7 Subgroup Analyses

在第5章中，我们讨论了研究间异质性的概念，以及为什么它在元分析中如此重要。我们还学习了一些方法，使我们能够确定哪些研究促成了观察到的异质性，通过离群值和影响力分析实现。但是这些分析全部是从统计学角度进行的。我们“测量”了数据中相当大的异质性，因此排除了统计性质上不拟合的研究（即离群和有影响力的研究），以提高模型的稳健性。

我们对研究间的异质性来源进行了敏感性分析（包括离群值分析和影响力分析）找到在统计学角度上拟合的不是那么好的观测效应量（离群值和异常值对异质性或效应量池化影响较大）。

- 上述的分析是post hoc analysis：Outlier and influence analyses are performed **after** seeing the data.
- influence analysis and outliers analysis is just focuse on the data, 但不探究异质性产生的原因是什么，即为什么有一些原始研究的观测效应量和我们所期望的（池化效应量）偏离过大，导致这种模式出现的原因是什么?
- 也就是根据现象进行统计学上的操作，但不问为什么，原因是什么：这可能是因为实验设计或者干预方式等因素具有细微差别（也可能是受试者特征），也就是未知的异质性来源，但是 influence analysis 和outliers analysis 可不会告诉你这些，它们只是察觉到这种模式的存在。

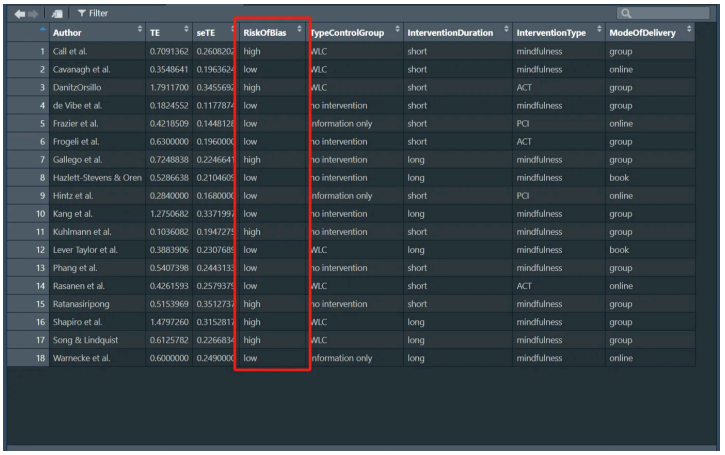
想象一下，你正在进行一项荟萃分析，调查一种药物治疗的有效性。你会发现，总的来说，治疗没有效果。然而，有三项研究发现了相当大的治疗效果。：影响力分析可能察觉这三项研究，但是为什么效果出现差异，不知道呀：可能是因为这三项研究有一个共性（例如干预手段和其他研究有微小的不同），导致了结果的差异，这种发现将非常突破性（所以亚组分析是必要的且重要的），还是如前所述，这是敏感性分析和影响力分析做不到的（这两种分析更关注异常值和强影响力研究是否会影响结果的稳健性，并探索剔除前后结果的变化是否显著），亚组分析是进一步探索导致差异（异质性模式）出现的原因。

- 影响力分析是看异质性会不会影响结果稳健性
- 亚组分析是找寻异质性来源（即影响因素分析）

`Subgroup analyses`, also known as `moderator analyses`, are one way to do this.
They allow us to test specific hypotheses, describing why some type of study produces lower or higher effects than another.

moderator的存在导致效应量存在显著差异，用先验假设探索和解释特定的异质性模式。所以 subgroup tests should be defined **a priori**。

我们应该定义可能影响观察到的效应大小的不同研究特征（干预特征，受试者特征），并相应地对每个研究进行编码。效应大小不同的原因数不胜数，但我们应该把注意力集中在我们分析的话题背景下重要的那些（干预特征，受试者特征）。



