**計畫名稱：**發展中風患者動作、平衡及行走能力之人工智慧評估系統

**英文計畫名稱：**Development of an artificial intelligence system for assessing motor function, balance, and mobility in patients with stroke

**二、研究計畫中英文摘要**

**（一）計畫中文摘要**

**背景：**上肢動作、下肢動作、平衡以及行走等4種功能為中風患者評估及復健之重要目標。然而上述4種功能之評估遭遇2項困境：(1) 多向度逐一評量費時費力；(2) 紙筆評量結果之運用效能低落。上述困境有望藉由人工智慧 (artificial intelligence, AI) 之優勢克服，以動作錄影及AI自動化評分可補強或取代臨床人員之4種功能評估。

**目的：**計畫目的有三：一、發展中風患者動作（含上/下肢）、平衡及行走能力之人工智慧評估系統 (an artificial intelligence system for assessing motor functions of upper and lower extremities, balance, and mobility in patients with stroke, AIFS) 而且可於臨床或居家完成4種功能之評估；二、驗證AIFS心理計量特性；三、建立患者功能恢復之預測模型。

**方法**：本研究含二階段，階段一（完成目的一）以4步驟發展AIFS：(1) 設計AIFS之3種評估情境（標準情境錄影、臨床情境錄影以及居家指定情境）的評估內容；(2) 收集影片：拍攝100位中風患者之3種評估情境影片及50位正常人之標準情境錄影/評估；(3)AI分析影片：結合「電腦視覺」與「深度學習」分析3種情境之影片，以訓練與建構AIFS；(4) 建立計分與輸出系統。階段二（完成目的二與三）先驗證AIFS之心理計量特性，含再測信度、同時效度（驗證於100位慢性期中風患者）、反應性（驗證於100位亞急性期中風患者，並追蹤1年），之後再以追蹤資料訓練AI建立患者4種功能恢復之預測模型。

**預期結果：**AIFS之創新處在於可分析3種錄影情境並完成快速、精準評估：一、標準情境錄影/分析適用於專業人員快速評估所需，可於5～8分鐘完成4種功能之評估；二、臨床情境錄影/分析適用於一般臨床情境，需2～3次之一般臨床治療時段之錄影/分析；三、居家指定情境錄影，病患/家屬可於10～15分鐘完成指定情境錄影再上傳供分析。AIFS將具備4大特點：(1) AI技術自動評分：以AI技術分析患者之肢體動作影片評量4種功能，節省評估人力及時間；(2) 提供評估報告給予醫療人員、患者及照顧者；(3) 提供4種功能恢復之AI預測模型；(4) 心理計量特性良好。

**對社會、經濟、學術發展等面向的預期影響性：**AIFS可能為世界首創以AI技術評估中風患者多項功能之AI系統。AIFS藉由AI技術有效降低中風患者評估之人力及時間。此外，AIFS具數位化之評估結果，其記錄之影像有助於呈現患者動作功能之改變，且評估結果有利於醫療人員、患者及家屬間共享，並與電子病歷結合。AIFS發展成功之經驗有助於申請者後續發展其它AI輔助之中風復健系統（如：「復健目標與計畫擬定輔助系統」與「復健治療監測系統」），以更全面地應用AI技術以提升中風復健之效能。

**關鍵詞：**中風、人工智慧、電腦測驗、職能治療/物理治療

**（二）計畫英文摘要**

**Background:** Upper extremity (UE) motor function, lower extremity (LE) motor function, balance, and mobility are important targets of rehabilitation in patients with stroke. Therefore, assessments and enhancement of these 4 functions are important for clinicians in rehabilitation. However, two drawbacks of assessments of the 4 functions have seriously limited the efficiency of rehabilitation. First, time- and manpower-consuming: face-to-face and one-on-one assessments of all 4 functions takes a lot of time and manpower for clinicians. Second, non-digitized results of assessments: the results of commonly used tests are recorded on papers, so the data are inefficient to manage and be shared among clinicians, patients, and caregivers. Theses 2 drawbacks of assessments are likely to be overcome by artificial intelligence (AI). The computer vision technology of AI can be used to “monitor” patients’ performance, assess the 4 functions, and record the results in digitized form.

**Purposes:** This project contains three objectives. First, we will develop an AI System for Assessing Motor Functions of Upper and Lower Extremities, Balance, and Mobility in Patients with Stroke (AIFS). Second, we aim to validate the AIFS. Third, we will establish an AI-based prediction model of the 4 functions’ prognoses.

**Methods:** This project has 2 phases. In the first phase (from 1st to 15th month), we will develop the AIFS through 4 steps: (1) designing the items of AIFS assessment in standardized, regular clinical, and assigned situations, (2) filming videos of 50 healthy adults in standardized situation and 100 patients in 3 assessment situations of AIFS, (3) establishing an AI system to analyze the videos of the 3 assessment situations, (4) establishing an AI scoring and output systems. In the second phase (from 16th to 36th month), we will examine the psychometric properties of the AIFS. The psychometric properties include the test-retest reliability, concurrent validity (these two psychometric properties are validated on 100 patients with chronic stroke), and responsiveness (validating on 100 patients with subacute stroke followed up to 1 year). Additionally, we will establish the AI prediction models of prognoses using the follow-up data.

**Expected results:** The AIFS will have 4 merits: (1) the AI assessment system can reduce time and manpower of assessments of the 4 functions, (2) data/results sharing among clinicians, patients, and caregivers, (3) AI prediction models of prognoses, and (4) sound psychometric properties.

**Expected impacts on the society, economy, and academic developments:** The AIFS may be the first measure using AI technology to assess motor functions, balance, and mobility in patients with stroke. The AIFS has great potentials to improve administrative efficiency, reduce manpower, and enhance data sharing among clinicians, patients, and caregivers. In addition, the experience of AIFS development can be helpful for researchers, in the future, to develop more AI-based systems (e.g., “AI-based Goal Setting and Treatment Planning System” and “AI-based Treatment Monitoring System”) to improve the efficiency and effect of stroke rehabilitation.

**三、研究計畫內容**

**（一）研究計畫之背景**

1. **上下肢動作、平衡及行走缺損普遍且嚴重影響中風個案與家屬**

國人中風之發生率高，每年約有1.7萬人因中風而導致失能，中風已成為國人失能的主要因素。1 台灣中風登錄資料庫顯示：初次中風之患者在中風一個月後的失能比例超過6成，即使中風半年後的失能比例仍超過5成 (51.7%)。1 失能不僅造成患者之生活品質不佳，更造成照護者之人力及經濟負擔，且使用大量的健保資源（高居前十名）。2 因此如何大幅提升中風復健效能一直是國際上復健相關研究人員極力探索的議題。

上肢動作、下肢動作、平衡以及行走能力缺損為中風個案最常見的後遺症，3-6 這4種功能缺損皆嚴重影響日常生活功能以及生活品質。3-5 因此上述4種功能復健一直是中風復健之重點、追蹤中風病人預後之重要項目、也是復健成效之重要指標。7因此評估中風病患這4種功能，一直是臨床與研究之重要議題。

1. **現有動作、平衡與行走評估工具之不足**

富格麥爾動作量表 (Fugl-Meyer motor scale, FM)8 是研究人員最常使用的動作功能量表。9 然而FM高達50題（表一），施測時間常超過20分鐘，因此FM甚少於國內臨床例行使用。國內復健各專業最常使用布朗斯壯動作恢復時期 (Brunnstrom motor recovery stages, BRS) 描述個案之動作復原狀態。布朗斯壯動作恢復時期僅有3題，評估快速/可行。然而布朗斯壯動作恢復時期僅能提供粗略的等級資料，難以量化個案之動作功能。因此FM與BRS皆難符合臨床人員/個案之評估/介入要求。

伯格平衡量表 (Berg Balance Scale, BBS)10, 11 是中風臨床與研究中最常被使用的平衡評估工具。然而BBS題目數多達14題，且每題評分標準不一。另BBS非專為中風病人設計（原為老人設計），坐立平衡僅有一題，故不適用於較嚴重的個案。另外，最近研究顯示：由於BBS部分站立項目（如單腳站、一腳往前跨）未明確規範由患側腳或健側腳執行，進而造成施測結果不穩定。12 因此BBS難以快速、精準評估個案之平衡功能。

後續Benaim等人發展出一個專為中風個案設計之平衡量表：「中風病人姿勢控制量表 (Postural Assessment Scale for Stroke patients, PASS)」。13 PASS總共有12個項目，內容可分為二部分：一、姿勢的維持；二、姿勢的變換。評分級距為4點量尺（0-3分），總分共36分。由於PASS專為中風設計，且心理計量特性良好，14 故PASS比BBS更適於中風病人。然而使用PASS評估之效率較差，如同上述一般傳統評估工具，包含評估者訓練時間（熟悉題目、指導語及計分規則）、評估時間（由治療師與患者一對一評估12題）及評估結果計分/歸檔時間（計算分數、紙本測驗結果歸檔或輸入電腦電子化）。因此 PASS 難於一般臨床情境例行使用。

功能性步態評估量表 (Functional Gait Assessment, FGA) 常用於評估中風患者行走功能。15 FGA共10題，每題為4點量尺計分，總分為0~30分，分數越高表示患者的行走能力越佳。FGA應用於中風患者具良好之信度、效度及反應性。16 然而使用FGA評估之效率差，包含評估者訓練時間、評估時間以及評估結果計分/歸檔時間。此外，各題之評分標準複雜（同時考量走路時間及走直線偏移距離），計分不易。因此一般臨床單位甚少例行使用FGA。

此外，現有測驗之評估結果的應用極其有限。中風患者之評估為臨床單位之例行性業務，長期下來累積資料相當可觀，可形成大數據以貢獻於後續研究分析，17, 18 例如：利用評估資料以「長期追蹤患者之功能變化」、「探索/驗證影響患者預後之重要影響因子」及「改良既有之評估工具」等。然而，目前中風復健領域對於這些評估結果之應用極其有限，主要原因為：一般臨床機構考量專業人員之業務量，且電子病歷紀錄內容多半較簡略，僅以檢核表方式記錄患者的主要問題，未詳細記錄患者於標準化工具之評量表現，致使後續分析困難，且不利於影像化地呈現患者之功能改變。19 故國內現行醫療單位中風病患之電子病歷內容不利後續研究分析。

