AI+CARE

ADVANCING DATA-DRIVEN HEALTHCARE

BRISBANE | 24-25 NOVEMBER 2025

AIDH
AUSTRALASIAN INSTITUTE OF DIGITAL HEALTH

More than accuracy: How artificial intelligence outputs are presented to clinicians, matters

Speaker **Monica Noselli**

PhD Candidate at The University of Queensland

Queensland Digital Healthcare Centre

UQ-QH-HDR Alliance Program

AI CARE ADVANCING DATA-DRIVEN HEALTHCARE

The University of Queensland Australia CREATE CHANGE

Queensland Digital Health Centre

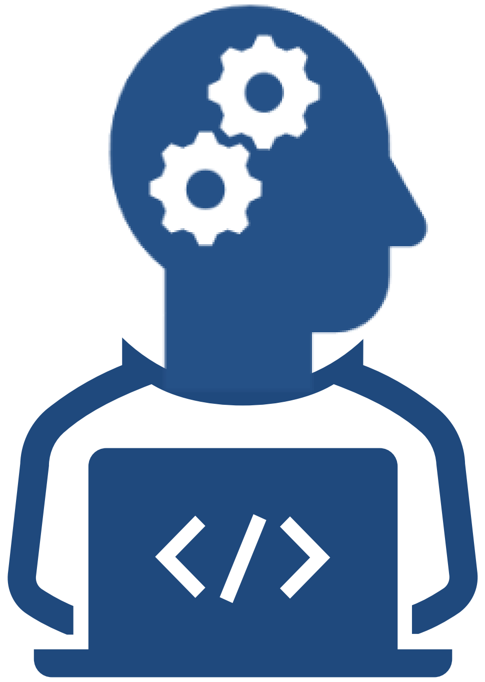
AIDH AUSTRALASIAN INSTITUTE OF DIGITAL HEALTH

AI in Healthcare: model development and evaluation

Accuracy

AI in Healthcare: Different Mindset

Computer Scientists



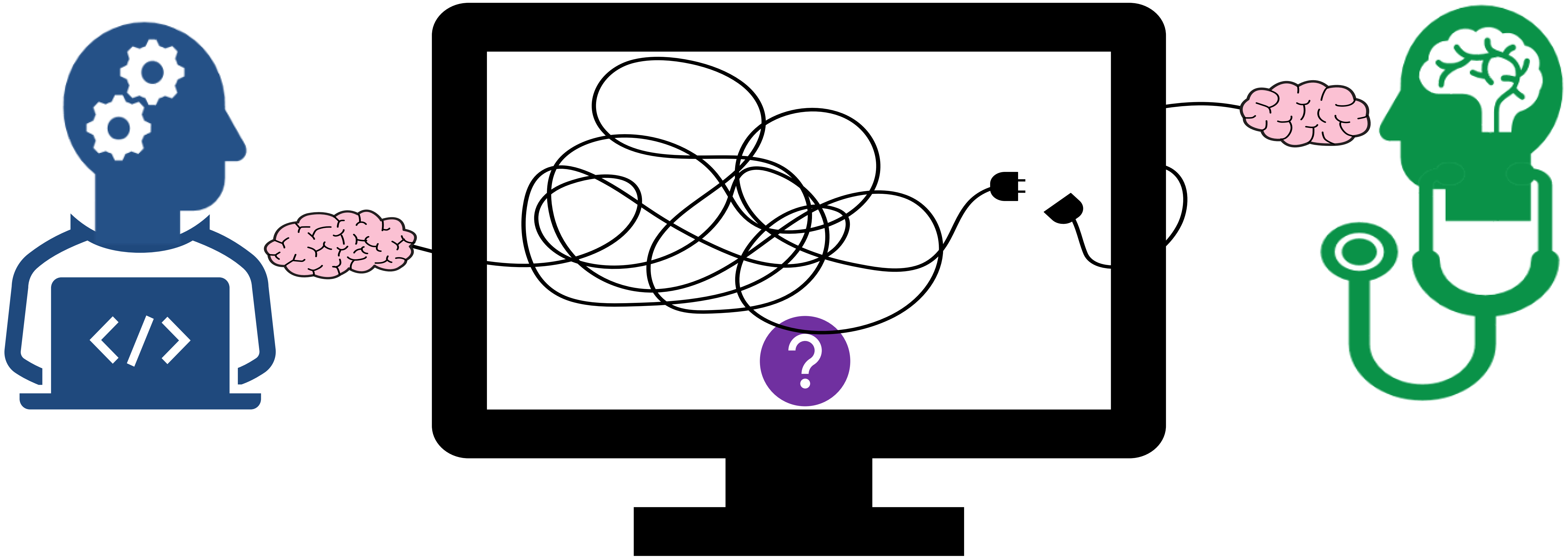
Clinicians

Data-centered
Exploration, discover new insight
Model

Patient-centered
Exploitation, Confirm old knowledge
Clinical Plausibility (clinical context)

Bienefeld, N., Boss, J.M., Lüthy, R. *et al.* **Solving the explainable AI conundrum by bridging clinicians' needs and developers' goals**. *npj Digit. Med.* **6**, 94 (2023). https://doi.org/10.1038/s41746-023-00837-4

AI in Healthcare: User interface (UI) - User Experience (UX)



to link algorithm performance and clinical impact

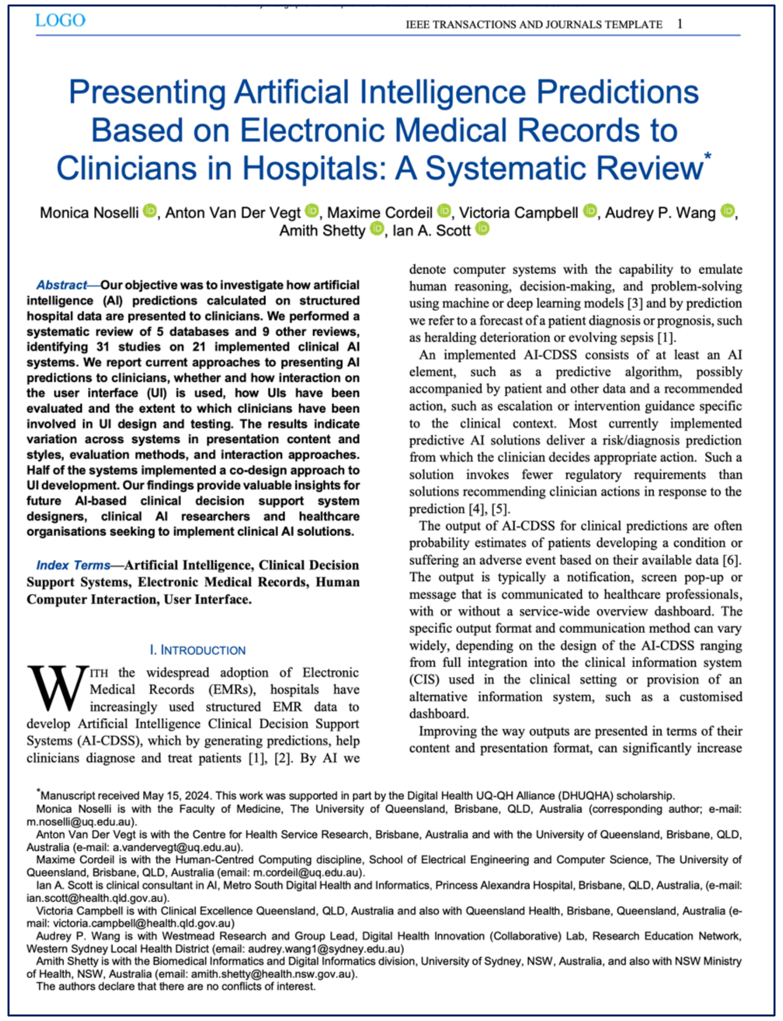
Systematic literature review

RQ-1 - How are **clinicians involved** in Artificial Intelligence Clinical Decision Support (AI-CDSS) **design**?

