

正本

檔 號：  
保存年限：

## 衛生福利部食品藥物管理署 函

地址：115021 臺北市南港區研究院路一段130巷  
109號

聯絡人：陳泰龍

聯絡電話：02-27878275

傳真：

電子郵件：tlchen@fda.gov.tw

108



台北市萬華區長沙街2段73號3F

受文者：臺北市儀器商業同業公會

發文日期：中華民國114年8月22日

發文字號：FDA器字第1141618407號

速別：普通件

密等及解密條件或保密期限：

附件：

主旨：修正之「人工智慧/機器學習技術之電腦輔助偵測(CADe)及  
電腦輔助診斷(CADx)醫療器材查驗登記技術指引」及相關  
問答集業經本署於114年8月12日以FDA器字第1141617758號  
公告，請查照並轉知所屬。

說明：旨揭公告及其附件請至本署全球資訊網站  
(<http://www.fda.gov.tw>)醫療器材專區(路徑：官網首頁>業務  
專區>醫療器材>最新消息)及智慧醫療器材資訊暨媒合平台  
(<http://aimd.fda.gov.tw/>)之公告訊息(路徑：網站首頁>公告事  
項>公告訊息)下載。

正本：新北市醫療器材商業同業公會、桃園市醫療器材商業同業公會、嘉義市醫療器材  
商業同業公會、台灣省醫療器材商業同業公會聯合會、臺南市醫療器材商業同業  
公會、臺北市醫療器材商業同業公會、臺中市醫療器材商業同業公會、高雄市醫  
療器材商業同業公會、高雄市直轄市醫療器材商業同業公會、臺南市直轄市醫  
療器材商業同業公會、中華民國醫療器材商業同業公會全國聯合會、彰化縣醫  
療器材商業同業公會、屏東縣醫療器材商業同業公會、台灣口腔生物科技暨醫療器材  
產業發展促進協會、台灣牙科器材同業交流與公益協會、台北市生物技術服務商

業同業公會、社團法人中華民國助聽器同業聯合協進會、中華民國助聽器商業同業公會、桃園縣助聽器商業公會、台中市助聽器商業同業公會、高雄市助聽器商業同業公會、彰化縣助聽器商業同業公會、中華民國儀器商業同業公會、嘉義市儀器商業同業公會、臺中市儀器商業同業公會、臺北市儀器商業同業公會、高雄市儀器商業同業公會、桃園市儀器商業同業公會、新竹市儀器商業同業公會、臺南市儀器商業同業公會、中華生物醫學工程協進會、台灣醫藥品法規學會、財團法人金屬工業研究發展中心、財團法人塑膠工業技術發展中心、財團法人台灣商品檢測驗證中心、財團法人醫藥品查驗中心、財團法人醫藥工業技術發展中心、財團法人工業技術研究院量測技術發展中心、財團法人生物技術開發中心、台灣生物產業發展協會、中華民國西藥商業同業公會全國聯合會、台灣醫療器材門市發展協會、中華民國生物產業發展協會、中華民國全國商業總會、中華民國全國工業總會、台灣先進醫療科技發展協會、台灣研發型生技新藥發展協會、經濟部產業發展署、國家科學及技術委員會新竹科學園區管理局、台灣科學工業園區工業同業公會、國家科學及技術委員會中部科學園區管理局、國家科學及技術委員會南部科學園區管理局、南港軟體工業園區二期管理委員會、台灣區電機電子工業同業公會、台灣臨床檢驗標準協會、台灣藥物臨床研究協會、財團法人資訊工業策進會、社團法人臺灣輔具暨長期照護產業發展協會、中華民國生物醫學工程學會、財團法人生技醫療科技政策研究中心、社團法人台灣醫學資訊學會、台灣醫院協會、社團法人台灣生醫電子工程協會、財團法人工業技術研究院生醫與醫材研究所、台灣智慧醫療聯盟計畫推動辦公室、成功大學人工智慧生技醫療創新研究中心、國立臺灣大學人工智慧技術暨全幅健康照護聯合研究中心、臺灣美國商會、歐洲在台商務協會、台北市日本工商會、臺灣醫療暨生技器材工業同業公會、中華民國資訊軟體協會、台北市電腦商業同業公會、台灣顯示器產業聯合總會、台灣健康資訊產業整合協會、台灣健康資訊交換第七層協定協會、台灣數位安全聯盟、財團法人中華民國國家資訊基本建設產業發展協進會、社團法人台灣電子設備協會、台灣生技醫療照護輔具協會、台灣醫療照護輔具協會、臺北醫學大學校級人工智慧醫療研究中心、中國醫藥大學附設醫院人工智慧中心、彰化基督教醫院人工智慧發展中心、輔仁大學人工智慧發展中心、清華大學人工智慧研發中心、陽明交通大學人工智慧普適研究中心、陽明交通大學數位醫學暨智慧醫療推動中心、陽明交通大學創新醫材轉譯研發中心

副本：

署長 姜至剛

## 衛生福利部食品藥物管理署 公告

發文日期：中華民國114年8月12日  
發文字號：FDA器字第1141617758號



主旨：公告修正「人工智慧/機器學習技術之電腦輔助偵測(CADe)及電腦輔助診斷(CADx)醫療器材查驗登記技術指引」

依據：行政程序法第165條

公告事項：

- 一、人工智慧技術發展快速，應用人工智慧/機器學習技術之醫療器材日益增加，在基於確保產品安全與有效性之前提下，特修正旨揭指引，以供業者作為評估產品開發及準備申請查驗登記/變更登記所須資料之參考。另就修正內容特訂定「人工智慧/機器學習技術之醫療器材獨立性能評估常見問答集」。
- 二、旨揭技術指引及相關問答集另載於本署全球資訊網站(<http://www.fda.gov.tw>)醫療器材專區(路徑：官網首頁>業務

專區>醫療器材>最新消息)及智慧醫療器材資訊暨媒合平台  
(<http://aimd.fda.gov.tw/>)之最新消息(路徑：網站首頁>公告訊  
息>最新消息)。

三、本署112年9月14日FDA器字第1121607549號公告自即日起停  
止適用。

署長 姜至剛



# 人工智慧/機器學習技術之 電腦輔助偵測(CADe)及電腦輔助診斷(CADx) 醫療器材查驗登記技術指引

111年8月公布  
112年9月修正  
114年8月修正

## 【前言】

為確保應用人工智慧(Artificial intelligence, AI)/機器學習(Machine learning, ML)技術之電腦輔助偵測(Computer-assisted detection, CADe)與電腦輔助診斷(Computer-assisted diagnosis, CADx)醫療器材之安全、效能及品質，本署參考美國、日本及韓國等國管理規範，且以本署「人工智慧/機器學習技術之醫療器材軟體查驗登記技術指引」為基礎，增加技術重點與舉例說明，提供外界作為該類型產品評估開發及查驗登記所需資料準備之參考，業者應視本身產品宣稱效能、適用範圍、對象及使用情境等評估本指引中的內容是否適用。

