



关联分析研究设计

易侬学院 陈驰

内容



研究设计分类



时相关系



病例对照研究



队列研究



横断面研究

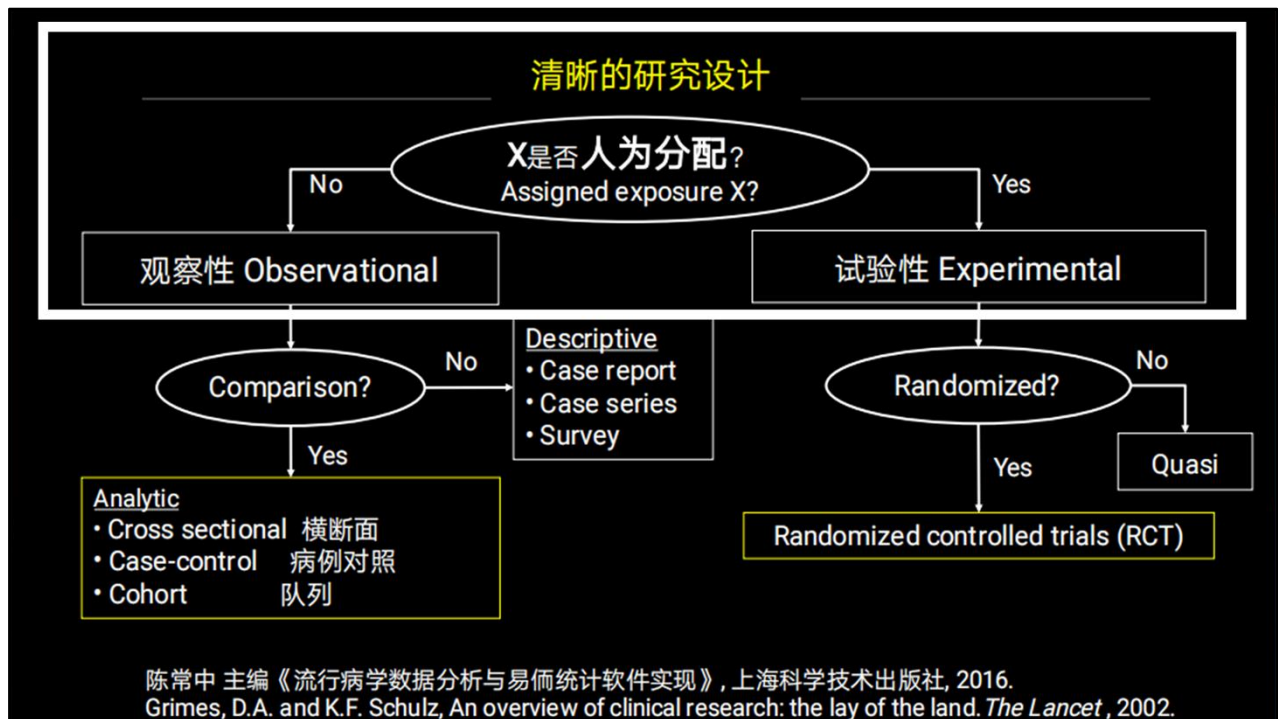


真实世界研究

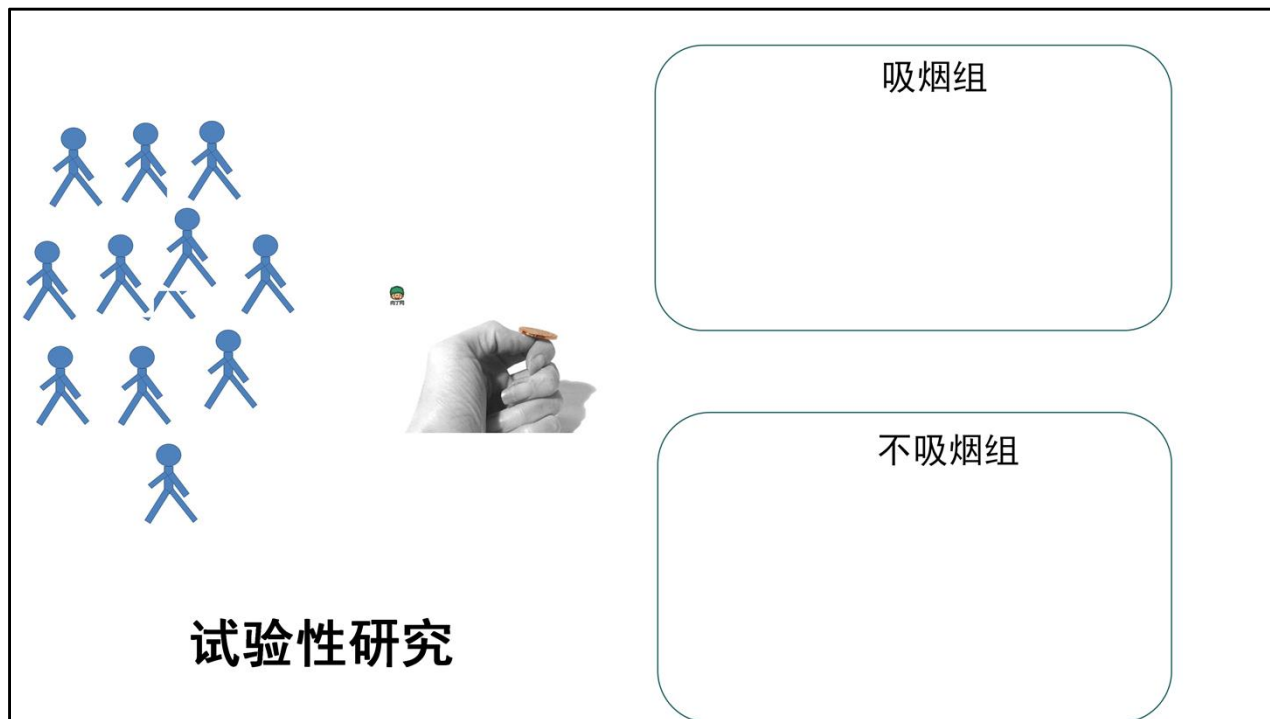


01

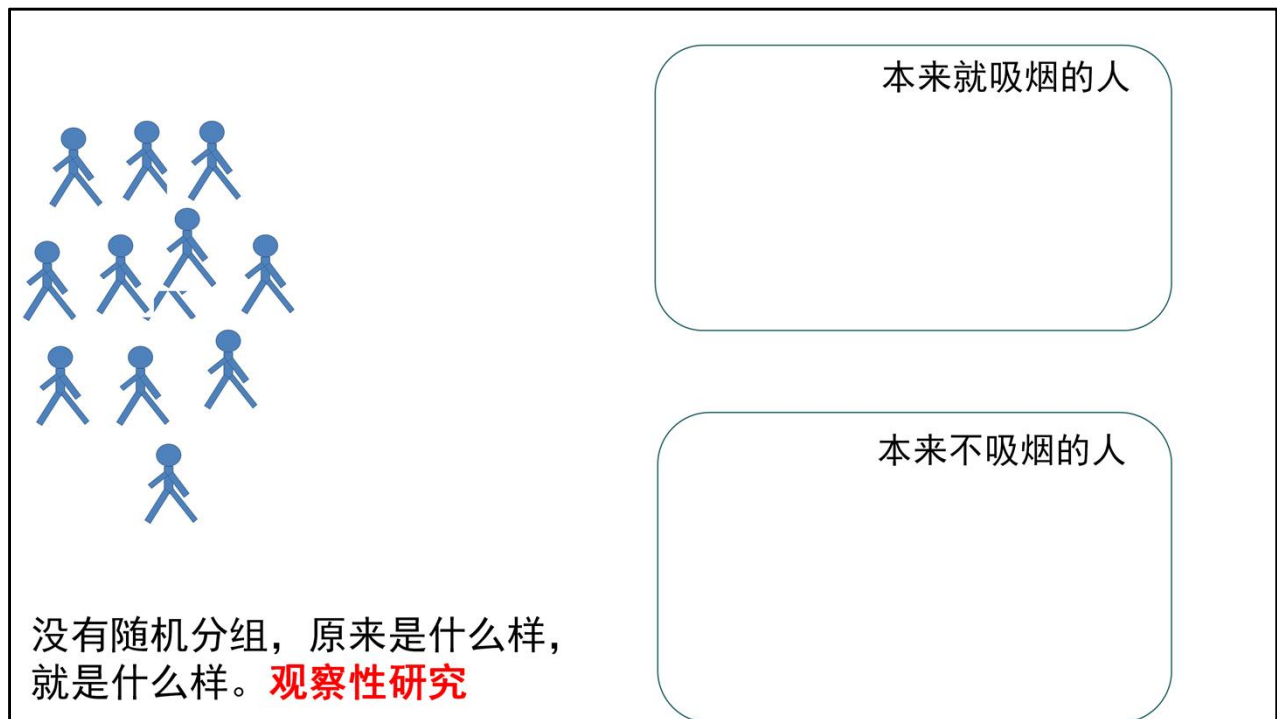
研究设计分类



临床研究设计类型，可根据X是否人为分配，分为观察性研究及实验性研究。区分两者的关键是分配X时的动机。是基于临床实际场景还是实验？其中，观察性研究根据是否比较分为描述性研究和分析性研究。分析性研究则是本节课重点阐述的内容：横断面、病例对照、队列。



