

US Intrastate and Interstate Opioid Crisis Prediction Model

—Predictive Analysis Based on Markov Chain Model

Bixuan Li¹, Kexin Zhu¹, Xuanlin Chen¹, Fang Wang^{1,2*}

¹College of Information Science and Technology, Hunan Agricultural University, Changsha Hunan

²Agricultural Mathematical Model and Data Process Center, Hunan Agricultural University, Changsha Hunan

Email: *topwang619@163.com

Received: Apr. 4th, 2019; accepted: Apr. 19th, 2019; published: Apr. 26th, 2019

Abstract

This article analyzes the annual use of synthetic opioids at the state and county levels in Kentucky, Ohio, Pennsylvania, Virginia, and West Virginia during the year 2010-2017. After carrying out the visual analysis of the data, the general transfer trend of two target substances can be initially determined, and the results show that the use of target substances is more serious at cities on the border. Then, we select the use of heroin as the reference sequence and make oxycodone, fentanyl, hydrocodone and buprenorphine represent the other synthetic opioids according to the results of correlation analysis. Using the Markov chain model, the use of the target substances in five studied states can be predicted, and the threshold level of the number of opioids in five states is given as reference for the governors. The results show that the US government needs to concentrate more on the control of fentanyl drugs, and the use of opioids in West Virginia is relatively optimal.

Keywords

Opioid Crisis, Correlation Analysis, Markov Chain Prediction Model, Monte Carlo Algorithm

美国州内与州际阿片危机预测模型

——基于马尔可夫链模型的预测分析

李碧璇¹, 朱可馨¹, 陈宣霖¹, 王 访^{1,2*}

¹湖南农业大学信息科学技术学院, 湖南 长沙

²湖南农业大学农业数学建模与数据处理中心, 湖南 长沙

Email: *topwang619@163.com

*通讯作者。

收稿日期：2019年4月4日；录用日期：2019年4月19日；发布日期：2019年4月26日

摘要

本文选取美国肯塔基州、俄亥俄州、宾夕法尼亚州、弗吉尼亚州和西弗吉尼亚州从2010年到2017年的县内阿片类药物鉴定数与州内阿片类药物鉴定数作为研究对象，并对该数据进行可视化分析，初步确定阿片类药物的大致转移趋势，结果显示阿片类药物主要集中在州的边境。然后利用关联度分析，以主要的阿片类药物海洛因作为参考标准，选取了海洛因、羟考酮、芬太尼、氢可待因和丁丙诺啡5种阿片类药物来衡量阿片类药物总体的使用情况。运用马尔可夫链模型，对美国5个州内阿片类药物的使用情况进行预测，给出5种阿片类药物数量的阈值水平，供美国政府参考。结果显示，美国政府需要加大力度对芬太尼药物进行控制，在西弗吉尼亚州的阿片类药物的使用情况相对最佳。

关键词

阿片危机，关联度分析，马尔可夫链预测模型，蒙特卡罗算法

Copyright © 2019 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着社会经济的不断发展，人类所承受的来自各个方面的压力不断增加。近年来，美国阿片类药物的滥用引起了美国政府的高度关注。阿片类药物指的是一类作用于阿片受体的化学物质，其作为一种镇痛药物，可以有助于减轻患者的疼痛。它能给病人一种舒适和愉悦的感觉，但长期使用会导致身心的双重依赖，一旦上瘾，似乎很难从根本上恢复。阿片类药物的滥用为美国现阶段阿片危机的主要根源。面对美国阿片类药物滥用所造成的种种问题[1]，2017年3月，美国总统特朗普签署行政令，成立打击药物成瘾与阿片类药物危机总统委员会。7月，国会要求总统宣布国家公共卫生紧急状态，以应对阿片类药物危机。孙丽[2]指出，美国阿片类药物滥用原因主要是病患缺乏知识、医生滥开处方以及药企的推广。针对阿片危机，2017年12月8日，加拿大卫生部发布公告[3]，修订阿片类药物说明书以减少滥用危害。可见，阿片危机对健康、经济和社会安全造成了严重危害。

针对如今社会的该种现象，本文利用关联度分析法提取出阿片危机的主要阿片类物质，再建立马尔可夫链模型给出了美国肯塔基州、俄亥俄州、宾夕法尼亚州、弗吉尼亚州和西弗吉尼亚州的州内与州际阿片类药物使用的影响，并预测阿片危机可能出现的时间与地点。

2. 数据的预处理

2.1. 数据来源和假设

本文选取美国肯塔基州、俄亥俄州、宾夕法尼亚州、弗吉尼亚州和西弗吉尼亚州从2010年到2017年的县内阿片类药物鉴定数与州内阿片类药物鉴定数作为研究对象，数据来源于美国大学生数学建模竞赛[4]。为了简化模型过程，本文做出以下假设：

