arXiv:2301.04578v1 [stat.AP] 11 Jan 2023

扩大入组资格背景下的精确剂量探索癌症临床试验

BY REBECCA B. SILVA[1,a], BIN CHENG[1,b], RICHARD D. CARVAJAL[2,d] AND SHING M. LEE[1,c]

[1]哥伦比亚大学梅尔曼公共卫生学院生物统计学系，a*rs4025@cumc.columbia.edu*；b*bc2159@cumc.columbia.edu*；c*sml2114@cumc.columbia.edu*

[2]哥伦比亚大学赫伯特·欧文综合癌症中心，d*rdc2150@cumc.columbia.edu*

扩大癌症临床试验的入选标准，有助于更准确地反映真实的患者群体。尽管这在普适性和患者招募方面的优势显而易见，但仍需新的剂量探索设计来确保患者安全性。如果存在毒性特征不同的亚群，这些设计应能够为其推荐精确剂量。虽然已有考虑患者异质性的剂量探索设计，但现有方法都假设异质性的来源已知，因此会预先指定亚群，或仅允许纳入少数患者特征。我们提出一种精确剂量探索设计，用于应对入组资格扩大后，I期癌症临床试验中患者异质性来源未知的情形；此时，多项入选标准可能导致患者亚组适用不同的最佳剂量。该设计采用一举两得的方法进行剂量探索：通过边际模型逐步纳入患者协变量，同时筛选出能够区分最大耐受剂量（MTD）的患者标准，并在必要时推荐亚群特异性最大耐受剂量。我们的模拟研究将所提设计与假设患者同质性的朴素方法进行比较。结果表明，存在异质性时，我们的设计会推荐多个剂量；不存在异质性时，则推荐单一剂量。所提出的剂量探索设计应对了癌症临床试验扩大入选标准所带来的挑战，也满足了早期临床试验中对更精确剂量的需求。

1. 引言。剂量探索型I期癌症临床试验的目标，是为后续II期和III期临床试验选出一种新疗法的安全剂量。假设治疗的毒性和疗效会随剂量增加，推荐剂量通常被称为最大耐受剂量（MTD）。大多数I期试验都会推荐一个最大耐受剂量，并假设所有符合入组资格的患者对新疗法具有相似的耐受性。过去十年来，人们一直主张扩大癌症临床试验的入选标准；这类试验最初是为细胞毒性化疗设计的。研究发现，入选标准过于严格，且无法反映可能从疗法中获益的实际患者群体 [Fehrenbacker 等(2009); Gerber 等(2016); Malik 和 Lu(2019年); Liu 等(2021年); Harvey 等(2021年)], 从而限制了研究结果的普适性，并显著减缓了患者入组 [Malik 和 Lu(2019年)]。一项研究考察了连续确诊的326名非小细胞肺癌（NSCLC）患者的入选不合格率，结果发现，依据传统入选标准，约80%的患者不符合入选资格 [Fehrenbacker 等(2009)]。

多位作者发现，纳入脑转移、HIV或既往癌症患者并不会影响临床试验结果，且能更准确地反映实际患者群体

关键词和短语：亚群特异性剂量探索、亚组剂量探索、入组资格标准、患者异质性、序贯设计、I期癌症试验。

[George（1996）；Kim等（2015、2017）；Beaver等（2017）；Malik和Lu（2019）]。2017年，美国临床肿瘤学会（ASCO）、癌症研究之友和美国食品药品监督管理局发布了关于排除标准的新建议，指出应重新审视包括脑转移、HIV和艾滋病、器官功能障碍、既往或并发癌症以及最低入组年龄在内的标准 [Kim等（2017）；美国食品药品监督管理局（2020）]。此外，合并用药、既往治疗、体能状态、肝功能或肾功能障碍等入选标准，以及胆红素、血小板、血红蛋白和碱性磷酸酶等实验室指标，仍在由ASCO和癌症研究之友持续质疑和评估 [Harvey等（2021）；Liu等（2021）；Magnuson等（2021）；Osarogiagbon等（2021）]。

虽然扩大入选标准具有明显优势，但也必须谨慎实施，以确保患者安全。具体而言，剂量探索方法应考虑人群可能存在异质性的情况，即不同患者亚群可能具有不同的耐受性，并采用精准医学方法进行剂量估计，在必要时推荐亚群特异性最大耐受剂量。已有研究提出了适用于I期试验的方法，以考虑预先指定的患者协变量或亚群。这些方法要么假定感兴趣的协变量或具有不同毒性特征的亚群事先已知，要么根据相似风险自适应地合并亚组，而不识别协变量。因此，在扩大入选标准的情境下，由于事先无法确定是否需要采用不同的最大耐受剂量，这些方法无法应对相关挑战。持续再评估方法（CRM）的扩展 [Pi- antadosiandLiu (1996)] 以及带过量毒性控制的剂量递增方法，允许在剂量-毒性模型中加入患者特异性协变量 [Babb and Rogatko (2001); Cheng et al. (2004)]。此外，还有一些方法允许预先指定亚群，包括双样本CRM [O'Quigley, Shen, and Gamst (1999)], 、有序组的双样本CRM [O'Quigley and Paoletti (2003)],、针对多个有序风险组 [Yuan and Chappell (2004)] 及部分有序风险组 [Conaway (2017)]的保序设计。近期，Morita, Thall, and Takeda (2017)提出了一种采用亚组特异性截距的分层贝叶斯方法；Chapple and Thall (2018)则提出对亚组参数采用尖峰与厚尾先验，以处理多个组的情况。

虽然精准医学通常被倡导应用于药物开发的后期阶段，并且需要更大的数据集，但在药物开发的早期阶段，理想的做法是采用一种方法，筛查扩展的入组标准、评估是否有必要将其纳入考量，并在需要时能够推荐多个最大耐受剂量。Beaver 等人（2017年）提出，“采用逻辑合理的方法来界定入组资格，或许能够在采用广泛入选标准的早期临床试验中检测到安全性信号”。随着肿瘤学试验的入选标准不断扩展，采用能够学习并评估毒性中患者异质性信号的方法，将是确保那些历来被认为不良结局风险较高的患者安全的关键。

作为一个说明性例子，我们采用一项 I 期试验，该试验评估了间歇给予丝裂原活化蛋白激酶激酶（MEK）抑制剂司美替尼治疗葡萄膜黑色素瘤患者的方案 [Khan 等人（2022）]。该研究在 28 名患者中评估了司美替尼的六种剂量（100、125、150、175、200、225 mg），并采用事件发生时间CRM（TITE-CRM）估计最大耐受剂量，目标毒性水平为 25%。由于葡萄膜黑色素瘤罕见，在三个大型研究中心开展入组历时三年多。该临床试验的主要研究者指出了三项标准（肌酐水平、器官和骨髓功能，以及是否存在已知或疑似脑转移），放宽这些标准或许能让更多患者受益于该临床试验。例如，肌酐水平反映肾功能，而 12% 至 53% 的癌症患者患有慢性肾病 [Kitchlu 等人 （2018）]。如果扩大入组标准，这项 I 期试验的设计就需要考虑

这三项标准，因为目前对它们与毒性之间关联的了解有限。其中一些标准可能会影响司美替尼的毒性概率，另一些则可能不会。例如，肾功能损害与毒性风险增加有关，因此研究者在确定剂量时会考虑肾功能障碍的程度 [Lameire （2014）；Leal 等人 （2011）；LoRusso 等人 （2012）]。然而，研究发现，剂量降低并非总是必要的 [Leal 等人 （2011）]。如果设计未考虑放宽后的标准，且至少有一项标准与毒性风险增加相关，那么就可能高估这些患者的最大耐受剂量。

