



数据分析结果的 临床意义解读

易侬学院 陈驰

各位学友，今天由我来分享“数据分析结果的临床意义解读”章节

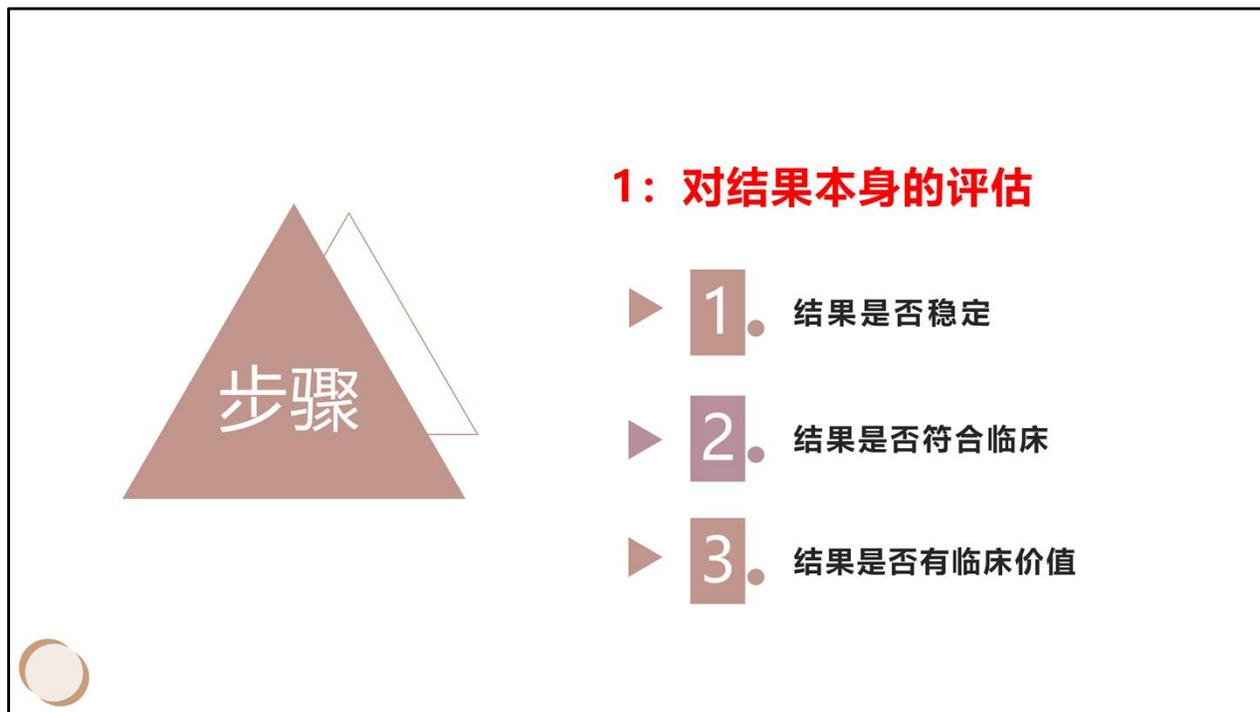
“

当数据分析结果出来以后，我们会下意识的直接开始写文章。然而，中间却少了一个步骤：
对“**证据力度的思辨**”。

”



数据分析结果和论文写作之间，尚有一个环节被忽视，即“对证据力度的思辨”



首先，我们需对数据分析所获结果进行评估，评估从稳定性、是否符合临床、是否具有临床价值三个方面进行评估。

结果是否稳定——阳性结果

只有XY的方程（不纳入其他的协变量）	Non-adjusted model
只调整人口学方程（只调整性别年龄）	Minimally-adjusted model
完全调整模型（调整所有的Z）	Fully-adjusted model

如果效应值大小，方向稳定，变化不大。结果稳。

如果效应值大小，方向不稳定，变化大。结果不稳。

针对阳性结果，我们评估结果是否稳定主要通过不同调整策略中，效应值及其95%可信区间变化趋势。

X = HIV * MDD
Y = 是否心衰

Table 3
Cox proportional hazard regression models examining the association between HIV/MDD group and incident heart failure

Characteristic	Model 1 HR (95% CI)	Model 2 HR (95% CI)	Model 3 HR (95% CI)
HIV / MDD Group			
HIV -- No MDD	1 [Reference]	1 [Reference]	1 [Reference]
HIV -- MDD	1.13 (1.00 – 1.27)	1.30 (1.15 – 1.47)	1.19 (1.05 – 1.35)
HIV + No MDD	1.25 (1.15 – 1.37)	1.32 (1.20 – 1.44)	1.28 (1.16 – 1.41)
HIV + MDD	1.54 (1.34 – 1.77)	1.87 (1.62 – 2.15)	1.68 (1.45 – 1.95)

Model 1: HIV MDD only
Model 2: Model 1+ age, sex, race/ethnicity
Model 3: Model 2+ all covariates

Abbreviations: BMI, body mass index (calculated as kg/m²); eGFR, estimated glomerular filtration rate; HCV, hepatitis C virus; HDL, high-density lipoprotein; HIV, human immunodeficiency virus; HR, hazard ratio; LDL, low-density lipoprotein; MDD, major depressive disorder

如这篇来自循环杂志的表格中，研究者呈现了三个模型，并根据不同调整策略观察效应值及其可信区间的变化趋势。可看到在三个模型中，效应值的数字、方向及可信区间的范围变化方向一致，且数字的变化幅度并不大。因此，可认为该结果稳定。

不稳

不稳：三个模型（未调整、微调整，完全调整模型）差别甚大

Table 3 Relationship between TG/HDL-C and baPWV in different models

Variable	Crude model (β , 95%CI, P)	Minimally adjusted model(β , 95%CI, P)	Fully adjusted model(β , 95%CI, P)
TG/HDL-C	14.9 (7.2, 22.5) <0.001	12.7 (5.8, 19.6) <0.001	1.8 (-5.0, 8.5) 0.609
TG/HDL-C(quartile)			
Q1	Ref	Ref	Ref
Q2	65.1 (20.6, 109.6) 0.004	41.4 (2.2, 80.5) 0.039	28.7 (-5.5, 63.0) 0.101
Q3	90.1 (45.6, 134.6) <0.001	69.2 (28.9, 109.6) 0.001	14.2 (-23.3, 51.7) 0.459
Q4	128.3 (83.9, 172.8) <0.001	108.1 (66.3, 149.9) <0.001	35.1 (-5.2, 75.4) 0.088
P for trend	<0.001	<0.001	0.171

Crude model: we did not adjust other covariants

Minimally adjusted model: we adjusted age and sex

Fully adjusted model: we adjusted age, sex, BMI, SBP, DBP, AST, ALT, GGT, uric acid, fasting glucose, TC, LDL, eGFR, smoking and exercise status, fatty liver, alcohol consumption, ABI

CI confidence interval, Ref reference

而这个例子中，与上篇循环杂志的结果完全不同，可看到不同调整策略中，效应值的大小降低了近10倍，且可信区间的范围及方向也发生巨大的变化

结果不稳定的原因：

有一个混杂因素，对XY的影响太大，甚至导致结果相反。

结果不稳定多提示有一个混杂因素，对于我们结果的影响太大，甚至可以导致结果完全相反，

$$Y = \text{截距} + aX + bz_1 + cz_2 + dz_3 + Z_4$$

Table 3 Relationship between TG/HDL-C and baPWV in different models

Variable	Crude model (β , 95%CI, P)	Minimally adjusted model(β , 95%CI, P)	Fully adjusted model(β , 95%CI, P)
TG/HDL-C	14.9 (7.2, 22.5) <0.001	12.7 (5.8, 19.6) <0.001	1.8 (-5.0, 8.5) 0.609
TG/HDL-C(quartile)			
Q1	Ref	Ref	Ref
Q2	65.1 (20.6, 109.6) 0.004	41.4 (2.2, 80.5) 0.039	28.7 (-5.5, 63.0) 0.101
Q3	90.1 (45.6, 134.6) <0.001	69.2 (28.9, 109.6) 0.001	14.2 (-23.3, 51.7) 0.459
Q4	128.3 (83.9, 172.8) <0.001	108.1 (66.3, 149.9) <0.001	35.1 (-5.2, 75.4) 0.088
P for trend	<0.001	<0.001	0.171

Crude model: we did not adjust other covariants

Minimally adjusted model: we adjusted age and sex

Fully adjusted model: we adjusted age, sex, BMI, SBP, DBP, AST, ALT, GGT, uric acid, fasting glucose, TC, LDL, eGFR, smoking and exercise status, fatty liver, alcohol consumption, ABI

CI confidence interval, Ref reference

可想而知，假设方程中巨大混杂Z4的出现，则无法观察到X与Y的关系。正如在很亮的房屋中点燃蜡烛，能觉得亮么？这个混杂怎么找？易侬的筛选协变量模块可以把大混杂给找出来。

注意：

- 低分（<5分）的SCI很少会注意到效应值变化不同。
- 因此，很多时候即使不稳，也会被审稿人视而不见。
- 教给大家，只不过是让大家有所防备。
- 无需刻意或特意去就这个Z做文章。因为这是高手玩的花活。

很多审稿人不太会注意到这种效应值的变化，所以很多时候即使你的结果不稳，也会被审稿人视而不见。在此提出，只是让大家有所防备，避免投稿后被审稿人问道，有应对方案。

解决方案1：在文章中说明结果受到了影响。既往的结果是假象。

应用场景：

- 一．既往的结果有争议
- 二．你调整的协变量，此前的文章没有调整。

如何利用结果不稳？第一种办法，如果既往结果有争议，而你找到的这个“大混杂”恰好对方都没调整，那么，这个结果反而是亮点。

RESEARCH

Antibiotics in fetal and early life and subsequent childhood asthma: nationwide population based study with sibling analysis

 OPEN ACCESS

Anne K Örtqvist *doctoral candidate*¹, Cecilia Lundholm *statistician*¹, Helle Kieler *associate professor*², Jonas F Ludvigsson *professor*^{1,3}, Tove Fall *associate professor*⁴, Weimin Ye *professor*¹, Catarina Almqvist *professor*^{1,5}

譬如这篇文章中，旨在探讨婴儿期抗生素使用与日后哮喘的关联

1.78 to 3.13) and disappeared for antibiotics for urinary tract and skin (0.85, 0.47 to 1.55).

Conclusions Previous positive associations between exposure to antibiotics in fetal and early life and subsequent childhood asthma could have been caused by confounding by shared familial factors, in addition to confounding by respiratory infections.

