



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《药物临床试验富集策略与设计指导原则（试行）》的通告（2020年第60号）

发布日期：20201231

为了促进临床试验各相关方正确地认识富集策略与设计的原理与方法，并指导其规范应用以达到提高临床试验效率的目的，药审中心组织制定了《药物临床试验富集策略与设计指导原则（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：

《药物临床试验富集策略与设计指导原则（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

药物临床试验富集策略与设计指导原则

（试行）

2020 年 12 月

目 录

一、概述	1
二、富集策略与设计的适用性	2
三、常用的富集策略与设计	3
(一) 同质化富集	3
(二) 预后型富集	5
(三) 预测型富集	6
(四) 复合型富集	11
(五) 适应性富集	11
四、富集策略与设计的相关考虑	12
(一) 生物标志物检测的灵敏度和特异度	12
(二) 是否纳入标志物阴性的受试者	13
(三) 入选人群和分析集	14
(四) 筛选富集人群对优效和非劣效试验的不同影响	14
(五) 控制 I 类错误	15
五、监管考虑	16
(一) 准确界定富集人群	16
(二) 不应忽视非富集人群的疗效	16
(三) 预先确定研究方案并与监管机构沟通	17
六、参考文献	18
附录 1: 词汇表	25
附录 2: 中英文对照表	27
附录 3: 富集设计的研究案例	28

药物临床试验富集策略与设计指导原则（试行）

一、概述

临床试验的目的是在入组的受试者中验证试验药物的有效性和安全性。但实际上，由于受试者病理生理学特点和药物作用机理的复杂性，不同受试者的药物治疗效果不尽相同，从而影响临床试验的效率。为了入组能够从试验药物中获益最大化的受试者，以提高临床试验的效率，富集策略的概念应运而生。

富集是指在临床试验中根据受试者的某些特征（如人口学、病理生理学、组织学、基因组和蛋白质组学等）前瞻性地精准定义从试验药物中获益最大化的目标人群。临床试验中有多种选择受试者的富集策略，例如，可以选择因具有一定特征而对研究药物最有可能应答的受试者，也可以选择那些对现有药物治疗效果不明显而可能对试验药物敏感的受试者，或者单纯选择更容易出现终点事件的受试者等。

本指导原则阐述了常用的富集策略与设计的原理与方法、各自的优缺点，并从实际应用和监管角度说明需要考虑的关键问题。本指导原则中，“富集策略”主要是指随机对照临床试验中用于选择最有可能获益的受试者的方法，但也可以扩展到使用外部（历史或平行）对照的单臂试验。

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使用。

二、富集策略与设计的适用性

广义来讲，所有临床试验设计都含有富集的概念，这主要反映在受试者的入选标准和排除标准的部分条目上，其目的是尽可能入选对试验药物有应答的受试者，从而提高临床试验的效率。例如，研究降胆固醇药物降低心血管事件发生率时，临床试验只入组血液中总胆固醇浓度高于某一阈值的患者。实际上，根据疾病领域、药物作用机理以及受试者的应答情况等，可选择不同的富集策略和设计。是否使用以及如何选择富集策略和设计，要从科学上的有效性、试验结果的可解释性和医疗实践中的可推广性等方面考虑。

1. 科学上的有效性：包括筛选受试者有科学依据，筛选工具的灵敏度和特异度符合一定要求，试验设计时使用避免偏倚的措施（如随机、盲法等），以及 I 类错误的控制等。

2. 试验结果的可解释性：指试验药物在富集人群中的疗效可以从疾病的病理生理学、基因组学、遗传学或者药物作用机理等方面进行解释；如果限于生物学、医学或者药理学等方面的知识而无法解释，则试验药物在同样类似的富集人群中的疗效需要具有一定程度的重现性。

3. 医疗实践中的可推广性：包括富集策略能够在临床实践中被广泛地使用，以便及时、准确地识别对试验药物有应

答或敏感的患者。有时，由于筛选患者的方法复杂、灵敏度偏低、成本高昂等原因而使其无法普及，或者筛选方法耗时较长而在治疗开始时无法富集患者，这些都会影响富集策略和方法的可推广性。

三、常用的富集策略与设计

根据临床试验关注的主要问题和实施过程，可以选择不同的富集策略，主要包括同质化富集、预后型富集、预测型富集、复合型富集和适应性富集五种策略类型。

实际应用中，通常依据与药物作用机理相关的标志物选择富集策略与设计。这里的“标志物”定义为与受试者预后或药物治疗应答有关的临床特征，包括人口学、既往病史、家族史、临床观测变量（如疾病严重程度）、实验室检查（如病理生理学、药物代谢）、生物标志物（如基因组学和蛋白质组学）等各种特征变量。根据标志物的不同作用，可分为预后型、预测型和混合型标志物。另外，在有些疾病领域，可能没有明显的标志物，这时一般根据受试者在筛选期间对治疗的应答情况，或者其他临床试验的数据以及相关文献报道选择富集的受试者。