譬如：研究者想知道吸烟会否对肺功能造成影响，其基于随机化（例如掷硬币）来决定谁被分到吸烟组，谁被分到不吸烟组，则试验性研究。



相反，同样的研究，研究者并未分组，而是基于研究对象本身是否有吸烟习惯自动的进行人群区分。即吸烟/不吸烟的分配动机并非为了试验，而是自然发生。

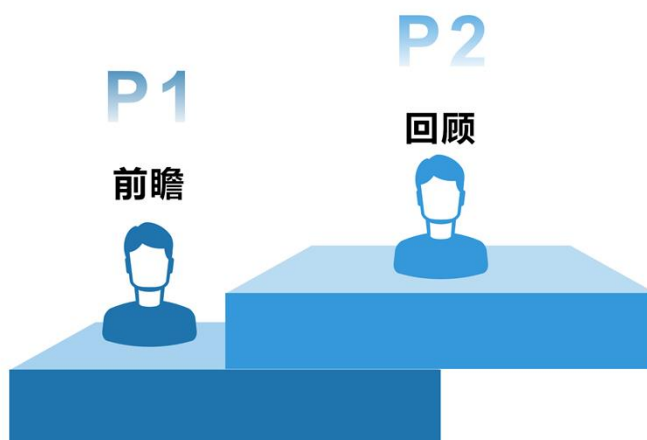


02

研究设计的时相

临床研究的时相

“



临床研究的时相只有两种：前瞻&回顾

关于时相的两个概念

- 一. 时相和研究设计类型无关。是 独立的存在。
- 二. 当你决定要做研究 的那一瞬间，XY的指标已经客观存在，则是回顾。如果X和Y任意一个未知，则为前瞻。

前瞻和回顾的区分与研究类型无关。其取决于研究者在进行研究时，X、Y数据是否客观存在。如X、Y数据任意一个未知，则为前瞻。如均存在，则是回顾。

03

病例对照研究



病例对照的三个基础概念：

01

概念1

病例对照≠将患者随机分成病例组及对照组

02

概念2

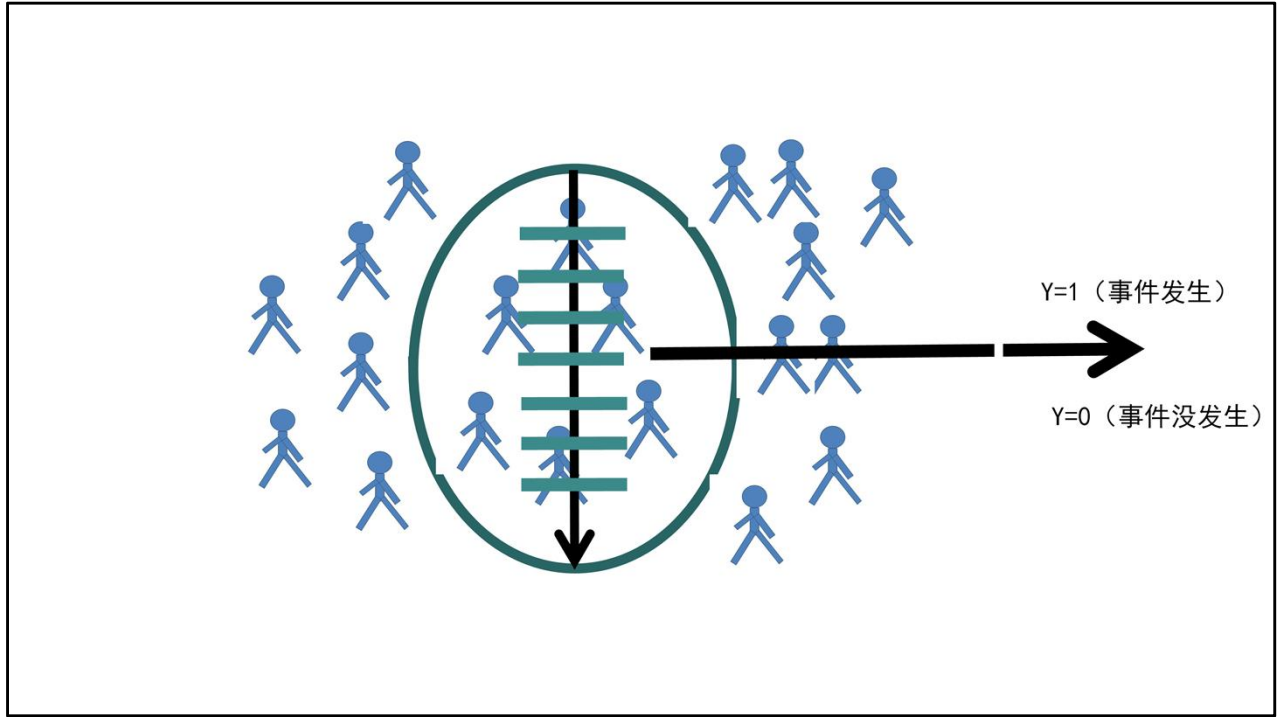
先知道结果，再反过来看X变化

03

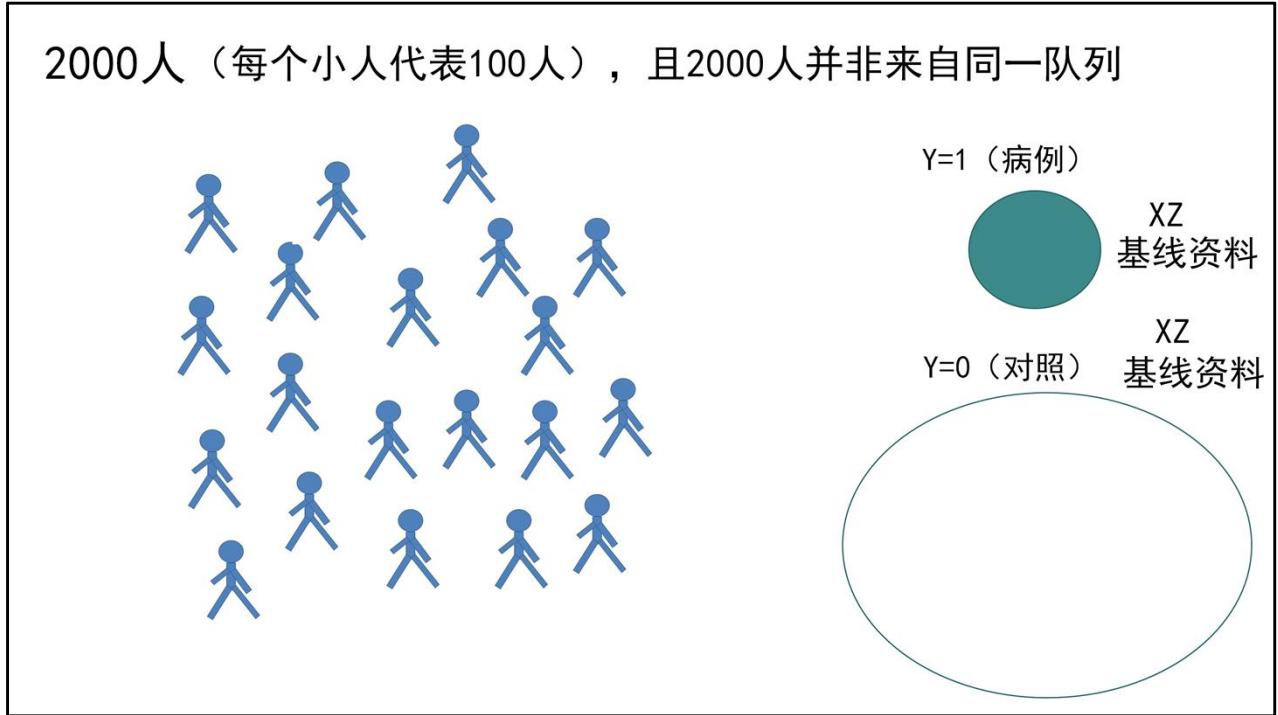
概念3

对照的选择是关键，一定要独立于X

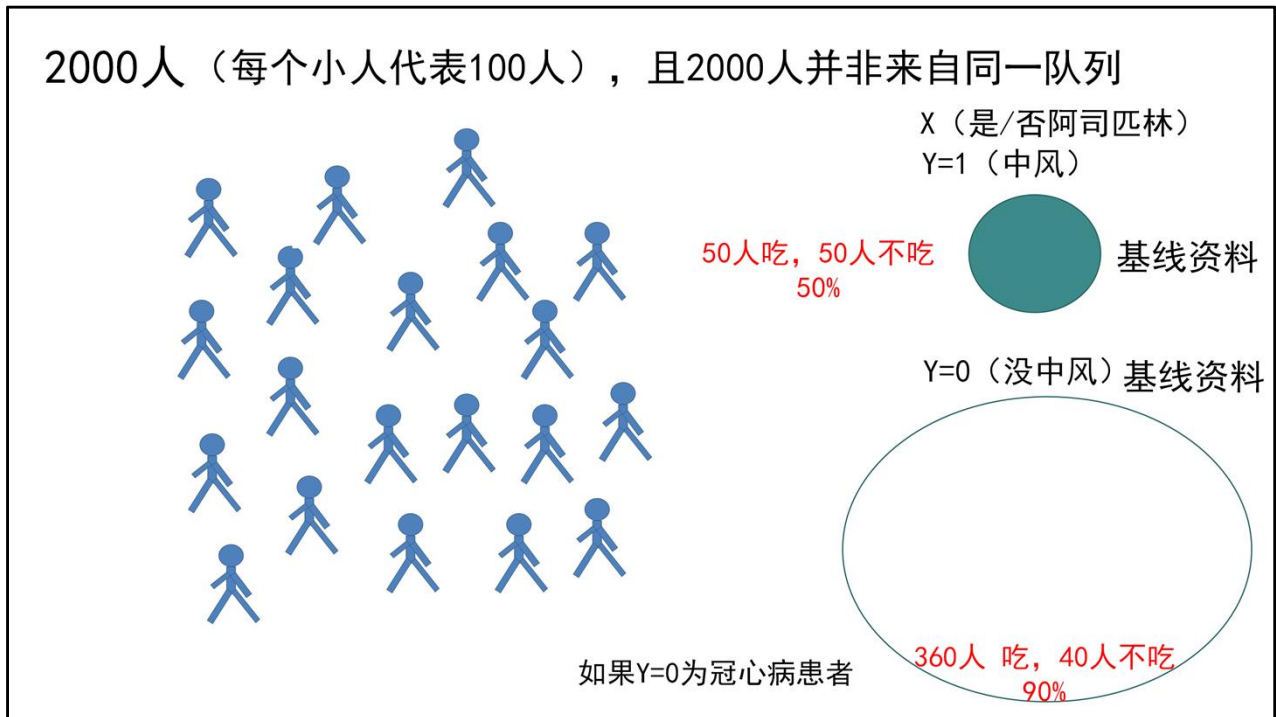




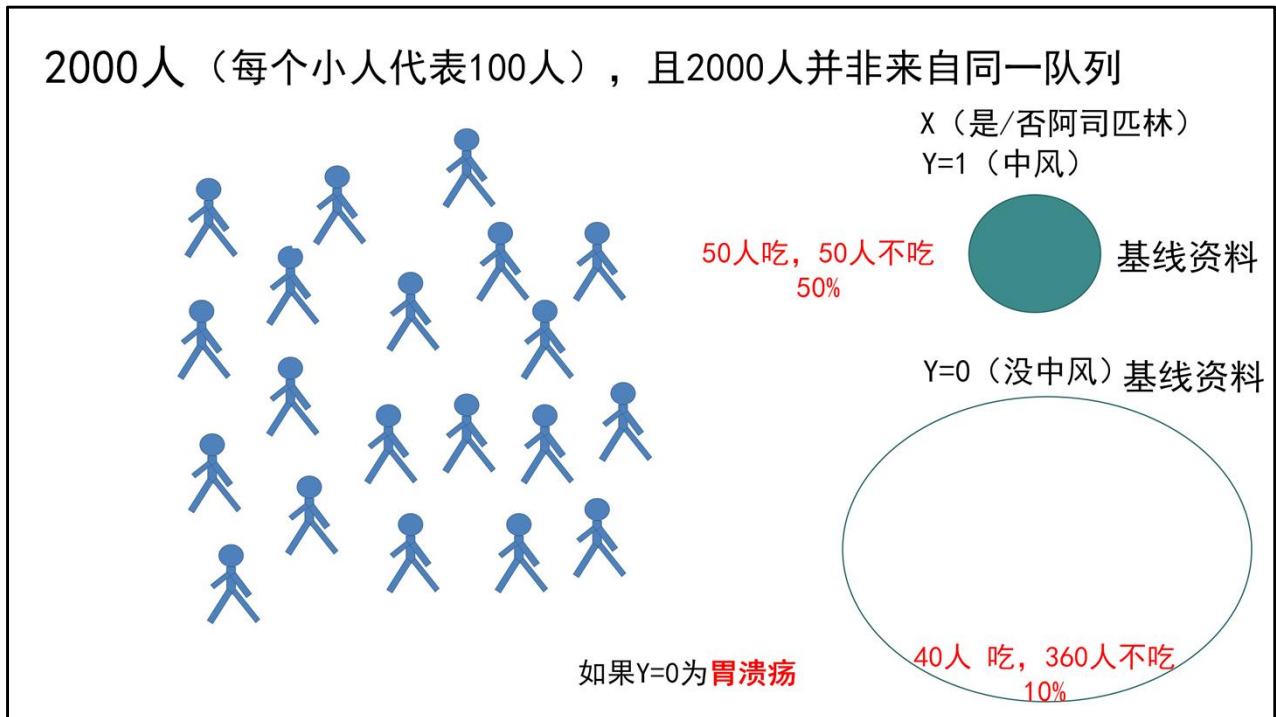
队列研究是先知道X，再来探讨Y。



而病例对照研究则不同。其在结局中找到发生事件者作为病例，在基于匹配的原则找到与其匹配因素相同，但未发生事件者作为对照。而最终所要研究的人群则变成了病例人群和对照人群。



如何理解对照独立于X？假设研究者想开展研究观察阿司匹林是否可预防中风。X为是否服用阿司匹林，Y=1的为中风，对照为没有中风者。此时，如果选择冠心病患者作为对照（未发生中风），则会得出阿司匹林可预防中风（因为冠心病患者需服用阿司匹林抗凝）