	Author	TE	seTE	RiskOfBias	TypeControlGroup	InterventionDuration	InterventionType	ModeOfDelivery
1	Call et al.	0.7091362	0.2608207	high	WLC	short	mindfulness	group
2	Cavanagh et al.	0.3548641	0.1963624	low	WLC	short	mindfulness	online
3	DanitzOrsillo	1.7911700	0.3455692	high	WLC	short	ACT	group
4	de Vibe et al.	0.1824552	0.1177874	low	no intervention	short	mindfulness	group
5	Frazier et al.	0.4218509	0.1448128	low	information only	short	PCI	online
6	Frogeli et al.	0.6300000	0.1960000	low	no intervention	short	ACT	group
7	Gallego et al.	0.7248838	0.2246641	high	no intervention	long	mindfulness	group
8	Hazlett-Stevens & Oren	0.5286638	0.2104609	low	no intervention	long	mindfulness	book
9	Hintz et al.	0.2840000	0.1680000	low	information only	short	PCI	online
10	Kang et al.	1.2750682	0.3371997	low	no intervention	long	mindfulness	group
11	Kuhlmann et al.	0.1036082	0.1947275	high	no intervention	short	mindfulness	group
12	Lever Taylor et al.	0.3883906	0.2307685	low	WLC	long	mindfulness	book
13	Phang et al.	0.5407398	0.2443133	low	no intervention	short	mindfulness	group
14	Rasanen et al.	0.4261593	0.2579375	low	WLC	short	ACT	online
15	Ratanasiripong	0.5153969	0.3512737	high	no intervention	short	mindfulness	group
16	Shapiro et al.	1.4797260	0.3152817	high	WLC	long	mindfulness	group
17	Song & Lindquist	0.6125782	0.2266834	high	WLC	long	mindfulness	group
18	Warnecke et al.	0.6000000	0.2490000	low	information only	long	mindfulness	online

- 上面是进行合适的编码的例子

例如，我们可以检查是否某种药物比另一种药物产生更高的效果（干预特征）。或者我们可以将随访期较短的研究与随访期较长的研究进行比较（干预周期也可比）。我们还可以检查观察到的效果是否因进行研究的文化区域而异（不同国家家或地区）。作为一名元分析专家，拥有一些特定学科的专业知识是有帮助的，因为这可以让他找到与该领域其他科学家或从业者实际相关的问题，所以，亚组分析要分析特定专业领域切实关注的问题（例如运动处方）。

- 是否干预效果会因受试者特征的不同而不同？（超重和肥胖个体对治疗的响应不同）。

- 存在的异质性模式可能可以被我们的科学假设所解释，亚组分析帮助我们更好的理解特定的话题或者至少产生指导未来决策的实用见解。

所以综上所述，本章主要关注：

1）describe the statistical model behind subgroup analyses,

2）how we can conduct one directly in *R*.

▼ 7.1 The Fixed-Effects (Plural) Model

我们假设 meta 分析中的研究并非来自一个整群。在亚组分析中，相反，我们假设它们属于不同的亚组，并且每个亚组都有自己的真实效应量。我们的目的是拒绝零假设（即亚组之间的效应大小没有差异）。
亚组分析的计算包括两个部分：

1. 首先，我们将每个子群的效应汇总。
2. 随后，使用 统计检验 比较亚组的效果（Borenstein和Higgins 2013）。

▼ 7.1.1 Pooling the Effect in Subgroups

- 即使我们把我们的原始研究切成亚组（更小的群体），研究间异质性（未知的异质性来源）仍然存在，建议选择是使用随机效应模型。这假设在一个亚组内的研究的真实效应也是一个分布，我们想要从中估计总体的均值。与常规荟萃分析的不同之处在于，我们进行了几个独立的随机效应元分析，每个亚组一个。从逻辑上讲，这会产生一种池化效应量μ_g，对于每个子组 g。
- tau^2 的估计是使用跨亚组的结果，而不是每个子组独有的。这意味着所有亚组被假设共享一个共同的被估计的“研究间异质性”。
- 👉因为：如果某些亚组研究数太少（e.g. ≤ 5），单独估 τ^2会很不稳定，所以实践中常用**合并** τ^2，让所有亚组共享一个统一的异质性估计值，这是**统计更稳、误差更小，更可靠**的处理方式。

▼ 7.1.2 Comparing the Subgroup Effects

我们现在要比较每个g组是否有差异，差异是否显著，我们的备择假设是 g 组之间的差异显著，至少有一个g组和其他组不同（来自不同的研究总体）。如果结果显著（p < 0.05），我们认为**亚组间存在真实差异**