簡言之，中風病人之動作、平衡與行走為最主要之復健目標，然而臨床人員若使用標準化評估工具（如FM、PASS與FGA等）例行評估這些功能，將花費許多時間，且評估結果之應用性不佳。如何改良評估工具、提升評估與後續應用價值，一直是中風復健之重要研究議題。

表一、常用於中風患者之動作、平衡與行走評估工具之不足

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 測驗名稱 | 評估結果 未數位化 | 評估時間長 | 評分不易a | 評估完整性 不足 |
| 富格麥爾動作量表  (Fugl-Meyer motor scale) | **√** | **√** |  |  |
| 布朗斯壯動作恢復時期 (Brunnstrom motor recovery stages) | **√** |  |  | **√**b |
| 伯格氏平衡量表  (Berg Balance Scale) | **√** | **√** | **√** | **√**c |
| 中風病人姿勢控制量表 (Postural Assessment Scale for Stroke patients) | **√** | **√** |  |  |
| 功能性步態評估量表 (Functional Gait Assessment) | **√** | **√** | **√** |  |

a每題評分標準不一

b僅能提供質性描述資料，難以量化個案之動作功能

c站立平衡僅有一題，故對於較嚴重的個案並不適用。

1. **臨床評估之瓶頸與其負面影響**

除上述評估工具本身之限制外，極其有限的臨床評估時間與人力，亦阻礙臨床人員使用標準化評估工具。

時間方面，職能治療/物理治療人員無特定、專用的評估時間，因此職能治療/物理治療人員只能在30分鐘的治療時間內，自行利用治療的空檔觀察、簡單評估或口語詢問個案的狀況。在如此有限且零碎的評估時間下，臨床人員幾乎不可能使用標準化評估工具，評估中風個案的多種能力（如：上、下肢動作、平衡、行走）。人力方面，一位職能治療/物理治療人員在同一治療時段常常需服務多位患者，且無專責的評估人員，因此沒有足夠的人力進行一對一的標準化評估。上述瓶頸導致大多數臨床單位未使用標準化評估工具評估與追蹤個案能力，僅依靠復健人員的質性觀察及少量且粗略的評估項目，如各個向度一題，且每題只有3點計分（不佳-尚可-良好）。非標準化且粗略的評估導致4項缺點，嚴重影響復健的效能：(1) 難以對症下藥：缺乏系統性且標準化評估，容易遺漏須評估之向度（能力），或某一向度中的評估不夠完整，故而職能治療/物理治療人員難以全面且完整地掌握個案狀況，給予個案最適合之治療計畫；(2) 難以呈現能力變化：粗略的評分容易導致多數個案集中於中間的分數點且難以進步，例如：3點計分（不佳-尚可-良好）中，個案的得分大多是 “尚可”，且經過治療也難以進步到 “良好”。此限制造成職能治療/物理治療人員難以依據個案之進步及時調整治療方案，也難以呈現個案進步的數據給個案、其他醫療人員或審核單位（如健保局或醫院評鑑小組）；(3) 難以預測預後：職能治療/物理治療人員無法掌握有效、精確且豐富的資訊，故無法建立精準之預後預測系統。此限制不利於職能治療/物理治療人員與個案討論復健目標、出院安置計畫或照護計畫；(4) 個案對於評估結果無感：由於評估向度少，且評估結果粗略，所以個案無法從評估結果得知詳細的能力狀況，例如：各項能力之百分等級。此限制造成個案不願意接受評估，進而導致職能治療/物理治療專業人員難以掌握個案問題與進展。

簡言之，現今評估之瓶頸為臨床人員之時間與人力不足，故無法執行多向度之標準化測驗。這些瓶頸導致臨床人員不易掌握個案之問題、進展與預後，很可能已間接限制當今復健的成效。

1. **人工智慧之介紹 (artificial intelligence, AI)**

AI 之目的為使電腦或機器具備類似人類的思考與行為模式（例如：問題解決及物品辨識）。20 AI機器/電腦為達到此目的，其學習能力扮演重要的角色。透過「深度學習 (deep learning)」學習方法，AI機器/電腦可從外界的訊息中提取並儲存重要之特徵，作為AI機器/電腦後行為表現（如：資料判讀、預測）之基礎。21 深度學習之所以能「學習」的基礎為類神經網路 (artificial neural network)，22 類神經網路由大量的人工神經元 (neurons) 聯結構成網絡進行計算，以模擬生物之神經網路。每個人工神經元可視為一個函數 (functions)，具備一組可由模型自行調節的權重（被學習演算法調節的數值參數，類似傳統線性模型中的迴歸係數值），以及一個激勵函數 (active function) 以控制人工神經元之輸出結果為線性或非線性。實務上，一個神經元可接收前一個（或多個）神經元之輸出值，並將輸入值乘以權重、加總、經過激勵函數後，形成新的輸出值並傳給下一個（或多個）神經元。由於神經元可以不斷地串接成一個非常「深層」之網絡，故而此方法稱為「深度學習」。22

類神經網路之「學習歷程」如下：22 輸入資料（即預測因子，如：中風個案之基本資料）通過類神經網路後得到一個輸出值（即預測值，如：中風個案於3個月後之上肢功能預測值），接著可藉由比較預測值與正確答案（如：實際上，中風個案於3個月後之上肢功能）之差異，自行調整類神經網路中的所有參數，並重複「估計🡪調整參數🡪估計…」之循環，使得預測值與正確答案持續逼近。當類神經網路之預測值與正確答案已非常逼近，則停止類神經網路之「學習歷程」。此時，以擬人化言之，該類神經網路已「學會」從輸入資料中獲得關鍵的特性，並藉以準確地預測正確答案。

「電腦視覺 (computer vision)」及「AI分類/預測」二項AI技術已廣泛應用於醫療領域，並已大幅提升部分醫療診斷及評估效能。23 電腦視覺主要處理/學習影像資訊（包含：照片及影片），使得AI機器/電腦具備視覺能力，可對影像資訊進行辨識、追蹤、標記及擷取特定影像。於影像醫學領域，幾乎皆可借助電腦視覺技術，提升影像判讀之精準度及降低醫療人員之人力負擔，例如：辨識X光之股骨頸骨折（正確率94.4%甚至優於人力辨識90.5%）、24 皮膚傷口之相片診斷是否有皮膚癌（正確率72.1%甚至優於人力辨識66.0%）25等。

AI分類/預測主要處理/學習數據資料（包含：名義、類別、連續變項資料），26 使得AI機器/電腦具備資料探勘 (data mining)、預測、分群、分類等功能。22 由於AI分類/預測模型可從資料中提取特徵，並透過自行調整模型參數，以選擇最適合用以分類/預測該筆資料之線性或非線性模型，因而可精準地預測線性或非線性資料，例如：AI模型區分中風患者及暫時性腦缺血 (transient ischemic attack) 之區辨正確率高達78.6%、27 AI模型藉由臨床資料（如：認知功能、跌倒史、視力評估、肌力評估等）評估老人之平衡功能（Pearson’s r = 0.69~0.89 vs. 平衡評估效標測驗）、28 腦癌患者接受放射治療之存活率預測 (areas under the curve: 0.84–0.75)。29此外，AI分類/預測模型與傳統迴歸分析相比，具有更佳之分類/預測精準度。30, 31 例如：AI模型預測腦外傷患者6個月後是否具有精神疾患之正確率高達0.91優於傳統迴歸之0.76。30

電腦視覺及AI分類/預測二者合用則具有更廣泛的應用領域，如以電腦視覺分析技術從影像及影片中抽出/標記重要特徵，再由AI模型分類/預測病人之診斷、功能高低或預後等臨床重要資訊。32, 33 例如：復健領域藉由分析自閉症兒童之眼神以診斷自閉症兒童（診斷正確率0.86–0.91）、34 藉由中風患者之個人資料及電腦斷層掃描診斷急性期中風患者之正確率高達85.2%、35 以及分析fMRI影像診斷是否具有認知缺損之患者〔如：輕微認知缺損 (mild cognitive impairment) 診斷正確率0.88、36 思覺失調症 (schizophrenia) 診斷正確率0.86、37 注意力不足過動症 (attention deficit hyperactivity disorder) 診斷正確率0.69〕。38

AI技術應用於動作評估與治療方面，雖然仍在起步階段，但已有初步的成果。過去研究中，常使用加速規或Kinect蒐集中風個案之動作資訊，再藉由AI模型判斷或預測個案之動作、平衡、行走等功能表現。39 近年來隨著AI技術發展，已有AI套件（如OpenPose）可分析普通攝影機或深度攝影機的影片，標註人體關節位置及隨著時間變化之參數（如加速度、改變量與運動軌跡等），再加上原已發展較成熟之AI模型判斷個案之動作、平衡、行走等功能表現，具有潛力應用於臨床及研究評估及介入個案功能。40 相關應用實例包含：錄影評估老年人行走時之髖、膝關節運動角度；41 錄影評估中風患者接受上肢動作復健時之身體代償動作；42 以及動作教學影片與摹仿動作之間的一致性，43 可應用於職能治療/物理治療確認個案從事活動的正確性。

AI相關理論與技術之精進，持續令人耳目一新。AI於醫療領域之應用亦必隨之暨廣又深。因此如何善用AI最新理論與技術應是大幅提升醫療/復健效能的關鍵利器。

1. **AI於中風患者功能評估之價值**

AI評估之潛在價值有四：(1) 提升評估之效率：AI技術可補強或取代臨床人員之肉眼，直接分析中風患者於影像中的上/下肢動作、平衡及行走等外顯行為。44, 45 如此將可減少評估時之人力（如：目前臨床上，一位治療師經常同時治療多位患者）及時間（如：減少施測者訓練、計分及資料整理時間），若患者無跌倒風險，甚至可以由患者自行施測（或在照護者些微保護下施測），大幅提升中風評估之效率。

價值(2) 拓展評估之適用情境：AI應用於中風患者之評估情境可分為三類（標準情境、臨床情境以及居家指定情境，表二）：(a) 標準情境：46 於研究室或醫療院所，研究及臨床人員設計一套固定之題目（即欲評估之動作）及評估流程，由研究人員或臨床人員施測及協助下，個案依據AI評估系統提供之示範影片/說明執行AI評估。標準情境評估之優點為可快速蒐集多種且不同難易度之動作資訊，因測驗題目皆為研究及臨床人員所精心設計、可包含多種功能、且具備重要之臨床意義。然而標準情境評估之缺點為須額外的時間及人力「刻意地」進行評估，且對於患者之可近性 (accesibility) 較低，因需與研究/臨床人員約定時間、地點才能評估。因此，標準情境適用於初次發展AI之評估，以提高評估效能；亦可用於患者初次評估以補強或取代現有常用工具評估。