RQ-2 - How are AI prediction presentation methods (UI) **evaluated**?

RQ-3 - How are AI predictions **presented** to clinicians?

RQ-4 - When are **interactive**, rather than static, visualisations used and what is their purpose and value?



LOGO IEEE TRANSACTIONS AND JOURNALS TEMPLATE 1

Presenting Artificial Intelligence Predictions Based on Electronic Medical Records to Clinicians in Hospitals: A Systematic Review*

Monica Noselli, Anton Van Der Vegt, Maxime Cordeil, Victoria Campbell, Audrey P. Wang, Amith Shetty, Ian A. Scott

Abstract—**Our objective was to investigate how artificial intelligence (AI) predictions calculated on structured hospital data are presented to clinicians. We performed a systematic review of 5 databases and 9 other reviews, identifying 31 studies on 21 implemented clinical AI systems. We report current approaches to presenting AI predictions to clinicians, whether and how interaction on the user interface (UI) is used, how UIs have been evaluated and the extent to which clinicians have been involved in UI design and testing. The results indicate variation across systems in presentation content and styles, evaluation methods, and interaction approaches. Half of the systems implemented a co-design approach to UI development. Our findings provide valuable insights for future AI-based clinical decision support system designers, clinical AI researchers and healthcare organisations seeking to implement clinical AI solutions.**

Index Terms—**Artificial Intelligence, Clinical Decision Support Systems, Electronic Medical Records, Human Computer Interaction, User Interface.**

I. Introduction

With the widespread adoption of Electronic Medical Records (EMRs), hospitals have increasingly used structured EMR data to develop Artificial Intelligence Clinical Decision Support Systems (AI-CDSS), which by generating predictions, help clinicians diagnose and treat patients [1], [2]. By AI we denote computer systems with the capability to emulate human reasoning, decision-making, and problem-solving using machine or deep learning models [3] and by prediction we refer to a forecast of a patient diagnosis or prognosis, such as heralding deterioration or evolving sepsis [1].

An implemented AI-CDSS consists of at least an AI element, such as a predictive algorithm, possibly accompanied by patient and other data and a recommended action, such as escalation or intervention guidance specific to the clinical context. Most currently implemented predictive AI solutions deliver a risk/diagnosis prediction from which the clinician decides appropriate action. Such a solution invokes fewer regulatory requirements than solutions recommending clinician actions in response to the prediction [4], [5].

The output of AI-CDSS for clinical predictions are often probability estimates of patients developing a condition or suffering an adverse event based on their available data [6]. The output is typically a notification, screen pop-up or message that is communicated to healthcare professionals, with or without a service-wide overview dashboard. The specific output format and communication method can vary widely, depending on the design of the AI-CDSS ranging from full integration into the clinical information system (CIS) used in the clinical setting or provision of an alternative information system, such as a customised dashboard.

Improving the way outputs are presented in terms of their content and presentation format, can significantly increase

*Manuscript received May 15, 2024. This work was supported in part by the Digital Health UQ-QH Alliance (DHUQHA) scholarship.

Monica Noselli is with the Faculty of Medicine, The University of Queensland, Brisbane, QLD, Australia (corresponding author; e-mail: m.noselli@uq.edu.au).

Anton Van Der Vegt is with the Centre for Health Service Research, Brisbane, Australia and with the University of Queensland, Brisbane, QLD, Australia (e-mail: a.vandervegt@uq.edu.au).

Maxime Cordeil is with the Human-Centred Computing discipline, School of Electrical Engineering and Computer Science, The University of Queensland, Brisbane, QLD, Australia (email: m.cordeil@uq.edu.au).

Ian A. Scott is clinical consultant in AI, Metro South Digital Health and Informatics, Princess Alexandra Hospital, Brisbane, QLD, Australia, (e-mail: ian.scott@health.qld.gov.au).

Victoria Campbell is with Clinical Excellence Queensland, QLD, Australia and also with Queensland Health, Brisbane, Queensland, Australia (e-mail: victoria.campbell@health.qld.gov.au)

Audrey P. Wang is with Westmead Research and Group Lead, Digital Health Innovation (Collaborative) Lab, Research Education Network, Western Sydney Local Health District (email: audrey.wang1@sydney.edu.au)

Amith Shetty is with the Biomedical Informatics and Digital Informatics division, University of Sydney, NSW, Australia, and also with NSW Ministry of Health, NSW, Australia (email: amith.shetty@health.nsw.gov.au).

The authors declare that there are no conflicts of interest.

Monica Noselli, Anton Van Der Vegt, Maxime Cordeil, Victoria Campbell, Audrey P. Wang, Amith Shetty, Ian A. Scott

AI CARE ADVANCING DATA-DRIVEN HEALTHCARE · The University of Queensland Australia CREATE CHANGE · Queensland Digital Health Centre · AIDH Australasian Institute of Digital Health

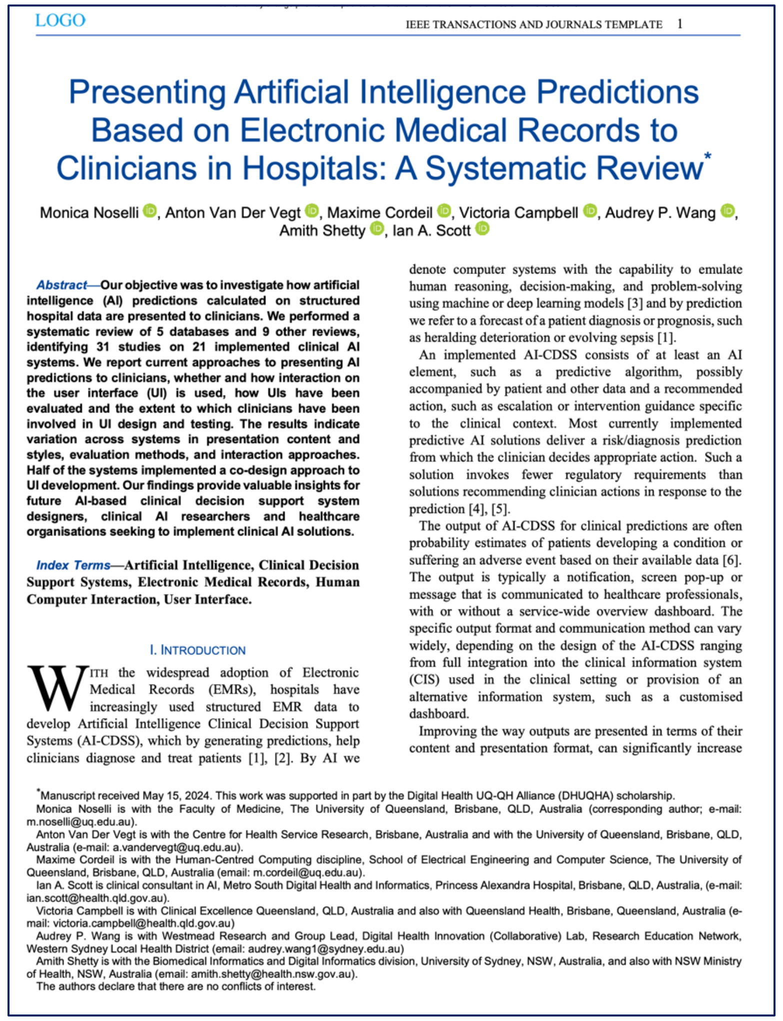
Systematic literature review

trialed or implemented in hospital settings

employing AI models and

EMR based (no imaging)

