

20190918	
发布通告	
名称	晚期非小细胞肺癌临床试验终点技术指导原则
适用范围	化学药,生物制品
专业分类	临床
版本状态	颁布
关键词	
附件 1	晚期非小细胞肺癌临床试验终点技术指导原则.pdf

晚期非小细胞肺癌临床试验终点技术指导原则

二〇一九年九月

晚期非小细胞肺癌临床试验终点技术指导原则

一、前言

肺癌的发病率和病死率居全球和中国恶性肿瘤之首^[1]。2015年中国新发肿瘤患者 392.9 万例，死亡 233.8 万例，新发肺癌 78.7 万例，死亡 63.1 万例^[2]。非小细胞肺癌（non-small cell lung cancer, NSCLC）占肺癌总体的 85%，大部分初诊时已为晚期。近年小分子酪氨酸激酶抑制剂（tyrosine kinase inhibitors, TKI）、抗血管生成药物和免疫检查点抑制剂的应用已显著提高了患者生存。

晚期 NSCLC 为抗肿瘤新药研发的热点领域，创新药物众多，临床证据链日趋复杂，涌现出了复杂的终点指标和研究设计——包括替代终点、中间临床终点和其它创新终点；并出现共同终点平行检验、复合终点序贯检验等复杂设计。现有指导原则尚不能涵盖。因此，本文旨在阐述当前晚期 NSCLC 临床试验终点的一般性设计与审评考虑，期望为抗肿瘤药物研发人员在晚期肺癌的临床试验设计和终点选择上提供参考，科学、高效地确定药物疗效，提高临床研发效率，使患者更早获益。

本指导原则适用于支持晚期 NSCLC 适应症注册的临床试验设计及终点选择。本指导原则所涉及的抗肿瘤药物试验设计同样应遵循临床试验设计的一般原则，包括但不限于人用药品注册技术要求国际协调会议（international conference for harmonization,

ICH) 所发布的 E8¹、E9²、E10³和 E17⁴等指导原则, 以及国家药品监督管理局 (National Medical Products Administration, NMPA) 已发布的《抗肿瘤药物临床试验终点技术指导原则》和《抗肿瘤药物临床试验技术指导原则》等相关内容。

本指导原则所涉及的观点代表当前 NMPA 对晚期 NSCLC 临床试验设计和终点选择的审评认识, 不能涵盖在抗肿瘤新药研发中遇到的所有情况。鼓励研发从业人员探索科学创新的终点和试验设计, 并及时与 NMPA 的审评部门沟通和交流。

二、背景

当前晚期 NSCLC 的治疗目标为延长生命和提高生活质量, 关键注册研究的试验终点应能有效反映临床获益的指标或事件。上世纪九十年代前批准的抗肿瘤药物多以总生存期 (overall survival, OS) 为主要终点。OS 的定义明确、客观稳健, 是反映患者生存获益的金标准, 但试验耗时长且需较大样本量。监管部门为加速药品上市、改善治疗可及性, 对难治疾病背景下具有突出临床获益的药物实施加速审批, 即允许使用可合理预测临床获益的替代终点, 如: 客观缓解率 (objective response rate, ORR) 和无进展生存期 (progression free survival, PFS) 为主要终点支持新药批准上市^[3-5]。

近年晚期 NSCLC 的治疗模式已从基于组织病理学的细胞毒化疗, 转变为基于驱动基因突变等疗效预测生物标志物的分子靶

¹ E8 《General Considerations for Clinical Trials》;

² E9 《Statistical Considerations in the Design of Clinical Trials》

³ E10 《Choice of Control Group in Clinical Trials》

⁴ E17 《General Principles for Planning and Design of Multi-Regional Clinical Trials》

向治疗，以及免疫、细胞毒和抗血管生成等药物联合治疗。监管部门通常基于直接的临床获益批准新药，如 OS 的延长和肿瘤相关症状的改善，也可接受可预测临床获益的替代终点批准新药，如较高的 ORR 和足够的缓解持续时间（duration of response, DOR），监管机构需要权衡替代终点和创新研究设计对临床获益评价的影响。本指导原则将对疗效评估的一般性监管要求及支持批准的终点选择进行讨论。

