



扫一扫上面的二维码图案，加我为朋友。

分析阿片类药物危机的蔓延

摘要

根据合成和可获得性，我们将 69 种阿片类药物分为四类。绘制随时间推移的每种类别的使用率表明，随着时间的推移，温和止痛药和天然生物碱的使用保持相对稳定，半合成药物的使用量略有下降，而芬太尼和海洛因等合成药物的使用量则急剧增加。这些发现与 CDC 的报告一致。我们基于 149 个社会经济变量的方差膨胀因子得分（多重共线性的一种常用度量）以及基于公共卫生文献的相关性，选择了 54 个。

为了模拟阿片类药物危机在肯塔基州，俄亥俄州，宾夕法尼亚州，西弗吉尼亚州和弗吉尼亚州的扩散，我们开发了两个完全不同的模型，然后进行比较。

我们的第一个模型建立在流行病学常用的建模方法之上：SIR / SIS 模型和随机模拟。我们设计了一种算法，使用阿片类药物滥用发生率和变化率的阈值来模拟六个离散类之间的随机游动，每个类代表阿片类药物危机的不同阶段。我们根据实际期望对某些类之间的过渡进行惩罚。由全局空间自相关统计量 Moran's I 启发的误差函数指导模型参数和系数的优化。通过误差计算和可视化映射测试我们的模型可以说明数百项试验的高精度。但是，该模型不能很好地了解社会经济因素对阿片类药物滥用的影响率，因为并入这些因素不会显著改变模型结果。

我们的第二个模型弥补了这一不足。在社会经济预测因素（包括总药物使用率）上运行空间回归模型，我们使用空间滞后，空间误差和空间杜宾模型探索由于溢出效应和与空间相关的风险因素而导致的阿片类药物危机的空间格局。尽管所有模型都确认了重要的空间信号，但空间杜宾模型表现最佳。我们还计算了每个预测变量对阿片类药物滥用率的直接，间接和总影响。在所有模型中，最重要的变量是每个县的总毒品使用率。（在整个七年中）的平均结果是，非法药物总使用率的单位提高将使阿片类药物滥用率提高 52%。鉴于 CDC 统计数据显示，在 2014 年，药物过量死亡人数中有 61% 与某种阿片类药物有关，这是非常现实的。相比之下，普通线性回归报告的单位总药物使用率增加中，阿片类药物滥用率仅增加 37%。诸如 Akaike 信息准则和似然比检验之类的统计量证明了我们空间模型的优越性。

为了预测五个州中每个州的阿片类药物流行病的起源，我们对 2000 年至 2010 年的随机游走模型进行了蒙特卡洛模拟。我们对这些县进行了制图，并在其他发现的背景下讨论了它们的分布。随机游走发现，阿片类药物危机最有可能始于肯塔基州的蒙哥马利，这一结论与关于阿片类药物滥用在农村社区中比在城市社区中更普遍的研究相吻合。

为了预测 2017-2020 年阿片类药物危机的蔓延，我们使用了随机游走模型和空间回归模型。两种模型之间的差异极小，尤其是在 2019 年和 2020 年。随机游走预测，在未来 7 年内，非法使用阿片类药物使用阈值以上的县将减少，这与阿片类药物流行的观点一致遵循在传染病流行病学中看到的溢出效应。

由于我们假设社会经济指标呈线性变化，因此第二个模型的误差在大约 4-5 年后显著增加。另一方面，随机游走是在想要聚类 and 随机分配类之间的健康压力下进行的。在初始日期附近，它聚集更多；但是随机性开始迅速复合。因此，随机游走的误差在 4-7 年大关附近最低。这意味着，预测未来的最佳策略是 1-3 年的空间回归，以及 4-7 年范围内的随机游走。预测超出此点的任何内容都将产生高误差。



扫一扫上面的二维码图案，加我为朋友。

介绍和问题陈述

滥用处方麻醉止痛药，海洛因和合成阿片类药物的致命后果正在影响所有 50 州和所有社会经济阶层的人们。阿片类药物的流行每天夺去 115 人的生命[卫生 & Services Administration 2019]。通过医疗保健费用，康复治疗，生产力下降，刑事司法介入，阿片类药物危机每年给美国联邦政府造成的损失估计为 785 亿美元[美国国家药物滥用研究所，2019 年]。我们的团队需要执行以下建模任务：

- 描述阿片类药物在肯塔基州，俄亥俄州，宾夕法尼亚州，弗吉尼亚州和西弗吉尼亚州的流行情况，并分析由此产生的模式；
- 将社会经济因素纳入我们的模型，并分析它们与阿片类药物滥用率之间的关联（如果有）；和
- 利用这些模型的结果来建议与阿片类药物流行作斗争的公共政策策略。

为了进行这些分析，我们仅限于提供社会经济指标的美国社区调查（ACS）和有关非法药物使用的国家法医实验室信息系统（NFLIS）提供的 2010-2016 年数据。所有数据均在县一级提供。

为了描述阿片类药物危机在这五个州的蔓延特征，我们开发了两个模型：

- 第一个模拟稳定，流行和流行阶段的随机游走。
- 第二个是空间回归模型的标准集合。

在描述了这两种建模方法及其结果之后，我们报告了这两种模型对阿片类药物流行病未来传播的预测，并比较了结果。这种双管齐下的方法为阿片类药物危机的性质提供了多种见解，并帮助我们确定了政府干预的策略。

阿片类药物危机的病因

流行病的简要时间表

阿片类药物在 1990 年代的研究表明，疼痛治疗不充分之后，出现在非癌症性疼痛市场。制药公司和医学团体确信阿片类药物不会使人上瘾（通过一些错误的研究）[Rummans 等。2018]。因此，第一波阿片类药物处方开始了。

随着上瘾的开始，第二波浪潮发生在 2010 年左右。政府组织对阿片类药物处方进行了限制。许多已经上瘾的人转而使用海洛因，因为海洛因经常不纯或与其他药物混合使用，导致死亡人数增加[国家药物滥用研究所，2018 年]。

第三波浪潮是在 2013 年随着诸如芬太尼之类的合成阿片类药物的出现而出现的。

风险因素

当前，年龄似乎对成瘾的易感性影响最大。一个人越年轻，越容易上瘾：接受过 18 至 30 岁治疗方案的个人中，有 74% 的人在 17 岁之前就开始滥用药物。但是，大多数因海洛因和处方止痛药成瘾而入院的人开始 25 岁以后使用毒品[药物滥用和精神卫生服务管理局，2014 年]。

常见的流行病学模型

公共卫生研究中的几种类型的模型和总体原则可指导我们完成给定任务的方法。

隔室模型

隔室模型通常用于简化传染病的数学建模。假设将每个种群的性质以及每个种群之间的转移时间的假设划分为多个种群。在 SIR 模型中，将人群分为三类：易感人群（S），感染人群（I）和迁出人群（R）。S 是尚未感染但容易感染的个体；我是那些被感染并能够传播疾病的人。R 族人由于已经恢复免疫力，被隔离或死亡而无

