



Molecular Imaging CRO Network

Micron's ViewPoint

---

## 医疗器械软件(SaMD)的开发 - FDA指南简介 -

---

Clinical Performance Assessment:  
Considerations for Computer-Assisted  
Detection Devices Applied to Radiology  
Images and Radiology Device Data in  
Premarket Notification (510(k))



Micron Inc.

## 目录

|                   |    |
|-------------------|----|
| 背景                | 3  |
| CADe的临床试验设计       | 4  |
| CADe 的临床试验设计的注意事项 | 6  |
| - 指标的选择           | 7  |
| - 阅片和随机化          | 8  |
| - Scoring         | 9  |
| - 培训              | 9  |
| - 研究对象            | 10 |
| 总结                | 12 |

### 免责声明

本文中记载的内容有可能在没有任何预告通知的情况下进行变更。

关于本文件，本公司 (Micron, Inc.) 不予任何形式的保证（包括但不限于商业性及对特定目的的适用性的暗示的保证等）。

如因本文件中的有关谬误而造成的任何损害，本公司不承担任何责任。

未经撰稿人事先书面许可，严禁复制、重新销售或修改本文件的任何部分。



Micron Inc.

## 背景

利用CT、核磁共振和超声波等医学影像辅助医生诊断的系统，即计算机辅助检测/诊断 (CAD) 系统，被期待来辅助医生诊断进而防止漏诊、缩短医生的阅片时间和减少医生之间的诊断差异等。将CAD系统引入临床实践的决定性因素之一是CAD系统的客观诊断性能。当CAD系统被认证/批准为医疗器械时，客观诊断性能的数据非常重要。

美国食品药品监督管理局 ("FDA") 于2022年9月28日发布了题为《Clinical Performance Assessment: Considerations for Computer-Assisted Detection Devices Applied to Radiology Images and Radiology Device Data in Premarket Notification (510(k))》的最终指南 ("FDA 指南") 。

FDA指南就应如何设计和开展临床试验 (可包括临床表现研究、性能评估研究、临床研究和非临床研究，但在本文件中统称为“临床试验”) 以证明CADe (计算机辅助检测) 系统的客观诊断性能提出了建议。这里的CADe是指在预期用户 (如医生或其他医疗保健专业人员，下称“临床医生”) 判读时，对患者的影像数据的某一部分进行识别、标记、突出显示或以其他方式引导注意。

本文就FDA指南，重点介绍了FDA指南中被认为对CADe系统临床试验设计开发特别重要的方面。

希望本文能给在计划开发CADe系统或进行临床试验的公司和机构在临床实践中实施此类系统提供帮助。



## CADe的临床试验设计

CADe系统以临床医生确认病人影像数据时断定有异常可能性为目的而被设计。CADe系统的输出与临床医生的诊断过程之间存在着复杂的关系，又因临床性能可能取决于以下研究设计中需要考虑的各种因素，所以临床试验设计时需多加考虑。

- 在CADe解释过程中的应用时间 (Concurrent Reader或Second Reader)
- CADe标记的特征 (大小、形状、边界类型 (如实线、虚线、圆线)、与异常点的距离)
- 临床医生对CADe检测到的疾病/异常的知识
- CADe标记的数量

临床性能评估的临床试验设计一定考虑排除或减轻有可能对结果有影响的偏差 (表 1、2)。尤其当临床性能评估在实验室环境 (而非临床环境) 中进行时，特别建议。

表 1 CADe系统临床试验设计实例

| 研究设计                               | 说明  |
|------------------------------------|---|
| 实地测验或<br>前瞻性临床试验<br><br>(例如随机对照试验) | 在实际临床环境中评估CADe系统。持病率较低的话，需要大量的病例，实际实施可能不切实际。                            |
| 回顾性临床试验                            | 病例数量充足的回顾性临床试验，有可能作为实地测验的替代方案。  |
| 压力测试                               | 回顾性临床试验的阅片困难的影像比列比临床诊断的阅片时更高。需要注意的是，临床试验和临床诊断环境之间的持病率和阅片难易度的差异可能会改变其性能。 |

表 2 : CADe系统的临床试验考量事项

| 考量事项                         | 说明   |
|------------------------------|--|
| 研究对象                         | 研究对象 (包括患病病例和正常病例) 是否充分代表了目标人群，或是从目标人群中抽取的样本 |
| 混淆因素                         | 旨在避免混淆。                                      |
| 样本量                          | 有足够的样本量来证明性能。                                |
| 参考标准<br>(Reference Standard) | 用于评估性能的“真相”定义是否恰当，是否分析了参考标准的不确定性 (如果适用的话)。   |
| 阅片医师                         | 选择能代表临床医师团体的医师。                              |
| 成像硬件                         | 选择和临床一样的硬件。                                  |