(b) 居家指定情境：47 此情境之評估題目同標準情境，但由患者/家屬/照護者自行錄影，再將錄影檔上傳雲端資料庫，再由AI完成評估。為確保患者/家屬/照護者可順利完成測驗，AI系統自動提供示範影片/說明，而研究/臨床人員可提供額外協助，如：操作說明及接受諮詢，其執行環境為患者之生活情境。居家指定情境之優缺點同標準情境，但居家指定情境評估之患者可近性高。因此，適用於出院後之定期追蹤、居家復健監測及線上篩檢。

(c) 臨床情境：48 於常規職能治療/物理治療的情境下，藉由固定之錄影設備拍攝醫療人員以「一般臨床評估工具」評估個案功能（如：初評、追蹤評估、結案評估）及從事治療活動（如：上下肢動作訓練、平衡訓練及行走訓練）之歷程，AI系統藉由分析錄影資料以完成評估。臨床情境評估之優點為可「自然地」評估個案功能，無須臨床人員及個案額外之評估時間。然而，由於無特定之測驗題目及拍攝掌鏡人員，因此可評估之功能向度及內容較侷限，且影片之拍攝品質不穩定。臨床情境評估可作為主要之動作評估方法，因評估負擔最小，但可能需要多次（如2～3）次臨床情境評估，方能累積可靠、豐富之評估資訊。另臨床情境之AI評估，若評估內容或精準度不足，可由AI提醒醫療人員補充不足，如AI提示部分標準情境之題目，藉以補足評估內容與提升精準度。

現有標準化工具之適用情境有限，僅能於標準情境下由研究人員評估或臨床情境下由臨床人員評估。相對地，AI以錄影方式評估可應用於更廣泛之評估情境，故能克服傳統評估受到環境（如：醫療院所）、人力（必需由臨床人員施測）、時間（僅臨床上初評、追蹤評估、結案評估）等限制，以蒐集患者更完整（不同情境下且時間上更連續）的資訊。

表二、AI系統評估之3種適用情境

|  |  |  |  |
| --- | --- | --- | --- |
|  | 標準情境 | 臨床情境 | 居家指定情境 |
| 評估內容 | 研究/臨床人員設計及  指定之動作 | 患者於常規臨床之表現，包含評估及治療 | 研究/臨床人員設計及 指定之動作 |
| 評估者/  影片拍攝者 | 研究/臨床人員 | 無：固定/自動錄影e | 患者/家屬/照護者 |
| 拍攝地點 | 研究室/醫療院所 | 研究室/醫療院所 | 患者生活環境 |
| 拍攝方式 | 系統呈現題目（指定之動作）、指導語及示範影片，請患者仿做並錄影，評估者必要時給予協助及糾正。 | 患者於醫療院所接受常規職能治療/物理治療評估或活動時拍攝影片。 | 同「標準情境」，但研究者額外提供評估相關補充資料（如系統操作說明），以利個案/家屬/照護者執行。 |
| 適用時機 | 1. 初評 2. 補充評估 | 1. 每次治療即自動評估 2. 長期追蹤 | 1. 長期追蹤 2. 篩檢評估 3. 居家復健監測 |
| 影片內容之訊息量a | 大 | 小 | 大 |
| 拍攝品質b | 高 | 不穩定 | 高 |
| 額外評估 時段/人力c | 需要 | 不用e | 需要 |
| 個案取得評估之可近性d | 低 | 中 | 高 |

a「標準情境」及「居家指定情境」均請患者從事指定動作，因此短時間的影片中即可獲得多種且不同難易度的動作影像資料；反之，「臨床情境」未指定特定之影片拍攝內容，故一段長時間（如：10分鐘）的影片中可能只包含少數幾種動作或動作難度較低，不易精準地評量患者之功能高低。

b「標準情境」及「居家指定情境」均有專職拍攝者掌鏡（前者為研究/臨床人員，後者為患者/家屬/照護者），故影片之拍攝品質可隨時監控及適時調整；而「臨床情境」運用固定之錄影器材，無專職拍攝者掌鏡監控，故無法確保拍攝對象於畫面中的大小及位置。

d「居家指定情境」由患者/家屬/照護者自行拍攝，故可於方便的時間及地點評估，不用到特定的醫療院所接受評估。「臨床情境」需特定之評估地點（即常規接受職能治療/物理治療之醫療院所），但無須額外且特定之評估時段（可於患者接受職能治療/物理治療評估及治療時自然地完成），且評估結果可透過AIFS分享系統輕易取得，故可近性中等。「標準情境」相較其它二情境需特定之評估地點及時間，故可近性較低。

e若評估內容或精準度不足，可由AI提醒醫療人員以部分「標準情境之題目」補足評估內容與精準度。

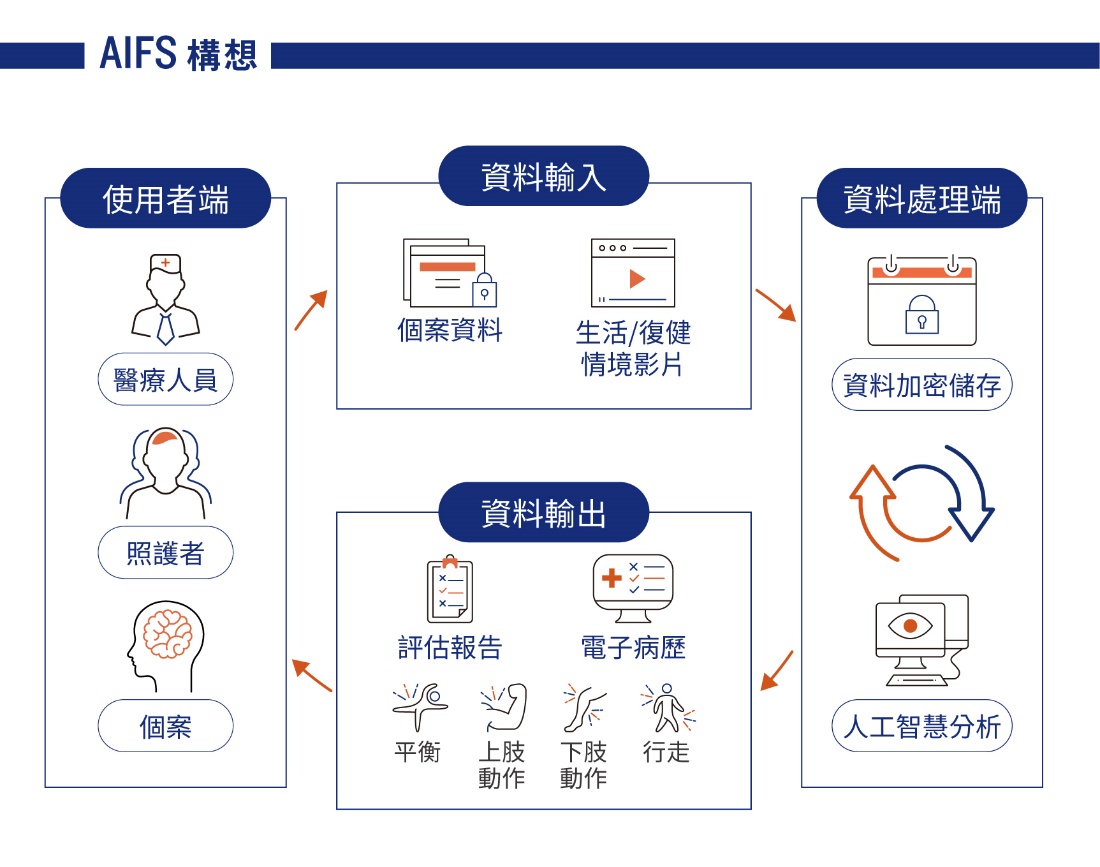
價值(3) 長期成效治療追蹤/預測： 以錄影方式評估可自然地於臨床情境中（臨床人員及患者無須額外之時間或心力投入於評估），大量且長期地蒐集患者於之影片累積形成大數據，用途廣大。如AI以視覺化地呈現患者功能之改變，且可將評估數據應用於後續研究分析。例如，建立患者各功能的變化曲線/預測以及追蹤治療成效等。49

價值(4) 增進評估之精準度：AI評估系統具備一致、客觀之評分標準，因此相較於現有評估工具之專業人員評估，將具備更佳之再測信度（或精準度）。48 此外，AI評估系統可掌握更多資訊（如：全身之關節角度、運動速度等）進行評分，相較於現有測驗僅專注於特定資訊（如：單獨上/下肢動作等），AI評估將具備更佳之反應性可更敏感地偵測患者功能之變化。

以上文獻回顧充分顯示：AI評估極可能克服傳統評估工具之瓶頸（如：耗費人力及時間、評估結果未數位化、後續分析之效率差），因此如何應用AI於中風個案之臨床評估，將是未來提升臨床效能/復健品質之關鍵。

