



国家药品监督管理局药品审评中心

CENTER FOR DRUG EVALUATION, NMPA
CENTRE FOR DRUG EVALUATION, NMPA

此页面上的内容需要较新版本的 Adobe Flash Player。



当前位置：新闻中心>>工作动态>>通知公告>>新闻正文

国家药监局药审中心关于发布《药物临床试验亚组分析指导原则（试行）》的通告（2020年第64号）

发布日期：20201231

为进一步探索药物临床试验中不同特征患者的疗效和安全性差异，评估不同亚组可能的获益-风险，为申办方提供药物临床试验中亚组分析的设计、实施和评价的指导性建议，药审中心组织制定了《药物临床试验亚组分析指导原则（试行）》（见附件）。根据《国家药监局综合司关于印发药品技术指导原则发布程序的通知》（药监综药管〔2020〕9号）要求，经国家药品监督管理局审查同意，现予发布，自发布之日起施行。

特此通告。

国家药品监督管理局药品审评中心

2020年12月31日

附件 1：

《药物临床试验亚组分析指导原则（试行）》.pdf

Copyright © 国家药品监督管理局药品审评中心 All Right Reserved.

地址：中国 北京市朝阳区建国路128号 邮编：100022

总机：8610-68585566 传真：8610-68584189 备案序号：京ICP备09013725号

药物临床试验亚组分析指导原则

（ 试行 ）

2020 年 12 月

目 录

一、概述.....	1
二、亚组的识别和定义.....	2
(一) 亚组的识别	2
(二) 亚组的定义	3
三、亚组分析的类型.....	3
(一) 探索性亚组分析.....	4
(二) 支持性亚组分析.....	4
(三) 确证性亚组分析.....	5
四、一般考虑	5
(一) 生物学合理性	5
(二) 异质性.....	6
(三) 一致性.....	6
(四) 可信度.....	6
五、确证性亚组分析.....	8
(一) 亚组的选择	8
(二) 试验设计	8
(三) 多重性.....	10
(四) 结果的解释	11
六、支持性亚组分析.....	11
七、其他考虑	13
(一) 非劣效试验中的亚组分析	13
(二) 安全性亚组分析.....	14
(三) 与监管机构的沟通.....	15
八、参考文献.....	15
附录1. 词汇表	19
附录2. 中英文对照表.....	21
附录3. 亚组分析与设计研究案例	22

药物临床试验亚组分析指导原则

（试行）

一、概述

参与临床试验的患者由于受各种因素（如遗传学、人口学、环境、病理生理学、合并症、合并用药、区域等）的影响，往往具有不同程度的异质性，从而可能导致试验药物在不同患者中的疗效不同。临床试验中将具有不同特征的患者分组，是探索不同患者人群之间疗效差异的直观方法，同时也是获益-风险评估不可缺少的一部分。

对于事先未设定亚组分析的临床试验，在对亚组结果进行解读和下结论时需要特别慎重。事后根据数据驱动寻找有统计学意义的亚组，会导致总 I 类错误率膨胀，其结果通常不能用于确证该亚组的有效性。此外，因亚组的样本量较少而导致检验效能不足，可能会影响试验药物在这些亚组人群中疗效的精确估计，或者无法得出各亚组间疗效一致的结论。

本指导原则中将目标适应症人群称为总体人群，将通过入排标准纳入临床试验的人群称为全人群。亚组人群（简称亚群）指总体人群中具有某些特征的一个子集，亚组是全人群中的一个子集。亚组分析是指针对试验药物在亚组中的疗效和/或安全性进行

试验设计与统计分析的过程，亚组是亚群的一个样本，将亚组结果推广到亚群时需要考虑亚组对亚群的代表性问题。

本指导原则主要阐述了亚组的识别和定义、亚组分析的类型、一般考虑以及确证性临床试验中的亚组分析等方面的内容，旨在为申办者能够在临床试验中对亚组分析进行正确地设计、实施和评价提供指导性建议。本指导原则主要适用于以支持药品注册上市为目的的确证性临床试验，也可供以非注册为目的的临床试验参考。

