

# AI輔助診斷與研究之倫理議題

陳祖裕

大林慈濟醫院人體研究保護計畫總負責人

醫學研究倫理基金會董事暨聯合人體試驗委員會執行秘書

台中慈濟醫院人體試驗委員會委員

中山醫學大學附設醫院受試者保護諮議委員會委員

教育部人體研究倫理審查委員會查核作業諮議會議委員暨查核委員

財團法人醫院評鑑暨醫療品質策進會人體研究倫理審查會查核委員

前美國人體試驗委員會專業人員 (CIP)



# 報告大綱

---

- 前言
- AI的優點與角色
- AI的問題與因應
- AI的倫理議題
- AI研究的倫理審查
- 結語



Jointly published with



September 2022

United States Government Accountability Office

---

Report to Congressional Requesters

TECHNOLOGY ASSESSMENT

# Artificial Intelligence in Health Care

Benefits and Challenges of Machine Learning  
Technologies for Medical Diagnostics

With content from the National Academy of Medicine



## PART TWO

---

# Meeting the Moment: Addressing Barriers and Facilitating Clinical Adoption of Artificial Intelligence in Medical Diagnosis

National Academy of Medicine

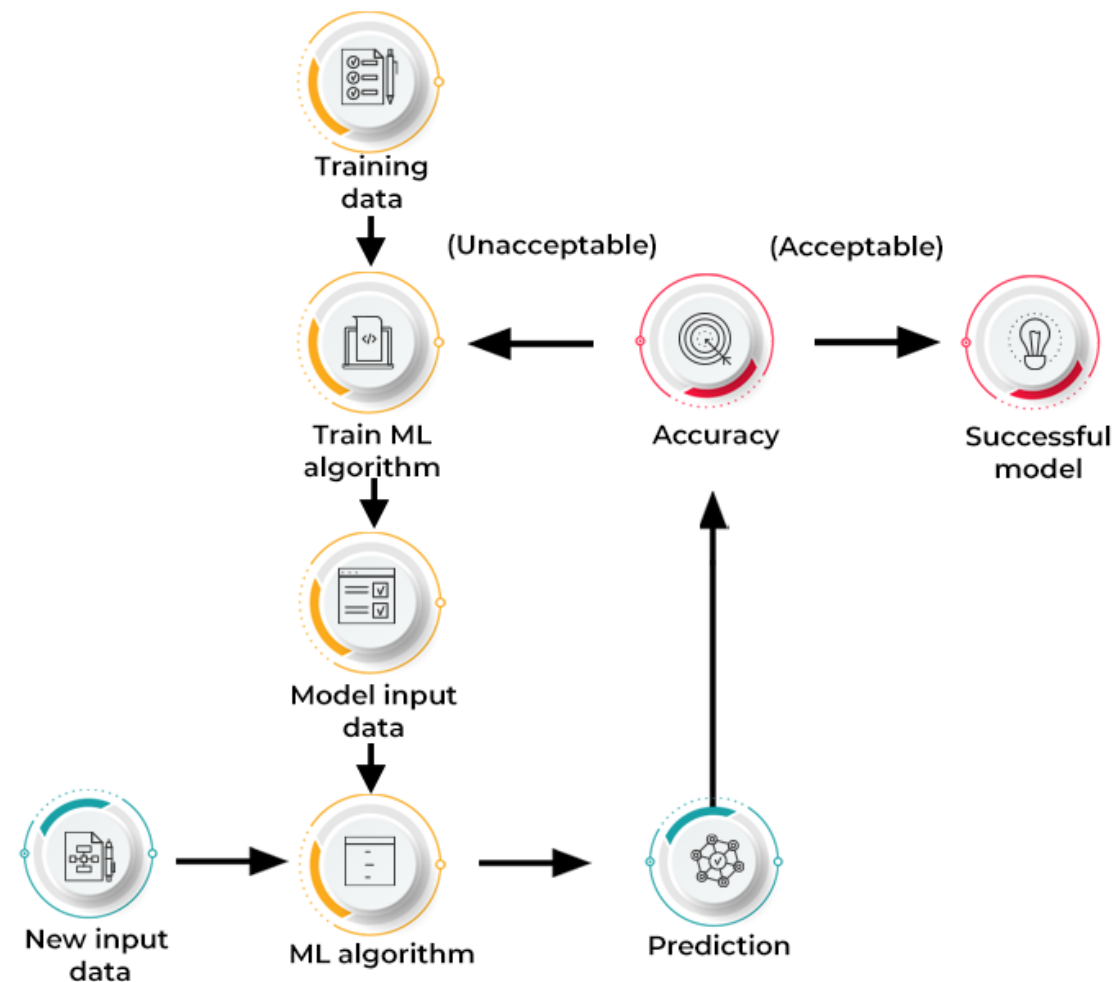
Part Two presents the NAM publication *Meeting the Moment: Addressing Barriers and Facilitating Clinical Adoption of Artificial Intelligence in Medical Diagnosis* discussing the factors influencing the adoption of non-autonomous point-of-care AI technology that can assist in the diagnosing of a disease. Although GAO and NAM staff consulted with and assisted each other throughout this work, reviews were conducted by GAO and NAM separately and independently, and authorship of the text of Part One and Part Two of the report lies solely with GAO and NAM, respectively.



# 前言

- **人工智慧 (AI)** 是指透過分析環境並採取行動 (具有一定程度的自主性) 來實現特定目標而表現出智慧行為的系統
- **AI** 結合了 **機器學習 (ML)** 技術、**機器人技術**、**演算法** 和 **自動化診斷決策支持 (diagnostic decision support, DSS)**
- **ML** 演算法和模型可以讓電腦系統根據模式和推理來執行任務，而不需要由人類輸入精確的指令

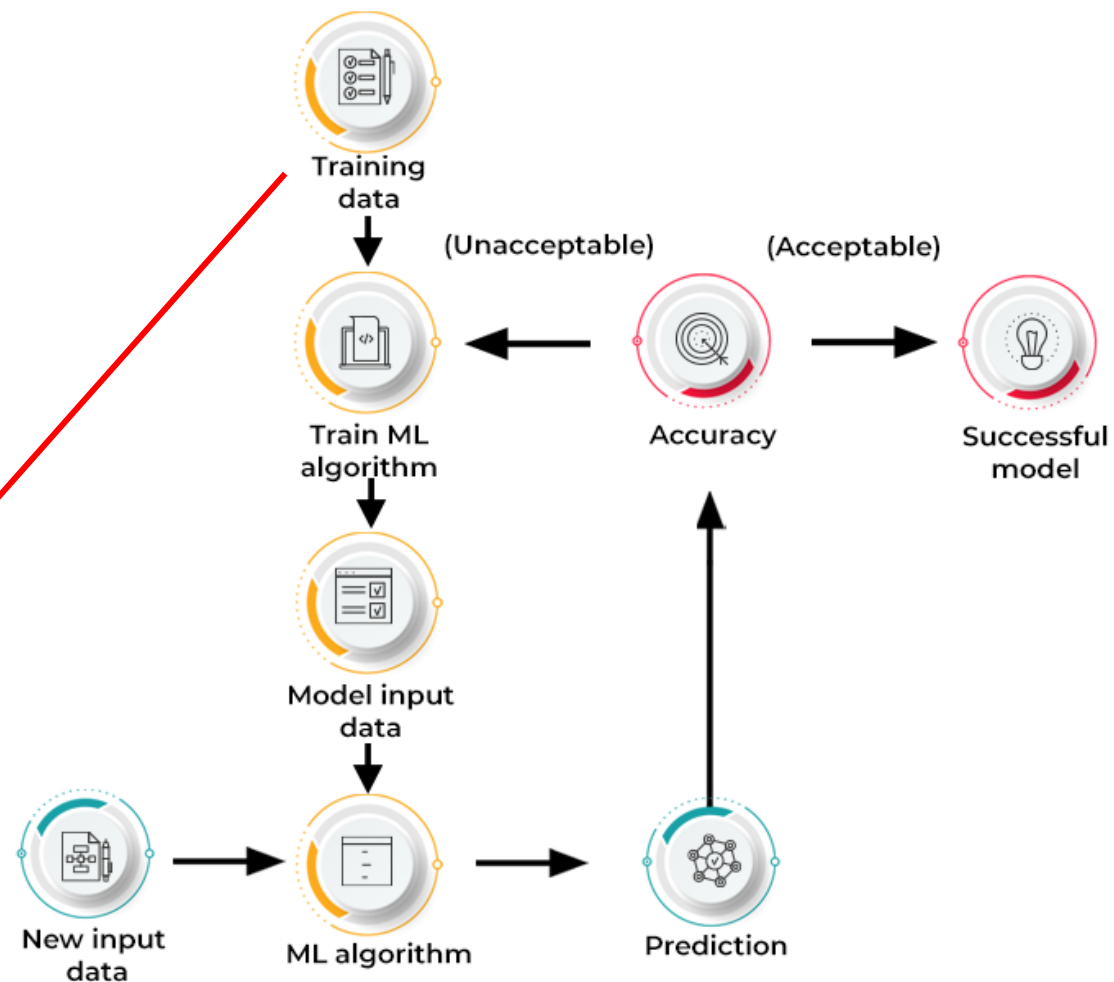
## HOW DOES MACHINE LEARNING WORK?



