

药物临床试验适应性设计指导原则 (试行)

2021 年 1 月

目 录

一、 概述.....	1
二、 适应性设计中需要考虑的因素	2
(一) 适用性	3
(二) 合理性	4
(三) 完整性	5
(四) 可行性	6
三、 常用的适应性设计	7
(一) 成组序贯设计	7
(二) 样本量重新估计	8
(三) 适应性无缝剂量选择的设计	10
(四) 适应性富集设计	12
(五) 两阶段适应性设计	13
(六) 适应性主方案试验设计	14
(七) 多重适应性设计	16
四、 其他考虑	16
(一) 仅基于外部数据的修改	16
(二) 监管的其他考虑	17
五、 参考文献	19
附录：词汇表	25

药物临床试验适应性设计指导原则

一、概述

确证性临床试验的设计一般基于前期探索性研究结果，很多时候仅依赖于非常有限的的数据，由此可能造成设计元素存在较大的偏差，从而直接影响试验的成败。随着药物研发的推动，临床研究的技术方法得到不断的发展，适应性设计也受到越来越多的研究与应用。适应性设计允许根据试验期间累积的数据对试验设计进行修改，以修正初始设计的偏差，从而增加试验的成功率，提高试验的效率。

成组序贯设计是最早应用于临床试验的适应性设计，其后，适应性设计较广泛地用于样本量的重新估计，现今逐步推广和发展到了多种类型的试验设计，例如两阶段设计、平台试验设计等更为复杂的设计。随着理论方法的不断成熟完善、模拟计算能力的进步，以及实践经验的积累，适应性设计在临床试验中得到越来越多的应用。

本指导原则对适应性设计的定义为：按照预先设定的计划，在期中分析时使用试验期间累积的数据对试验做出相应修改的临床试验设计。一方面，适应性修改是“按预先设定的计划”进行的，而不是临时提出的修改方案；另一方面，适应性修改是一个自我学习的过程，即通过对累积数据的不断学习，相应地修改试验方案，以适应不断变化的研究环境。

因此，适应性设计旨在更好地改进进行中的临床试验，而不是因设计本身缺陷而有极大可能导致临床试验失败所做的临时补救。

在实际当中有时会基于充分合理的外部数据对一个进行中的临床试验做出修改，如果这种修改仅仅基于外部数据，依据本指导原则的定义不将其归于适应性设计的范围。本指导原则着重于讨论适应性设计的基本概念和原则、常用的适应性设计类型、使用适应性设计时的考虑要点以及监管要求等，目的是指导和规范申办者如何采用以及实施适应性设计。申办者在设计适应性临床试验方案时，应同时参考其它相关的 ICH 指导原则和国内指导原则。本指导原则主要适用于药品的确证性临床试验，对于探索性研究也具有参考意义。本指导原则仅代表当前的观点和认识，随着研究和认识的深入将予以修订完善。

二、 适应性设计中需要考虑的因素

在决定是否采用适应性设计之前，应全面深入地权衡适应性设计和传统设计之间的优劣，尤其是适应性设计在设计、实施和统计分析方面的复杂性，以及由此而带来的在试验实施中可能会引入的、不可避免的操作偏倚以及其他各种挑战。采用适应性设计需综合考虑诸多因素，特别是适应性设计的适用性（fitness for purpose）、合理性（validity）、完整性（integrity）和可行性（feasibility）。

（一）适用性

适应性设计的适用性是指计划开展的试验是否适合采用适应性设计。一般而言，确证性临床试验需要良好合适的试验设计，包括试验目标、受试人群、入组分配、主要终点、分析方法等多个方面，其中每一个环节的偏差都可能导致试验的失败。适应性设计虽然可以实现自我学习，重新评估当前试验的计划，并可以调整设计时的偏差以寻求更好的方法来实现同一目标，但它并不是用于解决试验开始时设计上的错误。

