arXiv:2601.04966v1 [stat.AP] 8 Jan 2026

用于估计美国县级阿片类药物滥用患病率的贝叶斯多州数据整合方法

Zixuan Feng[1],Qiushi Chen*[1]、Paul Griffin[1], 和 Le Bao[2]

[1]宾夕法尼亚州立大学 Harold and Inge Marcus 工业与制造工程系[2]宾夕法尼亚州立大学统计系

摘要

药物过量死亡（包括阿片类药物所致死亡）仍是美国（US）面临的一项重大公共卫生威胁。为减轻阿片类药物滥用造成的危害，了解其在地方层面的患病率至关重要，这有助于社区利益相关方制定能够有效利用有限资源的应对策略。尽管已有若干州特定研究提供了县级患病率估计，但这类估计在全国范围内并不普遍可得，因为这些研究所用的数据集并非总能在其他州轻易获得，因此限制了现有模型的进一步应用。为弥补这一空白，我们提出一种贝叶斯多州数据整合方法，充分利用公开数据源，估计美国所有县的县级阿片类药物滥用患病率。该层级结构对阿片类药物滥用患病率和药物过量死亡结局进行联合建模，利用部分州已有的县级患病率估计及全国性调查得出的州级估计，并结合各县的协变量和混合效应来刻画县与州之间的异质性。此外，我们简约且具有普适性的建模框架采用马蹄先验，灵活地收缩系数并防止过拟合，从而确保在其他州新增县级患病率数据时仍能适用。利用真实世界数据，我们的模型通过交叉验证展现出较高的估计准确性，并首次提供了全国范围的县级阿片类药物滥用估计。

关键词： 贝叶　贝叶斯层级模型、小区域估计、阿片类药物流行、流行病学

*通讯作者。邮箱：q.chen@psu.edu

1 引言

过去二十年来，美国（US）的阿片类药物流行已成为最严重的公共卫生威胁之一（Salmond and Allread 2019、Hébert and Hill2024）。自1999年以来，这场流行已造成超过100万例药物过量死亡（Kennedy-Hendricksetal. 2024），并经历了多波演变：最初是处方阿片类药物，随后是海洛因、合成阿片类药物（Ciccarone2019），如今则是涉及阿片类药物和兴奋剂的多种物质使用，这被称为“第四波”（Ciccarone 2021）。除造成惨重的死亡负担外，阿片类药物流行还给医疗保健、刑事司法和生产力损失带来了沉重的经济负担（Luo etal. 2021、Florence et al.2021）；据估计，仅2020年，阿片类药物流行造成的经济损失总额就接近1.5万亿美元（Aboulenein2022）。尽管药物过量死亡人数近期有所下降，但药物过量造成的总体死亡负担依然惊人：仅2024年一年，就有超过8万人死于药物过量，其中约70%涉及阿片类药物（美国疾病控制与预防中心 2025年）。

为减轻阿片类药物和其他物质滥用造成的危害，了解社区层面的需求和负担，对于地方利益相关者利用有限资源，通过各类干预措施制定有效的应对策略至关重要（El-Bassel 等，2021，SAMHSA 2022年）。一项重要指标是患病率，即滥用阿片类药物且面临阿片类药物和其他药物过量死亡风险的人群比例。全国范围内与物质使用相关的患病率估计，通常来自全国药物使用与健康调查（NSDUH）。该调查是美国规模最大、内容最全面的年度普通人群调查之一，重点关注物质使用和心理健康（SAMHSA 2018年）。全国药物使用与健康调查通常会公布全国、州和州以下区域层面的阿片类药物及其他药物使用患病率估计（SAMHSA 2023年），但通常不提供县级估计（少数被视为独立州以下区域的大型县除外），因为全国调查采用多阶段抽样设计，每个县的样本量有限，难以覆盖相对罕见的事件（例如物质使用）（SAMHSA 2024年）。随着全国各地的县开始获得全国阿片类药物和解资金，以帮助减轻当地阿片类药物危机造成的危害（Rhubart 等，2022，Monnat 2018，Skinner 等，2024），人们比以往任何时候都更需要阿片类药物滥用的县级患病率估计，以帮助地方决策者了解社区中尚未满足的需求。

由于缺乏全国范围的县级阿片类药物滥用患病率估计，近期有多项研究开发了用于县级估计的州特定模型，这些模型需要州特定数据集。例如，Hepler 等 (2023利用贝叶斯分层时空模型，根据各县药物过量死亡人数和治疗入院人数，将全国药物使用与健康调查（NSDUH）的州级患病率估计重新分配，从而估计俄亥俄州2007—2019年的县级阿片类药物滥用情况

。Santaella-Tenorio 等 (2024)在此模型基础上进行扩展，除药物过量死亡和治疗数据外，还纳入阿片类药物过量急诊科（ED）就诊数据，以估计纽约州各县2007—2018年的阿片类药物滥用患病率。 Kline 等 (2025)根据北卡罗来纳州各县有关非法阿片类药物过量死亡和治疗数据的县级观测，采用了类似的丰度模型。此外，Doogan 等 (2022)根据个人层面的医疗补助行政理赔数据和生命统计数据，估计了俄亥俄州19个县医疗补助人群中的阿片类药物使用障碍（OUD）患病率。捕获-再捕获法也被用于估计马萨诸塞州（Barocas 等, 2018）和肯塔基州（Thompson 等， 2023）的县级阿片类药物使用障碍患病率。

然而，现有研究中用于估算阿片类药物滥用患病率的方法并不普遍适用于所有州。州特定统计模型依赖于临时性的措施，这些措施取自特定州的来源或项目（例如急诊就诊、理赔数据或物质使用治疗项目），因此难以在其他州复现该模型。捕获-再捕获分析需要个体级关联数据库；只有当某州已建立广泛的数据基础设施以支持该分析时，才具备可行性，而这类基础设施通常并不存在。乘数法对个体级关联数据集的要求虽较低，但通常用于调整更宏观层面的患病率估计（例如州或全国层面（Keyes 等， 2022， Mojtabai, 2022）），且难以捕捉更细粒度层面复杂的异质性。

为弥补现有文献的不足，我们提出一种估算县级阿片类药物滥用患病率的方法，该方法不受州特定数据来源的限制，适用于全美。我们提出一种贝叶斯多州数据整合方法，并遵循两项关键的数据原则。第一，我们纳入关于阿片类药物滥用患病率的最佳可用证据，涵盖多个层级，包括某些州基于流行病学研究（如捕获-再捕获研究）的县级估计，以及所有州的全国药物使用与健康调查州级估计，同时认识到全国药物使用与健康调查的估计可能存在低估（Reuter 等， 2021）。第二，我们的方法充分利用与阿片类药物滥用相关、且所有州均可按县级获取的公开数据源，包括美国疾病控制与预防中心（CDC）的流行病学研究在线数据（WONDER）中的药物过量死亡数据（美国疾病控制与预防中心, 2024a）、美国人口普查数据（美国人口普查局， 2024）以及县级健康排名（CountyHealth Rankings &Roadmaps 2024）中的县级人口统计和社会经济信息。我们提出的方法与现有州特定综合丰度模型（例如 Hepler 等

(2023年) 和 Kline 等 (2025年) 在估计问题的概念化方面有共同之处，即通过贝叶斯层级模型整合多个数据来源。我们的方法的主要区别在于，它整合了不同地理层级（即州级和县级）的患病率数据，这些数据来自多个来源；同时仅依赖公开数据，使我们能够推断全国各地的县级患病率，而不再局限于少数州。

据我们所知，这是首个联合估计美国各州县级阿片类药物滥用患病率的模型，填补了物质使用相关领域的一项关键知识空白。我们的贝叶斯多州数据整合方法的主要贡献总结如下。

- 数据整合：该方法无缝整合了多种公开数据源，包括全国药物使用与健康调查（NSDUH）的州级患病率数据、已发表的州特定研究中的县级患病率估计、药物过量死亡以及县级人口统计和社会经济特征，以推断县级阿片类药物滥用患病率。我们的模型专为捕捉现有数据源中的低估现象，并适应公共数据集中常见的数据抑制而设计。我们还明确构建了阿片类药物滥用与阿片类药物使用障碍（OUD）之间的关联，以协调定义上的差异，使本方法能够灵活整合不同数据源以不同方式报告的患病率估计。

- 层级结构：该方法同时对阿片类药物滥用患病率和阿片类药物过量死亡结局进行建模，并校正所有州中全国药物使用与健康调查（NSDUH）的低估。该方法利用县的协变量和混合效应，明确刻画不同县和州之间阿片类药物滥用患病率的异质性。

- 简约且具有普适性的模型框架：该框架可用于系统估计阿片类药物滥用患病率，并能轻松纳入额外的协变量数据。它采用马蹄形+ 先验——一种贝叶斯先验，可灵活收缩系数，从而防止过拟合。随着新证据的出现，例如未来新发表的研究可能提供其他州的县级患病率估计，该模型也可自然扩展。

本文其余部分安排如下。第2 节介绍广泛可用的阿片类药物及物质使用相关数据集，这些数据为我们提出的统计模型提供了依据，并指导了其结构。第3节介绍用于估计美国各州县级阿片类药物误用率的贝叶斯层级模型，并阐述我们的计算方法。第4 节展示基于真实世界

数据以及全部50个州的患病率估计结果。第5 节总结主要发现、实际意义及未来研究方向。

2 数据

本节介绍模型开发中使用的数据。这些数据来源在阿片类药物与物质使用研究中十分常见，并且所有州都可公开获取，这也促使我们开发贝叶斯层级模型。

县人口：这是阿片类药物滥用的基础人口数据，而阿片类药物滥用可能进一步导致阿片类药物过量死亡。为最大限度地覆盖各县，我们使用美国社区调查提供的各县总人口五年估算；单年度县人口估算仅适用于人口超过 65,000 的县（美国人口普查局，2024）。

阿片类药物相关过量死亡：每个县每年的阿片类药物相关过量死亡人数可通过 CDC WONDER 多重死因数据库查询（美国疾病控制与预防中心，2024a）。该数据库是阿片类药物与物质使用相关卫生政策研究中广泛使用的数据来源（Nam 等，2017、Buonora 等，2022、Feuerstein-Simon 等，2020、McClellan 等，2018）。尽管这些数据可通过在线查询公开获取，但为保护个人隐私，查询结果中所有低于 10 的数值都会被抑制。我们的初步分析发现，2010—2021 年间，全国约 45%—70% 的县存在被抑制的数值。由于数据抑制问题不容忽视，我们将在模型设定中明确加以考虑。

现有患病率估计：已发表的捕获-再捕获研究报告了2011-2015年马萨诸塞州（Barocas 等，2018）和2017-2018年肯塔基州（Thompson 等，2023）的阿片类药物使用障碍县级患病率估计。这些县级估计可作为参照点，使我们能够进一步建模其与其他县级和州级变量之间的关系。我们还纳入全国药物使用与健康调查的州级估计，以反映各州患病率的相对差异，以及阿片类药物滥用与阿片类药物使用障碍之间的联系。我们承认，使用全国药物使用与健康调查数据时，其估计存在低估（Reuter 等，2021，Keyes 等，2022），因此在模型中对这一数据局限性进行系统性调整。

阿片类药物配发率：各县每年每100名人口中配发的阿片类药物处方数量（CDC 2024b）反映了处方阿片类药物的可获得性。阿片类药物过度处方助长了阿片类药物流行的早期发展；既往研究表明，较高的配发率与误用和药物过量风险增加有关（Grigoras 等，2018，Sawyer 等，2021）。

县级社会人口因素：阿片类药物流行应放在

强调个体、社区与社会影响因素如何共同作用于健康结局的社会生态框架中加以理解（Jalali 等，2020）。以往研究考察了多种人口统计学、社会、经济及健康相关因素，将其作为县级预测因素，用于预测药物过量死亡率（Cano 等，2023），并将其用于患病率估计（Hepler 等，2023、Santaella-Tenorio 等，2024、Doogan 等，2022）。我们从县级健康排名（County Health Rankings）中获取县级社会人口因素。该综合数据集涵盖健康行为、临床护理、社会与经济因素以及物理环境等领域的 90 多项指标（CountyHealth Rankings & Roadmaps2024）。我们从该数据集中选取与物质使用广泛相关的教育、失业、风险行为及医疗服务可及性指标，作为县级协变量。案例研究中所纳入协变量的更详细说明见第 4 节。

3 方法

本节介绍用于利用多个数据来源的数据估计阿片类药物滥用县级患病率的多州贝叶斯层级模型（BHM）。下文将详细说明模型设定和计算方法。

3.1 贝叶斯层级模型

我们的贝叶斯层级模型由两部分组成：（1）针对直接观测到的阿片类药物滥用人群中阿片类药物过量死亡人数的模型；（2）针对现有阿片类药物使用相关患病率估计的模型，这些估计适用于县级（部分州）和州级（所有州）。整体模型结构如图所示1。

3.1.1 药物过量死亡模型

考虑一组县$\mathcal{I}$，其$\mathcal{I}$在 $i \in$ 的总人口为 N_i。令 p_i 表示县 i人口中阿片类药物滥用者所占的比例（即患病率），令 m_i表示县 i阿片类药物滥用人群中发生阿片类药物过量死亡的概率（即死亡率）。患病率和死亡率均以百分比表示，并采用逆逻辑变换参数化如下：

$$\text{logit}(p_i) = \beta_0^p + \boldsymbol{X}_i^p \boldsymbol{\beta}^p, \tag{1}$$

$$\text{logit}(m_i) = \beta_0^m + \boldsymbol{X}_i^m \boldsymbol{\beta}^m. \tag{2}$$

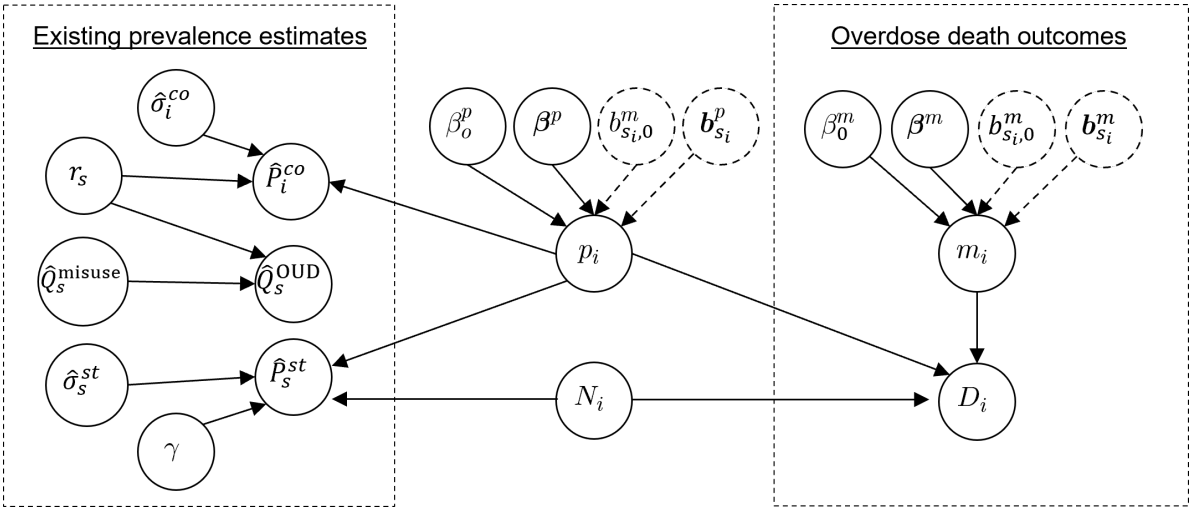


图 1：所提出的贝叶斯层级模型示意图。观测量用大写字母表示，模型参数用小写字母表示。虚线圆圈所示的潜在随机效应是否纳入模型，由残差分析决定。为使表述简洁，图中省略了协变量和模型先验。

