



## 一种评估孤立性肺结节患者肺癌验前概率的临床模型

*Michael K. Gould, MD, MS, FCCP; Lakshmi Ananth, MS; and Paul G. Barnett, PhD; for the Veterans Affairs SNAP Cooperative Study Group*

**背景** 在孤立性肺结节 (SPN) 患者中评估恶性肿瘤的临床概率有助于选择和解释随后的诊断方法。

**方法** 我们对伴有SPN的375例来自各地的退伍军人运用多因素逻辑回归分析识别恶性肿瘤的独立临床预测因子，并建立简约的临床预测模型以评估恶性肿瘤的验前概率。资料来源于退伍军人事务部的管理数据库和最近完成的协作研究，据此评估PET扫描诊断SPN的精确性。

**结果** 本研究中患者的平均年龄为 (65.9 ± 10.7) 岁。恶性SPN的患病率是54%。大多数患者目前吸烟 ( $n = 177$ ) 或既往吸烟 ( $n = 177$ )。恶性SPN的独立预测因子包括吸烟史 (OR 7.9, 95%CI 2.6 ~ 23.6)、年老 (OR 每增加 10 y 为 2.2, 95%CI 1.7 ~ 2.8)、大结节直径 (OR 每增加 1 mm 为 1.1, 95%CI 1.1 ~ 1.2) 和戒烟时间 (OR 每增加 10 y 为 0.6, 95%CI 0.5 ~ 0.7)。模型的精确度甚佳 (受试者工作特征曲线下面积为 0.79, 95%CI 0.74 ~ 0.84)，恶性SPN的预测概率与观察到的频率非常一致。

**结论** 我们的预测模型可用于评估SPN患者恶性肿瘤的验前概率，因此当选择诸如PET扫描等诊断方法并解释相应检查结果时有助于作出临床决策。

**关键词** 硬币病变 (coin lesion); 诊断 (diagnosis); 肺肿瘤 (lung neoplasms); 肺的 (pulmonary); 受试者工作特征曲线 (receiver operating characteristic curve); 统计模型 (statistical models); 退伍军人 (veterans)

**缩略语** AUC = area under the curve; CI = confidence interval; FDG = 18F-fluorodeoxyglucose; OR = odds ratio; PET = positron emission tomography; ROC = receiver operating characteristic; SPN = solitary pulmonary nodule; VA = Department of Veterans Affairs

孤立性肺结节 (SPN) 极为常见。对于肺结节患者的处理, 首要步骤是评估恶性肿瘤的临床验前概率。显然, 当验前概率高时, 须对其进行积极处理; 当验前概率低时, 将进行保守处理。大多数临床医师凭借临床经验和判断来评估验前概率, 部分医师则采用1个或多个已经建立的定量模型来评估验前概率<sup>[1-3]</sup>。验前概率可用于指导选择和解释随后的诊断方法, 例如<sup>18</sup>F-氟脱氧葡萄糖 (FDG)、PET 扫描和针吸活检等。

梅奥诊所的研究人员对SPN患者建立了广被引用的预测模型, 他们回顾性分析1984 - 1986年间419例新发肺结节患者的医疗记录和影像检查方法<sup>[2]</sup>。研究人员发现如下6个恶性肿瘤的独立预测因子: 年老、吸烟史、结节发现前至少有5 y胸腔外恶性肿瘤史、大结节直径、上叶病变和小锯齿状表现。模型的精确度很好, 受试者工作特征 (ROC) 曲线下面积 (AUC) 平均为  $0.83 \pm 0.02$ 。最近发现, 该模型与荷兰单中心评估106例诊断不明的肺结节患者的回顾性队列研究<sup>[4]</sup>有相似的精确度 (AUC为0.79), 但是模型往往低估恶性肿瘤的概率, 特别是在预测概率的下部区域。

梅奥诊所模型还有其他的局限性。因为其建模

人群来自美国中西部一所三级医疗中心20 y前处理过的肺结节患者。研究人员排除了5 y内有肺癌史和胸腔外恶性肿瘤史的患者, 从而限制了模型的概化。此外, 恶性肿瘤的患病率相对较低 (23%), 12% 的患者没有确诊。因此我们采用最近完成的退伍军人事务部合作研究<sup>[5, 6]</sup>提供的资料 (CSP 027)<sup>18</sup>F-FDG PET 成像在SPN患者处理中的应用”建立新的模型, 旨在来自各地伴有SPN退伍军人中评估恶性肿瘤的验前概率。

