项目立项依据：

1. 首先长期的retrospective data用于建立医院模型较少，利用医院现有的数据构造特征，优化目前的CVD~~风险~~转归预测模型。
2. 现有的国内外用机器学习方法预测风险的文献数量有限，而对于患者而言，预测转归可以更好地提高用药的依从性，对于医院而言可以更好的管理患者，对于社会而言可以避免不必要的资源浪费以及劳动力的损失。
3. 目前人口老龄化，CVD的流行率
4. 目前国外的特征并不一定适用于国内（加入用药，以及病人的基本特征，并发症，以及，日常检查项，比如LDL等等）

国内外研究现状：

1. 研究内容（背景）：

用于预测一段时间结果或者是不良反应（如5年后发生心衰）

1. 研究目标：

根据目前现有的SVM/ RANDOM FOREST/DESICONTREE/LOGIT REGRESSION以及集成等方法，对溯值数据库中三甲医院的患者（存在complete data）进行转归的预测。

（但是目前有个问题是CVD存在转归吗？，用什么指标来确定转归）

在长期随访中，以死亡/肿瘤进展作为观察结局的临床试验中，如果研究的中期评价发现新药在某些方面（如无疾病进展时间）显著的优于对照组，按照伦理学的规定，对照组的患者将可以再之后的治疗中选择改变治疗方案，接受新药的治疗，这种“选择性的治疗转换”破坏了随机化原则，导致试验中对象的生存时间与删失时间不独立。对于这样的北京问题，利用cox模型考察组间死亡风险比时，可能得到有偏的参数估计。另一方面，在观察性研究中，若存在一个随时间变化的协变量，是研究的混杂因素，且会受到前次暴露的影响，此时常用的分析方法在估计暴露效应的时候也是存在偏倚的，以上两个问题虽然都来自于临床试验和纵向数据两种背景，然而都可以通过逆概率加权的分析方法取解决

Inverse probability of censoring weighted IPCW,删失逆概率加权（将时依性混杂因素剔除）

**Cardiovascular Event Prediction**

1. Available Data
2. Objectives
3. Background
4. Subjects
5. Design