我们提出的设计——精确CRM（P-CRM）——适用于需要考察并了解多项患者特征的情形，这些特征共同构成了更广泛的患者群体。例如，如果司美替尼临床试验扩大了入选标准，我们就需要考虑评估三项患者标准，因为最大耐受剂量可能因其中一项或多项标准而异。我们的方法根据患者标准在控制剂量后的毒性边际相关性，依次选择纳入剂量-毒性模型的患者标准。每新增一个入组队列，就会利用有关毒性和患者特征的额外信息，评估各项标准是否应纳入或剔除。剂量分配依据当前数据和选定的患者标准进行；达到总样本量后，则使用最终模型中选定的协变量确定亚群特异性最大耐受剂量。该方法将剂量探索一分为二：既能了解各项标准及是否有必要区分最大耐受剂量，也能根据这些标准为每个患者亚群推荐剂量。

本文其余部分安排如下。第 2 节介绍精准剂量探索方法。第 3 节基于重新设计的司美替尼临床试验开展模拟研究，并比较 P-CRM 与不考虑任何患者异质性的朴素方法的表现。我们还讨论该设计的样本量考量。最后，第 4 节讨论所提设计的意义和局限性。

2. 方法。药物安全性通常通过首个治疗周期内是否发生严重不良事件（即毒性）来衡量，这种事件称为剂量限制性毒性（DLT）。在基于模型的设计中，会评估总计 J 个剂量，即 d_1、.....、d_J，并使用剂量-毒性模型估计发生毒性或 DLT 的概率。最大耐受剂量（MTD）是指毒性水平最接近 p_T 的剂量； p_T 是临床医生认为可接受的最高毒性发生率，也称为目标毒性水平。MTD 通过序贯方式进行估计。对于下文使用的所有符号，令 Y_i 表示第 i 位患者的二元毒性结局， X_i 表示分配给该患者的剂量。

2.1. 剂量-毒性模型。 传统上，通常假设整个患者群体在每个剂量水平下具有相同的毒性概率，因此具有相同的最大耐受剂量。假设采用逻辑斯蒂模型且患者完全同质， $\pi(d_j, \beta_0, \beta_1) = P(Y_i = 1|X_i = d_j)$ 可建模为

$$\text{logit}\{\pi(d_j, \beta_0, \beta_1)\} = \beta_0 + \beta_1 d_j, j = 1, \ldots, J. \tag{1}$$

20世纪90年代末，剂量探索模型中的患者同质性假设开始受到质疑。研究者提出了一种纳入患者协变量 z的剂量-毒性模型，其中$\pi(d_j, z_i, \boldsymbol{\theta}) = P(Y_i = 1|X_i = d_j, z_i)$ 可建模为，

$$\text{logit}\{\pi(d_j, z_i, \boldsymbol{\theta})\} = \beta_0 + \beta_1 d_j + \gamma z_i, \tag{2}$$

其中 $\boldsymbol{\theta} = (\beta_0, \beta_1, \gamma)^T$ [Piantadosi 和 Liu (1996年)]。方程（2）表示一个限制过强的模型；它假设我们事先知道是哪一个协变量导致了异质性，并且异质性完全由 z所致，这两项都是很强的假设。

本文考虑一组数量可能很大的患者协变量，$\boldsymbol{z}=(z_1, \ldots\ldots, z_M)^T$，这些协变量可能导致剂量异质性。我们将 $\pi(d_j, \boldsymbol{z}_i, \boldsymbol{\theta})=P(Y_i=1|X_i=d_j, \boldsymbol{z}_i)$ 建模为

$$\text{logit}\{\pi(d_j, \boldsymbol{z}_i, \boldsymbol{\theta})\} = \beta_0 + \beta_1 d_j + \sum_{m=1}^{M} \gamma_m z_{mi}, \tag{3}$$

其中 $\boldsymbol{\theta}=(\beta_0, \beta_1, \gamma_1, \ldots\ldots, \gamma_M)^T$. 此外，我们作出以下稀疏性假设

$$\sum_{m=1}^{M} 1_{\gamma_m \neq 0} = K \ll M. \tag{4}$$

在扩展的入组标准下，方程（3）表示一个更现实且可能具有高维度的模型，但我们假设只有其中一小部分标准会导致异质性。假设（4）符合临床医生的看法：只有少数扩展标准/协变量，或者根本没有任何此类标准/协变量，需要据此对最大耐受剂量进行区分。

因此，假设共有 M 个患者协变量可能与毒性相关，其中只有一小部分实际上能够区分最大耐受剂量，我们的目标有两个方面。一是确定导致最大耐受剂量异质性的 K 个协变量子集；二是实施一种能够确定亚群特异性最大耐受剂量的程序。

2.2. 精确剂量探索程序。令 $\Gamma=\{\gamma_i, i=1, \ldots, M\}$，且 $\Gamma_0=\{\gamma_i: \gamma_i \neq 0\}$。根据方程（3），$\Gamma_0$ 是 Γ 的一个未知小子集，必须对其进行估计，因为它会导致最大耐受剂量的异质性。估计 Γ 面临巨大挑战。由于 K 未知，因此必须进行估计，我们需要在 2^M 的所有子集中搜索 Γ，以找到 Γ_0。此外，在剂量探索研究中，患者按序贯队列分配剂量，而不像平行组设计那样同时分配至所有剂量。这意味着，随着毒性数据不断积累，模型（3）中参数的相关信息也会逐步累积，因此试图一次性估计Γ_0 未免过于雄心勃勃。

与剂量探索研究的序贯性质相一致，我们提出对 Γ_0进行序贯估计。我们提出的设计称为精准CRM（P-CRM），由两个阶段组成。第一阶段收集患者特征信息，但剂量分配并非针对个体患者进行。第二阶段，在获得足够信息后，该方法会在考虑剂量水平的基础上，根据患者协变量与毒性的边际相关性，依序选择患者协变量并将其纳入剂量-毒性模型。每新增一个入组队列，结合毒性和患者特征方面的新增信息，都会检验所选协变量是否应予以剔除。请注意，这些是入组队列，而非剂量队列，因此队列中的患者会根据各自的患者特征被分配到不同的剂量水平。随后，剂量分配将依据当前数据和所选患者协变量进行。达到总样本量后，将使用最终模型中选定的协变量确定协变量特异性最大耐受剂量。以下将更详细地介绍P-CRM的剂量探索阶段。

第一阶段：对于最初的 N_1 位患者，采用基于工作模型（1）的常规单样本CRM方法 [O'Quigley、Pepe和Fisher（1990年）]。收集协变量信息，但直到第二阶段才使用。骨架可由医生提供，或在指定目标毒性水平、p_T、队列规模和MTD初始猜测后，通过Lee和Cheung（2009）的方法进行校准。第 j个剂量的剂量标签（或标准化剂量）$\tilde{d_j}$ 通过方程logit$(p_{0j}) = \beta_0 + \beta_1 d_j$确定，其中 β_0 和 β_1 是参数的先验猜测

和 β_0 分别是 β_1 模型（1）中的参数，p_{0j} 是骨架，$j=1$、……、J。根据 N_1 位患者的毒性后验概率，使用新的骨架更新剂量标签。

第二阶段：在为 N_1 名患者分配剂量并观察到剂量和毒性后，每入组一个队列后，依次评估患者协变量是否纳入。

1. 在此阶段，患者按每组 L 人的队列入组。

2. 首个队列入组后，我们开始筛选协变量。对于每个协变量z_m，$m=1$，……，M，我们拟合以下工作模型

$$\text{logit}\{\pi(d_j, z_{m,i}, \boldsymbol{\theta_m})\} = \beta_{0m} + \beta_{1m} d_j + \gamma_m z_{m,i} \tag{5}$$

并将与协变量 z_m 相关的p值记为 $p_1(z_m)$。令 $m_1 = \text{argmin}_m\, p_1(z_m)$。如果对于某个预先指定的 $\alpha_1^* > 0$，$p_1(z_{m_1}) \geq \alpha_1^*$，则不纳入任何协变量，并按第一阶段的方式分配剂量。如果 $p_1(z_{m_1}) < \alpha_1^*$，则选中协变量 z_{m_1}，并将第 $(n+1)$ 个患者队列分配至最接近目标毒性水平p_T的剂量：

$$x_{[n+1]} = \text{argmin}_{x\in\{d_1,\ldots,d_J\}} |\pi(x; \boldsymbol{z_{m_1}}, n+1, \hat{\boldsymbol{\theta}}_{\boldsymbol{m_1},n}) - p_T|. \tag{6}$$