更多优秀论文（按模型或者时间分类），关注公众号“科研交流”，回复“优秀论文”
法再感染该疾病。



扫一扫上面的二维码图案，加我为朋友。

随机模拟

AA Brownlea 将时间因素纳入了澳大利亚卧龙岗市传染性肝炎的传播模型。他模拟了从感染起源开始以环形 式前进的随机扩散[Emch 等。2017]。

我们的模型

我们结合动态隔室和随机模拟的方法来模拟阿片类药物危机的扩散。

我们的需求与 Greenwood 模型有所不同，因为我们的人口规模和“被感染”的可能性都是动态的，尤其是当我们将社会经济因素作为预测因素时，这些预测因素会随着时间而变化。因此，我们第一个模型的基础是时间非均质的随机游走，其中每个县都充当“代理人”。

模型的基础

术语

- 患病率是人口中在特定时间点患病的人口比例。
- 流行病同时给人口，地区或社区内的大量个人带来负担。
- 在给定区域中持续存在的疾病称为地方病。
- 通过强烈传播发生的疾病表现出高且持续的发病率，称为高流行病。
- 如果一个采样单元中某些因素的存在或多或少地使该因素在相邻采样单元中的可能性出现，则现象表现出空间自相关性[Klinkenberg 2019]。

假设条件

1. 我们没有数据的所有县的滥用率低或没有阿片类药物的滥用；我们的数据集中报告了所有具有非法阿片类药物病例的县。
2. 三个月是一个县可以在阿片类药物流行的不同阶段之间进行过渡的最短时间。
3. 受教育程度是收入和医疗保健状况的代表，血统和语言共同表明种族，而退伍军人身份可以代替残疾人。
4. 在我们拥有数据的时间跨度（2010-2016 年）之后的五年内，可以接受社会经济指标的线性外推。

总体概念

使用聚合数据表征社会或生态现象的危险是生态谬论的诱惑：断言在一个分析规模上确定的关联在更大或更小的规模上都是有效的[Emch 等。2017]。我们将关于个人阿片类药物滥用的一些逻辑纳入我们的县阿片类药物使用率模型中，但并未声称县级模式代表个人的经历。

一个主要的问题是，健康结果的空间自相关是由于溢出（扩散）引起的，还是由彼此接近，具有相似社会，经济和环境特征的区域来解释，从而导致相似的健康结果。我们探索溢出效应，与空间相关的风险因素，以及两者的结合。

探索药物报告数据

跟踪在美国法医实验室信息系统（NFLIS）数据中标记的所有 69 种药物会很麻烦，而且可能不会产生有用的结果。因此，为了诊断 NFLIS 数据中有意义的趋势，我们基于化学合成和可用性将阿片类药物分为四类：

- 美沙酮，丁丙诺啡等：轻度止痛药有时用于治疗阿片类药物成瘾，在临床上很容易获得，并且不像其他阿片类药物那样受到严格的管制。



扫一扫上面的二维码图案，加我为朋友。

- 氢可酮，羟考酮等：半合成阿片类药物，是最易上瘾和致命的药物，是 21 年处方阿片类药物过量死亡的最大原因[疾病控制与预防中心，2019 年]。它们的合成特性使它们在实验室外生产变得更加困难。他们很可能被滥用为处方药可忽视的分销漏洞。

- 芬太尼，海洛因，U-48800 等：近年来类别最大，滥用和过量死亡人数增多。它主要包括合成阿片类药物。海洛因尽管属于半合成性质，但仍属于该类别，因为它是非法的，因此，与其他合成药物一样，如果没有专业实验室的协助，它就不可能完全制成。

- 吗啡，可待因等：罂粟的天然生物碱。它们的功能不及半合成和合成的表亲。可待因可以在各种泰诺止痛药中找到。

图 1 显示，在我们拥有数据的六年中，每种类别（合成药物除外）中阿片类药物的流行率保持相对恒定。根据疾病控制中心的报告 [Katz and Sanger-Katz 2018]，合成阿片类药物在 2011 年后变得越来越普遍。

建立模型

第一部分：流行病的特征

年度和县级汇总数据的离散性使我们考虑离散时间和离散空间（DTDS）马尔可夫模型。如果过程中未来状态的条件概率仅取决于当前状态，而不取决于过去状态，则随机过程具有 Markov 属性。我们的直觉是，就像一个从成瘾中康复的人一样，一个郡如果不做出持续的努力来解决药物滥用问题，可能会倒退。因此，一个县的现状直接影响其未来的状况。

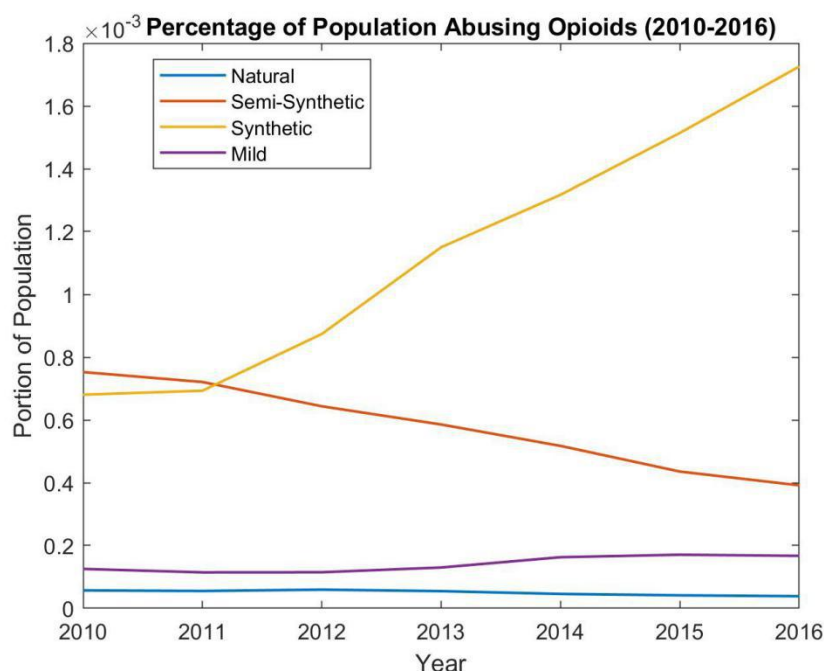


图 1. 2010-2016 年每种类别的药物趋势。

离散度的发生率和变化率

我们没有对每个县的阿片类药物滥用的确切患病率进行统计，而是将每个县分为“低”或“高”，“增加”或“减少”阿片类药物患病率。

我们选择中位数作为“低”和“高”患病率之间的临界值。

我们将大于均值一个标准偏差的所有值归为“增加”，而小于均值一半以下的值均归为“减小”。我们对差异的理由是，这些数据来自阿片类药物危机全面

更多优秀论文（按模型或者时间分类），关注公众号“科研交流”，回复“优秀论文”爆发的年份，因此数据集中的“正常”通常不是正常的。而且，一个县要比开发片类药物滥用问题更难进行恢复，因此负向的较小变化代表该县的较大变化。这个临界值之间的值被视为“稳定”。

没有数据的县，我们将其分类为“低稳定度”。

我们的数据类别

在图 2a 中，垂直线（红色）代表我们数据的中位数。在图 2b 中，左（橙色）和右（红色）垂直线代表减小和增大的截止值。水平轴的刻度为 10^{-3} 。我们有六类：低和稳定（LS），高和稳定（HS），低和递增（LI），高和递增（HI），低和递减（LD）和高且递减（HD）。