- 1) 概率转移矩阵每转移一次代表一年的变化；

2) 五个州阿片类药物数量每年为一固定数值，一个县或一个州的减少量可视为从该县或者该州转移出的阿片类药物数量。

3) 假定美国缉毒局针对每一年在当年进行药物鉴定，每年之间的数据相互独立。

2.2. 可视化统计分析

为了简化观察药物在州内县与县之间以及州与州之间的转移情况，基于维基百科中的 5 个州内主要的县与其邻近县的地理数据以及 2010~2017 年每年内各个县合成类阿片类药物以及海洛因数量求和得到的各州内按县对应的一组数据，由此得到阿片类药物在 5 个州内的分布图，如图 1~5 所示，其中颜色的深度表示所有阿片类药物的数量：下文中阿片类药物表示合成类阿片类药物以及海洛因。

由图 1~5 可知：

1) 肯塔基州的阿片类药物主要分布在东部，东南部的阿片类药物的数量随着时间的推移而减少，而东北部的数量基本不变。

2) 2010 年上半年俄亥俄州的阿片类药物分布非常分散，但 2014 年以后，阿片类药物越来越集中在西南部 and 东北部。

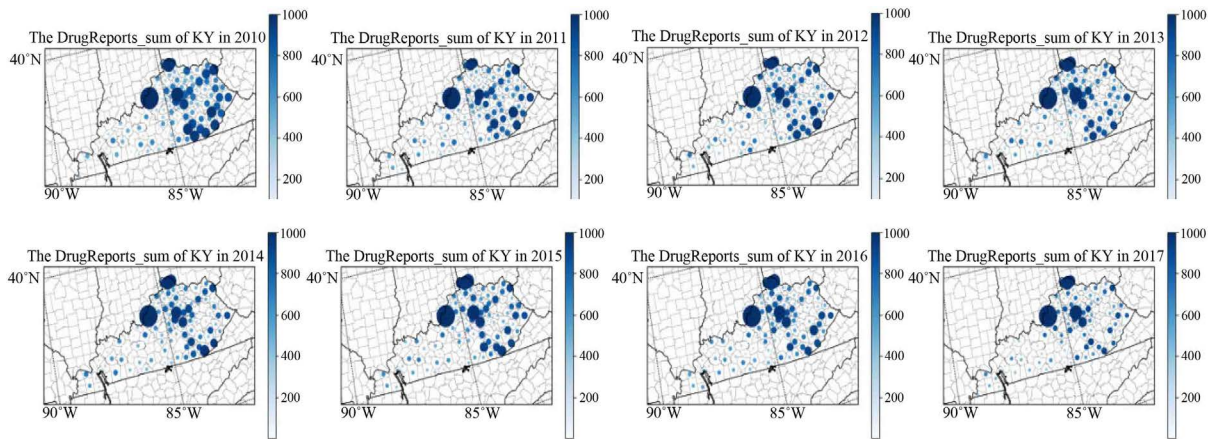


