中國醫藥大學附設醫院
醫療器材人工智慧軟體九大透明性宣告書

2024-12-23 訂定
2025-01-16 修訂

Source Attribute Categories for Predictive DSIs

(來源：Requirements for Decision Support Interventions and Predictive Models (Algorithmic Transparency) HTI-1 Final Rule)

1. Details and output of the intervention, including:
   - ▶ Name and contact information for the intervention developer;
   - ▶ Funding source of the technical implementation for the intervention(s) development;
   - ▶ Description of value that the intervention produces as an output; and
   - ▶ Whether the intervention output is a prediction, classification, recommendation, evaluation, analysis, or other type of output.
2. Purpose of the intervention, including:
   - ▶ Intended use of the intervention;
   - ▶ Intended patient population(s) for the intervention's use;
   - ▶ Intended user(s); and
   - ▶ Intended decision-making role for which the intervention was designed to be used/for (e.g., informs, augments, replaces clinical management).
3. Cautioned out-of-scope use of the intervention, including:
   - ▶ Description of tasks, situations, or populations where a user is cautioned against applying the intervention; and
   - ▶ Known risks, inappropriate settings, inappropriate uses, or known limitations.
4. Intervention development details and input features, including at a minimum:
   - ▶ Exclusion and inclusion criteria that influenced the training data set;
   - ▶ Use of variables in paragraph (b)(11)(iv)(A)(5)-(13) as input features;
   - ▶ Description of demographic representativeness according to variables in paragraph (b)(11)(iv)(A)(5)-(13) including, at a minimum, those used as input features in the intervention;
   - ▶ Description of relevance of training data to intended deployed setting.
5. Process used to ensure fairness in development of the intervention, including:
   - ▶ Description of the approach the intervention developer has taken to ensure that the intervention's output is fair; and
   - ▶ Description of approaches to manage, reduce, or eliminate bias.
6. External validation process, including:
   - ▶ Description of the data source, clinical setting, or environment where an intervention's validity and fairness has been assessed, other than the source of training and testing data;
   - ▶ Party that conducted the external testing;
   - ▶ Description of demographic representativeness of external data according to variables in paragraph (b)(11)(iv)(A)(5)-(13) including, at a minimum, those used as input features in the intervention; and
   - ▶ Description of external validation process.
7. Quantitative measures of performance, including:
   - ▶ Validity of intervention in test data derived from the same source as the initial training data;
   - ▶ Fairness of intervention in test data derived from the same source as the initial training data;
   - ▶ Validity of intervention in data external to or from a different source than the initial training data;
   - ▶ Fairness of intervention in data external to or from a different source than the initial training data;

- ▶ References to evaluation of use of the intervention on outcomes, including, bibliographic citations or hyperlinks to evaluations of how well the intervention reduced morbidity, mortality, length of stay, or other outcomes.

8. Ongoing maintenance of intervention implementation and use, including:
   - ▶ Description of process and frequency by which the intervention's validity is monitored over time;
   - ▶ Validity of intervention in local data;
   - ▶ Description of the process and frequency by which the intervention's fairness is monitored over time;
   - ▶ Fairness of intervention in local data.
9. Update and continued validation or fairness assessment schedule, including:
   - ▶ Description of process and frequency by which the intervention is updated; and
   - ▶ Description of frequency by which the intervention's performance is corrected when risks related to validity and fairness are identified.