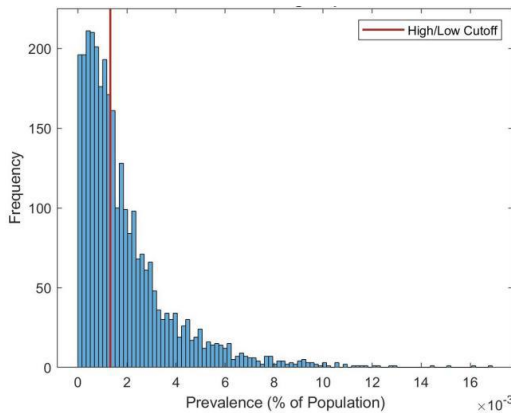


图 2a. 患病率。

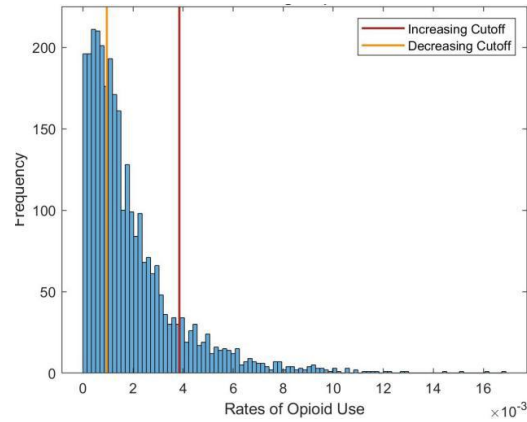


图 2b. 变化率。

图 2. 流行和减少/增加阿片类药物滥用的临界值。

在 2010 年至 2016 年期间，每个类别中县的比例保持不变，但是类别的空间分布却没有。实际上，这种模式使我们想起了布朗利肝炎模型的描述，它是一个环形的临床前部，从源头开始放射状扩展 [Emch 等。2017]，在这种情况下，显然是西南维吉尼亚州西南部。图 3 中的地图描绘了 2010 年和 2016 年每个县的阿片类药物滥用率。注意，深红色部分呈“环状”膨胀（表明阿片类药物滥用高且稳定的地区）。图例中的数字对应于以下六个类别：1 - 低稳定，2 - 高稳定，3 - 低递减，4 - 高递减，5 - 低增加，6 - 高增加。

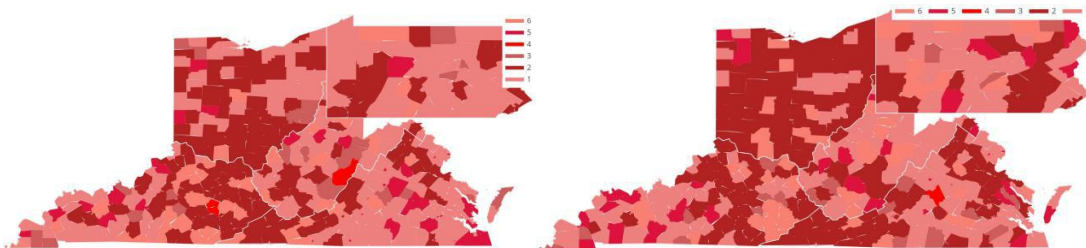


图 3. 从 2010 年（左）到 2016 年（右）阿片类药物的流行。

模型如何运作

在模型中下一步骤的县的州（类别）受其当前状态，邻近县的状态，噪声参数以及从概率向量中的随机选择影响。我们的 Python 代码执行以下任务：

初始化县概率向量，此向量为 $1 \rightarrow 6$ 向量，其中包含该县将位于概率中的预定类别中的概率

下一个时间步。通过将噪声均匀地添加到每个类别来进行初始化，以使一个国家从一个类别过渡到任何其他类别的概率大于零。优化了添加的噪声级别，以最大



扫一扫上面的二维码图案，加我为朋友。

程度地减少误差（在下面讨论）。

1. 每个邻近县都将计数添加到对应于邻近县类别的县概率向量的条目中。该计数包含反向距离权重。换句话说，一个直接相邻的县的计数相加为 1，一个县点之间的计数相加，源县之间的计数相加为 1/4，依此类推。

将概率向量中的每个条目除以下面的罚分矩阵中的相关转换得分，这反映了我们对类之间转换的难度的期望。如果我们的算法将一个县分类为高稳定度，但是数据表明该县应该是高度稳定的，那么我们希望该误差反映出我们的模型很接近并且没有错误地将那个县归为低患病率这一事实。这个需求解释了惩罚矩阵。

	<i>LS</i>	<i>HS</i>	<i>LI</i>	<i>HI</i>	<i>LD</i>	<i>HD</i>
<i>LS</i>	1	4	2	3	2	1
<i>HS</i>	4	1	3	2	3	2
<i>LI</i>	2	3	1	2	2	2
<i>HI</i>	3	2	2	1	4	2
<i>LD</i>	2	3	2	4	1	2
<i>HD</i>	3	2	2	2	2	1

2. 以与优化噪声相同的方式，通过预先优化的系数向量来缩放概率向量，以解决每个类别的总体重要性。
3. 归一化概率向量，使其成为离散的概率分布。
4. 最终概率向量定义了我们从中采样以确定下一个时间段的县级的分布。
5. 对一定数量的时间步长重复上述步骤，然后再进行一定数量的试验，以便每个试验都经过指定数量的时间步长。

在运行了详细的可能结果的蒙特卡洛模拟之后，我们在最感兴趣的分析时选择了郡中最可能的类别。

模型评估

误差分析对于该模型而言是棘手的，因为好的结果不一定意味着与数据的完美匹配。模型结果的聚类比实际数据大得多，如图 4。

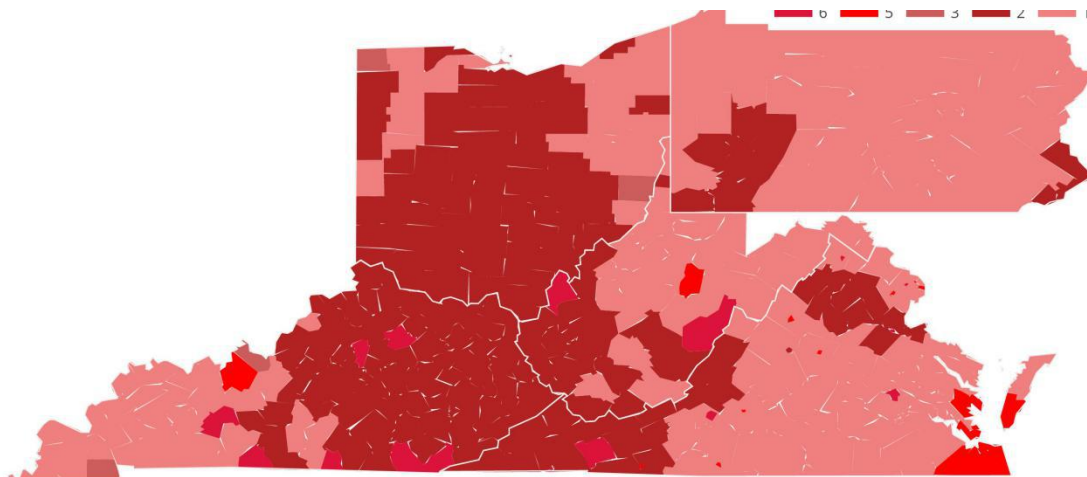


图 4. 2016 年的模型聚类。

而不是试图刻画每个位置的聚类，我们基于整体空间自相关设计了一个误差函数。我们期待 Moran 的 I 统计量[Emch 等。2017]（空间自相关的标准度量）以启发并提出以下公式。对于每个类别，对于该类别中的所有县，我们的空间自相关度量是具有相同类别的邻居的平均比例。然后，我们对模拟和实际空间自相关度量之间的差异类别进行求和。