# 什麼是ML?

- ❑ ML演算法處理大量的歷史資料，並從中找出資料的模式
- ❑ 這讓電腦能更精準地根據輸入的資料集預測結果
- ❑ 例子：  
科學家使用數百萬張X光片和對應的診斷結果來訓練一支ML醫療應用程式  
讓該程式可以僅透過X光片就能診斷病人是否罹患癌症

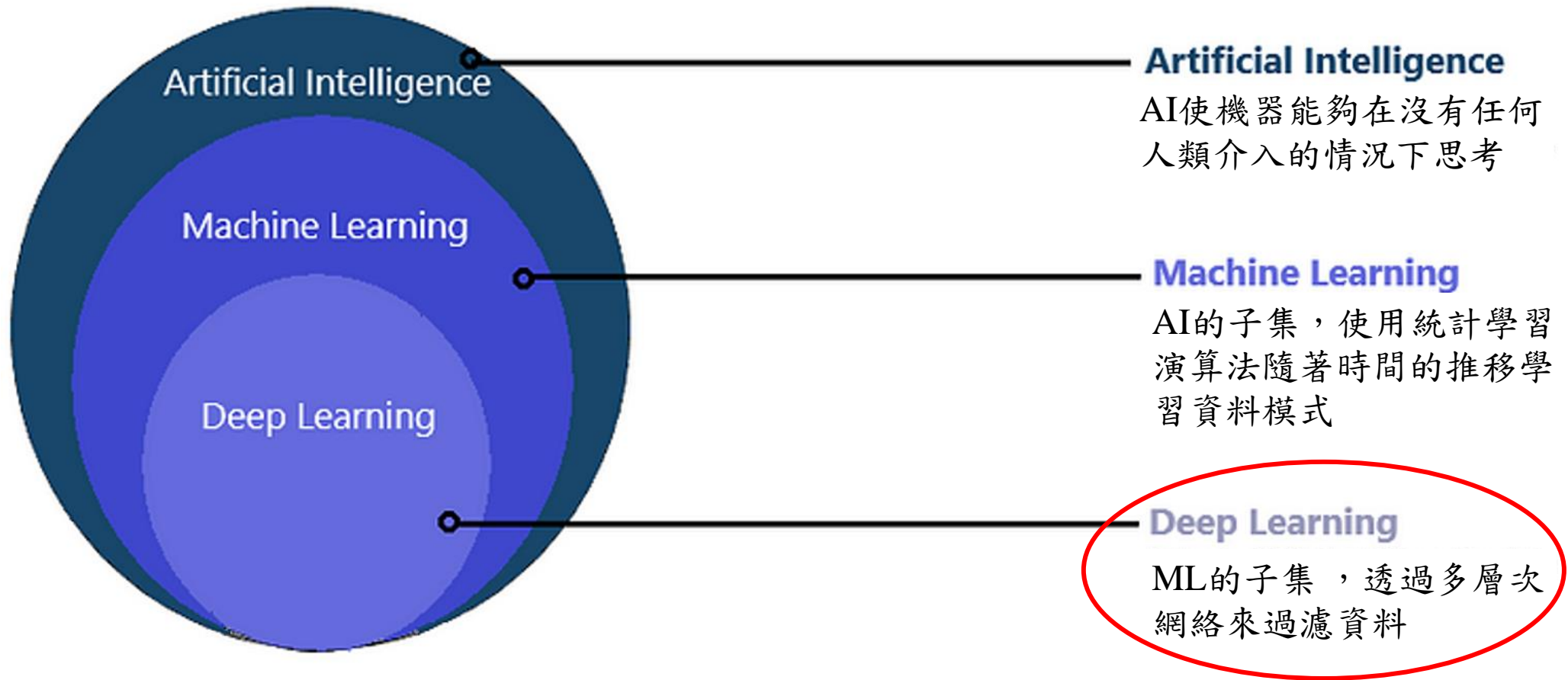
## HOW DOES MACHINE LEARNING WORK?



# ML演算法有哪些類型？

---

- 演算法可以根據輸入和輸出的類型分為四種：
  1. 監督式機器學習
  2. 無監督機器學習
  3. 半監督學習
  4. 強化機器學習



<https://www.analyticsvidhya.com/blog/2021/03/everything-you-need-to-know-about-machine-learning/>



# 機器學習有什麼優點和缺點？

---

## □ ML模型的優點：

- 可以辨識出人類可能會錯過的資料趨勢或模式
- 在設定後，可以自動作業，不需人類操作
- 可以在動態且複雜的資料環境下處理大量不同格式的資料
  - 例如：網路安全軟體中的ML可以在沒有管理員操作的情況下持續監控和抓出網路通訊中的違規行為
- 結果會隨著時間愈來愈精準

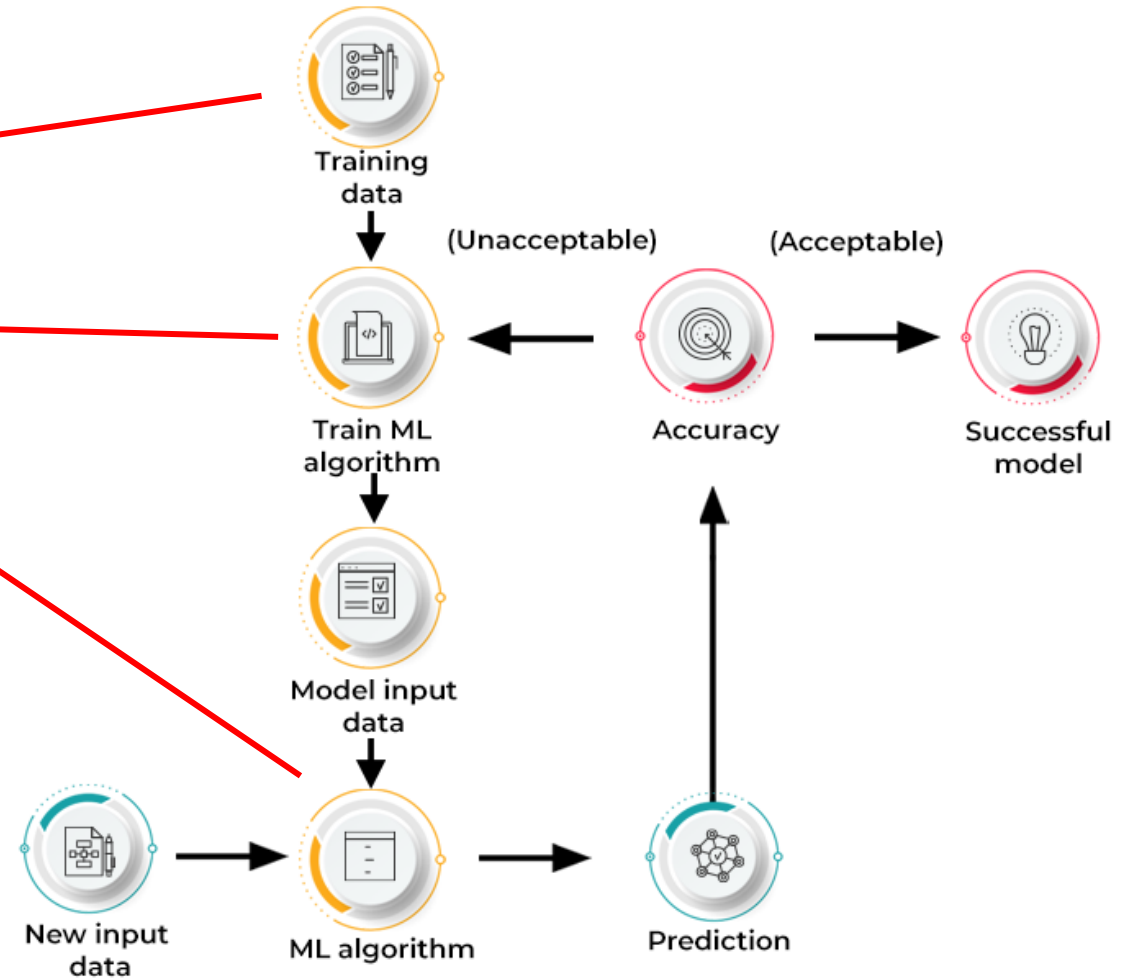
<https://aws.amazon.com/tw/what-is/machine-learning/>

