文章编号：1673-9493（2025）03-0016-11　中图分类号：TU984.3　文献标识码：A　DOI：10.19830/j.upi.2023.177

城市密度对美国大都市区常见慢性病的影响研究

Study on the Impacts of Urban Density on Common Chronic Diseases in American Metropolitan Areas

刘超 刘泽润 刁弥
LIU Chao, LIU Zerun, DIAO Mi

摘要：既有研究发现城市密度会影响常见慢性病的发病率，但其理论机制和实证研究尚不充分，城市大数据的涌现和人工智能的发展为探索这一复杂影响提供了新的思路和方法。笔者以美国大都市区内典型城市人口普查区为例，分析人口、建成环境和活动的多维度密度因素同常见慢性病之间的关系，通过可视化方法探明关系趋势，利用多种机器学习和可解释性方法厘清关键影响要素、非线性关系、关键阈值、要素之间的交互作用和空间上的分异，得出不同慢病之间的共性和特性。研究结果验证了城市密度对慢性病的重要作用，发现了密度与慢性病之间的复杂关系，以及在不同地区之间的规律差异。这对建设健康城市、引导城市可持续发展、提升宜居水平具有重要意义。

Abstract: Existed research has revealed that urban density would influence the incidence rates of common chronic diseases, but lack theoretical mechanism and empirical studies. The development of artificial intelligence (AI) and the urban big data provide new opportunities to explore original ideas and methods of complex correlations. Aiming at identifying the relationships between urban density and chronic diseases, this paper takes metropolitan census tracts in U.S. as the study areas and analyzes three dimensions of population density, built environment density, and activity density. With the help of data visualization, this paper initially investigates the trends between variables. By building several machine learning (ML) models and using interpretable methods, this paper then uncovers the key influencing features, describes non-linear relationships, finds key thresholds, explains the interactions of factors, and discovers spatial distinction. Finally, this paper summarizes similarities and differences of common chronic diseases. The results verify the significant impacts of urban density on chronic diseases, interpret the complicated relationships and regional disparity. The conclusions are of great importance for constructing health cities, guiding sustainable development, and promoting livability.

关键词：城市密度；公共健康；解释性机器学习；非线性关系；可持续发展

Keywords: Urban Density; Public Health; Interpretable Machine Learning; Nonlinear Relationship; Sustainable Development

国家重点研发计划项目（2022YFC3800804），北京大学林肯基金 2021-2022（FS01-20211215-LC），教育部产学合作协同育人专题（202102479005）

作者：刘超，同济大学建筑与规划学院城市规划系，副教授，硕士生导师；智慧社区分试验中心，负责人。liuchao1020@tongji.edu.cn
刘泽润，纽约大学坦登工程学院城市系统，博士研究生
刁弥（通信作者），同济大学建筑与规划学院城市规划系，教授

引言

公共健康是推进我国社会主义现代化建设和城市宜居建设的重要组成部分。在城市公共健康领域，慢性病（下称“慢病”）成为影响居民健康状况的首要杀手，由于长时间的暴露和负面效应的累积，慢病同城市环境的关系更加密切，尤其是反思城市建成环境对各种常见慢病的影响已经成为城市管理者和规划者关心的重要议题[1-3]。既有研究发现，密度作为城市规划设计与管理的重要指标和衡量方式，与健康效应有一定的关联，常见的结论有：城市密度增加带来了集聚效应，降低了城市经营成本，提高了基础设施服务水平，促进了资源共享和社会经济财富的发展，丰富了医疗资源；但高密度的城市建设也带来许多负面影响，伴随的拥堵、污染等降低了居民生活质量，导致慢病高发。虽然既有研究对密度影响健康的作用进行了诸多讨论，但是在验证密度作用、解释复杂影响机制、梳理多种疾病的特性和共性等方面仍需大量实证探索。

因此，研究城市密度对慢病的影响机制并开展多种慢病的实证研究，对完善城市发展理论、指导城市可持续发展、实现健康城市建设，具有极大的理论和实践意义。笔者关注城市密度和常见的慢病，旨在揭示城市密度如何影响常见慢病的发病，验证城市密度是否与常见慢病有关联并识别其关联机制；同时，从多维角度解释密度的影响作用，包括相关性、关键影响要素、非线性关系、关键阈值、要素之间的交互作用、多种疾病在地区上的异同性。为回答上述问题，笔者首先回溯了密度理论的演化过程，简要分析密度理论的发展趋势以及在城市公共

健康领域的应用；其次，对城市常见健康问题进行梳理，从多元健康风险角度出发，聚焦常见慢病；最后，基于理论的分析，综合运用一般性统计方法、传统机器学习方法和多种解释性机器学习（interpretable machine learning）方法，验证理论的合理性，完成对北美大都市区慢病实例的研究，探索实证研究和复杂关系解释的方法。

1 城市密度

城市密度有多种不同的定义[4]。广义上的密度涉及物理、地理、经济、社会等学科，涵盖社会交往、经济活动、能量流动、生态环境、文化历史等方方面面；狭义上，面向城市规划和建筑学科的实践环节，城市密度往往指物理层面的人口密度和建筑密度[5]，包括容积率和建筑覆盖率等。笔者以狭义的城市规划对城市密度实践定义为核心，综合考虑城市的人、建成环境以及人与建成环境的交互在空间上的分布强度，将城市密度解读为：人口密度、建成环境密度和活动密度。其中，人口密度指人口在单位空间（面积、房间等空间单元）上的分布强度，建成环境密度指城市空间的开发建设强度（建筑密度、建筑物覆盖率、容积率和开放空间比例等），活动密度指活动及活动设施在空间上的分布强度（POI密度、人口流动结构和社交网络密度等）。

城市诞生于空间集聚和经济活动的发展，作为创新和财富的高度聚集中心，城市同乡村的显著区别是用更少的资源实现更多的创新积累[6-9]，而密度理论是城市发展的核心动力之一。长久以来，对密度发展的主要争议是：城市密度的提升一方面有助于提高城市的效率，增加城市的多样性，实现资源的共享[10]；但另一方面，高密度也带来疾病骤增、污染蔓延和犯罪增加的后果，降低了城市中的生活质量[11-12]。国内外学者纷纷从城市密度的时空演变，以及密度与城市“功能”（住房、能源、创新、财富、交通、基础设施等）之间关系的角度解读密度：一方面，密度可以灵活描述多尺度、多区域内的城市空间特质，这使得在横向比较不同地区和纵向分析不同阶段的城市发展程度成为可能；另一方面，密度更加可控，通过控制城市基本单元的密度，城市规划管理者能够有效地引导城市的发展，评估和预测人群活动。密度的实践性和包容性展示了其理论发展和实证研究的巨大潜力，有关能源、碳排放和生产力等方面的城市密度研究也取得了有价值的成果，越来越多的研究捕捉到密度对人的作用力，特别是关注到人的行为、习惯、情绪和健康等问题，探索并总结有关密度的普适规律，仍令人倍受鼓舞并挑战重重。

2 城市中慢性病相关研究进展

在城市密度的有关讨论中，密度对人健康效应的影响逐渐引起更多的关注。城市中的健康问题涉及生理和心理健康、疾病和不适、症状和行为等多种类型，以及综合医学、公共健康、生物学、社会学等学科理论。结合不同国家地区和世界卫生组织对健康问题、健康风险等的定义，城市中的多元健康问题可被概括地分为各类疾病和健康负担[13]。其中，疾病主要包括慢病、传染病和心理疾病：慢病指不具传染性、长期积累造成的各类疾病，如各类心脑血管疾病等；传染病指能通过不同途径传播的疾病，常见有细菌性和病毒性的各类流感、肺炎等；心理疾病多与脑功能障碍和精神活动紊乱有关，如人格分裂、抑郁、焦虑、神经衰弱等。健康负担涵盖身体和心理上的各类非健康行为、亚健康状态和健康损失，包含肥胖、失眠、营养不良、行动障碍等。