其中，$\boldsymbol{X}_i^p = (X_{i,1}^p, X_{i,2}^p, \dots X_{i,P}^p)$是县 i的一组协变量，称为患病率决定因素；β_0^p 和 $\boldsymbol{\beta}^p$ 分别是患病率决定因素 $\boldsymbol{X}_i^p$ 的固定效应截距和斜率。同理，$\boldsymbol{X}_i^m = (X_{i,1}^m, X_{i,2}^m, \dots X_{i,M}^m)$、$\beta_0^m$和 $\boldsymbol{\beta}^m$ 分别表示死亡率决定因素、固定效应截距和死亡率决定因素的斜率。

我们将药物过量死亡人数建模为一个两阶段过程。首先定义辅助变量 O_i，表示县 i 中阿片类药物滥用人群的人数，其中 $O_i \mid N_i,\ p_i \sim$ 服从泊松$(p_i N_i)$分布。给定阿片类药物滥用人群的死亡率 m_i，药物过量死亡人数 $D_i \mid O_i,\ m_i \sim$ 服从二项$(O_i,\ m_i)$分布。泊松-二项模型等价于以分割概率 m_i 对泊松过程进行分割，由此得到药物过量死亡人数 D_i 的泊松分布，其期望值为 $m_i \cdot (p_i N_i)$。因此，我们用泊松分布对阿片类药物过量死亡人数 D_i 建模：

$$D_i \mid N_i, m_i, p_i \sim \text{Poisson}(m_i p_i N_i), \tag{3}$$

这些变量在各县之间条件独立。受抑制的药物过量死亡观测值被视为左删失值，在泊松模型下贡献 $P(D_i \le c)$。

3.1.2 纳入现有患病率估计

公开可获取的来源（如已发表文献和报告）提供了与阿片类药物使用和滥用相关的患病率估计。尽管这些估计仅涵盖某些县（例如来自特定州的捕获-再捕获研究），或仅以汇总层级提供（例如全国药物使用与健康调查发布的州级报告），但仍可能包含有价值的信息，可用于为所有县的患病率估计提供依据。值得注意的是，这些现有患病率估计的定义可能有所不同：有些关注阿片类药物滥用，另一些则仅限于符合特定诊断标准的一种临床疾病——阿片类药物使用障碍（OUD）。下文中，我们将多组现有患病率估计纳入模型，包括有限县的县级估计以及所有州的州级估计。

县级估计： 令 $\mathcal{I}_e \subset \mathcal{I}$ 表示已通过已发表的流行病学研究获得 OUD 人群县级患病率估计的县。由于本研究聚焦于阿片类药物滥用，其涵盖的概念范围比 OUD 更广，而 OUD 是根据更严格的临床标准定义的，因此我们将 OUD 人群建模为阿片类药物滥用人群的子集。具体而言，我们定义参数 r_s，表示州 $s \in \mathcal{S}$中阿片类药物滥用人口中符合 OUD 标准的比例。由于用于推断这些参数的 NSDUH 报告数据仅在州级层面提供，因此这些 OUD 与阿片类药物滥用之比 $\{r_s\}_{s\in\mathcal{S}}$ 因州而异。令 $\hat{} Q_s^{\text{OUD}}$和 $\hat{} Q_s^{\text{misuse}}$分别表示 NSDUH 州级报告中 OUD 人群和阿片类药物滥用人群的州级人数，其服从以下二项数据模型：

$$\hat{Q}_s^{\text{OUD}} \mid \hat{Q}_s^{\text{misuse}}, r_s \sim \text{Binomial}\left(\hat{Q}_s^{\text{misuse}}, r_s\right). \tag{4}$$

在正确定义比率 r_s 后，我们现在可以将$\hat{}$模型中的阿片类药物滥用患病率 p_i 与现有研究在县级估计的阿片类药物使用障碍患病率 P_i^{co}联系起来。$\hat{}$考虑到患病率的非负性，我们假设县 $i \in \mathcal{I}_e$（州 $s_i \in \mathcal{S}$）的现有估计值 P_i^{co}服从标准差为$\hat{}\sigma_i^{co}$的对数正态分布：

$$\hat{P}_i^{co} \mid p_i, r_{s_i}, \hat{\sigma}_i^{co} \sim \text{Lognormal}(u_i, v_i^2), \tag{5}$$

其中，分布参数 u_i 和 v_i 定义为

$$u_i = \log\left(p_i r_{s_i}\right) - v_i^2/2, \tag{6a}$$

$$v_i^2 = \log\left(1 + (\hat{\sigma}_i^{co})^2/(p_i r_{s_i})^2\right), \tag{6b}$$

这些参数通过矩估计法推导得出，即将对数正态分布的均值设为 $p_i r_{s_i}$ ，并将标准差设为$\hat{}$ σ_i^{co}。标准差$\hat{}\sigma_i^{co}$根据报告的$\hat{}$ P_i^{co}的95%置信区间（CI）计算得出，假设 $2 \times 1.96\hat{}\sigma_i^{co}$等于该置信区间的宽度。

州级估计： 将$\hat{}$ P_s^{st}记为根据NSDUH州级报告得出的州s $\in$的阿片类药物滥用州级估计。令$\mathcal{I}_s$ 表示州 s内的县。与县级估计的建模方式类似，我们假设州级估计量$\hat{}$ P_s^{st}服从对数正态分布，

$$\hat{P}_s^{st} \mid \{p_i\}_{i\in\mathcal{I}_s}, \{N_i\}_{i\in\mathcal{I}_s}, \gamma, \hat{\sigma}_s^{st} \sim \text{Lognormal}(\mu_s, \nu_s^2), \tag{7}$$

其中， γ 是一个乘数，用于调整NSDUH估计值的低估程度（$\gamma \leq 1$）；位置和尺度参数 μ_s 和 ν_s 通过使对数正态分布的均值与 $\bar{p}_s := \gamma \frac{\sum_{i\in_{\mathcal{I}}} {}_s p_i N_{i_{\mathcal{I}}}}{\sum_{i\in} {}_s N_i}$相匹配，并使标准差与$\hat{}\sigma_s^{st}$相匹配来确定，其中

$$\mu_s = \log(\gamma\bar{p}_s) - \nu_s^2/2, \tag{8a}$$
$$\nu_s^2 = \log\left(1 + (\hat{\sigma}_s^{st})^2/(\gamma\bar{p}_s)^2\right). \tag{8b}$$

与$\hat{}\sigma_i^{co}$类似，标准差$\hat{}\sigma_s^{st}$ 也根据NSDUH报告的州级患病率估计$\hat{}$ P_s^{st}的95%置信区间计算得出。所有分布假设均经实证验证，详细结果将在第3.4节介绍。

州级异质性：文献中普遍指出，全美各地社区受阿片类药物危机影响的程度并不相同（Monnat2018）。上述定义的基线模型（1)-（2）通过县的协变量捕捉县与县之间患病率和药物过量死亡率的差异。然而，由于各州的卫生、社会经济和地理因素（例如医疗服务可及性、失业和乡村性）可能存在显著差异，这些模型或许不足以解释州与州之间的差异（Cano 等，2023）。即使纳洛酮分发水平和治疗可及性相当，不同州的阿片类药物过量死亡降幅也可能不同（Chhatwal等，2023）。因此，我们建议在方程（1）和（2）中纳入州级随机效应，以捕捉基线模型无法解释的州级异质性。

我们通过残差分析进行一系列模型诊断，以确定需要添加哪些随机效应，同时避免可能的过拟合。首先，我们考察不含随机效应的基线模型对各县药物过量死亡人数的估计所对应的皮尔逊残差，发现超过半数州的平均残差显著不为零（附录图 A.1）。这清楚表明，州际

异质性尚未得到充分捕捉，因此有必要加入州级随机效应。考虑到与现有患病率估计相比，所有州的县级药物过量死亡信息都相对完整，我们首先在死亡率的截距（2）中加入随机效应，以捕捉其在各州之间的异质性。此举显著改善了模型拟合：42 个州中仅有 2 个州的州平均残差 95%置信区间未涵盖零（附录图 A.2）。在模型中为患病率（1）加入随机截距后，仅有一个州的平均残差显著异于零（附录图A.3）。随后，我们按州分别将残差对各县级协变量进行回归，以判断残差与各个协变量之间的相关性是否在各州间存在显著差异，从而确定是否有必要加入随机斜率。结果显示，超过 90% 的州中相关性不显著（附录表A.1），表明无需加入随机斜率。根据这些模型诊断结果，我们将方程（1）和（2）修改如下：

$$\text{logit}(p_i) = \beta_0^p + b_{s_i,0}^p + \boldsymbol{X}_i^p \boldsymbol{\beta}^p, \qquad b_{s,0}^p \sim \mathcal{N}\left(0, (\sigma_0^p)^2\right), \tag{9}$$

$$\text{logit}(m_i) = \beta_0^m + b_{s_i,0}^m + \boldsymbol{X}_i^m \boldsymbol{\beta}^m, \qquad b_{s,0}^m \sim \mathcal{N}\left(0, (\sigma_0^m)^2\right), \tag{10}$$

其中，标准差 σ_0^m和 σ_0^p 用于量化州际异质性，是需要估计的模型参数。下文的分析均采用基于方程（3）–（10）的最终模型设定。

3.2 似然与先验设定

给定模型参数$\boldsymbol{\theta} = (\beta_0^p$、 $\boldsymbol{\beta}^p$、 $b_{s,0}^p$、 σ_0^p、 β_0^m、 $\boldsymbol{\beta}^m$、 $b_{s,0}^m$、 σ_0^m、 γ、 $r_s)$ 时，我们的贝叶斯层级模型 $\mathscr{L}(\boldsymbol{\theta})$ 的联合似然由三个部分组成，分别对应县级药物过量死亡人数、现有的县级和州级患病率估计。在计算药物过量死亡的似然时，我们注意到，出于隐私保护目的，公开报告的药物过量死亡数据中的一些观测值 D_i 通常会被抑制，即实际计数小于或等于给定阈值 c （例如，对于通过 CDC WONDER 数据库公开查询获得的药物过量死亡数据， $c = 9$ ）。为体现这一点，我们定义指示变量 $\delta_i = 1$ ：如果县 i 中实际的药物过量死亡人数被抑制，则取相应值；否则取 $\delta_i = 0$。于是，县 i 的药物过量死亡数据的似然为：

$$\ell_i^D(\boldsymbol{\theta}) = \left[f_{pois}(D_i | m_i(\boldsymbol{\theta}) p_i(\boldsymbol{\theta}) N_i)\right]^{1-\delta_i} \left[F_{pois}(c | m_i(\boldsymbol{\theta}) p_i(\boldsymbol{\theta}) N_i)\right]^{\delta_i}, \tag{11}$$

其中，$f_{pois}(\cdot|m_i(\boldsymbol{\theta})p_i(\boldsymbol{\theta})N_i)\big)$ 和 $F_{pois}(\cdot|m_i(\boldsymbol{\theta})p_i(\boldsymbol{\theta})N_i)\big)$ 分别是期望值为 $m_i(\boldsymbol{\theta})p_i(\boldsymbol{\theta})N_i$ 的泊松分布的概率质量函数和累积分布函数。

现有患病率估计数据的似然函数$\{\hat{}P_i^{co}\}_{i\in\mathcal{I}_e}$在县级和$\{\hat{}P_s^{st}, \hat{}Q_s^{\mathrm{OUD}}\}_s$在州级的形式可直接由模型（4）、（5）和（7）得出，其中

$$\ell_i^{co}(\boldsymbol{\theta}) = f_{lognormal}\left(\hat{P}_i^{co} \mid u_i(\boldsymbol{\theta}), v_i^2(\boldsymbol{\theta})\right), \qquad \text{for } i \in \mathcal{I}_e, \quad (12)$$

$$\ell_s^{st}(\boldsymbol{\theta}) = f_{lognormal}\left(\hat{P}_s^{st} \mid \mu_s(\boldsymbol{\theta}), \nu_s^2(\boldsymbol{\theta})\right) \times f_{binomial}\left(\hat{Q}_s^{\mathrm{OUD}} \mid \hat{Q}_s^{\mathrm{misuse}}, r_s(\boldsymbol{\theta})\right), \quad \text{for } s \in \mathcal{S}, \quad (13)$$

其中，$f_{lognormal}(\cdot \mid u, v^2)$ 是对数正态分布的概率密度函数，其位置参数为 u，尺度参数为 v，定义于对数尺度上（定义见方程（6）和（8））；而 $f_{binomial}(\cdot \mid n, p)$ 是二项分布的概率质量函数，其 n 试验次数为 p，成功概率为。故完整似然为：

$$\mathscr{L}(\boldsymbol{\theta}) = \left(\prod_{i\in\mathcal{I}} \ell_i^D(\boldsymbol{\theta})\right) \times \left(\prod_{i\in\mathcal{I}_e} \ell_i^{co}(\boldsymbol{\theta})\right) \times \left(\prod_{s\in\mathcal{S}} \ell_s^{st}(\boldsymbol{\theta})\right). \quad (14)$$

我们为固定效应截距 $\beta_0^p, \beta_0^m \sim \mathcal{N}(0, 10^2)$以及标量和标准差 $\gamma, r_s, \sigma_0^p, \sigma_0^m \sim$采用半正态分布$(\sigma = 10)$作为弱信息先验。对于固定效应斜率 $\boldsymbol{\beta}^p$ 和 $\boldsymbol{\beta}^m$，我们采用 horseshoe+ 先验，以实现协变量的自动收缩和选择（Bhadra 等人，2017）。对于 β_j 中的每个元素$(\boldsymbol{\beta}^p, \boldsymbol{\beta}^m)$，其 horseshoe+ 先验由以下层级结构定义：

$$\begin{aligned}
\beta_j \mid \lambda_j &\sim \mathcal{N}(0, \lambda_j^2), \\
\lambda_j \mid \tau, \zeta_j &\sim \text{Half-Cauchy}(0, \tau\zeta_j), \\
\zeta_j &\sim \text{Half-Cauchy}(0, 1), \\
\tau &\sim \text{Half-Cauchy}(0, 1),
\end{aligned}$$

其中，λ_j 是各个组成部分的局部收缩参数 β_j，而 τ 则是多个 β_j的全局收缩参数。我们分别为 $\boldsymbol{\beta}^p$ 和 $\boldsymbol{\beta}^m$定义独立的全局收缩参数。作为一种诱导稀疏先验，这种层级结构能够使模型自适应地将较小的系数收缩至零附近，同时保留影响显著的较大系数，从而以数据驱动的方式进行变量选择。

为考察先验对模型表现的影响，我们通过一系列敏感性分析评估采用更强先验的模型。具体而言，我们（1）将

固定效应截距 β_0^p， β_0^m改为 $N(0, 5^2)$ 和 $N(0, 2.5_2)$；（2）将随机效应标准差的先验 σ_0^p，σ_0^m改为半正态分布(σ=5)和半正态分布(σ=2.5)；（3）用尺度为 0.5、正则化作用更强的半正态先验，替代马蹄先验+中全局和局部收缩参数所采用的重尾半柯西先验。