二、亚组的识别和定义

（一）亚组的识别

亚组的识别一般基于早期临床试验或确证性临床试验的探索性分析，可以使用定量的方法，如交互树和递归分割树等方法；也可以根据文献报道或者医疗实践积累的知识进行识别。

亚组识别主要关注不同亚组间疗效的差异及其临床意义，本质上属于探索性研究，一般无须控制总 I 类错误率。进行亚组识别时应考虑的因素包括但不限于以下几个方面：①临床上的可解释性；②临床上的可操作性；③药物的作用机理；④定义亚组的变量个数和类型（如连续变量、分类变量）；⑤误分或漏分亚组两种情况带来的风险；⑥亚组识别模型的选择（不宜过于复杂，以避免过拟合）。

（二）亚组的定义

亚组通常由患者的一个或多个内在和/或外在因素（见 ICH E5）来定义，而且应具有一定的临床意义。这些变量通常是基线变量，包括人口学特征（如年龄、性别等）、实验室检查指标、基因组相关标志物、疾病的严重程度或分型、临床状况（如合并症、伴随用药）、地区（如国家、试验中心）和环境因素等。

实际应用中，亚组大多由一到两个变量来定义，例如把研究对象按照年龄分为 18-65 岁和 ≥ 65 岁两个年龄组；又如在抗肿瘤药物研究中，把患者按照 ECOG 评分和基因突变（或基因表达）水平，分成不同的亚组。需要注意的是，使用多于两个变量定义亚组的情况比较少见，主要是因为：①定义亚组的变量过多，容易出现误分的情况；②使用多个变量定义亚组时，对亚组分析结果的解释往往出现困难。当亚组的定义确实需要涉及多个变量时，可以考虑将这些变量通过一定的方法转换成单一的风险分数，并以此来定义亚组。例如抗肿瘤药物试验中常用的 ECOG 评分，反应了患者自我保健、日常活动和体能状况等综合能力。

三、亚组分析的类型

在进行研究计划时，申办者应充分探索影响试验药物疗效的因素，并根据相关证据强度以及对药物疗效的影响程度等方面的知识确定相关影响变量。在制定方案时应充分讨论已知影响变量

并具体说明对这些变量的考虑，如作为统计学检验策略中的一部分、作为随机分层变量或作为事先规定的支持性亚组分析变量等。

根据研究目的，亚组分析分为探索性亚组分析、支持性亚组分析和确证性亚组分析。对于探索性亚组分析，亚组既可以在设计阶段事先定义、也可以在分析阶段事后定义（如根据数据驱动划分亚组）。对于支持性亚组分析，亚组一般应在临床试验的设计阶段事先定义，并在试验方案中详细描述。而对于确证性亚组分析，则必须在临床试验的设计阶段事先对亚组进行定义，并在试验方案中详细描述。

（一）探索性亚组分析

探索性亚组分析主要用于早期临床试验或在确证性临床试验的探索性分析中，其目的是发现药物在不同亚组间疗效和/或安全性方面的差异，进而提出研究假设，以待在后续的临床试验中进一步探索和验证。因此，探索性亚组分析主要关注的是其结果在生物学上的合理性或临床上的可解释性，是否进行多重性调整由申办者自行决定。

（二）支持性亚组分析

在以考察试验药物在全人群中的疗效为目的的确证性临床试验中，当全人群的主要终点同时具有统计学意义和临床意义时，通常还需要进行支持性亚组分析，目的是进一步考察试验

药物在各个亚组中疗效的一致性。如果试验药物在各亚组间的疗效一致，可为药物适用于全人群提供进一步支持性证据；如果各亚组间的疗效不一致，特别是方向相反时，则亚组分析结果的解释可能会出现困难，需要对其做进一步的分析和研究。当全人群的主要终点没有统计学意义或临床意义时，亚组分析结果只能为进一步研究提供线索。