1. **研究目的**

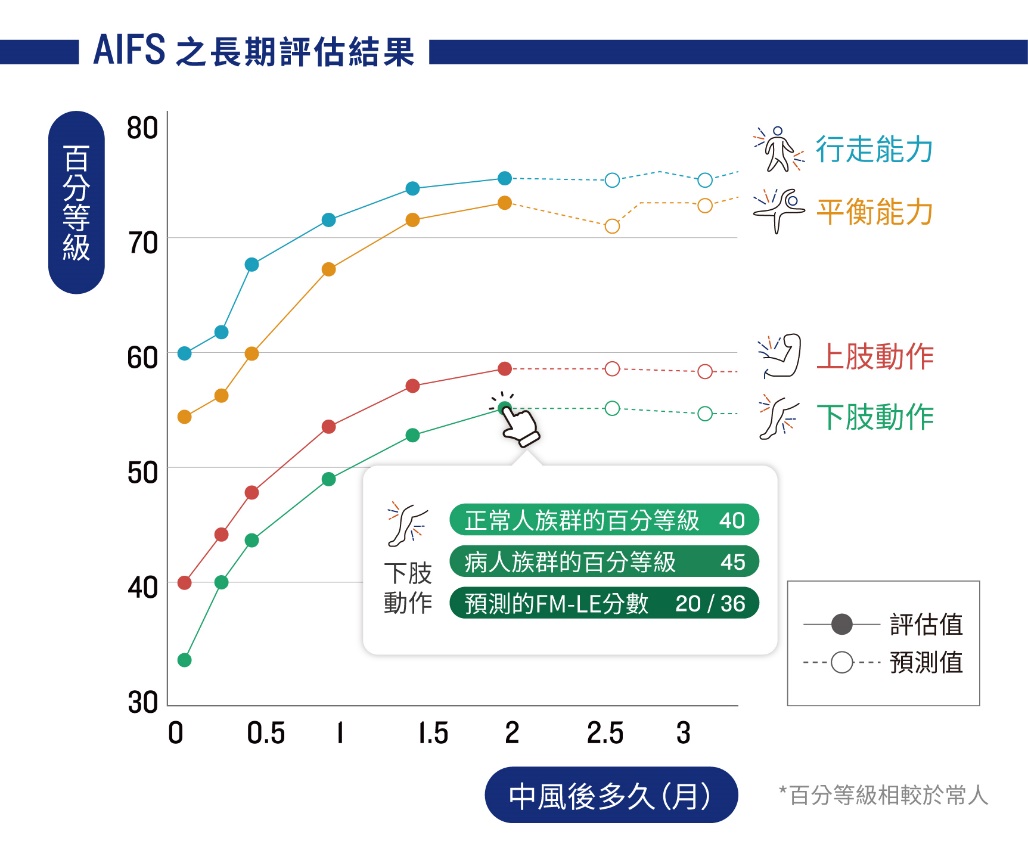
本計畫目的有三：一、應用AI系統發展中風患者動作、平衡及行走能力之人工智慧評估系統 (an **a**rtificial **i**ntelligence system for assessing motor **f**unctions of upper and lower extremities, balance, and mobility in patients with **s**troke, AIFS)，以應用於中風患者之評估、治療及追蹤成效（圖一～圖三）。考量臨床實務需求及目前AI技術發展，本計畫擬藉由中風患者於標準情境、臨床情境及居家指定情境之影像，評估患者之4種重要功能，包括上肢動作、下肢動作、平衡以及行走。二、驗證AIFS心理計量特性，包含再測信度 (test-retest reliability)、同時效度 (concurrent validity) 及反應性 (responsiveness)。三、建立AI患者功能恢復之預測模型。



圖一：AIFS構想



圖二：AIFS之AI分析與輸出架構



圖三：AIFS多次評估結果與預後預測

1. **預期結果與價值**

研究者預期 AIFS之評估及數位資料整理/輸出效率高，臨床及研究人員可精確、有效、且快速地評量中風患者之4種功能。具體之預期結果與價值如下：

預期結果

* 1. 自動評估與記錄中風患者之4種功能：AIFS將藉由分析患者影片，自動評估中風患者之四種功能（圖一～圖三）：上肢動作、下肢動作、平衡以及行走。AIFS自動且多向度之評估，將可大幅提升中風患者評估之效能。
  2. AIFS分數具體且具有臨床意義：AIFS輸出分數有3種：預測的效標分數、病人族群的百分等級以及與正常人族群的百分等級。「預測的效標分數」為AIFS依據個案之影片預測個案可能之黃金效標得分，例如：FM、PASS與FGA分數，使得研究及臨床人員得以將影片評估結果與黃金效標評估結果做比較。「病人族群的百分等級」為AIFS依據所有個案於同樣動作之影片，建立所有患者從事同一動作之標準動作（即患者平均而言的運動軌跡），接著比較患者與標準動作之動作差異，以取得患者動作於病人族群的百分等級，百分等級越高表示患者於中風族群中相對的動作能力越好。「正常人族群的百分等級」與「病人族群的百分等級」之概念與做法相同，但比較對象改為正常人族群。
  3. 全面評估患者於3種情境之功能：AIFS之適用情境包含標準情境、臨床情境、及居家指定情境。標準情境評估請患者從事AIFS指定之動作並錄影評估（預計題目8~12題，約5～8分鐘以內完成），因標準情境評估可於短時間內評估大量且多樣動作，適用於專業人員快速評估所需，如：初評及定期評估補充AIFS於臨床情境未評估到之項目/動作。臨床情境評估將為AIFS之主要評估情境，因分析之影片來自於患者真實情境，且無須額外增加患者及臨床人員之體力/時間負擔，最能發揮AIFS之價值（提升評估之效能）。居家指定情境之評估方式同標準情境，但因評估由患者/家屬自行完成，適用於患者出院後之長期追蹤、篩檢及居家復健監測。
  4. 3種錄影情境可擇一或依方便性使用：因AIFS評估結果相關性可能極高，可相互交替使用，故未來使用者可擇一或依方便性使用之。
  5. 自動產生評估報告（圖一～圖三）：AIFS自動將單次其長期追蹤評估之結果輸出並傳送給醫療人員，經醫療人員確認後，可傳送或列印給患者及照顧者，以利醫病溝通討論患者之功能狀態及改變。
  6. 患者功能恢復之AI預測模型（圖三）：追蹤且比較中風患者於不同時間或情境之表現，並以AI建立中風患者4種功能恢復之預測模型。
  7. 良好之心理計量特性：AIFS應用於中風患者具良好之再測信度、同時效度及反應性。
  8. 量化治療活動（圖四）：量化治療活動有助於探討職能治療/物理治療介入之種類、內容、頻率與持續時間等重要的治療參數，以發展個案適性、適量之介入方案，進一步提升治療成效。
  9. 加密系統保存影音資料：為保障受測者個資，AIFS之影音資料將保存於加密之雲端空間，除主持人及研究人員外，其他人員無法取得影音資料。

研究之價值與效益

AIFS達成上述預期結果後，可產生之臨床、研究效益多達6項，發展AIFS之成本效益高。6項效益詳述如下：

1. 減少評估時間及人力：AIFS具備多評估情境之特性，可在一般治療情境或居家情境即完成評估與長期追蹤，無須花費額外且特定之評估時間及人力。此外，AIFS可同時評估多人，以大幅減少現今臨床上一對一評估所花費的大量時間與人力。
2. 有助於累積大量、長期且數位化之數位化復健資料：礙於臨床上評估時間與人力不足，現今復健資料之評估內容粗略，評估內容往往只有簡短的質性描述或3~5點計分（如：平衡能力良好、尚可或不佳）的評估結果。這些評估結果難以用來客觀呈現療效，或作為臨床實證資料。而AIFS之評估結果包含影音資料與量化後的能力分數，且所有結果均已數位化可直接存入雲端資料庫（或電子病歷系統），易於未來資料管理與提取。AIFS累積長期且大量的數位化資料將成為重要的基石，以提升職能治療/物理治療成效或探討復健機轉。
3. 增加職能治療/物理治療評估的可近性，可應用於遠距醫療：中風個案或家屬可自行上傳動作/活動影片至AIFS線上系統，以獲得4向度之評估結果。此一特性有助於個案在接受職能治療/物理治療前即可讓職能治療/物理治療人員瞭解其初步的能力程度、個案接受職能治療/物理治療服務後自行掌握其能力進步情形、以及個案接受遠距醫療時的事前或即時評估。
4. 作為醫病溝通的重要參考數值：數位化評估結果有助於個案/家屬瞭解其能力程度與變化，且AIFS可自動輸出評估結果並傳送給個案/家屬。因此，AIFS之評估結果可做為醫療人員與個案/家屬討論治療目標、治療方案或長期安置計畫的重要依據，以克服個案/家屬因對於個案能力與變化不瞭解，而對於治療目標、治療方案或長期安置計畫的期待過高或過低，難與職能治療/物理治療人員達成共識。
5. 技術可類化於發展AI復健介入輔助系統：AIFS可評估中風個案之四種功能後，亦可用於監控個案從事職能治療/物理治療活動的質與量。質的方面可監控個案從事職能治療/物理治療活動的正確與否，是否與職能治療/物理治療人員建議的活動相符。量的方面可記錄個案從事復健活動的頻率與時間。上述質與量的紀錄有助於職能治療/物理治療人員掌握個案之職能治療/物理治療介入情形，適時調整治療活動或給予個案從事職能治療/物理治療活動的建議，以提升復健成效。
6. 未來可擴充評估向度，作為AI全面職能治療/物理治療評估的基礎：由於本計畫礙於時間、人力與經費有限，考量可行性與重要性後決定優先發展AI系統評估中風患者之上下肢動作、平衡與行走功能。本研究之執行經驗、技術與成果有助於未來擴充評估向度，如：認知、情緒、社交互動或語言能力等重要功能，以全面提升職能治療/物理治療評估的效能。