3.3 后验预测分布

我们使用 Stan 编程语言（Gelman 等人 2015）和 `rstan` R 中的接口（Stan 开发团队 2024）实现贝叶斯层级模型。我们使用哈密顿蒙特卡罗方法（Betancourt 和 Girolami 2015、Hoffman 和 Gelman 2011）结合无转向采样器（NUTS）生成后验样本。无转向采样器是一种高效算法，适用于高维且复杂的后验分布（Betancourt 2018）。以下各节中的所有计算研究均在配备 M2 芯片处理器和 16 GB 内存的 MacBook Air 笔记本电脑上完成。我们基于^潜在尺度缩减因子 $R < 1.1$ （Gelman 和 Rubin 1992）评估模型的收敛性，并通过目视检查轨迹图和自相关函数，确认并行链具有良好的混合性并实现收敛。我们使用 NUTS 运行 4 条并行链进行模型拟合。对于每条链，我们从 50,000 次迭代中收集样本，其中前 25,000 次迭代作为预热期，并采用 10 的抽稀参数，最终从 4 条链中获得总计 10,000 个后验样本，用于估计模型参数的后验分布。

为获得^ P_i^{co}和^ P_s^{st}的后验预测分布，在给定（$p_i^{(k)}$ 、$\gamma^{(k)}$、$r_{s_i}^{(k)}$）的每个后验样本的条件下，我们分别从分布（5）和（7）中抽取{ ^P^{co}、$_i$ $^{(k)}$ }和{ ^P^{st}、$_s$ $^{(k)}$}的样本。类似地，在给定 N_i 以及（$p_i^{(k)}$ 、$m_i^{(k)}$ ）的后验样本的条件下，根据（3）从泊松分布中抽取药物过量死亡数的后验预测样本{$D_i^{(k)}$ }。后验样本的95%预测区间根据其第2.5和第97.5百分位数定义。

3.4 模型诊断和验证

我们首先检查残差，以检验分布假设的兼容性。具体而言，我们分别基于参数后验均值，计算以下各项观测值与模型预测值之间的残差：参数（1）中的（1）药物过量死亡（未抑制）、（2）县级患病率估计和（3）州级患病率估计。鉴于对药物过量死亡数作出了泊松分布假设，我们将观测值与期望值之差除以期望值方差的平方根，从而得到药物过量死亡（未抑制）的皮尔逊残差。

后验预测检验将州级汇总的药物过量死亡总数的后验分布

与各州观测到的药物过量死亡数进行比较。我们重新查询 CDC WONDER 多重死因数据库，以获取州级阿片类药物过量死亡人数，该数据不存在受抑制数据。对于数值被抑制的县，我们计算后验预测分布D_i 低于抑制阈值 c的概率。

此外，我们采用 10 折交叉验证（CV）评估模型对未见过的数据进行预测的能力。对于 $\mathcal{I}_e$ 中已有阿片类药物使用障碍患病率 ˆ 估计值的县，我们将这些县随机分为 10 折，每次留出一折中的 P_i^{co}数据来拟合模型，然后将后验样本的均值和预测区间与留出观测值进行比较。对于药物过量结局，我们对所有 $\mathcal{I}$ 中的县进行分层 10 折交叉验证，并将 $\mathcal{I}_e$ 中的县随机分配到各折。我们每次留出一折中的药物过量死亡数据来拟合模型，并根据后验样本（$p_i^{(k)}, m_i^{(k)}$）获得预期药物过量死亡人数 $m_i^{(k)} p_i^{(k)} N_i$ 的后验样本。我们评估平均绝对百分比误差（MAPE）、重叠区间估计的比例，以及我们的区间估计涵盖患病率观测值 $\{\hat{}P_i^{co}\}_{i \in \mathcal{I}_e}$和未被抑制的药物过量死亡人数 $\{D_i : \delta_i = 0, i \in I$的比例。对于药物过量死亡人数被抑制的县，我们报告预测区间与受抑制数据范围重叠的比例（即 ≤ 9）。

为了进一步探究模型在不依赖州特定研究数据的情况下的可推广性，我们还通过按州留出观测到的阿片类药物使用障碍ˆ患病率数据 P_i^{co}来对模型进行交叉验证。由于我们的真实世界研究中纳入的县级阿片类药物使用障碍患病率数据仅来自两个州（有关数据的详细描述，请参见第4.1 节），ˆ留出一个州的数据后，模型只能根据另一个州观测到的患病率 P_i^{co}进行拟合。在探索性分析中，我们发现，鉴于当前模型包含多个随机效应，这种数据稀疏会导致完整模型（3）-（10）不可辨识；因此，我们移除患病率中的随机截距（9），以恢复模型的可辨识性，并在该留一州分析中使用修改后的模型。

4 结果

4.1 真实世界数据设置

在本节中，我们将所提出的贝叶斯层级模型应用于真实世界数据，以估计美国全部50个州（不包括华盛顿特区）的阿片类药物滥用县级患病率。为与一项关键研究保持一致，我们以2015年为参考年份，该研究

使用捕获-再捕获分析估计马萨诸塞州县级阿片类药物使用障碍患病率（Barocas 等，2018），同时利用年份最接近的其他数据来源，以确保时间上的接近。对于现有的县级估计值$\{\hat{P}_i^{co}\}_{i\in\mathcal{I}_e}$，除2015年马萨诸塞州研究中的估计值外（Barocas 等，2018），我们还纳入了2017年肯塔基州的估计值（Thompson 等，2023）。尽管相差两年，纳入多个州的县级患病率数据对于模型正确识别用于捕捉州间异质性的随机效应至关重要。为考虑年份差异，我们将肯塔基州报告的县级患病率估计值的标准差设为未知参数σ_{KY}^{co}，由数据推断，并用其替代根据报告的区间估计直接计算出的$\hat{\sigma}_i^{co}$。

对于州级患病率估计$\{\hat{P}_s^{st}\}_{s\in\mathcal{S}}$，我们结合全国药物使用与健康调查 2015年–2016年州报告中的州级止痛药使用障碍和海洛因使用估计值，以代表 2015年阿片类药物滥用患病率。此外，全国药物使用与健康调查 2021年–2022年州报告首次提供了州级 OUD 和阿片类药物滥用估计值，使我们能够利用这些数据按州推断模型参数r_s。

从 CDC WONDER 数据库查询了 2015年所有县的阿片类药物过量死亡人数$\{D_i\}$。我们采用国际疾病分类第十版（ICD-10）根本死因（UCD）代码，涵盖意外（X40–X44）、自杀（X60–X64）、他杀（X85）或未确定（Y10–Y14），以及多重死因（MCD）代码，涵盖海洛因（T40.1）、天然及半合成阿片类药物（T40.2）、美沙酮（T40.3）、美沙酮以外的合成阿片类药物（T40.4），以及其他和未明确的麻醉品（T40.6）；这些代码遵循文献中常用的筛选标准（Feuerstein-Simon 等，2020，Seth 等，2018）。

鉴于这些协变量具有广泛的相关性，并可能影响阿片类药物和物质使用，我们将以下协变量纳入考虑，作为患病率和死亡率的潜在决定因素：(1) 年度阿片类药物配发率（每100名人口），用于捕捉阿片类药物流行的供给侧效应（Vuolo 和 Kelly，2022、Thombs 等，2020）；(2) 成人吸烟率，这是一项与其他健康相关风险行为密切相关的风险因素，也可作为更广泛的物质使用行为的代理指标（Zale 等，2015、Young-Wolff 等，2017）；(3) 部分大学教育比例（25岁及以上成年人中至少完成一年大学教育者所占的比例），因为受教育程度可能影响风险行为以及获得医疗资源的机会（Powell，2023、Thombs 等，2020）；(4) 失业率，因为缺乏就业和经济压力可能导致物质滥用增加以及更高的药物过量风险（Rudolph 等，2020、Saba 等，2025）；(5) 每10万名人口中的初级保健医生数量，用于衡量获得医疗保健和满足基本健康需求的机会（Haffajee 等，2019、Eichmeyer 和 Zhang，2023）。阿片类药物配发

2015年的配发率数据根据零售处方配药数据库由美国疾病控制与预防中心获取（CDC 2024b），其他县级社会人口因素协变量则取自2015年的县级健康排名（County HealthRankings&Roadmaps 2024）。我们的完整数据集包含来自美国50个州的3,132个县。仅有 <0.5%的县的协变量存在缺失值，这些缺失值采用州均值进行插补。模型拟合前，所有变量均进行了标准化。

4.2 模型拟合优度

我们的模型诊断验证了所提出的层级模型对给定数据具有合理的拟合优度（见附录图A.4）。残差似乎以零为中心，与拟合值之间没有明显相关性；模型预测的期望值也与三种观测结果的观测值相吻合，这表明模型中的分布假设合理。

我们的后验预测检验进一步表明，在大多数州，州级药物过量死亡的观测人数落在基于后验预测分布确定的高概率区域内（见附录图A.5），这也说明所提出的模型假设合理且拟合恰当。我们还发现，在一些州，报告的药物过量死亡人数位于预测分布的较低端；这些州往往患病率较高，或拥有患病率估计较高的县。与模型预测样本相比，报告的药物过量死亡人数存在这种低估现象，可能是由于这些州中数据受抑制的县占比较高。由于模型参数的推断更多依赖未受抑制的死亡数据（≥ 10），而非受抑制的数据，因此将这些拟合后的参数应用于受抑制的县时，后验分布往往会偏向抑制阈值，导致预测值高于报告的观测值。

为进一步验证我们的模型对县级患病率的估计，我们通过 10 折交叉验证，将结果与马萨诸塞州（Barocas 等，2018）和肯塔基州（Thompson 等，2023）现有的估计进行比较。总体而言，在这两个州合计的 134 个县中，我们的模型平均绝对百分比误差为 45.2%，斯皮尔曼等级相关系数为 0.60。结果显示，我们对马萨诸塞州各县的预测准确度更高，平均绝对百分比误差为 12.8%，相关系数为 0.78（图 2）；相比之下，肯塔基州的平均绝对百分比误差为 49.0%，相关系数为 0.53（附录图 A.6）。此外，我们模型的 95 %预测区间涵盖了这两个州中 95.5%（134 个县中的 128 个县）的观测估计值，这支持了这些估计在实践中的有效性和相关性。

来自我们的模型。

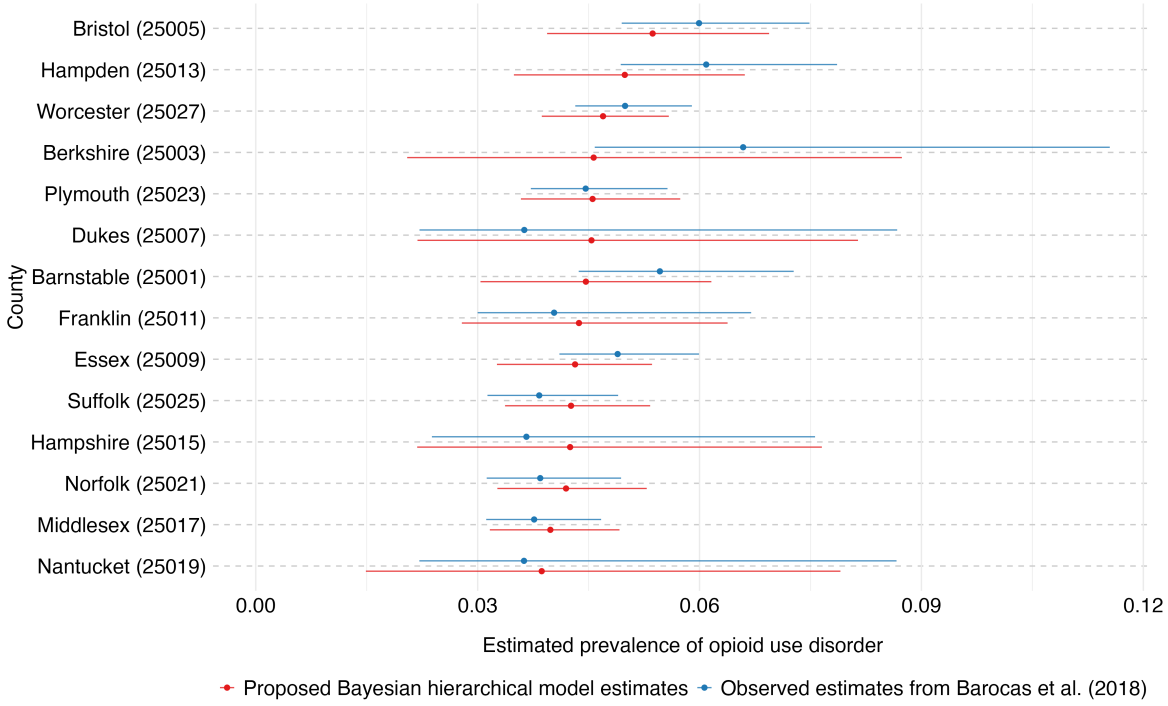


图 2：贝叶斯层级模型的 10 折交叉验证结果与Barocas 等人（2018年）的估计值进行比较。

我们还针对药物过量死亡结果进行了 10 折交叉验证，并分别分析未受抑制和受抑制药物过量死亡数据对应的县。在死亡人数未受抑制的县中，我们模型的 95%预测区间涵盖了 63.3%（615 个县中的 389 个县）的药物过量死亡观测人数，点估计的平均绝对百分比误差为 37.6%。对于数据受抑制的县（即观测死亡人数 ≤ 9），我们的模型估计的药物过量死亡人数在 90.8%（2,517 个县中的 2,286 个县）低于抑制阈值 $c = 9$，且这些县中药物过量死亡人数的后验预测分布低于阈值 c 的概率平均为 0.91，在 79.1%（2,517 个县中的 1,992 个县）中为 ≥ 0.95。

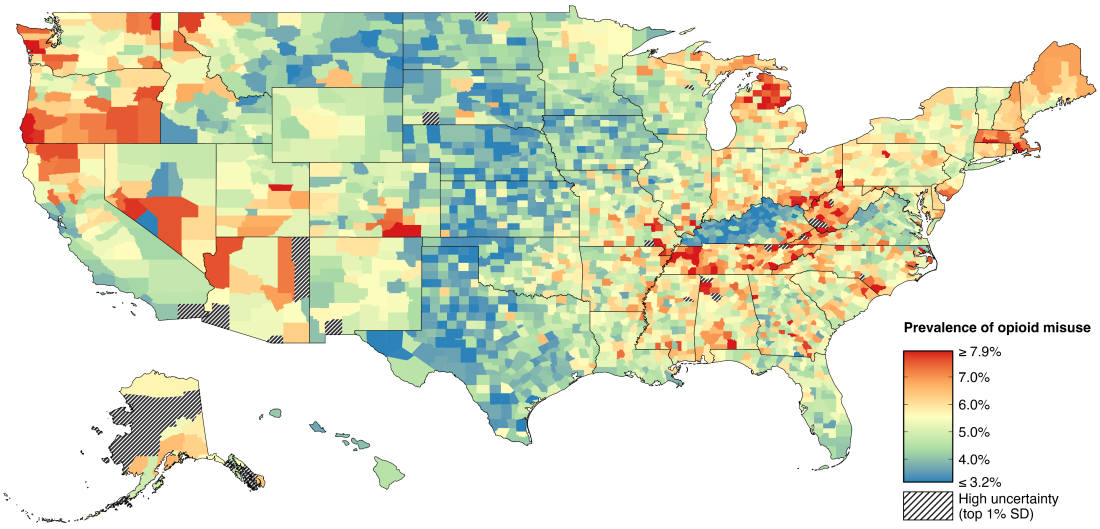
4.3 县级患病率估计

根据美国50个州的3,132个县的真实世界数据，我们的模型估计总体阿片类药物滥用患病率（$\sum_i p_i N_i / \sum_i N_i$）为5.1%（95%可信区间：3.7%，6.8%），相当于2015年美国11岁及以上人口中有1,410万（1,022万，1,866万）人滥用阿片类药物。在县级层面，阿片类药物滥用患病率估计值（p_i）的中位数为5.0%，第一和第三四分位数分别为4.3%和5.9%。