AI/ML CADe 與 CADx 醫療器材應符合醫療器材管理法等相關規定。本指引所述之產品描述、安全與效能驗證部分，係依據現行參考資料訂定，惟科技發展日新月異，法規更新未逮處，為確保國人健康安全，本署受理醫療器材查驗登記時，得視產品特性，要求申請者提供本指引所列項目外之安全與效能驗證評估資料。部分具有特殊需求或特殊風險考量的 AI/ML 醫療器材可能不完全適用本指引，該類產品將依個別狀況進行評估。

本指引將不定期更新。

本指引不用於界定醫療器材產品之屬性管理範圍，對於產品屬性管理範圍，可參考本署公布之「醫用軟體分類分級參考指引」。

## 一、適用範圍

本指引說明第二、三等級 AI/ML CADe 與 CADx 醫療器材查驗登記技術要點，適用於存在於其它儀器或以醫療器材軟體(Software as medical device, SaMD)形式單獨存在之醫療器材。

## 二、名詞定義

### (一)人工智慧/機器學習(Artificial intelligence / Machine learning, AI/ML)

人工智慧(AI)是電腦科學、統計學和工程學的一種分支應用，透過演算法或模型展現學習、決策、視覺感知及語音識別等行為。機器學習(ML)為人工智慧的子集合，使電腦系統(軟體)不需透過明確的程式設計，便能針對資料進行自主學習。不同類型的機器學習演算法包括但不限於神經網路[例如前饋式神經網路(Feedforward neural network, FNN)、循環神經網路(Recurrent neural network, RNN)、卷積神經網路(Convolutional neural network, CNN)等]、貝葉斯網路(Bayesian networks)、決策樹(Decision tree)和支持向量機(Support vector machine, SVM)等。

### (二)訓練資料集(Training dataset)

用於訓練機器學習模型的一組資料集，透過使用訓練資料集建立或改進機器學習模型參數。訓練資料集不可出現在測試資料集(Test dataset)中。

### (三)測試資料集(Test dataset)

從未於訓練期間用於機器學習演算法使用的一組資料集，用於評估訓練後機器學習模型的效能。

### (四)電腦輔助偵測(Computer-assisted detection, CADe)

透過電腦系統(軟體)自動提取臨床資料中的關鍵特徵，如醫學影像、檢驗數據、生理參數、基因特徵及心電圖等，輔助偵測病變或異常值。以醫學影像(超音波、X光攝影、磁振造影、電腦斷層掃描、正子造影、內視鏡等)為例，系統能自動辨識並標記可疑病變或異常位置。

### (五) 電腦輔助診斷(Computer-assisted diagnosis, CADx)

透過電腦系統(軟體)對醫學影像及其他臨床資料,如檢驗數據、生理參數、基因特徵、心電圖等,進行加工與處理,提供量化或分類的診斷支援訊息。例如以醫學影像(超音波、X光攝影、磁共振影、電腦斷層掃描、正子造影、內視鏡等)鑑別良性與惡性病變、評估病況發展情形(如嚴重度、期別),並針對診斷結果進行風險評估分析。

## 三、AI/ML CADe 與 CADx 醫療器材查驗登記申請文件技術要點

依據「醫療器材許可證核發與登錄及年度申報準則」規定,申請製造或輸入第二、三等級醫療器材查驗登記,應檢具附表二所列文件、資料,本章節係針對其中要求之技術文件、資料進行說明。

### (一) 產品之結構、材料、規格、性能、用途、圖樣等有關資料(醫療器材許可證核發與登錄及年度申報準則附表二之說明八)

#### 1. 產品概要

文件應包含產品名稱、版本、軟體使用之平台,以及產品功能性說明及敘述、軟體架構、產品採用演進式(Adaptive)或閉鎖式(Locked)演算法設計等內容,並根據產品預期用途、效能、禁忌症及使用限制,提供目標的功能值(如檢出率、偽陽性率、偽陰性率、和其它必要因素),可參考本署公布之「醫療器材中文說明書編寫原則」內容撰寫下列項目:

- (1) 產品敘述
- (2) 用途、效能或適應症
- (3) 預期可預見之副作用或併發症
- (4) 禁忌症
- (5) 警告、注意事項、使用限制

產品應於警告、注意事項、使用限制中載明醫療器材失效(Fail)或無法達到預期效能的情況(例如輸入臨床資料之品質不佳,或有

某些特定不適用之資料族群等)，並應評估產品輔助診斷帶來的影響(尤其是偽陽性、偽陰性結果可能導致的危害風險)。

文件應針對產品的預期用途有詳細描述，包含預期使用者及建議的使用者訓練方式、目標病症、病患族群、適用資料類型、使用情境[如與使用者同時判讀或在使用者判讀後提供第二判讀(Second read)]等，並提供詳細的使用說明、判讀流程、輸出結果、結果解讀的方法及應用。

另外，文件亦應提供資料產製(例如放射造影)器材之資訊、參數及資料產製流程，並說明產品輸出的檔案格式。

## 2. 演算法架構

文件應包含演算法架構說明及對應理論依據，並詳細描述演算法的輸入與輸出、演算法的流程圖、所有演算程序的說明、每個演算階段的設計與功能、輸出標記(或參數)的所有規格等。以醫學圖像類產品為例，其產品敘述及規格文件中，必須具體說明檢測原理、演算法設計與功能、輸入圖像處理程序、參考標準(Reference standard / Gold standard / Ground truth)及計分方式(Scoring)等。

演算法設計(建模)與輸出功能、輸入圖像處理程序、訓練方式及參考標準等說明如下：

### (1) 演算法設計(建模)與輸出功能

應提供用於AI/ML CADe與CADx醫療器材的網路架構及軟體程序的概述，建議包含以下(適用)之內容，詳如表一：

表一、AI/ML CADe 與 CADx 醫療器材的網路架構及軟體程序的概述  
建議內容

演算法名稱	例如支持向量機(SVM)/回歸、決策樹、卷積神經網路(CNN)、循環神經網路(RNN)等。
開發平台/框架	例如Tensorflow、PyTorch、Caffe等。