（一）同质化富集

同质化富集是指通过减少受试者间的异质性以提高临床试验的检验效能的一种研究策略。减少异质性最简单且实用的方法是尽量选择病情稳定的受试者，同时对入选受试者

进行精确定义，并对疾病的状态和有关变量进行精确测量。例如，在抗高血压药物试验中，为了筛选出血压相对稳定的受试者，可以在入组前对受试者的血压进行一段时间的观察，以排除血压变化较大的受试者。

一般来说，为了更加准确地定义富集人群，除常规的入选标准和排除标准外，还要考虑以下几方面问题：

（1）入选标准：精确定义入选标准，以确保入选受试者间的基线特征具有较好的一致性。

（2）排除标准：①对安慰剂过于敏感的受试者；②基线检测结果不稳定的受试者，如在初筛期病情或症状不稳定的受试者；③伴随某种疾病可能导致过早死亡的受试者；④服用与试验药具有相似治疗作用的药物的受试者；⑤可能无法耐受试验药物治疗的受试者；⑥可能因并发症提前退出研究的受试者。

（3）依从性：应尽可能入选依从性好的受试者，即选择不会因为非医学原因（如不便前往研究地点）而退出的受试者，以及能够坚持按照试验方案进行治疗的受试者，从而减少由于受试者过多退出或未能遵循方案规定的治疗而导致的差异。对患者的依从性识别和选择必须在随机化分组之前进行。

（4）培训：研究人员和临床试验协调员应接受相关培训，确保严格按照方案入选和排除受试者，并按照方案进行研究。

（二）预后型富集

预后型富集是指通过对预后型标志物的识别，入选更有可能观察到终点事件或疾病进展的高风险人群（特指更容易出现预后结局或疾病进展的人群），以增加检验效能的一种策略。该策略主要增加试验的绝对效应，而非相对效应。例如，在一项降低终点事件发生率的临床试验中，经过一段时间的治疗，高风险人群的终点事件发生率由 10% 降低至 5%，低风险人群的终点事件发生率由 1% 降低至 0.5%，虽然两者相对效应均降低 50%，但前者显然需要较少的样本量或者较短的随访时间就可以观察到试验药物的绝对疗效。常用的预后型富集设计有以下两种：

1. 基于终点事件的富集设计

在以降低终点事件发生率为主要评价指标的研究中，一般认为有效的试验药物在高风险人群中能够减少或者避免发生更多的终点事件。因此，应考虑在高风险人群中招募受试者。通常，在样本量不变时，与低风险人群相比，高风险人群更容易发生终点事件，经治疗后终点事件发生率降幅较大，因此检验效能更高。这种策略经常用在抗肿瘤和治疗心血管疾病的药物研究中，例如在基因 BRCA1/2 突变的女性人群中进行乳腺癌或卵巢癌预防的研究；又如在降血脂药物的研究中，选择血液中低密度脂蛋白偏高、高密度脂蛋白偏低和 C 反应蛋白偏高的患者试验。在有些疾病领域，如阿尔

茨海默症和各种癌症药物研究中，也可以通过基因组或蛋白质组学筛选高风险的患者。

2. 基于疾病进展的富集设计

预后型富集设计也可用于研究能够减缓疾病进展的试验药物，如在对阿尔茨海默症、帕金森病、类风湿性关节炎、慢性阻塞性肺疾病和恶性肿瘤等疾病开展药物临床试验时，可以选择疾病进展可能较快的受试者。例如，在类风湿性关节炎的患者中，具有以下特征的患者疾病往往进展较快，即类风湿因子阳性、具有某些临床特征（如多关节受影响、关节以外的病症、皮下结节、活动受限等）以及实验室指标异常（如血红蛋白降低）等；又如在慢性阻塞性肺病的患者中，具有近期发作史（过去一年中至少发作一次）或血浆纤维蛋白原较高的患者疾病进展较快。在抗肿瘤药物研究中，常见的预后标志物包括组织学分级、血管浸润、分子亚型以及转移性肿瘤结节等指标。

需要注意的是，如果预后型标志物与试验药物之间存在交互作用，即试验药物对该标志物阳性和阴性的患者具有不同的疗效，则该预后型标志物也可以起到预测的作用，此类标志物通常称为混合型标志物。

（三）预测型富集

预测型富集是指根据受试者的病理生理、应答史或与药物作用机制有关的疾病特征选择对试验药物最可能有应答

的受试者，以提高试验效率的一种研究策略。例如在肿瘤靶向治疗中，可根据药物相关的靶向基因或蛋白，选择可能有应答的受试者。采用这一策略既能增加试验药物的绝对效应，也能增加其相对效应，因此能够以较小的样本量获得较高的检验效能。当患有某种疾病的受试者中只有一小部分对试验药物有应答时（如只有部分受试者具有药物作用的受体），使用这种富集策略十分有效。在实践中，既可以基于研究者对疾病的认识（如各种标志物），也可以根据以往的试验数据和结果选择受试者。