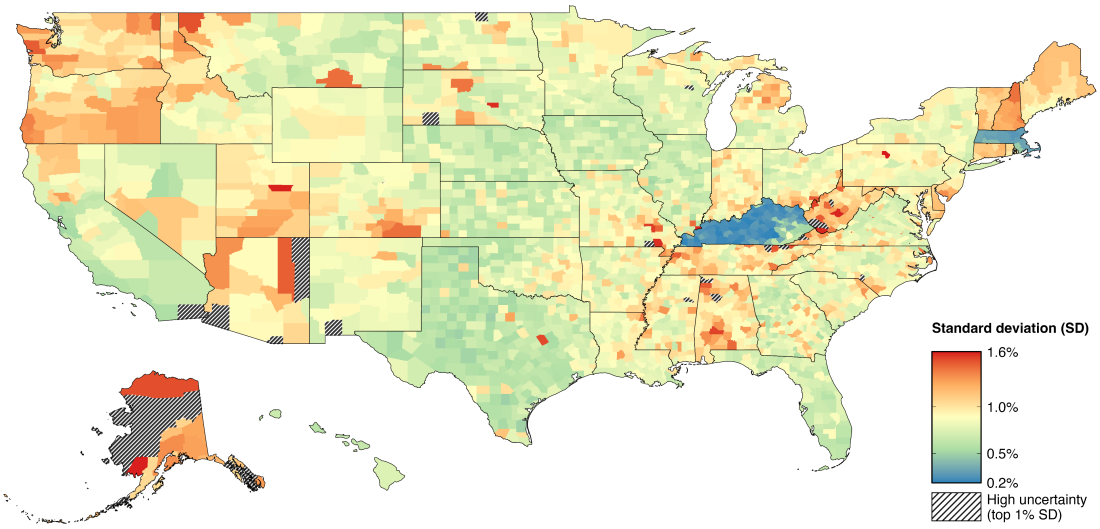
有1,672个县（53.4%）的估计患病率低于全国平均水平。所有县患病率估计值的直方图和分布见附录图 A.7。

图 3 展示了所有县的阿片类药物滥用患病率估计的后验预测均值（图 3a）和标准差（图3b）。鉴于不确定性较高，地图中未显示标准差最高的1%的县。正如预期，马萨诸塞州和肯塔基州各县的估计具有最低的不确定性（如图 3b中后验标准差较低的蓝色区域所示），因为这两个州提供了县级阿片类药物使用障碍患病率数据，这些数据有效地将后验估计集中在数据附近，从而带来了较高的置信度。患病率较高的县集中在阿巴拉契亚地区，包括肯塔基州东北部和俄亥俄州、西弗吉尼亚州的大部分地区以及田纳西州，同时也分布在新英格兰和太平洋西部地区；与肯塔基州和马萨诸塞州相比，这些地区估计值所对应的不确定性明显更高。由于这些地区的高患病率估计主要由县级协变量驱动，在缺乏直接县级患病率数据的情况下，模型还会将协变量系数的方差传播为患病率估计的高不确定性。州级患病率估计见附录图 A.8。县级患病率估计的完整结果见补充文件。

模型参数的估计值见表 1。对于患病率，固定效应截距（β_0^p ）估计为 -2.90（95% 可信区间：-3.09，-2.70），对应的全国基线患病率为 5.2%（4.4%，6.3%）。在县级协变量中，阿片类药物配发率与患病率呈显著正相关（with $\hat{}$ $\beta_1^p = 0.16$ [0.09, 0.22]），这与文献中有关阿片类药物配发在阿片类药物流行中作用的研究结果一致（Cano 等，2023；Monnat 等，2019）。失业率也与患病率呈正相关（with $\hat{}$ β_4^p =0.13 [0.05, 0.20]），这表明经济不稳定可能与阿片类药物滥用患病率升高相关（Azagba 等，2021）。其他协变量未显示显著效应；由于 horseshoe+ 先验的作用，其系数被收缩至接近 0（见第 3.2 节）。对于死亡率，成人吸烟率这一协变量与县级药物过量死亡率呈显著正相关（with $\hat{}$ $\beta_2^m = 0.15$ [0.07, 0.23]），这与以往研究一致：研究表明，吸烟率与阿片类药物滥用呈正相关（Rajabi 等，2019），且会增加药物过量死亡（Langabeer 等，2022），这可能是通过与其他社会和行为因素的复杂相互作用所致。我们观察到，在控制阿片类药物滥用后，失业率与死亡率呈负相关，这意味着它与药物过量死亡总人数的总体关联



（a）2015年美国所有县中，11 岁及以上人口的阿片类药物滥用患病率估计值。



(b) 美国所有县阿片类药物滥用患病率估计值的标准差 .

图 3：美国县级阿片类药物滥用患病率估计结果 .

死亡人数可能并不明确，这与文献中的发现一致：这些关联在不同地区和亚群体之间存在差异（Kerry等，2019年，Hollingsworth等，2017年，Frankenfeld和Leslie，2019年）。

重要的是，患病率和死亡率中的随机效应均如预期成功捕捉了各州之间的异质性，随机效应截距的标准差估计值显著不为零 $\sigma_0^m = 0.48$ （0.38，0.61），以及 $\sigma_0^p = 0.21$（0.14，0.28）。此外，我们的模型通过估计值 $\gamma = 0.21$（0.17，0.24）校正了全国药物使用与健康调查估计值的低估，这意味着平均而言，应用于全国药物使用与健康调查估计值的乘数约为4.9。由于模型中包含州级随机效应，我们的估计值相对于全国药物使用与健康调查的州级患病率估计值的实际相对尺度因州而异，介于0.15至0.36之间（附录图A.9）。这些尺度对应的乘数范围为2.8至6.7，与采用全国药物使用与健康调查作为基准数据来源的乘数法研究所报告的2.9至8.0的乘数范围相近（Mojtabai，2022，Lim等人，2024，Keyes 等，2022）。

表 1：模型参数的后验均值和 95% 可信区间。

参数	后验 mean	95% 可信区间
患病率		
固定效应截距（β_0^p）	-2.90	(-3.09, -2.70)*
阿片类药物配发系数（β_1^p）	0.16	(0.09, 0.22)*
成年人吸烟系数（β_2^p）	0.08	(-0.00, 0.16)
部分大学教育系数（β_3^p）	0.05	(-0.02, 0.12)
失业系数（β_4^p）	0.13	(0.05, 0.20)*
初级保健医生系数（β_5^p）	0.01	(-0.03, 0.04)
随机效应截距的标准差（σ_0^p）	0.21	（0.14，0.28）*
死亡率		
固定效应截距（β_0^m）	-6.01	（-6.23，-5.78）*
阿片类药物配发的系数（β_1^m）	0.06	(-0.00, 0.13)
成年人吸烟系数 (β_2^m)	0.15	(0.07, 0.23)*
部分大学教育系数 (β_3^m)	0.04	(-0.02, 0.12)
失业系数 (β_4^m)	-0.16	(-0.23, -0.07)*
初级保健医生系数 (β_5^m)	0.00	(-0.03, 0.03)
随机效应截距的标准差 (σ_0^m)	0.48	(0.38, 0.61)*
现有患病率估计		
NSDUH估计值的缩放系数（γ）	0.21	(0.17, 0.24)*
马萨诸塞州阿片类药物滥用者中 OUD 的比例（r_{MA}）	0.64	(0.63, 0.64)*
肯塔基州阿片类药物滥用者中 OUD 的比例（r_{KY}）	0.96	(0.95, 0.96)*
KY 估计值的标准差（σ_{KY}）	0.02	（0.01，0.02）*

* 95%可信区间不包含零，表明具有统计显著性。

最后，为进一步验证研究结果的实际意义，我们将其与其他州特定研究中县级阿片类药物滥用患病率的类似估计进行交叉比较。一项适合比较的研究采用贝叶斯层次模型，估计了2016年至2021年北卡罗来纳州的阿片类药物滥用患病率

(Kline 等， 2025). 尽管研究设计和背景有所不同，但我们发现，我们估计的患病率往往高于 Kline 的研究中2016年的估计值。不过，总体趋势一致，两组估计值之间的相关性显著，为 0.50（95% 可信区间：0.34-0.64）。在100个县中，我们的95%可信区间覆盖了其中61个县的点估计，并与其中97个县的区间估计重叠。我们还与一项俄亥俄州研究进行了比较，该研究估计了2019年医疗补助人群中的阿片类药物使用障碍患病率（Doogan 等, 2022）。虽然我们的估计值较低，这可能是因为医疗补助人群的阿片类药物使用障碍患病率往往高于一般人群，但两组估计值之间的斯皮尔曼等级相关仍高达0.66（0.29-0.86），表明两组估计值的相对大小一致。其他几项针对县级阿片类药物滥用或使用障碍患病率估计的州特定研究（例如纽约州（Santaella-Tenorio 等， 2024）和俄亥俄州（Hepler 等, 2023）的研究）未公布县级估计值的具体数值，因此未纳入比较。

4.4 敏感性分析

为进一步评估模型对没有县级患病率数据的州的预测能力，我们通过每次留出一个州（即马萨诸塞州或肯塔基州）的阿片类药物使用障碍患病率估计，进行留一州交叉验证。模型能够恢复留出州各县患病率估计的排序，表明其对没有县级患病率估计的州也具备预测能力。具体而言，基于 KY 的模型（留出马萨诸塞州各县的阿片类药物使用障碍患病率数据）估计的阿片类药物使用障碍患病率与马萨诸塞州的实际患病率数据呈显著正相关（斯皮尔曼相关系数 $\rho = 0.75$ 和皮尔逊相关系数 $r = 0.75$，p 值 < 0.05）；基于 MA 的模型对肯塔基州患病率的估计也呈现类似结果（$\rho = 0.47$ 和 $r = 0.61$，p 值 < 0.05）。另一方面，由于这些基于单州数据的模型没有用于捕捉州际异质性的州级随机效应，因此这些模型给出的患病率绝对值可推广性有限。例如，基于 MA 的模型估计全国患病率为 8.25%（95% 可信区间：6.98%，9.69%），显著高于完整模型估计的 5.1%（3.7%，6.8%）；而基于 KY 的模型给出的估计值则低得多，为 2.78%（95% 可信区间：2.32%，3.31%）。同样，在县级层面，与在两者之间取得平衡的完整模型相比，基于 MA 的模型始终高估患病率，而基于 KY 的模型则始终低估患病率（附录图 A.10）。这些差异也凸显了在估计阿片类药物滥用患病率时捕捉跨州差异的重要性，以及将县级阿片类药物使用障碍

患病率数据作为参考纳入多个州数据的必要性；这一做法已在我们的完整模型中实现。

我们对先验进行的敏感性分析表明，贝叶斯模型的结果对先验设定具有很强的稳健性。采用不同的先验设定时，计算时间仍大致相同，每˜5000 次迭代约需 1200–1400 秒，使用 `Stan` 软件。所有模型均已收敛，最大 R 为 1.08。如附录图A.11所示，与基准情形相比，即使采用更严格的先验设定，所有模型参数的估计仍保持一致，这证实推断结果主要由数据中的信息（所有协变量均已标准化）驱动，而非由先验的尺度决定。

5 讨论

本研究构建了一个多州贝叶斯层级模型，整合公开数据，以估计美国各县的阿片类药物滥用患病率。交叉验证表明，与现有患病率估计和观察到的药物过量死亡相比，我们提出的模型具有较高的预测准确性。本研究提出了一个统一框架，可无缝整合目前所有县均可获得的数据，无需依赖州特有的临时性 数据来源，从而为阿片类药物及物质滥用的县级患病率估计研究作出了贡献。

尽管有几个州利用州特定数据来源估算了阿片类药物滥用或阿片类药物使用障碍的县级患病率，但并非所有州都拥有类似的综合数据资源来支持此类分析；因此，这些州特定模型及其相应结果对其他州的借鉴意义有限。相比之下，我们的模型旨在广泛适用于所有州，且仅需使用公开且易于获取的信息，从而减轻数据基础设施薄弱的州获取和准备数据的负担。据我们所知，我们是首个对美国全部50个州的县级阿片类药物滥用情况进行联合建模的研究。我们的研究结果是迈向更深入了解阿片类药物危机的地方性负担，并为制定有效的资源分配策略以缓解这一危机提供参考的第一步。

近期一项由 Peterson 等人 （2025年）开展的研究也采用贝叶斯建模框架，旨在估计全美各县的阿片类药物使用障碍患病率，但我们的模型存在几项重要区别。Peterson 等人的时空模型采用自上而下的建模策略：首先利用全国药物使用与健康调查得出州级患病率估计，再将各州总量分配到每个县，并将药物过量死亡数据作为协变量之一。相比之下，我们的数据方法将全国药物使用与健康调查提供的州级患病率数据（涵盖所有州）、流行病学研究提供的县级患病率估计（针对

有限的几个州）以及药物过量死亡数据整合到统一的似然函数中。此外，我们将药物过量死亡建模为阿片类药物滥用的观测结果，从而能够以有原则的方式处理被抑制的计数，并纳入药物过量死亡数据中的不确定性。这两项研究对目标人群的定义也存在一个细微但重要的差异。 Peterson 等人 （2025年）聚焦于阿片类药物使用障碍人群，同时处理全国药物使用与健康调查中诊断标准随时间发生的变化；而我们的研究则估计面临药物过量风险的更广泛阿片类药物滥用人群的患病率。我们的模型还明确构建了阿片类药物使用障碍与阿片类药物滥用这两个概念之间的联系，因此能够灵活整合基于多个研究中不同定义得出的患病率估计。

在本研究中，我们利用全国性数据源提供的公开药物过量死亡数据，使模型能够广泛适用于全国各地；但需要指出的是，此类数据源存在内在异质性，这也构成了本研究的一项局限。尽管我们的模型仅纳入州级随机效应，以在较高层面上捕捉异质性，但药物过量死亡数据会受到当地非法毒品市场中毒品供应毒性的影响（例如芬太尼的渗透程度），而这种毒性在不同地区之间乃至同一州内都可能存在显著差异（Ciccarone 2019）。由于用于应对药物过量的可用资源存在差异，城乡差异也可能导致药物过量死亡率结果出现县级异质性。此外，药物过量死亡中所记录毒品类型的准确性取决于法医办公室的报告惯例，而这些惯例在不同州和县之间可能存在显著差异。既往研究也表明，生命统计数据中的阿片类药物过量死亡存在严重漏报和错误分类（Ruhm 2016）。受数据局限所限，我们的模型无法充分捕捉药物过量死亡数据中所有可能的异质性来源，这或许导致了后验预测检验中观察到的部分州的差异。尽管如此，交叉验证结果表明，模型平均而言具有合理的预测准确性，显示了模型能够从全国范围的公开数据中提取有用信息，以解决患病率估计问题。

本研究的另一项局限在于其横断面设计，以及所采用的数据来源并非取自完全相同的年份。尽管我们基于一项估计县级阿片类药物使用障碍患病率的开创性研究（Barocas 等, 2018）将2015年作为分析的参考年份，但受数据可得性限制，我们选择的其他数据来源均尽可能接近该年份。鉴于2017年肯塔基州的县级患病率估计被用作模型输入，以辅助推断另一个年份的情况，我们假定这些估计具有更高的不确定性，将其标准差作为待推断的模型参数，而不是根据报告值直接固定。为确定阿片类药物使用障碍与阿片类药物滥用之间的州特异性相对比率， r_s我们采用了$\hat{} \ Q_s^{\text{misuse}}$和$\hat{} \ Q_s^{\text{OUD}}$

