



疾病

预测

疾病的发生

目前,由于医院、诊所和有关行业如药房正在越来越多地使用电子记录,大量有关疾病发生率的资料被整理成数据被送到卫生部门。过去整理大量医疗文档需要花费数天或数周的时间,而在今天只需几小时就能收集到大量的有关患者情况的资料。日益增多的大量信息让人产生了对早期预测疾病爆发更高的期望,希望疾病能被尽快地引起关注和得到控制。但实际的情况是,说起来容易做起来难。

利用这些资料的用意之一是提供有参考价值的结果。2005年3月发表在 *PLoS Medicine* 的一篇文章介绍了一种新方法,该方法只要求研究者知道在一个指定的时期和地区疾病发生的实际数字。结果统计模型仅对某一种疾病的发生得出最有限的假设,并对疾病报告中发生的时间和地理的变量作校正。

这种方法较单因素统计方法提供了更详细的结果,后者纯粹根据时间和地理追踪疾病爆发。尤其是新的模式不需要有关当地人群和人群患病的相对危险性的信息,比如利用某一地区的一组人群如易患特殊疾病的婴儿和老年人是否高于平均比例。

透视疾病的窗口

PLoS Medicine 发表的文章要点是扫描统计量 (scan statistics) 的原理,即通过与一个较大地区或一段较长时期内收集的信息所确定的可能性比较,提出在一段严格定义的时间和空间内出现过量病例报告的可能性。扫描统计量并不是新的方法,但是 *PLoS Medicine* 的文章为它添加了一个步骤,即在不同时间和空间的样本中计算疾病爆发的可能性。在区域监测的情况下,突发性的局部爆发可能被隐藏在地区收集的資料中,而如果同时考虑到时间因素那么这种事件被揭示的可能性就会更大一些。以空间-时间排列的扫描统计量也许能成为一种优先的方法来表示疾病发生,如癌症,实际和预测的报告数量可以特殊的“视窗”方法计算。

这些视窗可以被看作是在所涉及的地理区域内一套数千或甚至数百万重叠的“圆筒”(cylinder),每一个视窗都会随着它所透视到的领域大小和时间跨度而变化。哈佛大学救护和预防系的副教授 Martin Kulldorff 是 *PLoS Medicine* 的文章的作者之一,他通过三维方式用圆筒作为一种审视资料样本的方法,在三维模式中,x轴和y轴代表被监测的地区,z轴代表时间。随着时间推移,样本叠加,通过数学分析每一个圆筒用一种可能性功能来比较其预测的和实际观察到的例数,这样就能挑选出观察例数意外增多的日期和地方。

这一新模式的检测是与纽约市健康和精神卫生部共同试验的,该部门是负责收集资料来估计疾病传播的主要机构之一。20世纪90年代后期该市发起了一个周密的综合监测计划,通过追踪救护车报告、急诊室就诊以及药房销售情况来发现相似病例聚集人群,这可能是一种疾病爆发的信号。

Kulldorff说研究者采取几个步骤来处理这个项目的计算负荷,圆筒底部依次由纽约市183个邮政编码的几种组合而成,半径为0~5公里。研究者定义的每一个圆筒高度为7天(理由是如果一次爆发已经持续超过一周,诊所的医生或实验室可能已经注意到疾病的爆发。)

但是尽管每个圆筒覆盖的地区相同,特定的日期会有所变化。例如,在一个月的30天中,第一次统计分析的圆筒高度根据第1天到第7天的资料而定义,接下来分析的圆筒高度根据第2天到第8天的资料而定义,依此类推,通过这扇移动的视窗可能发现在严格限定的时间和空间内发生的变化。

因此打个比方说,研究小组可能在一种疾病开始爆发的第七天捕捉到信息,而健康当局可能在更长天数内还未能确定。这个早期信号将会敦促官员们更快地查明情况并可能更有效地控制任何爆发。为了能前瞻性地保持这样的信号,统计分析也会参考前30天的情况,这样任何持续更长时间的趋势或变异都能与7天内的视窗所发现的情况作比较。

对该方法的评估

对该方法的评估是从腹泻病的历史报告开始的,报告来自2001年11月至2002年11月健康与精神卫生部每天从急诊室收集的资料。文档根据非特异性情况或症状如“腹泻”或“流感样”对个例进行分类,包括患者家庭的邮政编码和他们治疗的地点。

通过该模式产生了五组最具统计学意义的资料,其中四组资料与研究时期整个城市范围内轮状病毒、诺如病毒和流感的爆发相关,这意味着如果当时得到这个信息的话,它就可以被用于早期预警这些疾病的爆发。虽然该系统会产生一些与爆发不相关的信号,但是会比较少。

研究小组必须解决一个问题,即现实中的人们在一周特定时间内工作,而诊所和药房可能在固定的时间营业。这导致变异的产生,可能传递出假的信号,诸如星期天该地区其他大多数药房关闭时,营业药房的药物销量就会很高。

在计算疾病爆发的随机可能性时,可以通过将一周的每一天都考虑进去而避免这种偏差,打个比方,如果资料显示某一疾病在星期天聚集,那么评估可能的结果应该是相对于其它星期天而不是一周中的任何一天。

对于许多系统更成问题的是整个地区的报告在时间和地点上的不一致。“这是整个领域所面临的挑战,”Kulldorff注意到许多复杂的电子系统报告延迟的现象后指出,“这些挑战之一是报告要及时,否则部分资料可能被丢失。你可能仅仅获得90%的资料,其余部分几天以后才能得到。”不过他认为这些由空间-时间排列扫描统计量得出的初始结果足以保证更深入的研究。

