2024第六屆臺灣藥學聯合學術研討會
方法學是社會暨管理藥學的關鍵

應用機器學習方法學於臨床藥學服務

許茜甯PhD, RPh

高雄長庚紀念醫院藥劑部臨床總藥師

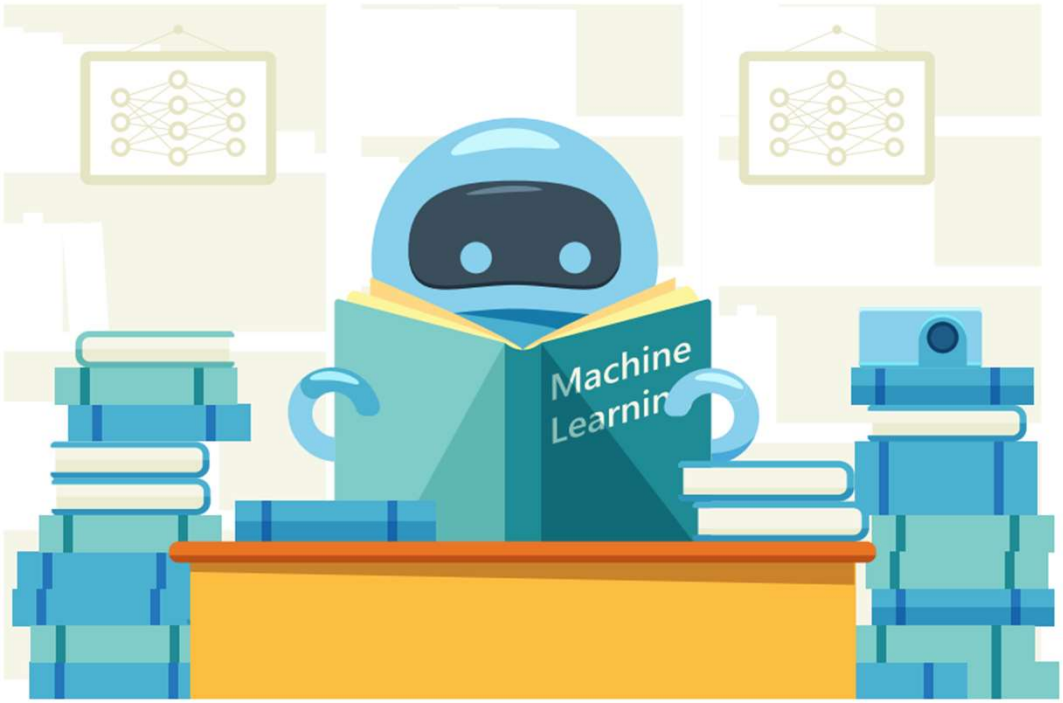
高雄醫學大學藥學系兼任副教授

Nov. 17, 2024

OUTLINE

- Applications of ML/AI in healthcare and process of ML/AI production
- Real-world example of ML predicted risk model (7 steps)
  - Community-acquired acute kidney injury (CA-AKI) hospitalization
- Summary of ML in kidney care: clinical applications and research

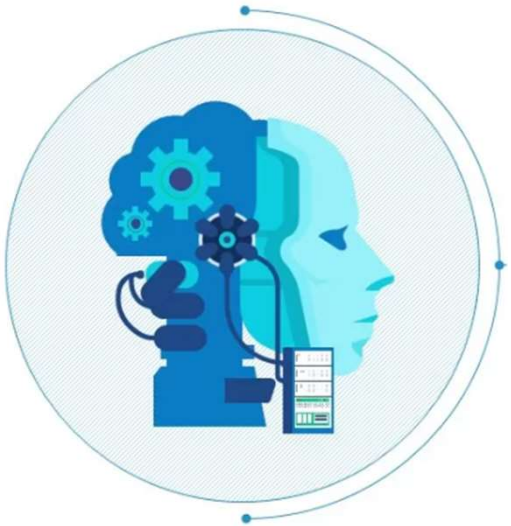
HOW IS ML/ AI HELPING US IN HEALTH CARE?



2024 第六屆「臺灣藥學聯合學術研討會」之歌

- Nature Medicine | VOL 27 | January 2021 | 2–3

APPLICATIONS OF ML/AI IN HEALTHCARE



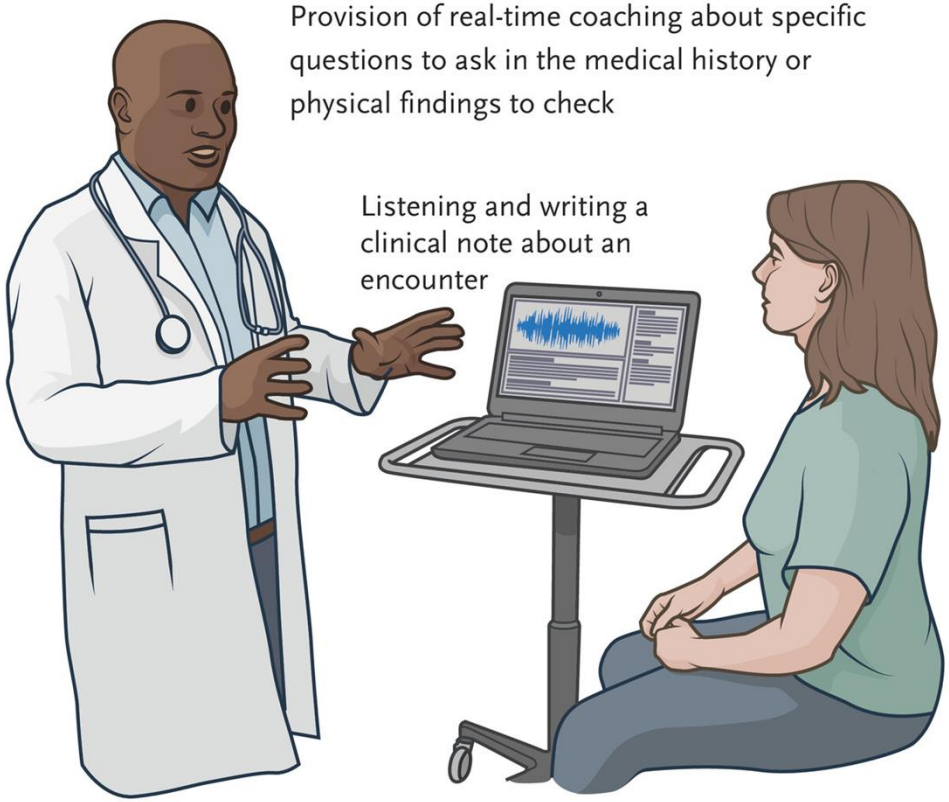
- **Developing diagnostics (diagnosis precision)**
  - 2018, the US FDA approved the first AI-based diagnostic, IDx-DR, for diabetic retinopathy by analyzing retinal images.
  - **950 AI/ML-enabled medical devices authorized by US FDA (US FDA Aug, 2024)**
- **Improving prognostics (clinical event prediction)**
  - A mathematical equation that calculates an individual's risk of an outcome event based on the characteristics (predictor) → A prediction tool: risk scores (scoring, nomograms)
  - Individualized treatment decision- making, risk stratification (disease prevention or early detection)
- **Patient monitoring (digital health, keep well and support)**
  - Smartphone or biological sensors
- **New drug discovery and precision medicine**
- **Smart hospital (automation → digitalization → AI)**
  - to avoid human error, improving human resources (e.g. dispensing, robotic surgery)
  - AI agent: medical history taking, medical education

醫學教育

https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices

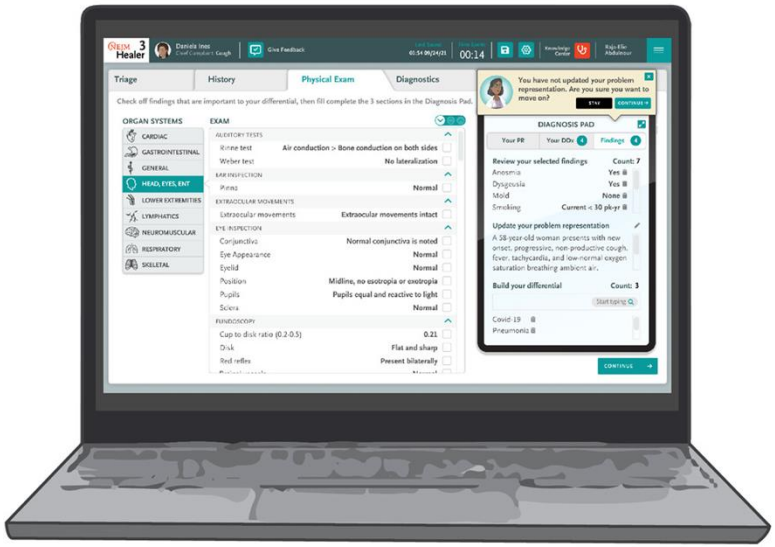
Provision of real-time coaching about specific questions to ask in the medical history or physical findings to check