- 既往研究说婴儿期与幼儿期接触抗生素与儿童期哮喘有关不对的。因为这个结果可能是由于家族因素和呼吸道感染混杂所造成的假象。

作者指出哮喘家族史和呼吸道感染既往文章未调整，而调整后则无法观察到抗生素与哮喘发病的关联，所以说此前同类研究得到的就是一个假象。

解决方案2：直接换X

用找到的这个影响非常大的Z，直接作为X重新写一篇。

第二方案直接换一个X。此前有一位学友他的文章被修回时，审稿人指出效应值不稳的原因很可能是巨大混杂因素的干扰，当找到后，审稿人直接给出换一个x的建议，即将这个巨大混杂当成X来写。

结果是否稳定——阴性结果

只有XY的方程（不纳入其他的协变量）	Non-adjusted model
只调整人口学方程（只调整性别年龄）	Minimally-adjusted model
完全调整模型（调整所有的Z）	Fully-adjusted model

如果效应值大小，方向稳定，变化不大。结果稳。

如果效应值大小，方向不稳定，变化大。结果不稳。

而当我们得到阴性结果时，结果稳定同此前的阳性结果。

“

- 样本量因素：很多时候，我们对“阴性”结果的描述不谨慎。我们写的是X对Y没有作用。然而，审稿人很可能会说：这种所谓的没有作用是否是因为样本量不够所致？你有多少把握说这种话？
- 正确的方法应该是：X is not significantly associated with Y, 或no substantial

”



但需注意，很多时候我们对阴性结果的描述并不严谨，在直接描述x和y没有关系时，审稿人会质疑说就阴性结果的原因很可能是样本量不够。这个时候我们就不能写x is not associated with y，而应该写 X is not significantly (并非显著相关)，或用 substantial这个词去描述。当你换一个话术，审稿人就不会有这样的质疑了。



陈老师，打扰您了，想请教您一个问题，就是我的一个文章回修了，是两组间率的比较（比较两个用药方案间的活产率），得到了一个阴性结果。审稿人给我的一个意见里面，让我计算power 看样本量够不够大，原话是 “Is the study powered for the primary outcome and the sample size big enough?”

我用软件算了一下，发现目前样本量下的power特别低，只有百分之十几；按目前两组的比率，要想得到两组差异，需要8000多的样本量，而我只有700多样本。想请教一下您，这个问题该怎么回复审稿人呀？麻烦老师啦~

这是此前的一个案例，审稿人提出：你的power不够大，因此，阴性结果是你的样本量不够导致的。而在原来的话术 x is not associated with y 中加上一个单词 significant 或者 substantial 这样的词，就可以规避掉审稿人的质疑。



“

此外，对于阴性结果一定要做交互作用检验，将所有的Z均做一次交互作用。如果在所有的层之间都看不到效应值的明显变化，则是真的阴性。

”



此外，必须对阴性结果在不同分层中进行检验，看这种“阴性”结果是否真的稳定存在。



举例：阴性结果的呈现

题目：

Statin Use and Survival After Colorectal Cancer: The Importance of Comprehensive Confounder Adjustment

X: 他汀使用

Y: 结肠癌的生存

特殊说明：本研究最大的亮点：confounder adjust

譬如这篇他汀和结肠癌的生存率的研究，作者得到的是个阴性结果，但是他为了证明这个阴性结果，专门的做了一个分层分析，在不同年龄不同性别分别比较X和Y的关系。而所有的结果都是阴性结果，也就是说该结果适用于所有的人群，它不是一个偶然发现，而是一个稳定的结果。

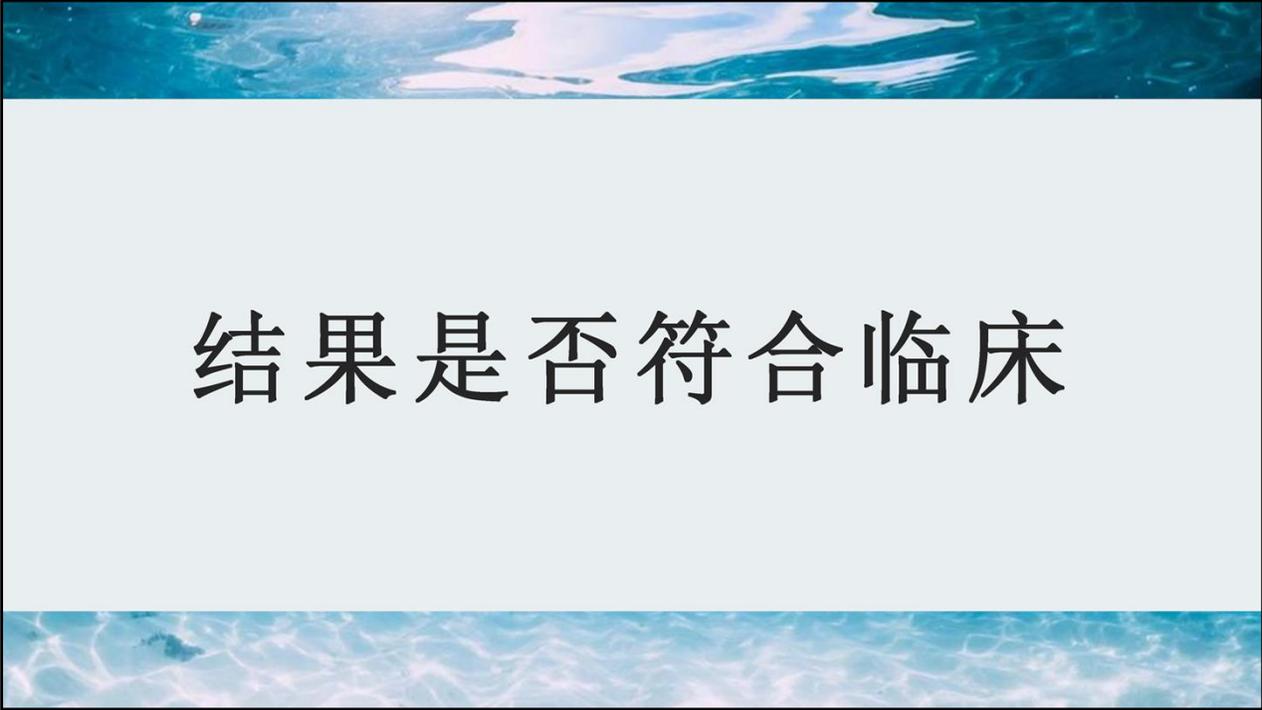
Table 3. Association of statin use at diagnosis with overall, colorectal cancer-specific, and recurrence-free survival stratified by age, sex, stage and location of colorectal cancer at diagnosis, and by conduct of chemotherapy

Subgroup	Statin use	Overall survival			CRC-specific survival		Recurrence-free survival*		
		No.	Events No. (%)	HR (95% CI)†	Events No. (%)	HR (95% CI)†	No.	Events No. (%)	HR (95% CI)†
Age at diagnosis, y									
<70	No	1257	290 (23)	1.00 (Ref.)	237 (19)	1.00 (Ref.)	1057	202 (19)	1.00 (Ref.)
	Yes	184	40 (22)	1.10 (0.70 to 1.74)	31 (17)	1.28 (0.77 to 2.15)	164	25 (15)	1.15 (0.65 to 2.03)
≥70	No	1028	361 (35)	1.00 (Ref.)	246 (24)	1.00 (Ref.)	880	211 (24)	1.00 (Ref.)
	Yes	228	78 (34)	1.19 (0.87 to 1.61)	49 (21)	1.13 (0.77 to 1.65)	204	39 (19)	0.85 (0.55 to 1.32)
Sex									
Male	No	1341	380 (28)	1.00 (Ref.)	276 (21)	1.00 (Ref.)	1125	237 (21)	1.00 (Ref.)
	Yes	272	74 (27)	1.00 (0.71 to 1.39)	46 (17)	0.93 (0.62 to 1.41)	243	37 (15)	0.80 (0.49 to 1.30)
Female	No	944	271 (29)	1.00 (Ref.)	207 (22)	1.00 (Ref.)	812	176 (22)	1.00 (Ref.)
	Yes	140	44 (31)	1.32 (0.89 to 1.96)	34 (24)	1.49 (0.94 to 2.35)	125	27 (22)	1.11 (0.66 to 1.84)