Figure 1. Distribution of Opioids in Kentucky from 2010 to 2017

图 1. 2010~2017 年肯塔基州阿片类药物分布情况

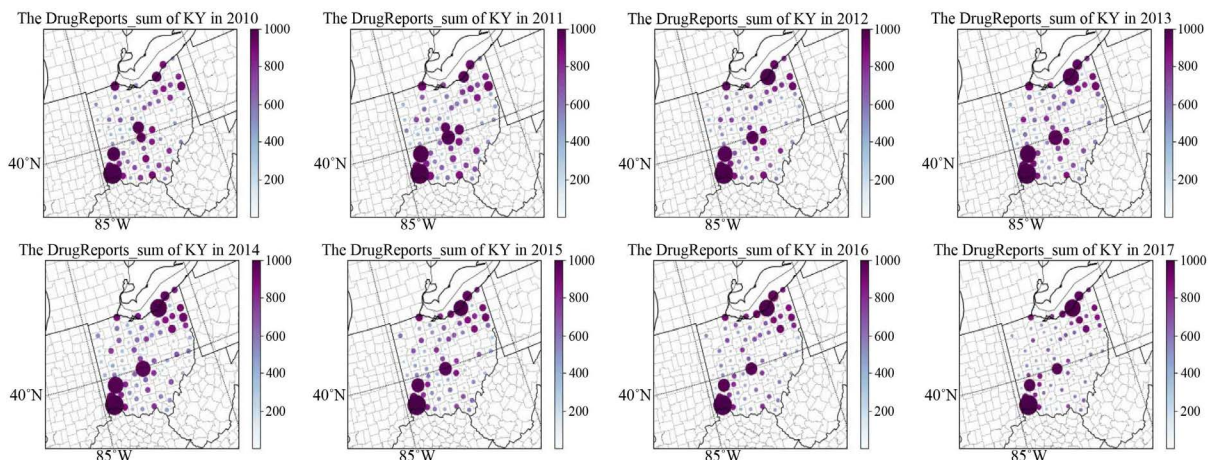


Figure 2. Distribution of Opioids in Ohio from 2010 to 2017

图 2. 2010~2017 年俄亥俄州阿片类药物分布情况

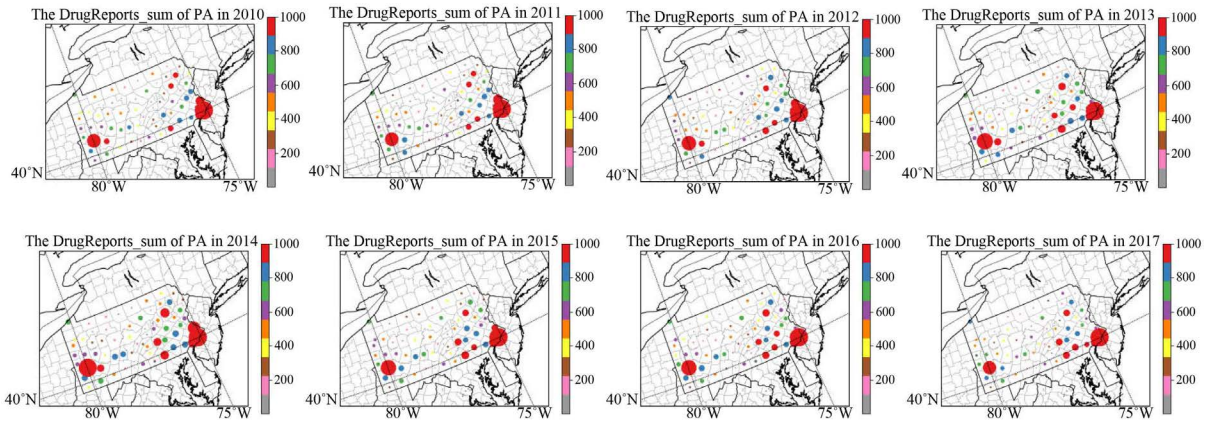


Figure 3. Distribution of Opioids in Pennsylvania from 2010 to 2017
图 3. 2010~2017 年宾夕法尼亚州阿片类药物分布情况

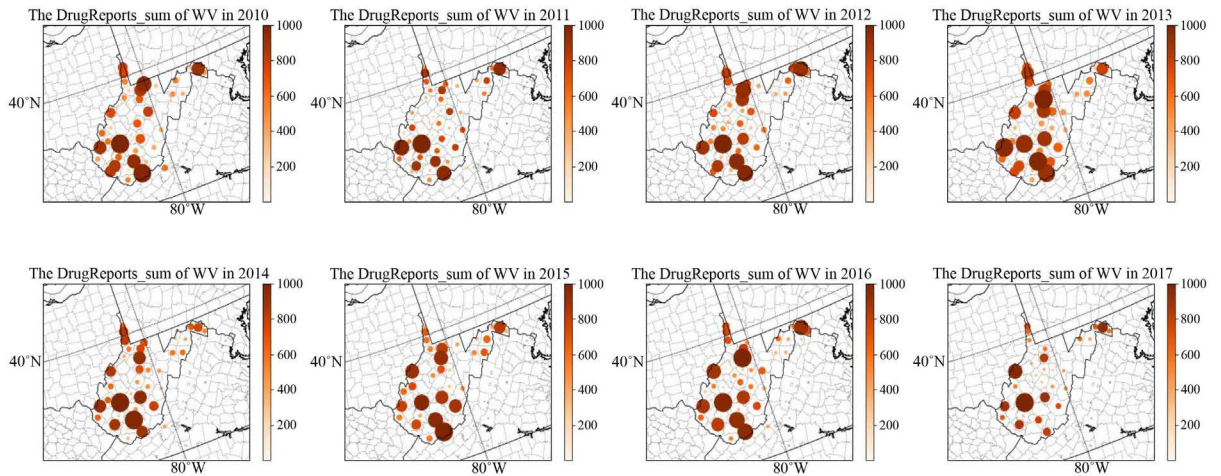


Figure 4. Distribution of Opioids in West Virginia from 2010 to 2017
图 4. 2010~2017 年西弗吉尼亚州阿片类药物分布情况