在多元健康问题中，慢病受到越来越多的关注，癌症、心脑血管疾病、糖尿病等已经成为造成城市居民死亡的主要疾病。大量研究证明慢病与环境作用和个体行为息息相关，自然环境、建成环境和社会环境共同作用影响慢病在人群中的发病。同传染病、健康负担、心理健康相比，慢病有较为一致的诊断标准、较长时间的观察期以及较为有限的遗传特性，这种可观测性和可控性使得慢病成为相对理想的衡量公共健康水平的视角，并在提升城市宜居性和指导建设健康城市上具有更强的可操作性和实践意义。

影响慢病的因素主要与基因、行为、社会经济、建成环境、气象等因素有关。然而，基因、行为因素在个体间差异巨大。在城市范围内，对慢病的关注往往从社会经济因素、建成环境因素、气象因素等角度出发。社会经济因素是当下健康问题研究的热点，众多学者发现了社会经济要素的主导作用，关注到社会经济要素中收入、种族、家庭的不同方面，并验证了社会经济维度的重要主导作用[14-18]。气象因素同样在很大程度上影响到健康风险的各个环节[19-23]，然而对于气象因素往往限制于特定的时空范围，尤其是对于长期积累的健康问题，会因为时空范围的差异得出相反的结论，不利于在较广空间范围和较长时间跨度上的分析。建成环境作为城市管理和建设的核心，越来越多的学者关注到城市建成环境同健康之间的关系[13,24-25]，而人与建成环境互动和密度控制已经成为新的热点和未来趋势[2,26-27]。

目前国内外学者多从建成环境维度出发，探究其密度或混合度与慢病的关系和影响机制。笔者通过对 Web of Science 和中国知网（CNKI）的文献数据库进行搜寻，选取了近几年内具有代表性的研究并汇总，表 1 总结了国内外相关研究的内容和结论，不同学科的专家学者针对城市化、土地使用、空间形态、道路交通、绿地与开放空间、人口密度、设施可达性等要素对健康水平的影响，用大量实证研究挖掘不同要素之间的关系，探究影响机制并提出应对策略[29,36]。其

中，中文文献在理论梳理和综述上做了许多工作，总结并展望了建成环境对慢病的影响模式及研究内容。舒平等利用CiteSpace对建成环境的健康效应热点进行了总结，发表了慢病为核心的城市环境健康效应研究[37]；王兰等对健康城市的研究和实践进行回顾，提出未来研究的展望，注重实践和实证[38-39]；张延吉等对建成环境影响健康的路径和机制进行了探索总结，明确了建成环境的作用地位[40]；肖扬等对建成环境影响健康的体系展开讨论，并展示已有的高密度健康城市研究中心在建成环境微观的成果[41]。

不同的研究实例运用了不同的研究方法，常见的有：一般统计方法，包括各类可视化方法，主成分分析，因子分析，线性回归和多元线性回归[16-19,28-30,32,34-35]；空间模型，包括空间回归模型（SRM），如空间滞后模型（SLM）、空间误差模型（SEM）、空间自回归模型（SAM）、空间杜宾模型（SDM）和地理加权回归模型（GWR）等[1,31,33,35]。尽管各类研究都取得了丰硕的成果，但国内仍多基于理论和综述提出研究的路径和方法，不同地区的研究发现了密度复杂的正负效应，同时，多元异构的影响要素和复杂晦涩的发病机理使得解释和预测健康问题困难重重，急需不同尺度面向具体慢病实例的实证研究，并对多种密度要素进行衡量，借助非线性的复杂模型揭示密度的复杂影响，发现并总结多种慢病的规律。

而可解释性机器学习在处理大量数据、建立复杂模型和解读辅助决策上展现出显著优势。可解释性的重点在于将复杂的计算模型转译为人类可理解的语言[42]，根据解释方式的不同，可解释性机器学习方法可以分为具有内在解释性的机器学习方法和本身不可解释模型的事后可解释性方法[43-44]。内在可解释模型指模型的算法结构和底层逻辑本身就可以很好的解释自己，因为这类模型的可解释性发生在训练前，又可称为事前可解释性模型。常见的内在可解释模型包括逻辑/线性回归、（浅层）决策树、k邻近算法、广义加性模型、（朴素）贝叶斯和基于规则的方法等[45]。不可解释模型往往具有较高的准确率和复杂性，无法通过模型结构简单推断，因此需要在训练完成的模型基础上，运用不同方法解释推断输入特征与预测结果之间的关系，也被称为事后可解释性。通过综合运用多种可解释性方法，可以有效地把抽象的模拟过程转化为通俗易懂的决策语言，帮助研究者和政策制定者得出结论以辅助实践。

综上所述，现有的研究对密度和慢性病的探讨广泛且深远，关注到人的行为、建成环境的影响、不同尺度的密度定义的方方面面，虽然理论的论述已有了丰富的成果，但对于实证的探索，尤其是多种疾病之间的比较和较大范围内的数据整合，仍需完善并进行方法和技术上的创新。因此，笔者在区域性的开放数据平台基础上，收集常见慢病如癌症、关节炎、冠心病、糖尿病、中风、慢性肾病、阻塞性肺病的数

表1 国内外有关城市常见慢病和密度要素的研究总结

来源	研究范围	主要慢病	城市密度要素	研究方法	主要发现
尤因等（Ewing et al.），2003[28]	美国的县和大都市地区	高血压、饮食、冠心病	基于人口普查和其他指标的主成分分析提取蔓延指数、郡县的紧凑程度	基于层次模型的横截面分析	县蔓延指数与高血压之间有显著的关联。与紧凑地区居民相比，开阔地区的居民在闲暇时间步行较少，体重更重，高血压患病率更高。在大都市水平上，蔓延指数与步行分钟指数相似，但与其他变量无关
王兰等，2018[29]	上海某城区	呼吸健康（肺癌患病率）	土地使用、空间形态、道路交通、绿地与开放空间	皮尔森相关系数、多元线性回归	工业用地、绿地、开放空间和道路密度与呼吸健康呈现显著的相关关系
马兰博等（Malambo et al.），2016[30]	南非	高血压	土地混合利用、道路连接、慢行系统、交通安全	Logistic回归分析	可感知的建筑环境属性与高血压显著相关。土地利用多样性与高血压显著相关
费等（Faeh et al.），2016[31]	瑞士	缺血性心脏病（IHD）	人口密度、路网密度	Weibull回归模型	建成环境要素的影响是独立的，且对心血管健康的影响是累积的
张延吉等，2018[32]	中国	慢病（高血压、心脏病、糖尿病等）、自我健康评价	土地利用密度、功能混合度、支路网肌理、设施可达性等	倾向值匹配法	高密度土地利用对健康状况有负向影响，不同要素在中高和中低阶层人群中的作用效果不同
希普和查利斯（Hipp & Chalise），2015[33]	美国	糖尿病	人口密度、骑行或步行通勤的人口比例	地理加权回归	与环境相关的变量中，骑自行车或步行上班的人口比例是与糖尿病相关的唯一显著因素，这种相关性的强度在全国不同地区有所不同
李等（Li et al.），2022[34]	武汉	缺血性心脏病	人口密度、容积率、归一化植被指数（NDVI）	Advanced Bayes Bayesian CAR model	开放空间比例较高、绿化增加、获得健康食品的机会增加以及体育设施的存在与缺血性心脏病风险降低有关
巴索拉斯等（Bassolas et al.），2019[35]	美国	缺血性死亡	活动密度、流动结构	Pearson和Spearman相关性，局部加权回归	流动性水平较高的城市更有可能广泛使用公共交通，提供更好的步行环境，降低人均污染物排放，并带来更好的健康评估

据，利用传统和前沿模型方法，解释城市密度同常见慢病之间的复杂关系，其成果用于比较多疾病、多地区的特性和共性，丰富实证研究的方法路径，并验证对密度理论的讨论。

3 研究数据和方法

3.1 研究对象和样本选择

本文研究的地区是美国（除夏威夷和阿拉斯加的大陆地区）的人口普查区（census tract），根据2010年人口普查区界定，美国共有近7万多个基本人口普查单元。为控制社会经济要素，聚焦主要城市地区，笔者基于美国社区调查（ACS: American Community Survey）的相关人口、经济和社会的数据[46]，借助K-means聚类方法对所有的人口普查去进行分组控制变量，以筛选出城市化水平较高、建设完备、经济活动频繁、贫富水平较为一致的地区作为主要研究对象，共14 304个样本。