是否采用适应性设计，首先应该考虑需要什么样的适应性修改、什么样的数据、验证什么样的假设、什么样的决策方法、什么样的条件使其能在实际中实施等。如果一个适应性设计并不能带来预期试验效率的增加、试验质量的提高，或者实施起来有极大的困难，则并不适合采用该设计。此外，适应性设计需要在设计阶段投入大量的时间进行深入地研究和仔细地计划。

大多数适应性设计方法是为满足临床试验的特殊需求而产生，可能不具备统计理论上的某些最优性，但它可能是解决临床试验某个特殊问题最为合适的方法，因此在考虑采用适应性设计时应主要基于需要解决的特殊问题。另外，适应性设计临床试验在操作和实施中较传统试验更为复杂和

困难，因此，试验设计方法的简易性有时也成为是否采用适应性设计的一个重要的考虑因素。

（二）合理性

适应性设计的合理性是指试验的总 I 类错误率能否得到控制，以及能否确保试验结果的可信度、可解释性和说服力。

判断适应性设计是否合理，最重要的标准是所使用的统计方法能否控制总 I 类错误率。适应性修改一般需要考虑统计检验的多重性问题，并需将试验的 I 类错误率控制在预先设定的水平。此外，对有些适应性设计来说，如果采用双侧检验，由于适应性修改前后阶段的 p 值不能反映组间比较的方向，有可能使得最终的整体 p 值的意义难以解释，为避免这种情况，可在试验方案中选择单侧检验；但对另一些适应性设计，例如不对称的双边假设，双侧检验会是更合适的选择。

保持试验的合理性还意味着应该有正确的统计推断方法，比如用于计算调整后的 p 值、估计效应量及置信区间，以及衡量不同阶段治疗效果的一致性。

适应性设计可能同时涉及多个目标人群、多个假设、多个终点或多重检验，故对统计分析方法的合理性有着很高的要求。如果对适应性修改没有相应合理的统计方法，则不宜采用该设计。由于适应性设计的复杂性，在某些情况下因没有适用的统计推断的理论公式或解析公式，需要基于模拟方

法验证统计方法的合理性，这在一定程度上增加了额外的不确定性。

如果适应性设计需要合并调整前后的数据，那么需要考虑数据合并的合理性（包括前后数据的差异以及合并方法等）以及合并后疗效估计的可解释性。如果适应性修改的最终统计检验结果虽为阳性但临床获益太小，也不足以支持所验证的药物疗效。

（三）完整性

适应性设计的完整性是指是否能够控制住试验操作所引入的偏倚。保持试验的完整性意味着需要按照预先设定的计划对方案进行调整和保持期中分析结果的盲态，以求最大限度地减少操作偏倚。

避免引入操作偏倚是所有临床试验的最基本要求。适应性设计由于涉及临床试验许多方面的修改，有可能影响后续试验的执行，对保持试验的完整性增加了额外的难度。因此在确证性试验中，适应性设计的期中分析一般应该由独立的数据监查委员会（Data Monitoring Committee, DMC）及其申办者以外的独立统计支持团队完成，并保证期中分析的结果不被申办者、研究者和受试者所知悉，以免影响后续试验的执行和引入操作偏倚。因适应性修改涉及多个环节，设立一个有效的防火墙以防止期中分析结果外泄而造成可能的操作偏倚是执行中最为重要的任务。为此，适应性设计的方案

应包含一个完善的操作流程，特别是关于如何设置相关信息的访问权限。同时，为避免不可控制因素对试验结果的影响，还要考虑怎样避免根据试验所做的修改而被间接地推出期中分析的结果。申办者应准备好试验所有需要的标准操作流程，并将涉及适应性修改的相关流程纳入其中，同时记录好实际操作的过程。以上这些都应该在试验的设计阶段仔细考虑，并需要在试验进行中严格地执行，以免影响试验的完整性。