## 材料和方 法

CSP 027为前瞻性研究, 在胸片结节直径为7 ~ 30 mm的退伍军人中, 比较CT扫描与FDG-PET扫描诊断SPN的精确度。关于研究方法已另有描述<sup>[5]</sup>。所有CSP 027研究参与者均签署了知情同意书, 研究方案获得所有10个研究中心的机构审查委员会认可。为了完成我们的分析, 我们接受放弃知情同意书的弃权书, 并得到斯坦福大学机构审查委员会的认可, 合并了退伍军人事务部的管理数据和CSP 027中未识别的数据。

### 参与者

1988年9月至2001年6月, 来自10个不同地方退伍军人事务部中心的532人入选了CSP 027研究 (安阿伯, MI; 亚特兰大, GA; 布法罗, NY; 达勒姆, NC; 印第安纳波利斯, IN; 麦迪逊, WI; 帕洛阿尔托, CA; 圣路易斯, MO; 西雅图, WA; 西部洛杉矶, CA)。CSP 027基本入选标准: 单个, 新发, 未经治疗, 胸片可见圆形或椭圆形不透明影以及结节直径为7 ~ 30 mm。排除标准: 年龄 < 21岁, 妊娠或哺乳期, 体重 > 350 ~ 400磅, 并发肺部感染, 6 mo内胸腔外科手术, 1 y内胸部放疗, 预期寿命 < 1 y。分析时, 我们排除了157名参与者, 因为他们没有合格的CT或PET扫描, 和(或) 根据我们的研究结果仍未明确SPN为恶性或良性。经外科手术或活检获

---

From the Veterans Affairs Palo Alto Health Care System (Dr. Gould), Palo Alto, CA; and the Veterans Affairs Health Economics Resource Center (Ms. Ananth and Dr. Barnett), Menlo Park, CA.

This study was supported by VA Cooperative Studies Program (CSP 027), “F-18 Fluorodeoxyglucose (FDG) Positron Emission Tomography (PET) Imaging in the Management of Patients with Solitary Pulmonary Nodules (SNAP).” Dr. Gould was supported by grant R01 CA117840-01A2 from the National Cancer Institute.

Correspondence to: Michael K. Gould, MD, MS, FCCP, Veterans Affairs Palo Alto Health Care System, 3801 Miranda Ave (111P), Palo Alto, CA 94304; e-mail: gould@stanford.edu

取结节、纵隔淋巴结或胸腔外临床可疑肿瘤转移灶的组织标本予病理检查，据此结果确立恶性诊断；经病理检查证实为良性病因学，或X线和临床上发现SPN至少稳定2y，据此确立良性诊断。

## 变量

CSP 027 研究收集的资料包括年龄、性别、结节直径、结节部位以及吸烟行为，后者又包括目前吸烟状况、吸烟年数、平均每天吸烟的包数和戒烟年数。尽管CSP 027 研究没有收集关于结节的小锯齿状或其他形态特征的信息，但读片者必须根据李克特5分量表(范围为良性~恶性)提供放射学诊断。所以，对于小锯齿状结节，我们以放射学诊断的恶性肿瘤代之，因为细刺结节恶变的可能性大约是良性的5倍<sup>[3, 6]</sup>。

我们从电子版退伍军人事务部数据库搜寻恶性肿瘤史的附注信息。通过比较医疗记录号和出生日期确认参与者的身份。复习住院及门诊患者的资料识别最近(如5y内)诊断为肺癌和胸腔外恶性肿瘤者。因为退伍军人事务部数据库的局限性，我们不能识别结节发现前已确诊为肿瘤至少5y的参与者。

## 统计分析

我们应用CSP 027研究和退伍军人事务部使用文档提供的数据以建立和内部验证验前概率模型，模型也许对伴有SPN的退伍军人和类似的老年群体特别有用<sup>[7]</sup>。我们以最终诊断为因变量，以下列因素为自变量，运用逐步逻辑回归建立了模型。这些因素包括：年龄、性别、吸烟史(从不或曾经)、吸烟量(包-y)、戒烟年数、肺癌史、胸腔外恶性肿瘤史、既往确诊为肺癌的时间、既往确诊为胸腔外恶性肿瘤的时间、结节大小、上叶病变、右肺病变以及胸片确诊的恶性肿瘤。应用自后淘汰变量法，经淘汰在0.05水平没有统计学意义的变量后，我们建立了最终的简约模型。当我们以更大的P值( $P = 0.10$ )作为排除标准，结果是相似的。所有临床上看似合理的交互作用均经检验，但没有

