

Project: Adverse Event Associated with Patient's Death

1. 基本**数据1**：每行代表一个trial arm, 每行有death和其他serious/non-serious events（心脏停跳，头晕，呕吐，等等）
2. 要做的分析 **Table 1**基于数据1：

其他 Event 名字	其他 Event Frequency $F(\text{Event})$	当有这些Events的时候, Death出现的次数: $F(\text{Death}) \text{ given Event}$	Death probability : $F(\text{Death})/F(\text{Event})$	
心跳停止	出现100次	10	10%	
头晕	200	20	10%	
恶心。。。	200	10	5%	

3. Association mining（Fp-tree）**数据3**结果: [Event 1, Event 2, Event 3] => [Death], 比如[头痛, 心跳停止]=>[死亡]
4. 要做的分析**Table 2**基于**数据3**和**数据1**：

几个Event集合 (从rule mining 结果来)	Event集合同时出现频率 $F(E1, E2, E3)$	当这些Events集体出现的时候, Death出现的频率 $F(\text{Death}), \text{ given } e1, e2, e3 \text{ 同时出现}$	Death的		
心跳停止, 头晕	20	5	25%		
恶心, 呕吐	10	4	40%		

5. 要做的分析**Table 3**基于数据1和数据2：

从数据1中，我们可以获取所有跟death一起出现的高频/高几率event, 比如心脏停跳，头疼

数据5：采用一个Frequency/proability阈值，比如出现次数高于10，proability>0.05, 可以选出一个Event列表：比如心脏停跳，肺炎，血液感染，呼吸困难，等等

数据6：用数据5，排除过滤数据1里面的低频（不重要）event，比如假设：头晕，头疼。等于数据6是数据1的高频重要部分

数据7：从数据6中读取所有可能的event对（event pair. 比如数据6有[死亡，心脏停跳，肺炎]，就会生成三个pair: [死亡，心脏停跳]，[死亡，肺炎]，[心脏停跳，肺炎]。基本算法就是从一个event集合（set）里面找出所有的event对(pair)

数据8: 统计所有event对在**数据1**中的出现频率

根据数据7，8生成Table 3，用于Gephi做网络图

st ar t	tar get	Freque ncy : 数据8 的频率	directi on	TOD O: dise ase	TODO: treatm ent
死亡	心脏 停跳	10	undirec ted	待继 续	待继续
心脏 停跳	肺炎	11	undirec ted	待继续	待继续

6. 把table 3导入到Gephi, 作图