# 機器學習有什麼優點和缺點？

## ❑ ML模型的缺點：

- ❑ 如果沒有充足的資料，難以訓練出可用的模型
- ❑ 一開始的訓練很耗費時間和成本
- ❑ 需要大量的運算能力，硬體設施需要大量的前期投資
- ❑ 需要專業人員協助，否則不容易正確地解釋輸出的結果和消除不確定性

<https://aws.amazon.com/tw/what-is/machine-learning/>



# 前言（續）

---

- ❑ 隨著大數據的使用不斷增加，AI也在快速成長，在安全、環境、研究、教育、醫療和貿易等領域找到了新的應用
- ❑ 醫療人員在診斷和治療使用AI-DSS（decision support system）：
  - ❑ 提高工作流程的效率
  - ❑ 增加精確度
  - ❑ 減少錯誤
- ❑ 儘管AI在醫療照護領域的應用率不斷增加，但目前還沒有普遍被接受的綜合框架來指導醫療照護領域基於AI的決策支援的開發和實施
- ❑ AI的引入伴隨著既廣泛又複雜的倫理問題
- ❑ 在引入AI時應採取更加謹慎的態度



# 前言（續）

---

- AI固有的幾個重要議題：
  - 避免偏差（例如，由於使用不充分或不良的測試和訓練資料集來開發AI演算法，便會導致不良或負面結果）
  - 保護病人隱私
  - 取得病人和醫師的信任
  - 實施中的後勤困難
  - 必要的社會文化或臨床路徑的改變

# 報告大綱

---

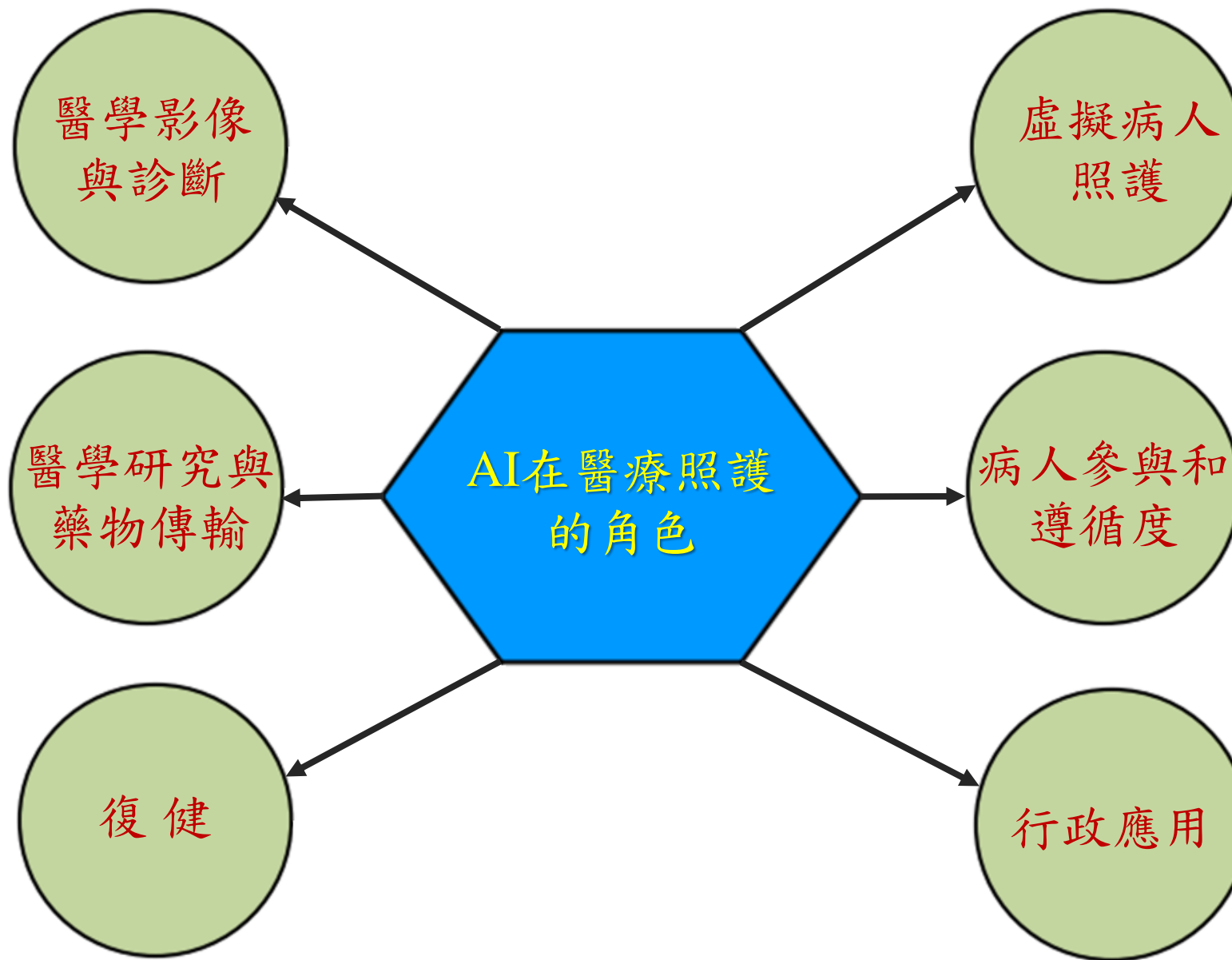
- 前言
- AI的優點與角色
- AI的問題與因應
- AI的倫理議題
- AI研究的倫理審查
- 結語

*Review*

# **A Review of the Role of Artificial Intelligence in Healthcare**

Ahmed Al Kuwaiti <sup>1,\*</sup> , Khalid Nazer <sup>2,3</sup>, Abdullah Al-Reedy <sup>4</sup>, Shaher Al-Shehri <sup>5</sup>, Afnan Al-Muhanna <sup>6,7</sup> ,  
Arun Vijay Subbarayalu <sup>8</sup> , Dhoha Al Muhanna <sup>9</sup> and Fahad A. Al-Muhanna <sup>10,11</sup> 





## AI在醫療護各方面的應用


# AI用於臨床照護

---

- ❑ 醫學影像
  - ❑ AI可減少分析掃描所需的成本和時間
  - ❑ AI在檢測肺炎、乳癌、皮膚癌以及眼部疾病等疾病方面已顯示出有希望的結果
- ❑ 超音波心臟檢查—牛津約翰拉德克利夫醫院的Ultromics系統用來分析超音波心動圖，從而檢測心跳模式並診斷冠心病
- ❑ 篩檢神經系統疾病—分析言語模式以預測精神病發作、識別和監測帕金森氏症等神經系統疾病的症狀
- ❑ 手術—由AI控制的機器人工具已用於研究中

# AI用於以病人和消費者為對象的應用程式

- 市場上已有個人化**健康評估**和**家庭照**的應用程式：
  - Ada Health Companion：聊天機器人，能將用戶的症狀資訊與其他資訊結合起來，提供可能的診斷
  - GP at Hand：正在倫敦進行試用，超過10萬民眾參與

Health.  
Powered by Ada.  ada

Supporting better health outcomes and clinical excellence with intelligent technology.

