arXiv:2603.02424v1 [stat.AP] 2 Mar 2026

关于使用时变模型评估口罩使用和超额死亡率的误用

Beny Spira[1] 和 Daniel V. Tausk[2]*

[1]微生物学系，生物医学研究所，圣保罗大学 e
圣保罗大学，圣保罗，巴西[2]数学系，圣保罗大学，圣保罗，巴西

*通讯作者，电子邮箱：tausk@ime.usp.br

摘要

口罩作为一种针对呼吸道病毒传播的人群层面干预措施，其有效性仍存在争议。新冠肺炎大流行期间发表的大量观察性研究报告了有益效果，而随机对照试验则一贯显示其影响有限或没有影响。一项针对欧洲国家的生态学分析报告称，2020年和2021年期间的平均口罩使用率与同期24 个欧洲国家的超额死亡率呈正相关（Tausk & Spira，2025a）。在多次尝试控制混杂变量后，这种关联依然存在。随后，其他作者对这一发现提出质疑，并将其归因于反向因果关系（Cerqueira-Silva等，2026年）。本文将详细重新评估这些批评。我们指出，他们的分析存在根本性缺陷：用于反驳原始发现的时间依赖回归框架会产生虚假结果，部分原因在于将累积超额死亡率作为结果变量，从而纳入干预前死亡，并在不可能的负时间滞后下也得出具有统计显著性的效应。诊断分析进一步表明，该模型的关键假设并不成立，因此任何关联或因果解释都无效。最后，我们提出一项关于口罩使用的原创纵向分析，旨在直接检验反向因果假说。通过构建多个指标，衡量疫情波次不同阶段（包括死亡率较低的波次间期）的口罩采用情况，我们发现口罩使用与超额死亡率之间的关联依然存在，并非由口罩使用的应对性增加所致。这些发现提供了有力的

证据表明，反向因果关系对观察到的关联至多只有轻微影响。这些结果凸显了在评估非药物干预时，恰当定义结局、验证模型以及确保证据严谨性的重要性。

1 引言

使用口罩作为限制呼吸道病毒传播的手段仍存在争议。一方面，新冠肺炎大流行期间发表的大多数论文都采取支持口罩的立场，并声称口罩能有效预防病毒传播（Spira，2026）。在此期间，世界卫生组织（WHO）也建议使用口罩，许多国家则强制要求佩戴口罩（Masks4All，2020）。此外，许多人，包括医护人员和科学家，凭直觉认为口罩能为个人提供保护，避免感染新冠肺炎。

另一方面，现有质量最高的证据——随机对照试验（RCTs）始终表明，口罩对病毒传播几乎没有影响，甚至完全没有影响。这些研究汇总于表1。在评估口罩使用情况的16项随机对照试验中（Cowling 等人，2008年，2009；Canini 等人，2010；Larson等人，2010；Simmerman等人，2011年；Suess 等人，2012年；MacIntyre等人，2016；Aiello等人，2012年；Barasheed，2014；Alfelali，2020年；Bundgaard，2020年；Solberg 等人，2024年；Abaluck 等人，2022年；Nanque 等人，2024年；Jacobs等人，2009；MacIntyre等人，2015），只有两项报告了具有统计学显著性的正面结果（Solberg 等人，2024年；Abaluck 等人，2022年），而且即使是这两项研究所报告的效果也较为有限，并存在重大局限。

表 1. 迄今为止关于口罩的随机对照试验

作者、年份，国家	研究场景	N	主要终点	风险比（95%置信区间）	备注
Cowling 等人（2008年）Hong Kong	社区	198	ILI LCI	0.88 [0.34，2.27] 1.16 [0.31，4.34]	医用/外科口罩与不戴口罩相比
Canini 等人（2010年）法国	社区	105	ILI	1.03 [0.52，2.00]	医用/外科口罩对比不戴口罩
Larson等人（2010年）美国	社区	509	ILI LCI	0.63 [0.31，1.29] 0.57 [0.17，1.87]	手卫生 + 医疗/外科口罩相比于手卫生
Simmerman等人（2011年）泰国	社区	442	ILI LCI	1.03 [0.61，1.73] 0.97 [0.62，1.52]	手卫生 + 医疗/外科口罩相比于手卫生
Suess 等人（2012年） 德国	社区	84	ILI LCI	0.61 [0.20 , 1.87] 0.39 [0.13 , 1.19]	医用/外科口罩与不戴口罩相比
MacIntyre。（2016）中国等人	社区	245	ILI	0.32 [0.03 , 3.11]	医用/外科口罩与不戴口罩相比
Aiello等人（2010年）美国	社区	1437	LCI ILI	2.43 [0.59, 10.12] 0.82 [0.65, 1.01]	医用/外科口罩与不戴口罩相比
Aiello等人（2012年）美国	社区	1178	ILI LCI	1.10 [0.88，1.38] 0.92 [0.59，1.42]	医用/外科口罩与不戴口罩相比
Barasheed 等等（2014年）沙特阿拉伯	社区	164	ILI	0.58 [0.32，1.04]	医用/外科口罩与不戴口罩相比
Alfelali 等人（2020年）沙特阿拉伯	社区	7687	ILI LCI	1.10 [0.90，1.35] 1.40 [0.92，2.14]	医用/外科口罩与不戴口罩相比
Bundgaard 等（2021年）丹麦	社区	4862	LCI* 实验室确诊的其他病毒感染	0.82 [0.54，1.23] 0.58 [0.25，1.31]	医用/外科口罩与不戴口罩相比
Solberg等人（2024年）挪威	社区	4467	SRRS	0.71 [0.58-0.87]	医用/外科口罩与不戴口罩相比。
Abaluck 等人 al.（2022年）孟加拉国	社区	336010	CLI	0.87 [0.81，0.94]	医用/外科口罩与不戴口罩相比
Nanque 等人 (2024年) 几内亚-比绍	社区	39574	CLI	0.81 [0.57，1.15]	布口罩与不戴口罩
Jacobs等人（2009年）日本	医院	33	ILI	0.88, [0.02, 32]	医用/外科口罩与不戴口罩相比
MacIntyre等人（2015年）越南	医院	1607	ILI	0.26 [0.03, 2.50]	医用口罩与布口罩相比口罩