（三）确证性亚组分析

确证性临床试验中，按照临床试验方案和/或统计分析计划中预先规定的亚组和多重性调整方法，考察试验药物在目标亚组和/或全人群中的疗效，其结果应同时具有临床意义和统计学意义，以支持药物说明书的撰写。确证性临床试验也可以对目标亚组进行确证性亚组分析，而对其它（非目标）亚组进行支持性或探索性亚组分析，以支持试验药物在目标亚组中的有效性和安全性的结论，并为非目标亚组的进一步研究提供线索。

四、一般考虑

除临床意义和获益-风险评估外，亚组分析中需要考虑的因素包括但不限于以下几个方面。

（一）生物学合理性

生物学合理性指亚组的生物学特征与研究终点（如主要疗效终点、不良事件等）之间的因果关联在生物学上的可解释性。例

如，不同患者之间潜在的病理生理学或遗传学的差异可能导致药物治疗效果的不同，亚组分析能够据此给出合理的解释。

（二）异质性

亚组分析的主要目的是为了更好地了解试验药物在各亚组和全人群中的疗效，而是否需要和如何设计亚组关键在于临床试验中目标人群的异质性。异质性与预后因素或预测因素对试验药物疗效的影响程度有关。虽然试验前可能无法识别所有潜在的异质性因素，但在计划临床试验时，申办者应充分讨论已知的预后因素和预测因素对药物疗效评价可能带来的影响。

临床试验方案中应明确研究的目标人群。通常，入排标准的限制条件越严格，招募的患者异质性就越小；反之，宽松的入排标准可能导致入组患者的异质性增加，此时进行亚组分析就显得非常必要。

（三）一致性

一致性是指不同亚组间显示出相同或相似的治疗效果。它反映了亚组结果对全人群疗效适用于试验总体人群的支持程度。在临床试验中需要考虑药物在所关心的亚组间的疗效差异，若亚组间结果不一致则需进一步评估不一致的原因和在特定亚组的疗效。

（四）可信度

可信度是指亚组分析结果的可靠性或证据强度。可信度评估包括但不限于以下几个方面：①亚组是否预先定义；②定义亚组的变量是否具有生物学上的合理性，包括对患者预后因素的选择或治疗应答的预测是否有科学依据；③划分亚组的依据是否充分；④亚组分析结果的可重现性，即在相同或相似条件下的其他临床试验中，具有相同或相似的亚组效应。

在研究设计阶段，有时基于先验知识指定用于亚组分析的变量。这种预先指定亚组变量的方法，通常用在确证性和支持性亚组分析中，蕴含了亚组之间疗效有差异的推测，因此得到的亚组分析结果具有一定的可信度。然而，即使亚组分析的变量不是预先指定的，也要予以重视，尤其是对在安全性亚组分析中发现的亚组之间的差异，要特别关注其生物学上的合理性和结果的可重现性。

当亚组的样本量不足、无法准确估计药物在亚组中的疗效时，应主要考察其生物学上的合理性和结果的可重现性。当亚组疗效在试验条件（如研究设计、目标人群、亚组定义、治疗方案、结局测量等）相同或相似的一系列临床试验中一致时，即使没有明确的临床和生物学方面的解释，亚组结果也具有一定的可信度。另外，虽然随机化可以使不同治疗组间入组患者的基线变量分布趋于平衡，但由于亚组内样本量的减少，可能会出现亚组内不同

处理组间基线不均衡的情况，因此必须检查药物在各亚组间的疗效差异是否是由于处理组间基线分布不均衡所致。

五、确证性亚组分析

确证性亚组分析是在确证性临床试验中对事先指定的目标亚组进行假设检验的分析，目标亚组可以作为主要或共同主要分析人群。确证性亚组分析的临床试验需要考虑的关键问题包括但不限于以下几个方面：①亚组的选择；②试验设计类型（如固定样本设计、适应性设计）；③多重性；④亚组分析结果的解释。