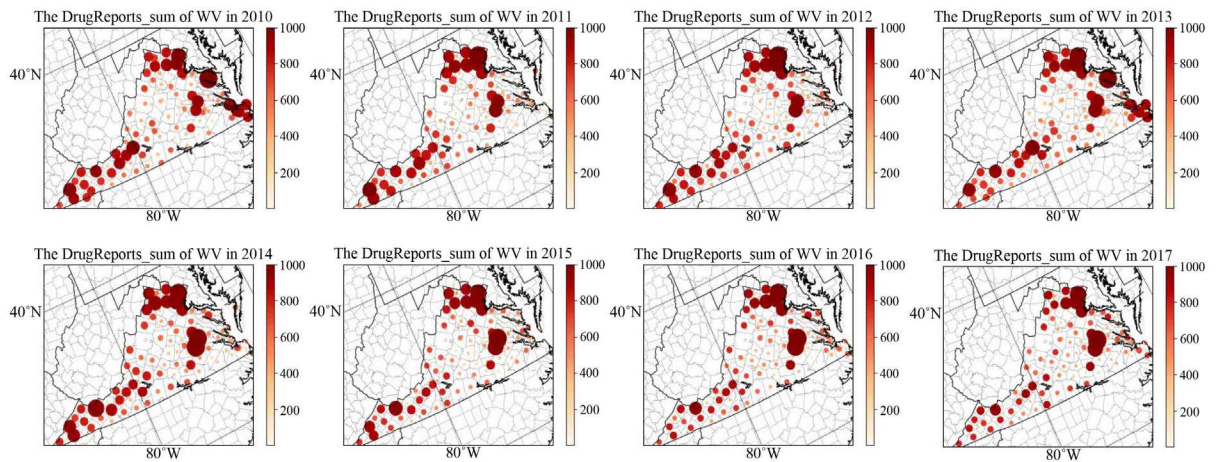


Figure 5. Distribution of Opioids in Virginia from 2010 to 2017
图 5. 2010~2017 年弗吉尼亚州阿片类药物分布情况

3) 在宾夕法尼亚州, 阿片类药物的数量基本不变。2010~2017 年, Philadelphia 和 Allegheny 两个县

的阿片类药物数量最多。

4) 西弗吉尼亚州西南部有大量阿片类药物, 尤其是 2013 年, 数量最多。但是, 该区域与该州的阿片类药物的数量呈下降趋势。

5) 弗吉尼亚州的阿片类药物主要分布在北部边境和西南部。随着时间的推移, 西南地区的阿片类药物的数量有所减少, 主要集中在少数县。

本文以上述同样的方法分析 2010~2017 年阿片类药物州际间的影响, 得到阿片类药物在 5 个州内的分布图:

由图 6 可知:

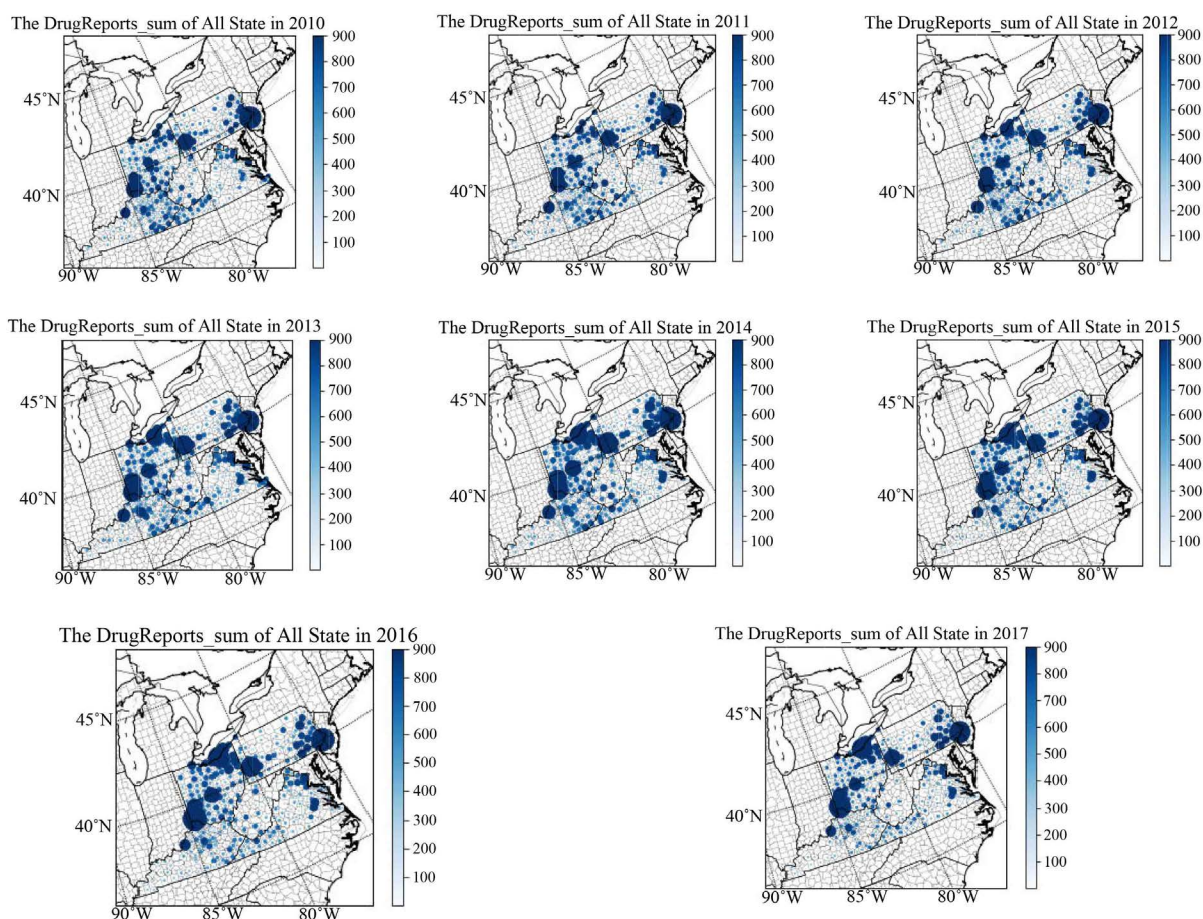


Figure 6. Distribution of all Opioids in five states from 2010 to 2017

图 6. 2010~2017 年 5 个州所有阿片类药物分布情况

1) 阿片类药物在所选取的五个州内分布极不均匀, 在州边境的县里有大量的阿片类药物, 而在远离边境的地方其数量少得多。本文作者认为可能的原因是: 由于边境缺乏政府监管, 阿片类药物以及毒品贸易猖獗, 并且由图可知, 此现象呈现不断增长的趋势。

2) 阿片类药物的分布特点表现为“大杂居, 小聚居”, 即在每个州的几个县的管辖区域内储存着大量的阿片类药物, 而在大多数县内只记录了少量的阿片类药物使用数量。由此可知, 阿片类药物的使用的聚集效应出现在某些特定的地方。

3) 西弗吉尼亚州的阿片类药物的数量从 2010 年到 2017 年呈下降趋势。

2.3. 关联度分析

由于在现实生活中阿片类药物的种类数量庞大, 本文考虑利用关联度分析[5] [6]选择几种典型的阿片类药物来衡量总体的阿片类药物使用情况。

第 1 步: 选取海洛因的使用数量作为参考序列:

$$A_0 = \{A_0(t) | t = 0, 1, 2, \dots, 7\} \quad (1)$$

其中 t 表示年份, $A_0(t)$ 表示第 t 年所有州的海洛因数量的总和。将剩下的阿片类药物作为比较序列:

$$A_k = \{A_k(t) | t = 0, 1, 2, \dots, 7\}, k = 1, 2, \dots, 68 \quad (2)$$

第 2 步: 计算比较序列 A_k 对参考序列 A_0 在 t 时刻的关联系数:

$$\gamma_k(t) = \frac{\min_s \min_r |A_0(r) - A_s(r)| + \rho \max_s \max_r |A_0(r) - A_s(r)|}{|A_0(t) - A_s(t)| + \rho \max_s \max_r |A_0(r) - A_s(r)|} \quad (3)$$

其中 $\rho \in [0, 1]$ 为分辨系数。在此, 选取 $\rho = 0.5$ 。

第 3 步: 计算每个比较序列与参考序列的关联度:

$$\phi_k = \frac{1}{8} \sum_{t=1}^8 \gamma_k(t) \quad (4)$$

第 4 步: 对所求的关联度进行排序, 如图 7 所示。

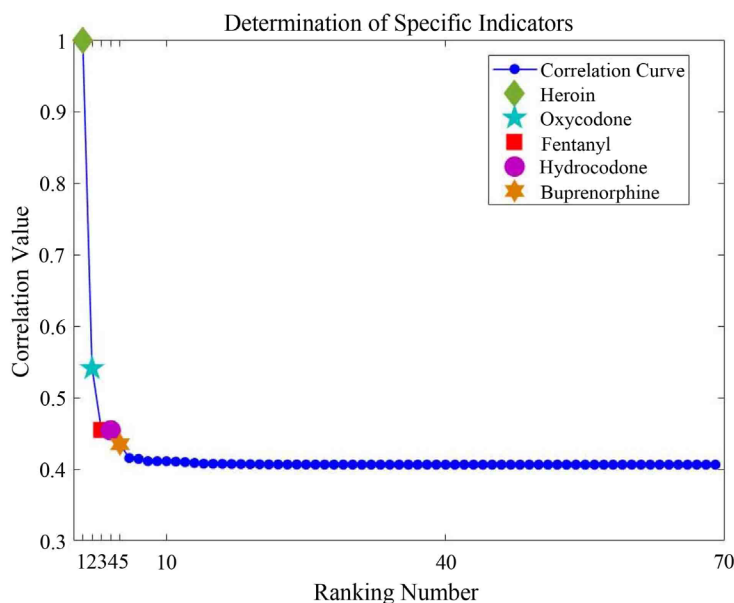


Figure 7. Correlation between substances referring to the amount of heroin
图 7. 参照海洛因数量的各物质之间的关联度