统计学意义，故未进入最终的模型。对于有统计学意义的预测因子，我们报告比值比(OR)和95%CI。

我们运用Hosmer-Lemeshow拟合优度统计数值( $P > 0.05$ )来评价模型的拟合度<sup>[8, 9]</sup>。我们应用最终的模型来计算每例研究参与者恶性肿瘤的估计概率，然后比较最终的诊断与恶性肿瘤的预测概率，构成ROC曲线。为了在CSP 027研究受试者中描述模型识别恶性肿瘤的精确度，我们报告了AUC和95%CI<sup>[10]</sup>。为了校准模型，我们五等分恶性肿瘤的预测概率并据此将研究队列分层，然后对每一区间预测概率内的患者描记观察到的恶性肿瘤概率<sup>[11]</sup>。

通过重新取样或交叉验证法，我们从内部验证了模型，这有助于我们运用全部的资料去建立模型<sup>[12]</sup>。为了达到这一目的，我们将研究人群随机分成10等份。对于每一组，我们运用逻辑回归模型估计的参数产生肿瘤转移的预测概率，模型使用的数据来自其他9组。10个逻辑回归分析有相同的标准；每个利用了90%的数据，然后我们对交叉验证法产生的概率计算AUC。为证实模型的临床应用价值，我们根据贝叶斯定理计算FDG-PET扫描后恶性肿瘤的验后概率，以此作为验前概率和FDG-PET扫描结果<sup>[13]</sup>。

## 结 果

375例良、恶性SPN患者的特征见表1。患者平均年龄为(65.9 ± 10.7)岁。恶性SPN的患病率是54%。大多数患者目前吸烟( $n = 177$ )或既往吸烟( $n = 177$ )。恶性结节患者年长些，多为目前吸烟、既往吸烟和吸烟量(包-y)大的患者。既往吸烟的良性SPN者戒烟年数长于既往吸烟的恶性SPN者。伴有恶性结节者多近期被诊断为肺癌，并可能有胸腔外恶性肿瘤病史，但后者的差异没有统计学意义。恶性结节多比较大，位于上叶，常为胸片确诊的恶性肿瘤。

运用多变量逻辑回归分析我们发现了恶性肿瘤的

表1 良、恶性SPN患者的特征<sup>(1)</sup>

特 征	SPN 组		P 值
	良性 (n = 171)	恶性 (n = 204)	
<b>临床</b>			
年龄 / 岁	62	68	< 0.0001
男性 / %	97	98.5	0.33
目前吸烟 / %	40	53	0.02
既往或目前吸烟 / %	91	97	0.01
吸烟史 / 包 - y	46	63	< 0.0001
戒烟时间 / y	8.1	5.6	0.03
其他恶性肿瘤史 / %	14	20	0.12
肺癌史 / %	9	16	0.04
<b>放射学</b>			
结节大小 / mm	14.8	18.9	< 0.0001
胸片“确诊恶性” / %	1	4	0.09
结节部位 / %			
上叶	46	54	0.11
右肺	59	55	0.42
诊断为肺癌的时间 / d	316	368	0.67
诊断为其他恶性肿瘤的时间 / d	460	620	0.11

<sup>(1)</sup> 除非另有说明，表中数值以 $\bar{x}$ 表示。

4个独立预测因子(表2)。其他预测因子与恶性肿瘤不相关，所以未纳入最终的模型分析。目前或既往吸烟者恶性结节的概率大约8倍于从不吸烟者(OR 7.9, 95%CI 2.6 ~ 23.6)。年龄每增加10岁恶性肿瘤的可能性增加2倍(OR 2.2, 95%CI 1.7 ~ 2.8)，结节直径每增加1 mm恶性肿瘤的可能性约增加10%(OR 1.1, 95%CI 1.1 ~ 1.2)，戒烟年数每增加10 y恶性肿瘤的可能性大约降低40%(OR 0.6, 95%CI 0.4 ~ 0.7)。

采用如下方程式对临床预测模型进行描述：

$$\text{恶性SPN的概率} = e^X / (1+e^X) \quad (1)$$

$$X = -8.404 + (2.061 \times \text{吸烟}) + (0.779 \times \text{年龄}10) + (0.112 \times \text{直径}) - (0.567 \times \text{戒烟年数}10) \quad (2)$$

e是自然对数的底数。目前或既往吸烟为1(其他为

0)；年龄10为年龄按每10年区分；直径指结节的最长直径，按mm计；戒烟年数10为戒烟年数以10年区分。

似然比拟合优度检验和Hosmer-Lemeshow检验表明，应用模型对结果的解释优于偶然性( $P < 0.001$ )，结果的预测可能性与观察到的可能性类似( $P = 0.61$ )。参数估计的相关矩阵显示对多重共线性无益。