同理，假设研究者想开展研究观察阿司匹林是否可预防中风。X为是否服用阿司匹林，Y=1的为中风，对照为没有中风者。但研究者选择胃溃疡患者作为对照（未发生中风），则会得出阿司匹林可导致中风（因为胃溃疡患者不能服用阿司匹林抗凝）

Cutaneous melanoma cases diagnosed in 1999 and 2000 were ascertained through the Iowa Cancer Registry (...). Controls, also identified through the Iowa Cancer Registry, were colorectal cancer patients diagnosed during the same time. Colorectal cancer controls were selected because they are common and have a relatively long survival, and because arsenic exposure has not been conclusively linked to the incidence of colorectal cancer.

病例 (Y=1) 的是在1999-2000于爱荷华州诊断的黑色素瘤

对照 (Y=0) 的也是1999-2000于爱荷华州诊断的结直肠癌患者

为什么选呢？因为结直肠癌活的够长，而且X（砷）**没有证据证明与结直肠癌的发病有关**

从上面的实例可看出，如果X与对照有关联，则结果会被歪曲，操纵。从这个例子可以看出，研究者选择结直肠癌患者作为对照的原因是因为“没有证据表明砷暴露与结直肠癌发病有关。即X独立于对照。

Tips:

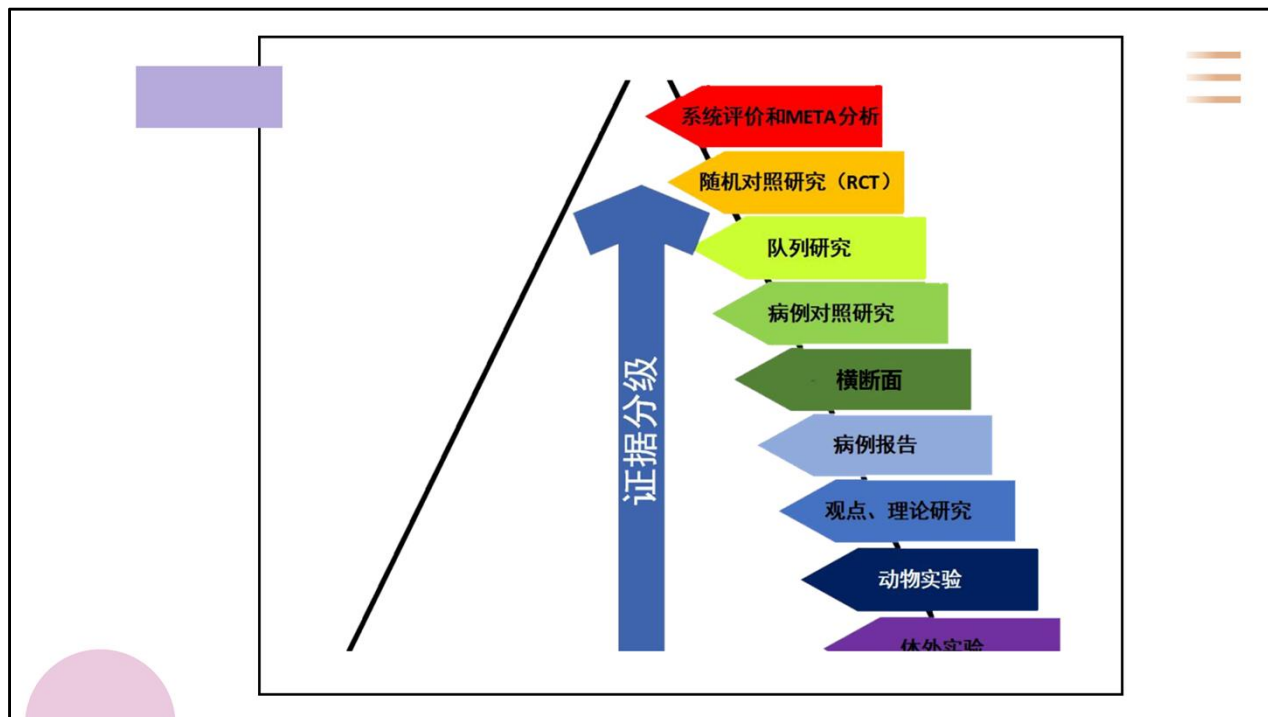
病例对照研究在设计上较简单，但对照的选择是非常难以判定且执行的。在有更好的选择前提下，何必？