Let's start with the symptom that's troubling you the most.

13,621,219  
users

32,827,218  
symptom assessments







*"Very friendly and easy to use, and turns out really accurate when you get a professional diagnosis later on. I really like how it takes to account a variety of factors when asking about symptoms. Don't expect it to replace professional medical diagnosis tho!"*



COVID-19 Vaccine Our NHS services What we treat GP clinic locations

## 24/7 appointments for free\*

-  On mobile, 24/7
-  In person at a choice of locations
-  Prescriptions delivered to your local pharmacy
-  An NHS GP Practice that works around you

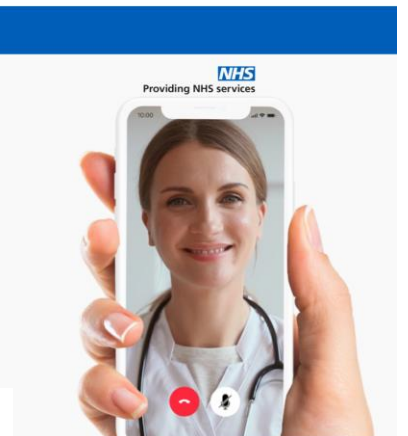
Get Started

Over 100,000 people already have GP at Hand as their NHS GP, with a 96%

satisfaction score



4.8 average user rating





# AI用於以病人和消費者為對象的應用程式

---

- 由AI驅動的資訊工具或聊天機器人用來幫助處理慢性病，如：
  - IBM為英國關節炎研究中心開發的關節炎虛擬助理，透過與病人互動來學習，以提供藥物、飲食和運動的個人化資訊和建議
  - 利用機器人系統支持患有早期失智症等疾病的病人家屬
  - 監測和支持肺結核病人遵循處方藥物和治療

IBM



IBM Newsroom

Announcements

Biographies

Media Contacts

The King and Honours

Global News Rooms

UK Think Blog

Subscribe

Arthritis Research UK Introduces IBM Watson-Powered 'Virtual Assistant' to Provide Information and Advice to People with Arthritis

Mar 14, 2017

# AI用於以病人和消費者為對象的應用程式

---

- 由AI驅動的資訊工具或聊天機器人用來幫助處理慢性病，如：
  - Sentrian分析病人在家中佩戴的感測器收集的資訊，以及早發現病情惡化，以便進行介入並減少住院

Sentrian, the World's First  
Remote Patient Intelligence  
Company, Launches to  
Eliminate Preventable  
Hospitalization



Company raises \$12 Million in financing led by Reed Elsevier Ventures,  
Frost Data Capital and TELUS Ventures

Leadership team includes former Chief Medical Scientist of IBM  
Research

NEWS PROVIDED BY  
**Sentrian** →  
10 Nov, 2014, 08:03 ET

# AI在醫學研究的優勢

- ❑ 可搜尋大量科學文獻，以及組合不同的資料集
- ❑ 能協助研究人員發現資料集中複雜的關聯
- ❑ 能高精度地觀察和處理大量資訊
- ❑ 以更快、更精確且更具成本效益的方式分析數據
- ❑ 可以幫助將合適的病人與臨床研究相匹配
- ❑ 可以澄清臨床醫師意見不一致的領域，如透過胸部X光識別結核病
- ❑ ML演算法使藥物發現過程更快、更經濟，如：
  - ❑ 德州大學MD Anderson Cancer Center的canSAR資料庫：能將病人的遺傳和臨床數據與科學研究資訊結合，並利用AI對癌症藥物的新標靶進行預測

# AI在醫學研究中的用途

---

- 醫療診斷
- 個人化/精準醫療
- 藥物發現
- 臨床試驗
  - 機器學習演算法可以優化臨床試驗研究
  - 為了提高試驗的安全性：對參與者的即時數據存取和遠端監控使研究人員能夠更準確地記錄生物變化，並確定參與者是否對治療做出不良反應。
- 流行病學

# 報告大綱

---

- 前言
- AI的優點與角色
- AI的問題與因應
- AI的倫理議題
- AI研究的倫理審查
- 結語



# AI在醫療照護領域的問題

---

- ❑ 必須擁有龐大的資料集才能正確地分類或預測各種任務
- ❑ 培訓醫療專業人員成功使用AI所投入的時間和資源成本高昂
- ❑ 演算法的預測不易作出實質解釋
- ❑ 專家很難理解數據與他們的預測之間的關係
- ❑ AI系統中的黑盒子問題會導致大眾對醫療保健系統失去信心
- ❑ 當推薦出錯時，演算法很難合法地保護自己
- ❑ 因為病人紀錄是保密的，醫療照護機構間在數據可造訪性複雜
- ❑ 臨床試驗中驗證AI藥物療效的實驗數據不足
- ❑ AI研究主要是在非臨床環境中進行，研究結果的推廣很困難

# AI在醫療照護領域的問題

---

- ❑ 某些機構因**缺乏經驗**導致數據和研究品質不佳，因而不願意執行基於AI的解決方案
- ❑ 系統可以不斷進步，但**企業內部的阻力**會阻礙進步
- ❑ AI的應用引發**資料安全和隱私**的問題（紀錄非常重要且容易受到駭客攻擊）
- ❑ AI的其他缺點包括：
  - ❑ 使人類變得懶惰
  - ❑ 造成失業
  - ❑ 缺乏情感和創造力

# 推展AI在醫療照護的策略

- 使用理由
  - 結盟
  - 激勵措施
- 使用手段
  - 基礎設施
  - 資源
- 使用方法
  - 工作流程
  - 效率
- 使用意願
  - 執行
  - 相信