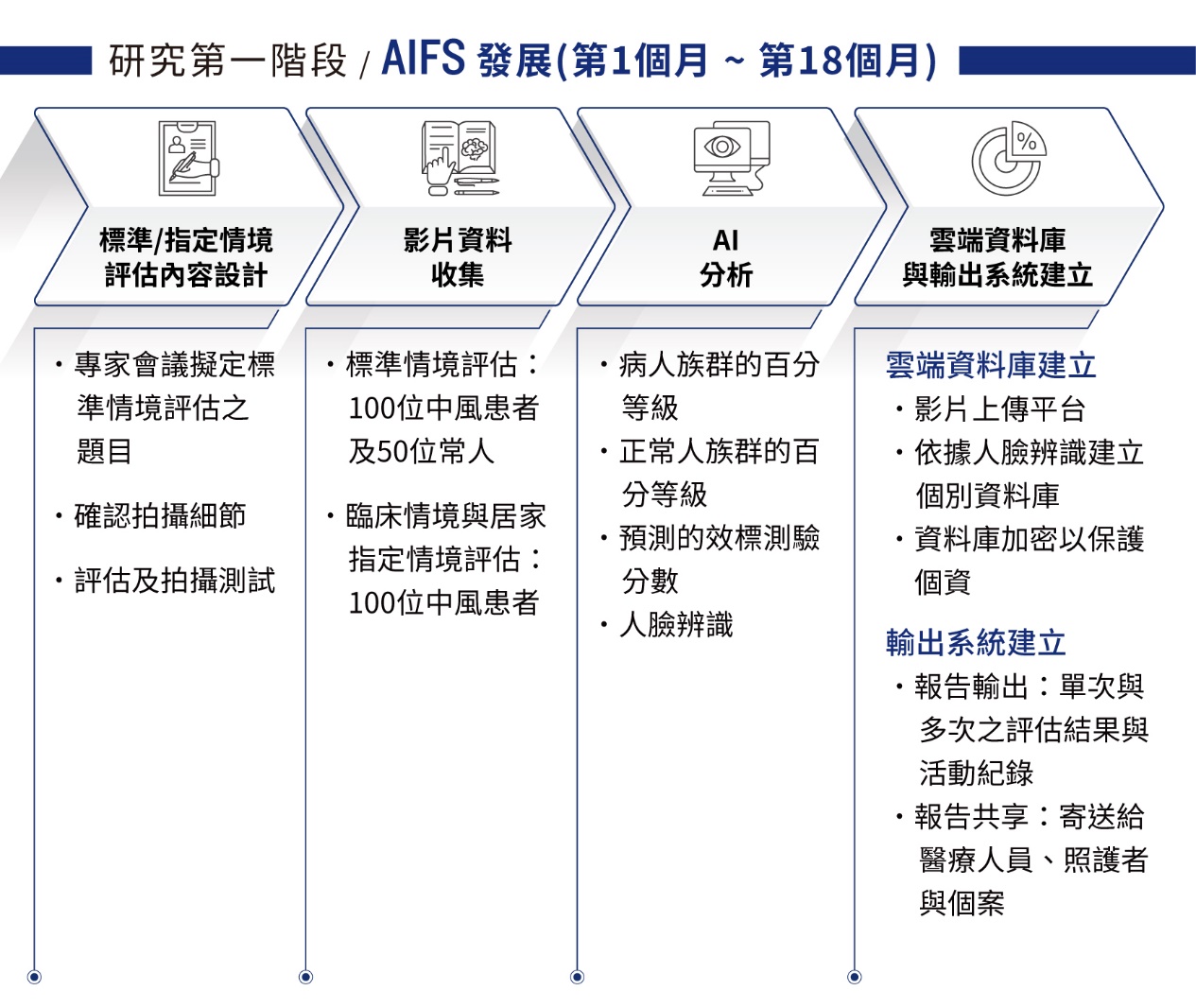
圖四：AIFS之治療活動量化紀錄

**（二）研究方法、進行步驟及執行進度**

本計畫預計以3年二階段完成AIFS之發展，含第一階段：AIFS發展（圖五）；第二階段：AIFS心理計量特性驗證及AI預測模型建立（圖六）。

**第一階段：AIFS發展**

AIFS發展包含4個步驟：(1) AIFS標準情境及居家指定情境評估內容設計、(2) 影片收集（影片包含3種評估情境：標準情境、臨床情境及居家指定情境）、(3) AI分析、與(4) 計分與輸出系統建立。具體之發展步驟詳述如下：



圖五：AIFS發展步驟

1. AIFS標準情境評估內容設計

此步驟中研究者將設計上肢動作、下肢動作、平衡以及行走等4種功能於標準情境及居家指定情境評估之題目，題目設計原則為：可同時呈現患者多種能力之動作（如：「從椅子上站起，並往前行走6步」，即可同時呈現患者之下肢動作、平衡以及行走能力）。此外，研究者亦設計AIFS之評估標準、輸出指標、及錄影拍攝細節（包括攝影機使用台數、拍攝之角度以及拍攝距離等）。研究者將先依據參考文獻及現有評估工具初步設計後，將召開2次專家會議，各邀請10位專家審核及修改之，專家將包括中風復健之臨床及研究人員，以及AI程式發展者。研究者預期先設計AIFS之標準情境題目約20題，而研究最終保留8~12題，以評量患者之4種功能。

1. 影片收集

研究者將分別招募50位正常人及100位中風患者接受AIFS之標準情境題目評估及錄影。此外，中風患者需額外拍攝其於其它2種情境（臨床情境及居家指定情境）之動作影片。本研究招募正常人從事標準情境題目之原因/目的有三： (1) 正常人可模擬多種程度之動作能力（如：不佳、中等、好），確保收集到廣泛功能程度之動作影片；(2) 資料收集初期可能須微調題目內容，正常人較可配合彈性調整；(3) 正常人體力較佳，可在較短時間內完成所有題目，增加資料蒐集的效率。

*樣本*

正常人之收案條件有3項：(a) 年齡 ≥ 20歲；(b) 無影響動作之疾患或外傷（如：中風、小兒麻痺或肢體骨折）；(c) 有意願參與研究（含錄影）。

中風患者之收案條件有3項：(a) 經醫師診斷為中風；(b) 年齡 ≥ 20歲；(c) 有意願參與研究。主持人將排除不願意接受錄影及無法理解指令之中風患者。

*評估流程*

正常人及中風患者均接受一回合之評估，期間須完成AIFS之標準情境題目（約20題）同時接受錄影，以及三項效標測驗 (FM, PASS, and FGA) 評估，作為AIFS分數輸出之學習目標。此外，中風患者接受評估後之一週內，需額外拍攝其於居家指定情境（由個案、家屬或照護者拍攝，並上傳至研究者加密之線上資料庫）及臨床情境（由研究人員拍攝）之動作影片。標準情境及居家指定情境之影片拍攝患者執行所有指定情境題目，臨床情境則為約30分鐘影片，拍攝患者接受職能治療/物理治療評估及活動情形。最後，由獨立之研究人員觀看中風患者於3種情境之影片進行評（依據上一步驟訂定之評分標準），包含：四種功能分數、肢體協調程度、治療師肢體協助之次數等，作為後續驗證AIFS評分與臨床人員評分一致性（步驟4）之驗證效標。

*效標測驗工具*

上/下肢動作：FM-motor8 主要評估中風患者之動作控制功能。包含上肢與下肢動作次量表，上肢次量表具有33題，下肢次量表具有17題，各題為3點量尺計分（0表示無動作，2分表示正常動作）。上肢次量表之分數介於0~66分，下肢介於0~34分，分數越高表示患者的動作功能越佳。FM-motor應用於中風患者具良好之信度、效度及反應性。50

平衡：PASS13 主要評估中風患者之平衡功能。PASS共12題，各題為4點量尺計分（0分表示無法從事該題，3分表示可以獨立從事），總分為0~36分，分數越高表示患者的平衡功能越佳。PASS應用於中風患者具良好之信度、效度及反應性。14

行走：FGA15 常用於評估中風患者行走功能。FGA共10題，每題為4點量尺計分（0分表示無法從事該題或表現極差，3分表示表現接近正常人），總分為0~30分，分數越高表示患者的行走能力越佳。FGA應用於中風患者具良好之信度、效度及反應性。16

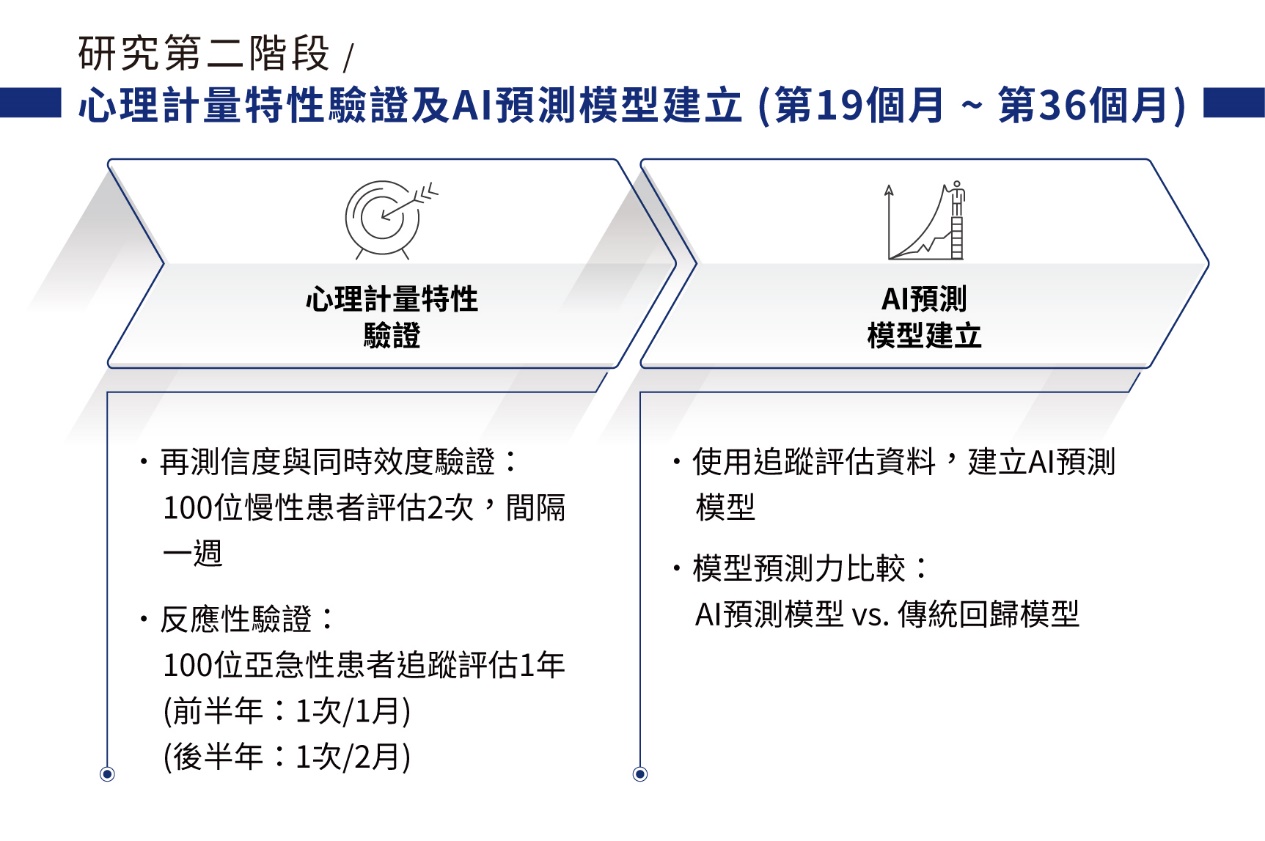
1. AI分析（結合電腦視覺與深度學習）

研究者使用電腦視覺與深度學習分析正常人於標準情境及中風患者於3種情境影片（圖二），AI分析可分成兩階段，第一階段研究者先以電腦視覺分析擷取影片中重要資訊；第二階段，研究者設計及訓練深度學習模型以建立計分系統，使數據更容易被職能治療/物理治療專業人員使用及解釋，包含：四種功能分數、肢體協調程度、治療師肢體協助之次數等。具體而言，第一階段中研究者應用電腦視覺之人體姿勢分析套件（如：OpenPose51 及YOLO252）擷取正常人及患者之肢體及環境資訊，肢體資訊包含：身體重要關節座標（如：各手指、肘、肩、髖以及膝關節等）、關節角度、肢體運動速度以及身體重心位置等，環境資訊包含：偵測是否使用輔具及輔具種類、是否需要他人協助。由於第一階段擷取之資訊均為醫療人員評估中風患者四種功能時主要參考/觀察之指標，因此有利於AIFS於第二階段藉由深度學習模型以建立計分系統。研究者將使用監督式學習模型，並以效標測驗（上/下肢FM、PASS、FGA分數）以及專業人員判斷分數作為學習目標，以輸出4種功能分數、肢體協調程度、治療師肢體協助之次數等分數。若AIFS之分數與效標測驗（上/下肢FM、PASS、FGA分數）評分之結果一致性高 (Cohen’s kappa ≥ 0.80)，53 則完成AIFS之計分系統。最後，研究者建立互聯網系統以匯出/分享AIFS評估結果。

