**作答要求：**

* **请根据题目要求提交报告作为正文**
* **除非在问题中明确指出不能使用已有程序之外， 其余均可使用已有程序**

**决赛题目**：

X公司计划在全球开展一项III期、随机、双盲的多中心临床试验，用于评估研究药物相比于安慰剂在某疾病患者人群中的治疗有效性和安全性。

该研究将会以1:1的比例将满足入排条件的受试者随机分配至研究药物或安慰剂组，并按照筛选访视时的体重(轻,重)和地理区域（亚洲，欧洲和美洲）进行分层随机分配。所有受试者将接受为期24周的药物治疗。

该临床试验的主要研究问题是研究药物较安慰剂组就下述主要终点有何比例差异(difference of proportions)：主要终点是受试者在第24周时通过患者评分标准A判定的疾病较基线的改善情况。具体来说，若经患者评分标准A判定为疾病较基线有改善，则称为应答者，否则称为非应答者。

另外，该III期研究还将考虑以下两个研究问题：

1. 关于患者生活质量指数评分在第24周改善，研究药物较安慰剂组有何比例差异(difference of proportions);
2. 关于患者症状评分B在第24周较基线有至少20%改善，研究药物较安慰剂组有何比例差异(difference of proportions）。

**问题1**： 在研究设计阶段假定第24周研究药物组和安慰剂组对应于主要终点的应答率分别为55%和30%。假设考虑到研究药物的治疗疗效可能与预期不一致，该研究考虑在试验过程中进行一次期中分析以此来评估是否需要进行样本量的再次估计。假设在该期中分析中

观测到试验药物组和安慰剂组的应答率分别为40%和25%，在I类错误率被控制在单侧0.025的条件下，请从统计方法的角度计算为了达到90%的条件把握度(conditional power)所需的样本量。

**问题2**： 假设在该国际多中心的试验中，各区域在主要终点上预期的药物疗效略有不同，见下表：

|  |  |  |
| --- | --- | --- |
| **区域(占总体样本量的百分比)** | **研究药物组应答率** | **安慰剂组应答率** |
| 亚洲(50%) | 53% | 35% |
| 欧洲(20%) | 54% | 27% |
| 美洲(30%) | 50% | 25% |

请根据按上述区域分层的Cochran-Mantel-Haenszel(CMH）的统计分析方法来计算总体样本量，将I类错误率控制在单侧0.025水平，达到95%把握度(power)。

**问题3：** 如果该研究同时也对研究**药物组和安慰剂**组关于主要终点的相对风险(relative risk)感兴趣，请对附件中的试验数据进行分析。模型中只考虑如下解释变量（不考虑交互作用项）：治疗组(研究药物和安慰剂组)，分层因素（体重和区域），以及连续变量年龄。请提供相对风险的点估计和对应的95%置信区间。

**问题4：** “辛普森悖论（英语：Simpson's paradox），是概率和统计中的一种现象，其中趋势出现在几组数据中，但当这些组被合并后趋势消失或反转”（[辛普森悖论 - 维基百科，自由的百科全书 (wikipedia.org)](https://zh.wikipedia.org/wiki/%E8%BE%9B%E6%99%AE%E6%A3%AE%E6%82%96%E8%AE%BA)）。请问在附件中的试验数据中关于相对风险(relative risk)是否存在辛普森悖论现象。 如果有，请找出一个例子。

**问题5：** 假设在试验进行的24周内中途停止治疗的比例是15%，第24周试验药物组和安慰剂组就主要终点的应答率分别为55%和30%。如果入组200的样本量， 假设在该国际多中心的试验中，某个国家需要足够的来自本国人群的样本量用于支持在该国的注册。其样本量的确定主要是基于一致性的评估标准。 这里一致性定义为：在该III期研究**基于200受试者**时主要终点显示统计学意义条件下，该国受试者的**疗效观测值是总体(即200)受试者疗效观测值的至少一半以上**。请通过模拟试验来计算该国需要多少样本量才能使得该一致性的概率至少为80%以上，并将程序附上(注：不能调用已有的计算一致性评估的软件包）。