方法：

把不同的亚组看作是不同的大型研究，并且这些研究已经被我们计算出来了效应量和SE，然后把这些研究进行比较看是否存在显著差异。

Therefore, we use the value of Q to determine if the subgroup differences are large enough to not be explainable by sampling error alone.和异质性检验的逻辑一样，我们看异质性到底是来源于抽样误差就可以解释，还是异质性大到不能仅用抽样误差解释，有其他的变异导致真实效应量的估计之间存在显著差异。

先计算Q值，如果组间（假设的大型研究间的Q值）比期望值$k-1$的卡方分布大且显著，那就证明三组之间的差异大到不能用抽样误差解释，所以真实效应量也不同（三者不共享同一真实效应量，是一个分布），也就代表研究间存在差异。

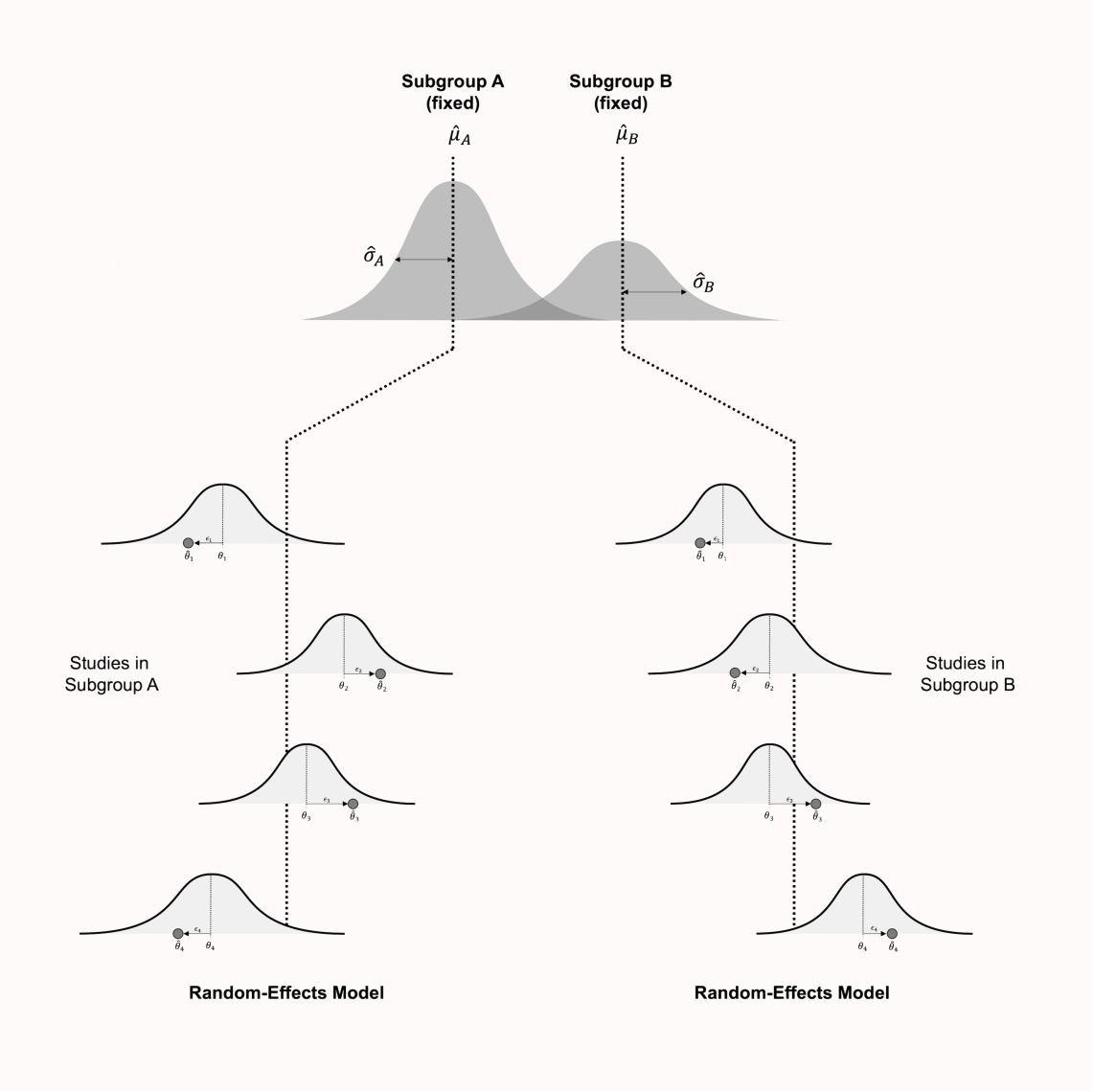
> 我们使用 Q 检验是为了判断：亚组间效应量的差异是否显著，是否可以归因于效应量的真实差异而非抽样误差。如果显著（p < 0.05），说明至少有一个亚组的真实效应量与其他不同。→ 📌 存在亚组间差异，我们假设的**研究特征可能是调节变量**。