1. 基于病理生理学特征的富集设计

疾病的病理生理学特征可以提示对试验药物有更好应答的受试者。基于病理生理学的富集指标可以是生物标志物（如影响肿瘤生长的基因突变、基因/蛋白表达水平）、影像学特征，以及与疾病表型相关的一些临床特征（如疾病分期、分型等）。根据标志物的性质，可将其分为以下几类：

（1）基因或蛋白标志物

对于治疗肿瘤的药物，由于通常针对肿瘤细胞表面或细胞内相关的受体、酶、激素或其他内源活性物质，因此可根据一个或多个相应的基因或蛋白标志物选择富集人群。例如曲妥珠单抗主要针对人表皮生长因子受体2蛋白阳性的乳腺癌患者。也有一些细胞受体最初作为蛋白标志物，但后来被确认为肿瘤基因标志物（如 EGFR 和 BRAF 基因突变），并

用该基因标志物定义病理生理状态、选择最有可能获益的受试者。

当在富集设计中使用基因或蛋白标志物时，标志物检测的准确性和精确性至关重要。如果检测不准确，不仅会影响富集的效果导致检验效能降低，而且在非劣效试验中可能会增加 I 类错误。

（2）药物代谢物

不同受试者对试验药物的代谢能力不同，入组能够产生足够数量活性代谢物的受试者，可以提高临床试验的效率。在某些情况下，给予活性代谢物产生能力较弱的患者更高的剂量，有助于产生足够的活性物质，从而更可能观察到药物的疗效。对于确定完全不能代谢出有效活性成分而无法从试验药物中获益的患者，不应纳入试验。

（3）肿瘤代谢物

抗肿瘤药物试验可以通过检测组织或血液中的肿瘤代谢物含量选择受试者。例如，只入选那些代谢反应较强的受试者，或者在肿瘤患者中按代谢反应程度进行分组，并对代谢反应较强的受试者进行主要分析。

2. 基于对试验药物应答证据的富集设计

此类富集设计根据受试者在筛选期内对试验药物（或既往对类似药物）的应答情况，选择可能合适的受试者。

（1）筛选有应答的受试者

对于在研究开始前无法根据标志物识别出可能对试验药物有应答的受试者的临床试验，需要设置合理的筛选期，将试验药物用于所有受试者。根据事先确定的主要终点或其替代终点（如症状、体征、实验室检查或疾病复发等）来筛选治疗有效的受试者；然后将筛选期内对试验药物有应答的受试者进行随机分组。选择有应答的受试者可以使用随机撤药设计，该设计一般分为两个阶段，第一阶段测试受试者是否对试验药物有应答（可以是单臂开放试验或随机对照试验），第二阶段对试验药物有应答的受试者，随机分配到试验组（继续使用试验药物）或安慰剂组（将试验药物撤出），无应答的受试者退出试验。例如，某降胆固醇药物的试验可以采用随机撤药设计，第一阶段将试验药物用于入组的高胆固醇患者，经过一段时间治疗后通过生化检查判断患者对药物是否有应答（即胆固醇降低）；第二阶段将对药物有应答的患者随机分配到试验组和对照组。

随机撤药设计由于筛选了对试验药物有应答的受试者，从而提高了临床试验的效率；同时，可利用已经入组的受试者研究药物的长期疗效或安全性，以及利用退出试验的受试者研究撤药效应。另一方面，这种设计更符合伦理学要求，即一旦治疗失败可及时终止试验，可用于儿童药物研究。采用这种先筛选、后随机化分组的方法可以首先入选较多的受试者，对这些受试者还可以按筛选期的应答程度进行分层，

将应答程度相对较好的亚组作为主要分析人群，同时可以进一步探索应答程度相对较弱的受试者。然而，这种设计并不适用于停药后药物作用持续时间相对较长（残留效应）或停药后会对受试者造成伤害的药物的研究，也不适用于对药物应答的终点衡量时间较长的研究。

（2）基于历史数据或文献报道确定入选受试者

根据既往研究中确定的亚组特征入选受试者，即在总体人群中很少或几乎没有观察到明显的治疗效果，但通过分析特定的亚组人群可能获得显著疗效，最后仅在亚组人群中招募受试者。例如，复方硝酸异山梨酯/盐酸肼屈嗪是一种治疗严重心力衰竭的药物，之前的研究发现其对非裔美国人的治疗效果明显优于高加索人，在后续设计的随机安慰剂对照试验中，入选了 1050 名心力衰竭的非裔美国人患者，由此验证了该复方药物在该心力衰竭患者亚组人群中的有效性。