扫一扫上面的二维码图案, 加我为朋友。

$$A_{\text{class}} = \frac{1}{\text{number of counties of that class}} \times \sum_{\text{counties of that class}} \left(\frac{\# \text{neighbors with same class}}{\# \text{neighbors}} \right),$$

$$\text{Err}_{\text{total}} = \sum_{\text{classes}} |A_{\text{predicted}} - A_{\text{real}}|.$$

使用此误差函数来优化模型中的噪声和其他参数水平, 可以消除聚类, 使我们得到的结果类似于图 5 中所示的结果。注意: 每次运行模型时, 由于随机因素, 我们会得到略有不同的结果噪音的性质。

模型测试

六个类别的直方图随时间或多或少保持恒定, 这取决于添加到初始概率分布的噪声级别。噪声较大的模型起初往往会出现分歧, 但到 2016 年, 噪声复合非常严重, 以至于噪声模拟的分布

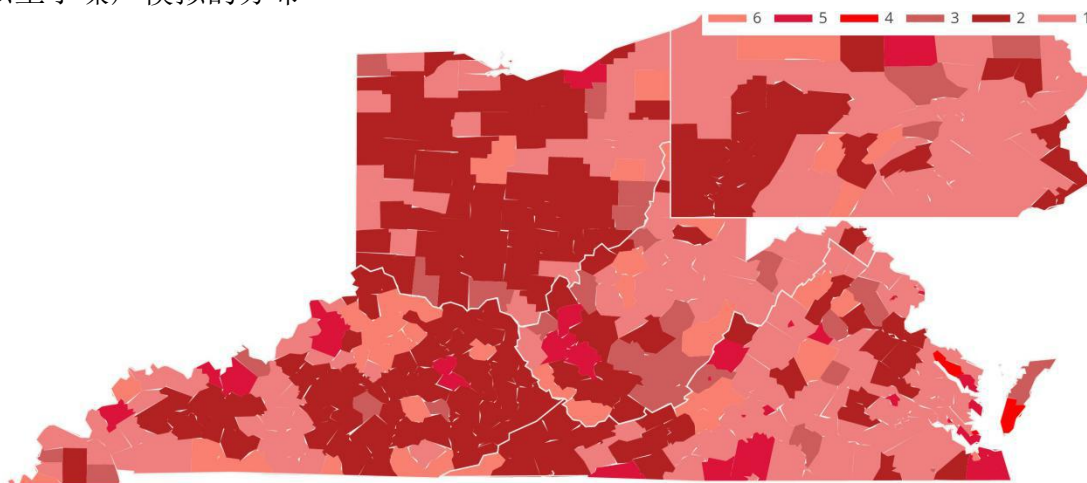


图 5. 2016 年该模型的最佳预测。

比没有噪声的模拟更紧密地匹配数据。我们发现最佳噪声水平为 0.3 (图 6)。

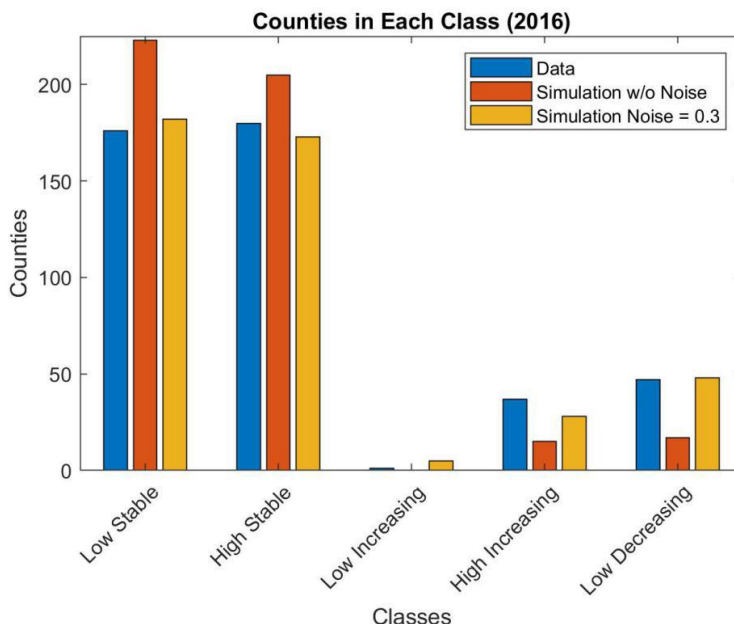


图 6. 2016 年数据与模拟的类别患病率。



扫一扫上面的二维码图案，加我为朋友。

由于随机游走是由 2010 年的数据初始化后自行产生的，因此没有明显的参考可以推断出最佳的时间步长。系数的优化将使模型适合数据，而与步长无关，因此我们选择三个月的步长（每年四个步长）以便于与年度数据进行比较，并认为三个月是最小的一个县可以切换类别的时间。最终的优化确定每个类别上的系数。

建立模型，第二部分：社会经济因素

文献

阿片类药物的流行因其在社会经济阶层中的范围而广为人知。种族是三个主要的风险因素[Berezow 2018; 药物滥用和精神卫生服务管理局 2018; Barbieri 2018]，教育[药物滥用和精神卫生服务管理局，2018；Scommegna 2018]和退伍军人身份[Recovery First 2019]。

其他风险因素包括性别（尽管近期有所增加，但妇女比男性更容易滥用阿片类药物）[国家药物滥用研究所，2018 年]，年龄（年轻人的风险更大）[国家药物滥用研究所，2018 年]和人口（阿片类药物滥用在农村社区比在城市环境中更为普遍）[Keyes 等。2014]。由于与慢性疼痛相关的阿片类药物处方的发生率较高，因此残疾可能是一个危险因素，但有关该主题的数据很少。

美国社区调查指标

为此问题提供的 2010 年至 2016 年美国社区调查（ACS）数据子集包括有关家庭规模和家庭结构，年龄和性别分布，受教育程度和受教育程度，退伍军人身份，残疾状况，居住流动性，出生地点的变量，在家里说的语言和祖先。它不包括我们的研究表明有用的某些统计数据，例如收入，失业率，医疗保健覆盖率和种族。

为了确保多年的一致性，我们会删除一年或一年以上缺少数据的变量。此过程删除了生育统计，残疾状况，公民身份，外国出生的世界出生地区以及几个其他家庭/家庭结构变量，共计 149 个社会经济因素中的 27 个。由于我们拥有的数据与我们想要的数据之间的差异，我们假设