注意：

This Q test is an **omnibus test，只告诉你组间存在差异，但是不告诉你哪两组之间有差异，类似于方差分析**

我在想是不是这里的固定效应指的是亚组的水平是固定的（所以称之为固定效应模型）

The Fixed-Effects (Plural) Model 假设每个子组有自己的共享真实效应（固定效应），因为亚组的水平是固定的，并且结论不需要从样本推广到总体，每个水平有自己真实的效应量。但在亚组内又假设是随机效应模型（异质性不仅来源于抽样误差，还有研究间异质性也就是真实效应量的分布）



具有固定水平的亚组变量的几个例子：

■年龄段:儿童、青年、成人、老年人。

■文化背景:西方、非西方。

■对照组:替代处理、最小处理、不处理。

■衡量结果的工具:自我报告,专家评估。

■学习质量:高、低、不清晰。

由于固定效应（复数）模型既包含随机效应(在亚组内)又包含固定效应(因为亚组假定水平是固定的),因此在文献中也称为**混合效应模型**。

计算每个亚组的效应量的时候，可以选择固定效应模型和随机效应模型。这时候一般**选择随机效应模型**，这和主效应分析选择固定模型还是随机效应模型的理由是一样的，尽管我们已经对研究进行了分组。

帮助辅助理解的段落：

> 最为关键的是第二部分，如何计算亚组间效应量的差异。这时候，可以将每个亚组的结果当作一个大的研究，比如你有三个亚组，就当作三个大的研究。随后，我们分析三个大的研究之间是否有异质性，如果有异质性则证明亚组间有差异，这是不是和主效应的异质性分析是一模一样的。但是问题来了，如果亚组内用的随机效应模型，那么这时候是叫做

> 混合效应模型的，为什么叫做混合效应模型呢?
> 这是因为我们不认为亚组是更大总体的一部分或者是总体的一个随机抽样，亚组是穷尽的（比如，你分析性别，只有男女两个水平，这时候包括所有的可能性，你不会再将此进行一进步的推论；其他类似的亚组变量有是否就业，是否有共患病），所以不需要推广到总体。因此，我们说它是固定效应（这和linear mixed modeling称呼预测变量为固定效应一样，所以，固定效应其实是在不同的统计检验里面有不同的含义），这仅仅是为了**强调它不是"随机"的**，不是总体的随机取样。当然，必须要承认的是，这里的固定效应和传统元分析中固定效应模型还是不一样的。这正是造成误解或者不好理解之处。
> 知道了这点，在亚组分析结果那里，也就知道了应该看看固定效应分析结果（Fixed effect analysis）还是混合效应分析（Mixed effectts analysis）结果了。如果亚组内用固定效应模型，那么看固定效应分析结果（Fixed effect analysis）；如果亚组内用随机效应模型，那么看混合效应分析（Mixed effectts analysis）结果。

也就是多个亚组，每个亚组都有自己的固定效应，**主要强调的是不是随机的，不是从更大的群体中随机抽出来了，而是固定的水平。**

- 在第8章中，我们将展示亚组分析只是元回归的一个特殊情况，为此我们也使用混合效应模型

此外，也有可能不能假定子组水平是固定的。想象一下，我们想要评估效应大小是否取决于观察到的效应的位置。一些研究评估了以色列、意大利、墨西哥和中国大陆的影响。有人可能会说，“原产国”不是一个固定水平的因素：世界上有很多很多国家，我们的研究只是“随机”选择。

