

《疾病监测》审稿意见与作者答复

题目：应用 log-binomial 回归和 logistic 回归分析法定传染病报告质量影响因素

作者：许意清 付志智

——审稿专家意见与答复——

初审专家意见及作者修改说明：

1. "1.1 调查对象 采用多阶段分层抽样的方法在广西 14 个地级市各随机抽取 2 个县作为调查地区，分别抽取 1 家县级以上综合医院和专科医院作为调查对象。",这段描述与“2.1 基本情况 共抽查 86 家医疗机构，均实现网络直报，其中省级 4 家，市级 28 家，县级 54 家。”描述矛盾。选取的县，怎么又有省级和市级医院？

感谢您的宝贵意见，表述可能不准确，在调查对象部分进行了重新进行了描述：采用多阶段分层抽样的方法共抽取 86 家县级以上医疗机构为研究对象：随机抽取 2 家省级人民医院和妇幼保健院，共 4 家省级医疗机构；分别从 14 个地级市随机抽取 1 家人民医院和妇幼保健院，共 28 家市级医疗机构；分别从 14 个地级市随机抽取 2 个县（市、区），各随机抽取 1 家县级人民医院和妇幼保健院，由于其中 2 个县（市、区）尚未设立妇幼保健院，共抽取 54 家县级医疗机构。

2."2.2.3 偏倚（表 3）采用多因素线性回归分析造成偏倚的影响因素"与表 3 的描述不清晰。如自变量、因变量是什么？多重线性回归的模型假设是如何检验的？将统计分析方法及编码放入方法学里介绍一下。

感谢您的宝贵意见，修改稿中，增加了偏倚回归分析的自变量和因变量的描述。多重回归模型假设检验在回归方程结果输出的同时即给出了假设检验 t 值及对应的 p 值。方法部分对 log-binomial 回归的命令进行了简单描述。

3.log-binomial 回归用于事件率大于 10%没有问题，模型应用合适。但如果用 logistic 回归，将完整率和报告率改为“不完整率”和“未报告率”为结果变量，拟合 logistic 回归是否可以。请讨论。

感谢您的宝贵意见，修改稿中增加了将“不完整率”和“未报告率”为结果变量的，logistic 拟合结果与 logbinomial 回归分析拟合结果的差异比较。

4.统计方法先写两个模型的原理、区别，再写模型拟合的策略和设置的检验水准，最后写应用拟合的软件及模块。

感谢您的宝贵意见，修改稿对统计方法部分按照您的意见进行了修改。

5.注意 STATA 是否有版权。如没有，可填一位有版权的单位作者。

感谢您的宝贵意见，本文作者没有 Stata 版权，作者使用免费软件 R v3.3.3 进行了拟合结果一致，所以修改稿软件替换为 R 语言。

6、本文中心思想是推广 log-binomial 模型，而不是要进行真正的数据分析，所

使用的数据可认为是举例，因此不对抽样等问题进行评论；

感谢您的宝贵意见，本文的首要中心思想是分析传染病报告质量的影响因素，这与作者的工作职责相关，次要中心思想是推广 log-binomial 回归模型，所以修改稿中，仍保留了较大篇幅的抽样部分以及报告质量影响因素方面的讨论。

7、既是推广，应该详细论述该模型，从理论和实践两方面论述模型优劣，本文仅仅只有公式不妥。

感谢您的宝贵意见，修改稿在方法部分和讨论部分增加了模型优劣的讨论。

8、本文是以 logistic 模型与 log-binomial 模型估计的关联强度大小进行模型优劣比较的，估得低就差吗？

感谢您的宝贵意见，模型估计值的高低并不等同于优劣，本文是从实际结果解释及应用的角度分析，log-binomial 回归估计的 PR 值才是通常意义来讲，解释的暴露值增加 1 个单位，结局危险度增加的比例的含义，OR 值是 PR 或 RR 真实值的估计值，所以 PR 被认为是金标准，而依据模型拟合结果，logistic 回归估计的 OR 值高于 PR，所以文章说高估。

9、OR 与 PR 两者不同，不能要求他们数值一致，公式需要推一推，看看还差多少？

感谢您的宝贵意见，OR 和 PR 两者具体的计算公式不同，但是大多数研究中 OR 值被解释为 PR 或者 RR 的含义，文章仅在统计方法部分和偏倚分析部分进行比较，由于所学专业为流行病学专业，尚不具备更为深入推导演算的能力。

复审专家意见及作者修改说明：

1、按照现有题目，是使用 log-b 模型进行数据分析，没有任何必要进行与逻辑回归的比较；

感谢您的宝贵意见！修改文章题目为《应用 log-binomial 回归和 logistic 回归分析法定传染病报告质量影响因素》。

2、可以同时列出两个模型的结果，但使用这个结果就说高估低估，继而进行模型比较，至少论据不充分；

感谢您的宝贵意见！RR/PR 值为直接关联强度，logistic 回归估计 OR 值本身没有错，然而大多数研究，解释其含义为患病率比/危险率比（自变量增加一个单位，患病或结局发生率增加的比例）。所以，本文作者参照相关学者做法，以 RR/PR 为真实值出发，比较 OR 与其大小的关系。修稿中增加相关参考文献加以佐证。

3、从 OR 是 RR 这个估计值这个角度来说，OR 是没有 RR 好，这需要论证吗？反之 RR 也可以是 OR 的估计值，从这个角度又怎么说？

感谢您的宝贵意见！RR/PR 和 OR 均为关联强度的估计值，但其含义不一样。然而在不少研究中 OR 值被解释为 RR/PR 值的含义。因此才有 OR 是 RR 估计值的这一观点。这一认识是建立在先验知识和学者们提出的基础上，并非自定义的以 RR/PR 为真值，而 OR 为估计值。所以反过来假设从实际意义的角度解释不成立，修改稿中增加了相关参考文献加以佐证。

文章修稿主要参照专家和编辑意见，增加了关于 OR 值为 RR 值估计值观点表述，

并引用了相关参考文献加以佐证。

-----定稿会意见与答复-----

定稿会意见：

本文经这次修改后，基本达到要求，可以发表，谢谢！