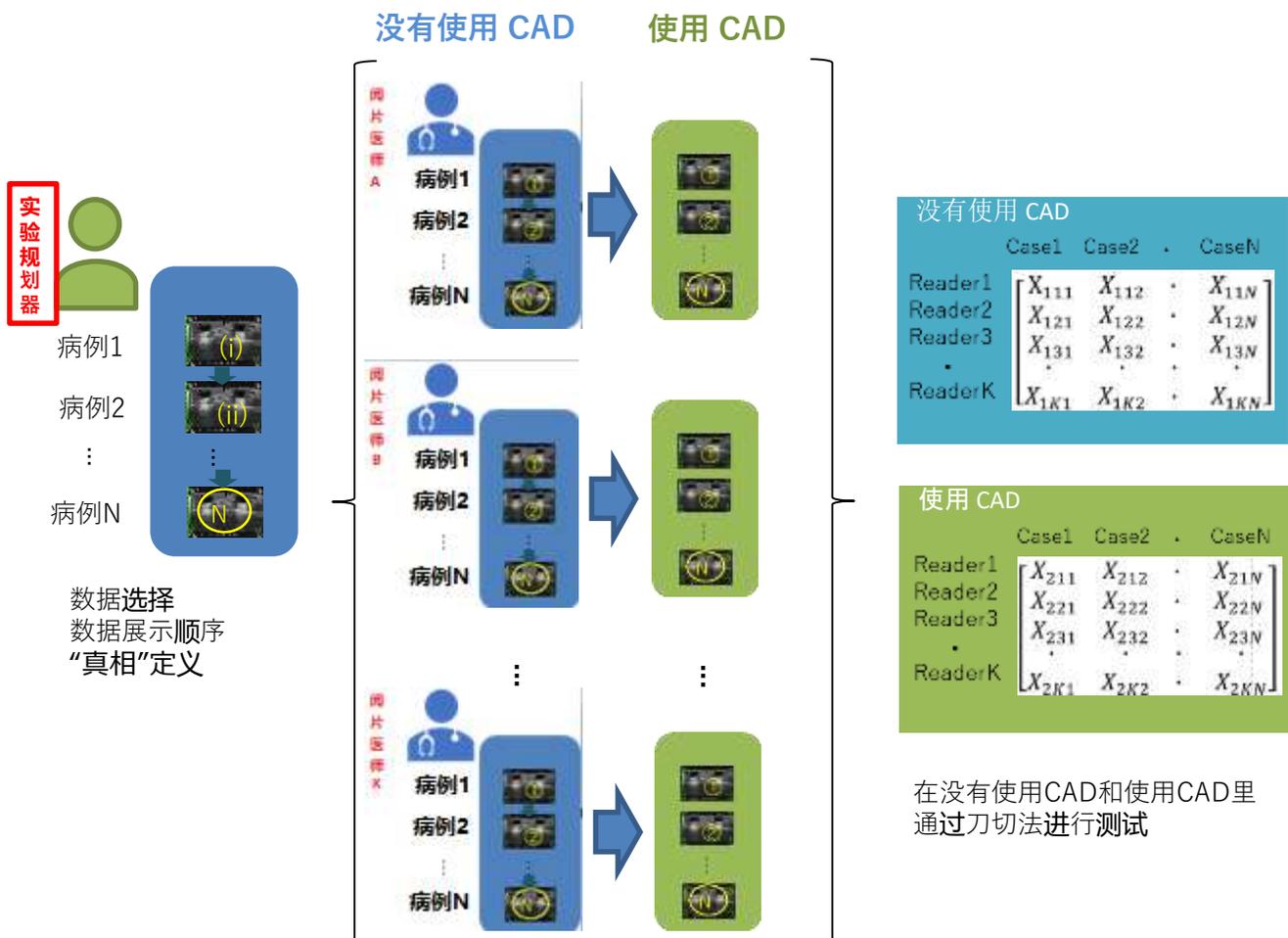
## CADe的临床试验设计

CADe系统的临床性能评估一般采用若干阅片医师评估若干病例的Multiple Reader Multiple Case (MRMC) 设计。通过实施“完全交叉”的所有阅片医师独立评估所有病例，可使病例数的统计检测能力达到最大限度。但是，即使是“非完全交叉”的设计，在前瞻性等研究中，也有若干阅片医师无法评估同一病例情形会被允许。