在这种情况下，不将子组作为固定的模型来建模是有意义的，但也让我们的模型将国家之间的可变性作为随机效应来估计。这导致了一个多层次的模型，我们将在第10章中介绍。

▼ 7.2 Limitations & Pitfalls of Subgroup Analyses

亚组分析是检验调节变量（moderator）的一个好手段

- 元分析所纳入的样本量通常比单个原始研究多几倍的数量级，提高了统计效力（因为合并了该话题的所有研究可用的证据）

Unfortunately, however, this does **not necessarily provide us with more statistical power** to detect subgroup differences. There are several reasons for this (L. V. Hedges and Pigott 2004):

1. 不一定能反映第 G 个组的真实效应，因为g组内**研究间异质性过大**，池化效应量的置信区间更大，可能找不出亚组之间的差异（即使真的有差异），由于**置信区间过宽泛而导致的重叠**：First, remember that in subgroup analyses, the results within subgroups are usually pooled using the random-effects model. If there is substantial between-study heterogeneity within the subgroup, this will decrease the precision (i.e. increase the standard error) of the pooled effect
2. 我们可以预期亚组分析的真实效应量差异比常规meta分析所预期的的小 → 统计功效低会导致 → 检测差异更困难 → 更容易得出“无显著差异”的结论，但这可能是因为**统计效能不足**，而不是差异不存在：In the same vein, statistical power is also often low because the effects we want to detect in subgroup analyses are much lower than in normal meta-analyses. Imagine that we want to examine if effects differed between studies assessing an outcome of interest through **self reports** versus **expert ratings**. Even if there is a difference, it is very likely to be small. It is often possible to find a significant difference between treatment and control groups. Yet, detecting effect size differences **between studies** is usually much harder, because the differences are smaller, and more statistical power is needed、

- 亚组之间检测的结果没有差异并不代表着真的没有差异，常因为统计效力不够或者单个亚组的研究间异质性过于大，导致我们不知道有没有差异：If we do **not** find a difference in effect sizes between subgroups, this does not automatically mean that the subgroups produce **equivalent** outcomes.
- 为此，我们可以通过 `subgroup power analysis` **检查是否存在统计效力的问题：**我们可以通过事先进行**亚组统计效能分析**来检查统计功率是否在亚组分析中存在问题。在这样的分析中，我们可以检查我们能够在亚组分析中检测到的最小效应大小差异。 In chapter 14.3 in the "Helpful Tools" section, as a general rule of thumb, that subgroup analyses only make sense when your meta-analysis contains at least K= 10，每组不少于3。

进行亚组分析要基于事前先验：

- 必须有**明确的理论基础、生物机制或先前研究**支持为什么该变量可能会影响干预效果。
- 不能只是"事后寻找显著差异"(即p-hacking)。
- 为了避免ecological bias，明确大于多少岁和小于多少岁的才被纳入进亚组分析。

👉：我们假设运动干预对青春期肥胖者的效果不同于儿童，是因为青春期脂肪代谢和激素水平不同。

3. 亚组分析实际上是观察性研究（即依然有混杂因素存在）：meta通常纳入的随机对照试验是完全随机的分配，组间的差异仅由不同的干预导致，Subgroup analyses, even when consisting `solely of randomized studies`, `cannot show causality` ，可能是由于对照组的不同导致的效应差异，而不是两个干预之间存在干预差异，这在方法学上是一个"混杂因素"(confounding) 的问题——你比较的不是"治疗本身的效果"，而是"治疗 + 方法差异"的综合效应。这里点明关键：**亚组之间差异≠治疗本身的差异**，可能是其他因素引起的（比如不同样本、对照方式、测量工具等）。
   所以：
   亚组分析不能用于得出**因果推论**；即使基础研究是 RCT，亚组之间的比较也具有**观察性性质**；观察到的亚组差异可能是由于**混杂变量（confounding factors）**；这提醒我们，在解释亚组结果时必须非常谨慎。如果你想要解释某种治疗是否真的比另一种好，亚组分析提供的是"提示"，不是"证据"。

4. 另外一个重大的缺陷就是，拿年龄举例子：我们想比较年龄亚组，根据汇总信息（平均值）将研究分类为子组可能很诱人。

   Schwarzer及其同事将研究的平均年龄作为一个常见的例子。假设你想评估老年人（65岁以上）和普通成年人之间的影响是否不同。因此，根据报告的平均年龄是高于还是低于65岁，将研究分为这两类。
   如果我们发现，在平均年龄较高的亚组中，影响更大，我们可能会直觉地认为，这表明，在老年人中，影响更大。但这种推理存在严重缺陷。当一项初步研究的平均年龄超过65岁时，仍有可能包括相当一部分年龄小于65岁的个体。反之亦然，即使在平均年龄较低的情况下，一项研究也完全有可能纳入了很大一部分65岁以上的人。

