1.20) 中风 HR = 0.79 (0.62 ~ 1.01)
Pereira 等 ^[36]	2003-2009	横断面研究	澳大利亚 11,404 位 25 岁以上成年人	1600 m 缓冲区内 NDVI	冠心病、中风入院	Logistic 回归	最高与最低三分位数比较: OR = 0.63 (0.43 ~ 0.92)
Villeneuve 等 ^[32]	1982-2004	队列研究	加拿大 35 岁以上的 574,840 人	500 m 缓冲区内 NDVI	缺血性心脏病、中风死亡	Cox 比例风险模型	每升高 1 个 IQR 的 NDVI: 缺血性心脏病 RR = 0.96 (0.95 ~ 0.98) 中风 RR = 0.96 (0.93 ~ 0.98)
Wilker 等 ^[33]	1999-2012	队列研究	美国 21 岁以上的 1645 位病人	250 m 缓冲区内 NDVI	缺血性中风死亡	Cox 比例风险模型	最高与最低四分位数比较: HR = 0.78 (0.63 ~ 0.97)
Wang 等 ^[37]	2010-2011	横断面研究	美国 65 岁以上的 249,405 人	街区尺度内的 NDVI 平均值	急性心肌梗死、缺血性心脏病、心脏衰竭发病	多层次 logistic 回归	最高与最低三分位数比较: 急性心肌梗死 OR = 0.75 (0.63 ~ 0.90) 缺血性心脏病 OR = 0.80 (0.77 ~ 0.83) 心脏衰竭 OR = 0.84 (0.80 ~ 0.88)
Chen 等 ^[34]	2000-2014	队列研究	加拿大 35 至 100 岁的 1,362,723 人	250 m 缓冲区内 NDVI	心血管疾病死亡、急性心肌梗死发病、心脏衰竭发病	多层次 Cox 比例风险模型	每升高 1 个 IQR 的 NDVI: CVD 死亡 HR = 0.90 (0.88 ~ 0.92) AMI 发病 HR = 0.93 (0.91 ~ 0.96) HF 发病 HR = 0.94 (0.93 ~ 0.96)

2.2.2 其他建成环境因素与心血管疾病关联的流行病学研究

前一小节对文献研究较为充分的绿地的心血管健康效应进行了综述,本小节聚焦于其他建成环境暴露因素。多篇综述文章表明,已有建成环境相关研究多是将建成环境因素与心血管疾病的风险因素(如高血压、糖尿病、肥胖症、BMI、体力活动不足等)进行关联,而直接对心血管疾病结局进行研究的流行病学证据较为匮乏^[41-43]。并且,关注的建成环境因素较为零散,多集中于交通环境、食物环境、城市形态等方面。主要文献信息已总结在表 2.2 中。

首先是交通环境层面。Kan 等人^[44]在 1987 年到 2002 年在美国进行了一项前瞻性队列研究,研究指出,居民住处附近道路的交通密度是冠心病发病的显著负面因素,到主要道路的距离越近也可能造成冠心病发病风险上升。Chum 等人^[45]在美国开展的一项横断面研究表明,距离道路 100 米以内是心血管疾病发病的显著风险因素,其风险是距离道路 100 米以上人群的 2 倍。Kwan 等人^[46]在马来西亚的一项生态学研究发现,道路交叉口密度越高,与越高的心血管疾病入院率显著相关,交叉口密度处于最高四分位区间的居民,其心血管发病风险是处于最低四分位区间居民的 3.5 倍。

其次是食物环境方面。Hamano 等人^[47]在瑞典一项 4,309,674 人的队列研究中,探究了快餐店可达性、酒吧可达性、体育活动设施可达性、医疗设施可达性对脑卒中发病及死亡的影响,在校正了年龄、收入等个人因素及教育水平、就业率、社会福利等邻里社区因素之后,只有快餐店可达性一项暴露因素显示会对男性和女性的脑卒中发病或死亡有显著促进作用,其余暴露因素的关联不显著。关于快餐店设施的此结论与 Lee 等人^[48]在韩国开展的横断面研究相符。Mobley 等人^[49]在美国开展的横断面研究中,探讨了杂货店、快餐店、餐厅、便利店设施人均拥有量与心血管发病的关系,均呈现较弱但不显著的保护关系。Cebrecos 等人^[50]在西班牙开展的横断面研究中,将食品店密度、食品店多样性、步行指数和烟酒店密度集合成为心脏健康指数的综合指标,并发现,心脏健康指数与心血管疾病发病风险呈显著负相关的关系。上文提到的 Chum 等人^[45]的横断面研究,也纳入了快餐店和食品店密度的指标,但均未呈现显著的健康效应。

最后一方面是城市形态层面。Griffin 等人^[51]在美国 45,376 位女性组成的队列研究中,探究了居住建筑密度、街道连通性、土地利用混合度、中心度这四个城市规划领域经常关注的指标对冠心病及心肌梗死的影响。研究发现,在校正了个人社会经济属性、健康水平及健康行为之后,越高的建筑密度、越高的土地利用混合度会对冠心病事件的发生起到显著的保护作用。土地利用混合度对冠心

表 2.2 其他建成环境因素与心血管疾病之间关联的流行病学研究

作者	研究年份	研究设计	研究人群	健康结局	统计方法	建成环境因素及其健康效应
Kan 等 ^[44]	1987-2002	队列研究	美国 13,309 名 中年人	冠心病发病	Cox 比例风险 模型	交通密度: HR = 1.03 (1.01 ~ 1.05) 到主干道距离小于 300 m: HR = 1.12 (0.95 ~ 1.32)
Hamano 等 ^[47]	2005-2007	队列研究	瑞典 35 至 80 岁 的 4,309,674 人	中风发病	多层次 logistic 回归	快餐店可达性: 男性 OR = 1.02 (1.00 ~ 1.05), 女性 OR = 1.03 (1.00 ~ 1.06) 酒吧可达性: 男性 OR = 1.01 (0.97 ~ 1.05), 女性 OR = 1.00 (0.96 ~ 1.05) 体育活动设施可达性: 男性 OR = 1.02 (0.99 ~ 1.05), 女性 OR = 1.02 (0.99 ~ 1.06) 保健设施可达性: 男性 OR = 1.02 (0.99 ~ 1.05), 女性 OR = 1.03 (1.00 ~ 1.07)
Griffin 等 ^[51]	1993-2005	队列研究	美国 45,376 名 女性	冠心病发病、冠 心病死亡或心 肌梗死	Cox 比例风险 模型	街道连通性: 发病 HR = 0.98 (0.94 ~ 1.02), 死亡 HR = 0.95 (0.89 ~ 1.01) 中心度: 发病 HR = 0.95 (0.91 ~ 1.00), 死亡 HR = 0.98 (0.91 ~ 1.05) 土地利用混合度: 发病 HR = 0.99 (0.94 ~ 1.04), 死亡 HR = 0.90 (0.84 ~ 0.97) 居住建筑密度: 发病 HR = 0.94 (0.91 ~ 0.97), 死亡 HR = 0.90 (0.86 ~ 0.95)
Cebrecos 等 ^[50]	2014	横断面研 究	西班牙的 128 个 邻里社区、2415 个调查区域	心血管疾病发 病	最小二乘回 归、地理加权 回归	心脏健康指数 (食品店密度、食品店多样性、步行指数、 烟酒店密度): 全局 $\beta = -0.037, p < 0.05$; 局部地理加权 $\beta =$ -0.002, $p < 0.05$

续表 2.2 其他建成环境因素与心血管疾病之间关联的流行病学研究

作者	研究年份	研究设计	研究人群	健康结局	统计方法	建成环境因素及其健康效应
Mobley 等 [49]	2001-2002	横断面研究	美国 2692 名女性	冠心病发病	最小二乘回归	10 年内 CHD 发病风险对数的回归系数(标准误): 土地利用混合度: $\beta = -0.22 (0.09)$, $p < 0.05$ 每 1k 位居民的健身设施数量: $\beta = -0.16 (0.08)$, $p < 0.05$ 每 1k 位居民的杂货店数量: $\beta = -0.01 (0.05)$ 每 1k 位居民的快餐店数量: $\beta = -0.01 (0.02)$ 每 1k 位居民的餐厅数量: $\beta = -0.25 (0.16)$ 每 1k 位居民的便利店数量: $\beta = -0.09 (0.08)$
Chum 等 [45]	2009-2011	横断面研究	加拿大 2411 人	心血管疾病、心肌梗死发病	多层次 logistic 回归	心血管疾病发病: 食品店密度(Q1 vs Q4): OR = 1.03 (0.92 ~ 1.14) 快餐店密度(Q1 vs Q4): OR = 1.01 (0.93 ~ 1.09) 公园和娱乐设施(Q1 vs Q4): OR = 1.03 (0.11 ~ 1.95) 暴力犯罪(Q1 vs Q4): OR = 0.82 (0.72 ~ 0.92) 需大修的房屋占比(Q1 vs Q4): OR = 0.99 (0.51 ~ 1.47) 距离道路 100 m 内: OR = 1.97 (1.67 ~ 2.49)
Servadio 等 [52]	2006-2016	横断面研究	美国亚特兰大州最大的 4 座城市	冠心病、中风发病	条件自相关回归模型	交叉口密度: 冠心病 $\beta = -0.0139$; 中风 $\beta = -0.0197$ 公园可达性: 冠心病 $\beta = 0.2581 (p < 0.05)$; 中风 $\beta = 0.2016 (p < 0.05)$
Jia 等 [53]	2015	横断面研究	中国 40 岁以上的 1944 人	冠心病、中风发病	混合效应 logistic 回归	中年人: NDVI: 冠心病 OR = 0.30 (0.12 ~ 0.66); 中风 OR = 0.80 (0.51 ~ 0.90) 步行得分: 冠心病 OR = 0.51 (0.33 ~ 0.79); 中风 OR = 0.57 (0.44 ~ 4.72) 邻里步行指数: 冠心病 OR = 0.98 (0.94 ~ 1.08); 中风 OR = 0.85 (0.70 ~ 0.98)

续表 2.2 其他建成环境因素与心血管疾病之间关联的流行病学研究

作者	研究年份	研究设计	研究人群	健康结局	统计方法	建成环境因素及其健康效应
Jia 等 ^[53]	续上表					老年人： NDVI：冠心病 OR = 0.31 (0.18 ~ 0.54)；中风 OR = 0.81 (0.58 ~ 0.94) 步行得分：冠心病 OR = 0.78 (0.58 ~ 0.99)；中风 OR = 0.89 (0.35 ~ 5.01) 邻里步行指数：冠心病 OR = 0.98 (0.91 ~ 1.07)；中风 OR = 0.88 (0.71 ~ 0.99)
Lee 等 ^[48]	2013-2014	横断面研究	韩国 40 岁以 上的 50,741 位 成人	中风、心肌 梗死发病	Logistic 回 归	40-50 岁： 快餐店密度(T3 vs T1)：中风 OR = 0.64 (0.44 ~ 0.92) 到公共交通的距离(T2 vs T1)：心肌梗死 OR = 1.36 (1.07 ~ 1.74) 60 岁以上： 人口密度(T3 vs T1)：中风 OR = 1.02 (0.78 ~ 1.32) 到公共交通的距离(T3 vs T1)：中风 OR = 0.91 (0.70 ~ 1.18) 街道连通度(T3 vs T1)：中风 OR = 0.93 (0.74 ~ 1.16)
Kwan 等 ^[46]	2004-2016	生态学研究	马来西亚吉隆 坡市	心 血 管 疾 病、缺血性 心脏病入院	广义线性混 合模型	(只展示显著的结果) 人口密度(Q3 vs Q1)：CVD OR = 0.29 (0.11 ~ 0.74) 土地利用混合度(Q3 vs Q1)：IHD OR = 0.57 (0.34 ~ 0.94) 休闲区域密度(Q4 vs Q1)：IHD OR = 0.54 (0.33 ~ 0.89) 道路交叉口密度(Q4 vs Q1)：IHD OR = 6.34 (2.25 ~ 17.91)；CVD OR = 3.46 (1.48 ~ 8.12) 注：CVD 为心血管疾病，IHD 为缺血性心脏病

病的显著保护作用也受到 Mobley 等人^[49]的研究结果的支持。Kwan 等人^[46]的生态学研究也表明土地利用混合度是缺血性心脏病入院事件的显著保护因素。

总体而言,其他建成环境因素当中,许多因素的心血管流行病学证据较少,且一致性较差,还需要未来进一步的深入研究。从已有证据看,交通密度越大或离主干道越近,以及土地利用混合度越低,与心血管疾病的发病风险越高有较为一致的显著关联。

2.2.3 空气污染与心血管疾病关联的流行病学研究

建成环境中的工业排放、交通尾气等因素都会造成空气污染。目前,世界卫生组织(World Health Organization, WHO) 2021 年更新的空气质量指南中,主要列出了细颗粒物($\text{PM}_{2.5}$)、可吸入颗粒物(PM_{10})、二氧化氮(NO_2)、二氧化硫(SO_2)、臭氧(O_3)和一氧化碳(CO)这六种主要空气污染物^[54]。由于空气污染的心血管流行病学研究较为丰富,此处仅对具有代表性的队列研究或系统综述进行文献梳理。

全球疾病负担项目(GBD)对风险因素暴露反应关系的纳入要求较为严苛,目前上述污染物仅纳入了 $\text{PM}_{2.5}$ 与缺血性心脏病、中风之间的暴露反应关系,该系统纳入了近 150 项 $\text{PM}_{2.5}$ 相关的队列研究,利用 IER(integrated exposure response)和 MR-BRT(meta-regression——Bayesian, regularized, trimmed)模型得到了非线性的暴露反应关系曲线^[55, 56]。例如,对于 50 岁的成年人,即使暴露在比《健康社区评价标准》要求还严格的 $30 \mu\text{g}/\text{m}^3$ 浓度下,缺血性心脏病和中风的相对危险度 RR 仍高达 1.61 (95% CI: 1.34 ~ 1.79) 和 1.56 (95% CI: 1.44 ~ 1.85) ^[56]。

而对于 $\text{PM}_{2.5}$ 以外的其余气态污染物对心血管健康的影响,尚未有非常成熟的理论基础,因此 GBD 项目暂时没有纳入考虑,但是也已经有大量学者予以关注。Chen 等人^[57]对发表于 1980 至 2019 年的 NO_2 、 SO_2 、 O_3 和 CO 四种大气污染物的健康效应文献进行了系统综述与荟萃分析,其中与心血管疾病的结果表明, SO_2 和 CO 对心血管疾病发病风险具有显著的危害作用(SO_2 : RR = 1.012, 95% CI: 1.007 ~ 1.018; CO : RR = 1.024, 95% CI: 1.011 ~ 1.038), 而 O_3 和 NO_2 的心血管健康影响效应并不显著,但是该研究同时纳入了队列研究、病例对照研究和大量横断面研究,对荟萃分析结果可能会造成影响。近年来也有一些新的队列研究支持这两种污染物的心血管健康影响。Niu 等人^[58]在中国开展了一项纳入 96,955 位成人的全国性队列研究,结果发现,五月至十月 O_3 的平均浓度每升高 $10 \mu\text{g}/\text{m}^3$ 时,心血管疾病的死亡风险会显著增加(HR = 1.093, 95% CI: 1.046 ~ 1.142),即使在调整了 $\text{PM}_{2.5}$ 浓度的情形下,该效应依旧显著。Strak 等人^[59]对共计 32.5 万人的 8 个欧洲

队列数据进行了整合分析, 研究发现, 即使是较低浓度的 NO_2 暴露, 仍会显著增加心血管疾病的死亡风险 ($\text{HR} = 1.089$, 95% CI: 1.060 ~ 1.120)。

总体而言, 空气污染对心血管疾病的发病和死亡风险都有着显著影响, 其中, 以 $\text{PM}_{2.5}$ 的健康效应为最强, 且得到国际公认。

2.3 暴露组学研究方法及其在建成环境领域的应用

2.3.1 暴露组学: 概念与方法

对各种环境风险因素暴露水平的测度, 是研究建成环境与人体健康结局关联的关键环节之一。近年来, 暴露的测度理论与方法也正在不断革新。人的健康水平是由先天的基因因素与后天的环境因素共同作用下决定的。然而, 在基因层面, 目前已经有了很成熟的全基因组关联研究技术 (genome-wide association studies, GWAS) 可以对人体基因组进行测序和分析, 也就是基因组学方法, 但是, 关于环境的全面暴露评价方法还较为匮乏。

在此背景下, Wild 在暴露 (exposure) 的基础上仿效基因组和基因组学, 提出了暴露组 (exposome) 和暴露组学 (exposomics) 的概念^[60, 61]。暴露组指的是一个个体从受孕到死亡之间所有环境暴露的集合, 可以从全面性 (scope) 与动态性 (dynamic) 两方面来解析。从全面性角度看, 暴露组将所有需要关注的环境暴露分为一般外源暴露 (general external exposures, 包括社会、经济、心理等因素)、特定外源暴露 (specific external exposures, 包括辐射、化学物质、饮食等)、人体内源暴露 (internal external exposures, 包括炎症、代谢、氧化应激等), 而传统的暴露科学只关注某一种或某一类暴露因素, 且集中在特定外源暴露层面。从动态性来看, 暴露组关注人从受精卵到死亡的整个时段, 或者是人生的多个关键时段, 而传统暴露只关注某个时间截面或是较短的一个时段, 比如前述建成环境领域内存在大量的横断面研究。此外, 随着高通量技术、大数据分析技术的出现, 暴露组学研究可以一次性同时测度多种环境因素的暴露水平, 而不再像传统暴露科学一样, 每种暴露分别用一种方法来进行测度^[60, 62]。已经有多个欧洲国家的队列研究采用暴露组学的研究方法对儿童或成人的环境暴露进行了采样分析^[63-65], 并分析了其时间变异性^[66-69]、不同时期各环境暴露因素之间的相关性^[70-72]、环境暴露因素与代谢物的相关性^[73-76], 以及各环境因素暴露水平的影响因素^[68, 77]。

暴露组学研究当中, 除了暴露测度方式出现革命性变化之外, 还有一部分研究着眼于暴露组与健康关联的方法创新。传统的暴露与健康的关联研究可认为是假设驱动型研究, 根据专业的先验认知提出假设之后, 选取 1 个关注变量与多个

根据认知确定的具有混杂效应的协变量,通过Cox比例风险模型、Logistic回归模型等模型得到风险比HR、相对危险度RR或比值比OR。与之对应,暴露组学研究方法是一种数据驱动型研究,它没有预先的假设,即没有代入研究者的先验知识,而将所有 N 个候选的潜在暴露因素都放入到统计模型内进行分析,此时就不再有固定的关注变量,且多种暴露因素也成为相互影响的协变量,该研究方法再通过数据驱动算法筛选出重要的暴露变量,并估计其效应值^[78, 79]。

面对众多可能具有多重共线性关系的环境暴露变量,暴露组学关联研究分析算法的难点之一,是如何有效地进行变量筛选。这类问题在统计学中被称为模型选择(model selection),常见的方法包括单变量-多变量回归方法、最优子集选择方法(如逐步回归)、收缩类方法(如岭回归、LASSO回归)、降维类方法(如主成分回归、偏最小二乘回归)^[80]。Agier等人^[81]选取了暴露组学中可以采用的几种变量选择算法——全环境关联研究方法(environment-wide association studies, EWAS)、全环境关联研究-多元线性回归方法(environment-wide association studies – multiple linear regression, EWAS-MLR,即对应上述的单变量-多变量回归方法)、弹性网模型(elastic net, ENET,即上述岭回归与LASSO回归的结合方法)、稀疏偏最小二乘回归(sparse partial least squares regression, SPLS,即上述偏最小二乘回归的延伸方法)、删除-替代-添加算法(deletion-substitution-addition algorithm, DSA,类似于上述的逐步回归算法)以及图形单元进化随机搜索(graphical unit evolutionary stochastic search, GUESS),并且采用交叉验证算法与敏感性验证方法,对这些算法的性能指标进行了评价与比较。该研究同时也指出,由于每组数据的内部相关性结构不同,这些变量选择算法的选优结果也可能不同,因此实际应用时,需要依据上述选优方法,对研究者自己的数据集进行分析,选出最佳的变量选择算法^[81]。

2.3.2 暴露组学在建成环境领域的应用

Andrianou等人^[82]进一步将暴露组学的概念引入建成环境领域,提出城市暴露组(urban exposome)的概念。该研究认为,城市暴露组研究需要对影响城市居民生活质量和健康水平的城市外部和内部因素进行的连续时空监测,并认为城市内部因素由于其宏观特性,应属于Wild提出的人类暴露组当中的一般外源暴露。

研究建成环境暴露相关的暴露组学研究目前较少,集中于欧洲HELIX、KABC、HERACLES等几个国际合作的队列研究当中,且主要是新生儿队列,如表2.3所示。Nieuwenhuijsen等人^[83]在HELIX项目内一项欧洲31,458名孕妇组成的队列研究当中,分别测度了孕期三大阶段的建成环境、大气污染、交通噪声、气象、紫

外辐射等暴露因素,并探究了各类暴露对婴儿出生体重的影响,其中,建成环境暴露包括了建筑密度、道路交叉口密度、土地利用混合度、公交站密度、设施密度与丰富度、人口密度、步行指数及绿地。研究利用 EWAS 算法,发现 NDVI、大型绿地存在、公交线路存在对出生体重有显著促进作用,步行指数、设施密度、建筑密度、道路交叉口密度、土地利用混合度对出生体重有显著降低作用,造成低出生体重的不良结局。当采用 DSA 算法时,NDVI、公交线路存在、土地利用混合度仍对结局产生显著影响。Vrijheid 等人^[84]在 HELIX 项目中利用上述类似方法,对 1301 对母子的上述暴露因素进行了产前孕妇和产后儿童两个阶段的暴露测量,来评价其对于儿童肥胖的影响。研究发现,产后儿童的学校附近设施密度越少,住所周边人口密度越大,会对超重与肥胖的发生以及 BMI 的增加造成显著促进作用。类似地,Agier 等人^[85]在 HELIX 项目中探究了产前 85 个暴露指标、产后 125 个暴露指标与儿童肺功能的关联性,其中肺功能采用 FEV₁ (最大深吸气后最大呼气第一秒呼出的气量容积)来表征。通过 EWAS 算法,研究结果表明,产前母亲住址离主干道路的距离越长、产后儿童学校周边的设施密度越高,与儿童 FEV₁ 的降低呈现显著关联性。Warembourg 等人^[86]在 HELIX 项目针对建成环境暴露对血压的影响开展了研究。研究发现,经过 DSA 算法筛选过后,建筑密度与收缩压呈现显著正相关,环境温度与收缩压、舒张压呈现显著负相关,而 PM_{2.5} 和 NO₂ 与舒张压呈现显著正相关关系。North 等人^[87]在 KABC 队列中则是采用包含建成环境因素在内的暴露组学研究方法,探究了产前孕妇、产后儿童暴露与儿童呼吸系统疾病的关系,然而在经过多元 Cox 比例风险模型分析之后,没有建成环境暴露因素呈现出显著关联关系。

上述研究主要集中在对新生儿队列中产前产后暴露对出生健康结局的影响,而对其他人群(如成年人、老年人)、其他健康结局(如本文关注的心血管疾病)的建成环境暴露组学研究目前仍然较为缺乏。Song 等人^[88]的一项中国横断面研究,探讨了建成环境暴露与 40 至 89 岁成年人血压的关系,并结合自然立方样条考察了非线性关联。研究发现,经过 DSA 算法筛选后,更高的 NDVI、更高的步行指数、更长的公交线路长度与收缩压降低呈现显著关联,而收缩压随着 PM_{2.5} 浓度的升高先上升后下降。关于舒张压,更高的人口密度和 NDVI、更长的公交路线长度、更远的到达主干道路的距离,以及更短的到达蓝色空间的距离,与舒张压的降低呈现显著关联性。

建成环境是一个由多种因素复合而成的复杂系统,所以很难通过先验知识来确定哪些建成环境暴露因素是显著的保护或风险因素。因此,暴露组学的方法为建成环境健康研究打开了崭新的领域。目前,建成环境领域的暴露组学研究还极

表 2.3 城市暴露组相关的流行病学研究

作者	研究年份	研究设计	研究人群	暴露因素 ^a	健康结局	统计分析方法
Nieuwenhuijsen 等 ^[83]	NA	队列研究	欧洲六国 31458 名孕妇 (HELIX 队列)	建成环境 (人口密度, 建筑密度, 道路交叉口密度, 公交站密度/路线长度, 设施密度, 设施多样性, 土地利用混合度, 步行指数); 空气污染 (PM _{2.5} , PM ₁₀ , NO ₂); 自然环境 (NDVI, 大型绿地/水面存在性, 到达绿地的距离, 最近绿地面积); 气象参数 (温度、湿度、紫外线); 交通 (总交通流量, 交通密度, 到主干道距离); 道路交通噪声	低出生体重	线性回归和 logistic 回归, 采用 EWAS 和 DSA 两种变量选择算法
Vrijheid 等 ^[84]	2013-2016	队列研究	欧洲六国 1301 个儿童 (HELIX 队列)	空气污染 (PM _{2.5} , PM ₁₀ , NO ₂); 紫外线; 自然环境 (NDVI, 绿色/蓝色空间存在性); 气象参数 (温度、湿度、气压); 建成环境 (人口密度, 建筑密度, 道路交叉口密度, 公交站密度/路线长度, 设施密度, 设施多样性, 土地利用混合度, 步行指数); 交通 (总交通流量, 交通密度, 到主干道距离); 道路交通噪声; 有机氯化物; 溴化合物; 全氟化烷基物质; 金属; 邻苯二甲酸酯代谢产物; 酚类化合物; 有机磷农药代谢产物; 水消毒副产物; 室内空气污染; 生活方式; 社会经济属性	超重/肥胖、BMI	线性回归和 logistic 回归, 采用 EWAS 和 DSA 两种变量选择算法
Agier 等 ^[85]	2003-2009	队列研究	欧洲六国 1033 对母子/母女 (HELIX 队列)	空气污染 (PM _{2.5} , PM ₁₀ , NO ₂); 紫外线; 自然环境 (NDVI, 绿色/蓝色空间存在性); 气象参数 (温度、湿度、气压); 建成环境 (人口密度, 建筑密度, 道路交叉口密度, 公交站密度/路线长度, 设施密度, 设施多样性, 土地利用混合度, 步行指数); 交通 (总交通流量, 交通密度, 到主干道距离); 道路交通噪声; 有机氯化物;	肺功能, 用 FEV ₁ 来表征	线性回归, 采用 EWAS 和 DSA 两种变量选择算法

续表 2.3 城市暴露组相关的流行病学研究

作者	研究年份	研究设计	研究人群	暴露因素 ^a	健康结局	统计分析方法
Agier 等 ^[85]	续上表			溴化合物；全氟化烷基化物质；金属；邻苯二甲酸酯代谢产物；酚类化合物；有机磷农药代谢产物；水消毒副产物；室内空气污染；生活方式；社会经济属性		
North 等 ^[87]	NA	队列研究	加拿大 560 名新生儿（KABC 队列）	住宅位置；住址周边主干道存在性；住址周边停车场存在性；社会经济属性；房屋类型；宠物；室内燃烧燃料和炊具类型；空气净化器；室内潮湿；吸烟；房屋年龄；地毯；车库类型等	呼吸系统疾病	Cox 比例风险模型，采用逐步回归算法
Song 等 ^[88]	2018-2019	横断面研究	中国 2002 名 40 至 89 岁的成人	空气污染（PM _{2.5} , NO ₂ , SO ₂ , CO, O ₃ ）；自然环境（NDVI, 到蓝色空间的距离）；建成环境（人口密度, 建筑密度, 道路交叉口密度, 公交站密度, 设施密度, 设施多样性, 土地利用混合度, 步行指数）；交通（到主干道距离）	血压（舒张压、收缩压、高血压）	线性回归和 logistic 回归，采用 EWAS 和 DSA 两种变量选择算法
Warembourg 等 ^[86]	1999-2010	队列研究	欧洲六国 1301 个儿童（HELIX 队列）	空气污染（PM _{2.5} , PM ₁₀ , NO ₂ ）；自然环境（NDVI, 大型绿色/蓝色空间存在性, 到达最近绿色/蓝色空间的距离, 最近绿色/蓝色空间的面积）；气象参数（温度、紫外线）；建成环境（人口密度, 建筑密度, 道路交叉口密度, 公交站密度/路线长度, 设施密度, 设施多样性, 土地利用混合度, 步行指数）；交通（总交通流量, 交通密度, 到主干道距离）；道路交通噪声；社会经济属性	血压（舒张压、收缩压）	线性回归，采用 EWAS 和 DSA 两种变量选择算法

^a 由于暴露因素数量大，此处仅详细列出本研究关注的城市建成环境及相关自然环境的暴露因素，其余类别的暴露因素仅展示大类信息。

不充分，一是对于测度哪些建成环境指标、如何测度没有统一的标准，二是研究的健康结局及人群类型还极为有限，这些都是后续研究需要深化解决的问题。

2.4 健康视角下的规划应对研究

现代城市规划与公共健康领域经历了“共生——分野——交汇”的历史进程。20 世纪初期，城市工业化带来的卫生问题催生了城市规划这一物质空间干预手段。30 年代，公共卫生走向更侧重个人风险因素的理论与实践，而城市规划走向更侧重外部空间因素的理论与实践。60 年代起，人们对慢性疾病的重视使得城市规划学科重新关注公共健康问题，使得健康城市理念被正式提出^[89, 90]。在此背景下，许多学者开始探讨城市规划与公共健康的作用关系，并基于此开展规划应对研究。

2.4.1 城市规划与公共健康的相互作用关系

现有研究中城市规划与公共健康相互作用关系的探讨，主要从健康传导机制和城市规划要素两个角度来考虑。

公共健康领域的学者从健康传导机制出发来理解两者的关系。Giles-Corti 等人^[91]认为，城市政策与规划干预手段跟健康之间的传导机制是较为间接的多环节过程，如图 2.1 所示。城市规划领域涉及的诸多建成环境因素首先影响了居民的交通出行方式、社会文化需求及个人健康行为（教育、工作、饮食、社交等），从而影响到人对空气污染、噪声、社交隔离、体力活动、饮食的暴露程度，再影响到一些中间的健康结局（肥胖、高血压等）或环境变化（气候变化、交通事故等），最终造成一系列慢性疾病的发生。Nieuwenhuijsen^[92]总结了城市与交通规划对心血管健康的传导路径（如图 2.2），其路径并非是一个完全的串联过程，而是一个多环节互相耦合影响的特征。城市设计要素一方面会影响人的室外出行意愿及交通出行方式，另一方面会影响到自然环境暴露水平（如空气污染、噪声、紫外线）、人的压力水平和社交活跃度，以及人的体力活动及肥胖水平。此外，交通出行方式也会作为中介因素影响到上述因素，最终这些因素共同造成心血管疾病的发生。

城市规划领域的学者则从城市规划要素出发对两者关系进行解读。王兰等^[93]将建成环境要素划分为土地使用、空间形态、道路交通、绿地和开放空间四大类，从而总结出健康城市规划的两条路径——减少空气污染对人体的影响和促进体育锻炼。李经纬等^[94]同样认为，建成环境对健康的影响机制可归纳为自然生态环境的污染问题，以及健康行为活动的促进（包括体力活动和社会交往活动）。田莉等^[95]总结了规划对健康发生作用的三方面机制，包括土地利用方面（土地利用的构成、混合度、开发强度对环境质量与个体行为的影响）、道路交通方面（交通

污染、交通事故，及步行可达性对体力活动的影响）和公共设施布局方面（卫生服务的空间布局与服务质量对健康的影响）。

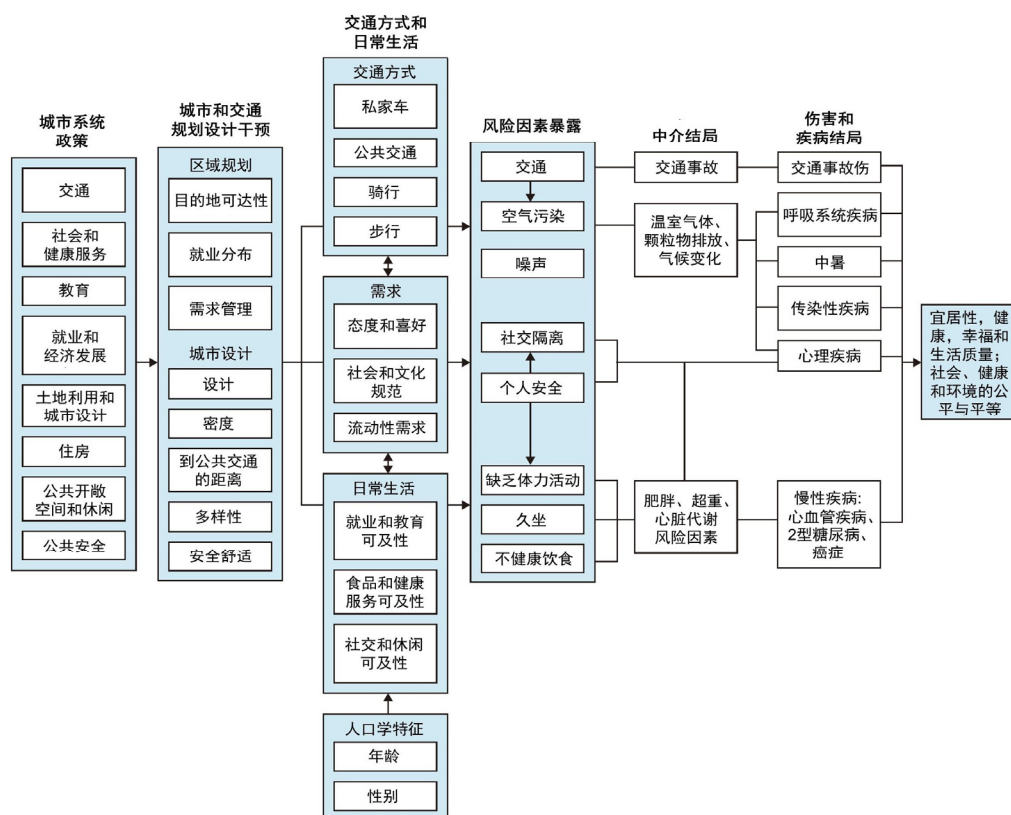


图 2.1 建成环境暴露的健康传导机制^[91]

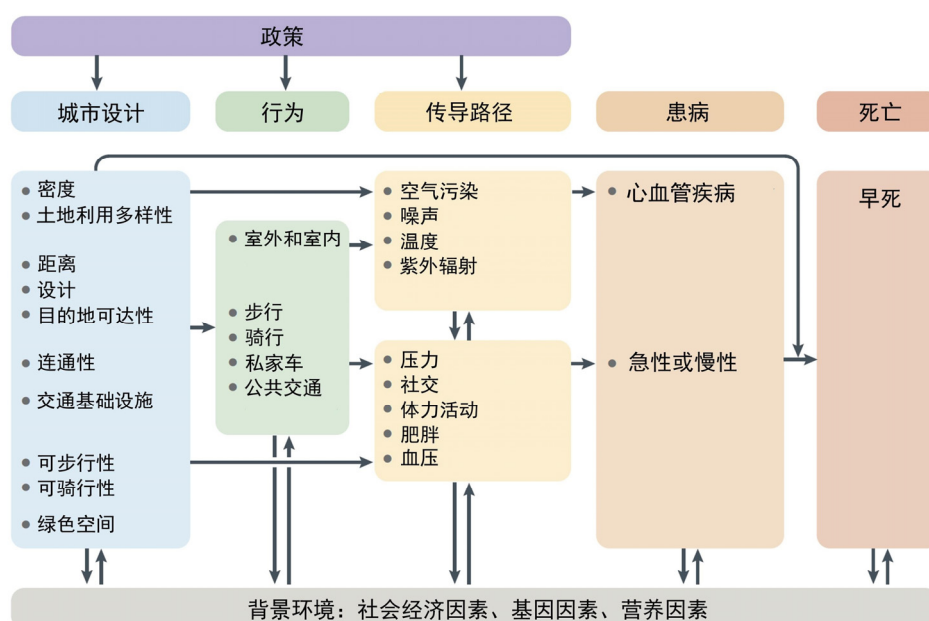


图 2.2 建成环境暴露对心血管疾病的影响机制^[92]

2.4.2 规划应对策略与技术

在健康促进机制的基础上,许多学者相继提出了相应的规划应对策略。王兰等^[93,96]梳理并归纳了土地使用、空间形态、道路交通、绿地和开放空间四个规划要素层面如何通过减少污染、促进锻炼的机制来增进健康。文中指出,土地利用中限制具有污染风险的用地可以减少人群对污染的暴露水平,而适当的土地利用强度与混合度可以使交通需求内部化,降低长距离交通带来的空气污染,促进居民选择慢行交通,增强体力锻炼。城区的肌理形态和街道空间几何形态可以影响污染物的扩散速率及人群选择慢行交通出行方式的舒适度。道路交通方面,应推广步行和骑行的方式,重点关注慢行基础设施的建设,以及避免机动车道与慢行系统的重叠。绿地和开放空间需要注意避免零星分散布局、促进集中化网络化布局来降低污染暴露,并且提高绿地的可达性以促进居民锻炼活动。王一^[97]提出,规划应从用地、公共设施、交通与能源四大要素出发来促进健康,包括以城市更新促进用地混合来提升社区就业与经济水平、注重学校的选址与联合共享以发挥社会效用、以交通可达性为纽带引导健康饮食及体育锻炼的健康生活方式、重视城市能源规划促进资源再利用。唐燕等^[98]在上述方面的建成环境要素之外,还着重介绍了《塑造邻里》一书中对于食物规划的内容,借助食物地图、食物雷达、食物里程等分析方法,提升食物商店步行可达性、增加当地食物生产提供、增加社区活动与食物的联系等,并融入了“社交园艺和园艺疗法”的新理念,促进居民身心健康。除此之外,还有部分学者致力于完善更加上游的健康城市理论,包括田莉等提出的区域/城市/社区层面的城乡健康理论模型^[95]、张天尧提出的结合弹性理论与社会生态学理论的健康城市规划框架^[99]等等。

在有了相应的规划应对策略基础上,还有一批规划应对技术已经开始影响或应用于实际的规划过程当中,包括健康城市指标的讨论与健康影响评估(Health Impact Assessment, HIA)。Webster 等人^[100]介绍了世界卫生组织欧洲健康城市网络项目中的健康指标,初期有 53 个,后来精简到 32 个,涵盖了健康、健康服务、环境指标、社会经济指标四个维度,其中建成环境涉及的绿地、工业用地、运动休闲设施、步行化、自行车路线、公共交通可达性及服务范围等都属于第三大类的环境指标。该项目试图通过城市之间的比较来认知各城市目前的健康水平,但是文中也指出,目前许多指标缺少标准化的国际定义,造成可比性的欠缺。除了健康指标之外,更具有影响力的是健康影响评估^[101]。HIA 由公共健康领域专家牵头,而规划师近年来才进入该行列。HIA 旨在评估规划、政策和项目促进健康的程度,从个体行为、社会经济属性、服务可达性、建成环境、自然环境多个方面来考量。在美国,HIA 已经有了较多的尝试,包括健康发展测量工具(Healthy

Development Measurement Tool, HDMT)、绿色社区认证(LEED-ND)及“为健康而设计”项目(Design for Health, DFH)。以“为健康而设计”项目为例,它是一套HIA分析的工具包,包括初步检查清单、快速评估工作坊、阈值分析、规划评审清单。其中的阈值分析部分,根据公共健康已有证据,对各个健康影响方面(如设施可达性、空气质量、房屋质量、体力活动、水质、食品环境等)给出阈值或定性判别标准,从而给出分项分数。HIA将为健康城市规划提供有力的支持和依据^[101]。

2.5 文献评述与研究方向

目前,建成环境暴露与心血管疾病之间关系的流行病学研究数量正在逐渐增加,但仍然存在以下几点局限。第一,因果推断能力较强的队列研究较少,而因果推断力较差的横断面研究较多,尤其是对于绿地、空气污染以外的其他建成环境因素,这可能与建成环境暴露关注较晚、队列研究成本较高有关。第二,大部分文献只关注建成环境暴露中的某个或某类因素,且集中在绿地、空气污染、食品环境、交通密度等几个指标。建成环境暴露对人体健康的影响机制是较为间接的,许多因素之间存在着潜在的相关性,虽然已有研究已经校正了许多个人的社会经济属性和健康数据,但是还有许多未考虑的建成环境要素可能在其中作为混杂因子未被发现,从而使研究结果偏离真实。另外,除了绿地、食品环境等因素之外,还有大量的建成环境要素值得被探究它们与心血管健康的联系。第三,健康结局大多关注到心血管疾病这样的大类疾病,少部分细化到冠心病、中风等详细的心血管疾病类别,但较少文献将本文关注的急性心肌梗死作为健康结局。并且,现有研究大多关注的是心血管类疾病死亡的风险因素,对于该类疾病的预后工作及复发风险仍是处于忽视的状态,这些都值得进一步的探究。

正是由于诸多环境暴露因素存在互相影响的原因,近年来暴露组学的新兴研究方法被提出,将原本假设驱动的环境流行病学研究向数据驱动方向推动。暴露组学方法对于建成环境这样的多因素复杂系统是一个非常有力的分析工具,但相关研究才刚刚起步。目前,建成环境暴露组学研究主要集中在产前产后暴露对新生儿健康结局的几个队列研究当中,对于其他人群、其他健康结局的该方面研究都极为有限。

城市规划学科与公共健康学科从诞生之初开始就存在着千丝万缕的联系,如今健康城市议题的兴起,为两个学科的再次交融提供了很好的契机。从城市规划学科角度来看,上述环境流行病学研究的探索,最终还需要落位到规划应对上。许多学者在城市规划如何影响公共健康的路径方面已经初步达成共识,但是,规

划与公共健康的联动刚刚起步，许多健康城市的探讨仍缺乏定量评估的科学研究支撑，多项健康要素之间的相对权重也常常基于经验而非科学的健康风险评估。

基于上述文献综述总结的诸多局限，本研究将基于急性心肌梗死患者的队列研究设计，聚焦于城市当中多种建成环境暴露对成年人的急性心肌梗死复发结局的影响，并采用暴露组学的分析方法进行关联分析。本研究还将基于上述关联分析得到的暴露反应关系，对北京市乡镇街道及居住小区进行急性心肌梗死复发的健康风险评估，并依据此定量评估结果提供相应的规划应对策略。

第3章 研究对象与数据基础

3.1 研究范围与研究对象

3.1.1 研究范围

本研究选取北京市全域作为研究范围，北京市下辖 16 个区，总面积为 16,410 km²。选择北京市的原因有二。第一，北京市既承担首都功能又是我国的特大城市之一，使得其成为我国具有代表性的城市之一。北京市的规模之大使得不同区域的城市发展状况有较大差异，从而也造就了居民截然不同的建成环境暴露水平。北京市较高的人口密度，使得建成环境改善带来的人群潜在健康收益将比较可观。第二，北京市的建成环境数据与心血管健康数据的可获得性较高，具有规模大、精度高的优势，便于后续建成环境与急性心肌梗死复发之间的个体水平的关联分析。此外，北京市五环内的范围还常被认为是北京的中心城区，该区域涵盖东城区、西城区，以及部分朝阳区、海淀区、丰台区、石景山区和大兴区，总面积为 667 km²。研究范围如图 3.1 所示。

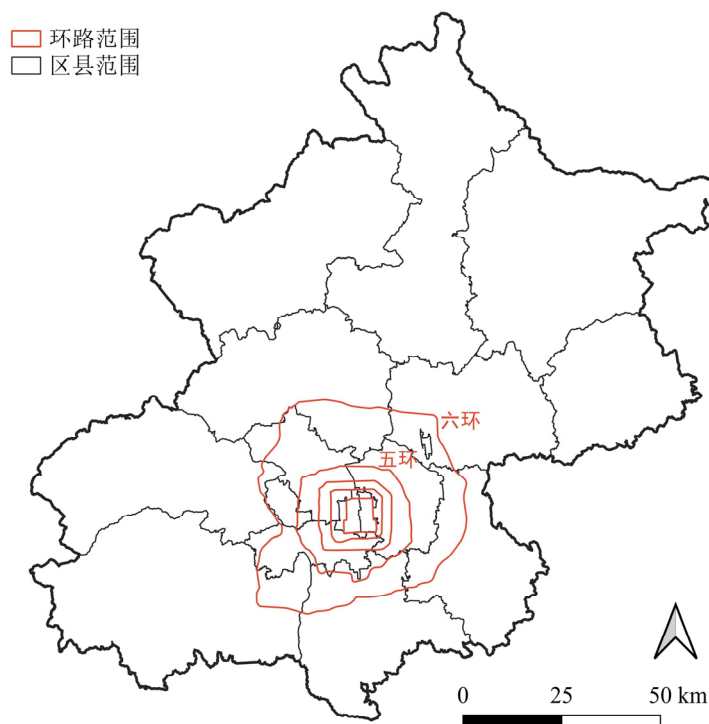


图 3.1 北京市全域及五环内区域

3.1.2 研究设计与研究对象

本研究采用队列研究设计，研究对象为北京市急性心肌梗死患者。急性心肌梗死患者数据来源于北京市心血管疾病监测系统，该系统通过患者的身份证号码，将北京医院出院信息记录与北京市人口登记监测系统进行匹配^[102]。系统内在 2013 年 1 月 1 日至 2019 年 12 月 31 日发生首次急性心肌梗死事件（第 10 版国际疾病编码中 I21 和 I22）并存活的北京市常住居民，被纳入本研究的患者队列。其中，在同一天出院并再次入院的患者被认为是同一次连续的治疗过程（ $N=4,870$ ）。此外，急性心肌梗死原因入院 1 天内即存活出院，且当天再无后续入院记录的患者被排除出该队列（ $N=963$ ），因为这些人群真正发生急性心肌梗死事件的概率极低。最终确定队列人数为 110,940 人。后续分析中，部分急性心肌梗死患者由于无法获得精确的患者家庭住址（ $N=13,393$ ）、无法获得患者住址的建成环境数据和社会经济属性（ $N=7,658$ ）、患者性别或婚姻状况未知（ $N=1,380$ ），被排除出队列之外。最终，研究对象确定为 88,509 位北京市急性心肌梗死患者，其中五环内有 43,374 人。

在该患者队列当中，随访从患者首次急性心肌梗死存活出院开始，到再一次急性心肌梗死复发结束，若在 2019 年 12 月 31 日之前未发生复发事件，也结束随访过程。随访过程中，首要关注的健康结局为急性心肌梗死的首次复发事件，次要健康结局为致死、非致死急性心肌梗死的首次复发事件，并根据每个个体的随访开始、结束时间，计算其复发间隔时长。本研究仅考虑患者急性心肌梗死的首次复发事件，而不考虑患者的多次复发情况。

3.1.3 研究人群的随访结果与基线特征

对 88,509 位急性心肌梗死患者的随访，随访人年数共计 261,296 人年，其中共计发生 12,216 件急性心肌梗死复发事件，比例为 13.8%。若进一步关注致死、非致死的急性心肌梗死复发事件，其随访人年数分别为 286,080 人年和 270,224 人年，复发事件频次分别为 3,762 次（占 4.3%）和 9,275 次（占 10.5%）。其中，有 821 位患者在随访期间先经历了非致死型复发，后经历了致死型复发事件。

从北京市心血管疾病监测系统中可以进一步提取出研究人群的基线水平，包括年龄、性别、婚姻状况以及患者的相关并发症情况。根据已有的急性心肌梗死发病或预后的流行病学研究，本研究选择了高血压、高血脂、糖尿病、中风、心力衰竭、心房颤动和肾功能障碍 7 类并发症^[103-110]。此外，为了反映患者的社会经济属性，本研究从已发表的数据库中选取了北京市交通分析小区尺度的平均教育年限、年均家庭收入这两个指标^[111]。个体水平的社会经济属性，由个体住址所在

交通分析小区的社会经济属性数值来确定。

北京市全域及五环内的队列人群的上述基线特征水平如表 3.1 所示。整个队列人群的年龄中位数为 66 岁，其中，70.5%为男性，91.1%为已婚人群，并发症中高血压、高血脂的发生概率较高，分别占到 68.0%和 66.8%。北京市全域队列和五环内子队列的基线特征除了社会经济属性之外，水平大致相似。社会经济属性层面，五环内由于地处城区，平均教育年限和年均家庭收入都显著高于北京市整体水平（两样本 Wilcoxon 检验 p 值均小于 0.001），从表 3.1 中可知，五环内人群该两项指标的 25%分位数，与北京市所有队列人群的中位数相近。

表 3.1 研究人群的基线特征水平

	北京市全域队列人群	北京市五环内队列人群
人数合计	88509 (100%)	43374 (49.0%)
年龄（岁）	66 (56 ~ 77)	67 (57 ~ 78)
性别		
男性	62391 (70.5%)	30821 (71.1%)
女性	26118 (29.5%)	12553 (28.9%)
婚姻状况		
目前未婚	7875 (8.9%)	3832 (8.8%)
目前已婚	80634 (91.1%)	39542 (91.2%)
平均教育年限（年）	12.2 (11.2 ~ 13.0)	12.7 (12.1 ~ 13.2)
年均家庭收入（万元）	9.9 (8.4 ~ 11.4)	11.2 (9.9 ~ 12.1)
并发症		
高血压	60145 (68.0%)	29440 (67.9%)
高血脂	59148 (66.8%)	25497 (58.8%)
糖尿病	35094 (39.7%)	17757 (40.9%)
中风	21478 (24.3%)	10530 (24.3%)
心力衰竭	34079 (38.5%)	17777 (41.0 %)
心房颤动	8614 (9.7%)	4605 (10.6%)
肾功能障碍	11690 (13.2%)	6363 (14.7%)

注：分类数据采用数量（百分比）的方式呈现，连续数据采用中位数（25%分位数 ~ 75%分位数）的方式呈现；五环内人群人数合计的百分比是五环内队列人数占北京市全域队列人数的比例，而其余百分比均为该亚组人数占五环内队列总人数的比例。

3.2 建成环境数据类型与来源

为了研究建成环境暴露与急性心肌梗死复发之间的关系，除了需要上一节的患者健康数据之外，还需要北京市建成环境数据的支持。本研究涉及的建成环境数据，包括基础城市空间数据、城市兴趣点（POI）数据、城市用地及建筑物数据、

街景图像数据、高分遥感影像数据等，下面逐一对这些数据类型及来源进行说明。

3.2.1 基础城市空间数据

基础城市空间数据主要包括：（1）北京市最新的区县和乡镇街道范围的面要素矢量数据，共计 16 个区县和 330 个乡镇街道；（2）北京市多年的道路网的线要素矢量数据，2014 年、2016 年、2017 年和 2019 年路网总长度分别为 36,805 km、37,536 km、38,048 km 和 41,891 km；（3）北京市六环内居住小区的面要素矢量数据，共计 13,083 个居住小区。上述数据如图 3.2 所示，可作为本研究空间落位和统计分析的基本单元，也可作为建成环境暴露水平及健康风险定评估的可视化基本对象。