此外，為提升AIFS評估資料管理之效能，研究者亦使用電腦視覺技術辨識人臉ID以自動存檔及管理受試者之評估資料，尤其當臨床情境評估時，畫面中可能同時出現多位患者及不相關之人員，此時人臉ID辨識及資料歸檔功能將顯得相當重要。人臉ID辨識之電腦視覺技術已相當成熟，如：FaceNet（Google所發展之深度學習網絡）之人臉辨識正確率高達0.99。54

**第二階段：AIFS心理計量特性驗證及功能恢復之AI預測模型建立**

本計畫於此階段將驗證AIFS之心理計量特性（再測信度、同時效度及反應性），並建立患者功能恢復之AI預測模型（圖六）。



圖六：AIFS心理計量特性驗證及AI預測模型建立

1. 再測信度及同時效度驗證

*樣本*

本階段蒐集100位慢性中風者之動作影像資料，驗證AIFS之再測信度及同時效度。慢性期中風患者之收案條件有3項：(a) 經醫師診斷為中風且中風後 ≥ 6個月；(b) 年齡 ≥ 20歲；(c) 有意願參與研究。研究者排除不願意接受錄影及認知功能太差無法理解指令之患者。

*評估流程*

於研究人員一對一施測下，中風患者接受二回合之評估，二回合皆須完成AIFS標準情境評估及三項效標測驗 (FM, PASS, and FGA) 評估，並接受拍攝影片。二回合評估間隔1週以確保患者於前後測期間肢體動作功能無明顯差異。此外，患者於每回合評估結束一週內，需拍攝居家指定情境（由患者、照護者或家屬依據研究人員提供之影片拍攝原則錄影，並上傳至本研究專用之加密雲端資料庫）及臨床情境之動作影片（由研究人員與患者約定時間、地點至醫療院所拍攝）。

*資料分析*

AIFS、FM、PASS及FGA於慢性中風患者2次評估間之再測信度驗證包含再測一致性、練習效應及隨機測量誤差。再測一致性以組內相關係數 (intraclass correlation coefficient, ICC) 驗證。ICC ≥ 0.75為良好；0.4~0.74為可接受；< 0.40為差。55

練習效應以2種統計方法驗證：(1) 以效應值分析二次評估間之平均分數差異大小。效應值之計算方式為：平均變化值/平均變化值之標準差。若效應值 ≥ 0.80表示練習效應大；0.50-0.79為中度；0.20-0.49為小；< 0.20為可忽略。56 (2) 以相依樣本t檢定驗證二次評估間之平均分數是否具顯著差異，若檢定結果顯示 p < 0.05，表示重複測量分數具統計顯著之練習效應。

隨機測量誤差以測量標準誤 (standard error of measurement, SEM) 驗證。SEM之估計公式為：SEM = SD × √(1-ICC)。SEM代表個別受測者每次評估分數之不穩定程度（即隨機測量誤差）。57 SEM越低，代表隨機測量誤差越小。另外，研究者亦計算SEM%（SEM除以所有評估分數之平均值），58 以比較不同評估工具之隨機測量誤差大小。SEM% < 5%為隨機誤差小，< 10%為可接受。58, 59

同時效度驗證方法為：研究者以Pearson’s r相關係數分別檢驗AIFS於3種評估情境之4種功能分數（上肢動作、下肢動作、平衡以及行走）分別與其對應之效標測驗（FM上肢動作、FM下肢動作、PASS以及FGA）的相關程度（如：AIFS標準情境之上肢動作 vs. FM上肢動作、AIFS臨床情境之上肢動作 vs. FM上肢動作…等）。若Pearson’s r ≥ 0.75為良好；0.40-0.74為中等；< 0.40為差。60

此外，研究者亦檢驗AIFS同一種功能於3種評估情境間的分數相關程度（如：AIFS標準情境、臨床情境以及居家指定情境評估之上肢動作分數間的相關），研究者預期AIFS同一功能於3種情境之評估分數極高（如：Pearson’s r ≥ 0.80），因AIFS之3種評估情境分數皆來自於類神經網路估計，且訓練類神經網路時之分數估計目標（即輸出值之逼近對象）均相同（效標測驗分數，如：FM上肢動作、FM下肢動作、PASS以及FGA）。

1. 反應性驗證並建立AI為基礎之預測模型

*樣本*

招募100位亞急性期中風患者，收案條件有3項：(a) 診斷為中風且中風後 < 1個月；(b) 年齡 ≥ 20歲；(c) 有意願參與研究。研究者排除不願意接受錄影及認知功能太差無法理解指令之亞急性及慢性期患者。

*評估流程*

研究人員追蹤亞急性中風患者1年，前半年每個月評估一次（因患者之功能變化可能較多，因此評估較密集），半年後每2個月評估一次，共10次評估（0、1~6、8、10、12）。AIFS之3種評估情境各有不同之資料收集方式，分別描述如下：

「標準情境評估」將由研究人員與個案約定時間及地點（可為研究室、醫療院所、或居家），由研究人員前往與患者一對一完成。

「臨床情境評估」由臨床人員於治療室架設固定之錄影機，拍攝患者於自然臨床情境從事職能治療/物理治療活動之影片，每次約拍攝30分鐘影片。然而若患者已無須再接受復健治療或轉院至本研究未申請人體試驗之醫療院所，則不繼續收集患者之AIFS臨床情境評，然而該患者仍須繼續完成AIFS之其它2種情境評估（居家指定情境及標準情境）。

「居家指定情境評估」則由研究者以電話與患者聯繫，提醒拍攝居家之動作影片（居家指定情境拍攝患者從事所有指定題目影片），並於加密之影片上傳系統確認患者之拍攝內容及品質，必要時請患者重新拍攝。

*資料分析*

AIFS之反應性以三種統計方法驗證：(1) 以「效應值」分析中風患者相鄰二次評估間的分數差異大小。效應值 ≥ 0.80表示反應性大；0.50-0.79為中度；0.20-0.49為小；< 0.20為差；61 (2) 以相依樣本t檢定驗證相鄰二次評估間的分數是否具顯著差異，若檢定結果顯著，表示AIFS可偵測中風患者功能之變化。(3) 使用bootstrap統計分析方法以檢定AIFS之反應性是否顯著優於效標測驗之反應性。62

AI技術建立4種功能恢復之預測模型：使用中風患者於各時間點之四種功能、人口學資料（性別、年齡、疾病史等）及中風病情資料（受損腦區、中風類型、藥物等），預測患者追蹤一年後之四種功能。研究者同時使用傳統之多元迴歸及類神經網路 (artificial neural network) 建立模型，並比較二種建模方式之預測力。最終，研究者選擇預測力較佳之模型，作為AIFS之預測模型。63, 64 類神經網路模型預計使用深度學習專門之軟體TensorFlow（Google 發展之深度學習軟體）分析，TensorFlow已是現今AI領域最為廣泛使用之分析軟體。65

**執行初探研究以支持本研究之可行性**

目的

研究者已執行一個初探研究以支持本研究計畫之可行性。此初探研究的主要目的為：使用AI方法分析受測者擦桌子的影片，以同時評估受測者的4種能力（上肢動作、下肢動作、平衡以及行走）。

研究對象

研究者招募9位職能治療師及1位護理師進行研究。

研究流程

每位受測者需依據臨床經驗模擬3位不同能力程度的中風病人執行擦桌子的任務。研究者設定8種不同能力程度包含8種：4種能力（上肢動作、下肢動作、平衡以及行走）\* 2種程度（良好或不佳），而每位受測者只需模擬其中3種的排列組合。擦桌子任務具體內容為：請受測者從椅子上站起來，走到3公尺前的桌子邊緣，開始使用模擬的患側手擦桌子（圖七）。

D:\Desktop\圖片1.tif

圖七：初探研究之影片拍攝場景。使用二台攝影機拍攝模擬個案執行擦桌子活動，並以AI分析模擬個案之4種能力（上肢動作、下肢動作、平衡以及行走）。

資料分析

本研究共拍攝30部影片（10位受測者\*3種不同動作能力），研究者抽出每位受測者的前2段影片作為訓練資料（共20部），其它10部為測試資料。由於受測者於每次擦桌子的時候，都事先給定「需模擬的4種功能之各別程度」，如：上、下肢動作良好，但平衡及行走不佳。因此研究者使用每部影片的模擬的條件，當作該影片分析後的正確輸出結果，以訓練AI模型並計算AI模型分析的正確率。

研究者設計AI模型以計算出受測者的身體關鍵點座標（如：圖二之電腦視覺分析），接著再以AI的支持向量機 (support vector machine) 依據受測者的身體關鍵點位置，判斷受測者所模擬的4種能力各別的程度為良好或不佳。

最後研究者計算10部測試資料（影片）中，AI模型於4種能力的判斷正確率。

結果

10部影片中，AI模型均可正確判斷受測者之4種能力程度，4種能力之程度判斷正確率均達100%。

結論

本初探研究已顯示AI模型可從個案從事日常生活活動中（如：擦桌子），同時分析個案之4種功能（上肢動作、下肢動作、平衡以及行走）各自的能力程度。此初步結果支持本研究之可行性，可藉由AI技術分析3種評估情境影片（標準情境錄影、臨床情境錄影以及居家指定情境錄影），以評估個案之上述4種功能。此外，此初步分析亦顯示：本研究小組已掌握AI分析的關鍵技術。