04

队列研究





队列研究性价比最高。是我们在目前能力范围内可达到的最高证据等级的研究类型。

队列研究的定义

A

队列：你的所有研究对象可以用某种共同的特征去概括。

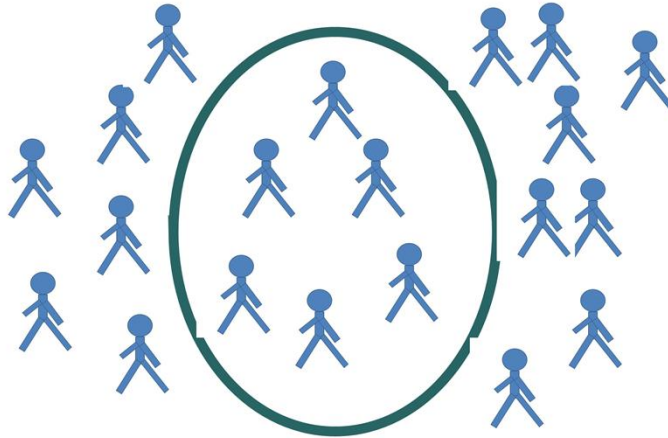
这种共同特征常见的是体征、疾病

B

队列研究：把这群具有共同特征的研究对象聚在一起做研究。

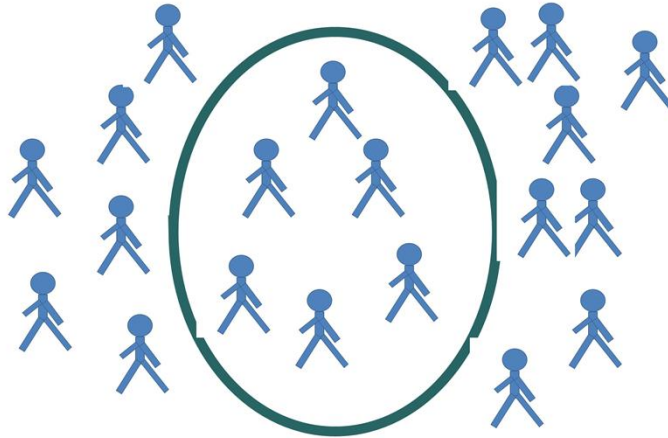
队列研究的步骤1——物以类聚人以群分

精髓：先想好你要做的这些研究对象有什么共同的特征。



队列研究的步骤2——设立纳排纯化队列

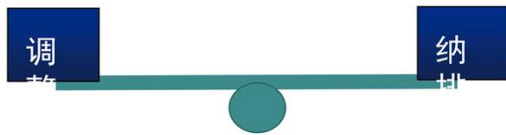
精髓：尽可能在保证样本量损失不大的前提下，纯化队列。



为什么都已经圈了人，还要设立纳排？

1000个研究对象（都是糖尿病）

5研究对象有艾滋病（995个没有艾滋病）



多1个Z（是否艾滋病），调整
样本量大，但统计学效能降低。
如果悬殊过大，数学调整往往失败。

艾滋病患者不纳入本研究
样本量减小，但统计学效能提高。
如果此类人过多，粗暴的排除，没人了。

疑问：我们如果设立纳排标准太多，是不是会降低人群的代表性？



01

在管好代表性之前，先保证内部正确性和可靠性。

02

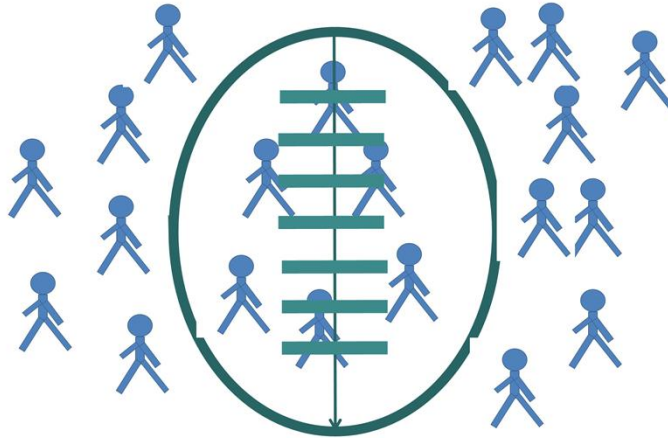
以我为主，你做你的，我做我的。

03

各为拼图，meta连接。

队列研究的步骤3——基线资料采集

精髓：只要第一次的。且要提前想好，要收集哪些基线资料





基线资料收集



什么叫基线资料?

患者进入队列时第一次产生的数据

该收集哪些?

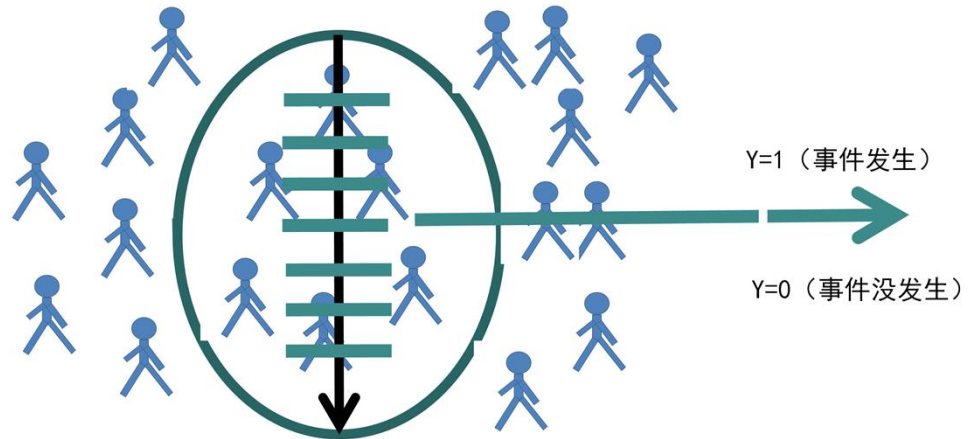
X, 必须收集;

Y, 必须收集;

Z: 临床经验+文献

队列研究的步骤4——随访

精髓：随访多久，随访结局的定义要说清楚



随访时间的判定

(1) 借鉴同类研究

同类（同Y）的研究，普遍的随访时间是多少？。

”

