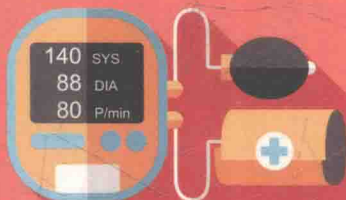


基于分层贝叶斯模型的高血压时空分布特征及其演化规律研究

徐 丽 / 著



厦门大学出版社
XIAMEN UNIVERSITY PRESS

国家一级出版社
全国百佳图书出版单位

基于分层贝叶斯模型的高血压时空分布特征及其演化规律研究

徐 丽 / 著



厦门大学出版社 国家一级出版社
XIAMEN UNIVERSITY PRESS 全国百佳图书出版单位

图书在版编目(CIP)数据

基于分层贝叶斯模型的高血压时空分布特征及其演化规律研究/

徐丽著. —厦门: 厦门大学出版社, 2018. 8

ISBN 978-7-5615-7088-3

I. ①基… II. ①徐… III. ①高血压-防治 IV. ①R544.1

中国版本图书馆 CIP 数据核字(2018)第 199445 号

出 版 人 郑文礼

责任编辑 吴兴友

封面设计 蒋卓群

技术编辑 朱 楷

出版发行 厦门大学出版社

社 址 厦门市软件园二期望海路 39 号

邮政编码 361008

总 编 办 0592-2182177 0592-2181406(传真)

营销中心 0592-2184458 0592-2181365

网 址 <http://www.xmupress.com>

邮 箱 xmup@xmupress.com

印 刷 厦门市金凯龙印刷有限公司

开本 889 mm × 1 194 mm 1/32

印张 8

插页 2

字数 173 千字

版次 2018 年 8 月第 1 版

印次 2018 年 8 月第 1 次印刷

定价 30.00 元

本书如有印装质量问题请直接寄承印厂调换



厦门大学出版社
微信二维码



厦门大学出版社
微博二维码

前 言

高血压是最常见的慢性病之一,也是心脑血管疾病最主要的危险因素。根据世界卫生组织(WHO)2013年的报告,全球每年有1 700万人死于心血管疾病,占死亡总数的1/3,其中有940万死于高血压的并发症。世界各国尤其是低收入和中等收入国家居高不下的高血压患病率严重消耗了医疗和社会资源,给家庭和国家造成了沉重的负担。

2017年中国科学院大学与昆明科技大学等多家科研机构联合发布的“中国心血管病死亡地图”显示,冠心病、脑血管病与高血压的死亡率均存在明显的空间差异。其中,南方高血压问题更严重,北方则是冠心病与脑血管病的重灾区。具体而言,高血压死亡率排名前四位的地区分别是湖北、广东、湖南和重庆,死亡率最

低的是青海、天津、宁夏和海南。

高血压病因复杂,且男性与女性高血压患者在生理、心理与社会经济状况等方面存在差异,其患病特点及其相关因素可能存在着差别。目前从性别视角对高血压分布模式及其关联因素的研究较少;且研究时大多采用基于个体层面的传统统计分析方法,忽略了疾病数据内在包含的空间与时间信息。

因此,尽管已有许多高血压病因相关研究,但受分析方法及研究视角等因素影响,已有研究具有一定的局限性。为此,本研究拟利用分层贝叶斯模型的灵活性、易拓展性等优点对其经典模型(BYM 模型与 SCM 模型)进行拓展,构建贝叶斯空间或时空模型,揭示高血压风险的人群分布特征、空间分布特征及其随时间演变规律,挖掘出男性与女性高血压患病风险的异同,为相关部门制订高血压防治规划及针对性地开展以人群为基础的干预策略提供参考依据,进而为其他慢性病实行多疾病联合监测与防控提供新的分析视角。

本研究的基本思路是:首先,考虑到面板数据特点,对不同性别分别采用传统的混合 logistic 回归法与面板 logistic 模型从个体层面筛选影响高血压患病的可能因素;然后,分别采用标化患病比、BYM 空间模型与 SCM 空间模型从空间视角分年份研究中国 7 省市不同性别高血压患病的空间变异,了解不同地区高血压风险的地区差异;其次,将贝叶斯样条分别引入 BYM 空间模型与 SCM 空间模型,构建 BYM 时空模型与 SCM 时空模型,揭示高血压空间分布模式随时间推移的持久性并突出不寻常的模式,为高血压病因探索提供