21 Systems



Presenting Artificial Intelligence Predictions Based on Electronic Medical Records to Clinicians in Hospitals: A Systematic Review*

Monica Noselli, Anton Van Der Vegt, Maxime Cordeil, Victoria Campbell, Audrey P. Wang, Amith Shetty, Ian A. Scott

Abstract—**Our objective was to investigate how artificial intelligence (AI) predictions calculated on structured hospital data are presented to clinicians. We performed a systematic review of 5 databases and 9 other reviews, identifying 31 studies on 21 implemented clinical AI systems. We report current approaches to presenting AI predictions to clinicians, whether and how interaction on the user interface (UI) is used, how UIs have been evaluated and the extent to which clinicians have been involved in UI design and testing. The results indicate variation across systems in presentation content and styles, evaluation methods, and interaction approaches. Half of the systems implemented a co-design approach to UI development. Our findings provide valuable insights for future AI-based clinical decision support system designers, clinical AI researchers and healthcare organisations seeking to implement clinical AI solutions.**

Index Terms—**Artificial Intelligence, Clinical Decision Support Systems, Electronic Medical Records, Human Computer Interaction, User Interface.**

I. Introduction

With the widespread adoption of Electronic Medical Records (EMRs), hospitals have increasingly used structured EMR data to develop Artificial Intelligence Clinical Decision Support Systems (AI-CDSS), which by generating predictions, help clinicians diagnose and treat patients [1], [2]. By AI we denote computer systems with the capability to emulate human reasoning, decision-making, and problem-solving using machine or deep learning models [3] and by prediction we refer to a forecast of a patient diagnosis or prognosis, such as heralding deterioration or evolving sepsis [1].

An implemented AI-CDSS consists of at least an AI element, such as a predictive algorithm, possibly accompanied by patient and other data and a recommended action, such as escalation or intervention guidance specific to the clinical context. Most currently implemented predictive AI solutions deliver a risk/diagnosis prediction from which the clinician decides appropriate action. Such a solution invokes fewer regulatory requirements than solutions recommending clinician actions in response to the prediction [4], [5].

The output of AI-CDSS for clinical predictions are often probability estimates of patients developing a condition or suffering an adverse event based on their available data [6]. The output is typically a notification, screen pop-up or message that is communicated to healthcare professionals, with or without a service-wide overview dashboard. The specific output format and communication method can vary widely, depending on the design of the AI-CDSS ranging from full integration into the clinical information system (CIS) used in the clinical setting or provision of an alternative information system, such as a customised dashboard.

Improving the way outputs are presented in terms of their content and presentation format, can significantly increase

*Manuscript received May 15, 2024. This work was supported in part by the Digital Health UQ-QH Alliance (DHUQHA) scholarship.
Monica Noselli is with the Faculty of Medicine, The University of Queensland, Brisbane, QLD, Australia (corresponding author; e-mail: m.noselli@uq.edu.au).
Anton Van Der Vegt is with the Centre for Health Service Research, Brisbane, Australia and with the University of Queensland, Brisbane, QLD, Australia (e-mail: a.vandervegt@uq.edu.au).
Maxime Cordeil is with the Human-Centred Computing discipline, School of Electrical Engineering and Computer Science, The University of Queensland, Brisbane, QLD, Australia (email: m.cordeil@uq.edu.au).
Ian A. Scott is clinical consultant in AI, Metro South Digital Health and Informatics, Princess Alexandra Hospital, Brisbane, QLD, Australia, (e-mail: ian.scott@health.qld.gov.au).
Victoria Campbell is with Clinical Excellence Queensland, QLD, Australia and also with Queensland Health, Brisbane, Queensland, Australia (e-mail: victoria.campbell@health.qld.gov.au)
Audrey P. Wang is with Westmead Research and Group Lead, Digital Health Innovation (Collaborative) Lab, Research Education Network, Western Sydney Local Health District (email: audrey.wang1@sydney.edu.au)
Amith Shetty is with the Biomedical Informatics and Digital Informatics division, University of Sydney, NSW, Australia, and also with NSW Ministry of Health, NSW, Australia (email: amith.shetty@health.nsw.gov.au).
The authors declare that there are no conflicts of interest.

Monica Noselli, Anton Van Der Vegt, Maxime Cordeil, Victoria Campbell, Audrey P. Wang, Amith Shetty, Ian A. Scott

RQ-1 - How are **clinicians involved** in Artificial Intelligence Clinical Decision Support (AI-CDSS) **design**?

~50% (n=11) Co-designed

RQ-2 Co-design	**System Count N**	**%**
Methodology		
Requirements/design meetings	8	38
Prototyping stages	8	38
Pilot study	8	38
Retrospective cognitive interview	1	5
Semi structured focus group interview	1	5
Survey/interview	6	29
Focus group	3	14
Sorting card	1	5
Concurrent think-aloud	2	10
Not specified	7	33

Systems	xAI	interact	**co-design**	UI evaluated
CarePre	yes	yes	yes	yes
Trews	yes	yes	yes	yes
E	yes	yes	yes	yes
G	yes	yes	yes	yes
Q	yes	yes	yes	yes
N	yes	yes	yes	yes
RetainVis	yes	yes	yes	no
F	yes	no	yes	yes
D	yes	no	yes	yes
I	yes	no	yes	no
B	yes	no	unclear	yes
C	yes	no	unclear	yes
H	yes	no	unclear	yes
L	yes	no	unclear	no
Sepsis Watch	no	yes	yes	yes
O	no	yes	unclear	no
M	no	yes	unclear	no
K	no	no	unclear	no
AUTO	no	no	unclear	no
P	no	no	unclear	no
Previse	no	no	unclear	no

RQ-1 - How are **clinicians involved** in Artificial Intelligence Clinical Decision Support (AI-CDSS) **design**?