函式庫	例如Keras等。
模型	<p>例如神經網路之層數(Number of Layers)、權重(Weights)、激發函數(Activation)、優化器(Optimizer) [如Adam、隨機梯度下降(Stochastic gradient descent, SGD)等]、損失函數(Loss function) [如交叉熵(Cross-entropy)、L1範數(L1-norm)等]及模型評價指標(Metric)等</p> <p>若採用多個模型融合(Model ensembling)，應說明不同模型訓練與推理的方式，以及模型融合的策略，如多數投票。</p>
演算法功能 (例如提供影像標記、病變風險的量化診斷支援訊息等)	<ul style="list-style-type: none"> <li>■ AI/ML CADe醫療器材應說明器材進行圖像標記的種類與代表意義，包含圖像標記幾何形狀與相關參數，例如形狀、尺寸、關注位置(重疊或相鄰)、界線(實心或虛線)及顏色等。</li> <li>■ AI/ML CADx醫療器材應說明器材量化病變風險的評估方法、評估級距(例如風險高低、百分比及評分等級等)、代表意義、演算法的限制等，並應說明量化方法在臨床上的科學依據，以及演算法輸出結果對臨床實務、病患照護的影響與相關性。</li> </ul>
演算法處理流程	<p><b>建議</b>利用流程圖方式說明演算法處理流程，包括所有手動操作和相關的預設值，例如由醫師選擇規則(Rules)或閾值(Thresholds)；及所有半自動操作及相關的預設值，例如選擇種子點(Seed points)進行區域分割；及所有不需要與臨床醫師互動的自動化操作，例如影像自動由儀器/PACS系統導入並進行正規化處理。</p>

## (2)輸入圖像之處理程序

建議包含以下內容：

- I. 圖像處理步驟，例如濾波(Filtering)、分割(Segmentation)、正規化(Normalization)、對位(Registration)、假影或位移校正(Artifact or motion correction)等。
- II. 圖像正規化校準(Calibration)原理與參考依據。

## (3)演算法訓練方法

AI/ML CADe與CADx醫療器材其效能來自於良好設計的訓練方式及大量合適的資料，為了確認產品之預期用途及其安全及效能，應明確說明訓練之方法、架構、流程及所使用資料之相關性質、維護品質之方法等細節。

由於訓練方法必須依靠大量資料之分析，其訓練結果易受使用資料之性質、品質等影響，因此取得資料及維護其品質之作法亦須說明。資料可區分為建立演算法模組的訓練資料集、驗證產品初步功能的驗證用資料及測試產品最終效能的測試資料集。評估效能的測試用資料必須嚴格與上述其它資料區分，避免驗證結果之偏差。然而不論是何種資料，皆須考慮與產品模組之宣稱及預期用途之對應與其相關用途對應性。相關重點如下：

### I. 訓練方法、架構及流程描述

應說明用於訓練AI/ML模組之方法、架構及流程，包含使用的基礎模型、調整部分、訓練過程中使用之預訓練(Pre-training)之內容，以及用來確認訓練參數的性能指標(Performance metrics)與其允收標準(Criteria)等。

### II. 訓練資料集(Training dataset)

應說明用於訓練AI/ML模組之資料，包含資料之族群及臨

床意義、產製資料之形式、產製方法、附帶之其它資訊，以及資料收集與使用之歷程，請參考表二。

表 二、資料說明

<p>族群及用途對應性</p>	<p>資料的來源、性質，亦必須考慮族群取樣的偏差及資料本身之用途對應性。</p> <p>應包含以下(適用)之內容：</p> <ul style="list-style-type: none"> <li>■ 資料來源或資料庫說明</li> <li>■ 資料納入和排除之準則</li> <li>■ 病患分佈(如人種、年齡、性別)</li> <li>■ 與產品應用相關之病患病史</li> <li>■ 病患疾病狀態和造影檢查之適應症</li> <li>■ 資料來源機構(Collection sites)</li> <li>■ 案例數量</li> </ul> <p>-AI/ML CADe醫療器材應含疾病與正常之案例數及決定疾病狀態、位置與範圍的方法</p> <p>-AI/ML CADx醫療器材應含決定疾病狀態、良性/惡性、嚴重程度、期別的方法</p> <ul style="list-style-type: none"> <li>■ 依據干擾因子(Confounders)或修飾因子(Effect modifiers)分層的案例分佈，例如病變型態(如大腸瘻肉為增生型或腺瘤型)、病變大小、病變位置、疾病分期、器官特徵、共病(Concomitant disease)等</li> <li>■ 比較病患資料與產品目標族群間於臨床造影資料及疾病之特性(Characteristics)</li> </ul>
<p>產製形式</p>	<p>使用何種設備或臨床實務方式產生資料，如CT、MRI、PET、超音波、內視鏡或其它數位化醫學圖像及數值等。</p>

產製方法	資料的產製過程，包含使用之儀器廠牌與型號、資料獲取參數、資料格式、圖像擷取程序(Protocol)、造影檢查之方式，例如在有無顯影劑(Contrast)的情況下進行檢查、病患的顯影劑類型和劑量、病患的身體質量指數(Body mass index, BMI)、輻射曝露(Radiation exposure)、不同加權的MRI圖像、加工(如病患資料的數位化)處理場域(Processing sites)等。
附帶之其它資訊	資料之額外註解和臨床判斷，及產生這些註解與判斷的方法。例如：病灶位置、大小、良性/惡性或其它診斷資訊等。

#### (4) 參考標準(Reference Standard)

參考標準係客觀、確定的基準，用於比較、評估、訓練等之預期結果，可用於決定病患資料中疾病/病症/異常是否存在，亦或是疾病/病症/異常的良性/惡性及期別等屬性，應在評估前即完成參考標準的決定。參考標準**建議**包含以下內容：

- I. 參考標準依據，例如同類型器材的輸出、臨床檢驗結果、活體組織切片(Biopsy)或臨床醫師判讀等。
- II. 若參考標準為由臨床醫師判讀，應說明臨床醫師數量、資格、經驗、訓練要求等。
- III. 當有多位臨床醫師參與制定參考標準時，應描述如何結合多個判讀結果來完成一個總結的參考標準之程序，及多個判讀結果不一致時之處理程序。另外，參與制定參考標準之判讀醫師不可與參與產品臨床評估或臨床試驗之判讀醫師相同。

#### (二) 臨床前測試及原廠品質管制之檢驗規格與方法、原始檢驗紀錄及檢

驗成績書(醫療器材許可證核發與登錄及年度申報準則附表二之說明七)

### 1. 網路安全(Cybersecurity)

醫療器材網路安全係針對醫療器材因網路行為或資料傳輸引起的安全問題，防止醫療器材被未經授權的存取、修改、誤用或拒用，使功能減損而導致病患傷害，或避免資訊係經由醫療器材被未經授權的存取或轉移至外部接受者。可參考本署公布之「適用於製造業者之醫療器材網路安全指引」，提供醫療器材網路安全相關文件。