Listening and writing a clinical note about an encounter



Serving as a teacher and an assessor in medical education

Creating realistic “flight simulators” for simple and complex patient encounters

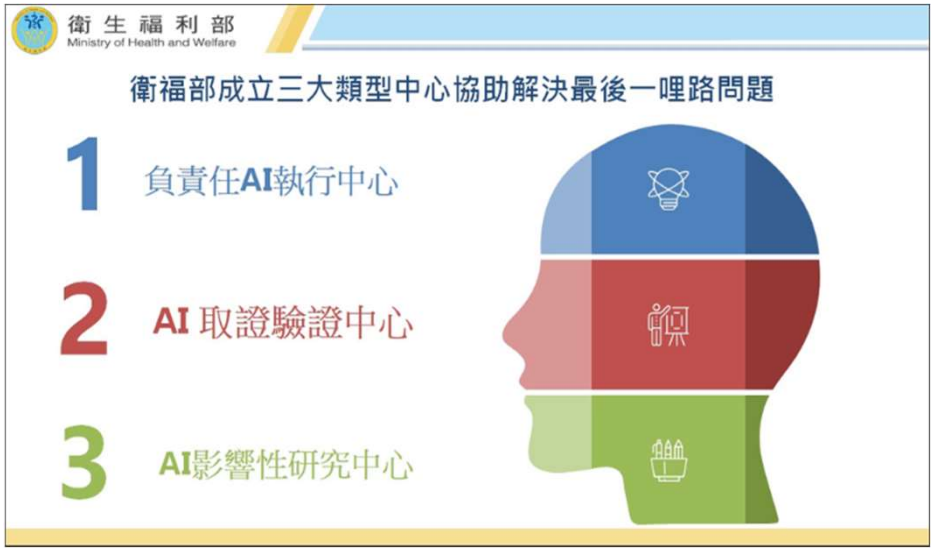


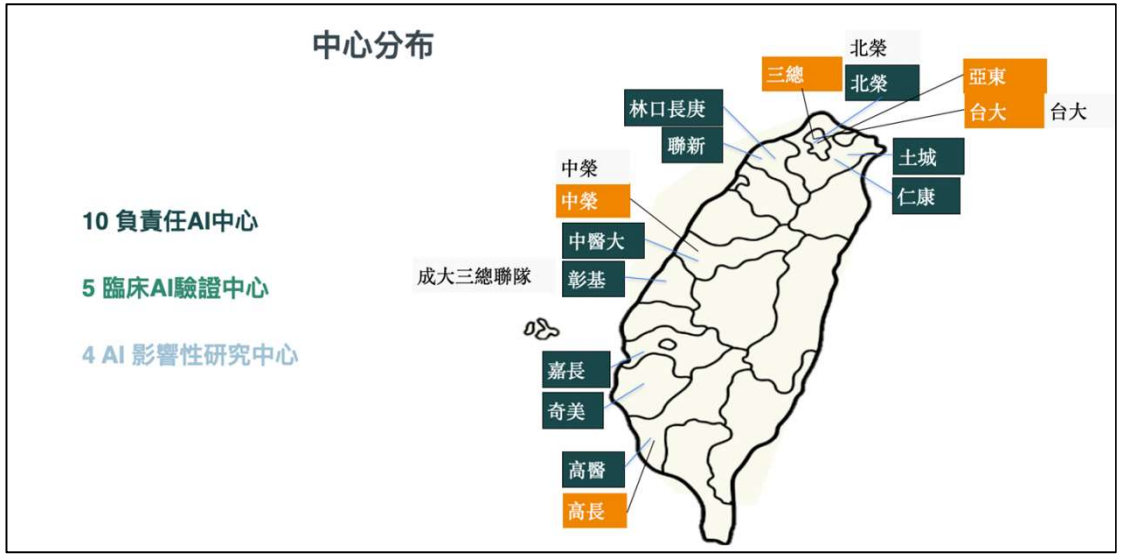
N Engl J Med 2023;388:1201-1208

衛福部3大AI中心啟動！促智慧醫療落地 完成最後1哩路

2024-10-07

跨部門解決「落地」、「取證」、「給付」瓶頸

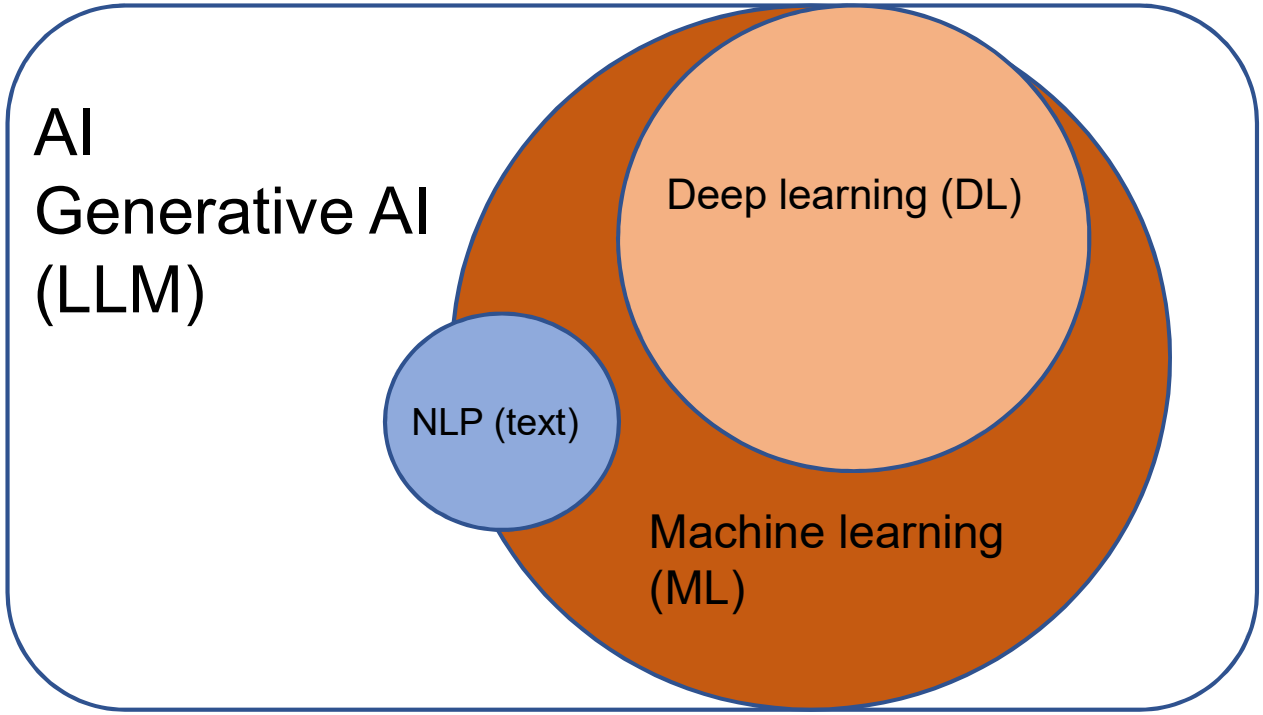




https://technews.tw/2024/10/07/three-major-ai-centers/

AI VS ML VS GENERATIVE AI

- Artificial intelligence (AI) primary is a branch of computer science (data science) that concerns building “intelligent” (thinking like human) smart machines to performing intelligent tasks.
- LLM (large language model), a large DL model designed for NLP tasks, such as human language generation.

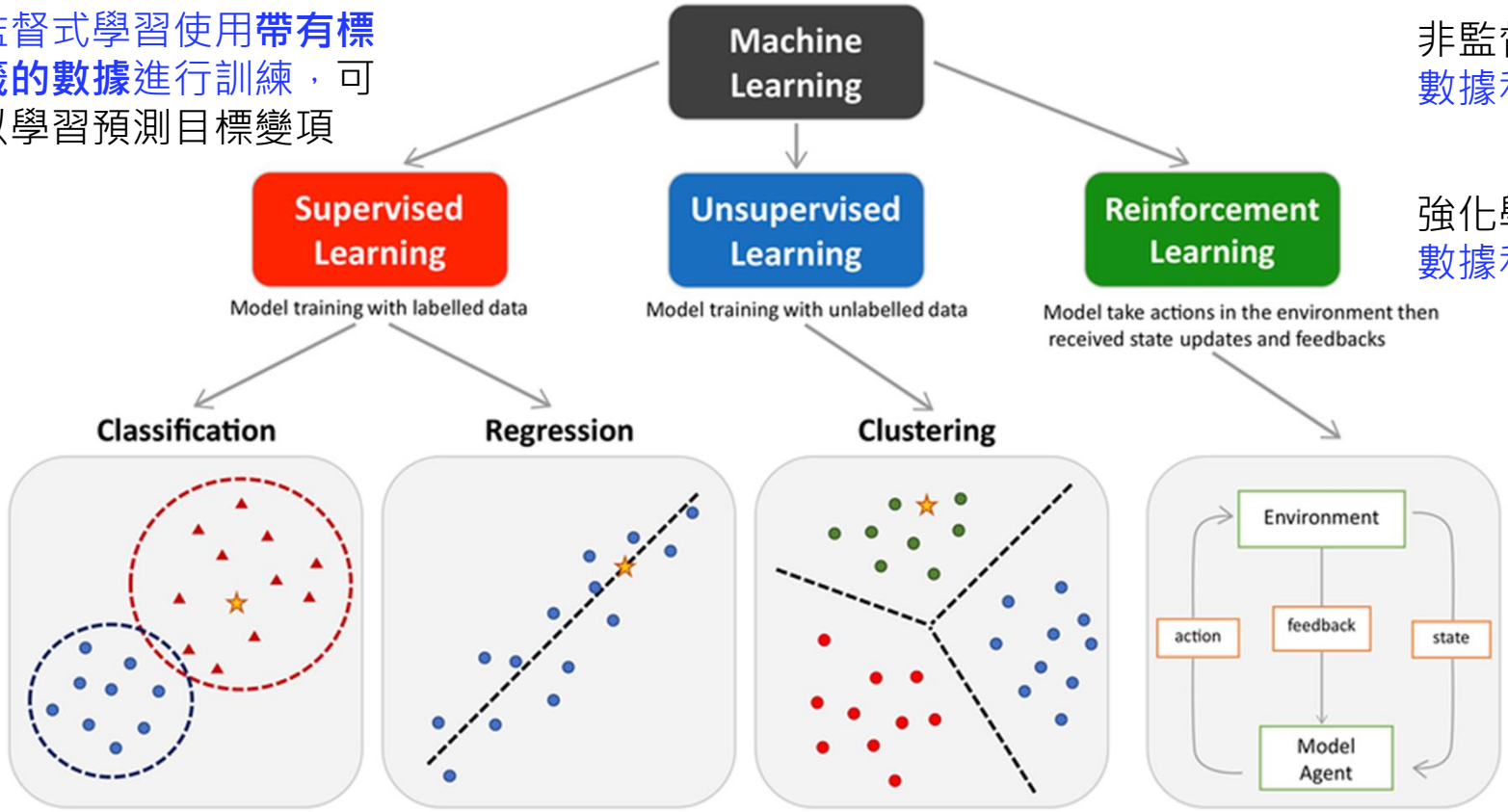


MAIN TYPES OF MACHINE LEARNING

監督式學習使用**帶有標籤的數據**進行訓練，可以學習預測目標變項

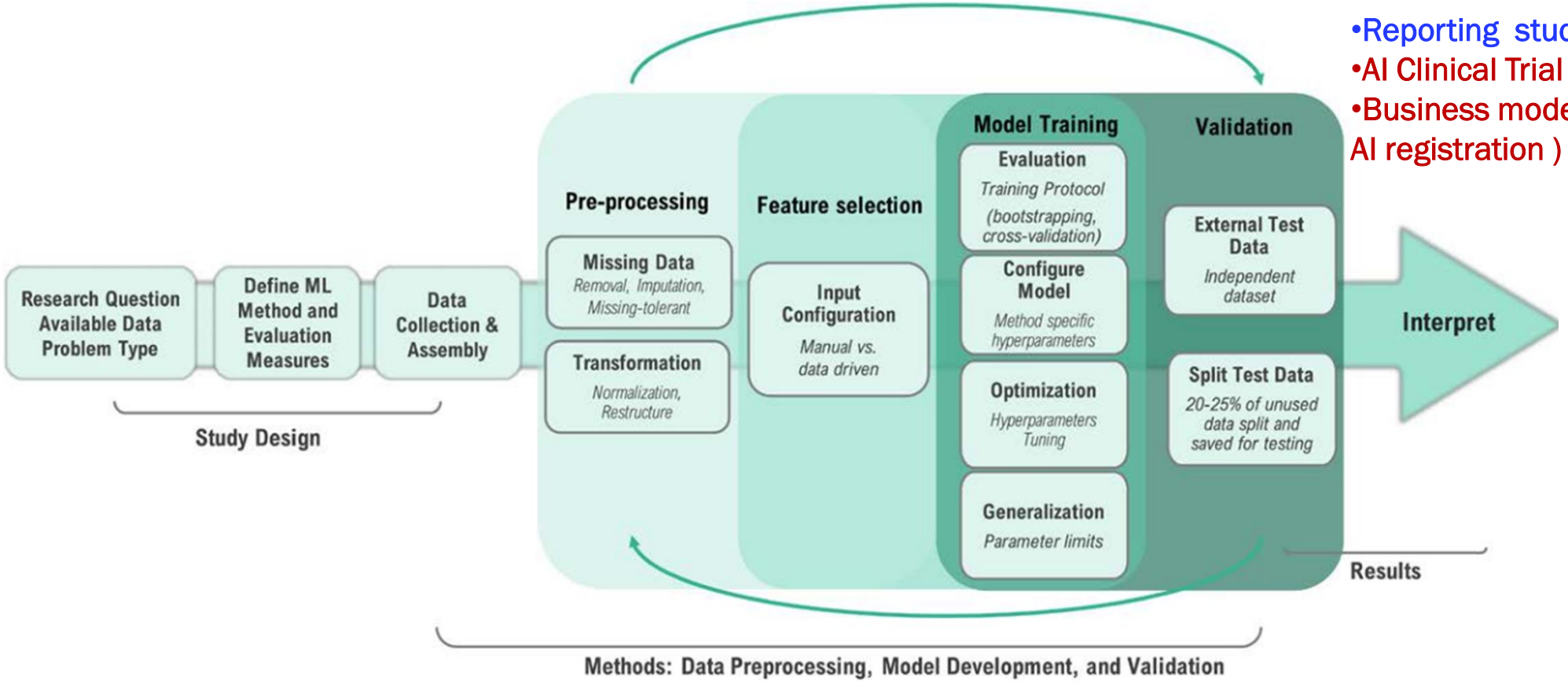
非監督式學習處理無標籤數據和發現隱含結構模式

強化學習可以處理無標籤數據和發現隱含結構模式



Frontiers in pharmacology, 2021, 12: 720694.

