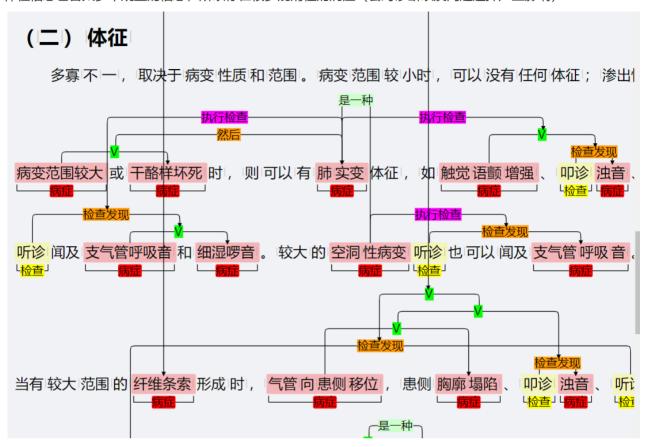
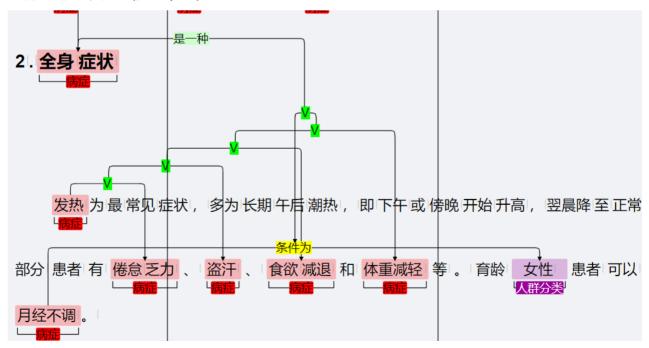
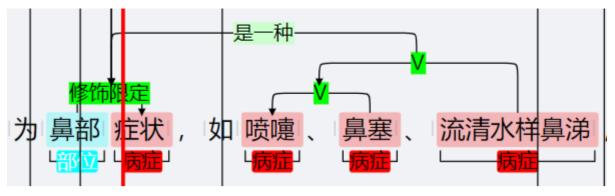
现有标记数据的问题

1. 体征信息包含太多不规整的信息,所以存在很多混淆性的病症(会对诊断以及问题选择产生影响)

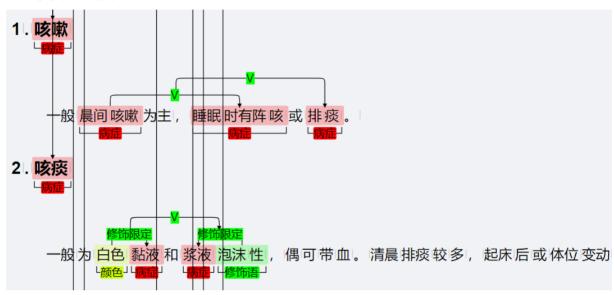


2. 导致关系标记不准确 (见FAQ 5.1)

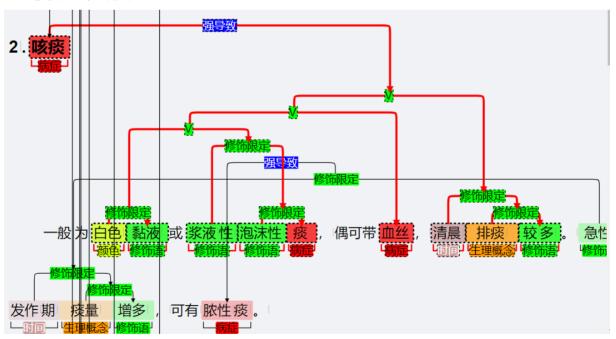




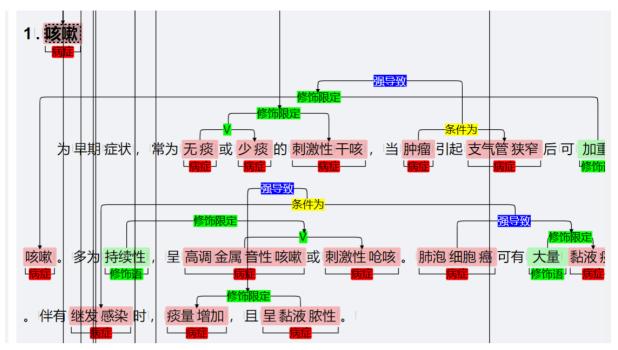
- 3. 症状细分标记有问题,可能会影响鉴别诊断
 - 慢性支气管炎3



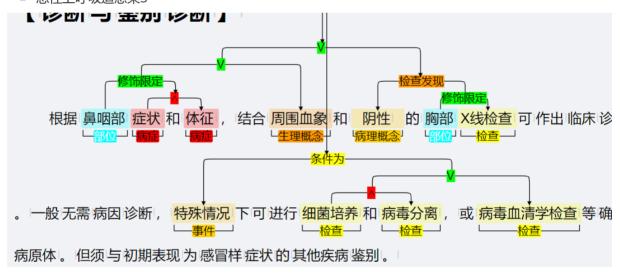
■ 慢性阻塞性肺疾病3



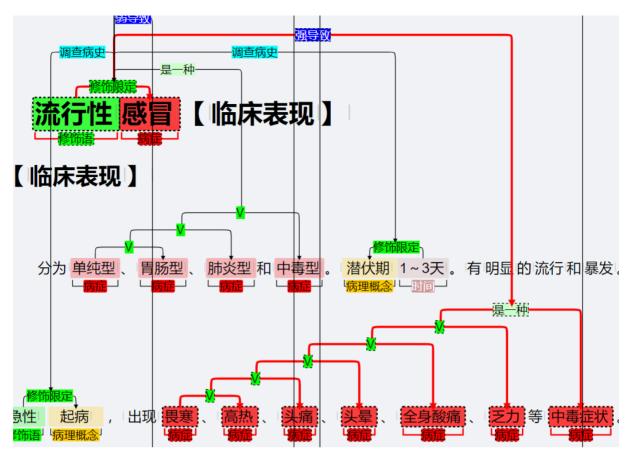
■ 原发性支气管肺癌1



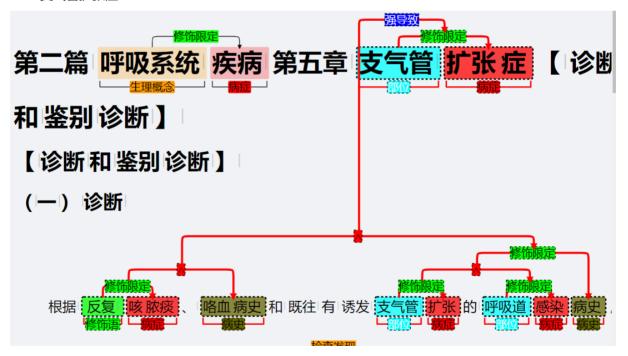
- 4. 诊断小节的数据可能作用不大,反而可能引起病症数据不准确(影响也不算特别大)
 - 急性上呼吸道感染3



- 5. 标记质量有问题
 - 急性上呼吸道感染4



■ 支气管扩张症1



之前数据已经标注了所有必须的信息,但是也包含了很多无用以及错误的知识,所以我建议用现有的文本模型重新标记《呼吸系统疾病》中"临床表现"和"诊断及鉴别诊断章节",或者直接使用外部文档进行统计