填寫說明

- ▶ 輸出型態：預測 prediction、分類 classification、推薦 recommendation、評估 evaluation、分析 analysis。
- ▶ 輸出結果：範例如：(a) 電腦輔助標記檢測到可疑鈣化或軟組織病變的位置。(b) 決策支持通過區域評分提供，範圍從 0 到 100，分數越高表示懷疑程度越高。(c) 連接不同乳房視圖中相應區域的鏈接，可用於增強用戶界面和工作流程。(d) 檢查評分將檢查分類為 1 到 10 分，分數越高表示患癌可能性越大。該評分經過校準，使得在無癌症的乳房 X 光照片人群中，每個分類大約占 10%。
- ▶ 預期用途：範例如：這套人工智慧系統旨在輔助放射科醫師辨識乳房 X 光片上的潛在乳癌病灶，幫助提高早期發現的機率，減少診斷遺漏的風險，並提升工作流程。
- ▶ 預期病人族群：範例如：接受乳房 X 光檢查之女性病患。
- ▶ 預期使用者：範例如：放射科醫師。
- ▶ 預期用法：告知 informs、增強 augments、取代臨床管理 replaces clinical management。
- ▶ 警語：範例如：這套人工智慧系統不應獨立用作診斷工具。它不是要取代放射科醫師，而是輔助他們進行診斷。
- ▶ 已知不適用情境：範例如：請勿將其應用於乳房 X 光片以外的影像，例如 MRI 或超音波圖像。另外如果使用的目標人群和 AI 訓練人群有顯著差異性，也應進行機構正確性評估。
- ▶ 本透明性宣告涉及三種資料集。訓練資料 training data 與測試資料 testing data 描述於第四節。外部驗證資料 external data 描述於第六節。
- ▶ 訓練/測試資料集的挑選：請敘述挑選條件或排除條件。範例如：這套人工智慧系統是基於大量來自不同人群的匿名乳房 X 光片資料進行開發的。乳房 X 光片包含 2D 影像 (FFDM 全數位乳房攝影) 和 3D 影像 (DBT 數位乳房斷層攝影) 切片。
- ▶ 輸入特徵：種族 race、民族 ethnicity、使用語言 language、性別傾向 sexual orientation、性別認同 gender identity、性別 sex、年齡 date of birth、健康資料的社會決定因素 social determinants of health data、健康狀況評估 health status assessments 等為 ONC certification criteria for Health IT 170.315(b)(11)(iv)(A)(5)-(13) 特別強調之輸入特徵。一般而言，真正常用之特徵請於「其他」說明。範例如：輸入特徵包括可疑的鈣化點和軟組織病變 (包括密度、腫塊、結構變形和不對稱)、病患年齡和乳房密度。
- ▶ 輸入特徵的分布及合理性：範例如：開發過程中使用了先進的機器學習算法，如卷積神經網絡 (CNN)。這些算法使用大量經病理切片證實的乳癌樣本、良性異常樣本和正常組織樣本進行訓練。

- 確保結果公平性的方法：公平性 fairness。範例如：我們使用了檢測和緩解偏差的方法，確保訓練資料的平衡，並進行了公平性審查。
- 管理、減少或消除偏見的方法：偏見 bias。範例如：在開發過程中，對資料集進行了分析，以確保其代表不同年齡、種族和乳房密度的人群。
- 與訓練/測試資料差異：範例如：這套人工智慧系統在不同醫療機構使用的外部驗證資料集上進行了驗證，這些資料集不包含在初始訓練資料中。
- 參與單位/資料來源：範例如：外部驗證資料來自七個歐盟國家和美國的多個臨床中心。
- 外部驗證資料輸入特徵的分布：範例如：外部驗證資料包含來自不同製造商的 2D 和 3D 乳房 X 光片 (2D: Hologic、GE、Philips、Siemens and Fujifilm，3D: Hologic、Siemens and Fujifilm)，代表了常規乳癌篩檢和無症狀患者。外部驗證資料包含 7882 例無癌症檢查和 1240 例有癌症檢查。在無癌症檢查中，4797 例是 2D 影像，3085 例是 3D 影像。在有癌症檢查中，819 例是 2D 影像，421 例是 3D 影像。總體而言，外部驗證資料中有癌症的檢查中，61%的病灶被歸類為腫塊，33%為可疑鈣化點，6%為結構變形或不對稱。三種主要的組織學癌症類型分別是浸潤性導管癌(60.5%)、原位導管癌(25.9%) 和浸潤性小葉癌(9.0%)。病灶範圍中位數 (定義為兩維度的最大直徑) 在 2D 資料中為 16mm (四分位距：11-24)，在 3D 資料中也是 16mm (四分位距：11-25)。
- 驗證程序：範例如：驗證過程確認了系統在各種臨床環境和病患群體中的準確性和適用性。
- 量化指標：請考量到測試/外部驗證資料的有效性 validity 與公平性 fairness。範例如：通過計算至少在一個視角 (MLO 或 CC) 中正確定位的癌症比例來計算基於檢查的性能指標如下：準確率：95%、靈敏度：94.7%、特異性：90%、精確度：91%、AUC: 0.949、召回率：92%。
- 參考文獻：如：評估如何降低發病率、死亡率、住院時間、...的參考文獻超連結。
- 持續維護：請考量有效性與公平性的監控程序與頻率。範例如：我們有專門的團隊負責在系統部署後的監控，處理任何問題。系統會定期更新，我們也提供使用者支持和故障排除服務。
- 更新與持續驗證：請考量定期更新與發現瑕疵修正時的情境。範例如：此人工智慧系統每年會隨機選 300 例以上真實世界影像進行機構獨立性能指標評估，確保系統有穩定的正確性。更新和驗證結果會定期提供給臨床醫師，當靈敏度低於預設 85%服務會暫停，系統會補充新影像資料訓練，性能提升至 90%後重新恢復使用。
- 相關認證說明書：已取得 TFDA, FDA 等認證的說明書。