提示性信息;再次,将区域协变量引入 BYM 空间模型与 SCM 空间模型,探究高血压时空变异潜在来源,丰富高血压病因学研究;最后,将传统的 logistic 回归与贝叶斯疾病制图结果进行对比,分析其异同点,为制定和评价高血压卫生政策与干预措施提供依据。

目 录

第一章 导论	1
第一节 研究背景与研究意义	1
(一) 研究背景	1
(二) 研究意义	3
第二节 研究思路、基本框架与创新点	3
(一) 研究思路	3
(二) 基本框架	4
(三) 创新点	5
第三节 研究方法	6
(一) 基于个体层面的传统方法	6
(二) 基于空间层面的传统方法	6
(三) BYM 空间与时空模型	6
(四) SCM 空间与时空模型	6
第二章 基本概念与理论基础	8
第一节 基本概念	8
(一) 空间自相关	8
(二) 时空分析常用方法	8
(三) 分层贝叶斯模型与疾病制图	9
(四) 贝叶斯疾病制图	10

(五) 标化患病比	10
第二节 理论基础	11
(一) 全局 Moran 指数	11
(二) 先验分布相关问题	12
(三) BYM 基本模型	18
(四) SCM 基本模型	22
(五) MCMC 常见算法概述	23
 第三章 文献综述	28
第一节 高血压危险因素研究现状	28
第二节 疾病制图的常用方法	29
(一) “内插”制图法	30
(二) “平滑”制图法	33
(三) 小结	37
第三节 贝叶斯疾病制图研究进展	38
(一) 疾病制图的贝叶斯方法	39
(二) 贝叶斯疾病制图	41
(三) 展望	48
第四节 BYM 研究现状	48
(一) BYM 空间分析	48
(二) BYM 时空分析	50
第五节 SCM 研究现状	52
(一) SCM 空间分析	52
(二) SCM 时空分析	54
第六节 MCMC 算法研究现状	57

第四章 描述统计分析	66
第一节 资料来源与分析指标	66
(一)资料来源	66
(二)分析指标定义	67
(三)数据缺失情况	70
第二节 分析方法与分析工具	71
第三节 不同特征调查样本情况	73
(一)总体	73
(二)性别	75
(三)地区	76
(四)年份	83
(五)居住地	89
第四节 高血压患病特征	90
(一)患病率及其变化趋势	90
(二)不同特征的患病率	93
第五章 基于传统方法的高血压风险变异 来源研究	97
第一节 单因素分析与多因素分析	97
(一)单因素分析	97
(二)多因素分析	98
第二节 基于面板数据的分析	104
(一)面板数据描述性统计	104
(二)固定效应 logit 分析	108
(三)随机效应 logit 分析	112

(四)混合 logit 分析	117
第三节 模型比较	120
(一)固定效应与随机效应	120
(二)混合效应与随机效应	121

第六章 高血压时空分布特征及其演化规律研究——空间模型

123

第一节 标化患病比与空间自相关	123
(一)标化患病比分析	123
(二)空间自相关分析	127
第二节 BYM 模型	128
(一)先验参数的选择	128
(二)模型诊断与敏感性分析	135
(三)实证分析	143
第三节 SCM 模型	145
(一)先验参数的选择	145
(二)模型诊断与评估	145
(三)实证分析	146

第七章 高血压时空分布特征及其演化规律研究——时空模型

151

第一节 BYM 时空模型	151
(一)研究动机	151
(二)模型构建	152
(三)模型选择	155

(四)模型诊断与敏感性分析	156
第二节 SCM 时空模型	161
(一)研究动机	161
(二)模型构建	162
(三)模型选择	164
(四)模型诊断与敏感性分析	166
第三节 实证分析	167
(一)BYM 时空模型分析结果	167
(二)SCM 时空模型分析结果	170
 第八章 高血压时空变异来源研究 ...	176
第一节 BYM 协变量模型	176
(一)研究动机	176
(二)模型构建	176
(三)参数的选择	178
(四)BYM 模型变量选择	185
(五)模型诊断与敏感性分析	191
第二节 SCM 协变量模型	194
(一)研究动机	194
(二)模型构建	194
第三节 实证分析	195
(一)BYM 协变量模型分析结果	195
(二)SCM 协变量模型分析结果	198