   - 这就是**生态学偏差的核心问题：不能从群体统计指标（如均值）推断个体因果关系。**
   - 每当我们想要使用聚合（宏观）级别上的关系来预测个体（微观）级别上的关联时，就会出现这种情况（生态学谬误）。

   ### ✅ 那什么情况下可以说"没有生态偏差"?

   只有当你**明确知道这个研究的全部样本都在某一年龄段内**，比如：

   - "研究招募标准为仅限12-17岁青少年"(说明所有人 <18)；
   - "仅纳入成年人 ≥18岁"(说明所有人 ≥18)；

   这时候，才可以放心地将这项研究归入"青少年组"或"成人组"，不会产生**生态学偏差**。

   但是！！！也有可能出现confounding，例如干预方式、结局测量等混杂因素的影响。

Subgroup Analysis: Summary of the Dos & Don'ts

1. Subgroup analyses depend on the statistical power, so it usually makes no sense to conduct one when the number of studies is small (i.e. K < 10).
2. If you do not find a difference in effect sizes between subgroups, this does **not** automatically mean that the subgroups produce **equivalent** results.
3. Subgroup analyses are purely **observational**, so we should always keep in mind that effect differences may also be caused by confounding variables.
4. It is a bad idea to use aggregate study information in subgroup analyses, because this may introduce ecological bias.

▼ 7.3 Subgroup Analysis in *R*

在`{meta}`中的每个元分析函数中，都可以指定subgroup argument：

This tells the function which effect size falls into which subgroup and runs a subgroup analysis.

- only thing we have to take care of is that studies in the same subgroup have absolutely identical labels.

Here, we want to examine：

if there are effect size differences between studies with a high versus low risk of bias. The risk of bias information is stored in the `RiskOfBias` column.

- Let us have a look at this column first. In our code, we use the `head` function so that only the first few rows of the data set are shown.

```
head(ThirdWave[,c("Author", "RiskOfBias")])

##            Author RiskOfBias
## 1     Call et al.       high
## 2 Cavanagh et al.        low
## 3   DanitzOrsillo       high
## 4  de Vibe et al.        low
## 5  Frazier et al.        low
## 6  Frogeli et al.        low
```

- 当我们使用`metagen`计算meta分析时，这些信息（ROB的标识符，为了正确的进入亚组）被内部保存在m.gen对象中。

> When we calculated the meta-analysis using `metagen`, this information was saved internally in the `m.gen` object.

In our example, we want to apply the fixed-effects (plural) model and assume that studies within subgroups are pooled using the random-effects model. Given that `m.gen` contains results for the random-effects model (because we set `comb.fixed` to `FALSE` and `comb.random` to `TRUE`), there is nothing we have to change. Because the original meta-analysis was performed using the random-effects model, `update` automatically assumes that studies within subgroups should also be pooled using the random-effects model.

Therefore, the resulting code looks like this:

```
update(m.gen,
       subgroup = RiskOfBias,
       tau.common = FALSE)
```

```
## Review:     Third Wave Psychotherapies
##
## Number of studies combined: k = 18
##
##                              SMD            95%-CI    t  p-value
## Random effects model (HK) 0.5771 [ 0.3782; 0.7760] 6.12 < 0.0001
## Prediction interval              [-0.0572; 1.2115]
##
## Quantifying heterogeneity:
##  tau^2 = 0.0820 [0.0295; 0.3533]; tau = 0.2863 [0.1717; 0.5944]
##  I^2 = 62.6% [37.9%; 77.5%]; H = 1.64 [1.27; 2.11]
##
## Test of heterogeneity:
##      Q d.f. p-value
##  45.50   17  0.0002
##
## Results for subgroups (random effects model (HK)):
##                     k    SMD           95%-CI  tau^2    tau     Q   I^2
## RiskOfBias = high   7 0.8126 [0.2835; 1.3417] 0.2423 0.4922 25.89 76.8%
## RiskOfBias = low   11 0.4300 [0.2770; 0.5830] 0.0099 0.0997 13.42 25.5%
##
## Test for subgroup differences (random effects model (HK)):
##                   Q d.f. p-value
## Between groups 2.84    1  0.0917
##
## Details on meta-analytical method:
## - Inverse variance method
## - Restricted maximum-likelihood estimator for tau^2
## - Q-Profile method for confidence interval of tau^2 and tau
## - Hartung-Knapp (HK) adjustment for random effects model (df = 17)
## - Prediction interval based on t-distribution (df = 16)
```