3. 收集到新入组队列的数据后，我们继续选择下一个协变量，并考虑移除已选中的协变量。对于剩余的每个协变量 $z_m, m \neq m_1$，我们拟合工作模型（5）。令 $p_2(z_m)$ 为与协变量 $z_m, m \neq m_1$相关的p值，并令 $m_2 = \arg\min^{m\neq m_1} p_2(z_m)$。当且仅当对于某个预先指定的 α_2^*，z_{m_2} 满足 $p_2(z_{m_2}) < \alpha_2^*$ 时，才选择协变量 z_{m_2}。

同样，给定已选中的 q 个协变量 $z_{m_1},\ldots,z_{m_q}$，要选择第 $(q+1)$ 个协变量，则在模型（5）下逐一检验剩余的 $M-q$ 个协变量，其中$m \notin \{m_1,\ldots,m_q\}$。令 $p_{q+1}(z_m)$ 为模型（5）中与协变量 z_m 相关的p值，并令 $m_{q+1} = \text{argmin}^{m\notin\{m_1,\ldots,m_q\}} p_{q+1}(z_m)$。那么，当且仅当对于某个 $\alpha_{q+1}^* > 0$，$z_{m_{q+1}}$ 满足 $p_{q+1}(z_{m_{q+1}}) < \alpha_{q+1}^*$ 时，才选择协变量 $z_{m_{q+1}}$。

4. 为评估是否应移除选定的协变量，假设已选定 q协变量 $z_{m_1},\ldots,z_{m_q}$，我们拟合一个工作模型

$$\text{logit}\{\pi(d_j, z_{m,i}, \boldsymbol{\theta_m})\} = \beta_{0m} + \beta_{1m} d_j + \sum_{l=1}^{q} \gamma_{m_l} z_{m_l,i}. \tag{7}$$

令 $m_q^* = \arg\max_m\{p_q(z_{m_1})\ldots,p_q(z_{m_q})\}$。若协变量 $z_{m_q^*}$ 满足 $p_q(z_{m_q^*}) > \alpha_q^{**}$，则将其移除

5. 给定 q 尚未移除的剩余协变量，根据工作模型（7），按照方程（6）进行剂量分配。若移除后不再有剩余协变量，则按照第一阶段进行剂量分配。

6. 重复步骤 3 至 5，直到以这种方式为 $N_{\max}$ 患者分配了推荐剂量。

达到 $N_{\max}$ 后，按如下方式估计最大耐受剂量：

如果没有选定的协变量剩余，则使用单样本CRM估计最大耐受剂量，并为所有患者推荐一个统一的最大耐受剂量。如果模型中仍有 $k>0$ 协变量 $z_{m_1},\ldots,z_{m_k}$，令 $\boldsymbol{z}_k^* = (z_{m_1},\ldots,z_{m_k})^T$，则最大耐受剂量为

$$\text{MTD}_{\boldsymbol{z}_k^*} = \text{argmin}_{x\in\{d_1,\ldots,d_J\}} |\pi(x; \boldsymbol{z}_k^*, \hat{\boldsymbol{\theta}}_{\boldsymbol{m},N_{\max}}) - p_T|, \tag{8}$$

其中 $\pi(x; \boldsymbol{z}_k^*, \hat{\boldsymbol{\theta}}_{\boldsymbol{m},N_{\max}})$ 来自最终模型

$$\text{logit}\{\pi(d_j, \boldsymbol{z}_{k,i}^*, \boldsymbol{\theta_m})\} = \beta_{0m} + \beta_{1m} d_j + \sum_{l=1}^{k} \gamma_{m_l} z_{m_l,i}. \tag{9}$$

关于队列规模的选择， L，我们建议根据总样本量和所考虑的标准数量，使用三至五名患者， M。纳入阈值 α_q^*和排除阈值 $\alpha_q^{**}, q = 1, \ldots, M$, 应取决于为新队列拟合的边际模型数量。如果 M 是临床试验中正在检验的协变量总数，而 q 是模型中已选出的协变量数量，我们建议使用 $\alpha_q^* = \alpha(M - q)/M$ 作为纳入阈值，使用 $\alpha_q^{**} = \alpha/q$ 作为排除阈值，其中 $\alpha = 0.20$。我们建议的纳入阈值 α_q^* 随 q增加而降低，表明纳入更多协变量时将施加更大的惩罚，这与稀疏性假设一致（4）。

3. 数值研究。

3.1. 模拟设定：重新设计司美替尼研究。我们进行了模拟研究，以考察 P-CRM 剂量探索设计的运行特征。我们重新设计了司美替尼临床试验，纳入三项扩展后的患者标准，并评估是否需要亚群特异性最大耐受剂量。我们将该设计与单样本CRM进行比较，以展示在患者异质性确实存在时忽略这种异质性会如何影响最大耐受剂量的推荐。

我们测试了三项二元患者标准（$M = 3$），并考察了六个剂量水平（$J = 6$）下剂量-毒性关系的五种情景。对于每种情景，我们分别考虑每项标准的患病率为50%和25%， $P(z_m = 1) = 0.50,\ 0.25$，适用于 $m = 1,\ 2,\ 3$。各情景的剂量-毒性关系见图1 ，每个剂量水平下发生剂量限制性毒性的确切概率见附录。这些情景均非根据前文所述的剂量-毒性模型生成。

这些情景包括以下情况：三个入选标准中有零个或一个与毒性密切相关，因此分别形成一个或两个剂量-毒性关系不同的亚组。在情景1至4中， z_2 与毒性相关；对于符合放宽后的标准 $z_2 = 1$的新符合入选条件的患者，其最大耐受剂量不同于符合最初发表的标准 $z_2 = 0$的患者的最大耐受剂量。在场景 1和场景 2中，亚群的最大耐受剂量相差一个剂量水平；但在场景 1中，较低的最大耐受剂量为剂量水平一，而在场景 2中，较低的最大耐受剂量为剂量水平二。这些情景用于考察：当最大耐受剂量位于剂量水平范围的边界上而非范围内部时，该方法的表现如何。在情景3和4中，亚群的最大耐受剂量分别相差两个和三个剂量水平，这代表亚群的异质性逐渐增大，也意味着区分最大耐受剂量的必要性更高。最后，在场景 5中，没有任何标准与毒性相关，因此不存在异质性，所有患者的最大耐受剂量均相同。

在 P-CRM 的第一阶段和单样本CRM中，我们按照原始试验的设定，从剂量水平二开始进行 CRM，并将目标毒性水平 p_T 设为0.25。我们使用逻辑斯蒂剂量-毒性模型，固定截距为 3，并假设剂量协变量 β 服从均值为0 、方差为1.34 的正态先验。对剂量-毒性模型进行校准，使其选择的剂量对应的剂量限制性毒性发生率介于 17% 和 33% 之间 [Lee和Cheung (2009)]。队列规模为三。在 P-CRM 中，我们设置 $N_1 = 15$ ；在 N_1 名患者通过 CRM 分配剂量后，将更新后的标准化剂量设为 $d_j = \log(\frac{p_{0j}^*}{1-p_{0j}^*}) - 3;\ j = 1, \ldots, J$, 其中 p_{0j}^* 是在假设剂量效应系数为 1、截距为3时，剂量 j 下的毒性后验概率。在第二阶段，我们在每个边际逻辑斯蒂回归模型中使用固定截距3 ，队列规模为三。尽管每次增加三名患者，但剂量分配取决于各自的协变量模式，以及剂量寻找算法选出的协变量。此外，我们不允许跳过尚未分配给患者的剂量。我们对司美替尼临床试验进行了 2,000 次模拟，设置了四种不同的总样本量： $N_{\max} = 30,45,60,72$， 以评估样本量的影响。