### 2. 軟體確效(Software validation)

可參考本署公布之「醫療器材軟體確效指引」所述檢附相關資料，其重點包含：

- (1) 醫療器材軟體之風險等級(Level of concern)
- (2) 醫療器材軟體之描述(Software description)
- (3) 醫療器材之危害分析(Device hazard analysis)
- (4) 醫療器材軟體之要求規格(Software requirements specification, SRS)
- (5) 醫療器材軟體設計架構圖(Architecture design chart)
- (6) 醫療器材軟體設計規格(Software design specification, SDS)
- (7) 追溯性分析(Traceability analysis)
- (8) 醫療器材軟體開發環境(Software development environment description)
- (9) 查證與確認文件(Verification and Validation Documentation)
- (10) 醫療器材軟體修訂歷史紀錄(Revision Level History)
- (11) 未解決的異常(Unresolved Anomalies / Bugs or Defects)

### 3. 產品宣稱功能測試(Product claimed functional test)

產品宣稱之量測、偵測或輸出功能應進行測試，並提供相

應之功能測試資料。

#### 4. 獨立性能評估(Standalone Performance Assessment)

獨立性能評估是指在測試資料未再新增醫事人員介入判讀(如人工圈選、標註)的情況下，僅用既有之資料，評估AI/ML醫療器材的性能。以CADE為例，獨立性能評估旨在評估CADE在未再新增醫事人員介入判讀的情況下，可正確偵測及標註病灶的能力。獨立性能評估是評估人工智慧/機器學習醫療器材性能的重要方法，它有助於評估醫療器材在不同臨床子群組(例如年齡、性別或疾病嚴重程度的病患族群)、資料來源機構和影像獲取設備下的泛化能力(Generalizability)，泛化能力是指在未判讀過的資料或條件下仍能維持穩定性能的能力。相關重點包含但不限於以下事項：

##### (1) 測試資料集(Test dataset)

獨立性能評估常使用回溯性資料，應考量資料之族群代表性及其臨床意義、產製資料之形式、產製方法、附帶之其它資訊，以及資料收集與使用之歷程，請參考表二。若使用資料庫，應自行評估是否符合產品說明書，並應確認資料是否為原始資料或經過預處理。

##### (2) 測試項目舉例如下：

- I. 偵測準確性測試(Detection accuracy testing): 適用於CADE，評估CADE在偵測出異常或病變的準確性。
- II. 定位準確性測試(Localization accuracy testing): 適用於CADE，評估產品於醫學圖像中，對異常或病變位置進行正確標示的能力。
- III. 診斷準確性測試(Diagnostic accuracy testing): 適用於CADx，評估CADx在區分疾病與非疾病、區分不同疾病類型上的準確性。
- IV. 泛化性測試(Generalizability testing): 評估產品在面對不同

條件時的性能穩定性，常透過來自不同機構、設備或參數設定等的資料來測試產品在不同條件下的性能穩定性。

V. 壓力測試(Stress testing): 採用在預期用途或範圍內之罕見或特殊的真實資料的演算法性能測試。

VI. 重複性與再現性測試:在相同或不同的資料擷取設備及操作條件下，對同一病患重複檢查，以評估產品輸出結果之差異。

### (3) 統計分析方法

在進行獨立性能評估之前，應事先建立參考標準的定義、計分過程及分析方法。應使用適當的統計方法評估產品性能，並於評估前確認主要及次要評估指標。常見之評估指標舉例如下：

- I. 靈敏度(Sensitivity)與特異度(Specificity)
- II. 偽陽性率(False Positive Rate)與偽陰性率(False Negative Rate)
- III. 陽性預測值(Positive Predictive Value)與陰性預測值(Negative Predictive Value)
- IV. 接受者操作特徵曲線(Receiver Operating Characteristic Curve, ROC Curve)或自由反應接受者操作特徵曲線(Free-Response Receiver Operating Characteristic Curve, FROC Curve)

### 5. 其它

依產品特性提供本指引所列項目外之安全及功能性驗證資料。

(三)臨床證據資料(Clinical evidence information) (醫療器材許可證核發與登錄及年度申報準則附表二之說明九)

本項資料包含學術理論依據與有關研究報告及資料(Academic and theoretical bases, relevant research reports and data)、臨床評估報告或臨床試驗報告(Clinical evaluation report or clinical trial report)，

用於驗證產品於臨床使用之安全性及有效性。相關重點包含但不限於以下事項：

1. 應提供產品在其預期用途下，佐證臨床使用安全性及有效性之臨床評估報告或臨床試驗報告，可採用前瞻式(Prospective)或回溯式(Retrospective)之臨床設計，在其輔助使用者判讀之實際使用情境下(如與臨床使用者同時判讀、在使用者判讀後提供第二判讀、或在臨床使用前先行單獨判讀)，執行產品臨床使用安全性及有效性之驗證。若產品係於國內執行臨床試驗者，應符合本署「醫療器材優良臨床試驗管理辦法」相關規定。

2. 執行臨床評估或臨床試驗時，須注意如下重點：

(1) 產品設計規格

產品設計規格須詳細界定適用或搭配使用之器材(硬體及軟體)詳細規格。產品宣稱適用之搭配使用器材，其規格須與設計開發所使用之訓練、驗證及測試資料所搭配之使用器材一致。倘若產品宣稱適用搭配之使用器材規格超出原產品設計開發、性能驗證所搭配之使用器材規格，則須提供規格適用性之科學證據。

(2) 測試資料集(Test dataset)

請參考表二。

- I. 臨床評估或臨床試驗之測試資料取得來源及方式，例如醫學圖像產製設備與產製參數(程序)應與產品說明書所載一致。
- II. 測試資料須如前述提供病患資料之相關訊息，且須具備足夠的代表性，並應根據臨床相關的干擾因子、修飾因子及共病等定義測試資料之子群組來分別進行評估。
- III. 當使用非我國人口族群的測試資料時，應說明用於我國人口族群的適用性，且若從多個醫療機構收集測試資料，應說明資料的整合性。

### (3) 臨床設計

- I. 用於臨床評估或臨床試驗之臨床設計應考量產品輔助使用者判讀之實際使用情境、輸出結果(如疾病良性/惡性、嚴重度、期別等)，以及使用者對於分析結果進行評估確認的專業能力。
- II. 臨床評估或臨床試驗應採用控制良好的臨床設計，建議盡量採用多判讀者/多案例(Multiple reader multiple case, MRMC)方法進行判讀，即一組臨床判讀人員在多種判讀情境進行案例判讀，且每位判讀人員皆獨立完成判讀所有案例(Fully-crossed)，以在給定的案例數量下，達到最大的統計檢定力。若某些情形下，每位判讀人員不易完成所有案例的判讀(Non-fully crossed)，該判讀結果仍是可接受的，惟應於報告中說明理由。
- III. 進行判讀之使用者/臨床醫師(即判讀人員)應符合產品的預期使用者族群，申請時並應說明判讀人員之資格與經驗。針對產品的使用方式與臨床設計的方法及流程，應給予判讀人員適當的訓練，並於申請時提供判讀人員訓練的詳細說明。
- IV. 臨床評估或臨床試驗之臨床結果判讀方法應與臨床實務具有一致性。