- because biased studies are more likely to overestimate the effects of a treatment.
- the test is 0.09, which is larger than the conventional significance threshold, but still indicates a difference on a trend level（可能因为样本量不足、效应量较小、功效不足等原因未达显著）

We can also check:

the results if were to assume a **common** `tau^2` estimate in both subgroups. We only have to set `tau.common` to `TRUE`.

```
update(m.gen, subgroup = RiskOfBias, tau.common = TRUE)
```

```
## [...]
##                     k    SMD           95%-CI  tau^2    tau     Q   I^2
## RiskOfBias = high   7 0.7691 [0.2533; 1.2848] 0.0691 0.2630 25.89 76.8%
## RiskOfBias = low   11 0.4698 [0.3015; 0.6382] 0.0691 0.2630 13.42 25.5%
##
## Test for subgroup differences (random effects model (HK)):
##                    Q d.f. p-value
## Between groups  1.79    1  0.1814
## Within groups  39.31   16  0.0010
##
## Details on meta-analytical method:
## - Inverse variance method
## - Restricted maximum-likelihood estimator for tau^2
##   (assuming common tau^2 in subgroups)
## [...]
```

- In the output, we see that the estimated between-study heterogeneity variance is $\tau 2 = 0.069$, and identical in both subgroups.

- The test of subgroup differences again indicates that there is not a significant difference between studies with a low versus high risk of bias (p = 0.181)

> **reporting the Results of a Subgroup Analysis**
> The results of subgroup analyses are usually reported in a table displaying the estimated effect and heterogeneity in each subgroup, as well as the p-value of the test for subgroup differences.

1. 当我们不确定或假设两组之间的研究间异质性不同时，使用各个G组的独立tau^2方是更合适的选择，但是需要足够的研究数量：每个亚组研究数量 ≥ 6–7。
2. 也可以都做：`common = TRUE / FALSE` 然后观察两种方法对结果的解释是否相似。

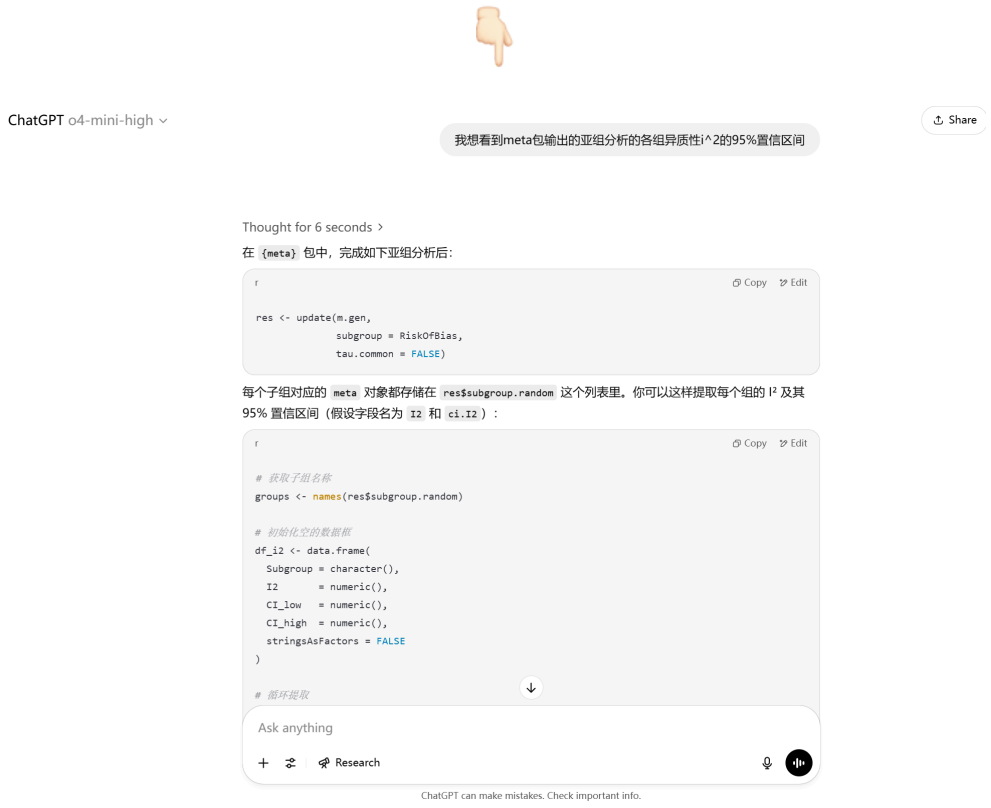
Risk of Bias	g	95%CI	p	I^2	95%CI I^2	p(subgroup)
High	0.81	0.28–1.34	0.009	0.77	0.51–0.89	0.092
Low	0.43	0.28–0.58	< 0.001	0.25	0.00–0.63	