- 受教育程度代表收入和医疗保健状况，
- 血统和语言表示种族，并且
- 退伍军人身份代表残疾。

由于调查数据中关于家庭规模和家庭结构的信息太多，因此我们将随后的研究重点放在家庭/家庭结构与阿片类药物成瘾之间的任何联系上。尽管该流行病影响着各种家庭，但我们注意到在阿片类药物流行的地区，祖父母抚养孙子的人数有所增加

受到最严重的打击[BAART Programs 2018]。这是由于成瘾导致父母自愿与非子女分离（例如死亡）而发生的。

具有社会经济地位的模型优化

为了将社会经济因素纳入我们的随机游走模型，我们首先在 sklearn 中运行随机森林算法[Pedregosa 等。[2011 年]根据 23 个社会经济预测因素将每个县划分为我们的六个类别之一。使用 sklearn 中的特征重要性属性，我们发现 10 个最重要的社会经济因素是非法药物的总使用率，总人口，在美国出生的人，在美国的血统，在爱尔兰的血统，只有在家讲英语的人，有大学学历但没有学位，高中毕业率，波兰血统以及具有研究生或专业学位的人。不幸的是，这些功能排名是基于绝对量级的，因此它们无法提供对阿片类药物滥用率影响方向的见解。

我们的适应模型使用此随机森林分类器生成的概率来初始化随机游走的概率向量。然后，算法将像以前一样进行。

进行此修改后，我们使用误差函数将新模型与旧模型的性能进行比较。一旦我

更多优秀论文（按模型或者时间分类），关注公众号“科研交流”，回复“优秀论文”
们优化了系数，这两个模型就变得更加准确，并且它们的性能也变得可比。



扫一扫上面的二维码图案，加我为朋友。

误差和灵敏度分析

仿真开始后，随机游走模型立即出现高误差。由于噪声，误差在 2013 年左右到峰值，然后开始急剧下降；在没有噪音的情况下，错误会继续向上传播。

我们的模型支持前面提到的聚类与后来添加的随机性（包括噪声）之间的健康关系。但是，达到这种平衡大约需要 24 个时间步（6 个“年”），这就是为什么在模拟开始时误差要比在结束时误差大的原因。

可能的原点位置

为了避免依赖于推断的社会经济因素而给我们的模型增加偏见，我们使用不依赖社会经济因素的旧模型进行了原产地识别分析。我们进行了蒙特卡洛模拟，以寻找可能的起源地，并在 2000 年第一波阿片类药物危机最严重的时期开始在每个县中进行流行。然后，模拟会逐步传播到 2010 年，然后我们将结果与给定数据进行比较。与 2010 年数据差异最小的模拟表明流行开始了。图 7 突出显示了可能已经开始流行的县（紫色），县首先在各自的州感染了该流行病（蓝色：宾夕法尼亚州，西弗吉尼亚州和弗吉尼亚州阴影最高的县，俄亥俄州的阴影最低的县；金：肯塔基州三个相邻阴影县中最左边的）。根据随机游走模型，危机最可能的起因是肯塔基州的蒙哥马利（黄金）。

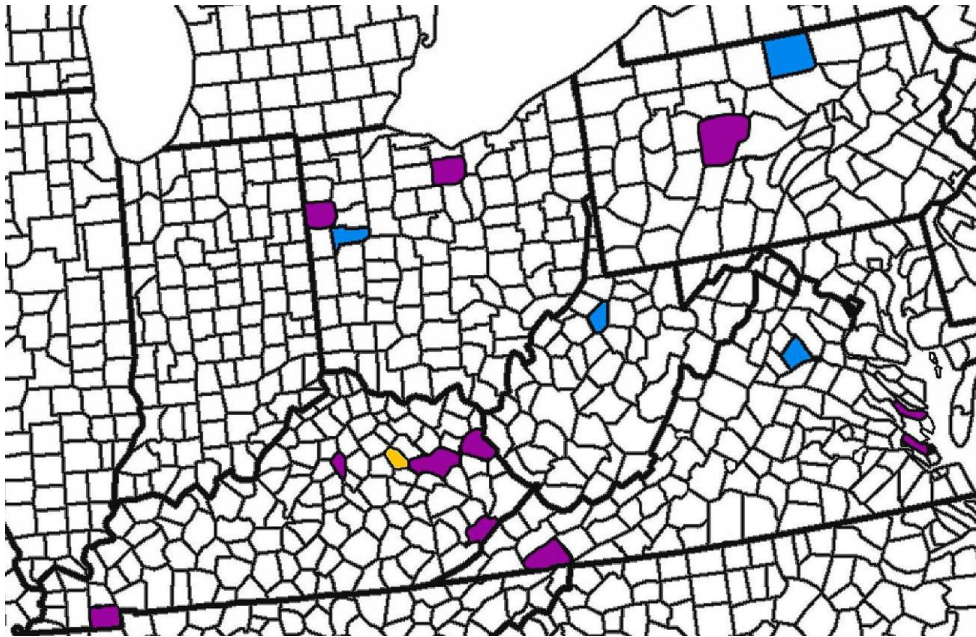


图 7. 流行病的可能起源位置。

根据该结果模型得出的流行病起源与西南维吉尼亚州相距不远，我们早先将其假设为起源。

第二种方法：空间回归

我们探索空间回归模型在 NFLIS 和 ACS 数据中的应用[Spielman 2015;Ver Hoef 等。2017]。考虑到空间关系非常重要，因为空间相关会极大地减少独立数据样本中包含的信息，通常减少 2 倍[Waller and Gotway 2004]。

在空间回归中，存在三种常见的模型类型：

- 空间自回归模型（SAR），也称为空间滞后模型，用于量化因变量 y 之间的空间相关性

相邻区域， y 的扩散或“溢出”效应[Sparks 2015]：



扫一扫上面的二维码图案，加我为朋友。

$$y = \rho W y + X \beta + u,$$

其中 W 表示空间权重（相邻性或区域之间的距离）， ρ 是空间自回归系数，且假定误差 u 是经典的（与 y 无关）[Viton 2010]。

• **空间误差模型（SEM）** 量化了残差的空间依赖性。他们将溢出误差概念化为一个或多个未识别的预测变量中的空间相关性，而不是溢出效应：

$$y = X \beta + u,$$

$$u = \lambda W u + v,$$

其中，除上面定义的变量外， λ 是空间自回归系数，且 $N = (0, \sigma^2)$ [Viton 2010]。

• **空间杜宾模型（SDM）** 是空间滞后模型和空间误差模型的组合 [Anselin 2003]。通过使用 W 滞后模型中的预测变量，除常规预测变量外，我们还可以获得空间预测变量的集合：

$$y = \rho W y + X \beta + W X \theta + \epsilon,$$

除上述定义的变量外， θ 是滞后预测变量的回归系数的向量， ϵ 是误差 [Sparks 2015]。也可以将滞后的预测变量包括在空间误差模型中，从而产生 Durbin 误差模型（DEM），但是该模型似乎不太常见，因此我们不进行探讨。