（四）可行性

适应性设计的可行性是指试验的适应性修改能否在实际中实施。由于适应性设计比传统设计更为复杂，并且实施和分析更加困难，在计划采用适应性设计之前，可能需要考虑以下因素：适应性调整策略能够保障试验的合理性和完整性；相对于试验周期，有充裕的时间根据试验累积数据的分析结果进行适应性修改和开展后续试验的操作；期中数据收集和数据清理可以快速完成，以便按预定计划完成期中分析和调整；具备能够快速修改随机化程序/药物供应系统；具备足够的药物供应管理的能力以及能够负担增加的药物供应；提前准备好适应性设计的数据采集系统；保证与各相关方的沟通顺畅有效；能够配备专业软件来完成复杂设计和相关分析的计算等。同时，在试验设计阶段，申办者也可以与研究者的沟通，评估所考虑的适应性设计在实际中能够顺利进行的

可行性。如果相关适应性修改难以实施，则应该考虑其它设计。

综上所述，若计划采用适应性设计，需要仔细地评估其是否确有优势。若无法决策，可以采用模拟方法以评估适应性设计的效率。如果评估后适应性设计没有体现出太多的优势，建议谨慎考虑适应性设计。

三、常用的适应性设计

（一）成组序贯设计

成组序贯设计是指方案中预先计划在试验过程中进行一次或多次期中分析，依据每一次期中分析的结果做出后续试验的决策，决策通常有四种可能：①依据优效性终止试验；②依据无效性终止试验；③依据安全性终止试验；④继续试验。期中分析的时间一般基于累积数据的占比，如受试者入组比例或发生目标事件数的比例，或日历时间。如果期中分析至少有一次优效性分析，且有提前终止试验的可能，则应调整分析的 I 类错误率以将总 I 类错误率控制在事先设定的水平。调整 I 类错误率的常用方法包括 Pocock 方法、O'Brien & Fleming 方法和 Lan & DeMets 方法等。由于期中分析仅使用了部分数据，结果仍有较大的不确定性，评估早期优效性时一般建议使用较为保守的方法以便终止试验时增加优效结论的可靠性。无效性边界的设定分为绑定和非绑定。绑定边界在期中分析结果一旦跨越无效性边界时必须终止试验。

非绑定边界在期中分析结果跨越无效性边界时，一般会终止试验，但在有些情形下独立数据监查委员会基于全面评估后仍然可以建议试验继续进行。对于非绑定边界，无需调整最终分析的 I 类错误率。

选择期中分析的时间点也要仔细考虑。如果成组序贯调整计划中存在以优效性提前终止试验的可能，时间点的选择应该考虑期中分析时的数据量是否充分以及随访时间是否足够以便能够提供可靠的疗效估计和安全性评价的结果，也包括重要的次要终点以及一些重要的亚组结果的估计。若期中分析是要验证药物的安全性和无效性，时间点则应该侧重于如何最大程度地保护受试者。

（二）样本量重新估计

样本量重新估计是指依据预先设定的期中分析计划，利用累积的试验数据重新计算样本量，以保证最终的统计检验能达到预先设定的目标或修改后的目标，并同时能够控制 I 类错误率。

初始样本量的估计通常取决于效应量、主要终点的变异性、试验随访时间、受试者脱落率等诸多因素，而这些常常基于以往的研究数据。多数情况下，试验设计阶段样本量的估计所需要的参数信息往往不够充分，可能会导致样本量估算的不够准确。适应性设计中的样本量重新估计为此类问题提供了有效的解决方案。

样本量重新估计的方法可以分为盲态方法和非盲态方法。

盲态方法，也称为非比较分析方法（**non-comparative analysis**），是指期中分析时不使用实际试验分组的信息，或者未做任何涉及组间比较的分析。

盲态方法的样本量重新估计是指根据累积的数据，计算样本量的重要参数（如合并方差或合并事件发生率），然后对样本量进行重新估计。因期中分析时不涉及组间的疗效比较，故一般不需要调整 I 类错误率。该方法比较容易实施，一般不会引入操作偏倚，而且相关的统计方法也较为完善，只需要在试验设计的阶段预先做好规划。盲态方法的样本量重新估计也可由申办者完成。

非盲态方法，也称比较分析方法（**comparative analysis**），是指期中分析时使用了试验分组信息（包括各组的真实名称或可区分的分组代码）的分析，分析内容涉及组间的比较。