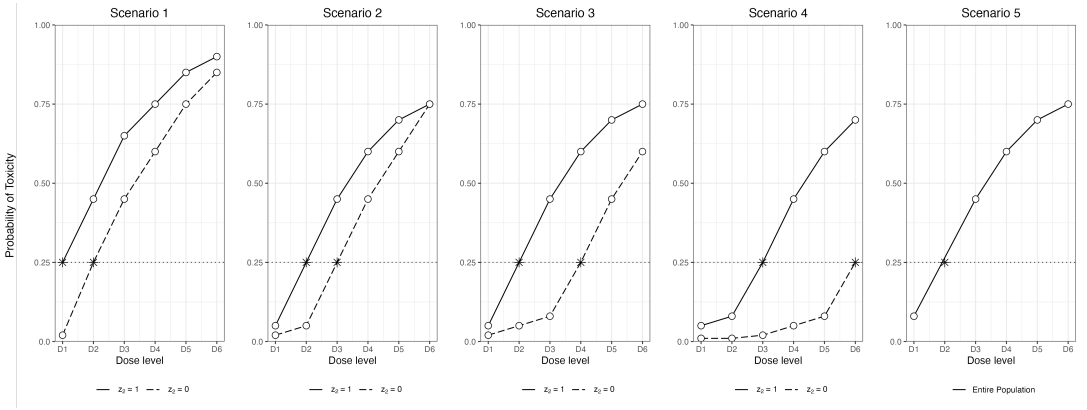


图1 1。情景*1-5*的剂量-毒性关系。情景*1*、*2*、*3*和*4*各有一个与毒性相关的标准，按一个、两个或三个剂量水平区分最大耐受剂量。场景 *5*中没有与毒性相关的标准，因此人群只有一个最大耐受剂量。水平虚线表示目标毒性水平*0.25*。星形点表示协变量特异性最大耐受剂量。

对于每种情景和标准患病率，我们关注三项性能指标：协变量正确选择概率、最大耐受剂量正确选择概率（PCS）以及最大耐受剂量加权选择概率（WPS） [Morita等人(2017)]。协变量正确选择概率定义为：最终模型（公式9）中保留的 k 个协变量确为与毒性相关标准的频率。对于情景1至4，正确的最终模型除剂量 z_2 外还包含一个协变量；对于场景 5，则不包含其他协变量。PCS 的计算方法是：基于最终模型中选定的协变量，在 $N_{\max} = 30,45,60,72$ 名观察患者中估计最大耐受剂量。对于每种观察到的患者协变量模式，我们使用剂量探索程序得到的最终模型，估计各剂量下协变量特异性的毒性概率，并按照第2节所述估计协变量特异性最大耐受剂量。如果未选出任何协变量，则使用单样本CRM估计最大耐受剂量。为计算加权选择概率，我们采用Morita等人(2017)定义的指标，该指标通过一种权重考虑接近最大耐受剂量的剂量选择情况；该权重由所选剂量下的毒性概率与目标毒性水平之间的距离确定。设 $r_{j,k}$ 为第 k 个亚组在剂量水平 j、 $j = 1,\ldots,6$ 下真实的剂量限制性毒性概率。亚组根据真实最大耐受剂量定义，如图 1 所示；对于情景1至4， $k = 1,2$ ；对于场景 5， $k = 1$ 。亚组按真实最大耐受剂量递增排序， p_T 为目标毒性水平。随后，亚组 k 的 WPS（WPS_k ）定义为

$$\text{WPS}_k = \sum_{j=1}^{J} w_{j,k} P(d_j \text{ is selected as the MTD for subgroup } k),$$

其中

$$w_{j,k} = \frac{(\max_{j'=1,\ldots,J} |r_{j',k} - p_T|) - |r_{j,k} - p_T|}{(\max_{j'=1,\ldots,J} |r_{j',k} - p_T|) - (\min_{j'=1,\ldots,J} |r_{j',k} - p_T|)}.$$

WPS 对最不理想的剂量赋予零权重，而对真实毒性率更接近目标毒性水平的剂量赋予相对较高的权重；PCS 则仅对真实的最大耐受剂量赋予权重，对其他所有剂量均不赋予权重。我们将 P-CRM 的亚组 PCS 和 WPS 与单样本 CRM 的亚组 PCS 和 WPS 进行了比较。

表1 总计三个协变量时的标准选择概率，*M = 3*，每个协变量的患病率分别为 *50%* 和 *25%*，对应于 $N\max = 45$

患病率 = 50%	情景	无标准	正确标准	正确标准及其他标准	错误标准
	1	30	**48**	6	16
	2	30	**44**	6	19
	3	11	**68**	10	11
	4	6	**73**	14	7
	5	**56**	NA	NA	44
患病率 = 25%	情景	无标准	正确标准	正确标准及其他标准	错误标准
	1	38	**43**	4	16
	2	39	**41**	3	16
	3	15	**67**	8	10
	4	7	**79**	10	4
	5	**63**	NA	NA	37

* 由于四舍五入，选择概率之和可能不等于 100。

3.2. 模拟结果。 当每项标准的患病率分别为 50% 和 25% 时，P-CRM 的标准选择概率见表 1 （针对 $N_{\max} = 45$）。每项标准的患病率为 50% 和 25% 时，P-CRM 和单样本 CRM 在不同总样本量下的 PCS 和 WPS 分别见图 2 和3。关于 $N_{\max} = 60$ 的标准选择的其他模拟结果，72 以及患病率为 50% 时 P-CRM 和单样本 CRM 下各剂量的选择概率，见附录。

表1 给出了当 $P(z_m = 1) = 0.50,0.25, m$ =1,2,3 和 $N_{\max} = 45$时的标准选择概率。正确标准的选择概率以粗体显示；在情景1至4中，应仅选择一个标准，即 z_2；在场景 5 中，则不应选择任何标准。亚群之间的剂量-毒性关系异质性越高，P-CRM选择标准的概率越高。例如，当患病率为 50% 且最大耐受剂量相差至少两个剂量水平时，该设计选择真实标准 z_2 的概率较高，约为 70%；但在更具挑战性的情景1和2中，最大耐受剂量仅相差一个剂量水平，该方法选择 z_2 的概率为 40% 至 50%。在这些情况下，如果该方法未能正确选择真实标准，通常会选择不设任何标准，采用朴素方法，即在剂量-毒性模型中不纳入任何协变量，并假设人群是同质的。在场景 5 中，当没有标准与毒性相关时，该方法约有 56% 的概率不会选择任何标准。当每个标准的患病率为 25% 时，正确标准的选择比例基本相似；情景1至3略有下降，平均下降 3%。当样本量增加至 $N_{\max} = 60$ 和 $N_{\max} = 72$时，正确标准的选择概率会提高，尤其是在异质性较小的情况下，详见附录。

鉴于该设计未预先假设哪些标准与毒性相关，因此对于最佳剂量相差超过一个剂量水平的异质性人群，该设计仍能以理想的比率选出正确标准。除场景 5 中人群同质的情况外，该设计对错误标准选择或假阳性率（FPR）的概率进行了控制。如果研究者希望降低假阳性率，可以考虑对 α 采用更保守的调整，但这样一来，该设计筛查患者异质性的可能性也会降低。当患病率减半时，该方法在选择标准方面的表现几乎与标准患病率为 50% 时一样好，这表明只有当标准极为罕见时，正确选择率才会大幅下降。

图2 （A 和 B）分别比较了在每项标准患病率均为 50% 时，P-CRM 与单样本CRM的亚群特异性 PCS 和 WPS，见$N_{\max} = 30$，45，60，72。随着亚组异质性增大以及总样本量