这就是作者分层关系验证的图表

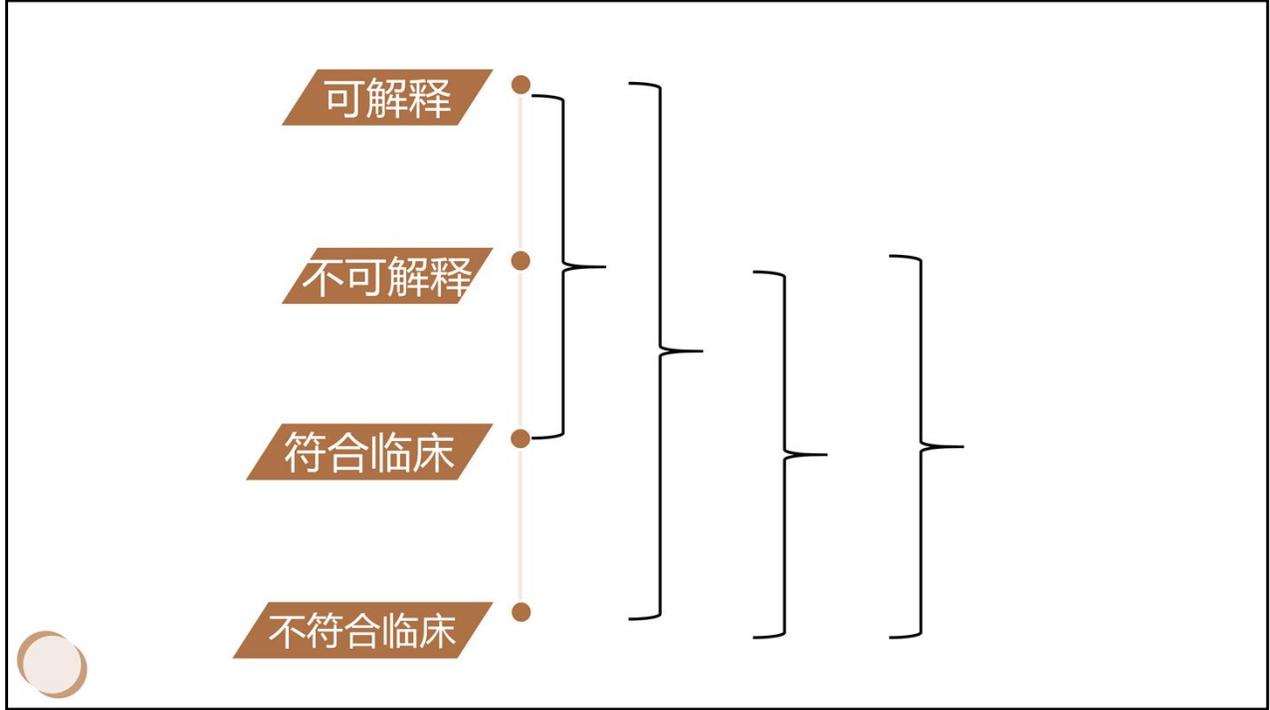
UICC stage									
Stage I + II	No	1165	174 (15)	1.00 (Ref.)	82 (7)	1.00 (Ref.)	1165	158 (14)	1.00 (Ref.)
	Yes	243	34 (14)	1.07 (0.66 to 1.72)	10 (4)	0.97 (0.43 to 2.19)	241	16 (7)	0.50 (0.26 to 0.95)
Stage III	No	771	210 (27)	1.00 (Ref.)	155 (20)	1.00 (Ref.)	771	255 (33)	1.00 (Ref.)
	Yes	126	48 (38)	1.14 (0.75 to 1.75)	35 (28)	1.22 (0.75 to 2.00)	126	48 (38)	1.25 (0.82 to 1.92)
Stage IV	No	348	267 (77)	1.00 (Ref.)	246 (71)	1.00 (Ref.)	--	--	--
	Yes	43	36 (84)	1.04 (0.67 to 1.63)	35 (81)	1.07 (0.68 to 1.70)	--	--	--
Location of CRC									
Proximal colon	No	701	205 (29)	1.00 (Ref.)	141 (20)	1.00 (Ref.)	604	110 (18)	1.00 (Ref.)
	Yes	145	47 (32)	1.08 (0.71 to 1.64)	32 (22)	1.01 (0.60 to 1.69)	124	18 (15)	0.78 (0.48 to 1.26)
Distal colon	No	631	175 (28)	1.00 (Ref.)	129 (20)	1.00 (Ref.)	523	103 (20)	1.00 (Ref.)
	Yes	110	27 (25)	0.95 (0.57 to 1.59)	17 (15)	1.01 (0.53 to 1.95)	98	17 (17)	1.24 (0.71 to 2.14)
Rectum	No	946	267 (28)	1.00 (Ref.)	209 (22)	1.00 (Ref.)	805	198 (25)	1.00 (Ref.)
	Yes	156	44 (28)	1.25 (0.84 to 1.87)	31 (20)	1.35 (0.84 to 2.15)	145	29 (27)	1.16 (0.76 to 1.76)
Chemotherapy									
No	No	1325	330 (25)	1.00 (Ref.)	221 (17)	1.00 (Ref.)	1195	199 (17)	1.00 (Ref.)
	Yes	272	66 (24)	1.29 (0.93 to 1.79)	36 (13)	1.26 (0.82 to 1.93)	256	30 (12)	0.99 (0.67 to 1.47)
Yes	No	954	317 (33)	1.00 (Ref.)	258 (27)	1.00 (Ref.)	738	212 (29)	1.00 (Ref.)
	Yes	139	52 (37)	1.03 (0.69 to 1.52)	44 (32)	1.08 (0.70 to 1.67)	112	34 (30)	1.21 (0.82 to 1.78)

同上



结果是否符合临床

第二是我们需要判定一下结果是否符合临床



结果做出来无非就只有4种可能，即：可解释不可解释 vs 符合临床不符合临床的两两组合。

“

第一种情况：可解释且符合临床

皆大欢喜。然而，反过来说，你并没有什么新的，令人眼前一亮的发现。这种结果多见于对于此前研究的换人群验证等。

”

第一种情况：可解释而且又符合临床。这种结果文章很好写，但意味着亮点太少，只适合对别人及结果进行验证，比如说前人是美国人群，我们用中国人验证。这只能起到锦上添花、验证的作用。

“

第二种情况：不可解释且不符合临床

一定要投稿。因为这种无非就是两种结果：

第一种：审稿人指出你的这种结果是因为……，改后重投。

第二种：审稿人给你找到了原因。你直接抄他的。

”

第二种结果：不符合临床也不能解释。但对于这种情况我们一定要投稿

第一：审稿人会指出你的结果这么怪，是因为什么原因，即使拒稿，但你可照着他的建议修改，以后重新做数据分析再投。相当于至少咱可找到业内的高手审稿。

第二种可能性：审稿人帮你找到了这个原因，你直接照抄。

“

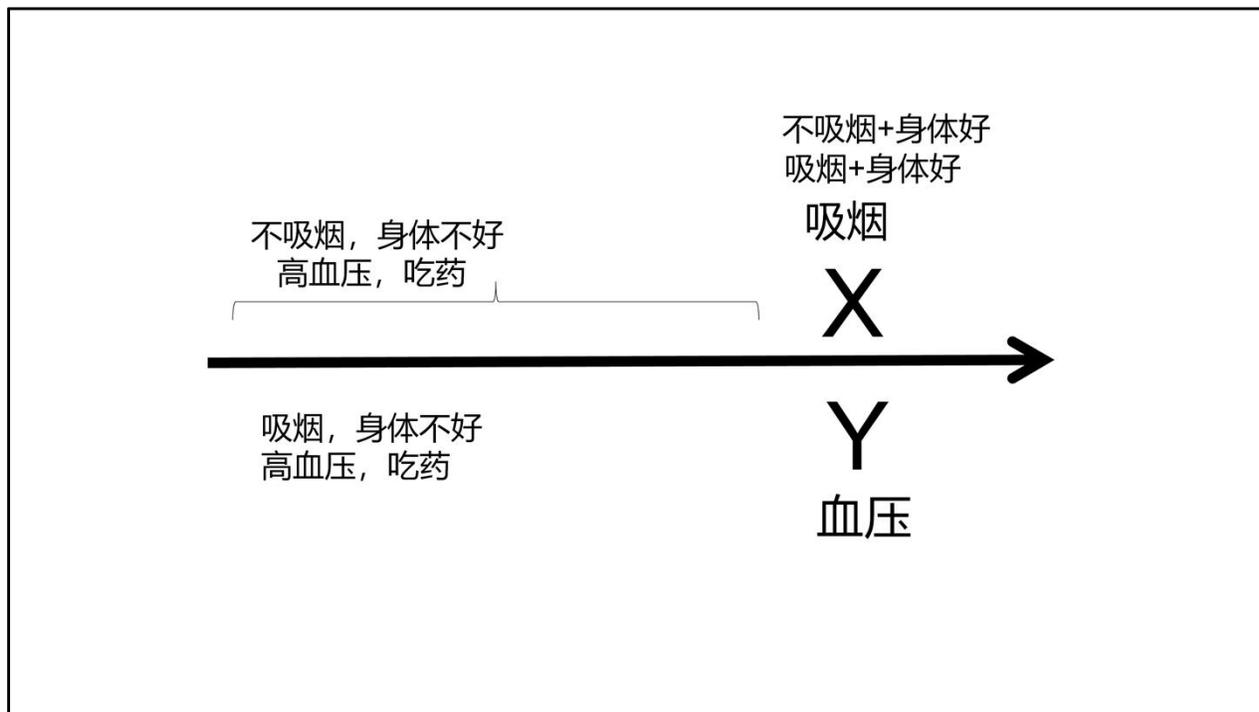
第三种情况：不符合临床但可解释

场景1：横断面数据

横断面数据往往存在很大程度的偏向性。好比在横断面的人群中，我们发现抽烟的患者反而活的比不抽烟的人长。

”

第三种情况不符合临床但可以解释的。比如横断面数据，我们有时候会发现抽烟的患者反而活得比不抽烟的人长，这个结果完全是反的，为什么会出现这样的情况？就可能是幸存者效应所致。



大家想一想，比如说我们在一个横断面的人群中，我问一下这个病人的吸烟情况，然后再看一下他的血压，会发现什么？要么吸烟对于血压没有影响，要么就是吸烟的人血压会更好，因为在进入这个横断面研究前，那些不吸烟/吸烟的身体不好者，已被提前淘汰。

“

第三种情况：不符合临床但可解释

场景2：特殊人群的干扰

我们很多情况下，会发现结果和临床相反，如何解释（规避）？

”

还有一种情况就是特殊人群的干扰，我们在很多情况会发现结果和临床完全相反，但是可以解释。这就需要利用易侬的诊断关联关系模块。

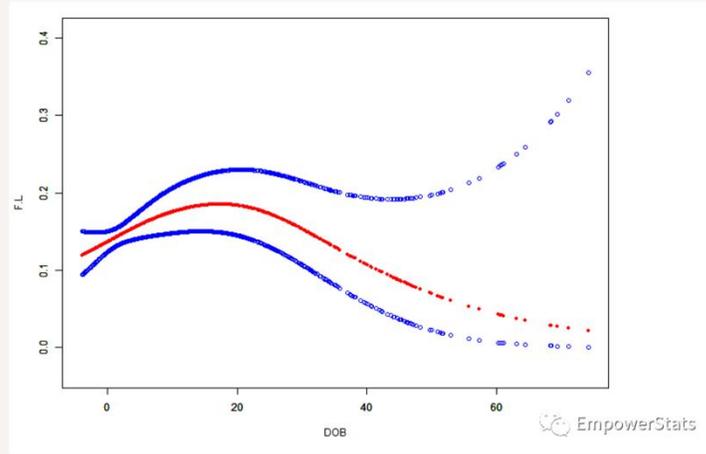
“

诊断关联关系模块

”