然而，由於本初探研究的受測者非真正的中風患者，因此僅能要求受測者在各種能力下模擬2種不同的能力程度（良好及不佳），因此AI模型僅將各種能力粗分為2等級（良好及不佳）。未來研究中將採用效標測驗作為AIFS之學習對象，使得AIFS可更精準地評估中風患者之4種功能。

**預期可能遭遇之困難及解決途徑**

1. 本計畫須使用大量之AI技術，故而本研究小組已與成功大學資訊工程學系之連震杰教授合作，以提供AI技術支援，應能順利克服本研究可能遭遇之AI技術困難。連震杰教授之團隊致力於產學合作開發AI系統已近20年，近年來更致力於應用電腦視覺及深度學習於醫療器材領域，如：與成功大學前瞻醫療器材科技中心之長照2.0研發、與成功大學肌肉骨骼研究中心之精準醫療復健系統研發。於AI技術之應用端，連教授之團隊可提供豐富之經驗與技術。本研究小組與連教授已合作約一年，2019年底已共同發表多篇學術研討會論文（如：人工智慧方法發展短版傅格-梅爾上下肢動作評估量表於中風患者66、兒童人工智慧評估平台之發展方向與挑戰67 以及人工智慧評估應用於兒童著色畫之初探68）。本研究將有AI專家協助，主持人預期將能順利完成AIFS之發展與應用。
2. 資料上傳應用的平台需程式設計師協助建立，本研究小組長期合作之電腦工程師可協助及支援。
3. 本計畫需患者/照護者協助錄影患者之居家復健表現，並需定期、長期追蹤，可能降低受測者之意願或增加流失率。因此，本計畫編列受試者參與費，以提升患者/照護者參與動機願意參與研究，並降低追蹤之流失率。
4. 本計畫欲量化治療活動並輸出可使治療師、患者、其他醫療相關專業人員相互溝通之評估報告。因研究者與實務工作者之觀點可能存在落差，使得AIFS發展後未必能切合實務工作者之需求。為同時兼顧二方之需求，本計畫事先將舉辦多次的專家會議以確認雙方對於平台功能以及輸出之評估報告之需求。並編列專家諮詢費，使臨床實務工作者能對此系統提出建議。並且，系統發展後，本計畫亦將舉辦研討會，與臨床工作者與學術研究人員分享研究成果，並交流建議。
5. AIFS未來的錄影功能，可能和近來逐漸提倡的病人隱私和病人自主權會有衝突。未來推行AIFS初期，可優先應用於取得同意或自願之患者，待AIFS之成效日漸顯著且無個案資料外洩之風險，則可更進一步推廣至更多患者，或吸引更多自願接受評估之患者。
6. AIFS之AI分析可能遭遇之技術困難有4，包括 (1) 對多人之人臉辨識後，如何對個別樣本影像動作分流處理與辨識；(2) 在複雜背景下個案穿著各式顏色衣褲之上下肢及轉身後左右側區別；(3) 肩部與前臂的旋轉動作辨識；(4) 手指的細緻動作分辨等。為解決上述困難，本研究採用3台攝影機拍攝不同角度下一位或多位（2~4位）受測者之影像，並採用具備手指關節位置分析之AI人體動作分析套件（如OpenPose），以取得身體各關節座標與分析結果可信度（0~1，分數越高表示分析結果可信度越高）。之後，研究者先去除可信度較差之數值後，設計一個深度學習模型將3台攝影機所得之資訊共同當作輸入資料，以判斷受測者之「上下肢位置及轉身後左右側區別（困難2）」、「肩部與前臂的旋轉動作辨識（困難3）」以及「手指的細緻動作分辨（困難4）」。此外，研究者採用人臉辨識套件（如OpenFace）辨識個案身份及人臉在影片中的位置，即可配對人體動作分析套件所得之肢體位置，以穩定地於多人情境下配對受測者身份與動作。

**預期完成之工作項目及成果**

第一年：

1. 完成AIFS之3大功能設計。
2. 完成AIFS之資料蒐集：中風患者100位及正常人50位。
3. 發表 “Application of artificial intelligence in patients with stroke: a systematic review” 於國際學術研討會與學術期刊。

第二年：

1. 完成AIFS之電腦視覺分析。
2. 完成AIFS之計分系統建立。
3. 完成臨床測試於100位慢性中風患者，並完成再測信度驗證。
4. AIFS之追蹤評估：100位亞急性患者之追蹤評估，每位患者至少完成3次。
5. 發表 “Development of an artificial intelligence assessment system for patients with stroke” 於國際學術研討會與學術期刊。

第三年：

1. 完成AIFS之100位亞急性患者追蹤評估，每位患者完成6次。
2. 建立AIFS 4種功能恢復之預測模型。
3. 發表 “Responsiveness of the artificial intelligence assessment system for patients with stroke” 於國際學術研討會與學術期刊。
4. 發表 “Artificial neural network for predicting progression of 4 functions in patients with stroke” 於國際學術研討會與學術期刊。
5. 舉辦小型研討會 “1st mini-symposium and workshop of artificial intelligence in rehabilitation”以分享/交流本計畫之研究成果與方法，並聽取臨床與學術專家之建議。

**創新與預期貢獻**

創新：

AIFS可能為世界首創以電腦視覺技術評估中風患者之AI系統，可做為未來研究人員投入發展AI系統與復健領域之參考。

臨床貢獻有六：

1. 改變傳統評估方式以提升評估效能：AIFS可藉由電腦視覺技術自動且同時量化多位中風患者不同功能表現，可有效降低評估之人力、時間，提升評估效能。
2. 有利成效追蹤：錄影不須耗費太多人力且方便，有助於密集且持續的追蹤患者不同時間點的表現，進而確認其治療進步情況。
3. 提升評估結果之準確度：AIFS可分析患者於3種情境之功能（標準情境、臨床情境及居家指定情境之影片）且可多次錄影，並藉由持續學習多次錄影之巨大資料庫，修正原本之評分/預測模型，進而提升評估結果之準確度。
4. 擴大評估適用情境：錄影之適用性高，可將過去僅限於醫院情境之評估擴大至診所、家庭、長照機構等，增加評估與追蹤評估之可及性。
5. 提升治療成效：臨床人員可經由每次自動輸出之評估報告，確認患者功能是否如預期，以調整介入方式及目標，進而改善中風復健之品質與成效。並且，AIFS降低臨床人員之評估及病歷撰寫負擔後，臨床人員將可專注心力於治療活動中，因而提升治療品質。
6. 推廣至其它需評估肢體動作之族群：AIFS之評估方法具極高潛力應用於廣大需動作評估之族群/患者，如：兒童（動作發展）、老人（動作老化）以及脊髓損傷（動作/行動功能）等。

學術貢獻有二：

1. 建立大數據評估資料庫，提升評估結果應用價值：AIFS可藉由AI技術整合所有評估資料，包括患者之個人資訊、病情資料、多次且不同情境之評估結果等，累積形成大型電子資料庫，作為後續研究分析及精準醫療之基礎。如：療效驗證（特定治療方案最適用之患者特質及劑量）、功能復原之機轉及相關因子。
2. 建立跨專業研究團隊：本計畫將促進資訊工程與職能治療領域人員共同合作，不同領域的人各自貢獻自己的專長，將能共同提升復健醫療的服務品質。

**對於參與之工作人員，預期可獲之訓練**

1. 學習AI之理論與實務。
2. 獲得AIFS平台設計之經驗。
3. 增進職能治療活動分析之功能，更具體量化治療活動以及欲評估之功能，並連結評估功能與治療活動。
4. 學習執行研究計畫之統合能力、舉辦專家討論會議及評量工具心理計量特性驗證之理論與實務（包含資料收集與分析技巧等）。
5. 學習研究結果呈現及論文報告書寫。

**參考文獻**

1. 邱弘毅. 腦中風之現況與流行病學特徵. *腦中風會訊*. 2008;15:2-4

2. 衛生福利部中央健康保險署. 全民健康保險就醫疾病資訊. 2018

3. Langhorne P, Coupar F, Pollock A. Motor recovery after stroke: A systematic review. *Lancet Neurology*. 2009;8:741-754

4. Kim SS, Lee HJ, You YY. Effects of ankle strengthening exercises combined with motor imagery training on the timed up and go test score and weight bearing ratio in stroke patients. *Journal of Physical Therapy Science*. 2015;27:2303-2305

5. Lum PS, Burgar CG, Shor PC, Majmundar M, Van der Loos M. Robot-assisted movement training compared with conventional therapy techniques for the rehabilitation of upper-limb motor function after stroke. *Archives of Physical Medicine Rehabilitation*. 2002;83:952-959

6. Pandian S, Arya KN, Kumar D. Does motor training of the nonparetic side influences balance and function in chronic stroke? A pilot RCT. *The Scientific World Journal*. 2014;2014:769726

7. Langhorne P, Bernhardt J, Kwakkel G. Stroke rehabilitation. *Lancet*. 2011;377:1693-1702

8. Fugl-Meyer AR, Jääskö L, Leyman I, Olsson S, Steglind S. The post-stroke hemiplegic patient: A method for evaluation of physical performance. *Scandinavian Journal of Rehabilitation Medicine*. 1975;7:13-31

9. Bushnell C, Bettger JP, Cockroft KM, Cramer SC, Edelen MO, Hanley D, et al. Chronic stroke outcome measures for motor function intervention trials: Expert panel recommendations. *Circulation: Cardiovascular Quality Outcomes*. 2015;8:S163-S169

10. Berg K, Wood-Dauphinee S, Williams J. The Balance Scale: Reliability assessment with elderly residents and patients with an acute stroke. *Scandinavian Journal of Rehabilitation Medicine*. 1995;27:27-36

11. Berg KO, Wood-Dauphinee SL, Williams JI, Maki B. Measuring balance in the elderly: Validation of an instrument. *Canadian Journal of Public Health*. 1992;83:S7-S11

12. Kwong PW, Ng SS, Liu T-W, Chung RC, Ng GY. Effect of leg selection on the Berg Balance Scale scores of hemiparetic stroke survivors: A cross-sectional study. *Archives of Physical Medicine*. 2016;97:545-551

13. Benaim C, Pérennou DA, Villy J, Rousseaux M, Pelissier JY. Validation of a standardized assessment of postural control in stroke patients: The Postural Assessment Scale for Stroke Patients (PASS). *Stroke*. 1999;30:1862-1868