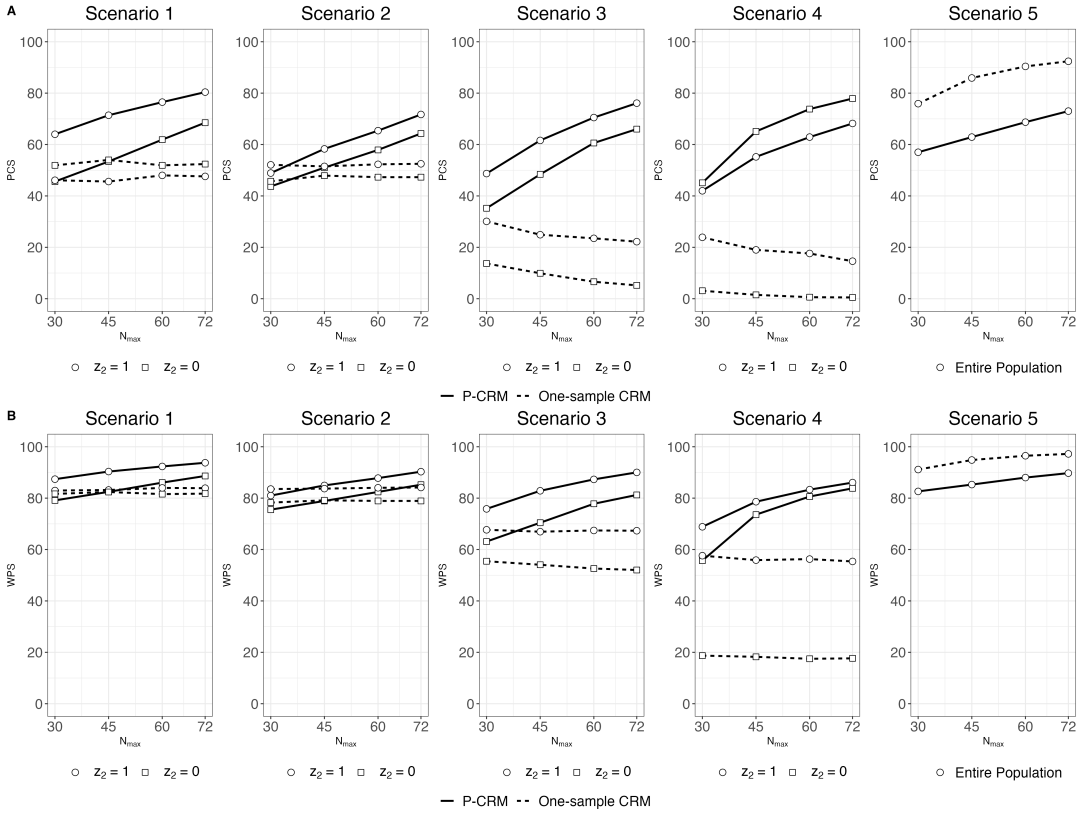


图 2。*A*）*P-CRM*和单样本*CRM*在 N最大 $=30$、45、60、72 条件下，对标准患病率为*50%*的各亚组的*MTD*正确选择比例（*PCS*）；*B*）加权选择概率（*WPS*）。

增加，P-CRM 的 PCS 和 WPS 均显著优于单样本CRM。单样本CRM的表现表明，当存在多个真实最大耐受剂量时，忽略患者标准会导致至少一个亚组、且往往两个亚组的给药剂量选择错误。当亚组间真实最大耐受剂量相差至少两个剂量水平时，单样本CRM最常给一个亚组分配可能的亚治疗剂量、给另一个亚组分配不安全剂量，从而错误识别各自的真实最大耐受剂量；而 P-CRM 最常分配的是亚组的真实最大耐受剂量。对于异质性更高的人群，P-CRM 下的 PCS 增幅在 $N_{\max}$ 后更大。在存在异质性的情景中，到 $N_{\max}=60$ 时平均 PCS 超过 60%。对于同质人群，到 $N_{\max}=30$ 时 PCS 接近 60%。所有情景下较高的 WPS 表明，P-CRM 很少分配与真实最大耐受剂量相差超过一个剂量水平的剂量。附录表3和表4列出了标准患病率为 50% 时，P-CRM 和单样本CRM在各剂量水平上的详细选择率；结果显示，当最大耐受剂量判断有误时，最大耐受剂量较低的亚组更可能被推荐使用低于未纳入扩展入组资格标准的亚组所获推荐的剂量。

图 3 （A 和 B）分别比较了在每项标准的患病率均为 25% 的情况下，P-CRM 与单样本CRM的PCS和WPS，适用于 $N_{\max}=30$、45、60、72。与患病率为 50% 时相比，满足扩大后入选标准的亚组（$z_i=1$、$i=1$、2）的PCS平均降低了 14%。然而，尽管估算该亚组的精确亚群特异性最大耐受剂量更具挑战性，该设计更有可能为耐受性较差的亚组分配较低剂量，因此推荐更接近真实最大耐受剂量的剂量，这一点可从其WPS仍优于单样本CRM看出

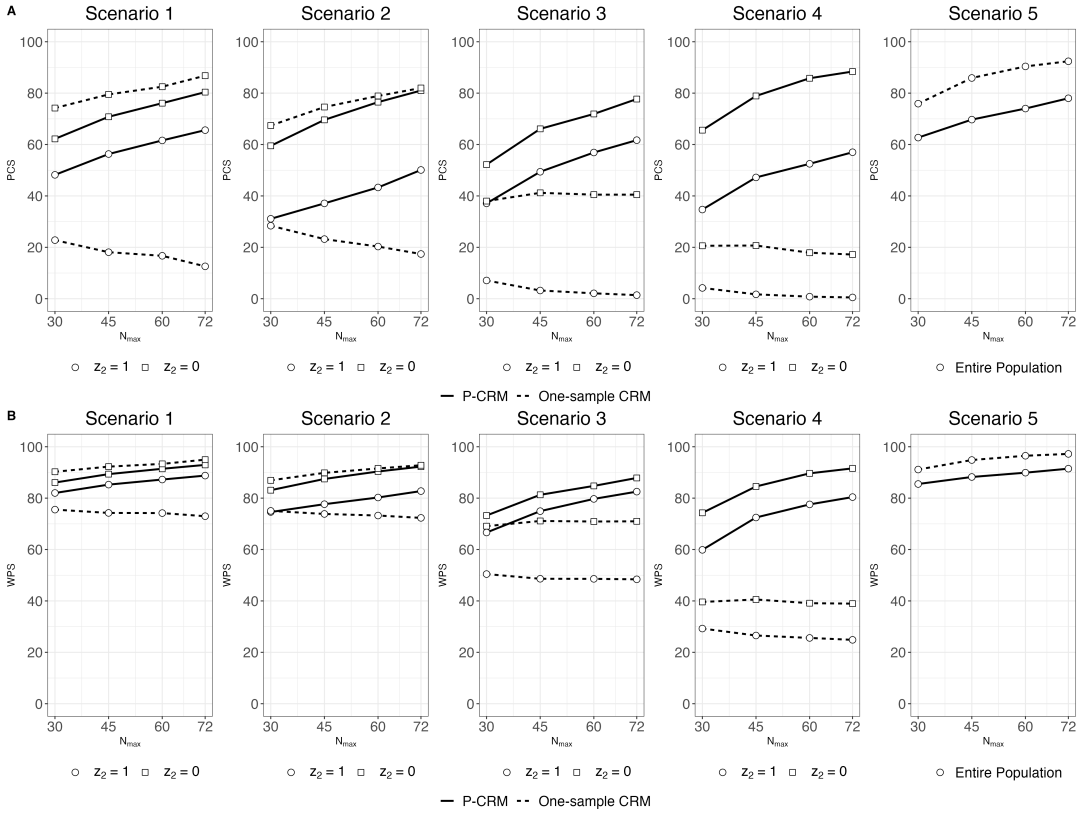


图3 。*A*）*P-CRM* 和单样本 *CRM* 在 N最多 $=30$,45、60、72 种标准患病率为 *25%* 的情况下，各亚组的 *MTD*正确选择比例（*PCS*）和 *B*）加权选择概率（*WPS*）。