（一）亚组的选择

确证性亚组分析应在临床试验方案中预先规定目标亚组。关于亚组的选择，如果是基于医疗知识或实践，例如按照疾病严重程度、人口学特征（性别、年龄等）或已知的能够分辨亚组的生物标志物（如基因突变）进行分类，通常具有一定的临床意义。另一方面，亚组的定义是否合理取决于亚组分类器（如标志物）是否能够可靠地识别最有可能从药物中获益的亚组人群。亚组分类器的确定通常基于早期临床研究数据，由于样本量往往不足，因此分类器的性能可能有限，在研究设计时要考虑到这一问题。

（二）试验设计

把亚组作为主要或共同主要分析人群的试验设计应考虑亚组的样本量以及是否使用分层随机化等关键问题。

在进行确证性亚组分析的样本量估计时，除了常用的试验设计参数（如期望治疗效应大小及其变异度、I 类错误和II类错误等）之外，还必须考虑亚组分类器分辨亚组的准确度，以及亚组人群在总体人群中所占的比例。

如果有可靠的证据表明试验药物在不同亚组中的疗效不同，且其具有生物学上的合理性和重要的临床意义，可以采用固定样本量设计，以验证药物的疗效。此时，临床试验的目标人群可以是亚组人群和/或总体人群，设计时至少有以下三种方案（以标志物举例）：

（1）如果只有标志物阳性的患者才能从试验药物中获益，入组患者可以仅限于这个亚组。

（2）如果标志物阳性和阴性的患者都能从试验药物中获益，但阳性患者获益高于阴性患者，而试验的主要目的是验证药物在阳性患者中的疗效，则样本量和药物疗效的估计可以主要针对标志物阳性的亚组，但标志物阴性的亚组也可纳入试验，以便更好地了解试验药物在该人群中的疗效，用于获益-风险评估或后续研究设计。

（3）如果标志物阳性和阴性的患者都能从试验药物中获益，但不能确定哪个亚组人群的获益更大，而试验的主要目的是验证药物在总体人群中的疗效，则患者可以在总体人群中招募，也可

以在两个亚组人群中进行分层随机化。如果是后者，则需要注意各亚组样本量的比例应与总体人群中各亚组人群患者的比例相近，以避免由于过多纳入疗效较好的亚组人群而夸大了药物在总体人群中的疗效。

如果没有充分的证据表明不同亚群间的疗效差异具有临床意义时，可以采用适应性设计的方法进行亚组的选择。例如，当试验药物在标志物阳性和阴性患者中的疗效不确定时，可以考虑两阶段适应性设计。第一阶段试验的数据可用于估计药物在亚组中的疗效，然后据此调整第二阶段的入组人群及其样本量。对于适应性设计中的亚组选择，应重点考虑试验设计和统计分析方法的有效性。

（三）多重性

多重性是确证性亚组分析中需要重点关注的问题之一。对于将全人群和亚组作为共同主要分析人群的临床试验，由于要进行多次检验，因此如果不进行多重性调整，会增加总 I 类错误率。为了将总 I 类错误率控制在预设的水平，有多种多重性调整方法，不同的多重性调整方法各有其优缺点。多重性调整的方法应在临床试验方案和/或统计分析计划中事先指定并阐明方法的合理性。

值得注意的是，由于亚组属于全人群的一部分，亚组和全人群的统计量呈正相关，因此使用统计量的联合分布来确定检验界

值，可以提高检验效能；但通常由于数据有限，估计统计量的相关系数往往不可靠，并可能增加 I 类错误。因此，使用基于统计量的联合分布来确定检验界值的多重性调整方法需要特别谨慎。

（四）结果的解释

亚组分析的结果提示不同亚组人群可能的获益与风险，因此直接影响决策和说明书的撰写。

在一项以预先规定的某一目标亚组和全人群作为共同主要分析人群的确证性临床试验中，与对照组相比，试验药物的疗效经过多重性调整后的统计分析可以得出下述四个统计学结论之一：①全人群中的疗效差异有统计学意义而目标亚组的疗效差异无统计学意义；②目标亚组的疗效有统计学意义而全人群的疗效无统计学意义；③全人群和目标亚组的疗效都有统计学意义；④全人群和目标亚组的疗效都没有统计学意义。需要注意的是，如果结论③是由于药物在目标亚组中的疗效较大所致，而其余亚组很少甚至不获益时，将其使用限定于该目标亚组可能更合适。