这些估计值仅在2021—2022年NSDUH州级报告或更晚的报告中可获得，这隐含地假设了这些比率在时间上的稳定性。对时间动态进行明确建模，将为整合来自不同时间点的数据来源提供一种原则性方法。随着更多研究的开展，新获得的数据更有可能具有更好的时间一致性，这也可能减轻当前模型中由时间差异引起的局限。

我们提出的公共数据整合方法不应取代各州特定研究的工作；这类研究利用州内可获取的临时数据集，重点估计阿片类药物滥用的县级患病率。事实上，未来若能利用各州特定研究发表的更多估计结果，我们的方法还可以进一步完善。在此案例研究中，我们采用了马萨诸塞州和肯塔基州现有的阿片类药物使用障碍患病率县级估计值；这些估计值是基于各州内多个关联数据集，通过捕获-再捕获分析系统估算得出的。如果其他州也能采用类似设计开展捕获-再捕获分析，其估计结果便可直接纳入我们当前的模型。此外，随着相关估计结果不断见诸文献，我们的方法还可扩展，以纳入通过其他统计方法（例如综合丰度模型、乘数法）得出的州特定估计值，以及针对不同目标人群（例如阿片类药物滥用与阿片类药物使用障碍）的估计值。比如，可以对模型进行调整，分别构建与各州特定研究所用估计方法相匹配的模型，再通过集成框架综合其推断结果。整合现有但分散在各项州特定研究中的估计值，不仅有助于我们对不同方法得出的估计值进行交叉验证，还能最大限度地发挥这些估计值的价值，从而增进对其他州和县阿片类药物使用与滥用情况的整体认识，并据此加以完善。

我们的研究为未来研究提供了多个方向。对当前模型一个颇有前景的拓展，是纳入县与县之间的空间相关性，以考虑地理和社会邻近性，使模型能够借鉴其他县的信息，并有望减少各县估计结果之间的差异。尽管空间模型功能强大，但可能面临计算方面的挑战，尤其是在全国县级尺度上，因此有必要进一步开发和完善计算方法。另一种拓展当前模型、且无需进行重大结构调整的方向，是将其他相关的社会、人口统计学和新兴流行病学因素作为协变量纳入模型；随后可通过特征选择方法对这些因素进行系统筛选，从而更好地捕捉县与州之间的异质性。还可以进一步完善模型，调整毒理学报告中药物特异性数据缺失所导致的阿片类药物涉入药物过量死亡情况的低估（Gutkind等，2025年）。虽然药物分类信息的完整性已有所提升

近年（Boslett等，2019年），鉴于当前阿片类药物危机阶段的多种物质使用现象以及新兴药物威胁，死亡记录中的药物特异性变得日益重要。最后，政策制定者或许也有兴趣进一步评估不同数据来源的贡献，以及获取额外数据的潜在价值；例如，可开展信息价值分析（Keisler等，2014年、Parsons和Bao，2022年），以帮助政策制定者确定优先领域，开展额外的数据收集工作，从而改进对全国阿片类药物滥用患病率的估计。

致谢

本研究获美国国家科学基金会（NSF）资助，资助协议编号为 CMMI-2240408；并获美国国立卫生研究院（NIH）/美国国家过敏和传染病研究所（NIAID）资助，项目编号为 R01AI136664 和 R01AI170249。本材料中所表达的任何意见、研究结果、结论或建议均代表作者个人观点，不一定反映 NSF 或 NIH 的立场。

参考文献

Aboulenein, A.（2022年）。2020年阿片类药物危机给美国造成近1.5万亿美元损失——国会报告。https://www.reuters.com/world/us/opioid-crisis-cost-us-nearly-15-trillion-2020-congressional-report-2022-09-28/

访问日期：2024年8月1日。Azagba, S., Shan, L., Qeadan, F., and Wolfson, M. (2021). 失业率、阿片类药物滥用及其他物质滥用：来自治疗入院数据的准实验性证据。《*BMC*精神病学》，21(1):22. Barocas, J. A., White, L. F., Wang, J., Walley, A. Y., LaRochelle, M. R., Bernson, D., Land, T., Morgan, J. R., Samet, J. H., and Linas, B. P. (2018). 马萨诸塞州2011年至2015年阿片类药物使用障碍的患病率估计：一项捕获-再捕获分析。《美国公共卫生杂志》, 108(12):1675–1681. Betancourt, M. (2018). 哈密顿蒙特卡洛方法概念导论。arXiv预印本 arXiv:1701.02434. Betancourt, M. and Girolami, M. (2015). 用于层级模型的哈密顿蒙特卡洛方法。载于贝叶斯方法论的当前趋势及其应用，第79–101页。Chapman and Hall/CRC. Bhadra, A., Datta, J., Polson, N. G., and Willard, B. (2017). 超稀疏信号的马蹄估计量+。Boslett, A. J., Denham, A., Hill, E. L., and Adams, M. C. B. (2019). 未分类药物过量

死亡与阿片类药物危机：不平等的新兴模式。美国医学信息学协会杂志, 26(8-9):767–777. Buonora, M. J., Hanna, D. B., Zhang, C., Bachhuber, M. A., Moir, L. H., Salvi, P. S., Cunningham, C. O., and Starrels, J. L. (2022). 处方阿片类药物危机高峰期美国各州的阿片类药物处方政策：与阿片类药物过量死亡率的关联。国际毒品政策杂志, 110(103888):103888. Cano, M., Oh, S., Osborn, P., Olowolaju, S. A., Sanchez, A., Kim, Y., and Moreno, A. C. (2023). 美国药物过量死亡率的县级预测因素：系统综述。《药物与酒精依赖》, 242(109714):109714. 美国疾病控制与预防中心（2024a）。CDC流行病学研究广泛在线数据（WONDER）多死因数据。美国疾病控制与预防中心（2024b）。美国药物配发率地图。访问日期：2022年10月1日。美国疾病控制与预防中心（2025）。2024年美国药物过量死亡人数下降近27%。 https://www.cdc.gov/nchs/pressroom/nchs_press_releases/2025/20250514.htm。美国疾病控制与预防中心、国家卫生统计中心（NCHS）。Chhatwal, J., Mueller, P. P., Chen, Q., Kulkarni, N., Adee, M., Zarkin, G., LaRochelle, M. R., Knudsen, A. B., and Barbosa, C. (2023). 在美国4个州持续实施公共卫生干预措施可减少的阿片类药物过量死亡估计数。《美国医学会杂志·网络开放》, 6(6):e2314925. Ciccarone, D. (2019). 三波流行病：美国阿片类药物过量危机的供需驱动因素。国际毒品政策杂志, 71:183–188. Ciccarone, D. (2021). 非法芬太尼类药物、兴奋剂的兴起与阿片类药物过量危机的第四波。《精神病学当前观点》, 34(4):344–350. 县级健康排名与路线图（2024）。2024年县级健康排名与路线图技术文件。 https://www.countyhealthrankings.org/sites/default/files/media/document/2024%20CHRR%20Technical%20Document_1.pdf。访问日期：2022年10月1日。Doogan, N. J., Mack, A., Wang, J., Crane, D., Jackson, R., Applegate, M., Villani, J., Chandler, R., and Barocas, J. A. (2022). 俄亥俄州医疗补助人群中的阿片类药物使用障碍：利用乘数法对19个县进行的患病率估计。《美国流行病学杂志》, 191(12):2098–2108. Eichmeyer, S. and Zhang, J. (2023). 初级保健提供者对阿片类药物使用及其不良后果的影响。《公共经济学杂志》, 217(104784):104784. El-Bassel, N., Gilbert, L., Hunt, T., Wu, E., Oga, E. A., Mukherjee, T. I., Campbell, A. N. C., Sabounchi, N., Gutnick, D., Kerner, R., Venner, K. L., Lounsbury, D., Huang, T. T. K., and Rapkin, B. (2021). 通过社区参与实施阿片类药物使用障碍循证实践：数据驱动范式与系统科学方法。《药物与酒精依赖》, 222(108675):108675. Feuerstein-Simon, R., Lowenstein, M., Sharma, M., Dupuis, R., Luna Marti, X., and Cannuscio, C. C. (2020). 地方卫生部门与旨在应对阿片类药物过量死亡率的循证政策实施。《物质滥用》, 41(4):468–474. Florence, C., Luo, F., and Rice, K. (2021). 2017年美国阿片类药物使用障碍和致命性阿片类药物过量的经济负担。《药物与酒精依赖》,218(108350):108350.

Frankenfeld, C. L. 和 Leslie, T. F. (2019). 2013—2017年美国药物过量死亡与县级社会经济因素以及居住地种族、西班牙裔、贫困和失业隔离的关联. 《年鉴：流行病学》, 35:12–19. Gelman, A., Lee, D. 和 Guo, J. (2015). Stan：一种用于贝叶斯推断和优化的概率编程语言. 《教育与行为统计学杂志》, 40(5):530–543. Gelman, A. 和 Rubin, D. B. (1992). 使用多个序列进行迭代模拟推断. 《统计科学》, 7(4):457–472. Grigoras, C. A., Karanika, S., Velmahos, E., Alevizakos, M., Flokas, M.-E., Kaspiris-Rousellis, C., Evaggelidis, I.-N., Artelaris, P., Siettos, C. I. 和 Mylonakis, E. (2018). 阿片类药物死亡与处方药及社会决定因素的相关性：一项针对医疗保险参保者的横断面研究. 《药物》, 78(1):111–121. Gutkind, S., Marziali, M. E., Bruzelius, E., Mannes, Z. L., Martins, S. S., Hasin, D. S. 和 Mauro, P. M. (2025). 美国药物相关过量死亡中阿片类药物涉入的误分类：范围综述. 《年鉴：流行病学》, 102:8–22. Haffajee, R. L., Lin, L. A., Bohnert, A. S. B. 和 Goldstick, J. E. (2019). 阿片类药物过量死亡率高且提供阿片类药物使用障碍治疗药物的能力有限的美国县的特征. 《美国医学会杂志·网络开放》, 2(6):e196373. Hébert, A. H. 和 Hill, A. L. (2024). 按人口群体和兴奋剂共同涉入情况分析阿片类药物过量对美国预期寿命和寿命损失年数的影响：基于2019年至2022年死亡率数据的分析. 《柳叶刀区域健康—美洲》, 36. Hepler, S. A., Kline, D. M., Bonny, A., McKnight, E. 和 Waller, L. A. (2023). 用于估计俄亥俄州县级阿片类药物滥用患病率的综合丰度模型. 《英国皇家统计学会杂志：*A*辑》, 186(1):43–60. Hoffman, M. D. 和 Gelman, A. (2011). 无转向采样器：在哈密顿蒙特卡洛中自适应设定路径长度. arXiv 预印本 arXiv:1111.4246. Hollingsworth, A., Ruhm, C. J. 和 Simon, K. (2017). 宏观经济状况与阿片类药物滥用.《健康经济学杂志》, 56:222–233. Jalali, M. S., Botticelli, M., Hwang, R. C., Koh, H. K. 和 McHugh, R. K. (2020). 阿片类药物危机：情境化的社会生态框架. 《卫生研究、政策与系统》, 18(1):87. Keisler, J. M., Collier, Z. A., Chu, E., Sinatra, N. 和 Linkov, I. (2014). 信息价值分析：应用现状. 《环境系统与决策》, 34(1):3–23. Kennedy-Hendricks, A., Ettman, C. K., Gollust, S. E., Bandara, S. N., Abdalla, S. M., Castrucci, B. C. 和 Galea, S. (2024). 美国成年人因药物过量而经历的个人丧失.《美国医学会健康论坛》, 5(5):e241262. Kerry, R., Yoo, E. 和 Ingram, B. (2019). 美国西部药物中毒死亡的空间分析：采用轮廓回归调整共线性和空间相关性的比较研究. 《药物与酒精依赖》, 204(107598):107598. Keyes, K. M., Rutherford, C., Hamilton, A., Barocas, J. A., Gelberg, K. H., Mueller, P. P., Feaster, D. J., El-Bassel, N. 和 Cerdá, M. (2022). 2010年至2019年美国阿片类药物使用障碍的患病率和趋势如何？采用乘数法估计患病率

以估计未知人口规模中的患病率. 《药物与酒精依赖报告》, 3(100052):100052. Kline, D. M., White, B. N., Lancaster, K. E., Egan, K. L., Murphy, E., Miller, W. C. 和 Hepler, S. A. (2025). 采用综合丰度模型估计2016年至2021年北卡罗来纳州各县的阿片类药物滥用患病率. 《流行病学》, 36(3):310–318. Langabeer, J. R., Chambers, K. A., Cardenas-Turanzas, M. 和 Champagne-Langabeer, T. (2022). 美国阿片类药物死亡率背后的县级因素. 《物质滥用》, 43(1):76–82. Lim, T. Y., Keyes, K. M., Caulkins, J. P., Stringfellow, E. J., Cerdá, M. 和 Jalali, M. S. (2024). 改进美国阿片类药物使用障碍患病率估计：修订 Keyes 等人的研究. 《成瘾医学杂志》, 18(6):705–710. Luo, F., Li, M. 和 Florence, C. (2021). 美国各州阿片类药物使用障碍和致命性阿片类药物过量的经济成本，2017年. 《发病率与死亡率周报》, 70(15):541–546. McClellan, C., Lambdin, B. H., Ali, M. M., Mutter, R., Davis, C. S., Wheeler, E., Pemberton, M. 和 Kral, A. H. (2018). 阿片类药物过量法律与阿片类药物使用及过量死亡率的关联. 《成瘾行为》Mojtabai, R. (2022). 使用基准乘数法估计美国物质使用障碍的患病率. 《美国医学会精神病学杂志》, 79(11):1074–1080. Monnat, S. M. (2018). 与美国县级药物相关死亡率差异相关的因素. 《美国预防医学杂志》, 54(5):611–619. Monnat, S. M., Peters, D. J., Berg, M. T. 和 Hochstetler, A. (2019). 利用人口普查数据了解各县总体药物死亡率和按阿片类药物类型划分的阿片类药物相关死亡率差异. 《美国公共卫生杂志》, 109(8):1084–1091. Nam, Y. H., Shea, D. G., Shi, Y. 和 Moran, J. R. (2017). 州处方药监测项目与致命性药物过量. 《美国管理式医疗杂志》, 23(5):297–303. Parsons, J. 和 Bao, L. (2022). 统一处理离群值和影响性数据检测的方法——回顾信息价值. 《统计学（国际统计学会）》, 11(1). Peterson, E. N., Edwards, A., Wetzel, M., Waller, L. A., Cooper, H. 和 Yarbrough, C. (2025). 数据稀疏情况下估计阿片类药物使用障碍风险的贝叶斯时空自上而下框架. arXiv 预印本 arXiv:2506.02303. Powell, D. (2023). 受教育程度与美国药物过量死亡. 《美国医学会健康论坛》, 4(10):e233274. Rajabi, A., Dehghani, M., Shojaei, A., Farjam, M. 和 Motevalian, S. A. (2019). 吸烟与阿片类药物使用之间的关联：荟萃分析. 《成瘾行为》, 92:225–235. Reuter, P., Caulkins, J. P. 和 Midgette, G. (2021). 普通人群调查无法充分测量海洛因使用. 《成瘾》, 116(10):2600–2609. Rhubart, D., Chen, Q., Sterner, G., Newton, R., Shaw, B. 和 Scanlon, D. (2022). 将阿片类药物流行的缓解概念化并加以测量：宾夕法尼亚州的一项案例研究. 《米尔班克季刊》, 100(4):991–1005. Rudolph, K. E., Kinnard, E. N., Aguirre, A. R., Goin, D. E., Feelemyer, J., Fink, D. 和 Cerda, M. (2020). 相对经济状况与药物过量死亡. 《流行病学》, 31(4):551–558. Ruhm, C. J. (2016). 美国药物中毒死亡，1999–2012年：一种统计调整