Systems	xAI	interact	co-design	UI evaluated
CarePre	yes	yes	yes	yes
Trews	yes	yes	yes	yes
E	yes	yes	yes	yes
G	yes	yes	yes	yes
Q	yes	yes	yes	yes
N	yes	yes	yes	yes
RetainVis	yes	yes	yes	no
F	yes	no	yes	yes
D	yes	no	yes	yes
I	yes	no	yes	no
B	yes	no	unclear	yes
C	yes	no	unclear	yes
H	yes	no	unclear	yes
L	yes	no	unclear	no
Sepsis Watch	no	yes	yes	yes
O	no	yes	unclear	no
M	no	yes	unclear	no
K	no	no	unclear	no
AUTO	no	no	unclear	no
P	no	no	unclear	no
Previse	no	no	unclear	no

Codesigned systems implemented

explainability techniques (xAI)

interactions

only 3 (dark blue) High Quality User Interface Evaluation

RQ-2 - How are AI prediction presentation methods (UI) **evaluated**?

>50% (n=12) were user interface evaluated

QL n=6	Systems that were High Quality **User Interface Evaluated** (6/21)

Mixed Methods Appraisal Tool*

a critical appraisal tool for evaluating the quality of qualitative, quantitative, and mixed-methods studies

Systems	xAI	interact	co-design	UI evaluated	other evaluation
CarePre	yes	yes	yes	yes	yes
Trews	yes	yes	yes	yes	yes
E	yes	yes	yes	yes	yes
G	yes	yes	yes	yes	yes
Q	yes	yes	yes	yes	yes
N	yes	yes	yes	yes	no
RetainVis	yes	yes	yes	no	yes
F	yes	no	yes	yes	yes
D	yes	no	yes	yes	no
I	yes	no	yes	no	yes
B	yes	no	unclear	yes	no
C	yes	no	unclear	yes	no
H	yes	no	unclear	yes	no
L	yes	no	unclear	no	yes
Sepsis Watch	no	yes	yes	yes	yes
O	no	yes	unclear	no	yes
M	no	yes	unclear	no	yes
K	no	no	unclear	no	yes
AUTO	no	no	unclear	no	yes
P	no	no	unclear	no	yes
Previse	no	no	unclear	no	yes

AI+CARE ADVANCING DATA-DRIVEN HEALTHCARE · The University of Queensland Australia CREATE CHANGE · Queensland Digital Health Centre · AIDH Australasian Institute of Digital Health

* Q. N. Hong *et al.*, "The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers," *Educ. Inf.*, vol. 34, no. 4, pp. 285–291, 2018, doi: 10.3233/EFI-180221.

RQ-2 - How are AI prediction presentation methods (UI) **evaluated**?

>50% (n=12) were user interface evaluated

QL n=6	Systems that were High Quality **User Interface Evaluated** (6/21)

~75% (n=16)

outcome (e.g., mortality rate and ICU transfer)

algorithm performance evaluation (e.g., Accuracy)

Systems	xAI	interact	co-design	UI evaluated	other evaluation
CarePre	yes	yes	yes	yes	yes
Trews	yes	yes	yes	yes	yes
E	yes	yes	yes	yes	yes
G	yes	yes	yes	yes	yes
Q	yes	yes	yes	yes	yes
N	yes	yes	yes	yes	no
RetainVis	yes	yes	yes	no	yes
F	yes	no	yes	yes	yes
D	yes	no	yes	yes	no
I	yes	no	yes	no	yes
B	yes	no	unclear	yes	no
C	yes	no	unclear	yes	no
H	yes	no	unclear	yes	no
L	yes	no	unclear	no	yes
Sepsis Watch	no	yes	yes	yes	yes
O	no	yes	unclear	no	yes
M	no	yes	unclear	no	yes
K	no	no	unclear	no	yes
AUTO	no	no	unclear	no	yes
P	no	no	unclear	no	yes
Previse	no	no	unclear	no	yes

So what?

Leverage codesign wisely:

Codesign helps **align** the system with **clinical workflow** and user comprehension but may also **introduce subjective preferences** without objective validation.

Rigorous Evaluation:

Ensure thorough UI evaluations with clinicians are conducted and transparently reported before implementation.

RQ-3 - How are AI predictions **presented** to clinicians?

Structure

Centralised: predictions are presented to a **single dedicated recipient**, such as a nurse monitoring multiple patient cohorts

Distributed: predictions are sent to **multiple recipients**, such as bedside nurses caring for a set number of patients or doctors who are in care of the patient.

System Label	Care Setting	Structure
B	ICU	Distributed
C	ED	Distributed
D	unclear	Centralised
E	ED, W	unclear
F	unclear	Centralised
G	unclear	Distributed
H	W	Distributed
I	ED, W	Centralised
K	unclear	Centralised
L	W	Distributed
M	W	Distributed
N	ICU	Distributed
O	ICU, W	Centralised
	ICU	Centralised
P	ED, W	Centralised
	ED, W	unclear
Q	ICU	Distributed
	ICU	Distributed
	ED	Distributed
CarePre	unclear	unclear
Retainvis	unclear	unclear
Previse	ICU, ED	Centralised
	ED, W	Centralised
AUTO	ICU, ED, W	Distributed
	ICU	Distributed
SepsisWatch	ED	Centralised
	ED	Centralised
	ED	Centralised
Trews	ED, W	Distributed
	ED, W	Distributed
	ICU	Distributed

RQ-3 - How are AI predictions **presented** to clinicians?

Structure

Centralised: predictions are presented to a **single dedicated recipient**, such as a nurse monitoring multiple patient cohorts

Distributed: predictions are sent to **multiple recipients**, such as bedside nurses caring for a set number of patients or doctors who are in care of the patient.

System Label	Care Setting	Structure
B	ICU	Distributed
C	ED	Distributed
D	unclear	Centralised
E	ED, W	unclear
		lised
		uted
		uted
		lised
		lised
		uted
		uted
		uted
		lised
		lised
		lised
		ear
		uted
		ibuted
		istributed
		unclear
Retainvis	unclear	unclear
Previse	ICU, ED	Centralised
	ED, W	Centralised
AUTO	ICU, ED, W	Distributed
	ICU	Distributed
SepsisWatch	ED	Centralised
	ED	Centralised
	ED	Centralised
Trews	ED, W	Distributed
	ED, W	Distributed
	ICU	Distributed

Devices

Human resources
Infrastructure

Personalisation/Customisation

RQ-3 - How are AI predictions **presented** to clinicians?

Care Setting

Intensive Care Unit - Emergency Department - Ward

Different

- ❖ clinical workflow and decision timeline
- ❖ amount, type, and quality of data collected
- ❖ roles and team structures
- ❖ alert tolerance and cognitive load
- ❖ clinical priorities
- ❖ infrastructure and integration constraints

System Label	Care Setting	Structure
B	ICU	Distributed
C	ED	Distributed
D	unclear	Centralised
E	ED, W	unclear
F	unclear	Centralised
G	unclear	Distributed
H	W	Distributed
I	ED, W	Centralised
K	unclear	Centralised
L	W	Distributed
M	W	Distributed
N	ICU	Distributed
O	ICU, W	Centralised
	ICU	Centralised
P	ED, W	Centralised
	ED, W	unclear
Q	ICU	Distributed
	ICU	Distributed
	ED	Distributed
CarePre	unclear	unclear
Retainvis	unclear	unclear
Previse	ICU, ED	Centralised
	ED, W	Centralised
AUTO	ICU, ED, W	Distributed
	ICU	Distributed
SepsisWatch	ED	Centralised
	ED	Centralised
	ED	Centralised
Trews	ED, W	Distributed
	ED, W	Distributed
	ICU	Distributed

RQ-3 - How are AI predictions **presented** to clinicians?