对亚组分析结果解释时，除了要考虑统计学意义，还需要考虑临床意义，只有同时具有统计学意义和临床意义，才可支持药物的上市和说明书的撰写。

六、支持性亚组分析

一般情况下，确证性临床试验的目的是验证药物在全人群中

的有效性以及各主要亚组之间疗效的一致性，后者通过支持性亚组分析来实现。常用的支持性亚组分析方法包括但不限于：①描述性分析（如疗效的点估计）和区间估计；②图形显示（如森林图）；③模型法（包括协变量校正分析在内的估计、亚组与处理交互作用的检验等）。使用何种方法应在临床试验方案及统计分析计划中充分考虑和说明，当不同方法间结果不一致时需要进一步分析其不一致的原因。

需要指出的是，当各亚组的样本量较少或亚组间样本量分布不均衡时，药物与亚组变量交互作用的检验效能往往不足。为此，可以考虑选取大于 0.05 的检验水准进行检验。如果交互作用显著，则提示试验药物在各亚组间的疗效可能不同。

在确证性临床试验中，如果试验药物的疗效在全人群中具有统计学意义，通常需要报告亚组分析的结果（如各亚组疗效的点估计、置信区间、森林图、交互作用等）。需要注意的是，如果在全人群疗效和各亚组的疗效同时进行分析，则各亚组疗效与全人群疗效的方向应该一致，否则需要进一步分析其不一致的原因。

如果各亚组间的疗效差异具有临床意义，应考虑按照以下几个步骤探索其可能的原因：

（1）定义亚组的变量是否与相应的预后或预测因素有关。这些因素通常从早期临床试验、文献报告或医疗实践中获知，可

以是内在因素（如遇药物代谢酶相关的基因多态性等遗传因素）、外在因素（如疾病严重程度、吸烟状况和 BMI 等）或治疗方法（如不同剂量的合并用药）等。

（2）如果定义亚组的变量与相应的预后或预测因素有关，则应进行进一步分析。如首先分析预后因素是否具有预测作用，如果有预测作用说明该因素会影响患者对药物的应答或疗效；然后，分析预后因素在各亚组的试验组和对照组中的分布是否均衡，如果不均衡则其可能是造成亚组间疗效差异的原因。

（3）如果通过上述分析，亚组间疗效的差异仍无法解释，则需考虑进一步的探索性分析以识别造成此差异的其他可能原因，对试验结果给出合理的解释。

七、其他考虑

（一）非劣效试验中的亚组分析

在只有阳性对照的双臂非劣效临床试验中的亚组分析需十分谨慎。当在各亚组间对试验药疗效进行非劣效评价时，其结果依赖于阳性对照药相对于安慰剂的疗效在各亚组中的一致性。若阳性对照药相对于安慰剂的效应与定义亚组的变量有关时，则非劣效界值的选择应考虑这些特征变量在研究人群中的分布。由于非劣效界值的确定通常使用历史数据，因此，如果定义亚组的变量在试验人群中的分布与历史数据不同，则会影响非劣效检验结

论的正确性。要强调的是，非劣效界值的确定应当基于阳性对照药疗效（相对于安慰剂）尽可能多的历史数据的 meta 分析结果，而不是基于主观选择部分试验数据或者部分亚组人群数据的分析结果。

（二）安全性亚组分析

安全性亚组分析主要用于研究与药物安全性相关的风险因素，即探究具有一定特征的亚组人群可能对药物产生的不良反应。安全性亚组分析可以由预后因素（如年龄、疾病分期、是否有并发症等）或预测因素（如基因分型等）定义亚组。

