What's needed to mainstream AI into clinical practice?

Ian Scott MBBS FRACP MHA Med

Director of Internal Medicine and Clinical Epidemiology,

Princess Alexandra Hospital

Professor of Medicine, University of Queensland

Adjunct Professor, Queensland University of Technology

Chair, Metro South Clinical AI Working Group

Digital Health Grand Rounds
23rd November 2021

The University Of Queensland Australia

Queensland University of Technology QUT

Metro South Health

Queensland Government

Meet the team

Stephen Canaris
BAppSci - Health Information Management
Senior Director, Office of CIO, Digital Health & Informatics – Metro South Health

Ahmad Abdel-Hafez
PhD Data Science
Principal Data Analyst, Digital Health & Informatics – Metro South Health

Oscar Bonilla
Binfo Tech
Manager, Data & Analytics Service, Office of the CIO, Digital Health & Informatics – Metro South Health

Nazanin Ghahreman-Falconer
PhD, Dip Grad Med Ethics, BPharm
Lecturer
School of Pharmacy
The University of Queensland

Michael Barras
BPharm PhD
Director of Pharmacy (PAH),
Research Conjoint (UQ)

Sven Marxen
BPharm
Acting Director of Pharmacy
Logan and Beaudesert Hospitals

Aaron Van Garderen
BPharm, GCClinPharm, GCProfComp
Pharmacist, Logan Hospital

David Cook
PhD FANCICS
Intensivist, PAH
Assoc Professor UQ

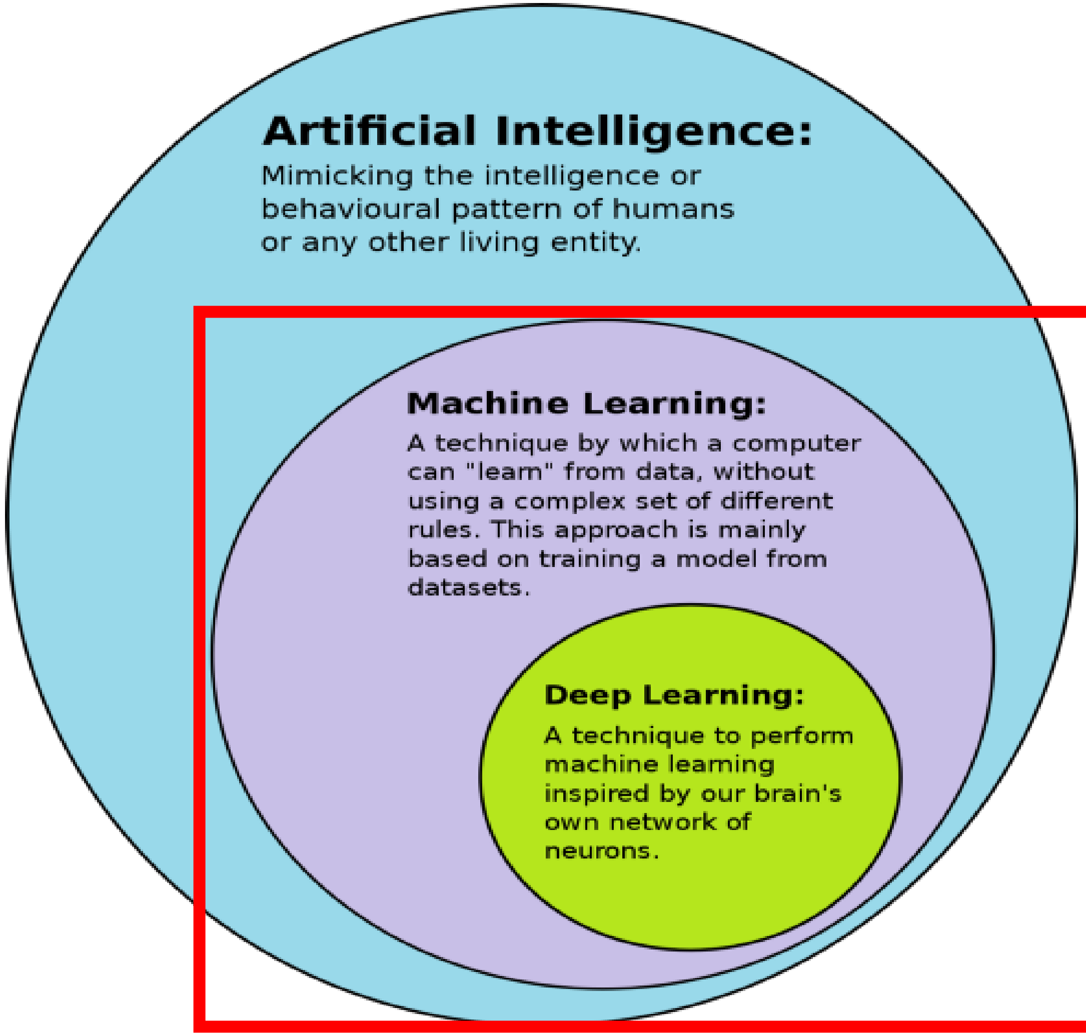
Paul Lane
MBBS FANZICS
FRACMA
Director of Patient Safety, Townsville Hospital

