

以一项肿瘤研究的设计讨论看 RCT 与 RWS 的联系与区别

Relationship and Distinction between RCT and RWS with Design of An Oncology Project as Example

姚向明, 杭州泰格医药科技股份有限公司

摘要

在 2021 年, 研究者发起设计一项关于食管癌的研究, 统计师参与这研究设计。根据研究者感兴趣的医学问题和执行该研究的可能条件, 统计师和研究者一起思考了诸如基线特征、治疗策略、伴发事件和终点评估的临床要素。“随机对照研究”和“真实世界研究”被作为两个备选方案, 设计人员分别考虑了其受试者入组和分配、基线信息的均衡性和偏倚控制方式(如倾向评分法)、样本量计算和预估、关键研究终点选择和统计分析方法、分析时间节点设计和 alpha 控制策略及研究持续时间等要素。比较了 RCT 和 RWS 的优势和劣势后, 制定了一个包括两部分的初步设计。在设计考虑中, 使用 SAS 9.4 和 excel 表格进行了一些模拟运算。

An IIT on esophageal carcinoma was designed in 2021, statistician took part in the design of this trial. Dependent on the clinical question of interest and the possible conditions to conduct this trial, statistician went through the clinical information factors such as baseline characters, treatments, intercurrent events, and endpoints measurement, together with investigator. Two scenarios, RCT and RWS, were both considered, including the subject enrollment/assignment, baseline characters balance and bias control (such as PS method), sample size calculation/expectation, key endpoints selections and statistical analysis methods, analysis schedules and alpha control, and study duration etc. Comparing the advantage and weakness of both RCT and RWS, an initial protocol was drafted as two parts. Some simulations via SAS 9.4 and Excel were conducted during the consideration.

简介

本文介绍了一个研究者发起的研究设计的过程。在这个过程中, 研究者考虑了随机对照研究和真实世界研究两大设计策略。首先统计师协助研究者制定了一个设计流程: 背景理论沟通→理论与具体医学问题的结合→在 RCT 和 RWS 策略下研究设计各要素的具体设计→结合开展研究的条件对比两种策略的优劣性和可行性→制定初步研究设计→在一个小样本研究中模拟演习初步设计并进行必要调整。在背景理论沟通部分, 简要介绍了各因素逻辑关系、混杂和偏倚, 以及常用研究设计的构架和相互关系; 在具体设计部分, 统计师使用 SAS 工具进行了各种模拟(样本量、持续时间等)。为方便研究者及助手结合临床问题进行决定, 统计师使用 excel 中的公式功能, 制作样本量预估和研究持续时间模拟小工具。

临床试验中相关因素的逻辑关系

临床试验从某种意义上说是发现医疗实践中各因素之间关联的过程, 我们往往自然而然会把我们的注意力放在我们关心的若干因素上, 而忽略了其它一些因素。我们可以尝试着将因素归纳为“基线”、“变化”、“结局”三类, 这三类因素之间是一个相互派生, 环环相扣的过程。然后我们尝试将这些因素用下图来表示。



图 1：临床试验中相关因素的逻辑关系示意图

在实际的研究中，我们会通过收集数据或利用现有数据去探索因素与因素之间关系。我们可以把我们感兴趣的因素及其之间联系看成是主干，而把其它因素看成是主干之外的树枝和树叶（其中一些是会带来偏倚的混杂因素）。它们的关系就像一个网络一样，但我们可以把这个网络图化整为零成一个个结点进行探索，如下图。