CLI，COVID样疾病；ILI，流感样疾病；LCI，实验室确诊感染；LCI*，实验室确诊的SARS-CoV-2感染；LCI**，实验室确诊的其他病毒感染；SRRS，自我报告的呼吸道症状。

由 Solberg等人开展的(2024社区型随机对照试验报告称，在公共场所使用外科口罩与呼吸道感染风险相对降低29%相关；然而，这仅相当于绝对风险降低3.2%。此外，主要终点范围宽泛且具有主观性，即“自我报告的呼吸道症状”。若将结局限定为通过PCR或抗原检测确诊新冠肺炎这一更严格的标准，估计优势比为1.07，且无统计学意义。此外，亚组分析显示，表面上的获益仅见于自认为口罩能够降低感染风险的参与者。

Abaluck 等人开展的大型整群随机试验 (2022年)报告称，口罩使用者中类新冠疾病的发生率显著降低（相对风险降低13%）。然而，在随后一篇关于减少呼吸道病毒传播的物理干预措施的Cochrane综述中，该研究在六个领域中的五个被评为偏倚风险高（Jefferson等人，2023），原因包括基线不平衡、结局评估主观、随访不完整，以及偏离预先规定的方案。值得注意的是，方案中规定的基线血清学检测并未实施，且未提供充分理由。总的来说，质量较高的随机研究几乎不支持口罩能显著预防呼吸道病毒传播（见第3节），这与更广泛的大流行文献中对口罩压倒性正面的描述形成鲜明对比。

随机对照试验的证据并未减少大量将口罩描绘为抗击病毒感染的重要且有效工具的文章（Spira，2026年）。尽管如此，仍有少数既非随机对照试验、也非其综述或荟萃分析的论文，对这一主流叙事提出了质疑。其中一项研究是对欧洲口罩使用与超额死亡率之间关系所做的生态学观察分析（Tausk & Spira，2025a）。该研究使用双变量和多变量回归分析，考察了24个欧洲国家的平均口罩使用情况与各类超额死亡率指标之间的相关性。在双变量分析（Spearman 相关系数 $\rho = 0.477$，$p = 0.018$）和多变量模型（标准化系数 $= 0.52$，$p = 0.0012$）中，口罩使用率与年龄调整后的超额死亡率之间均存在具有统计学显著性的正相关。相比之下，在双变量分析（Spearman 相关系数 $\rho = -0.659$，$p < 0.001$）和多变量分析（标准化系数 $= -0.48$，$p = 0.0016$）中，疫苗接种率均与年龄调整后的超额死亡率呈显著负相关，这与疫苗接种具有保护作用的结论一致。研究对潜在混杂因素进行了全面考量，包括社会经济因素、基线人群健康状况指标、城市人口密度、疫苗接种率以及其他非药物干预措施的严格程度。此外，研究还开展了多项敏感性分析，以评估研究结果的稳健性。在所有纳入考量的分析中，口罩使用与超额死亡率之间具有统计学显著性的正相关均持续存在。

文章发表后，期刊 *BMC*公共卫生 受到要求撤回该文章的压力（Jornal_da_USP，2026年）。这促使施普林格·自然的研究诚信办公室启动发表后审查；数月后，审查以发布一份勘误 （Tausk& Spira，2025年）告终。该勘误 包含一些方法学澄清，并对讨论部分的三个句子作了小幅修订，采用了更谨慎的措辞，但未改变研究结果或结论。

随后，显然对这一结果并不满意，一群作者在《柳叶刀区域健康—美洲》（Cerqueira-Silva等，2026年发表了一篇文章，明确意在削弱Tausk& Spira （2025a）的研究发现。
作者认为，关于口罩使用与超额死亡率之间所谓关联的结论，建立在有缺陷的方法学假设之上，尤其是采用总体平均值，忽略了口罩采用的时间波动。他们主张，口罩使用实际上是对疫情加剧和政府管控趋严的反应。为反驳Tausk &Spira（2025a），作者使用双向固定效应模型对同一组数据进行了时间依赖分析，结果显示口罩使用与累积超额死亡率之间存在统计学显著负相关。作者还认为，原研究的发现缺乏临床和生物学合理性，因为据他们所说，科学已证明口罩能够阻挡病毒传播。此外，他们的批评带有政治视角，援引该研究据称被用于否认主义论调，因此将讨论政治化，而非严格保持其科学性。

在本文中，我们通过说明他们用于评估口罩使用与超额死亡率之间关系的双向固定效应模型存在根本性缺陷，驳斥Cerqueira-Silva 等人(2026年的论断。我们还对其文章中其他重要的失实描述作出回应。最后，我们提供了一项原创分析，表明Tausk & Spira (2025a) 报告的口罩使用与超额死亡率之间的正相关，不太可能由反向因果关系解释，也就是说，这种相关关系并非由大流行期间新冠肺炎死亡率上升后口罩使用随之增加所致。

2 结果与讨论

1 Cerqueira-Silva 等人(2026年所使用的双向固定效应模型显然存在缺陷

Cerqueira-Silva等人（2026年）使用双向固定效应模型，估算口罩使用对超额死亡率以及随后几周的累积超额死亡率的影响。更具体地说，对于 t 2020年1月1日至2022年1月2日期间的每一周，系数 β 衡量其模型中平均口罩使用对第 t 周的影响，具体而言是对

第 $t+d$ 周超额死亡率以及累积超额死亡率（从2020年1月1日至第 $t+d$周）的影响，其中滞后期 d 介于1至4之间（关于累积超额死亡率变量的定义，参见Our World in Data（2024年））。点估计值 β 及其对应的置信区间报告于补充材料部分，为方便读者，现转载如下（表2）。

表2。 转载自Cerqueira-Silva等人（2026年）

Lag	每周超额死亡率 β（95%置信区间）	累积超额死亡率 β（95%置信区间）
1周	0.0524 [−0.1008; 0.2057]	−0.0674 [−0.1270; −0.0077]
2 周	0.0426 [−0.1015; 0.1867]	−0.0688 [−0.1269; −0.0107]
3 周	0.0269 [−0.1107; 0.1644]	−0.0702 [−0.1263; −0.0140]
4周	0.0134 [−0.1207; 0.1476]	−0.0713 [−0.1253; −0.0172]