第九章 研究结论与建议	201
第一节 研究结论	201
(一)方法学讨论	201
(二)高血压时空分布特征及其演化规律	206
(三)高血压时空变异潜在来源	207
第二节 政策建议	212
(一)推广健康生活方式,控制高血压危险因素流行	212
(二)加强健康教育,开展健康体检和筛查,推广高血压自我管理	213
(三)因时制宜,因地制宜,制订高血压防治中长期规划	213
第三节 总结与展望	214
参考文献	216

第一章 导论

第一节 研究背景与研究意义

(一) 研究背景

高血压是最常见的慢性病之一,也是心脑血管疾病最主要的危险因素。根据世界卫生组织(WHO)2013年的报告,全球每年有1700万人死于心血管疾病,占死亡总数的1/3,其中有940万死于高血压的并发症。具体而言,至少45%由于心脏病造成的死亡与高血压有关,而中风与之相关的比例则占到了51%^[1]。

世界各国尤其是低收入和中等收入国家居高不下的高血压患病率严重消耗了医疗和社会资源,给家庭和国家造成了沉重的负担(刘力生,2010)^[2]。据WHO报告预测,在2011年至2025年期间,低收入和中等收入国家与非传染性疾病相关的累计产出损失预计将达到7.28万亿美元,而年度产出损失共计约为这些国家国内生产总值的4%,其中心血管疾病(包含高血压)占了约一半的成本^[1]。

作为中等收入国家之一,我国高血压患病形势同样不容乐观。报告显示,我国18岁及以上居民高血压患病率为33.5%,城市略高于农村(34.7% vs. 32.9%),男性略高于女性(35.1% vs. 31.8%),东、中、西部地区的高血压患病率依次降低,分别为36.2%、34.1%、28.8%。无论城乡、男女及东

中西部,高血压患病率均随年龄增大而增加。2010 年高血压标化患病比 (standardized prevalence rate, SPR) 为 30.7%,比 2004 年和 2007 年分别增加了 49.8%和 22.3%,且各年龄组居民的患病率均有明显的升高趋势^[3],一方面反映了我国人口老龄化程度日益严峻的现实,另一方面与人们现代化的生活方式紧密相关,如不规律作息及不合理膳食引起的肥胖、应酬引致的吸烟饮酒等需求,以及日益增大的生活和工作压力等。

另据 2017 年中国科学院大学与昆明科技大学等多家科研机构联合发布的“中国心血管病死亡地图”显示,冠心病、脑血管病与高血压的死亡率均存在明显的空间差异。其中,南方高血压问题更严重,北方则是冠心病与脑血管病的重灾区。具体而言,高血压死亡率排名前四位的地区分别是湖北、广东、湖南和重庆,死亡率最低的是青海、天津、宁夏和海南^[4]。

高血压病因复杂,且男性与女性高血压患者在生理、心理与社会经济状况等方面存在差异,其患病特点及其相关因素可能也存在着差别。目前从性别视角对高血压分布模式及其关联因素的研究较少;且研究时大多采用基于个体层面的传统统计分析方法,忽略了疾病数据内在包含的空间与时间信息。因此,本研究拟以高血压为研究对象,以其分布模式和影响因素为基础,利用拓展后的分层贝叶斯模型揭示高血压风险的人群分布特征、空间分布特征及其随时间演变规律,挖掘男性与女性高血压风险的异同,为相关部门制订高血压防治规划及针对性地开展以人群为基础的干预策略提供参考依据,进而为其他慢性病实行多疾病联合监测与防控提供新的分析视角。

（二）研究意义

本研究的学术价值主要体现在以下两个方面：①将时空协变量引入贝叶斯 BYM 与 SCM 模型，模型不仅包含了随机效应还包含了固定效应，从而能够量化单种或多种疾病潜在危险因素与结果变量的协变量效应，为区域层面上疾病风险防控提供较为可靠的信息；②将贝叶斯 B 样条引入 BYM 与 SCM 模型，使得模型能够同时对空间随机效应、时间随机效应及时空交互随机效应进行建模，充分利用样本的空间与时间信息，有助于研究疾病风险的持久性模式与不寻常模式，有助于考察导致疾病风险的系统性因素与偶然性因素。