模型规格的问题取决于我们是否认为阿片类药物滥用是扩散引起的，还是仅受空间变化风险因素的影响。幸运的是，我们可以依靠统计检验以及我们的研究和直觉的结合来选择模型。

模型拟合

为了执行统计回归，我们使用 R 包 `spdep`。为了量化空间权重（模型中的 W ），我们使用了所有美国县的 shapefile [美国人口普查局 2017]，将其子集化为仅包括五个感兴趣的状态，然后将其转换为空间权重对象，如图所示。图 8。

我们使用方差膨胀因子（VIF）的方法识别并删除高度相关的变量。VIF 的公式如下：

$$VIF_k = \frac{1}{1 - R_k^2},$$

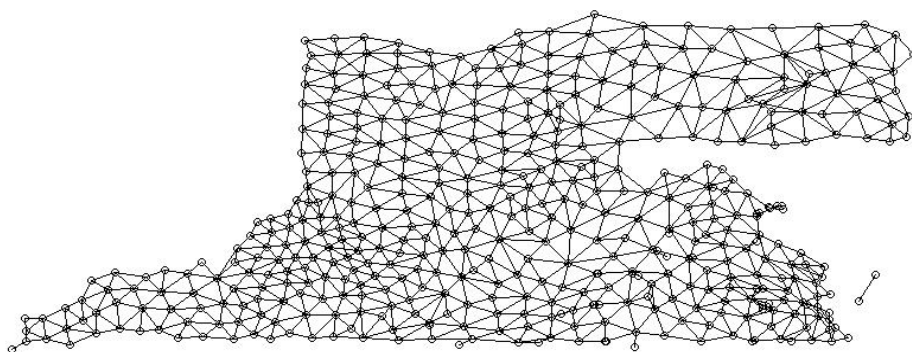


图 8. 县之间空间权重的可视化

其中 R_k^2 是通过对剩余预测变量上的第 k 个预测变量进行回归获得的 R^2 值 [Simon 等. 2018]。因此，如果一个预测变量与其他预测变量高度相关，则其 R^2 值将很大，VIF 表达的分母将变小，而 VIF 得分将变大。解决多重共线性问题的一种常见启



发式方法是删除 VIF 得分超过 10 的预测变量。在我们的数据中, VIF 得分超过 10 的变量组包括家庭/家庭结构的许多重叠指标, 包括教育入学率和受教育程度, 居民流动性, 居住地点出生和语言。删除该组变量后, 我们得到了 54 个预测因子, 主要来自 ACS, 但也包括总非法药物使用率 (NFLIS 的 TotalDrugReportsCounty 除以总人口), 以及我们的因变量: 非法阿片类药物的使用率 (NFLIS DrugReports 以总人口)。

我们使用类型为的函数 `lagsarlm`, `errorsarlm` 和 `lagsarlm` 对空间滞后, 空间误差和空间 Durbin 模型进行回归, 分别为“混合”。

统计结果

每个模型 (跨越七年) 都在数据中确认了高度重要的空间信号。似然比检验 (LRT) 是来自两个嵌套模型的似然函数的比率, 一个模型比另一个模型具有更多的参数 [White 2017]。在这种情况下, 轻铁会比较空间线性

回归与常规线性回归。空间自回归系数 γ 和 δ (分别针对 SAR 和 SEM) 的似然比检验和渐近 t 检验的所有 p 值均为 < 0.05 。这意味着空间

回归为我们提供的信息比常规回归要多得多。

关于空间模型规范, 最直接的比较可以基于最大化的对数似然性 [Anselin 2003]。多年来, 空间杜宾模型的对数似然率最高。这很有道理

因为我们期望基于人类的互动以及社会经济和/或监管模式, 阿片类药物的流行会同时显示空间滞后和空间误差模式。

每个 SAR 和 Durbin 模型还针对残差自相关执行 Lagrange 乘数 (LM) 测试。大约一半的 SAR 和 Durbin 模型显示出显著的残余自相关 ($p < 0.01$)。这表明可能还有其他空间自相关变量可以改善我们的模型, 和/或我们的误差项是异方差 (具有不一致的方差) [Spielman 2015]。为了测试后者, 我们对每个空间 Durbin 模型执行了 Breusch-Pagan (BP) 测试。所有 BP 测试均显示残差中的高度异方差。在空间模型中, 这通常是由于空间单元具有不同的人口规模 [Spielman 2015]。

与普通线性回归产生的系数不同, 空间滞后模型的系数不便于解释, 因为一个区域中一个预测变量的变化会影响其他区域的响应, 进而影响发生初始变化的区域中的响应。 [Sparks 2015]。为了说明直接 (局部) 影响和间接 (溢出) 影响, 我们在 `spdep` 中使用了影响函数来计算每个预测变量中单位增加的总体平均影响。我们的结果表明, 到目前为止, 一个县的总非法药物使用率是该县阿片类药物滥用水平的最强预测指标。

疾病控制中心报告称, 2014 年, 过量用药导致的死亡中有 61% 涉及某种类型的阿片类药物, 包括海洛因 [Rudd 等。2016]。到 2017 年, 这一比例为 67% [疾病控制与预防中心 2019]。因此, 我们的发现发现, 将非法药物总使用率提高一个单位会使阿片类药物的滥用率提高 52%, 这似乎是现实的。为了进行比较, 在相同预测变量集上进行的常规线性回归得出 $R^2 = 0.74$, 但总非法药物使用率的系数仅为 0.374。值得注意的是, 我们的平均直接影响率为 0.372。这一结果证实了我们的空间模型在预测阿片类药物滥用方面远远优于常规回归模型, 因为它考虑了空间扩散的间接影响。`sarlm` 模型的另一个输出结果就是 Akaike 信息准则 (AIC), 正式量化了所有模型的比较结果。

除了非法药物的使用率外, 其他变量均没有可比的系数大小, 它们似乎也无助于制定有助于解决阿片类药物危机的政策。

敏感性分析

为了测试分析对所选社会经济变量的敏感性, 我们使用了 23 个变量的较小子集, 这些变量强调教育程度和血统, 但还包括负责孙子孙女, 文职退伍军人, 去年



扫一扫上面的二维码图案，加我为朋友。

搬迁者，出生地，而且在家里只会说英语。结果与较大数据集的结果没有显著差异。空间杜宾

型号均表现最佳。非法药物总使用率单位增加对阿片类药物使用率的平均间接，间接和总影响为

分别为.377，.154 和.531，与较大模型的相应平均影响相比，每个值均大于.01。对于精简数据集，空间模型的共识是：

- 教育程度：拥有学士学位的百分比都与非法使用阿片类药物成正相关；高中毕业生百分比和某些大学但没有学位的百分比大多呈负相关；
- 家庭：负责孙子的祖父母百分比正相关；
- 语言：在家中只会说英语的比例大多呈正相关；和
- 祖先：荷兰人和爱尔兰人的百分比均呈正相关；挪威语，乌克兰语和匈牙利语的百分比大多呈负相关。