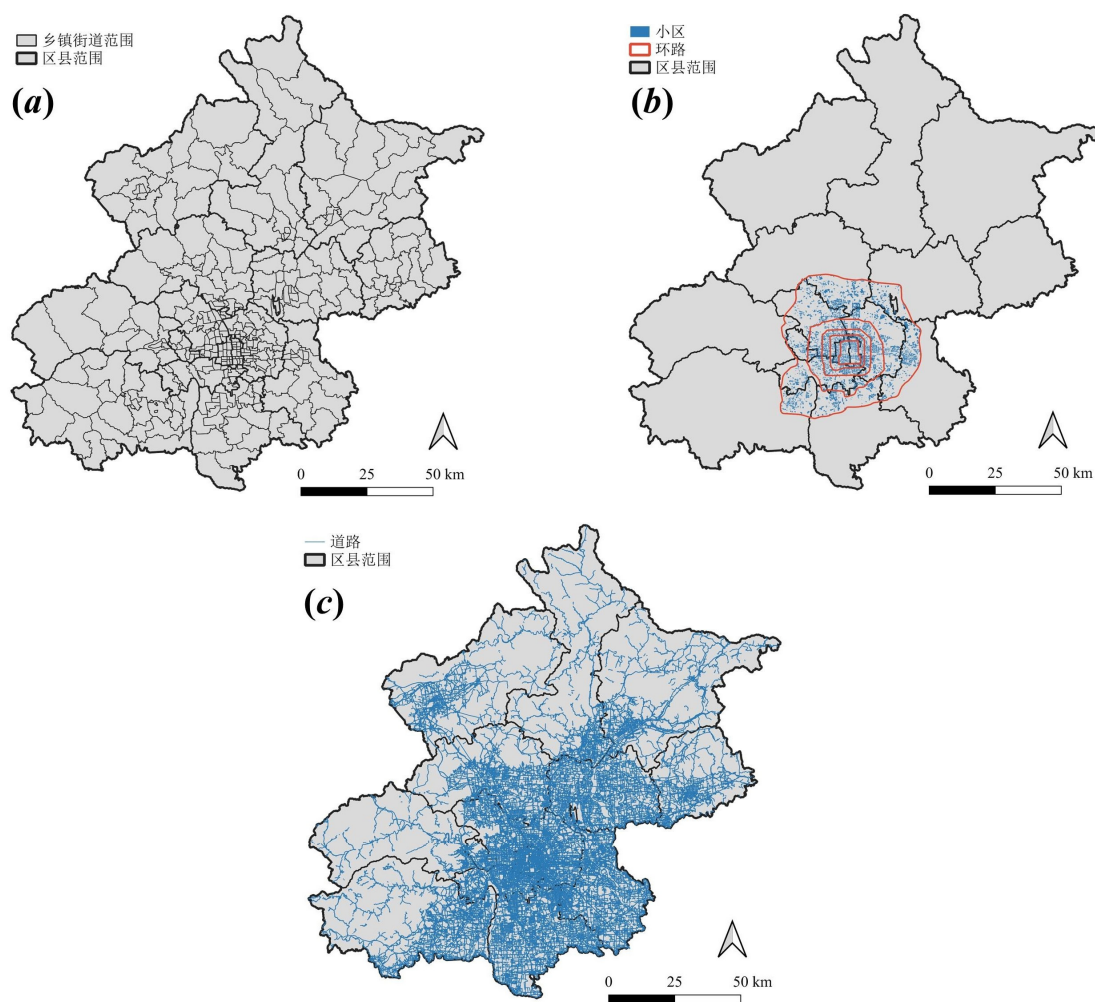


图 3.2 北京市基础城市空间数据

(a) 北京市区县及乡镇街道数据 (b) 北京市六环内小区数据 (c) 北京市道路数据

3.2.2 城市兴趣点（POI）数据

本研究从高德地图 API 抓取了多年的北京市 POI 数据，其数据类型为点要素的矢量数据，2011 年、2014 年、2015 年、2016 年、2017 年、2018 年和 2019 年的 POI 总数分别为 18.6 万个、28.1 万个、79.2 万个、137.0 万个、161.1 万个、176.3 万个和 184.8 万个。高德地图将所有 POI 按照功能属性分成 20 余个大类，若干个中类及小类，可供后续不同功能的 POI 空间分析。其中，大类包括汽车服务、汽车销售、汽车维修、摩托车服务、餐饮服务、购物服务、生活服务、体育休闲服务、医疗保健服务、住宿服务、风景名胜、商务住宅、政府机构及社会团体、科教文化服务、交通设施服务、金融保险服务、公司企业、道路附属设施等。由于数据较多，图 3.3 仅展示了北京市 2019 年的所有 POI 数据分布情况，及公园这一小类的 POI 分布情况。

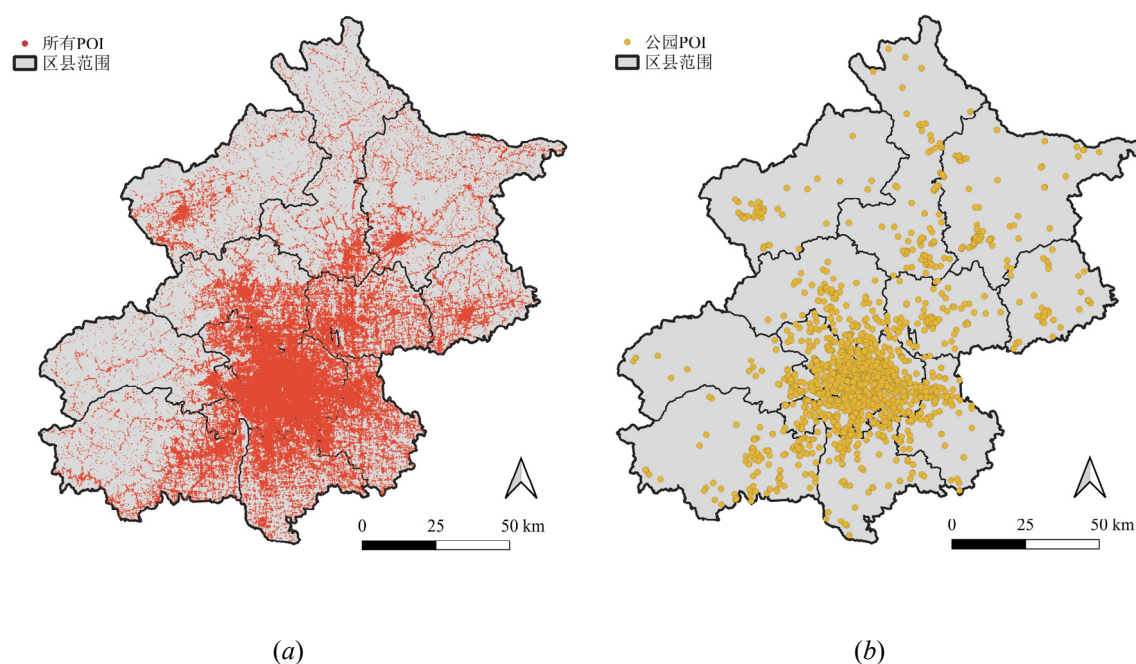


图 3.3 北京市 POI 数据

(a) 北京市 2019 年所有 POI 数据 (b) 北京市 2019 年公园 POI 数据

3.2.3 城市用地及建筑物数据

本研究使用的城市用地数据由北京市城市规划设计研究院提供，为 2014 年北京市全域的调研数据。城市用地类型根据北京市地方标准《城乡规划用地分类标准》（DB 11/996-2013）^[112]，主要分为公共管理与公共服务设施用地（A）、商业服务业设施用地（B）、军事外事安保用地（D）、多功能用地（F）、绿地与广

场用地（G）、工业用地（M）、保护区用地（P）、居住用地（R）、城市交通设施用地（S）、区域交通设施用地（T）、公用设施用地（U）、物流仓储用地（W）、村庄建设用地（C）、非建设用地（E）等。为了后续分析方便，根据国家标准《城市用地分类与规划建设用地标准》（GB 50137-2011）^[113]，将所有用地分为主要 8 个大类，用地类型合并依据如表 3.2 所示，合并后的用地分布情况如图 3.4 所示。

表 3.2 城市用地类型分类依据

用地分类	用地编码	北京市全域该类用地面积 (km ²)
公共管理与服务	A, D2	190.3
商业服务	B, C3 (旧版 E62)	513.4
交通设施	S, T	109.5
绿地和广场	G, E3 ~ E5	9503.7
公用设施	U	45.5
居住	R, C1 (旧版 E61)	1049.6
工业仓储	W, M	317.8
其他	剩余用地	4251.2

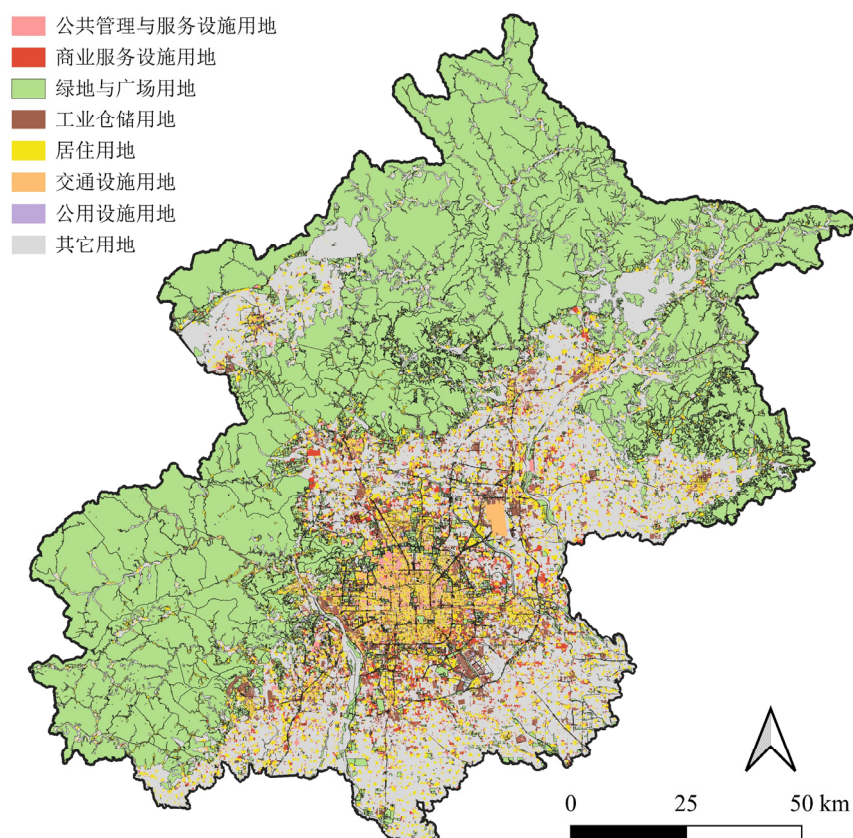


图 3.4 北京市用地数据

北京市建筑物数据来自于百度地图，为面要素矢量数据，包含建筑物占地面积、楼板面积（或层数）等信息，可获得年份为 2014 年和 2018 年，两个年份的建筑物总数分别为 31.6 万个和 61.8 万个。数据主要分布在五环以内。图 3.5 展示了五环范围内的建筑物矢量数据分布情况。

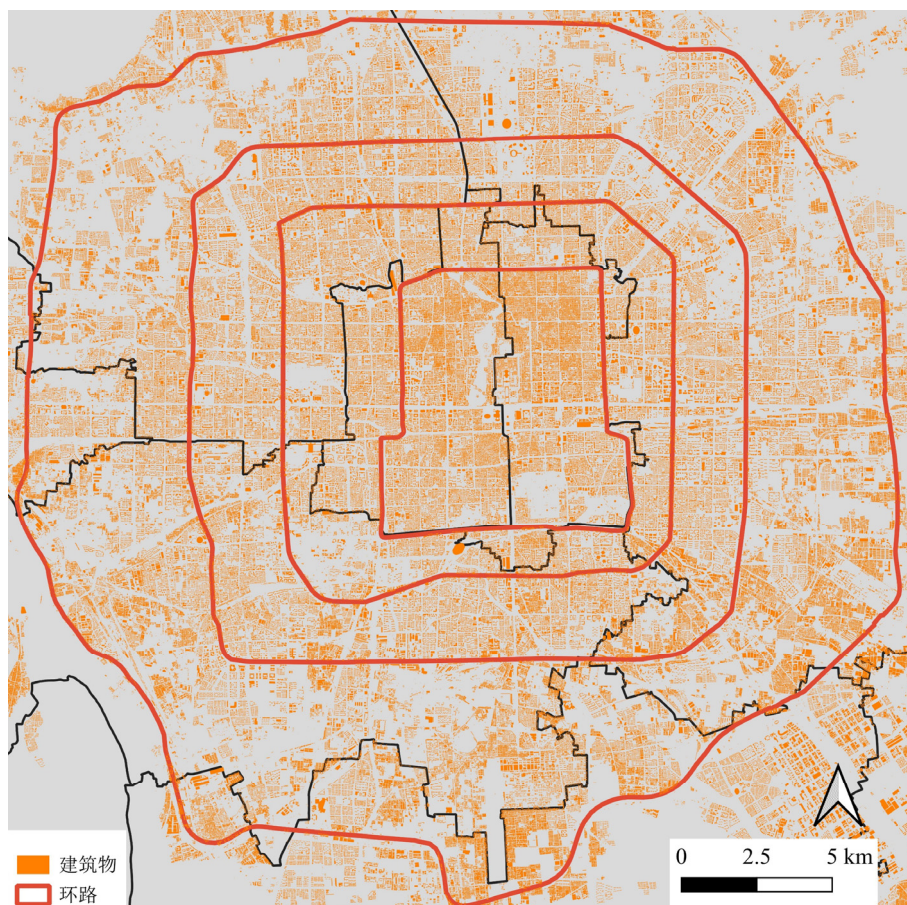


图 3.5 北京市五环内 2014 年建筑物数据

3.2.4 街景图像数据

本研究所采用的街景图像数据来自于百度地图街景，并利用百度地图时光机功能获得多年历史数据，涵盖 2013 年、2015 年、2016 年、2017 年和 2019 年，数据范围为北京市五环范围内。街景图像数据依据 Chen 等人^[14]发表研究中的数据采集方式进行，现简述如下：

（1）以 50 米为间隔距离沿街道进行等间距采样，得到各个街景采样点的空间位置，并在 ArcGIS 软件中确定这些采样点的经纬度坐标，获得的 2013 年、2015 年、2016 年、2017 年和 2019 年的街景点分别有 5.7 万个、4.3 万个、3.4 万个、4.2 万个和 4.6 万个；

(2) 针对每个街景点, 在百度地图街景上进行四个方向(平行于道路的前、后方向以及垂直于道路的左、右方向)的街景图像抓取, 每个方向的图像边缘与相邻图片重合, 视线俯仰角度为 0° , 焦距为 85 mm (如图 3.6 所示)。

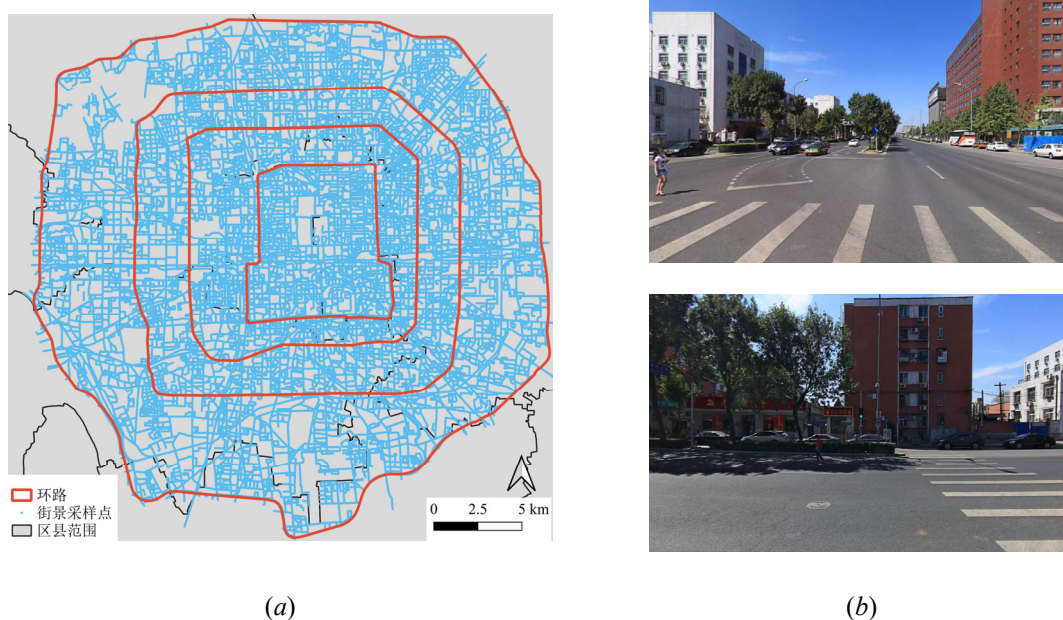


图 3.6 北京市五环内街景图像数据
(a) 2019 年街景采样点空间分布 (b) 街景图片

3.2.5 高分遥感影像数据

本研究采用的高分遥感影像数据, 下载于美国地质调查局 (United States Geological Survey, USGS) 官网上的 Landsat8 Level1 的 30 米分辨率遥感影像。为了保证多年比较的公平性和准确性, 以及满足 NDVI 指标测度需求, 本研究选取拍摄于夏季 (6 月至 8 月) 的云量小于 5% 的遥感影像, 最终符合要求的有 2013 年、2014 年、2015 年、2017 年和 2019 年五年的高分遥感影像数据。

3.2.6 其他建成环境数据

除了上述主要的建成环境数据之外, 本研究还使用了 WorldPop 提供的人口数量的 100 米栅格数据 (根据联合国人口估计调整后的无约束版本), 时间覆盖 2013 年至 2019 年间的每一年, 该数据为人口密度指标的测度提供数据基础^[115]。此外, 本研究从已发表文献中获取了大气 $\text{PM}_{2.5}$ 浓度的 1 km 栅格数据, 该数据由遥感数据和地表监测站数据建模得到, 数据时间覆盖 2013 年至 2018 年间的每一年, 可为空气污染的暴露水平提供数据基础^[13]。

第4章 建成环境暴露评价

4.1 患者住址的空间化

急性心肌梗死患者住址的精确经纬度坐标，是后续个体水平的建成环境暴露评价的重要基础。因此，本节先介绍患者住址空间化的基本方法与解析结果。需要说明的是，本节住址空间化的对象为北京市 2013 年至 2019 年所有急性冠心病的住院及死亡人群（ $N=239,653$ ，事件数共计 269,810 条），本研究后续分析的急性心肌梗死患者群体（ $N=110,940$ ）仅为该人群的子集。

4.1.1 患者住址空间化方法

1. 基本流程

本研究需要将北京市心血管疾病监测系统内的患者家庭住址文本信息，转化为患者住址的经纬度坐标信息，从而完成患者住址的空间落位，这个过程被称为地址解析，或者地理编码（geocoding）。对于北京市急性冠心病住院及死亡事件，总体上遵循先软件自动地址解析、再人工核查剩余地址信息的基本原则。整体流程如图 4.1 所示^[116]。

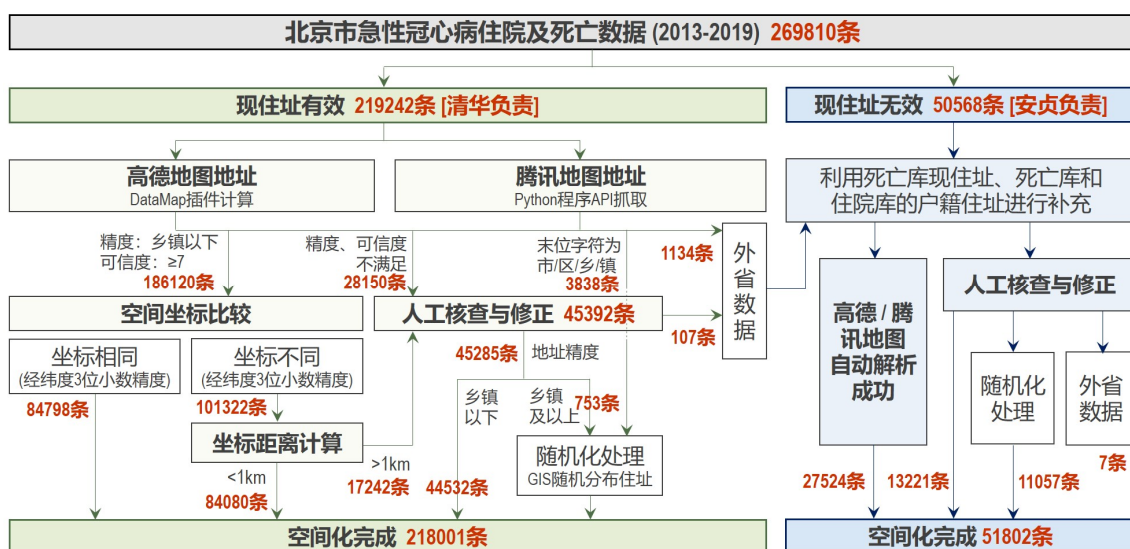


图 4.1 患者住址空间化的基本流程

本研究先对现住址有效的数据开展地址解析。在软件自动地址解析的部分，考虑高德地图、腾讯地图分别进行独立地址解析，以控制地址经纬度坐标的精准度。具体而言，可通过 Excel 中的 DataMap 插件，对住址文本信息进行高德地图

的经纬度解析，并通过 Python 代码，实现腾讯地图的地址解析。软件自动解析结果可能存在以下两种可能出现的问题：

(1) 其中某一个在线地图不能精准解析患者住址经纬度。本研究给出以下此类问题的识别准则：高德地图参考 DataMap 插件给出的精度信息，若识别精度达到门牌号、兴趣点、村庄、道路、道路交叉口的，认为地址解析成功，若识别精度为省、市、区县、乡镇、开发区、热点商圈的，认为地址解析失败；腾讯地图参照地址解析给出的可信度 (reliability) 与精度 (level) 信息，若可信度 ≥ 7 且精度 > 3 ，则认为地址解析成功，否则认为地址解析失败。

(2) 两个在线地图都能精准解析患者住址经纬度，但两者的解析结果差异过大。需要说明的是，只有不存在上述问题 (1) 的住址才能进入此问题的识别，否则将进入到后续的人工核查与修正环节。高德地图、腾讯地图都采用 GCJ-02 坐标系，因此不存在坐标转换问题，可直接对两个坐标结果进行比较。若两个坐标无论是经度还是纬度，均相差 < 0.001 度，则认为获得的经纬度有效，这部分患者住址被认为解析成功。若两个坐标的经度或是纬度相差 > 0.001 度，则需要进一步计算投影坐标系下的空间距离^[117]，若空间距离 < 1 km，则同样认为解析成功，否则将进入人工核查与修正环节。

对于软件自动解析失败、需要人工核查与修正的患者住址，具体方法由于较为复杂，将在第 2 部分详细介绍。人工核查与修正完成后，对于能够解析到比乡镇水平更细精度的坐标，认为解析成功，否则认为该地址文本较为模糊，无法获得该患者的住址精确经纬度，该患者不再纳入后续队列的分析。

基于上述基本流程，本研究再对现住址无效的患者，参考其死亡库中的现住址以及死亡库、住院库中的户籍住址，开展了地址解析工作。整个地址解析流程中各个步骤涉及的地址数目，已总结在图 4.1 中。

2. 人工核查与修正方法

人工核查与修正环节，针对的是前述无法通过软件自动解析得到精确经纬度的现住址。这些地址解析失败的原因主要包括：患者住院库内的现住址格式不规范、住址本身填写内容较为模糊以及电子信息录入失误。因此，需要综合考察住院库、死亡库内的现住址、户籍住址信息，来辅助判断患者住址坐标。基于此，人工核查与修正方法将分为两大步骤：地址信息来源选择、地址规范化及解析。

(1) 地址信息来源选择。住院库、死亡库两大患者数据库可以提供多种来源的患者住址数据，为了下文指代方便，将不同来源的地址信息予以不同编号，如表 4.1 所示。

表 4.1 不同来源的地址信息编号

地址信息来源	编号	地址信息来源	编号
住院库现住址详细地址	1A	住院库户籍详细地址	3A
地址 1A 医疗代码对应地区	1B	地址 3A 医疗代码对应地区	3B
死亡库生前常住详细地址	2A	死亡库户籍详细地址	4A
地址 2A 医疗代码对应地区	2B	地址 4A 医疗代码对应地区	4B

地址信息来源的选择总原则，是现住址优于户籍住址、住院库优于死亡库，即优先级从高到低排序为 1A、1B、2A、2B、3A、3B、4A、4B。具体而言：

①对于现住址（1A、1B、2A、2B）优于户籍住址（3A、3B、4A、4B）的原则，如果现住址的精度不够，则需要先判断现住址与户籍住址的一致性。如果现住址和户籍住址明显来自于同一地区，只不过户籍住址更为详细，那么可以选用户籍住址来补全住址信息（表 4.2 例 1）；如果户籍住址明显与现住址不是同一个地方，那么保留现住址较为模糊的地址信息，而不采用户籍住址（表 4.2 例 2）。

表 4.2 地址信息来源选择示例

	例 1	例 2	例 3	例 4
1A	北京市昌平区昌平	北京昌平小汤山	北京市昌平区沙河	北京市草厂村
1B	昌平区	昌平区	昌平区	通州区
2A			北京市昌平区沙河地区友谊家园 6-2-1	
2B			北京市昌平区沙河（地区）镇	
3A	北京市昌平区小汤山镇尚信村 201 号	北京市海淀区清河街道北京轮胎厂新西区 5 楼	昌平区沙河	潮县镇
3B	昌平区	海淀区	昌平区	通州区
4A			北京市海淀区西北旺镇永丰东玉河村 151 号	
4B				
地址选择	3A	1A	2A	1A+1B+3A
规范地址	北京市昌平区小汤山镇尚信村	北京市昌平区小汤山镇	北京市昌平区沙河地区友谊家园 6 号楼	北京市通州区潮县镇草厂村委会

②在现住址（1A、1B、2A、2B）内部，对于住院库（1A、1B）优于死亡库（2A、2B）的原则，指的是优先采用住院库数据。如果住院库的现住址数据不够详细，只能够精确到区县、乡镇级别，而死亡库的现住址数据更加精细，即使这两个地址不是指代同一处，也可以采用死亡库较为详细的现住址数据来补全地址信息（表 4.2 例 3）。

③在现住址的住院库数据（1A、1B）内部，如果 1A 不够精确，可以采用 1B 信息补全，因为 1A、1B 本身来自同一数据源。如果有极小概率遇到 1A、1B 指代不同地方，则哪个详细就用哪一个。其他同样数字编号的 A/B 数据也同理处理（表 4.2 例 4）。

（2）地址规范化及解析。为了准确获得住址经纬度，需要将上一步确认的患者住址进行规范化，规范标准格式为：北京市 XX 区 XX 镇/街道 XX 村/社区/小区/道路（XX 号楼/门牌号），括号里为非必须项，视提供的地址信息精度而定。然后将规范化的地名输入到高德地图坐标拾取系统

（<https://lbs.amap.com/tools/picker>），按关键字搜索，点击搜索后得到地址解析的经纬度坐标。将鼠标停留在地图中蓝色标签位置，会弹出解析地址的中文名称，需要审计人员确保是否与输入地址一致。核实无误后，将经纬度坐标存储到指定文件中。表 4.3 展示了部分地址规范化的示例。

表 4.3 地址规范化示例

序号	原地址信息	规范化后的地址信息
1	北京市 10 街坊 19 楼 23 号	北京市朝阳区酒仙桥街道十街坊 19 号楼
2	北京市峪口镇胡营村	北京市平谷区峪口镇胡家营村
3	北京市石景山区西井 2-3-47 号	北京市石景山区苹果园街道西井 2 区 3 号楼
4	北京市朝阳区垡头东里南院 8 排 5 号	北京市朝阳区垡头街道东里社区居民委员会
5	北京市通州区露朝家园	北京市通州区潞潮佳苑
6	北京市净土胡同 5 号	北京市东城区净土胡同 5 号院
7	北京市石景山模南 6-40	北京市石景山区金顶街街道模式口南里 6 号楼

由于人工核查与修正工作量较大，本环节共招募了 20 位志愿者共同完成。志愿者在进行人工核查工作之前，发放了工作指南（如图 4.2 所示）进行操作培训，培训过后通过测试者方可开展工作。工作期间，每位志愿者每 3-4 天提交 300 个地址审查结果，接受 5% 的抽样核查，审查结果有问题者会进行进一步一对一的沟通与指导，并被要求整改之前的审查结果。

4.1.2 患者住址空间化结果

按照上一小节的地址解析流程，本研究获得了北京市 2013 年至 2019 年急性冠心病患者的住址经纬度坐标，为了后续空间计算方便，统一转换为 WGS84 坐标。从中筛选出急性心肌梗死患者共计 110,940 人，其中，能够获得精确坐标信息（即比乡镇街道尺度更精细）的共计 97,547 人。将所有急性心肌梗死患者住址空间落位，结果如图 4.3 所示。总体而言，北京市城区的急性心肌梗死患者人数显著高于郊区，北京市东南部的急性心肌梗死患者人数显著高于西北部。其中，人数最多

的是丰台区新村街道，为 2,810 人，最少的是密云区檀营地区，仅为 81 人。



图 4.2 人工核查与修正工作指南部分页面

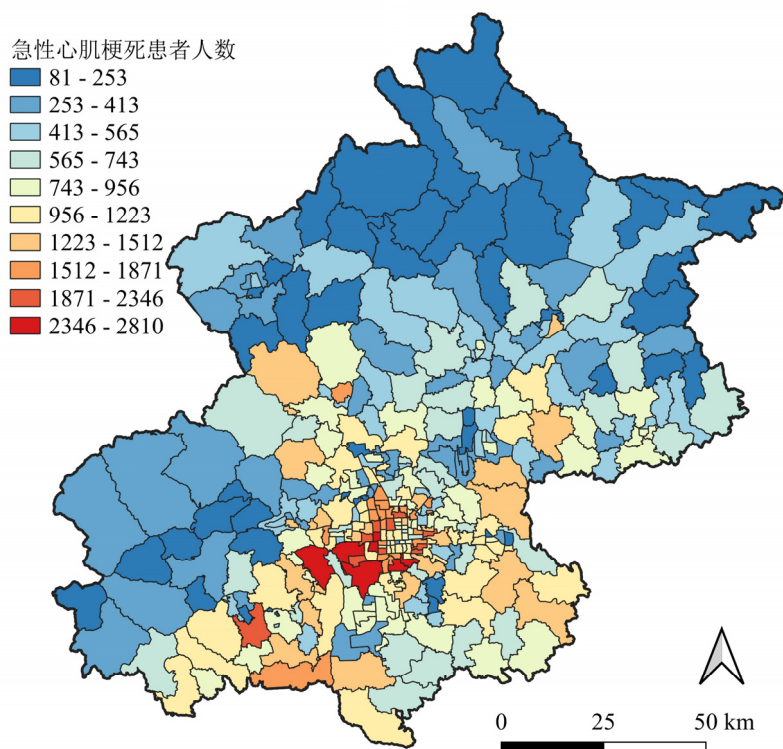


图 4.3 2013 年至 2019 年北京市急性心肌梗死患者住址空间分布

4.2 建成环境暴露因素的选择

第2章文献综述中已提及,本研究将采用新兴的暴露组学方法来研究建成环境暴露的健康效应问题,因此,需要尽可能多地选择潜在的建成环境暴露因素进行测度。根据2.1节中提及的建成环境5D理论、2.2节中提及的多种绿地暴露测度指标、其他与心血管疾病显著关联的建成环境要素,以及2.3节中暴露组学文章涉及的建成环境要素,结合空气污染因素,本研究提出了6大类共计26种建成环境暴露因素的清单(如表4.4所示),体现出暴露组的全面性特征。这些建成环境暴露因素中,建筑密度、建筑容积率、空间失序、步行指数、街道绿视率由于数据所限仅能在五环内范围开展测度,NDVI在北京市西南侧边界处有极少量缺失数据无法进行测度,其余建成环境暴露因素均能在北京市全域进行测度。

表 4.4 本研究选择的建成环境暴露因素

分类	暴露因素	数据空间范围
1.密度	(1) 人口密度	全北京市
	(2) 建筑密度	五环内
	(3) 建筑容积率	五环内
	(4) 城市功能密度	全北京市
2.多样性	(5) 土地利用混合度	全北京市
	(6) 城市功能混合度	全北京市
3.设计	(7) 路网密度	全北京市
	(8) 到主干道的距离	全北京市
	(9) 空间失序	五环内
	(10) 步行指数	五环内
	(11) 绿地比例	全北京市
	(12) 街道绿视率	五环内
	(13) NDVI	全北京市
4.到公共交通的距离	(14) 地铁站密度、到地铁站的距离	全北京市
	(15) 公交站密度	全北京市
5.目的地可达性	(16) 公园密度、到公园的距离	全北京市
	(17) 到大型绿地的距离	全北京市
	(18) 烟酒店密度	全北京市
	(19) 餐厅密度	全北京市
	(20) 快餐店密度	全北京市
	(21) 甜品/饮品/糕饼店密度	全北京市
	(22) 蔬果店密度	全北京市
	(23) 运动场馆密度	全北京市
	(24) 综合医院密度、到综合医院的距离	全北京市
	(25) 药房密度	全北京市
6.空气污染	(26) 大气 PM _{2.5} 浓度	全北京市

4.3 患者住址周边建成环境测度方法与结果

本节着重介绍患者住址周边上述建成环境暴露因素的测度方法及相应结果。首先，除了距离类指标之外，其余建成环境指标均需依托于某个空间单元来进行统计，本研究中采用急性心肌梗死患者住址的 300 米、500 米、1000 米缓冲区作为不同空间尺度的测度范围，分别对应居民的 5 分钟、10 分钟、15 分钟生活圈的概念。其次，为了反映暴露组在时间尺度上动态性的特征，各建成环境指标在 2013 年至 2019 年间每一个可获取数据的年份均开展了测度。上述多年的建成环境测度结果可为 4.4 节患者个体层面的建成环境暴露评价提供数据基础。

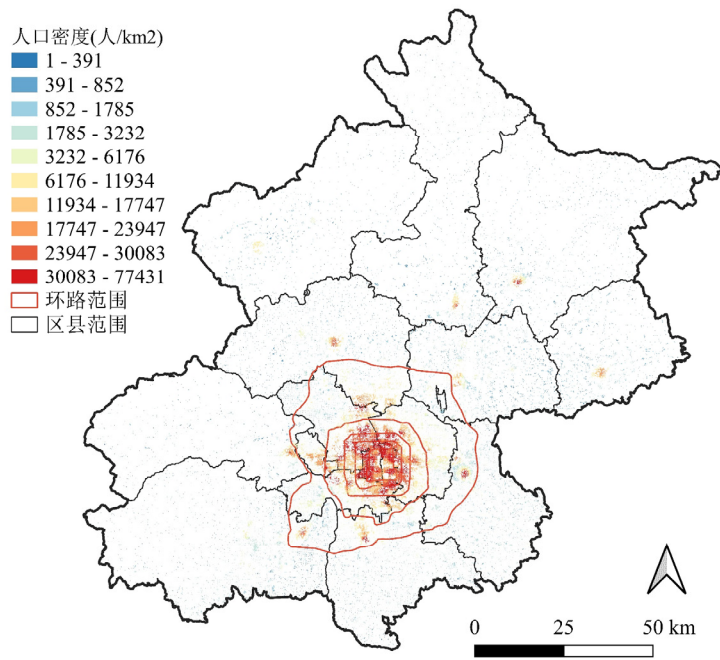
为了保证计算精度的前提下降低计算复杂度，北京市全域被划分为 $100\text{ m} \times 100\text{ m}$ 的网格，每个网格的中心点位置将代表这个网格的空间信息。将所有急性心肌梗死患者住址的空间位置连接到该网格上，可以得到全北京市共有 61,368 个网格存在急性心肌梗死患者。接下来的建成环境测度，便不再基于患者住址的精确经纬度，而是基于患者住址所在网格的中心点来开展测度，这是因为 100 米的精度在 4.1 节中地址解析环节当中是可以容忍的。所有建成环境测度均在 QGIS 软件（3.16.7 版本）内完成，NDVI 的测度还借助了 ENVI 软件（5.3 版本）。下面按照表 4.4 的六大类指标，逐一对各个建成环境因素的测度方式及结果进行详细说明。

4.3.1 第 1 类建成环境暴露因素：密度

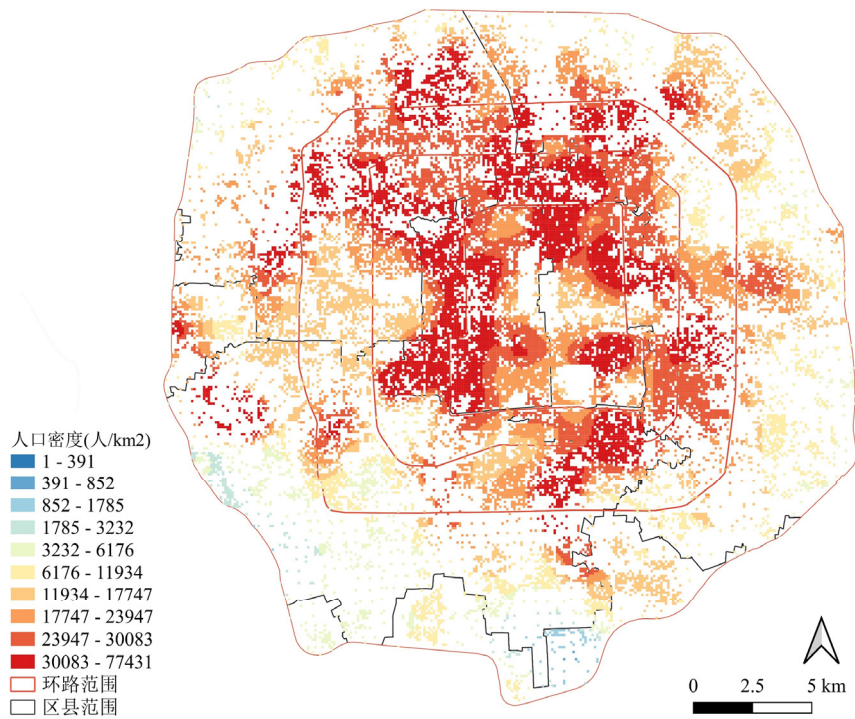
（1）人口密度

人口密度的测度方式为：首先，需要在患者住址（即住址所在 100 m 网格中心点，下文同，不再注释）周边建立 300 m、500 m、1000 m 缓冲区；然后，将 WorldPop^[115]提供的人口 100 m 栅格数据转化为点要素矢量数据，再将人口点要素与患者住址缓冲区进行空间连接，并统计缓冲区内的总人口数；最终，将缓冲区内总人口数除以缓冲区面积，即得到人口密度指标，单位为人/ km^2 。

该指标在 2013 年至 2019 年每一年均完成了测度，测度范围是北京市全域。图 4.4 展示了北京市全域及五环内患者住址 1000 m 缓冲区内 2019 年人口密度测度的可视化结果。结果表明，北京市五环内的人口密度远高于五环之外的郊区，郊区仅各区县的中心城区人口密度较高；而在五环范围内，市中心和北城的人口密度较其他区域更高。



(a)



(b)

图 4.4 北京市 AMI 患者住址 1000 m 缓冲区内 2019 年人口密度测度结果

(a) 北京市全域 (b) 五环内

（2）建筑密度与建筑容积率

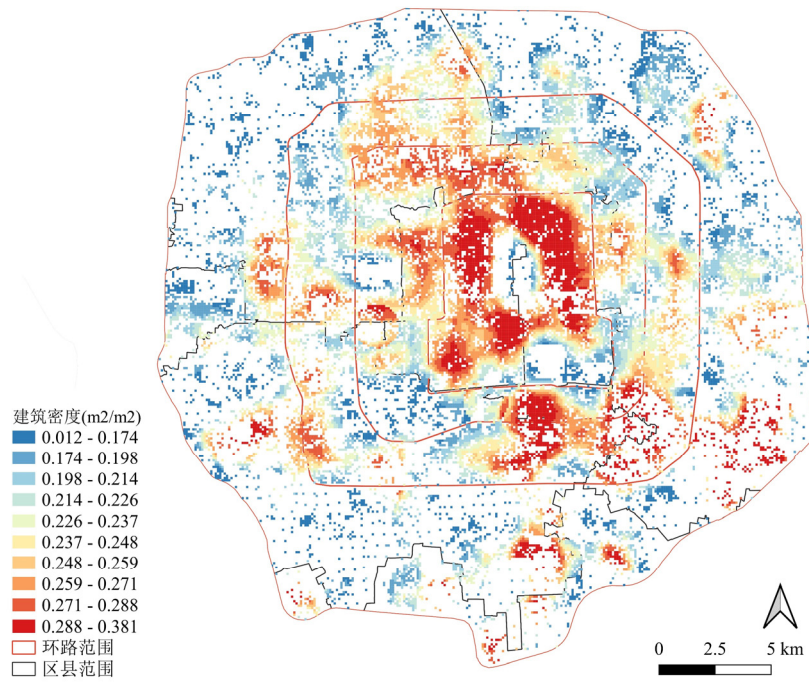
建筑密度与建筑容积率的测度基于的是五环内建筑物的多边形面要素矢量数据，它们分别是二维和三维的视角来度量城市中建筑物密度水平的。测度过程中，首先建立患者住址的 300 m、500 m、1000 m 缓冲区，然后计算各个多边形的面积来获得各建筑物的投影面积，再将建筑物面要素转换为点要素，把这些点要素空间连接到缓冲区内，统计缓冲区内建筑投影面积和楼板面积的总和。对于建筑密度指标，将缓冲区内建筑投影面积总和除以缓冲区面积即可获得，单位为 m^2/m^2 。而对于建筑容积率，需要用缓冲区内建筑楼板面积总和除以缓冲区面积来获得，单位为 m^2/m^2 。由于 2018 年建筑物数据未直接给出楼板面积，因此 2018 年的这两个指标测度时，假设建筑层高为 3 米，由建筑物高度推测出层数之后，利用建筑投影面积乘以层数来估算建筑楼板面积。

该两项指标在 2014、2018 年完成了测度，测度范围是北京市五环内区域。图 4.5 展示了北京市五环内患者住址 1000 m 缓冲区内 2018 年建筑密度、建筑容积率测度的可视化结果。结果表明，北京市五环内的建筑密度最高区域集中在二环旧城内，此区域是旧城的胡同四合院片区，但该片区建筑层高普遍较为低矮，因此建筑容积率最高的区域位于二环旧城以外，主要集中在东、西、北三个方向的二环与三环之间。

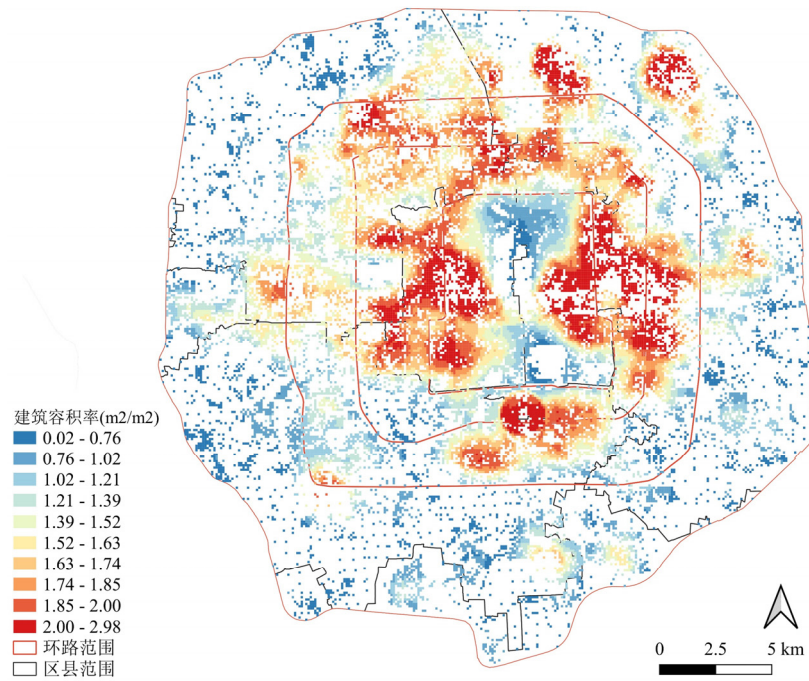
（3）城市功能密度

本研究中，城市功能密度由缓冲区内的 POI 密度指标来衡量。建立患者住址缓冲区后，将 POI 点要素与缓冲区进行空间连接，统计出缓冲区内的 POI 个数。由于在同一尺度缓冲区下，缓冲区面积相同，因此 POI 个数与 POI 密度是等价的。

该两项指标在 2011 年以及 2014 年至 2019 年的每一年完成了测度，测度范围是北京市全域，2013 年的结果采用 2011 年和 2014 年的测度结果进行线性插值得。图 4.6 展示了北京市患者住址 1000 m 缓冲区内 2019 年城市功能密度测度的可视化结果。结果表明，北京市四环以内的城市功能密度远高于四环以外的区域，且中心城区呈现东、北方向功能密度略高的分布情况。

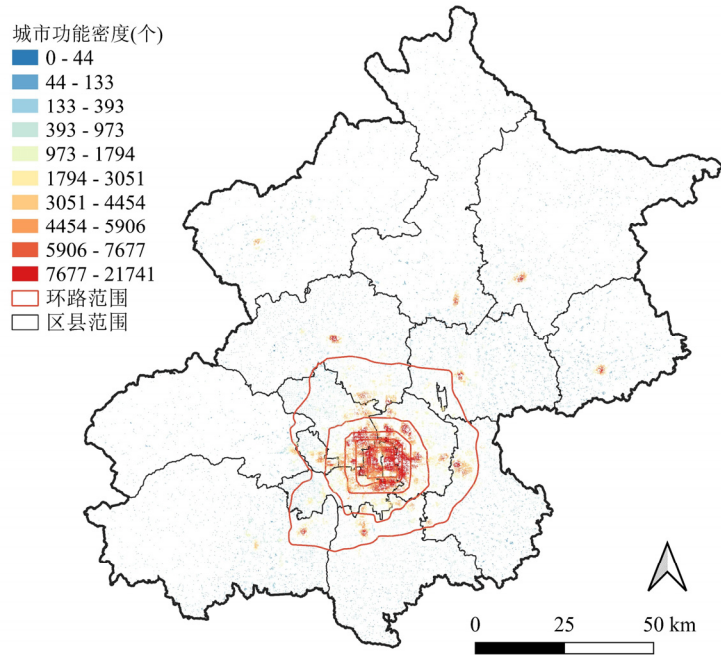


(a)

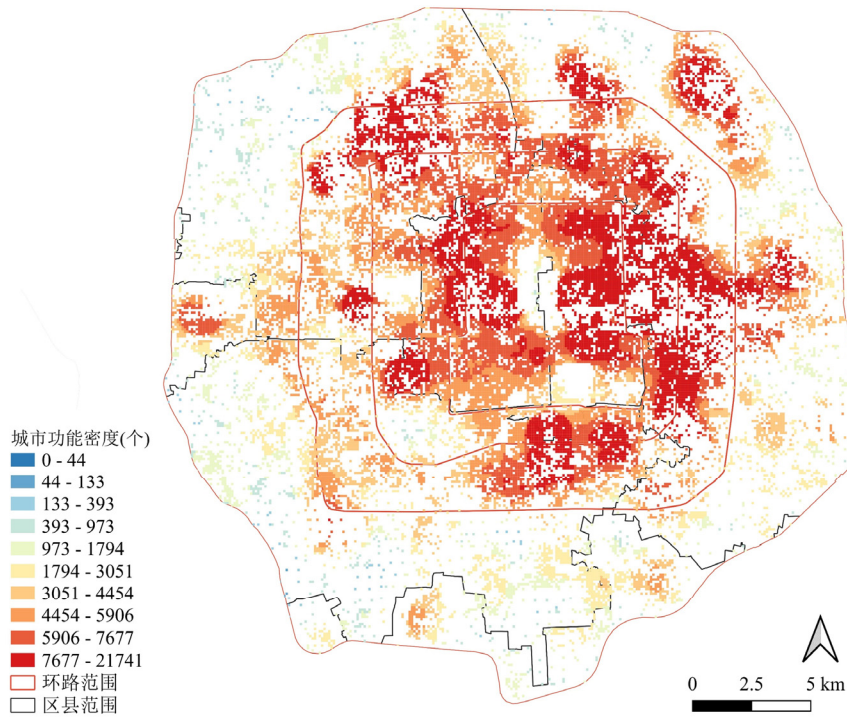


(b)

图 4.5 北京市五环内 AMI 患者住址 1000 m 缓冲区内 2018 年(a)建筑密度与(b)建筑容积率测度结果



(a)



(b)

图 4.6 北京市 AMI 患者住址 1000 m 缓冲区内 2019 年城市功能密度测度结果

(a) 北京市全域 (b) 五环内

4.3.2 第2类建成环境暴露因素：多样性

(1) 土地利用混合度

土地利用混合度是基于2014年的北京市用地数据开展测度的。首先，建立患者住址的300 m、500 m、1000 m缓冲区。然后，将按照3.2.3节合并为八类用地类型的用地数据分别存为八个用地矢量图层。将不同性质的用地分别与缓冲区进行相交运算，再按照缓冲区序号融合，计算出该缓冲区内各类用地的面积 F_i ($i = 1, 2, \dots, 8$)，并按字段连接到相应的住址缓冲区。接着，计算缓冲区内各类用地的归一化面积：

$$A_i = \frac{\text{缓冲区内该类用地面积 } F_i}{\text{全北京市该类用地总面积}} \quad (4.1)$$

然后，可计算缓冲区内各类用地的归一化占比：

$$p_i = A_i / \sum_{i=1}^8 A_i \quad (4.2)$$

最后，采用香农信息熵指标（Shannon's entropy index）来计算土地利用混合度，无量纲^[118]：

$$M = -\sum_{i=1}^8 p_i \ln p_i \quad (4.3)$$

该指标由于数据所限，仅在2014年完成了测度，测度范围是北京市全域。图4.7展示了北京市患者住址1000 m缓冲区内2014年土地利用混合度测度的可视化结果。结果表明，北京市二环内旧城及城北的土地利用性质较为单一，这可能与西城区、东城区的文化历史保护，以及海淀区为大片高校单位的用地规划相关。

(2) 城市功能混合度

城市功能混合度的测度方式类似于土地利用混合度，不同之处在于城市功能混合度基于的是POI数据，而非用地数据。本研究先将所有POI根据高德的分类信息合并为12大类，具体分类依据如表4.5所示。