OVERVIEW OF ML STUDY DESIGN FLOW



- Reporting study results
- AI Clinical Trial (跨院驗證)
- Business model (medical AI registration)

Circ Cardiovasc Qual Outcomes. 2020;13:e006556

MAIN IDEA OF THE TALK

1.Clear study question and circumstance of its intended use

2.Data, Study design and ML analytic methods (7 steps)

- TRIPOD+AI checklist for reporting prediction models

Published on 4.8.2020 in Vol 22, No 8 (2020):August

Preprints (earlier versions) of this paper are available at https://preprints.jmir.org/preprint/16903, first published November 04, 2019.

Machine Learning Model for Risk Prediction of Community-Acquired Acute Kidney Injury Hospitalization From Electronic Health Records: Development and Validation Study

Chien-Ning Hsu [1, 2]; Chien-Liang Liu [3]; You-Lin Tain [4]; Chin-Yu Kuo [3]; Yun-Chun Lin [3]

Citation

Please cite as:

Hsu C, Liu C, Tain Y, Kuo C, Lin Y
Machine Learning Model for Risk Prediction of Community-Acquired Acute Kidney Injury Hospitalization From Electronic Health Records: Development and Validation Study
J Med Internet Res 2020;22(8):e16903
doi: 10.2196/16903
PMID: 32749223
PMCID: 7435690

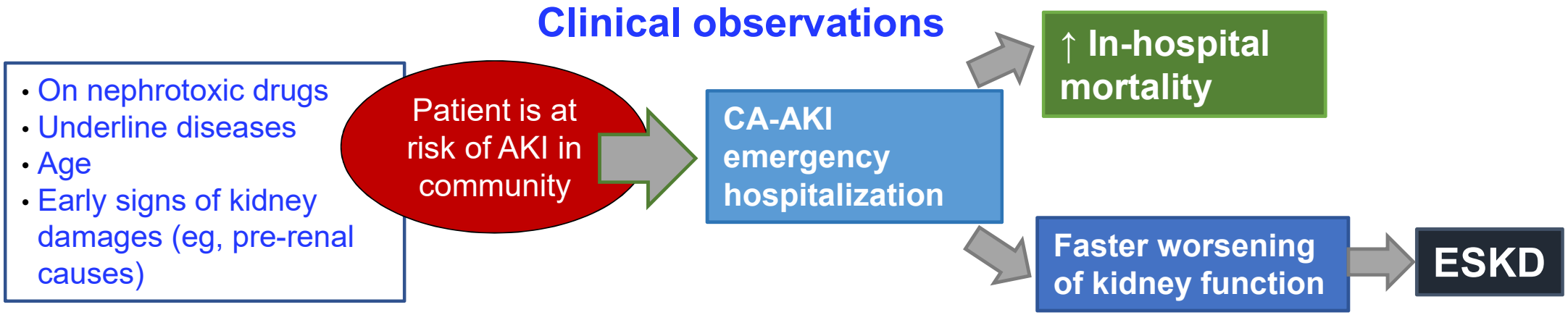
Copy Citation to Clipboard

2023 HEALTH CARE SCIENCES & SERVICES Ranking <5%

The risk calculator is qualified as a SaMD (software as medical device) for the clinical decision support system (US FDA) and Taiwan FDA guidelines

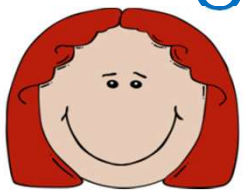
CAAKI-related hospitalization

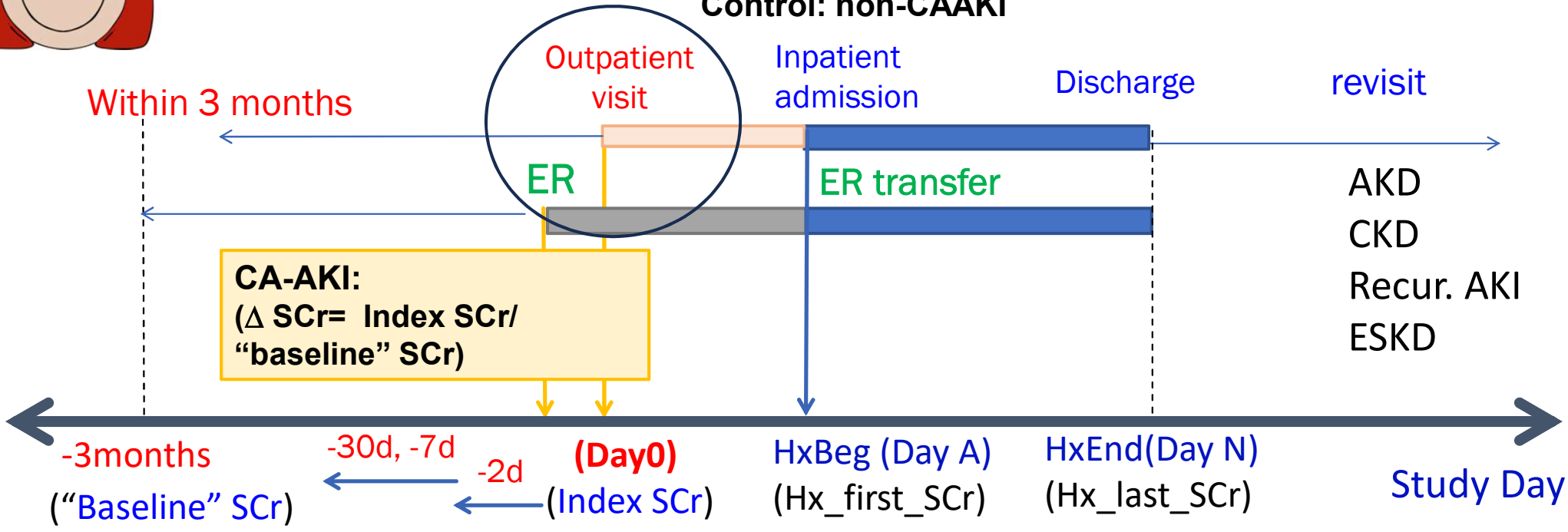
Objective: to develop a risk calculator to early identification and prevention of AKI hospitalization in the outpatient setting



AKI= acute kidney injury; CA-AKI= community-acquired AKI; CKD= chronic kidney disease; ESKD=end-stage kidney disease

Case-control study design to label training samples





- CA-AKI=Community-acquired acute kidney injury
- AKD=Acute kidney disease , CKD=Chronic kidney disease
- LOS= length of stay in hospitalization , SCr=serum creatinine

$$\text{CA-AKI} = \frac{(SCr_{index})}{(SCr_{baseline})} \geq 1.5 \; or$$

$$SCr\ baseline\ increase\ to \geq 4\ (SCr\ index)$$

Machine Learning Model for Risk Prediction of Community-Acquired Acute Kidney Injury Hospitalization From Electronic Health Records: Development and Validation Study

Chien-Ning Hsu, Chien-Liang Liu, You-Lin Tain, Chin-Yu Kuo, Yun-Chun Lin

J Med Internet Res 2020;22(8):e16903

01 **Data preprocessing**
Machine learning algorithms

02 **Important variable selection**
Model evaluation: AUROC
(Discriminatory ability)

03 **Logistic regression: scoring function**

15 files in CGRD repository

Training

CGRD (2010-2017)

Derivation Cohort (2010-2016)

(n=204,064)

(n=30,803)

Validation Cohort (2017)

Temporal validation

Data preprocessing

Feature Selection

Model Training

Scoring Function

DeLong's test

High risk (Precision care)

Low risk (usual care)

1.DATA PREPARATION FOR MACHINE LEARNING

- **Data Source:** CGMHs (CGRD repository, 2010-2017)
- **Data dimension (total 15 files)**
  - Laboratory results
  - Medical encounters (i.e., outpatient, inpatient, emergency department)
  - Diagnoses
  - Medication uses
- **Data cleaning/missing imputation and transformation**
  - 80 variables (initially)
  - 234,867 hospitalized adults: Case:(n=19,448) vs Controls (n=215,419), 8.3% (CAAKI Hx)
- **Internal validation**
  - 2010-2016 (training and internal validation in the development dataset)
  - 2017 (temporal validation in the testing dataset)

■ Quality of data
- Volume
- Dimension
- Currency/timing
- Completeness/ correctness/ concordance

2. DATA PRE-PROCESSING (1)

- Misclassification
- Data leakage (prevalent vs incident case)

Data cleaning and labelling

- Remove the data with NA data at admission, including patients who died before ER transfer
- $Ratio = \frac{Index\_SCR}{Basline\_SCR}$

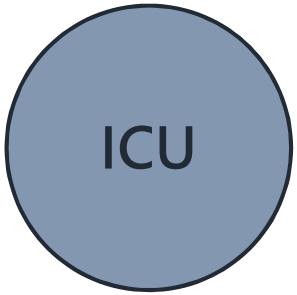
$$\boldsymbol{Label} = \begin{cases} 0, & Ratio < 1.5\ (non - AKI\ at\ admision) \\ 1, & Ratio \geq 1.5\ (AKI\ at\ admission) \end{cases}$$

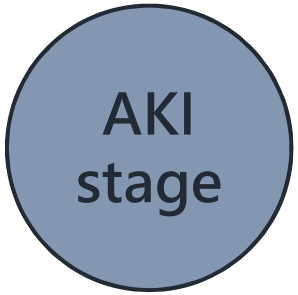
Dropping features

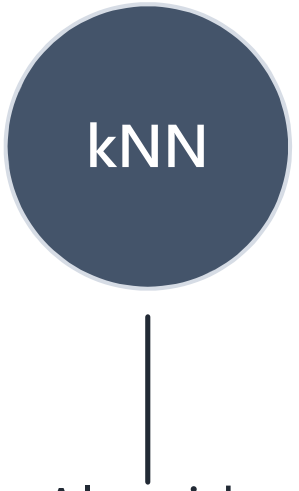
- XGBoost could handle missing data to complete the preliminary selection variables