分析。 人口健康指标，14(1):2。Saba, S. K., Lavista Ferres, J. M., and Weeks, W. B. (2025年)。利用公共数据：2000—2019年县级地方经济困境和药物过量死亡的变化。 《国际公共卫生杂志》， 70:1607991。Salmond, S. and Allread, V. (2019年)。美国阿片类药物流行病的人口健康方法。《骨科护理》， 38(2):95–108。SAMHSA (2018年)。2016年全国药物使用与健康调查（NSDUH-2016-DS0001）。检索自 https://datafiles.samhsa.gov/。访问日期：2024年10月1日。SAMHSA (2022年)。社区参与：构建有效且公平的物质使用预防体系的重要组成部分。SAMHSA 出版物 PEP22-06-01-005，国家心理健康与物质使用政策实验室，SAMHSA，马里兰州罗克维尔。SAMHSA (2023年)。全国药物使用与健康调查（NSDUH）：全国发布。 https://www.samhsa.gov/data/nsduh/national-releases。访问日期：2024年10月1日。SAMHSA (2024年)。如何获取 NSDUH 县级估计? https://www.samhsa.gov/data/faq/finding-commonly-requested-data-products/where-can-i-get-county-level-estimates-nsduh。访问日期：2024年10月1日。Santaella-Tenorio, J., Hepler, S. A., Rivera-Aguirre, A., Kline, D. M., and Cerda, M. (2024年)。纽约州各县阿片类药物滥用患病率估计，2007—2018年：贝叶斯时空丰度模型方法。《美国流行病学杂志》， 193(7):959–967。Sawyer, J. L., Shrestha, S., Pustz, J. C., Gottlieb, R., Nichols, D., Van Handel, M., Lingwall, C., and Stopka, T. J. (2021年)。描述地方社区的阿片类药物相关过量风险：2017年印第安纳州阿片类药物过量脆弱性评估。《预防医学报告》， 24(101538):101538。Seth, P., Rudd, R. A., Noonan, R. K., and Haegerich, T. M. (2018年)。量化处方阿片类药物过量死亡的流行病。《美国公共卫生杂志》， 108(4):500–502。Skinner, A., Nolen, S., Cerdá, M., Rich, J. D., and Marshall, B. D. L. (2024年)。一种简单的启发式方法：分配阿片类药物和解资金，以降低美国的药物过量死亡率。《美国药物与酒精滥用杂志》， 50(3):269–275。Stan 开发团队 (2024年)。RStan：Stan 的 R 接口。R 软件包版本 2.32.6。Thombs, R. P., Thombs, D. L., Jorgenson, A. K., and Harris Braswell, T. (2020年)。是什么导致了美国的药物过量流行病? 《健康与社会行为杂志》， 61(3):275–289。Thompson, K., Barocas, J. A., Delcher, C., Bae, J., Hammerslag, L., Wang, J., Chandler, R., Villani, J., Walsh, S., and Talbert, J. (2023年)。肯塔基州各县阿片类药物使用障碍的患病率：为期两年的多样本捕获-再捕获分析。《药物与酒精依赖》， 242(109710):109710。美国人口普查局 (2024年)。美国人口普查数据。 https://www.census.gov。访问日期：2024年10月1日。Vuolo, M. and Kelly, B. C. (2022年)。县级阿片类药物配发率对个体处方阿片类药物和海洛因消费模式的影响：来自美国全国数据的证据。《美国精神病学杂志》， 179(4):305–311。Young-Wolff, K. C., Klebaner, D., Weisner, C., Von Korff, M., and Campbell, C. I. (2017年)。吸烟

慢性阿片类药物治疗男性和女性的吸烟状况、阿片类药物相关问题及担忧。《临床疼痛杂志》，33(8):730–737。Zale, E. L., Dorfman, M. L., Hooten, W. M., Warner, D. O., Zvolensky, M. J., and Ditre, J. W. (2015年)。吸烟、尼古丁依赖与处方阿片类药物滥用模式：来自全国代表性样本的结果。《尼古丁与烟草研究》，17(9):1096–1103。

A 附录：补充结果

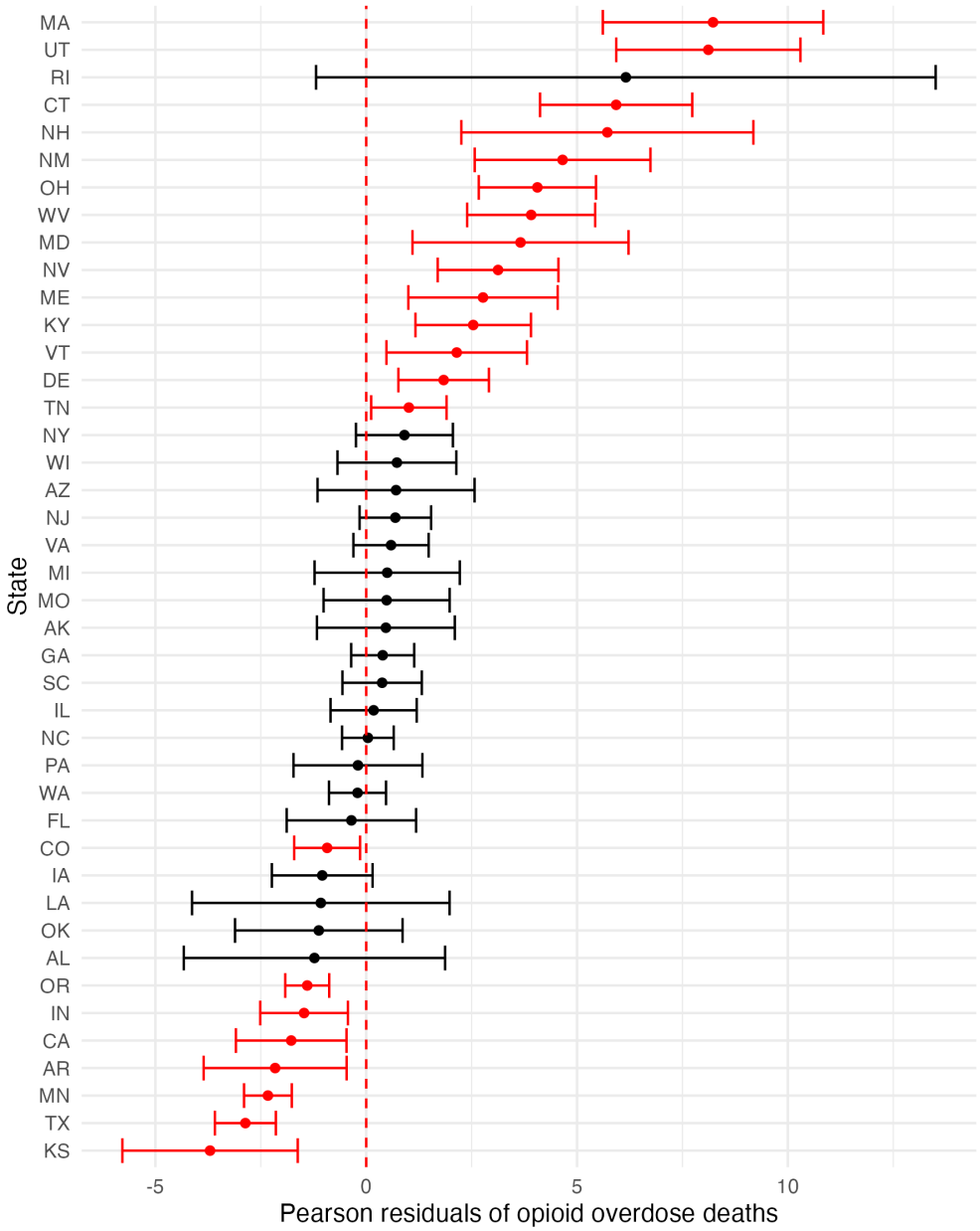


图 A.1：按州划分的、基线模型中未被抑制观测值县的阿片类药物过量死亡皮尔逊残差。点表示相应州的平均残差，线表示相应的 95%置信区间（CI）。在至少有 3 个县存在未被抑制的阿片类药物过量死亡的 42 个州中，超过半数（22 个州）的残差显著不为零，如红色标示的置信区间所示。这表明该模型未能充分捕捉州际异质性。

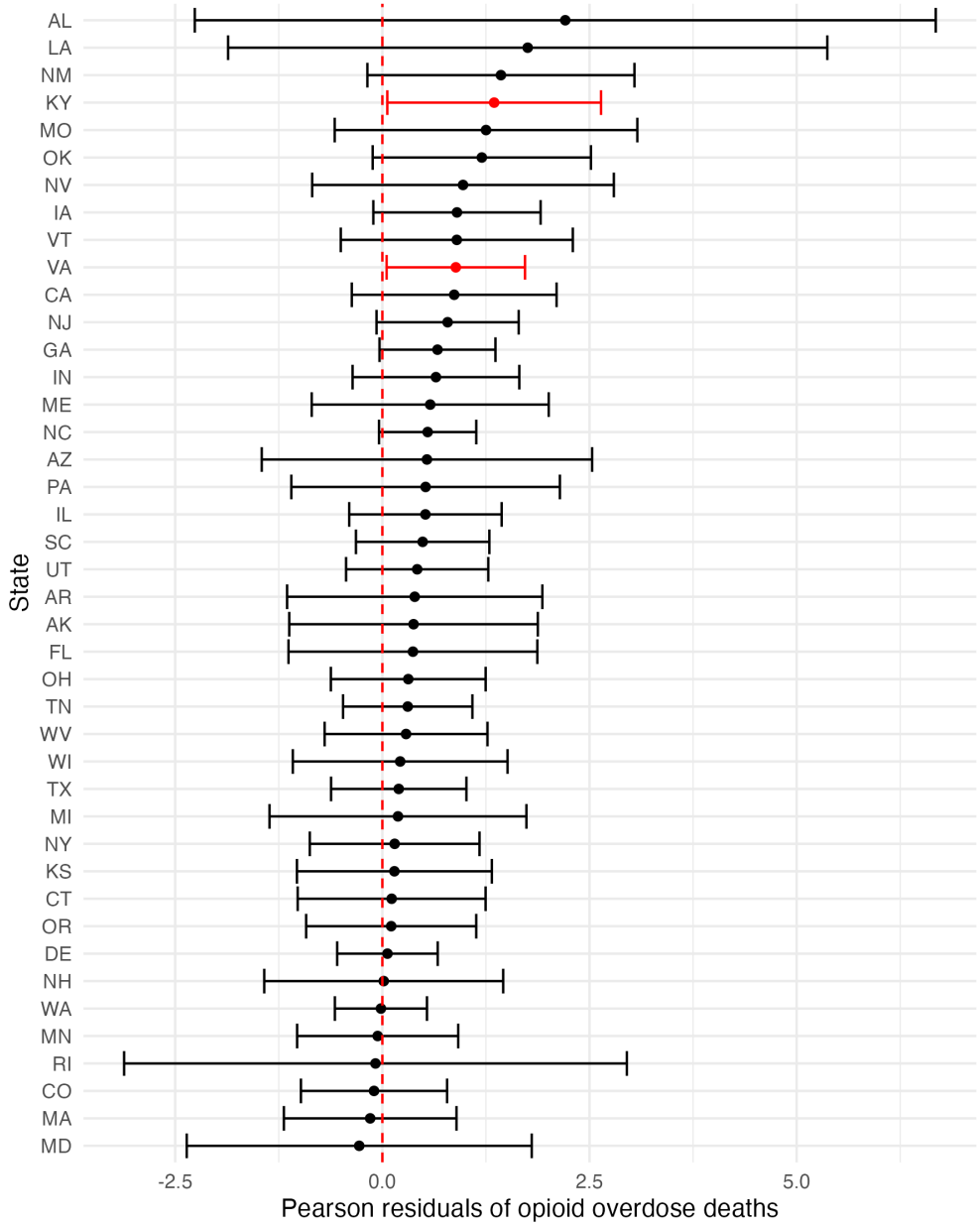


图 A.2：在按州划分、具有未被抑制观测值的县的模型中，为死亡率加入随机截距后，阿片类药物过量死亡的皮尔逊残差。点表示相应州的平均残差，线表示对应的 95%置信区间。对于大多数州（42 个至少有 3 个县存在未被抑制的药物过量死亡人数的州中有 40 个），平均残差与零无显著差异。

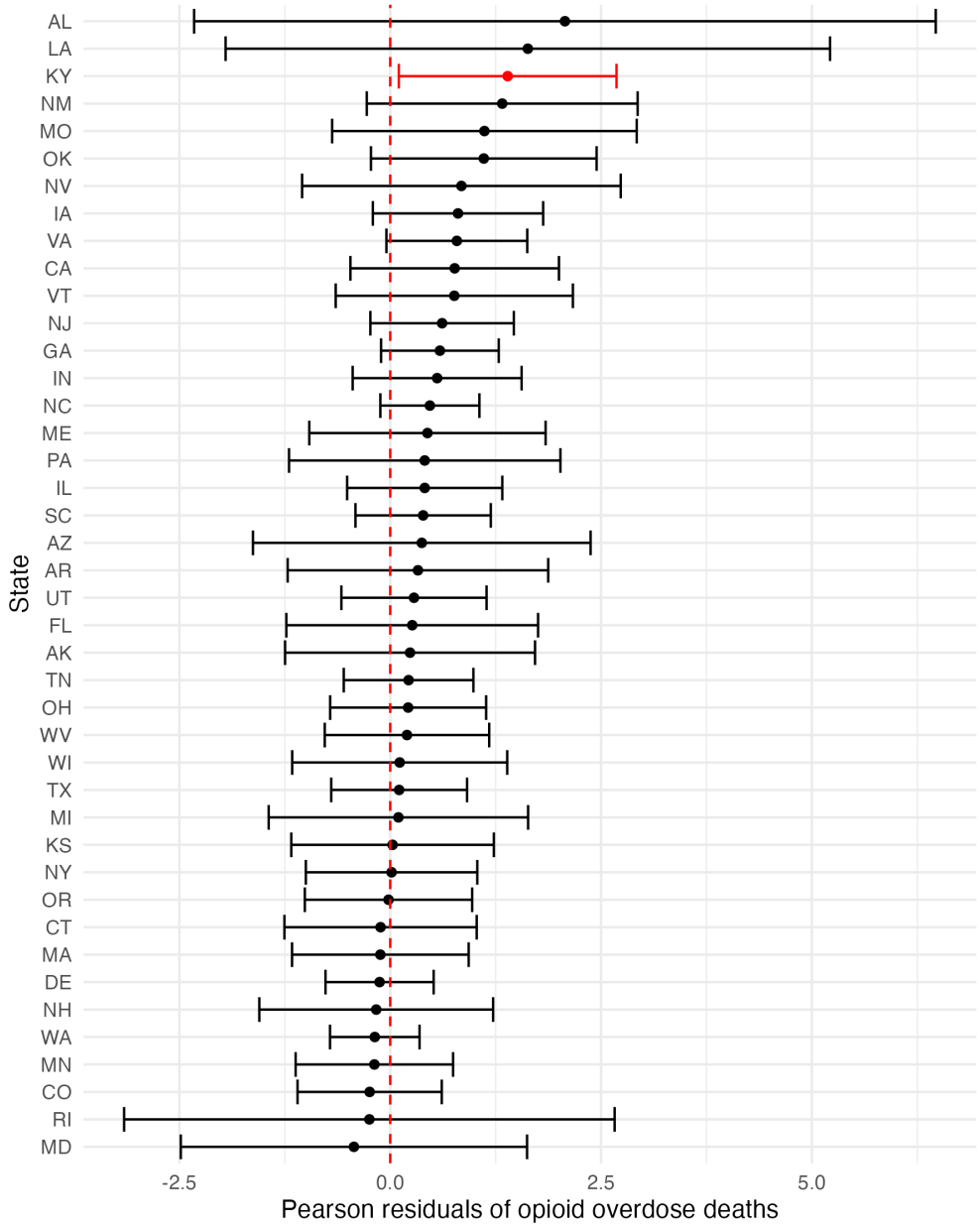


图 A.3：在模型中为死亡率和患病率添加随机截距后，各州阿片类药物过量死亡的皮尔逊残差。点表示相应州的平均残差，线表示相应的 95%置信区间。只有一个州的平均残差显著异于零。