### (4) 臨床結果分析方法

臨床結果分析之主要及次要評估指標應符合產品的預期用途，且於臨床評估或臨床試驗開始前即應確認不可變動。臨床評估或臨床試驗多會觀察產品的靈敏度、特異度及ROC(或其變形)曲線下之面積(Area under the curve, AUC)，並常採用ROC(或其變形)曲線的AUC來總結臨床結果作為主要評估指標，以靈敏度及特異度，或以病患召回率(Recall rate)作為次要評估指標，若有臨床使用之特殊需要，亦得將靈敏度或特異度改列

入主要評估指標。前述臨床評估指標結果應符合產品之宣稱效能。

撰寫臨床評估或臨床試驗計畫時，**建議**應考慮包含但不限於下列重點：

- I. 產品宣稱及預期用途、禁忌等
- II. 研究對象(Objectives)
- III. 病患族群(Patient population)
- IV. 參與臨床評估或臨床試驗之醫事人員數量及資格
- V. 資料取得方式
- VI. 統計分析方式
- VII. 臨床分析結果

#### (四)宣稱可獨立診斷之AI/ML CADx醫療器材之特殊要求

宣稱可獨立診斷之AI/ML CADx醫療器材，除前章所述資料外，尚須提供下列符合要求之資料：

1. 須詳細評估其產品失效、無法達到預期效能、或輸出偽陽性、偽陰性結果時，可能導致的危害風險及臨床上相應的處置方式。
2. 測試資料須依據產品宣稱之預期用途，盡可能涵蓋較多的資料來源機構(例如不同的地區、不同的代表性機構)及資料產製設備(例如不同種類及產製參數)。
3. 宣稱可獨立診斷，不需依賴醫事人員判讀之AI/ML CADx醫療器材，執行臨床評估或臨床試驗時仍應將AI/ML CADx醫療器材獨立判讀與臨床醫師傳統判讀進行統計分析之比較，以驗證臨床上AI/ML CADx醫療器材獨立進行診斷之效能。

前述各項之要求可能不完全適用於全部的產品，最終仍須依產品本身特性(例如宣稱效能、風險程度等)而定。

#### 四、參考資料

1. 醫療器材管理法，109 年 1 月 15 日
2. 醫療器材許可證核發與登錄及年度申報準則，110 年 4 月 29 日
3. 醫療器材優良臨床試驗管理辦法，110 年 4 月 9 日
4. 醫用軟體分類分級參考指引，FDA 器字第 1111609211 號公告，111 年 9 月 15 日
5. 人工智慧/機器學習技術之醫療器材軟體查驗登記技術指引，FDA 器字第 1101608354 號公告，110 年 8 月 17 日
6. 適用於製造業者之醫療器材網路安全指引，FDA 器字第 1101603391 號公告，110 年 5 月 3 日
7. 醫療器材中文說明書編寫原則，FDA 器字第 1101603096 號公告，110 年 4 月 29 日
8. 醫療器材軟體確效指引，FDA 器字第 1061607211 號公告，106 年 12 月 15 日
9. Machine Learning-enabled Medical Devices: Key Terms and Definitions. IMDRF/AIMD WG/N67, May 2022
10. Principles and Practices for Medical Device Cybersecurity. IMDRF/CYBER WG/N60FINAL, March 2020
11. US FDA Computer-Assisted Detection Devices Applied to Radiology Images and Radiology Device Data – Premarket Notification [510(k)] Submissions, September 2022
12. US FDA Clinical Performance Assessment: Considerations for Computer-Assisted Detection Devices Applied to Radiology Images and Radiology Device Data in - Premarket Notification [510(k)] Submissions, September 2022
13. 中國 NMPA 深度學習輔助決策醫療器械軟件審評要點，2019 年
14. 人工智慧應用於醫學影像輔助系統審查基準（人工知能技術を利用した醫學影像診断支援システムに関する評価指標），厚生勞

働省醫藥・生活衛生局醫療機器審査管理課, 藥生機審発 0523 第 2 號, 2019

15. 韓國 MFDS 빅데이터 및 인공지능 ( AI ) 기술이 적용된 의료기기의 허가·심사 가이드라인 ( 민원인 안내서 ) [Guideline on Review and Approval for Artificial Intelligence (AI) and Big Data-Based Medical Devices , October 2019

16. 中國 NMPA 人工智慧醫療器械註冊審査指導原則, 2022 年

## 附錄一、肺結節電腦斷層影像電腦輔助偵測醫療器材軟體

本附錄補充本文內容，以舉例方式說明肺結節 CT 影像 CADe 醫療器材的技術重點，提供外界作為該類型產品評估開發及查驗登記所需資料準備之參考，業者應視本身產品宣稱效能、適用範圍、對象及使用情境等評估本附錄中的內容是否適用。

產品概要	<p>文件應針對產品的預期用途有詳細描述，包含預期使用者(臨床急診/健檢等不分科單位或是胸腔相關專科)及建議的使用者訓練方式、目標病症、病患族群(確定/疑似有肺結節之患者)、適用資料類型[如低劑量/一般電腦斷層，掃描區域(如從肺尖到底部)等]、使用情境[如與使用者同時判讀或在使用者判讀後提供第二判讀(Second read)]等，並提供詳細的使用說明、判讀流程、輸出結果、結果解讀的方法及應用。應說明工作原理，例如運用深度學習技術進行醫學影像之分析處理、對肺結節進行影像偵測(Detection)、分割(Segmentation)等。</p> <p><b>【範例】</b>本產品為採用閉鎖式人工智慧演算法之電腦輔助偵測軟體(CADe)，適用於20歲以上成年人，供經培訓合格的醫師使用，不能單獨作為臨床診療決策依據。本產品可以在醫師進行常規檢查時，輔助識別與標記雙肺(從肺尖到底部)之低劑量電腦斷層影像的結節，適用之結節尺寸為4mm (含)到30mm (含)，適用結節樣態包含實質性結節(solid nodules)、非實質結節(subsolid nodules)、與毛玻璃病灶(ground-glass opacities)，偵測有疑似結節存在時，</p>
------	---