无论是否选择完全交叉设计，本指南都建议使用MRMC设计来评估CADe系统的临床性能。

图1展示了在Second Reader型CADe系统临床试验中经常被采用的连续序列设计的MRMC设计实例（此例不是出自FDA指南）。

图 1 MRMC实验示例（连续序列设计）。



(从Junji Shiraishi 的《医学影像视觉评估方法》，Ohmsha略改后引用)

## CADe的临床试验设计的注意事项

以下是FDA指南中列出的CADe系统临床试验设计注意事项的汇总清单。

在下面列出的项目中，“指标的选择”、“阅片和随机化”、“Scoring”、“培训”和“研究对象人群”将在接下来几页详细介绍。

日本語の意味を確認

### 指标的选择

- 在灵敏度和特异性之上，基础指标也有可能作为指标的候选

### 试验对照组

- 建议与对照方法相比来评估CADe系统的性能
- 如果将传统的临床解释方法（如不使用CADe系统的阅片）设为研究对照组，则在CADe性能评估中选择最相关的方法

### 阅片和随机化

- 临床试验中的阅片方案（如按顺序阅片、交叉阅片）的应选择符合CADe系统的预期用途

### 阅片的评估尺度

- 对于基于ROC的指标，阅片医师在评估患者病变情况时，使用评分表（如7分制或100分制）收集数据
- 建议对阅片医师进行评分表使用方法的培训

### Scoring

- 确定阅片医师的理解和参考标准之间关系的程序
- Scoring程序和定义是CADe系统临床试验里的一项重要因素，建议加以说明

### 培训

- 针对CADe系统的使用方法以及参与临床试验的细节，建议对参与临床试验的阅片医师进行明确的指导和培训

### 研究对象人群

- 依据被明确定义的“入选”和“排除”标准，进行前瞻性或回顾性的数据收集
- 建议将数据收集程序纳入方案里

### 参考标准 Reference Standard

- 参考标准（又称黄金标准、基准真实值）是因真理过程而被决定，用于展示是否存在疾病/异常及其范围和位置
- 建议记述用于确定参考标准的方法（如病理诊断等）



## 指标的选择

为了展示CADe系统的有效性，需要选择指标。指标（主要指标和次要指标）的选择应基于CADe系统的预期用途，并在评估开始前被确定。Receiver operative characteristic (ROC) 或其衍生系列（如LROC或FROC等），或灵敏度和特异性是替补的指标（表3，图2）。

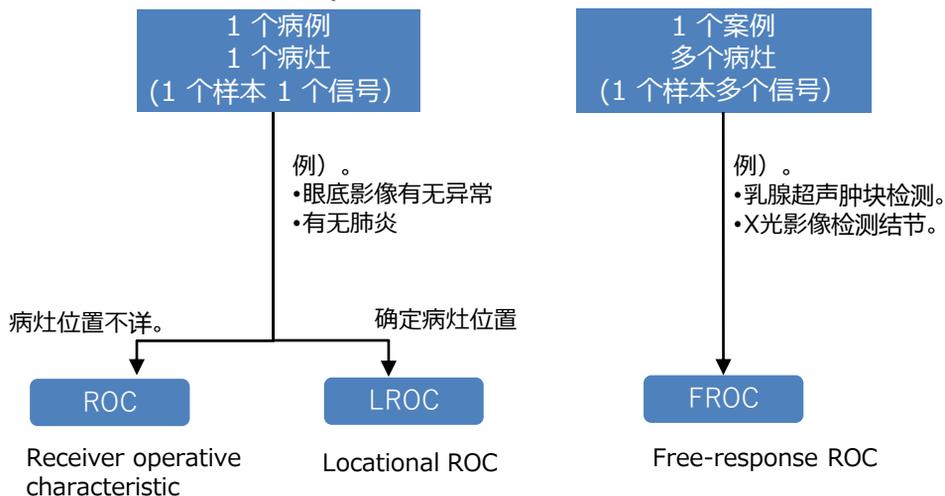
ROC、灵敏度和特异性的数据收集可在一次的阅片中同时进行。例如，阅片医师可先评估阳性/阴性来评估灵敏度/特异性，然后立即确认阳性/阴性判断的可信度等评估来对ROC评估。