(2) 根据疾病的中位生存期

譬如，肝癌，半年，1年。 譬如乳腺癌：随访5年，10年。





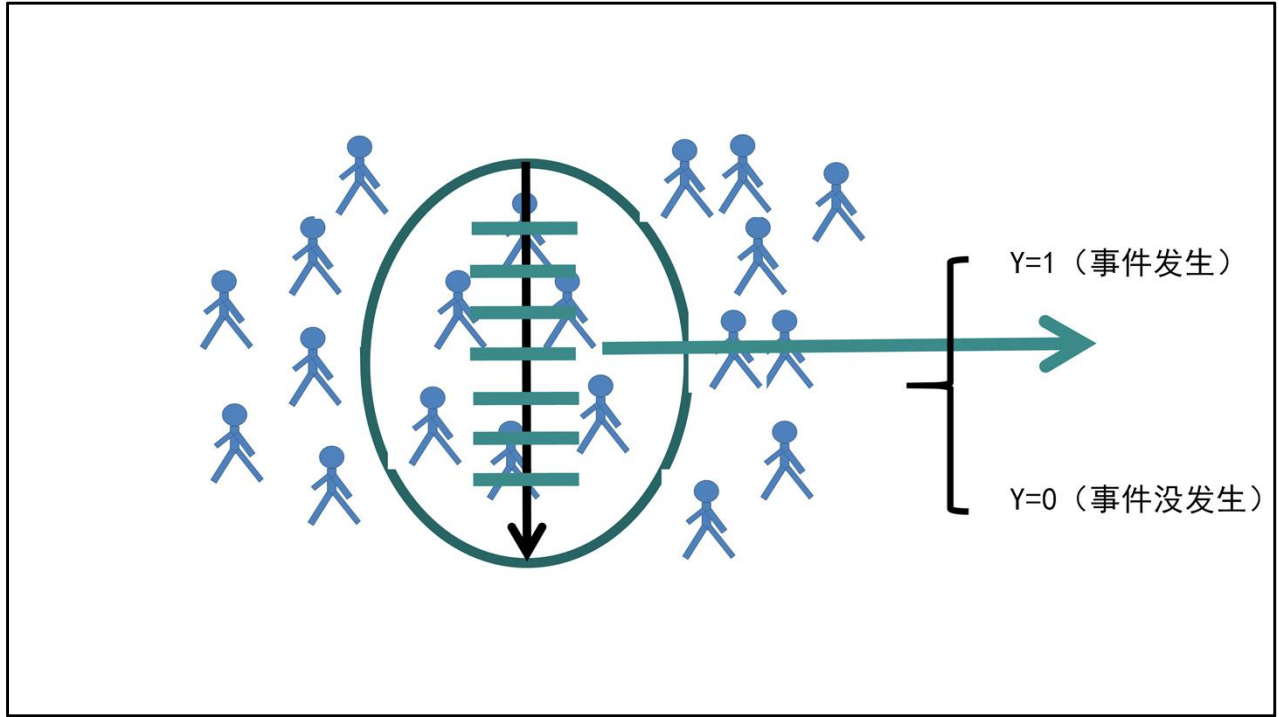
随访结局定义

01 选择那种可以客观测量，评估的。

02 常见的包括死亡，复发，治疗应答，副作用。

03 必要时，可以选择次级结局。

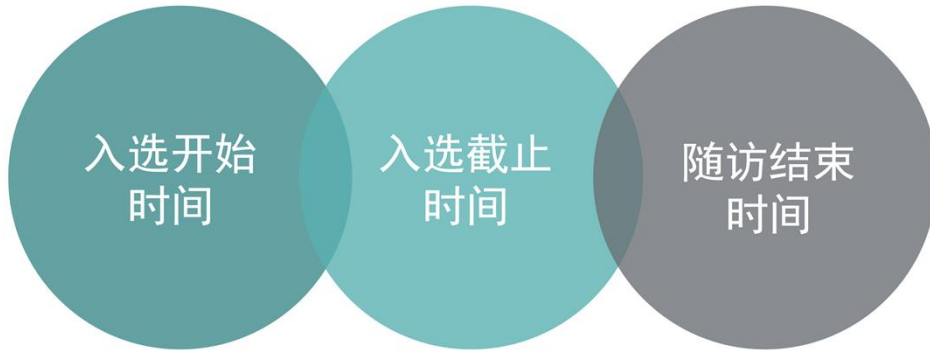




| 队列研究的参数



队列研究参数——时间

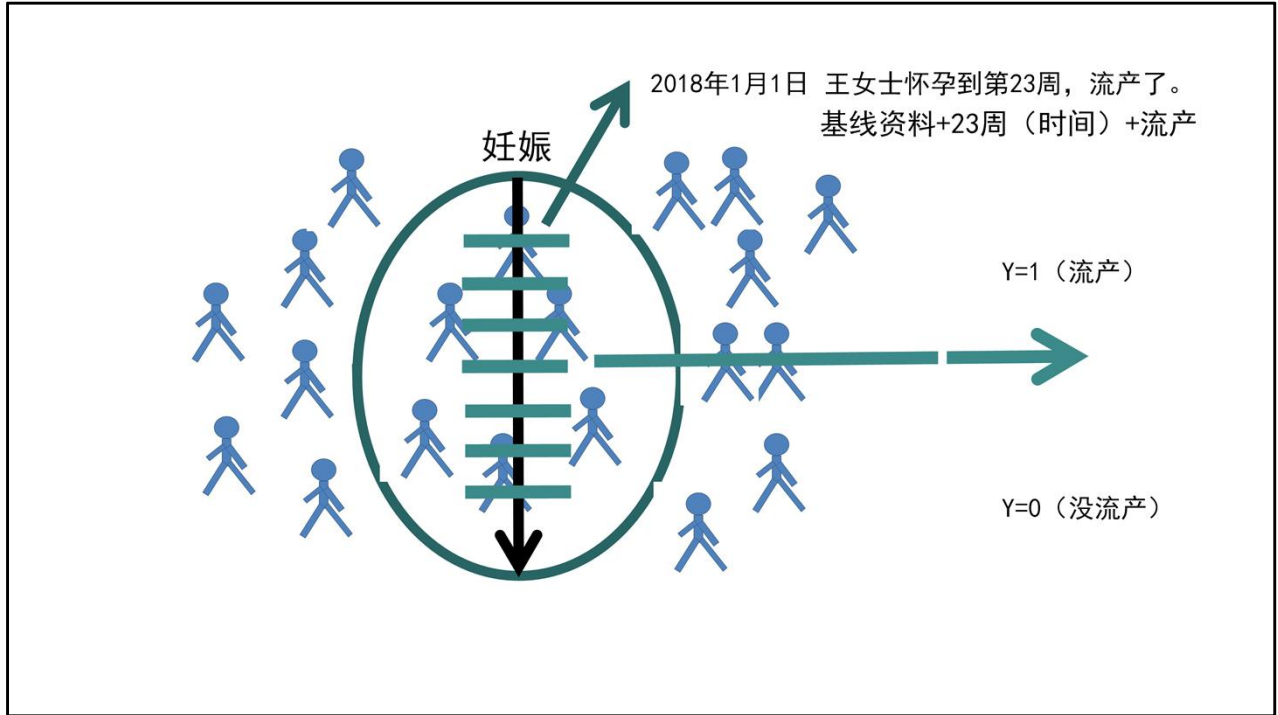


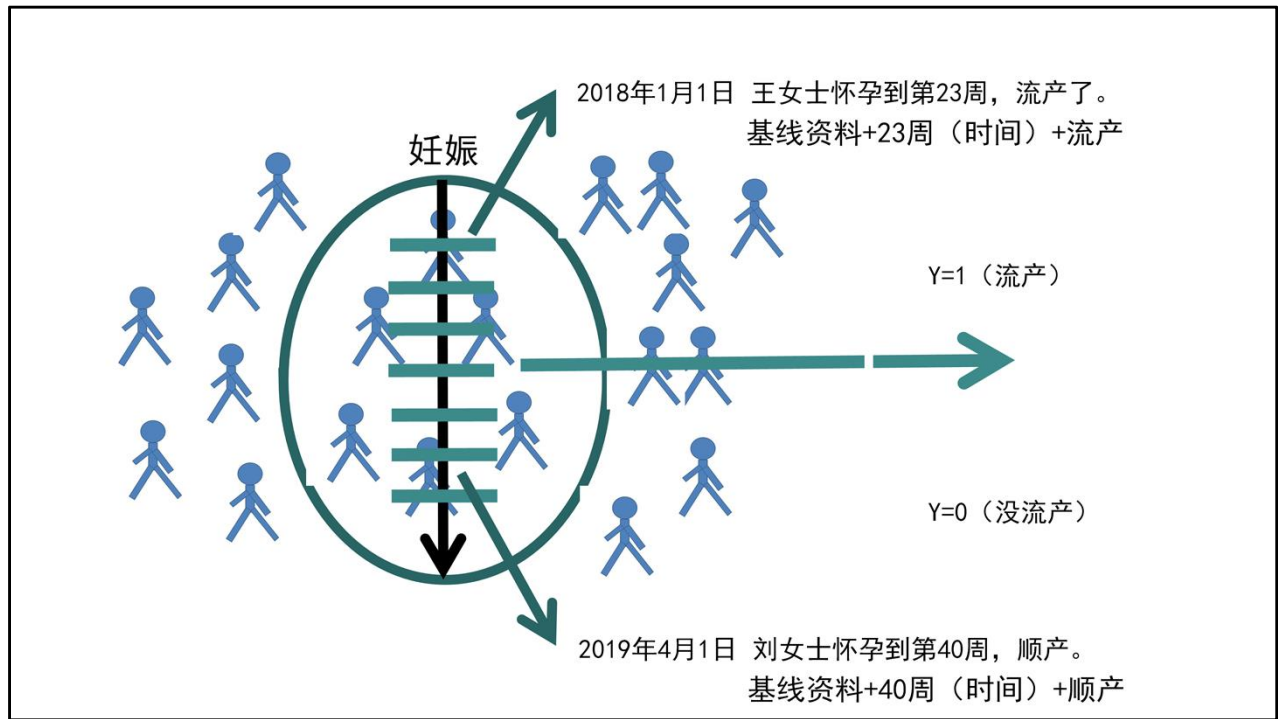
入选开始时间：第一个患者进入队列的时间 (entry time)

入选结束时间：最后一个患者进入队列的时间

随访结束时间：

随访结束时间不是具体日期，而是一个标准
妊娠结束、出ICU，出院。产后28天。





A	B	C	D	E	F	G	H	I
ID	name	Sex	age on pregnancy	G	P	流产史	怀孕时间	怀孕结局
1	王女士	2	24	1	1	0	23	1
ID	name	Sex	age on pregnancy	G	P	流产史	怀孕时间	怀孕结局
1	王女士	2	24	1	1	0	23	1
2	刘女士	2	28	1	2	1	40	0
3	陈女士	2	33	2	2	1	36	0