表 A.1：在贝叶斯层级模型中为死亡率和患病率加入州级随机截距后，按州对药物过量死亡的残差与各县协变量进行回归所得的单变量回归分析结果。本表汇总了单变量回归中截距和斜率显著的州的数量（在至少有 3 个县拥有未屏蔽的药物过量死亡数据的 42 个州中）。结果表明，当前模型的残差仅在少数几个州仍与协变量显著相关，这意味着在模型中为这些协变量加入随机斜率并无必要。

协变量	**显著截距 ($p < 0.05$)**	**显著斜率 ($p < 0.05$)**
阿片类药物配发	1	3
成年人吸烟	0	4
部分大学教育	2	3
失业	3	4
初级保健医生	0	1

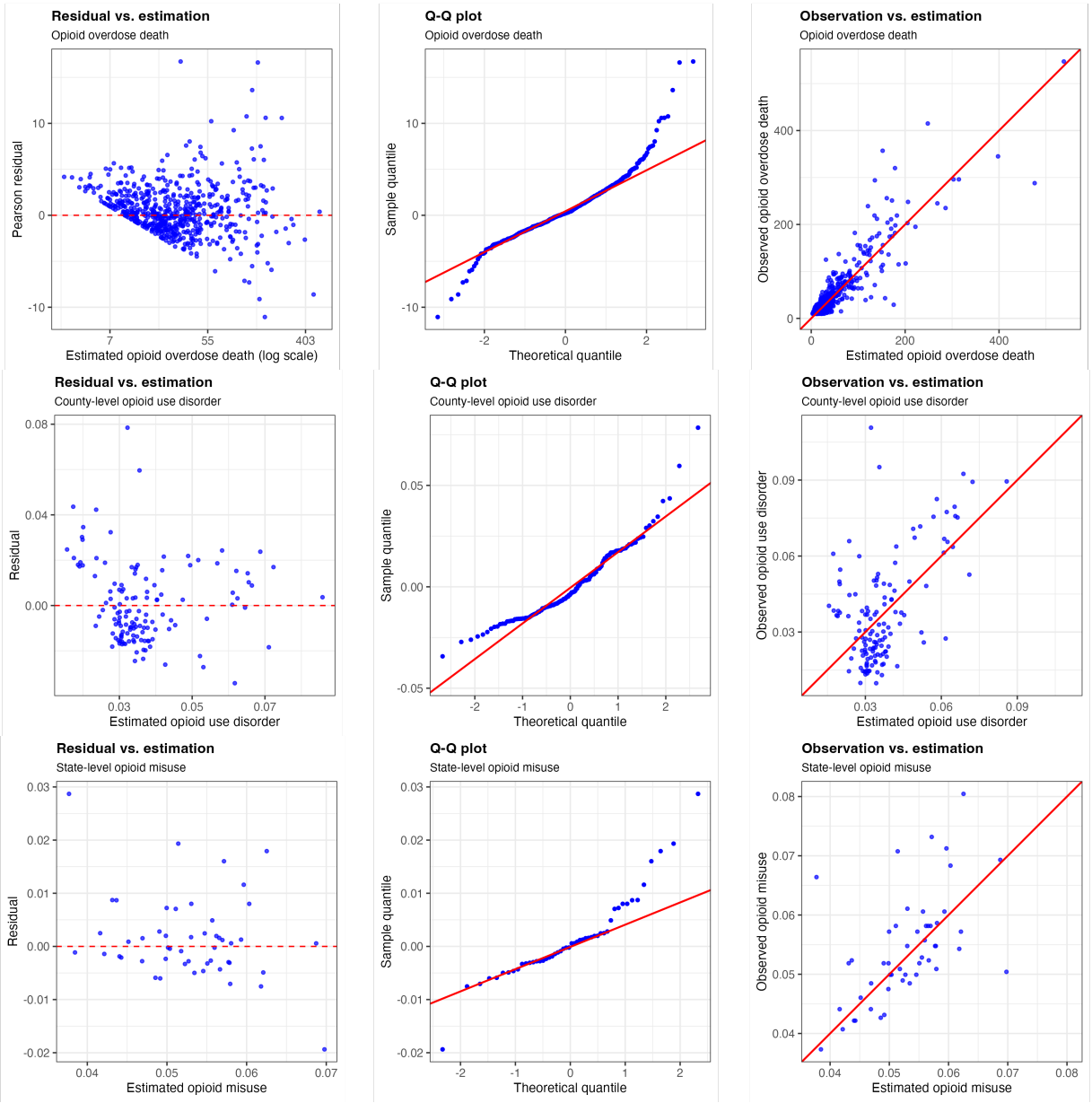


图 A.4：用于评估所提出的贝叶斯层级模型的分布假设和拟合优度的模型诊断图。各行对应不同的输入数据，分别为：(1) 县级药物过量死亡人数（针对数值未被抑制的县）、(2) 县级阿片类药物使用障碍患病率估计（马萨诸塞州和肯塔基州）以及 (3) 州级阿片类药物滥用率估计。从左到右，各列分别显示：(1) 残差与估计值的关系图、(2) 残差的正态 Q-Q 图以及 (3) 观测值与估计值的关系图。

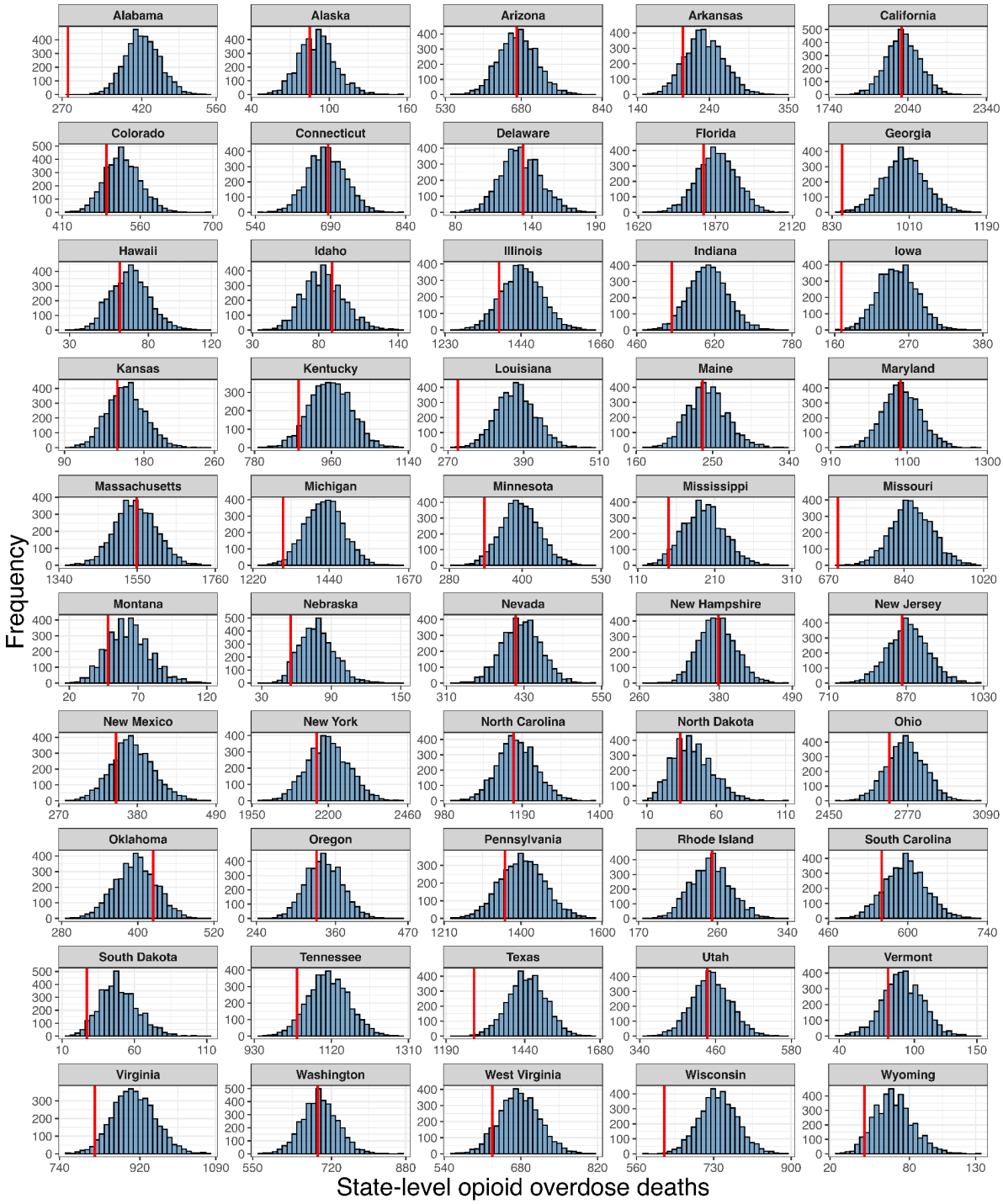


图 A.5：州级阿片类药物过量死亡的后验预测检验。直方图表示贝叶斯分层模型按州给出的药物过量死亡后验预测分布，红线表示各州观察到的药物过量死亡总数。

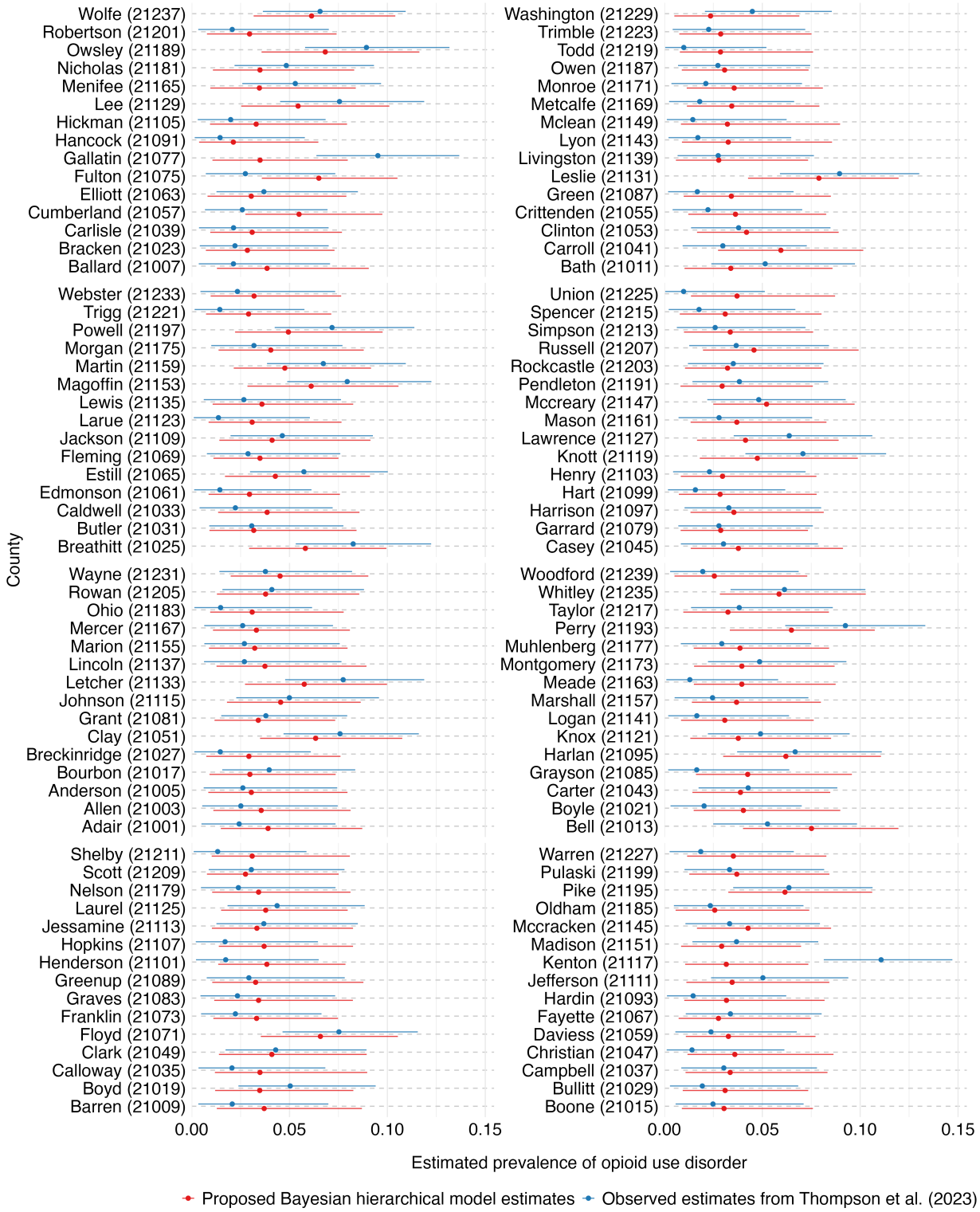


图 A.6：贝叶斯层级模型的 10 折交叉验证结果与现有估计值的比较，现有估计值来自 Thompson 等人 (2023年)。图中所示的 Thompson 等人 (2023年) 的 OUD 估计值可信区间基于估计的 $\hat{\sigma}_{KY}$（详见第4.1节）。