将上述12类POI点要素分别空间连接到患者住址缓冲区，并统计缓冲区内各类POI的个数，进一步计算缓冲区内各类POI点的归一化个数：

$$N_i = \frac{\text{缓冲区内该类POI个数}}{\text{全北京市该类POI总个数}} \quad (4.4)$$

然后，可计算缓冲区内各类POI的归一化占比：

$$p_i = N_i / \sum_{i=1}^{12} N_i \quad (4.5)$$

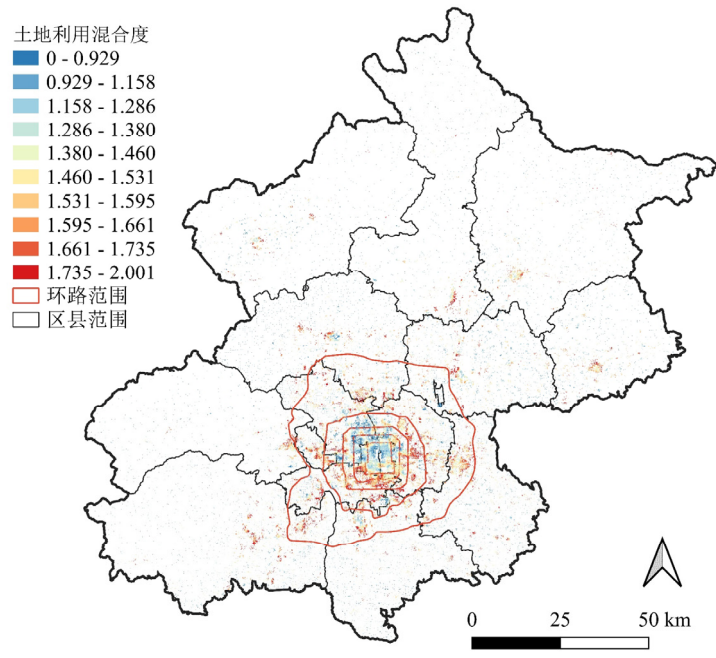
最后，同样采用香农信息熵指标来计算城市功能混合度，无量纲^[118]：

$$M = -\sum_{i=1}^{12} p_i \ln p_i \quad (4.6)$$

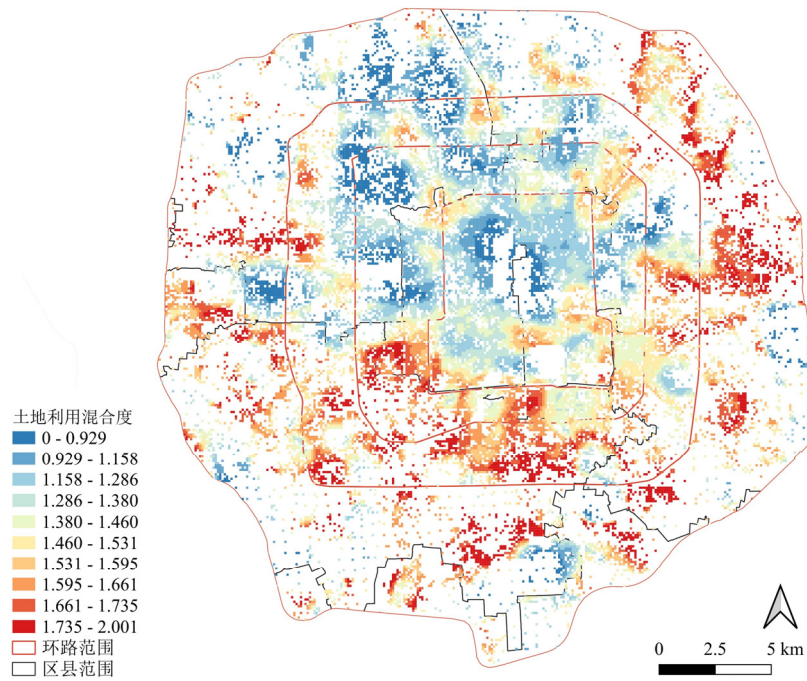
表 4.5 POI 分类依据

POI 分类	高德编码	POI 分类	高德编码
餐饮	05XXXX	办公	1201XX/1202XX/17XXXX
购物	06XXXX	住宅	1203XX
生活服务	07XXXX	政府机构	13XXXX
体育休闲	08XXXX	科教文化	14XXXX
医疗保健	09XXXX	交通设施	15XXXX/18XXXX
绿地公园	11XXXX	其他类型	其他编号

该指标在 2015 年至 2019 年每一年都完成了测度，测度范围是北京市全域，2013 年和 2014 年的结果采用 2015 年和 2016 年的测度结果进行线性外推获得。图 4.8 展示了北京市患者住址 1000 m 缓冲区内 2019 年城市功能混合度测度的可视化结果。结果表明，北京中心城区，尤其是四环内的空间范围，城市功能多样性显著高于外围郊区。特别是二环内及二环周边地区，虽然用地性质较为单一，但 POI 类型较丰富。



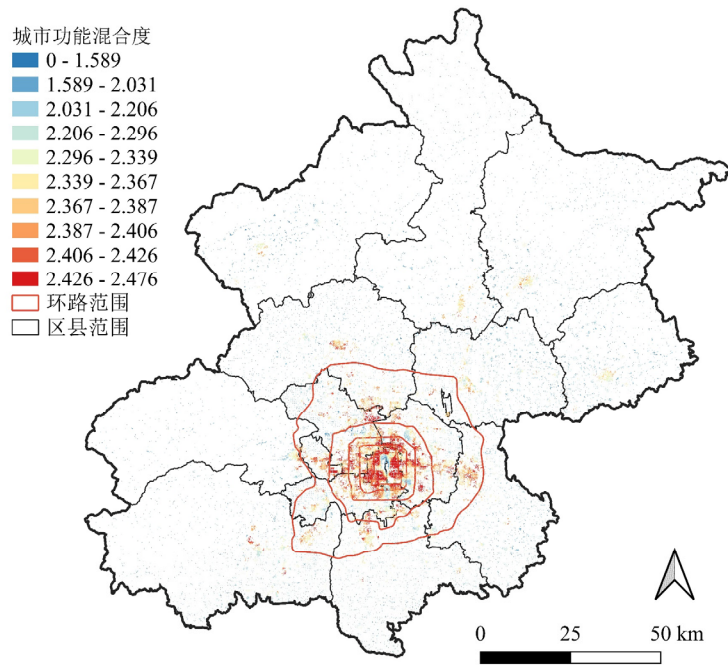
(a)



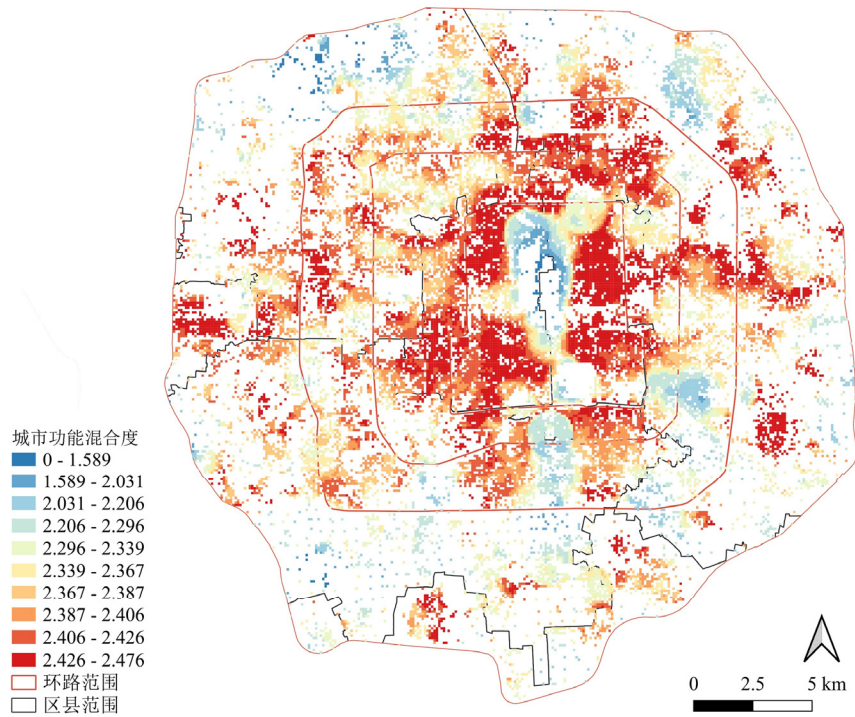
(b)

图 4.7 北京市 AMI 患者住址 1000 m 缓冲区内 2014 年土地利用混合度测度结果

(a) 北京市全域 (b) 五环内



(a)



(b)

图 4.8 北京市 AMI 患者住址 1000 m 缓冲区内 2019 年城市功能混合度测度结果

(a) 北京市全域 (b) 五环内

4.3.3 第3类建成环境暴露因素：设计

(1) 路网密度

路网密度由基础城市空间数据中的道路网数据测度得到。由于原有道路网数据存在多线道路的重复数据，因此先对道路数据进行简化：

(1) 对道路做 15 米缓冲区并融合，填充为黑色后，导出为高清 tiff 文件，颜色设置为 1 位单色阈值；

(2) 将 tiff 文件导入 ArcGIS，使用唯一值可视化地图，并新建折线图层存放生成的单线道路线要素；

(3) 利用 ArcGIS 中的 ArcScan 矢量化工具，矢量化选项中采用中心线，并将栅格捕捉选项的最大宽度设为 80 像素，点击生成要素后便可提取出道路中心线，从而简化为单线路网。

接着，将道路网分别与患者住址缓冲区进行相交运算，再按照缓冲区序号融合，计算每个缓冲区内的道路总长度，并按字段连接到相应缓冲区上。最终，将各缓冲区内的道路总长度除以缓冲区面积，就能得到路网密度，单位为 km/km^2 。

该指标在 2014 年、2016 年、2017 年和 2019 年完成了测度，测度范围是北京市全域，其余年份由线性插值或外推获得。图 4.9 展示了北京市患者住址 1000 m 缓冲区内 2019 年路网密度测度的可视化结果。结果表明，北京中心城区，尤其是四环内的空间范围，路网密度显著高于外围郊区。

(2) 到主干道的距离

由于多年路网数据中，只有 2016 年高德路网有较为清晰的道路等级分类，因此本指标仅测度 2016 年的结果。将道路等级分类中高速、国道、省道、县道、乡道、市区一级道路和快速路作为主干道，计算各患者住址到主干道的最近距离，单位为 m。该指标测度范围是北京市全域，图 4.10 展示了北京市患者住址 1000 m 缓冲区内 2016 年测度结果。

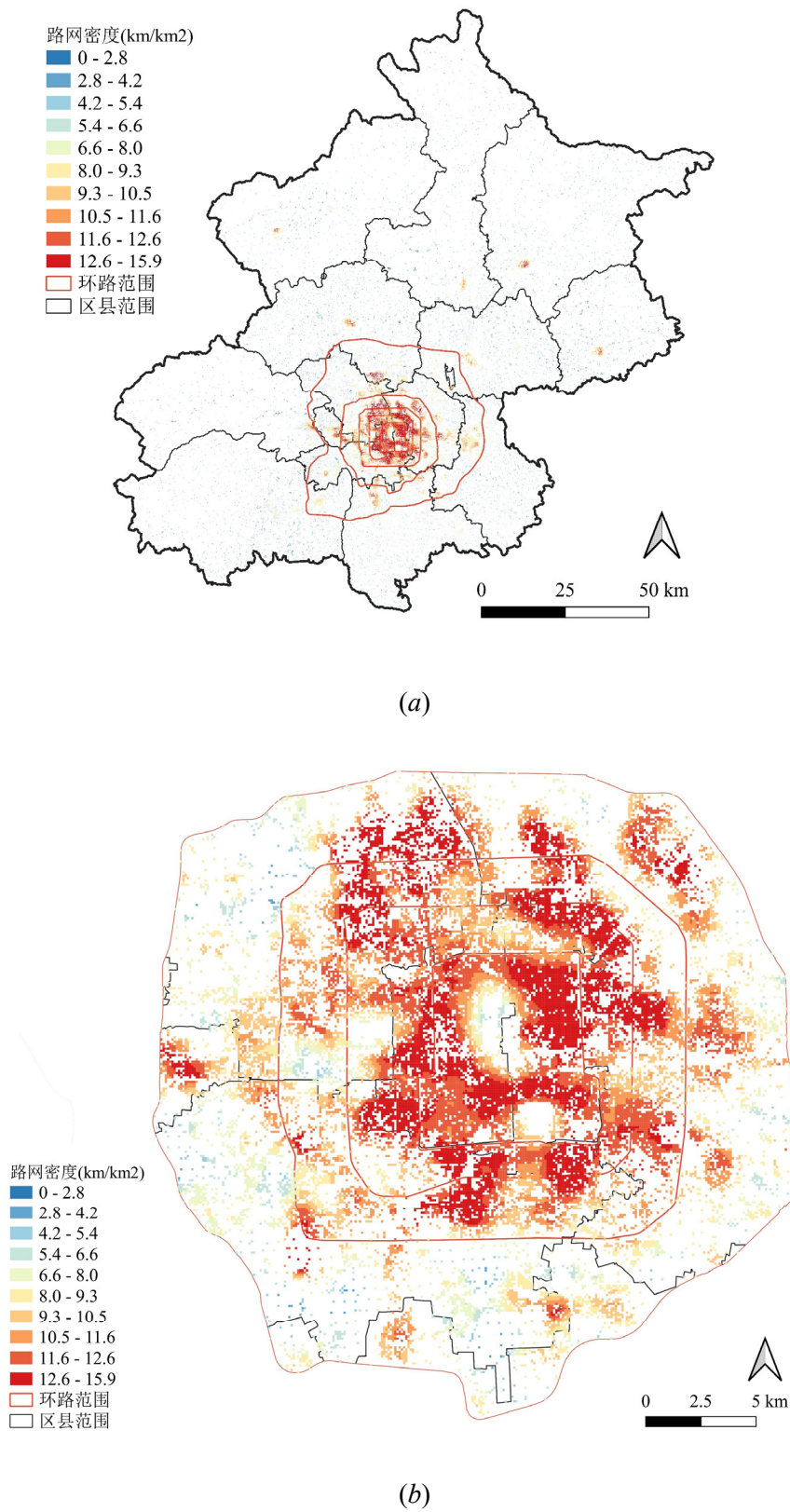
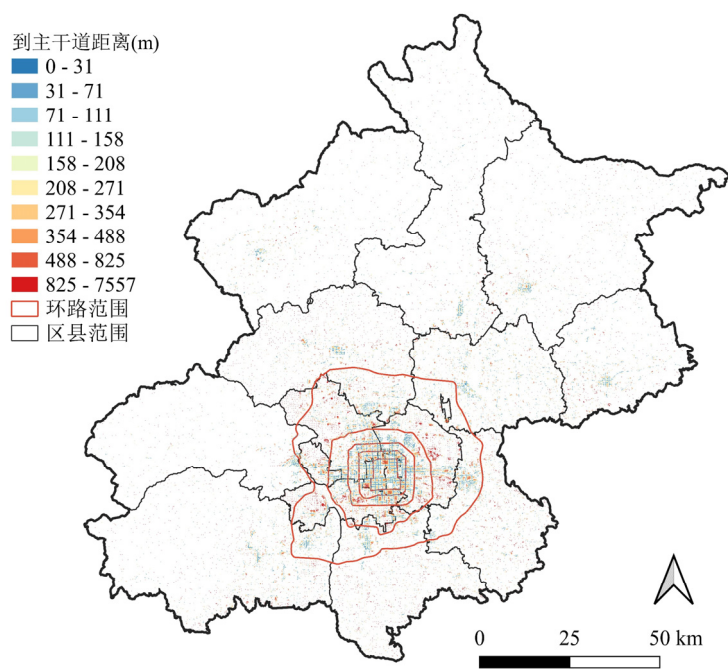
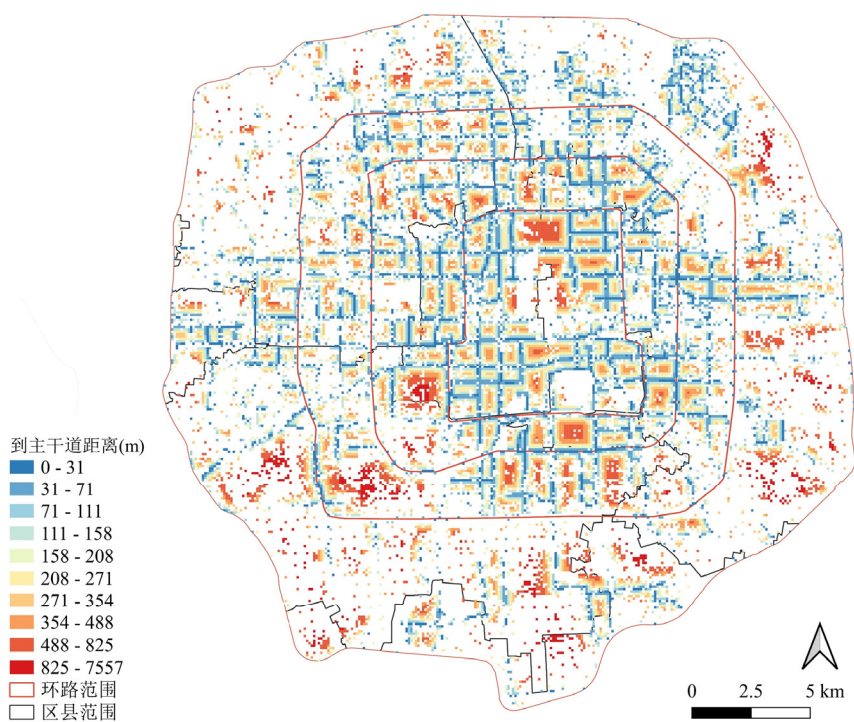


图 4.9 北京市 AMI 患者住址 1000 m 缓冲区内 2019 年路网密度测度结果
(a) 北京市全域 (b) 五环内



(a)



(b)

图 4.10 北京市 AMI 患者住址 1000 m 缓冲区内 2016 年到主干道距离测度结果

(a) 北京市全域 (b) 五环内

(3) 空间失序

空间失序是城市公共空间品质不足的物质环境表征，对城市公共安全、居民公共健康造成威胁。Chen 等人^[114]借鉴了西方建成环境空间失序已有研究，并结合中国城市自身特点，提出了五大类共计 15 项空间失序要素的审计清单，包括建筑类、沿街商业类、环境绿化类、道路类、其他公共设施类，具体审计清单如图 4.11 所示。Chen 等人^[114]结合人工虚拟审计以及图像识别的深度学习算法（包括 Faster R-CNN、SSD、SegNet 算法），实现了对街景图像的各个空间失序要素的识别。对于每张街景图像，存在某一空间失序要素即加 1 分，不存在该要素记为 0 分，从而得到该街景图像的空间失序总分。每个街景点的空间失序得分，为该街景点平行于道路（前、后）和垂直于道路（左、右）四个方向的街景图像空间失序得分的平均值。本研究使用了 Chen 等人对北京市五环内空间失序的各个街景点的测度结果，然后将各个街景点空间连接到对应的患者住址缓冲区，计算出各缓冲区内的街景点空间失序平均值。



图 4.11 空间失序要素审计清单^[114]

该指标在 2013 年、2015 年、2016 年、2017 年和 2019 年完成了测度，测度范围是北京市五环内区域，其余年份由线性插值或外推获得。图 4.12 展示了北京市患者住址 1000 m 缓冲区内 2019 年空间失序测度的可视化结果。可以发现，北京市五环内整个南城以及旧城的东北侧（鼓楼以东至交道口、国子监一带）的空间失序程度较高。

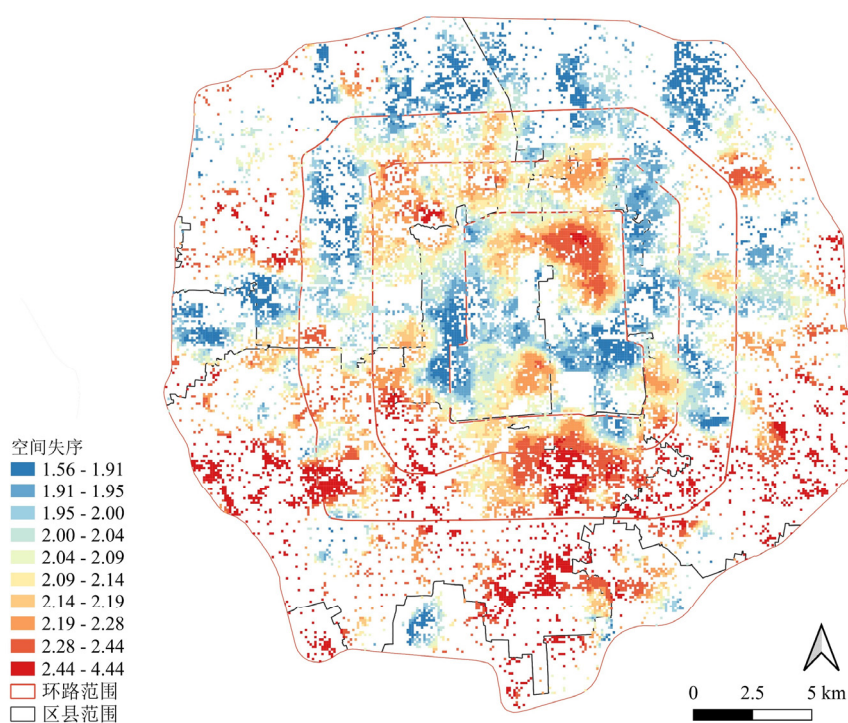


图 4.12 北京市五环内 AMI 患者住址 1000 m 缓冲区内 2019 年空间失序测度结果

(4) 步行指数

步行指数 (Walkability) 是建成环境的一个综合性指标, 本研究参考 Sallis 等人的研究^[119], 选取居住建筑物密度、道路交叉口密度以及土地利用混合度三个单项指标来评估步行指数。其中, 居住建筑物密度的计算方法为: 将几何中心落在居住用地内的建筑物视为居住建筑, 转为点要素后空间连接到缓冲区, 从而统计出居住建筑物的个数。道路交叉口密度, 采用计算路网密度时得到的单线简化路网, 获得道路交叉口的点要素, 再空间连接到相应缓冲区进行计数。土地利用混合度指标, 前文已经计算获得。按照 Sallis 等人^[119]的研究算法, 三个指标进行 z-score 归一化之后相加, 即可得到步行指数。

由于用地数据仅在 2014 年可获得, 因此步行指数指标也仅在 2014 年完成测度。由于建筑物数据是五环范围内的, 因此步行指数的测度范围是北京市五环内区域。图 4.13 展示了北京市患者住址 1000 m 缓冲区内 2014 年步行指数测度的可视化结果。结果表明, 北京市五环内, 二环内旧城的东北角、西南角, 以及南二环到南三环的区域, 步行指数较高。

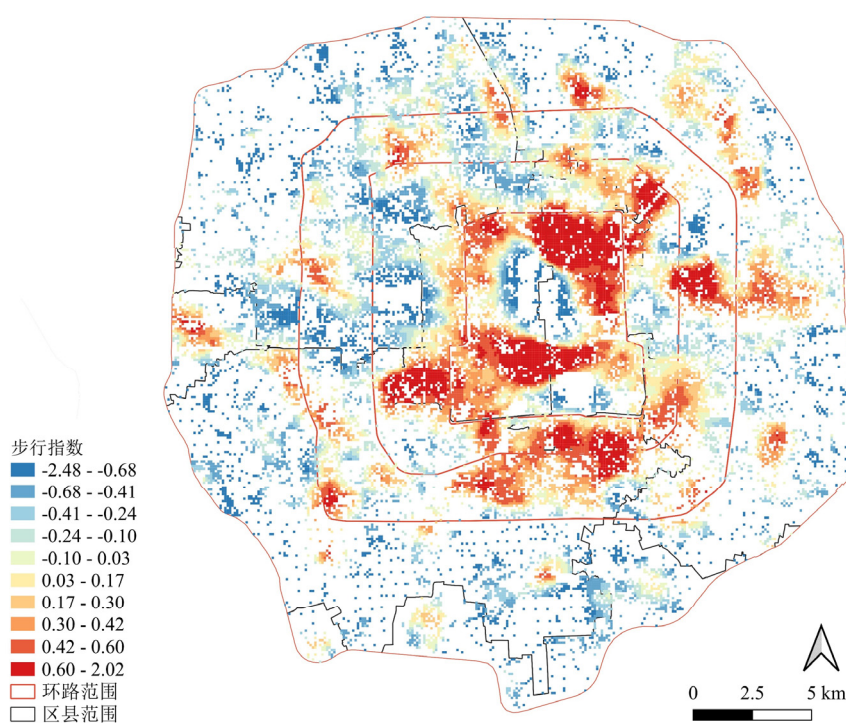


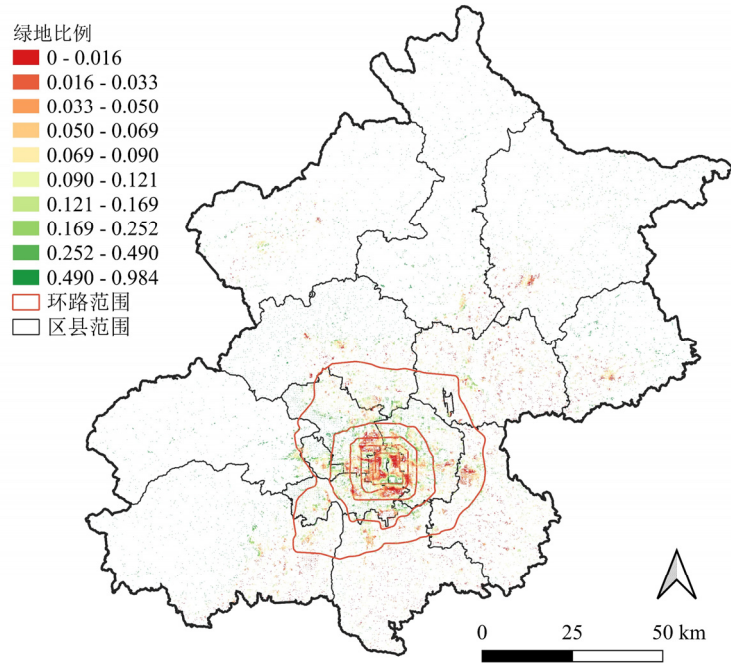
图 4.13 北京市五环内 AMI 患者住址 1000 m 缓冲区内 2014 年步行指数测度结果

(5) 绿地比例

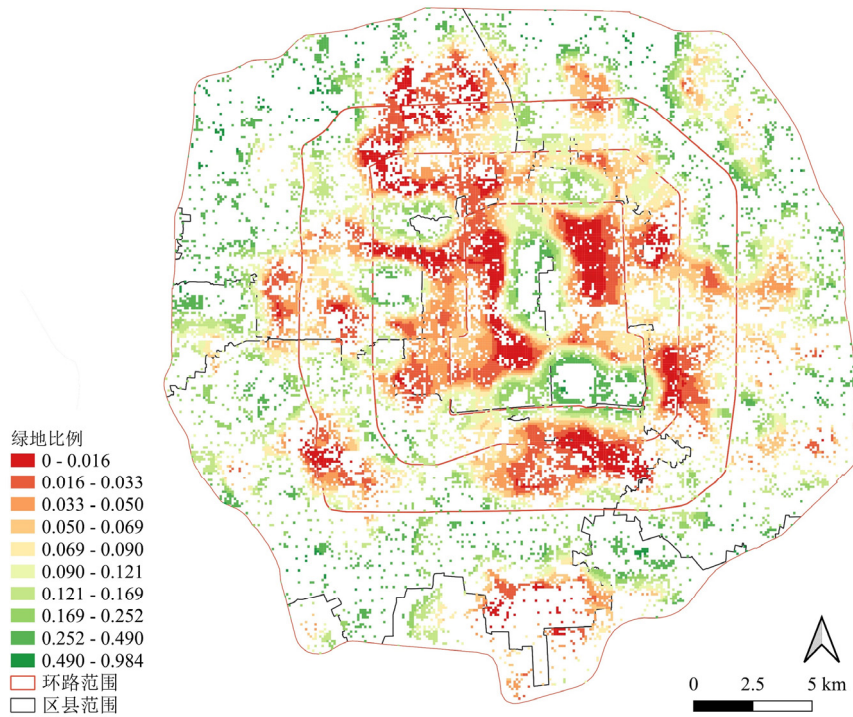
为了借鉴 2.1 节文献综述的内容,有关绿地的暴露指标在本研究中共选择了绿地比例、街道绿视率、NDVI、公园密度、到公园的距离以及到大型绿地的距离共 6 种,此处先介绍绿地比例的测度方式。

绿地比例采用的是 2014 年用地数据中的绿地用地,将其与缓冲区进行相交运算,再按照缓冲区序号融合,计算相应的面积,再按字段连接到相应缓冲区,最后将其面积除以缓冲区总面积,得到绿地比例,无量纲。

由于用地数据仅在 2014 年可获得,因此绿地比例也仅在 2014 年完成测度,测度范围是北京市全域。图 4.14 展示了北京市患者住址 1000 m 缓冲区内 2014 年绿地比例测度的可视化结果。可以发现,北京市中心城区及东南郊区的绿地比例较低,而以山地为主的西北郊区较高,与图 3.4 的绿地空间分布一致。北京市五环内只有在一些大型公园附近的区域,绿地比例相对较高。



(a)



(b)

图 4.14 北京市 AMI 患者住址 1000 m 缓冲区内 2014 年绿地比例测度结果

(a) 北京市全域 (b) 五环内

(6) 街道绿视率

街道绿视率指标的测度基于百度街景图片，本研究采用 SegNet 算法，计算每个百度街景点四个朝向的街景图片中绿色的面积比例^[17]。在 SegNet 算法可识别的 150 个对象中，以下对象被识别为绿色：树、草、山、植物、花、棕榈。类似于 4.3.8 节中空间失序的指标计算，将每个街景点平行于道路（前、后）和垂直于道路（左、右）四个朝向的街景图片的绿色比例取平均值，作为该街景点的绿视率，然后再计算各个患者住址缓冲区内所有街景点的绿视率平均值，作为该缓冲区的绿视率。

该指标在 2013 年、2015 年和 2017 年完成了测度，测度范围是北京市五环内区域，其余年份由线性插值或外推获得。2016 年和 2019 年的街景图片数据未被使用，是因为这两年的街景图片拍摄于冬季，无法真实反应绿化水平。图 4.15 展示了北京市五环内患者住址 1000 m 缓冲区内 2017 年街道绿视率测度的可视化结果。可以看出，北京市五环内，东南侧的绿视率水平要显著低于西北侧，尤其是东二环至东五环一带，街道绿视率大多小于 5%，整体偏低。

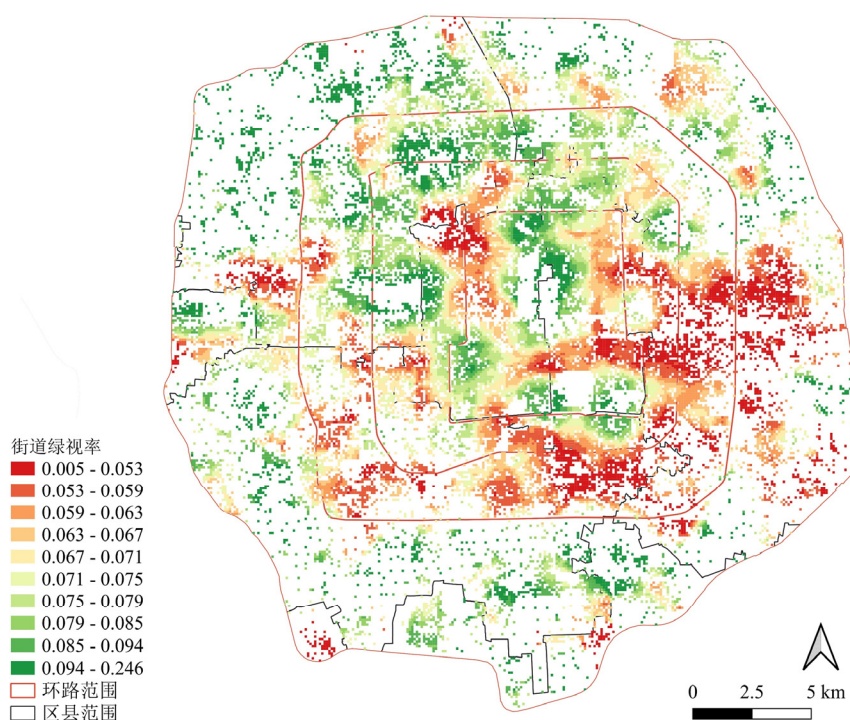


图 4.15 北京市五环内 AMI 患者住址 1000 m 缓冲区内 2017 年街道绿视率测度结果

(7) NDVI

NDVI 的测度基于的数据为 3.2.5 节介绍的 Landsat8 Level1 的 30 米分辨率遥感影像数据,测度需要在 ENVI 软件中完成。首先,需要对遥感影像进行辐射定标和大气校正。其次,可以根据式(4.7)对 NDVI 进行栅格计算:

$$NDVI = \frac{NIR - R}{NIR + R} \quad (4.7)$$

其中, NIR 为近红外波段的反射值, R 为红光波段的反射值。NDVI 的取值范围是从-1 到 1, 因此, 在该区间范围以外的异常值(大约占 0.5%)被强制设定为-1 和 1。建立患者住址缓冲区之后, 利用 QGIS 栅格分析中的分区统计功能, 对缓冲区内 NDVI 求平均值, 即可得到患者住址周边多年的 NDVI 的测度结果。

该指标在 2013 年、2014 年、2015 年、2017 年和 2019 年完成了测度, 测度范围是北京市全域(除去西南角边界缺失部分, 涉及患者比例仅 0.5%), 其余年份由线性插值或外推获得。图 4.16 (a)展示了由 ENVI 分析得到的 2019 年北京市全域 NDVI 空间分布, 图 4.16 (b)展示了北京市五环内患者住址 1000 m 缓冲区内 2019 年 NDVI 测度的可视化结果。

4.3.4 第 4 类建成环境暴露因素: 到公共交通的距离

(1) 地铁站密度及到地铁站的距离

地铁站数据来自于高德地图 POI 数据中分类编码为 150500 的 POI, 并根据北京地铁每年的开通情况进行了人工核查与修正, 最终获得了 2013 年至 2019 年每一年的地铁站 POI 点数据。地铁站密度的计算, 可先将地铁站 POI 点要素与患者住址缓冲区进行空间连接, 再统计出缓冲区内的地铁站 POI 个数即可。到地铁站的距离, 可直接采用近邻分析, 计算出患者住址到地铁站 POI 点要素集合的最近距离。该指标对北京市全域进行了测度, 图 4.17 仅展示了北京市五环内患者住址 1000 m 缓冲区的 2019 年地铁站密度、到地铁站的距离的可视化结果。

(2) 公交站密度

公交站数据来自于高德地图 POI 数据中分类编码形如 1507XX 的 POI, 时间覆盖 2014 年、2016 年、2017 年、2018 年和 2019 年, 空间覆盖北京市全域。公交站密度的计算与地铁站密度类似, 先将公交站 POI 点要素与患者住址缓冲区进行空间连接, 再统计出缓冲区内的公交站 POI 个数即可。图 4.18 仅展示了北京市五环内患者住址 1000 m 缓冲区的 2019 年公交站密度的可视化结果。

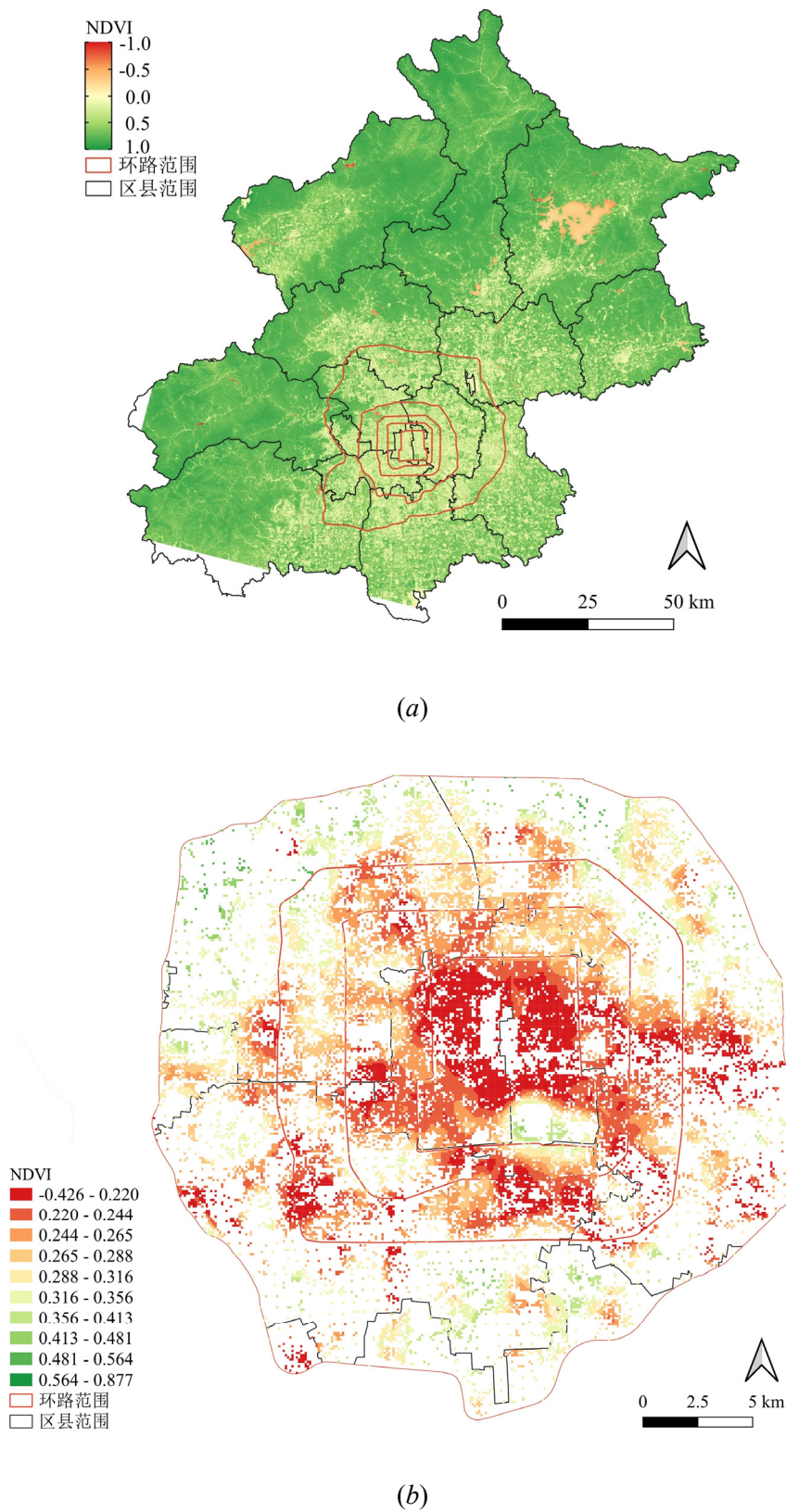
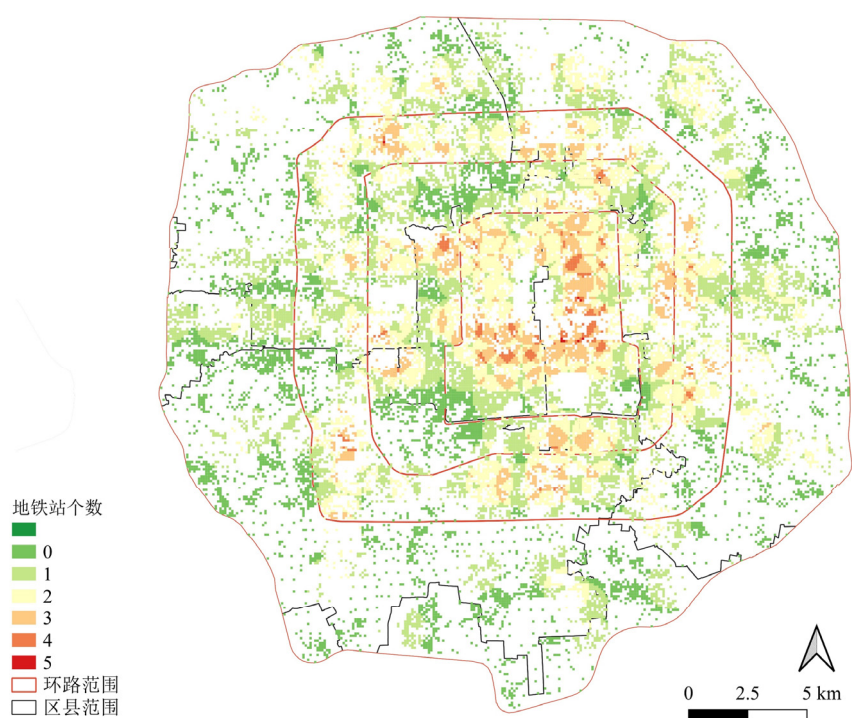
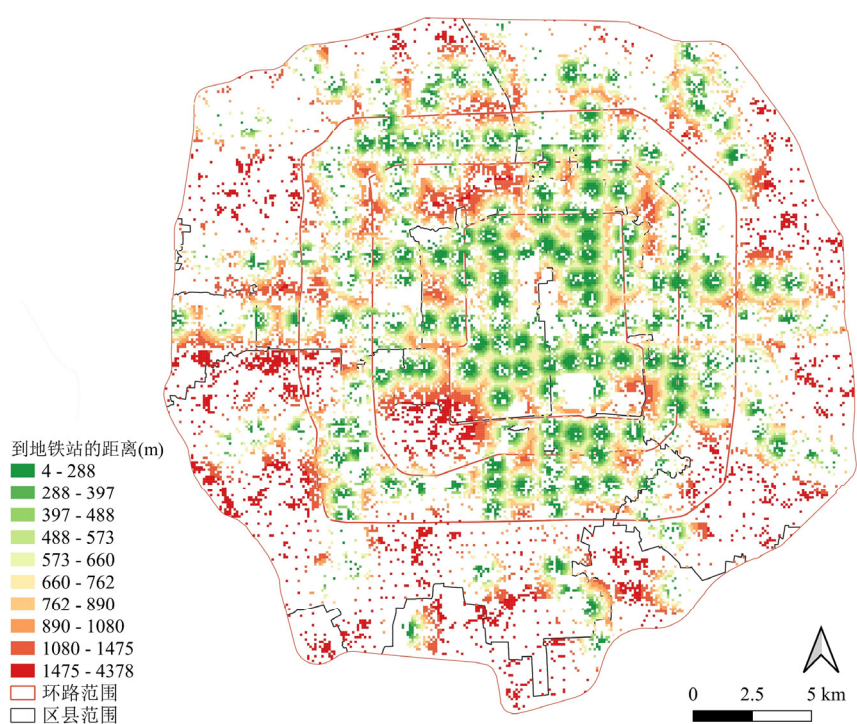


图 4.16 北京市 2019 年 NDVI 测度结果
(a) NDVI 空间分布, (b) 五环内 AMI 患者住址 1000 m 缓冲区内测度结果



(a)



(b)

图 4.17 北京市五环内 AMI 患者住址 1000 m 缓冲区内 2019 年地铁站测度结果：

(a) 地铁站密度，(b) 到地铁站的距离

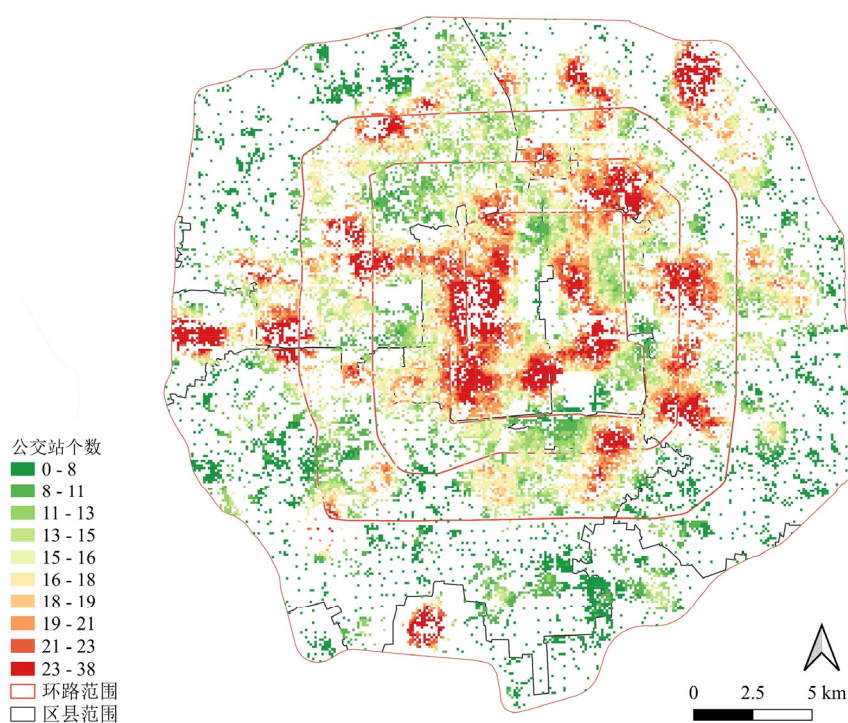


图 4.18 北京市五环内 AMI 患者住址 1000 m 缓冲区内 2019 年公交站密度测度结果

4.3.5 第 5 类建成环境暴露因素：目的地可达性

目的地可达性类别中的众多建成环境暴露因素，大多依赖于各类型的 POI 数据，包括公园、烟酒店、餐厅、快餐店、甜品/饮品/糕饼店、蔬果店、运动场馆、综合医院和药房。因此，先要对 POI 数据按照上述功能分类成为单独的矢量图层，分类依据如表 4.6 所示。

表 4.6 各功能的 POI 分类依据

POI 功能	高德分类编码	可获得的数据年份
公园	1101XX	2011, 2014, 2016-2019
烟酒店	0602XX, 0604XX, 061210	2011, 2014-2019
快餐店	0503XX 中的西式油炸类快餐店	2011, 2014-2019
餐厅	050XXX	2011, 2014-2019
甜品/饮品/糕饼店	050700, 050800, 050900	2011, 2014-2019
蔬果店	0604XX, 060704, 060705	2011, 2014-2019
运动场馆	0801XX, 0802XX	2011, 2014-2019
综合医院	0901XX	2011, 2014, 2016-2019
药房	090601	2011, 2014, 2016-2019

POI 相关的测度，包括密度和距离两类指标。对于密度类指标，将对应功能的 POI 点要素与患者住址缓冲区进行空间连接，再统计出缓冲区内的该功能 POI 个数即可。对于距离类指标，只需要计算患者住址到对应功能 POI 点要素集合的最近距离即可确定。在目的地可达性相关的建成环境暴露因素中，只有公园和综合医院这两种数量较为稀少的 POI 才计算距离类指标。

除上述 POI 相关数据之外，到大型绿地的距离指标的计算方式为：将用地数据中面积大于 5000 m² 的绿地筛选出来，作为大型绿地的数据集，再计算患者住址到大型绿地数据集的最近距离。

目的地可达性相关指标分别在表 4.6 中提供的数据年份下完成了测度，测度范围是北京市全域。由于涉及的建成环境暴露因素较多，图 4.19 至图 4.30 仅展示了这些暴露因素在 2019 年北京市五环内 1000 m 缓冲区内的测度可视化结果，而未展示北京市全域的计算结果。