如果目标是研究口罩使用对死亡率随时间变化的影响，那么自然的结果变量选择是：在口罩使用期间之后若干周测得的每周超额死亡率，也就是口罩使用可能产生影响的时间窗口内的超额死亡率。累积超额死亡率作为结果变量显得有些奇怪，因为它涵盖了口罩使用期间之前数量庞大（且不固定）的若干周内的超额死亡率。作者似乎也明白这一点，正如他们自己所说，累积超额死亡率“可能会因平滑掉异质性流行病阶段而引入不同的偏差”。

我们注意到，当使用每周超额死亡率作为结果变量时，参数 β 的点估计在所有滞后期均为正。这与我们的发现一致：更多的口罩使用与更高的超额死亡率相关。然而，由于结果在统计学上并不显著，因此无法据此得出太多结论。

只有在将累积超额死亡率用作结果变量时，β 的估计值才会变为负值且具有统计显著性，这表明口罩使用正在降低超额死亡率。然而，观察到的效应仅滞后一周，这是一个警示信号，因为如果口罩能够减少传播，那么对死亡率的任何影响都应在更长时间后才会显现。由于文献中对病例与死亡之间间隔的估计各不相同（Fritz，2022年），我们决定使用Cerqueira-Silva等人提供的代码（2026年），计算参数 β 在滞后为0和 -1 时的估计值（表3）。

表3。扩展滞后分析

Lag	每周超额死亡率 β （95%置信区间）	累积超额死亡率 β （95%置信区间）
-1 周	**0**.0438 [−0.1246； 0.2123]	**−0.0639 [−0.1251； −0.0027]**
0周	**0**.0523 [−0.1093; 0.2140]	**−0.0658 [−0.1264; −0.0051]**
1周	**0**.0524 [−0.1008; 0.2057]	−0.0674 [−0.1270; −0.0077]
2 周	**0**.0426 [−0.1015; 0.1867]	−0.0688 [−0.1269; −0.0107]
3 周	**0**.0269 [−0.1107; 0.1644]	−0.0702 [−0.1263; −0.0140]
4周	**0**.0134 [−0.1207; 0.1476]	−0.0713 [−0.1253; −0.0172]

表3显示， β 及其相应置信区间在滞后0 和 −1 时的估计值，与滞后1至4周时所得的估计值非常相似。某一周的口罩使用不可能影响同一周的累积超额死亡率，更不可能影响前几周的超额死亡率，因此，这一结果有力地证明，Cerqueira-Silva 等人（2026年）报告的效应，是由于将干预前死亡率纳入结果变量而产生的假象。因此，所报告的效应并不真实，使得分析在根本上站不住脚。此外，我们将在下文表明，Cerqueira-Silva等人（2026年）所用模型的重要假设并不成立。这或许可以解释，为何该模型会预测在口罩干预前时期，口罩使用对超额死亡率存在虚假的影响。

1.1 模型的基本假设无效

Cerqueira-Silva 等人使用的Cerqueira-Silva 等人(2026年双向固定效应模型形式如下

$$Y_{it+d} = a_i + b_t + \beta M_{it} + \varepsilon_{it},$$

其中， Y_{it} 是国家 i 在第 t周的结果变量（超额死亡率或累积超额死亡率）， d 是在1至4周之间固定选定的时间滞后， a_i 是代表国家特异性时间不变效应的参数， b_t 是代表时间特异性国家无关效应的参数， β 是我们关注的参数， M_{it} 是国家 i在第 t周的平均口罩使用情况， ε_{it} 是随机误差项。简而言之，除口罩使用的影响和随机噪声外，结果变量是国家特异性时间不变项 a_i 与时间特异性国家无关项 b_t之和。

众所周知，要在双向固定效应模型中恰当地估计感兴趣的参数，主效应具有 $a_i + b_t$ 有效的加性结构至关重要（Imai&Kim， 2020年）。事实上，很容易证明，若可加性假设不成立，那么通常用于估计 β 的普通最小二乘估计量往往会产生偏差，而且随着样本量趋于无穷大，偏差也不会趋于零；换言之，该估计量是不一致的。

就当前情况而言，可加性假设显然并不现实。至少，不同国家感染浪潮之间的时间滞后会使时间效应因国家而异。此外，通常不能指望国家特异性效应以加性的方式发挥作用。例如，病毒再生数的微小差异就会对感染浪潮产生复杂的非线性效应。此外，年龄结构、人口整体基线健康状况以及卫生基础设施质量方面的差异，也会改变感染病死率，从而产生国家特异性的乘性效应。

为了评估模型假设的有效性，我们绘制了24个国家各自的残差随时间变化图。这是诊断线性模型的标准图形方法：当模型假设成立时，残差应表现为围绕零值波动的随机噪声。然而，本例中观察到的情况恰恰相反。以累积超额死亡率为结果变量并采用时间滞后 d介于1至4周的24张残差图见 GitHub 仓库中的文件（Tausk，2026年）。这些图有力地证明，模型假设遭到了严重违反。如果Cerqueira-Silva 等人(2026年)进行过这项基本的诊断检查，就会发现他们的模型存在根本性缺陷。

2 解决反向因果问题

Cerqueira-Silva 等人 (2026年提出的核心论点之一是：口罩使用率较高的国家之所以如此，是为了应对更严重的疫情波次；这一事实也解释了Tausk& Spira (2025a)发现的口罩使用与超额死亡率之间的正相关。我们将这一质疑称为反向因果问题：因果方向（如果存在的话）并非从干预（口罩使用）指向结果（死亡率），而是相反。Cerqueira-Silva 等人 (2026年)将图1d作为唯一的直接证据，支持口罩采用率在疫情暴发加剧后出现反应性上升这一主张。他们写道：

> 为进一步提供研究发现的背景信息，我们纳入了巴西的数据（图 *1d*），展示超额死亡率（月度）和口罩使用情况（*30*天移动平均值）。图中显示，口罩采用通常发生在死亡率达到峰值之后，这表明人们是在疫情暴发恶化后作出应对，而非口罩产生了因果效应；这一点与 *Tausk* 和 *Spira* 的结论相矛盾。