3. 基于对现有药物无应答的富集设计

在富集设计中，除上述选择对试验药物有应答的受试者外，还可以考虑选择对现有对照药物无应答的受试者，目的是更好地显现出试验药物的治疗效果。

无应答的富集设计适用于满足一定条件的临床试验，即试验药物与现有对照药物具有不同的作用机制，或试验药物的疗效至少略优于现有对照药物。如果对受试者未加选择，则需要较大的样本量才能显示出试验药物的疗效；然而，如

果只选择对现有对照药物无应答的受试者，由于对照组的应答率很低，则可能只需要较小的样本量，就能得出试验组优于对照组的结论。必须指出的是，对于某些可能危及生命、并不断进展的疾病来说，采用无应答的富集设计会将受试者随机分配到无应答的对照组，可能存在伦理问题。

（四）复合型富集

复合型富集指同时使用多个标志物（如同时使用预后型和预测型标志物）以减少受试者异质性的富集策略。对有些疾病领域，疾病的发生、发展和预后机制复杂、个体异质性强或伴有混合疾病风险，使用单一标志物不大可能富集最有可能获益的受试者，而使用复合标志物（例如综合评分）进行富集可以有效地降低受试者的异质性，从而提高试验效率。

需要注意的是，使用复合标志物评分时应列出其构成的单个标志物并阐明它们之间的关系，或其与临床疾病特征的关联；如果对不同的单个标志物赋予不同的权重，应详细说明其生物学上的原理。

（五）适应性富集

适应性富集策略是指按照预先制定的计划，根据临床试验期中分析结果，在保证试验的合理性和完整性的前提下，对目标人群进行调整，如改变入组标准或仅纳入一个亚组的受试者等。

当试验药物在标志物阳性和阴性的受试者中的疗效不

确定时，试验可以同时入组标志物阳性和阴性的受试者，根据期中分析结果适应性地调整需要入选的受试者。当标志物阳性受试者的疗效比较为主要分析时，如果期中分析结果显示标志物阴性受试者的疗效远低于标志物阳性受试者，则应减少或完全停止标志物阴性的受试者入组。当标志物阳性受试者的疗效高于标志物阴性受试者的证据不够充分时，也可以考虑首先入组标志物阳性的受试者。如果期中分析结果表明试验药物在该标志物阳性的受试者中有疗效，则再考虑入组标志物阴性的受试者；否则，终止试验。

一般来说，如果标志物与疗效的关系越不确定，越需要包含标志物阴性的受试者，此时可以评估药物在全人群中使用时的获益与风险。当不确定一个标志物的预测性时，主要分析可以是全人群中的疗效比较；如果标志物阳性人群和全人群的疗效同时作为主要分析时，需要按照一定的规则将检验水准 α 进行分配。无论何种情况，都应事先在方案中明确规定检验假设，并需要对 I 类错误进行控制。

四、富集策略与设计的相关考虑

（一）生物标志物检测的灵敏度和特异度

当采用筛检试验选择受试者时，必须考虑检测方法的可靠性，以便能更准确地选择高风险或者对试验药物有应答的受试者。理想情况下，用于筛选受试者的检测方法，应该对选择高风险或对试验药物有应答的受试者有较高的灵敏度，

同时对鉴别低风险或对试验药物无应答的受试者有较高的特异度。

当利用生物标志物筛选入组受试者时，如果不能准确给出预测标志物的阈值，可以通过受试者诊断特征曲线分析，即对预测标志物不同阈值点的灵敏度和特异度进行分析，并用受试者诊断特征曲线下面积衡量其筛检效果。一般可在早期研究阶段通过适当样本量的试验对预测标志物阈值进行探索，为确证性试验提供准确阈值。

（二）是否纳入标志物阴性的受试者

富集设计既可以只纳入标志物阳性的受试者，也可以同时纳入阳性和阴性的受试者。然而，富集设计的关键问题是纳入标志物阴性受试者的比例。一般来说，可以考虑如下的富集设计：

1. 只纳入标志物阳性的受试者

如果作用机制或已有数据表明，试验药物在标志物阳性的受试者中有明显的疗效，而在标志物阴性的受试者中疗效较小甚至没有疗效，则应考虑不纳入标志物阴性的受试者。

2. 同时纳入标志物阳性和阴性的受试者

如果作用机制或已有数据表明，标志物阳性的受试者较阴性受试者的疗效更好，在已知试验药物毒性相对较小的情况下，可同时纳入标志物阳性和阴性的受试者。这种策略的优点是，能在非富集人群中提供合理的获益-风险估计。

如果在试验开始之前就能够明确标志物，则可以通过其分组实施分层随机化，主要分析可限制在标志物阳性的受试者中。实际中，也可以在全人群中进行主要分析，或者在全人群和标志物阳性受试者中同时进行主要分析，并适当地控制 I 类错误。