	<p>會在低劑量電腦斷層影像中標記該 ROI (Region Of Interests)位置，供醫師診斷時參考。</p>
<p>演算法功能</p>	<p>用於肺結節CT影像電腦輔助偵測之演算法功能舉例如下：</p> <ul style="list-style-type: none"> <li>■ 肺結節標記，應說明器材進行影像標記的種類與代表意義，包含影像標記幾何形狀與相關參數，例如幾何形狀(如方框等)、尺寸、邊界框線(如實心或虛線)、顏色、肺結節中心點等。</li> <li>■ 肺結節密度(Density)分類，應說明密度類別[如實質(Solid)、非實質(Sub-solid)(含部份實性和毛玻璃樣(Ground glass)結節)]、密度類別制定依據以及臨床適用性。</li> <li>■ 肺結節量測，應說明量測原理，量測內容可能包含體積、最大軸向平面最長直徑、短徑、平均直徑、最大三維直徑、有效直徑、平均密度、面積等。</li> <li>■ 肺結節的三維或最大密度投影(Maximum intensity projection, MIP)視覺化。</li> </ul>
<p>資料集之族群及用途對應性</p>	<p>資料的來源、性質，亦必須考慮族群取樣的偏差及資料本身之用途對應性。</p> <p>應包含以下(適用)之內容：</p> <ul style="list-style-type: none"> <li>■ 資料來源(健檢或者常規門診)或資料庫說明</li> <li>■ 資料納入和排除之準則</li> </ul> <p>例如：納入雙肺胸部CT影像、掃描區域需要包括從肺尖到底部的整個胸部、影像有正規化等；排除輸入品質不佳之影像(應定義並列舉”品質不佳之情形)，或有某些特定不適用之資料族群(如胸腔植入物之影像)等。</p> <ul style="list-style-type: none"> <li>■ 病患分佈(如人種、年齡、性別)</li> <li>■ 與產品應用相關之病患病史</li> </ul>

	<ul style="list-style-type: none"> <li>■ 病患疾病狀態和造影檢查之適應症</li> <li>■ 資料來源機構(Collection sites)</li> <li>■ 案例數量，含疾病與正常之案例數及決定疾病狀態、位置與範圍的方法</li> <li>■ 依據干擾因子(Confounders)或修飾因子(Effect modifiers)分層的案例分佈，例如病變型態、病變大小(如肺結節大小分為4-6mm、6-8mm以及8-10mm)、病變位置(如肺葉)、器官特徵等</li> <li>■ 比較病患資料與產品目標族群間於臨床造影資料及疾病之特性(Characteristics)</li> </ul>
產製形式	使用何種設備或臨床實務方式產生資料，如使用低劑量電腦斷層產生資料。
產製方法	<p>資料的產製過程，應說明下列事項:</p> <ul style="list-style-type: none"> <li>■ 使用之儀器廠牌與型號、偵測器排數、CT掃描參數(如管電壓、管電流、層厚/層間距、窗寬、窗位、重建方式、顯示方式等)。</li> <li>■ 資料格式(如DICOM等)、CT影像品質要求(如解析度、偽影)。</li> <li>■ 影像擷取程序(Protocol)、造影檢查之方式，例如在有無顯影劑(Contrast)的情況下進行檢查、病人的顯影劑類型和劑量。</li> </ul>
資料集附帶之其它資訊	資料之額外註解和臨床判斷，及產生這些註解與判斷的方法。例如：病變位置、大小或其它資訊等。
臨床結果分析方法	臨床結果分析之主要及次要評估指標應符合產品的預期用途，且於臨床評估或臨床試驗開始前即應確認不可變動。臨床評估或臨床試驗多會觀察產品的靈敏度、特異度及ROC(Receiver Operating Characteristic, ROC)或是ROC曲線的變形，如FROC

	<p>曲線 (Free-Response Receiver Operating Characteristic, FROC)等。產品的效能可採用前述不同的總結臨床評估指標，其結果應符合產品宣稱之效能。</p> <p>例如，對於肺結節偵測演算法，需分層統計演算法對於不同大小和不同密度的肺結節的偵測效能，可採用FROC曲線、靈敏度及特異度等指標來評估，同時結合臨床需求論述結果的合理性。對於肺結節大小，可將肺結節進行分層統計(例如分為4-6mm、6-8mm以及8-10mm)，一般而言，肺結節尺寸越大，偵測的難度越低。對於肺結節密度，應說明是二元分類 (Binary classification) 還是多類別分類 (Multi-class classification)，需定義密度種類及採用該標準的依據，以進行密度類型的分層統計。</p>
--	---

## 附錄二、糖尿病視網膜病變眼底影像電腦輔助偵測醫療器材軟體

本附錄補充本文內容，以舉例方式說明糖尿病視網膜病變眼底影像 CADe 醫療器材的技術重點，提供外界作為該類型產品評估開發及查驗登記所需資料準備之參考，業者應視本身產品宣稱效能、適用範圍、對象及使用情境等評估本附錄中的內容是否適用。

本附錄適用之產品，係運用人工智慧/機器學習技術對由眼底攝影機(採用照相技術用於眼底攝影成相的常規眼底攝影機，不包括雷射、螢光等特殊方式的眼底攝影機)拍攝的彩色眼底照片進行分析。

產品概要	<p>文件應針對產品的預期用途有詳細描述，包含預期使用者及建議的使用者訓練方式、目標病症、病患族群、適用資料類型[如影像類型、影像像素(Pixel)、拍攝範圍、拍攝視野(Field of view)等]、使用情境[如與使用者同時判讀或在使用者判讀後提供第二判讀(Second read)]等，並提供詳細的使用說明、判讀流程、輸出結果、結果解讀的方法及應用。另應說明工作原理，例如運用深度學習技術對眼底影像進行分析處理、對糖尿病視網膜病變進行影像偵測、分類、分割(Segmentation)等。</p> <p><b>【範例】</b>本產品為採用閉鎖式人工智慧演算法之糖尿病視網膜病變電腦輔助偵測軟體(CADe)，適用於20歲以上且已確診為糖尿病患者，供經培訓合格的醫師使用，不能單獨作為臨床診療決策依據。本產品可連結眼底攝影機以取得彩色眼底影像，該影像須為DICOM、JPEG 或 PNG 格式，輸入影像解析度至少1000×1000像素、視野(Field of View, FOV) 為45度。本產品可以在醫師進行常規檢查時，提供糖尿病視網膜病變四種主要病徵的位置標示，包含微血管</p>
------	---