然而，除了该图表呈现的是单个国家——巴西——的数据，而不是一篇批评针对欧洲国家开展的研究的论文所应呈现的欧洲数据之外，我们很难看出该图表如何支持反向因果关系这一主张。例如，第二波疫情期间的口罩使用率在2021年4月达到峰值，但与2020年6月第一波疫情期间几乎完全相同，始终维持在70%左右，尽管第二波疫情

的致死性高得多，超额死亡率约为80%，而第一波疫情约为27%。如果反向因果关系确实在起作用，那么预计第二波疫情期间的口罩使用率应显著更高，但并未观察到这种增加。由Cerqueira-Silva 等人 (2026年)呈现的数据因此否定了而非支持其关于反向因果关系的主张，削弱了其解释的可信度。

尽管 Cerqueira-Silva 等人 （2026年）提出的关于反向因果关系的论据并不成立，但我们承认，反向因果问题确实值得关注，而且我们在原始文章中对此讨论得不够充分。因此，我们在此详细探讨这一问题。

我们首先定义五个指标——`maskall`、`maskinwave`、`maskinterwave`、`maskbeginwave`和`maskpeakwave` 这些指标对应于不同时段的口罩使用情况。我们注意到，由于各国疫情波次的起止时间并不相同，在时间依赖分析中，按相同时间段比较各国的口罩使用情况和死亡率意义不大。因此，这五个指标所对应的时段是根据其与疫情波次的关系来界定的，而非根据日历时间。我们为每个国家识别出三个疫情波次（图 3）。方法部分介绍了用于确定各国每个疫情波次起止时间的具体方法。下面我们定义这五个口罩使用指数：

- maskall：整个分析期间（从2020年02月04日至2021年12月31日）的平均口罩使用情况；
- maskinwave：疫情波次对应期间的平均口罩使用情况；
- maskinterwave：波次间期的平均口罩使用情况，波次间期定义为maskinwave对应期间的补集；
- maskbeginwave：疫情波次早期阶段的平均口罩使用情况，早期阶段定义为该波次持续时间的第一个十分位；
- maskpeakwave：疫情波次峰值附近的平均口罩使用情况，定义为从该波次新冠肺炎死亡率最高日开始的一个十分位。

指标 `maskall` 与 Tausk & Spira (2025a) 中使用的指标基本完全相同（唯一的区别是，该研究在计算平均值之前，剔除了大流行初期的零值）。我们假设死亡率对口罩使用的影响在时间上是局部的，即某一时间段的口罩使用只能受到近期死亡率的影响。在这一假设下，指标 `maskinterwave` 只能受到以下因素的影响

波次间期附近发生的死亡率，而这一死亡率与波次中期的死亡率相比非常低。因此，在整个分析期间，可能影响 maskinterwave 的死亡率部分在新冠肺炎死亡率中所占比重很小，这意味着反向因果关系在整个期间内至多只会对 maskinterwave 与新冠肺炎死亡率之间相关性的形成起到较小作用。相比之下，反向因果关系很可能对新冠肺炎死亡率与指标 maskall、maskinwave 和 maskpeakwave 之间相关性的形成起到重大作用。指标 maskbeginwave 处于 maskinterwave 与这些指标之间：与涉及 maskinterwave 的相关性相比，反向因果关系对涉及 maskbeginwave 的相关性的影响可能更大，但小于对涉及其他指标的相关性的影响。

在定义了相关的口罩使用指标后，我们通过计算两两之间的皮尔逊相关系数，考察了这些指标之间的相互关系。

表4. 口罩使用指标之间的皮尔逊相关及相应的95%置信区间。

指标	相关性
maskall 与 maskinterwave	$0.95\ [0.92, 1]$
maskall 与 maskinwave	$0.99\ [0.99, 1]$
maskall 和 maskbeginwave	$0.94\ [0.90, 1]$
maskall 和 maskpeakwave	$0.94\ [0.91, 1]$
maskinterwave 和 maskinwave	$0.93\ [0.89, 1]$
maskinterwave 和 maskbeginwave	$0.92\ [0.87, 1]$
maskinterwave 和 maskpeakwave	$0.86\ [0.79, 0.98]$
maskinwave 和 maskbeginwave	$0.96\ [0.93, 0.99]$
maskinwave 和 maskpeakwave	$0.96\ [0.94, 1]$
maskbeginwave 和 maskpeakwave	$0.92\ [0.87, 0.98]$

表 4 显示，成对的皮尔逊相关系数接近 1，这表明这些指标大致是彼此的仿射递增变换（即形如 $y = ax + b$ 且 $a > 0$ 的变换）。简单来说，这些不同的口罩使用量度以相似的方式对国家进行排序，而从一个指标切换到另一个指标，本质上只是尺度的变化。

皮尔逊相关系数的高值也意味着，就多元回归分析而言，这些指标在很大程度上可以相互替代。在我们的原始文章中，我们进行了多元回归分析，以年龄调整后的超额死亡率作为结果变量，并以疫苗接种率、口罩使用比例、人类发展指数以及心血管疾病死亡率和预期寿命的主成分作为解释变量（见 Tausk & Spira (2025a) 中的表3）；

在敏感性分析中，我们还检验了其他回归模型，相关结果报告于该文章的补充材料 2 中）。表5 报告了我们原始文章中所使用的多元回归结果，其中口罩变量依次由上述五个指标中的每一个替代。对于每个口罩指数，均列出了口罩变量的标准化回归系数及其相应的95%置信区间和 p 值。正如预期，各口罩使用指标的结果相似。

表 5. 复现的多元回归分析来自Tausk&Spira （2025a），

将口罩变量替换为五个口罩使用指数中的每一个。

口罩指数	标准化系数 [置信区间]	p 值
maskall	0.51 [0.23, 0.79]	0.001
maskinterwave	0.44 [0.16, 0.72]	0.005
maskinwave	0.51 [0.23, 0.78]	0.001
maskbeginwave	0.51 [0.25, 0.78]	0.001
maskpeakwave	0.52 [0.28, 0.76]	0.0002

