

Probability of Cancer in Pulmonary Nodules Detected on First Screening CT

第一次CT扫描中检测到的肺结节是癌症的概率

背景

- 肺扫描的主要问题是：
 - 定义什么样的结果是阳性结果
 - 对检查出的结节怎么处理
- 肺扫描最关心的是是否有癌症

论文中的方法

- 分析了两组数据：PanCan (Pan-Canadian) 和 BCCA (British Columbia Cancer Agency)
- 给出了两个多变量分类器（逻辑回归模型）来估算肺癌的可能性
- 研究中确定了一些影响因素，这些因素与有肺癌或即将得肺癌有着决定性的联系。

结果和结论

- 分类器的效果非常好，两个分类器（parsimonious & full）的AUC都在0.9以上。甚至对于10mm以下的结节的判定都很准确。而这些结节正是在临床中最难处理的那一部分。
- 结节的大小和有肺癌的概率之间的关系是非线性的。
- 所有 perifissural 结节都不是癌症
- 结节是否有爪这一特征并没有给分类效果带来很大提升，但是这个特征确实有带来提升（ $P=0.09$ ）
- 可以根据分类器的结论来帮助医生确定复诊（什么时间复诊、怎么复诊）

存在的问题

- 没有考虑结节的形状问题（因为有一个数据集里没有是否有爪的信息，所以最后的模型里没包含这个变量）
- 使用的是简单的逻辑回归。
- 并不直接处理CT图像数据，而是使用总结下来的变量（在实际中需要医生给出这些变量，再由分类器分类）作为分类器的输入。
例如：是否为女性、年龄、是否吸烟、家族病史、结节大小、结节所在位置（肺的上半叶还是下半叶）、是实心还是半实心还是perifissural。