	<p>瘤(Microaneurysms) 、出血(Hemorrhages) 、軟滲出物(Soft Exudates) 、硬滲出物(Hard Exudates) ，供醫師診斷時參考。</p>
演算法功能	<p>用於糖尿病視網膜病變眼底影像電腦輔助偵測演算法功能舉例如下：</p> <ul style="list-style-type: none"> <li>■ 器材可偵測眼底影像有無異常，並可針對微血管瘤(Microaneurysms)、硬滲出物(Hard Exudates)、軟滲出物(Soft Exudates)、出血(Hemorrhages)等進行影像標記。</li> </ul>
族群及用途對應性	<p>資料的來源、性質，亦必須考慮族群取樣的偏差及資料本身之用途對應性。</p> <p>應包含以下(適用)之內容：</p> <ul style="list-style-type: none"> <li>■ 資料來源(如常規門診)或資料庫說明</li> <li>■ 資料納入和排除之準則</li> </ul> <p>例如：納入至少1000×1000像素且對焦在視神經盤中心 (Optic disc centered )或黃斑中心 (Macula centered)之彩色眼底影像);排除輸入品質不佳之醫學影像(應定義並列舉”品質不佳之情形)，或有某些特定不適用之資料族群、不適用之影像類型(如眼底螢光血管攝影等))之準則</p> <ul style="list-style-type: none"> <li>■ 病患分佈(如人種、年齡、性別)</li> <li>■ 與產品應用相關之病患病史</li> <li>■ 病患疾病狀態和檢查之適應症</li> <li>■ 資料來源機構(Collection sites)</li> <li>■ 案例數量，含疾病與正常之案例數及決定疾病狀態、位置與範圍的方法</li> <li>■ 依據干擾因子(Confounders)或修飾因子(Effect modifiers)分層的案例分佈，例如病變型態、病變</li> </ul>

	<p>大小、病變位置、共病(Concomitant disease)等</p> <p>■ 比較病患資料與產品目標族群間於臨床造影資料及疾病之特性(Characteristics)</p>
產製形式	使用何種設備或臨床實務方式產生資料，如使用眼底攝影機產生彩色眼底照片影像資料。
產製方法	資料的產製過程，應說明使用之儀器廠牌與型號、資料格式(如DICOM、JPEG等)、拍攝範圍(如包括黃斑部和視神經盤 (Optic disc))、拍攝視野(如45°)、影像類型(如彩色眼底照片)、影像像素(如1000×1000像素)、影像品質要求(如輸入影像解析度至少1000×1000像素)、影像擷取程序(Protocol)等。
資料集附帶之其它資訊	資料之額外註解和臨床判斷，及產生這些註解與判斷的方法。例如：病變位置、大小或其它資訊等。
臨床結果分析方法	臨床結果分析之主要及次要評估指標應符合產品的預期用途，且於臨床評估或臨床試驗開始前即應確認不可變動。臨床評估或臨床試驗多會觀察產品的靈敏度、特異度及ROC(或其變形)曲線下之面積(Area under the curve, AUC)，並常採用ROC(或其變形)曲線的AUC來總結臨床結果作為主要評估指標，以靈敏度及特異度，或以病患召回率(Recall rate)作為次要評估指標。若有臨床使用之特殊需要，例如，產品用於提供糖尿病視網膜病變狀況之分類建議(如是否需轉診或有無微血管瘤等)，以靈敏度、特異度作為主要評估指標。前述臨床評估指標結果應符合產品之宣稱效能。

## 附錄參考資料

1. Syngo.CT Lung CAD (VD20) 510(k) Summary (K203258), 2021
2. 中國 NMPA 肺結節 CT 圖像輔助檢測軟件註冊審查指導原則,2022 年
3. AI-Rad コンパニオン(認証番号 302AABZX00047000)(2022 年 4 月)
4. US FDA De Novo Classification Request For IDX-DR (DEN180001), 2018
5. EyeArt 510(k) Summary (K200667), 2020
6. 中國 NMPA 糖尿病視網膜病變眼底圖像輔助診斷軟件註冊審查指導原則，2022 年

人工智慧/機器學習技術之  
醫療器材獨立性能評估  
常見問答集

114 年 8 月

## 一、前言

為協助我國醫療器材業者瞭解人工智慧／機器學習（AI/ML）技術之醫療器材在申請查驗登記時之獨立性能評估（Standalone Performance Assessment）的審查重點及獨立性能評估之使用時機，衛生福利部食品藥物管理署特別撰擬「人工智慧／機器學習技術之醫療器材獨立性能評估常見問答集」，針對其常見問題進行說明，以協助業者準備申請查驗登記之資料。

## 二、常見問答集

### (一) 請問什麼是 AI/ML 醫療器材的獨立性能評估(Standalone Performance Assessment)?

獨立性能評估是指在測試資料未再新增醫事人員介入判讀(如人工圈選、標註)的情況下，使用既有之資料進行該醫療器材之性能測試，這種評估主要聚焦於產品性能。

上述所稱「既有之資料」，係指於資料取得時即已具備可用於評估之關鍵資訊(如標註結果)之資料，不涉及再次由醫事人員進行臨床判讀。若資料原本即已包含標註資訊，後續進行影像清晰度確認、排除資料錯誤等處理，若未更動原始標註資訊或重新標註，仍屬「既有之資料」範疇。

「既有之資料」常見來源包括公開可取得且去識別化之資料庫、機構已建置完成並取得使用授權之回溯性臨床資料集。

### (二) 請問獨立性能評估與臨床評估(Clinical evaluation)有何差異？

獨立性能評估是指在測試資料未再新增醫事人員介入判讀(如人工圈選、標註)的情況下，使用既有之資料進行該醫療器材之性能測試，這種評估主要聚焦於產品性能；臨床評估之目的為建立 AI/ML 醫療器材預期使用臨床症狀與臨床實證之關連，其資料特徵應由醫事人員進行確認或判定，以確保 AI/ML 醫療器材能提供預期之技術層面與臨床層面資訊。

### (三) 請問執行 AI/ML 醫療器材進行獨立性能評估時，是否需要倫理審查委員會(Institutional Review Board, IRB)的審查？

執行獨立性能評估是否需要倫理審查委員會(Institutional Review Board, IRB)的審查，取決於是否進行「人體研究」且是否屬主管機關公

告的「得免倫理審查委員會審查之人體研究案件範圍」，並應符合《人體研究法》、《醫療器材管理法》、《醫療器材優良臨床試驗管理辦法》等相關規定。依據衛署醫字第 1010265075 號公告（101 年 7 月 5 日）「得免倫理審查委員會審查之人體研究案件範圍」第 2 款「使用已合法公開週知之資訊，且資訊之使用符合其公開週知之目的」得免送倫理審查委員會審查或由倫理審查委員會核發免審證明。如果獨立性能評估之測試資料來源倘涉人體研究法之規定，則需要 IRB 的審查。