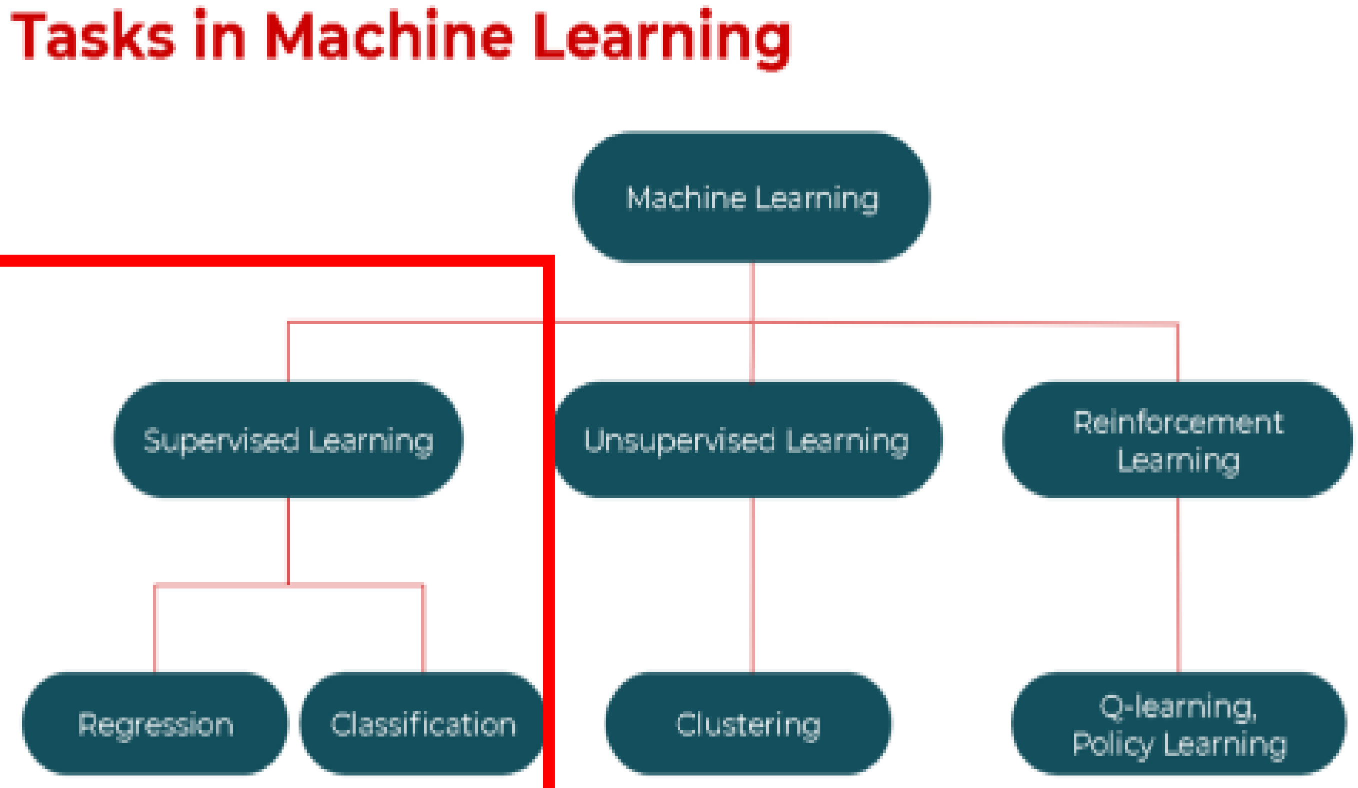
Steven McPhail
BPhys PhD
Professor in Health Services Research, Faculty of Health, School of Public Health & Social Work, QUT