三、抗肿瘤药物临床试验的常用终点

抗肿瘤药物临床试验常用的终点依据来源可分为三类：基于死亡事件的终点，如 OS 及 OS 率；基于肿瘤测量的终点，如采用实体瘤疗效评价标准（response evaluation criteria in solid tumors, RECIST）评估的 ORR，或基于 RECIST 和访视的至进展时间（time to progression, TTP）、PFS 和至治疗失败时间（time to treatment failure, TTF）等；基于症状评估的终点如疼痛的减轻、生活质量（quality of life, QoL）等患者报告结果（patient report outcome, PRO）。

终点的选择应结合肿瘤分期、既往治疗和起效特点等因素综合考量。

表 1、晚期 NSCLC 药物常用临床试验终点比较

终点	研究设计	优点	缺点
OS	· 随机对照试验（randomized controlled trials, RCT） · 推荐盲法	· 广泛应用的临床获益终点 · 易于测量且精确	· 随访时间较长，通常需较大样本 · 易受交叉和后续治疗影响 · 包括非肿瘤死亡

终点	研究设计	优点	缺点
PFS	RCT ·推荐盲法 ·推荐盲态独立中心评估 (Blinded Independent Central Review, BICR)	·与 OS 相比通常需要较少的样本量、较短的随访时间 ·包括疾病稳定 ·不受交叉和后续治疗影响 ·通常基于客观、定量评估	·统计学意义或不具有临床意义，不总能转化为总生存获益 ·非精确测量，评价可存偏倚，特别是在开放标签试验中 ·不同试验可有不同的评估间隔，结果受评估间隔影响，需频繁影像学评估
TTP	·同 PFS	·较 PFS 能更加客观体现药物的抗肿瘤作用	·不包括死亡，与 OS 的相关性低于 PFS
TTF	·同 PFS	·综合了药物的有效性和耐受性特征，是综合的临床终点	·由于治疗失败停止治疗的停药判断可受研究者的主观影响，不如 RECIST 判断的 PD 客观
ORR	·单臂试验 (single arm trial, SAT), ·单臂有条件批准的常用替代终点 ·独立影像评估 ·推荐 BICR	·更小的样本量、更快达到终点 ·更适于富集人群试验 ·疗效归因于药物，排除疾病自然进程	·不是临床获益的直接测量 ·不是对药物活性的综合测量 ·需结合缓解持续时间和疾病控制判断
症状终点	RCT ·推荐盲法	·患者临床获益的直接感受	·数据缺失和不完整较普遍 ·小变化的临床意义不明确 ·缺乏经过验证的测量工具 ·可能有偏差，在不同治疗组间评估时间上的可比性十分关键

(一) 基于死亡事件的终点

OS 定义为从随机化到任何因素导致患者死亡的时间。OS 的判定精确可靠，不易偏倚，常作为首选终点。以 OS 为主要终点的临床试验需采用随机对照设计，常需较大样本量和更长的随访时间，易受到交叉和后续治疗影响。

OS 率定义为自随机化至指定时间节点同一试验组内生存的受试者所占的比例，为 OS 的中间临床终点，在既往研究中可作为次要终点，随机临床试验中，可通过 OS 率的比较观察到治疗组的获益，如 1 年 OS 率。虽然 OS 率的计算精确可靠并将更早达到，但 OS 率通常作为描述性终点，OS 率的评估更多受到时

间选择的影响，其临床意义和统计学意义尚不明确，提前分析 OS 率将导致破盲不利于生存获益的检验，现阶段 OS 率多作为次要终点和支持性证据。

（二）基于肿瘤测量的终点

肿瘤临床治疗常基于病灶的影像学评估结果决策，肿瘤测量的终点被视为具有临床获益相关性。RECIST（目前为 1.1 版本）是目前广泛应用于 NSCLC 的疗效评价标准。随着免疫检查点抑制剂的使用，免疫治疗疗效评价标准（immune response evaluation criteria in solid tumors, iRECIST）开始应用于临床试验，但对照组为标准化疗时不适用 iRECIST 判断 PFS。目前在以新药注册为目的的肿瘤免疫治疗临床试验中，RECIST 1.1 仍为最常用的肿瘤疗效判定标准，推荐在单纯免疫治疗及“头对头”设计的免疫治疗临床试验中增加 iRECIST 评估标准，并与传统 RECIST 结果进行对比。