一般来说，如果标志物的阈值或标志物阴性的受试者应答程度不确定，则有必要纳入标志物阴性的受试者。

（三）入选人群和分析集

使用富集策略的主要问题是研究结果的适用性和外推性，即采用富集设计时，要重点考虑这种富集策略是否能够在医学实践中用于识别对研究药物应答的人群，以及该药在更广泛的患者人群中是否也有类似的疗效。因此，对不符合富集入选标准的患者人群进行研究同样重要。需要注意的是，试验确定的入选受试者和主要分析集可以不同（后者可以是前者的子集），但这些必须在研究方案中明确定义。在基因或其它检测结果不能立即获得，而患者需要及时接受治疗时，选择以全人群入组，以提供更多的安全性信息，但主要疗效分析可以是其中的一个子集。

（四）筛选富集人群对优效和非劣效试验的不同影响

使用标志物选择受试者时，对优效和非劣效试验有着不同的影响。对于采用预后型富集策略的优效试验，如果筛检方法的灵敏度不高，则需要招募更多的受试者从中进行筛选，

才能获得规定的富集样本量；如果特异度不高，则需要增加富集样本量或延长试验时间才能获得足够的终点事件数。对于采用预测型富集策略的优效试验，如果筛检方法的灵敏度不高，则会导致符合入组条件的受试者不足；如果特异度不高，则会纳入较多的不符合入组条件的受试者。但无论采用预后型或者预测型富集策略，都不会增加优效试验的 I 类错误。

然而，对于非劣效试验，筛检的准确度不仅会影响研究所需的样本量或持续时间，还可能增加 I 类错误。例如，采用预后型富集策略进行非劣效试验，如果阳性对照的筛选方法与以往研究不同，则可能导致阳性对照组的疗效低于以往研究的疗效，从而增加 I 类错误。另外，对于基于预测型富集策略的非劣效试验，对 I 类错误的影响更为复杂，它取决于标志物是与试验药物和阳性对照药物的治疗都相关，还是仅与其中一种治疗相关。因此，非劣效试验中选择受试者的检测方法最好与阳性对照以往研究的筛检方法一致，或者两种筛检方法有相似的灵敏度和特异度。

（五）控制 I 类错误

对于同时入选富集人群和非富集人群的富集设计，可根据筛检方法的特性和受试者对治疗的应答情况，选择不同的假设检验策略。如果有多个假设检验，如在标志物阳性人群和全人群中分别进行假设检验，则需要考虑多重性调整的问题。

题；如果仅有一个假设检验，如在标志物阳性人群中进行假设检验，则无需考虑这一问题。在不同假设下 I 类错误 α 的分配，可以根据标志物阳性人群对药物的应答程度、阳性人群在全人群中的比例以及按照预先设定的检验效能所需要的样本量进行设置。对全人群和富集人群进行假设检验时，可以采取平行策略或者序贯策略进行假设检验。

五、监管考虑

（一）准确界定富集人群

临床试验是否使用、何时使用以及使用何种富集策略，主要取决于能否准确界定富集人群，因为这会对产品说明书的撰写和后续的医疗实践产生影响。说明书应该严格按照富集人群的界定条件定义治疗有效的患者人群，不能随意外推。如果使用富集策略和设计无法准确界定富集人群，则可能导致在说明书中无法准确地定义治疗有效的患者人群，进而无法准确指导临床合理用药。

（二）不应忽视非富集人群的疗效

试验药物在富集人群中的有效性和安全性得到确证后，其在非富集人群中相应的信息也应受到重视。通过非富集人群中开展进一步研究，可以更全面地描述药物的获益-风险状况，为药物能够在更广泛的患者人群中应用提供依据。

对基于高风险人群进行预后型富集分析获批上市的药物，在随后的低风险人群中试验，可能会使用不同的结局指

标，如在高风险人群中使用病死率，而在低风险人群中可以使用一个有临床意义的复合结局指标，从而有助于提高试验效率。

（三）预先确定研究方案并与监管机构沟通

一般来说，在研究开始之前，应预先计划并确定受试者的选择方案。如果已知可用于富集的特征变量或标志物，可在筛选受试者时对其进行测量。而当特征变量或标志物在研究人群中富集的效果或分布不确定时，可以考虑进行适应性富集，即在试验过程中依据积累的数据，对试验设计的要素进行调整。无论采用何种策略和设计，调整方法和过程应在研究方案中事先说明，确保其合理性和正确性，并与监管机构进行充分的沟通。