2. DATA PRE-PROCESSING (2)

- Imputation of missing data (continuous data)







Group	Calculating the Median of Variables from each Group		
ICU	base_BUN	base_K	base_P
0 (No)	29.4	4.1	4.25
1 (Yes)	24	4	3.95

2. DATA PRE-PROCESSING (IMPUTATION METHODS)

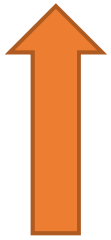
- Results of different imputation methods
- Using XGBoost to compare different method of filling missing values

XGBoost with all features			
Methods	**AUC**	**Sensitivity**	**Specificity**
Statistic			
AKI stage	0.805 ± 0.012	0.770 ± 0.049	0.682 ± 0.037
ICU	**0.807 ± 0.004**	**0.747 ± 0.058**	**0.712 ± 0.055**
Algorithm			
KNN	0.801± 0.013	0.764±0.049	0.684±0.043

3. FEATURE (VARIABLE) SELECTION METHODS

- Advantages of feature selection:

- Complexity
- Over-fitting



- Interpretability
- Efficiency

Feature selection List		
Filter method	**Wrapper method**	**Embedded method**
• Symmetrical Uncertainty	• Recursive Feature Elimination (RFE) based on Random Forest (RF) • Stepwise Logistic Regression (LR)	• Least Absolute Shrinkage and Selection Operator(LASSO) • XGBoost • Light Gradient Boosting Machine (LightGBM)

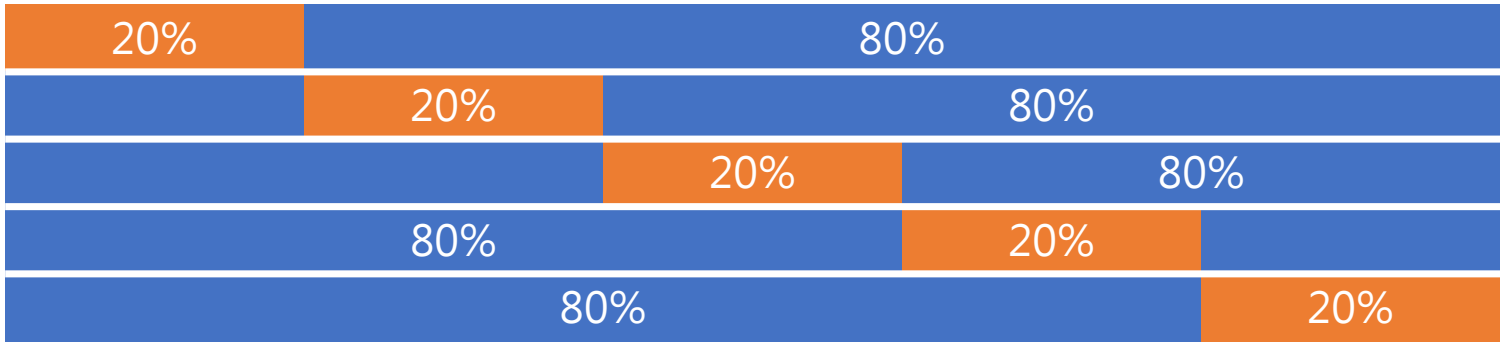
4. MACHINE LEARNING ALGORITHMS SELECTION FOR CLASSIFICATION

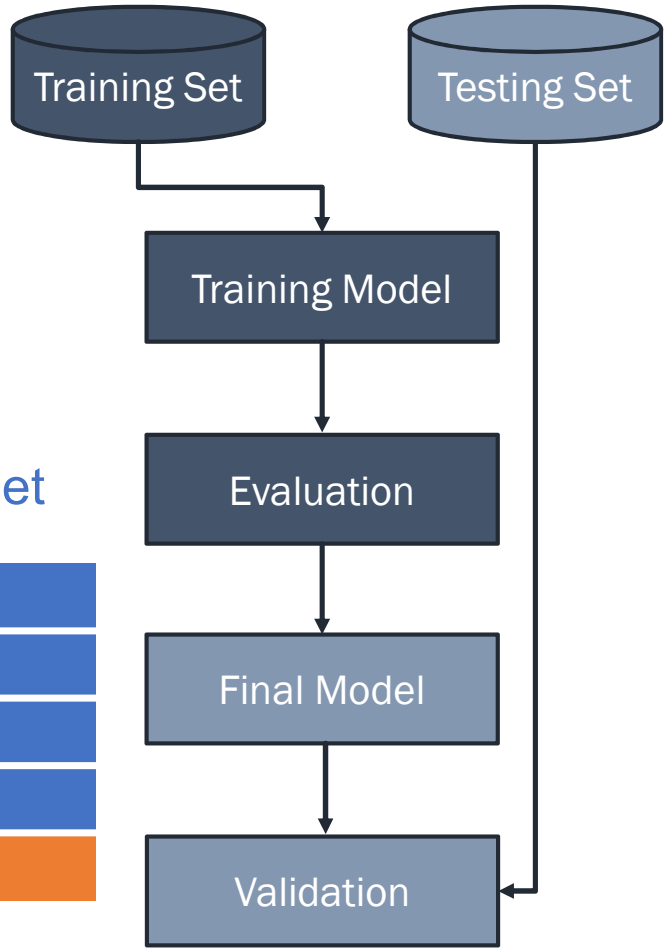
- **Prediction:** with/without CA-AKI at admission in inpatients
- **Machine learning (ML) algorithms**
  - XGBoost (eXtreme Gradient Boosting)
  - LASSO (Least Absolute Shrinkage and Selection Operator)
    - Ridge Regression
  - Random forest
    - Decision Tree
  - **Logistic regression (referent method for classification)**

- Deep learning (subarea of ML)
  - Artificial neural networks (ANNs)
  - convolutional neural networks (CNNs),
  - recurrent neural networks (RNNs),
  - long short-term memory networks (LSTMs)

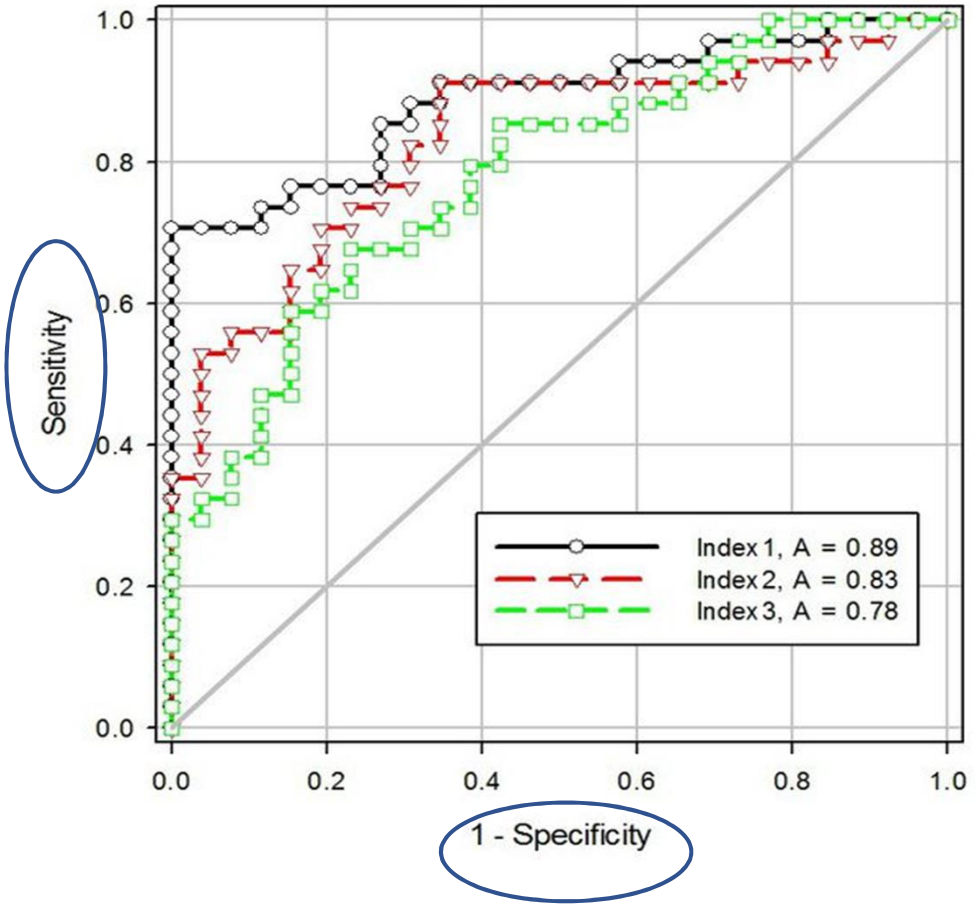
4. MODEL CONSTRUCTION

- **Data: Training set (70 or 80% all data)**
- **Algorithms:**
  a) **XGBoost**
  b) Logistic Regression
  c) LASSO
  d) Random Forest
  e) LightGBM
- Internal validation (5- or 10-fold cross validation) in training set





5. AUROC (CLASSIFICATION)

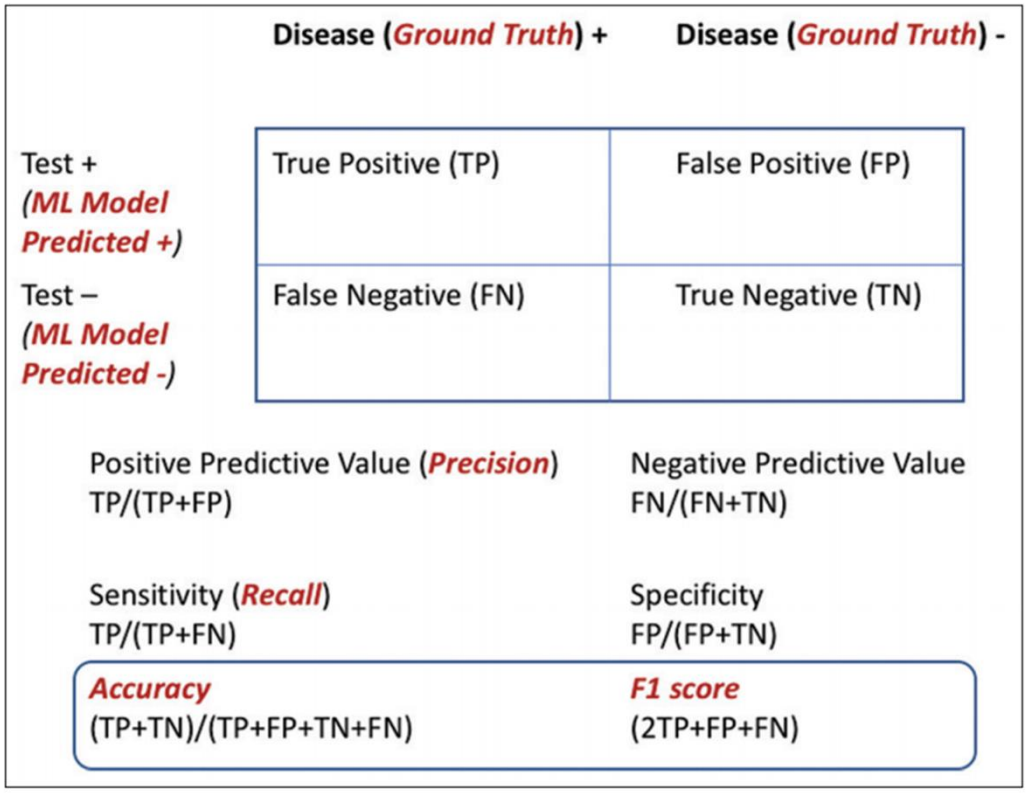


- AUROC 取決sensitivity/specificity (cut-off point)，自行設定或由Youden index

  AUROC= area under receiver operating characteristics curves

  - **AUROC = 0.5 (no discrimination 無辨識力)**
  - 0.7≦AUC≦0.8 (acceptable discrimination 可接受的鑑別力)
  - 0.8≦AUROC≦0.9 (excellent discrimination 優良的鑑別力)
  - 0.9≦AUROC≦1.0 (outstanding discrimination 極佳的辨識力)