入选开始时间：2018年1月1日，

入选结束时间：2019年4月1日，

随访结束时间：随访至**每一位**研究对象妊娠结束（自然结束，不正常）

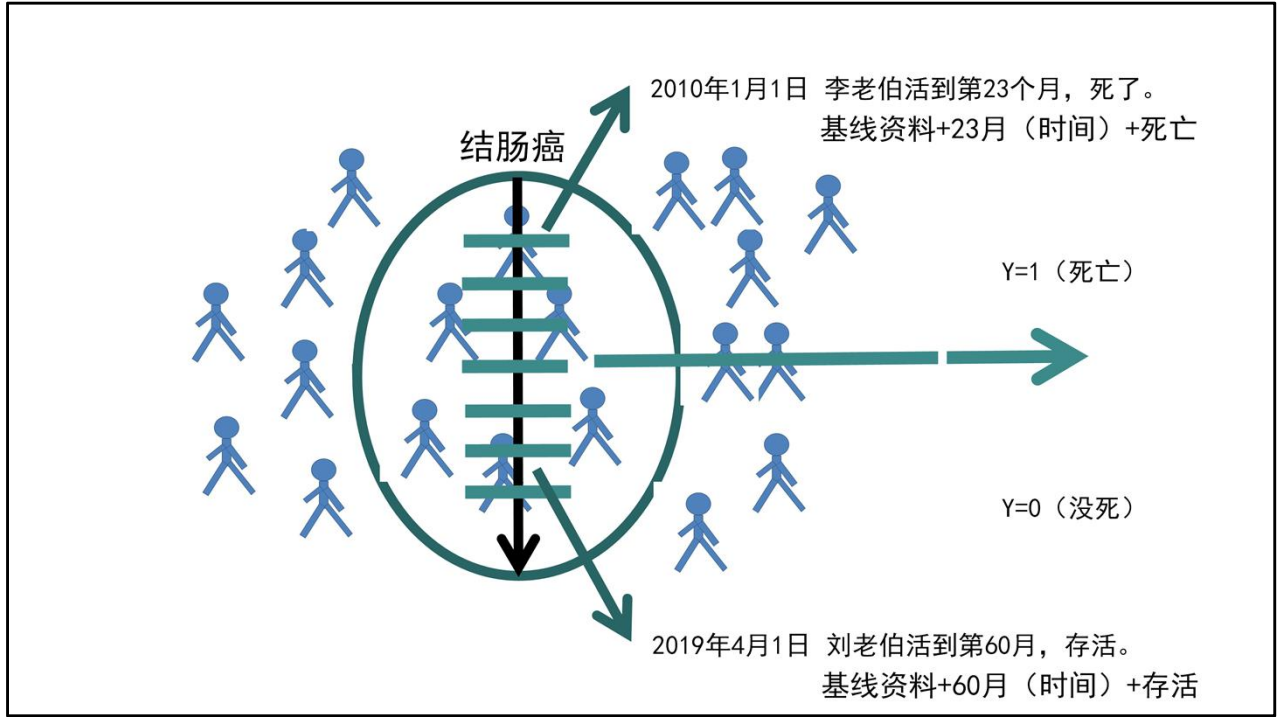
入选开始时间：第一个患者进入队列的时间（entry time）

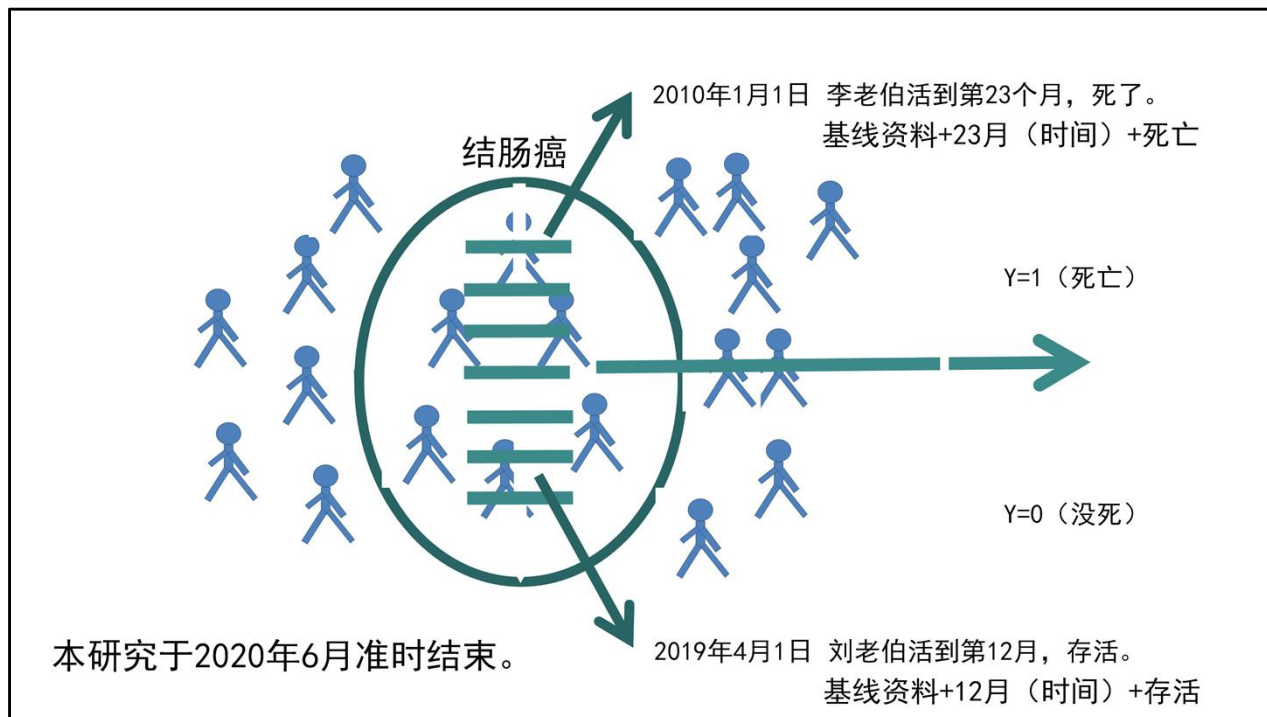
入选结束时间：最后一个患者进入队列的时间

随访结束时间：

随访结束时间不是具体日期，而是一个标准（少数）
妊娠结束、出ICU，出院。产后28天。

随访结束时间是具体日期（最常见）





考虑最短随访时间来定。

- 根据参考文献，确定最短随访时间，和入选结束时间
- 相加。
 - 入选开始时间：2010年1月1日，活了10年（120个月），没死
 - 入选结束时间：2019年4月1日，活了1年（14个月），没死
 - 随访结束时间：2020年6月1日

假设你现在准备写标书

- 一. 这个研究从多久开始做 (入组开始时间)
- 二. 我们科一年可以收大概多少人?
- 三. (入组开始时间+收到理想样本量所用的时间) = 入组结束时间
- 四. 所需最短随访时间是多久 (文献, 经验)
- 五. 最短随访时间+入组结束时间=随访结束时间。

队列参数——地点

- 按科室，医院（机构）名称，城市，国家的顺序表示
- Department of Hematology of the People' s Hospital of Guizhou province, Guiyang city, China
- 盲审目的，因此，只写大概，譬如 a hospital of China



队列参数——人物

- 队列特征
- 纳排标准，文献+实际情况，且可以随时修改
- 样本量：不可低于最低标准，最好大于最高标准



XYZ

- 一. 所测量的指标是否公认的指标，如果有很多，尽可能全要。
- 二. XYZ是何时测量的？（基线还是重复测量？）
- 三. XYZ的测量步骤（需要时才写）
- 四. 有无使用盲法
- 五. 有无使用一致性
- 六. 有无写清楚Z的纳入理由



Brachial-Ankle Pulse Wave Velocity: Myths, Misconceptions, and Realities

所测量的指标是否公认的指标，
如果有很多，尽可能全要



- 一. 以高分的同类研究作为参照，必须一致。
- 二. 如有很多，全要，其目的是敏感性分析，增加结果的稳定性。



测量指标是否公认的指标，需要谨慎的进行思考。如果有很多，则全部纳入而不挑选。譬如研究肥胖与死亡的关系，肥胖可用BMI，腰围，腰臀比等表示，该如何选择？以高分文献作为参照进行选择，如果有很多，则全要，最终全部呈现，以增加结果稳定性。譬如，如果要研究某指标与血管硬度的关系，如何代表血管硬度？查综述可知BAPWV是目前用于代表血管硬度的指标，可放心使用。



XYZ是何时测量的？（基线还是重复测量？）



新手建议：只要基线时测量的Z，因为只有基线的Z才能用常规的回归方程调整。



新手建议：X尽可能为基线，如果有条件，将其设置为重复测量亦可，并用变化率来做文章。



新手建议：Y尽可能选择二分类变量。多分类变量可进行合并。





Mean platelet volume at baseline and immune thrombocytopenia relapse in Chinese newly-diagnosed patients: a retrospective cohort study

Patients were followed at 3, 6, and 9 months, at 1 year, and every 6 months thereafter, at which times data on vital signs, including body weight,



譬如上述两篇例文，一篇对平均血小板体积的获取时间进行了阐明，为基线测量。另一篇则对体重进行了测量，但并非基线时测量，而是入院后（基线后）3，6，9，12个月。

XYZ的测量步骤(需要说明的才写)

A

对于采用了特殊方法或特殊形式测量的XYZ，写在正文中。

B

对于常规的普通指标，无需在正文中占篇幅，写在附件文件中。



有无用盲法

01

盲X

标本冻冰箱，最后统一同批次由独立第三方检测机构进行检测。

02

盲检测者

检测者（数据分析人员）并不知道患者的其他临床情况（诊断）



有没有一致性

一些主观的检查，譬如阅片，B超，病理等。

可以写成：由三位资历丰富的老专家一起诊断。只有当三个老专家一致，诊断才……，否则，少数服从多数，且老专家不知道患者其他临床数据。

一些特殊的检测指标：可以反复多次测量，并计算一致性，或者反复多次，如果误差不大，取平均值。



举例：

Published in final edited form as:

N Engl J Med. 2003 April 17; 348(16): 1517–1526. doi:10.1056/NEJMoa022848.

Intellectual Impairment in Children with Blood Lead Concentrations below 10 μg per Deciliter

Lead values were calculated as the means of six analyses of each sample (SD, 0.03 μg per deciliter [0.001 μmol per liter]). The results of repeated analyses, separated by five days.