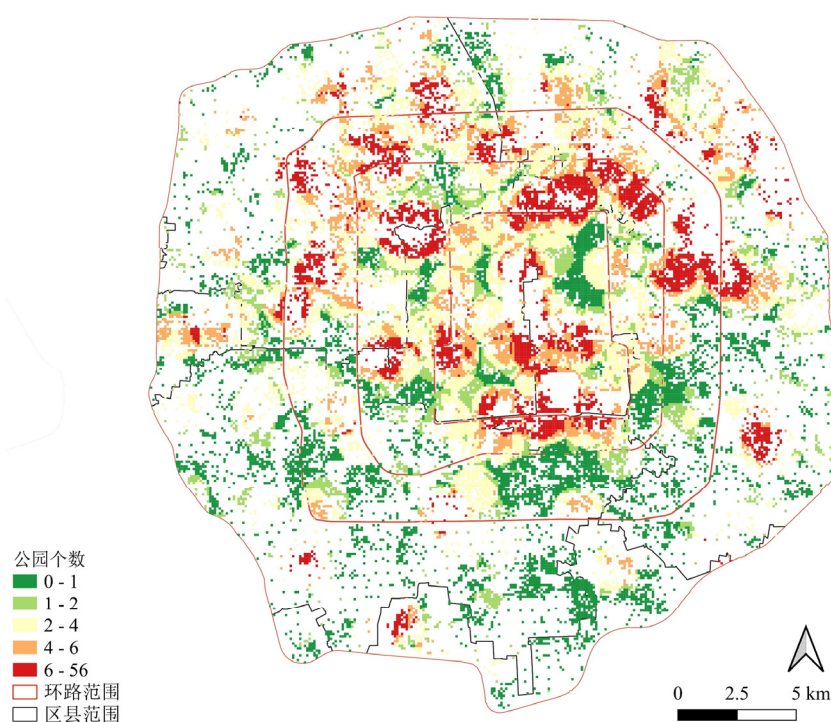


图 4.19 北京市五环内 AMI 患者住址 1000 m 缓冲区内 2019 年公园密度测度结果

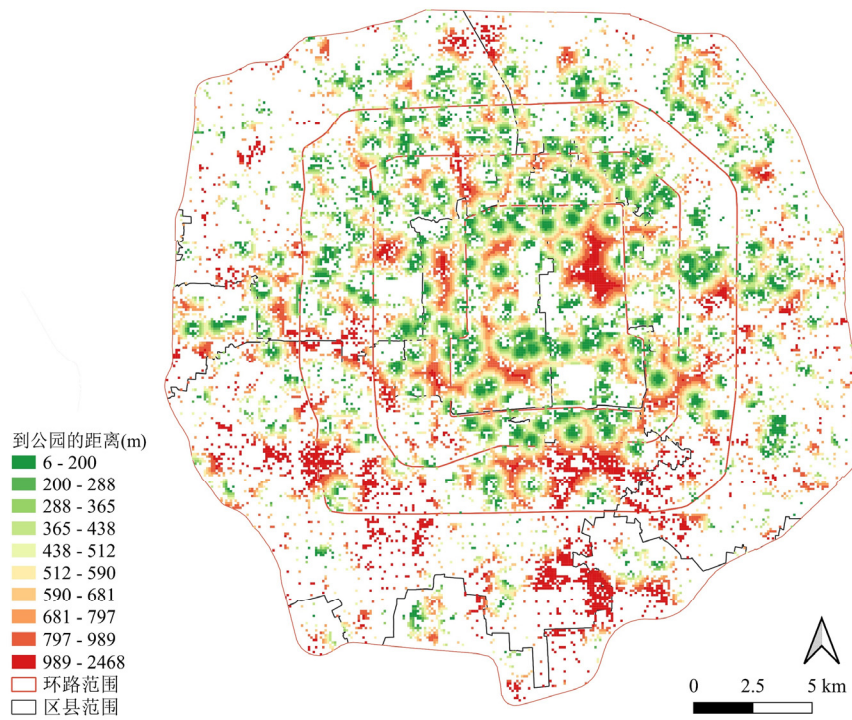


图 4.20 北京市五环内 AMI 患者住址 1000 m 缓冲区内 2019 年到公园的距离测度结果

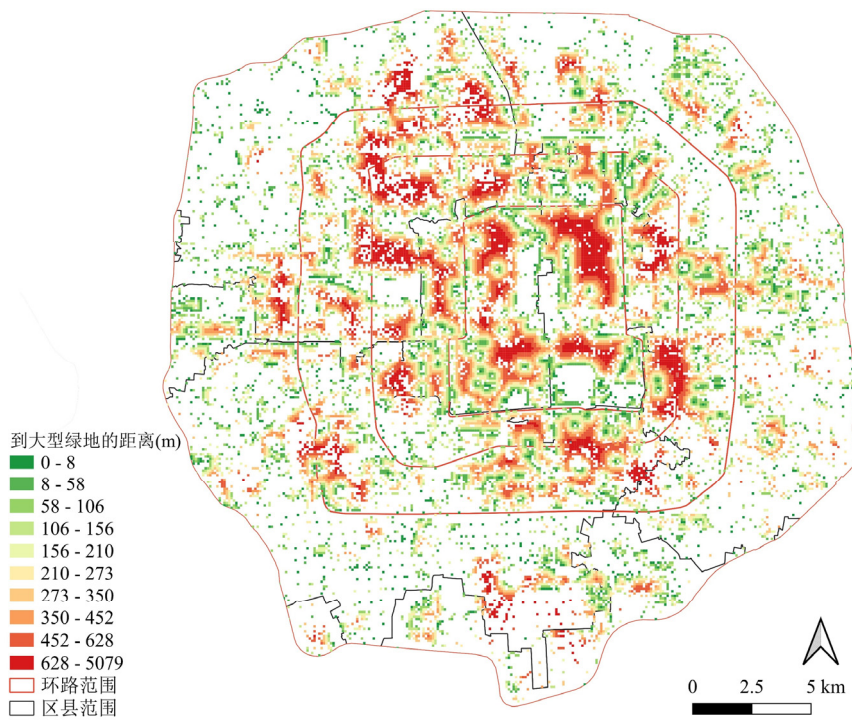


图 4.21 北京市五环内 AMI 患者住址 1000 m 缓冲区内 2019 年到大型绿地的距离的测度结果

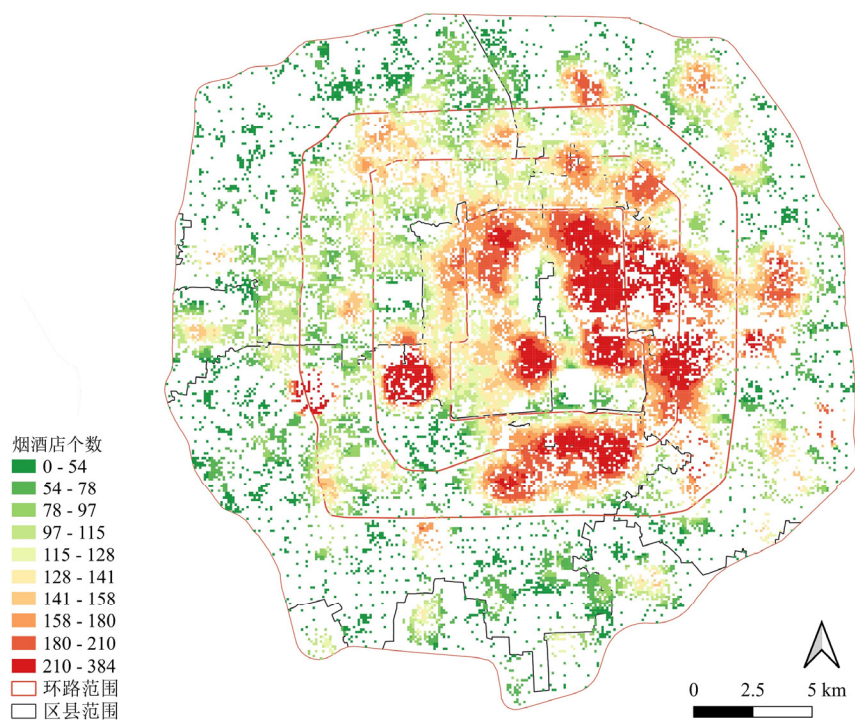


图 4.22 北京市五环内 AMI 患者住址 1000 m 缓冲区内 2019 年烟酒店密度测度结果

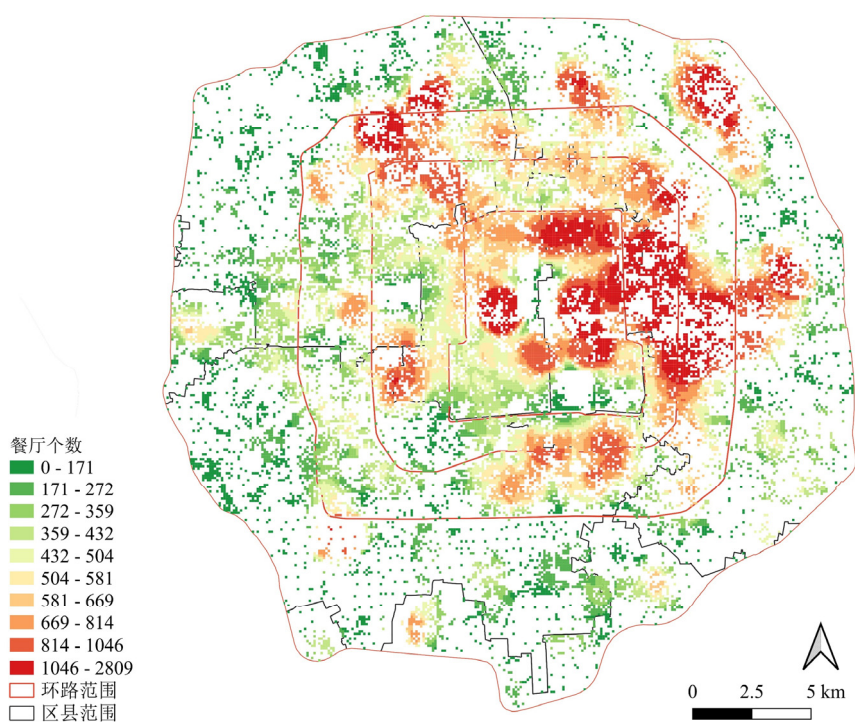


图 4.23 北京市五环内 AMI 患者住址 1000 m 缓冲区内 2019 年餐厅密度测度结果

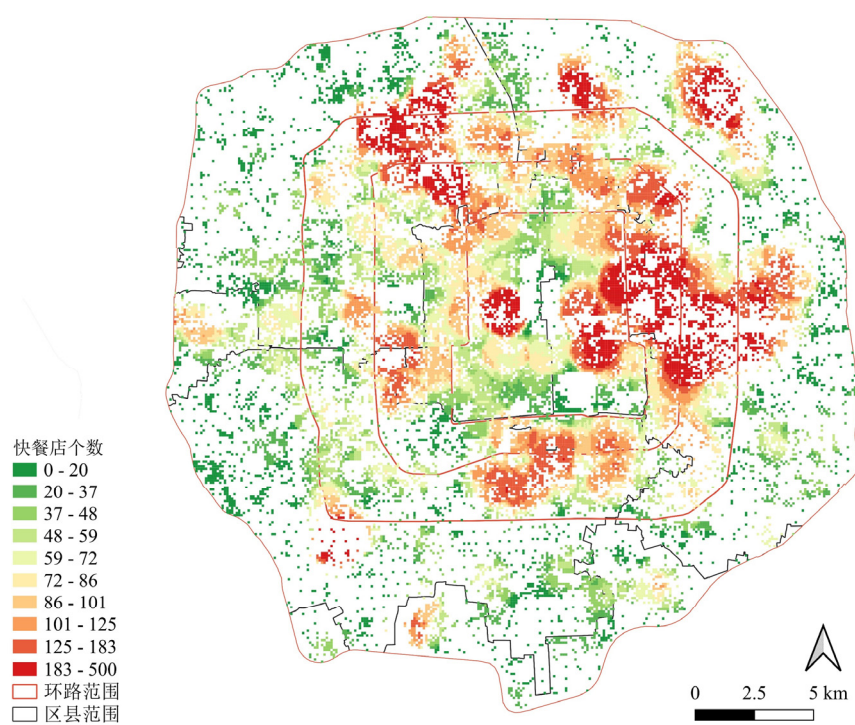


图 4.24 北京市五环内 AMI 患者住址 1000 m 缓冲区内 2019 年快餐店密度测度结果

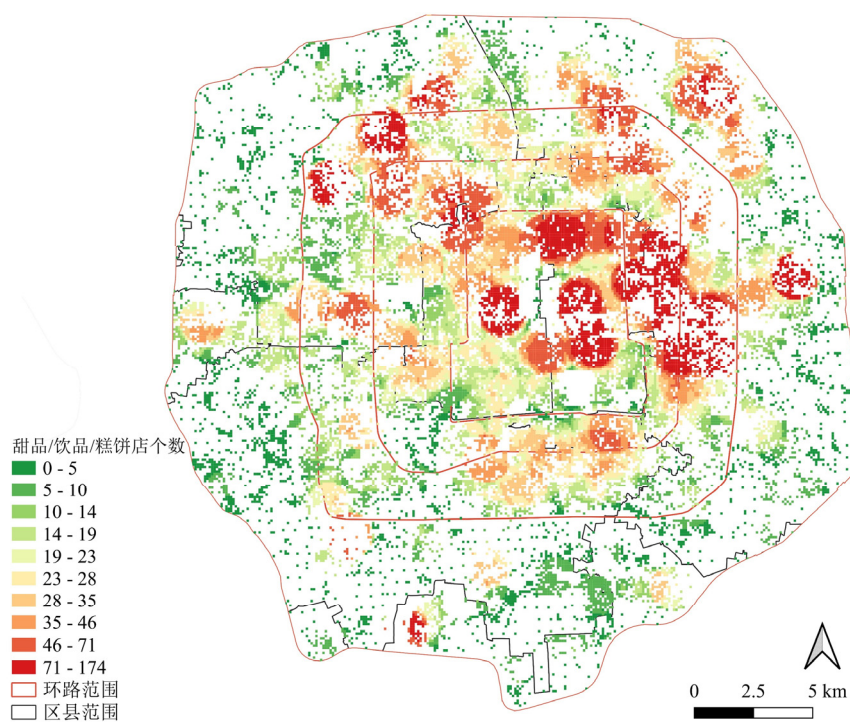


图 4.25 北京市五环内 AMI 患者住址 1000 m 缓冲区内 2019 年甜品/饮品/糕饼店密度测度结果

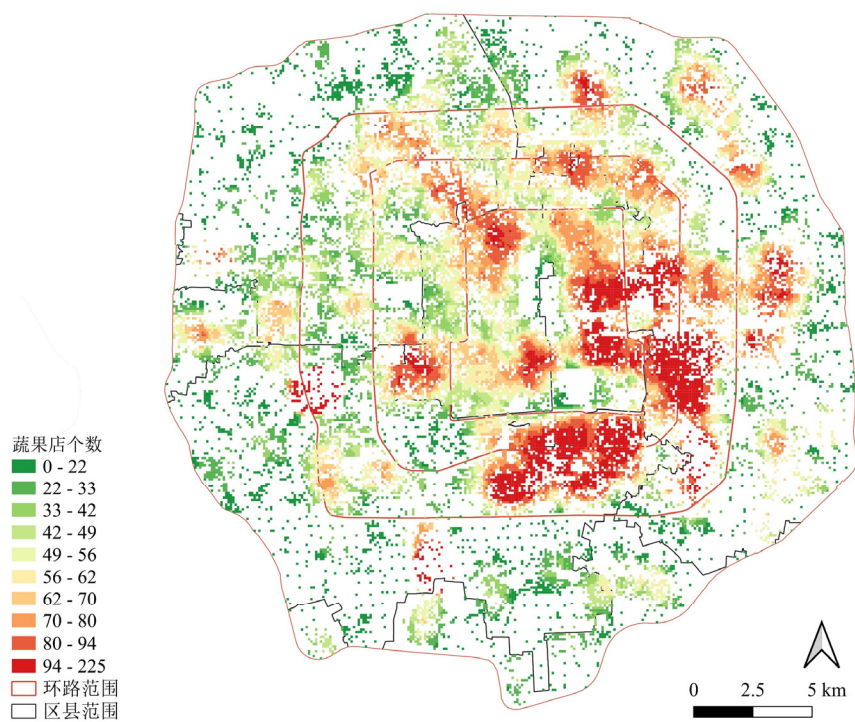


图 4.26 北京市五环内 AMI 患者住址 1000 m 缓冲区内 2019 年蔬果店密度测度结果

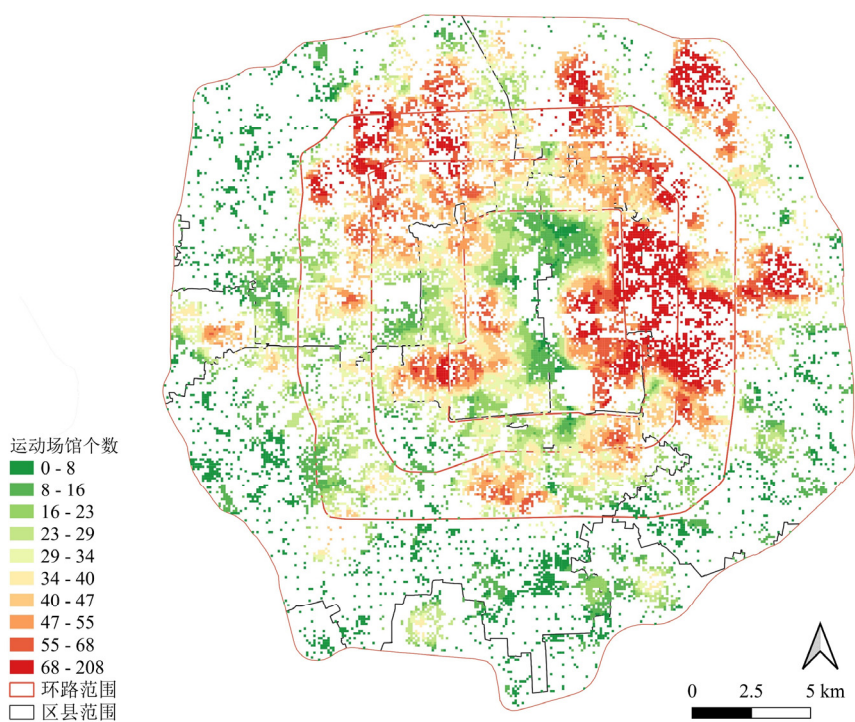
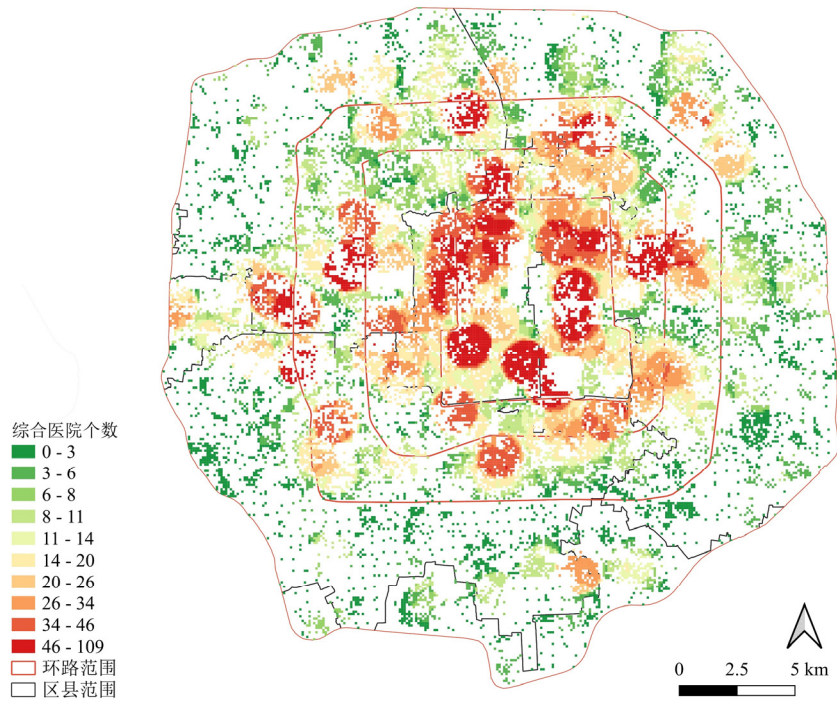
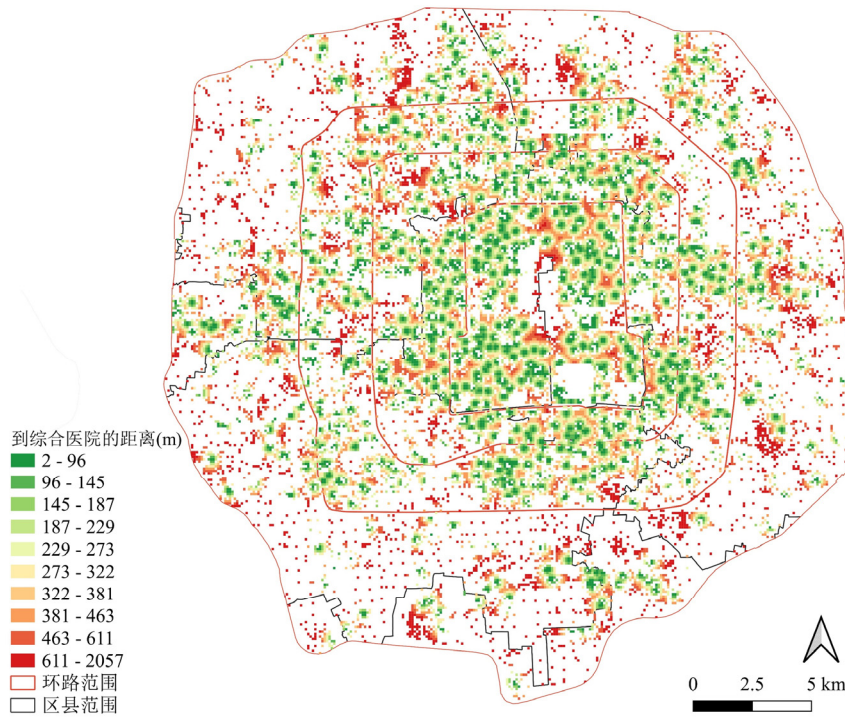


图 4.27 北京市五环内 AMI 患者住址 1000 m 缓冲区内 2019 年运动场馆密度测度结果



(a)



(b)

图 4.28 北京市五环内 AMI 患者住址 1000 m 缓冲区内 2019 年综合医院测度结果
(a) 综合医院密度 (b) 到综合医院的距离

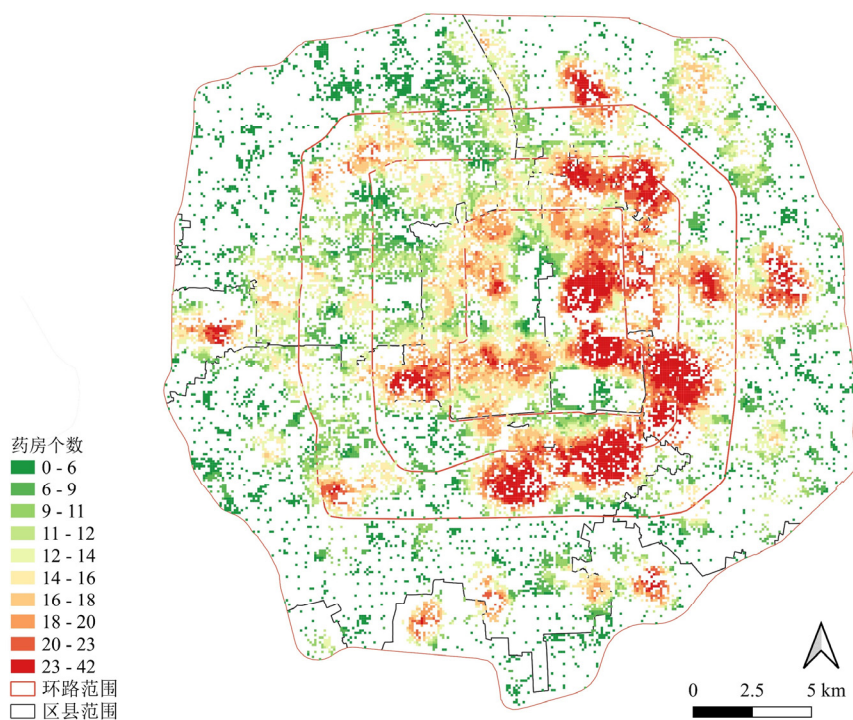
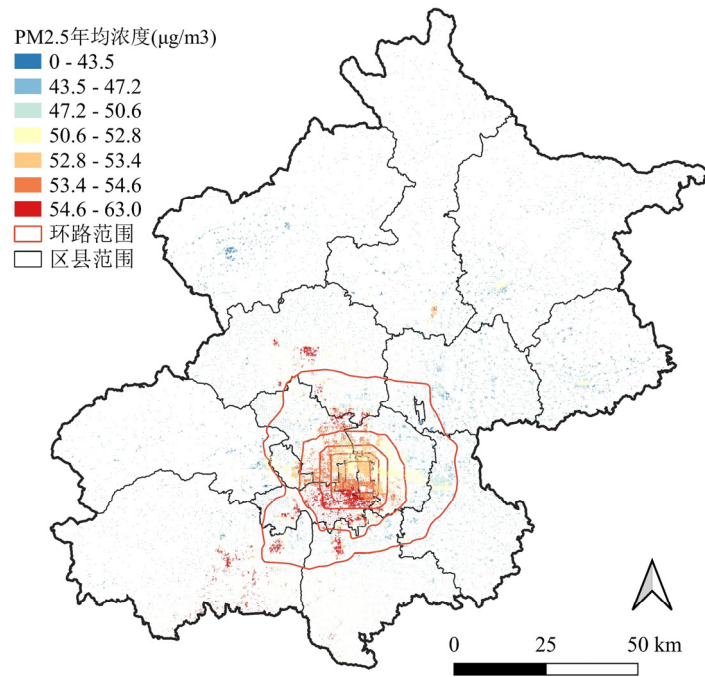


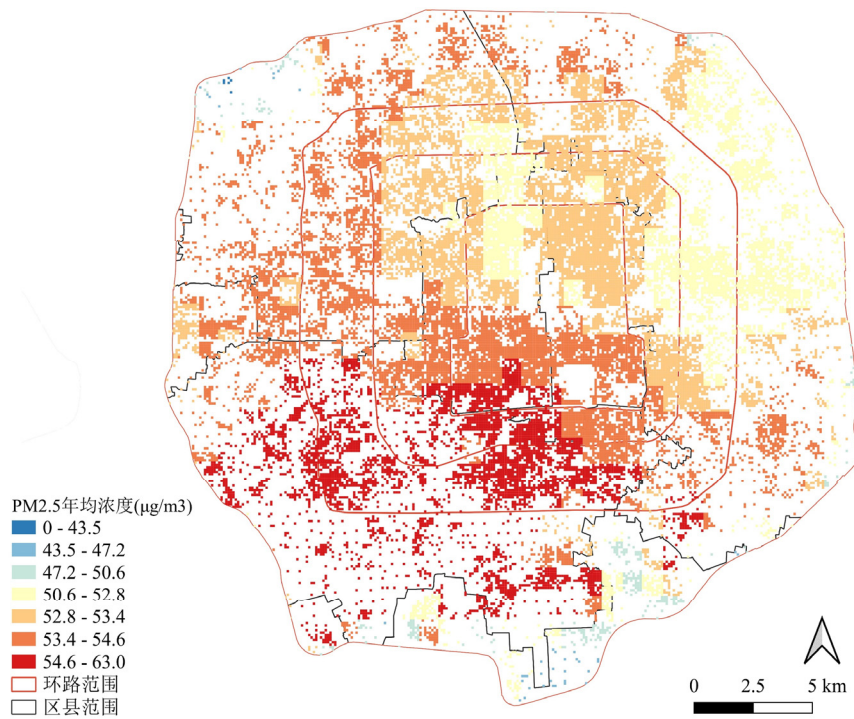
图 4.29 北京市五环内 AMI 患者住址 1000 m 缓冲区内 2019 年药房密度测度结果

4.3.6 空气污染

PM_{2.5} 浓度的测度基于 3.2.6 节介绍的大气 PM_{2.5} 浓度的 1 km 栅格数据，只需要对患者住址所在的大气质量栅格进行栅格采样，即可获得该住址处的 PM_{2.5} 年均浓度。该指标在 2013 年至 2018 年的每一年都完成了测度，测度范围是北京市全域，2019 年由线性外推获得。图 4.28 展示了 2018 年北京市急性心肌梗死患者住址所在位置的 PM_{2.5} 浓度的可视化结果。



(a)



(b)

图 4.30 北京市 AMI 患者住址 1000 m 缓冲区内 2018 年 $\text{PM}_{2.5}$ 浓度测度结果

(a) 北京市全域 (b) 五环内

4.4 患者建成环境暴露评价方法与结果

4.4.1 建成环境暴露评价方法

为了反映暴露组在时间尺度上动态性的特征，基于 4.3 节获得的多年建成环境测度结果，本研究分别测度了 AMI 患者对建成环境的基线暴露水平、累积平均暴露水平两种暴露评价方式^[120, 121]。基线暴露采用的是患者急性心肌梗死首次发病康复出院当年的建成环境暴露水平，而累积平均暴露水平采用的是患者从首次发病康复出院到首次急性心肌梗死复发（或 2019 年 12 月 31 日）之间的多年暴露水平的平均值。对于部分建成环境暴露因素缺失个别年份数据的情况，采用相邻年份数据进行线性插值来完成填补。由于数据所限，土地利用混合度、步行指数、绿地比例和到大型绿地的距离这四个指标仅能在 2014 年进行测度，到主干道的距离仅能在 2016 年进行测度，因此上述五个指标仅进行基线暴露评价（近似采用数据所在年份的测度结果），而不进行累积平均暴露评价。

如 2.3 节所述，环境暴露变量之间可能存在较强的多重共线性。因此，在上述建成环境暴露评价结果基础之上，本研究进一步进行了相关性分析。考虑到变量间可能存在非线性相关，本研究选用 spearman 相关系数：设 x_i 在 X 样本中的秩为 R_i ， y_i 在 Y 样本中的秩为 S_i ，则两个样本的 spearman 相关系数为

$$r_s = \frac{\sum_{i=1}^n (R_i - \bar{R})(S_i - \bar{S})}{\sqrt{\sum_{i=1}^n (R_i - \bar{R})^2} \sqrt{\sum_{i=1}^n (S_i - \bar{S})^2}} \quad (4.8)$$

其中， $\bar{R}$ 和 $\bar{S}$ 为 R_i 和 S_i 的均值。计算出建成环境各因素之间的两两 spearman 相关系数之后，可绘制成相关系数矩阵热力图。

4.4.2 建成环境暴露评价结果

基于上述建成环境暴露评价方法，北京市急性心肌梗死患者的建成环境基线暴露水平与累积平均暴露水平的统计结果如表 4.7 所示。

患者建成环境暴露水平之间的相关系数矩阵热力图如图 4.31 所示。从图中可以发现，无论是基线暴露水平，还是累积平均暴露水平，密度类指标之间、目的地可达性指标之间，以及同一种指标的不同缓冲区测度结果之间，均存在较高的相关系数（绝对值大于 0.6），说明建成环境暴露因素之间的相关性较高，这也体现出后续变量选择算法的必要性，否则多重共线性的存在将会使回归模型结果出现估计有偏、方差很大的问题。

表 4.7 北京市急性心肌梗死患者的建成环境暴露评价结果

暴露因素	基线暴露水平			累积平均暴露水平		
	均值±标准差	中位数	IQR	均值±标准差	中位数	IQR
人口密度 (人/km²)						
300 m 缓冲区	15244 ± 14403	13103	19482	15896 ± 14977	13780	20347
500 m 缓冲区	15152 ± 13572	13535	19832	15808 ± 14114	14174	20680
1000 m 缓冲区	14895 ± 12195	14151	20608	15544 ± 12681	14868	21538
建筑密度 (m²/m²) 注：五环内						
300 m 缓冲区	0.25 ± 0.07	0.25	0.08	0.25 ± 0.06	0.25	0.08
500 m 缓冲区	0.25 ± 0.06	0.24	0.07	0.25 ± 0.05	0.25	0.06
1000 m 缓冲区	0.24 ± 0.05	0.24	0.05	0.24 ± 0.04	0.24	0.05
建筑容积率 (m²/m²) 注：五环内						
300 m 缓冲区	1.51 ± 0.57	1.49	0.73	1.60 ± 0.58	1.60	0.71
500 m 缓冲区	1.46 ± 0.50	1.48	0.64	1.56 ± 0.51	1.58	0.62
1000 m 缓冲区	1.40 ± 0.43	1.43	0.58	1.50 ± 0.43	1.56	0.57
城市功能密度 (个)						
300 m 缓冲区	281 ± 332	162	373	366 ± 356	294	430
500 m 缓冲区	750 ± 834	433	1055	979 ± 892	813	1225
1000 m 缓冲区	2774 ± 2929	1586	4091	3621 ± 3108	3175	4857
土地利用混合度 (-)						
300 m 缓冲区	1.01 ± 0.34	1.05	0.45	NA	NA	NA
500 m 缓冲区	1.23 ± 0.30	1.27	0.39	NA	NA	NA
1000 m 缓冲区	1.44 ± 0.26	1.49	0.33	NA	NA	NA
城市功能混合度 (-)						
300 m 缓冲区	1.93 ± 0.58	2.15	0.46	2.02 ± 0.46	2.19	0.34
500 m 缓冲区	2.09 ± 0.49	2.26	0.31	2.15 ± 0.37	2.29	0.23
1000 m 缓冲区	2.22 ± 0.37	2.35	0.20	2.27 ± 0.28	2.36	0.13
路网密度 (km/km²)						
300 m 缓冲区	10.7 ± 3.7	11.1	5.3	10.8 ± 3.6	11.2	5.0
500 m 缓冲区	10.0 ± 3.5	10.6	5.2	10.1 ± 3.4	10.7	5.0
1000 m 缓冲区	9.2 ± 3.5	10.0	5.7	9.3 ± 3.4	10.0	5.5
到主干道的距离 (m)						
	259.9 ± 296.8	183.6	240.2	NA	NA	NA
空间失序 (-) 注：五环内						
300 m 缓冲区	1.95 ± 0.41	1.90	0.50	2.00 ± 0.35	1.94	0.41
500 m 缓冲区	1.94 ± 0.32	1.91	0.41	1.99 ± 0.27	1.94	0.32
1000 m 缓冲区	1.93 ± 0.25	1.91	0.34	1.98 ± 0.21	1.95	0.26
步行指数 (-) 注：五环内						
300 m 缓冲区	0.14 ± 0.55	0.16	0.65	NA	NA	NA
500 m 缓冲区	0.14 ± 0.54	0.17	0.63	NA	NA	NA
1000 m 缓冲区	0.13 ± 0.50	0.18	0.60	NA	NA	NA

续表 4.7 北京市急性心肌梗死患者的建成环境暴露评价结果

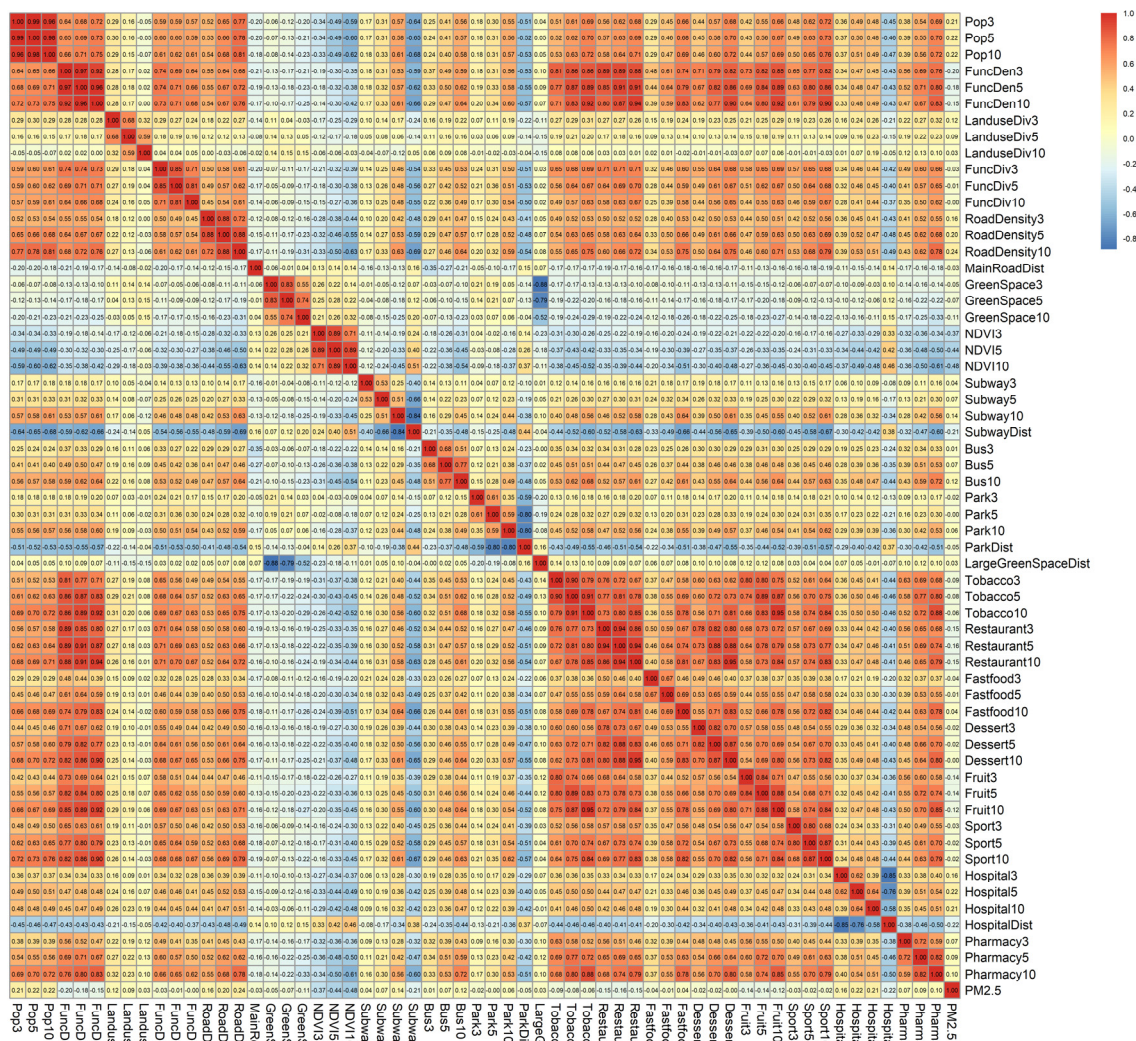
暴露因素	基线暴露水平			累积平均暴露水平		
	均值±标准差	中位数	IQR	均值±标准差	中位数	IQR
绿地比例 (-)						
300 m 缓冲区	0.07 ± 0.12	0.02	0.08	NA	NA	NA
500 m 缓冲区	0.09 ± 0.13	0.05	0.10	NA	NA	NA
1000 m 缓冲区	0.12 ± 0.14	0.07	0.11	NA	NA	NA
街道绿视率 (-) 注：五环内						
300 m 缓冲区	0.21 ± 0.08	0.20	0.10	0.23 ± 0.09	0.22	0.11
500 m 缓冲区	0.21 ± 0.07	0.20	0.08	0.23 ± 0.07	0.22	0.08
1000 m 缓冲区	0.21 ± 0.05	0.20	0.06	0.23 ± 0.05	0.22	0.06
NDVI (-)						
300 m 缓冲区	0.26 ± 0.10	0.25	0.12	0.30 ± 0.10	0.26	0.10
500 m 缓冲区	0.28 ± 0.10	0.27	0.12	0.29 ± 0.09	0.27	0.10
1000 m 缓冲区	0.30 ± 0.11	0.28	0.13	0.32 ± 0.11	0.29	0.12
地铁站密度 (个)						
300 m 缓冲区	0.1 ± 0.2	0	0	0.06 ± 0.23	0	0
500 m 缓冲区	0.2 ± 0.4	0	0	0.19 ± 0.40	0	0
1000 m 缓冲区	0.8 ± 1.0	0	1	0.81 ± 0.96	0.33	1.17
到地铁站的距离 (m)	5204.4 ± 9395.8	1080.3	3849.7	4984.85 ± 9207.42	1032.8	3323.36
公交站密度 (个)						
300 m 缓冲区	1.6 ± 1.7	1	2	1.6 ± 1.6	1	1
500 m 缓冲区	4.2 ± 3.5	4	4	4.3 ± 3.2	4	4
1000 m 缓冲区	15.0 ± 10.3	14	12	15.3 ± 9.8	15	12
公园密度 (个)						
300 m 缓冲区	0.2 ± 0.5	0	0	0.2 ± 0.5	0	0
500 m 缓冲区	0.6 ± 1.1	0	1	0.6 ± 1.1	0	1
1000 m 缓冲区	2.4 ± 3.2	2	4	2.7 ± 3.2	2	4
到公园的距离 (m)	1160.4 ± 1541.5	686.4	768.5	1040.9 ± 1249.3	639.2	672.6
到大型绿地的距离(m)	297.9 ± 264.2	238.4	287.3	NA	NA	NA
烟酒店密度 (个)						
300 m 缓冲区	8.2 ± 7.4	7	10	9.8 ± 7.9	9	10
500 m 缓冲区	20.2 ± 16.8	17	24	24.3 ± 18.1	23	27
1000 m 缓冲区	69.6 ± 55.7	61	85	83.6 ± 60.1	83	95
餐厅密度 (个)						
300 m 缓冲区	29.6 ± 42.5	14	39	37.9 ± 44.8	26	47
500 m 缓冲区	77.7 ± 98.5	41	106	99.5 ± 102.5	78	123
1000 m 缓冲区	279.3 ± 321.2	156	409	358.2 ± 327.8	305	458

续表 4.7 北京市急性心肌梗死患者的建成环境暴露评价结果

暴露因素	基线暴露水平			累积平均暴露水平		
	均值±标准差	中位数	IQR	均值±标准差	中位数	IQR
快餐店密度 (个)						
300 m 缓冲区	0.4 ± 0.9	0	0	0.5 ± 1.0	0	1
500 m 缓冲区	1.1 ± 1.8	0	2	1.4 ± 2.0	1	2
1000 m 缓冲区	4.2 ± 4.7	3	6	5.0 ± 5.2	4	8
甜品/饮品/糕饼店密度 (个)						
300 m 缓冲区	2.1 ± 4.3	1	2	2.4 ± 4.2	1	3
500 m 缓冲区	5.6 ± 9.5	2	7	6.5 ± 9.2	4	8
1000 m 缓冲区	20.5 ± 28.5	10	26	23.7 ± 26.1	17	32
蔬果店密度 (个)						
300 m 缓冲区	3.1 ± 3.6	2	4	4.0 ± 3.9	3	5
500 m 缓冲区	7.5 ± 7.5	5	9	9.7 ± 8.1	8	11
1000 m 缓冲区	25.2 ± 23.3	19	32	32.6 ± 25.4	31	39
运动场馆密度 (个)						
300 m 缓冲区	1.8 ± 2.9	1	2	2.2 ± 3.2	1	3
500 m 缓冲区	4.9 ± 6.4	3	7	6.0 ± 7.1	4	8
1000 m 缓冲区	18.9 ± 20.8	13	27	23.2 ± 22.9	19	32
综合医院密度 (个)						
300 m 缓冲区	0.5 ± 0.5	1	1	0.5 ± 0.5	1	1
500 m 缓冲区	0.7 ± 0.4	1	1	0.7 ± 0.4	1	1
1000 m 缓冲区	0.9 ± 0.3	1	0	0.9 ± 0.3	1	0
到综合医院的距离 (m)	514.9 ± 724.3	283.1	338.6	534.6 ± 735.0	296.3	347.0
药房密度 (个)						
300 m 缓冲区	1.3 ± 1.2	1	2	1.4 ± 1.3	1	2
500 m 缓冲区	3.2 ± 2.4	3	4	3.4 ± 2.5	3	4
1000 m 缓冲区	10.7 ± 7.4	11	12	11.6 ± 7.8	12	12
PM_{2.5} (μg/m³)	66.9 ± 17.6	63.8	32.2	57.3 ± 11.8	54.9	17.0

注：建筑密度、建筑容积率、空间失序、步行指数、街道绿视率五个指标由于数据所限，为五环内的暴露评价结果，其余建成环境暴露因素为北京市全域的暴露评价结果。

第 4 章 建成环境暴露评价



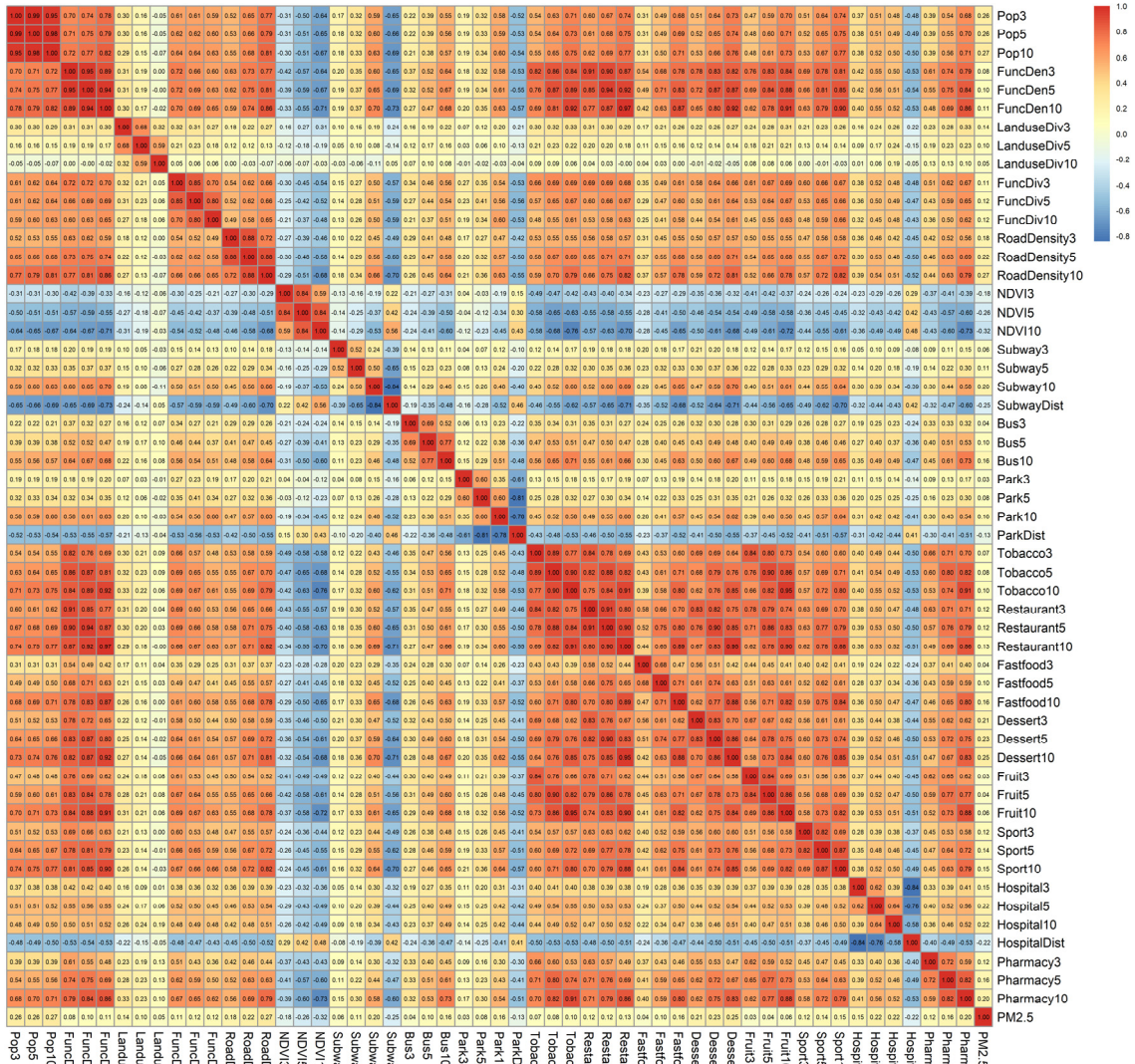
(a)

图 4.31 建成环境暴露因素之间的 Spearman 相关系数矩阵

(a) 基线暴露 (b) 累积平均暴露

注：标签中的数字“3”、“5”、“10”代表 300 米、500 米、1000 米缓冲距；Pop 为人口密度，FuncDen 为城市功能密度，LanduseDiv 为土地利用混合度，FuncDiv 为城市功能混合度，RoadDensity 为路网密度，MainRoadDist 为到主干道的距离，GreenSpace 为绿地比例，NDVI 为归一化植被指数，Subway 为地铁站密度，SubwayDist 为到地铁站的距离，Bus 为公交站密度，ParkDist 为到公园的距离，LargeGreenSpaceDist 为到大型绿地的距离，Tobacco 为烟酒店密度，Restaurant 为餐厅密度，Fastfood 为快餐店密度，Dessert 为甜品/饮品/糕饼店密度，Fruit 为蔬果店密度，Sport 为运动场馆密度，Hospital 为综合医院密度，HospitalDist 为到综合医院的距离，Pharmacy 为药房密度，PM2.5 为 PM_{2.5} 年均浓度。

第4章 建成环境暴露评价



(b)

续图 4.31 建成环境暴露因素之间的 Spearman 相关系数矩阵

(a) 基线暴露 (b) 累积平均暴露

注：标签中的数字“3”、“5”、“10”代表 300 米、500 米、1000 米缓冲区；Pop 为人口密度，FuncDen 为城市功能密度，FuncDiv 为城市功能混合度，RoadDensity 为路网密度，NDVI 为归一化植被指数，Subway 为地铁站密度，SubwayDist 为到地铁站的距离，Bus 为公交站密度，ParkDist 为到公园的距离，Tobacoo 为烟酒店密度，Restaurant 为餐厅密度，Fastfood 为快餐店密度，Dessert 为甜品/饮品/糕饼店密度，Fruit 为蔬果店密度，Sport 为运动场馆密度，Hospital 为综合医院密度，HospitalDist 为到综合医院的距离，Pharmacy 为药房密度，PM2.5 为 PM_{2.5} 年均浓度。

4.5 本章小结

本章主要介绍了建成环境暴露评价的具体步骤与主要结果：

（1）根据北京市心血管疾病监测系统提供的 2013 年至 2019 年间急性冠心病患者的住址文本信息，本章通过软件自动解析、招募并培训 20 位志愿者开展人工核查的方法，耗时 4 个月，完成了北京市近 24 万位急性冠心病患者住址的地址解析，确定了患者住址的精确经纬度，实现空间落位；

（2）根据已有文献综述，确定了本研究关注的建成环境暴露因素，包括密度、多样性、设计、到公共交通的距离、目的地可达性五大类以及空气污染，共计 26 个小类的环境暴露因素；

（3）针对上述建成环境暴露因素，本章对急性心肌梗死患者住址周边 300 m、500 m、1000 m 三种缓冲区内进行了多年份的建成环境指标测度，4.3 节对各个指标的测度方式逐一进行了详细介绍；基于此，本章对急性心肌梗死患者个体层面的建成环境暴露水平开展评价，采用了基线暴露、累积平均暴露两种暴露评价方式，并对各个建成环境暴露因素的评价结果进行了相关性分析，发现多数建成环境暴露因素之间存在较高的相关性。

本章的建成环境暴露评价结果，为第 5 章与急性心肌梗死复发的关联分析提供了暴露数据基础。

第 5 章 建成环境暴露与 AMI 复发的关联分析

5.1 暴露组变量选择算法的比较

5.1.1 变量选择算法的原理

暴露组学的核心思想之一，是将尽可能多的潜在环境暴露因素纳入研究的考察范围，但是，众多的环境暴露因素之间难免存在相关性，若将所有环境暴露变量纳入到回归模型中，会造成回归系数估计的有偏性，以及回归系数估计方差的显著膨胀，导致众多变量出现不显著的现象。因此，暴露组学研究中，需要首先进行变量选择。

如第 2 章文献综述所言，Agier 等人^[81]对暴露组学中的变量选择算法开展了综述，共涉及 EWAS、EWAS-MLR、ENET、sPLS、DSA 等算法。

上述模型均基于线性模型，而线性模型中最简单的是线性回归模型，下文以线性回归模型为例，简要介绍上述变量选择算法的基本原理。但是，这些变量选择算法并不仅限于线性回归模型，也适用于自变量 X 相关部分呈线性形式的其他模型，如 logistic 回归、Cox 比例风险模型等。

1.EWAS 算法

EWAS 方法是对所有待选环境暴露因素逐个进行单因素回归的方法，以线性回归模型为例：

$$Y = \beta_0 + \beta_i X_i + \sum_k \gamma_k Z_k + \varepsilon, i=1, \dots, N \quad (5.1)$$

其中， β_0 为截距， X_i 为第 i 个待选环境暴露因素的观测值， β_i 为对应的回归系数， Z_k 和 γ_k 为模型的第 k 个控制变量（如年龄、性别等）及其对应的回归系数， ε 为随机误差项，服从正态分布 $N(0, \sigma^2)$ ， N 为环境暴露因素的总个数。

因此，EWAS 算法会建立 N 个线性回归模型，每个模型中分别对 β_i 进行双侧 t 检验，得到其 p 值。由于 N 个假设检验同时进行，为了控制第 I 类错误概率，需对 p 值进行多重比较的校正，本研究根据 Agier 等人^[81]的研究结果，采用 BH

（Benjamini and Hochberg）方法进行 p 值校正。对于校正后的 p 值小于 0.05 的环境暴露因素，认为其对观测值 Y 具有显著影响，从而该变量被筛选出来。

2.EWAS-MLR 算法

EWAS-MLR 算法先通过 EWAS 算法初步筛选出环境暴露变量，再将这些变量

纳入到一个多元线性回归模型中进行分析，模型形式如下：

$$Y = \beta_0 + \sum_{i \in EWAS} \beta_i X_i + \sum_k \gamma_k Z_k + \varepsilon \quad (5.2)$$

其中， $i \in EWAS$ 代表该模型中的 X 来自于 EWAS 算法中筛选出的变量集合。

从多元线性回归模型中，挑选出双侧 t 检验的 p 值小于 0.05 的变量（不需要再作 p 值校正），则将这些变量当作算法选择出的最终变量。

在介绍下面的算法之前，需简要介绍前面的线性回归模型的拟合原理。线性回归模型的参数估计采用的是最小二乘法，即选择使得误差平方和最小的参数作为拟合值，其原理可由下式表达：

$$\hat{\beta}^{OLS} = \arg \min_{\beta} \|Y - X\beta\|_2^2 \quad (5.3)$$

其中， $\hat{\beta}^{OLS}$ 为回归系数的最小二乘估计值， Y 为观测值构成的向量， X 为设计矩阵， β 为系数向量， $\|\cdot\|_2$ 为向量的二范数。上述最优化问题具有如下的解析解：

$$\hat{\beta}^{OLS} = (X^T X)^{-1} X^T Y \quad (5.4)$$

由统计学知识可以推得，上述解析解是一个无偏估计。

3.ENET 算法

ENET 算法集合了岭回归与 LASSO 回归两种正则化方法的优势。这两种正则化方法的思想，是在公式(5.3)追求的无偏性基础上，兼顾回归系数估计量的估计方差尽可能小的原则。因此，这两种方法都在公式(5.3)的优化目标函数上增加了惩罚项，以同时控制估计偏差和方差。