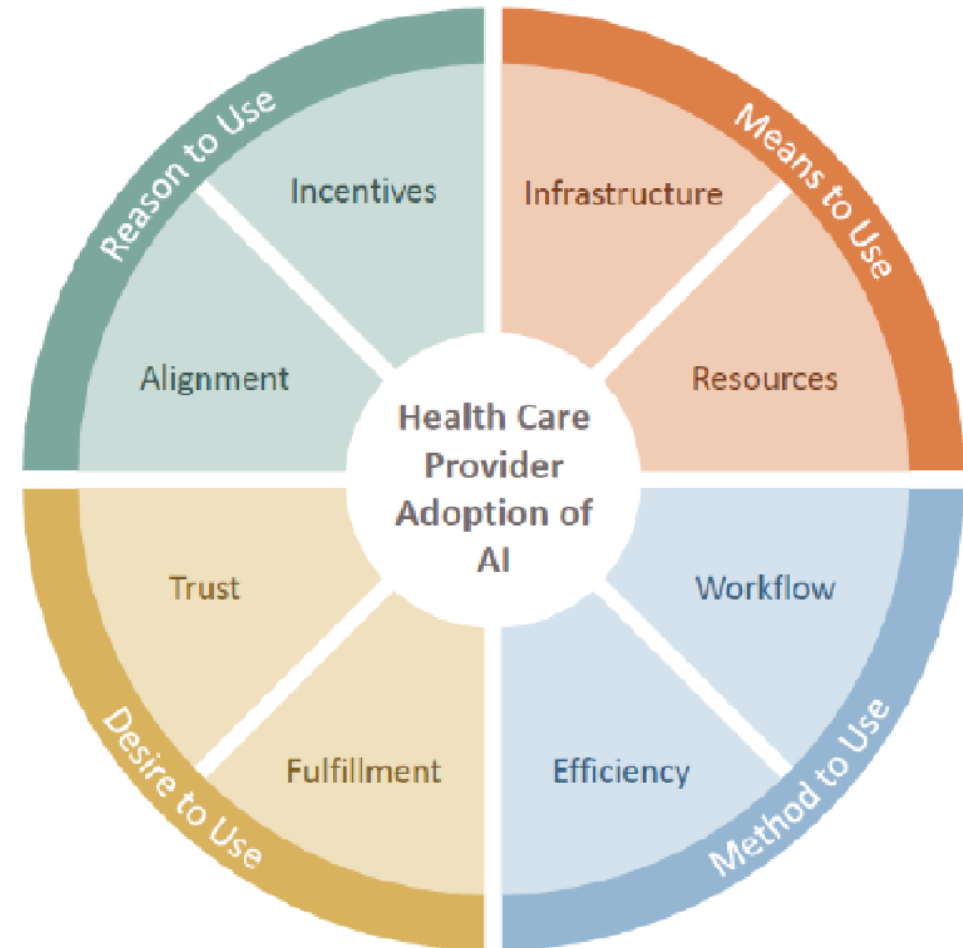


Figure 2: Core Domains to Support Provider Adoption of AI-DDS Tools

# 報告大綱

---

- 前言
- AI的優點與角色
- AI的問題與因應
- AI的倫理議題
- AI研究的倫理審查
- 結語

### Artificial intelligence (AI) in healthcare and research





# AI用於醫療照護引發的倫理和社會問題

---

- ❑ 可靠性和安全性
- ❑ 透明度和問責制
- ❑ 數據偏差、公平性和平等
- ❑ 信任
- ❑ 對病人的影響
- ❑ 對醫療專業人員的影響
- ❑ 資料隱私和安全
- ❑ 惡意使用AI

# 可靠性和安全性

- 可靠性和安全性
- 透明度和問責制
- 數據偏差、公平性和平等
- 信任
- 對病人的影響
- 對醫療專業人員的影響
- 資料隱私和安全
- 惡意使用AI

- 當AI用於控制設備、提供治療或做出醫療決策時，可靠性和安全性是關鍵問題
- AI可能會犯錯，如果錯誤難以發現或產生連鎖反應，則可能會產生嚴重影響
- 如2015年的一項臨床試驗，AI應用程式被用來預測哪些病人可能在肺炎後出現併發症而應該住院治療，由於無法考慮上下文訊息，該應用程式錯誤地指示醫生將氣喘患者送回家
- 使用AI的症狀檢查應用程式的效能受到質疑，例如，人們發現應用程式的推薦可能過於謹慎，可能會增加對不必要的檢查和治療的需求

# 透明度和問責制

- 可靠性和安全性
- 透明度和問責制
- 數據偏差、公平性和平等
- 信任
- 對病人的影響
- 對醫療專業人員的影響
- 資料隱私和安全
- 惡意使用AI

- 判定產生AI輸出的底層邏輯可能很困難或不可能
- 有些AI是專有的並且故意保密，但有些AI對於人類來說太複雜而難以理解
- ML技術特別不透明，因為它們在學習時不斷調整自己的參數和規則
- 驗證AI系統的輸出以及識別數據中的錯誤或偏差十分困難
- 相關問題：
  - 誰對AI做出的決策負責
  - 因使用AI而受到傷害的人如何尋求補償及救助

# 數據偏差、公平性和平等

- 可靠性和安全性
- 透明度和問責制
- 數據偏差、公平性和平等
- 信任
- 對病人的影響
- 對醫療專業人員的影響
- 資料隱私和安全
- 惡意使用AI

- 因訓練數據的不足或偏差導致有瑕疵的ML模式，
- 因AI開發人員的信念和偏差可導致演算法本身存在偏差
- AI可能會以隱藏的方式對性別、種族、殘疾和年齡等受特徵引發歧視
- 因AI系統的偏差可能會做出不公平的決定
- AI在醫療照護領域的好處可能分佈不均，
  - 在資料稀缺或更難以數位方式收集或呈現的情況下，AI的效果可能會較差
  - 如患有罕見疾病的人或少數族裔人口是AI醫療照護的弱勢族群

# 信任

- 可靠性和安全性
- 透明度和問責制
- 數據偏差、公平性和平等
- 信任
- 對病人的影響
- 對醫療專業人員的影響
- 資料隱私和安全
- 惡意使用AI

- DeepMind（一家英國AI公司）與倫敦皇家自由醫院之間的合作引發了關於商業公司是否有權存取病患資料的公開辯論
- 如果人們無法相信AI是為了公共利益而開發的，公眾可能會強烈反對AI
- 在醫療執業層面上，如果AI系統要在醫療保健領域成功實施，病人和醫療保健專業人員都需要能夠信任它們
- IBM的Watson for Oncology（一種提供癌症治療建議的AI工具）的臨床試驗在一些診所被叫停，因為美國以外的醫師對其建議沒有信心，認為該模型反映了美國特有的癌症治療方法



# 對病人的影響（正面）

- 可靠性和安全性
- 透明度和問責制
- 數據偏差、公平性和平等
- 信任
- 對病人的影響
- 對醫療專業人員的影響
- 資料隱私和安全
- 惡意使用AI

- AI可協助人們評估自己的症狀，並在可能的情況下照顧自己
- 旨在支持患有慢性疾病或殘疾的人的AI系統，可以提高人們的尊嚴、獨立性和生活品質；並使原本可能被送往護理機構的人們能夠在家中待更長時間

# 對病人的影響（負面）

- 可靠性和安全性
- 透明度和問責制
- 數據偏差、公平性和平等
- 信任
- 對病人的影響
- 對醫療專業人員的影響
- 資料隱私和安全
- 惡意使用AI

- 如果使用AI技術取代病人與家屬或照護者相處的時間，可能會增加人際接觸的喪失和社交孤立
- AI系統可能會對個人自主權產生負面影響：例如，如果它們根據風險計算或用戶最佳利益的計算來限制選擇
- 如果AI系統用於進行診斷或製定治療計畫，但醫療人員無法解釋這些是如何得出的，這可能會被視為限制病人對其健康做出自由、知情決定的權利
- 旨在模仿人類伴侶或照護者的應用程式可能會導致用戶無法判斷他們是在與真人還是在與技術進行交流，這可能被視為一種欺騙或欺詐。

# 對醫療專業人員的影響

- 可靠性和安全性
- 透明度和問責制
- 數據偏差、公平性和平等
- 信任
- 對病人的影響
- 對醫療專業人員的影響
- 資料隱私和安全
- 惡意使用AI

- AI可以實現以前由人類執行的任務的自動化，這可以讓醫療人員騰出更多時間與病人接觸
- 醫療人員的專業知識受到AI的挑戰，其自主權和權威會受到威脅
- 醫療人員對個別病人的倫理義務可能會受到使用AI決策支援系統的影響
- AI系統的引入可能會導致解僱低技能員工
- 如果技術失敗且工作人員在沒有電腦指導的情況下無法識別錯誤或執行必要的任務
- AI可能會讓醫療保健專業人員沾沾自喜，不去檢查結果和挑戰錯誤

# 資料隱私和安全

- 可靠性和安全性
- 透明度和問責制
- 數據偏差、公平性和平等
- 信任
- 對病人的影響
- 對醫療專業人員的影響
- 資料隱私和安全
- 惡意使用AI

- 醫療照護中的AI應用程式利用敏感和私密的數據，這些都受到法律控制
- 其他類型的與健康狀況無關的數據，如社交媒體活動和網路搜尋歷史紀錄，也可以用來揭示有關用戶及其周圍人的健康狀況的資訊
- AI可用於偵測網路攻擊並保護醫療電腦系統，但AI系統有可能被駭客攻擊以獲取敏感數據，或以不易察覺的方式發送虛假或有偏見的數據