非盲态方法的样本量重新估计是指根据累积数据以及分组信息，计算样本量的重要参数（如试验效应量），然后对样本量进行重新估计，因期中分析涉及组间的疗效比较，通常需要对 I 类错误率进行相应调整。

非盲态分析的样本量重新估计需要预先在研究方案中阐明，包括重新估计的时点、决策时使用的标准、重新估计时使用的方法、调整检验水准 α 的方法、执行非盲态分析的

人员，以及执行整个操作过程的人员等。应该注意的是，一个试验中不宜做过多次数的样本量重新估计。当重新估计的样本量少于最初设计的样本量时，通常不接受样本量减少的调整。

是否采用非盲态样本量重新估计需要考虑多种因素。例如，若有比较可靠的前期数据，非盲态下样本量重新估计是否必要；采用非盲态样本量重估所付的代价（如检验水准调整）与初始设计时略微放大样本量相比，是否有利；期中分析能否很快完成，是否可能因为入组较快而导致没有充足时间用来调整试验；期中分析的时间节点和推断方法是否合理；现有数据能否支持进行计划内的期中分析等。因此，应根据试验本身的特点，仔细考虑各种因素，然后做出合适的决策。

（三）适应性无缝剂量选择的设计

适应性无缝剂量选择的设计是指将两个试验无缝连接，在前期试验结束时做剂量选择，并将所选剂量用于后期试验。最终分析时则同时包含前期和后期两个试验入组的所有受试者的数据。本指导原则以 II/III 期试验为例对适应性无缝剂量选择的设计予以阐述，其他无缝设计情形可以此为参考。

在传统的设计中，独立的 II 期剂量选择通常包括多个剂量组，目的是选出合适的剂量并用于 III 期试验。III 期试验是一个独立于 II 期的试验，其最终分析并不包含 II 期试验的数据。以此为特定目标的 II/III 期试验也常称为 II/III 期操作

无缝设计。操作无缝设计将 II 期试验的受试者排除在 III 期的最终分析之外，且不需要在 III 期的最终分析时对 I 类错误率进行调整。另一种被称为 II/III 期推断无缝设计，是指在最终分析时包含了选中剂量和未选中剂量的 II 期试验的所有受试者。适应性 II/III 期推断无缝剂量选择的设计是推断无缝设计的特例。这种设计具有很多优点，例如可以缩短通常由 II 期试验结束时到 III 期试验开始时的时间间隔、减少试验的总样本量、缩短试验的时长、减少试验的费用等。同时，因 II 期入组的受试者有更长的随访时间，有时可以更早地观察到药物的长期安全性。

采用适应性 II/III 期无缝剂量选择的设计需要考虑多种因素。由于期中分析时对 II 期数据可能无法进行全面深入地分析，如果对试验药物了解甚少，一般应慎重选择采用适应性 II/III 期无缝剂量选择的设计，因 II 期试验的数据要包含在最终分析中加之 III 期试验已经在进行之中，而如果使用两个单独的试验可以有更多的选择方式。还有一些因素，例如，III 期试验的主要终点需要较长的随访时间，II 期或许只能够用替代终点进行判断，当替代终点与主要终点关联性不高甚至较差时，用替代终点选择 III 期试验的剂量会带来很大的不确定性。又如，也应考虑是否有足够的生产能力在短时间内提供 III 期所需的药物。

以上讨论的适应性 II/III 期无缝剂量选择的设计也可以直接应用于其他类似的试验，例如联合用药和单药的选择，或者不同药物之间的选择等。

（四）适应性富集设计

适应性富集设计是指试验将根据期中分析的结果，依据预先设定的标准对目标人群进行适应性调整，以决定试验后续阶段的目标人群。试验的后续阶段可能继续在全人群中进行，或者仅入组亚群并有可能需要做一些相应的适应性调整，或者加大样本量继续入组全人群，这同时也自然地加大了亚群的入组人数。试验的最终分析目标可能仅是全人群、亚群，或者全人群和亚群都包含。试验的最终分析将包含试验的两个阶段入组的所有受试者的数据，并有相应的调整方法以控制 I 类错误率。