岭回归方法增加的是回归系数向量的二范数惩罚项^[80]：

$$\beta^{Ridge} = \arg \min_{\beta} \|Y - X\beta\|_2^2 + \lambda \|\beta\|_2^2 \quad (5.5)$$

其中， λ 为惩罚系数，可用交叉检验法确定。上述最优化问题也存在解析解：

$$\hat{\beta}^{Ridge} = (X^T X + \lambda I)^{-1} X^T Y \quad (5.6)$$

当多个自变量之间存在较为严重的多重共线性时， $X^T X$ 将接近奇异矩阵，导致不可逆，使得公式(5.4)存在严重的数值不稳定性。而岭回归中，公式(5.6)括号里的矩阵实现了主对角元素占优，从而能够实现求逆的数值稳定性。这对于建成环境研究极为重要，因为从 4.4.2 节可知，众多建成环境因素之间存在较高的相关性。然而，岭回归方法并不能实现变量选择的功能。

而 LASSO 回归采用的是回归系数向量的一范数惩罚项^[80]：

$$\beta^{LASSO} = \arg \min_{\beta} \|Y - X\beta\|_2^2 + \lambda \|\beta\|_1 \quad (5.7)$$

其中, λ 为惩罚系数, 可用交叉检验法确定。上述最优化问题没有解析解, 需要通过坐标下降法等数值优化算法得到。LASSO 回归可以大幅缩减变量个数, 实现变量选择的效果。

ENET 算法同时包含二范数和一范数的惩罚项, 得到如下优化目标函数^[122]:

$$\beta^{ENET} = \arg \min_{\beta} \|Y - X\beta\|^2 + \lambda_1 \|\beta\|_1 + \lambda_2 \|\beta\|_2^2 \quad (5.8)$$

该算法既具备了岭回归抵抗多重共线性的功能, 又具备了 LASSO 回归变量选择的效果。

4.sPLS 算法

sPLS 算法可理解为主成分分析 (principle component analysis, PCA) 的延伸方法。主成分分析是面对高维数据共线性的问题而提出的一种降维方法, 它通过选择已有自变量的线性组合形式 (称为主成分), 从而使得少量主成分便能挖掘出数据集的绝大部分信息量 (一般大于 80% 即可)。主成分, 相当于是依赖于 X 的信息完成的一次变量选择过程。依托于主成分分析得到的主成分变量, 将其与观测值 Y 一起建立回归模型, 则得到主成分回归模型 (principle component regression, PCR)^[80]。然而, 主成分回归中的主成分 (即变量选择的关键) 仅仅依赖于 X 的信息, 而没有用到因变量 Y 的信息。于是, 偏最小二乘回归 (partial least squares, PLS) 模型同时考虑了 X 和 Y 的信息, 即额外考虑了各主成分对 Y 的贡献大小, 来进行变量选择。sPLS 算法则是在 PLS 算法基础上更进一步, 类似于 ENET 算法的思想, 在参数估计时加入了回归系数的惩罚项, 来实现变量选择^[123]。

5.DSA 算法

DSA 算法与逐步回归的思想类似, 模型的初始状态是一个仅有截距项的空模型, 变量集合为一个空集, 接下来的每一步迭代过程, 即是依次开展删除、替代、增加变量的三种操作^[124]。

(1) 删除变量操作, 指的是在当前模型中删除一个变量, 使得均方误差 (root mean squared error, RMSE) 减小最多。RMSE 的定义式如下:

$$RMSE = \sqrt{\frac{\|Y - X\beta\|^2}{n}} \quad (5.9)$$

其中, n 为观测样本量。RMSE 的计算采用交叉检验法来完成。若能找到这样的变量, 则进入下一轮迭代 (变量数减少 1 个), 否则进入替代变量操作环节。

(2) 替代变量操作, 指的是在当前模型中将一个变量替换为另一个未纳入的

变量，使得 RMSE 减小最多。若能实现，则进入下一轮迭代（变量数不变），否则进入增加变量操作环节。

(3) 增加变量操作，指的是在当前模型中加入一个未纳入的变量，使得模型的 RMSE 减小最多，然后进行下一轮迭代过程（变量数增加 1 个）。

当纳入模型的变量个数已经达到用户设定值，或者两步迭代之间的 RMSE 差值小于设定精度，则迭代终止。

5.1.2 变量选择算法的性能比较方法^[81]

1. 暴露数据 X 的生成

本研究选取了居住在五环内的子队列人群作为变量选择算法性能比较的数据集，并计算了其样本 Pearson 相关系数矩阵，确定与该矩阵最接近的正定矩阵 Σ 。然后，将该正定阵作为多元正态分布的协方差矩阵，生成以下多元正态的暴露数据集：

$$X \sim N(0, \Sigma) \quad (5.10)$$

根据相关文献^[81]，生成暴露数据 X 的样本量可选为数据集样本个数的 10%。

2. 结局数据 Y 的生成

代表健康结局的响应变量 Y 由下式生成：

$$Y = \sum_i \beta_i X_i + \varepsilon, \varepsilon \sim N(0, \sigma^2) \quad (5.11)$$

其中， X_i 为所有样本第 i 个环境暴露因素观测值构成的向量， ε 为回归模型的残差。随机选择所有环境暴露因素中的 k 个变量，将这 k 个变量的回归系数 β_i 设为 1，其余变量的回归系数 β_i 设为 0。这些被随机选中的变量被记为“真实变量”（true predictors）， k 可以选择 0、1、2、3、5、10、15 或 25 这几种情形。

3. 主分析（情形 1）

在主分析当中，真实变量之间的相关系数不作任何要求。用 5.1.1 节中的五种变量选择算法分别对上述结局数据 Y 和暴露数据 X 构成的数据集进行拟合，然后利用 8 项指标来评价每种算法的统计学性能。进行真实变量随机选择和模型拟合的整套流程会独立进行 100 次，从而得到统计学性能指标的变化范围。8 项统计学性能评价指标如下：

(1) 灵敏度（sensitivity）：在所有真实变量中，被算法选择出的真实变量的占比。该指标越高代表算法性能越好，其定义式为：

$$Sensitivity = N_{11}/N_{1+} \quad (5.12)$$

其中, N_{11} 为被算法选择出的真实变量个数, N_{1+} 为设定的真实变量个数, 即为 k 。

(2) 可替代的灵敏度 (alternative sensitivity): 为修正版本的灵敏度指标, 其考虑到算法可能选择了一些错误变量, 但这些错误变量与真实变量高度相关。该指标越高代表算法性能越好, 其定义式为:

$$AltSens = \frac{1}{k} \sum_{i \in A} \max_{j \in B} |corr(X_i, X_j)| \quad (5.13)$$

其中, A 为真实变量构成的集合, B 为算法选择的变量集合。

(3) 特异度 (specificity): 在所有错误变量中, 没有被算法选择出的错误变量的占比。该指标越高代表算法性能越好, 其定义式为:

$$Specificity = N_{22}/N_{2+} \quad (5.14)$$

其中, N_{22} 为没有被算法选择出的错误变量个数, N_{2+} 为设定的错误变量个数。

(4) 错误发现率 (false discovery proportion, FDP): 在所有被算法选择出的变量当中错误变量的占比。该指标越低代表算法性能越好, 其定义式为:

$$FDP = N_{21}/N_{+1} \quad (5.15)$$

其中, N_{21} 为被算法选择出的错误变量个数, N_{+1} 为被算法选择出的所有变量个数。

(5) 可替代的错误发现率 (alternative FDP): 为修正版本的 FDP 指标, 其考虑到算法可能选择了一些错误变量, 但这些错误变量与真实变量高度相关。该指标越低代表算法性能越好, 其定义式为:

$$AltFDP = 1 - \frac{1}{n_B} \sum_{j \in B} \max_{i \in A} |corr(X_i, X_j)| \quad (5.16)$$

(6) n_B/k : 算法选择出的变量个数, 与设定的真实变量个数的比值。该指标越接近 1 代表算法性能越好。

(7) 所有变量的平均绝对误差 (mean absolute error, MAE): 所有变量回归系数的拟合值与真实值之差的绝对值求和。该指标越小代表算法性能越好, 其定义式为

$$MAE = \frac{1}{n_{total}} \sum_i |\beta_i - \hat{\beta}_i| \quad (5.17)$$

其中, β_i 为真实的回归系数, $\hat{\beta}_i$ 为算法拟合的回归系数, n_{total} 为变量总个数。

(8) 真实变量的 MAE: 设定的真实变量的回归系数拟合值与真实值之差的

绝对值求和。该指标越小代表算法性能越好，其定义式为

$$MAE\ TP = \frac{1}{k} \sum_{i \in A} |\beta_i - \hat{\beta}_i| \quad (5.18)$$

4. 敏感性分析（情形2至情形7）

敏感性分析的目标是检验上述变量选择算法在不同情形下的性能稳健性。在情形2和情形3当中，设定的真实变量的相关系数被分别限制在（0, 0.2）和（0.5, 1）的范围内。在情形4和5当中，暴露数据的相关系数矩阵 Σ 被分别乘以0.5和2。为了符合相关系数矩阵的定义，情形4当中， Σ 乘以0.5之后，需要保持主对角元素为1；情形5当中， Σ 乘以2之后，需要将大于1的矩阵元素强制设为1。在情形6中，暴露数据不再从多元正态分布中进行随机抽样来生成，而是在原始队列暴露数据中进行重抽样，以探究偏态分布对算法性能的影响。在情形7当中，设定的真实变量的回归系数不再为常数1，而设定为服从一个0.5至1.5之间的均匀分布。每一种情形都需要分别计算100次。

5.1.3 变量选择算法的性能比较结果

变量选择算法的性能比较主分析（情形1）结果如表5.1和图5.1所示。结果表明，对于敏感度与修正后的敏感度指标，五种算法得分都很高，并没有显著差异。但是，对于FDP和修正后的FDP指标，EWAS和EWAS-MLR算法的指标结果显著高于其他三种算法，说明这两种算法选择的变量中错误变量的比例显著高于其他算法。相比之下，ENET和DSA的该两项指标比sPLS更优一些，特别是当 k 比较大的时候。而对于特异度指标，EWAS的指标结果显著低于其他四种算法，也再次说明EWAS会选出太多的错误变量。从 n_B/k 这个指标可以看出，EWAS算法选出的变量个数太多，平均是真实变量个数的19倍。最后，从真实变量的回归系数的拟合偏差上来看，EWAS首先偏差远远大于其他算法，而DSA最佳，EWAS-MLR和sPLS效果也不错，但ENET偏差稍大，这是因为5.1.1节中所述及，ENET算法本身追求的是兼顾估计偏差和方差更小，而无法保证估计完全无偏。

由于变量选择算法更看重变量集合的选择，选完变量后可以再放入正常的多元回归模型中进行参数估计，因此，MAE和MAE TP这两个指标并没有那么重要。总体而言，主分析（情形1）的结果表明，五种算法的性能从高到低排序为DSA、ENET、sPLS、EWAS-MLR、EWAS。该排序也表明，传统研究中常用的EWAS-MLR方法在暴露组学研究中并不是那么可靠。

表 5.1 变量选择算法在主分析中 100 次计算的统计学性能指标

注：数值为均值（最小值，最大值）

	EWAS	EWAS-MLR	ENET	sPLS	DSA
Sensitivity	0.99 (0.98, 1.00)	0.98 (0.96, 1.00)	1.00 (0.99, 1.00)	0.99 (0.98, 1.00)	1.00 (0.99, 1.00)
AltSens	1.00 (0.99, 1.00)	0.99 (0.98, 1.00)	1.00 (0.99, 1.00)	1.00 (0.99, 1.00)	1.00 (1.00, 1.00)
FDP	0.89 (0.69, 0.97)	0.36 (0.13, 0.57)	0.16 (0.00, 0.29)	0.32 (0.10, 0.55)	0.05 (0.01, 0.11)
AltFDP	0.54 (0.29, 0.73)	0.23 (0.05, 0.45)	0.04 (0.00, 0.08)	0.12 (0.06, 0.17)	0.03 (0.01, 0.09)
Specificity	0.27 (0.04, 1.00)	0.96 (0.93, 1.00)	0.96 (0.86, 1.00)	0.88 (0.54, 1.00)	1.00 (0.99, 1.00)
n_B/k	18.74 (3.22, 46.92)	1.88 (1.14, 3.21)	1.23 (1.00, 1.42)	1.72 (1.19, 2.31)	1.10 (1.01, 1.34)
MAE	0.91 (0.07, 2.72)	0.09 (0.00, 0.13)	0.02 (0.00, 0.06)	0.04 (0.00, 0.14)	0.01 (0.00, 0.02)
MAE TP	0.92 (0.04, 2.62)	0.12 (0.09, 0.15)	0.34 (0.17, 0.60)	0.13 (0.05, 0.26)	0.06 (0.04, 0.08)

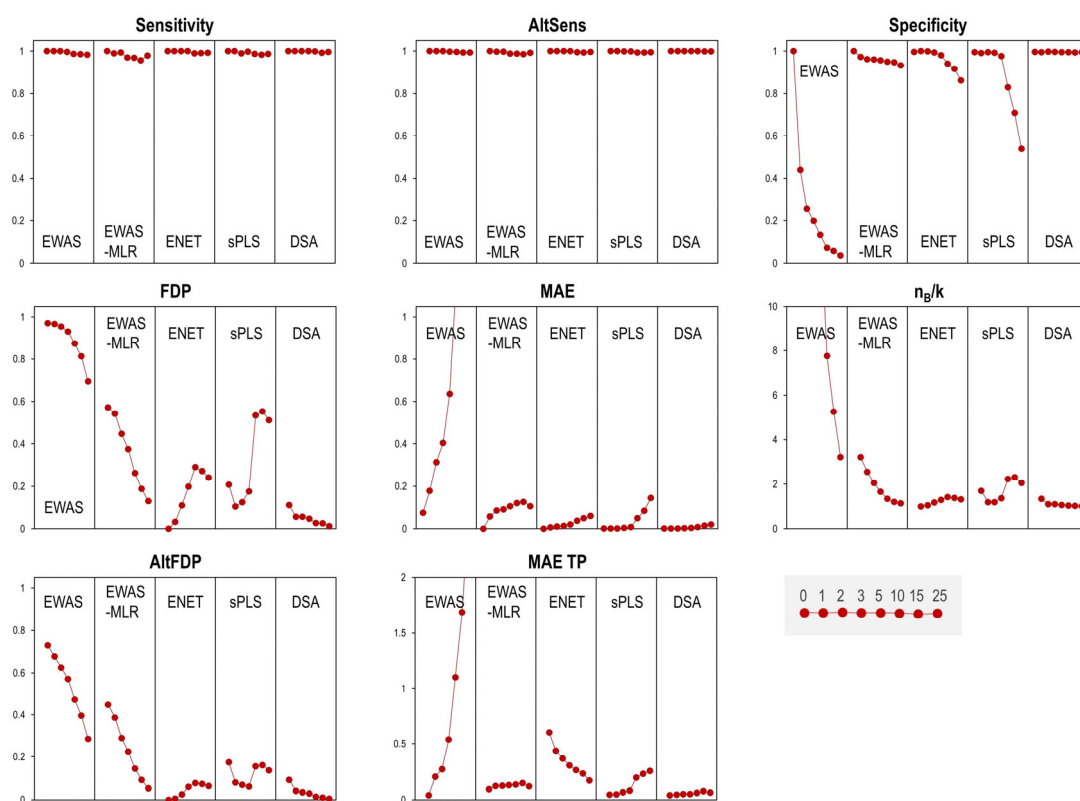


图 5.1 变量选择算法在主分析中 100 次计算的统计学性能指标均值

敏感性分析的结果如图 5.2~图 5.5 所示。图 5.2 展示了情形 2 和情形 3 改变真实变量相关系数范围后的算法性能结果。结果显示，当真实变量相关系数越高时，算法的敏感度指标会下降、FDP 和 MAE 指标会上升。然而，真实变量相关系数的改变，整体上对这些统计学性能指标的影响并不大，特别是对 ENET、sPLS 和 DSA 这三种算法而言。

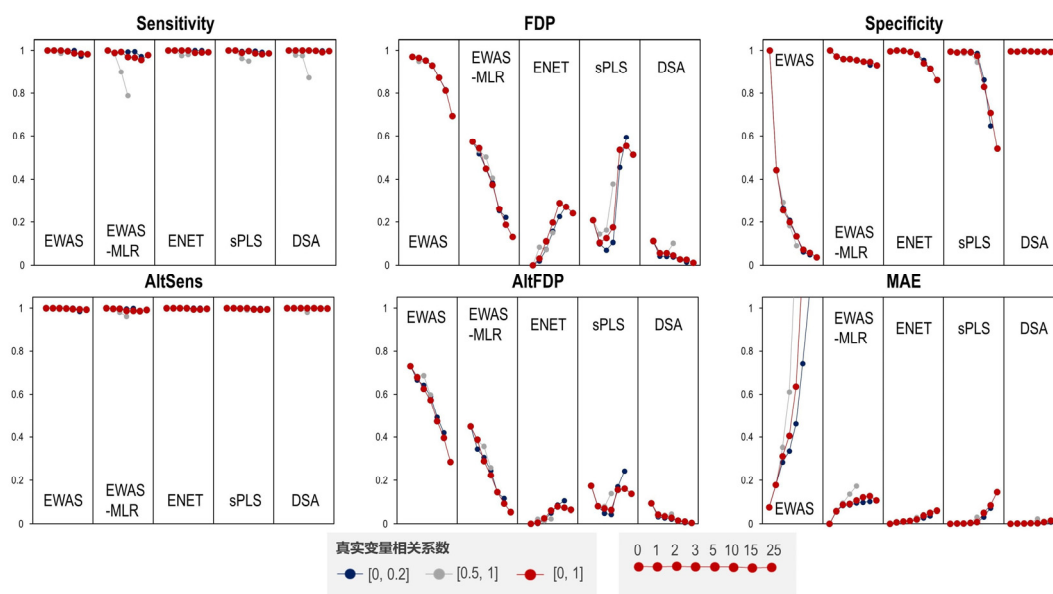


图 5.2 变量选择算法在情形 2 和 3 中 100 次计算的统计学性能指标均值

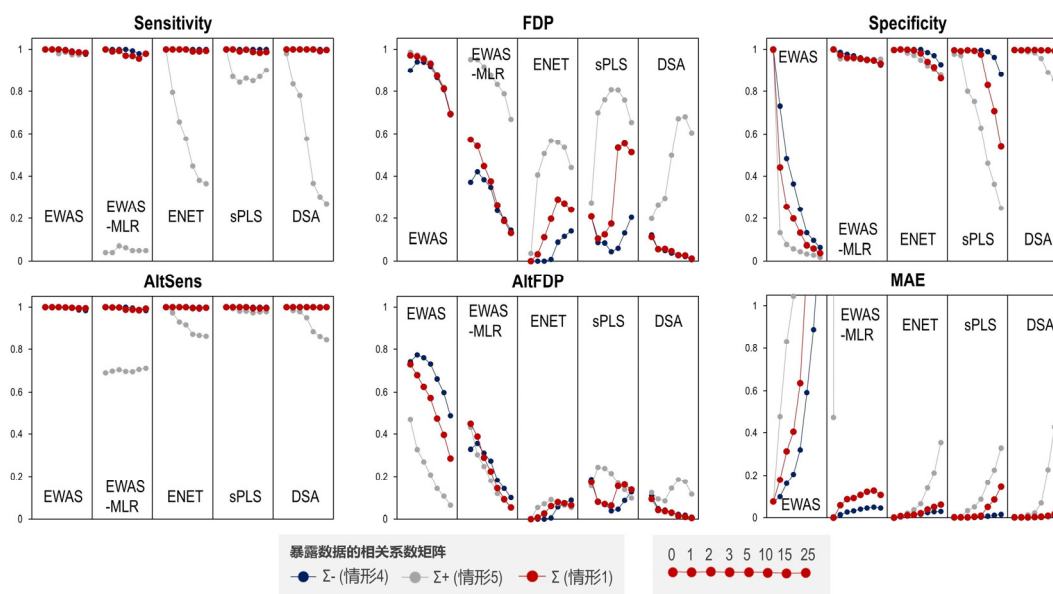


图 5.3 变量选择算法在情形 4 和 5 中 100 次计算的统计学性能指标均值

情形 4 和 5 中暴露变量相关系数矩阵的改变，会对算法性能造成显著影响，如图 5.3 所示。相关系数矩阵中的元素越大，所有指标的统计学性能都会变差。敏感度指标，只有 EWAS 和 sPLS 算法保持相对稳定。所有算法的 FDP 指标在情形 5 中都显著变差，因为情形 5 就是一个变量间高度共线性的情境，但是修正后的 FDP 指标呈现出很好的稳健性。其中，ENET 算法的修正 FDP 结果最优。从特异度来看，EWAS-MLR 和 ENET 算法的变量选择效果也非常稳健。而从 MAE 来看，ENET 和 sPLS 算法表现最优。情形 4 和 5 的讨论，反映的是变量间共线性程度对算法性能的影响，也是暴露组学非常关注的方面，从这两个情形来看，算法性能是 ENET 最优，DSA 和 sPLS 次之，EWAS 和 EWAS-MLR 处于最后位置。

图 5.4 展示了数据偏态分布对算法性能的影响，结果表明，数据是否正态对算法性能的影响很小，因此可以忽略此维度的影响。最后，图 5.5 展示了真实变量回归系数的波动对算法性能的影响。整体而言，回归系数波动对算法性能的影响也可忽略不计，除了 ENET 算法的 FDP 和特异度指标进一步变优了。

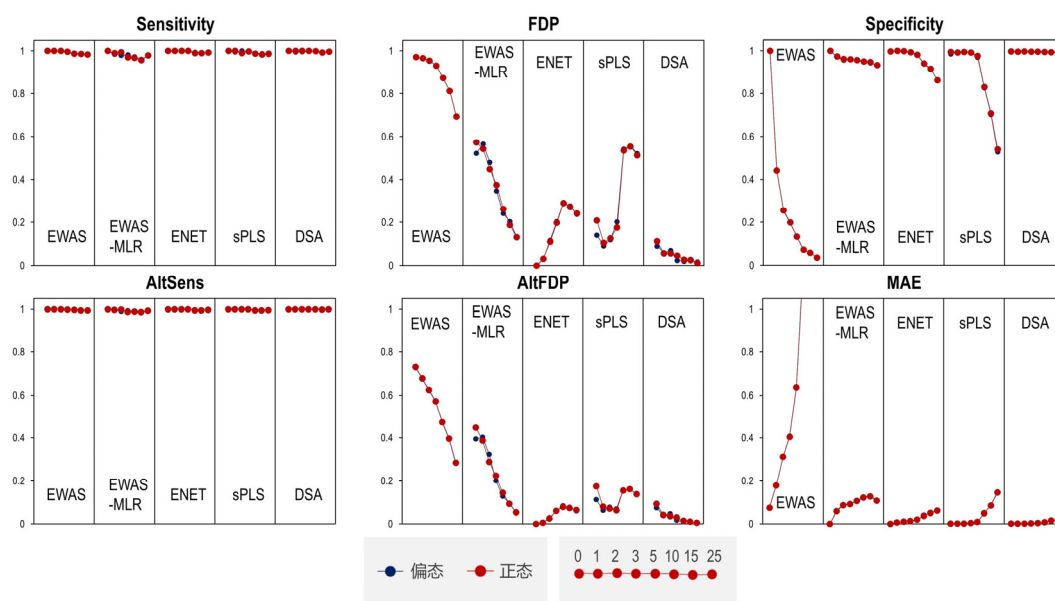


图 5.4 变量选择算法在情形 6 中 100 次计算的统计学性能指标均值

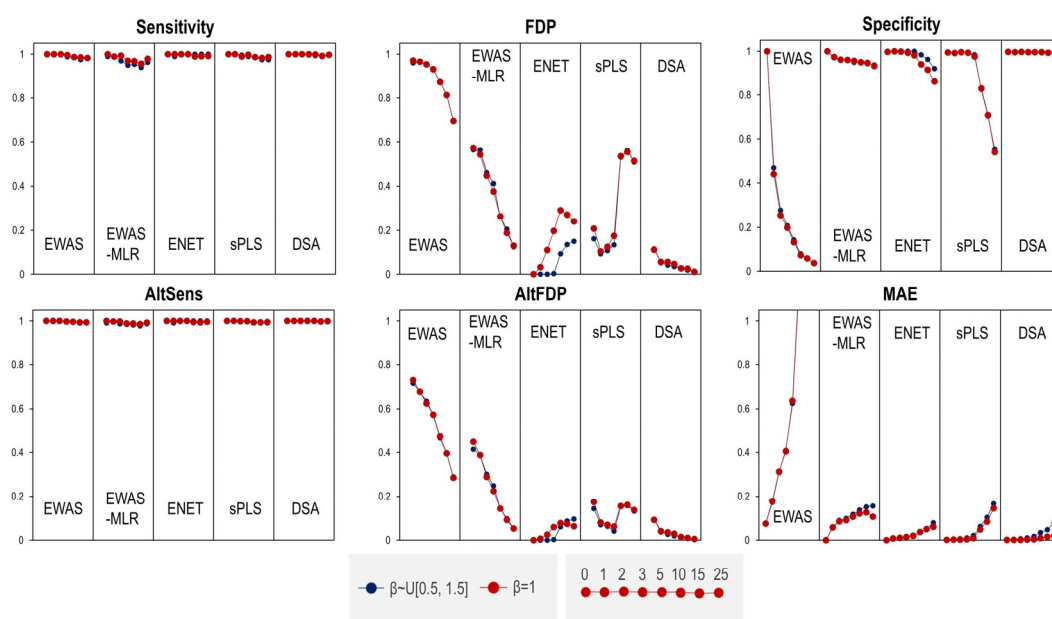


图 5.5 变量选择算法在情形 7 中 100 次计算的统计学性能指标均值

综上所述，DSA 算法在主分析（情形 1）中具有最优的统计学性能，而 ENET 算法在敏感性分析中，特别是多重共线性较为严重的情形 5 中表现最优。此外，在主分析中，ENET 算法的性能与 DSA 其实差距不大。在 R 语言的现有程序包当中，DSA 算法常常能与线性回归模型和广义线性模型相结合，但与 Cox 比例风险模型结合的功能尚缺乏。因此，在后续的统计分析当中，本研究选取 ENET 变量选择算法，与生存分析的 Cox 比例风险模型进行嵌套使用，来分析建成环境暴露与 AMI 复发的关联性。

5.2 统计分析方法

5.2.1 暴露变量、协变量、结局变量的选择

本研究用到的建成环境暴露变量为 4.2 节列出的 6 大类（密度、多样性、设计、到公共交通的距离、目的地可达性、空气污染）共计 26 种环境暴露因素，详见表 4.4。其暴露水平的测度考虑空间和时间两个维度——空间尺度上，每种建成环境暴露因素均测度患者住址周边 300 米、500 米、1000 米缓冲区内的暴露水平；时间尺度上，既测度患者首次患得 AMI 康复当年的基线暴露水平，又测度患者从康复到首次复发之间多年的累积平均暴露水平。

本研究选择的协变量及选择原因在 3.1.3 节已经作了详细说明。协变量包括患者个人属性（年龄、性别和婚姻状况）、患者的相关并发症（高血压、高血脂、

糖尿病、中风、心力衰竭、心房颤动和肾功能障碍），以及北京市交通分析小区尺度的平均教育年限、年均家庭收入这两个社会经济属性指标。

本研究选择的健康结局变量在 3.1.2 节已经做了详细解释。本研究首要关注的健康结局为所有急性心肌梗死的首次复发事件，次要健康结局为致死、非致死急性心肌梗死的首次复发事件，并根据每个个体的随访开始、结束时间，计算其随访时长，根据随访结束时是否复发判断数据是否删失。

5.2.2 生存分析的基本原理^[125]

生存数据（survival data 或 time-to-event data）是在随访期间关注某事件是否发生而记录的一类数据，包括从随访开始到随访结束的时长（连续变量）和随访结束时事件是否发生（分类变量）这一组变量。由于随访成本、个人因素等原因，许多研究对象在随访结束时并未发生关注事件，因此这类数据也被称为删失数据（censored data）。生存分析（survival analysis）是一系列针对生存数据的统计分析方法。

生存分析关心的是研究人群的生存函数（survivor function），其定义为生存时间大于 t 的概率，即

$$S(t) = P(T > t) = 1 - F(t) = \int_t^{\infty} f(t) dt \quad (5.19)$$

其中， $S(t)$ 为生存函数， $F(t)$ 为生存时间的累积分布函数（cumulative distribution function, CDF）， $f(t)$ 为生存时间的概率密度函数（probability density function, PDF）。

除了生存函数之外，生存分析往往还关注风险函数（hazard function），即在生存时间能够到达 t 的前提条件下，个体在 t 时刻瞬间“死亡”的风险，其公式为

$$h(t) = \lim_{h \rightarrow 0^+} \frac{P(t \leq T < t+h | T \geq t)}{h} = \frac{f(t)}{S(t)} = -\frac{d \ln S(t)}{dt} \quad (5.20)$$

其中， $h(t)$ 为风险函数。

生存分析中最经典的 Cox 比例风险模型（Cox proportional hazard model），即采用多个解释变量 X 与风险函数去进行回归建模，来分析各个因素对事件发生风险的影响。Cox 比例风险模型的形式为

$$h(t) = h_0(t) \exp \left(\sum_i \beta_i X_i \right) \quad (5.21)$$

其中， $h_0(t)$ 为所有 X 均为 0 时的基准风险。Cox 比例风险模型有两条重要的假设：第一，风险函数与基准风险的比值不随时间变化，即各个协变量对结局的影响不随时间变化，称为比例风险假定（proportional hazard assumption, 简记为 PH 假定）；

第二，模型中协变量与风险函数呈现对数线性关系，称为对数线性假定。本研究采用 Cox 比例风险模型来对 AMI 复发的生存数据进行分析。

5.2.3 统计分析方法与流程

本研究在进行建成环境暴露与 AMI 复发的关联分析时，采用的统计分析详细步骤如下：

（1）单因素分析：如同大部分暴露组学研究，在进行变量选择之前，先进行单因素分析（即 EWAS 方法），将每个建成环境暴露因素与所有 AMI 复发事件的生存时间分别建立 Cox 比例风险模型，并调整相应协变量，最后对各个变量对应的回归系数 p 值进行 BH 方法校正，校正 p 值小于 0.05 者认为显著，并汇报单因素分析结果^[83-85]。

（2）变量选择：将第（1）步中筛选出的建成环境暴露因素纳入含有多个自变量的 Cox 比例风险模型，结局事件为所有 AMI 的首次复发事件，并嵌套使用 ENET 变量选择的算法，调整相应协变量。ENET 算法为 5.1 节性能比较选出的最优的变量选择算法，其中包含两个待定参数——总惩罚系数、岭回归与 LASSO 回归惩罚的比例系数，它们由 10 折交叉检验来实现参数估计，交叉检验的优化目标是让模型的 RMSE 最小化。为了避免过拟合现象，最优参数由 RMSE 在其最小值附近 1 个标准误区间内的众多模型中变量最稀疏（即变量数最少）的那个模型来确定^[81]。

（3）重复变量剔除：对于 ENET 算法筛选出的变量，如果存在同一建成环境指标在不同缓冲区内的结果都被选出的情形（比如 300 米和 1000 米缓冲区内的蔬果店密度都被算法选出），保留单因素 p 值较小的那个变量。

（4）首要结局关联分析：将第（3）步中最终确定的建成环境暴露因素纳入多因素的 Cox 比例风险模型中，结局事件为所有 AMI 复发事件，并对协变量进行调整。回归系数 p 值小于 0.05 的变量（无需校正）被认为是对所有 AMI 复发有显著影响的建成环境暴露因素。关联效应采用暴露因素每升高一个 IQR 水平的危险比 HR（hazard ratio）及其 95% 置信区间。第 i 个暴露因素的 HR 可表达为

$$HR_i = \exp(\beta_i) \quad (5.22)$$

（5）次要结局关联分析：将第（3）步中最终确定的建成环境暴露因素纳入多因素的 Cox 比例风险模型中，结局事件选择致死型 AMI 复发事件和非致死型 AMI 复发事件，分别进行模型拟合，并都对协变量进行调整。回归系数 p 值小于 0.05 的变量（无需校正）被认为是对致死型/非致死型 AMI 复发有显著影响的建成环境暴露因素。

(6) 模型检验：每个选择出的建成环境暴露变量是否满足比例风险假定，需采用 Schoenfeld 残差检验来诊断。若某个暴露变量不满足比例风险假定，则对该变量采用含时变系数的 Cox 模型，即添加该自变量与对数时间的交互项^[126-129]：

$$h(t) = h_0(t) \exp \left(\sum_i \beta_i X_i + \gamma_k \ln(t) X_k \right) \quad (5.23)$$

其中， k 代表第 k 个变量不满足比例风险假定， γ_k 为交互项的回归系数。此时，变量 X_k 的 HR 变成一个随时间变化的函数：

$$HR_k(t) = \exp(\beta_k + \gamma_k \ln(t)) \quad (5.24)$$

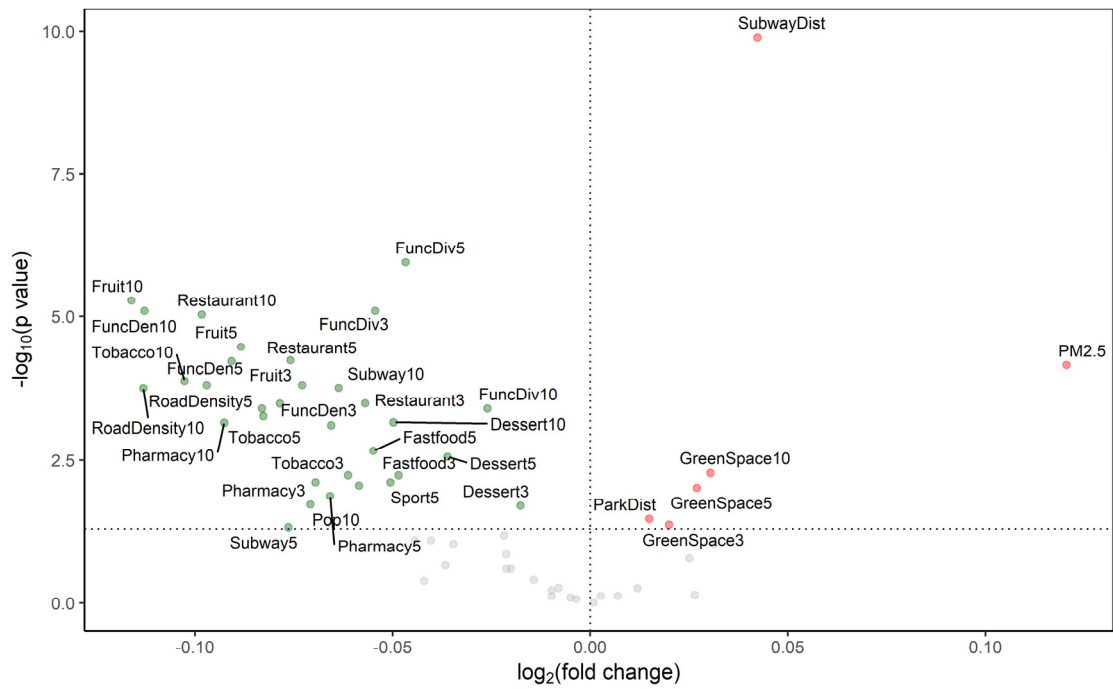
而对数线性假定采用绘制 Martingale 残差图来进行图像化检验，并在残差图中添加数据点的局部加权回归平滑曲线（locally weighted scatterplot smoothing, Lowess），若 Lowess 曲线与直线比较接近，则认为对数线性假定近似成立。

以上所有的统计分析均在统计软件 R 语言（4.1.0 版本）中完成。采用 ENET 算法进行 Cox 模型的变量选择过程使用的是 glmnet 包中的 glmnet 函数；单因素及多因素的 Cox 模型采用 survival 包中的 coxph 函数；Schoenfeld 残差检验使用 survival 包中的 cox.zph 函数；Martingale 残差图使用 survminer 包中的 ggcoxfunctional 函数来实现绘制。

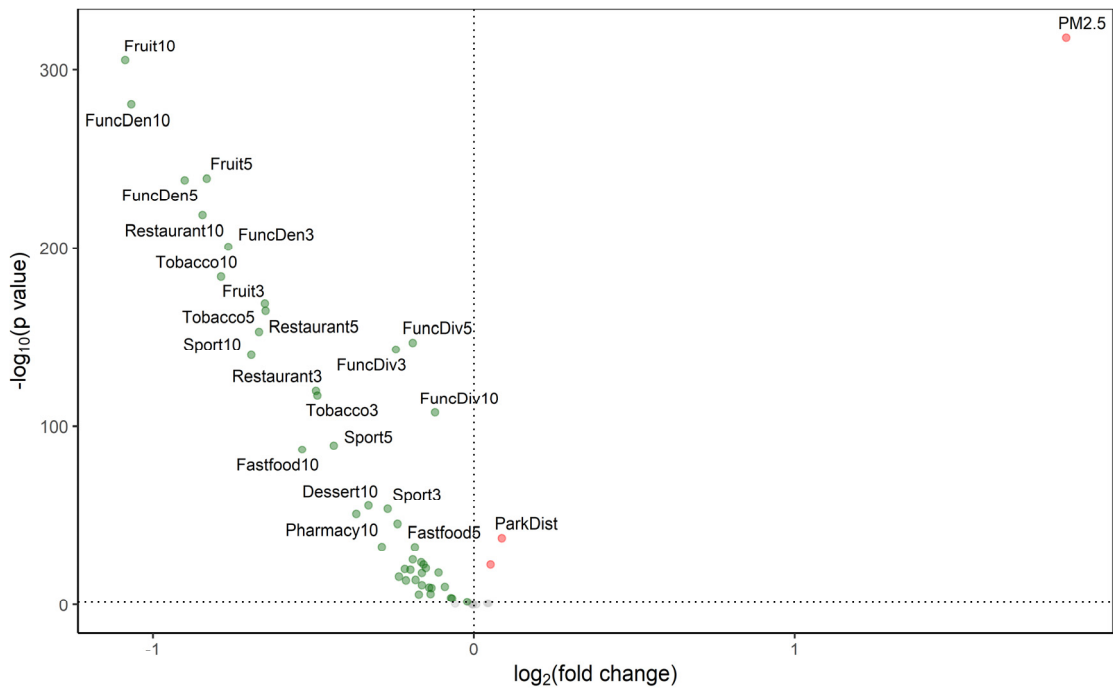
5.3 单因素关联分析结果

图 5.6 采用火山图形式，展示了利用 EWAS 方法进行单因素 Cox 比例风险模型的分析结果。纵坐标为负的 p 值的以 10 为底的对数，横坐标为 HR 的以 2 为底的对数值。在水平虚线以上的建成环境暴露因素表明其在单因素分析中为显著因素，在垂直虚线以左和以右的建成环境暴露因素分别代表保护因素和风险因素。

对于基线暴露来说，涉及 15 种建成环境暴露因素的 38 个变量均呈现出显著性（ $p < 0.05$ ），其中，PM_{2.5}、到地铁站的距离、到公园的距离和绿地比例与 AMI 复发风险的提高显著正相关。对于累积平均暴露水平来说，涉及 16 种建成环境暴露因素的 47 个变量均呈现出显著性，其中，PM_{2.5}、到地铁站的距离与 AMI 复发风险的提高显著正相关。上述结果表明，在单因素关联分析中，统计学上具有显著意义的建成环境暴露因素占到了绝大多数，这与 5.1 节中 EWAS 算法性能评价结果相符。单因素关联分析的具体结果如表 5.2 所示。



(a)



(b)

图 5.6 单因素分析结果的火山图：(a) 基线暴露，(b) 累积平均暴露

注： p 值已经采用 BH 方法校正，水平虚线代表 $p=0.05$ 的阈值；英文缩写对应指标详见图 4.31

表 5.2 单因素分析的 HR 值及 95%置信区间

暴露因素	基线暴露水平			累积平均暴露水平		
	HR (95% CI)	p 值	IQR	HR (95% CI)	p 值	IQR
人口密度 (人/km²)						
300 m 缓冲区	0.972 (0.945 ~ 1.000)	0.054	19482	0.908 (0.882 ~ 0.935)	< 0.001	20347
500 m 缓冲区	0.970 (0.940 ~ 1.000)	0.053	19832	0.893 (0.865 ~ 0.922)	< 0.001	20680
1000 m 缓冲区	0.952 (0.917 ~ 0.989)	0.011	20608	0.850 (0.818 ~ 0.883)	< 0.001	21538
城市功能密度 (个)						
300 m 缓冲区	0.956 (0.932 ~ 0.979)	< 0.001	373	0.588 (0.568 ~ 0.609)	< 0.001	430
500 m 缓冲区	0.939 (0.913 ~ 0.965)	< 0.001	1055	0.535 (0.516 ~ 0.556)	< 0.001	1225
1000 m 缓冲区	0.925 (0.897 ~ 0.954)	< 0.001	4091	0.477 (0.458 ~ 0.497)	< 0.001	4857
土地利用混合度 (-)						
300 m 缓冲区	1.005 (0.981 ~ 1.030)	0.698	0.45	NA	NA	NA
500 m 缓冲区	0.997 (0.973 ~ 1.020)	0.774	0.39	NA	NA	NA
1000 m 缓冲区	0.998 (0.975 ~ 1.021)	0.832	0.33	NA	NA	NA
城市功能混合度 (-)						
300 m 缓冲区	0.963 (0.949 ~ 0.977)	< 0.001	0.46	0.845 (0.834 ~ 0.856)	< 0.001	0.34
500 m 缓冲区	0.968 (0.957 ~ 0.979)	< 0.001	0.31	0.876 (0.867 ~ 0.885)	< 0.001	0.23
1000 m 缓冲区	0.982 (0.973 ~ 0.991)	< 0.001	0.20	0.919 (0.912 ~ 0.926)	< 0.001	0.13
路网密度 (km/km²)						
300 m 缓冲区	0.960 (0.934 ~ 0.987)	0.004	5.3	0.912 (0.887 ~ 0.939)	< 0.001	5.0
500 m 缓冲区	0.935 (0.906 ~ 0.965)	< 0.001	5.2	0.881 (0.854 ~ 0.910)	< 0.001	5.0
1000 m 缓冲区	0.925 (0.890 ~ 0.960)	< 0.001	5.7	0.863 (0.832 ~ 0.896)	< 0.001	5.5
到主干道的距离 (m)						
	0.994 (0.980 ~ 1.009)	0.459	240.2	NA	NA	NA
绿地比例 (-)						
300 m 缓冲区	1.014 (1.002 ~ 1.026)	0.026	0.08	NA	NA	NA
500 m 缓冲区	1.019 (1.006 ~ 1.032)	0.005	0.10	NA	NA	NA
1000 m 缓冲区	1.021 (1.008 ~ 1.035)	0.002	0.11	NA	NA	NA
NDVI (-)						
300 m 缓冲区	1.008 (0.986 ~ 1.031)	0.472	0.12	0.901 (0.882 ~ 0.921)	< 0.001	0.10
500 m 缓冲区	1.018 (0.995 ~ 1.040)	0.125	0.12	0.897 (0.878 ~ 0.916)	< 0.001	0.10
1000 m 缓冲区	1.022 (0.999 ~ 1.046)	0.064	0.13	0.876 (0.855 ~ 0.898)	< 0.001	0.12
地铁站密度 (个)						
300 m 缓冲区	1.018 (0.942 ~ 1.101)	0.646	0	0.960 (0.888 ~ 1.038)	0.308	0
500 m 缓冲区	0.948 (0.904 ~ 0.995)	0.03	0	0.888 (0.846 ~ 0.932)	< 0.001	0
1000 m 缓冲区	0.957 (0.937 ~ 0.977)	< 0.001	1	0.893 (0.871 ~ 0.916)	< 0.001	1.17
到地铁站的距离(m)						
	1.030 (1.021 ~ 1.038)	< 0.001	3849.7	1.036 (1.029 ~ 1.044)	< 0.001	3323.36

续表 5.2 单因素分析的HR值及95%置信区间

暴露因素	基线暴露水平			累积平均暴露水平		
	HR (95% CI)	p 值	IQR	HR (95% CI)	p 值	IQR
公交站密度 (个)						
300 m 缓冲区	0.993 (0.973 ~ 1.014)	0.524	2	0.985 (0.974 ~ 0.997)	0.016	1
500 m 缓冲区	0.986 (0.965 ~ 1.007)	0.201	4	0.954 (0.931 ~ 0.978)	< 0.001	4
1000 m 缓冲区	0.985 (0.963 ~ 1.008)	0.202	12	0.952 (0.928 ~ 0.976)	< 0.001	12
公园密度 (个)						
300 m 缓冲区	0.975 (0.940 ~ 1.011)	0.168	0	0.910 (0.878 ~ 0.944)	< 0.001	0
500 m 缓冲区	0.985 (0.968 ~ 1.003)	0.104	1	0.939 (0.922 ~ 0.957)	< 0.001	1
1000 m 缓冲区	0.976 (0.952 ~ 1.002)	0.066	4	0.871 (0.847 ~ 0.897)	< 0.001	4
到公园的距离 (m)	1.010 (1.002 ~ 1.019)	0.02	768.5	1.062 (1.052 ~ 1.071)	< 0.001	672.6
到大型绿地的距离(m)	0.990 (0.971 ~ 1.010)	0.318	287.3	NA	NA	NA
烟酒店密度 (个)						
300 m 缓冲区	0.958 (0.932 ~ 0.985)	0.003	10	0.713 (0.693 ~ 0.734)	< 0.001	10
500 m 缓冲区	0.944 (0.916 ~ 0.973)	< 0.001	24	0.628 (0.607 ~ 0.650)	< 0.001	27
1000 m 缓冲区	0.931 (0.901 ~ 0.963)	< 0.001	85	0.579 (0.558 ~ 0.601)	< 0.001	95
餐厅密度 (个)						
300 m 缓冲区	0.967 (0.946 ~ 0.988)	0.003	39	0.711 (0.691 ~ 0.731)	< 0.001	47
500 m 缓冲区	0.963 (0.941 ~ 0.984)	< 0.001	106	0.637 (0.617 ~ 0.658)	< 0.001	123
1000 m 缓冲区	0.947 (0.921 ~ 0.973)	< 0.001	409	0.556 (0.536 ~ 0.577)	< 0.001	458
快餐店密度 (个)						
300 m 缓冲区	0.961 (0.942 ~ 0.980)	< 0.001	0	0.892 (0.873 ~ 0.912)	< 0.001	1
500 m 缓冲区	0.949 (0.927 ~ 0.971)	< 0.001	2	0.848 (0.829 ~ 0.867)	< 0.001	2
1000 m 缓冲区	0.934 (0.909 ~ 0.960)	< 0.001	6	0.690 (0.665 ~ 0.716)	< 0.001	8
甜品/饮品/糕饼店密度 (个)						
300 m 缓冲区	0.988 (0.978 ~ 0.997)	0.011	2	0.926 (0.911 ~ 0.942)	< 0.001	3
500 m 缓冲区	0.975 (0.961 ~ 0.990)	0.001	7	0.880 (0.862 ~ 0.899)	< 0.001	8
1000 m 缓冲区	0.966 (0.948 ~ 0.984)	< 0.001	26	0.796 (0.774 ~ 0.819)	< 0.001	32
蔬果店密度 (个)						
300 m 缓冲区	0.951 (0.928 ~ 0.974)	< 0.001	4	0.636 (0.616 ~ 0.657)	< 0.001	5
500 m 缓冲区	0.941 (0.916 ~ 0.965)	< 0.001	9	0.561 (0.543 ~ 0.581)	< 0.001	11
1000 m 缓冲区	0.923 (0.895 ~ 0.951)	< 0.001	32	0.471 (0.452 ~ 0.490)	< 0.001	39
运动场馆密度 (个)						
300 m 缓冲区	0.985 (0.971 ~ 0.999)	0.042	2	0.830 (0.811 ~ 0.850)	< 0.001	3
500 m 缓冲区	0.966 (0.943 ~ 0.989)	0.004	7	0.739 (0.717 ~ 0.761)	< 0.001	8
1000 m 缓冲区	0.944 (0.917 ~ 0.972)	< 0.001	27	0.618 (0.595 ~ 0.642)	< 0.001	32

续表 5.2 单因素分析的HR值及95%置信区间

暴露因素	基线暴露水平			累积平均暴露水平		
	HR (95% CI)	<i>p</i> 值	IQR	HR (95% CI)	<i>p</i> 值	IQR
综合医院密度 (个)						
300 m 缓冲区	0.993 (0.957 ~ 1.031)	0.721	1	1.029 (0.991 ~ 1.068)	0.135	1
500 m 缓冲区	1.001 (0.957 ~ 1.046)	0.98	1	1.032 (0.987 ~ 1.080)	0.168	1
1000 m 缓冲区	0.971 (0.915 ~ 1.031)	0.337	0	0.998 (0.940 ~ 1.059)	0.94	0
到综合医院的距离 (m)	1.002 (0.993 ~ 1.011)	0.699	338.6	0.991 (0.981 ~ 1.000)	0.062	347.0
药房密度 (个)						
300 m 缓冲区	0.953 (0.922 ~ 0.984)	0.004	2	0.861 (0.834 ~ 0.888)	< 0.001	2
500 m 缓冲区	0.955 (0.924 ~ 0.988)	0.007	4	0.819 (0.793 ~ 0.846)	< 0.001	4
1000 m 缓冲区	0.938 (0.906 ~ 0.971)	< 0.001	12	0.775 (0.750 ~ 0.801)	< 0.001	12
PM _{2.5} (μg/m ³)	1.087 (1.047 ~ 1.128)	< 0.001	32.2	3.594 (3.502 ~ 3.689)	< 0.001	17.0

注：HR 及 95%CI 均是各个变量每升高 1 个 IQR 水平下的，个别 IQR=0 的变量则报告 IQR=1 的 HR 及 95%CI。表中的 *p* 值为未校正的版本。