基于上述原因，如果Tausk&Spira （2025a）所报告的口罩使用与死亡率之间的关联是由反向因果关系驱动的，则一旦将多元回归中的口罩变量替换为`maskinterwave`指标，该关联应基本消失。此外，人们会预期涉及`maskbeginwave` 的关联应明显弱于超额死亡率与总体口罩使用之间的原始关联。然而，我们观察到，使用`maskinterwave` 指标时，标准化系数仅略有减小，且 p 值仍然很低。此外，将`maskall` 替换为`maskbeginwave` 几乎不会使结果发生任何变化。这表明，所观察到的口罩使用与死亡率之间的关联中，至多只有一小部分可归因于反向因果关系。

接下来，我们寻找反向因果关系的直接证据。为此，我们对各项口罩使用指标进行了额外比较，并考察这些指标之间的关系如何随COVID-19死亡率而变化。首先，我们将maskbeginwave与maskpeakwave进行比较。在本分析中，我们将各波次分开考虑，因此共有 $3 \times 24 = 72$个数据点（24个国家各自的三个波次）。在此研究场景中，皮尔逊相关系数为0.85（95%置信区间 $[0.80, 0.93]$）。指标maskpeakwave大于maskbeginwave（$p < 0.0001$，单侧配对 Wilcoxon 检验），即波次峰值附近的口罩使用高于波次起始附近。

maskbeginwave 与 maskpeakwave 之间较高的皮尔逊相关系数意味着，这两个变量的散点图可以由相应的回归线很好地近似。

在初步近似中忽略回归线的截距后，这一观察结果促使我们用 maskpeakwave与 maskbeginwave的比值来表示疫情波次对口罩使用的影响，该比值可解释为回归线斜率的估计值。该比值的中位数为 1。52 ，第一和第三四分位数分别为 1。19 和 2。12。图1 展示了比值 maskpeakwave/maskbeginwave 与相应波次期间每百万人平均 COVID-19死亡率之间的散点图。斯皮尔曼相关略呈负值，且在统计学上不显著 −0.13 （95% 置信区间 $[-0.37, 0.09]$），这表明，一个波次期间佩戴口罩的增加不受该波次期间每百万人 COVID-19死亡率的影响。

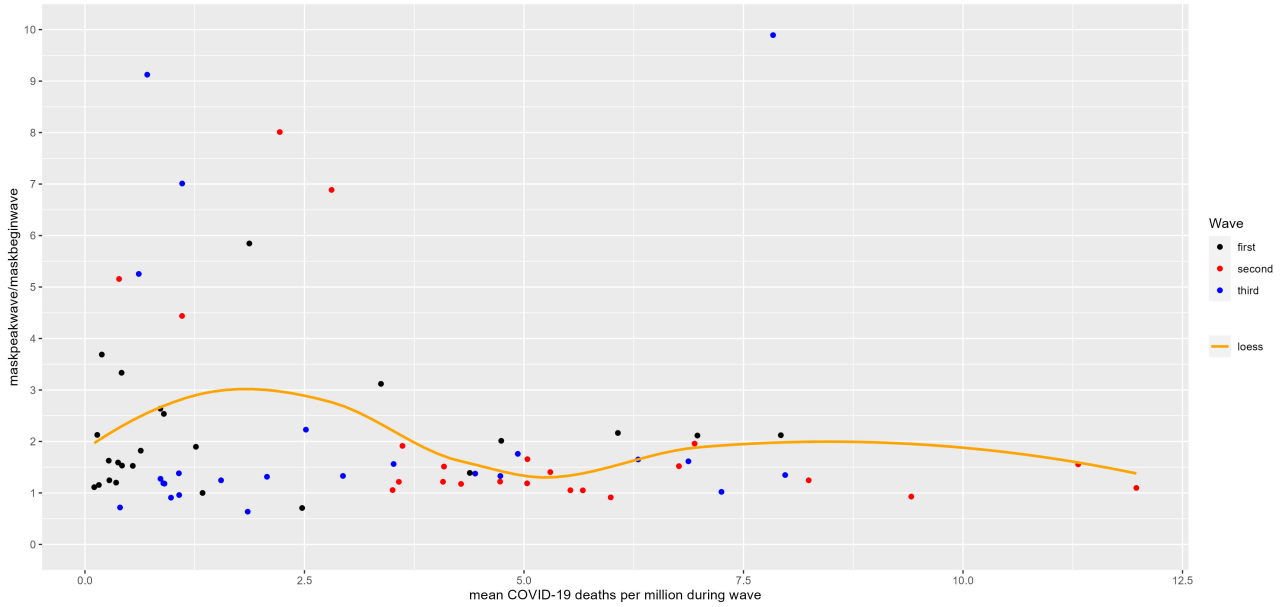

图1. 散点　图，显示 maskpeakwave/maskbeginwave 比值相对于平均新冠肺炎相应波次期间的每百万人口死亡率。

该问题也可通过另一种方式考察：将 maskpeakwave 对 maskbeginwave 以及该波次期间平均每百万人口COVID-19死亡率进行回归分析。此处，平均COVID-19死亡率的系数表示在控制 maskbeginwave 的情况下，COVID-19死亡率对 maskpeakwave 的效应。标准化系数为 0.03，95% 置信区间为 $[-0.10, 0.15]$ （p = 0.69）。这提供的证据表明，一旦考虑了基线口罩使用，某一波次期间佩戴口罩的增加并不受该波次COVID-19死亡率高低的影响。

通过比较maskinwave 与maskinterwave （此处每个国家计一个数据点），可以进行类似的分析。指数maskinwave 大于maskinterwave （p <0。0001 ，单侧配对 Wilcoxon 检验）。比值 maskinwave/maskinterwave的中位数为1。94，第一和第三四分位数分别为1。59 和2。23。图 2 显示了散点图

该图绘制了整个分析期间maskinwave/maskinterwave与每百万人新冠肺炎死亡人数的关系。斯皮尔曼相关为负，但无统计学显著性 -0.25（95%置信区间 $[-0.81, 0.17]$）。以maskinterwave和每百万人新冠肺炎总死亡率为自变量对maskinwave进行回归分析，得到死亡率变量的标准化系数为0。11，其95%置信区间为 $[-0.06, 0.28]$ （$p = 0.20$）。因此，几乎没有证据表明，疫情波次期间佩戴口罩的增加会受到新冠肺炎死亡率水平的影响。