模型的精确度非常好，ROC曲线的AUC为0.79(95%CI 0.74 ~ 0.84)。我们用交叉验证法产生的预测概率与其有近似的AUC(0.78, 95%CI 0.73 ~ 0.83)。模型校准良好；对于每一区间预测概率内的患者，观察到的恶性SPN频率与预测概率近似(图1)。例如，对于第4区间预测概率内的75例患者，恶性肿瘤的预测概率为66% ~ 79%，观察到的恶性SPN频率为72%。

为了评价FDG-PET扫描后恶性肿瘤的验后概率，我们假设这一检查方法鉴别恶性SPN的敏感性为94%，特异性为83%，正如荟萃分析所显示的<sup>[14]</sup>。当FDG-PET扫描结果为阴性且验前概率相对较低(20%)时，则恶性肿瘤的验后概率 < 2%；相反，当FDG-PET扫描结果为阴性且验前概率相对较高(65%)时，计算的验后概率 > 10%。

## 讨 论

SPN的处理仍是棘手的难题<sup>[15]</sup>。在本研究中，我们运用最近完成的退伍军人事务部合作研究提供的数据和退伍军人事务部的管理数据，建立并验证一种评估验前概率的新模型，从而便于SPN患者的处理。我们发现恶性SPN的4个独立预测因子，包括吸烟史、年老、大结节直径和戒烟时间短。重要的是，我们建立了简约的临床预测方程式，因其极佳的精确度和良好的校准度而用于评估特定患者

表2 恶性SPN的预测因子

预测因子	OR	95%CI
吸烟史 <sup>(1)</sup>	7.9	2.6 ~ 23.6
年龄每增加10岁	2.2	1.7 ~ 2.8
结节直径每增加1 mm	1.1	1.1 ~ 1.2
戒烟年数每增加10 y	0.6	0.4 ~ 0.7

(1) 曾经或从不。

恶性SPN的概率。

我们模型的精确度类似 Swensen 等<sup>[2]</sup>在梅奥诊所建立的模型。正如以前的模型，我们模型中纳入的吸烟史、年老和大结节直径为恶性SPN的独立预测因子。然而，在我们模型中，吸烟史校正后的OR值高于梅奥诊所模型的相应值。与梅奥诊所模型不同的是，我们发现戒烟时间短是恶性肿瘤的另一独立预测因子。此外，经校正其他可能的预测因子后，我们不能证实他们的发现即上叶结节更可能为恶性肿瘤。梅奥诊所的研究人员排除了5 y内有肺癌史和胸腔外恶性肿瘤史的患者，或许他们认为这些是恶性SPN的有力预测因子，但我们发现新近诊断的肺癌或胸腔外恶性肿瘤不是恶性肿瘤的独立预测因子。遗憾的是，因为CSP 027没有收集这些变量的资料(例如小锯齿状和很久前的胸腔外恶性肿瘤史)，所以我们不能评价它们，但梅奥诊所模型中有。由此，我们不能进行梅奥诊所模型的外部验证，也不能在CSP 027研究参加者中将我们模型的精确度与梅奥诊所模型的精确度相提并论。

我们和其他学者的研究<sup>[16, 17]</sup>均提示严格根据恶性肿瘤的验前概率处理SPN是一种有效并符合成本-效率的策略。例如：Cummings 等<sup>[16]</sup>报道对孤立性肺结节的处理可选择外科手术、针吸活检及密切观察。当恶性肿瘤的概率 > 68% 时外科手术显得更有效；当恶性肿瘤的概率为3% ~ 68% 时首选针吸活检；当恶性肿瘤的概率 < 3% 时首选密切观

察<sup>[16]</sup>。以往研究<sup>[17]</sup>中，我们发现当恶性肿瘤的概率为20% ~ 69%时PET扫描比针吸活检更具成本-效果，以此来证实和补充这些结果。

采用我们的模型能得出恶性SPN的验前概率，这种评估有助于解释诊断性检查结果。正如我们计算结果所表明的，当恶性肿瘤的验前概率相对较低(20%)且FDG-PET扫描结果为阴性时，则验后概率 < 2%，密切观察的策略是合理的。然而，当PET扫描结果为阴性且验前概率相对较高(65%)时，则验后概率 > 10%，在这种情况下须考虑针吸活检或电视辅助胸腔镜下活检。实例表明我们的预计方程式是如何有助于选择和解释SPN处理方法的临床决策。方程式也可并入正式的决策分析或成本-效果分析。