5.4 多因素关联分析结果

5.4.1 所有 AMI 复发事件的关联分析结果

通过 ENET 算法的变量选择，基线暴露水平下共选出了 8 个变量，包括 500 米缓冲区内的城市功能多样性、500 米缓冲区内的路网密度、1000 米缓冲区内的绿地比例、到地铁站的距离、300 米缓冲区内的快餐店密度、300 米缓冲区内的蔬果店密度、300 米缓冲区内的药房密度和 PM_{2.5}。多因素 Cox 模型的分析结果如表 5.3 所示。经过检验，到地铁站的距离和 PM_{2.5} 不符合比例风险假定 ($p < 0.05$)，大部分变量符合对数线性假定。对于那些符合比例风险假定的变量，Cox 比例风险模型和含时变系数的 Cox 模型得到的 HR 值计算结果非常接近；而对于那些不符合比例风险假定的变量，表 5.3 中分别给出了 1 个月、3 个月、6 个月、1 年、2 年、3 年和 5 年时的 HR 值。从表 5.3 中可以得出，更高的城市功能混合度与 AMI 复发风险的降低呈现显著相关性 (HR = 0.986, 95% CI: 0.972~0.999)。在 Cox 比例风险模型当中，AMI 复发风险的增加与更远的到地铁站距离 (HR = 1.030, 95% CI: 1.021~1.040)、更高的年均 PM_{2.5} 浓度 (HR = 1.136, 95% CI: 1.090~1.184) 显著相关。然而，当考虑时变系数的效应之后，到地铁站距离的 HR 值随着时间不断减弱，仅在 3 个月内的短期阶段对 AMI 复发有显著影响；而 PM_{2.5} 对 AMI 复发的影响仍然具有显著意义，且效应值随时间不断增强。

累积平均暴露水平下共选出了 5 个变量，包括 500 米缓冲区内的城市功能多

样性、到地铁站的距离、到公园的距离、1000米缓冲区内的蔬果店密度和PM_{2.5}。结果表明,AMI复发风险与500米缓冲区内的城市功能混合度(HR = 0.961, 95% CI: 0.949~0.974)、到地铁站距离(HR = 1.115, 95% CI: 1.106~1.124)以及PM_{2.5}(HR = 3.743, 95% CI: 3.643~3.845)之间存在显著关联。此外,有其他两个建成环境暴露因素被识别出与AMI复发显著相关,分别是到公园的距离(HR = 1.045, 95% CI: 1.032-1.057)和1000米缓冲区内的蔬果店密度(HR = 0.665, 95% CI: 0.636-0.696)。除了到公园距离这个因素之外,其余四个建成环境暴露因素均不符合比例风险假定。但是,含时变系数的Cox模型分析结果表明,除了PM_{2.5}的HR随时间显著升高之外,其余变量的HR值随时间变化幅度很小,与Cox比例风险模型得出的结果极为相似。不论是基线暴露还是累积平均暴露,PM_{2.5}的HR都会随时间明显上升,表明PM_{2.5}暴露对急性心肌梗死患者的健康影响可能具有累积效应。

此外,许多EWAS算法中识别出的显著变量,在ENET算法中并未被选择出来,或者即使被选择出来,在最后的多元分析中也不显著,再次表明传统的流行病学关联分析方法可能会导致较多环境暴露因素的“假阳性”现象。

综上所述,本研究发现,更高的AMI复发风险与更低的城市功能混合度、更远的到地铁站距离、更远的到公园距离、更低的蔬果店密度及更高的PM_{2.5}年均浓度之间存在显著关联关系。

5.4.2 致死型、非致死型AMI复发事件的关联分析结果

本节分析次要结局——致死型和非致死型AMI复发事件的多因素关联分析结果,相关结果如表5.3和表5.4所示。

先分析基线暴露水平。在Cox比例风险模型中,到地铁站的距离为致死型和非致死型复发事件的显著风险因素,该结论与所有复发事件的主要结局中的结论是一致的。然而,考虑时变系数效应之后,到地铁站的距离仅在短期3个月内与非致死型AMI复发事件显著关联。在考虑时变系数下,PM_{2.5}也仅与致死型AMI复发事件呈现显著相关,与非致死型复发结局关系并不显著。除此之外,500米缓冲区内的城市功能混合度越高,与致死型AMI复发风险越低呈显著相关(HR = 0.945, 95% CI: 0.923~0.967)。

其次分析累积平均暴露水平。从表5.4中可以得出,致死型复发风险的增加与500米缓冲区内城市功能混合度的降低以及年均PM_{2.5}浓度的升高显著相关;而非致死型复发风险的增加,则与更远的到地铁站距离(3个月内)及更高的年均PM_{2.5}浓度显著相关。

表 5.3 建成环境基线暴露水平与所有、致死型、非致死型 AMI 复发事件之间的 HR 值及 95%置信区间

变量	缓冲区	IQR	所有 AMI 复发事件		致死型 AMI 复发事件		非致死型 AMI 复发事件	
			HR (95% CI) ^a	p 值	HR (95% CI) ^a	p 值	HR (95% CI) ^a	p 值
城市功能混合度 (-)	500 m	0.314	0.986 (0.972 ~ 0.999)^b	0.037	0.945 (0.923 ~ 0.967)^b	< 0.001	0.999 (0.983 ~ 1.015)	0.887
路网密度 (km/km ²)	500 m	5.232	0.970 (0.933 ~ 1.008)	0.124	0.988 (0.921 ~ 1.060)	0.729	0.960 (0.919 ~ 1.004)	0.073
绿地比例 (-)	1000 m	0.110	1.009 (0.994 ~ 1.025)	0.222	1.009 (0.981 ~ 1.037)	0.539	1.012 (0.995 ~ 1.030)	0.169
到地铁站的距离 (km)	——	3.849	1.030 (1.021 ~ 1.040)^c	< 0.001	1.030 (1.013 ~ 1.048)^c	< 0.001	1.028 (1.018 ~ 1.038)^c	< 0.001
1 个月			1.038 (1.018 ~ 1.058)^b	< 0.001	1.037 (0.995 ~ 1.081)	0.087	1.036 (1.014 ~ 1.059)^b	0.001
3 个月			1.028 (1.006 ~ 1.050)^b	0.012	1.027 (0.980 ~ 1.075)	0.262	1.026 (1.002 ~ 1.051)^b	0.037
6 个月			1.022 (0.998 ~ 1.046)	0.070	1.020 (0.971 ~ 1.072)	0.424	1.020 (0.994 ~ 1.046)	0.143
1 年			1.015 (0.990 ~ 1.041)	0.234	1.014 (0.962 ~ 1.069)	0.610	1.013 (0.985 ~ 1.042)	0.363
2 年			1.009 (0.982 ~ 1.036)	0.510	1.008 (0.952 ~ 1.066)	0.795	1.007 (0.977 ~ 1.037)	0.664
3 年			1.005 (0.978 ~ 1.034)	0.705	1.004 (0.947 ~ 1.064)	0.898	1.003 (0.972 ~ 1.035)	0.854
5 年			1.001 (0.972 ~ 1.031)	0.955	0.999 (0.940 ~ 1.063)	0.980	0.998 (0.966 ~ 1.031)	0.918
快餐店密度 (个)	300 m	1	0.988 (0.965 ~ 1.011)	0.306	0.980 (0.939 ~ 1.022)	0.348	0.991 (0.965 ~ 1.018)	0.500
蔬果店密度 (个)	300 m	4	0.996 (0.968 ~ 1.025)	0.800	1.014 (0.962 ~ 1.069)	0.599	0.988 (0.956 ~ 1.021)	0.470
药店密度 (个)	300 m	2	0.994 (0.957 ~ 1.033)	0.773	0.969 (0.903 ~ 1.040)	0.386	1.020 (0.976 ~ 1.065)	0.378
PM _{2.5} (μg/m ³)	——	32.200	1.136 (1.090 ~ 1.184)^c	< 0.001	1.330 (1.232 ~ 1.436)^c	< 0.001	1.107 (1.056 ~ 1.161)^c	< 0.001
1 个月			1.116 (1.007 ~ 1.236)^b	0.036	1.222 (0.983 ~ 1.520)	0.071	1.093 (0.975 ~ 1.226)	0.127
3 个月			1.148 (1.023 ~ 1.289)^b	0.019	1.308 (1.025 ~ 1.669)^b	0.031	1.111 (0.976 ~ 1.264)	0.112
6 个月			1.169 (1.032 ~ 1.325)^b	0.014	1.365 (1.051 ~ 1.773)^b	0.020	1.122 (0.976 ~ 1.291)	0.107
1 年			1.191 (1.041 ~ 1.363)^b	0.011	1.426 (1.076 ~ 1.889)^b	0.013	1.134 (0.974 ~ 1.319)	0.104
2 年			1.213 (1.049 ~ 1.402)^b	0.009	1.488 (1.101 ~ 2.011)^b	0.010	1.145 (0.973 ~ 1.348)	0.103
3 年			1.226 (1.053 ~ 1.426)^b	0.008	1.526 (1.115 ~ 2.087)^b	0.008	1.152 (0.972 ~ 1.365)	0.103
5 年			1.242 (1.059 ~ 1.456)^b	0.008	1.574 (1.133 ~ 2.188)^b	0.007	1.161 (0.971 ~ 1.388)	0.103

表 5.4 建成环境累积平均暴露水平与所有、致死型、非致死型 AMI 复发事件之间的 HR 值及 95%置信区间

变量	缓冲区	IQR	所有 AMI 复发事件		致死型 AMI 复发事件		非致死型 AMI 复发事件	
			HR (95% CI) ^a	<i>p</i> 值	HR (95% CI) ^a	<i>p</i> 值	HR (95% CI) ^a	<i>p</i> 值
城市功能混合度 (-)	500 m	0.226	0.961 (0.949 ~ 0.974)^c	< 0.001	0.951 (0.925 ~ 0.978)^b	< 0.001	0.995 (0.978 ~ 1.013)	0.594
1 个月			0.966 (0.938 ~ 0.995)^b	0.023				
3 个月			0.955 (0.924 ~ 0.988)^b	0.007				
6 个月			0.948 (0.915 ~ 0.983)^b	0.004				
1 年			0.942 (0.906 ~ 0.979)^b	0.002				
2 年			0.935 (0.897 ~ 0.974)^b	0.001				
3 年			0.931 (0.892 ~ 0.972)^b	0.001				
5 年			0.926 (0.885 ~ 0.969)^b	< 0.001				
到地铁站的距离 (km)	——	3.323	1.115 (1.106 ~ 1.124)^c	< 0.001	1.020 (1.004 ~ 1.036)^c	0.012	1.025 (1.015 ~ 1.034)^c	< 0.001
1 个月			1.122 (1.100 ~ 1.143)^b	< 0.001	1.035 (0.996 ~ 1.076)	0.082	1.031 (1.013 ~ 1.049)^b	< 0.001
3 个月			1.121 (1.097 ~ 1.146)^b	< 0.001	1.022 (0.979 ~ 1.067)	0.313	1.021 (1.001 ~ 1.041)^b	0.042
6 个月			1.121 (1.095 ~ 1.148)^b	< 0.001	1.014 (0.969 ~ 1.062)	0.54	1.014 (0.993 ~ 1.036)	0.184
1 年			1.121 (1.093 ~ 1.150)^b	< 0.001	1.006 (0.958 ~ 1.057)	0.796	1.008 (0.985 ~ 1.031)	0.487
2 年			1.121 (1.091 ~ 1.152)^b	< 0.001	0.999 (0.948 ~ 1.053)	0.963	1.002 (0.978 ~ 1.027)	0.877
3 年			1.121 (1.090 ~ 1.153)^b	< 0.001	0.994 (0.941 ~ 1.050)	0.836	0.998 (0.973 ~ 1.024)	0.897
5 年			1.121 (1.088 ~ 1.155)^b	< 0.001	0.989 (0.934 ~ 1.047)	0.694	0.994 (0.968 ~ 1.021)	0.647
到公园的距离 (km)	——	0.673	1.048 (1.035 ~ 1.061)^b	< 0.001	0.993 (0.969 ~ 1.016) ^c	0.538	0.995 (0.981 ~ 1.010)	0.515
1 个月					0.997 (0.940 ~ 1.057)	0.915		
3 个月					0.994 (0.932 ~ 1.060)	0.859		
6 个月					0.993 (0.927 ~ 1.063)	0.831		
1 年					0.991 (0.921 ~ 1.066)	0.806		
2 年					0.989 (0.915 ~ 1.070)	0.786		

续表 5.4 建成环境累积平均暴露水平与所有、致死型、非致死型 AMI 复发事件之间的 HR 值及 95%置信区间

变量	缓冲区	IQR	所有 AMI 复发事件		致死型 AMI 复发事件		非致死型 AMI 复发事件	
			HR (95% CI) ^a	<i>p</i> 值	HR (95% CI) ^a	<i>p</i> 值	HR (95% CI) ^a	<i>p</i> 值
3 年					0.988 (0.911 ~ 1.072)	0.775		
5 年					0.987 (0.907 ~ 1.074)	0.763		
蔬果店密度 (个)	1000 m	39	0.665 (0.636 ~ 0.696)^c	< 0.001	1.019 (0.955 ~ 1.089)	0.565	0.961 (0.922 ~ 1.002)	0.059
1 个月			0.667 (0.596 ~ 0.746)^b	< 0.001				
3 个月			0.647 (0.570 ~ 0.733)^b	< 0.001				
6 个月			0.634 (0.554 ~ 0.726)^b	< 0.001				
1 年			0.621 (0.537 ~ 0.718)^b	< 0.001				
2 年			0.609 (0.521 ~ 0.712)^b	< 0.001				
3 年			0.602 (0.512 ~ 0.708)^b	< 0.001				
5 年			0.593 (0.501 ~ 0.704)^b	< 0.001				
PM _{2.5} (μg/m ³)	——	16.995	3.743 (3.643 ~ 3.845)^c	< 0.001	1.148 (1.084 ~ 1.217)^b	< 0.001	1.037 (1.002 ~ 1.073)^b	0.036
1 个月			3.672 (3.427 ~ 3.935)^b	< 0.001				
3 个月			4.399 (4.066 ~ 4.758)^b	< 0.001				
6 个月			4.929 (4.527 ~ 5.367)^b	< 0.001				
1 年			5.536 (5.048 ~ 6.071)^b	< 0.001				
2 年			6.204 (5.616 ~ 6.853)^b	< 0.001				
3 年			6.631 (5.976 ~ 7.357)^b	< 0.001				
5 年			7.211 (6.463 ~ 8.046)^b	< 0.001				

^a HR 为各个建成环境暴露因素每增加 1 个 IQR 水平下的结果，已经调整了年龄、性别、婚姻状况、平均教育年限、家庭年均收入、并发症史（包括高血压、高血脂、糖尿病、中风、心力衰竭、心房颤动和肾功能障碍）。

^b $p < 0.05$ 。

^c 提供的 HR 为 Cox 比例风险模型的结果，而非含时变系数的 Cox 模型的结果。

综上所述,与致死型、非致死型AMI复发事件显著相关的建成环境暴露因素,是所有AMI复发事件识别出的显著因素的一个子集。除了PM_{2.5}的累积平均暴露水平之外,其余在主要结局中呈现显著性的因素,最多与致死型或非致死型复发中的一种结局存在显著关联性。

5.4.3 北京市五环内子队列的关联分析结果

单独进行北京市五环内子队列的关联分析,是为了探究五个五环内才可进行评估的建成环境暴露因素(即建筑密度、建筑容积率、空间失序、步行指数、街道绿视率),是否会对AMI复发事件产生显著影响。然而,这五个建成环境暴露因素并未被ENET算法选择出来,从而未被纳入最终的多因素Cox模型进行分析。因此,从目前的证据来看,还不足以支持这五个因素对AMI复发存在健康效应的结论。此外,五环内的多因素关联分析结果,可作为前述全队列分析的一个稳健性检验来对待——对于基线暴露情景,主要结局分析中显著的城市功能混合度和到地铁站的距离在此处都未被算法选择出来,而PM_{2.5}暴露也仅与致死型AMI复发事件呈现显著相关(HR = 1.238, 95% CI: 1.076~1.423);而对于累积平均暴露水平,只有PM_{2.5}这一个因素被ENET算法选择出,其他建成环境暴露因素都未被选出,与主要结局分析类似,PM_{2.5}的累积平均暴露与所有、致死型、非致死型AMI复发结局都呈现较强的显著相关性。上述差异,可能主要来自于北京市五环内和五环外的建成环境暴露水平存在较大差异,这是因为五环以外的地区,城镇化水平相对较低,城市建设与空间品质相对落后,而五环内的区域往往被认为是北京市的中心城区。

5.5 结果讨论

上述多因素关联分析结果表明,AMI复发风险与城市功能混合度、到地铁站的距离、到公园的距离和蔬果店密度这四个建成环境暴露因素以及PM_{2.5}年均浓度显著关联。本节将这些研究结果与现有流行病学证据进行对比,并尝试对这些建成环境要素对居民心血管健康的传导机制开展初步探讨。

城市功能混合度、到地铁站的距离、到公园的距离和蔬果店密度这四个显著的建成环境指标都与城市和交通设计方面息息相关。Gile-Corti等人^[91, 130]认为城市和交通设计干预可以间接地通过影响居民交通出行方式、居民生活习惯以及带来的相关风险因素暴露(如空气污染、噪音污染、体力活动不足、饮食不健康等),来影响到慢性疾病的发生风险。Nieuwenhuijsen^[92]认为城市设计可以影响到人们的步行及其他交通出行行为,而这些与居民的诸多心血管代谢风险因素(如

肥胖和高血压)、体力活动水平、自然环境暴露等密切相关,从而影响到心血管健康水平。

本研究揭示了无论是从基线暴露还是累积平均暴露的角度,离地铁站更近都有利于降低AMI复发风险的关联规律。直接关注到地铁站距离的流行病学证据较为匮乏,但是,已经有来自多个国家的多项横断面研究显示,到公共交通的距离更近,或是更高的公共交通密度,可以显著降低人们的静坐时间、增加人们的步行行为^[91, 119, 131]。同时,高体力活动水平已被认为与较低的心血管疾病发病风险显著相关^[132, 133]。GBD 2019 项目指出,全球归因于体力活动不足的缺血性心脏病的疾病负担占到了4.16%^[56]。这些已有研究有助于解释本研究中到地铁站距离和AMI复发风险之间的关联关系。

除了到地铁站的距离之外,城市功能混合度被识别为AMI复发的显著保护因素。城市功能混合度是从POI的角度测度的功能多样性,而土地利用混合度是从用地的角度测度的多样性。尽管目前关注城市功能混合度的流行病学研究仍然缺乏,但是,一些关注紧凑型城市的研究已经表明,混合用地的城市发展模式将有助于降低心血管疾病的疾病负担水平^[134, 135]。背后原因为,混合用地的发展模式将居住用地、商业用地等多种用地类型集中在相近的邻里社区尺度内,促进了居民的步行出行的近距离交通意愿,而减少了以小汽车为主的远距离出行行为,从而促进了居民体力活动水平,预防了相关非传染性慢性疾病(如心血管疾病)的发生^[91]。未来需要更多流行病学研究去增加城市功能多样性对AMI复发的证据强度,并探索与深化其影响机制。

除此以外,本研究还发现到达公园的距离越近,将可能降低AMI的复发风险。尽管公园与AMI复发的流行病学研究仍然缺乏,但是已经有许多队列研究揭示了居住在公园附近或者住址周边有更高的公园比例的居民,其急性心肌梗死或者心血管疾病的发病风险更低^[31, 34, 136]。仅有一项横断面研究观察到了公园可达性对AMI的负面健康效应,这可能与横断面研究设计较弱的因果推断能力有关^[45]。公园可达性可能会影响居住在周边的居民到达公园的意愿,从而影响体力活动水平,此外公园也可以缓解邻里环境的空气污染和噪声污染问题,而这些都是心血管疾病公认的风险因素^[91, 137, 138]。本研究的发现与以往研究一致,均支持公园对心血管健康的积极效应。

最后,1000米缓冲区内的蔬果店密度在本研究中被发现是AMI复发的显著保护因素。Giles-Corti等人^[91, 130]认为这类型商店通常是通过影响人们的饮食习惯来影响居民健康。西方国家的几项横断面研究显示,超市的邻近性或者可达性与居民蔬菜水果的摄入量呈现显著正相关关系^[139, 140]。已有荟萃分析指出,蔬菜水果摄

入不足的饮食习惯与缺血性心脏病发病风险的提高显著相关^[141]。GBD 2019 评估结果显示,全球归因于蔬菜、水果摄入不足的缺血性心脏病疾病负担分别占到4.61%和5.62%^[56]。本研究得到的结果与上述文献结果一致,认为应该在居民可步行到达的区域内多增加蔬果店的城市设计应对策略,以降低急性心肌梗死的复发风险。

除了上述四个显著的建成环境暴露因素之外,本研究发现的PM_{2.5}暴露与急性心肌梗死复发的显著关联在已有的流行病学研究中也得到了印证。Liu 等人^[142]和Chen 等人^[103]通过队列研究分别揭示了更高的PM_{2.5}浓度暴露与ST段抬高型心肌梗死的医院复诊率以及急性心肌梗死的致死型复发事件呈现显著的密切关联。Zhu 等人^[143]通过一项荟萃分析研究,揭示了大气PM_{2.5}暴露与心肌梗死预后死亡之间的显著关联(HR = 1.07, 95% CI: 1.04~1.09, 每升高10 μg/m³ PM_{2.5}),并解释了心肌梗死预后死亡大部分归因于心血管疾病。这些关联可以进一步被PM_{2.5}相关的毒理学证据支持。多项毒理学研究表明,AMI患者可能通过系统氧化应激及炎症、动脉粥样硬化、心率变异性和血液凝固性等相关机制,对PM_{2.5}暴露体现出易感性^[144, 145]。因此,降低PM_{2.5}的暴露水平可能可以降低AMI的复发风险。

5.6 本章小结

本章基于第4章建成环境暴露的评价结果,比较选择了适用于本研究数据集的暴露组变量选择算法,并通过生存分析方法探究了建成环境暴露与急性心肌梗死复发风险之间的关联规律,也为第6章的复发风险现状评估提供了暴露反应关系的数据基础。本章的主要结论为:

(1) 采用多种统计学性能指标对比了五种变量选择算法的性能,在综合考虑变量选择算法的准确性与稳健性的准则下,ENET算法的统计学性能在本研究数据集下最优,因此本研究选择ENET算法作为变量选择方法;

(2) 单因素关联分析中,基线暴露水平下关于15种建成环境暴露因素的38个变量呈现出显著性,累积平均暴露水平下关于16种建成环境暴露因素的47个变量均呈现出显著性;

(3) 多因素关联分析中,关于所有AMI复发事件这一主要结局,在基线暴露水平下,500米缓冲区内的城市功能混合度为显著保护因素,到地铁站的距离及PM_{2.5}为显著危险因素;在累积平均暴露水平下,500米缓冲区内的城市功能混合度、1000米缓冲区内的蔬果店密度为显著保护因素,到地铁站的距离、到公园的距离及PM_{2.5}为显著危险因素。

(4) 多因素关联分析中, 关于致死型复发事件, 无论是在基线暴露水平下, 还是在累积平均暴露水平下, 500 米缓冲区内的城市功能混合度均为显著保护因素, 到地铁站的距离及 $\text{PM}_{2.5}$ 为显著危险因素; 关于非致死型复发事件, 除了城市功能混合度不再显著之外, 其余结论与致死型复发事件相同。

第6章 北京市建成环境的归因 AMI 复发风险 现状评估及规划应对

6.1 北京市建成环境归因 AMI 复发风险评估

本章基于第5章获得的建成环境与AMI复发之间的暴露反应关系，对北京市建成环境现状开展AMI复发风险评估，从而进一步识别出待提升小区，探讨其规划应对策略。

6.1.1 建成环境归因 AMI 复发风险评估方法

建成环境的归因AMI复发风险评估参照GBD 2019项目的风险因素归因疾病负担的评估流程来进行，其主要分为四个步骤：环境风险因素暴露估计、暴露反应关系估计、PAF估计及归因疾病负担估计^[56]。由于本研究中关注的健康结局均为AMI复发事件，因此只要PAF相同，归因疾病负担就应该相同。此外，对于AMI复发结局暂无相应的总疾病负担数据。综上，本章的风险评估工作仅进行GBD 2019项目评估流程的前三个步骤，得到的PAF越高，就代表该空间区域的AMI复发风险较高。下面逐一对评估流程每一个环节予以介绍。

(1) 环境风险因素暴露估计

本章风险评估中关注的建成环境风险因素，为第5章探究建成环境与AMI复发关联时识别出的显著因素。此处采用累积平均暴露情形下的显著因素，即500米缓冲区内的城市功能混合度、到地铁站的距离、到公园的距离、1000米缓冲区内的蔬果店密度及PM_{2.5}浓度，这是因为风险现状评估是以同一年份的不同空间为对象进行评估，而非围绕每一个预后的患者开展，因此并不能确切地获得所谓的基线暴露水平。本研究的研究范围为北京市全域，研究年份是最近数据较全的2021年。本节所用环境数据集为上述五个指标所需的2021年北京市POI数据（来自于高德地图抓取）及PM_{2.5}大气浓度数据^[146, 147]。

开展建成环境暴露评价时，与4.3节介绍相同，将北京市全域划分为100 m×100 m的网格，取其中心点来进行空间分析。500米缓冲区内的城市功能混合度、到地铁站的距离、到公园的距离、1000米缓冲区内的蔬果店密度这四个指标均采用的是POI数据，其测度方式已分别在4.3.2节、4.3.4节和4.3.5节详细阐述。PM_{2.5}浓度的测度方式也已在4.3.6节详细介绍。此外，建成环境暴露数据时间变异性相对较小，因此，本研究将2021年的暴露数据近似看作累积平均暴露来处理。

(2) 暴露反应关系估计

本章使用的暴露反应关系，为第5章中获得的累积平均暴露情景下有统计学显著意义的 HR 值。由于第5章后来采用了含时变系数的 Cox 模型，但此处考虑到距离患者康复的时间未知，且大部分指标 HR 随时间变化不大，因此本章采用原 Cox 比例风险模型获得的暴露反应关系，现总结在表 6.1 中。

表 6.1 建成环境与 AMI 复发之间具有显著性的 HR 及其 95%置信区间

环境暴露因素	缓冲区	HR ₀ (95% CI)	单位水平 E_0	TMREL	最大值 $E_{\max}$
城市功能混合度	500 m	0.839 (0.793 ~ 0.890)	1	——	2.437
到地铁站的距离	——	1.033 (1.031 ~ 1.036)	1 km	0.145 km	39.471 km
到公园的距离	——	1.072 (1.052 ~ 1.092)	1 km	0.093 km	6.384 km
蔬果店密度	1000 m	0.901 (0.890 ~ 0.911)	10 个	——	104 个
PM _{2.5}	——	2.174 (2.140 ~ 2.209)	10 $\mu\text{g}/\text{m}^3$	32.875 $\mu\text{g}/\text{m}^3$	89.7 $\mu\text{g}/\text{m}^3$

注：表中 HR₀ 指的是对应每增加 1 个单位水平 E_0 下的 HR，不再是每增加一个 IQR 水平的 HR。

基于 Cox 模型的回归系数解读方式，暴露反应关系为一个指数函数形式。考虑到第5章队列研究中各个建成环境因素的暴露水平存在上限，因此，取队列研究中各个因素暴露水平的 99%分位数作为最大限制 $E_{\max}$ （如表 6.1 所示），高于 $E_{\max}$ 的暴露水平区段不再进行暴露反应关系的简单外推，而是采用 $E_{\max}$ 下对应的 HR 值进行保守估计。上述暴露反应关系可以定量表达为^[148]

$$HR(E) = \begin{cases} HR_0^{\frac{E}{E_0}}, & E \leq E_{\max} \\ HR_0^{\frac{E_{\max}}{E_0}}, & E > E_{\max} \end{cases} \quad (6.1)$$

其中， E 为当前关注的暴露水平， E_0 为 HR₀ 对应的单位水平， $E_{\max}$ 为队列研究中该因素暴露水平的 99%分位数。

(3) PAF 估计

具备了建成环境因素暴露水平和暴露反应关系之后，便可以进行 PAF 的估计。PAF 指的是由于该环境因素暴露导致的额外健康风险占总健康风险的比例。举例说明，若某因素的 PAF 为 20%，则该健康结局的疾病负担中认为有 20%是该因素的暴露造成的。对于显著的风险因素，PAF 的定义式可由公式(6.2)表达^[56, 149]：

$$PAF = \frac{\int_{sp} [HR(E) - HR(TMREL)] p(E) dE}{\int_{sp} HR(E) p(E) dE} = 1 - \frac{HR(TMREL)}{\overline{HR}} \quad (6.2)$$

其中， E 为该建成环境因素的暴露水平；TMREL 为理论最小风险暴露水平

(theoretical minimum risk exposure level)，即低于此暴露水平认为无额外健康风险，取队列研究中各因素暴露水平的 1%分位数； sp 为选定的空间单元； $p(E)$ 为该空间单元内该因素暴露水平的概率密度函数在 E 处的取值； $\overline{HR}$ 为该空间单元内各点 HR 的人口加权平均值。风险因素的 PAF 为一个正数。

对于显著的保护因素，PAF 的定义式可由公式(6.3)表达^[56]：

$$PAF = \int_{sp} [HR(E) - 1] p(E) dE = \overline{HR} - 1 \quad (6.3)$$

此处，保护因素的 PAF 定义为一个负数。

公式(6.2)和(6.3)中都涉及空间单元 sp ，本章考虑了北京市乡镇街道尺度和居住小区尺度两种空间单元。北京市乡镇街道尺度是根据 2020 年我国常住人口七普数据，选择了北京市人口密度大于 100 人/km²（即人口密集区）的乡镇街道作为分析的空间单元，这是因为对于人口稀疏的乡镇街道（如密云区、怀柔区、延庆区一些以山地为主的乡镇街道），大部分地区无人居住，对每个 100 m 网格开展 AMI 复发风险评估意义不大。北京市居住小区尺度是根据目前可获得的数据资源，选择了北京市六环内的居住小区作为分析的空间单元。空间单元所用到的数据为 3.2.1 节介绍的北京市乡镇街道和六环内小区的面要素矢量数据。

上面介绍了单一的建成环境因素的 PAF 评估方法，若要进一步综合考虑多个建成环境因素对 AMI 复发风险的总影响，可采用如下方法进行评价。多个风险因素的总 PAF 不是单项 PAF 的简单加和，而需要用公式(6.4)进行计算^[56]：

$$PAF_h = 1 - \prod_i (1 - PAF_i) \quad (6.4)$$

其中， PAF_h 为多个风险因素的总 PAF， PAF_i 为单个风险因素 i 的造成的 PAF。而多个保护因素的总 PAF 采用公式(6.5)进行计算：

$$PAF_b = \prod_j (1 + PAF_j) - 1 \quad (6.5)$$

其中， PAF_b 为多个保护因素的总 PAF， PAF_j 为单个保护因素 j 的造成的 PAF。则，所有建成环境暴露因素造成的 AMI 复发的 PAF 为

$$PAF = PAF_h + PAF_b \quad (6.6)$$

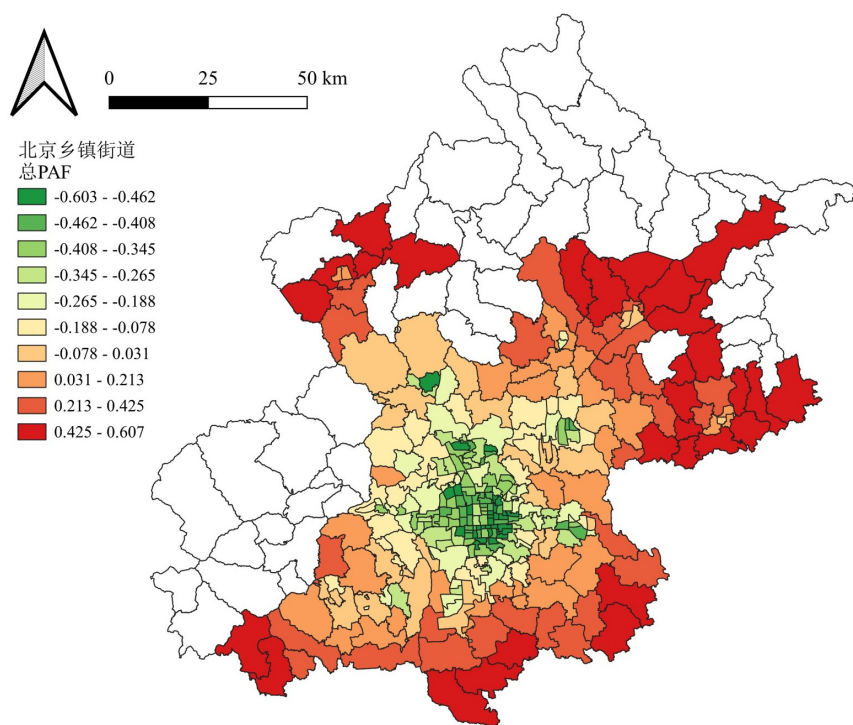
上述 PAF 指标可从公共健康视角确定建成环境归因 AMI 复发风险较高的空间区域，从而为识别优先待提升区域提供指导。从城市规划视角，还关注建成环境归因复发风险的干预潜力大小，即城市规划与设计干预能够贡献的风险降低比例。由于上述 PAF 指标会受保护因素与风险因素的正负抵消，本研究将建成环境归因风险干预潜力定义为

$$Potential = \sum_{i \in \text{风险因素}} PAF_i + \sum_{i \in \text{保护因素}} (PAF_i - PAF_{i,\max}) \quad (6.7)$$

其中, $PAF_{i,\max}$ 为保护因素 i 在暴露水平 $E_{\max}$ 下对应的 PAF 值。

6.1.2 乡镇街道尺度复发风险评估结果

本节介绍北京市人口密度大于 100 人/ km^2 的乡镇街道的建成环境归因 AMI 复发风险的评估结果。图 6.1 展示的是五个显著的环境暴露因素的 AMI 复发总 PAF 及干预潜力的空间分布情况。结果表明,北京市整体上呈现中心城区的建成环境 PAF 较低, 偏远乡镇街道 PAF 较高的分布; 而在郊区区县, 区县城关所在乡镇街道的 PAF 相对较低。PAF 最低的为东城区崇文门外街道(-60.3%), PAF 最高的为大兴区榆垓镇(60.7%)。图中黄色到绿色的空间单元表示的是建成环境造成的 AMI 复发 PAF 小于 0 的区域, 约为北京市六环内的空间范围, 该结果同时也表明, 北京市近 60% 的乡镇街道内, 五项建成环境因素对 AMI 复发整体上呈现出保护作用。建成环境归因风险干预潜力的空间分布情况与总 PAF 类似, 呈现中心城区干预潜力较低、偏远地区干预潜力较高的分布特征, 潜力最低的仍为东城区崇文门外街道(0.192), PAF 最高的为怀柔区桥梓镇(2.014)。



(a)

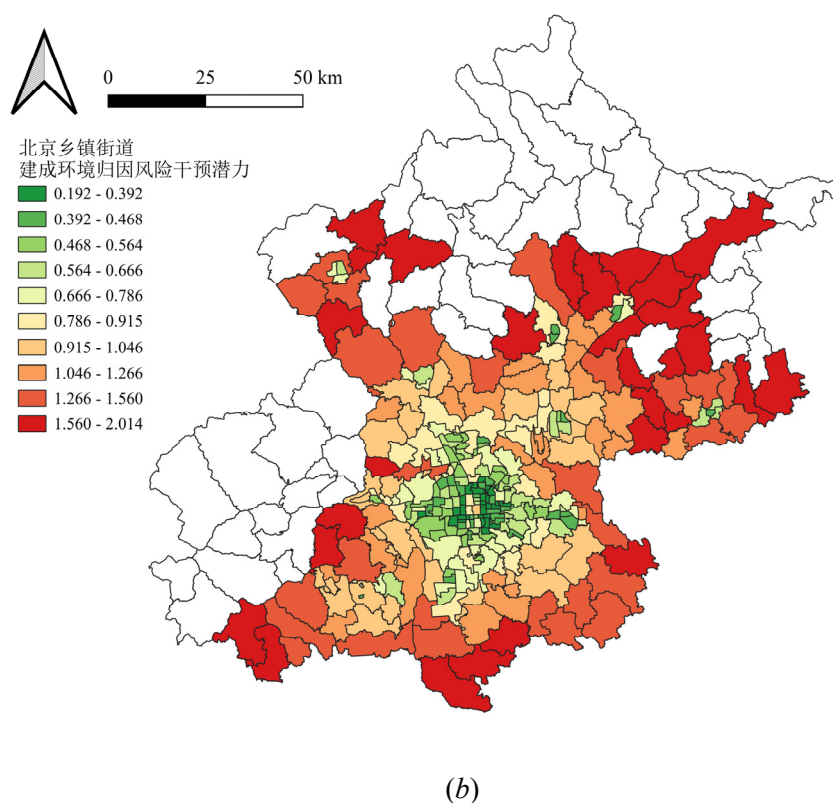


图 6.1 北京市乡镇街道尺度建成环境的总归因复发风险结果
(a) 总 PAF (b) 建成环境归因风险干预潜力

其次考察五个建成环境因素的单项 PAF 空间分布结果。图 6.2 是城市功能混合度对 AMI 复发的 PAF 评估结果，其空间分布模式与图 6.1 比较接近——中心城区和郊区区县城关地区的 PAF 较低，其余郊区区域 PAF 较高。其中，PAF 最优的为东城区朝阳门街道（-34.7%），PAF 最差的为密云区巨各庄镇（-14.2%）。出现上述结果的原因，是中心城区与郊区区县城关地区的 POI 数量和种类都相对较多，导致有较高的城市功能混合度，从而促进附近居民的步行出行意愿、提高其体力活动水平。

图 6.3 是到地铁站距离对 AMI 复发的 PAF 评估结果，呈现中心城区向外围郊区 PAF 逐渐增加的趋势，但是整体而言，PAF 的差异和绝对水平相较前两个因素并不明显。背后原因，一方面是近年来，北京市的地铁线路网已经深入到各个郊区区县，另一方面是到地铁站距离的暴露反应关系相比其他因素来说更弱一些。从全北京市来看，PAF 最优的为东城区交道口街道（0.7%），PAF 最差的为密云区太师屯镇（72.4%）。

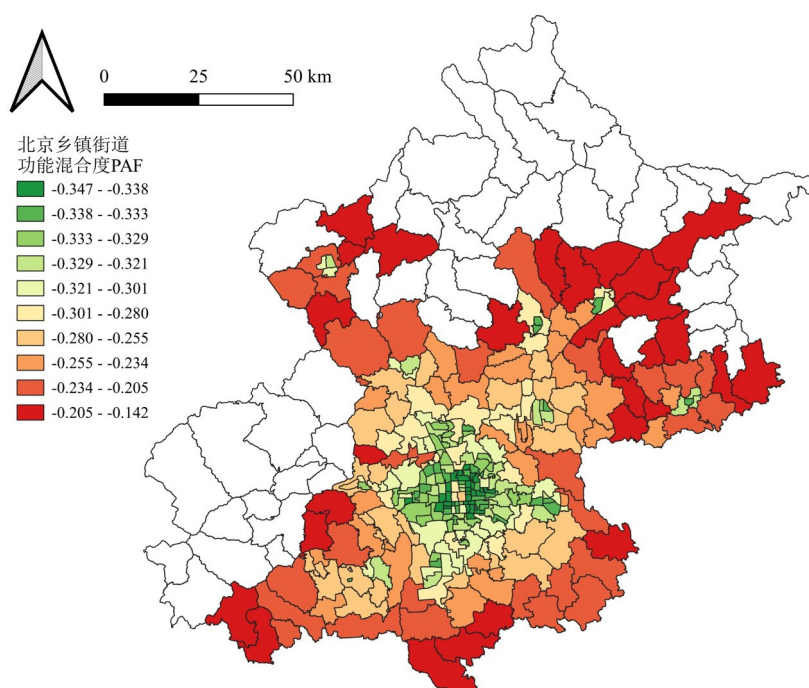


图 6.2 北京市乡镇街道尺度城市功能混合度对 AMI 复发的 PAF 结果

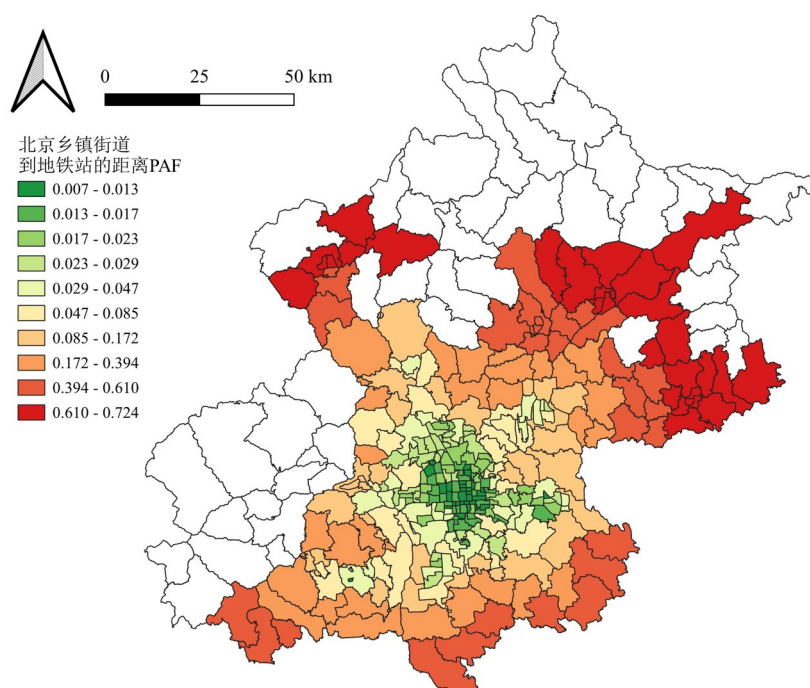


图 6.3 北京市乡镇街道尺度到地铁站距离对 AMI 复发的 PAF 结果

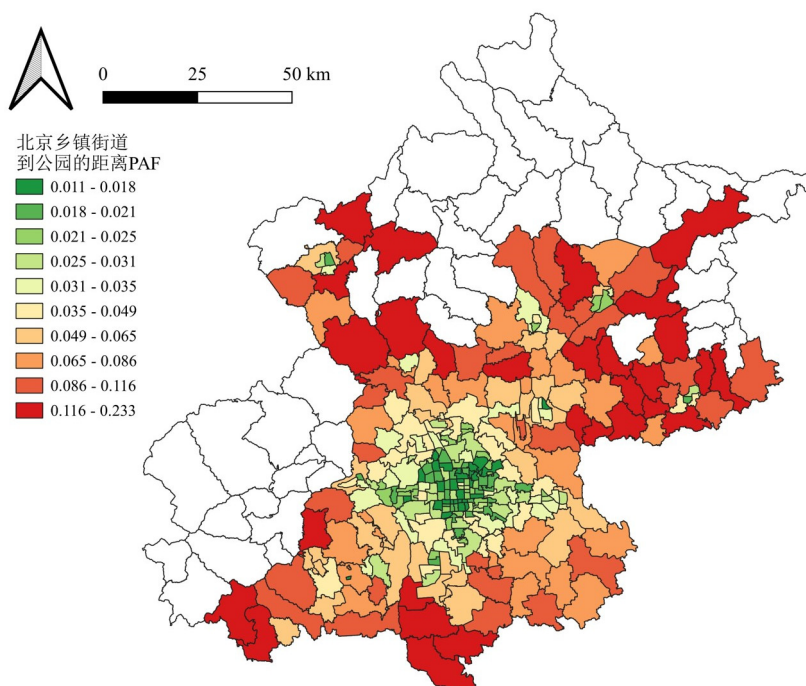


图 6.4 北京市乡镇街道尺度到公园距离对 AMI 复发的 PAF 结果

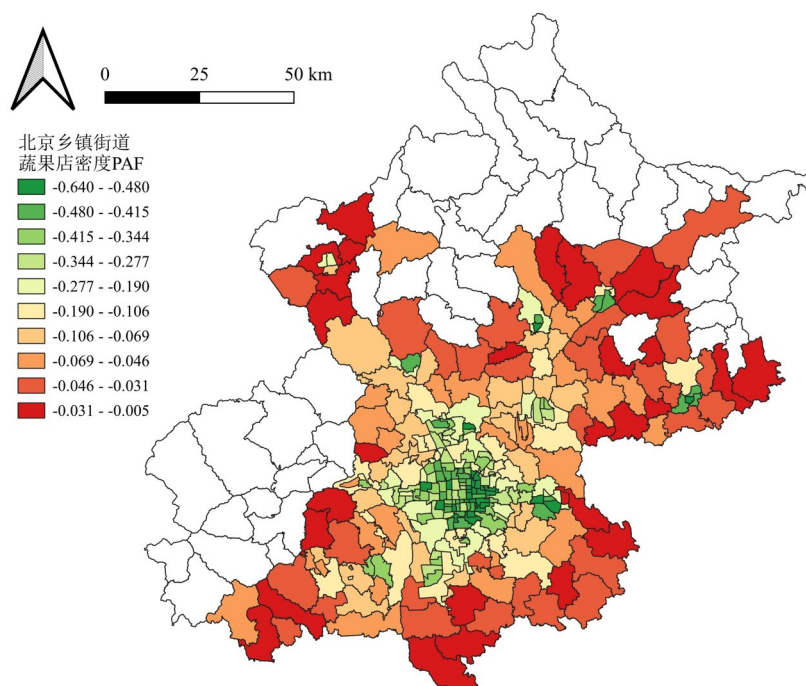


图 6.5 北京市乡镇街道尺度蔬果店密度对 AMI 复发的 PAF 结果

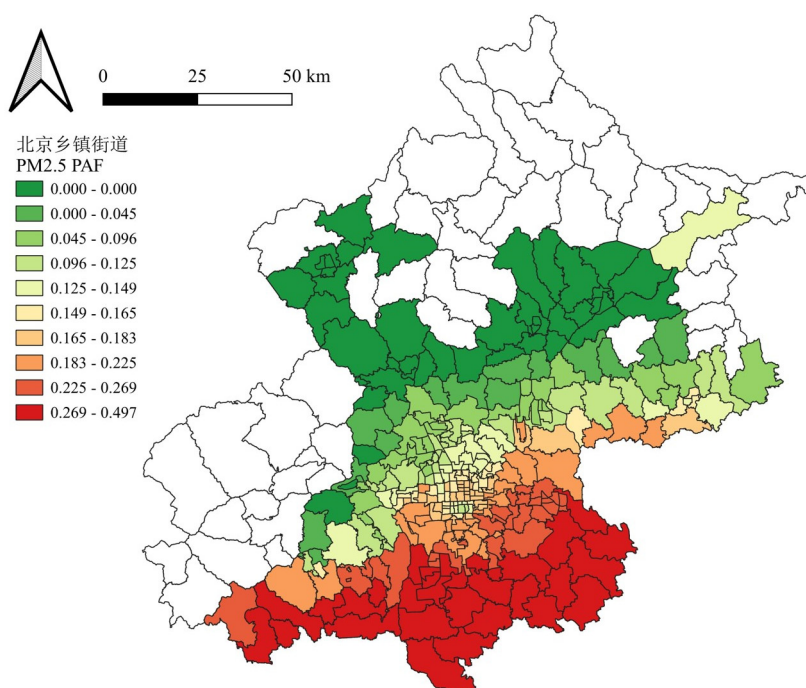
图 6.6 北京市乡镇街道尺度 PM_{2.5} 对 AMI 复发的 PAF 结果

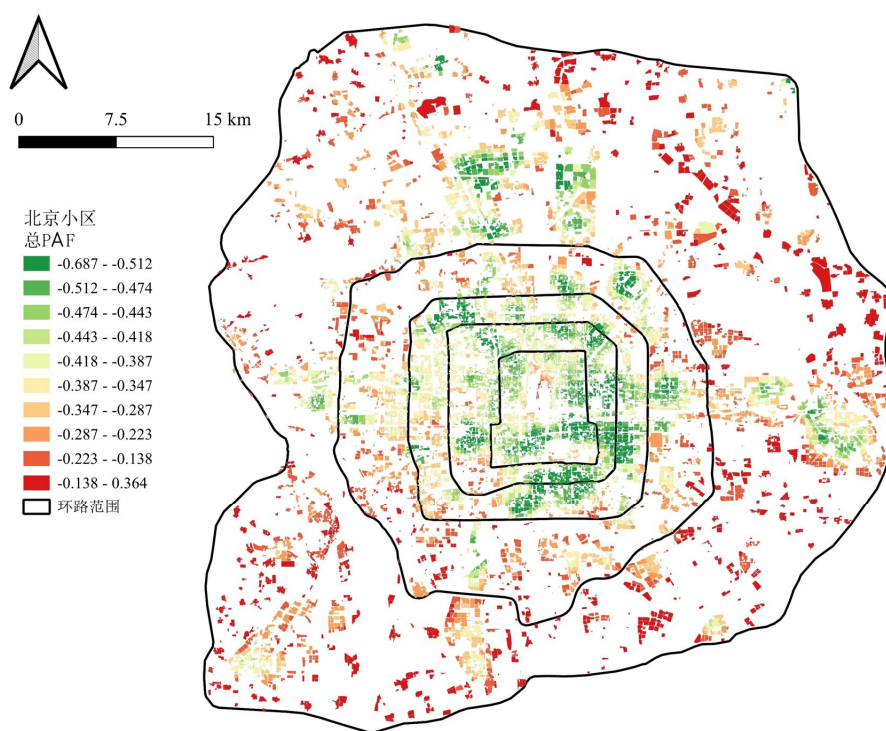
图 6.4 是到公园距离对 AMI 复发的 PAF 评估结果，空间分布规律呈现中心城区向外围郊区 PAF 逐渐增加的趋势，而郊区区县城关区域以及东南方向的通州区，到公园距离引起的 PAF 也相对较低。大约 90% 的乡镇街道的 PAF 均小于 10%，表明到公园距离带来的复发风险问题相对来说并不严重，这也跟近年来北京市注重绿色空间建设（如口袋公园建设）有关。从全北京市来看，PAF 最优的为东城区大栅栏街道和房山区星城街道（1.1%），PAF 最差的为顺义区北务镇（23.3%）。

图 6.5 是蔬果店密度对 AMI 复发的 PAF 评估结果，空间分布规律与前述指标类似，将近一半的乡镇街道该指标的保护效应 PAF 达到了 -20% 以下。从全北京市来看，PAF 最优的为东城区崇文门外街道（-64.0%），PAF 最差的为延庆区八达岭镇（-0.5%）。

图 6.6 是 PM_{2.5} 年均浓度对 AMI 复发的 PAF 评估结果，空间分布规律与前述指标不同，呈现出西北方向到东南方向逐渐升高的趋势，这主要是因为北京南部靠近河北省重工业区域的污染源，因此南部的 PM_{2.5} 污染更为严重，而北京市西北部为山区地形，有效地阻隔了大气污染物的迁移，所以空气污染状况较轻。从全北京市来看，PAF 最差的为大兴区榆垓镇（49.7%），而密云、怀柔、延庆的大部分乡镇街道因为浓度低于 6.1.1 节中设定的 *TMREL*，因此 PAF 目前认为是 0。

6.1.3 居住小区尺度复发风险评估结果

本节介绍北京市居住小区的建成环境归因 AMI 复发风险的评估结果，由居住小区数据所限，本节的复发风险评估局限在北京市六环范围以内。图 6.7 展示的是五个显著的环境暴露因素的 AMI 复发总 PAF 和建成环境干预潜力的空间分布情况。从图中可以发现，两者空间分布特征极为相似，因此下面仅以总 PAF 为例进行结果分析与讨论。PAF 较低的区域（偏绿色区域）呈现出单中心、放射状的布局——四环内区域 PAF 普遍较低，四环外 PAF 较低区域呈现多轴放射状。其中，城北往回龙观、天通苑方向的放射轴最为明显，且绿色区域分布范围最广，这可能与北京市南北发展不对称、北部发展更为良好有关，使得这一片区的建成环境整体品质较好；城东往通州城关方向的放射轴也较为明显，这可能与北京市近年将市政府搬迁至通州区相关，使得通州区的建成环境品质也有所提升；而城西往石景山方向、城南往大兴黄村地区也有相应的 PAF 较低的轴向，但是没有上述两个方向更为明显。整体而言，六环内有近 90% 居住小区，该五项建成环境指标带来的 AMI 复发 PAF 小于 0，即这几个建成环境暴露因素整体上呈现保护作用。