# 惡意使用AI

- 可靠性和安全性
- 透明度和問責制
- 數據偏差、公平性和平等
- 信任
- 對病人的影響
- 對醫療專業人員的影響
- 資料隱私和安全
- 惡意使用AI

- AI也可能被用於惡意目的，例如，AI可能被用於秘密監視或篩檢
- 分析運動行為（例如某人在鍵盤上打字的方式）以及透過追蹤智慧型手機檢測到的行動模式的AI技術可以在一個人不知情的情況下揭示有關其健康的資訊
- AI可用於以更低的財務成本和更大規模的網路攻擊

# THE BELMONT REPORT

Ethical Principles and Guidelines for the  
Protection of Human Subjects of Research

## Respect for Persons

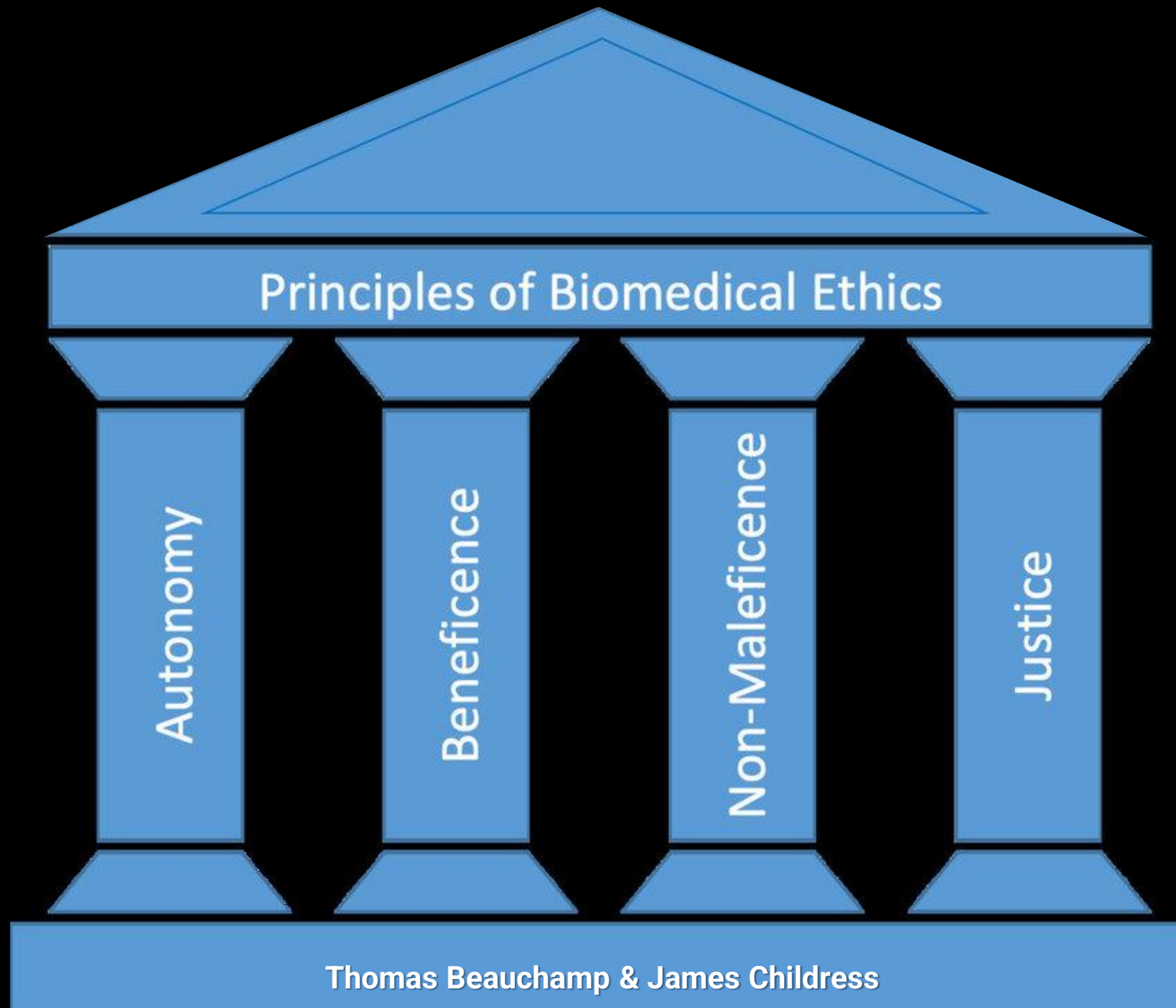
**Beneficence**

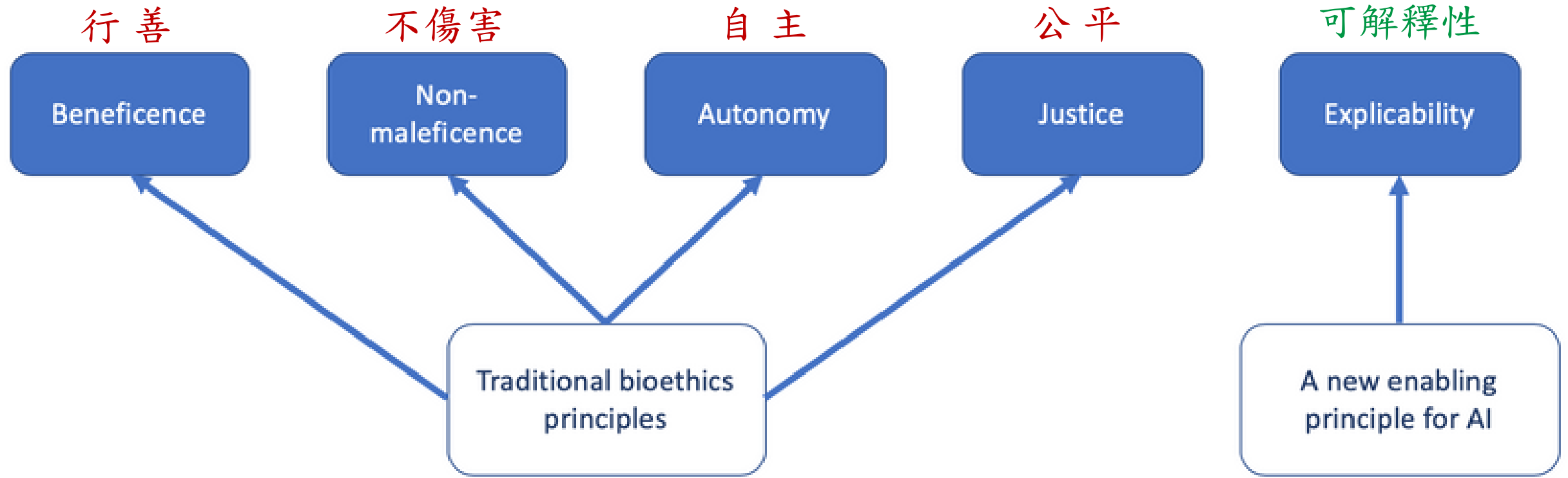


**Justice**

基本原則不變？









# Ethical challenges regarding artificial intelligence in medicine from the perspective of scientific editing and peer review

Seong Ho Park<sup>1</sup>, Young-Hak Kim<sup>2</sup>, Jun Young Lee<sup>3</sup>, Soyoung Yoo<sup>4</sup>, Chong Jai Kim<sup>5</sup>

<sup>1</sup>Department of Radiology and Research Institute of Radiology, <sup>2</sup>Cardiology Division, Asan Medical Center, University of Ulsan College of Medicine, Seoul; <sup>3</sup>National IT Industry Promotion Agency, Jincheon; <sup>4</sup>Health Innovation Big Data Center, Asan Medical Center, Seoul; <sup>5</sup>Department of Pathology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Korea