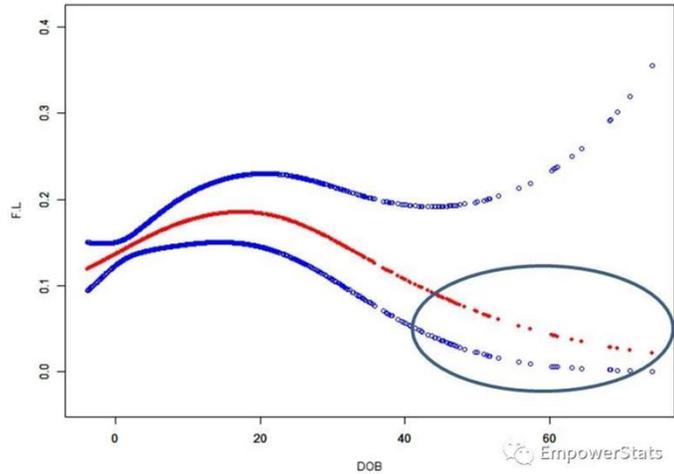


例1. 研究指标A和非酒精性脂肪肝的关系。临床上认为A越高，非酒精性脂肪肝风险越高，即A是危险因素。然而数据结果却是相反的，平滑曲线拟合如下：



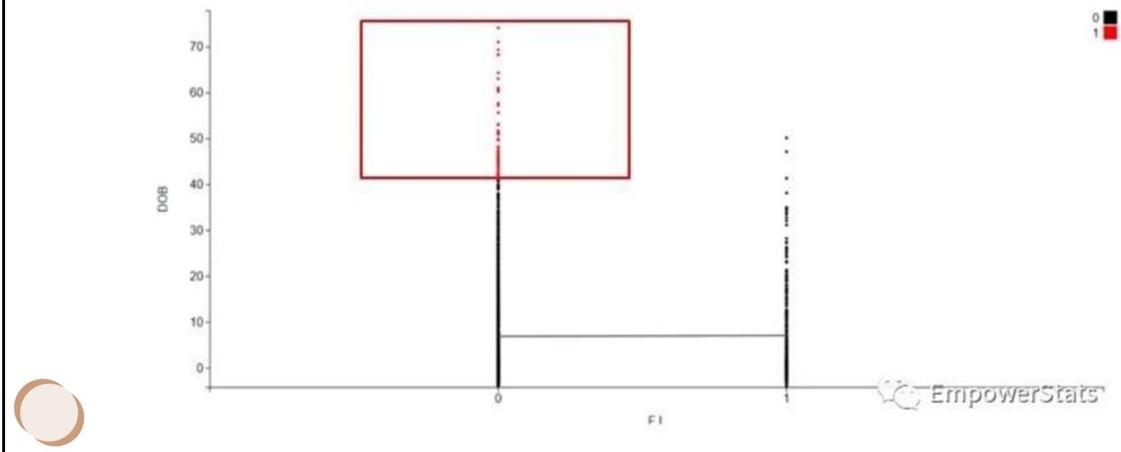
比如研究DOG和非酒精性脂肪肝的关系，临床上已公认dog越高，非酒精性脂肪肝效果就风险也越高。

问题出在右边圈出来的研究对象，是他们导致结果很奇怪。下一步，想把这些人拎出来看看，到底有什么特殊性。想法很简单，实现起来也不难。



但结果却相反，在图右边，为什么 Dob 高了以后，它的脂肪肝的风险反而还低，不符合临床。

按非酒精性脂肪肝分成两组，每个点是一个研究对象。可以在图上用鼠标拖拽选出左上方这些奇怪的人（指标A高且脂肪肝发生率低的）。



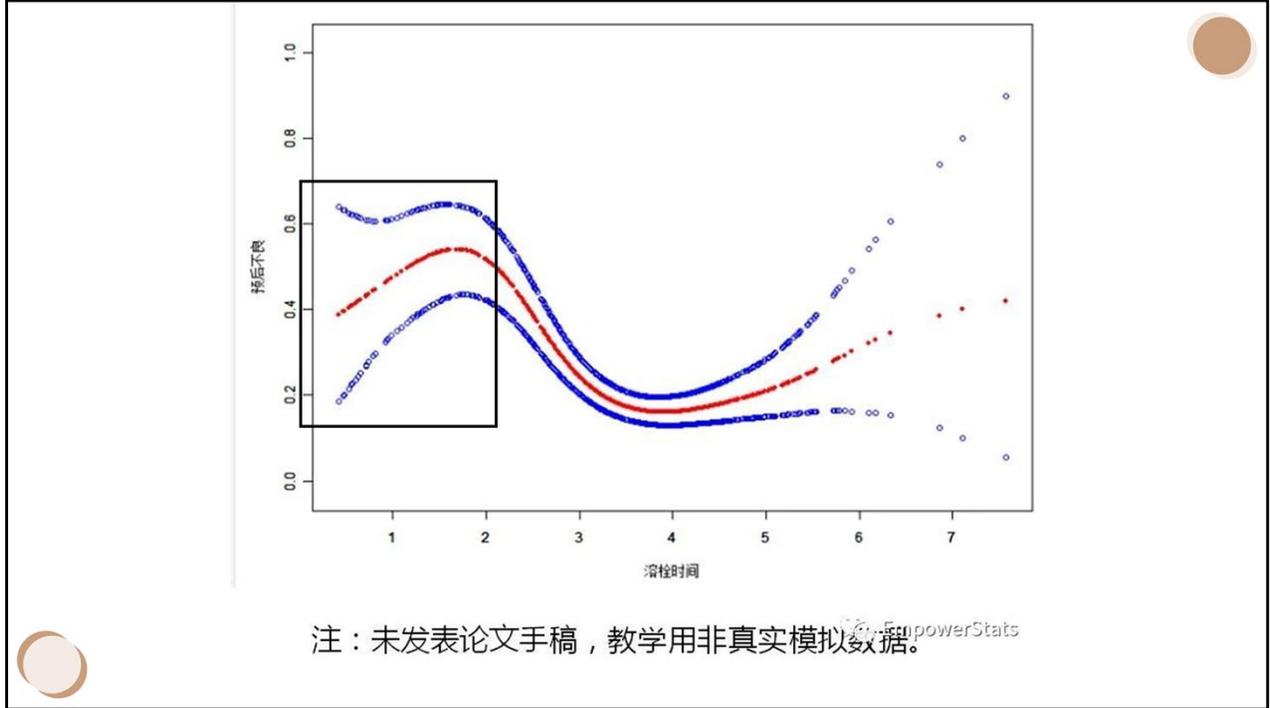
这是，易侬的诊断关联关系模块可发现这群人的特点就是他们的肝功能都很好。因为在肝功能很好的情况下，就算这个指标高了，它也不会发生非酒精性脂肪肝。

“

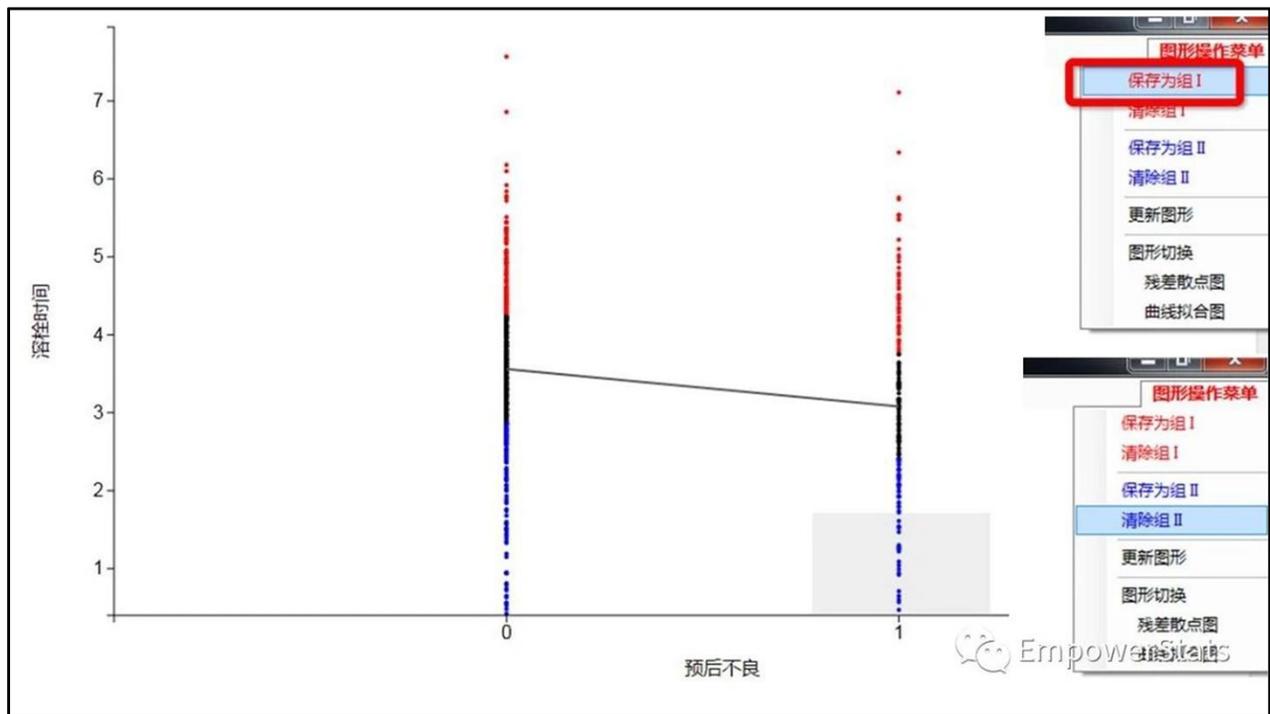
临床上解释为：对于肝功能很好的患者，就算指标A很高，也不会发生非酒精性脂肪肝。这就提示在研究的排除标准里面加上：排除肝功能情况很好的。至于如何定义ALT的范围，可以结合临床参考值范围确定

”

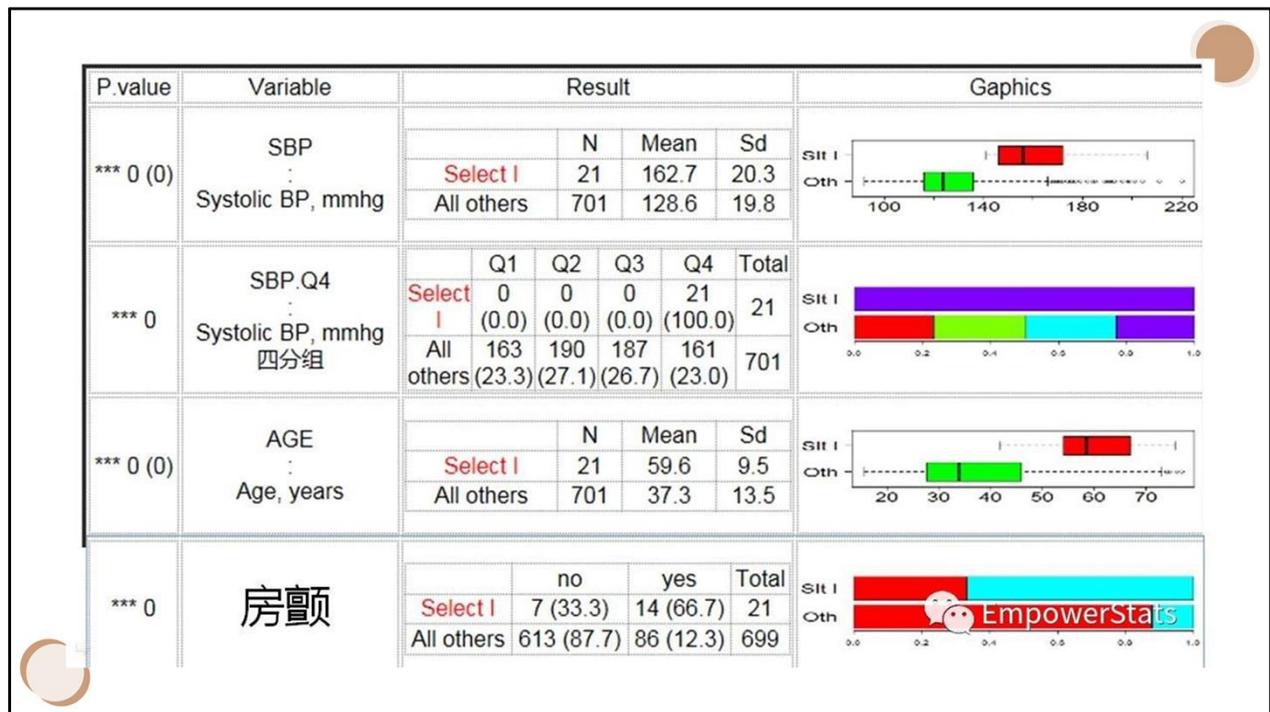




同理。该研究是溶栓时间和愈后不良的关系，我们都知道溶栓的时间越早效果就越好，但图中框选的人群很奇怪，他的溶栓的时间很早，但是愈后很差。



为什么这群人不符合，于是通过易侬的“诊断关联关系模块”把这群人给他找出来，



结果这群人的特征是多数者为房颤，临床上可解释为房颤患者就算早做溶栓治疗，效果依旧不好。因此，在纳排标准设置时排除房颤患者即可。所以大家要是发现临床上的结果和临床相反，我们就用诊断关联关系模块就可以解释它。