ORR 定义为肿瘤体积缩小达到预先规定值并能维持最低时限要求的患者比例。ORR 为完全缓解（complete response, CR）与部分缓解（partial response, PR）的比例之和，ORR 不包括疾病稳定（stable disease, SD），排除了疾病自然病程的影响，相比疾病控制率（disease control rate, DCR），ORR 可更可靠反映药物的抗肿瘤活性，是单臂临床试验常用替代终点。

DOR 定义为肿瘤第一次评估为客观缓解至第一次评估为 PD 或 PD 前任何原因死亡的时间，反映了 ORR 的持续时间。

PFS 定义为从随机化至出现肿瘤客观进展或全因死亡的时

间，是 OS 的替代终点。与 TTP 相比，PFS 包括了任何原因导致的死亡，与 OS 相关性更高，且不受后续治疗影响，是随机对照设计临床试验最常用的替代终点。

TTP 定义为从随机化至出现肿瘤客观进展的时间，不包括死亡。TTP 能精确反映治疗带来的近期生存获益，由于排除了死亡，TTP 对治疗临床获益的相关性差于 PFS 和 TTF。

TTF 定义为从随机化至治疗失败或退出试验的时间，退出试验的原因可为患者要求、疾病进展、死亡或不良事件等。与 PFS 相比，TTF 覆盖了非疾病进展导致的退出，并可包括疾病进展后的继续治疗，是综合的临床终点。因 TTF 不能充分将药物的疗效和耐受性等因素区分，当前，不常用于抗肿瘤药物确证性研究的主要研究终点。

TTP、PFS 和 TTF 结果的判定均受访视间隔设计和试验质量的影响，如由失访或在研究期间未观察到终点事件而产生的删失和截尾值过多，将影响以上终点结果的分析。

肺癌脑转移患者的纳入及终点指标考虑

脑转移是导致晚期 NSCLC 疾病进展和治疗失败的重要原因，有症状患者需进行临床干预。既往晚期 NSCLC 临床试验多排除脑转移患者，或仅入组经局部和系统治疗病情稳定的无症状脑转移患者，因此临床试验结果不能反映药物对脑转移患者的疗效，将导致重要有效性信息的缺失。考虑肺癌脑转移高发，且控制脑转移为转移性 NSCLC 重要治疗目标，鼓励基于药物前期的临床研究结果，在试验中纳入脑转移患者。注意在基线收集完整

准确的脑转移信息，在疗效评估中增加局部转移获益的评价，如颅内缓解率、颅内缓解持续时间和出现新的脑转移至颅内疾病进展时间等。

（三）基于症状评估的终点

肿瘤患者症状和体征的改善被认为是直接的临床获益，而非替代终点。监管当局可能基于显著的症状改善（如恶性积液的控制、癌性乏力的改善和骨相关事件的改善等）批准新药上市。复合症状终点中的不同症状指标应具有相似的临床重要性，其临床获益不仅归因于单个指标的改善，对复合终点进行分析时应针对具体单个指标进行分析。

当以症状和体征的改善作为支持抗肿瘤药物审批的主要终点时，应当能够区分是肿瘤相关症状的改善还是试验组药物毒性的减小或缺失。如选择具体症状缓解作为终点，受试者需在基线时存在该症状并且症状由疾病导致。

数据缺失、评价不充分将增加症状终点的评价复杂性，应严格执行访视计划使访视完成率均衡和最大化，统计分析计划（statistical analysis plan, SAP）应说明如何处理缺失数据，当患者停止治疗时应该继续收集可供分析的信息。应进行多种症状的前瞻性数据收集。存在多重检验情况时，需要在 SAP 中说明必要的一类错误控制方法。

四、晚期非小细胞肺癌新药注册临床试验的设计及终点考虑

我国晚期 NSCLC 具有高比例的驱动基因突变，其临床试验设计可根据有无疗效预测生物标志物分为富集人群试验和非富

集人群试验。

（一）以肿瘤生物标志物富集人群的临床试验

以肿瘤生物标志物富集人群的试验设计，应考虑生物标志物和治疗分线等因素：

1. 生物标志物

生物标志物选择：需明确具体的生物标志物命名，例如明确某表皮生长因子受体酪氨酸激酶抑制剂（epidermal growth factor receptor-tyrosine kinase inhibitor, EGFR-TKI）适用于 EGFR 第 19 外显子缺失突变和第 21 外显子 L858R 突变。晚期肺癌相关生物标志物的命名应遵循 ICH E15⁵。