# 透明度的層面

---

- 訓練資料的透明度
- 測試數據和測試結果的透明度
- 解釋研究結果的透明度
- 透過共享演算法和數據提高透明度

## 在評估AI演算法效能的研究中提高倫理透明度所需提出的問題

---

### 要問的問題

---

#### 關於訓練數據

作者是否徹底解釋了他們如何收集、處理和組織資料？

作者是否徹底描述了資料/患者的特徵，包括人口特徵、臨床特徵和技術特徵？

作者是否明確揭露了數據中預期的偏差以及偏差可能導致的意外後果和陷阱？

#### 關於測試數據和結果

作者是否明確說明測試資料是前瞻性收集的還是回顧性收集的？

作者是否清楚說明測試資料是從中提取訓練資料的初始資料集的子樣本還是獨立的外部資料？

對於外部數據，作者是否明確說明測試數據代表便利性系列還是臨床隊列？

對於臨床隊列，作者是否清楚地解釋了患者的資格標準以及他們代表的具體臨床環境？

如果使用多個機構的測試資料集，作者是否報告每個機構的結果？

作者是否澄清（透過提供註冊中心名稱和研究識別碼）他們是否有意在可公開訪問的註冊中心註冊該研究？

#### 關於研究結果的解釋

作者是否明確解釋結果並避免過度解釋？

#### 關於演算法和數據的共享

如果作者願意分享，他們是否解釋瞭如何在報告中存取他們的演算法和資料（例如，放置網頁連結以供下載）？

對於共享數據，作者是否解釋了他們如何確保患者隱私和數據保護？

# 報告大綱

---

- 前言
- AI的優點與角色
- AI的問題與因應
- AI的倫理議題
- AI研究的倫理審查
- 結語



# AI研究的倫理審查重點

---

## The challenges of AI integration in medical research

- ❑ Privacy 隱私
- ❑ Regulation 法規
- ❑ Transparency 透明度
- ❑ Culture 文化
- ❑ Recruitment 招募

## Considerations for IRB Review of Research Involving Artificial Intelligence

2022年7月21日核准

SACHRP (Secretary's Advisory Committee on Human Research Protections) 被要求對隨後的質疑作出回應。鑑於人工智慧 (AI) 和機器學習 (ML) 快速發展的性質，以及這些術語的不精確定義和不同理解，委員會並未就其認為具有權威性的簡明背景框架達成共識。另一方面，考慮到這些技術的使用同樣的演變和普遍性，委員會認為必須立即對指控做出回應，這一點很重要。

## Considerations for IRB Review of Research Involving AI

大眾向SACHRP 之質疑：

1. 在什麼條件下，AI或AI驗證活動的資料收集符合Common Rule  
「旨在開發或貢獻通用知識」的研究定義？
2. 當AI涉及涉及私人身份資訊的研究時，這些人甚麼時候是研究參與者？該研究是否抓住了研究參與者的定義中「關於誰」？對他們還有其他倫理考量嗎？
3. 根據Common Rule，AI或AI驗證活動的資料收集通常何時可以獲得豁免？



## Considerations for IRB Review of Research Involving AI

4. 對於需要根據Common Rule進行審查的研究，對於其資訊包含在用於AI開發和共享的資料集中的人類來說，哪些人類受試者保護考慮因素是最突出的？當研究重點是AI的測試或驗證時，這些考慮因素是否有所不同？其他倫理考量與非人類受試者相關嗎？
5. 資助機構、研究人員、HRPP工作人員和IRB是否可以使用現有的框架或工具來闡明和減輕針對人類的AI研究和開發的倫理關注？
6. 對於研究知情同意書，是否針對AI對研究活動揭露充分性的衝擊有特別的考量？
7. 涉及AI的研究有何「獨特」之處，需要IRB考慮並判定Common Rule的適用性尚未考慮到所有人類受試者研究？

## Considerations for IRB Review of Research Involving Artificial Intelligence

8. 在AI研究中需要特別注意 45 CFR 46.111 的哪些具體部分；即：隱私和保密、知情同意、風險？
9. 就機構/HRPP的責任而言，與其他人類研究的責任相比，針對AI要有哪些具體考量？
10. 在研究中使用AI是否存在更大的偏差和/或缺陷的可能性？IRB 在審查中應如何考慮這種可能性？（即：臉部辨識演算法可能很大程度上基於白人男性，但「使用該演算法」的研究人員可能沒有意識到這一點。）

## 1. 在什麼條件下，AI或AI驗證活動的資料收集符合Common Rule 「旨在開發或貢獻通用知識」的研究定義？（1/2）

如果資料收集是研究計畫的一部分，則這種收集很符合Common Rule的研究定義。但AI經常使用為其他目的收集的數據，例如醫療紀錄或社交媒體貼文。在目前的法規框架下，此類收集本身並不是研究，而此類數據的後續二次使用通常被視為屬於 45 CFR 46.104(d)(4) Common Rule中的豁免範圍。當使用從社交媒體帖子收集的材料進行研究時，這種豁免的使用尤其成問題，這些材料被認為是「公開可用的」，或者，如「去識別化」醫療紀錄的情況，它是「以無法輕易確定人類受試者身份的方式記錄。」



## 1. 在什麼條件下，AI或AI驗證活動的資料收集符合Common Rule 「旨在開發或貢獻通用知識」的研究定義？（2/2）

這種法規方式不一定是錯誤的，但它是在大數據和AI普及之前制定的；目前使用這些工具的研究正在利用並非為此目的而制定的研究豁免。2018年Common Rule明確承認了這一限制，並在第102(e)(7)(i)條中承諾定期重新審視可識別性的概念。因此，許多AI研究都是符合規定的，但卻不一定有充分保護研究參與者的權利和福祉。

2. 當AI涉及涉及私人身份資訊的研究時，這些人甚麼時候是研究參與者？該研究是否抓住了研究參與者的定義中「關於誰」？對他們還有其他倫理考量嗎？（1/2）

Common Rule第102(e)(1)條將人類受試者定義為「研究人員...對其進行研究的活體個體：(i) 透過介入或與個體互動來獲取資訊或生物樣本，並使用、研究或分析資訊或生物樣本，或(ii) 獲取、使用、研究、分析或產生可識別的私人資訊或可識別的生物樣本。」



## 2. 當AI涉及涉及私人身份資訊的研究時，這些人甚麼時候是研究參與者？該研究是否抓住了研究參與者的定義中「關於誰」？對他們還有其他倫理考量嗎？（2/2）

該問題明確假設AI研究涉及私人身份資訊。在這種情況下，此類個體應被視為人類受試者。如同對問題1的答覆所述，當所謂的私人資訊被視為公開時，法規語言中存在含糊之處。「私有」和「公共」的傳統定義不應被假定適用於互聯網、萬維網和（在可預見的未來）元宇宙創建的新環境，因為傳統上私有的信息必須以如下方式共享：參與成本，這正在成為一種預期的社會規範。



3. 根據Common Rule，AI或AI驗證活動的資料收集通常何時可以獲得豁免？

如問題1。



4. 對於需要根據Common Rule進行審查的研究，對於其資訊包含在用於AI開發和共享的資料集中的人類來說，哪些人類受試者保護考慮因素是最突出的？當研究重點是AI的測試或驗證時，這些考慮因素是否有所不同？其他倫理考量與非人類受試者相關嗎？

根據目前的法規解釋，只有一部分研究需要根據Common Rule進行審查。此類研究的特點是數據被視為「可識別的私人資訊」，但不受二次使用的豁免。換句話說，數據不會「公開」，研究人員可輕易知道受試者的身份，並且數據無法受到其他法規制度的保護，特別是HIPAA或聯邦隱私行為。符合這些標準的研究可能被認為風險微小，因為資料的收集和使用已經成為日常生活中普遍存在的現實，並且符合簡易審查標準。



4. 對於需要根據Common Rule進行審查的研究，對於其資訊包含在用於AI開發和共享的資料集中的人類來說，哪些人類受試者保護考慮因素是最突出的？當研究重點是AI的測試或驗證時，這些考慮因素是否有所不同？其他倫理考量與非人類受試者相關嗎？

鑑於研究的微小風險性質，保護措施可能是考慮到 111(a)(3) - 公平選擇受試者和 111(a)(7) - 隱私保護而產生的保護措施。AI研究即使滿足了將其置於法規主動監督之下的所有要求，也很可能有資格獲得免除知情同意，因為如果沒有這種豁免，研究就無法實際進行。請注意，111(a)(2) 的風險/效益計算所提供的保護將是有限的，因為該研究可能被視為風險微小，大多數 IRB 會不考慮應用所獲得知識的長期影響，排除因分析或強化既有偏差而造成群體傷害的保護措施。