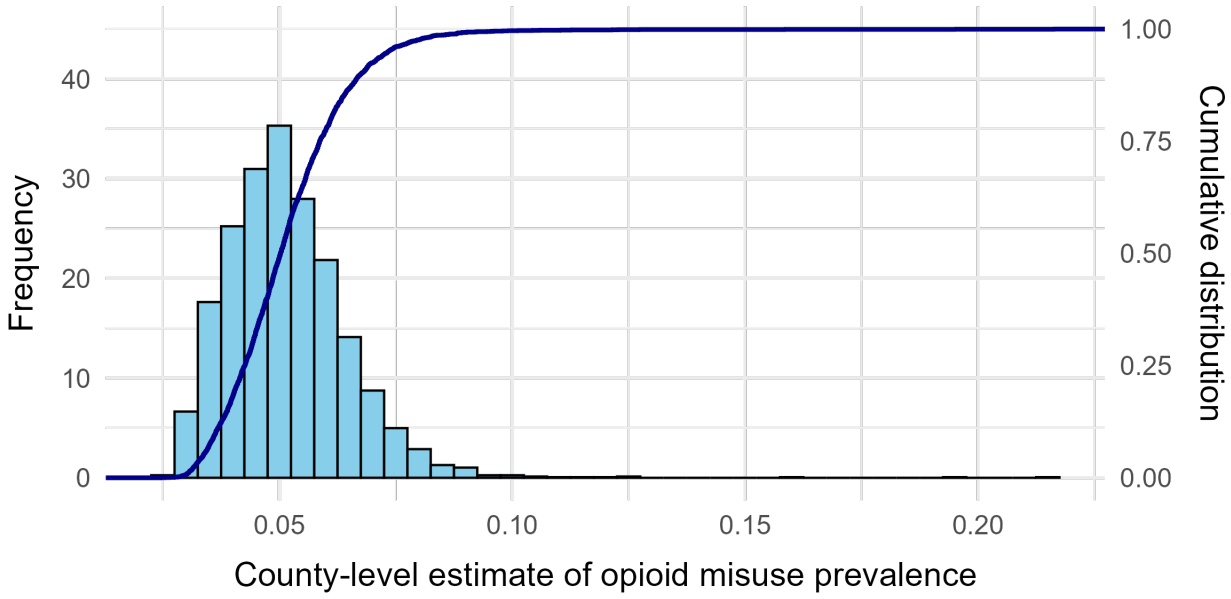


图 A.7：县级阿片类药物滥用估计患病率的分布。

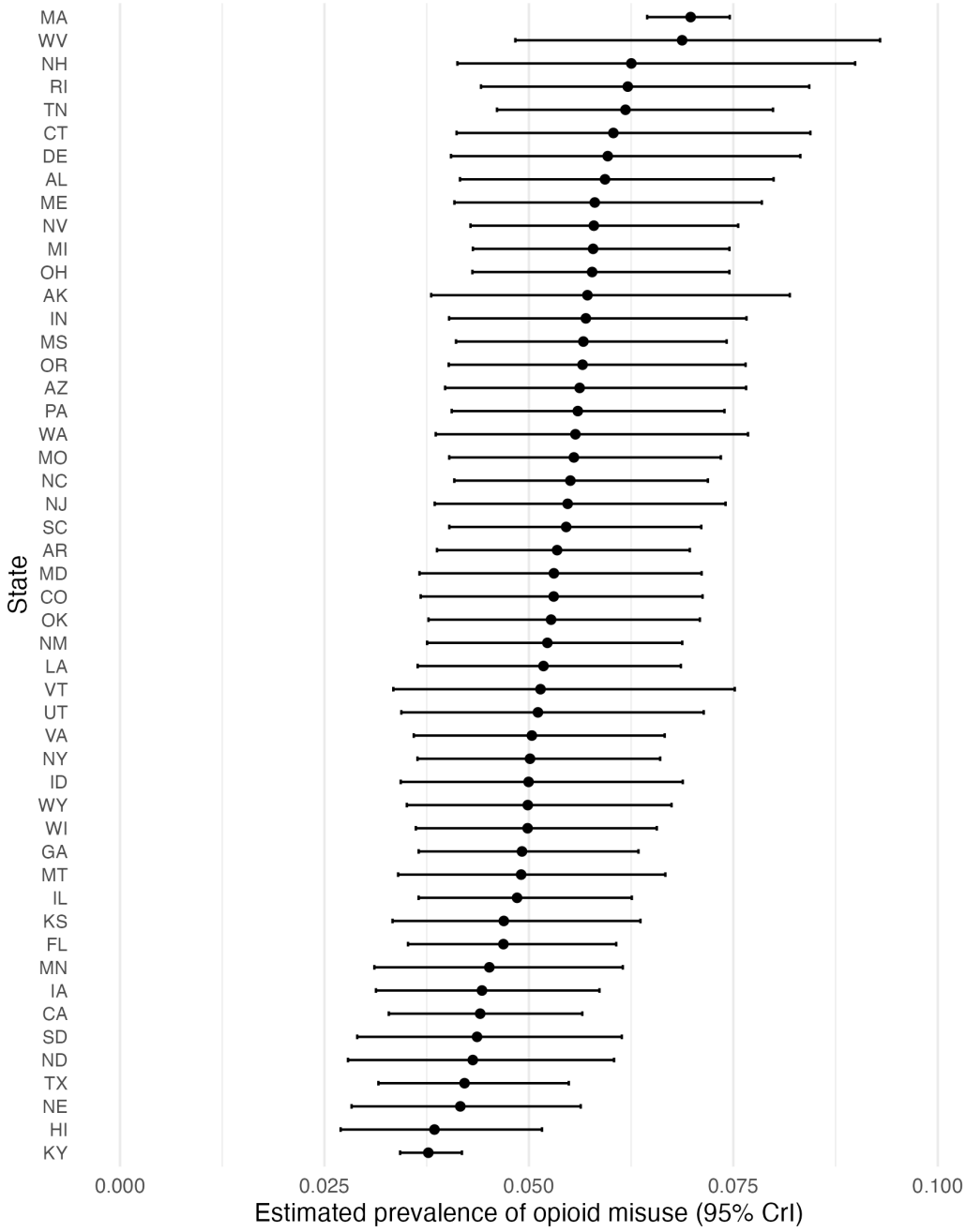


图 A.8：阿片类药物滥用患病率的州级估计。

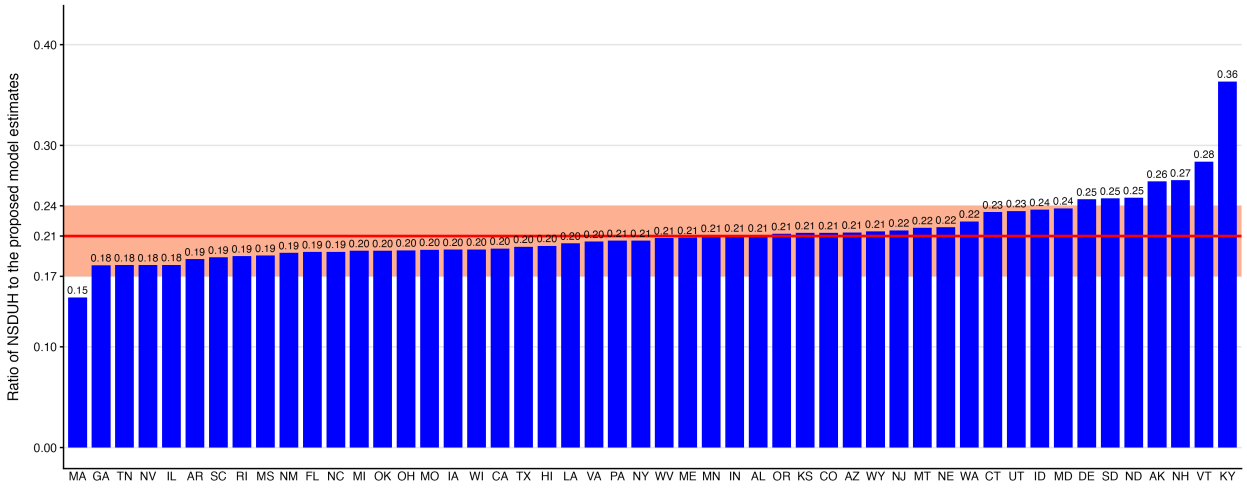


图 A.9：全国药物使用与健康调查州级阿片类药物滥用估计值与拟议贝叶斯层级模型在州级汇总后的估计值之比。红线及其阴影区域表示缩放系数 $\gamma = 0.21$（0.17，0.24）的后验均值和95%可信区间。

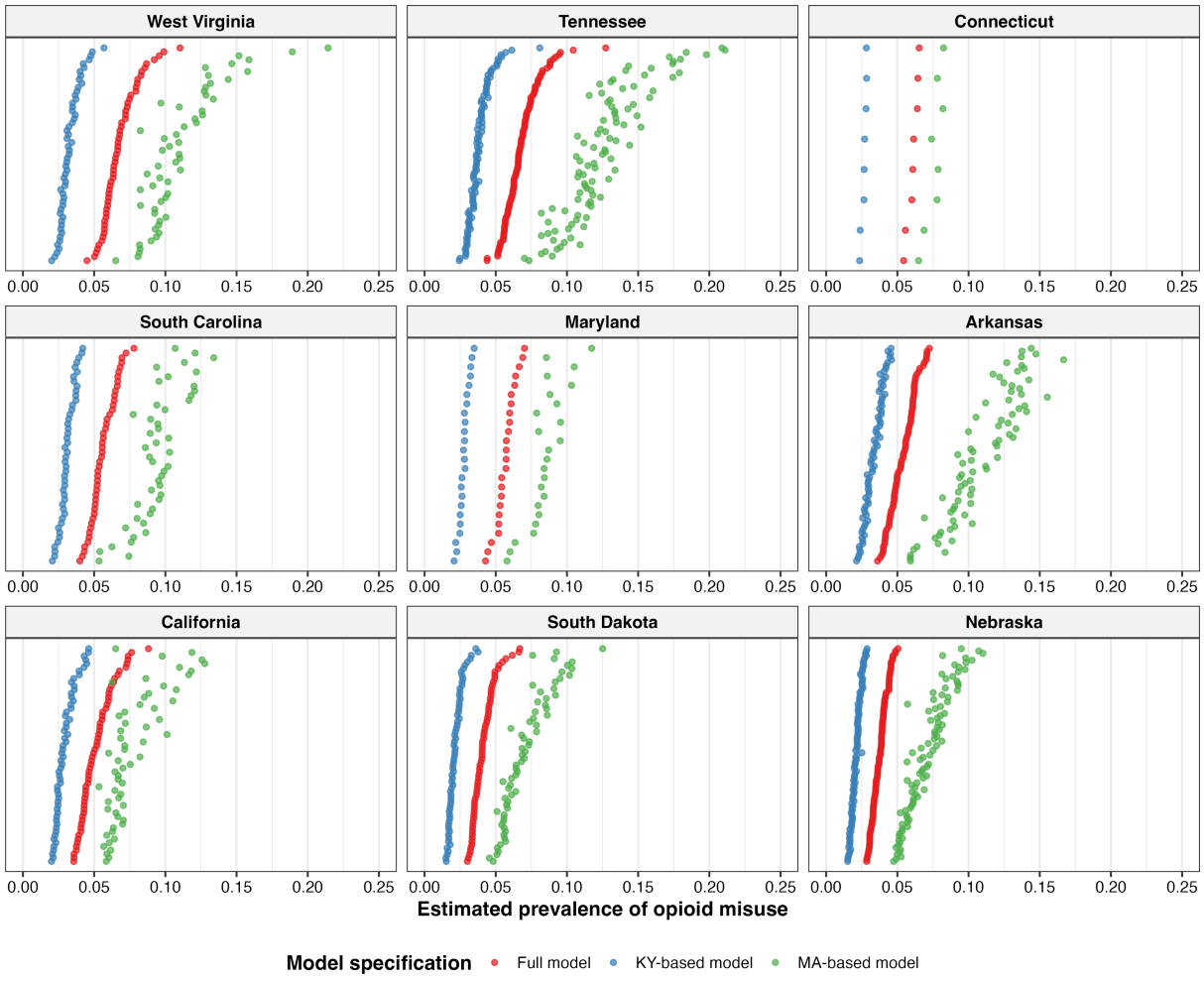


图 A.10：留一州交叉验证中，九个选定州的单州模型与完整模型所得县级患病率估计的比较。基于 MA 和 KY 的模型估计值，分别通过留出肯塔基州 (KY) 和马萨诸塞州 (MA) 的县级患病率数据得出。为便于说明，选取九个州，分别代表患病率高（西弗吉尼亚州、田纳西州、康涅狄格州）、中等（南卡罗来纳州、马里兰州、阿肯色州）和低（加利福尼亚州、南达科他州和内布拉斯加州）的州。每行中不同颜色的三个数据点代表同一县在不同模型下的患病率估计；为便于阅读，图中省略了县名。

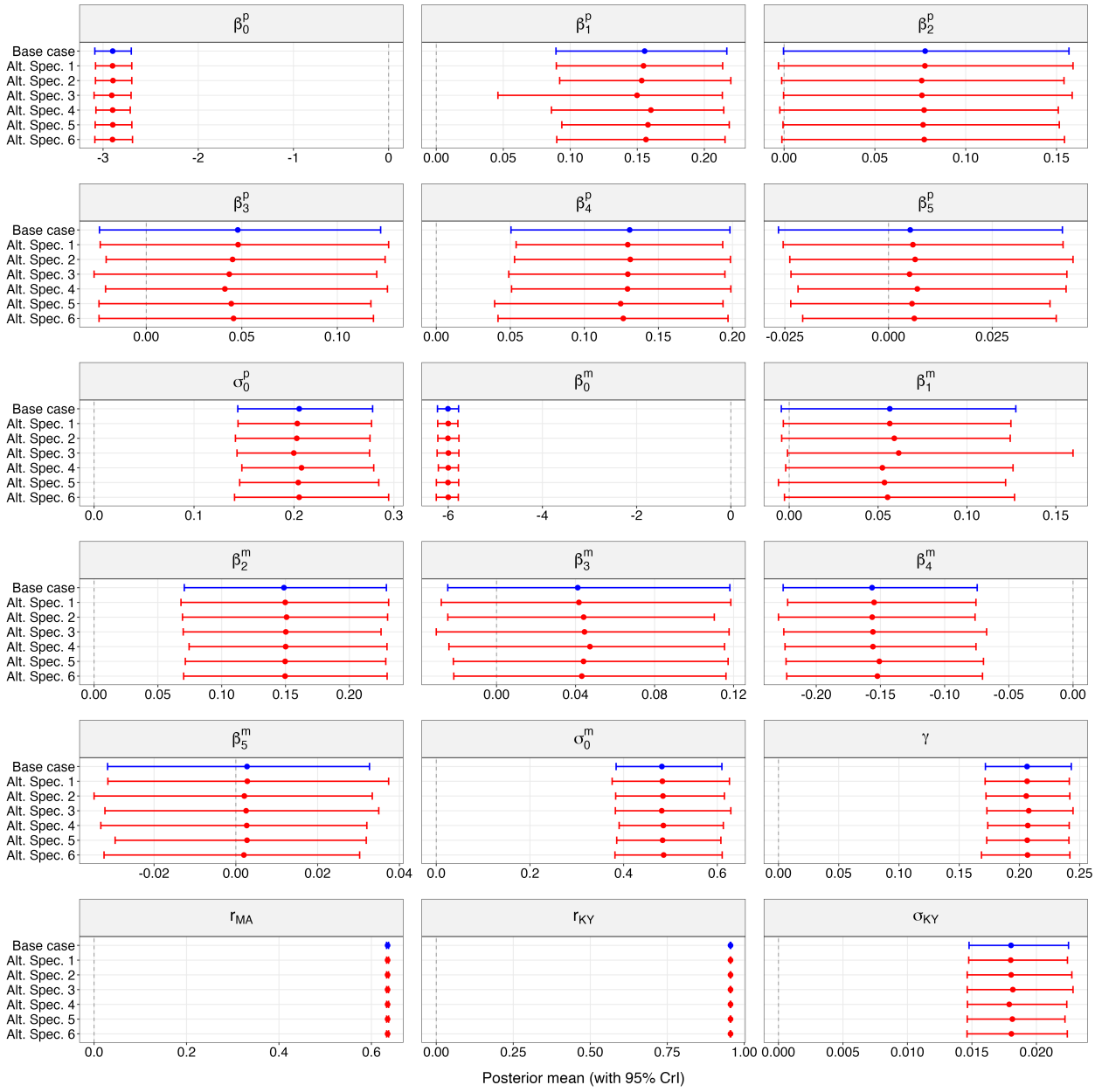


图 A.11：贝叶斯模型先验设定的敏感性分析结果。点和水平线分别表示每种先验设定下各参数的后验均值和95%可信区间。替代设定（Alt. Spec.）1和2分别将固定效应截距 β_0^p 和 β_0^m的先验从基准情形 $\mathcal{N}(0, 5^2)$ 改为$\mathcal{N}(0, 5^2)$和$\mathcal{N}(0,\ 2.5_2)$；替代设定3和4分别将 σ_0^p，σ_0^m的先验从基准情形的Half-Normal($\sigma = 10$)改为Half-Normal($\sigma = 5$)和Half-Normal($\sigma = 2.5$)；替代设定5采用尺度为0.5的更强收缩Half-Normal Horseshoe+ 先验（基准情形中的尺度为1.0）；替代设定6结合了替代设定2、4和5。