- if you want to extract the subgroup-specific p values, it is necessary to save the results of `update` to an object, and then extract the `pval.random.w` element from that object using the `$` operator.

▼ TIPS

经过非常长的一段时间的摸索（大约1h），我想弄清楚**两个亚组分别的I^2的95%CI**（因为根据代码，并不现实，显示的仅仅是I^2本2）。

1. 问了AI



结果完全无法用（完全瞎扯，甚至函数的命令都是错的）

2. 使用$符号进行extraction of info，奈何没有i^2的95CI这一信息，仅有的是

```
> res$I2.w
     high       low
0.7682416 0.2548068
>
```

3. 找到了meta的官方文档（meta: metabin – R documentation – Quantargo），在value列找到了这个函数命令的调用：`lower.I2.w, upper.I2.w =` 亚组内异质性统计量I^2的置信下限和置信上限。

`I2.w`	Heterogeneity statistic I^2 within subgroups - if `byvar` is not missing.
`lower.I2.w, upper.I2.w`	Lower and upper confidence limit for heterogeneity statistic I^2 within subgroups - if `byvar` is not missing.

4. 也可以通过写代码实现：相当于将亚组作为主效应进行分析（因为主效应会自动在输出的模型结果中显示I^2的95%CI）

```
library(dplyr)

revise_version <- ThirdWave %>%
  filter(RiskOfBias == "high")

A <- meta::metagen(TE = TE,
                   seTE = seTE,
                   studlab = Author,
                   data = revise_version,
                   sm = "SMD",
                   fixed = FALSE,
                   random = TRUE,
                   method.tau = "REML",
                   method.random.ci = "HK",
                   title = "Third WaIe Psychotherapies with low"

)

A

## Quantifying heterogeneity (with 95%-CIs):
## tau^2 = 0.2423 [0.0569; 1.5610]; tau = 0.4922 [0.2386; 1.2494]
## I^2 = 76.8% [51.5%; 88.9%]; H = 2.08 [1.44; 3.01]
```

▼ summary

- Although there are various ways to assess the heterogeneity of a meta-analysis, these approaches do not tell us **why** we find excess variability in our data. Subgroup analysis allows us to test hypotheses on why some studies have higher or lower true effect sizes than others.
- For subgroup analyses, we usually assume a **fixed-effects (plural) model(水平是固定的，g组内是随机效应的)**. Studies within subgroups are pooled, in most cases, using the random-effects model. Subsequently, a Q test based on the overall subgroup results is used to determine if the groups differ significantly.
- The subgroup analysis model is called a “fixed-effects” model because the different categories themselves are assumed to be fixed（不是在更大的样本中被随机抽取的）. The subgroup levels are not seen as random

draws from a universe of possible categories. They represent the only values that the subgroup variable can take.

- When calculating a subgroup analysis, we have to **decide whether separate or common estimates of the between-study heterogeneity should be used** to pool the results within subgroups（我们不能假定亚组间共享异质性，使用共同估计的亚组异质性仅出于实用性考虑，即亚组内研究过少，当研究数量够时，优先独立亚组内异质性）.
- Subgroup analyses are not a panacea. They often lack the statistical power needed to detect subgroup differences. Therefore, a non-significant test for subgroup differences does **not** automatically mean that the subgroups produce equivalent results.（请看有用工具的功效分析）