3.2 研究数据和预处理

本文根据影响范围和死亡率排名，选取主要慢病为研究重点，数据来源于美国国家疾病控制与预防中心（CDC: Centers for Disease Control and Prevention）2018年PLACE项目[47]，该项目由罗伯特·伍德·约翰逊基金会（Robert Wood Johnson Foundation）和美国国家疾病控制与预防中心基金共同资助，数据集人口普查区尺度的各种慢病的发病率。密度数据主要来自于2010年人口普查以及2014—2018年的美国社区调查[46]，美国国家土地覆盖数据库（NLCD: National Land Cover Database）的开发强度数据，以及开放街道地图（OSM: Open Street Map）的兴趣点数据（表2）。本文收集了全美7万多个人口普查区的数据，通过数据清洗，删除异常值和缺失值，得到27 210条有效记录。为了综合表示密度指标，笔者利用熵值法对多种密度要素进行综合，得出了综合密度指标，用于可视化分析。对数据分布进行检查后发现，原始的密度数据呈现右侧的长尾分布，大多数数据密集分布在左侧，数据分布不平均，这样利用原始数据绘制散点图，会使得多数数据点密集聚积在一侧，难以观察其中的关系。为了更好地稳定方差，减轻数据分布倾斜的影响，使得原本密集的区间中的值更加分散，整体分布接近正态分布，研究将密度数据进行对数变化（log transformation）。

3.3 模型分析

笔者综合运用可视化、多种机器学习和解释性方法探索、归纳和演绎不同数据背后的规律。具体来说，运用可视化方法，初步探索多种慢病发病率同综合密度指标之间的关系；利用决策树、随机森林、极度提升树模型的多种方法，建立慢病发病率同多维度密度要素之间的机器学习模型，比较选择最优模型；并在机器学习模型基础上，运用置换特征重要性、部分依赖图、个体期望和SHAP值分析等可解释性机器学习方法，对模型结果和机制进行解读，解释复杂的关系，基本流程如图1所示。

4 研究结果

4.1 慢病发病率同综合密度相关性分析

简单的可视化能快速有效地发现变量之间的关系和数据中的趋势，通过绘制多种慢病发病率与城市综合密度指标的散点图，并进行线性拟合（图2上）和曲线拟合（图2下），发现不同慢病之间的共性。可视化的结果显示，对数变换过后，多种慢病发病率同城市综合密度之间均呈现出较好的线性拟合，并表现出相同的趋势，即随着城市密度指标的增长，发病率逐渐减小，说明慢病的发病率同城市综合密度之间表现出近似幂律分布（power law），且指数为负。同时，曲线拟合在多种慢病的尾部呈现收敛趋势，说明这种幂律关系的指数在高密度时绝对值变小。以上发现提示如下结果：慢病同其他复杂的城市现象一样，同密度呈现具有长尾效应的幂律分布，绝大多数城市地区密度较低，随着密度的增高，慢病发病率快速下降，少数城市地区密度极速增长，这种密度

表2 主要变量与数据来源

类型			变量	数据源
健康风险	疾病	慢病	冠心病、癌症、糖尿病、关节炎、慢性肾病、阻塞性肺病	美国国家疾病控制与预防中心
城市密度	建成环境密度	开放空间	开放空间比例	美国国家土地覆盖数据库
		开发强度	低、中、高密度开发	美国国家土地覆盖数据库
	人口密度	人口密度	人口密度	美国人口普查和美国社区调查
		拥挤程度	拥挤程度	美国人口普查和美国社区调查
	活动密度	POI密度	POI密度	开放街道地图（OSM）
社会经济要素	人口特征	年龄	大于65岁人口比例、小于17岁人口比例	美国人口普查和美国社区调查
		种族	少数种族人口比例	美国人口普查和美国社区调查
		收入	贫困人口比例、人均资产	美国人口普查和美国社区调查
		就业	失业人口比例	美国人口普查和美国社区调查

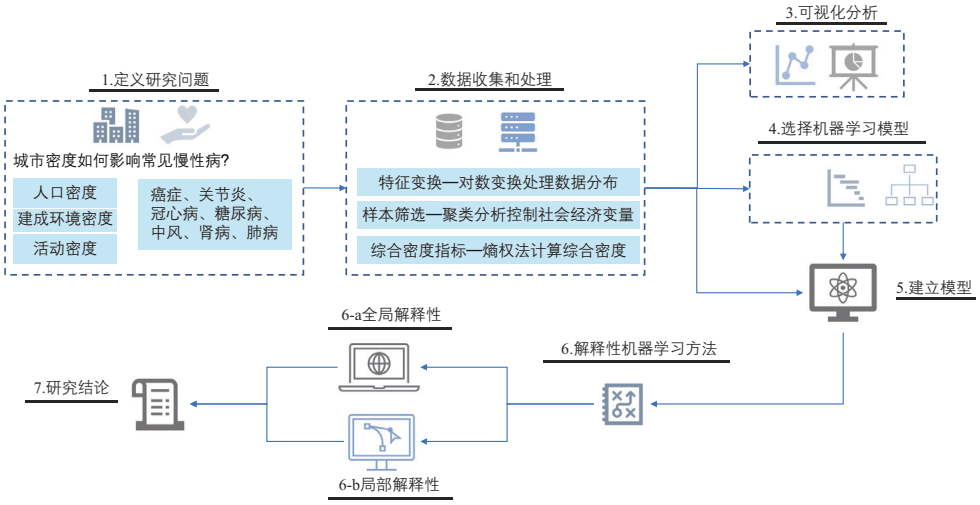
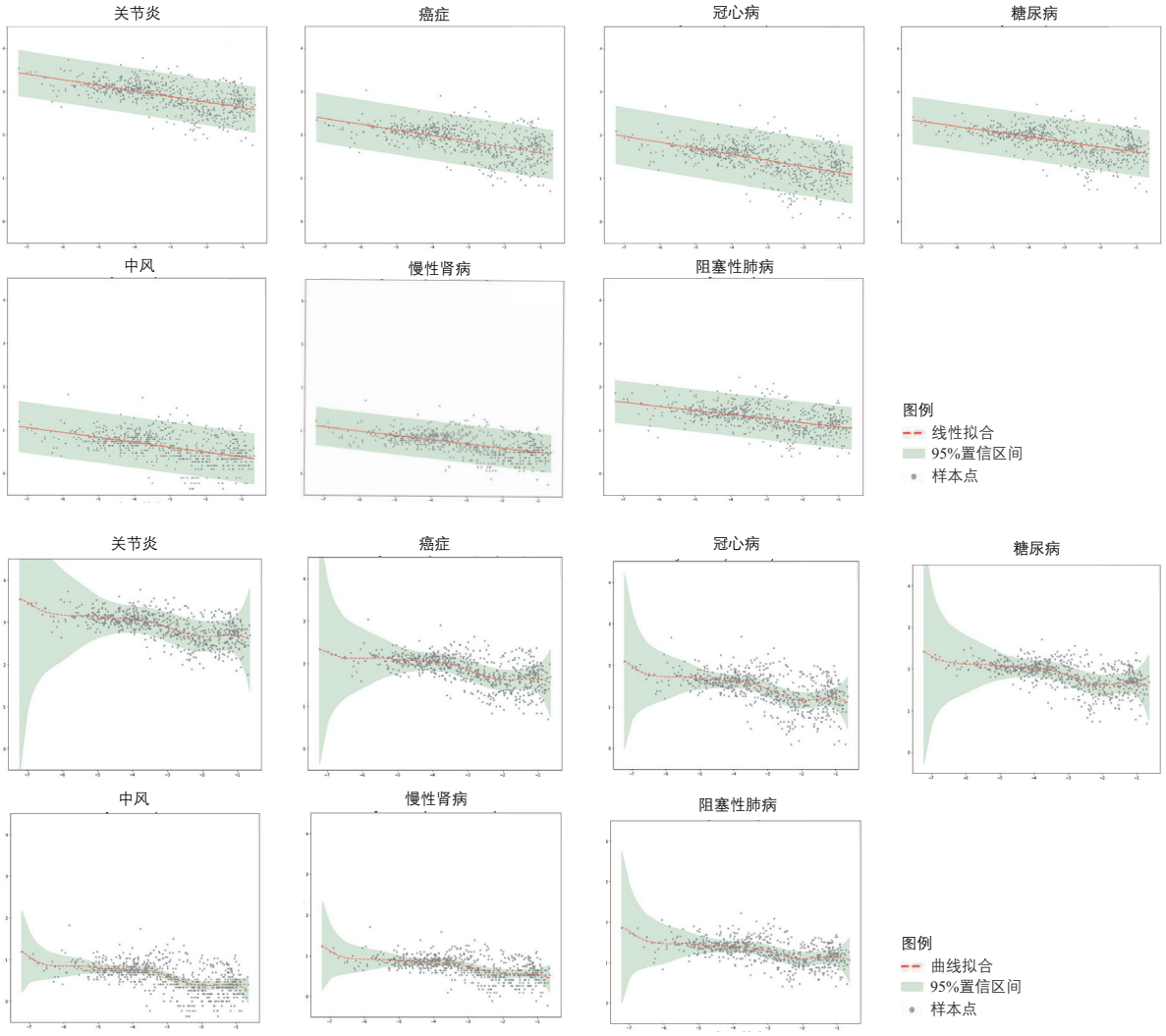