安全性亚组分析与有效性亚组分析略有不同，在有效性评价中划分亚组的因素用于划分安全性评价的亚组可能并不合适，因为：①药物的有效性和安全性一般使用不同的临床终点或者替代终点，其机理可能不同；②定义有效性亚组的变量（例如靶点受体）和安全性亚组的变量（例如年龄等因素）往往不同；③对可能导致潜在危害的药物，不同亚组患者的耐受程度可能不同（例如病情严重的患者为了治疗获益可能更容易耐受一定程度的不良反应），此时亚组的获益-风险评估非常重要。

亚组人群的安全性分析具有一定的挑战，特别是对低发生率或者潜伏期较长的安全性事件，由于样本量较小或随访时间较短，较难在试验期间发现和验证。如果现有数据提示试验药物在特定

亚组人群中可能与某一严重不良事件（SAE）有潜在的关联，则需进一步评估试验药物在该亚组人群中是否会引起严重的不良反应，并且需要对其进行获益-风险评估。

（三）与监管机构的沟通

当临床试验设计包含确证性亚组分析时，鼓励申办者就临床试验设计中的关键问题与监管机构进行沟通，沟通的内容包括但不限于试验设计的类型、亚组的选择、I 类错误控制、一致性和可信性等方面。对于支持性亚组分析中的关键问题鼓励申办者与监管机构进行沟通。

八、参考文献

1. ICH. ICH E5: Ethnic Factors in the Acceptability of foreign Clinical Data E5 (R1). ICH Harmonised Tripartite ([https://database.ich.org/sites/default/files/E5_R1 Guideline.pdf](https://database.ich.org/sites/default/files/E5_R1_Guideline.pdf)), 1998.
2. EMA. Guideline on the investigation of subgroups in confirmatory clinical trials. European Medicines Agency (EMA) (https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-investigation-subgroups-confirmatory-clinical-trials_en.pdf), 2019.
3. Tanniou J, Van Der Tweel I, Teerenstra S, et al. Subgroup analyses in confirmatory clinical trials: time to be specific about their purposes. BMC Medical Research Methodology. 2016,

16(1):20.

4. Alosh M, Huque MF, Bretz F, et al. Tutorial on statistical considerations on subgroup analysis in confirmatory clinical trials. *Statistics in Medicine*. 2017, 36(8): 1334-1360.
5. Lipkovich I, Dmitrienko A, B. R. D' Agostino S. Tutorial in biostatistics: data-driven subgroup identification and analysis in clinical trials. *Stat Med*. 2017, 36(1): 136-196.
6. Burke JF, Sussman JB, Kent DM, et al. Three simple rules to ensure reasonably credible subgroup analyses. *BMJ*. 2015, 351: h5651.
7. Loh WY, Cao L, Zhou P. Subgroup identification for precision medicine: A comparative review of 13 methods. 2019, 9(5): e1326.
8. Ondra T, Dmitrienko A, Friede T, et al. Methods for identification and confirmation of targeted subgroups in clinical trials: A systematic review. *Journal of Biopharmaceutical Statistics*. 2016, 26(1): 99-119.
9. Rothwell, PM. Treating individuals 2. Subgroup analysis in randomised controlled trials: importance, indications, and interpretation. *Lancet*. 2005, 365(9454): 176-86.
10. Dmitrienko A, Muysers C, Fritsch A, et al. General guidance on exploratory and confirmatory subgroup analysis in late-stage clinical trials. *Journal of Biopharmaceutical Statistics*. 2016,

26(1): 71-98.

11. Dmitrienko A, D'agostino R. Traditional multiplicity adjustment methods in clinical trials. *Statistics in Medicine*. 2013, 32(29): 5172-5218.
12. Zhao YD, Dmitrienko A, Tamura R. Tamura, Design and Analysis Considerations in Clinical Trials With a Sensitive Subpopulation. *Statistics in Biopharmaceutical Research*. 2010, 2(1): 72-83.
13. Woodcock J, Lavange LM. Master Protocols to Study Multiple Therapies, Multiple Diseases, or Both. *N Engl J Med*. 2017, 377(1): 62-70.
14. Stallard N, Todd S, Parashar D, et al. On the need to adjust for multiplicity in confirmatory clinical trials with master protocols. *Ann Oncol*. 2019, 30(4): 506-509.
15. Cui L, Hung HM, Wang SJ, et al. Issues related to subgroup analysis in clinical trials. *J Biopharm Stat*. 2002, 12(3): 347-58.
16. Bradley, CA. Pembrolizumab improves OS across PD-L1 subgroups. *Nat Rev Clin Oncol*. 2019, 16(7): 403.
17. Ransohoff, RM. Natalizumab for multiple sclerosis. *N Engl J Med*. 2007, 356(25): 2622-9.
18. Kachuck, NJ. Registries, research, and regrets: is the FDA's post-marketing REMS process not adequately protecting patients? *Ther Adv Neurol Disord*. 2011, 4(6): 339-47.