但我们的研究也有局限性。因CSP 027研究没有收集或只收集了少许胸片或CT扫描关于结节形态方面的信息，我们检测了相对少量的临床预测因子。所以，模型只解释了 > 30% 的病因学差异，表明还存在其他未发现的恶性SPN预测因子。因本

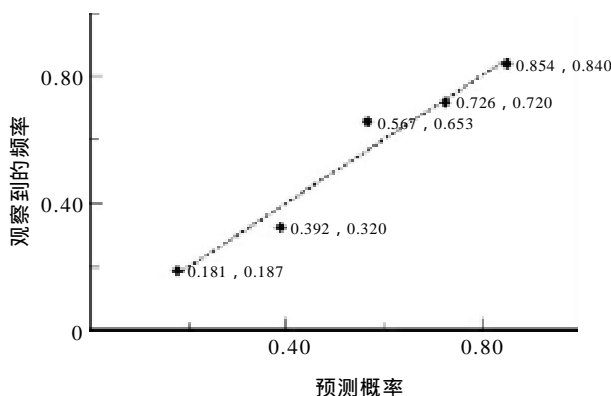


图1 临床预测模型的校准曲线

对于每一区间预测概率内的患者，以所观察到的恶性肿瘤频率来表示恶性肿瘤预测概率。曲线表明，对于每一区间预测概率内的患者，观察到的恶性肿瘤频率与恶性肿瘤的预测概率近似。对于第1 ~ 5区间预测概率内的患者，其预测概率范围分别为 < 0.29、0.29 ~ 0.48、0.48 ~ 0.66、0.66 ~ 0.79 和 > 0.79。

模型的建模人群为来自不同地方伴有SPN的老年白人男性，所以在该群患者中可能特别适用。然而，模型是否适用于有孤立小结节的女性尚不清楚。另外，模型不适用于结节直径 < 7 mm 的患者。因为模型的精确度很可能依赖于目标人群恶性肿瘤的患病率，与本研究相比恶性肿瘤患病率较低或较高的人群也许不能得到很好的校准。然而，我们认为本研究人群恶性肿瘤的患病率具有良好的代表性，因为退伍军人事务部CSP 027研究入选标准包括入选患者的胸片上有SPN，以此尽量减少选择偏倚。既往一些关于SPN患者的研究入选了较多根据胸部CT扫描结果或仅接受外科治疗的患者，结果人为导致恶性肿瘤的高患病率。

尽管模型的精确度非常好，但我们仍强调宁愿将模型作为工具来指导选择和解释诊断方法，并不试图将其作为单独的方法。虽然模型的交叉验证法产生近似的AUC，在其他SPN患者独立队列中我们的结果仍然需要外部验证。尽管本研究存在这些局限性，但是我们能采用简约的临床预计方程式估计恶性肿瘤的验前概率，有助于作出临床决策来处理SPN患者。目前吸烟或近期戒烟、有较大SPN的

年老患者发生恶性肿瘤的风险最高。

(朱韧 译 徐建芳 校)

## 参考文献

- 1 Cummings SR, et al. Am Rev Respir Dis 1986;134:449-452
- 2 Swensen SJ, et al. Arch Intern Med 1997;157:849-855
- 3 Gurney JW. Radiology 1993;186:405-413
- 4 Herder GJ, et al. Chest 2005;128:2490-2496
- 5 Kymes SM, et al. Clin Trials 2006;3:31-42
- 6 Siegelman SS, et al. Radiology 1986;160:307-312
- 7 Laupacis A, et al. JAMA 1997;277:488-494
- 8 Hosmer DW, et al. Stat Med 1997;16:965-980
- 9 Pulkstenis E, et al. Stat Med 2002;21:79-93
- 10 Hanley JA, et al. Radiology 1982;143:29-36
- 11 Katz MH. Multivariable analysis: a practical guide for clinicians. Cambridge, UK: Cambridge University Press, 1999
- 12 Efron B, et al. An introduction to the bootstrap. Boca Raton, FL: Chapman and Hall, 1993
- 13 Sox HCJ, et al. Medical decision making. Boston, MA: Butterworth-Heinemann, 1988
- 14 Gould MK, et al. JAMA 2001;285:914-924
- 15 Ost D, et al. N Engl J Med 2003;348:2535-2542
- 16 Cummings SR, et al. Am Rev Respir Dis 1986;134:453-460
- 17 Gould MK, et al. Ann Intern Med 2003;138:724-735

【英文原件请参阅 *CHEST 2007;131:383-388*】