5.MODEL EVALUATION (CLASSIFICATION)



- Is the model performance too good to be true?
  - Data leakage
  - Overfit

- Sensitivity (True Positive Rate , TPR), Recall (「靈敏性, 有病的偵測率」。
- Specificity (True negative rate, TNR)，「特異性, 沒病的偵測率」。
- Accuracy (判斷的正確率)，越高越好。
- Precision，Positive predictive value (PPV) : 模型結果呈現有病，而且確實有病的比率，越高越好。(與群體中疾病發生率有關)
- Negative predictive value (NPV) : 模型結果呈沒病，而且確實沒有病的比率，越高越好。
- F1: 合併recall and precision

5.MODELING-BUILDING STRATEGY

- Testing dataset: internal temporal validation (2017)
- Size of predictive model

XGBoost model with all (47) variables	
AUC	0.8011±0.0093
Sensitivity	0.6706±0.064
Specificity	0.7715±0.0661
Names	importance
Baseline SCr	100
Baseline eGFR	59.67
Baseline BUN	16.02
Baseline Calcium	12.41
Age at admission	9.46
CKD	8.55
RASI/Diuretics	6.77
Baseline Phosphorus	5.97
Diabetes without complications	3.53
Severe liver disease	3.18
Baseline SUA	2.59
Any tumor	2.37
Baseline HbA1c	2.26
Baseline TG	1.65
Anti-cholinergic agents	1.57
Sum of class of anti-microbial agents (1-7)	1.5
……	

XGBoost model with top 10 variables	
AUC	0.789±0.004
Sens	0.651±0.017
Spec	0.7764±0.0205
Names	importance
Baseline SCr	100
Baseline eGFR	18.18
Baseline BUN	11.12
Baseline Calcium	8.47
CKD	5.18
Age at admission	4.08
RASI/Diuretics	3.62
Baseline Phosphorus	3.38
Diabetes without complications	0.62
Severe liver disease	0

XGBoost model with top 5 variables	
AUC	0.7438±0.0045
Sens	0.542±0.0235
Spec	0.8321±0.0278
Names	importance
Baseline SCr	100
Baseline eGFR	57.98
Age at admission	30.57
Baseline Calcium	6.48
Baseline BUN	0

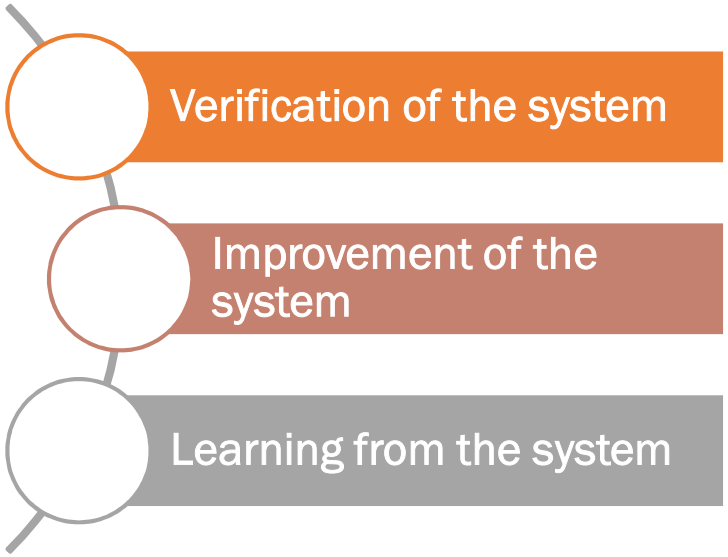
- Is the model performance too good to be true?
  - ➢ Unseen new association

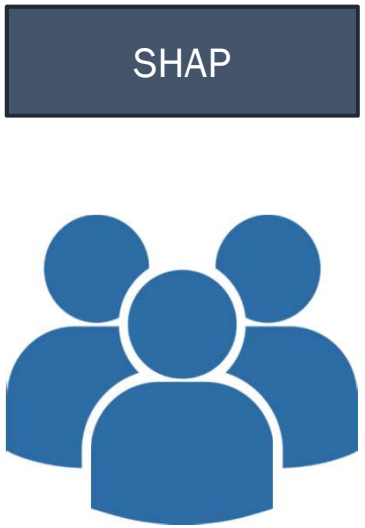
5. 4-DIMENSION,10-IMPORTANT FACTORS

Dimension (47)	Important feature (10)
Demographic information	Age at index hospital admission
Charlson comorbid condition (19 conditions) (<12 months, including Day 0)	• Diabetes without complication • Chronic kidney disease • Severe liver disease
Recent use of medicines (19 classes) (<3 months, not including Day 0)	RASI/K-sparing diuretics (2 classes)
Baseline laboratory results (<3 months, not including Day 0)	• SCr • eGFR • BUN • Calcium • Phosphorus

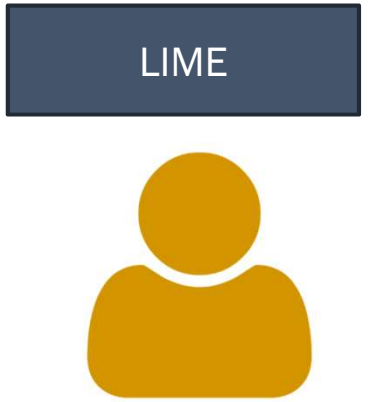
6.MODEL EXPLANATION

- To improve the explainability of prediction model





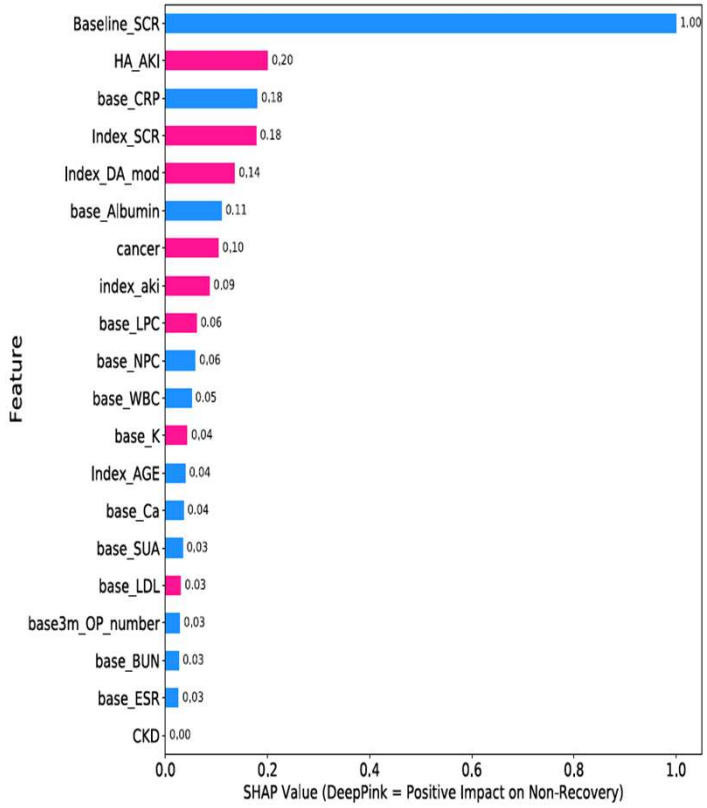
Overall Explanation

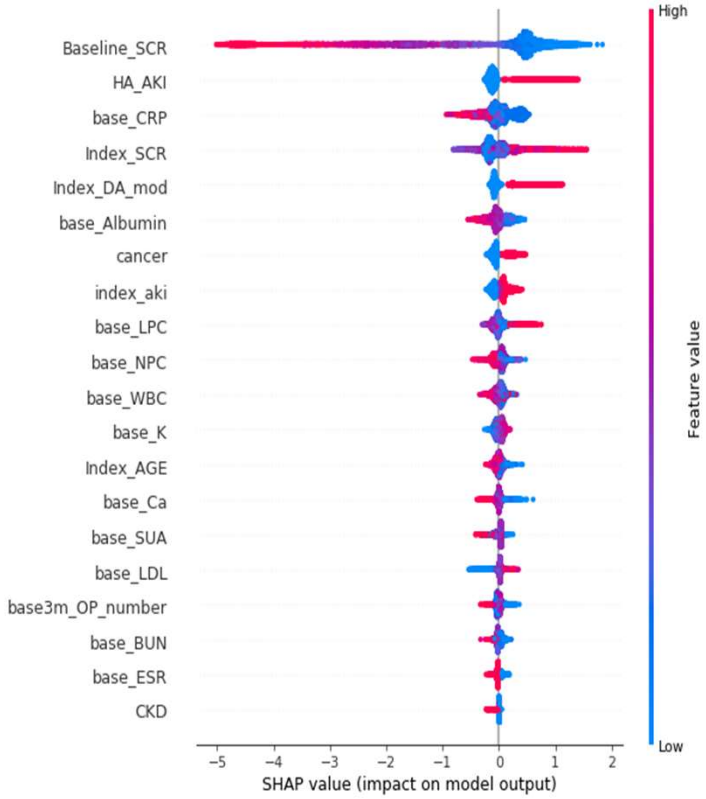


Individual Explanation

6.MODEL EXPLANATION (explainability):

SHAP (SHapley Additive exPlanations) framework to explain the models' predictions