由此考虑选择前 5 种物质, 它们分别是海洛因、羟考酮、芬太尼、氢可待因和丁丙诺啡。5 种阿片类药物其总计数占有所有目标物质总计数的 90.10%。4 种合成类阿片类药物的总计数占有合成类阿片总数的 79.87%。下文中将该 5 种物质称为特定的阿片类药物。

3. 马尔可夫预测模型

为了研究阿片类药物在五个州内未来的发展趋势, 在五个州内特定的阿片类药物总量保持不变的假

设前提下, 可以视为各州(县)合成类阿片类药物和海洛因数量的增减视为相互转移。根据缉毒局年度报告的数据相互独立的假设, 因此, 用马尔可夫链模型[7] [8] [9]来进行预测。

记状态空间为 $E = \{X_1, X_2, X_3, X_4, X_5\}$, 在第 $t+1$ 年, 阿片类药物从第 i 个州转移至第 j 个州的转移概率为:

$$P\{\xi_{t+1} = X_j | \xi_t = X_i\} = p_{ij} \quad (5)$$

因此, 5 个州之间特定的阿片类药物转移的一步转移概率矩阵可记为:

$$P = \begin{pmatrix} p_{11} & p_{12} & p_{13} & p_{14} & p_{15} \\ p_{21} & p_{22} & p_{23} & p_{24} & p_{25} \\ p_{31} & p_{32} & p_{33} & p_{34} & p_{35} \\ p_{41} & p_{42} & p_{43} & p_{44} & p_{45} \\ p_{51} & p_{52} & p_{53} & p_{54} & p_{55} \end{pmatrix} \quad (6)$$

其中, 该矩阵满足以下条件:

- 对于任意的 $X_i, X_j \in E$, 有 $p_{ij} \in [0, 1]$ 成立;
- 对于任意的 $X_i \in E$, $\sum_{j \in E} p_{ij} = 1$ 。

本文将按如下步骤来确定一步转移概率矩阵:

1) 记对角线上的元素为保留概率, 将下一年的某一特定的阿片类药物总数与当年的该种特定的阿片类药物总数之比与和 1 进行比较, 如果比值大于或等于 1, 则保留概率记为 1, 否则记为 0。

2) 保留概率为 1 的行, 该行其他列的元素记为 0。基于前提假设 2, 将五个州特定的阿片类药物总量视为衡量, 针对比值小于 1 的行, 其他列的元素将按如下公式计算:

$$p_{ij} = \frac{[(1 - p_{ii})] * |X_j(t+1) - X_j(t)|}{|X_1(t+1) - X_1(t)| + \dots + |X_{i-1}(t+1) - X_{i-1}(t)| + |X_{i+1}(t+1) - X_{i+1}(t)| + \dots} \quad (7)$$

即每个元素的值是根据 $1 - p_{ii}$ 按 t 年和 $t+1$ 年特定的阿片类药物的州之间差异的比例来分配。

从上述两个步骤中, 可以得到 2010~2017 年 7 个一步概率转移矩阵, 将该 7 个矩阵元素之间的平均值作为最终一步概率转移矩阵中的元素, 由此得到的一步转移概率矩阵是齐次的。在构建州际间概率转移矩阵的过程中, 以下对几种特殊情况作解释:

a) 在部分州中, 由于政府在对其数据进行统计时没有检测出特定的阿片类药物。因此, 将缺失年份的数量记为 0。

b) 如果在 $t+1$ 年特定的阿片类药物数量减少至 0, 则将其保留概率记为 0。

c) 如果在 t 和 $t+1$ 年特定的阿片类药物数量都为 0, 则将其保留概率记为 1。

在分析州内县与县之间的传播特征时, 同样按上述方法求解, 本文不再重复。本文将马尔可夫链的一步转移概率矩阵的图像如表 1 形式表示:

由表 1 中图像可知:

1) 对于海洛因, 只有西弗吉尼亚州的马尔可夫链是遍历的, 这表明海洛因可以转移到该州的每个县。根据其平稳分布可知: 三个较大的转移概率对应的县是 Logan, Kanawha 以及 Harrison。表明它们是潜在的海洛因交易中心。对于其他四个州而言, 肯塔基州的海洛因药物滥用县分别是 Caldwell, Fulton, Knott, Monroe, Morgan, Washington 以及 Wayne; 俄亥俄州的海洛因滥用县是 Summit; 宾夕法尼亚州的海洛因滥用县是 Lancaster; 弗吉尼亚州的海洛因药物滥用县分别 Dinwiddie, Halifax, King and Queen, Lee, Mecklenburg, Nelson, Bedford City 以及 Martinsville City。