Kulldorff想用其它来源的信息来评价该模式的性能。“现在还不清楚究竟哪些资料实际上最适合用来预测感染性疾病爆发—急救、求诊、实验室检测结果等等,”他说。是否应限制该模式仅用于追踪自然疾病爆发也还不清楚。他建议可以成功地将该模式应用于其他公共卫生如抗生素耐药细菌的传播,以及其他完全不同的领域如犯罪学、生态学或工程学。

使用软件

为了推广该模式的应用,它已经被制作成SaTScan™软件,自从几年前Kulldorff开始在健康和公共卫生部工作之前他就已经在研发SaTScan™。该程序可以在<http://www.satscan.org>免费下载。

SaTScan的目的是确定与疾病有关资料反常的聚集—也就是说聚集在统计学上与健康信息的常规变异是不同的,这些一天天不断集中的信息来自一个特殊区域。这种反常具有时间形式(在短期内发生)或地理形式(在一个小的区域发生)。

Kulldorff在哈佛大学的同事之一已经对SaTScan用于评价疾病爆发了述评,包括可能是由于故意的行为所导致的爆发。Katherine Yih是研究小组成员之一,从事国立生物恐怖综合监测鉴定计划(National Bioterrorism Syndromic Surveillance Demonstration Program)的工作,该计划是由美国疾病预防控制中心与覆盖不同州的2千万人群的保健组织共同合作发起的。这项计划的目标是利用保健计划和执业机构的资料来监测局部爆发,并负责快速的公共卫生随访。

该计划的研究者检查求诊者的记录—它显示了由诊所提供的医疗护理记录,而不是医院的记录—这种医疗护理是按照象纽约市项目那样由邮

政编码组合起来的。该系统每个晚上寻找聚集的疾病,记录以24小时为一个周期加以汇总,它可能提示统计学意义上的爆发,那样就可以向波及地区的官员们发出警告。

在2004年9月24日由CDC出版的《发病率和死亡率周报》(*Morbidity and Mortality Weekly Report*)的一篇文章中,Yih和她的同事报道在科罗拉多、德克萨斯和马塞诸塞3个州发现呼吸道疾病异常聚集发病与严重的流感爆发有关。他们现在正在利用明尼苏达州卫生署所确定的自然爆发的资料,把注意转向评价监测系统是否能检测到胃肠道疾病自然发生的爆发,他们同时也在进行模拟操作以确定该监测系统能否敏感地检测生物恐怖行动。

虽然Yih认可这种方法,但是她注意到该方法仍有一些局限。“收集资料还有其他目的,”她说。例如她最近的一些研究是从明尼苏达一个组织取得信息,由该组织的一名预检护士通过电话提供保健护理信息的。提供这些服务的机构最优先考虑的是试图将一个患者自我报告的症状与一组容易识别的疾病联系起来,而不是仅仅记录患者报告的症状。如果给一名打电话来询问的人提供咨询的话这种服务有其意义,但像Yih这样的研究者就可能要花费很多时间去翻阅这些组织记录的资料去查找病人的症状。

“这种方式的好处在于你不必要求这些保健组织的人员花很多时间按照你的监测格式或者数据格式将资料输入,他们可在常规病人护理过程中输入资料。”她说。她又补充道,相同的症状,数据可能并不同一格式,也可能报告不够及时,所以对于确定疾病爆发并不是最为适用。而且,从部分人群中获得的资料是否足以发出警报也难以确定,例如,一项可能正在支持这些资料的保健计划所覆盖的一个大都市人群可能只有5%~15%。

Kulldorff和他的同事引用了对他们的疾病监测策略造成影响的事情作为补充说明,该策略是在高度局限的水平检测疾病特殊的爆发,而不是在一个广泛监测区域同时检测到的。例如,如果一种感染因子在地铁中发生传播,被感染的人群未必是附近的居民,他们可以是住在任何地方的人,他们也不可能都涌向同一个急诊室求诊。

而且他们采用的方法是视病情是否严重到需送急诊室,因而较难检测到不太严重疾病的小规模爆发。因此,他们不以这种方法作为唯一的选择。他们建议,鉴于每种检测系统各有其利弊,有效而全面的监测是将不同的系统相结合。

Yih建议用现行的SaTScan和相似系统进行评估,以确定用于监测计划的疾病资料所消耗的资源,以及公共卫生官员为跟踪这些由项目发出的警报信号而投入的资源是否合理。最理想的是这些能信息使卫生官员的工作顺利进行,保证他们所作的努力更加有效。

这种保证就是利用扫描统计量系统给这些官员们提供真正的服务。Emory大学的生物统计学家Lance Waller如此说,他评估了疾病监测模式所使用的方法。“这些数据本来被分别收集并储存在各自的电脑中,而现在可以被汇总起来,”他说,“对公共卫生工作人员来说,能够从多家医院获得多项健康记录,并从多家药房获知药物销售情况是一件新生事物。这种方法非常好。”

纽约市也认为这个系统是件好事。从2003年起扫描软件已经被应用于日常的城市监测系统,对市区38家医院急诊室报告的呼吸道症状、发热、流感和腹泻进行监测。很快这个系统将能在日常的检索中查出流行病的爆发先兆,包括有高度传染性的布鲁斯病毒。

—Tim Loughheed

译自 *EHP* 114:A610-A613 (2006)

参 考 读 物

Kleinman KP, Abrams AM, Kulldorff M, Platt R. 2005. A model-adjusted space-time scan statistic with an application to syndromic surveillance. *Epidemiol Infect* 133(3):409-419.

Kulldorff M, Heffernan R, Hartman J, Assunção R, Mostashari F. 2005. A space-time permutation scan statistic for disease outbreak detection. *PLoS Med* 2(3):e59.

National Bioterrorism Syndromic Surveillance Demonstration Project homepage, <http://btsurveillance.org/>.

SaTScan homepage, <http://www.satscan.org/>.