(a)

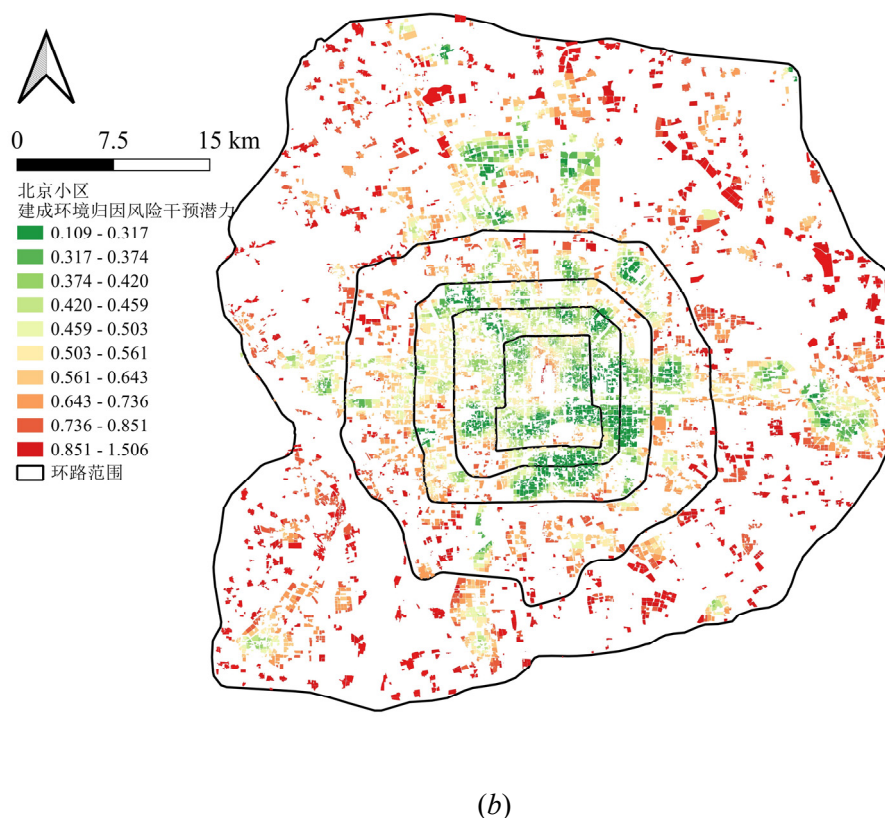


图 6.7 北京市六环内小区尺度建成环境的总归因复发风险结果
(a) 总 PAF (b) 建成环境归因风险干预潜力

图 6.8 至图 6.12 分别展示了五个环境因素的单项 PAF 空间分布结果。城市功能混合度、到地铁站的距离、到公园的距离和蔬果店密度的分布结果与图 6.7 相似，呈现出中心城区普遍较优、外围呈现放射状的 PAF 空间布局，其中，到公园的距离指标 PAF 的空间异质性较低，六环以内各个小区结果更偏均匀。而 $\text{PM}_{2.5}$ 由于污染源主要在南部的特征，与乡镇街道尺度分析结果相同，呈现出西北较优、东南较差的 PAF 空间分布结果。

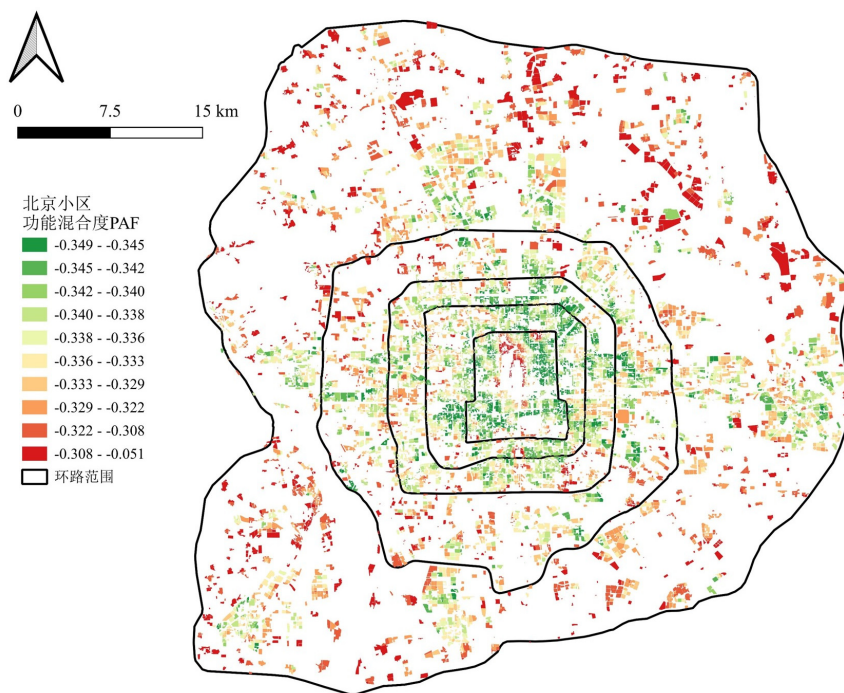


图 6.8 北京市六环内小区尺度城市功能混合度对 AMI 复发的 PAF 结果

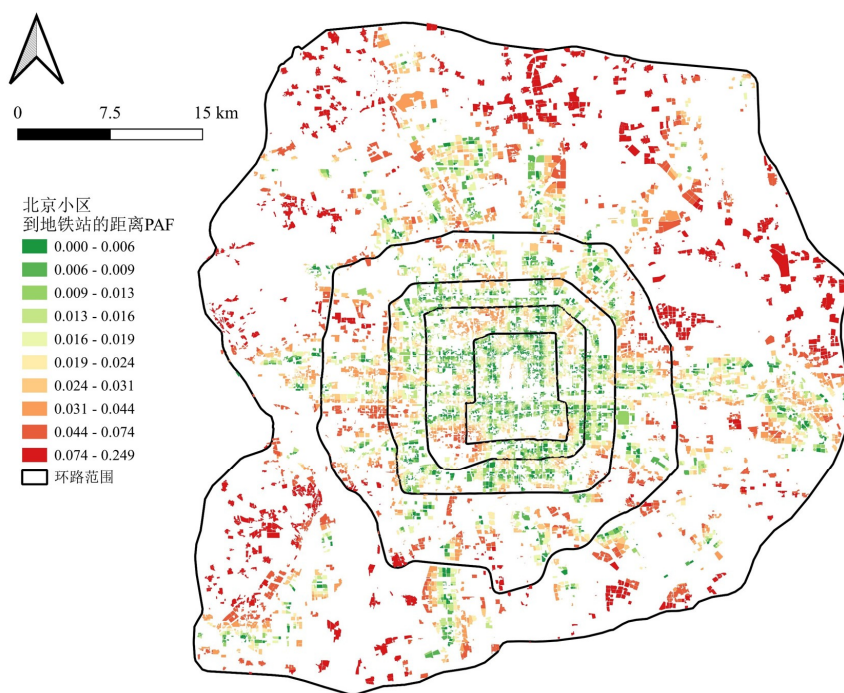


图 6.9 北京市六环内小区尺度到地铁站距离对 AMI 复发的 PAF 结果

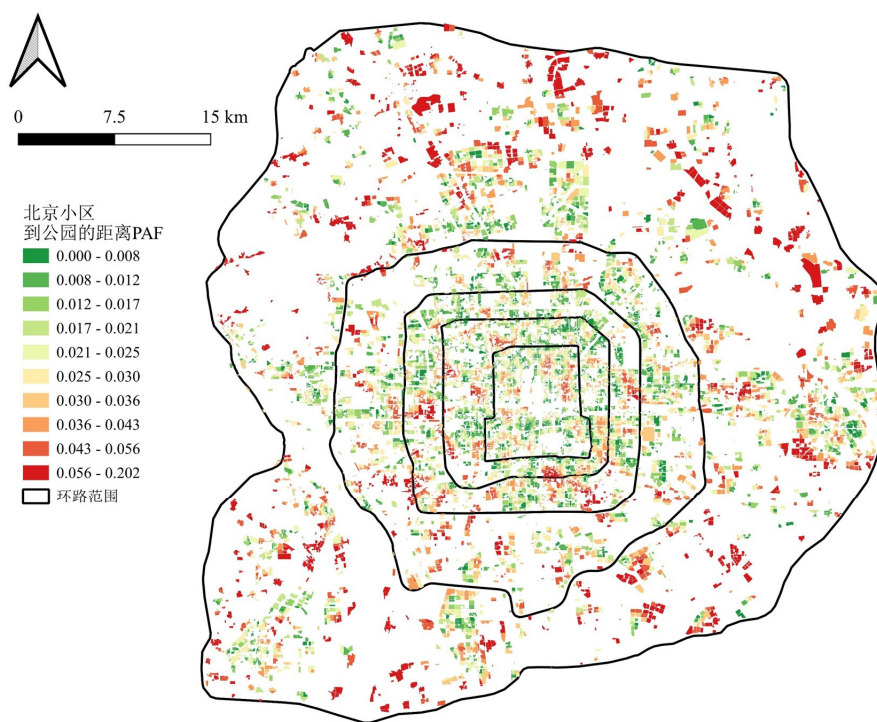


图 6.10 北京市六环内小区尺度到公园距离对 AMI 复发的 PAF 结果

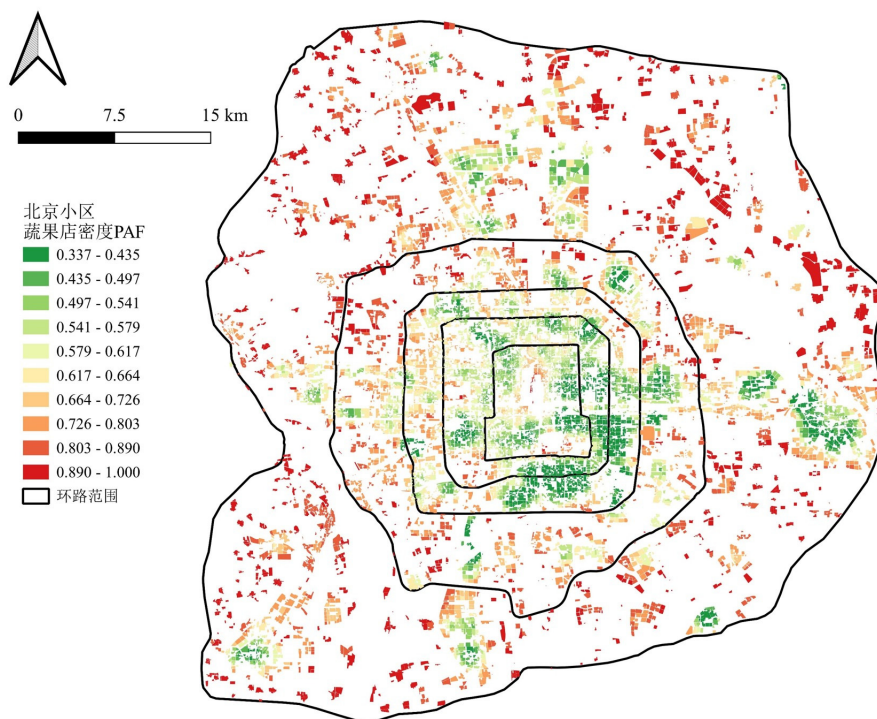
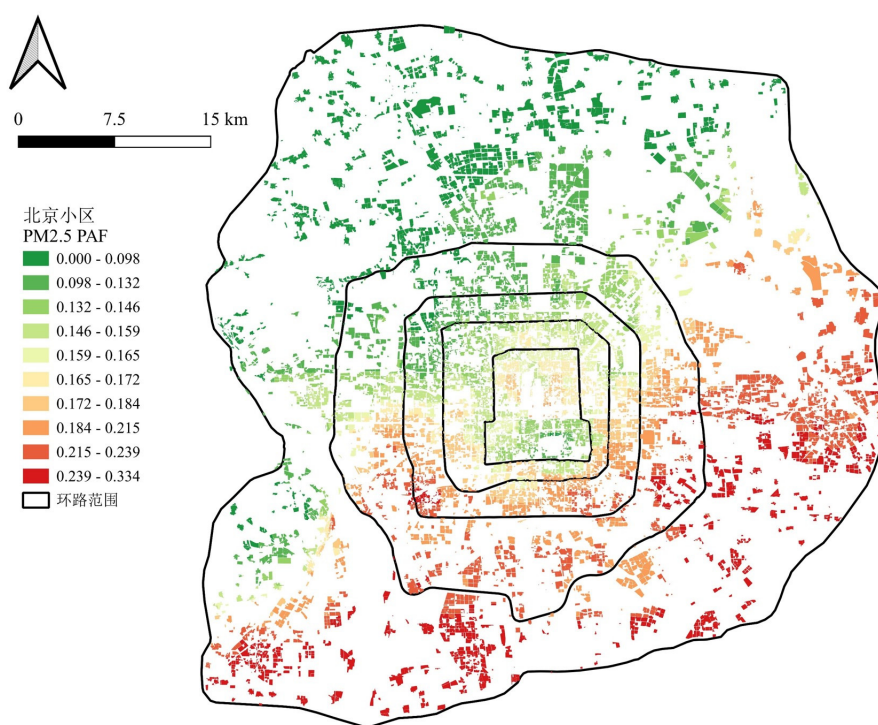


图 6.11 北京市六环内小区尺度蔬果店密度对 AMI 复发的 PAF 结果

图 6.12 北京市六环内小区尺度 PM_{2.5} 对 AMI 复发的 PAF 结果

6.2 北京市居住小区 AMI 预后的规划应对策略

由于地铁站、公园、功能混合度、蔬果店这些建成环境暴露因素更偏向于中微观的城市设计层面，因此，本节选择居住小区的空间尺度，探究如何基于风险评估结果确定规划应对策略。6.2.1 节首先从北京市相关标准与导则出发，梳理总结与本研究关注的五个显著的环境暴露因素有关的条文与建议；6.2.2 节基于风险评估结果，对六环内待提升小区进行识别，并依据出现问题的建成环境要素以及这些因素的改建难易程度，对待提升小区进行分类；最后，6.2.3 节对识别出的几类待提升小区，选取典型小区案例，结合遥感影像、街景图片等数据，进行规划应对策略的分析与探讨。

6.2.1 相关标准与导则

本节选择了《北京市城市设计导则》^[150]、《健康社区评价标准》^[2]以及《城市一刻钟便民生活圈建设指南》^[151]这三项标准或导则，来梳理相应建成环境因素的规划设计策略。这三项标准、导则，分别侧重城市设计、社区健康和便利设施这三个维度，以覆盖城市功能混合度、地铁站、公园、蔬果店、PM_{2.5}这些建成环境因素。

（1）城市功能混合度

城市功能混合度考虑的是城市 POI 的多功能混合程度。《健康社区评价标准》的人文指标的评分项中，8.2.3 节为“社区设有文体、商业及社区服务综合体”，8.2.4 节为“社区设置集展览展示、图书阅览、书画室、网络信息服务、休闲娱乐、教育培训、团队活动于一体的综合性、多功能、公益性文体活动中心”。上述策略表明，居住小区内部可以通过社区服务综合体、文体活动中心的方式，集约更多种城市服务功能，从而实现城市功能混合度的提升，促进小区居民的步行出行意愿及体力活动水平。

《城市一刻钟便民生活圈建设指南》考察的空间尺度比社区稍大，其中提到了“拓展空间”、“丰富业态”的两大建设指南。“拓展空间”指的是“结合城市更新行动，充分利用腾退空间和地下空间”，“挖掘交通场站、公园、园区等空间资源，来布设更多便民设施”；“对于一些不便于建设便民服务设施的空间，还可以投放移动便民商业服务设施，如自助售卖机、蔬菜直通车等”。“丰富业态”意见中，提出要“巩固提升基本保障类业态，将蔬菜零售、便利店（社区超市）、早餐等纳入资金支持范围”；并且“推进一店多能，实现高、低消费频次服务集约布局”。

（2）到地铁站的距离

该建成环境因素改变较难，因为地铁站涉及整体城市的规划设计范畴，因此，上述标准导则涉及较少。《北京市城市设计导则》提到可以“通过促进轨道交通与城市的协调融合发展，加强一体化建设，形成站城一体的布局模式，增强区域可达性”。城市规划领域中非常经典的 TOD 发展模式（以公共交通为导向的开发模式），即为一种轨道交通与城市社区融合发展的范例。

《健康社区评价标准》中虽然没有直接提及轨道交通，但是在健身指标的评分项中，7.2.5 和 7.2.9 条提出应“鼓励社区人员采用健康的出行方式”，对健身步道、自行车停车位、自行车骑行道、共享单车停车位等提出了具体要求。5.5 节中曾经提及，到地铁站的距离越近，其实是促进居民能够选择更为健康的交通出行方式（比如步行、骑自行车至附近的地铁站），增加体力活动水平，从而改善心血管健康。因此，《健康社区评价标准》的相关条文，可以认为是在地铁站建设难以改变的情形下，促进附近居民使用地铁这种健康出行方式的一种策略。

（3）到公园绿地的距离

公园绿地一直是城市规划与设计关注的对象。《北京市城市设计导则》在总体指导要求中，提出要“充分利用城市内一定规模的留白增绿空间，种植本地树种，降低维护成本，形成高低错落、亲切自然的城市森林群落”。在街区指导要

求的绿色空间体系栏目当中，提出以下四点要求：

“（1）结合地区整体景观格局，构建有利于生态保护、休闲游憩的绿色空间体系，包括公园绿地、绿道、通风廊道、交通廊道、河流生态廊道等；

（2）重视绿色空间的功能性和实效性，鼓励公园绿地与公共广场用地、室外体育运动场地进行复合设置；支持绿色开放空间与其他经营性功能的合理复合设置；

（3）分级分类对城市公园、河道绿廊、社区公园等周边的建筑形体关系、开放空间形态进行引导，提升公园的可达性、舒适性和观赏性；

（4）鼓励利用腾退资源、闲置空间建设小微绿地、口袋公园，为市民提供家门口的绿色空间。”

《健康社区评价标准》从多个维度多次提及公园绿地相关的条文：空气指标 4.2.14 条提供了社区内及周围环境的绿地率指标（不低于 30%、35%、40%分别对应不同评分），4.2.15 条指出“公园绿地乔灌木比例达到 25%，或防护绿地乔灌木比例达到 60%，或其他建设用地乔灌木比例达到 40%”；舒适指标 6.2.13 条，“社区规划兼顾当地地理位置、气候、地形、环境、全年主导风向等条件，规划建设利用湿地、绿地、街道等形成连续的开敞空间和通风廊道”；以及健身指标 7.1.1 条，“社区内应设置健身运动场地，室外人均健身场地面积不应少于 0.1 m²，室内人均健身场地面积不应少于 0.05 m²”。

（4）蔬果店密度

关于蔬果店，除了上文《城市一刻钟便民生活圈建设指南》中“拓展空间”、“丰富业态”的策略同样适用之外，《健康社区评价标准》在服务指标的评分项中，9.2.8 条提出“社区应具有安全便利的食品销售渠道，可提供充足的粮食、水果、蔬菜等食品”；“利用公有闲置物业或以回购廉租方式保障老旧社区基本商业业态用房需求，用于支持社区菜店、菜市场、农副产品平价商店、便利店、早餐店等居民生活必备的商业网点建设”；“社区能够具备满足社区居民食品消费需求的便民服务网络，打造‘一刻钟’食品便民服务商圈”。

（5）PM_{2.5}

关于 PM_{2.5} 污染，其主要来自于城市或区域尺度的工业污染源，少部分可能来自于小区内部或附近的一些由焚烧、垃圾堆放、施工造成的污染源。在小区层面，对于工业污染源，往往只能从防护的角度来考虑，而对于小区内部的一些小型污染源，则可以进行污染源防控措施。《健康社区评价标准》空气指标中，提出社区内应实施环境绿化措施、降低施工扬尘、设置禁烟标识、禁止露天焚烧处理，并且社区边界与工业污染源之间需满足防护距离、与机动车道间设置不小于 3 米

宽度的乔灌结合绿化带。此外，对于绿化率及乔灌木比例的要求，在公园绿地部分已经说明，此处不再重复。

6.2.2 待提升小区的识别与分类

按照 6.1 节的建成环境归因 AMI 复发风险评估结果，选择总 PAF 值位于后 40% 的六环内小区作为待提升小区。对于这些待提升小区，五个建成环境因素均设置相应的 PAF 临界值，将这些待提升小区的建成环境暴露水平分成优、差两档。PAF 临界值设置及其背后含义如表 6.2 所示。

表 6.2 建成环境因素的 PAF 临界值

	功能混合度	到地铁站距离	到公园距离	蔬果店密度	PM _{2.5}
PAF 临界值	-20%	5%	5%	-20%	15%
含义	功能混合度达到 1.0	到地铁站的距离约 2 km，大约为大型地铁站服务半径	到公园的距离约 1 km，即 10 分钟生活圈	1000 米缓冲区内蔬果店有 20 家	PM _{2.5} 浓度为 35 $\mu\text{g}/\text{m}^3$ ，为《健康社区评价标准》中的浓度限值

然而，在这五个指标当中，PM_{2.5} 和到地铁站的距离属于较难改变的建成环境因素，因为前者受区域尺度的大气环境影响，后者受城市整体规划的约束；功能混合度属于中等难度，因为涉及到各类功能设施的配置；蔬果店密度和到公园的距离是相对较容易改变的因素。因此，优先考虑较易优化的功能混合度、蔬果店密度、到公园的距离这三项指标，可将待提升小区分成表 6.3 所示的十类，各类待提升小区的空间分布如图 6.13 所示。其中，第 1 至 3 类待提升小区为公园、蔬果店、功能混合度指标均判别为优的待提升小区，此时这些小区进一步考虑改建难度较高的地铁站和 PM_{2.5} 因素，即如何通过其他间接手段缓解这两个环境暴露因素带来的健康风险。而第 4 至 10 类待提升小区为公园、蔬果店、功能混合度至少有一项指标被判为差的待提升小区，因此应着重对该短板设计规划应对策略，此时不再考虑地铁站和 PM_{2.5} 这两个因素。

从表 6.3 可以看出，占比最高的待提升小区为第 5 类，占到 32.5%，该类小区由于蔬果店密度较低导致总 PAF 较高；其次是第 1 类，占到 30.6%，该类小区的主要短板是 PM_{2.5} 浓度较高；接下来依次是第 7 类、第 4 类 and 第 3 类，分别占 16.9%、9.7% 和 6.8%，其余类别占比均小于 5%，可以忽略。

表 6.3 待提升小区分类

类别	到公园距离	蔬果店密度	功能混合度	到地铁站距离	PM _{2.5}	个数(比例)
1	√	√	√	√	×	1602 (30.6%)
2	√	√	√	×	√	110 (2.1%)
3	√	√	√	×	×	356 (6.8%)
4	×	√	√			509 (9.7%)
5	√	×	√			1702 (32.5%)
6	√	√	×			6 (0.1%)
7	×	×	√			886 (16.9%)
8	×	√	×			0 (0.0%)
9	√	×	×			42 (0.8%)
10	×	×	×			20 (0.4%)
合计						5233 (100%)

注：√表示 PAF 比临界值更优，×表示 PAF 比临界值更差。

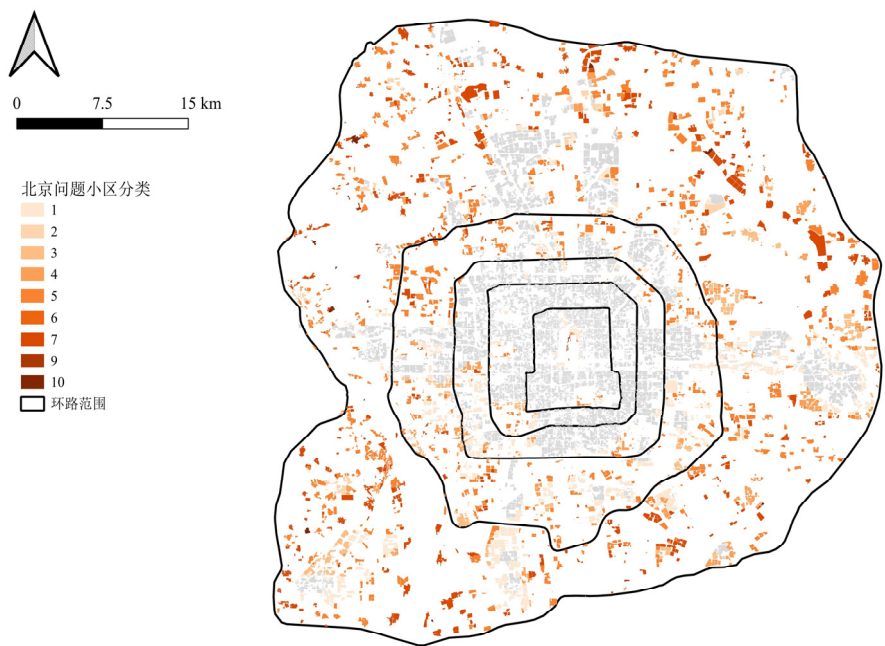


图 6.13 北京市六环内待提升小区的空间分布及分类结果

6.2.3 各类待提升小区的规划应对策略

本节针对十类待提升小区中排名前五（即出现频率>5%）的类别，即第 1、3、4、5、7 类分别给出规划应对策略及应对模式图，并举例典型小区进行具体分析。

第 1 类待提升小区

这类待提升小区的特征是建成环境指标整体良好，但 PM_{2.5} 浓度较高，导致 AMI 复发的总 PAF 较高。受 PM_{2.5} 浓度分布影响，这类待提升小区主要分布在北

京市的东郊与南郊，如图 6.14 所示。

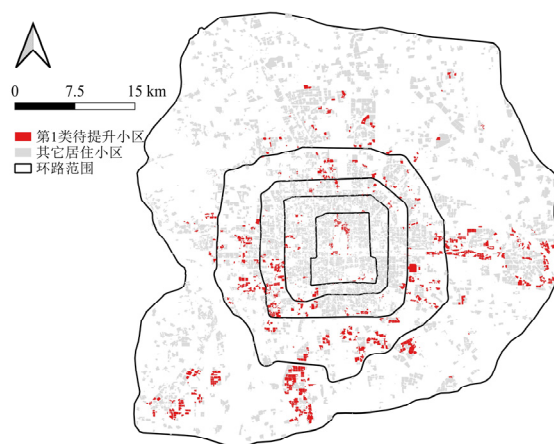


图 6.14 第 1 类待提升小区的空间分布

针对此类待提升小区，可以通过增设小区与污染源（如工业、主干道、垃圾场等）之间的防护绿地，来缓解 $PM_{2.5}$ 的大气污染迁移问题。余新晓等人^[152]将此防护绿地进一步分为颗粒物滞纳型和扩散型绿地，认为居住小区不同区域应该配置不同属性的防护绿地：

- （1）居住区外围边界绿地：选择滞尘能力强、耐瘠薄、耐干旱、寿命长、病虫害少的树种，特别是对于工业污染源的来流方向，需要额外增强绿地配置；
- （2）居住区道路绿地：选择滞尘能力强、分枝较高的乔木乡土树种，乔木下配植灌木；
- （3）居住区宅旁绿地：滞尘能力中等的林木为主，草本覆盖地面；
- （4）居住区广场游园绿地：半封闭，颗粒物来源方向乔灌木结合，下游方向留出通风口。

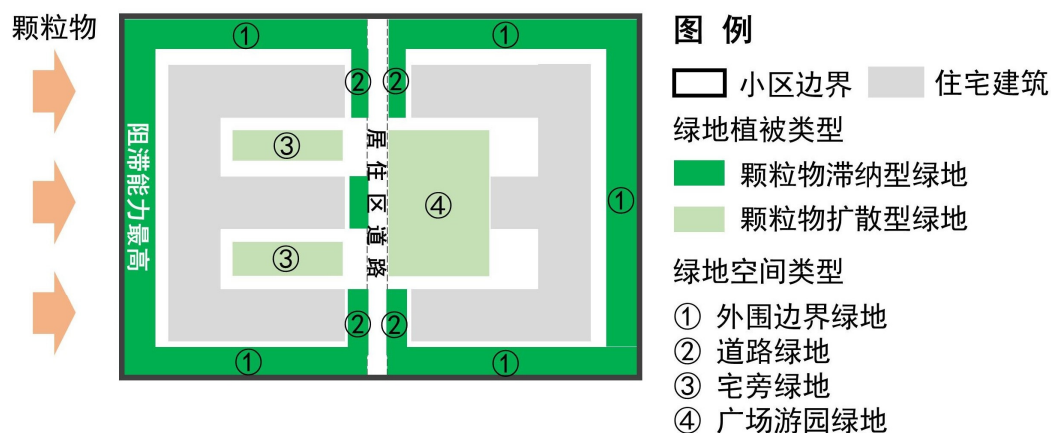


图 6.15 第 1 类待提升小区的规划应对模式图

分项	指标参数	
植物种类	上层	侧柏、圆柏、毛白杨、洋白蜡
	中层	金银木
	下层	草坪
群落特征	常绿阔叶比	1：3
	乔灌木比	2：1
	乔木郁闭度	>80%
	地被层盖度	>60%

图 6.16 居住区外围边界绿地配置方案^[152]

上述不同类型的居住小区防护绿地可以构成图 6.15 所示的应对模式图。余新晓等人^[152]还对上述各类小区内的绿地的植物配置给出了推荐组合，比如居住区外围边界绿地，给出了图 6.16 所示的植物配置方案。

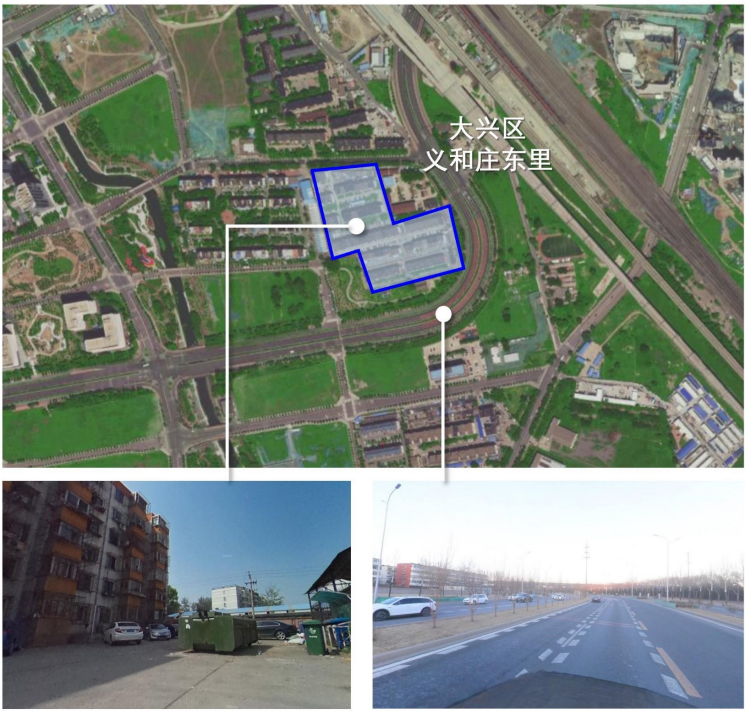


图 6.17 第 1 类待提升小区案例：大兴区义和庄东里

第 1 类待提升小区选取的案例是大兴区的义和庄东里小区，其遥感影像如图 6.17 所示。此处东南方向即为河北省及部分大兴区的众多工业污染源，PM_{2.5}污染较为严重。然而从小区东南方向兴华大街的街景图像可以发现，颗粒物来源方向的道路两侧视野极为开阔，多为空地或者草地，缺乏乔木树种，阻滞颗粒物能力较低，而小区与主干道之间也缺乏防护绿地配置。另外，从小区内部图片可以看

出，小区中心广场基本为水泥硬质铺装，缺乏宅边绿地或者广场游园绿地对颗粒物进行进一步阻滞，或引导通风方向，导致对 $\text{PM}_{2.5}$ 缓解作用较弱，并且，广场旁存在一处疏于管理的垃圾房，可能会造成额外的 $\text{PM}_{2.5}$ 污染源。因此，该小区可在兴华大街两侧配置高大乔木为主的防护绿地，在小区内部增设中等乔木与灌木结合的宅边绿地，并对垃圾房进行密闭管理、及时清运垃圾。

第3类待提升小区

这类待提升小区的特征是建成环境指标整体较好，但 $\text{PM}_{2.5}$ 浓度较高且到达地铁站的距离较远，导致 AMI 复发的总 PAF 较高。该类待提升小区的空间分布没有显著特征，主要聚集在通州区、房山区、大兴区的部分地区，如图 6.18 所示。

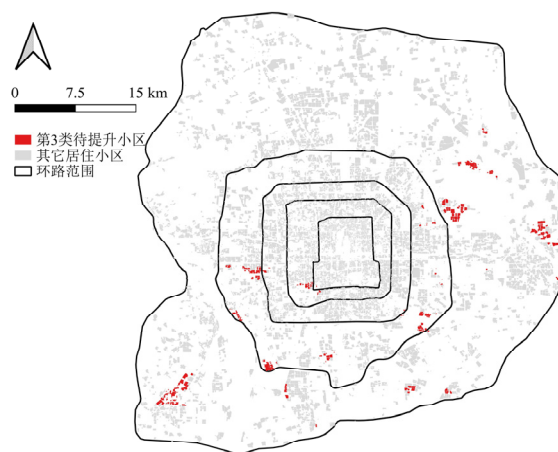


图 6.18 第3类待提升小区的空间分布

该类小区的 $\text{PM}_{2.5}$ 暴露缓解的规划应对策略，与第1类待提升小区相同，所以此处主要探讨对于地铁站距离较远的应对策略。对于地铁站距离较远的问题，为了鼓励社区人员采用健康出行方式到达就近的地铁站（或地铁站实在太远时到达就近的公交站），可以在社区出入口、地铁站口增设共享单车与自行车的停车位，并在主要道路上设置独立的自行车专用道，一方面实现机动车与非机动车分离，保障骑行车的交通安全，另一方面，提升自行车道及步行道的设施品质，增强居民的骑行或步行意愿。应对模式图如图 6.19 所示。

第3类待提升小区选取的案例是房山区碧桂园小区 A 区，其遥感影像如图 6.20 所示。该小区到达最近的地铁站——房山线广阳城站的路途距离是 2.8 km，骑行需要大约 16 分钟。然而，目前该小区到达广阳城地铁站的骑行环境较不友好，可能会大幅降低居民骑行到地铁站的意愿。小区大门所在支路原香路的街景图片中可以看到，道路宽度较为狭窄，且原先并未预留非机动车道，导致机非分离的栅

栏增设之后，非机动车车道异常狭窄，极不便于骑行。该区域到达广阳城地铁站的主干道长阳路的街景图片显示，自行车与机动车处于混行状态，无机非分离设施，并且部分路段没有专门设置的步行道，居民只能在机动车道上贴沿路边行走，交通安全隐患较大。因此，未来可改善长阳路主干道及附近支路的步行及骑行环境，或者增加短途接驳公交的发车频次，以解决到达地铁站的距离问题。

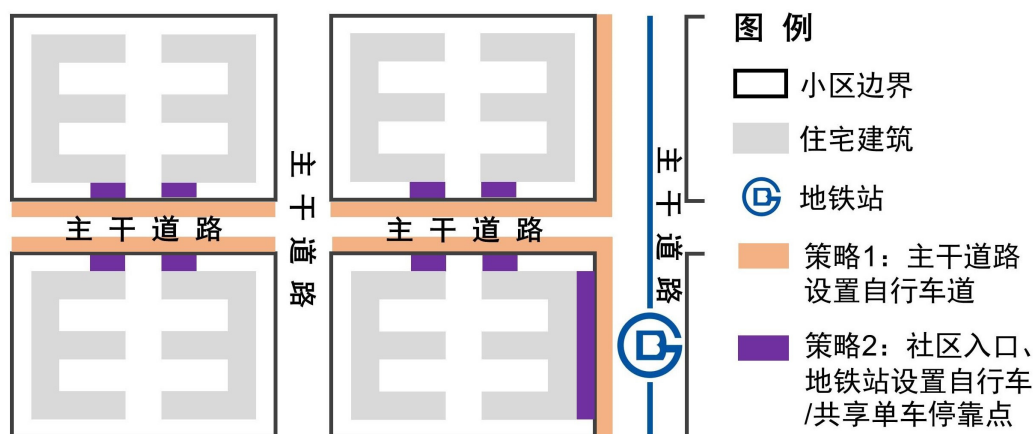


图 6.19 第3类待提升小区的规划应对模式图



图 6.20 第3类待提升小区案例：房山区碧桂园小区A区

第4类待提升小区

该类待提升小区的特征是建成环境指标整体较好，但到达公园的距离较远，导致 AMI 复发的总 PAF 较高。该类待提升小区的空间分布没有显著特征，主要聚集在各个方向的五环以外地区，五环以内的西南侧即丰台区也有部分该类待提升小区，如图 6.21 所示。对于这类待提升小区，可以考虑：

- (1) 完善绿地层级结构，除了乡镇街道层面的大型公园之外，注意增设街头绿地、口袋公园等小微绿地，及居住区内部社区层级的绿地和慢行步道；
- (2) 充分利用滨河、街角、街边等低效闲置用地，营造绿色空间
- (3) 注入更多种类功能的活动用地，采用慢行步道、骑行道进行串联，并可以构建绿廊、绿带将多级活动用地连接到邻近的大型城市公园。

该类待提升小区的规划应对模式图如图 6.22 所示。

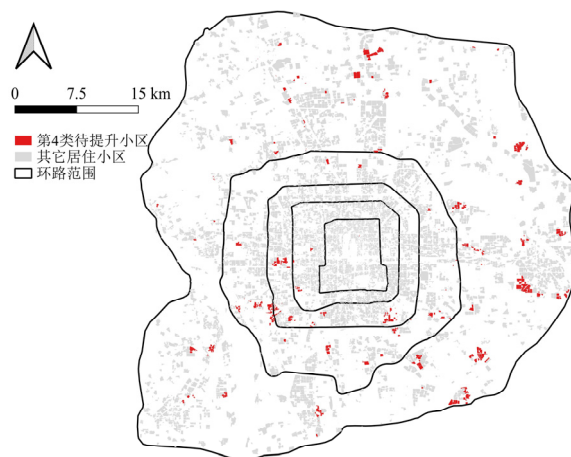


图 6.21 第4类待提升小区的空间分布

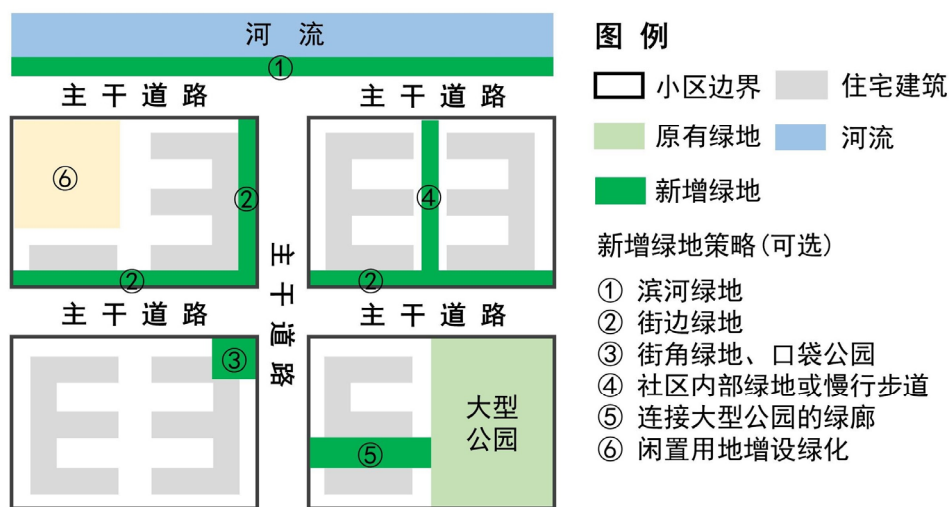


图 6.22 第4类待提升小区的规划应对模式图

第4类待提升小区选取的案例是朝阳区东坝家园B区，其遥感影像如图6.23所示。该小区附近的潜在绿色空间资源较为丰富。例如，东坝片区的西北角为大型城市公园——东坝千亩湖公园，风景优美，但是距离较远，与东坝中心片区缺乏联系。此外，小区北侧即为坝河，然而从街景图片可以看出，坝河的滨河空间用铁丝网进行封闭，居民无法到达滨水空间，且铁丝网后的绿化空间也缺乏规划设计，品质较差，绿地周边乱停车现象普遍，影响到了步行环境。此外，小区东南角为两条主干道——东坝中路与东坝中街的交叉口，交叉口附近存在闲置用地，也影响了城市美观。

因此，东坝家园附近区域，首先可以充分利用交叉口旁的小微空地，改造为居民休憩的街角或者口袋公园，从而弥补该地区缺乏微观层次绿地的不足。此外，可以充分利用坝河的滨水空间，将其开放给居民，优化提升绿化设计及品质，并以坝河滨水空间为绿廊纽带，连接东坝的中心住区与西北侧的千亩湖公园，为当地居民提供一个良好的步行环境，也促进当地绿化层级结构的完善。

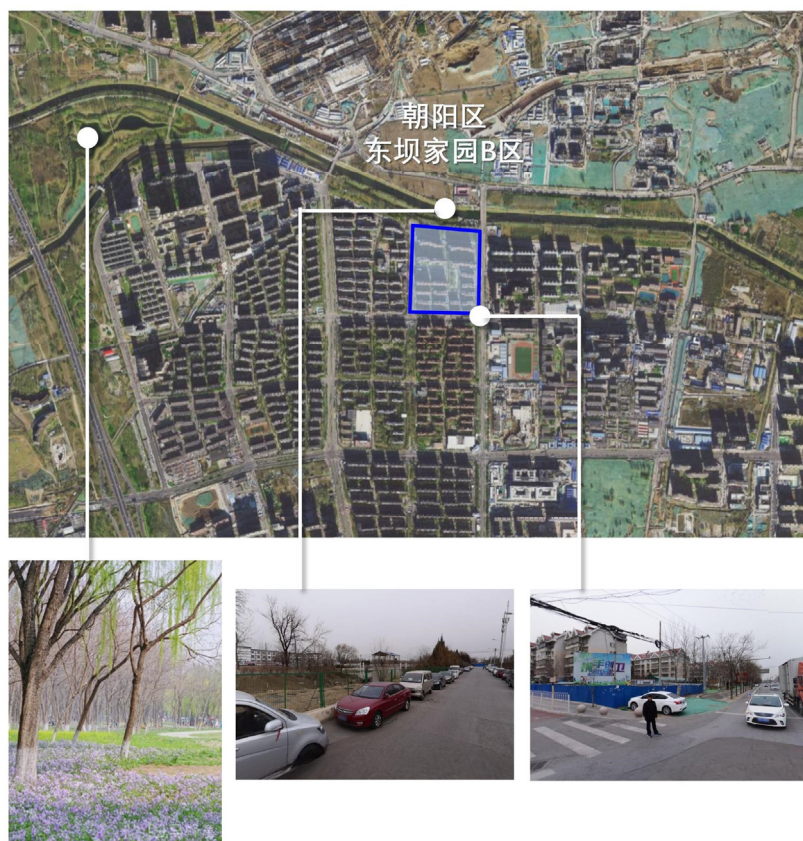


图 6.23 第4类待提升小区案例：朝阳区东坝家园B区

第5类待提升小区

该类待提升小区的特征是建成环境指标整体较好，但周边蔬果店密度较低，导致 AMI 复发的总 PAF 较高。该类待提升小区的占比最高，广泛分布于四环以外的大部分区域，如图 6.24 所示。对于该类待提升小区，结合已有标准导则给出的建议，可以考虑：

（1）小区内部可以采用已有的或者增设新的社区综合服务中心，通过“一店多能”的策略，在综合服务中心集约设置蔬果店；

（2）社区周边可以打开街道两侧空间，增设底商店铺，或利用周边腾退及闲置空间增设蔬果店等便民设施；

（3）若实在没有空间增设蔬果店，可以考虑在小区内部公共空间按时段增设临时蔬果集市或摊位。

该类待提升小区的规划应对模式图如图 6.25 所示。

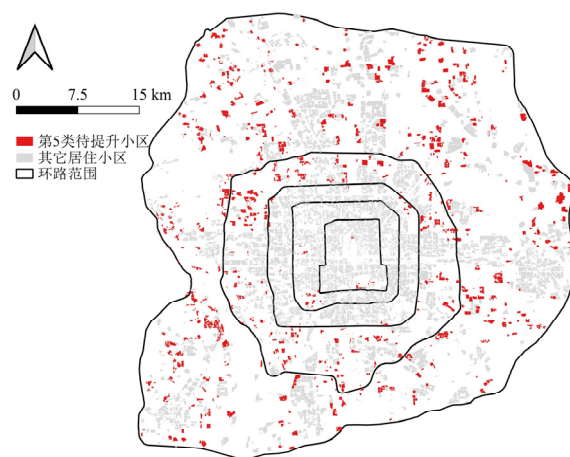


图 6.24 第5类待提升小区的空间分布

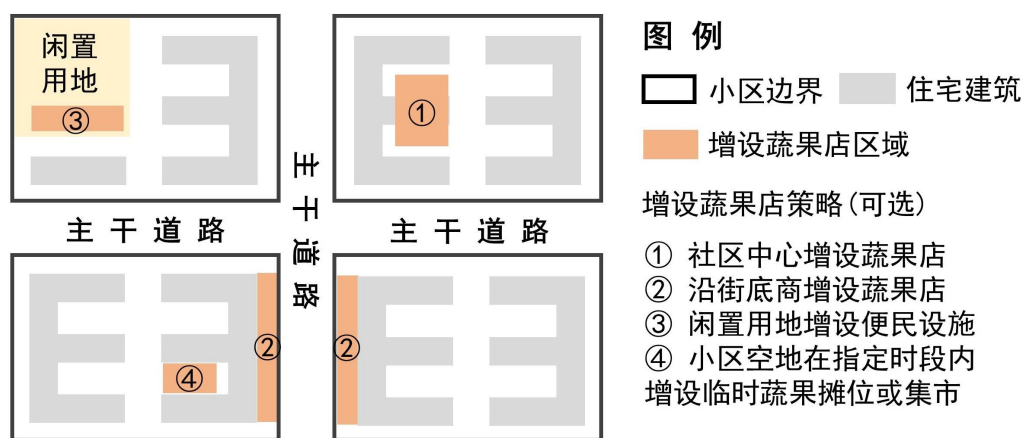


图 6.25 第5类待提升小区的规划应对模式图



图 6.26 第 5 类待提升小区案例：海淀区门头馨村北二区

第 5 类待提升小区选取的案例是海淀区门头馨村北二区，其平面图如图 6.26 所示，绿色圆点即为蔬果店 POI 的空间分布情况。整体而言，该小区周边的蔬果店密度相对较低，多分布在西侧主干道香山南路和南侧主干道闵庄路上，以及东侧门头馨园小区，但是该小区与门头馨园小区中间被西五环隔断，因此东部的蔬果店 POI 无法服务到该案例小区。从街景图片可以看出，小区南北两侧的支路均为交通性道路，两侧空间较为封闭，缺乏沿街商铺或公共空间设置蔬果店等功能设施。因此，可考虑在小区内部，或该小区与门头馨村北一区中间的草坪空地，增设蔬果店底商或者临时蔬果店摊位，以缓解该小区蔬果店可达性较差的问题。

第 7 类待提升小区

该类待提升小区的特征是到公园的距离较远，且周边蔬果店密度较低，导致 AMI 复发的总 PAF 较高。该类待提升小区可视为第 4 类和第 5 类待提升小区的复合版，如图 6.27 所示，其广泛分布于五环以外的大部分区域。对于该类待提升小区，公园绿地的提升策略可参考第 4 类待提升小区的应对策略，而蔬果店密度的

提升策略可参考第5类待提升小区的应对策略，以及这两类规划应对策略还可以进一步作整合考虑。比如，充分利用社区低效用地设置小微绿地的同时，可以在此处设置小型蔬果店或定期集市摊位，集约使用土地的同时还可以提升设施所在空间品质。该类待提升小区的规划应对模式图如图6.28所示。

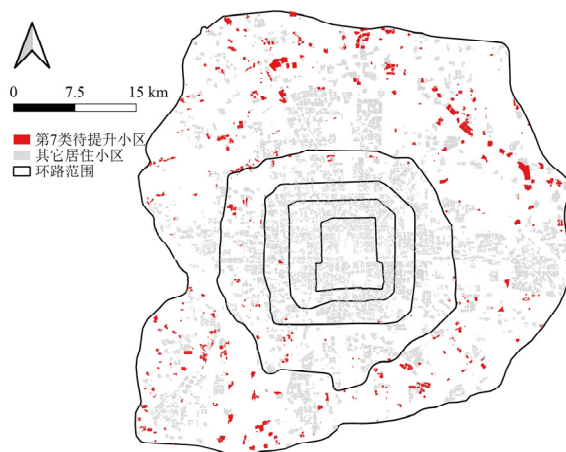


图 6.27 第7类待提升小区的空间分布

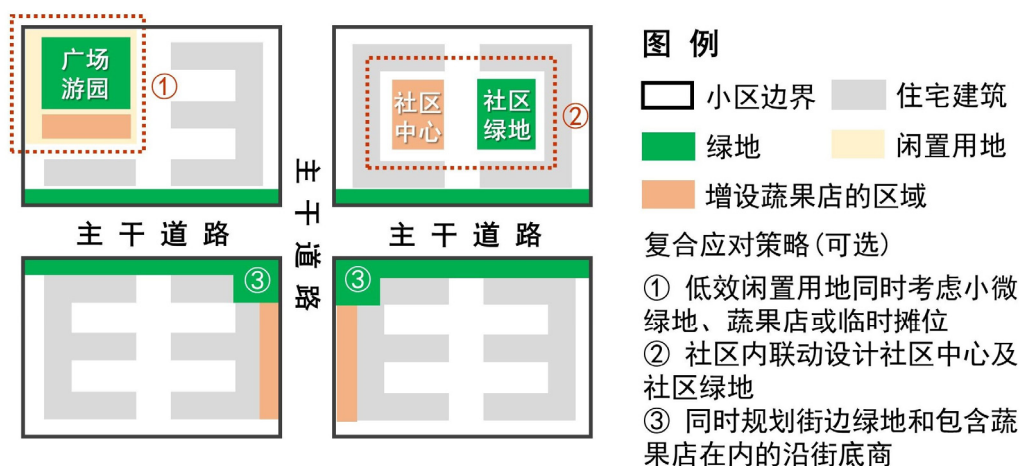


图 6.28 第7类待提升小区的规划应对模式图

第7类待提升小区选取的案例是昌平区二拨子新村，其遥感影像如图6.29所示。可以看到，二拨子新村处于西侧京新高速、南侧铁路及地铁13号线、东侧京藏高速的半包围地带当中，与周边地区交通联系较为割裂。小区西侧跨过京新高速的唐家岭地块有大块绿地，但是限于高速阻隔无法到达。小区南部有部分闲置空地，但目前被围墙阻隔，没有被充分利用，导致二拨子新村周边既没有大型城市公园，也没有小微绿地。现有蔬果店主要集中在小区东北侧的二拨子新村农贸市场，市场内蔬果店数量较少，且所在的公共空间虽然功能丰富，但是空间品质