Andrew Wheaton
BA BNurs
Clinical ethicist, PAH

Ian Scott
FRACP MHA MEd
Director, Internal Medicine & Clin Epi, PAH
Professor in Clinical Decision-making, UQ

The AI domains





AI/ML applications already proven in clinical trials

IDx-DR: Detects diabetic retinopathy

This software analyzes images of the eye to determine whether the patient should be referred to an eye professional because the images portray more than mild diabetic retinopathy or the patient should be rescreened in a year because the images were negative for more than mild diabetic retinopathy.[42]

OsteoDetect: Detects and diagnoses wrist fractures

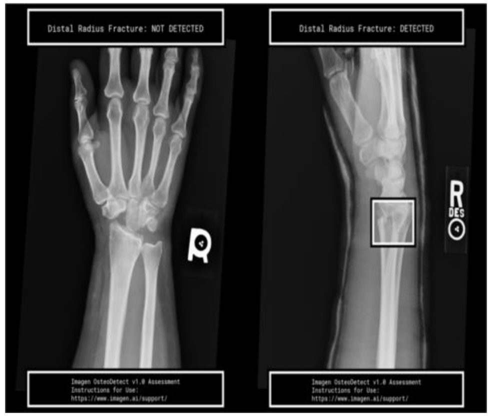
This software analyzes X-rays for signs of distal radius fracture and marks the location of the fracture to aid in detection and diagnosis.[43]

ContaCT: Detects a possible stroke and notifies a specialist

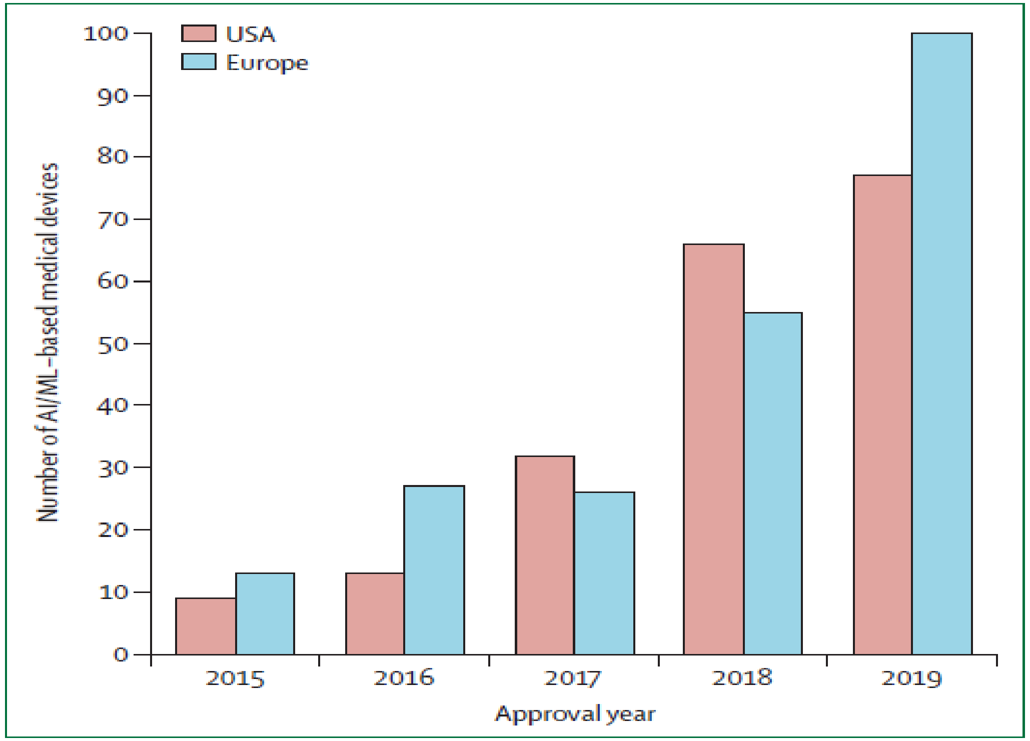
This software analyzes CT images of the brain for indicators usually associated with a stroke, and immediately texts a specialist if a suspected large vessel blockage is identified, potentially involving the specialist sooner than the usual standard of care.[44]