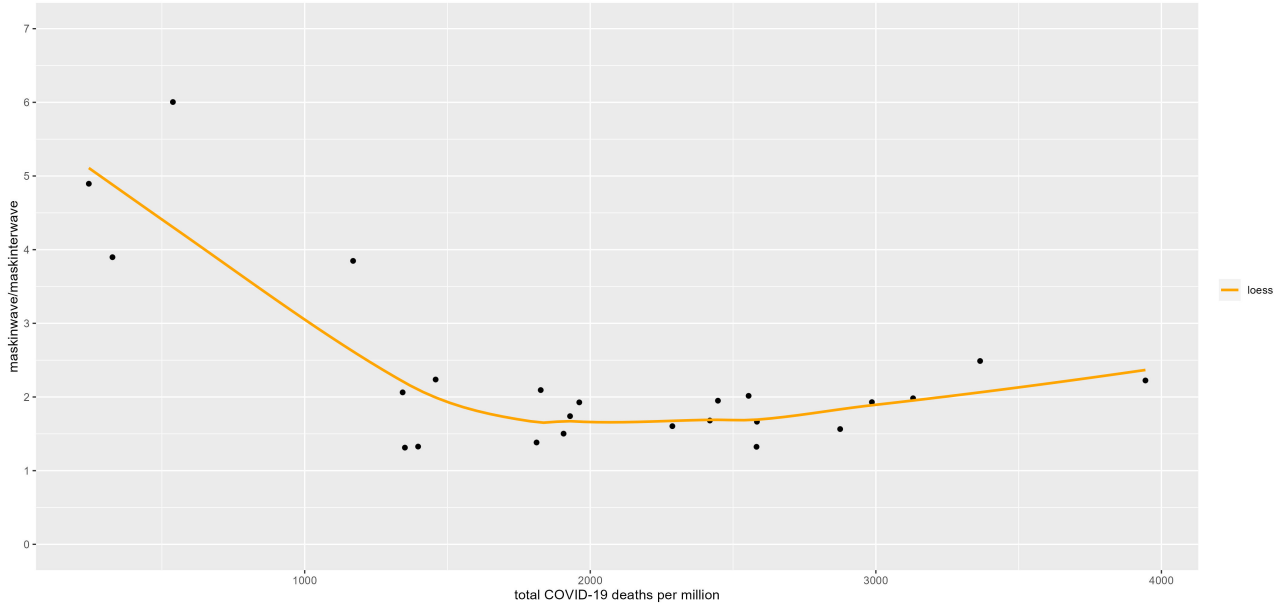

图2. maskinwave/maskinterwave 比率与整个分析期内 COVID-19 总死亡率的散点图

每百万人死亡率。

总之，尽管在大流行的关键时期，口罩使用确实会增加——疫情波次期间高于波次之间，且在波峰附近高于波次初期——但跨国家和跨波次的比较表明，这种增幅的大小不太可能与新冠肺炎死亡率水平有关。此外，一项考察2020—2021年冬季第二波新冠疫情期间口罩使用情况的既往研究报告称，平均口罩使用率与新冠肺炎死亡率呈正相关（Spira，2022年）。关键在于，第二波疫情开始时，口罩强制令及其执行措施早已实施数月，并于2020年上半年开始推行。这些措施在法国、德国、意大利、葡萄牙和西班牙等国家得到严格执行，而在斯堪的纳维亚国家，执行力度则相对较弱。因此，在冬季波次期间感染和死亡率随后上升之前，口罩使用就已十分普遍，而非因疫情上升才开始。总而言之，这些发现有力地表明

反向因果关系无法解释我们在原始文章中报告的口罩使用与死亡率之间的正相关关系。

3 错误表征 不实陈述 ks

Cerqueira-Silva等人(2026年) 指出

> 大量证据表明，口罩，尤其是正确使用的高等级口罩，是抵御病毒传播的有效屏障；或者至少不会增加感染或死亡率。[6–11]

支持这一说法的参考文献之一是关于非药物干预的Cochrane综述（参考文献编号8）。然而，Cochrane作者明确指出：

> 随机对照试验的汇总结果并未显示，使用医用外科口罩能够明确降低呼吸道病毒感染。在医护人员日常护理中，为降低呼吸道病毒感染风险，使用医用外科口罩与使用*N95-P2*呼吸器相比，并无明确差异。

因此，目前尚不清楚这篇综述为何能被引作口罩有效性的证据。此外，2023年综述中关于口罩的结论，在定性上与2020年末发表的上一版Cochrane综述（Jefferson 等人，2020年）完全相同，后者得出了实质上相同的评估。鉴于Cochrane综述在国际上被公认为评估医疗干预效果的高质量证据金标准，因此不难理解为何Cerqueira-Silva 等人(2026年)试图将一篇Cochrane综述纳入据称支持其论点的研究之中。然而，Cerqueira-Silva 等人(2026年)将这篇Cochrane综述描述为支持使用口罩减少病毒传播，根本上歪曲了其结论，从而误导读者，使其无法了解证据的实际情况。此外，他们在句子后半部分声称口罩显然不会增加感染，这一说法也并非完全属实。Cochrane系统综述（Jefferson 等人，2023）分析的三项结局，其置信区间均包含高于1.0的数值。具体而言，合并风险比（[95%，CI]）分别为：流感或COVID样疾病为0.95（[0.84,1.09]），实验室确诊的流感或SARS-CoV-2感染为1.01（[0.72,1.42]），以及实验室确诊的其他呼吸道病毒感染为0.58（[0.25,1.31]）。此外，纳入Cochrane综述的若干随机对照试验报告称，口罩组的感染率更高；也就是说，各研究中的效应方向并不一致（见表1）。

3 方法

在需要判断变量之间是否近似存在仿射变换关系时，采用皮尔逊相关系数进行两两比较 $y = ax + b$。否则，采用斯皮尔曼相关。在所有情况下，均使用基本自助法计算置信区间（从而避免对二元正态性的假设）。对于多元回归，使用 betaDelta R 软件包计算标准化回归系数的置信区间（假设采用多元正态近似），并针对非标准化系数计算 p 值（因为标准化系数与非标准化系数的原假设相同）。