4. 對於需要根據Common Rule進行審查的研究，對於其資訊包含在用於AI開發和共享的資料集中的人類來說，哪些人類受試者保護考慮因素是最突出的？當研究重點是AI的測試或驗證時，這些考慮因素是否有所不同？其他倫理考量與非人類受試者相關嗎？

儘管至少有一個領域為了AI開發而定期收集資料：mHealth（行動醫療），但對於那些資訊包含用於開發AI的人與用於驗證和測試AI的人的保護沒有明顯差異。如果資料收集本身是研究的一部分，就像在mHealth的開發中一樣，那麼研究參與者將有望獲得額外的保護，因為他們的參與須要知情同意。

最後，對於那些根據法規不被視為人類受試者的人，仍有倫理方面的考量。這些問題包括持續存在的群體傷害，以及可能移除解決疾病的公共資源及邊緣化。

## 5. 資助機構、研究人員、HRPP 工作人員和 IRB 是否可以使用現有的框架或工具來闡明和減輕針對人類的AI研究和開發的倫理關注？

以下文件提供一些有用的工具和觀點，用於理解AI的倫理問題：

✧ Keans, M., Roth, A. (2020) *The Ethical Algorithm: The Science of Socially Aware Algorithm Design*. Oxford University Press

這本書以非技術性討論AI/ML演算法的原則和價值。

✧ Hutson, M. (2022, Feb 26). The Future of Computing. *Science News*, 201(4), 16-22.  
<https://www.sciencenews.org/century/computer-ai-algorithm-moore-law-ethics>

一篇非專業文章，回顧電腦技術的歷史、AI/ML背後的歷史和基本概念，以及提出的一些倫理問題。

✧ Wolfram U. *Zero to AI in 60 inutes*. <https://www.wolfram.com/wolfram-u/machine-learning-zero-to-AI-60-minutes/>

這是一門簡短的線上課程，旨在說明任何人只要能夠使用適當的工具，無需深入了解軟體設計或編碼知識，即可輕鬆使用 AI/ML。

## 5. 資助機構、研究人員、HRPP 工作人員和 IRB 是否可以使用現有的框架或工具來闡明和減輕針對人類的AI研究和開發的倫理關注？

✧ Fjeld, J., Achten, N., Hilligoss, H., Nagy, A. & Srikumar, M. *Principled artificial intelligence: Mapping consensus in ethical and rights-based approaches to principles for AI*. Berkman Klein Center Research Publication (2020). [https://dash.harvard.edu/bitstream/handle/1/42160420/HLS%20White%20Paper%20Final\\_v3.pdf?sequence=1&isAllowed=y](https://dash.harvard.edu/bitstream/handle/1/42160420/HLS%20White%20Paper%20Final_v3.pdf?sequence=1&isAllowed=y) - PDF

一份白皮書，介紹AI治理文件的國際調查。

✧ Bernstein, M. S. et al. Ethics and society review: *Ethics reflection as a precondition to research funding*. Proc Natl Acad Sci U S A 118, e2117261118 (2021).

描述一所大學對AI/ML研究潛在危害的解決方法。

## 6. 對於研究知情同意書，是否針對AI對研究活動揭露充分性的衝擊有特別的考量？

需要知情同意的研究可能只佔所有AI研究的一小部分。對於需要知情同意的AI研究，當前人體研究同意要素的要求並不適合此類研究的風險和利益的性質。特別是116(b)(2) 要求披露「對受試者造成的任何可預見的風險或不適」及116(b)(3) 要求披露「研究對受試者或其他人可能合理預期的任何利益」。IRB的核准標準反映了這種對風險和利益的不對稱考量。傷害的風險可能涉及所有人，而潛在的利益預計亦會惠及所有人，但現行法規只允許考慮後者。目前的AI研究最重要的是傷害的是影響群體，而不是個人。受監管的研究是公共事務；風險和利益應在私人利益和公共利益之間得到平衡。目前的法規結構僅將部分任務交給IRB。



## 6. 對於研究知情同意書，是否針對AI對研究活動揭露充分性的衝擊有特別的考量？

更新後的Common Rule中新增的第 116(b)(9) 條的揭露要求也不適合AI研究，因為它反映了過於簡化的可識別性概念。刪除標識符不再意味著個人無法被識別，也不意味著私人和敏感資訊不會被披露並可能在未來與個人聯繫。應明確揭露該風險。

## 7. 涉及AI的研究有何「獨特」之處，需要 IRB 考慮並判定 Common Rule 的適用性尚未考慮到所有人類受試者研究？

法規中 IRB 確定 Common Rule 是否適用的「迴旋餘地」很小；更好的問題可能是，目前對人類受試者和人類受試者研究的定義是否允許 IRB 在包括 AI 在內的研究背景下充分保護個人和群體。



8. 在AI研究中需要特別注意 45 CFR 46.111 的哪些具體部分；即：隱私和保密、知情同意、風險？

請參閱上面問題 4 和 6 的答案。



## 9. 就機構/HRPP的責任而言，與其他人類研究的責任相比，針對AI要 有哪些具體考量？

AI引發了群體傷害問題，這主要與人們對資料集的限制知之甚少以及AI工具的使用可能掩蓋疾病、邊緣化和不平等的潛在和可解決原因的可能性有關。此外，BD 還提出了目前法規中未有很好解決的隱私和身分問題。就機構對其服務或所在社區負責而言，這些機構應考慮這些因素，並可能透過其HRPP予以考慮。然而，許多可預見的危害遠遠超出了任何單一機構的管轄範圍，最好在聯邦層級解決。此外，將這項責任留給各個機構可能會造成保護不一致的拼湊，這將不可避免地使保護較好的人受益，而犧牲保護較差的人的利益。



10. 在研究中使用AI是否存在更大的偏差和/或缺陷的可能性？IRB 在審查中應如何考慮這種可能性？（即：臉部辨識演算法可能很大程度上基於白人男性，但「使用該演算法」的研究人員可能沒有意識到這一點。）

AI的潛在危害源於資料集中未被認識到的限制或偏差，例如系統性種族主義和歧視以及數據不代表其結論所適用的人群的其他情況所引起的限制或偏差。此外，在大多數AI研究中，初始資料集的組裝與AI研究分開進行，這使得研究人員更有可能沒有意識到他們的結論的普遍性是有限的。

REVIEW ARTICLE

## Artificial intelligence and the Institutional Ethics Committee

A balanced insight into pros and cons, challenges, and future directions  
in ethical review of clinical research

### 摘要

將人工智慧（AI）與IRB研究倫理審查的結合帶來了重大的前景和問題。這項研究探討了AI在該領域的各種影響。AI的好處包括提高生產力、標準化評估、經濟高效的解決方案、徹底的分析 and 全球協作。這些改進可以透過加速和標準化審核來徹底改變審核流程。然而，其缺點包括人類洞察力的喪失、數據品質依賴、AI特定的倫理問題、與各種倫理框架的潛在不相容、實施成本和過度依賴。這個關鍵領域的AI整合的未來路線圖強調人與AI的協作、開放性、倫理設計、持續學習、可造訪性和利益相關者參與。此次探索最終呼籲跨學科合作和謹慎管理，以平衡技術創新和倫理廉正。這項綜合研究指導學者、倫理學家、政策制定者和實務工作者了解臨床研究倫理中AI的複雜世界。

# 報告大綱

---

- 前言
- AI的優點與角色
- AI的問題與因應
- AI的倫理議題
- AI研究的倫理審查
- 結語

# 結語

---

- ❑ 在不久的將來，AI在臨床研究的各個方面發揮愈來愈重要的作用
- ❑ 將受益於AI的領域包括：醫學影像、超音波心動圖、神經系統篩檢和手術
- ❑ 在將新技術引入醫學研究之前，新技術需要能夠與現有工作流程整合並獲得監管部門的批准
- ❑ AI的早期應用並沒有迎來機器統治人類的反烏托邦時代，而是補充和優化了人類研究人員和臨床醫生的重要工作