如果已知试验药物只对某特定亚群有效，那么，临床试验应该只在该亚群中招募受试者。但实际中更为常见的情形是试验药物有可能对某亚群有较大的疗效，但不清楚对全人群是否也有足够大的疗效。在这种情况下，如果试验药物对全人群有足够大的疗效，只入组亚群受试者就会失去显示对全人群有效的机会；如果试验药物对全人群疗效较小但对某亚群有效，入组全人群受试者极有可能得不到预期的阳性结果，同时也失去了显示对亚群有效的机会。采用适应性富集设计来选择目标人群可以同时兼顾两者，利用试验本身的结

果以便可以更科学地选择出目标人群，增加药物研发的成功率。

由于适应性富集设计中目标人群的选择涉及全人群和亚群，以及期中分析时采用非盲态的组间比较，因此应分别明确定义两个人群的统计假设和相应的统计方法，并控制 I 类错误率。

对于目标人群的选择标准，可以基于疾病特征、预后生物标志物或预测生物标志物等各种标准。一般而言，采用公认的疾病相关特征或预后相关生物标志物来选定目标人群，试验的设计和操作会相对简单。目前，采用预测生物标志物来选择目标人群的研究日趋增多，但许多预测生物标志物的临床价值尚不明确。如果试验要用一个全新的预测生物标志物来选择目标人群，必须要有对应的诊断方法。所用诊断方法必须已经被监管部门批准上市，如果没有，可能需要同时研发。

（五）两阶段适应性设计

两阶段适应性设计，是指将一个试验分为两个阶段，适应性调整前是第 1 阶段，适应性调整后是第 2 阶段。在第 1 阶段结束时进行期中分析，依据预先设定的修改计划，对第 2 阶段的试验进行适应性修改。

以上所讨论的成组序贯设计（若仅有一次期中分析）、样本量重新估计、适应性 II/III 期无缝剂量选择的设计、适应性

富集设计都是两阶段适应性设计。两阶段适应性设计也包括其他常见的设计，例如在第 1 阶段结束期中分析时，从第 1 阶段选择一个合适的主要终点用于第 2 阶段；从第 1 阶段的两个或多个目标子群中选择一个合适的目标子群用于第 2 阶段；将第 1 阶段的单一主要假设修改为多个主要假设等。

有两点需要注意：一是成组序贯设计和适应性成组序贯设计之间存在区别。两者仅在期中分析、提前终止试验和样本量重新估计时类似；若一个适应性成组序贯设计包含了其他的适应性修改，则成组序贯设计中标准的分析方法就不适用。另一点是当两阶段适应性设计在分析以生存期为终点的试验时，无论第 1 阶段入组受试者的终点事件发生在哪个阶段，计算时均应将其归在第 1 阶段的结果内，否则两阶段的独立性假设将不再成立，导致 I 类错误率增加。

大多数适应性设计都属于两阶段范畴。两阶段适应性设计的原理和方法可类似地推广到多阶段或多重适应性设计。

（六）适应性主方案试验设计

主方案试验设计是指一个整体临床试验方案含有多个子方案，不同的子方案可同时检验一种药物对于多种疾病的临床效果，也可同时检验多种药物对于一种疾病的临床效果，或者同时检验多种药物对于多种疾病的临床效果。每一个子方案可以是单臂试验，也可以是随机对照试验。如果有子方案是随机对照试验且病人群体相同，这些随机对照试验有可

能共用一个对照组，也可能有各自的对照组。主方案试验也用来泛指由患者特定特征（如疾病、组织学类型、分子标记物）为标志的临床试验。主方案试验具有很多优点，例如能够为患者提供最大的入组机会并选择最合适的受试药物的机会。常见的主方案设计包括篮式试验、伞式试验和平台试验设计。

篮式设计旨在评估一种药物治疗具有同一种生物学特征的不同疾病类型的临床效果，每一个子方案都针对一种或多种疾病类型。伞式设计旨在评估多种药物针对同一种疾病或生物标记物类型的靶向治疗的临床效果。平台设计旨在评估多种药物针对多种疾病的临床疗效。平台试验通常会维持试验长期进行，并允许新的试验药物随时加入试验平台，同时，对照药物随着时间推移也可能发生变更。