本例为发表与新英格兰的范文。其主要研究儿童血铅含量与智商的关联。研究中，作者对于特殊指标（血铅）进行了6次测量，并最终取平均值代表儿童血铅。

写清楚纳入Z的理由

基于数据分析结果

基于文献

基于临床经验



随访策略

- 谁来随访（管理随访数据）
- 数据存放在哪里？
- 随访方式（电话，家访，随诊）
- 随访哪些指标
- 随访问隔



“

总结：

- 队列研究六要素，时间地点和人物，XYZ随访术。
- 地点根据杂志要求，略或详。排列是从小到大。
- XYZ要注意是否公认测量的方法，是否可靠（盲法，一致性等）
- 随访策略要讲清楚要点。

”





队列参数：伦理和知情同意

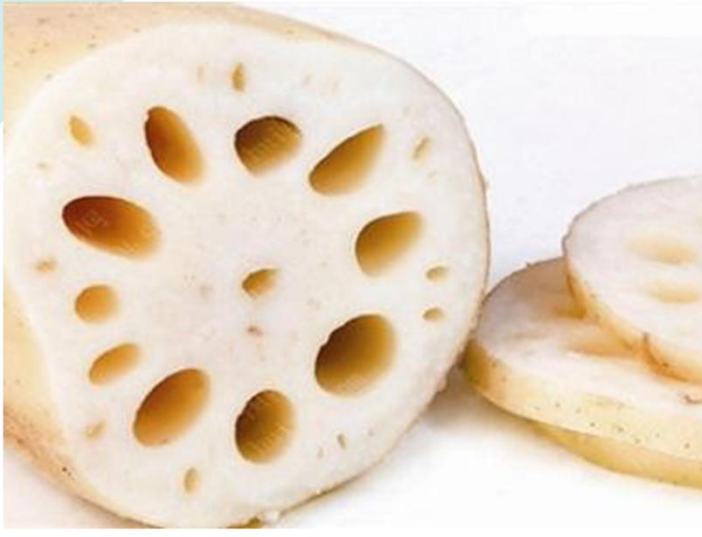
- 一. 所有队列必须获取伦理同意
- 二. 回顾性队列中，知情同意可豁免
- 三. 前瞻性队列中，必须获取知情同意





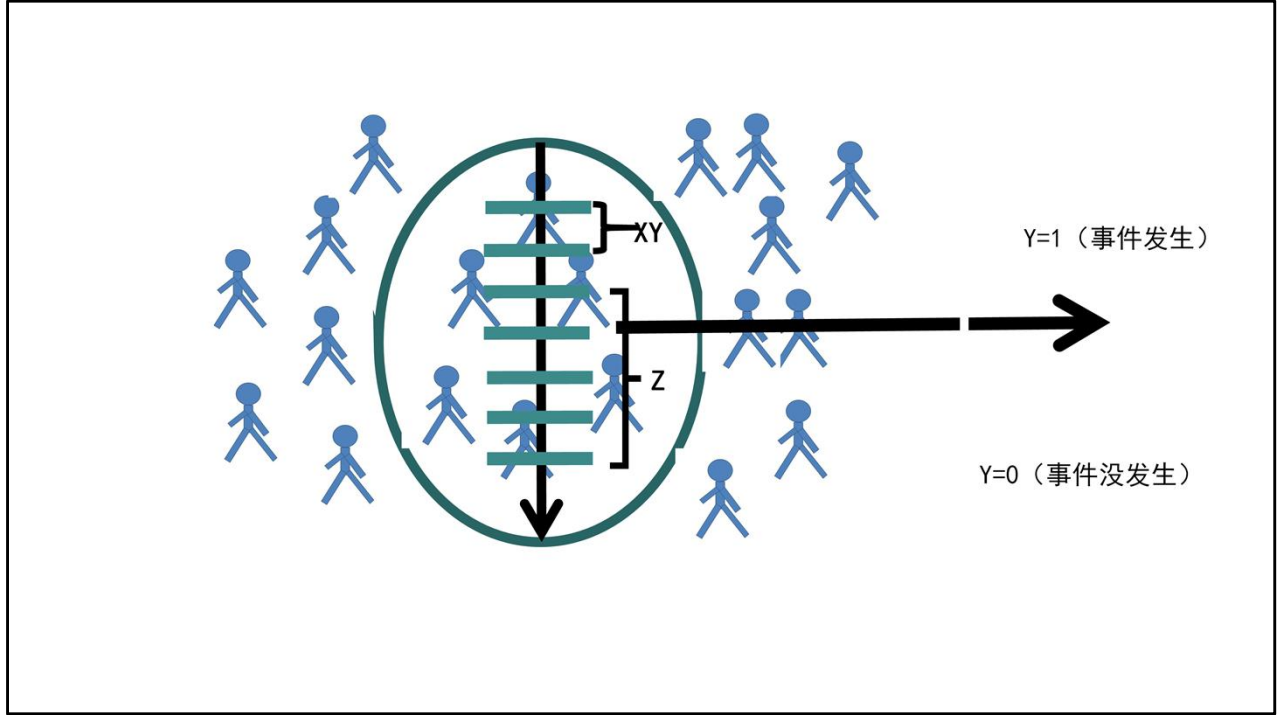
05

横断面研究

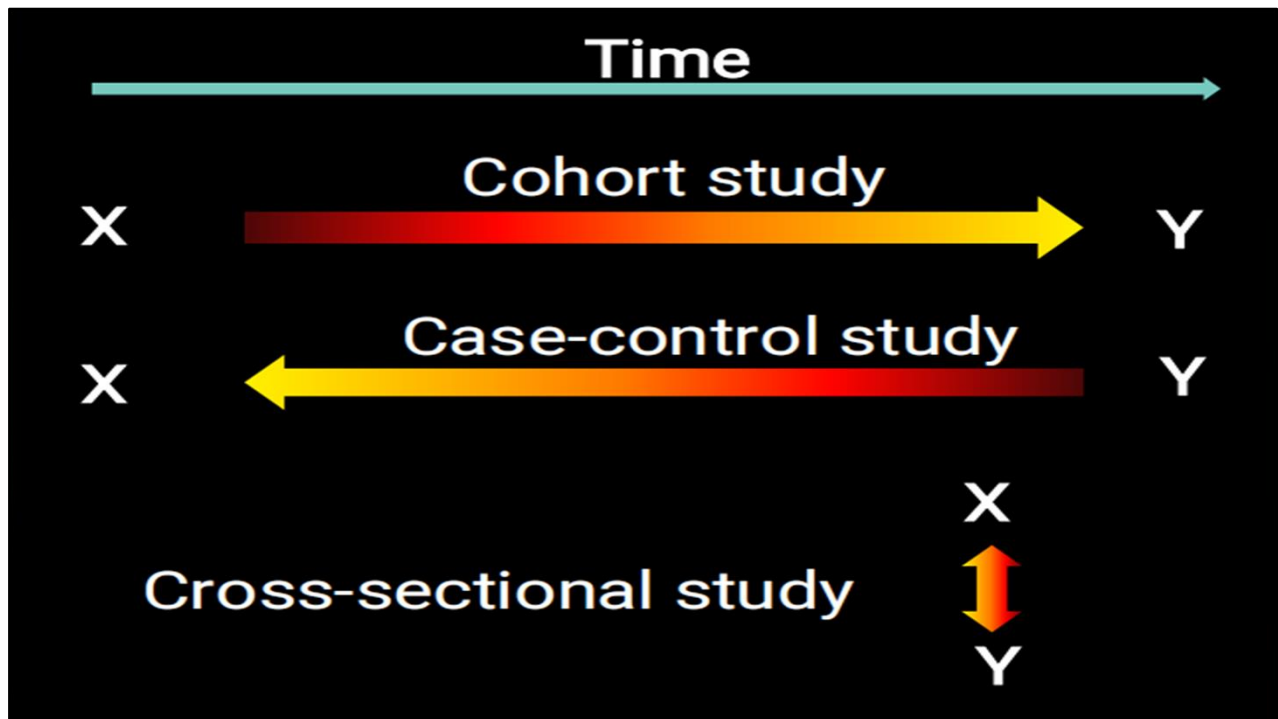


横 断 面

对横断面研究最直观的了解，就是理解“横断面”。何为横断面？这个剖开的莲藕就是横断面。



横断面的表现形式之一，为队列形式，但未随访，仅观察队列基线之间的相关关联，即XYZ均为基线测量，从而发生于同一时间点。



这幅图从时相上阐明了三种研究类型的区别。其中，横断面研究的XY发生于同一时间点。

优点：提高产出，更为灵活。

缺点1

证据级别弱（基因除外）

缺点2

有时候很难自圆其说。



横断面研究因为XY发生于同一时间点，因此，其证据级别弱。除非是基因相关类研究。此外，横断面研究的数据有时会出现不符合临床实际的发现，甚至背离常识。这是因为“健康者”效应导致的。在结果解读的课程中，我们会详细介绍。

- 一. 横断面研究，文章里面只能说X is associated with Y
- 二. 绝对不能出现 X is the potential risk factor之类的话。（除非你的XY是基因）

如何写关于横断面研究的研究理由



体重指数与CVD密切相关，是独立危险因素；LDL与CVD密切相关，是独立危险因素。因此，了解体重指数与 LDL之间的关联有助于帮助我们更好的了解CVD的发生发展机制，为立即CVD提供新的new insight



横断面研究面临的一个难点是在撰写论文时，如何在前言中提供一个理由，你为何要观察X和Y的联系？譬如，我们研究患者入院时体重指数（基线）与LDL的关联（基线）。此时，我们的话术可为：体重指数与CVD相关，LDL与CVD相关。那么，LDL与BMI是否有关呢？这可以……。