的。在标准患病率为 25% 时，采用单样本CRM会完全忽略满足 $z_2 = 1$ 的亚组；只要亚组间的最大耐受剂量相差超过一个剂量水平，无论异质性如何，PCS都接近 0%，且WPS会随着异质性增加而下降。

总体而言，当有一个标准会使患者群体的最佳剂量出现差异，且各亚组最大耐受剂量处的毒性概率至少相差 20% 时，P-CRM 的表现良好。对于异质性更强的患者群体，该设计更常能够识别出真实标准。就剂量分配而言，我们发现，即使该设计未能正确识别真实最大耐受剂量，尤其是在标准患病率较低时，它仍倾向于推荐接近真实最大耐受剂量的剂量，而不是为患者分配与患病率较高亚组所推荐的相同剂量。最后，单样本 CRM 表现不佳，凸显了假设患者同质性并忽略患者标准会导致剂量建议毒性过高或治疗剂量不足。

3.3. 样本量考量。 该设计所需的总样本量取决于所考虑的入选标准数量（ M）、需要区分最大耐受剂量的标准数量（ K）、剂量水平数量（ J）以及这些标准的患病率。在 $J = 6$、 $M = 3$、标准患病率为50%，且只有一个协变量与毒性相关的情况下，若要使各情景的平均正确选择概率至少达到50%，样本量至少需要45。当三个协变量中有一个以上与毒性相关，导致出现三个或更多具有不同最大耐受剂量的亚组时，所需样本

量约为60（数据未显示）。患病率降低时，总体正确选择概率会提高，但患病率较低亚组的正确选择概率会下降，因此需要更大的样本量，才能使亚组正确选择概率至少达到50%。在相同情景下，若只考虑两个标准（ $M = 2$），则样本量约为35即可使平均正确选择概率至少达到50%（数据未显示）。若只考虑一个标准，也应考虑其他现有方法。也就是说，该方法的独特优势在于可以同时考虑多个可能的亚群，从而减少寻找一个或多个最大耐受剂量所需的样本量。尽管了解异质性需要增加样本量，但若使用单样本CRM，针对每个潜在亚组分别开展剂量探索临床试验，则需要大得多的样本量，因为可能的患者亚组很多（2^M），其中一些亚组可能因患病率较低而样本量不足。

4. 讨论。本文提出的精确剂量探索设计 P-CRM，旨在患者群体可能存在异质性、但尚不清楚哪些患者协变量确实会影响毒性，因而必须在最大耐受剂量确定中加以考虑时，进行剂量探索。随着癌症临床试验的入选标准不断扩大，新符合入选条件的患者可能具有不同的安全性特征，因为许多患者最初被纳入时是出于安全性考虑；因此，需要一种能够考虑入选标准的方法来确保患者安全。针对早期试验中如何应对入选标准扩大的建议之一，是在初始传统剂量探索临床试验后采用扩展队列 [Malik 和 Lu(2017年);Kim 等人(2017年)]。然而，潜在队列的数量可能很多，而且哪些具体队列应纳入尚不明确。P-CRM 的第二阶段与扩展队列类似，能够考虑潜在亚群；但 P-CRM 还能够通过持续的剂量探索程序，准确调整这些患者群体的剂量，并进一步了解哪些亚群确实存在异质性。

P-CRM 使用边际逻辑回归模型来确定每个协变量是否与毒性相关。如果协变量与毒性表现出显著关联，则会依次将其纳入剂量-毒性模型；在获得每个队列的患者数据后，并在控制其他已选协变量的情况下，再评估是否将其移除。剂量分配和最大耐受剂量推荐基于剂量以及模型中选定的任何患者协变量。当未加入任何患者协变量时，该设计会恢复为原始 CRM，并假设患者同质。该设计无需假设哪些协变量能够区分患者群体，并且可以同时筛选多个潜在协变量。考虑到我们的模型可能具有较高维度且带有稀疏性约束，一种自然的方法是最小绝对收缩和选择算子（LASSO）。然而，在样本量较小且数据稀疏的剂量探索场景中使用 LASSO，会导致协变量选择的假阳性率较高，并使最大耐受剂量正确选择概率整体降低。

我们的模拟研究发现，该设计能够正确识别与毒性相关的协变量，同时控制假阳性率。尽管当患者异质性较大时，该设计表现最佳，但我们发现，在任何患者异质性情景下，P-CRM仍优于忽略患者异质性的朴素方法，而且区分最大耐受剂量是必要的。当亚群异质性较低导致正确选择协变量的概率不那么理想时，由于此时真实最大耐受剂量彼此接近，选择理想剂量的概率仍然很高。此外，P-CRM对占比较低的亚组的真实最大耐受剂量选择得不够准确，但所选剂量比占比较高亚组的真实最大耐受剂量更接近该亚组各自的真实最大耐受剂量。最后，模拟研究表明，忽略患者特征可能会导致至少一个、甚至所有患者群体亚群出现毒性过高或治疗剂量不足的剂量建议，但

可能影响患者群体中的所有亚群。因此，对于异质性人群，忽略患者特征的后果远比针对同质人群纳入患者特征严重。

由于我们的设计是首个考虑扩展的入组标准且假设对异质性一无所知的I期设计，因此我们不将其与其他考虑已知患者异质性的方法进行比较。例如，O'Quigley等人（1999年）提出的双样本CRM，或Morita等人（2017年）提出的分层贝叶斯CRM，都假设患者亚群已知。P-CRM必须先识别亚群，然后再为每个亚群确定剂量。

与传统I期试验相比，为了纳入更多患者标准而扩大样本量是必要的；但即使典型的I期试验样本量为30，考虑异质性的优势也已十分明显。此外，随着入组资格的扩大，可招募入组并从安全剂量中获益的患者也会更多。因此，在实践中应用该方法时，必须权衡希望评估的标准数量、预期的亚组数量、样本量的增加以及符合这些标准的患者比例。正在开展的工作包括纳入疗效，以便同时考虑毒性和疗效的异质性，从而确定最佳剂量。精准医学方法主要用于晚期试验和大样本量场景，而我们的方法则为样本量有限的早期试验提供了一种精准剂量探索方法。

所提出的设计 P-CRM 能够筛选出与毒性相关的患者特征，并以简单的方式确定亚群特异性最大耐受剂量。该设计只需使用逻辑回归，因此在这种情况下纳入精准医学无需设定多个先验分布。更精准地确定最大耐受剂量并识别患者异质性的来源，从长远来看会带来优势，有助于提高II期和III期试验的成功率，并可直接应用于真实的患者群体。

附录

表格 2情景*1-5*中各剂量的毒性概率。

情景	亚群	D1	D2	D3	D4	D5	D6
1	$z_2 = 1$	**0.25**	0.45	0.60	0.75	0.85	0.90
	$z_2 = 0$	0.02	**0.25**	0.45	0.60	0.75	0.85
2	$z_2 = 1$	0.05	**0.25**	0.45	0.60	0.75	0.85
	$z_2 = 0$	0.02	0.05	**0.25**	0.45	0.60	0.75
3	$z_2 = 1$	0.05	**0.25**	0.45	0.60	0.75	0.85
	$z_2 = 0$	0.02	0.05	0.08	**0.25**	0.45	0.60
4	$z_2 = 1$	0.05	0.08	**0.25**	0.45	0.60	0.70
	$z_2 = 0$	0.01	0.01	0.02	0.05	0.08	**0.25**
5	总体人群	0.08	**0.25**	0.45	0.60	0.70	0.75

表3 三个协变量总数下的标准选择概率，*M = 3*，每个协变量的患病率分别为 *50%* 和 *25%*，适用于 $N_{\text{最大值}} = 60$

患病率 = 50%	情景	无标准	正确标准	正确标准及其他标准	错误标准
	1	21	**57**	8	14
	2	24	**52**	7	16
	3	4	**76**	14	6
	4	3	**76**	19	2
	5	**56**	NA	NA	44
患病率 = 25%	情景	无标准	正确标准	正确标准及其他标准	错误标准
	1	29	**50**	6	16
	2	31	**48**	5	16
	3	8	**73**	11	8
	4	2	**82**	14	2
	5	**58**	NA	NA	42

