

第一天上午	关于数据预处理的方式与思路	<p>常见的缺失值有哪几种情况？ 常用的缺失值处理方法</p> <p>案例1：审稿人的要求：对缺失数据采用MI方法进行填补 什么情况下有必要将连续资料划分为为分类资料？ 如何把一个连续资料划分为分类资料？常用的方法有哪些？应注意哪些事项？</p> <p>案例2：审稿人的质疑：为什么该指标按1.0划分而不是按1.2划分？依据是什么？ 变量太多了怎么办？——介绍几种常见的降维技术</p> <p>主成分分析的思路与注意事项 案例3：审稿人的质疑：这些变量之间是否存在较高的相关性，是如何处理的？</p>
第一天下午	关于组间比较的分析方法与思路	<p>常见的组间比较方法介绍 组间比较方法的选择思路与建议</p> <p>案例1：两组等级资料比较，到底用卡方检验还是秩和检验？ 关于多重比较问题的探讨 案例2：这么多两两比较方法，到底应该选择哪个？ 案例3：利用FDR法从1万个基因中筛选出有意义的基因</p> <p>生存曲线比较的常用方法介绍 案例4：两条生存曲线有交叉，此时log-rank检验还能用吗？ 什么是bootstrap法？有什么用处？ 案例5：利用bootstrap法进行组间比较</p>
第二天上午	关于预测预警的分析方法与思路	<p>常见的预测预警方法介绍 如何利用logistic回归建立预测模型？</p> <p>案例1：审稿人的质疑：为什么选择这几个变量预测，而不是其它几个变量？有什么依据？ 什么是交叉验证？有什么用处？ 如何利用交叉验证思想选择模型和变量 案例2：审稿人的要求：建议利用交叉验证法重新建立模型。 什么是分类树？如何利用分类树建立预测模型 案例3：用分类树还是logistic回归？</p>
第二天下午	关于变化趋势的分析方法与思路	<p>变化趋势分析的常用方法介绍</p> <p>案例1：儿童生长发育情况的变化趋势分析 随机效应模型/混合效应模型思路介绍 线性发展模型/曲线发展模型介绍 案例2：比较不同人群的变化趋势是否一致</p>