Care Setting

Intensive Care Unit - Emergency Department - Ward

Different

- clinical workflow and decision timeline
- amount, type, and quality of data collected
- roles and team structures
- alert tolerance and cognitive load
- clinical priorities
- infrastructure and integration constraints

System Label	Care Setting	Structure
B	**ICU**	Distributed
C	**ED**	Distributed
D	unclear	Centralised
E	**ED, W**	unclear
F	unclear	Centralised
		Distributed
		Distributed
		Centralised
		Centralised
		Distributed
		Distributed
		Distributed
O		Centralised
	ICU	Centralised
P	**ED, W**	Centralised
	ED, W	unclear
Q	**ICU**	Distributed
	ICU	Distributed
	ED	Distributed
CarePre	unclear	unclear
Retainvis	unclear	unclear
Previse	**ICU, ED**	Centralised
	ED, W	Centralised
AUTO	**ICU, ED, W**	Distributed
	ICU	Distributed
SepsisWatch	**ED**	Centralised
	ED	Centralised
	ED	Centralised
Trews	**ED, W**	Distributed
	ED, W	Distributed
	ICU	Distributed

Personalisation
Customisation
Communication

So what?

Support customisation and communication:

Allow UI tailoring to clinician needs: the UI is the critical to link algorithm performance and clinical impact.

RQ-3 - How are AI predictions **presented** to clinicians?

RQ-1 Presentation	Systems Count	
	N	%
Integrated in EHR	6	29
Stand-alone web-based (e.g., dashboard)	11	52
Integrated & web-based	4	19
Unclear	1	5

RQ-3 - How are AI predictions **presented** to clinicians?

RQ-1 Presentation	Systems Count	
	N	%
Integrated in EHR	6	29
Stand-alone web-based (e.g., dashboard)	11	52
Integrated & web-based	4	19
Unclear	1	5

Fragmented information

So what?

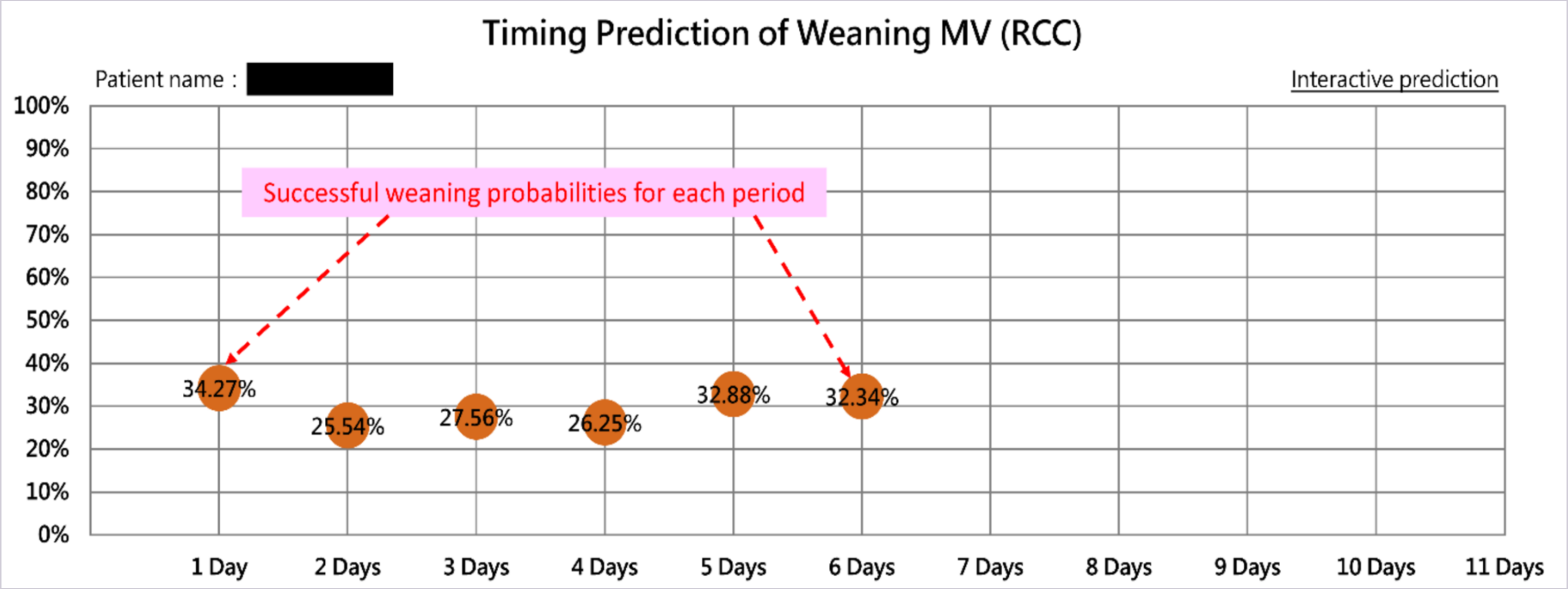
Decide integration early:

Clarify whether the system will be inside the EMR (minimal addition or advanced app) or external (e.g., dashboard), as this **affects cost, usability, and data access**.

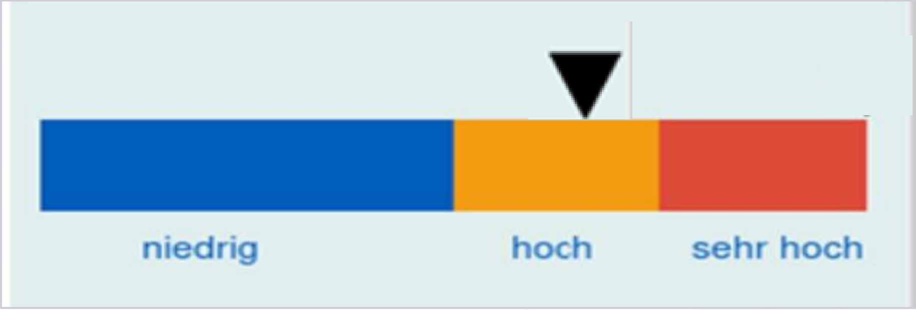
RQ-3 - How are AI predictions **presented** to clinicians?