图 1 模型分析流程



注：横坐标表示对数变换后的城市综合密度指标，纵坐标表示对数变换后的慢病发病率，图中每一个点表示每个人口普查区的数据，红色虚线分别代表对样本点基于最小二乘法（OLS）的线性拟合，以及对样本点基于局部加权回归（LOWESS）的曲线拟合。

图 2 常见慢病与综合密度指标可视化分析线性拟合（上）与曲线拟合（下）

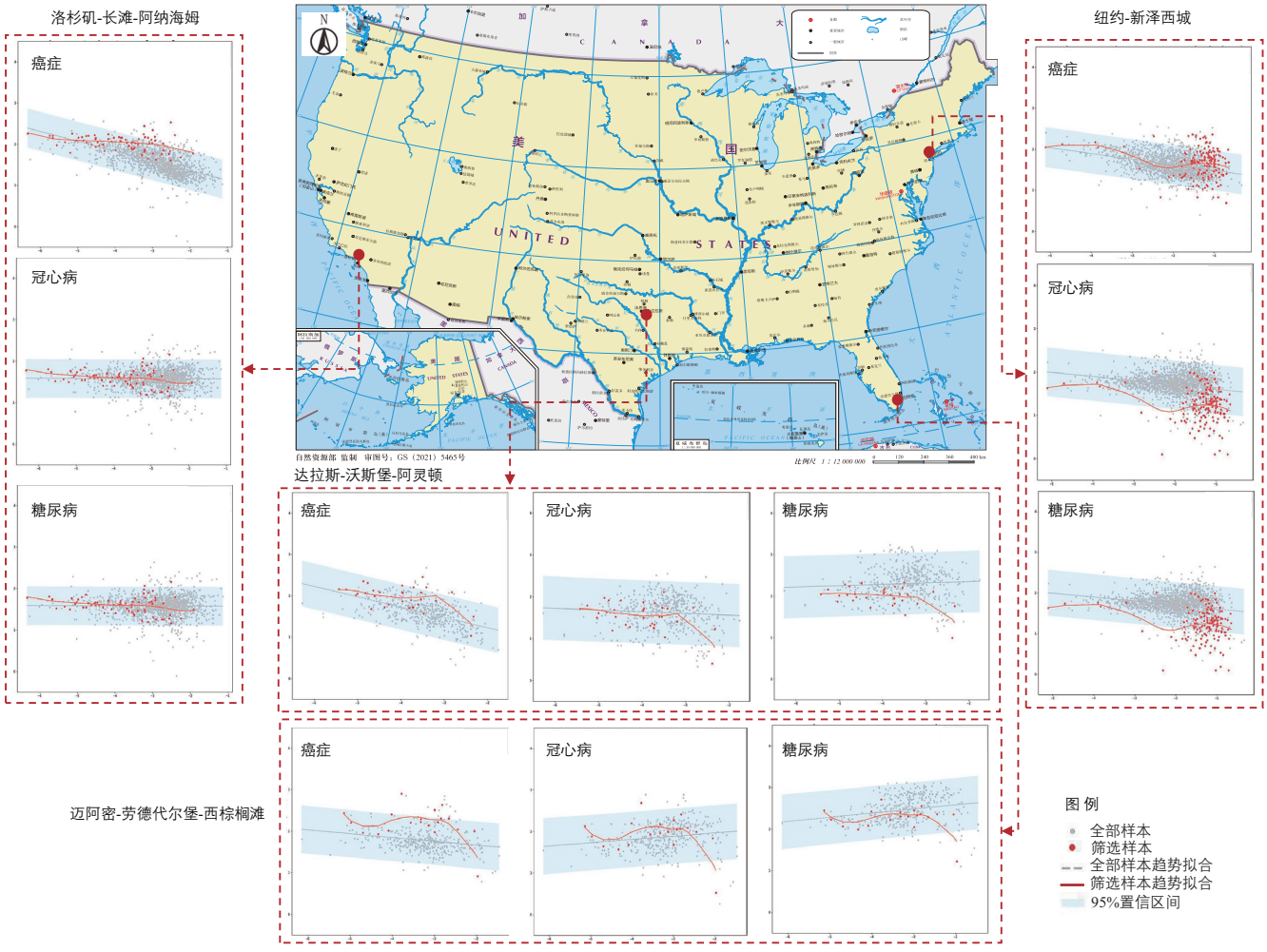
增长对慢病的降低作用减小，并在极高密度的尾部趋于平缓。然而，这种幂律规律并不是完美的，曲线拟合发现了尾部收敛的趋势，仍有大部分样本分布在完美幂律拟合线的两侧，说明仍有未知的非线性关系有待挖掘。

4.2 前十大都市区慢病发病率同综合密度相关性分析

笔者选择人口规模排名前 10 的大都市区为代表，绘制慢病发病率同综合人口密度的散点图（图 3），来表述密度与慢性病发病率关系的空间差异。总的来说，癌症和关节炎在不同区域中呈现和总体一致的趋势，发病率同城市综合密度负相关；冠心病在洛杉矶、凤凰城和迈阿密地区呈现与总体相反的趋势，发病率同城市综合密度不显著正相关；而糖尿病、中风在洛杉矶、凤凰城、迈阿密、休斯敦和达拉斯地区的相反趋势更显著；慢性肾病、阻塞性肺病也呈现出地区之间的差异。

4.3 多维度密度要素解读

关注到人口密度、建成环境密度和活动密度三个维度的多个密度要素，并为进一步明晰不同维度密度要素的贡献，识别其中的重要要素，笔者将各类慢病发病率作为因变量，将多种密度要素作为自变量，运用多种机器学习方法建立模型，并选择其中表现最好的模型，分析其中要素的贡献。由简单到复杂，本文用到的模型方法有逐步多元线性回归（MLR: Stepwise Multiple Linear Regression）、决策树（DT: Decision Tree）、随机森林（RF: Random Forest）、极度提升决策树（XGBoost: Extreme Gradient Boosted Trees），通过网格搜索结合交叉验证（cross-validation）确定决策树、随机森林和极度提升决策树的最佳参数，通过 MAE、MSE、RMSE、R^2 等指标对不同模型性能进行比较，选择其中表现最优模型。表 3 总结了多种模型基于 R^2 的表现结果，可以看到，模型在不同慢病上的性能相似，随着