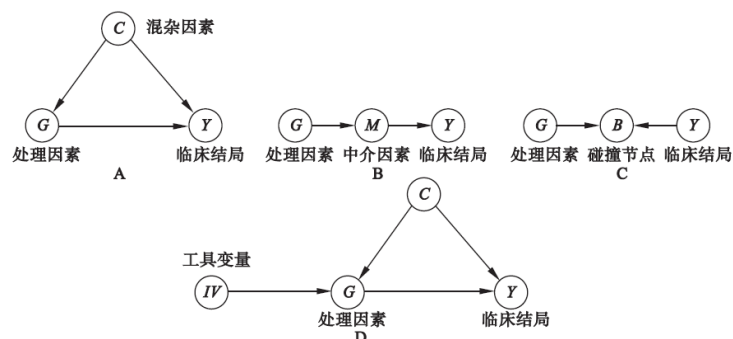


图 2：有向无环示意图（DAG）——摘自《倾向性评分方法及其规范化应用的统计学共识》

根据研究者提供的研究目的：研究预期手术组能够取得优于根治性同步放化疗的生存率，统计师和研究者一起整理了本研究可能的相关因素逻辑关系如下图：



图 3：计划进行的临床试验中相关因素的逻辑关系示意图

前瞻性研究设计的几种类型

盲态随机对照干预性研究结果的证据力是最强的，如果条件允许的话，研究者希望能设计成这样一个研究。

结合这个研究开展的实际情况，我们对于总体设计的各个要素进行了下面的梳理：

双盲、随机、对照、干预性研究
双盲：手术治疗与根治性放化疗完全不同
选择：开放性试验
随机：伦理学问题及可行性问题
选择：意向性入组
对照：手术治疗 VS 根治性放化疗
干预：实用性临床试验
 （非干预：观察性临床试验）

实际上，对于总体设计的考虑经历了从证据力最强到较强的结合实际开展条件的选择如下图。



图 4：前瞻性研究设计的类型

经过上述梳理，研究者实际上选择了实用性临床试验的设计（Pragmatic Clinical Trial, PCT，实操/实效临床试验）：尽可能接近真实世界临床实践的临床试验，是介于 RCT 和观察性研究之间的一种研究类型。

实际上已经在真实世界研究的范畴之内：针对预设的临床问题，在真实世界环境下收集与研究对象健康状况和/或诊疗及保健有关的数据（真实世界数据）或基于这些数据衍生的汇总数据，通过分析，获得药物的使用情况及潜在获益-风险的临床证据（真实世界证据）的研究过程。

即通过下面的路径，以期获得有效的真实世界证据。

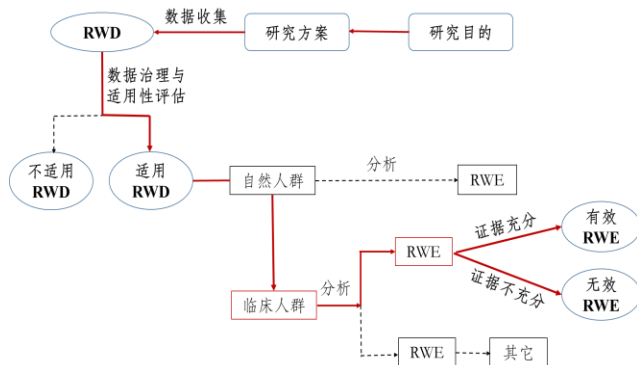


图 5：RWE 获得路径

对于随机对照研究 RCT 和实用临床研究 PCT，统计师制作了一个对比表供研究者理解：

	随机对照研究 RCT	实用临床研究 PCT
干预处理	标准化	标准化或非标准化
人群	严格入排	入选标准宽泛
终点指标	临床终点/替代指标	临床终点

对照组	阳性药物、安慰剂	不同的标准化治疗
分组方式	随机化	随机化或自然选择
盲法设计	双盲、单盲、开放	开放
样本量	相对较小	大

表 1: 随机对照研究 RCT 和实用临床研究 PCT 对比表

研究设计的基本构想

虽然是一个前瞻性的总体设计，但研究者希望能充分利用已有的数据库，并希望能够在 RWE 的获得中，能把前瞻性收集的数据和已有的数据库数据结合起来，从而更早获得 RWE。

从研究者感兴趣的医学问题角度，研究者和统计师确定了如下的疗效终点：

- Time to Event, 事件的定义：复合策略
 - OS 包括 1)因研究疾病进展死亡；2)因不良事件死亡；3)因研究疾病和不良事件以外原因死亡
 - PFS 包括 1)疾病进展；2)因各种原因死亡
 - LDFS 包括 1)死亡；2)局部复发；3)喉切除；4)气管造口或管饲
- Best Response:
 - pCR
 - MPR

结合上面的情况，初步确定研究设计分两部分：

第一部分是使用现有数据库内的数据进行分析：

第一步：分别以 OS、PFS、LDFS、pCR、MPR 为终点，探索影响这些终点的基线因素和伴发事件。计划基于 SAS 9.4 的 proc logistic/glimmix、proc phreg 分别进行 stepwise 逐步回归挑选出影响 OS、PFS、LDFS、pCR、MPR 的因素，并结合临床进行挑选；