生物标本要求：标本的取材、固定、储运条件等可显著影响判定结果，需明确研究可采纳的生物标志物来源——原发灶、转移灶或外周血（如 ctDNA）等；明确标本的固定要求——石蜡包埋或新鲜冰冻组织；明确标本的储运条件等。以上内容应要在研究方案给予说明或细化为相应的标准操作规程。生物标本的采集、处理、运输、储运和处置的一般性原则应遵循 ICH E18⁶的相关要求。

伴随诊断方法：在确定生物标志物和标本要求后，需明确伴随诊断方法。

部分应用广泛、效能已被证实的生物标志物伴随诊断，如检测 EGFR 敏感突变的测序法或突变扩增系统（amplification

⁵ E15 《Definitions for Genomic Biomarkers, Pharmacogenomics, Pharmacogenetics, Genomic Data and Sample Coding Categories》

⁶ E18 《Genomic Sampling and Management of Genomic Data》

refractory mutation system, ARMS) 法, 可在本地实验室检测, 不需明确伴随诊断/诊断试剂盒。

尚未广泛应用, 正在临床开发过程中的生物标志物伴随诊断, 如 PD-L1 表达、c-MET 扩增和微卫星不稳定性高/错配修复基因缺陷 (high microsatellite instability/DNA mismatch repair deficiency, MSI-H/dMMR), 需明确伴随诊断方法/诊断试剂盒, 并推荐采用中心实验室 (接受本地中心实验室)。

鼓励新药和体外伴随诊断试剂同步开发。如开发一个新生物标志物或检测方法, 在递交上市申请时应提交充分的关于所选定生物标志物的验证资料。建议在开展关键注册临床试验前与监管沟通讨论所需提交的资料要求。应遵循 ICH E16⁷对递交材料的背景、结构以及格式要求准备资料, 逐步推进。

2. 对治疗分线的考虑: 通常初治患者体能状态好, 耐受性优于经多线治疗患者。因此, 试验设计时需考虑对既往治疗线数进行分层。

应综合疾病背景、药物可及性、临床需求、前期临床试验数据等因素, 综合决定以 RCT 或 SAT 支持新药注册。

3. RCT 的设计和终点考虑: RCT 是确证药物疗效最为可靠的方法, OS 为反映抗肿瘤药物临床获益的金标准, 通常作为 RCT 的主要终点。随着治疗手段的丰富, OS 不断延长增加了评价难度, 因此, 单独 PFS 或 PFS 与 OS 共终点可被接受作为初治晚期 NSCLC 注册研究的主要终点。纳入复发难治或无标准治疗患者

⁷ E16 《Biomarkers Related to Drug or Biotechnology Product Development Context, Structure and Format of Qualification Submissions》

的 RCT 则仍推荐 OS 为主要终点。

选定主要研究终点的同时应确定试验设计的设计——优效、非劣效或是等效设计。当对照组为标准治疗时，可接受优效、非劣效或等效设计，选择等效或非劣效假设时，申请人需要与监管部门沟通等效或非劣效界值；当为安慰剂对照或试验设计为加载设计（Add on）时，仅接受优效设计，安慰剂对照临床试验设计推荐联合最佳支持治疗（best support care, BSC），以提高对受试者的保护。

在 RCT 试验中，监管机构鼓励以合理的替代终点和预设的期中分析支持批准，在多臂 RCT、预设期中分析、共同主要终点或拟对主要终点多重检验的情况下，需保障复杂设计下主要终点的总体一类错误值，控制在双侧 0.05 以下。

4.SAT 的设计及终点考虑：在难治的疾病背景下观察到新药具有突破性疗效时，申请人可与监管机构沟通，考虑采用替代终点的 SAT 作为支持有条件批准的关键注册研究，并通过后续确证性随机对照试验获得完全批准。

SAT 最大的局限是无对照带来的偏倚，导致药物疗效评价的不确定性。因此，如考虑以 SAT 作为支持新药加速批准的关键研究，应具备如下条件：

充分的历史数据：SAT 的历史对照数据应来自高级别循证医学证据——具备相近年代、相同疾病背景和足够样本量的特点。数据可来自单个 RCT、真实世界证据、系统回顾和 meta 分析的结果，并分析现阶段下该数据的可靠性。