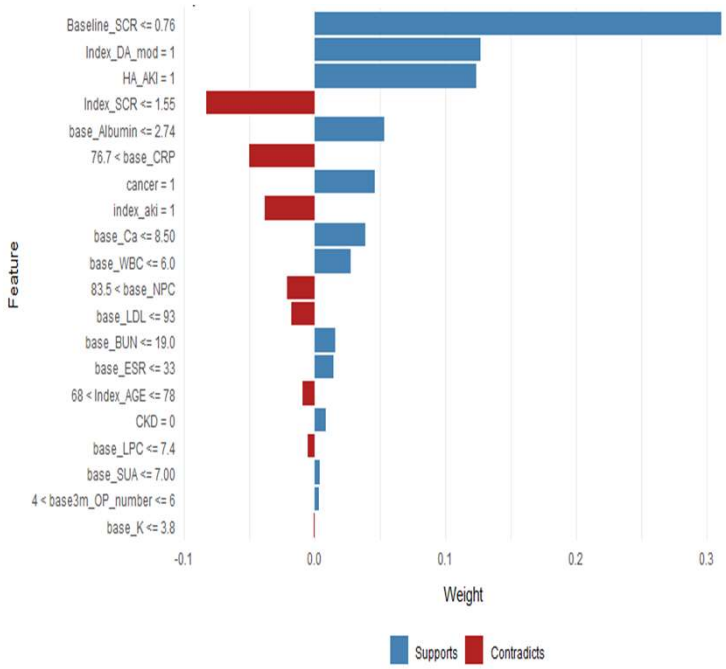


Front Med 2022; 8:789874

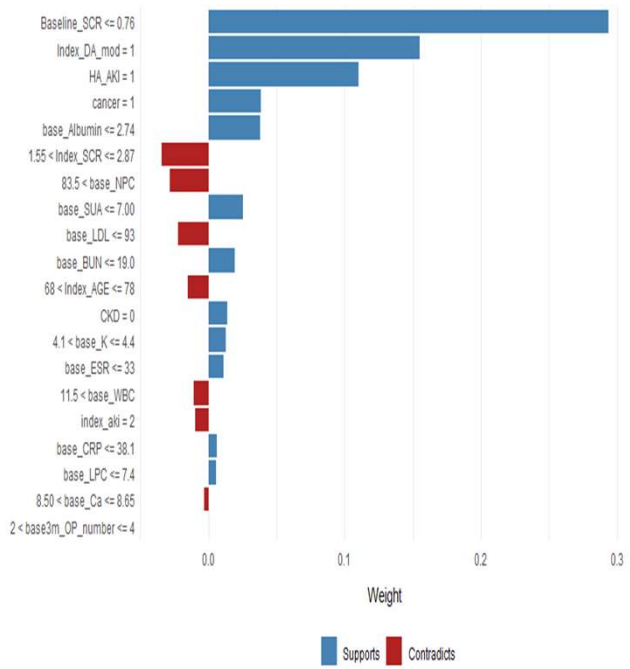
6.MODEL EXPLANATION (explainability)

Local Interpretable Model-agnostic Explanations (LIME)

A: True AKI non-recovery



B: False AKI non-recovery



A: True AKI non-recovery		B: False AKI non-recovery	
Predication probability		**Predication probability**	
non-recovery (true)	0.960	non-recovery (true)	0.948
reovery (false)	0.040	recovery (false)	0.052
ID	2231	ID	2388
Feature	**Actual value**	**Feature**	**Actual value**
base_SCR	0.59	base_SCR	0.68
index_SCR	0.98	index_SCR	1.99
HA_AKI	yes	HA_AKI	Yes
base_BUN	13.3	base_BUN	10.1
index_DA_mod	Yes	index_DA_mod	Yes
base_CRP	166.72	base_CRP	30.76
base_Albumin	1.93	base_Albumin	2.01
base_LPC	2	base_LPC	7
Cancer	Yes	Cancer	Yes
index_aki	Yes	index_aki	YES
base_WBC	1.6K	base_WBC	14K
CKD	No	CKD	No
index_AGE	70	index_AGE	76
base_NPC	85	base_NPC	88
base_Ca	8.01	base_Ca	8.62
base_SUA	7	base_SUA	4.62
base_K	3.1	base_K	4.3
base3m_OP_number	6	base3m_OP_number	3
base_LDL	93	base_LDL	93
base_ESR	33	base_ESR	33

Front Med 2022; 8:789874

7. REPORTING ML PREDICTIVE MODEL

- April 18, 2024
- 27-item checklist

RESEARCH METHODS AND REPORTING

TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods

Gary S Collins,[1] Karel G M Moons,[2] Paula Dhiman,[1] Richard D Riley,[3,4] Andrew L Beam,[5] Ben Van Calster,[6,7] Marzyeh Ghassemi,[8] Xiaoxuan Liu,[9,10] Johannes B Reitsma,[2] Maarten van Smeden,[2] Anne-Laure Boulesteix,[11] Jennifer Catherine Camaradou,[12,13] Leo Anthony Celi,[14,15,16] Spiros Denaxas,[17,18] Alastair K Denniston,[4,9] Ben Glocker,[19] Robert M Golub,[20] Hugh Harvey,[21] Georg Heinze,[22] Michael M Hoffman,[23,24,25,26] André Pascal Kengne,[27] Emily Lam,[12] Naomi Lee,[28] Elizabeth W Loder,[29,30] Lena Maier-Hein,[31] Bilal A Mateen,[17,32,33] Melissa D McCradden,[34,35] Lauren Oakden-Rayner,[36] Johan Ordish,[37] Richard Parnell,[12] Sherri Rose,[38] Karandeep Singh,[39] Laure Wynants,[40] Patricia Logullo[1]

BMJ. 2024;385:e078378.

TRIPOD+AI CHECKLIST FOR PREDICTION MODEL STUDIES

- **Methods: item 5 to 17**

Sample size	10	D;E	Explain how the study size was arrived at (separately for development and evaluation), and justify that the study size was sufficient to answer the research question. Include details of any sample size calculation
Class imbalance	13	D;E	If class imbalance methods were used, state why and how this was done, and any subsequent methods to recalibrate the model or the model predictions

D;E=items applicable to both the development and evaluation of a prediction model.

A class imbalance where disease categories are not equally represented in the development dataset (i.e. incidence, background risk),

- Over- or under presented class
- Solutions: find the better data or larger size of dataset. if it is possible, tuning techniques

.

ML IN KIDNEY CARE

腎臟病

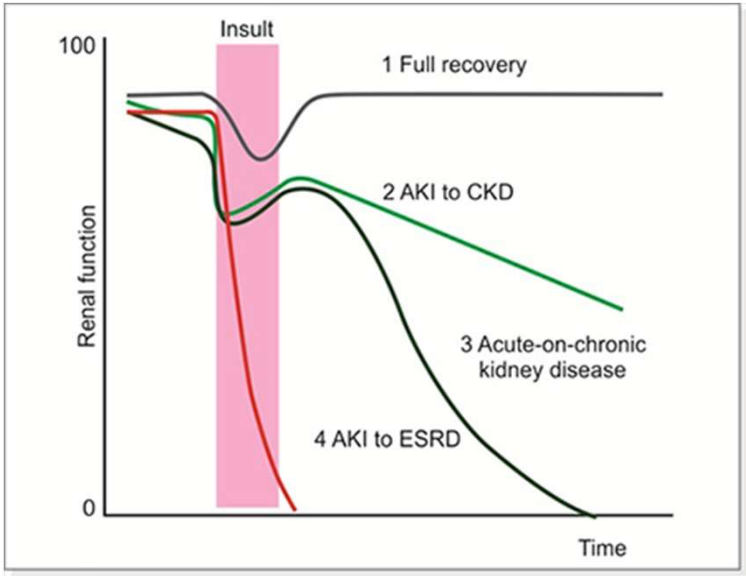
- 2023年 KDIGO 估計全球超過8億5千萬人具有某一類型腎臟病，而透析人口隨著老年化社會持續增加，是全球共同面對的挑戰
- 藥物 (藥品、顯影劑、成藥、中草藥)引起的腎毒性範圍廣泛，適當用藥管理策略才能避免藥物造成腎臟功能受損
- AKI(急性腎損傷)是短時間內快速腎功能下降；若不能恢復，提高CKD發生風險或加速腎功能惡化形成末期腎病，加重透析治療的需求以及病人、社會的負擔

KDIGO 2012 Clinical Practice Guidelines for AKI

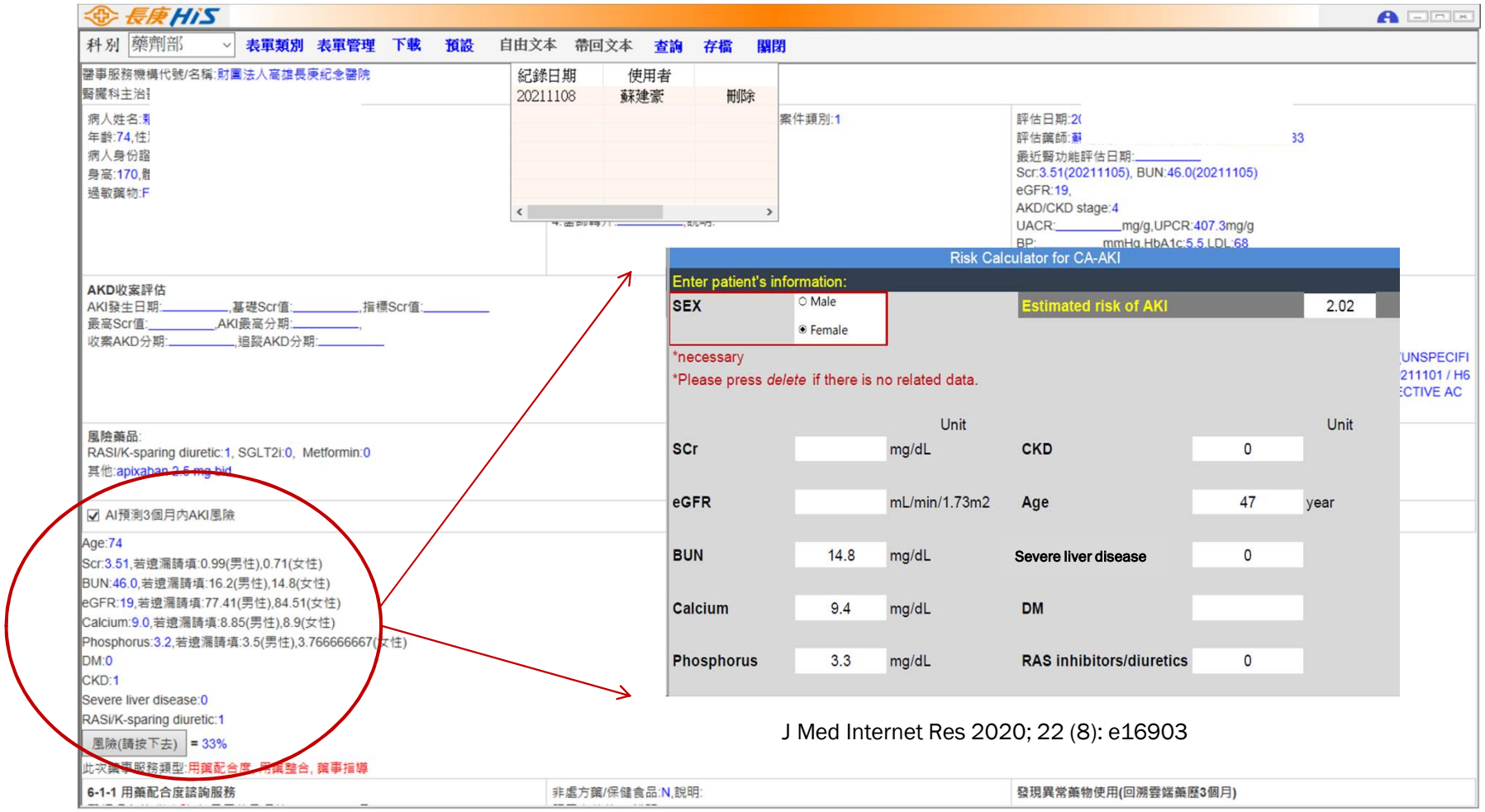
Kidney disease	Functional criteria
AKI (Acute kidney disease)	↑SCr by 50% within 7 days, or ↑SCr by 0.3 mg/dl within 2 days, or oliguria
AKD (Acute kidney disease)	AKI, or Kidney damage for <3 months, or ↓GFR by 35% or ↑SCr by 50% for <3 months
CKD (Chronic kidney disease)	GFR<60 for >3 months