**(四) 請問執行 AI/ML 醫療器材進行獨立性能評估後，是否一定還要再提供完整之臨床證據資料？**

依《醫療器材許可證核發與登錄及年度申報準則》第 6 條及附表二說明九，已有類似品經中央主管機關核准上市者，除須以臨床證據佐證其安全及有效性者外，得免附臨床證據資料。

若 AI/ML 醫材為無類似品，則應提供臨床證據資料(包含學術理論依據與有關研究報告及資料、臨床評估報告或臨床試驗報告)，用以驗證產品在臨床使用時的安全性及有效性。於此情況，若僅進行獨立性能評估，無法完整驗證產品之安全及效能。

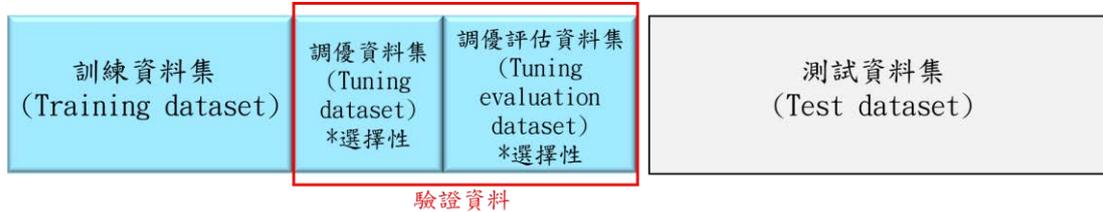
若提供之獨立性能評估不足以佐證產品安全及效能，則須檢附臨床證據資料加以佐證。

**(五) 人工智慧醫療器材於產品開發與測試階段常見的資料集有哪些？**

在進行人工智慧醫療器材之開發與測試時，通常至少需準備訓練資料集(Training dataset)與測試資料集(Test dataset)。視開發需求，亦可增加驗證訓練結果之資料，如模型調優資料集(Tuning dataset)與調優評估資料集(Tuning evaluation dataset)等其他資料集。

訓練資料集(Training dataset)是用於訓練機器學習模型的一組資料集，透過使用訓練資料集建立或改進機器學習模型參數。模型調優資料

集(Tuning dataset)與調優評估資料集(Tuning evaluation dataset)用於調整其非可學習參數，並評估訓練模型。測試資料集(Test dataset)是用於評估訓練後機器學習模型的效能。上述資料集之概念如下圖：



(六)目前僅於《人工智慧/機器學習技術之電腦輔助偵測 (CAdE) 及電腦輔助診斷 (CAdx) 醫療器材查驗登記技術指引》中提及「獨立性能評估」，請問此評估方式是否僅適用於 CAdE/CAdx 類型的人工智慧醫療器材？

獨立性能評估為評估人工智慧醫療器材產品性能之方法，適用範圍不限於 CAdE/CAdx。雖目前僅於 CAdE/CAdx 技術指引中明確提及，但對於其他如醫學電腦輔助分流(Triage)與通知軟體等運用人工智慧/機器學習技術之醫療器材，同樣可透過獨立性能評估驗證產品性能。

(七)請問泛化性測試(Generalizability testing)、重複性與再現性測試有何不同？

泛化性測試是評估產品在面對不同條件時(例如：相同規格於不同品牌之設備)的性能穩定性。重複性與再現性測試則評估在相同或不同的資料擷取設備及操作條件下，對同一病患重複檢查，以衡量輸出結果差異。

(八)請問是否可以使用公開資料庫進行獨立性能評估？

使用公開資料庫進行獨立性能評估時，應確認符合產品的預期用途，建議注意下列議題：

- 資料代表性：雖然這些資料庫提供了大量資料，但它們可能不能完全代表所有群體或臨床情況，評估時應考慮資料的來源和

特點，例如是否涵蓋不同人種、性別、年齡等

- 資料品質：公開資料庫的品質可能參差不齊。在使用前，應確認符合業者產品的預期用途與輸入格式，例如確認資料是否為原始檔或經過處理。
- 時效性：某些資料庫可能包含較舊的資料，評估時應考慮資料的時效性。
- 法律和倫理考慮：使用這些資料庫時，必須遵守相關的資料使用協議和隱私保護規定。

常見之資料庫如下(僅供參考):

- The Cancer Imaging Archive (TCIA)
- NIH Chest X-ray Dataset
- MIMIC (Medical Information Mart for Intensive Care)
- MIT-BIH(The Massachusetts Institute of Technology – Beth Israel Hospital Arrhythmia Database)
- 全民健康保險學術研究資料庫
- 國家高速網路與計算中心生科雲醫療影像資料庫

**(九)請問使用 AI/ML 醫療器材進行獨立性能評估時，應注意哪些事項？**

- 進行獨立性能評估所使用的測試資料集，應與建立演算法模組的訓練資料集、驗證產品初步功能的驗證用資料不同，且應能代表目標族群及目標疾病、病症(Condition)或異常(Abnormality)，測試資料集應包含疾病案例、異常案例與正常案例。
- 應提供資料案例收集的計畫書(Protocol)。
- 可進行壓力測試 (Stress testing)，刻意納入符合預期用途的病

例，但具備較罕見之資料特徵，這些資料是已知會對臨床醫師構成挑戰。

- 建議在執行獨立性能評估時，勿使用虛擬、合成資料。
- 為提升產品泛化能力(Generalizability)，建議測試資料來自不同的臨床機構。

**(十) 請問獨立性能評估通常包括哪些主要評估指標？**

應依據產品特性訂定主要評估指標，常見的主要評估指標包括靈敏度(Sensitivity)、特異度(Specificity)、陽性預測值(Positive predictive value)、陰性預測值(Negative predictive value)、AUC-ROC 曲線(Area Under the Receiver Operating Characteristic curve)等。

### 三、參考資料

- (一) 《醫療器材管理法》
- (二) 《醫療器材優良臨床試驗管理辦法》
- (三) 《醫療器材許可證核發與登錄及年度申報準則》
- (四) 《人工智慧/機器學習技術之醫療器材軟體查驗登記技術指引》
- (五) 《人工智慧/機器學習技術之電腦輔助偵測(CADe)及電腦輔助診斷(CADx)醫療器材查驗登記技術指引》
- (六) 《人工智慧/機器學習技術之電腦輔助分流(Computer Aided Triage)醫療器材軟體查驗登記技術指引》
- (七) US FDA, Computer-Assisted Detection Devices Applied to Radiology Images and Radiology Device Data – Premarket Notification [510(k)] Submissions (2022)
- (八) US FDA, Artificial Intelligence-Enabled Device Software Functions: Lifecycle Management and Marketing Submission Recommendations (Draft Guidance for Industry and Food and Drug Administration Staff) (2025)
- (九) The EU Artificial Intelligence Act (2024)