主方案试验虽然具有很多优点，但由于其复杂性，在计划、执行、统一管理结构的建立、尤其是统计分析等方面都面临着较大地挑战。如果计划采用主方案试验，应对试验各个方面可能涉及的各种问题做全面、深入和细致的研究后，再慎重选择。

适应性主方案设计是指在主方案设计中包含了一种或多种适应性调整的设计，它可以灵活地采用多种适应性调整，例如添加一个或多个新的子方案，提前结束一个或多个子方

案，重新估计样本量，调整检验的假设、主要终点和主要统计方法，或对不同的子方案设计做不同的适应性调整等。

（七）多重适应性设计

多重适应性设计是指一个试验中采用了多于一种适应性调整方法的试验设计。以上所讨论的适应性设计方法都可以同时用于同一个临床试验。例如，一个临床试验在第1阶段结束时确定了下阶段的用药剂量，其后可以选择目标人群，再其后可以做样本量重新估计。

原则上讲，如果一个临床试验设计包含了多种适应性调整，只要符合适用性、合理性、完整性和可行性的要求，多重适应性设计都可以考虑。但由于多重适应性设计的复杂性，在一个试验中是否有必要引入过多的适应性调整，建议申办者予以慎重考虑。

四、 其他考虑

（一）仅基于外部数据的修改

仅基于外部数据的修改是指在试验进行过程中仅仅基于外部数据对于一个进行中的临床试验做出某些修改，在本指导原则中，不将其归于所定义的适应性修改。

在试验进行过程中，常常会有与本试验相关的新的信息出现，而这些信息一般都是基于在当前试验设计时还不存在的新近完成的试验或研究。基于外部数据对于一个进行中的临床试验做出某些修改，必须有充分的依据且不应破坏试验

的合理性及完整性，并需提前与监管机构进行沟通确认后方可通过试验方案的修正案来体现。申办者尤其要注意这些修改是仅基于外部数据，而非基于进行中的试验本身的结果。

在当前试验设计时可能同时会有其它尚未完成的相同药物的其它试验正在进行，其与当前试验相关，并在当前试验的设计时预先设定，这同样被视为外部数据并可通过试验方案的修正案来体现。

如果需要用一个 II 期试验的结果来决定当前试验的生物标志物的阈值，同时结合当前试验期中分析的累积数据做出修改，则在本指导原则中被归为适应性修改，此时需在适应性修改计划中预先设定。

（二）监管的其他考虑

作为试验方案的一部分，适应性修改计划应在临床试验开始前的试验方案中预先设定。

申办者如果计划在确证性试验中采用适应性设计，或在统计推断中用到贝叶斯方法或模拟方法等，申办者应在试验方案设计阶段与监管部门进行沟通交流。

申办者在沟通交流中递交的资料应包含重要的用来支持采用适应性设计的文献和数据以便监管部门审评。资料的准备应主要围绕预先设定的适应性调整计划的细节，包括其适用性、合理性和完整性等。

申办者在沟通交流资料中应讨论采用适应性设计的理由，包括与传统设计相比的优势、需要采用适应性设计解决的具体问题以及解决的方式、适应性调整后结果的可解释性等；还应包括预先设定的调整计划的细节，比如，期中分析的时间和目的、决定适应性调整的统计规则、最终分析的统计检验方法、控制 I 类错误率的方法等；以及关键的实施适应性设计的操作流程，保证试验完整性的具体措施等。

