

# 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

111.8

## 【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，以作為評估開發產品及查驗登記所需資料之參考。

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神經網絡[例如前饋式神經網絡(Feedforward neural network, FNN)、循環神經網絡(Recurrent neural network, RNN)、卷積神經網絡(Convolutional neural network, CNN)等]、貝葉斯網絡(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)

從未於訓練期間向機器學習演算法顯示的一組資料集，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)自動提取醫學圖像(Images)上的可疑病變或異常，並標記其位置、分析醫學圖像及資料或醫學檢測結果等，以輔助偵測病變或異常值，其醫學圖像可能來自超音波、X光攝影、磁振造影(Magnetic resonance imaging, MRI)、電腦斷層

掃描(Computed tomography, CT)、正子造影(Positron emission tomography, PET)、內視鏡或其它數位化醫學圖像等。

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

透過電腦系統(軟體)對於醫學圖像或其資料、醫學檢測結果等進行加工、處理，進而提供量化或分類之診斷支援訊息[例如鑑別良性/惡性病變、病況發展情形(嚴重度、期別)]，並針對診斷結果進行風險評估分析，其醫學圖像可能來自超音波、X光攝影、磁振造影、電腦斷層掃描、正子造影、內視鏡或其它數位化醫學圖像等。

### 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、檢測所需時間和其它必要因素)，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(例如輸入醫學圖像之品質不佳，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，文件亦應提供數據資料產製(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醫療器材的網絡架構及

軟體程序的概述，**建議**包含以下(適用)之內容：

演算法名稱	例如支持向量機(SVM)/回歸、決策樹、卷積神經網絡(CNN)、循環神經網絡(RNN)等。
開發平台	例如Tensorflow、Caffe等。
函式庫	例如Keras等。
模型	例如神經網絡之層數(Number of Layers)、權重(Weights)、激發函數(Activation)、優化器(Optimizer)、損失函數(Loss function)及模型評價指標(Metric)等。
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	<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。

## (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)、假影或移動校正(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其它兩項資料區分，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模組之方法、架構及流程，包含使用的基礎模型、調整部分、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，以及數據資料收集與使用之歷程。

<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>■ 病人分佈(如人種、年齡、性別)</li> <li>■ 與產品應用相關之病人病史</li> <li>■ 病人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 <ul style="list-style-type: none"> <li>-AI/ML CADe醫療器材應含決定疾病狀態、位置與範圍的方法</li> <li>-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</li> </ul> </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(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>■ 比較病人資料與產品目標族群間於臨床造影數據及疾病之特性(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據等。</p>
<p>產製方法</p>	<p>資料的產製過程，包含使用之儀器廠牌與型號、</p>

	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、病人的顯影劑類型和劑量、病人的身體質量指數(Body mass index, BMI)、輻射曝露(Radiation exposure)、不同加權的MRI圖像、加工(如病人資料的數位化)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，可用於決定病人資料中疾病/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. 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)

- 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數據之相關訊息，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、特異性及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，以靈敏度及特異性(或病患召回率(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性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

### 3. 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)及數據產製設備(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 1091610954 號公告，109 年 12 月 24 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, July 2012
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, January 2020
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），厚生労働省醫薬・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. 빅데이터 및 인공지능 (AI) 기술이 적용된 의료기기의 허가·심사 가이드라인 (민원인 안내서) [Big data and artificial intelligence (AI) technology approved medical device permission /examination Guidelines (Complaint Guide) ], National Institute of Food and Drug Safety Evaluation, October 2019