Table 1. Image of one-step probability transition matrix, The horizontal white lines in each image represent the absorption state in the Markov process, Under each image, the state of absorption or traversal is shown. The numbers in the absorption brackets indicate the number of crisis centers, and the numbers in the traversal brackets indicate the time when the crisis occurred

表 1. 一步转移概率矩阵的图像，其中每幅图像中的横排列的白线表示马尔可夫过程中的吸收状态，在每幅图像下，显示了吸收或遍历的状态。吸收边括号内的数字表示危机中心的数目，遍历旁边括号内的数字表示危机发生的时间

	肯塔基州	俄亥俄州	宾夕法尼亚州	弗吉尼亚州	西弗吉尼亚州
海洛因					
遍历/吸收	吸收(7)	吸收(1)	吸收(1)	吸收(8)	遍历(113)
羟考酮					
遍历/吸收	吸收(1)	遍历(20)	遍历(26)	遍历(40)	遍历 (11)
芬太尼					
遍历/吸收	吸收(17)	吸收(21)	吸收(13)	吸收(11)	吸收(3)
氢可待因					
遍历/吸收	遍历(7)	遍历(6)	遍历(3)	吸收(1)	吸收(1)
丁丙诺啡					
遍历/吸收	吸收(1)	吸收(1)	吸收(1)	吸收(5)	遍历(4)

2) 对于羟考酮，除肯塔基州外，其他四个州的马尔可夫链都是遍历的。表明羟考酮可以转移到这四个州的每个县。在肯塔基州，唯一的“吸收县”是 Hancock。根据这四个州的平稳分布可知：它们的较大概率县是 Franklin, Philadelphia, Henrico, Raleigh。

3) 对于芬太尼，由图可知该药物对所有的州都不是遍历的，而是分别具有多种吸收态。表明美国对该阿片类药物依赖情况严重。

4) 对于氢可待因，肯塔基州、俄亥俄州和宾夕法尼亚州的马尔科夫过程是遍历的，而弗吉尼亚州和西弗吉尼亚州的马尔科夫过程存在吸引子，弗吉尼亚州和西弗吉尼亚州的唯一“吸引县”分别是 Covington City 和 Tucker。

5) 对于丁丙诺啡，只有西弗吉尼亚州的马尔科夫过程是遍历的，其中概率最大的是 Marshall。前三个州的“吸收县”分别是 Edmonson, Monroe, Crawford。弗吉尼亚州的“吸收县”有 Bath, King And Queen, Montgomery, Westmoreland 以及 Fairfax City。

为了确定特定的阿片类药物在州际间的传播特征，对其以上述相同的方法进行描述，求出它们的马尔可夫链都是遍历的。根据表 2 中结果可知，这些特定的阿片类药物在俄亥俄州的出现可能性最大，即 5 种特定类阿片物质以较大的概率将全部扩散到俄亥俄州。反映出如果俄亥俄州政府不采取措施积极应对，该州将来这 5 种阿片类药物将出现滥用灾难。针对宾夕法尼亚州而言，其政府应该加强对海洛因和羟考酮用量的控制；肯塔基州政府应该在芬太尼控制上加大力度；而对于弗吉尼亚州政府需要特别关注氢可待因和丁丙诺啡的使用情况。

Table 2. Stationary distribution of five specific Opioid drugs

表 2. 五种特定阿片类药物的稳态分布

	肯塔基州	俄亥俄州	宾夕法尼亚州	弗吉尼亚州	西弗吉尼亚州
海洛因	0.1278	0.3914	0.3594	0.0759	0.0455
羟考酮	0.1886	0.2950	0.2311	0.2058	0.0795
芬太尼	0.2111	0.4375	0.1436	0.1469	0.0608
氢可待因	0.2524	0.2847	0.1284	0.2718	0.0627
丁丙诺啡	0.1851	0.3436	0.2057	0.2259	0.0397

最后，本文引入阿片类药物数量阈值的概念，政府通过参考此数据来把控社会的稳定与经济的发展。由于若存在一个正数 q ，使得 P^q 的每个元素都是正数时， P^q 被称为常返态，则本文提出计算阈值以及产生危机的时间的方法如下：本文将同种阿片类药物通过转移能返回到当前州的所需转移次数 q 视为危机在 q 年后可能出现。将保留概率乘以相应特定的阿片类药物数量作为出现阿片危机时药物数量的阈值。