横断面福利-XY可以互换

- 我们目前常用的数据分析方法，很难分析多分类变量作为结局（好中差）。万一你的Y是这种，干脆和X互换。



总结

- 一. 横断面可以看成是没有随访的队列（常用）。
- 二. 不用担心失随访，数据利用率更高。
- 三. 证据级别弱，无因果（基因除外），只能说associated
- 四. 逻辑可以扯结局的危险因素
- 五. XY可以互换



06

真实世界研究





真实世界研究目前尚无完善的指南。但GRACE原则可供参考。

Target study population

1. 设置纳排标准时，需考虑排除的患者，会对结果带来怎样的影响。

例：当我们在研究某种药物是否可预防冠心病时，我们排除了吸烟者。

X=是否吃药； Y: Incidence of MI

人群：non-smoker population

(该药物与吸烟密切相关)

Target study population

2. 治疗组 vs 对比组的人群是否存在系统偏差。

例：当我们在研究放疗+化疗 vs 单纯化疗，

X=放疗加化疗 vs 化疗 alone; Y: Mortality of CRC

人群：newly-diagnosed CRC patients

放疗加化疗的人群 和 只用化疗 的人群，是有选择偏倚的。

Target study population

3. 是否存在“健康者效应”

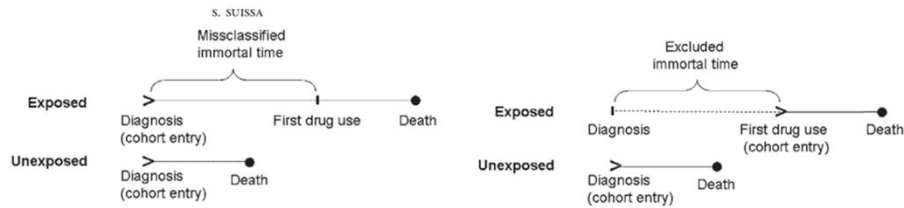
例：当我们在研究某种新药治疗肿瘤的疗效，人群是长期使用该药物的人。

既往使用过程中，所谓的药物使用者已经经过筛选（不适合这个药的，已经死了）。

Grace 指南：To the extent feasible, it is desirable to consider focusing CE studies on new users of the treatments of interest.

X (治疗方案)

1. Time to diagnosis



对于治疗方案，需要注意“不朽偏倚”。不朽偏倚指的是患者的随访时间从诊断开始进行计算，但诊断和用药之间尚有一段时间，因此，真正的随访时间应该从首次服药时间开始算起。对于随访时间较短的结局，不朽偏倚是一定要排除的。

X (治疗方案)

2. Time vary recording

好的疗效比比较研究通常反映了在实际实践环境中使用的干预措施的复杂性。

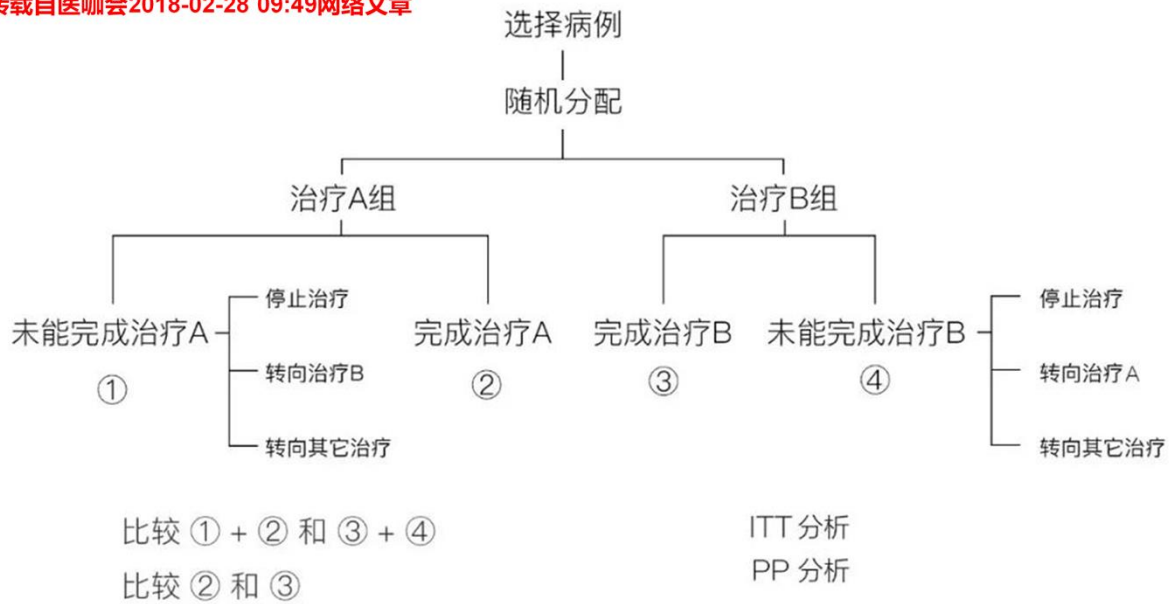
因此，如何兼顾实际临床场景下干预措施的复杂性？time-vary，或者

采用敏感性分析

- (1) 现有方法学不支持time-vary的真实世界研究（PSM, IPTW等）
- (2) 从ITT中获得灵感

既往讲队列时，我们对于X往往为一次性测量，基线获取。然而，临床变化复杂，常涉及到药物使用状态的更迭，变化。incidence，采用随时间变化的测量策略，可更好的兼顾。

转载自医咖会2018-02-28 09:49网络文章



譬如，可以参照RCT中的ITT分析及PP分析原理，对其进行敏感性分析：即按计划完成治疗者计算一个结果，全部患者（未能按计划完成者+按计划完成治疗者）再算一次，结果是否会有不同？

X (治疗方案)

3. 在理想和现实找到平衡点

好对于每个比较剂，考虑品牌、剂量、给药方法、使用时间、是否标注适应症、目前正在使用的治疗方案、必要信息被准确记录和获取的可能性等。

但需注意：必须兼顾可完成性与指标的完整性。一般对于单人，需尽快出结果的团队，可采用“是否”用药，而无需记录太复杂（品牌、剂量等）。

X (治疗方案)

4. 错分类

Non-differential misclassification is unrelated to the outcome and will usually bias the estimated effect of intervention towards the null

Differential misclassification occurs when misclassification of intervention status is related to the outcome or the risk of the outcome, and is likely to lead to bias

真实世界研究，往往会面临错分类。但要区分是有意识的错分类还是无意的。如果错分类的发生对于不同治疗策略组是平等的，则该结果不会导致偏倚。如只是针对某一种治疗策略，则需要分析其偏性的方向。



譬如调查青少年吸烟状态（吸烟与不吸烟）对肺癌的影响。如果是用调查问卷，则很可能出现误分类。很多年轻人怕家长责备等，明明有吸烟习惯，却谎报为不吸烟。最终，导致吸烟的人比实际估计的少，造成偏性。因为这种错分类主要出现于吸烟者。

Y (结局变量)

1. 阐明其类型、临床代表的意义，用于评估疗效是否合适

可参照RCT的结局变量以及其他同样研究此类疾病的文献，并根据临床实际情况进行选择。直接终点优于中间终点、替代终点或复合终点。

2. 针对中间终点：

当优质数据可连接中期和长期结果时、时间或成本的限制，研究长期结果不可行时，中期终点则是有用的。

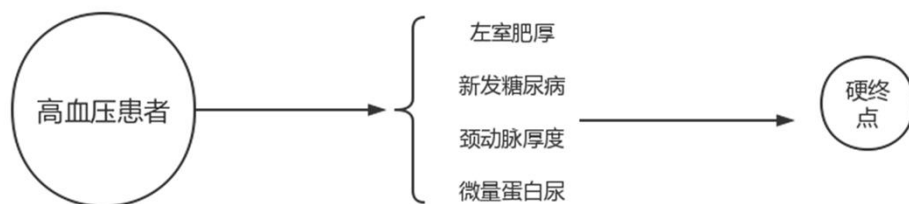
3. 针对替代终点：

不管干预的作用机制如何，结果对所有比较者都是有意义和有效的。当使用有效性的替代指标时(例如，骨密度作为骨折的替代指标)，考虑这一点尤为重要。

4. 如果使用中间终点，替代终点一定要有高分参考文献，有人用过。

5. 标准化、可重复和可独立验证的结果的定量测量比临床印象和/或其他测量更可取。只要这些测量在临床上是重要的或者是经过验证的替代测量。

中间终点



理由：硬终点的获益是建立在血压控制基础上的。

譬如对于高血压患者，如果要随访至其发生高心病，高血压肾病、中风，可能需要较长时间。因此，可以用中间终点（左室肥厚等）进行替代，从而降低随访的难度和花费。

替代终点



注意：比较的药物都会对EGFR造成影响。

再譬如，如果随访糖尿病肾病患者进展为ESRD，很可能在短时间内不能看到结果。如果随访指标变为EGFR，则更容易实施。

Z (协变量)

常见需考虑因素:

- a. 疾病诊断确定性、严重程度、显著的共病、治疗史
- b. 患者特征——年龄、性别、种族、社会经济状况、地理位置、获得医疗服务的机会(例如, 健康保险覆盖范围)
- c. 非处方药的信息, 这些产品可能是重要的混杂因素。



谢谢聆听

谢谢