FDA指南**建议选择ROC作为主要指标**，但也指出选择其他指标也有合适的时候。如果选择ROC作为主要指标，建议按照上述方式将灵敏度和特异性作为次要指标进行评估。还建议报告基于病变类型或成像条件等混淆因素的分层解析结果。

表 3：CADe 的指标示例

| 指标  | 说明   |
|---|--|
| <b>灵敏度</b><br>Sensitivity                       | 当参考标准为阳性时，诊断结果为阳性（真阳性）的百分比。                        |
| <b>特异性</b><br>Specificity                       | 当参考标准为阴性时，诊断结果为阴性（真阴性）的百分比。                        |
| <b>ROC</b><br>Receiver Operating Characteristic | 将纵轴作为真阳性率（灵敏度），横轴作为假阳性率（1 - 特异性）时的所有可能的阈值布局后得到的曲线。 |

图 2 ROC解析的类型  
(此处不是出自FDA指南)





Micron Inc.

## 阅片和随机化

临床试验中的阅片方案的选择需要和CADe系统的预期使用类型一致。表4是阅片方案每种使用类型的一例。在和Concurrent Reader一样的同一影像进行若干次阅片的方案中，需要避开因记住第一次阅片结果而造成的偏见（memory bias）时，建议阅片间隔至少保持在4周以上。

此外，为了减少性能评估的偏见，阅片医师，病例和阅片顺序需要进行随机化（Latin Square Design等）。建议解释说明随机化的方法以及展示相关联的流程。

表 4 阅片方案示例

| 使用类型  | 设计                  | 流程   |
|---|---------------------|--|
| Second Reader :<br>阅片医生在没有CAD的情况下阅片后，参照CAD的结果再次阅片 | 按顺序阅片<br>Sequential | <pre> graph LR     Case[Case] --&gt; without_CAD[without CAD]     without_CAD --&gt; with_CAD[with CAD]             </pre>   |
| Concurrent Reader :<br>阅片医师一边参考CAD的结果一边阅片         | 交叉阅片<br>Crossover   | <pre> graph LR     Case[Case] --&gt; S1[Session 1]     Case --&gt; S2[Session 2]     S1 --&gt; W1[without CAD]     W1 --&gt; S2     S2 --&gt; W2[with CAD]     S1 -.-&gt; &gt;= 4 weeks  S2             </pre> |



## Scoring

FDA指南将确定阅片医生的诠释和参考标准之间关联的程序称为Scoring。Scoring的程序和定义是CAdE系统临床试验的重要因素，建议加以说明。Scoring程序应在评估开始前被确定并阐述。建议对记载相关的Scoring过程时包含以下内容。

- 明确方法是电子或非电子
- 阅片医师对于参考标准的边界、面积和体积进行评估的边界、面积和体积的重叠。
- 阅片医师对于参考标准的边界或空间位置进行评估重心的关系。
- 阅片医师对于参考标准的重心进行评估的边界或空间位置的关系。
- 最终阅片医师的决定
- 其他

如果选择最终阅片医师的决定，建议说明参与阅片的人数、资历、经验和专业、事前指示的内容以及在Scoring中使用的特定标准。如果参与Scoring的最终阅片医师不止一人，还应说明评估的综合过程以及如何处理差异的方法。

## 培训

FDA指南建议对参与试验的读阅片医师进行以下方面的培训。

- CAdE系统的细节和使用方法
- 参与临床试验的相关指示
- 试验特有的阅片指示或规则
- 评分表相关培训
- 典型的正常/异常病例的评估的实践。

对于特殊的阅片指示或规则（例如，在使用CAD阅片时，改变无CAD的初始解释的规则），建议从阅片、临床工作流程和医疗行为的观点来展示临床上的有效性。也建议培训时不使用实际试验的影像。



## 研究对象

病例数据应依据被明确定义的“入选”和“排除”标准，进行前瞻性或回顾性的数据收集。建议将数据收集程序纳入方案里。为临床试验收集的病例数据应独立于系统开发过程中使用的病例数据。病例数据收集方法的一例是，从各研究中心连续收集入选标准内和排除标准外的病例数据。

收集病例数据时应注意以下事项

- 从CAdE系统的预期使用人群中取样。
- 包括适当比例的疾病/有异常例和正常病例
- 样本量应足够大，以便有足够的力量显示有效性。
- 必要时，从考虑到临床重要因素（如临床重要混杂因素、效应修饰因子、并发症）的病例组包含足够数量的病例，按每个亚组来获得临床性能的估计值和置信区间。
- 如果对每个分组提出多项绩效要求，则需要对每个分组的测试进行统计调整。