“

临床上解释为：就算很早接受了溶栓治疗，因为房颤的原因预后还是很差的。这就提示我们在研究的排除标准里面加上一条：排除房颤患者，因为在房颤患者中观察不到早期溶栓对预后的改善作用。完美解决！

”



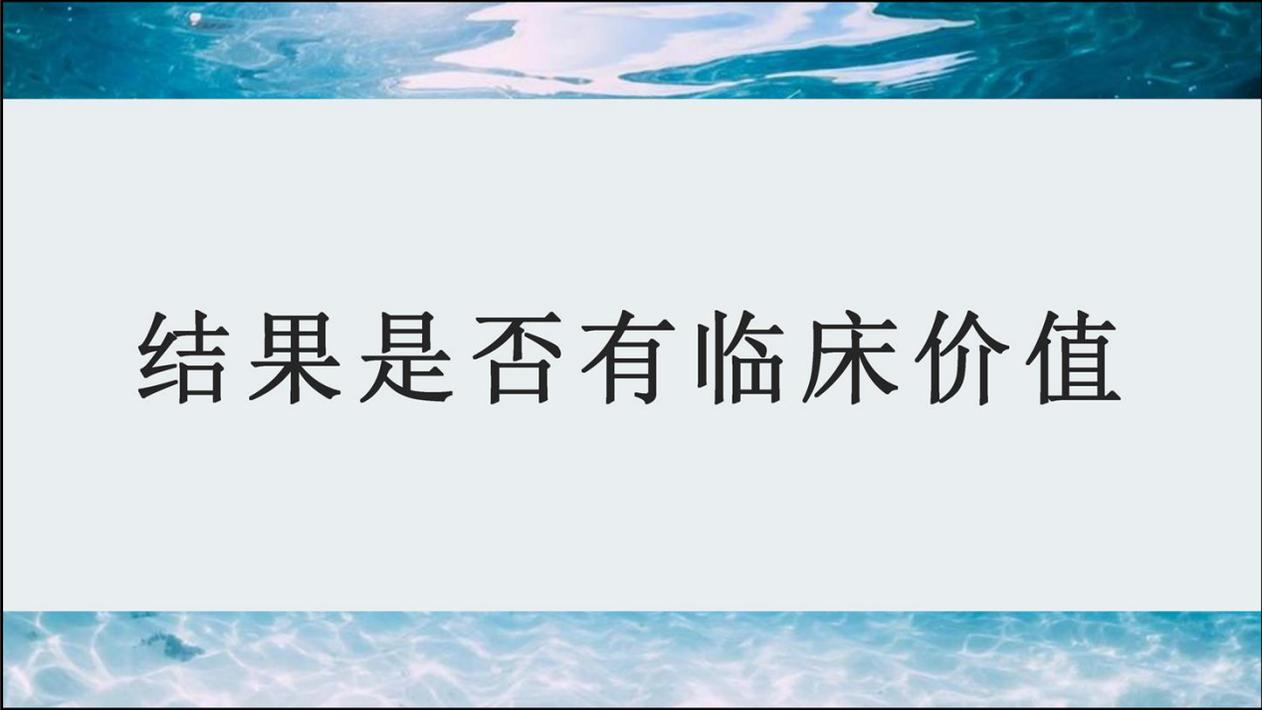
“

第四种情况：不可解释但符合临床

基础研究——寻找潜在机制去解释

”

后一种情况最极端，不可以解释，但是我们在临床上观察到的相符。这个时候就只有从潜在的机制上去解释了



结果是否有临床价值

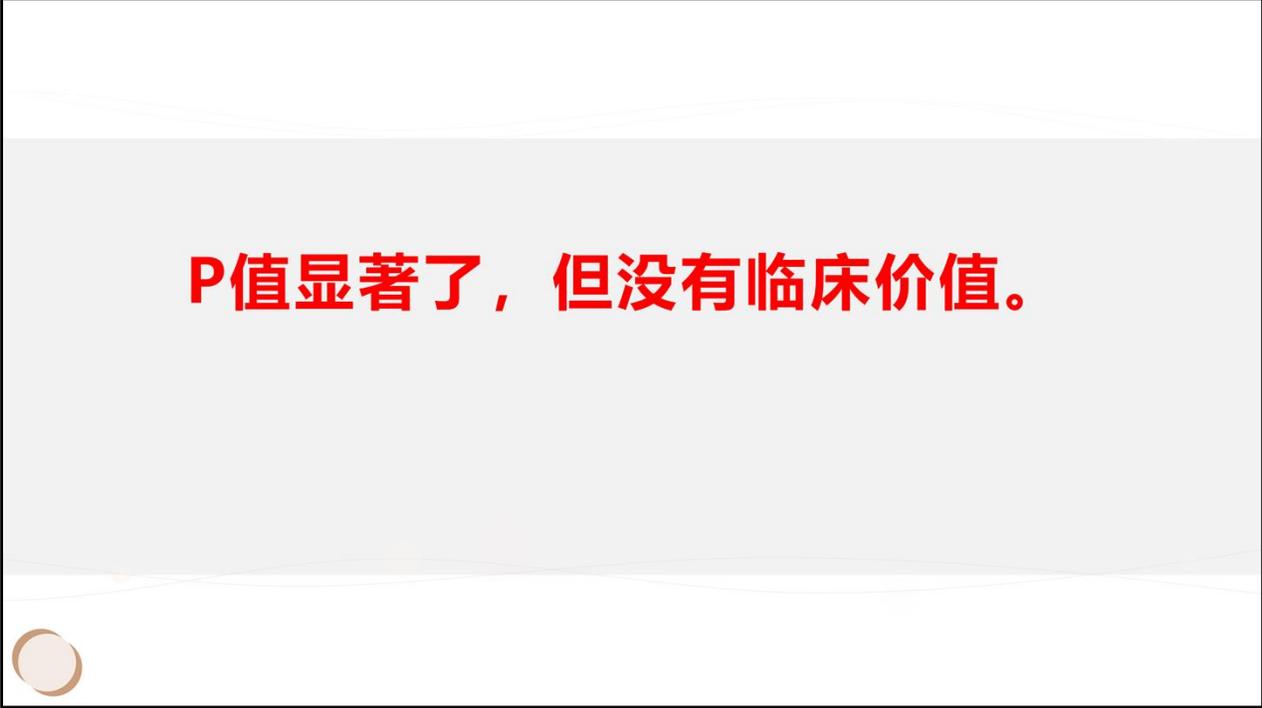
接下来，我们看一下如何判定结果是否有临床价值



”

**从效应值的大小，而非
P值去判定临床价值**

临床价值不能够靠P值评估，而只能从效应值去判断它到底有没有临床价值。



P值显著了，但没有临床价值。

因为很多时候你的p值显著，但没有临床价值。

OR=0.98 (95%CI: 0.14-0.99), P=0.032

X=1: 常规化疗组

X=2: 常规化疗+二代靶向药物

Y=是否死亡 (0=没死, 1=死)

譬如观察常规化疗组vs常规化疗组加二代靶向药物哪一个治疗肿瘤的效果好，最后得到这么一个结论，效应值0.98，且P值显著。如果解读0.98的意义，可知常规化疗加上二代靶向药物可以降低死亡率2%。但没有临床价值。因为我加了一个这么贵的二代靶向药物，病人的死亡才下降了2%，证明该靶向药物效果有限。



”

**从偏性的方向去评估临
床价值**

此外，我们还可以从偏性的方向去评估

“

任何的临床研究
都会有偏性

”

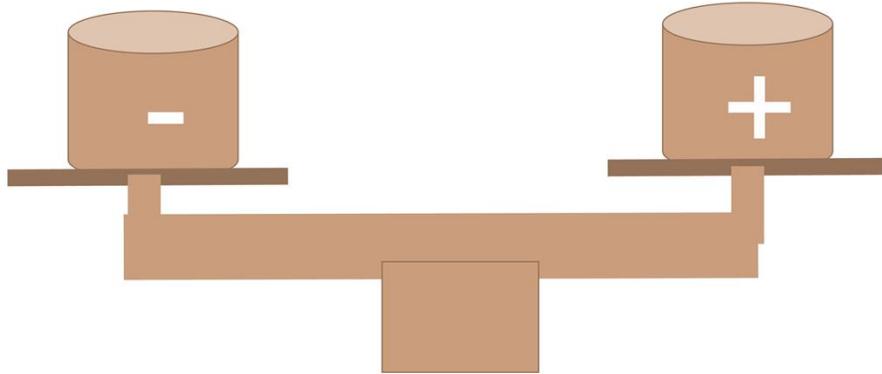
1 观察偏性

2 测量偏性

3 选择偏性

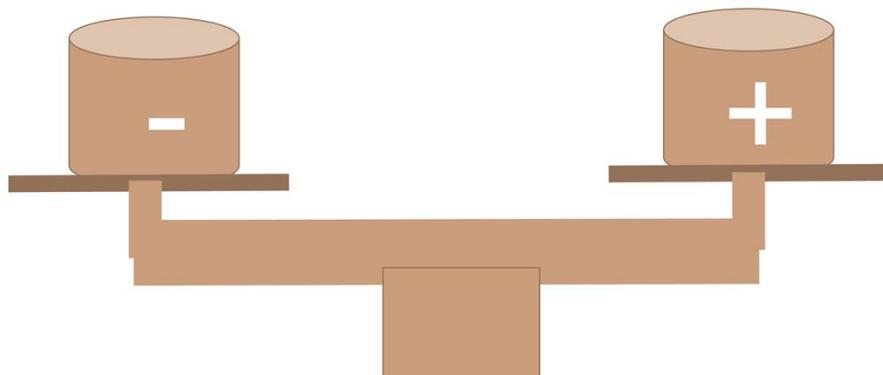
我们做任何的临床研究，它都会有偏性，包括观察偏性、测量偏性，还有选择偏性这三种，这个时候我们要评估一下偏向性它会对结果带来什么样的影响？