QL n=6	Systems that were High Quality **User Interface Evaluated** (6/21)

UI element	QL n=6	Codesign n=11	Count of AI-CDSS associated with UI elements
Prediction presentation			1 2 3 4 5 6 7 8 9 10 11
Single value (updated over time/interaction)	B, C, F, H	CarePre, SepsisWatch, E, F, G, I, Retainvis, TREWS	F F
Trend	D, N	D, N, Q, G, Retainvis	D N D N
Number	B, C, D, F, N	CarePre, Q, D, E, F, G, I, N, Retainvis	D F N D F N
Category (e.g., low, medium, high) / use of colour	B, D, F, H, N	CarePre, SepsisWatch, Q, D, F TREWS, N, E, G, I, Retainvis	D F N D F N



1



2

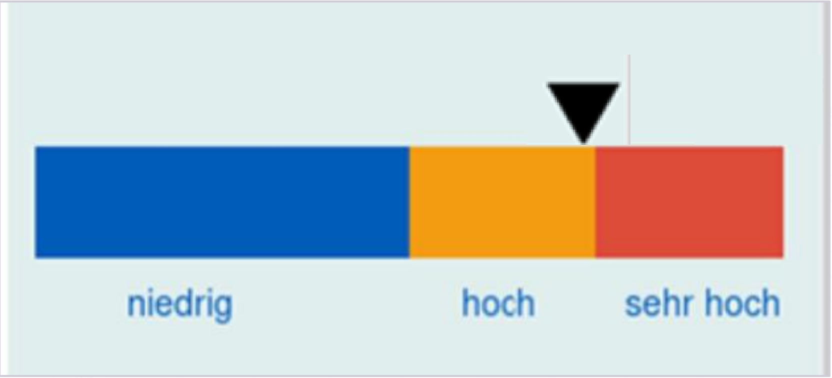
AI CARE ADVANCING DATA-DRIVEN HEALTHCARE · The University of Queensland Australia CREATE CHANGE · Queensland Digital Health Centre · AIDH Australasian Institute of Digital Health

1 - K. M. Liao *et al.*, “Development of an Interactive AI System for the Optimal Timing Prediction of Successful Weaning from Mechanical Ventilation for Patients in Respiratory Care Centers,” *Diagnostics*, vol. 12, no. 4, Apr. 2022, doi: 10.3390/diagnostics12040975
2 - S. Jauk, D. Kramer, A. Avian, A. Berghold, W. Leodolter, and S. Schulz, “Technology Acceptance of a Machine Learning Algorithm Predicting Delirium in a Clinical Setting: a Mixed-Methods Study”, doi: 10.1007/s10916-021-01727-6/Published..
3 - S. V. Kovalchuk, G. D. Kopanitsa, I. V. Derevitskii, G. A. Matveev, and D. A. Savitskaya, “Three-stage intelligent support of clinical decision making for higher trust, validity, and explainability.” *J. Biomed. Inform.*, vol. 127, Mar. 2022, doi: 10.1016/j.jbi.2022.104013..

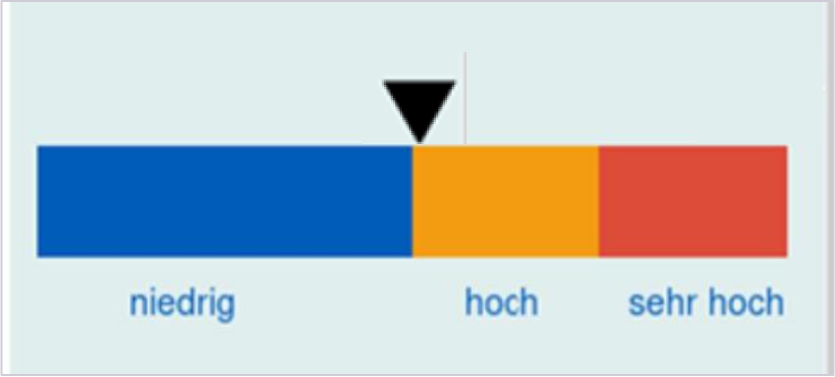
RQ-3 - How are AI predictions **presented** to clinicians?

QL n=6	Systems that were High Quality **User Interface Evaluated** (6/21)

UI element	QL n=6	Codesign n=11	Count of AI-CDSS associated with UI elements
Prediction presentation			1 2 3 4 5 6 7 8 9 10 11
Single value (updated over time/interaction)	B, C, F, H	CarePre, SepsisWatch, E, F, G, I, Retainvis, TREWS	F F
Trend	D, N	D, N, Q, G, Retainvis	D N D N
Number	B, C, D, F, N	CarePre, Q, D, E, F, G, I, N, Retainvis	D F N D F N
Category (e.g., low, medium, high) / use of colour	B, D, F, H, N	CarePre, SepsisWatch, Q, D, F TREWS, N, E, G, I, Retainvis	D F N D F N
Threshold presentation			1 2 3 4 5 6 7 8 9 10 11
Visually implemented	B, D, H, N	D, G, N	D N D N
Disclosed value	B		
Category (low, medium, high; use of colour)	H, N	G, N	N N
Content: XAI			1 2 3 4 5 6 7 8 9 10 11
Standard technique (Shapley)	B, C, D, F, H, N	TREWS, D, E, F, Q, I, N, Retainvis	D F N D F N
Custom technique (metrics, patient similarity, …)	B, D, H	CarePre, D, E, G, Q, I	D D
Standard medical score	F	F	F F



2



1 - K. M. Liao *et al.*, “Development of an Interactive AI System for the Optimal Timing Prediction of Successful Weaning from Mechanical Ventilation for Patients in Respiratory Care Centers,” *Diagnostics*, vol. 12, no. 4, Apr. 2022, doi: 10.3390/diagnostics12040975
2 - S. Jauk, D. Kramer, A. Avian, A. Berghold, W. Leodolter, and S. Schulz, “Technology Acceptance of a Machine Learning Algorithm Predicting Delirium in a Clinical Setting: a Mixed-Methods Study”, doi: 10.1007/s10916-021-01727-6/Published..
3 - S. V. Kovalchuk, G. D. Kopanitsa, I. V. Derevitskii, G. A. Matveev, and D. A. Savitskaya, “Three-stage intelligent support of clinical decision making for higher trust, validity, and explainability.” *J. Biomed. Inform.*, vol. 127, Mar. 2022, doi: 10.1016/j.jbi.2022.104013..

RQ-3 - How are AI predictions **presented** to clinicians?

Threshold is the value that determines when the AI's continuous **prediction** (usually a probability between 0 and 1) **becomes** a **positive** vs. **negative** classification

for example, when the model decides that a patient *does* or *does not* have sepsis risk.

UI element			
Prediction presentation			
Single value (updated over time/interaction)			
Trend			
Number			
Category (e.g., low, medium, high) / use of colour	F, H, N	TREWS, N, E, G, I, Retainvis	D F N
Threshold presentation			1 2 3 4 5 6 7 8 9 10 11
Visually implemented	B, D, H, N	D, G, N	D N D N
Disclosed value	B		
Category (low, medium, high; use of colour)	H, N	G, N	N N
Content: XAI			1 2 3 4 5 6 7 8 9 10 11
Standard technique (Shapley)	B, C, D, F, H, N	TREWS, D, E, F, Q, I, N, Retainvis	D F N D F N
Custom technique (metrics, patient similarity, …)	B, D, H	CarePre, D, E, G, Q, I	D D
Standard medical score	F	F	F F