注意：“大多数”正面/负面表示在七年模型中该类别中至少有 5：2 的多数。

未来趋势和与第一个模型的比较

R 中的 `predict.sarlm` 函数使我们可以将趋势从初始结果应用于新数据。我们用它来预测空间滞后模型在未来五年内的非法阿片类药物的使用。因为我们试图对模型中相同的预测变量但对变量的观察结果不同（推断的社会经济因素）进行预测，所以我们使用趋势预测类型[Bivar and nd]。

在对阿片类药物滥用流行率做出预测后，我们通过六阶段分类器对这些数字进行运算，以使结果与我们的随机游走模型可比。图 9 说明了我们的第一个模型（随机游走）和第二个模型（空间滞后）的预测之间的偏差。

空间回归模型在很大程度上取决于社会经济因素的准确性。因为我们假设 SES 指标随时间呈线性变化，所以外推指标最接近初始数据是最准确的，这意味着空间回归模型在 2016 年之后的几年内最准确。如前所述，`randomwalk` 模型具有显著的意义。从初始点开始的几年中的错误。几年后，我们先前对社会经济指标的假设使空间回归越来越不准确；随机游走模型的准确性随着随机性和聚类之间的张力达到平衡而增加。

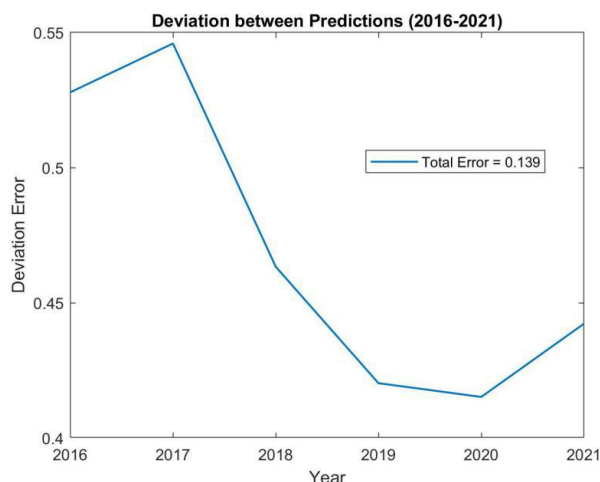


图 9. 随机游走和空间回归预测之间的误差。

药物鉴定阈值

如果一个县的稳定性较高，较高且不断增加，或者较低且持续时间超过 1.5 年（6 个时间步长），则将其分类为有问题的县。图 10 描述了到 2028 年有问题县的预计数量。



扫一扫上面的二维码图案，加我为朋友。

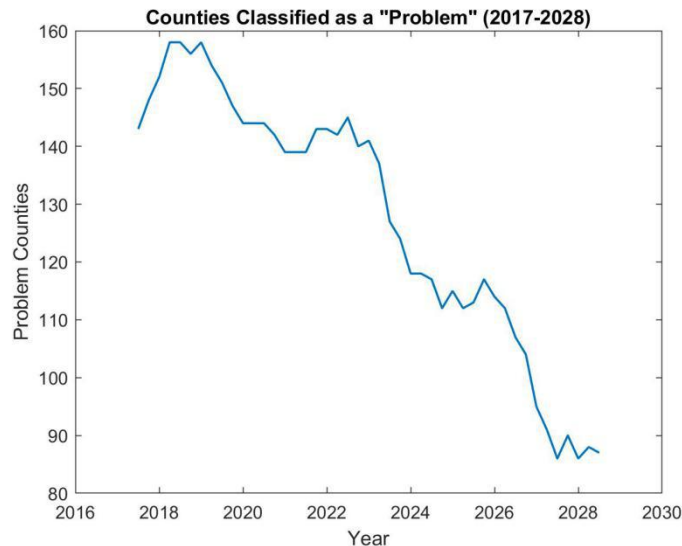


图 10. 被分类为有问题的县, 2017-2028 年。

我们的模型表明, 有问题的县的数量将在未来 10 年中稳定减少。但是, 这可能是由于溢出效应所致: 随着危机扩散到该国其他地区, 问题的严重性可能在我们的研究区域内减小。

对抗阿片类药物流行的策略

减少非法药物使用

我们的空间回归和随机森林模型的结果都表明, 在每个县中, 总的非法药物使用率是阿片类药物使用率的最重要预测指标。在所有其他空间常数模型中, 过去 7 年的空间回归模型的平均结果是, 毒品总使用量减少 0.1% 会使阿片类药物的总使用量减少 0.05%, 几乎是各县阿片类药物滥用率中位数的一半在这个地区。减少毒品的使用总量将减少阿片类药物的使用量, 因为通过毒品等级制进行的人数会减少。

解决非法药物使用的推荐方法包括投资于康复基础设施[Chandler 等人, 2003 年。2009 年], 向愿意与警察合作的社区伸出援手并为其提供资源[Moore 和 Kleiman 1989 年], 并改善青少年对非法物质成瘾性的教育[Gerstein and Green 1993 年]。我们关于祖先和教育的模型的结果表明, 针对所有种族和社会经济地位的人们开展教育计划是必要的。

直接阿片类药物干预

在 1970 年代初, 美国宣布了“禁毒战争”, 扩大了联邦禁毒机构的规模和存在, 同时停止了对这些药物的医疗功效的调查。自[Drug Policy Alliance 2018]以来的几十年里, 毒品滥用的犯罪率一直保持不变或上升。发动另一场“毒品战争”将无法解决阿片类药物的流行。相反, 我们必须将其视为公共卫生危机, 需要基于科学和卫生的解决方案, 而不是好斗的方法。朝这个方向迈出的第一步是对阿片类药物处方的限制; 此过程的下一步包括预防, 减少过量死亡和改善刑事司法系统内外的成瘾护理[国家成瘾和物质滥用中心 2017 年]。



扫一扫上面的二维码图案，加我为朋友。

- 促进全面的公众意识和教育运动。应该针对那些已经滥用阿片类药物的人于一系列运动，以传达紧急的安全问题，例如共用针头会带来血源性疾病的危险。对青少年的预防措施尤为重要，因为他们对成瘾的敏感性增加[哥伦比亚大学国家瘾与物质滥用中心 2011 年]。指导教育工作者以公共卫生科学为基础的有效预防和干预策略，而不是使学生远离资源的惩罚，这可能会在无意中加剧这种行为。使用学校网站向父母和孩子分发有关阿片类药物滥用的信息。

- 降低阿片类药物的可用性和可及性。各国应扩大和鼓励使用处方药监测计划（PDMP）来跟踪受控处方药的处方和分发，以发现可疑的滥用，购物或转移用途。应制定计划，以告知医疗专业人员有关疼痛管理的安全处方做法，并促进遵守 CDC 处方指南。在美国，每年约有 61 % 的处方药未消费 [Kennedy-Hendricks 等。2016]。药品回收计划将减少这些毒品上街销售 [国家成瘾和物质滥用中心 2017 年]。

- 减少过量服药死亡。将可拯救生命的阿片类药物过量的解毒剂（如纳洛酮）送入急救人员手中。给药已经挽救了成千上万人的生命 [Wheeler 等。2015]。此外，应尽快将有关危机的新信息分发给利益相关者，包括有关新兴合成材料以及如何治疗接触其的人的新闻。