KDIGO: Kidney Disease: Improving Global Outcomes; CJASN 2008; 3:881-886

Consequence of AKI



3個月內CA-AKI住院風險指數計算器 (PRE-ESRD 藥師門診HIS)

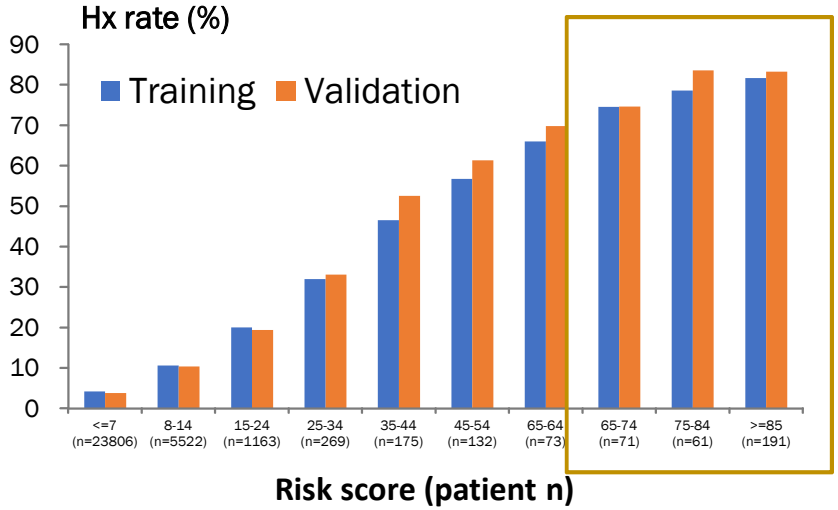


J Med Internet Res 2020; 22 (8): e16903

MODEL IMPLEMENTATION AND ASSESSMENT

- 預測能力高低之外 (AUROC)，比較風險指數 (risk categorization)與藥物使用型態
  - ➢腎病照護門診重點用藥管理 (pre-ESRD 藥師門診: Jan –Nov, 2023)
  - ➢腎臟保護藥品使用與風險指數變化

	重點用藥型態變化 (n=58)	
	第一次評估	3個月再評估
用藥≥10項人數	30 (51.72)	31 (53.45)
10類重點用藥管理		
RASI/K-sparing diuretics	52 (89.66)	47 (81.03)
SGLT2i	20 (34.48)	24 (41.38)
Metformin	12 (20.69)	10 (17.24)
Anti-thrombotic agents	26 (44.83)	26 (44.83)
NSAIDs	6 (10.34)	5 (8.62)
Anti-viral agents	0 (0.00)	2 (3.45)
CAAKI 住院風險指數		
>55	4 (6.9)	3 (5.17)

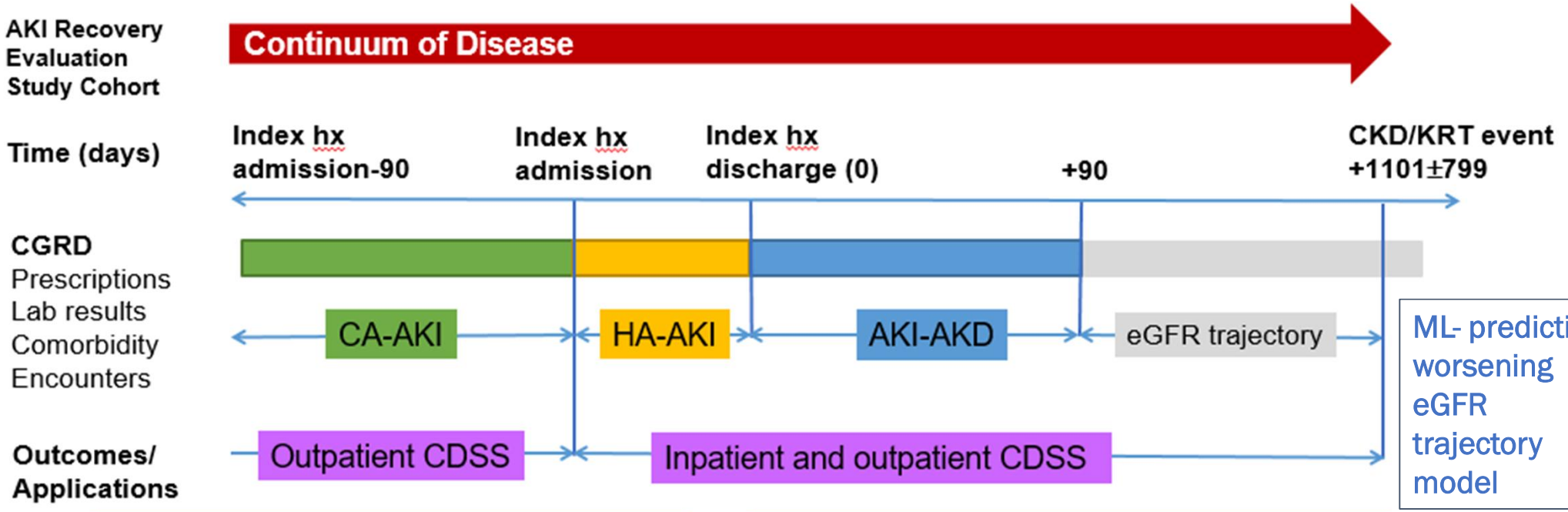


Risk score<7, predicted prob. <4%

Risk score ≥45, predicted prob. ≥61%

Risk score ≥55, predicted prob. ≥70%

ML+EHR DATA FOR SUSTAINED KIDNEY CARE

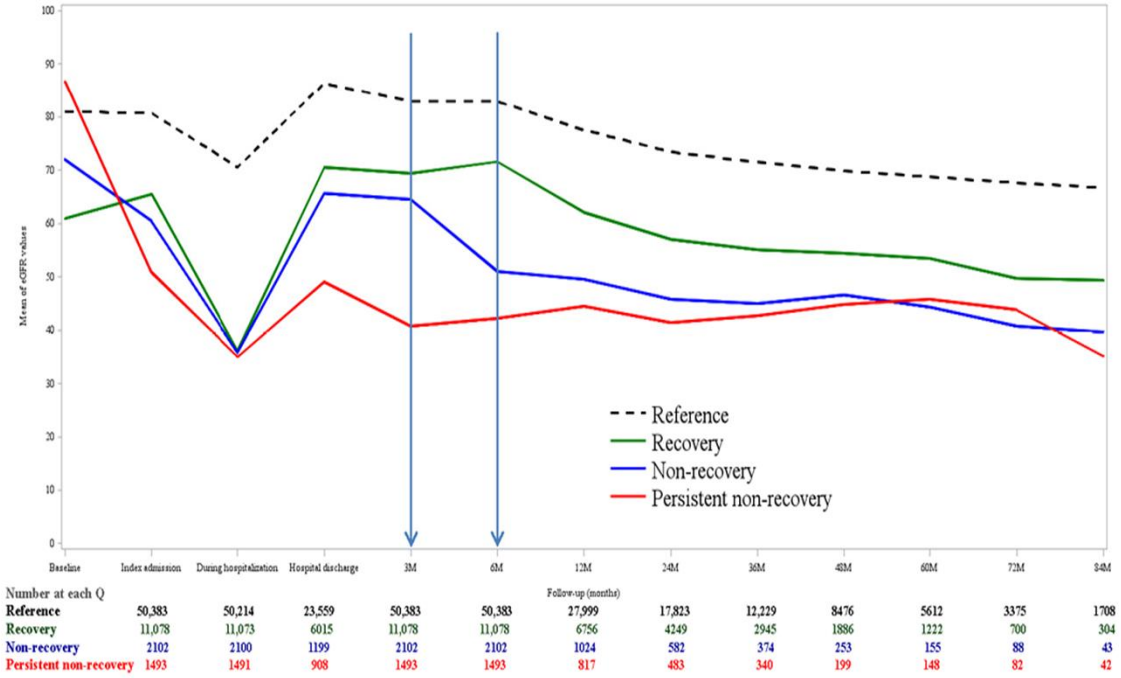


ML- predictive worsening eGFR trajectory model

- Unmet need,社區型急性腎損傷(CA-AKI)住院風險評估與預防
- 科學性: 2020 SCI 1 篇 (<5% ranking)
- 應用: 10-variable risk scoring function , Pre-ESRD outpatient setting (KCGMH)

- Post-AKI care, 辨識AKI-nonrecovery風險因子並預測高風險病情變化與CKD&KRT預防
- 科學性:長期性eGFR時間序列資料; SCI 3篇
- 應用: 20-variable risk scoring function, 出院準備與門診追蹤

IMPACTS OF AKI NON-RECOVERY ON CKD/ESKD PREVENTION



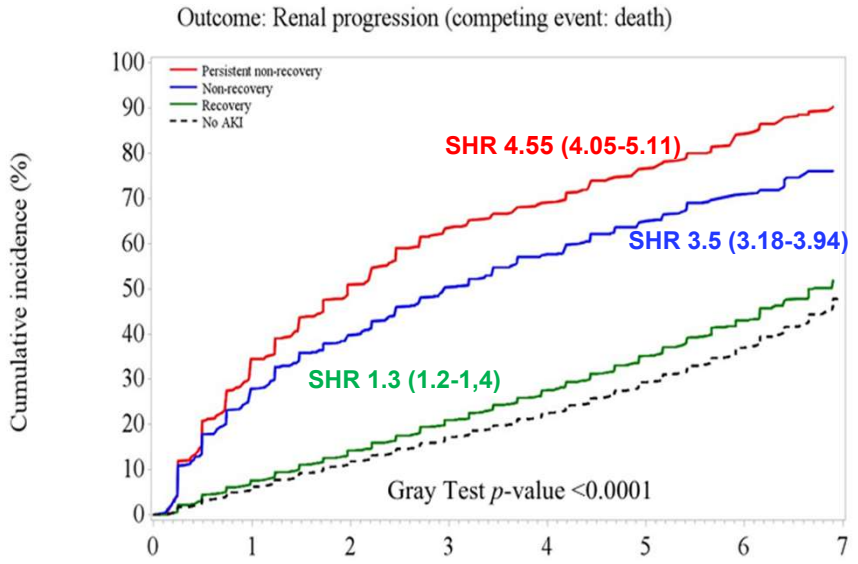


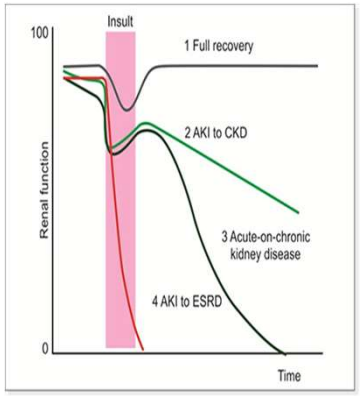
Figure 3. Cumulative risk of adverse kidney outcomes by acute kidney injury recovery patterns at 6 months after the index hospital discharge in the naïve 6-month survival cohort (n = 48,357).