注：横坐标表示对数变换后的城市综合密度指标，纵坐标表示对数变换后的慢病发病率，灰色的点表示大都市区中所有的人口普查区样本，红色的点表示筛选出的样本点，蓝色的线是对所有样本点的线性拟合，红色的线是对筛选样本点的曲线拟合。

图 3 典型代表都市区部分慢病发病与城市综合密度散点图

模型复杂度提高，预测的准确性不断提高，然而过于复杂的模型容易过拟合，其泛化能力较差。总的来说，随机森林模型在对不同慢病拟合中达到最佳的准确性和泛化能力的平衡，因此，笔者将选择随机森林作为后续各种解释性方法的学习模型。

在随机森林模型结果的基础上，本文利用置换特征重要性（permutation feature importance）和特征的 SHAP 值（shapley values）来测度不同特征的重要性，结果如表 4 所示。置换特征重要性能更好地洞察特征的全局贡献，置换特征重要性的数值越大，说明替换该特征对模型性能的扰动更大。SHAP 值能提供基于局部和各样本的视角，各个特征 SHAP 值的平均绝对值反映了特征对结局变量的贡献，数值越大说明贡献程度越大，各个样本的 SHAP 值总结显示了在每一个样本上特征的贡献，其绝对值表示贡献程度，正负表示贡献方向，颜色表示特征值本身的大小（图 4）。结果显示，不同方法测算的不同慢病的重要特征排序具有一致性，首要影

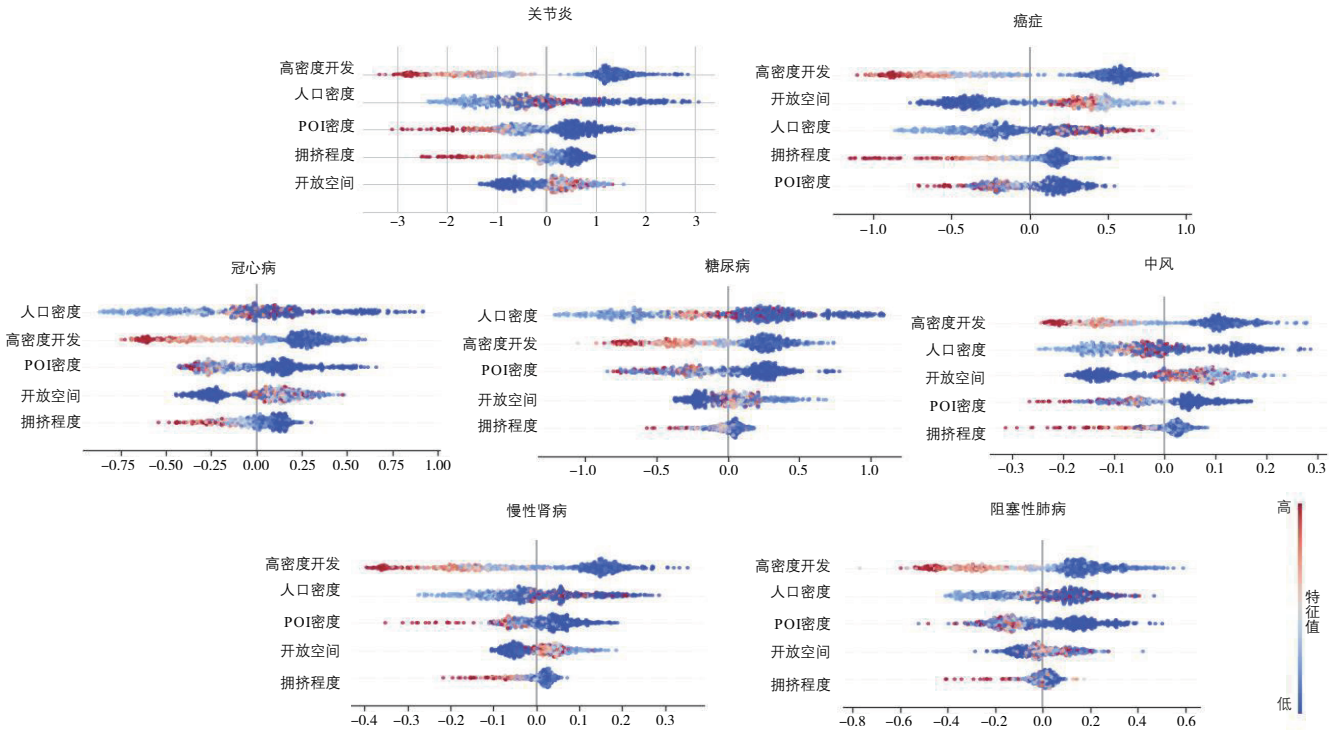
表 3 多模型结果 R^2 比较

疾病	MLR 逐步多元线性回归	DT 决策树	RF 随机森林	XGBoost 极度提升决策树
关节炎	0.40	0.40	0.46	0.48
中风	0.31	0.25	0.28	0.27
阻塞肺病	0.28	0.29	0.32	0.21
冠心病	0.33	—	0.43	0.39
糖尿病	0.31	0.38	0.44	0.39
癌症	0.39	0.43	0.47	0.45
慢性肾病	0.32	0.27	0.23	0.23

注：当模型 R^2<0 时，表中记为“—”。

表 4 常见慢病不同模型特征重要性结果

疾病	特征重要性（两个首要特征）	
	置换特征重要性	SHAP 特征重要性
关节炎	高密度开发	高密度开发
	人口密度	人口密度
癌症	高密度开发	高密度开发
	人口密度	人口密度
冠心病	人口密度	人口密度
	高密度开发	高密度开发
糖尿病	人口密度	人口密度
	高密度开发	高密度开发
中风	人口密度	高密度开发
	高密度开发	人口密度
慢性肾病	高密度开发	高密度开发
	人口密度	人口密度
阻塞性肺炎	人口密度	高密度开发
	POI 密度	人口密度