(1) 就数据分析的结果而言，无非就是两个趋势，正/负(增加/减少)

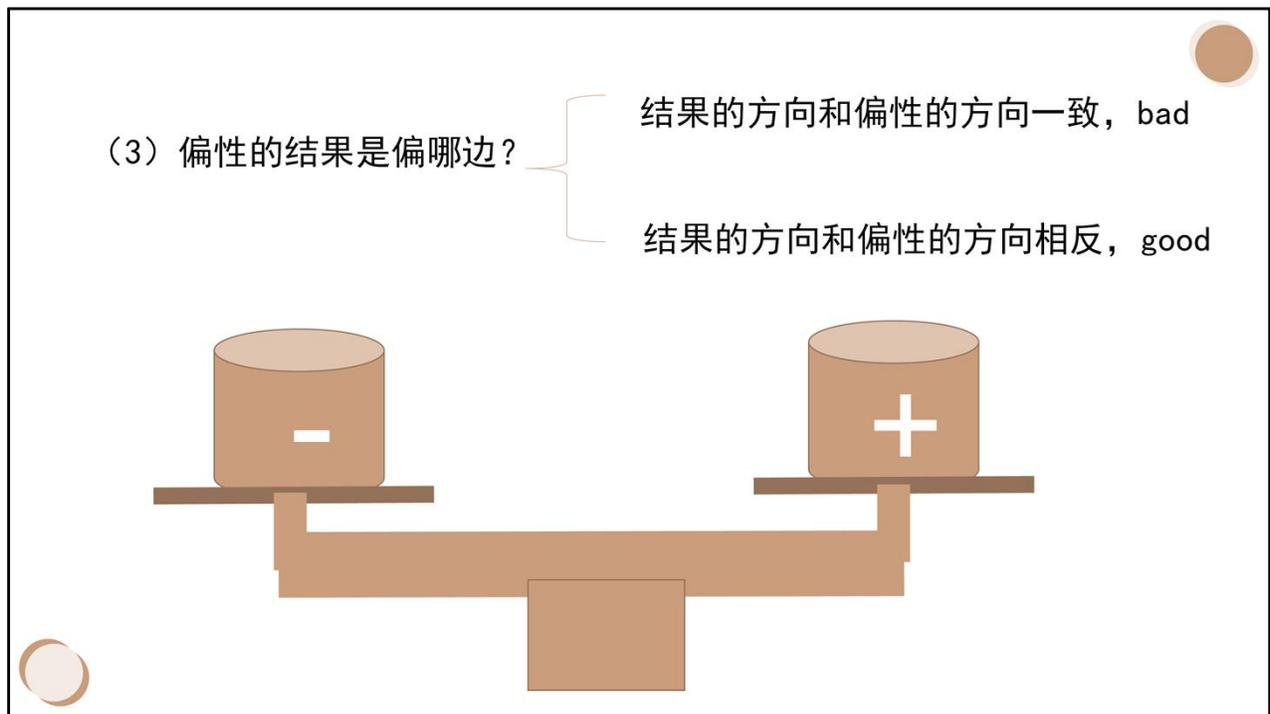


就数据分析的结果而言，无非就是两个趋势，
要么是正相关，要么是负相关

(2) 你的结果是偏哪边?



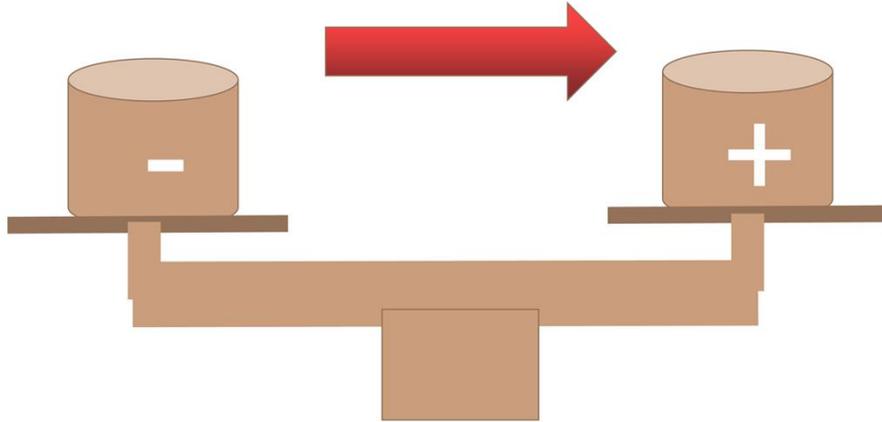
第一步你要先判断你得到的结果方向是什么



第二步，是判断偏性对结果造成的影响。如果结果的方向和偏性的方向是一致的，这个结果就很不好。比如说苏炳添顺风跑出了9秒58最好世界成绩，但这个成绩不能被承认。因为他跑步的方向和风的方向一致。而这个风的方向就是偏性的方向。再譬如博尔特逆风跑了9秒8，但是他的结果很可能被低估，即他跑的会更快。因为逆风的情况下，他都能够跑进9秒8，如果顺风只会更快。

我想研究吸烟与肺癌的关联（X, 吸烟，Y, 是否肺癌）
偏性是人群中，年轻人太多

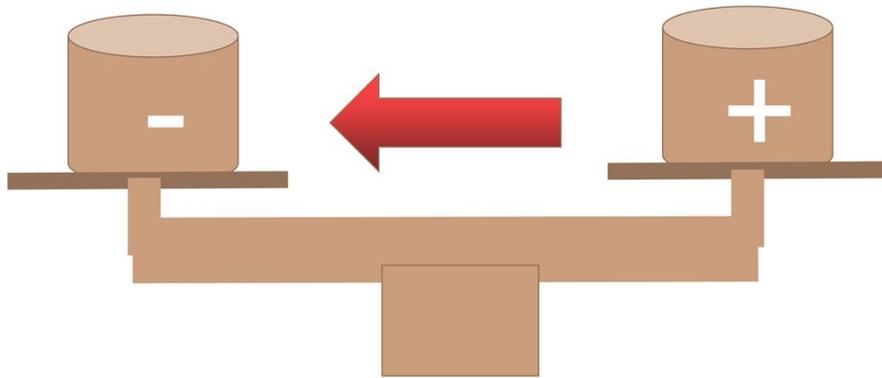
第一步：XY的方向



假设现在想开展吸烟和肺癌关系的研究，我发现吸烟和不吸烟的人相比，肺癌的风险是增加的，那我的结果的方向是正相关。

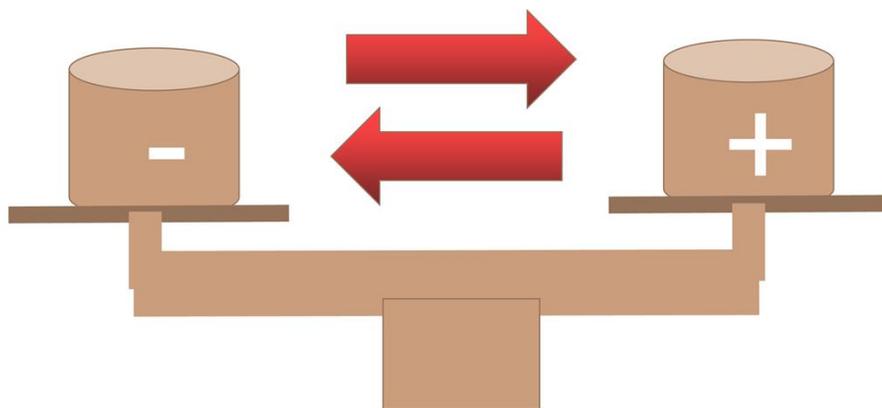
我想研究吸烟与肺癌的关联（X, 吸烟，Y, 是否肺癌）
偏性是纳入人群中，年轻人太多（年龄）