第二步：基于第一步选择出的因素，计算倾向性评分（PS），分别尝试匹配、分层、协变量加权、逆概率加权方法控制混杂，计算 OS、PFS、LDFS、pCR、MPR 在两种处理间的差异量及可信区间；

第三步：基于第一步选择出的因素，在控制共线性的基础上，进行 OS、PFS、LDFS、pCR、MPR 各个终点的多元回归分析；

第四步：结合第二步和第三步的结果，1）选择第二部分（前瞻性研究）使用 PS 的方式，2）计划第二部分（前瞻性研究）的样本量，3）选择第二部分多个终点 step-down 的检验顺序，控制总体 alpha。

第二部分是基于实用临床研究的思路平行数据进行分析（暂时不考虑合并第一部分数据）：

第一步：使用第一部分确定的 PS 使用方式，确保干预组间的基线因素的均衡可比性；

第二步：计算 OS、PFS、LDFS、pCR、MPR 在两种处理间的差异量及可信区间；

第三步：基于第一部分选择出的因素，在控制共线性的基础上，进行 OS、PFS、LDFS、pCR、MPR 各个终点的多元回归分析。

但由于医疗实践发展，存在对照治疗组样本量不足的可能性，除了合并第一部分数据的备选方案外，还有一个备选方案就是以第一部分的对照治疗组的点估计为对照，进行试验干预组的单臂试验。

倾向评分及控制混杂的方法

为协助研究者设计方案，统计师简要介绍了倾向评分及通过处理倾向评分控制混杂的方法。倾向性评分可以理解为在观察到的协变量条件下，观察对象接受某种处理（或暴露）的概率。PS 本身不能控制混杂，但基于对于 PS 的处理可以不同程度地提高干预组间的均衡性（可比性），类似于“事后随机化”。

PS 处理的基本思路如下：

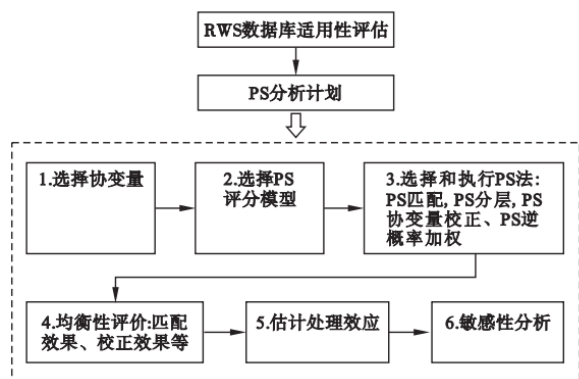


图 6 PS 方法的基本步骤——摘自《倾向性评分方法及其规范化应用的统计学共识》

汇总匹配、分层、协变量加权、逆概率加权方法优劣性如下表：

方法	优势	劣势
匹配	相较 PS 分层和 PS 校正控制混杂因素的优势更加明显匹配得当分析结论相较于其他 PS 方法更接近于随机化研究	未能匹配的对象从分析中剔除降低了疗效的估计精确度和外部推广的验证力度
分层	利用所有数据	控制混杂的效果低于匹配和加权生存分析中无法很好的运用当两组的 PS 评分分布偏离较大时可能有的层中只有对照组个体而有的层只有试验组的个体此时这些层不能提供信息
协变量加权	利用所有数据	可在设计时考虑也可在分析时进行没有明确的界定需要明确定义 PS 评分与结局变量关系的回归模型协变量校正策略易受分析结果的影响
逆概率加权	利用所有数据相较 PS 分层和 PS 校正控制混杂因素的优势更加明显分析结论相较于其他 PS 方法更接近于随机化研究	极端倾向评分值易导致分析结果不稳定

表 2：四种基本 PS 方法总结

样本量的估计方法

研究者希望用 5 年生存率为主要的终点，计划通过计算风险率 hazard rate 的比值（HR）并进行组间差异的假设检验，然后通过指数分布，计算两个组的 5 年生存率。