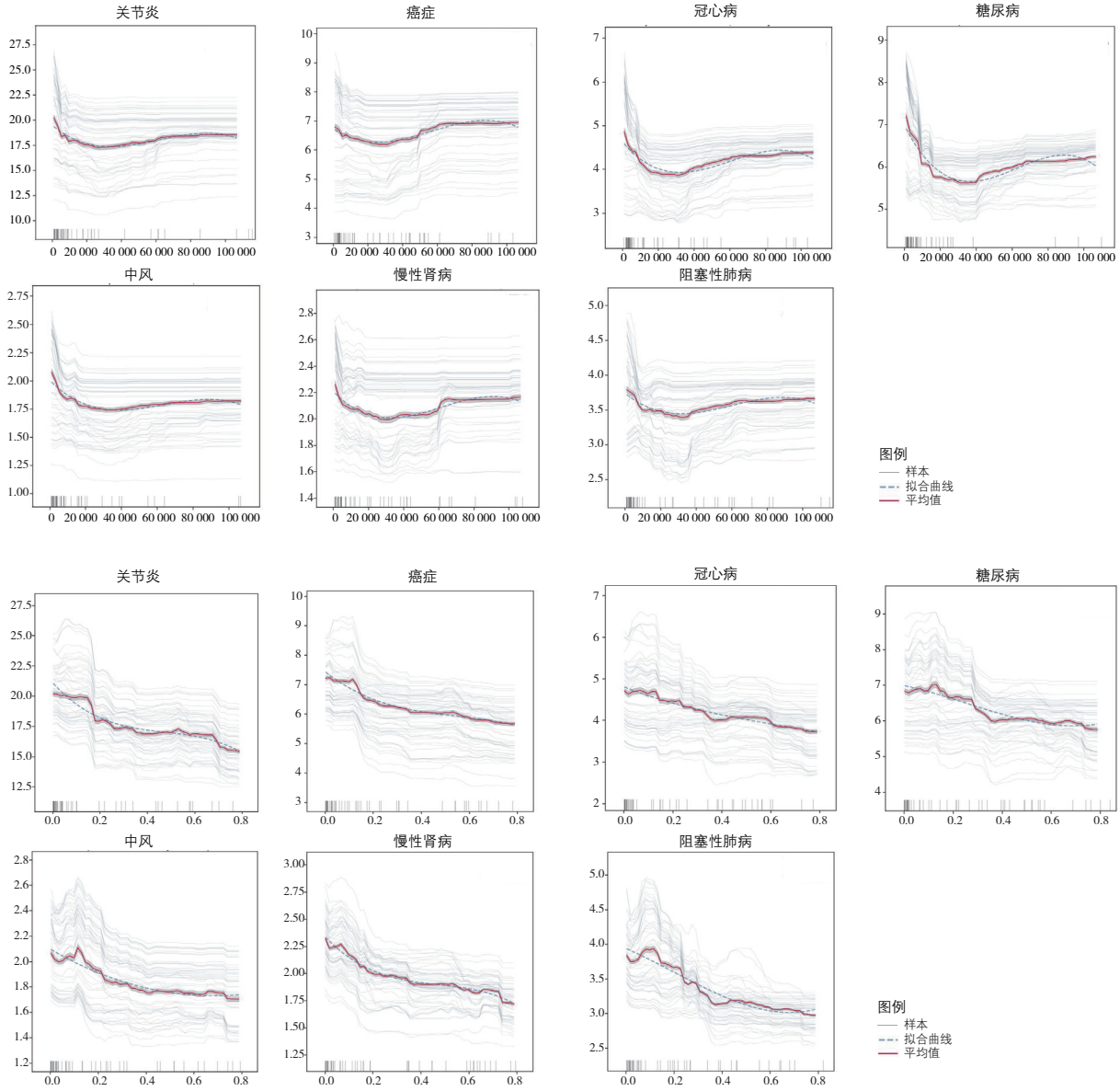


注：此 SHAP 总结图展示了每个特征对模型输出的影响。x 轴表示 SHAP 值，表明特征对预测结果的影响大小；y 轴列出了特征；每个点代表一个数据点，颜色表示特征值，蓝色表示低值，红色表示高值。

图 4 常见慢病的 SHAP 特征重要性分析

响要素基本均为人口密度和高密度开发，其次是 POI 密度和开放空间，再者是拥挤程度。从样本点分布来看，人口密度更高的样本点的特征贡献较为均匀地集中在 y 轴附近，说明高人口密度的正负影响并不显著；而高密度开发更高的样本点的特征贡献多位于 y 轴左侧，表明更高的高密度开发对慢病发病呈现负向影响。

为更加直观地观测首要影响要素同发病率之间的关系，本文借助部分依赖图和个体条件期望图来展示人口密度（图 5 上）和高密度开发（图 5 下）与多种慢病预测的依赖关系。多种慢病的部分依赖图表现出相似的规律：随着人口密度的增大，发病率呈现先降低后变缓，而后略有升高并趋于平稳的趋势，在人口密度小于 2 万人 /km^2 时，下降趋势显著，人口密度越高，发病率越低，在 2 万 ~4 万人 /km^2 之间，下降趋势基本消失，并在数值达到 4 万人 /km^2 左右出现转折，随着人口密度进一步提高，发病率缓慢上升，但这种增长倾向在更高密度区间逐渐平稳；多种慢性发病率同高密度开发呈现负向关系，即更高的高密度开发比例对应更低的发病率，且这种关系趋于线性，仅在 0.4 比例左右



注：上图中横坐标表示人口密度，纵坐标表示人口密度对机器学习模型预测结果的边际效应。下图中横坐标表示高密度开发比例，纵坐标表示高密度开发比例对机器学习模型预测结果的边际效应。通过计算每个给定特征的边际效应，得到个体条件期望，如图中灰色线所示（随机选择 50 个样本的个体条件期望进行可视化），将所有样本的边际效应取平均得到整体的部分依赖，如图中红线所示，并用蓝色虚线对部分依赖结果进行平滑曲线拟合。

图 5 人口密度（上）和高密度开发（下）的部分依赖图和个体期望分析

有不显著的凹点出现①。此外，不同密度要素之间是交互作用②的，这些交互作用也影响到非线性关系的产生。

4 结论和讨论

笔者运用多种模型方法研究城市密度对慢性病发病率的影响，其研究结果显示：(1) 慢病发病率与城市综合密度近似呈现幂律分布，指数为负，表明慢病大体上符合城市幂律解释的规律，但仍存在更复杂的非线性关系。(2) 通过可解释性机器学习，发现人口密度和高密度开发是主要影响因子。在低密度区，人口密度增长对健康产生正面效应，而在高密度区，这种影响转为负向，主要由拥挤引起的活动空间压缩和身心污染所致。(3) 不同慢病在不同空间分布上呈现出异质性，以东海岸较北部、发展阶段更完整的城市（如纽约、芝加哥等地区）与西海岸较南部、历史较短的城市（如洛杉矶、凤凰城等地区）相比，城市密度的负向影响更显著。(4) 在不同慢病中，这种地区差异的显著性不同，受环境长期影响、发作时间更长的长效损害累积造成的疾病空间差异不显著，与环境交互更明显、发作时间相对较短的短期损害累积的疾病空间差异更显著，表现在本文研究的疾病案例中，从癌症、关节炎，到冠心病、糖尿病、中风，再到慢性肾病、阻塞性肺病，地区之间的差异逐渐扩大。这进一步佐证了环境对慢病起到重要作用[48-50]，许多疾病风险由于环境的差异所导致[51]，不同地区之间包括空间结构、人口和环境特征以及健康行为的不同，导致了慢病在地区之间的差异，诸如癌症、心脏病和糖尿病受到更多环境要素的驱动。

总的来说，本研究证明了城市密度对慢性病的影响机制是复杂累积的，城市密度与公共空间可达性、开放空间比例、城市绿地比例、步行环境友好性、设施密度、空气质量、噪声水平等指标息息相关，与其他地区的相关研究结论一致[52-53]。一方面，城市的高度发展带来了更好的可达性和更综合的医疗设施资源，完善的城市系统带给人们更优质的服务，降低了医疗服务的获取难度和成本，提升了整体生活质量；另一方面，高密度带来的开放空间、绿地、步行系统资源的紧张，影响到人们的生活方式，如体力活动的减少、生活习惯的改变[54-55]，阻碍了人们对健康生活方式的追求；同时，高密度直接作用于环境压力，如空气污染、噪声污染等[56-57]，降低了城市的宜居性和环境质量，影响慢病的发病机率。此外，高密度环境可能增大居民的心理压力，压抑的空间和有限的资源进一步加剧社会隔离，影响心理健康，从而增加慢性疾病的风险[58-60]。这种复杂的相互作用导致城市密度与慢性疾病之间存在非线性的关系，之前的许多研究也揭示了人口密度与典型慢性病的非线性，在人口相对稀疏的地区，如欧洲部分地区和多伦多，研究发现人口密度与心脏病和糖尿病呈现负相关[48-50]，但在人口相对密集的地区，如武汉、台湾和上海，研究发现人口密度与心血管、癌症等发病呈现正相关[51]；而对于北美地区，多数研究讨论到慢性病受到人口密度、设施密度和开放空间可达性的影响[39,41]，但鲜有人聚焦全美范围的大都市区样本，关注非线性的阈值。本研究在支持前人结论的基础上，一定程度上证明人口密度的多重作用实际上是密度与慢性病的非线性关系，在低密度区间负相关，在高密度区间正相关，支持密度倒 U 型作用。这也启示城市规划在引导健康城市和可持续发展的过程中，应当建立更加全面的评估体系，更多考虑环境对疾病的风险影响指标，测算更精确的密度指标；也要对密度进行控制，将密度管控在合理的范围内，实现经济、环境和健康发展的最佳平衡[5,34]；目前关于城市密度和健康风险的研究方兴未艾，以地方案例统计分析为主流。大规模多病种的综合分析和临床病理推断，是未来辅助规划决策者调控城市密度促进健康的必要条件。

慢病发病机制复杂且多变，本研究关注到常见疾病同环境之间的关系，利用多种机器学习方法，展开城市密度影响常见慢病的初步研究，建立了一套完整的利用多种机器学习和解释性方法分析复杂城市问题的方法体系，对数据驱动的可持续建成环境领域的实证研究有借鉴意义，并得出了一些城市密度作用慢病的共性和特性的结论，进一步揭示密度的正负效应，佐证环境要素对典型慢病的影响。但现阶段研究缺少空间上对自相关问题和可塑性面积单元问题的考虑，对机制解释不足，未来研究将更多考虑空间相关性和单元选择，需更多临床实验和案例进行验证，通过对疾病发作机理的挖掘，深入探索其中的原理。同时考虑到在高密度区间的发现，仅仅依靠北美地区的数据说服力有限，未来应该不断扩充到更多高密度城市地区，特别是中国等亚洲地区数据。此外，影响疾病的因素甚广，涉及基因、行为、生活习惯、气候等方方面面，未来也应建立更完备的要素体系，综合多种空间模型和分析路径，完善研究的方法和理论。UPI

注：文中未注明资料来源的图表均为作者绘制。

参考文献

[1] 李琳，曾伟平 . 中国城市公共健康水平的时空演化及其影响因素 [J]. 地理研究 , 2022, 41(10): 2760-2776.