下文将详细说明，为了对第 2 小节进行反向因果分析，如何定义疫情波次。为了确定各国疫情波次的开始和结束时间，我们首先将2020年02月04日至2021年12月31日的分析时段划分为三个阶段。所有国家的阶段划分均相同。阶段划分步骤如下：

(i) 将每百万人新冠肺炎死亡人数除以整个分析时段内每百万人新冠肺炎死亡人数的总和，对曲线进行归一化；(ii) 将所有国家归一化后的每百万人新冠肺炎死亡人数相加；(iii) 将步骤 (ii) 所得曲线的局部极小值定义为各阶段的边界。

归一化可确保死亡率较低的国家在求和时也获得不可忽略的权重。图 3 展示了所有国家经归一化的每百万人新冠肺炎死亡人数，并将分析时段划分为三个阶段。第一阶段为 2020年/02/04 至 2020年/08/02，第二阶段为 2020年/08/03 至 2021年/07/11，第三阶段为 2021年/07/12 至 2021年/12/31。

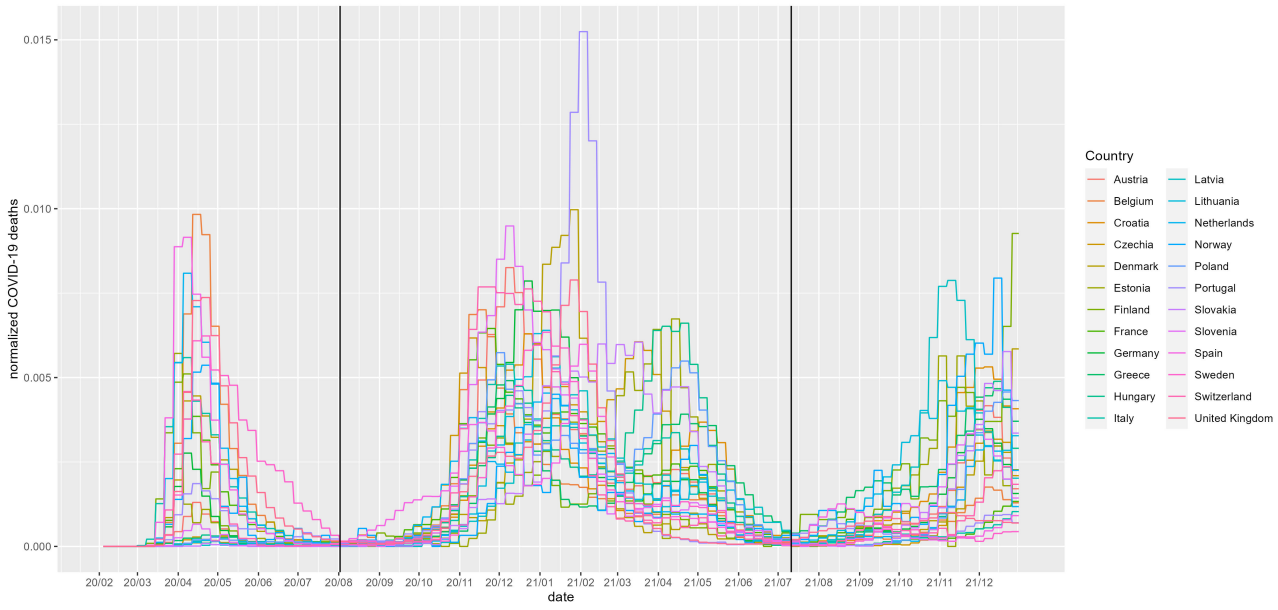

图 3. 所有24个国家每百万人新冠肺炎死亡人数的标准化值。竖线表示各疫情阶段的分界线。

定义各阶段后，针对每个国家的每个阶段，将波次定义为该阶段内包含至少 99% 该阶段新冠肺炎死亡人数的最短时间区间。此定义借鉴了最高后验密度区间的概念，该区间常在贝叶斯统计学中用作可信区间。展示疫情波次以及波次内与指标`maskbeginwave`和 `maskpeakwave`对应时段的图表见输出文件 `Waves.pdf` 和 `WavesMarked.pdf` ，这些文件可在 GitHub 仓库（Tausk，2026年）中获取。用于第 2 小节反向因果分析的所有 R代码、数据集和输出文件也可在该仓库中获取。第1小节的分析采用了作者在（Cerqueira-Silva等人，2026年）中提供的代码，并在 Cerqueira-Silva （2025年）基础上作了显而易见的调整。

参考文献

Jason Abaluck、Laura H. Kwong、Ashley Styczynski、Ashraful Haque、Md. Alamgir Kabir 等。2022年。社区口罩佩戴对新冠肺炎的影响：孟加拉国的一项整群随机试验。《科学》，**375**(6577)。

Allison E. Aiello, Vanessa Perez, Rebecca M. Coulborn, Brian M. Davis, Monica Uddin 等。2012年。年轻成年人中的口罩、手卫生与流感：一项随机干预试验。*PLoS ONE*，**7**(1)，e29744。

Mohammad Alfelali, Elizabeth A. Haworth, Osamah Barasheed, Al-Mamoon Badahdah, Hamid Bokhary 等。2020年。口罩对朝觐者病毒性呼吸道感染的防护作用：一项颇具挑战的整群随机试验。*PLOS ONE*，**15(10)**，e0240287。

Osamah Barasheed、Nedal Almasri、Al-Mamoon Badahdah、Leon Heron、Janette Taylor 等。2014年。测试外科口罩在预防2011年澳大利亚朝觐者中类流感疾病传播方面有效性的随机对照试验先导研究。*Infectious Disorders - Drug Targets*，**14(2)**，110–116。

Henning Bundgaard、Johan Skov Bundgaard、Daniel Emil Tadeusz Raaschou-Pedersen、Christian von Buchwald、Tobias Todsen 等。2020年。在其他公共卫生措施的基础上增加口罩建议，以预防丹麦口罩佩戴者感染SARS-CoV-2的效果：一项随机对照试验。*Ann IntMed*，**174(3)**，335–343。