使疾病/有异常例丰富化可以有效的收集病例数据，但丰富化的程度应考虑到研究设计中可能出现的偏差。也可以用临床医生阅片困难的阅片结果来丰富研究对象。这就是所谓的“压力测试”，因为研究设计时会产生偏差所以在临床试验中应用压力测试时需谨慎。

适当的压力测试研究人群范例：

检测结肠息肉CAdE系统，如果系统旨在检测各种大小的息肉的话，那么对息肉较小的病例进行丰富化比较适当。

FDA指南还记载了在CAdE系统的临床试验中的数据再利用时的注意事项，并提及需解释部分测试数据的重用不会对性能估计造成不合理的偏差，而且测试数据的完整性会得到维护。



Micron Inc.

## 研究对象

在描述研究对象时，建议酌情包括以下信息。

表 5 研究对象的描述事项

|      | 信息   | 例  |
|------|--|--|
| 病例特征 | 人口统计数据   | 年龄、民族、种族和性别  |
|      | 与CAdE系统应用相关的病史   |  |
|      | 患者的疾病状况和放射检查的适应症   |  |
| 影像特征 | 成像条件   | 是否使用造影剂、造影剂类型和每位患者的剂量、患者体重指数、辐射剂量、核磁共振成像的T-weighting和成像视野等 |
|      | 影像数据的详情  | 成像设备的品牌和型号以及扫描协议   |
|      | 成像中心   |  |
|      | 需要的话，数据处理中心  | 病人数据电子化  |
| 病例数量 | <ul style="list-style-type: none"> <li>• 疾病/有异常病例数</li> <li>• 正常病例数</li> <li>• 确定疾病状况、位置和范围程度（参考标准）的方法</li> </ul>  |  |
| 混淆因素 | 按混淆因素或影响因素分层的疾病组分布 <ul style="list-style-type: none"> <li>• 病灶类型（如增生性和腺瘤性结直肠息肉）</li> <li>• 病灶的大小，位置</li> <li>• 疾病阶段、并发症</li> <li>• 器官特征（如乳房的构成）</li> <li>• 成像设备（如品牌和型号、扫描协议）</li> <li>• 成像中心</li> <li>• 需要的话，数据处理中心</li> </ul> |  |
| 对比   | 将收集到的病例数据的临床、影像和病理特征与目标人群进行比较  |  |



Micron Inc.

## 总结

为了以适当的方式评估CAD系统的诊断准确性，确立适当的临床试验非常重要，如果这些数据会被用作申请批准的话，则需要更加合理和有效的设计。

本文介绍了FDA指南概述的进行CADe系统临床试验时的注意事项。在日本开发CAD系统时需要注意的一般要点，请参阅“[医用画像診断支援システムの臨床試験～医療AIの社会実装に向けて～](#)”。此外，关于本文未提及的“参考标准”的制定过程，请参阅“[医用画像診断支援システムの臨床評価：参照基準の重要性～医療AIの社会実装に向けて～](#)”。

Micron是日本和亚洲领先的Imaging CRO，为许多基于影像的临床试验提供支持。Micron有支持多项本文所介绍的SaMD的项目（[プログラム医療機器 \(SaMD\) 承認取得支援実績のお知らせ](#)），并提供依据FDA指南的临床试验建议。如需了解更多信息，欢迎随时咨询·交流。

### 会社概要

東京本社 : 〒108-0073 東京都港区三田三丁目13番16号 三田43MTビル9階  
TEL.03-6631-3693

大阪支社 : 〒532-0003 大阪市淀川区宮原四丁目5-36 ONEST新大阪スクエア6階  
TEL.06-6399-0007

名古屋事業所 : 〒474-8511 愛知県大府市森岡町七丁目430番地  
国立研究開発法人 国立長寿医療研究センター 南病棟2階  
TEL.0562-46-2105

事業内容

- ① イメージング技法を活用した医薬品、診断薬、医療機器、バイオマーカーの開発支援
- ② 臨床開発支援  
(モニタリング、品質管理、イメージング・コアラボ業務、画像解析、読影支援等)
- ③ PET薬剤の治験薬GMP製造支援
- ④ 臨床開発に係るコンサルティング

Homepage : <https://micron-kobe.com>

LinkedIn : <https://www.linkedin.com/company/micron-imaging/>

Email : [imagingbiomarker@micron-kobe.com](mailto:imagingbiomarker@micron-kobe.com)