① 更多展示不同密度区间典型地区的卫星图片可扫描文末二维码获取，以便更形象地理解不同的密度区间。

② 更多交互作用的细节分析可扫描文末二维码获取。

[2] 宋玢，任云英，李薇，等．英文文献中城市规划与公共健康跨学科研究的知识图谱分析 [J]. 西安建筑科技大学学报 (自然科学版), 2021, 53(4): 568-576.
[3] 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.
[4] 张潇越．城市密度特征演变及相关因子分析 [D]. 深圳：深圳大学，2019.
[5] 李和平，刘志．中国城市密度时空演变与高密度发展分析——从 1981 年到 2014 年 [J]. 城市发展研究，2019, 26(4): 46-54.
[6] BECKER G S, GLAESER E L, MURPHY K M. Population and economic growth[J]. American economic review, 1999, 89 (2): 145-149.
[7] BETTENCOURT L, WEST G. A unified theory of urban living[J]. Nature, 2010, 467 (7318): 912-913.
[8] BETTENCOURT L M A, LOBO J, HELBING D, et al. Growth, innovation, scaling, and the pace of life in cities[J]. Proceedings of the national academy of sciences, 2007, 104(17): 7301-7306.
[9] MILGRAM S. The experience of living in cities[J]. Science, 1970, 167: 1461-1468.
[10] KATES R W, PARRIS T M. Long-term trends and a sustainability transition[J]. Proceedings of the national academy of sciences, 2003, 100 (14): 8062-8067.
[11] GLAESER E L, KOLKO J, SAIZ A. Consumer city[J]. Journal of economic geography, 2001, 1(1): 27-50.
[12] PAN W, GHOSHAL G, KRUMME C, et al. Urban characteristics attributable to density-driven tie formation[J]. Nature communications, 2013, 4: 1961.
[13] LEE C, MOUDON A V. Physical activity and environment research in the health field: implications for urban and transportation planning practice and research[J]. Journal of planning literature, 2004, 19(2): 147-181.
[14] ALMAGRO M, COVEN J, GUPTA A, et al. Racial disparities in frontline workers and housing crowding during COVID-19: evidence from geolocation data[J/OL]. [2023-11-19]. https://econpapers.repec.org/RePEc:fip:fedmoi:88803.
[15] YOU H, WU X, GUO X. Distribution of COVID-19 morbidity rate in association with social and economic factors in Wuhan, China: implications for urban development[J]. International journal of environmental research and public health, 2020, 17(10): 3417.
[16] SANNIGRAHI S, PILLA F, BASU B, et al. Examining the association between socio-demographic composition and COVID-19 fatalities in the European region using spatial regression approach[J]. Sustain cities society, 2020, 62: 102418.
[17] BARBIERI T, BASSO G, SCICCHITANO S. Italian workers at risk during the COVID-19 epidemic[J]. Italian economic journal, 2022, 1(8): 175-195.
[18] CHEN J T, KRIEGER N. Revealing the unequal burden of COVID-19 by income, race/ethnicity, and household crowding: US county versus zip code analyses[J]. J public health manag pract, 2021, 27(1): 43-56.
[19] XIE J, ZHU Y. Association between ambient temperature and COVID-19 infection in 122 cities from China[J]. Science of the total environment, 2020, 724: 138201.
[20] SHAHZAD F, SHAHZAD U, FAREED Z, et al. Asymmetric nexus between temperature and COVID-19 in the top ten affected provinces of China: a current application of quantile-on-quantile approach[J]. Science of the total environment, 2020, 736: 139115.
[21] BASHIR M F, MA B, BILAL, et al. Correlation between climate indicators and COVID-19 pandemic in New York, USA[J]. Science of the total environment, 2020, 728: 138835.
[22] AHMADI M, SHARIFI A, DOROSTI S, et al. Investigation of effective climatology parameters on COVID-19 outbreak in Iran[J]. Science of the total environment, 2020, 729: 138705.
[23] 崔秀云，邢银花，张彩霞，等．兰州市气象因素对心脑血管疾病就诊人次的影响研究 [J]. 气象与环境科学，2021, 44(6): 87-92.
[24] DILL J. The Role of health and physical activity in the adoption of innovative land use policy: findings from surveys of local planners[J]. Journal of physical activity and health, 2016, 8(supplement 1): 116-124.
[25] NUTSFORD D P A, KINGHAM S. An ecological study investigating the association between access to urban green space and mental health[J]. Public health, 2013, 127(11): 1005-1011.
[26] 程晗蓓．城市社区建成环境对流动人口健康的影响研究 [D]. 武汉：武汉大学，2021.
[27] 王淼，王南．基于 CiteSpace 软件的国内外建成环境健康效应对比研究 [J]. 当代建筑，2022(11): 117-120.
[28] EWING R, SCHMID T, KILLINGSWORTH R, et al. Relationship between urban sprawl and physical activity, obesity, and morbidity[J]. American journal of health promotion, 2003, 18(1): 47-57.
[29] 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.
[30] MALAMBO P, KENGNE A P, LAMBERT E V, et al. Association between perceived built environment and prevalent hypertension among South African adults[J]. Advances in epidemiology, 2016: 1-11.
[31] FAEH D, MOSER A, PANCZAK R, et al. Independent at heart: persistent association of altitude with ischaemic heart disease mortality after consideration of climate, topography and built environment[J]. Journal of epidemiology and community health, 2016, 70(8): 798-806.
[32] RADER B, SCARPINO S V, NANDE A, et al. Crowding and the shape of COVID-19 epidemics[J]. Nature medicine, 2020, 26(12): 1829-1834.
[33] HIPP J A, CHALISE N. Peer reviewed: spatial analysis and correlates of county-level diabetes prevalence, 2009-2010[J/OL]. [2023-11-19]. https://doi.org/10.5888/pcd12.140404.
[34] LI X, ZHOU L, LIU X, et al. Community built environment and the associated ischemic heart disease risk: evidence from multi-source data in Wuhan, China[J]. Journal of transport & health, 2022, 25: 101371.
[35] BASSOLAS A, BARBOSA-FILHO H, DICKINSON B, et al. Hierarchical organization of urban mobility and its connection with city livability[J]. Nature communications, 2019, 10(1): 4817.
[36] 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.
[37] 舒平，舒瑞桓．基于 CiteSpace 的城市建成环境健康效应研究热点分析 [J]. 河北工业大学学报 (社会科学版), 2021, 13(3): 77-82.
[38] 蒋希冀，叶丹，王兰．健康城市规划的研究热点与展望：基于英文文献计量分析 [J]. 城市发展研究，2022, 29(1): 104-112.
[39] 王兰，蒋希冀．2019 年健康城市研究与实践热点回眸 [J]. 科技导报，2020, 38(3): 164-171.
[40] 张延吉，邓伟涛，赵立珍，等．城市建成环境如何影响居民生理健康？——中介机制与实证检验 [J]. 地理研究，2020, 39(4): 822-835.
[41] 肖扬，钦莫伊·萨卡尔，克里斯·韦伯斯特．建成环境与健康的关联——来自香港大学高密度健康城市研究中心的探索性研究 [J]. 时代建筑，2017(5): 29-33.
[42] 陈珂锐，孟小峰．机器学习的可解释性 [J]. 计算机研究与发展，2020, 57(9): 1971-1986.
[43] 纪守领，李进锋，杜天宇，等．机器学习模型可解释性方法、应用与安全研究综述 [J]. 计算机研究与发展，2019, 56(10): 2071-2096.
[44] 杨丰春，郑思，李姣．可解释机器学习方法在疾病预测中的应用：脓毒血症患者死亡风险研究 [J]. 首都医科大学学报，2022, 43(4): 610-617.
[45] MOLNAR C. Interpretable machine learning: a guide for making black box models explainable (2nd ed.)[EB/OL]. [2023-11-19]. https://dlib.hust.edu.vn/handle/HUST/2499.
[46] United States Census Bureau. American community survey 5-year data (2009-2021)[DB/OL]. [2023-11-19]. https://www.census.gov/en.html.
[47] Centers for Disease Control and Prevention, Division of Population Health. PLACES: Local Data for Better Health, Place Data (2016-2022)[EB/OL]. [2023-11-19]. https://www.cdc.gov/places/index.html.
[48] GLAZIER R H, CREATORE M I, WEYMAN J T, et al. Density, destinations or both? a comparison of measures of walkability in relation to