J. Pers. Med. 2022, 12, 1606

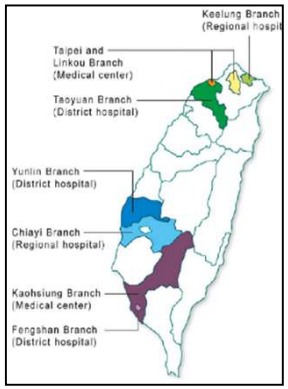
從發掘問題到研發與落地

- 2016-2023: 7 AKI SCI papers published (3 審稿中)
- 2023:國家生技醫療品質SNQ銅獎 (**藥護腎:創新智慧領航照護模式**)
- 2020:由蘇建豪藥師代表獲國家醫療品質獎佳作暨潛力獎(智慧醫療類:**AKI智能照護系統**)

2014
Drug-induced AKI ?

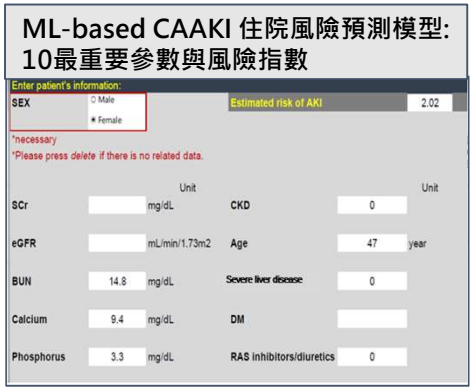


2016, 2020
1st AKI epi/ML paper



2019, 2020
雙專利

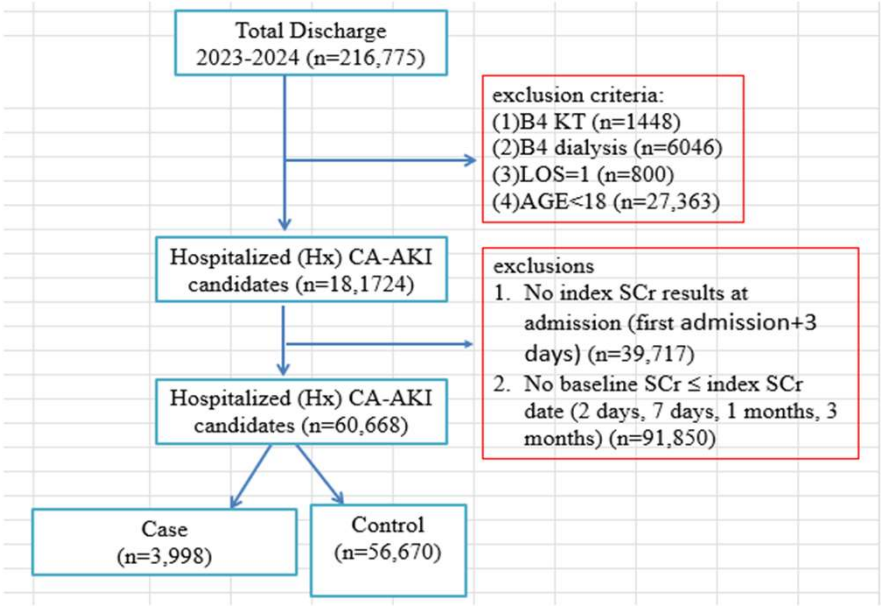
2022
pre-ESRD 藥師門診



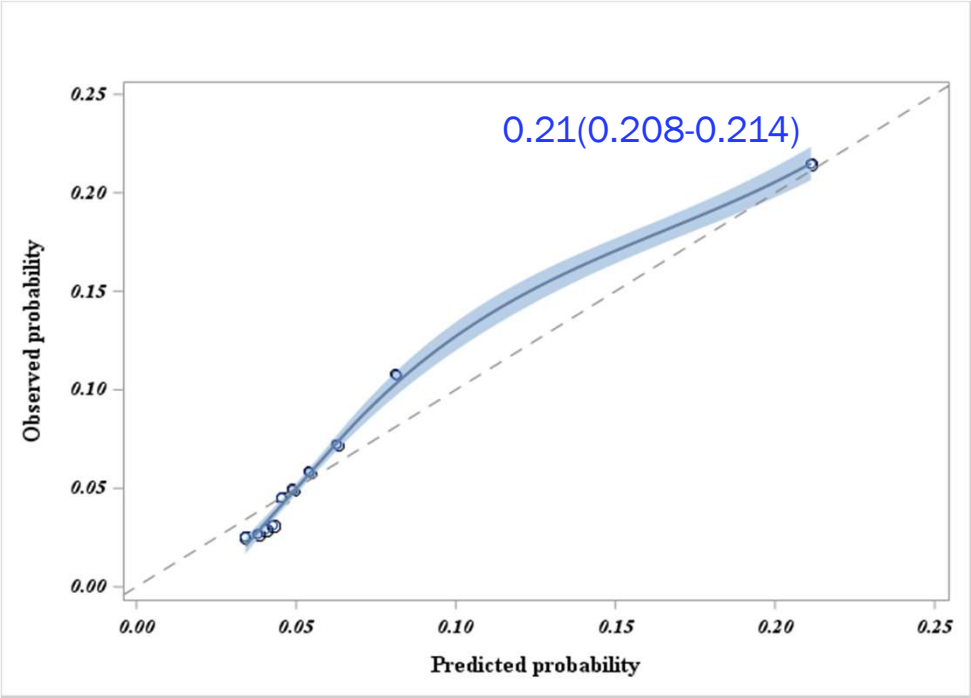
2023
SNQ銅獎(醫事組唯一)

MODEL UPDATING FOR CA-AKI HOSPITALIZATION RISK PREDICTION

Model updating in post-COVID 19 period: CGRD Jan 2023 to Aug 2024



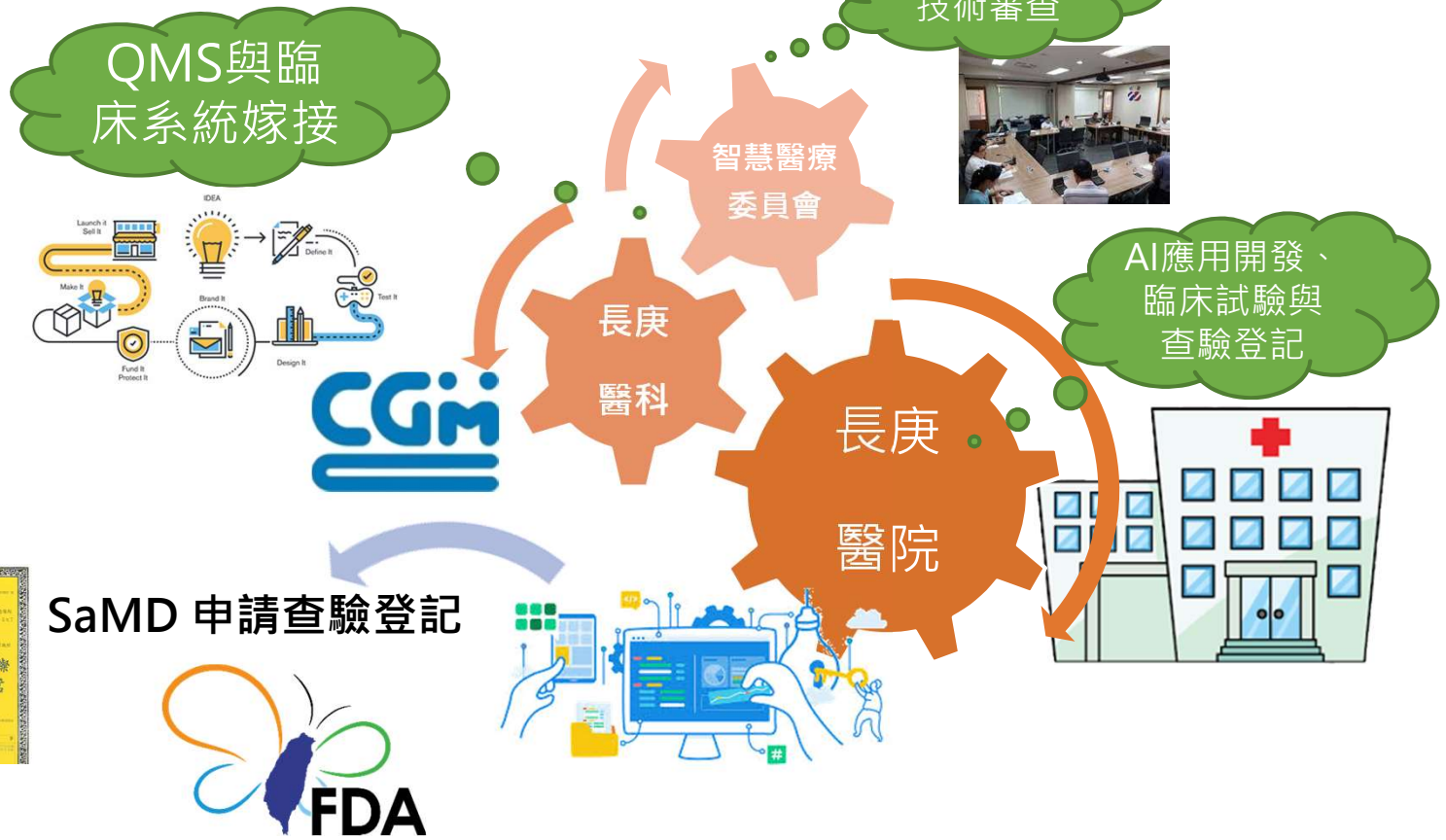
	Total (n=60,668)	CAAKI hospitalization (n=3,998)	Control (n=56,670)	
Risk score				
Mean(SD)	8.12(10.56)	14.91(16.16)	7.64(9.88)	<.0001
Median(IQR)	5.34(3.92-8.08)	8.95(5.45-16.89)	5.22(3.87-7.71)	<.0001



- CAAKI hospitalization rate: from 8.3% to 6.6%
- Calibration intercept: 0.04
- Calibration slope: 0.95
- AUROC (AKI≥ stage 2): 0.748

未來取證規劃

電腦輔助診斷類型(CADe) SaMD



關鍵合作夥伴

- 長庚醫學科技公司
- 邑泰科技公司
- 陽明交通大學
- 台灣腎臟科學會與
- 台灣臨床藥學會會員醫院與醫療人員
- 東南亞國家聯盟醫師團體

醫療器材軟體 (software as medical device)/品質管理系統 (Quality management systems, QMS)

TAKE-HOME MESSAGE

- **The first thing to think for ML applications in healthcare:** What is the problem to solve? where is (a specified clinical setting or care pathway) the model intended to use ?
- Although multiple ML/AI techniques have been achieved great prediction results, system reviews highlight the limitations of existing ML/AI prediction models.
- Learning、Applications and Research
  - NEJM AI、JAMA+AI, BMJ, etc.
- Collaborations and Sharing (**Pharmacists can help!**)

謝謝聆聽

致謝

- 高雄長庚研究計畫 (2014-2022)
- 高雄長庚藥劑部、檢驗醫學部及(兒童)腎臟專科醫師
- 陽明交大工業工程學系劉建良教授實驗室 (2018迄今)

許茜甯 PhD, RPh
Email: cnhsu10@cgmh.org.tw

高雄長庚紀念醫院
Kaohsiung Chang Gung Memorial Hospital