五、 参考文献

1. 国家药品监督管理局药品审评中心. 药物临床试验数据监查委员会指导原则（试行）. 2020.
2. Bauer P, Köhne K. Evaluation of experiments with adaptive interim analyses. *Biometrics*.1994; 50:1029–1041. corrections. *Biometrics*.1996; 52:380.
3. Bauer P, Kieser M. Combining different phases in the development of medical treatments within a single trial. *Statistics in Medicine*.1999; 18: 1833-1848.
4. Bauer P, Posch M. Letter to the Editor. Modification of the sample size and the schedule of interim analyses in survival trials based on data inspections by H. Schaefer and H.-H. Mueller, *Statistics in Medicine* 2001; 20: 3741–3751. *Statistics in Medicine*.2004; 23: 1333–1335.
5. Chang M. Adaptive design method based on sum of p-values. *Statistics in Medicine*.2007; 26:2772–2784.
6. Chen C, Li X, Li W, Beckman RA. Adaptive Expansion of Biomarker Populations in Phase 3 Clinical Trials. *Contemporary Clinical Trials*.2018; 71:18-85.
7. Chen C, Anderson K, Mehrotra DV, Rubin EH and Tse A. A 2-in-1 Adaptive Phase 2/3 Design for Expedited Oncology Drug Development. *Contemporary Clinical Trials*.2018; 64:238-242.

8. Chen JYH, DeMets DL, Lan GKK. Increasing the sample size when the unblinded interim results is promising. *Statistics in Medicine*.2004; 23:1023-1038.
9. Chow SC. Complex innovative design for NASA clinical trials. *Academic Journal of Gastroenterology & Hepatology*.2020; 2: 1-9.
10. Chow SC, Chang M. *Adaptive Design Methods in Clinical Trials*. CRC Press, 2nd edition.2011.
11. Chow SC, Lin M. Analysis of two-stage adaptive seamless trial design. *Pharmaceutica Analytica Acta*.2015; 6: 341-440.
12. Chow SC, Shao J, Wang H, Lokhnygina Y. *Sample size calculations in clinical research*. Chapman & Hall/CRC, 3rd edition.2018.
13. Chow SC, Tu YH. On two-stage seamless adaptive design in clinical trials. *Journal of Formosan Medical Association*.2008; 107: s1-59.
14. Cui L, Hung HMJ, Wang SJ. Modification of sample size in group sequential clinical trials. *Biometrics*.1999; 55:853-857.
15. Cui L, Zhang L. On the efficiency of adaptive sample size design. *Statistics in Medicine*.2019;38:933-944.
16. Friede T, Kieser M. Sample size recalculation in internal pilot study designs: a review. *Biometrical Journal*.2006; 48:537–555.
17. Friede T, Parsons N, Stallard N. A conditional error function approach for subgroup selection in adaptive clinical trials. *Statistics in Medicine*.2012; 31:4309–4320.

18. Friede T, Stallard Nigel. A comparison of methods for adaptive treatment selection. *Biometrical Journal*.2008; 50:767–781.
19. Gould AL. Interim analyses for monitoring clinical trials that do not materially affect the Type I error rate. *Statistics in Medicine*.1992; 14:1039-1051.
20. Hochberg Y. A sharper Bonferroni procedure for multiple tests of significance. *Biometrika*.1988; 75: 800–802.
21. Howard DR, Brown JM, Todd S, Gregory WM. Recommendations on multiple testing adjustment in multi-arm trials with a shared control group. *Statistical Methods in Medical Research*.2018 ; 27:1513-1530.
22. Jenkins M, Stone A, Jennison C. An adaptive seamless phase II/III design for oncology trials with subpopulation selection using correlated survival endpoints. *Pharmaceutical Statistics*.2011; 10:347–356.
23. Lan KG, DeMets DL. Discrete Sequential Boundaries for Clinical Trials. *Biometrika*.1983; 70:659–663.
24. Lehman W, Wassmer G. Adaptive sample size calculations in group sequential trials. *Biometrics*.1999; 55: 1286–1290.
25. Maca J, Bhattacharya S, Dragalin S. et al. Adaptive Seamless Phase II/III Designs Background, Operational Aspects, and Examples. *Drug Information Journal*.2006; 40: 463-474.