- 改善刑事司法系统内外的成瘾治疗。增加治疗能力，扩大药物辅助治疗（MAT）计划的可用性以及扩大成瘾治疗计划的保险范围将帮助刑事司法系统之外的人们获得他们所需的克服成瘾的帮助。提供足够的康复支持系统将鼓励康复的成瘾者保持清洁。改善刑事司法系统的待遇将减少吸毒者的再犯，其中四分之三的人在释放后的五年内因另一种犯罪而被捕 [Durose 等人。2014]。鼓励各国对执法人员进行关于成瘾作为一种慢性健康状况的教育，并实施和支持诸如替代监禁（ATI）之类的转移方案，该方案使刑事司法系统（CJS）中的人员有更多的治疗选择。此外，各州应在重返平民生活时支持对 CJS 中的人的治疗选择。

错误的局限性和来源

总体

- 数据限制：此报告的局限性在于数据仅限于所提供的 ACS 和 NFLIS 数据。
- 可修改的面积单位问题：MAUP 描述了在以不同尺度进行的空间分析中的差异，因为一个人分析信息或分组方案的尺度可以产生不同的结果 [Emch 等。2017]。如果为我们提供了不同规模的数据（例如，在个人级别上），或者在人口普查区域或邮政编码区域内进行了汇总，则我们的调查可能会得出不同的结果。
- 邻接与距离：在我们的两个模型中，我们都基于县邻接而不是物理距离来合并空间权重。在模型中，尽管某些县的大小和形状差异很大，但似乎所有县的间隔均等。改变空间权重的类型可能会改变我们的结果 [Waller and Gotway 2004]。
- SES 变量的线性外推：预测 2017 年至 2021 年的阿片类药物滥用率需要对给定数据时间范围以外的社会经济指标进行外推。为简单起见，我们假设每个变量的线性趋势。实际上，这些趋势可能会更加多样化。

模型 I：随机漫步

- 马尔可夫性质：马尔可夫过程的定义特征是未来状态仅依赖于当前状态。我们名义上假设马尔可夫性质可模拟每个县的阿片类药物滥用率的变化。之所以采用“名义上”，是因为选择一个县的级别取决于该县及其邻居的当前级别，而该级别又取决于该县的过去状态。在这两种情况下，从长远来看，做出这种半马尔可夫假设都可能是不现实的，因为诸如政策实施之类的长期变化在很大程度上表明该地区



过去的药物滥用问题, 并将促使一个县变得更加一致。从长远来看, 滥用毒品行为率很高。这种假设可能会降低我们长期预测的有效性。

马尔可夫模型中变量的线性插值: 与上述外推类似, 对于每个“三个月”时间步长的社会经济变量线性插值是一种方便但可能不准确的捷径。但是, 内插法引入的误差比外插法引入的误差要小得多, 因为我们正在处理给定数据点之间的时间步长, 并且由于社会经济因素(例如具有一定家庭结构或血统的人的百分比)在一年的时间内不太可能发生巨大变化, 而在五年之内可能会发生显著变化。

- 确定潜在的原产地: 在 2000 年, 我们将所有县初始化为低稳定级。但是, 考虑到该模型固有的随机性, 低稳定县在一个时间步中跳到高增加, 高减少或高稳定类别的可能性很小。这有些不切实际, 并且可能会导致我们估算 2010 年空间分布时出现误差。

模型 II: 空间回归

- 预测: 在预测 2017 年至 2021 年的阿片类药物滥用率时, 我们采用了 2016 年的空间滞后模型。尽管使用最新模型似乎是最佳选择, 但鉴于我们每年必须分别进行空间回归, 因此 2016 年模型不一定代表 2010-2016 年。这意味着我们的预测几乎肯定会因 2016 年的社会经济因素数据而产生偏差。

- 空间模型检验: 空间自回归和空间误差回归是渐近的, 也就是说, 它们仅对大量区域给出近似有效的推论[Waller and Gotway 2004]。每年我们都有大约 460 个县的数据, 这似乎足够好。

结论

阿片类药物流行是史诗般的危机, 需要公众和政策制定者立即关注。我们使用随机游走模型和一系列空间回归模型来描述这场危机在美国阿巴拉契亚地区的蔓延特征, 该地区的流行最为普遍。我们的模型表现出色, 并且相辅相成。这些模型为与该地区阿片类药物流行和未来空间动态相关的社会经济变量提供了有用的见解, 使我们能够为公共政策干预提供明智的建议。

致首席行政官的备忘录

十月 3, 2019

发送至: Uttam Dhillon-毒品执法机构代理管理员抄送: Nora D. Volkow 博士-国家药物滥用研究所所长来自: MCM 小组 # 1901679

主题: 对抗阿片类药物流行的策略

阿片类药物的流行病在 2017 年[疾病控制与预防中心 2019 年]夺走了 4.7 万人的生命, 如果不解决, 在 2027 年之前还可以再夺走 50 万人的生命[Blau 2017]。我们的团队分析了给定的数据并构思了两个模型, 可以准确地表征和预测阿片类药物在肯塔基州, 俄亥俄州, 宾夕法尼亚州, 弗吉尼亚州和西弗吉尼亚州的流行情况。为了更好地了解这场危机以及我们有可能减少其影响的任何机会, 我们在模型中探索了当前的研究并调查了社会经济状况对阿片类药物滥用的影响。

结果

我们利用两种模型来更全面地了解阿片类药物的流行情况。两者都将成瘾定性为从最初来源向外传播的传染病。我们确定该地区的可能起源点是肯塔基州的蒙哥马利县, 这是一个没有太多治疗支持的农村地区。问题变得越来越严重, 蔓延到附近的县。我们预计该流行病将继续在全国范围内蔓延, 特别是影响那些已经大量使用非法毒品的县。减缓这一流行病的传播至关重要, 应立即加以解决。



扫一扫上面的二维码图案，加我为朋友。

提案

为了减轻阿片类药物流行的影响并减少未来几年过量用药的死亡人数，我们出以下政策：

1. 限制阿片类药物急性疼痛的处方，而不限那些有残疾或慢性严重疼痛的使用阿片类药物。对医疗专业人员进行安全的处方教育，并鼓励他们遵守 CDC 阿片类药物处方指南。

实施针对性的教育和公众意识运动。其中一组应针对最容易被加成的青少年，并通过学校系统和网站执行。另一个应解决那些已经滥用的阿片类药物，警告他们继续使用的潜在危险，并促进靠近他们的治疗计划和诊所。

2. 将阿片类药物过量的解毒剂纳洛酮分发给急救人员，并可能分发给上瘾者的家人和朋友，并告知他们如何使用它。

3. 扩大有成瘾问题者的保险范围和治疗能力。扩大利用美沙酮和丁丙诺啡等药物的药物辅助治疗计划的可用性，尤其是在农村地区，也将使处境艰难的人克服其成瘾性。

4. 在刑事司法系统中改善成瘾治疗。这将降低成瘾者的再犯率，成瘾者更容易被捕，其中四分之三将在其释放后的五年内因另一种犯罪而被捕[Durose 等人。2014]。对执法人员进行有关成瘾的慢性健康状况教育，并支持重返平民生活后继续接受治疗。

参考文献

省略