六、参考文献

- [1] Amur S, LaVange L, Zineh I, et al. Biomarker Qualification: Toward a Multiple Stakeholder Framework for Biomarker Development, Regulatory Acceptance, and Utilization. *Clinical Pharmacology and Therapeutics*, 2015; 98 (1): 34-46.
- [2] Barker AD, Sigman CC, Kelloff GJ, et al. I-SPY 2: an adaptive breast cancer trial design in the setting of neoadjuvant chemotherapy. *Clinical pharmacology and therapeutics*, 2009; 86 (1): 97-100.
- [3] Bibbins-Domingo K, Fernandez A. BiDil for heart failure in black patients: implications of the U.S. Food and Drug Administration approval. *Annals of Internal Medicine*, 2007; 146 (1): 57-62.
- [4] Chapman PB, Hauschild A, Robert C, et al. Improved survival with vemurafenib in melanoma with BRAF V600E mutation. *The New England Journal of Medicine*, 2011; 364 (26): 2507-2516.
- [5] D'Agostino RB Sr. The Delayed-Start Study Design. *The New England Journal of Medicine*, 2009; 361 (13): 1304-1306.
- [6] D'Amico AV, Chen MH, Roehl KA, et al. Preoperative PSA velocity and the risk of death from prostate cancer after radical prostatectomy. *The New England Journal of Medicine*, 2004;

351 (2): 125-135.

- [7]Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. *Lancet*, 2005; 365 (9472): 1687-1717.
- [8]European Medicines Agency (EMA). Points to consider on multiplicity issues in clinical trials.
- [9]European Medicines Agency (EMA). Guideline on the investigation of subgroups in confirmatory clinical trials.
- [10]European Medicines Agency (EMA). Qualification Opinion of Alzheimer's Disease Novel Methodologies/biomarkers for BMS-708163. 2011.
- [11]European Medicines Agency (EMA). Qualification Opinion of Low Hippocampal Volume (Atrophy) By MRI for Use in Regulatory Clinical Trials - in Pre-Dementia Stage of Alzheimer's Disease. 2011.
- [12]Fan C, Oh DS, Wessels L, et al. Concordance among Gene-Expression-Based Predictors for Breast Cancer. *The New England Journal of Medicine*, 2006; 355 (6): 560-569.
- [13]Freidlin B, Korn EL. Biomarker enrichment strategies: matching trial design to biomarker credentials. *Nature*

- Reviews Clinical Oncology, 2014; 11 (2): 81-90.
- [14]Havel JJ, Chowell D, Chan TA. The evolving landscape of biomarkers for checkpoint inhibitor immunotherapy. Nature Reviews Cancer, 2019; 19 (3): 133-150.
- [15]Hughes RA, Donofrio P, Bril V, et al. Intravenous immune globulin (10% caprylate-chromatography purified) for the treatment of chronic inflammatory demyelinating polyradiculoneuropathy (ICE study): a randomised placebo-controlled trial. The Lancet Neurology, 2008; 7 (2): 136-144.
- [16]Institute of Medicine (US) Committee on Strategies for Small-Number-Participant Clinical Research Trials. Small Clinical Trials: Issues and Challenges. Washington, DC. The National Academies Press, 2001.
- [17]International Conference on Harmonization (ICH). E5 guideline “Ethnic Factors in the Acceptability of foreign Clinical Data E5 (R1)”.
- [18]International Conference on Harmonization (ICH). E5 guideline “Implementation Working Group Questions & Answers (R1)”.
- [19]Jiang W, Freidlin B, Simon R. Biomarker-adaptive threshold design: a procedure for evaluating treatment with possible biomarker-defined subset effect. Journal of the National

Cancer Institute, 2007; 99 (13): 1036-1043.

- [20] Kowanetz M, Zou W, Gettinger SN, et al. Differential regulation of PD-L1 expression by immune and tumor cells in NSCLC and the response to treatment with atezolizumab (anti-PD-L1). *Proceedings of the National Academy of Sciences of the United States of America*, 2018; 115 (43): E10119-E10126.
- [21] Liu A, Liu C, Li Q, et al. A threshold sample-enrichment approach in a clinical trial with heterogeneous subpopulations. *Clinical trials*, 2010; 7(5): 537-545.
- [22] Loo E, Khalili P, Beuhler K, et al. BRAF V600E Mutation Across Multiple Tumor Types: Correlation Between DNA-based Sequencing and Mutation-specific Immunohistochemistry. *Applied immunohistochemistry & molecular morphology*, 2018; 26 (10): 709-713.
- [23] Parkinson Study Group. A controlled, randomized, delayed-start study of rasagiline in early Parkinson disease. *Archives of neurology*, 2004; 61 (4): 561-566.
- [24] Priscilla Velengtas, Penny Mohr, Messner DA. Making informed decisions: Assessing the strengths and weaknesses of study designs and analytic methods for comparative effectiveness research. National Pharmaceutical Council ed.

Washington, DC. 2012.