Guardian Connect System: Continuous glucose monitoring system

This product monitors glucose levels in the tissues of a diabetic patient, using a sensor inserted under the skin, either on an arm or on the abdomen. A transmitter processes and sends this information wirelessly to an application installed on a mobile device. Patients can use the program to monitor whether their glucose levels are too low or high.[45]



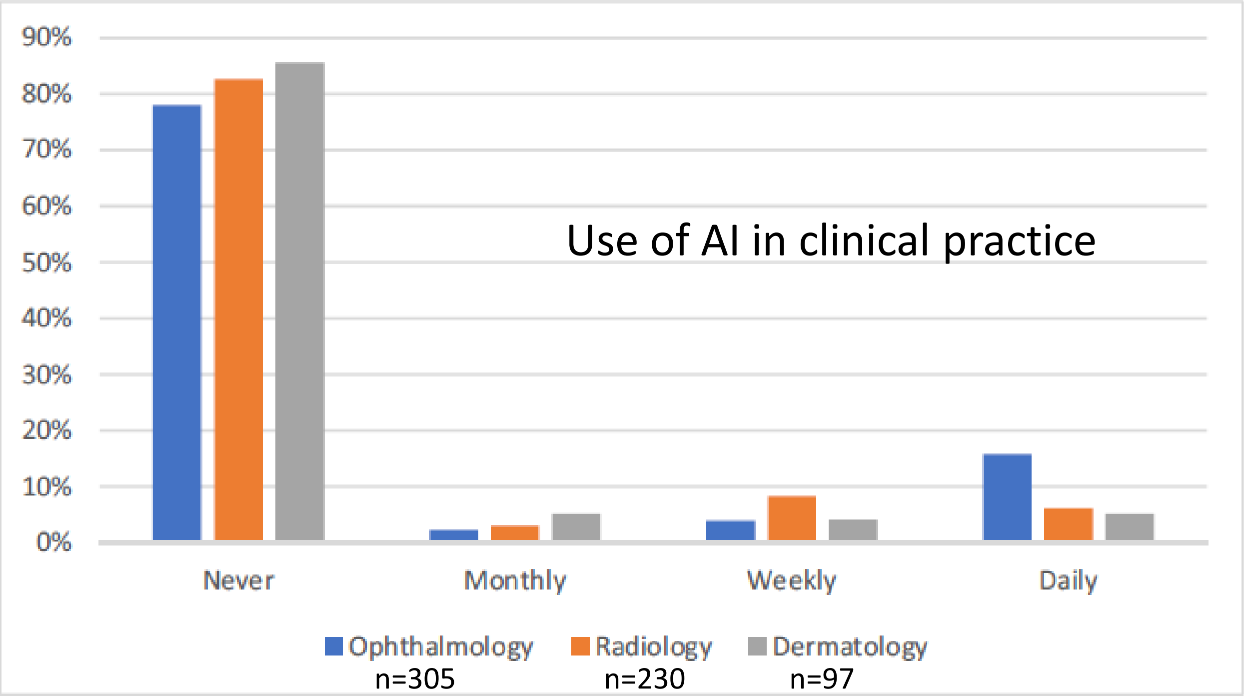
Real world gap in AI/ML adoption



Muehlematter et al. Lancet Digital Health 2021

TGA approvals

- Automated retinopathy analysis systems
- Cardiac electrophysiology software
- Image segmentation and lesion identification systems



Scheetz et al Sci Rep 2021

Potential explanations

- Limited evidence base
- Limited literacy
- Divergent stakeholder perceptions and expectations
- Fragmented AI/ML ecosystems
- Lagging organisational readiness and capacity
- Constraints in the AI/ML application pipeline
- Under-developed regulatory guidance

Limited evidence base

JOURNAL OF MEDICAL INTERNET RESEARCH — Yin et al

Review

Role of Artificial Intelligence Applications in Real-Life Clinical Practice: Systematic Review

Jiamin Yin[1*], BA; Kee Yuan Ngiam[2*], MBBS; Hock Hai Teo[1*], PhD

‘Currently insufficient level 1 evidence to advocate the routine use of healthcare AI for decision support’