26. Marcus R, Peritz E, Gabriel KR. On closed testing procedures with special reference to ordered analysis of variance. *Biometrika*.1976; 63:655–660.
27. Mehta CR, Pocock SJ. Adaptive increase in sample size when interim results are promising: a practical guide with examples. *Statistics in Medicine*.2001; 30:3267- 3284.
28. Müller HH, Schäfer H. Adaptive group sequential designs for clinical trials: Combining the advantages of adaptive and of classical group sequential approaches. *Biometrics*.2001; 57:886–891.
29. Müller HH, Schäfer H. A general statistical principle for changing a design any time during the course of a trial. *Statistics in Medicine*.2004; 23:2497–2508.
30. O'Brien PC, Fleming TR. A Multiple Testing Procedure for Clinical Trials. *Biometrics*.1979; 549–556.
31. Pocock SJ. Group Sequential Methods in the Design and Analysis of Clinical Trials. *Biometrika*.1977; 64:191–199.
32. Proschan MA, Hunsberger SA. Designed extension of studies based on conditional power. *Biometrics*.1995; 51:1315–1324.
33. Rosenblum M, Van Der Laan MJ. Optimizing randomized trial designs to distinguish which subpopulations benefit from treatment. *Biometrika*.2011; 98: 845-860.

34. Shih WJ. Sample size re-estimation – a journey for a decade. *Statistics in Medicine*.2001; 20:515-518.
35. Shih WJ, Li G, Wang Y. Methods for flexible sample-size design in clinical trials: Likelihood, weighted, dual test, and promising zone approaches. *Contemporary Clinical Trials*.2016; 47: 40-48.
36. Simes RJ. An improved Bonferroni procedure for multiple tests of significance. *Biometrika*.1986; 73:751–754.
37. Stallard N, Hamborg T, Parsons N, Friede T. Adaptive designs for confirmatory clinical trials with subgroup selection. *Journal of Biopharmaceutical Statistics*.2014; 24:168–187.
38. Wang SJ, Hung HMJ, O’Neill RT. Adaptive patient enrichment designs in therapeutic trials. *Biometrical Journal*.2009; 51:358–374.
39. Wassmer G, Brannath W. Group sequential and confirmatory adaptive designs in clinical trials. Springer,2016.
40. Wu PS, Lin M, Chow SC. On sample size estimation and re-estimation adjusting for variability in confirmatory trials. *Journal of Biopharmaceutical Statistics*.2016; 26:44-54.
41. Zhang JJ, Blumenthal G, He K, Tang S, Cortazar P, Sridhara R. Overestimation of the effect size in group sequential trials. *Clinical Cancer Research*.2012; 18: 4872-4876.

42. Zheng J, Chow SC. Criteria for dose-finding in two-stage seamless adaptive design. *Journal of Biopharmaceutical Statistics*. 2019; 29: 908-919.

附录：词汇表

适应性设计 (Adaptive design)：按照预先设定的计划，在期中分析时使用试验期间累积的数据对试验做出相应修改的临床试验设计。

成组序贯设计 (Group sequential design)：是指方案中预先计划在试验过程中进行一次或多次期中分析，依据每一次期中分析的结果做出后续试验决策的试验设计。

盲态/非比较分析方法 (Blinded/Non-comparative analysis)：是指期中分析时不使用实际试验分组信息，或者未做任何涉及组间比较的分析。

非盲态分析/比较分析方法 (Unblinded/Comparative analysis)：是指期中分析时使用试验分组信息（包括各组的真实名称或可区分的分组代码）的分析，分析内容涉及组间比较。

适应性主方案设计 (Master protocol with adaptive designs)：是指在主方案设计中包含了一种或多种适应性调整的设计。

多重适应性设计 (Multiple adaptive design)：是指一个试验中采用了多于一种适应性调整方法的试验设计。

贝叶斯方法 (Bayesian method)：贝叶斯方法一般是指在对未知参数做出统计推断时，它先使用先验信息（先验分布函数）对未知参数做一初始判断，在搜集到新数据后，它根据

贝叶斯原理将先验信息和新数据总结在另一个函数中（后验分布函数），并基于此后验分布做出统计推断。

模拟方法 (Simulation method)：是指使用计算机技术通过创建虚拟患者数据并根据预先指定的模型预测患者的临床结果来模拟临床试验的进行。