1

2

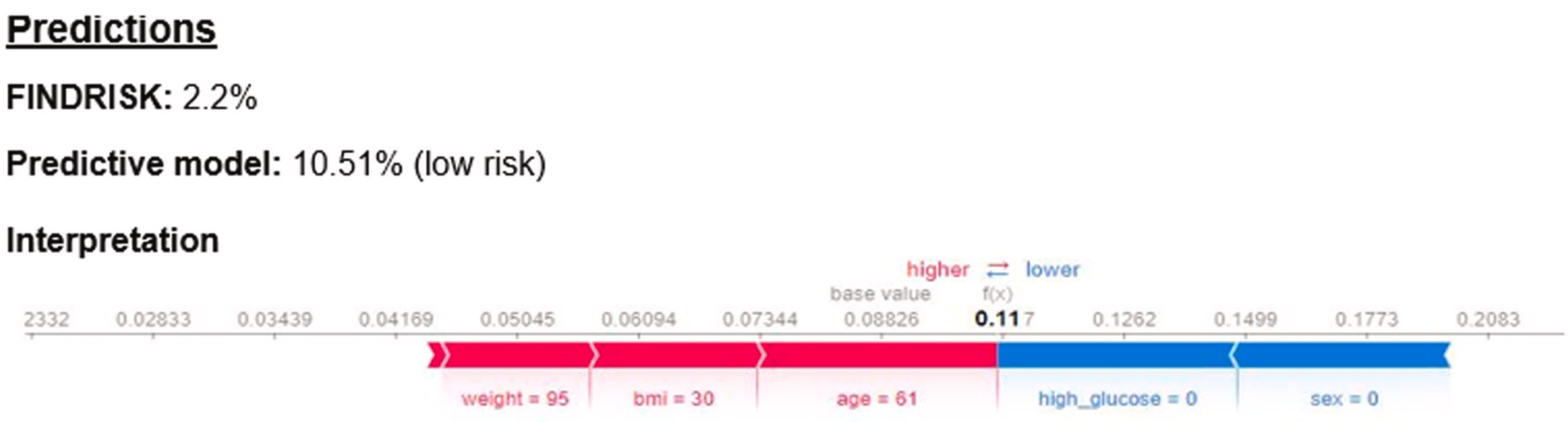
3

AI CARE ADVANCING DATA-DRIVEN HEALTHCARE · The University of Queensland Australia CREATE CHANGE · Queensland Digital Health Centre · AIDH Australasian Institute of Digital Health

1 - K. M. Liao *et al.*, "Development of an Interactive AI System for the Optimal Timing Prediction of Successful Weaning from Mechanical Ventilation for Patients in Respiratory Care Centers," *Diagnostics*, vol. 12, no. 4, Apr. 2022, doi: 10.3390/diagnostics12040975
2 - S. Jauk, D. Kramer, A. Avian, A. Berghold, W. Leodolter, and S. Schulz, "Technology Acceptance of a Machine Learning Algorithm Predicting Delirium in a Clinical Setting: a Mixed-Methods Study", doi: 10.1007/s10916-021-01727-6/Published..
3 - S. V. Kovalchuk, G. D. Kopanitsa, I. V. Derevitskii, G. A. Matveev, and D. A. Savitskaya, "Three-stage intelligent support of clinical decision making for higher trust, validity, and explainability." *J. Biomed. Inform.*, vol. 127, Mar. 2022, doi: 10.1016/j.jbi.2022.104013..

RQ-3 - How are AI predictions **presented** to clinicians?

QL n=6	Systems that were High Quality **User Interface Evaluated** (6/21)



Content: XAI			1 2 3 4 5 6 7 8 9 10 11
Standard technique (Shapley)	B, C, D, F, H, N	TREWS, D, E, F, Q, I, N, Retainvis	D F N D F N
Custom technique (metrics, patient similarity, …)	B, D, H	CarePre, D, E, G, Q, I	D D
Standard medical score	F	F	F F

3

ADVANCING DATA-DRIVEN HEALTHCARE
THE UNIVERSITY OF QUEENSLAND AUSTRALIA CREATE CHANGE
Queensland Digital Health Centre
AUSTRALASIAN INSTITUTE OF DIGITAL HEALTH

1 - K. M. Liao *et al.*, "Development of an Interactive AI System for the Optimal Timing Prediction of Successful Weaning from Mechanical Ventilation for Patients in Respiratory Care Centers," *Diagnostics*, vol. 12, no. 4, Apr. 2022, doi: 10.3390/diagnostics12040975
2 - S. Jauk, D. Kramer, A. Avian, A. Berghold, W. Leodolter, and S. Schulz, "Technology Acceptance of a Machine Learning Algorithm Predicting Delirium in a Clinical Setting: a Mixed-Methods Study", doi: 10.1007/s10916-021-01727-6/Published..
3 - S. V. Kovalchuk, G. D. Kopanitsa, I. V. Derevitskii, G. A. Matveev, and D. A. Savitskaya, "Three-stage intelligent support of clinical decision making for higher trust, validity, and explainability." *J. Biomed. Inform.*, vol. 127, Mar. 2022, doi: 10.1016/j.jbi.2022.104013..

RQ-3 - How are AI predictions **presented** to clinicians?

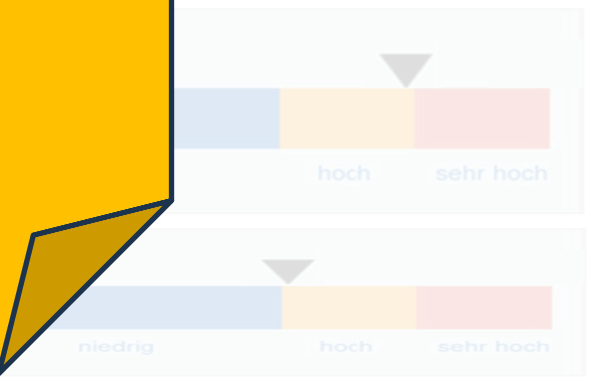
What level of explainable information (xAI) should we present?

- patient-specific details / threshold
- patient similarity
- feature-based information
- metrics of model confidence
- model performance evidence from validation studies

*To **complement** or **challenge** clinician's reasoning*

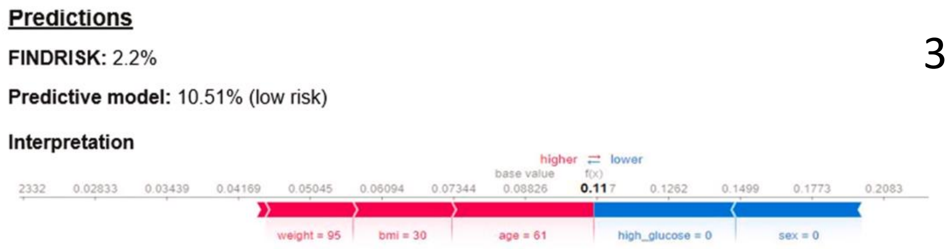
UI element
Prediction presentation
Single value (updated over time/interaction)
Trend
Number
Category (e.g., low, medium, high) / use of colour
Threshold presentation
Visually implemented
Disclosed value
Category (low, medium, high; use of colour)

1



2

Content: XAI			1 2 3 4 5 6 7 8 9 10 11
Standard technique (Shapley)	B, C, D, F, H, N	TREWS, D, E, F, Q, I, N, Retainvis	D F N D F N
Custom technique (metrics, patient similarity, …)	B, D, H	CarePre, D, E, G, Q, I	D D
Standard medical score	F	F	F F



3

1 - K. M. Liao *et al.*, "Development of an Interactive AI System for the Optimal Timing Prediction of Successful Weaning from Mechanical Ventilation for Patients in Respiratory Care Centers," *Diagnostics*, vol. 12, no. 4, Apr. 2022, doi: 10.3390/diagnostics12040975
2 - S. Jauk, D. Kramer, A. Avian, A. Berghold, W. Leodolter, and S. Schulz, "Technology Acceptance of a Machine Learning Algorithm Predicting Delirium in a Clinical Setting: a Mixed-Methods Study", doi: 10.1007/s10916-021-01727-6/Published..
3 - S. V. Kovalchuk, G. D. Kopanitsa, I. V. Derevitskii, G. A. Matveev, and D. A. Savitskaya, "Three-stage intelligent support of clinical decision making for higher trust, validity, and explainability." *J. Biomed. Inform.*, vol. 127, Mar. 2022, doi: 10.1016/j.jbi.2022.104013..