- 19.Chen J, Heyse J, Lai TL. Medical Product Safety Evaluation: Biological Models and Statistical Methods. New York: Chapman and Hall/CRC (<https://doi.org/10.1201/9781351021982>). 2018.
- 20.Hemmings, R. An Overview of Statistical and Regulatory Issues in the Planning, Analysis, and Interpretation of Subgroup Analyses in Confirmatory Clinical Trials. Journal of Biopharmaceutical Statistics. 2014, 24(1): 4-18.

附录 1. 词汇表

重现性 (Reproducibility): 临床试验的结果可在相同或相似的研究条件下重现的能力。

分类器 (Classifier): 分类器是对样本进行分类算法的统称，包括决策树、回归模型、支持向量机、神经网络等多种算法。其主要特点是可以使用已有的具有分类标签的数据，拟合一个函数或分类模型，使其能够对新的无标签数据所属类别进行预测。

风险分数 (Risk Score): 将患者多个基线特征变量通过一定的方法转换成能够预测患者结局的单一变量，后者通常称为风险分数。

I类错误 (Type I Error): 当原假设为真而检验结果拒绝了原假设的错误，其概率通常用 α 表示，并用其作为检验水准。

内在因素 (Intrinsic Factors): 指患者个体的遗传、生理和病理特征，是一个人“内在”的生物特征，而不是由个人所处的环境决定。

确证性亚组分析 (Confirmatory Subgroup Analysis): 按照预先制定的计划，以特定的亚组为研究的主要目标人群，能够对I类错误进行控制的亚组分析，以确证目标亚群患者的疗效。

探索性亚组分析 (Exploratory Subgroup Analysis): 通过亚组分析探索药物对不同亚组患者可能存在的疗效和/或安全性上的差异，主要关注的是其结果在生物学或者临床上的合理性。

异质性 (Heterogeneity) : 临床试验中的异质性体现在个体和群体两个水平，前者通常是指患者间具有不同的特征，个体性质或状态的不同可能会导致不同的患者对治疗有不同的应答；后者通常是指不同中心、种族、地域等患者群体具有不同的特征，有可能导致不同患者群体对治疗有不同的应答。

外在因素 (Extrinsic Factors) : 指与患者居住地的环境和文化相关的因素。外在因素往往与遗传和基因无关，更多是由文化和行为决定。

亚组分析 (Subgroup Analysis) : 将患者根据其特征变量值分成不同的亚组，并估计各亚组的疗效和/或安全性的分析策略。

支持性亚组分析 (Supportive Subgroup Analysis) : 通过亚组分析探讨药物的疗效和/或安全性在各亚组中的一致性，为试验药物适用于全人群的结论提供进一步的支持。

附录 2. 中英文对照表

中文	英文
重现性	Reproducibility
递归分割树方法	Recursive Partitioning Tree Method
多重性调整	Multiplicity Adjustment
分层随机化	Stratified Randomization
交互树方法	Interaction Tree Method
可信度	Credibility
内在因素	Intrinsic Factors
适应性设计	Adaptive Design
外在因素	Extrinsic Factors
异质性	Heterogeneity
一致性	Consistency
准确度	Accuracy
总 I 类错误率	Familywise Error Rate (FWER)