- [25] Ridker PM, Danielson E, Fonseca FA, et al. Rosuvastatin to prevent vascular events in men and women with elevated C-reactive protein. *The New England Journal of Medicine*, 2008; 359 (21): 2195-2207.
- [26] Scandinavian Simvastatin Survival Study Group. Randomised trial of cholesterol lowering in 4444 patients with coronary heart disease: The Scandinavian Simvastatin Survival Study (4S). *Lancet*, 1994; 344 (8934): 1383-1389.
- [27] Schrock AB, Ouyang C, Sandhu J, et al. Tumor mutational burden is predictive of response to immune checkpoint inhibitors in MSI-high metastatic colorectal cancer. *Annals of oncology*, 2019; 30 (7): 1096-1103.
- [28] Singh BN. Comparative efficacy and safety of bepridil and diltiazem in chronic stable angina pectoris refractory to diltiazem. The Bepridil Collaborative Study Group. *The American journal of cardiology*, 1991; 68 (4): 306-312.
- [29] Simon R. Clinical trial designs for evaluating the medical utility of prognostic and predictive biomarkers in oncology. *Personalized medicine*, 2010; 7 (1): 33-47.
- [30] Taylor AL, Ziesche S, Yancy C, et al. Combination of isosorbide dinitrate and hydralazine in blacks with heart

failure. The New England Journal of Medicine, 2004; 351 (20): 2049-2057.

[31] Temple R. Enrichment of clinical study populations. Clinical Pharmacology & Therapeutics, 2010; 88 (6): 774-778.

[32] Temple RJ. Special study designs: early escape, enrichment, studies in non-responders. Communications in Statistics - Theory and Methods, 1994; 23 (2): 499-531.

[33] U.S. Food and Drug Administration (FDA). Enrichment Strategies for Clinical Trials to Support Determination of Effectiveness of Human Drugs and Biological Products – Guidance for Industry.

[34] U.S. Food and Drug Administration (FDA). Qualification of Biomarker–Plasma Fibrinogen in Studies Examining Exacerbations and/or All-Cause Mortality in Patients with Chronic Obstructive Pulmonary Disease. 2016.

[35] U.S. Food and Drug Administration (FDA). Qualification of Biomarker–Total Kidney Volume in Studies for Treatment of Autosomal Dominant Polycystic Kidney Disease. 2016.

[36] Wang S, Hung HMJ, O'Neill RT. Genomic Classifier for Patient Enrichment: Misclassification and Type I Error Issues in Pharmacogenomics Noninferiority Trial. Statistics in Biopharmaceutical Research, 2011; 3 (2): 310-319.

- [37] Wang SJ, Hung HM, O'Neill RT. Adaptive patient enrichment designs in therapeutic trials. *Biometrical Journal*, 2009; 51 (2): 358-374.
- [38] Wang SJ, O'Neill RT, Hung HM. Approaches to evaluation of treatment effect in randomized clinical trials with genomic subset. *Pharm Stat*, 2007; 6 (3): 227-244.

附录 1：词汇表

灵敏度 (Sensitivity)：评价诊断试验和筛检试验准确度的基本指标之一。在药物临床试验的富集研究中，灵敏度表示对于发生终点事件具有高风险或对药物有应答的受试者，能够将其正确识别出的概率。

适应性富集设计 (Adaptive Enrichment Design)：按照预先制定的计划，根据临床试验数据的期中分析结果，在保证试验合理性和完整性的前提下，允许在试验过程中适应性地更新入排标准，选择最有可能从治疗中获益的受试者入组的适应性设计。

随机撤药设计 (Randomized Withdrawal Design)：在此类设计中，所有受试者都在最初的开放标签期间接受试验药物的治疗，之后对药物无应答的受试者退出试验，有应答的受试者（富集亚群）在试验第二阶段中随机接受试验药物或安慰剂。

特异度 (Specificity)：评价诊断试验和筛检试验准确度的基本指标之一。在药物临床试验的富集研究中，特异度表示对于发生终点事件具有低风险或对药物无应答的受试者，能够将其正确识别出的概率。

异质性 (Heterogeneity)：在临床试验中，异质性体现在个体和群体两个水平，前者通常是指受试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受试者对治疗有

不同的应答；后者通常是指不同中心、种族、地域等受试者具有不同的特征，有可能导致不同受试者对治疗有不同的应答。

预测型富集 (Predictive Enrichment)：是指选择性地纳入可能会对治疗有应答的受试者的一种研究策略或设计，这些受试者具有共同的有预测意义的生物学和组织病理学特征，可以更敏感地显示试验药物的疗效。

预后型富集 (Prognostic Enrichment)：是指选择性地纳入更可能发生终点事件（如死亡或疾病恶化）的受试者，从而降低为达到统计学显著疗效所需样本量的一种研究策略或设计。