意向性入组的 PCT，样本量计算仍然参照 RCT，但由于两组比例不定，对于疗效的估计需要更加谨慎（通过需要适当扩大样本量），且需要把波动因素考虑进来（假设在 0.5-2.0 之间波动）。下面通过样本量计算软件计算。

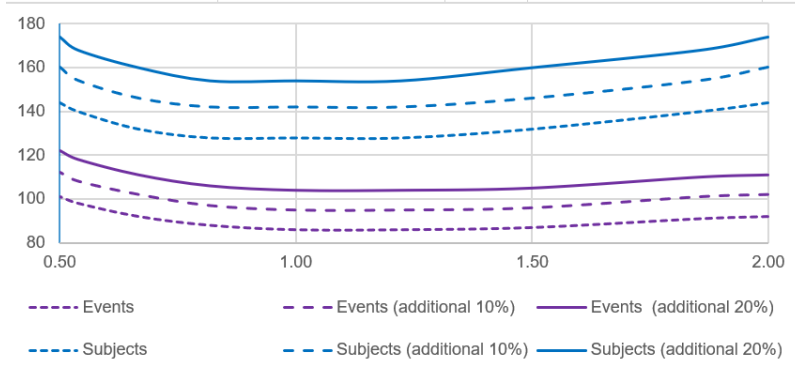
5 years OS rate=28% vs 50%)

HR=-ln(0.5)/-ln(0.28)=0.5445

OS=32.67 months vs 60 months

Enroll=26 months / Follow-up=60 months

Alpha=0.025 one-tailed / Beta=0.20



如果采用了备选的单臂试验方案，将通过计算中位 OS 并与作为对照的点估计进行比较。在这种情况下，需要计算单组生存时间估计的样本量，通过下面的公式写 SAS code 进行计算。

单位时间事件发生风险 λ 的自然对数 γ 来代表平均水平：

$$\gamma = \ln(\lambda) = \ln\left(\frac{-\ln(1-0.5)}{T_m}\right) = \ln\left(\frac{\ln(2)}{T_m}\right)$$

n 是样本量， λ 为单位时间内发生事件的风险， T 为受试者入组持续时间， τ 为入组结束后随访观察时间。

$$\text{var}(\gamma) = \frac{1}{n \left[1 - \frac{1}{\lambda T} e^{-\lambda \tau} (1 - e^{-\lambda T})\right]} = \frac{1}{E(d)} = \frac{1}{n \text{Prob}_d}$$

$$\gamma_1 - Z_{1-\beta} \sqrt{\text{var}(\gamma_1)} > \gamma_0 + Z_{1-\alpha} \sqrt{\text{var}(\gamma_1)},$$

计算样本量 n 如下：

$$n > \frac{1}{\ln\left(\frac{T_{m0}}{T_{m1}}\right)^2} \frac{(Z_{1-\alpha} + Z_{1-\beta})^2}{\left[1 - \frac{1}{\lambda_1 T} e^{-\lambda_1 \tau} (1 - e^{-\lambda_1 T})\right]}$$

结论

虽然盲态随机对照研究是结果证据力最高的研究设计，但在研究条件限制情况下，在某些要素下实事求是地进行调整，但仍然需要尽可能接近盲态随机对照研究欲达到的控制基线混杂因素带来偏倚的目标，进行研究的总体设计。倾向性评分本身并控制混杂，但通过使用倾向性评分，可以达到基线的最大程度的均衡可比性。倾向性评分的处理有多种方式，需要根据实际情况进行选择。在估计样本量时，仍需要按照 RCT 的原则进行，并且需要考虑受试者比例并不能肯定接近统计检验效能最高的 1:1。在一些医疗条件下，对照组由于意向性可能无法获得足以进行统计检验的样本量，在这种情况下，单臂设计也是备选方案之一。

参考文献

黄丽红, 王永吉, 王素珍, 赵杨, 王彤. 倾向性评分方法及其规范化应用的统计学共识. 中国卫生统计, 2020, 37(6): 952-958

致谢

感谢泰格医药首席统计专家魏朝晖博士对于本文及本文所基于的研究者发起研究设计的指导和提供的参考文献。

联系方式

请提供你宝贵的意见和指导，如有疑问或指正，请联系我：

姚向明/Joe Yao

Joe.yao@tigermedgrp.com