本研究的应用价值主要体现在以下三个方面：①以性别为研究视角，探讨高血压风险空间（时空）变异及其潜在来源，为相关部门分人群进行高血压防控提供理论依据；②利用分层贝叶斯模型从空间或时空视角进行研究，为疾病病因探索提供了新的研究思路，对指导高血压的有效防控具有必要性和现实意义；③本研究的关注点从以往的罕见疾病转为非罕见疾病（高血压），丰富分层贝叶斯模型的相关研究，为高血压或其他慢性病防控与干预策略的制定提供了指导。

第二节 研究思路、基本框架与创新点

（一）研究思路

本书试图基于中国健康与营养调查（China Health and Nutrition Survey, CHNS）的数据分性别研究中国七省份的高血压患病特点及其随时间变化的趋势，并探讨其可能原因。首先，考虑到面板数据特点，对不同性别分别采用传统的混合 logistic 回归法与面板 logistic 模型从个体层面筛选

影响高血压患病的可能因素；然后，分别采用标化患病比、BYM 空间模型与 SCM 空间模型从空间视角分年份研究中国七省市不同性别高血压患病的空间变异，了解不同地区高血压疾病负担的差异；其次，将贝叶斯样条分别引入 BYM 空间模型与 SCM 空间模型，构建 BYM 时空模型与 SCM 时空模型，用于揭示高血压空间分布模式随时间推移的持久性并突出不寻常的模式，为高血压病因探索提供提示性信息；再次，将区域协变量引入 BYM 空间模型与 SCM 空间模型，探究高血压时空变异潜在来源，丰富高血压病因学研究；最后，将传统的 logistic 回归与贝叶斯疾病制图的结果进行对比，比较其异同点，为制定和评价高血压卫生政策与干预措施提供依据。

(二) 基本框架

本书的基本框架如图 1-1 所示。

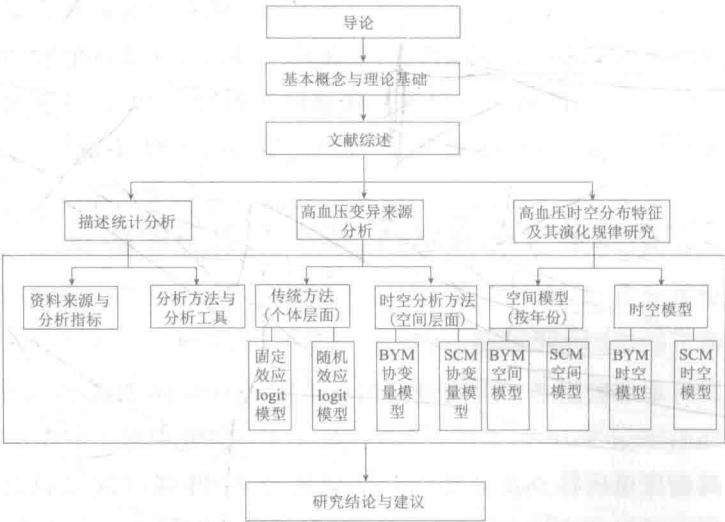


图 1-1 框架结构图

(三) 创新点

当前高血压分布模式及其关联因素相关研究集中于个体层面的传统统计分析,没有考虑疾病的空间与时间背景;且分析时仅针对总体人群,忽略不同性别高血压可能存在的患病特点差异。另外,当前贝叶斯疾病制图模型相关研究主要侧重于从空间维度建模,较少同时关注时空维度的情形。因此,本研究拟基于现有的分层贝叶斯模型框架构建时空疾病制图模型,探讨高血压风险的空间变异、时空变异及其潜在来源,无论从理论还是应用角度都具有其适用性和独特性。

1. 理论模型创新

当前研究主要关注空间建模,无法充分挖掘数据中的信息。在“大数据”时代背景下,具有分层结构的数据越来越普遍,多维度数据的可获得性大大增加,这给疾病时空数据分析带来了新的挑战。本研究探讨了同时存在时空维度时BYM与SCM模型的构建问题,丰富贝叶斯疾病制图模型相关研究,为疾病时空数据分析提供新的研究思路。

2. 研究视角创新

当前研究主要采用传统统计分析方法从个体层面分析心脑血管疾病的分布模式及其影响因素,且贝叶斯疾病制图应用研究主要关注癌症等罕见疾病。本项目从高血压等常见疾病出发,从时空分析视角探讨高血压风险变异及其潜在危险因素的效应,为高血压及其他慢性病有效防控与综合监测提供理论指导。