Laetitia Canini、Laurent Andréoletti、Pascal Ferrari、Romina D'Angelo、Thierry Blanchon 等。2010年。预防家庭内流感传播的外科口罩：一项整群随机试验。*PLoS ONE*，**5(11)**，e13998。

Thiago Cerqueira-Silva。2025年。口罩再分析。网址 https://github.com/csthiago/reanalysis_口罩，访问日期：2026年2月21日。

Thiago Cerqueira-Silva、Felipe Argolo、Gabriel Gonçalves da Costa、Felipe Nogueira Barbara、Pedro Hallal 等。2026年。揭示偏倚：为何所谓的口罩与超额死亡关联站不住脚。*The Lancet Regional Health -Americas*,**55**（3月），101360。

Benjamin J Cowling、Rita OP Fung、Calvin KY Cheng、Vicky J Fang、Kwok Hung Chan 等。2008年。预防家庭中流感传播的非药物干预整群随机试验初步结果。*PloSone*,**3**（5），e2101。

Benjamin J Cowling、Kwok-Hung Chan、Vicky J Fang、Calvin KY Cheng、Rita OP Fung 等。2009。口罩和手卫生对预防家庭中流感传播的作用：一项整群随机试验。*Annals of internal medicine*,**151**（7），437–446。

Manuela Fritz. 2022年. 波次接波次：利用德国空间面板数据确定COVID-19感染与死亡的时间滞后. *Journal of SpatialEconometrics*, **3(1)**.

Kosuke Imai, In Song Kim. 2020年. 论双向固定效应回归模型在面板数据因果推断中的应用. *Political Analysis*, **29**(3), 405–415.

Joshua L. Jacobs、Sachiko Ohde、Osamu Takahashi、Yasuharu Tokuda、Fumio Omata 等。2009年。使用外科口罩降低日本医护人员普通感冒发病率：一项随机对照试验。美国感染控制杂志，**37**(5)，417–419。

Tom Jefferson、Chris B Del Mar、Liz Dooley、Eliana Ferroni、Lubna A Al-Ansary 等。2020年。通过物理干预阻断或减少呼吸道病毒的传播。*Cochrane*系统评价数据库，**2020年**(11)。

Tom Jefferson、Liz Dooley、Eliana Ferroni、Lubna A Al-Ansary、Mieke L van Driel 等。2023年。通过物理干预阻断或减少呼吸道病毒的传播。*Cochrane*系统评价数据库，**2023**(1)。

期刊_USP_大学。2026年。针对研究人员，认为口罩导致大流行期间死亡人数增加的说法毫无根据。网址 https://jornal.usp.br/ciencias/para-pesquisadores-ideia-de-que-mascaras-aumentaram-mortes-na-pandemia-nao-tem-fundam

Elaine L Larson、Yu-Hui Ferng、Jennifer Wong-McLoughlin、Shuang Wang、Michael Haber 等。2010。非药物干预对拥挤城市家庭中上呼吸道感染和流感的影响。公共卫生报告，**125**’(2)，178–191。

C Raina MacIntyre、Holly Seale、Tham Chi Dung、Nguyen Tran Hien、Phan Thi Nga 等。2015。比较医护人员使用布口罩与医用口罩的整群随机试验。*BMJ Open*, **5**(4)，e006577。

Chandini Raina MacIntyre、Yi Zhang、Abrar Ahmad Chughtai、Holly Seale、Daitao Zhang 等。2016。整群随机对照试验：考察呼吸道疾病患者使用医用口罩进行源头控制的效果。*BMJ Open*，**6**(12)，e012330。

Masks4All。2020年（4月）。哪些国家要求或建议在公共场所佩戴口罩？网址 https://masks4all.co/what-countries-require-masks-in-public/ （访问于2023-05-21）。

Line M. Nanque、Andreas M. Jensen、Arthur Diness、Sebastian Nielsen、Carlos Cabral 等。2024年。《在几内亚比绍城市地区分发本地生产的布制口罩对类新冠疾病和全因死亡率的影响——一项整群随机对照试验》。*PLOS*全球公共卫生，**4**(2)，e0002901。

Our World in Data。2024年。*Our World in Data* 的新冠肺炎（冠状病毒）数据。网址：https://github.com/owid/covid-19-data/tree/master/public/data，访问日期：2026年2月21日。

James M. Simmerman、Piyarat Suntarattiwong、Jens Levy、Richard G. Jarman、Suchada Kaewchana 等。2011年。泰国曼谷一项旨在通过洗手和佩戴口罩减少流感传播的家庭随机对照试验结果：洗手和佩戴口罩的家庭随机对照试验。《流感与其他呼吸道病毒》，5(4)，256–267。

Runar Barstad Solberg、Atle Fretheim、Ingeborg Hess Elgersma、Mette Fagernes、Bjørn Gunnar Iversen 等。2024年。在公共场所佩戴外科口罩对成人自我报告的呼吸道症状的个人防护效果：实用性随机优效性试验。《英国医学杂志》，7月，e078918。

Beny Spira。2022年。欧洲口罩佩戴依从性与新冠肺炎结局之间的相关性。《*Cureus*》，14(4月)，e24268。

Beny Spira。2026年。《揭示新冠肺炎大流行中的发表偏倚》。*MetaArxiv*，1月。

Thorsten Suess、Cornelius Remschmidt、Susanne B Schink、Brunhilde Schweiger、Andreas Nitsche 等。2012年。口罩和手卫生在预防家庭流感传播中的作用：一项整群随机试验的结果；德国柏林，2009—2011年。《*BMC*传染病》，12(1)。

Daniel Tausk。2026年。口罩反向因果分析。网址：https://github.com/danieltausk/MasksReverseCausalityAnalysis。

Daniel V. Tausk、Beny Spira。2025a。口罩使用是否与超额死亡率相关？来自24个欧洲国家的发现。*BMC*公共卫生，25(1)。

Daniel V. Tausk、Beny Spira。2025年。更正：口罩使用是否与超额死亡率相关？来自24个欧洲国家的发现。*BMC*公共卫生，25(1)。