附录 3. 亚组分析与设计研究案例

示例 1 确证性亚组分析案例:帕博利珠单抗治疗非小细胞肺癌的亚组分析

帕博利珠单抗在获批用于 PD-L1 肿瘤细胞阳性比例分数 (tumor proportion score, TPS) $\geq 50\%$ 、且 EGFR 基因敏感突变阴性和 ALK 融合基因阴性的转移性非小细胞肺癌患者的一线治疗后, 申办者进行了一项确证性亚组分析的 III 期临床试验确证该药在 TPS $\geq 1\%$ 患者中的疗效。

试验的目标人群为既往未经治疗的局部晚期或转移性非小细胞肺癌 (EGFR 和 ALK 基因均无突变) 且 PD-L1 表达水平 TPS $\geq 1\%$ 的患者; 试验目的是验证与铂类化疗相比帕博利珠单抗的优效性; 主要终点指标为总生存率 (OS); 预先定义三个确证性亚组: TPS $\geq 50\%$ 、 $20\% \leq \text{TPS} < 50\%$ 和 $1\% \leq \text{TPS} < 20\%$; 检验水准为单侧 0.025, 采用序贯检验法, 即首先检验 TPS $\geq 50\%$ 的亚组; 如果该亚组有统计学意义, 则再检验 $20\% \leq \text{TPS} < 50\%$ 亚组, 否则终止检验; 如果 $20\% \leq \text{TPS} < 50\%$ 亚组有统计学意义, 则再检验 $1\% \leq \text{TPS} < 20\%$ 亚组。随机分层变量包括地区、ECOG 评分、肿瘤组织类型和 PD-L1 TPS 比例分数, 按照 1:1 的比例将 1274 名患者随机分配至帕博利珠单抗或铂类化疗。TPS $\geq 50\%$ 亚组中, 299 例接受帕博利珠单抗, 300 例接受铂类化疗; $20\% \leq \text{TPS} < 50\%$ 亚组

中, 114 例接受帕博利珠单抗, 105 例接受铂类化疗; $1\% \leq \text{TPS} < 20\%$ 亚组中, 224 例接受帕博利珠单抗, 232 例接受铂类化疗。该研究的中位随访时间为 12.8 个月。

结果显示, $\text{TPS} \geq 50\%$ 亚组: 帕博利珠单抗与铂类化疗中位 OS (月) 分别为 20.0 和 12.2, $P=0.0003$, HR 及其 95%CI 为 0.69 (0.56-0.85), 达到统计学假设。 $20\% \leq \text{TPS} < 50\%$ 亚组: 帕博利珠单抗与铂类化疗中位 OS (月) 分别为 17.7 和 13.0, $P=0.0020$, HR 及其 95%CI 为 0.77 (0.64-0.92), 达到统计学假设。 $1\% \leq \text{TPS} < 20\%$ 亚组: 帕博利珠单抗与铂类化疗中位 OS (月) 分别为 16.7 vs 12.1, $P=0.0018$, HR 及其 95%CI 为 0.81 (0.71-0.93), 达到统计学假设。据此, 帕博利珠单抗的适应症从 $\text{TPS} \geq 50\%$ 扩展到 $\text{TPS} \geq 1\%$ 。

示例 2 获益-风险评估案例: 那他珠单抗治疗复发-缓解型多发性硬化症的获益-风险评估

那他珠单抗 (Tysabri) 是用于治疗多发性硬化症 (MS) 的单克隆抗体, 该药获批后不久就被迫退市, 其原因是两名患者在服用了该药后发生了进行性多灶性白质脑病 (PML), 这是一种罕见的不可治愈且易于致命的脑部感染疾病 (18 个月的随访期内, 约有 1/1000 的患者并发感染 PML)。经过对患者使用该药的获益-风险进行重新评估后, 认为对于其他药物治疗失败或不耐受

的复发-缓解型硬化症（RRMS）亚型患者，与现有的其他治疗方法相比，该药更有效，患者完全可以接受用药带来的风险。于是，该药重新获批仅用于在其他药物治疗失败或不耐受的 RRMS 亚型患者。