4. 模型检验

在本文 2.3 节中，关联度分析中的分辨系数 ρ 取 0.5 并对其结果进行了展示。由前文可知：第 5 个和第 6 个物质之间关联度的差值对于主要物质的选择起确定性作用，若后 64 个阿片类物质的关联度差异较小则满足模型成立条件。借用蒙特卡洛算法的思想，对 ρ 产生 10,000 个 0 到 1 之间的随机数，通过对每一个可能的 ρ 进行 10,000 次实验，发现对于任意 $\rho \in [0,1]$ ，其结论依旧成立。令 ρ 以 0.2 的步长，从 0.1 取至 0.7，并计算其相关度，其结果如图 8 所示。

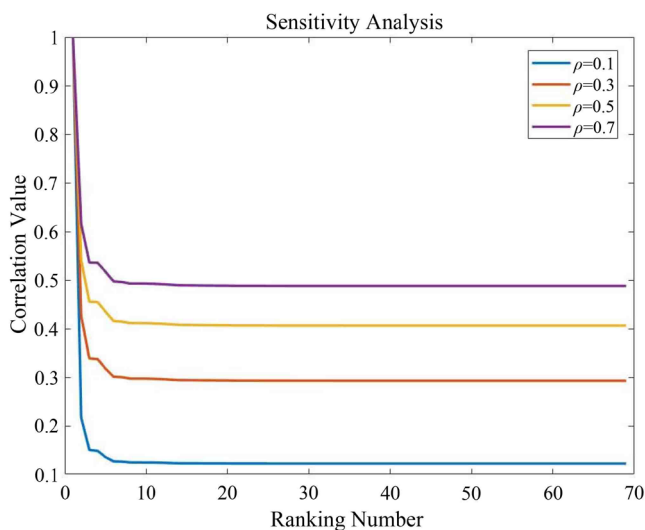


Figure 8. Sensitivity test in correlation analysis

图 8. 关联度分析中的灵敏度检验

5. 结语

本文选取美国肯塔基州、俄亥俄州、宾夕法尼亚州、弗吉尼亚州和西弗吉尼亚州从 2010 年到 2017 年的县内阿片类药物鉴定数与州内阿片类药物鉴定数作阿片类药物的大致转移趋势图, 给出了州际与州内之间阿片类药物的影响情况。然后利用关联度分析, 由于海洛因在所有的阿片类药物中都具有竞争优势, 因此, 以海洛因的使用数量作为参考标准, 最终选取了 5 种阿片类药物来衡量阿片类药物总体的使用情况。运用马尔可夫链模型, 对美国 5 个州内阿片类药物的使用情况进行预测, 给出 5 种阿片类药物数量的阈值水平以及各州可能产生危机的时间。模型检验结果显示, 该模型与关联度分析中分辨系数的选取无关。

基金项目

本文由湖南农业大学大学生创新性实验计划项目(SCX1802, XCX18007)资助。

参考文献

- [1] 郝伟, 赵敏. 美国阿片类药物危机根源: 芬太尼类药物是压死骆驼的最后一根稻草[EB/OL]. <https://xw.qq.com/SOT/20181202009033/SOT201812020090330E>, 2018-12-02/2019-03-02.
- [2] 加拿大修订阿片类药品说明书以减少滥用危害[J]. 中国医药导刊, 2018, 20(2): 108.
- [3] 孙丽. 美国阿片类药物滥用及其管控措施[J]. 中国药物滥用防治杂志, 2018, 24(4): 219-224.
- [4] <http://www.comap.com/undergraduate/contests/mcm>
- [5] 影响城市土壤重金属污染因子的关联度分析[J]. 土壤学报, 2003, 40(6): 921-928.
- [6] 王瑞玲, 陈印军. 我国“三废”排放的库兹涅茨曲线特征及其成因的灰色关联度分析[J]. 中国人口·资源与环境, 2005, 15(2): 42-47.
- [7] 夏乐天. 马尔可夫链预测方法及其在水文序列中的应用研究[D]: [博士学位论文]. 南京: 河海大学, 2005.
- [8] Ethier, S.N. and Kurtz, T.G. (2008) Markov Processes: Characterization and Convergence.
- [9] Tierney, L. (1994) Markov Chains for Exploring Posterior Distributions. *Annals of Statistics*, **22**, 1701-1728. <https://doi.org/10.1214/aos/1176325750>

知网检索的两种方式:

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>
下拉列表框选择: [ISSN], 输入期刊 ISSN: 2325-2251, 即可查询
2. 打开知网首页 <http://cnki.net/>
左侧“国际文献总库”进入, 输入文章标题, 即可查询

投稿请点击: <http://www.hanspub.org/Submission.aspx>

期刊邮箱: sa@hanspub.org