* 选择概率之和可能因四舍五入而不等于 100。

表4：在协变量总数为 *3* 的情况下，标准选择概率，$M = 3$，每个协变量的患病率分别为 *50%* 和 *25%*，适用于 $N\max = 72$

患病率 = 50%	情景	无标准	正确标准	正确标准及其他标准	错误标准
	1	18	**64**	8	10
	2	21	**58**	9	13
	3	2	**78**	16	4
	4	2	**77**	20	1
	5	**56**	NA	NA	44
患病率 = 25%	情景	无标准	正确标准	正确标准及其他标准	错误标准
	1	26	**54**	6	14
	2	27	**52**	6	14
	3	6	**76**	13	5
	4	1	**83**	15	1
	5	**55**	NA	NA	45

* 选择概率之和可能因四舍五入而不等于 100。

表5在标准患病率为 *50%* 时，*P-CRM* 下每个 N最大剂量选择的比例、正确选择概率（*PCS*）和 *WPS*。

P-CRM	情景	$N_{\max}$	真实最大耐受剂量	D1	D2	D3	D4	D5	D6	PCS	WPS
	场景 1	30	1	**0.64**	0.30	0.05	0.01	0.00	0.00	0.64	0.87
		30	2	0.27	**0.46**	0.22	0.04	0.01	0.00	0.46	0.79
		45	1	**0.71**	0.25	0.03	0.00	0.00	0.00	0.71	0.90
		45	2	0.20	**0.53**	0.23	0.03	0.00	0.00	0.53	0.82
		60	1	**0.76**	0.22	0.02	0.00	0.00	0.00	0.76	0.92
		60	2	0.16	**0.62**	0.21	0.02	0.00	0.00	0.62	0.86
		72	1	**0.80**	0.19	0.01	0.00	0.00	0.00	0.80	0.94
		72	2	0.12	**0.69**	0.18	0.01	0.00	0.00	0.69	0.89
	场景 2	30	2	0.13	**0.49**	0.32	0.05	0.01	0.00	0.49	0.81
		30	3	0.03	0.28	**0.44**	0.20	0.05	0.00	0.44	0.76
		45	2	0.13	**0.58**	0.25	0.04	0.00	0.00	0.58	0.85
		45	3	0.01	0.22	**0.51**	0.22	0.04	0.00	0.51	0.79
		60	2	0.10	**0.65**	0.22	0.02	0.00	0.00	0.65	0.88
		60	3	0.00	0.19	**0.58**	0.20	0.02	0.00	0.58	0.82
		72	2	0.08	**0.72**	0.19	0.01	0.00	0.00	0.72	0.90
		72	3	0.00	0.17	**0.64**	0.17	0.01	0.00	0.64	0.85
	场景 3	30	2	0.10	**0.49**	0.31	0.08	0.02	0.00	0.49	0.76
		30	4	0.01	0.13	0.28	**0.35**	0.18	0.05	0.35	0.63
		45	2	0.11	**0.62**	0.22	0.04	0.01	0.00	0.62	0.83
		45	4	0.00	0.06	0.20	**0.48**	0.21	0.04	0.48	0.70
		60	2	0.09	**0.70**	0.17	0.03	0.00	0.00	0.70	0.87
		60	4	0.00	0.03	0.16	**0.61**	0.17	0.03	0.61	0.78
		72	2	0.07	**0.76**	0.15	0.01	0.00	0.00	0.76	0.90
		72	4	0.00	0.02	0.14	**0.66**	0.17	0.01	0.66	0.81
	场景 4	30	3	0.01	0.10	**0.42**	0.32	0.11	0.04	0.42	0.69
		30	6	0.00	0.01	0.11	0.18	0.24	**0.45**	0.45	0.56
		45	3	0.00	0.12	**0.55**	0.26	0.05	0.01	0.55	0.79
		45	6	0.00	0.00	0.03	0.07	0.24	**0.65**	0.65	0.74
		60	3	0.00	0.11	**0.63**	0.23	0.02	0.00	0.63	0.83
		60	6	0.00	0.00	0.01	0.04	0.21	**0.74**	0.74	0.81
		72	3	0.00	0.10	**0.68**	0.20	0.01	0.00	0.68	0.86
		72	6	0.00	0.00	0.01	0.03	0.19	**0.78**	0.78	0.84
	场景 5	30	2	0.20	**0.57**	0.19	0.04	0.00	0.00	0.57	0.83
		45	2	0.17	**0.63**	0.17	0.02	0.00	0.00	0.63	0.85
		60	2	0.15	**0.69**	0.15	0.01	0.00	0.00	0.69	0.88
		72	2	0.12	**0.73**	0.14	0.01	0.00	0.00	0.73	0.90

表6 当标准患病率为 *50%* 时，单样本*CRM*下每个 N最大值 对应的剂量选择比例、正确选择概率和 *WPS*

单样本CRM	情景	$N_{\max}$	真实最大耐受剂量	D1	D2	D3	D4	D5	D6	PCS	WPS
	场景 1	30	1	**0.46**	0.52	0.02	0.00	0.00	0.00	0.46	0.83
		30	2	0.46	**0.52**	0.02	0.00	0.00	0.00	0.52	0.82
		45	1	**0.46**	0.54	0.00	0.00	0.00	0.00	0.46	0.83
		45	2	0.46	**0.54**	0.00	0.00	0.00	0.00	0.54	0.82
		60	1	**0.48**	0.52	0.00	0.00	0.00	0.00	0.48	0.84
		60	2	0.48	**0.52**	0.00	0.00	0.00	0.00	0.52	0.82
		72	1	**0.48**	0.52	0.00	0.00	0.00	0.00	0.48	0.84
		72	2	0.48	**0.52**	0.00	0.00	0.00	0.00	0.52	0.82
	场景 2	30	2	0.00	**0.52**	0.46	0.02	0.00	0.00	0.52	0.84
		30	3	0.00	0.52	**0.46**	0.02	0.00	0.00	0.46	0.78
		45	2	0.00	**0.52**	0.48	0.00	0.00	0.00	0.52	0.84
		45	3	0.00	0.52	**0.48**	0.00	0.00	0.00	0.48	0.79
		60	2	0.00	**0.52**	0.47	0.00	0.00	0.00	0.52	0.84
		60	3	0.00	0.52	**0.47**	0.00	0.00	0.00	0.47	0.79
		72	2	0.00	**0.52**	0.47	0.00	0.00	0.00	0.52	0.84
		72	3	0.00	0.52	**0.47**	0.00	0.00	0.00	0.47	0.79
	场景 3	30	2	0.00	**0.30**	0.56	0.14	0.00	0.00	0.30	0.68
		30	4	0.00	0.30	0.56	**0.14**	0.00	0.00	0.14	0.55
		45	2	0.00	**0.25**	0.65	0.10	0.00	0.00	0.25	0.67
		45	4	0.00	0.25	0.65	**0.10**	0.00	0.00	0.10	0.54
		60	2	0.00	**0.23**	0.70	0.07	0.00	0.00	0.23	0.67
		60	4	0.00	0.23	0.70	**0.07**	0.00	0.00	0.07	0.53
		72	2	0.00	**0.22**	0.72	0.05	0.00	0.00	0.22	0.67
		72	4	0.00	0.22	0.72	**0.05**	0.00	0.00	0.05	0.52
	场景 4	30	3	0.00	0.01	**0.24**	0.52	0.21	0.03	0.24	0.58
		30	6	0.00	0.01	0.24	0.52	0.21	**0.03**	0.03	0.19
		45	3	0.00	0.00	**0.19**	0.57	0.22	0.01	0.19	0.56
		45	6	0.00	0.00	0.19	0.57	0.22	**0.01**	0.01	0.18
		60	3	0.00	0.00	**0.18**	0.61	0.20	0.01	0.18	0.56
		60	6	0.00	0.00	0.18	0.61	0.20	**0.01**	0.01	0.17
		72	3	0.00	0.00	**0.15**	0.66	0.19	0.00	0.15	0.55
		72	6	0.00	0.00	0.15	0.66	0.19	**0.00**	0.00	0.18
	场景 5	30	2	0.13	**0.76**	0.11	0.00	0.00	0.00	0.76	0.91
		45	2	0.08	**0.86**	0.06	0.00	0.00	0.00	0.86	0.95
		60	2	0.06	**0.90**	0.04	0.00	0.00	0.00	0.90	0.96
		72	2	0.04	**0.92**	0.03	0.00	0.00	0.00	0.92	0.97