14. Mao H-F, Hsueh I-P, Tang P-F, Sheu C-F, Hsieh C-L. Analysis and comparison of the psychometric properties of three balance measures for stroke patients. *Stroke*. 2002;33:1022-1027

15. Wrisley DM, Marchetti GF, Kuharsky DK, Whitney SL. Reliability, internal consistency, and validity of data obtained with the Functional Gait Assessment. *Physical Therapy*. 2004;84:906-918

16. Lin J-H, Hsu M-J, Hsu H-W, Wu H-C, Hsieh C-L. Psychometric comparisons of 3 functional ambulation measures for patients with stroke. *Stroke*. 2010;41:2021-2025

17. Raghupathi W, Raghupathi V. Big data analytics in healthcare: Promise and potential. *Health Information Science and Systems*. 2014;2:3

18. Murdoch TB, Detsky AS. The inevitable application of big data to health care. *Journal of the American Medical Association*. 2013;309:1351-1352

19. 謝清麟, 林郁芬, 侯孟真, 李淑君, 唐世芬. 中風病人職能治療之病歷資料分析與改良. *職能治療學會雜誌*. 2014;32:269-280

20. Li X, Jiang H. Artificial intelligence technology and engineering applications. *Applied Computational Electromagnetics Society Journal*. 2017;32:381-388

21. Zhang L, Tan J, Han D, Zhu H. From machine learning to deep learning: Progress in machine intelligence for rational drug discovery. *Drug Discovery Today*. 2017;22:1680-1685

22. Amato F, López A, Peña-Méndez EM, Vaňhara P, Hampl A, Havel J. Artificial neural networks in medical diagnosis. *Journal of Applied Biomedicine*. 2013;11:47-58

23. Lisboa PJ. A review of evidence of health benefit from artificial neural networks in medical intervention. *Neural Networks*. 2002;15:11-39

24. Adams M, Chen W, Holcdorf D, McCusker MW, Howe PD, Gaillard F. Computer vs human: Deep learning versus perceptual training for the detection of neck of femur fractures. *Journal of Medical Imaging and Radiation Oncology*. 2019;63:27-32

25. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*. 2017;542:115–118

26. Chollet F. *Deep learning with python*. Manning Publications Co.; 2017.

27. Singh MS, Choudhary P. Stroke prediction using artificial intelligence. *Industrial Automation and Electromechanical Engineering Conference (IEMECON), 2017 8th Annual*. 2017:158-161

28. Lugade V, Lin V, Farley A, Chou L-S. An artificial neural network estimation of gait balance control in the elderly using clinical evaluations. *PloS One*. 2014;9:e97595

29. Oermann EK, Kress M-AS, Collins BT, Collins SP, Morris D, Ahalt SC, et al. Predicting survival in patients with brain metastases treated with radiosurgery using artificial neural networks. 2013;72:944-952

30. Shafiei E, Fakharian E, Omidi A, Akbari H, Delpisheh A, Nademi A. Comparison of artificial neural network and logistic regression models for prediction of psychological symptom six months after mild traumatic brain injury. *Iranian Journal of Psychiatry and Behavioral Sciences*. 2017;11:e5849

31. Sargent DJ. Comparison of artificial neural networks with other statistical approaches: Results from medical data sets. *Cancer*. 2001;91:1636-1642

32. Jiang J, Trundle P, Ren J. Medical image analysis with artificial neural networks. *Computerized Medical Imaging and Graphics*. 2010;34:617-631

33. Egmont-Petersen M, de Ridder D, Handels H. Image processing with neural networks-A review. *Pattern Recognition*. 2002;35:2279-2301

34. Pusiol G, Esteva A, Hall SS, Frank M, Milstein A, Fei-Fei L. Vision-based classification of developmental disorders using eye-movements. *International Conference on Medical Image Computing and Computer-Assisted Intervention*. 2016:317-325

35. Abedi V, Goyal N, Tsivgoulis G, Hosseinichimeh N, Hontecillas R, Bassaganya-Riera J, et al. Novel screening tool for stroke using artificial neural network. *Stroke*. 2017;48:1678–1681.

36. Hu C, Ju R, Shen Y, Zhou P, Li Q. Clinical decision support for Alzheimer's disease based on deep learning and brain network. *2016 IEEE International Conference on Communications (ICC)*. 2016:1-6

37. Kim J, Calhoun VD, Shim E, Lee J-H. Deep neural network with weight sparsity control and pre-training extracts hierarchical features and enhances classification performance: Evidence from whole-brain resting-state functional connectivity patterns of schizophrenia. *Neuroimage*. 2016;124:127-146

38. Zou L, Zheng J, Miao C, Mckeown MJ, Wang ZJ. 3D CNN Based Automatic diagnosis of attention deficit hyperactivity disorder using functional and structural MRI. *IEEE Access*. 2017;5:23626-23636

39. Jung H-T, Kim H, Oh MY, Ryu T, Kim Y. Learning classifier to evaluate movement quality in unassisted pick-and-place exercises for post-stroke patients: A preliminary study. *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 2017:2490-2493

40. Clark RA, Mentiplay BF, Hough E, Pua YH. Three-dimensional cameras and skeleton pose tracking for physical function assessment: A review of uses, validity, current developments and Kinect alternatives. *Gait and Posture*. 2019;68:193-200

41. Dawe RJ, Yu L, Leurgans SE, Truty T, Curran T, Hausdorff JM, et al. Expanding instrumented gait testing in the community setting: A portable, depth-sensing camera captures joint motion in older adults. *PloS One*. 2019;14:e0215995

42. Zhi YX, Lukasik M, Li MH, Dolatabadi E, Wang RH, Taati B. Automatic detection of compensation during robotic stroke rehabilitation therapy. *IEEE Journal of Translational Engineering in Health and Medicine*. 2017;6:1-7

43. Qiao S, Wang Y, Li J. Real-time human gesture grading based on OpenPose. *2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*. 2017:1-6

44. Julianjatsono R, Ferdiana R, Hartanto R. High-resolution automated Fugl-Meyer Assessment using sensor data and regression model. *Science and Technology-Computer (ICST), 2017 3rd International Conference on*. 2017:28-32

45. Prakash C, Kumar R, Mittal N. Recent developments in human gait research: Parameters, approaches, applications, machine learning techniques, datasets and challenges. *Artificial Intelligence Review*. 2018;49:1-40

46. Kim W-S, Cho S, Baek D, Bang H, Paik N-J. Upper extremity functional evaluation by Fugl-Meyer assessment scoring using depth-sensing camera in hemiplegic stroke patients. *PloS One*. 2016;11:e0158640

47. Antón D, Kurillo G, Goñi A, Illarramendi A, Bajcsy R. Real-time communication for Kinect-based telerehabilitation. *Future Generation Computer Systems*. 2017;75:72-81

48. Bailon C, Damas M, Pomares H, Banos O. SPIRA: An automatic system to support lower limb injury assessment. *Journal of Ambient Intelligence and Humanized Computing*. 2018

49. Chiu H-W, Huang Y-T, Cheng C-A. Using artificial neural network to predict functional recovery of patients treated by intravenous thrombolysis in acute ischemic stroke. *World Congress on Medical Physics and Biomedical Engineering 2018*. 2019:331-334

50. Gladstone DJ, Danells CJ, Black SE. The Fugl-Meyer assessment of motor recovery after stroke: A critical review of its measurement properties. *Neurorehabilitation Neural Repair*. 2002;16:232-240

51. Cao Z, Simon T, Wei S-E, Sheikh Y. Realtime multi-person 2D pose estimation using part affinity fields. *arXiv*. 2016;1611.08050v2

52. Redmon J, Farhadi A. YOLO9000: Better, faster, stronger. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2017:7263-7271

53. Berry KJ, Mielke Jr PW. A generalization of Cohen's kappa agreement measure to interval measurement and multiple raters. *Educational and Psychological Measurement*. 1988;48:921-933

54. Schroff F, Kalenichenko D, Philbin J. Facenet: A unified embedding for face recognition and clustering. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2015:815-823

55. Shrout PE, Fleiss JL. Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*. 1979;86:420-428

56. Cohen J. Statistical power analysis for the behavioral sciences. 2nd. 1988

57. Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Medicine*. 1998;26:217-238

58. Flansbjer UB, Holmback AM, Downham D, Patten C, Lexell J. Reliability of gait performance tests in men and women with hemiparesis after stroke. *Journal of Rehabilitation Medicine*. 2005;37:75-82

59. Huang SL, Hsieh CL, Wu RM, Tai CH, Lin CH, Lu WS. Minimal detectable change of the timed "Up & Go" test and the dynamic gait index in people with Parkinson disease. *Physical Therapy*. 2011;91:114-121

60. Dodrill CB, Troupin AS. Effects of repeated administrations of a comprehensive neuropsychological battery among chronic epileptics. *Journal of Nervous and Mental Disease*. 1975;161:185-190

61. Cohen J. *Statistical power analysis for the behavioural sciences*. Hillsdale, NJ: erlbaum; 1988.

62. Efron B, Tibshirani R. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical science*. 1986:54-75

63. Chiu C-C, Lee K-T, Lee H-H, Wang J-J, Sun D-P, Huang C-C, et al. Comparison of models for predicting quality of life after surgical resection of hepatocellular carcinoma: A prospective study. *Journal of Gastrointestinal Surgery*. 2018;22:1724-1731

64. Venta A, Ball Cooper E, Shmueli-Goetz Y, Sharp C. Artificial neural network coding of the child attachment interview using linguistic data. *Attachment and Human Development*. 2018;20:62-83

65. Abadi M, Barham P, Chen J, Chen Z, Davis A, Dean J, et al. Tensorflow: A system for large-scale machine learning. *2th USENIX Symposium on Operating Systems Design and Implementation*. 2016;16:265-283

66. 林恭宏, 李士捷, 黃千瑀, 謝清麟. 人工智慧方法發展短版傅格-梅爾上下肢動作評估量表於中風患者. *Taiwan Occupational Therapy Association 2019 Annual Meeting and International Conference*. 2019

67. 黃千瑀, 陳官琳, 謝清麟, 林恭宏. 兒童人工智慧評估平台之發展方向與挑戰. *Taiwan Occupational Therapy Association 2019 Annual Meeting and International Conference*. 2019

68. 李士捷, 陳官琳, 黃千瑀, 連震杰. 人工智慧評估應用於兒童著色畫之初探. *Taiwan Occupational Therapy Association 2019 Annual Meeting and International Conference*. 2019