较差。因此，该小区可以兼顾公园绿地设计及蔬果店两个方面。首先，农贸市场地块可以进行更新改造，将原先一至二层的农贸市场老旧建筑群改建为容积率更高的中高层农贸市场，将包括蔬果店在内的各类功能集约设置在其中，空余出的土地可进行社区绿地的规划设计，发挥此处位于周边社区中心地带的地理位置优势，同时可以优化农贸市场周围的空间品质。其次，小区南侧闲置空地也可以进行绿化设计与改造，并可以兼顾周边二拨子社区卫生服务站附近的公共区域一同更新设计，适当增设蔬果店等便民设施。



图 6.29 第 7 类待提升小区案例：昌平区二拨子新村

6.3 本章小结

本章基于第 5 章建成环境暴露与 AMI 复发之间的暴露反应关系，开展了北京市 2021 年建成环境的归因 AMI 复发风险评估，并依据此开展待提升小区的识别，提出对应的规划应对策略。本章的主要结论为：

(1) 北京市 2021 年归因于建成环境暴露的 AMI 复发风险，在乡镇街道尺度呈现出中心城区及郊区区县城关地区风险较低、其余偏远地区风险较高的空间分布，近 60% 的乡镇街道内建成环境对 AMI 复发呈整体保护作用；

(2) 北京市 2021 年归因于建成环境暴露的 AMI 复发风险，在六环内的居住小区尺度呈现出中心城区及四个方向放射状区域风险较低、其余地区风险较高的空间分布，近 90% 的六环内居住小区内建成环境对 AMI 复发呈整体保护作用；

(3) 本研究将 PAF 排在后 40% 的居住小区识别为待提升小区，按照五个环境因素的优劣情况，将待提升小区分为 10 类，其中前五名分别为第 5 类（蔬果店密度较低）、第 1 类（PM_{2.5} 浓度较高）、第 7 类（到公园距离较远、蔬果店密度较低）、第 4 类（到公园距离较远）和第 3 类（PM_{2.5} 浓度较高、到地铁站距离较远）；

(4) 本研究针对上述排名前五的待提升小区类别存在的问题，借鉴和参考了北京市相关标准与导则，分别探讨了其规划应对策略，并各选取一个典型居住小区案例开展了具体问题分析，以更好地支持城市管理者开展城市更新设计工作，有效降低建成环境暴露造成的 AMI 复发风险。

第7章 研究结论与展望

7.1 研究结论

本研究重点关注建成环境暴露对急性心肌梗死这种重要健康结局的复发的定量影响，试图回答第1章中提出的两个科学问题：建成环境暴露与急性心肌梗死复发之间的定量关联规律是什么？如何将以上定量规律应用到规划实践当中，从而降低居民的急性心肌梗死复发风险？

以往相关的文献研究往往存在以下问题：（1）建成环境的心血管流行病学队列研究较为匮乏，且多关注发病、死亡结局，而忽视预后复发结局；（2）建成环境的心血管流行病学研究未考虑多个建成环境要素同时暴露的影响；（3）目前健康城市建设仍缺乏定量的健康风险评估过程以辅助指导规划应对策略。为此，本论文针对以上科学问题开展相关研究，并尝试突破现有文献研究的局限性，主要研究贡献与结论为：

（1）采用地址解析、暴露组学、空间分析等手段，提出了根据患者住址文本信息确定其建成环境暴露评价的方法。本研究通过软件自动解析与人工审查修正相结合的方法，将患者住址文本信息转化为经纬度坐标，实现住址的空间落位；再根据已有文献综述，构建了密度、多样性、设计、到公共交通的距离、目的地可达性、空气污染共六大类、26个小类的建成环境暴露指标体系；结合暴露组学“全面性”、“动态性”的特征，选择了300米、500米、1000米三种不同缓冲区的空间尺度，以及基线暴露、累积平均暴露两种不同的时间尺度，对上述建成环境暴露指标进行了空间测度与暴露评价。本研究发现，多种建成环境暴露因素之间存在较高的相关性。

（2）采用回顾性队列研究设计及暴露组学分析方法，识别了与急性心肌梗死复发显著关联的建成环境暴露因素，揭示了建成环境与急性心肌梗死复发之间的暴露反应关系。患者队列以2013年至2019年北京市心血管疾病监测系统内的急性心肌梗死患者为研究对象，以患者的首次急性心肌梗死复发为首要结局开展分析。本研究首先通过统计学性能对比，选择ENET算法作为暴露组学变量选择方法，以克服建成环境因素多重共线性的问题。本研究进一步通过Cox模型分析发现：校正了个人属性、社会经济属性及相关并发症后，在基线暴露水平下，500米缓冲区内的城市功能混合度为显著保护因素，到地铁站的距离及PM_{2.5}是显著危险因素；在累积平均暴露水平下，500米缓冲区内的城市功能混合度、1000米缓冲区内的蔬果店密度为显著保护因素，到地铁站的距离、到公园的距离及PM_{2.5}是显

著危险因素。在致死型以及非致死型复发的次要结局中，无论是在基线暴露水平下，还是在累积平均暴露水平下，到地铁站的距离及 PM_{2.5} 均为显著危险因素；而 500 米缓冲区内的城市功能混合度仅在致死型复发结局中为显著保护因素。

(3) 借鉴风险因素归因疾病负担的评估方法，建立了建成环境的归因 AMI 复发风险评估方法，并依据此提出了待提升小区分类原则及相应的规划应对策略。本研究以北京市 2021 年的建成环境状况为例，基于暴露反应关系，确定了建成环境造成 AMI 复发的人口归因分数，研究发现：近 60% 的乡镇街道和近 90% 的六环内居住小区内，建成环境对 AMI 复发呈整体保护作用。本研究进一步对排名后 40% 的待提升小区进行分类，发现占比较高的待提升小区类别主要面临的是蔬果店缺乏、PM_{2.5} 浓度较高、到公园距离较远等问题。本研究分别给出了利用社区服务中心及腾退空间增设蔬果店、设置颗粒物滞纳型防护绿地、利用闲置小微空地及滨水空间营造多层次绿地结构等规划应对策略。

7.2 研究展望

本论文仍存在如下局限性，值得未来进一步深入研究：

(1) 本研究遵循以往建成环境流行病学的范式，采用研究对象住址周边不同的缓冲区范围进行暴露评价。然而，某类型城市功能在住址周边密度高，或者距离住址近，并不代表个人到达或使用这些功能设施的真实暴露水平就高，这两者之间可能存在偏差。此外，由于建成环境大数据精度所限，一些小微尺度的建成环境因素可能难以捕集，如绿地用地数据为全市尺度，而不能涵盖人可能经常到达的小区内部绿地。因此，随着穿戴式相机等新兴穿戴设备的出现，未来研究可以采用穿戴设备来进行个人层面建成环境真实暴露水平的动态评价。

(2) 本研究聚焦于建成环境暴露，因此暴露组学的“全面性”体现在尽可能多地选择潜在的建成环境暴露因素。然而，人还同时暴露在建成环境以外的众多环境因素中，如室内空气化学污染、室外气象参数、噪声等。本研究由于队列数据有限，未能对这些环境暴露因素都纳入考虑。未来研究可以考虑一个更为全面的环境暴露组，来系统评价这其中建成环境在某些健康结局发生风险中扮演的角色及其重要性。

(3) 本研究受数据所限，调整的协变量中未考虑患者首次急性心肌梗死的治疗方案、患者重要的行为习惯（如饮食习惯）、患者遗传因素及生理指标因素（如吸烟、BMI 等），可能会对关联分析结果带来偏差。未来研究可以在队列随访过程中补充这些协变量，以获得更为准确的暴露反应关系。

(4) 本研究基于的暴露组学分析方法主要是针对线性模型而言的，因此本研

究也只探讨了建成环境暴露与急性心肌梗死复发风险之间的线性关联关系。关于非线性关联，如常用的自然立方样条方法，怎样与暴露组学分析方法进行结合，实现方法学的范式创新，是值得未来探索的研究方向。

（5）本研究采用的建成环境的归因风险评估方法，目前仅应用于急性心肌梗死复发事件，未来可以在丰富建成环境暴露反应关系（如其他心血管疾病、呼吸系统疾病等）的数据基础之后，考虑多种健康结局的 PAF 加和，从而构建定量的健康城市指标评分体系。本研究开展的规划应对策略，目前仅探讨了居住小区尺度，未来可基于城市整体评估结果，围绕乡镇街道或街区尺度的归因风险较高区域，提出针对性和整合性的规划应对策略和建议。

参考文献

- [1] 中国建筑学会. 健康建筑评价标准 T/ASC02-2021: [S]. 北京: 中国建筑工业出版社, 2021.
- [2] 中国工程建设标准化协会, 中国城市科学研究会. 健康社区评价标准 T/CECS 650-2020: [S]. 北京: 中国计划出版社, 2020.
- [3] 中国建筑学会. 健康小镇评价标准 T/ASC12-2020: [S]. 北京: 中国建筑工业出版社, 2020.
- [4] 国家统计局. 人口出生率、死亡率和自然增长率 [Z]. 2023
- [5] Vos T, Lim S S, Abbafati C, et al. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019 [J]. *Lancet*, 2020, 396(10258): 1204-1222.
- [6] 《中国心血管健康与疾病报告 2021》编写组. 《中国心血管健康与疾病报告 2021》要点解读 [J]. *中国心血管杂志*, 2022, 27: 305-318.
- [7] 孙佳艺, 张倩, 赵冬, 等. 北京市 2007-2012 年急性心肌梗死住院患者 30 天病死率变化趋势分析 [J]. *中华流行病学杂志*, 2018, 39(3): 363-367.
- [8] Xu H Y, Yang Y J, Wang C S, et al. Association of Hospital-Level Differences in Care With Outcomes Among Patients With Acute ST-Segment Elevation Myocardial Infarction in China [J]. *Jama Netw Open*, 2020, 3(10): e2021677.
- [9] Chen L, Zhang Z X, Long Y. Association between leisure-time physical activity and the built environment in China: Empirical evidence from an accelerometer and GPS-based fitness app [J]. *Plos One*, 2021, 16(12): e0260570.
- [10] Li X, Li Y, Jia T, et al. The six dimensions of built environment on urban vitality: Fusion evidence from multi-source data [J]. *Cities*, 2022, 121: 103482.
- [11] Carlson T N, Ripley D A. On the relation between NDVI, fractional vegetation cover, and leaf area index [J]. *Remote Sens Environ*, 1997, 62(3): 241-252.
- [12] Weng Q H. Thermal infrared remote sensing for urban climate and environmental studies: Methods, applications, and trends [J]. *Isprs J Photogramm*, 2009, 64(4): 335-344.
- [13] Hammer M S, Van Donkelaar A, Li C, et al. Global Estimates and Long-Term Trends of Fine Particulate Matter Concentrations (1998-2018) [J]. *Environ Sci Technol*, 2020, 54(13): 7879-7890.
- [14] Cooper M J, Martin R V, McInden C A, et al. Inferring ground-level nitrogen dioxide concentrations at fine spatial resolution applied to the TROPOMI satellite instrument [J]. *Environ Res Lett*, 2020, 15(10): 104013.
- [15] Han Q Z, Yin Q, Zheng X, et al. Remote sensing image building detection method based on Mask R-CNN [J]. *Complex Intell Syst*, 2022, 8(3): 1847-1855.

- [16] Helbich M, Yao Y, Liu Y, et al. Using deep learning to examine street view green and blue spaces and their associations with geriatric depression in Beijing, China [J]. *Environ Int*, 2019, 126: 107-117.
- [17] Li X J, Zhang C R, Li W D, et al. Assessing street-level urban greenery using Google Street View and a modified green view index [J]. *Urban for Urban Gree*, 2015, 14(3): 675-685.
- [18] 何宛余, 李春, 聂广洋, 等. 深度学习在城市感知的应用可能——基于卷积神经网络的图像判别分析 [J]. *国际城市规划*, 2019, 34(1): 8-17.
- [19] 城乡规划学名词审定委员会. 城乡规划学名词 [M]. 北京: 科学出版社, 2021.
- [20] 谭纵波. 城市规划 (修订版) [M]. 北京: 清华大学出版社, 2016.
- [21] Ewing R, Cervero R. Travel and the Built Environment [J]. *J Am Plann Assoc*, 2010, 76(3): 265-294.
- [22] Zhang Y, Liu N, Li Y, et al. Neighborhood infrastructure-related risk factors and non-communicable diseases: a systematic meta-review [J]. *Environment Health*, 2023, 22: 2.
- [23] Maas J, Verheij R A, De Vries S, et al. Morbidity is related to a green living environment [J]. *J Epidemiol Commun H*, 2009, 63(12): 967-973.
- [24] Mitchell R, Popham F. Effect of exposure to natural environment on health inequalities: an observational population study [J]. *Lancet*, 2008, 372(9650): 1655-1660.
- [25] Richardson E A, Mitchell R. Gender differences in relationships between urban green space and health in the United Kingdom [J]. *Soc Sci Med*, 2010, 71(3): 568-575.
- [26] Bixby H, Hodgson S, Fortunato L, et al. Associations between Green Space and Health in English Cities: An Ecological, Cross-Sectional Study [J]. *Plos One*, 2015, 10(3): e119495.
- [27] Lachowycz K, Jones A P. Does walking explain associations between access to greenspace and lower mortality? [J]. *Soc Sci Med*, 2014, 107: 9-17.
- [28] Richardson E, Pearce J, Mitchell R, et al. The association between green space and cause-specific mortality in urban New Zealand: an ecological analysis of green space utility [J]. *Bmc Public Health*, 2010, 10: 240.
- [29] Richardson E A, Mitchell R, Hartig T, et al. Green cities and health: a question of scale? [J]. *J Epidemiol Commun H*, 2012, 66(2): 160-165.
- [30] Hu Z Y, Liebens J, Rao K R. Linking stroke mortality with air pollution, income, and greenness in northwest Florida: An ecological geographical study [J]. *Int J Health Geogr*, 2008, 7: 20.
- [31] Tamosiunas A, Grazuleviciene R, Luksiene D, et al. Accessibility and use of urban green spaces, and cardiovascular health: findings from a Kaunas cohort study [J]. *Environ Health-Glob*, 2014, 13(1): 20.
- [32] Villeneuve P J, Jerrett M, Su J G, et al. A cohort study relating urban green space with mortality in Ontario, Canada [J]. *Environ Res*, 2012, 115: 51-58.
- [33] Wilker E H, Wu C D, Mcneely E, et al. Green space and mortality following ischemic stroke [J]. *Environ Res*, 2014, 133: 42-48.

-
- [34] Chen H, Burnett R T, Bai L, et al. Residential Greenness and Cardiovascular Disease Incidence, Readmission, and Mortality [J]. *Environ Health Persp*, 2020, 128(8): 087005.
- [35] James P, Hart J E, Banay R F, et al. Exposure to Greenness and Mortality in a Nationwide Prospective Cohort Study of Women [J]. *Environ Health Persp*, 2016, 124(9): 1344-1352.
- [36] Pereira G, Foster S, Martin K, et al. The association between neighborhood greenness and cardiovascular disease: an observational study [J]. *Bmc Public Health*, 2012, 12: 466.
- [37] Wang K F, Lombard J, Rundek T, et al. Relationship of Neighborhood Greenness to Heart Disease in 249 405 US Medicare Beneficiaries [J]. *J Am Heart Assoc*, 2019, 8(6): e010258.
- [38] Twohig-Bennett C, Jones A. The health benefits of the great outdoors: A systematic review and meta- analysis of greenspace exposure and health outcomes [J]. *Environ Res*, 2018, 166: 628-637.
- [39] Gascon M, Triguero-Mas M, Martinez D, et al. Residential green spaces and mortality: A systematic review [J]. *Environ Int*, 2016, 86: 60-67.
- [40] James P, Banay R F, Hart J E, et al. A Review of the Health Benefits of Greenness [J]. *Curr Epidemiol Rep*, 2015, 2(2): 131-142.
- [41] Malambo P, Kengne A P, De Villiers A, et al. Built Environment, Selected Risk Factors and Major Cardiovascular Disease Outcomes: A Systematic Review [J]. *Plos One*, 2016, 11(11): e0166846.
- [42] Salgado M, Madureira J, Mendes A S, et al. Environmental determinants of population health in urban settings. A systematic review [J]. *Bmc Public Health*, 2020, 20(1): 853.
- [43] Chandrabose M, Rachele J N, Gunn L, et al. Built environment and cardio-metabolic health: systematic review and meta-analysis of longitudinal studies [J]. *Obes Rev*, 2019, 20(1): 41-54.
- [44] Kan H D, Heiss G, Rose K M, et al. Prospective Analysis of Traffic Exposure as a Risk Factor for Incident Coronary Heart Disease: The Atherosclerosis Risk in Communities (ARIC) Study [J]. *Environ Health Persp*, 2008, 116(11): 1463-1468.
- [45] Chum A, O'campo P. Cross-sectional associations between residential environmental exposures and cardiovascular diseases [J]. *Bmc Public Health*, 2015, 15: 438.
- [46] Kwan S C, Ismail R, Ismail N H, et al. An ecological study of the relationship between urban built environment and cardiovascular hospital admissions (2004-2016) in an Asian developing country [J]. *Soc Sci Med*, 2021, 276: 113868.
- [47] Hamano T, Kawakami N, Li X J, et al. Neighbourhood Environment and Stroke: A Follow-Up Study in Sweden [J]. *Plos One*, 2013, 8(2): e56680.
- [48] Lee E Y, Choi J, Lee S, et al. Objectively Measured Built Environments and Cardiovascular Diseases in Middle-Aged and Older Korean Adults [J]. *Int J Env Res Pub He*, 2021, 18(4): 1861.
- [49] Mobley L R, Root E D, Finkelstein E A, et al. Environment, obesity, and cardiovascular disease risk in low-income women [J]. *Am J Prev Med*, 2006, 30(4): 327-332.

- [50] Cebrecos A, Escobar F, Borrell L N, et al. A multicomponent method assessing healthy cardiovascular urban environments: The Heart Healthy Hoods Index [J]. *Health Place*, 2019, 55: 111-119.
- [51] Griffin B A, Eibner C, Bird C E, et al. The relationship between urban sprawl and coronary heart disease in women [J]. *Health Place*, 2013, 20: 51-61.
- [52] Servadio J L, Lawal A S, Davis T, et al. Demographic Inequities in Health Outcomes and Air Pollution Exposure in the Atlanta Area and its Relationship to Urban Infrastructure [J]. *J Urban Health*, 2019, 96(2): 219-234.
- [53] Jia X J, Yu Y, Xia W N, et al. Cardiovascular diseases in middle aged and older adults in China: the joint effects and mediation of different types of physical exercise and neighborhood greenness and walkability [J]. *Environ Res*, 2018, 167: 175-183.
- [54] World Health Organization (Who). WHO global air quality guidelines: Particulate matter (PM_{2.5} and PM₁₀), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide [M]. Bonn, Germany: WHO European Centre for Environment and Health, 2021.
- [55] Burnett R T, Pope C A, 3rd, Ezzati M, et al. An integrated risk function for estimating the global burden of disease attributable to ambient fine particulate matter exposure [J]. *Environ Health Perspect*, 2014, 122(4): 397-403.
- [56] Murray C J L, Aravkin A Y, Zheng P, et al. Global burden of 87 risk factors in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019 [J]. *Lancet*, 2020, 396(10258): 1223-1249.
- [57] Chen Z R, Liu N R, Tang H, et al. Health effects of exposure to sulfur dioxide, nitrogen dioxide, ozone, and carbon monoxide between 1980 and 2019: A systematic review and meta-analysis [J]. *Indoor Air*, 2022, 32(11): e13170.
- [58] Niu Y, Zhou Y C, Chen R J, et al. Long-term exposure to ozone and cardiovascular mortality in China: a nationwide cohort study [J]. *Lancet Planet Health*, 2022, 6(6): E496-E503.
- [59] Strak M, Weinmayr G, Rodopoulou S, et al. Long term exposure to low level air pollution and mortality in eight European cohorts within the ELAPSE project: pooled analysis [J]. *British Medical Journal*, 2021, 374: n1904.
- [60] Wild C P. The exposome: from concept to utility [J]. *Int J Epidemiol*, 2012, 41(1): 24-32.
- [61] Wild C P. Complementing the genome with an "exposome": The outstanding challenge of environmental exposure measurement in molecular epidemiology [J]. *Cancer Epidebm Biomar*, 2005, 14(8): 1847-1850.
- [62] 白志鹏, 陈莉, 韩斌. 暴露组学的概念与应用 [J]. *环境与健康杂志*, 2015, 32(1): 1-9.
- [63] Louis G M B, Schisterman E F, Sweeney A M, et al. Designing prospective cohort studies for assessing reproductive and developmental toxicity during sensitive windows of human reproduction and development - the LIFE Study [J]. *Paediatr Perinat Ep*, 2011, 25(5): 413-424.
- [64] Maitre L, De Bont J, Casas M, et al. Human Early Life Exposome (HELIX) study: a European population-based exposome cohort [J]. *BMJ Open*, 2018, 8(9): e021311.

-
- [65] Timmermans E J, Lakerveld J, Beulens J W J, et al. Cohort profile: the Geoscience and Health Cohort Consortium (GECCO) in the Netherlands [J]. *BMJ Open*, 2018, 8(6): e021597.
- [66] Casas M, Basagana X, Sakhi A K, et al. Variability of urinary concentrations of non-persistent chemicals in pregnant women and school-aged children [J]. *Environ Int*, 2018, 121: 561-573.
- [67] Chung M K, Kannan K, Louis G M, et al. Toward Capturing the Exposome: Exposure Biomarker Variability and Coexposure Patterns in the Shared Environment [J]. *Environ Sci Technol*, 2018, 52(15): 8801-8810.
- [68] Jiang C, Wang X, Li X Y, et al. Dynamic Human Environmental Exposome Revealed by Longitudinal Personal Monitoring [J]. *Cell*, 2018, 175(1): 277-291.
- [69] Paglia G, Miedico O, Tarallo M, et al. Evaluation of Seasonal Variability of Toxic and Essential Elements in Urine Analyzed by Inductively Coupled Plasma Mass Spectrometry [J]. *Expos Health*, 2017, 9(2): 79-88.
- [70] Haug L S, Sakhi A K, Cequier E, et al. In-utero and childhood chemical exposome in six European mother-child cohorts [J]. *Environ Int*, 2018, 121: 751-763.
- [71] Robinson O, Basagana X, Agier L, et al. The Pregnancy Exposome: Multiple Environmental Exposures in the INMA-Sabadell Birth Cohort [J]. *Environ Sci Technol*, 2015, 49(17): 10632-10641.
- [72] Tamayo-Uria I, Maitre L, Thomsen C, et al. The early-life exposome: Description and patterns in six European countries [J]. *Environ Int*, 2019, 123: 189-200.
- [73] Jamin E L, Bonvallot N, Tremblay-Franco M, et al. Untargeted profiling of pesticide metabolites by LC-HRMS: an exposomics tool for human exposure evaluation [J]. *Anal Bioanal Chem*, 2014, 406(4): 1149-1161.
- [74] Maitre L, Lau C H E, Vizcaino E, et al. Assessment of metabolic phenotypic variability in children's urine using H-1 NMR spectroscopy [J]. *Sci Rep-Uk*, 2017, 7: 46082.
- [75] Maitre L, Robinson O, Martinez D, et al. Urine Metabolic Signatures of Multiple Environmental Pollutants in Pregnant Women: An Exposome Approach [J]. *Environ Sci Technol*, 2018, 52(22): 13469-13480.
- [76] Yao Y Y, Wang P G, Shao G, et al. Nontargeted analysis of the urine nonpolar sulfateome: a pathway to the nonpolar xenobiotic exposome [J]. *Rapid Commun Mass Sp*, 2016, 30(21): 2341-2350.
- [77] Robinson O, Tamayo I, De Castro M, et al. The Urban Exposome during Pregnancy and Its Socioeconomic Determinants [J]. *Environ Health Persp*, 2018, 126(7): 077005.
- [78] Haddad N, Andrianou X D, Makris K C. A Scoping Review on the Characteristics of Human Exposome Studies [J]. *Curr Pollut Rep*, 2019, 5(4): 378-393.
- [79] Santos S, Maitre L, Warembourg C, et al. Applying the exposome concept in birth cohort research: a review of statistical approaches [J]. *Eur J Epidemiol*, 2020, 35(3): 193-204.
- [80] Hastie T, Tibshirani R, Friedman J. The elements of statistical learning: Data mining, inference, and prediction [M]. New York: Springer, 2009.

- [81] Agier L, Portengen L, Chadeau-Hyam M, et al. A Systematic Comparison of Linear Regression-Based Statistical Methods to Assess Exposome-Health Associations [J]. *Environ Health Persp*, 2016, 124(12): 1848-1856.
- [82] Andrianou X D, Makris K C. The framework of urban exposome: Application of the exposome concept in urban health studies [J]. *Sci Total Environ*, 2018, 636: 963-967.
- [83] Nieuwenhuijsen M J, Agier L, Basagana X, et al. Influence of the Urban Exposome on Birth Weight [J]. *Environ Health Persp*, 2019, 127(4): 047007.
- [84] Vrijheid M, Fossati S, Maitre L, et al. Early-Life Environmental Exposures and Childhood Obesity: An Exposome-Wide Approach [J]. *Environ Health Persp*, 2020, 128(6): 067009.
- [85] Agier L, Basagana X, Maitre L, et al. Early-life exposome and lung function in children in Europe: an analysis of data from the longitudinal, population-based HELIX cohort [J]. *Lancet Planet Health*, 2019, 3(2): E81-E92.
- [86] Warembourg C, Nieuwenhuijsen M, Ballester F, et al. Urban environment during early-life and blood pressure in young children [J]. *Environ Int*, 2021, 146: 106174.
- [87] North M L, Brook J R, Lee E Y, et al. The Kingston Allergy Birth Cohort Exploring parentally reported respiratory outcomes through the lens of the exposome [J]. *Ann Allerg Asthma Im*, 2017, 118(4): 465-473.
- [88] Song J, Du P, Yi W Z, et al. Using an Exposome-Wide Approach to Explore the Impact of Urban Environments on Blood Pressure among Adults in Beijing-Tianjin-Hebei and Surrounding Areas of China [J]. *Environ Sci Technol*, 2022, 56(12): 8395-8405.
- [89] 王兰, 凯瑟琳·罗斯. 健康城市规划与评估: 兴起与趋势 [J]. *国际城市规划*, 2016, 31(4): 1-3.
- [90] 杨瑞, 欧阳伟, 田莉. 城市规划与公共健康的渊源、发展与演进 [J]. *上海城市规划*, 2018, 3: 79-85.
- [91] Giles-Corti B, Vernez-Moudon A, Reis R, et al. City planning and population health: a global challenge [J]. *Lancet*, 2016, 388(10062): 2912-2924.
- [92] Nieuwenhuijsen M J. Influence of urban and transport planning and the city environment on cardiovascular disease [J]. *Nat Rev Cardiol*, 2018, 15(7): 432-438.
- [93] 王兰, 廖舒文, 赵晓菁. 健康城市规划路径与要素辨析 [J]. *国际城市规划*, 2016, 31(4): 4-9.
- [94] 李经纬, 欧阳伟, 田莉. 建成环境对公共健康影响的尺度与方法研究 [J]. *上海城市规划*, 2020, 2: 38-43.
- [95] 田莉, 李经纬, 欧阳伟, 等. 城乡规划与公共健康的关系及跨学科研究框架构想 [J]. *城市规划学刊*, 2016, 2: 111-116.
- [96] 王兰, 蒋希冀, 孙文尧, 等. 城市建成环境对呼吸健康的影响及规划策略——以上海市某城区为例 [J]. *城市规划*, 2018, 42(6): 15-22.
- [97] 王一. 健康城市导向下的社区规划 [J]. *规划师*, 2015, 31(10): 101-105.

- [98] 唐燕, 梁思思, 郭磊贤. 通向“健康城市”的邻里规划——《塑造邻里:为了地方健康和全球可持续性》引介 [J]. 国际城市规划, 2014, 29(6): 120-125.
- [99] 张天尧. 生态学视角下健康城市规划理论框架的构建 [J]. 规划师, 2015, 31(6): 20-26.
- [100] 普蕾米拉·韦伯斯特, 丹尼丝·桑德森, 徐望悦, 等. 健康城市指标——衡量健康的适当工具? [J]. 国际城市规划, 2016, 31(4): 27-31.
- [101] 安·福赛思, 卡丽莎·希弗利·斯洛特巴克, 凯文·克里泽克, 等. 健康影响评估之规划师版:哪些工具有用? [J]. 国际城市规划, 2016, 31(4): 32-43.
- [102] Xie W X, Li G, Zhao D, et al. Relationship between fine particulate air pollution and ischaemic heart disease morbidity and mortality [J]. Heart, 2015, 101(4): 257-263.
- [103] Chen H, Burnett R T, Copes R, et al. Ambient Fine Particulate Matter and Mortality among Survivors of Myocardial Infarction: Population-Based Cohort Study [J]. Environ Health Persp, 2016, 124(9): 1421-1428.
- [104] Kojima S, Michikawa T, Ueda K, et al. Asian dust exposure triggers acute myocardial infarction [J]. Eur Heart J, 2017, 38(43): 3202-3208.
- [105] Hu C Y, Schottker B, Venisse N, et al. Bisphenol A, Chlorinated Derivatives of Bisphenol A and Occurrence of Myocardial Infarction in Patients with Type 2 Diabetes: Nested Case-Control Studies in Two European Cohorts [J]. Environ Sci Technol, 2019, 53(16): 9876-9883.
- [106] Bergkvist C, Berglund M, Glynn A, et al. Dietary exposure to polychlorinated biphenyls and risk of myocardial infarction in men - A population-based prospective cohort study [J]. Environ Int, 2016, 88: 9-14.
- [107] Bhaskaran K, Hajat S, Armstrong B, et al. The effects of hourly differences in air pollution on the risk of myocardial infarction: case crossover analysis of the MINAP database [J]. British Medical Journal, 2011, 343: d5531.
- [108] Bai L, Shin S, Burnett R T, et al. Exposure to ambient air pollution and the incidence of congestive heart failure and acute myocardial infarction: A population-based study of 5.1 million Canadian adults living in Ontario [J]. Environ Int, 2019, 132: 105004.
- [109] Lipsett M J, Ostro B D, Reynolds P, et al. Long-Term Exposure to Air Pollution and Cardiorespiratory Disease in the California Teachers Study Cohort [J]. Am J Resp Crit Care, 2011, 184(7): 828-835.
- [110] Bally M, Dendukuri N, Rich B, et al. Risk of acute myocardial infarction with NSAIDs in real world use: bayesian meta-analysis of individual patient data [J]. British Medical Journal, 2017, 357: j1909.
- [111] Li Y, Zhang Y, Suel E, et al. Beijing neighborhood socioeconomic status dataset [DS]. 2022,
- [112] 北京市城市规划设计研究院. 城乡规划用地分类标准 DB 11/996-2013: [S]. 北京: 中国建筑工业出版社, 2013.
- [113] 中华人民共和国住房和城乡建设部. 城市用地分类与规划建设用地标准 GB 50137-2011: [S]. 北京: 中国建筑工业出版社, 2010.

-
- [114] Chen J J, Chen L, Li Y, et al. Measuring Physical Disorder in Urban Street Spaces: A Large-Scale Analysis Using Street View Images and Deep Learning [J]. *Ann Am Assoc Geogr*, 2022, 113(2): 469-487.
 - [115] Worldpop (Www.Worldpop.Org - School of Geography and Environmental Science, University of Southampton; Department of Geography and Geosciences,, University of Louisville; Departement De Geographie,, Universite De Namur), Center for International Earth Science Information Network (Ciesin), Columbia University. Global High Resolution Population Denominators Project (Funded by The Bill and Melinda Gates Foundation (OPP1134076)) [DS]. 2018,
 - [116] Su Y W, Liu Z Y, Chang J, et al. Measuring Accessibility to Health Care Using Taxi Trajectories Data: A Case Study of Acute Myocardial Infarction Cases in Beijing (Online ahead of print) [J]. *Int J Health Policy*, 2022.
 - [117] 黎珍惜, 黎家勋. 基于经纬度快速计算两点间距离及测量误差 [J]. *测绘与空间地理信息*, 2013, 36(11): 235-237.
 - [118] Long Y, Liu X J. Featured graphic. How mixed is Beijing, China? A visual exploration of mixed land use [J]. *Environ Plann A*, 2013, 45(12): 2797-2798.
 - [119] Sallis J, Cerin E, Conway T L. Physical activity in relation to urban environments in 14 cities worldwide: a cross-sectional study [J]. *Lancet*, 2016, 387(10034): 2207-2217.
 - [120] Crouse D L, Pinault L, Balram A, et al. Urban greenness and mortality in Canada's largest cities: a national cohort study [J]. *Lancet Planet Health*, 2017, 1(7): E289-E297.
 - [121] Ji J S, Zhu A, Bai C, et al. Residential greenness and mortality in oldest-old women and men in China: a longitudinal cohort study [J]. *Lancet Planet Health*, 2019, 3(1): E17-E25.
 - [122] Zou H, Hastie T. Regularization and variable selection via the elastic net [J]. *J R Stat Soc B*, 2005, 67: 301-320.
 - [123] Chun H, Keles S. Sparse partial least squares regression for simultaneous dimension reduction and variable selection [J]. *J R Stat Soc B*, 2010, 72: 3-25.
 - [124] Sinisi S E, Van Der Laan M J. Loss-Based Cross-Validated Deletion/Substitution/Addition Algorithms in Estimation [J]. UC Berkeley Division of Biostatistics Working Paper Series, 2004, Working paper 143.
 - [125] Klein J P, Moeschberger M L. Survival analysis: Techniques for censored and truncated data (second edition) [M]. New York: Springer, 2003.
 - [126] Baulies S, Belin L, Mallon P, et al. Time-varying effect and long-term survival analysis in breast cancer patients treated with neoadjuvant chemotherapy [J]. *Brit J Cancer*, 2015, 113(1): 30-36.
 - [127] Syn N L, Chua D W, Chen L R, et al. Time-varying prognostic effects of primary tumor sidedness and grade after curative liver resection for colorectal liver metastases [J]. *Surg Oncol*, 2021, 38: 101586.
 - [128] Therneau T, Crowson C, Clinic M. Using time dependent covariates and time dependent coefficients in the Cox model [R], 2014.

- [129] Zhang Z H, Reinikainen J, Adeleke K A, et al. Time-varying covariates and coefficients in Cox regression models [J]. *Ann Transl Med*, 2018, 6(7): 121.
- [130] Giles-Corti B, Moudon A V, Lowe M, et al. What next? Expanding our view of city planning and global health, and implementing and monitoring evidence-informed policy [J]. *Lancet Glob Health*, 2022, 10(6): E919-E926.
- [131] Cerin E, Sallis J F, Salvo D, et al. Determining thresholds for spatial urban design and transport features that support walking to create healthy and sustainable cities: findings from the IPEN Adult study [J]. *Lancet Glob Health*, 2022, 10(6): E895-E906.
- [132] Dinu M, Pagliai G, Macchi C, et al. Active Commuting and Multiple Health Outcomes: A Systematic Review and Meta-Analysis [J]. *Sports Med*, 2019, 49(3): 437-452.
- [133] Lear S A, Hu W, Rangarajan S, et al. The effect of physical activity on mortality and cardiovascular disease in 130 000 people from 17 high-income, middle-income, and low-income countries: the PURE study [J]. *Lancet*, 2017, 390(10113): 2643-2654.
- [134] Shen Y S, Lung S C C. Multiple impacts and pathways of urban form and environmental factors on cardiovascular mortality [J]. *Sci Total Environ*, 2020, 738: 139512.
- [135] Stevenson M, Thompson J, De Sa T H, et al. Land use, transport, and population health: estimating the health benefits of compact cities [J]. *Lancet*, 2016, 388(10062): 2925-2935.
- [136] Seo S, Choi S, Kim K, et al. Association between urban green space and the risk of cardiovascular disease: A longitudinal study in seven Korean metropolitan areas [J]. *Environ Int*, 2019, 125: 51-57.
- [137] Diener A, Mudu P. How can vegetation protect us from air pollution? A critical review on green spaces' mitigation abilities for air-borne particles from a public health perspective-with implications for urban planning [J]. *Sci Total Environ*, 2021, 796: 148605.
- [138] Margaritis E, Kang J. Relationship between green space-related morphology and noise pollution [J]. *Ecol Indic*, 2017, 72: 921-933.
- [139] Michimi A, Wimberly M C. Associations of supermarket accessibility with obesity and fruit and vegetable consumption in the conterminous United States [J]. *Int J Health Geogr*, 2010, 9: 49.
- [140] Thornton L E, Pearce J R, Macdonald L, et al. Does the choice of neighbourhood supermarket access measure influence associations with individual-level fruit and vegetable consumption? A case study from Glasgow [J]. *Int J Health Geogr*, 2012, 11: 29.
- [141] Gan Y, Tong X Y, Li L Q, et al. Consumption of fruit and vegetable and risk of coronary heart disease: A meta-analysis of prospective cohort studies [J]. *Int J Cardiol*, 2015, 183: 129-137.
- [142] Liu H, Tian Y H, Cao Y Y, et al. Fine particulate air pollution and hospital admissions and readmissions for acute myocardial infarction in 26 Chinese cities [J]. *Chemosphere*, 2018, 192: 282-288.
- [143] Zhu W T, Cai J J, Hu Y C, et al. Long-term exposure to fine particulate matter relates with incident myocardial infarction (MI) risks and post-MI mortality: A meta-analysis [J]. *Chemosphere*, 2021, 267: 128903.

- [144] Brook R D, Rajagopalan S, Pope C A, et al. Particulate Matter Air Pollution and Cardiovascular Disease An Update to the Scientific Statement From the American Heart Association [J]. *Circulation*, 2010, 121(21): 2331-2378.
- [145] Sacks J D, Stanek L W, Luben T J, et al. Particulate Matter-Induced Health Effects: Who Is Susceptible? [J]. *Environ Health Persp*, 2011, 119(4): 446-454.
- [146] Wei J, Li Z Q, Lyapustin A, et al. Reconstructing 1-km-resolution high-quality PM_{2.5} data records from 2000 to 2018 in China: spatiotemporal variations and policy implications [J]. *Remote Sens Environ*, 2021, 252: 112136.
- [147] Wei J, Li Z Q, Cribb M, et al. Improved 1 km resolution PM_{2.5} estimates across China using enhanced space-time extremely randomized trees [J]. *Atmos Chem Phys*, 2020, 20(6): 3273-3289.
- [148] Liu N R, Bu Z M, Liu W, et al. Health effects of exposure to indoor volatile organic compounds from 1980 to 2017: A systematic review and meta-analysis [J]. *Indoor Air*, 2022, 32(5): e13038.
- [149] Zhou B, Zhao B. Population inhalation exposure to polycyclic aromatic hydrocarbons and associated lung cancer risk in Beijing region: Contributions of indoor and outdoor sources and exposures [J]. *Atmos Environ*, 2012, 62: 472-480.
- [150] 北京市规划和自然资源委员会. 北京市城市设计导则 [Z]. 2021, http://ghzrzyw.beijing.gov.cn/biaozhunganli/bz/cxgh/202112/t20211207_2555119.html.
- [151] 中华人民共和国商务部. 城市一刻钟便民生活圈建设指南 [Z]. 2021, <http://www.mofcom.gov.cn/aarticle/h/redht/202107/20210703179512.html>.
- [152] 余新晓, 张志强, 夏新莉. 森林对 PM_{2.5} 等颗粒物的调控功能与技术 [M]. 北京: 科学出版社, 2017.

致 谢

清华园九年时光转瞬即逝，论文完成之际，不禁回想这四年的城乡规划兼硕的独特经历，感慨良多。与其他兼硕的同学不同，我的博士课题跟城乡规划关系不大。选择城乡规划作为兼硕，是我想圆自己一个本科学习城乡规划的梦想，不留遗憾。因此，当我博士生入学时在新生手册的角落里惊喜发现建筑学院直博兼硕的招生宣传之后，我战战兢兢地向博士生导师提出申请意愿，最终有幸如愿。

衷心感谢导师龙瀛老师过去四年来对我的不断鼓励与悉心指导。与龙老师的缘分，始于大四时期偶然在选课系统里看到了龙老师开设的《城市模型概论》课程，这也成为我初次接触城市大数据分析的契机。兼硕期间，龙老师一直鼓励我从跨学科的视角，批判性地看待城乡规划学科，并且给我提供多样的科研与合作项目机会，让我最终从其中挑选到适合自己且感兴趣的课题内容。龙老师严谨治学的科研态度、多元交叉的探索风格，大大拓展了我的研究视野，提高了我的研究素养与能力。

感谢我的父母一直支持我选择城乡规划的直博兼硕项目，你们的关怀、鼓励和陪伴使我有勇气、有信心选择这条路，并坚持到最后。也感谢你们在我面对各方压力感到困苦时，为我排忧解难、纾解心情。

感谢首都医科大学附属安贞医院的邓秋菊老师、刘静老师、常杰同学、胡飘飘同学、黄煜琳同学提供的流行病学研究分析方法的指导。感谢龙瀛老师课题组的张雨洋博士、李彦博士、李文越博士、李文竹博士、张恩嘉学姐、侯静轩学长、赵慧敏同学、王新宇同学、李伟健同学等众多学长学姐和同学在我完成硕士课题期间给予的大力帮助。

感谢工业工程系的王琛副教授和统计学研究中心的周在莹副教授、邓婉璐副教授和王江典老师，你们的统计学系列课程为我打开了通向统计学的大门，化解了我在本论文研究期间遇到的诸多统计分析问题。

最后感谢我的挚友——康旭源和陈威铭，感谢你们一直以来的陪伴，我不开心的时候听我吐槽诉苦，我获得成果的时候与我共同庆祝，周末假期的时候与我一同畅游京城古迹，是你们的鼓励与支持让我能够兼顾两方学业走到现在。

本课题承蒙国家自然科学基金（资助号：82073635 和 82103962）、首都卫生发展科研专项（资助号：PXM2020_026272_000010）以及惠康信托基金会“Pathways to Equitable Healthy Cities”项目（资助号：209376/Z/17/Z）资助，特此致谢。

声 明

本人郑重声明：所呈交的学位论文，是本人在导师指导下，独立进行研究工作所取得的成果。尽我所知，除文中已经注明引用的内容外，本学位论文的研究成果不包含任何他人享有著作权的内容。对本论文所涉及的研究工作做出贡献的其他个人和集体，均已在文中以明确方式标明。

签 名：_____日 期：_____

个人简历、在学期间完成的相关学术成果

个人简历

1996 年 10 月 29 日出生于江苏省无锡市。

2014 年 9 月考入清华大学建筑学院建筑技术科学系建筑环境与能源应用工程专业，2018 年 7 月本科毕业并获得工学学士学位。

2018 年 9 月免试进入清华大学建筑技术科学系攻读土木工程博士学位至今，师从张寅平教授。

2019 年 9 月入选清华大学建筑学院直博兼硕项目，兼修城乡规划学硕士至今，师从龙瀛副教授。

在学期间完成的相关学术成果

学术论文（+ 为共同第一作者）：

- [1] **Liu N R**⁺, Fang L⁺, Liu W, et al. Health effects of exposure to indoor formaldehyde: A systematic review and meta-analysis on the literature in the past 40 years [J]. Building and Environment, 2023, 233: 110080.
- [2] Zhang Y Y⁺, **Liu N R**⁺, Li Y⁺, et al. Neighborhood infrastructure-related risk factors and non-communicable diseases: a systematic meta-review [J]. Environmental Health, 2023, 22(1): 2.
- [3] **Liu N R**, Zhang X, Wang L Y, et al. Early-Stage Emissions of Formaldehyde and Volatile Organic Compounds from Building Materials: Model Development, Evaluation, and Applications [J]. Environmental Science & Technology, 2022, 56: 14680-14689.
- [4] Chen Z R⁺, **Liu N R**⁺, Tang H, et al. Health effects of exposure to sulfur dioxide, nitrogen dioxide, ozone and carbon monoxide between 1980 and 2019: a systematic review and meta-analysis [J]. Indoor Air, 2022, 32(11): e13170.
- [5] **Liu N R**, Bu Z M, Liu W, et al. Indoor exposure levels and risk assessment of volatile organic compounds in residences, schools, and offices in China from 2000 to 2021: A systematic review [J]. Indoor Air, 2022, 32(9): e13091.
- [6] **Liu N R**, Zhao Z Z, Mandin C, et al. Why should we target the burden of disease for indoor air pollutants? [J]. Indoor Air, 2022, 32(11): e13147.
- [7] **Liu N R**, Bu Z M, Liu W, et al. Health effects of exposure to indoor volatile organic compounds (VOCs) from 1980 to 2017: A systematic review and meta-analysis [J]. Indoor Air, 2022, 32(5): e13038.

- [8] **Liu N R**, Cao J P, Huang J, et al. Role of Clothing in Skin Exposure to Di(n-butyl) Phthalate and Tris(1-chloro-2-propyl) Phosphate: Experimental Observations via Skin Wipes [J]. Environmental Science & Technology Letters, 2021, 8(3): 270-275.
- [9] Li W Z, **Liu N R**, Long Y. Assessing carbon reduction benefits of teleworking: A case study of Beijing [J]. Science of the Total Environment, 2023, 889: 164262.
- [10] Fang L, **Liu N R**, Liu W, et al. Indoor formaldehyde levels in residences, schools, and offices in China in the past 30 years: A systematic review [J]. Indoor Air, 2022, 32(10): e13141.
- [11] Cao J P, **Liu N R**, Zhang Y P. SPME-Based Ca-History Method for Measuring SVOC Diffusion Coefficients in Clothing Material [J]. Environmental Science & Technology, 2017, 51: 9137-9145.
- [12] Su C X, Pan M Y, **Liu N R**, et al. Lung cancer as adverse health effect by indoor radon exposure in China from 2000 to 2020: A systematic review and meta-analysis [J]. Indoor Air, 2022, 32(11): e13154.
- [13] Liu W, Sun Y X, **Liu N R**, et al. Indoor exposure to phthalates and its burden of disease in China [J]. Indoor Air, 2022, 32(4): e13030.
- [14] Li W Y, Long Y, Kwan M P, **Liu N R**, et al. Measuring individuals' mobility-based exposure to neighborhood physical disorder with wearable cameras? [J]. Applied Geography, 2022, 145: 102728.
- [15] Zhang Y Y, Deng Q J, Guo M N, ..., **Liu N R**, Liu J, Long Y. Using street view imagery to examine the association between urban neighborhood disorder and the long-term recurrence risk of patients discharged with acute myocardial infarction in central Beijing, China [J]. Cities, 2023, 138: 104366.
- [16] 莫金汉, 曹彬, **刘宁睿**, 等. 建筑空气环境人因工程学: 问题、思考与探索 [J]. 科学通报, 2022, 67(16): 1729-1743.
- [17] 张雨洋, **刘宁睿**, 龙瀛. 健康居住小区评价体系构建探析—基于城市规划与公共健康的结合视角 [J]. 风景园林, 2020, 27(11): 96-103.

参与学术专著:

- [1] Yang T, **Liu N R**, Zhou X J, et al. Source/Sink Characteristics of VVOCs and VOCs[M]//Zhang Y P, Hopke, P K, Mandin C. Handbook of Indoor Air Quality. Singapore: Springer, 2022: 647-694.
- [2] Zhang J S, Chen W H, **Liu N R**, et al. Testing and Reducing VOC Emissions from Building Materials and Furniture[M]//Zhang Y P, Hopke, P K, Mandin C. Handbook of Indoor Air Quality. Singapore: Springer, 2022: 1591-1636.

- [3] Hänninen O, Mandin C, Liu W, **Liu N R**, et al. Disease Burden of Indoor Air Pollution[M]//Zhang Y P, Hopke, P K, Mandin C. Handbook of Indoor Air Quality. Singapore: Springer, 2022: 1325-1368.

指导教师评语

缺血性心脏病已经成为影响中国人群健康水平的重要因素之一，其中急性心肌梗死的预后复发问题无论从普遍性还是从严重程度来看都尤为突出。从城市规划/建成环境角度探索急性心肌梗死复发并识别其作用，具有重要意义，总体上属于健康城市（规划）的范畴。

在既有研究多数关注单一建成环境暴露因素的健康效应识别的背景下，本论文创新性地采用暴露组学（urban exposome）方法，利用大规模的北京多年急性心肌梗死患者的病例数据，对这一科学问题进行回答，并提出了相应的规划应对策略。总体上，该论文具有重要的理论方法探索价值和一定的实践支持意义。

论文的工作量大，一、大量的病例地址不规范，无法完全通过自动的地理编码实现，需要进行大量人工校核工作；二、构建了涵盖六大类的 26 个小类的诸多建成环境暴露指标，并采用了三种缓冲区的空间尺度与与基线暴露、累积平均暴露两种时间尺度，实现对患者的建成环境暴露进行全面测度；三、采用 EWAS、EWAS-MLR、ENET、SPLS、DSA 多种算法对建成环境暴露的复发风险贡献进行识别。

作为兼修城乡规划学的硕士学位论文，研究在关联识别基础上还对北京市建成环境的归因 AMI 复发风险进行了现状评估，并给出了相应的规划应对策略，如利用社区服务中心及腾退空间增设蔬果店、设置颗粒物滞纳型防护绿地、利用闲置小微空地及滨水空间营造多层次绿地结构等。

论文结构合理，概念明确，资料翔实，论述充分，图文规范，达到了工学硕士学位论文的学术水平。

答辩委员会决议书

论文研究建成环境暴露与急性心肌梗死复发间的关联与规划应对策略，选题具有重要的理论意义和实践价值。

论文构建了包括 6 大类 26 小类的建成环境指标体系，并对各指标进行多时空尺度的测度和全面评价；进而基于北京急性心肌梗死患者的队列数据，定量地揭示了建成环境暴露与急性心肌梗死复发风险之间的关联；对北京市建成环境的归因急性心肌梗死复发风险进行了评估，分类提出了规划应对策略。论文有独到的新见解，成果突出。

论文工作表明作者对相关领域具有深入的了解，并具有较强的科研能力。论文概念清晰，层次分明，论证深入，文笔流畅，文字图表规范，学风严谨。

答辩阐述清楚，重点突出，很好地回答了答辩委员提出的问题。

答辩委员会表决，一致同意该论文达到了硕士论文的水平，通过论文答辩，并建议授予刘宁睿工学硕士学位。同时，答辩委员会认为本论文是一篇优秀的硕士论文。