第二步：偏性的方向（年轻人，不容易观察到肺癌）

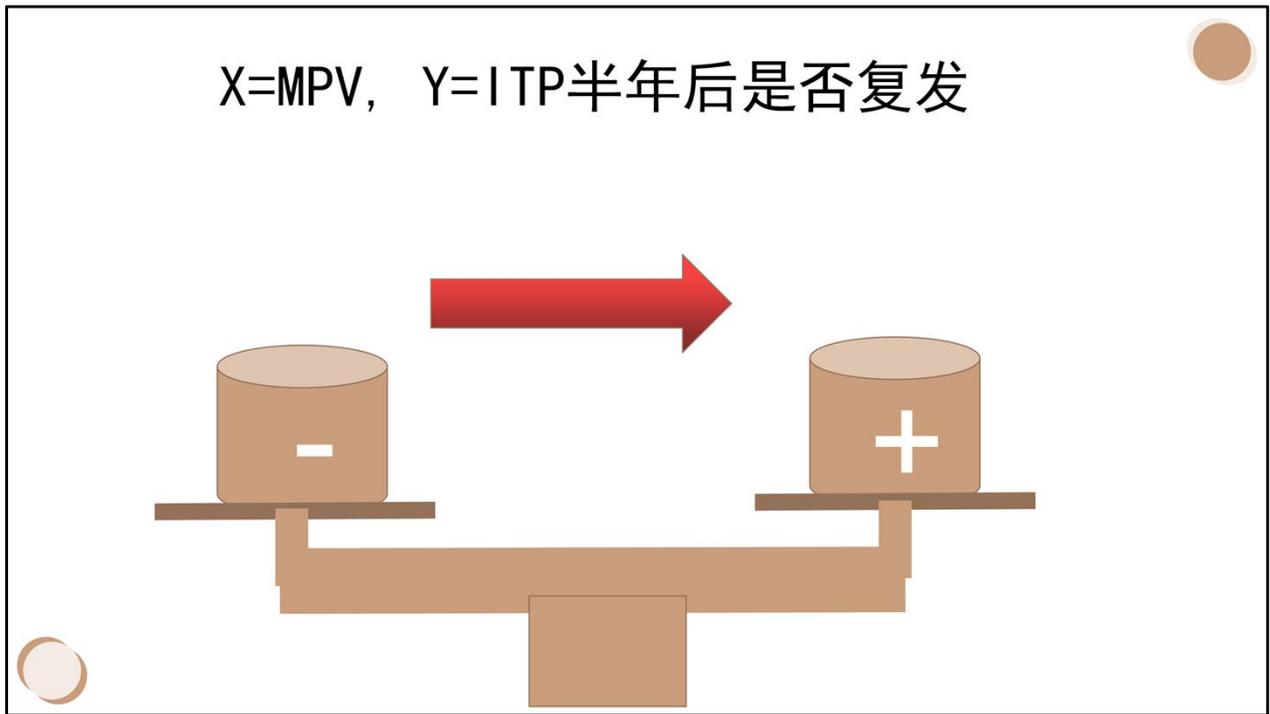


偏性是人群中年轻人太多，而年轻人是不容易观察到吸烟与肺癌的关系的。即偏性的方向偏负相关

在年轻人这么多的情况下，我都能观察到吸烟导致肺癌的现象。如果把这些人排除，只要老人，那结果只会更牛逼。



$X=MPV$, $Y=ITP$ 半年后是否复发



在这个例子中，我们的结果是正相关，



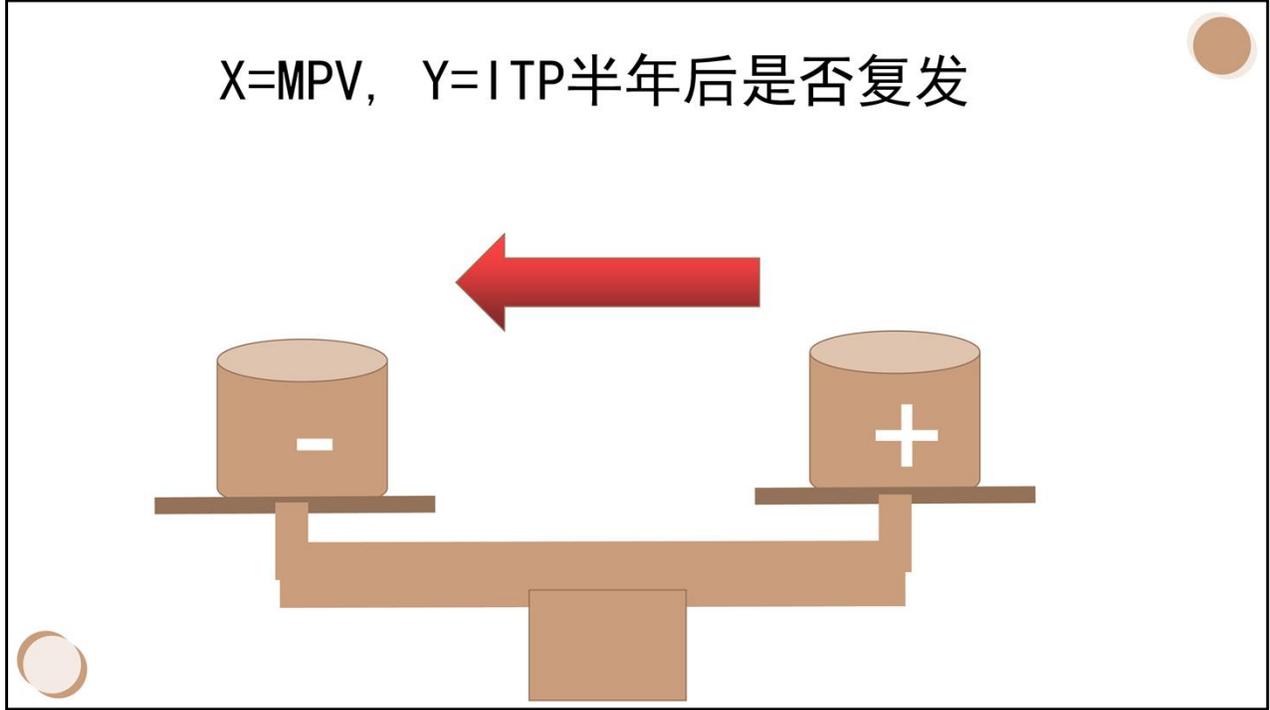
The exclusion of patients who did not respond to front line therapy is a substantial omission. These may be infrequent patients (5-20%) but they are very important because they may end up being the hardest to treat patients and their outcome is important to predict at diagnosis (if possible).

译：作者把对治疗无应答的患者排除出本研究可能是一个很明显的纰漏。尽管这些病人很少（5-20%），但这些患者会很重要因为他们往往治疗效果极差。



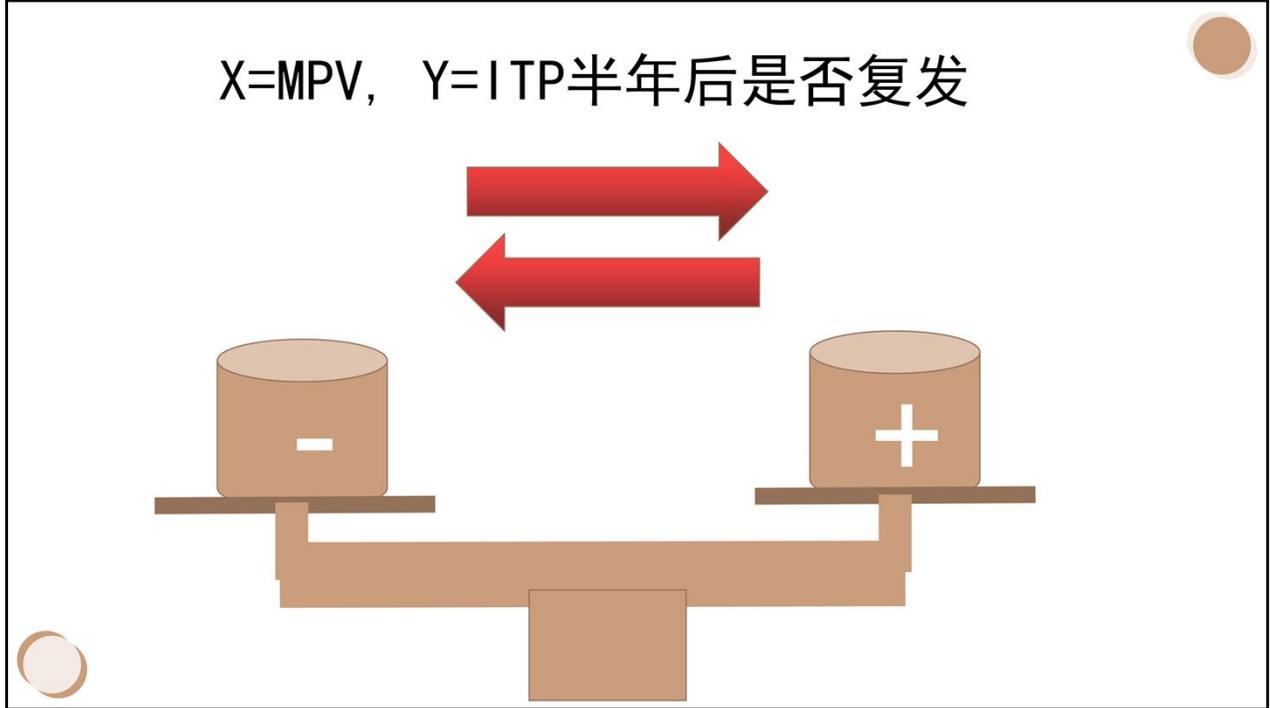
审稿人却指出偏性是排除对治疗无应答的患者。

$X=MPV$, $Y=ITP$ 半年后是否复发



因此，把这些患者排除出本研究，会导致结果偏向负相关

$X=MPV$, $Y=ITP$ 半年后是否复发



然而，两者的方向相反，我们的结果更强。



”

**从实际的变化幅度去评
估临床价值**

X: 蛋白表达IPP值。

Y: Gensini评分

```
percentile          |-----distribution-----
```

Exposure	Non-adjusted	Adjust
BECLIN1.1	4191.32 (3031.57, 5351.07) <0.0001	4039.23 (2871.87, 5206.59) <0.0001
NEW. VAR	4.19 (3.03, 5.35) <0.0001	4.04 (2.87, 5.21) <0.0001

50%	0.002	0.002	****
75%	0.004	0.003	***
90%	0.005	0.003	***
95%	0.005	0.004	***
mean	0.002	0.005	***
std	0.002	0.006	**
n	518	0.006	
.	0		-----distribution-----

beclin1蛋白每增加 0.001单位, Gensini评分 增加4.19分, 调整后, 也是增加4.04分。

此前我们做过一个研究, 某蛋白的表达IPP值和 Gensini评分的关联, 该评分主要评估冠脉堵塞程度, 但我们发现IPP每增加一个单位, 评分会增加4191分, 而Gensini评分最高也就是160分。其原因就是 IPP的值太小, 所以把该值1000, 即每变化0.001个单位, 评分会增加4.19分。这是, 结果的临床价值就好解释。

X: 新生儿出生体重 (g) 。
Y: 阿帕评分

- 新生儿体重每增加1g, 阿帕评分增加0.04分
- 将原始的出生体重 (g)/100。
- per 100g change, XX评分增加4分。
- 这个结果更符合临床, 更容易解释。

同样一个道理, 新生儿的体重每增加一克, 阿帕评分会增加0.04分, 解决办法是原始的出生体重除以100, 变成每变化100克, 阿帕评分会增加4分, 这个结果就更符合临床也更容易解释。

单击此处添加大标题内容

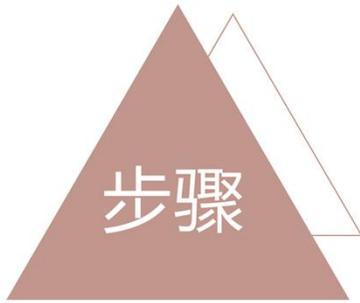
A

效应值太大，可扩大X的倍数（ $\times 10, 100$ ）。自然变小

B

效应值太小，可缩小X的倍数（ $\div 10, 100$ ）。自然变大

所以说效应值见大化小，见小化大。如此而已



2: 敏感性分析

- ▶ 1. 易侻可轻松做到的
- ▶ 2. 其他常见的敏感性分析



敏感性分析：让结果稳定可靠的不二法门

敏感性分析不是单独的一种统计学方法，而是通过各种形式证实你所得到的结果是可信赖的，稳健的（robustness）。

敏感性分析它不是一种单独的统计学方法，它是用各种各样的方法去证明你得到的这个结果是稳健的，只要你能够想得到、其你掌握的统计学方法或其他方法，这些方法可让结果变得稳定，均可使用。