So what?

Personalisation and Drill-down information:

Allow UI tailoring to clinician needs: moving from a general overview to progressively more detailed and specific levels as they need more information.

RQ-4 - When are **interactive**, rather than static, visualisations used and what is their purpose and value?

QL n=6	Systems that were High Quality **User Interface Evaluated** (6/21)

UI element	**QL n=6**	**Codesign n=11**	**Count of AI-CDSS associated with UI elements**
			1 2 3 4 5 6 7 8 9 10 11
Explore and interpret patient data	B, D, N	CarePre, SepsisWatch, Q, D TREWS, E, G, I, N, Retainvis	D N D N
Analyse data	D, N	CarePre, SepsisWatch, Q, D TREWS, E, G, N, Retainvis	D N D N
User feedback	D	SepsisWatch, Q, D TREWS, Retainvis	D D
Accept/reject prediction		CarePre, SepsisWatch, Q, TREWS, Retainvis	

RQ-4 - When are **interactive**, rather than static, visualisations used and what is their purpose and value?

Critical thinking

UI element			Count of AI-CDSS associated with UI elements (1 2 3 4 5 6 7 8 9 10 11)
Explore and interpret patient data	B, D, N	CarePre, SepsisWatch, Q, D	D N
		TREWS, E, G, I, N, Retainvis	D N
Analyse data	D, N	CarePre, SepsisWatch, Q, D	D N
		TREWS, E, G, N, Retainvis	D N
User feedback	D	SepsisWatch, Q, D TREWS,	D
		Retainvis	D
Accept/reject prediction		CarePre, SepsisWatch, Q,	
		TREWS, Retainvis	

RQ-4 - When are **interactive**, rather than static, visualisations used and what is their purpose and value?

Critical thinking

UI element			Count of AI-CDSS associated with UI elements
			1 2 3 4 5 6 7 8 9 10 11
Explore and interpret patient data	B, D, N	CarePre, SepsisWatch, Q, D TREWS, E, G, I, N, Retainvis	D N D N
Analyse data			N N
User feedback			
Accept/reject prediction		TREWS, Retainvis	

Supporting System

RQ-4 - When are **interactive**, rather than static, visualisations used and what is their purpose and value?

UI element			Count of AI-CDSS associated with UI elements
			1 2 3 4 5 6 7 8 9 10 11
Explore and interpret patient data	B, D, N	CarePre, SepsisWatch, Q, D TREWS, E, G, I, N, Retainvis	D N D N
Analyse data			N N
User feedback			
Accept/reject prediction		TREWS, Retainvis	

Critical thinking

Supporting System

Trust

AI CARE
ADVANCING DATA-DRIVEN HEALTHCARE
The University of Queensland Australia
CREATE CHANGE
Queensland Digital Health Centre
AIDH
AUSTRALASIAN INSTITUTE OF DIGITAL HEALTH

RQ-4 - When are **interactive**, rather than static, visualisations used and what is their purpose and value?

Critical thinking

Time Constraint

UI element			Count of A
			1 2
Explore and interpret patient data	B, D, N	CarePre, SepsisWatch, Q, D TREWS, E, G, I, N, Retainvis	D N D N
Analyse data			N N
User feedback			
Accept/reject prediction		TREWS, Retainvis	

Supporting System

Trust

RQ-4 - When are **interactive**, rather than static, visualisations used and what is their purpose and value?

Critical thinking

Time Constraint

UI element			Count of A
			1 2
Explore and interpret patient data	B, D, N	CarePre, SepsisWatch, Q, D TREWS, E, G, I, N, Retainvis	D N D N
Analyse data			N N
User feedback			
Accept/reject prediction		TREWS, Retainvis	

Supporting System

Cognitive Overload

Trust

RQ-4 - When are **interactive**, rather than static, visualisations used and what is their purpose and value?

Critical thinking

Time Constraint

UI element			Count of A
			1 2
Explore and interpret patient data	B, D, N	CarePre, SepsisWatch, Q, D TREWS, E, G, I, N, Retainvis	D N D N
Analyse data			N N
User feedback			
Accept/reject prediction		TREWS, Retainvis	

Supporting System

Cognitive Overload

Trust

Deskilling

So what?

Interactivity Trade-offs:

While often added through codesign, the benefits vs. time-cost of interactive UIs remain unclear.

So what?

Implication for Designers/Developers

- **Use evidence-based design:** Build on UI findings to guide component selection and establish a starting point for the design.
- **Decide integration early:** Clarify whether the system will be inside the EMR (minimal addition or advanced app) or external (e.g., dashboard), as this affects cost, usability, and data access.
- **Leverage codesign wisely:** Codesign helps align the system with clinical workflow and user comprehension but may also introduce subjective preferences without objective validation.
- **Support customisation and communication:** Allow UI tailoring to clinician needs: the UI is the critical to link algorithm performance and clinical impact.
- **Incorporate feedback loops:** Clinician feedback during and after design promotes usability, engagement, and continuous improvement.

Implication for Researchers

- **No UI Evaluation Standards:** Unlike algorithm accuracy (TRIPOD-AI) or intervention impact (CONSORT-AI), there are no equivalent standards for AI-CDSS UI evaluation.
- **Limitations of XAI:** Heavy reliance on Shapley techniques: need to test and refine new approaches like evaluative AI.
- **Single-Task Systems Risk:** Most AI-CDSS are dedicated to single tasks, raising concerns about neglecting differential diagnoses and workflow fragmentation.
- **Interactivity Trade-offs:** While often added through codesign, the benefits vs. time-cost of interactive UIs remain unclear.
- **Unclear Codesign Effectiveness:** It is unknown which codesign methods work best for improving clinician satisfaction, efficiency, and overall system success.

Implication for Healthcare Organisations

- **Rigorous Evaluation:** Ensure thorough UI evaluations with clinicians are conducted and transparently reported before implementation.
- **Ongoing Feedback:** Establish continuous evaluation protocols that integrate clinician feedback to refine AI-CDSS over time.
- **Clinician Involvement:** Promote active clinician participation in design and implementation to create more effective tools.
- **Training & Safety:** Provide staff training on interpreting outputs and XAI to prevent automation bias and over-reliance on AI.
- **Standardisation:** Encourage more standardised UI designs (with local customisation) to support clinician usability across diverse systems.

[champion’s citation]

“I think people underestimate how complicated this (UI) is.
The complexity is in the machine learning and the algorithm.
But that's almost the tip of the iceberg.
The rest of it is really complicated.
How do you get people to change their behaviour?”

AI CARE

ADVANCING DATA-DRIVEN HEALTHCARE

BRISBANE | 24-25 NOVEMBER 2025

Thank you

AIDH
AUSTRALASIAN INSTITUTE OF DIGITAL HEALTH