（下转 41 页）

究其收缩背后差异化的社会经济驱动机制；针对乡村地区开展多尺度、多层级的增长与收缩情景规划，以明确不同情景下的规划治理路径和精明发展策略。UPI

注：文中未注明资料来源的图表均译自参考文献 [2]。

参考文献

[1] ESPON. ESPON policy brief on shrinking rural regions[R/OL]. [2022-12-15]. https://www.espon.eu/rural-shrinking.

[2] ESPON. European Shrinking Rural Areas Challenges, Actions and Perspectives for Territorial Governance (ESCAPE)[R/OL]. [2022-12-15]. https://www.espon.eu/escape.

[3] COPUS A, KAHILA P, THOMAS D, et al. European shrinking rural areas: key messages for a refreshed long-term European policy vision[J]. TERRA: Revista de Desarrollo Local, 2021(8): 280-309.

[4] THOMAS D. Shrinking in rural areas-mitigation and adaptation as the basis for intervention logics[R]. (2019-10-18)[2022-12-15]. https://ec.europa.eu/regional_policy/rest/cms/upload/18102019_105551_brussels_euregionsweek_2019_dax.pdf.

[5] 杨春华，姚逸苇．何谓“农村空心化”?——一个结构化的概念分析视角 [J]. 农村经济，2021(7): 79-86.

[6] 刘彦随，刘玉，翟荣新．中国农村空心化的地理学研究与整治实践 [J]. 地理学报，2009, 64(10): 1193-1202.

[7] 刘彦随，刘玉．中国农村空心化问题研究的进展与展望 [J]. 地理研究，2010, 29(1): 35-42.

[8] 赵民，游猎，陈晨．论农村人居空间的“精明收缩”导向和规划策略 [J]. 城市规划，2015, 39(7): 9-18, 24.

[9] 焦林申，张敏．收缩乡村的空废成因与精明收缩规划策略——基于豫东典型乡村的田野调查 [J]. 经济地理，2021, 41(4): 221-232.

[10] 杨忍，刘彦随，陈秧分．中国农村空心化综合测度与分区 [J]. 地理研究，2012, 31(9): 1697-1706.

[11] 潘竟虎，李瑶．甘肃省县域单元农村空心化差异的 PCA-ESDA 测度 [J]. 人口与发展，2013, 19(6): 65-73.

[12] 谭雪兰，于思远，欧阳巧玲，等．快速城市化区域农村空心化测度与影响因素研究——以长株潭地区为例 [J]. 地理研究，2017, 36(4): 684-694.

[13] 李玉红，王皓．中国人口空心村与实心村空间分布——来自第三次农业普查行政村抽样的证据 [J]. 中国农村经济，2020(4): 124-144.

[14] 王良健，陈坤秋，李宁慧．中国县域农村人口空心化程度的测度及时空分异特征 [J]. 人口学刊，2017, 39(5): 14-24.

[15] 陈涛，陈池波．中国农村人口空心化测量指标改进研究 [J]. 中国地质大学学报 (社会科学版), 2017, 17(1): 149-155.

[16] 刘振，戚伟，刘盛和．中国人口收缩的城乡分异特征及形成机理 [J]. 地理科学，2021, 41(7): 1116-1128.

[17] 周恺，钱芳芳，严妍．湖南省多地理尺度下的人口“收缩地图”[J]. 地理研究，2017, 36(2): 267-280.

[18] 贾铠阳，乔伟峰，王亚华，等．乡村振兴背景下村域尺度国土空间规划：认知、职能与构建 [J]. 中国土地科学，2019, 33(8): 16-23.

[19] 郑殿元，文琦，王银，等．中国村域人口空心化分异机制及重构策略 [J]. 经济地理，2019, 39(2): 161-168, 189.

[20] ZHOU Y, SHEN Y, YANG X, et al. Where to revitalize, and how? a rural typology zoning for china[J]. Land, 2021, 10(12): 1336.

[21] 张立，李雯骐，张尚武．国土空间规划背景下建构乡村规划体系的思考——兼议村庄规划的管控约束与发展导向 [J]. 城市规划学刊，2021(6): 70-77.

（本文编辑：顾春雪）

（上接 25 页）

transportation behaviors, obesity and diabetes in Toronto, Canada[J]. PLOS One, 2014, 9(1): e85295.

[49] 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.

[50] KONISHI M, MATSUZAWA Y, EBINA T, et al. Impact of population density on mortality in patients hospitalized for heart failure–JROAD-DPC registry analysis[J]. Journal of cardiology, 2020, 75(4): 447-453.

[51] YANG C Y, HSIEH Y L. The relationship between population density and cancer mortality in Taiwan[J]. Japanese journal of cancer research, 1998, 89(4): 355-360.

[52] LAWRENCE D F, ADHIKARI B, WHITE K R, et al. Chronic disease and where you live: built and natural environment relationships with physical activity, obesity, and diabetes[J]. Environment international, 2022, 158: 106959.

[53] LAWRENCE D F, WALI B. Treating two pandemics for the price of one: chronic and infectious disease impacts of the built and natural environment[J]. Sustainable cities and society, 2021, 73: 103089.

[54] LESTAN K A, ERŽEN I, GOLOBIČ M. The role of open space in urban neighbourhoods for health-related lifestyle[J]. International journal of environmental research and public health, 2014, 11(6), 6547-6570.

[55] MICHEL O D, MARQUES M L, SANTOS D D, et al. Spatial index relating urban environment to health lifestyle and obesity risk in men and women from different age groupS[J]. PLOS One 15, 2020, 3: e0229961.

[56] RAINALD B, SCHRAUTH P. Population density and urban air quality[J]. Regional science and urban economics, 2021, 86: 103596.

[57] NATALIE M, ROJAS-RUEDA D, BASAGAÑA X, et al. Health impacts related to urban and transport planning: a burden of disease assessment[J]. Environment international, 2017, 107: 243-257.

[58] SHIMA H, SABOURI S, EWING R. Does density aggravate the COVID-19 pandemic? early findings and lessons for planners[J]. Journal of the American Planning Association, 2020, 86(4): 495-509.

[59] TSAI W L, MELISSA R M, JENNINGS V, et al. Relationships between characteristics of urban green land cover and mental health in US metropolitan areas[J]. International journal of environmental research and public health, 2018,15(2): 340.

[60] GRUEBNER O, MICHAEL A R, MAZDA A, et al. Cities and mental health[J] Deutsches ärzteblatt international, 2017, 114(8): 121.

（本文编辑：许玫）

本文更多增强内容扫码进入