附录 2：中英文对照表

中文	英文
低风险人群	Low-risk Population
多重性	Multiplicity
复合结局指标	Composite Endpoint
富集策略	Enrichment Strategy
富集人群	Enriched Population
高风险人群	High-risk Population
复合型富集策略	Mixed Enrichment Strategy
获益-风险比	Benefit-risk Ratio
可推广性	Generalizability
灵敏度	Sensitivity
目标人群	Target Population
筛检试验	Screening Test
适应性富集策略	Adaptive Enrichment Strategy
受试者诊断特征	Receiver Operating Characteristic, ROC
随机撤药	Randomized Withdrawal
特异度	Specificity
同质化富集策略	Reducing Heterogeneity Strategy
异质性	Heterogeneity
预测型富集策略	Predictive Enrichment Strategy
预后型富集策略	Prognostic Enrichment Strategy

附录 3：富集设计的研究案例

示例 1：预后型富集——心血管病研究

在心血管病研究中，选择高风险人群（如 AMI、中风、极高的胆固醇水平、非常严重的 CHF 或接受血管成形术等）进行临床研究，能够获得更多的结局事件。斯堪的纳维亚辛伐他汀生存研究（4S）是一项研究降脂药物的试验，主要目的是评估辛伐他汀能否通过降低血清胆固醇改善冠心病患者的生存率。该研究为随机双盲安慰剂对照的多中心临床试验，招募了 4444 名患有心绞痛或先前有心肌梗死（MI）的患者，这些受试者都具有较高的总胆固醇（TC）水平。在平均 5.4 年的随访期内，辛伐他汀相对于安慰剂的心血管病死亡率能够得到显著降低（相对危险度 RR 为 0.70，95% CI: 0.58-0.85）。

示例 2：预测型富集

黑色素瘤研究：BRAF 激酶抑制剂是一种治疗黑色素瘤的靶向药物，BRAF 基因第 15 外显子（V600E）可以作为预测型生物标志物。已知 BRAF 基因编码细胞质丝氨酸/苏氨酸激酶，该酶调节丝裂原活化蛋白激酶信号转导通路，控制包括细胞生长和分裂（增殖）在内的几个重要细胞功能。现已发现，BRAF V600E 在多种肿瘤中发生突变，如黑色素瘤、大肠癌、甲状腺乳头状癌、毛细胞白血病（Hairy Cell Leukemia）和朗格汉斯细胞增生症等。一项黑色素瘤 III 期临床试验研

究，入选了 675 例转移性或不可切除、BRAFV600E 突变的受试者，分别给予 BRAF 激酶抑制剂维罗非尼或化疗药物达卡巴嗪，结果发现：给予维罗非尼靶向药治疗的受试者其应答率为 48%，而使用达卡巴嗪化疗的受试者其应答率仅为 5%；两者相比较，经维罗非尼治疗的受试者相对死亡风险降低了 63%。

MSI 研究：微卫星不稳定性（MSI）是对免疫检查点抑制剂产生应答的生物标志物。PD-1/PD-L1 通路是调节 T 细胞活化的信号通路，在肿瘤发生和进展中起重要作用。实际中，通常用免疫组化方法检测 PD-L1 蛋白的表达水平，将其作为预测标志物并选择高表达的受试者，但其对 PD-1/PD-L1 抑制剂的应答率仅为 10-20%。然而，在高度微卫星不稳定性（MSI-HIGH）的肿瘤受试者中的应答率却能达到 50%。基于此，帕博利珠单抗通过临床试验确证了在 MSI-HIGH 型或存在错配修复缺陷的结直肠癌和子宫内膜癌患者人群中的疗效。

示例 3：随机撤药设计——普瑞巴林用于治疗纤维性肌痛症的研究

一项研究普瑞巴林（Pregabalin）对治疗纤维性肌痛症受试者的疗效的临床试验使用了两阶段随机撤药设计，其目的是比较普瑞巴林相对于安慰剂在接受治疗到疗效丧失的时间上的差异（Time to Loss of Therapeutic Response, TLTR）。

第一阶段是开放性试验，患有纤维性肌痛症受试者全部接受普瑞巴林治疗，并观察 6 周。其中，在 1-3 周，受试者接受递增剂量的普瑞巴林以决定其最佳剂量；在 4-6 周，受试者维持在此最佳剂量。第一阶段开放性治疗结束后，受试者必须具有至少 50% 以上的疼痛减轻、且在 PGIC 量表上的自我评价至少是“显著改善”才能进入第二阶段的双盲、安慰剂对照的试验。在入组的 1051 名受试者中，经第一阶段的治疗，符合以上条件进入第二阶段的有 566 人，被随机分配到安慰剂组或普瑞巴林组。经过第二阶段 26 周的治疗，两组间在治疗应答丧失时间（Time to Loss of Therapeutic Response, LTR）上有显著差异（ $P < 0.0001$ ）。试验结束时，安慰剂组有 61%（178）、普瑞巴林组有 32%（90）达到了 LTR。