难治的疾病背景：SAT 应在复发难治的、医疗需求亟待满足的疾病中开展。SAT 受试者应是目前标准治疗失败或无标准治疗的患者人群。

以缓解率/客观指标为主要终点：由于无对照设计及历史对照的不确定性，SAT 需慎重考虑主要研究终点。ORR 定义准确、排除肿瘤自然病程并直接反映药物的抗肿瘤疗效。因此晚期 NSCLC 的 SAT 常采用 ORR 为主要终点。

预设的缓解率目标值：应预设 ORR 的目标值。目标 ORR 值 95%的置信区间下限应在充分历史数据基础上结合产品自身早期临床数据后，与监管机构沟通共同确定。通常期望目标 ORR 值与历史对照相比有较大提高，目标 ORR 值 95%的置信区间下限需高于标准治疗历史数据上限^[10]。

持续的缓解时间：在确定 ORR 达到既定标准基础上，评价缓解持续时间非常重要，持续的缓解时间将增加治疗获益的稳健性，较长的用药时间也反映药物的安全耐受性信息。

独立评审委员会 (independent review committee, IRC)：SAT 必须使用独立影像评审委员会，以减少研究者偏倚，并采用研究者评价的 ORR 进行敏感性分析。

SAT 上市后要求和模式：由于 SAT 为有条件的加速批准，申请人需在开展 SAT 前与监管机构沟通完全批准的条件。通常情况下，在同一线患者中开展的 RCT 被视为严谨的确证性临床^[11]，因分子靶向治疗不过分强调化疗的分线，因此在前一线人群中具有严谨统计假设的确证性临床试验也可能被接受。监管部门

将结合目标疾病人群发病率、疗效数据的强度、伦理要求等方面综合考虑确证性研究的具体要求。无法完成 RCT 的罕见情况下，可依据真实世界数据支持完全批准。

（二）不以肿瘤生物标志物富集人群的临床试验

晚期 NSCLC 中，因尚无明确的疗效预测因子，细胞毒和抗血管药物的临床试验常在非富集人群中开展（考虑组织学区分），部分免疫检查点抑制剂也选择非富集人群开展研究。

非富集试验需着重考虑影响预后的因素（如病理组织学类型、ECOG 体能状态评分、既往治疗和基线脑转移等），并将重要预后因素分层。在非富集试验中，治疗分线是试验主要终点选择的重要依据。

现阶段无驱动基因突变的晚期 NSCLC 的一线标准治疗为以含铂双药为基础的联合治疗方案，可联合抗血管或免疫治疗。与富集人群的 RCT 设计相似，非富集人群一/二线治疗的 RCT 可以选择 PFS 为主要研究终点；二线后的研究推荐选择 OS 为主要终点，但需考虑后续治疗及交叉可能带来的影响。免疫检查点抑制剂的疗效主要体现为部分患者的长期生存获益，需在一线治疗中关注 OS 获益的评价，推荐将 OS 和 PFS 作为共同主要终点，或将 OS 作为唯一主要终点。免疫检查点抑制剂的疗效普遍具有延滞拖尾效应，因此不建议对主要终点在过少事件数（如不足计划 OS 事件数 50%）时开展期中分析。

维持治疗的目的是巩固一线治疗获益，延长至疾病进展时间。因此维持治疗临床试验的主要终点宜与治疗目标契合，可选

择 PFS 或 TTF 为主要终点；当维持治疗的疗效为远期生存获益时，应选择 OS 为主要终点。

除生物标志物的考虑外，非富集人群与富集人群遵循相同的试验设计和终点选择。

五、晚期非小细胞肺癌新药注册临床试验主要终点获益考虑

在现阶段晚期 NSCLC 的药物研发中，不同研究终点的 RCT 设计，均在主要终点达到统计学显著性后评价临床价值。

（一）总生存期

现阶段晚期 NSCLC 的治疗目标是延长患者生命，提高生活质量。OS 是反映生存获益的直接终点，其临床意义重于 PFS 等替代终点。后续治疗是 OS 最大的影响因素，因此，在后续治疗均衡的情况下，现阶段国内外专家考虑认为，在转移性肺鳞癌中，2.5 个月以上 OS 获益被认为具有显著的临床价值，非鳞癌中 3.25 个月以上的 OS 获益被认为具有显著的临床价值，目标 HR 的点估计值通常不高于 0.80^[12]。通常将风险比和中位获益结合评价，认为两者均达到具有更稳健的 OS 获益。