用易侖可轻松实现的敏感性分析



用易侖做数据分析时常见的敏感性分析思维

SEX	Male	Female	P-value	P-value*
N	417	415		
Age, years	38.4 ± 14.7	37.3 ± 13.1	0.243	0.484
Height, m	1.6 ± 0.1	1.5 ± 0.1	<0.001	<0.001
Weight, kg	56.4 ± 7.1	50.4 ± 6.8	<0.001	<0.001
Body mass index, kg/m ²	20.9 ± 2.0	21.5 ± 2.5	0.001	<0.001
Systolic BP, mmhg	132.1 ± 22.0	128.5 ± 23.0	0.027	<0.001
Diastolic BP, mmhg	69.6 ± 11.7	69.4 ± 10.9	0.750	0.969

在做表1时，连续变量分别用参数检验和非参数检验均各算有一次，各获取一个P值

如果你们用易侖的研究人群模块去做表一，你就会发现有两个p值，这两个p值的话是怎么回事？P值是正态分布的时候做出来的，p值是用偏态分布的时候做出来的。这也相当于敏感性分析。

用易侬做数据分析时常见的敏感性分析思维

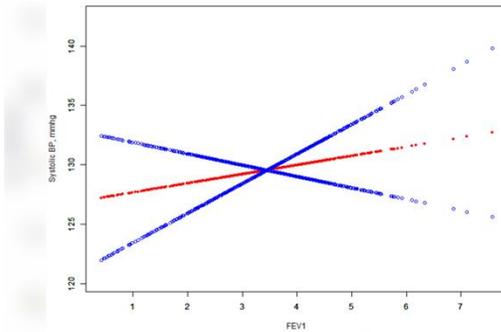
Exposure	Non-adjusted	Adjust I	Adjust II
FEV1	-3.74 (-5.06, -2.41) <0.0001	-3.17 (-4.61, -1.72) <0.0001	-1.41 (-2.97, 0.15) 0.0778
FEV1 三分组			
低	0	0	0
中	-11.17 (-14.76, -7.57) <0.0001	-10.12 (-13.80, -6.44) <0.0001	-5.66 (-9.37, -1.96) 0.0028
高	-9.04 (-12.64, -5.43) <0.0001	-6.94 (-10.92, -2.96) 0.0007	-2.49 (-6.93, 1.95) 0.2725
FEV1 三分组	-4.49 (-6.32, -2.67) <0.0001	-3.50 (-5.51, -1.48) 0.0007	-1.40 (-3.63, 0.83) 0.2204

探讨XY独立作用时，连续变量转变为分类变量，再多算一次。

第二点我们都知道x它有三种形式，连续变量，分类变量，分类变量里面又分有序分类变量和无序分类变量。在这种情况下我们将其进行有序和无序分类变量转化，结果就是不管x的数据形式如何变化，其得到的结果都是稳定的。

用易侖做数据分析时常见的敏感性分析思维

Adjust II		
-1.41	(-2.97, 0.15)	0.0778
0		
-5.66	(-9.37, -1.96)	0.0028
-2.49	(-6.93, 1.95)	0.2725
-1.40	(-3.63, 0.83)	0.2204



探讨XY独立作用时，已做完直线关系，再做曲线拟合，证实XY存在非线性关联的可能较小。

用易侬做数据分析时常见的敏感性分析思维

Adjust II
-1.41 (-2.97, 0.15) 0.0778
0
-5.66 (-9.37, -1.96) 0.0028
-2.49 (-6.93, 1.95) 0.2725
-1.40 (-3.63, 0.83) 0.2204

Y: Systolic BP, mmhg	N	Coeff.	95%CI Low	95%CI High	P value	P(interaction)

SEX						0.4138
Male	365	-0.12	-1.83	1.58	0.8860	
Female	357	1.15	-1.43	3.73	0.3809	
Total	722	0.26	-1.16	1.69	0.7161	
OCCU						0.3251
1	368	-0.23	-2.12	1.65	0.8079	
2	44	0.60	-4.40	5.61	0.8140	
3	123	0.87	-2.32	4.07	0.5920	
4	187	3.33	0.01	6.66	0.0500	
Total	722	0.68	-0.72	2.08	0.3441	

探讨XY独立作用时，得到阴性结果，再做交互检验，证实XY的阴性关联在所有亚组中都存在。结果稳定。

其他的敏感性分析



剩余的敏感性分析，
基于临床知识、思
辨能力、以及审稿
人建议

logo

一. 在X上面做文章

- 不同的测量方法
- 不同的定义原则
- 不同的临床切点
- 随/不随时间变

通过上面的提示，我们可以做的是

剩余的敏感性分析，
基于临床知识、思
辨能力、以及审稿
人建议

- 一．在X上面做文章
 - 不同的测量方法
 - 不同的定义原则
 - 不同的临床切点
 - 随/不随时间变

- 一．睡眠（午觉、睡眠时间、睡眠 质量）的测量方式是否只有一种？如果有多种，当初在设计队列
 - 的时候可否进行多种问卷，以便后续进行敏感性分析
- 二．睡眠质量的切点：匹兹堡就是5。
- 三．本研究涉及了重复测量，time-vary设计？

我们国家东风队列发表的一篇顶刊，其研究假设是睡眠（X）和中风（Y）。其中，睡眠研究者用了睡眠时间，睡眠质量，午睡时间来表示。即不同测量方法全用。



剩余的敏感性分析，
基于临床知识、思
辨能力、以及审稿
人建议

logo

一. 在Y上面做文章

- 不同的测量方法
- 不同的定义原则
- 不同的竞争终点
- 随不随时间变

通过上面的提示，我们可以做的是

剩余的敏感性分析，
基于临床知识、思
辨能力、以及审稿
人建议

logo

- 一．在Y上面做文章
 - 不同的测量方法
 - 不同的定义原则
 - 不同的竞争终点
 - 随不随时间变

- 一．中风的定义
- 二．文中有死亡结局。如果患者随访过程中死亡，且无中风？
- 三．疑似病例与确诊病例的分别对结果影响

还是东风队列的例子，对中风的定义，研究者就定义了三种，第一种是全部中风患者，第二种是经过影像学确诊的患者，第三种是未确诊，仅是依靠病历上的记录。最终，对三种结局均进行计算，并得出无论如何定义中风，最终结果都是一样的。

剩余的敏感性分析，
基于临床知识、思
辨能力、以及审稿
人建议

logo

一. 在Z上面做文章

- 要/不要
- 调整不调整
- 不同的时间点

通过上面的提示，我们可以做的是

剩余的敏感性分析，
基于临床知识、思
辨能力、以及审稿
人建议

logo

一．在Z上面做文章

- 要/不要
- 调整不调整
- 不同的时间点

- 一． 助眠药？纳入不纳入，调整不调整对结果的影响
- 二． 有没有从事昼夜颠倒的工作
- 三． 有没有其他神经精神疾病

在这篇文章里面没有服用安眠药这个变量，此外，有没有从事昼夜颠倒的工作？有没有其他的神经精神疾病？这些对于睡眠有很大的影响，而这篇文章里面都没有调整。

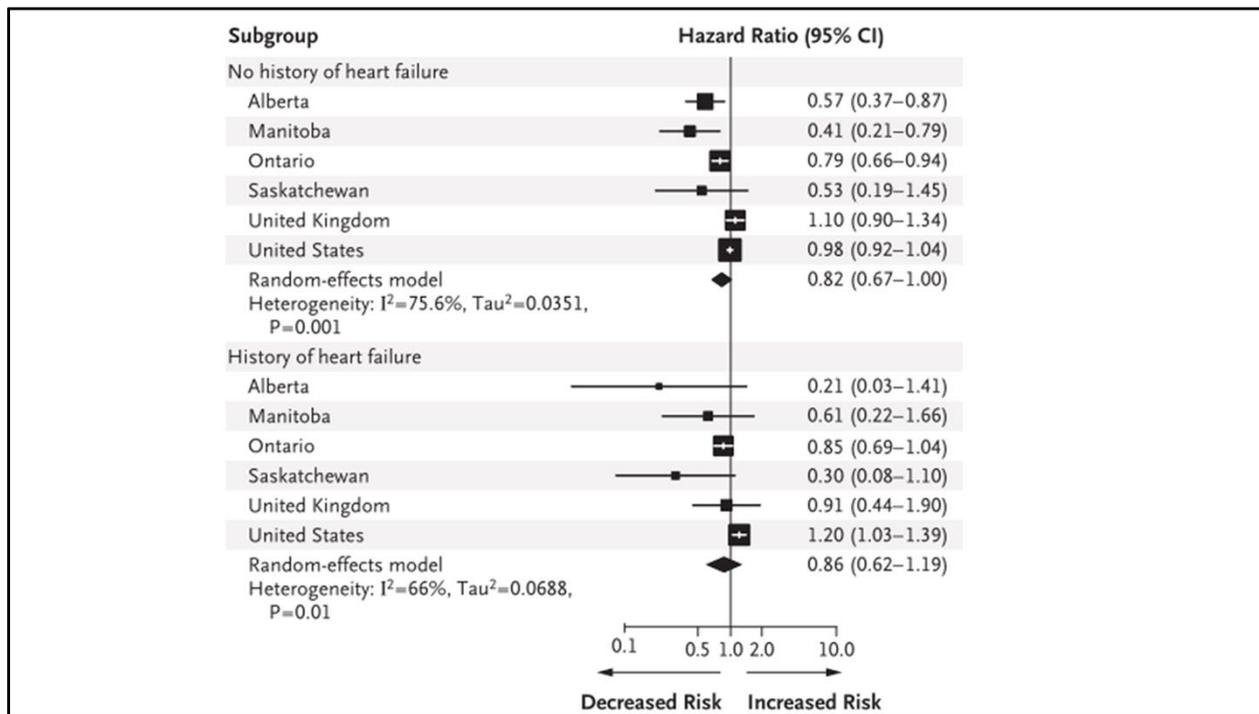


剩余的敏感性分析，
基于临床知识、思
辨能力、以及审稿
人建议

logo

地点上面做文章

- 不同中心影响



例文中，6个中心数据用meta分析的形式进行合并。



剩余的敏感性分析，
基于临床知识、思
辨能力、以及审稿
人建议

logo

人物上面做文章

- 不同纳排标准

通过上面的提示，我们可以做的是

剩余的敏感性分析，
基于临床知识、思
辨能力、以及审稿
人建议

logo

- 一．人物上面做文章
 - 不同纳排标准

我们排除了9,213名基
线时患有冠心病
(CHD) (n = 5,468)、
中风(n = 1972)或癌
症(n = 2,182)

- 一．排除肿瘤者
- 二．排除有冠心病及其他基础疾病者

东风队列例文，因为排除了本身有基础疾病的患者，对结果造成的影响是？因此，作者把这些排除的人加进来，又再做了一个敏感性分析。



THANKS

年中总结报告

BUSINESS REPORT