致谢。 本出版物获得了美国国立卫生研究院国家转化科学促进中心的资助，资助编号为 UL1TR001873 和 TL1TR001875。

参考文献

BABB, J. S. AND ROGATKO, A. (2001). 癌症I期临床试验中的患者个体化给药。《医学统计学》**20,** 2079–2090. BEAVER, J. A., ISON, G., AND PAZDUR, R. (2017). 重新评估入选标准——在肿瘤学试验中平衡患者保护与参与。《新英格兰医学杂志》**16,** 1504–1505. CHAPPLE, A. D. AND THALL, P. F. (2018). 基于毒性发生时间、允许自适应亚组组合的I期临床试验亚组特异性剂量探索。《药物统计学》**17,** 734–749. CHENG, J. D., BABB, J. S., LANGER, C., AAMDAL, S., ROBERT, F., ENGELHARDT, L.R., ET AL. (2004). I期临床试验中的患者个体化给药：过量控制递增法在PNU- 214936中的作用**22,** 602 609. 《临床肿瘤学杂志》：美国临床肿瘤学会官方期刊– . CONAWAY, M. (2017). 偏序群I期试验的保序设计。《临床试验》**14,** 491–498. FEHRENBACHER, L., ACKERSON, L., AND SOMKIN, C. (2009). 基于人群的新诊断非小细胞肺癌（NSCLC）患者队列中，化疗（CT）和抗血管生成治疗（AAT）随机临床试验的入选率。《临床肿瘤学杂志》**27,** 6538. GEORGE, S. L. (1996). 放宽癌症临床试验的患者入选标准。《临床肿瘤学杂志》：美国临床肿瘤学会官方期刊**14,** 1364–1370. GERBER, D. E., LACCETTI, A.L., XUAN, L., HALM, E. A., AND PRUITT, S. L. (2016). 既往癌症对肺癌临床试验入组资格的影响。《美国国家癌症研究所杂志》**106,** 302. HARVEY, R. D., MILEHAM, K. F., BHATNAGAR, V., BREWER, J. R., RAHMAN, A., MORAVEK, C., ET AL. (2021). 扩大晚期非小细胞肺癌患者临床试验入选标准的影响：对部分ASCO-Friends建议的真实世界分析。《临床癌症研究》**27,** 2430 2434– . KHAN, S., PATEL, S. P., SHOUSHTARI, A. N., AMBROSINI, G., CREMERS, S., LEE, S. M., ET AL. (2022) 间歇性MEK抑制治疗转移性葡萄膜黑色素瘤。《肿瘤学前沿》**12**. doi: 10.3389/fonc.2022.975643. KIM, E. S., BERNSTEIN, D., HILSENBECK, S. G., CHUNG, C. H., DICKER, A. P., ERSEK, J. L., ET AL.(2015). 推进分子驱动型试验入选标准现代化。《临床肿瘤学杂志》：美国临床肿瘤学会官方期刊**33,** 2815–2820. KIM, E. S., BRUINOOGE, S. S., ROBERTS, S., ISON, G., LIN, N. U., GORE, L., ET AL. (2017). 扩大入选标准，使临床试验更具代表性：美国临床肿瘤学会与癌症研究之友联合研究声明。《临床肿瘤学杂志》**35,** 3737. KITCHLU, A., SHAPIRO, J., AMIR, E., GARG, A. X., KIM, S. J., ET AL. (2018). 慢性肾病患者的代表性。《美国医学会杂志》**319**, 2437-2439. LACCETTI, A. L., PRUITT, S. L., XUAN, L., HALM, E. A., AND GERBER, D. E. (2016). 既往癌症不会对局部晚期肺癌患者的生存产生不利影响：一项全国性SEER-Medicare分析。《肺癌》（阿姆斯特丹，荷兰）**98,** 106. LAMEIRE, N. (2014). 新型抗癌药物的肾毒性。《临床肾脏病学杂志》**7,** 11–22. LEAL, T. B., REMICK, S. C., TAKIMOTO, C. H., RAMANATHAN, R. K., DAVIES, A., EGORIN, M. J., ET AL. (2011). 肾功能受损成年癌症患者使用硼替佐米的剂量递增及药理学研究：美国国家癌症研究所器官功能障碍工作组研究。《癌症化疗与药理学》**68,** 1439–1447. LEE, S. M. AND CHEUNG, Y. K. (2009). 持续再评估方法中的模型校准。《临床试验》（英国伦敦）**6,** 227–238. LIU, R., RIZZO, S., WHIPPLE, S., PAL, N., PINEDA, A. L., LU, M., ET AL. (2021). 使用真实世界数据和人工智能评估肿瘤学试验的入选标准。《自然》**592,** 629–633. LORUSSO, P. M., VENKATAKRISHNAN, K., RAMANATHAN, R. K., SARANTOPOULOS, J., MULKERIN, D., SHIBATA, S. I., ET AL. (2012). 硼替佐米在晚期恶性肿瘤及不同程度肝功能障碍患者中的药代动力学与安全性：美国国家癌症研究所器官功能障碍工作组I期研究NCI-6432。《临床癌症研究》：美国癌症研究协会官方期刊**18,**2954–2963. MAGNUSON, A., BRUINOOGE, S. S., SINGH, H., WILNER, K. D., JALAL, S., LICHTMAN, S. M., ET AL. (2021). 临床试验入选标准现代化：ASCO-癌症研究之友体能状态工作组的建议。《临床癌症研究》：美国癌症研究协会官方期刊**27,** 2424–2429.

MALIK, L. AND LU, D. (2019). I期临床试验的入选标准：严格还是宽松？《癌症化疗与药理学》**83,** 999–1002. MORITA, S., THALL, P. F., AND TAKEDA, K. (2017). I期试验中亚组特异性剂量选择方法的模拟研究。《药物统计学》**16,** 143–156. O'QUIGLEY, J. AND PAOLETTI, X. (2003**59,** 430 440). 有序组的持续再评估方法。《生物统计学》– . O'QUIGLEY, J., PEPE, M., AND FISHER, L. (1990). 持续再评估方法：癌症I期临床试验的实用设计。《生物统计学》**46,** 33–48. O'QUIGLEY, J., SHEN, L. Z. AND GAMST, A. (1999). 双样本持续再评估法。《生物医药统计学杂志》**9,** 17–44. OSAROGIAGBON, R. U., VEGA, D. M., FASHOYIN-AJE, L., WEDAM, S., ISON, G., ATIENZA, S., ET AL. (2021). 临床试验入选标准现代化：ASCO-癌症研究之友既往治疗工作组的建议。《临床癌症研究》：美国癌症研究协会官方期刊**27,** 2408–2415. PIANTADOSI, S. AND LIU, G. (1996). 使用药代动力学测量改进剂量递增研究设计。《医学统计学》**15,** 1605–1618. 美国食品药品管理局。(2020). 提高临床试验人群的多样性——入选标准、入组实践和试验设计行业指南。
h
t
t
p
s://
w
ww.fda.gov/regulatory-information/search-fda-guidance-documents/enhancing-diversity-clinical-trial-populations-eligibility-criteria-enrollment-practices-and-trial。YUAN, Z. AND CHAPPELL, R. (2004). 多个风险组癌症I期临床试验的保序设计。《临床试验》（英国伦敦）**1,** 499–508.