（二）无进展生存期

PFS 为 OS 的替代终点，PFS 获益能否转化为 OS 具有不确定性，多种非治疗因素可使 PFS 的差异达到统计学意义——预后因素分层不均衡、选择了较弱的对照治疗、甚至疗效评价时间点的设计等。现阶段晚期一线 NSCLC 中，肺鳞癌中 3 个月以上的 PFS 获益具有显著的临床价值，非鳞癌中 4 个月以上的 PFS 获益具有显著的临床价值，目标 HR 的点估计值通常不高于

0.70^[12]。

（三）客观缓解率

以 ORR 为主要终点时，核心考虑为全分析集 IRC 评价的 ORR 是否达到了预设目标值，并兼顾足够的缓解持续时间。现阶段下晚期 NSCLC 中，ORR 95%的置信区间下限需达到预设目标值。

六、结语

晚期 NSCLC 是抗肿瘤药物的研发热点，伴随新药研发，晚期 NSCLC 治疗的证据链日益丰富，药物临床试验的设计和终点选择趋于复杂。现阶段下，延长生存时间和提高生活质量仍是晚期肺癌治疗的核心目标，临床试验的终点选择均以能够客观高效反映肿瘤治疗临床获益为原则。科学的进步将推动抗肿瘤产品的研发，鼓励申请人与监管机构沟通并探索创新的试验设计和研究终点。本指导原则将基于晚期非小细胞肺癌临床试验的研究进展情况适时进行更新。

参考文献

- [1] Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. CA Cancer J Clin. 2016. 66 (1) : 7-30.
- [2] 郑荣寿, 孙可欣, 张思维, 等. 2015 年中国恶性肿瘤流行情况分析, 中华肿瘤杂志, 2019, 41 (1): 19-28.
- [3] Downing NS, Aminawung JA, Shah ND, Braunstein JB, Krumholz HM, Ross JS. Regulatory review of novel therapeutics--comparison of three regulatory agencies. N Engl J Med. 2012. 366 (24) : 2284-93.
- [4] Carpenter D, Zucker EJ, Avorn J. Drug-review deadlines and safety problems. N Engl J Med. 2008. 358 (13) : 1354-61.
- [5] Greene JA, Podolsky SH. Reform, regulation, and pharmaceuticals--the Kefauver-Harris Amendments at 50. N Engl J Med. 2012. 367 (16) : 1481-3.
- [6] Nishino M, Giobbie-Hurder A, Gargano M, Suda M, Ramaiya NH, Hodi FS. Developing a common language for tumor response to immunotherapy: immune-related response criteria using unidimensional measurements. Clin Cancer Res. 2013. 19 (14) : 3936-43.

- [7] Seymour L, Bogaerts J, Perrone A, et al. iRECIST: guidelines for response criteria for use in trials testing immunotherapeutics. *Lancet Oncol.* 2017. 18 (3) : e143-e152.
- [8] Lin NU, Lee EQ, Aoyama H, et al. Response assessment criteria for brain metastases: proposal from the RANO group. *Lancet Oncol.* 2015. 16 (6) : e270-8.
- [9] Reck M, Rodríguez-Abreu D, Robinson AG, et al. Pembrolizumab versus Chemotherapy for PD-L1-Positive Non-Small-Cell Lung Cancer. *N Engl J Med.* 2016. 375 (19) : 1823-1833.
- [10] Khozin S, Weinstock C, Blumenthal GM, et al. Osimertinib for the treatment of metastatic EGFR T790M mutation-positive non-small cell lung cancer. *Clin Cancer Res.* 2017, 23 (9) : 2131-35.
- [11] Odogwu L, Mathieu L, Goldberg KB, et al. FDA benefit-risk assessment of osimertinib for the treatment of metastatic non-small cell lung cancer harboring epidermal growth factor receptor for T790M mutation. *Oncologist.* 2018, 23 (3) : 353-59.
- [12] Ellis LM, Bernstein DS, Voest EE, et al. American Society of Clinical Oncology perspective: Raising the bar for clinical trials by defining clinically meaningful outcomes. *J Clin Oncol.* 2014. 32 (12) : 1277-80.