‘AI applications should provide solid scientific evidence of its effectiveness relative to [current] standard of care’

‘AI applications [that are not coupled with] effective patient-specific interventions….may be ineffective in improving patient outcomes’

‘We urge the healthcare AI research community to work closely with healthcare providers and institutions to demonstrate the potential for AI in real-life clinical settings’

51 studies 2010-2020
- 20 observational; 13 RCTs; 14 experimental
- Most studies moderate to high risk of bias
- No comparison group (17); small samples (14); limited information (29)

Decision support in four categories of tasks
- Disease screening/triage (n=16)
- Disease diagnosis (n=16)
- Risk analysis (n=14)
- Treatment (n=7)

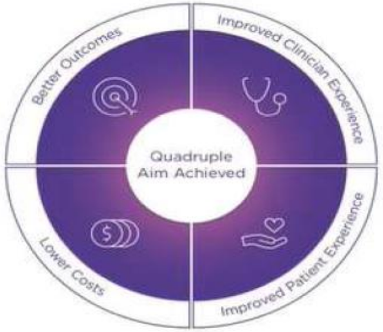
Targeted diseases/conditions
- Sepsis (n=6)
- Breast cancer (n=5)
- Diabetic retinopathy (n=4)
- Colonic polyp/adenoma (n=4)
- Cataracts (n=2)
- Stroke (n=2)

Application performance (n=26)
- 24/26 reported acceptably satisfactory performance

Clinician benefits (n=25)
- 16/18: enhanced decision-making
- 6/7: improved workflow efficiencies

Patient benefits (n=14)
- 8/11: improved clinical outcomes
- 3/3: better patient experience

Cost-effectiveness (n=1)
- 1/1: reduced costs



Strategy: Define and apply appropriate research designs for AI/ML

- Sociotechnical innovations
  - Mixed methods analyses are required
    - qualitative as well as quantitative inquiry into model performance and use
    - informed by theories of technology acceptance
- Rapidly evolving technologies
  - Traditional linear RCTs using fixed study protocols may not work
    - Adaptive study designs with built-in PDSA
    - Hybrid effectiveness-implementation trials
    - Stepped wedge roll-out
- Task and context-sensitive
  - Single phase III intervention trials aim to prove reproducible ‘average’ effect in large heterogenous populations across multiple settings --- generalisability
    - AI/ML applications are designed for specific tasks within specific contexts
    - Multiple local prospective replication trials are necessary for external validation
- Choice of comparator
  - Human expert performance within current system of care delivery

Strategy: Encourage methodological rigour in AI/ML research

JAMA | Users' Guides to the Medical Literature

How to Read Articles That Use Machine Learning
Users' Guides to the Medical Literature

Yun Liu, PhD; Po-Hsuan Cameron Chen, PhD; Jonathan Krause, PhD; Lily Peng, MD, PhD

CONSENSUS STATEMENT
https://doi.org/10.1038/s41591-020-1034-x
nature medicine
Check for updates
OPEN

Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension

Xiaoxuan Liu[1,2,3,4,5], Samantha Cruz Rivera[5,6,7], David Moher[8,9], Melanie J. Calvert[4,5,6,7,10,11,12], Alastair K. Denniston[2,3,4,5,6,13] ✉ and The SPIRIT-AI and CONSORT-AI Working Group*

BMJ Open

Protocol for development of a reporting guideline (TRIPOD-AI) and risk of bias tool (PROBAST-AI) for diagnostic and prognostic prediction model studies based on artificial intelligence

Developing specific reporting guidelines for diagnostic accuracy studies assessing AI interventions: The STARD-AI Steering Group

nature medicine
CONSENSUS STATEMENT
https://doi.org/10.1038/s41591-020-1037-7
Check for updates
OPEN

Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension

Samantha Cruz Rivera[1,2,3], Xiaoxuan Liu[3,4,5,6,7], An-Wen Chan[8], Alastair K. Denniston[1,3,4,5,6,9,10], Melanie J. Calvert[1,2,3,4,6,10,11,12], The SPIRIT-AI and CONSORT-AI Working Group*, SPIRIT-AI and CONSORT-AI Steering Group and SPIRIT-AI and CONSORT-AI Consensus Group

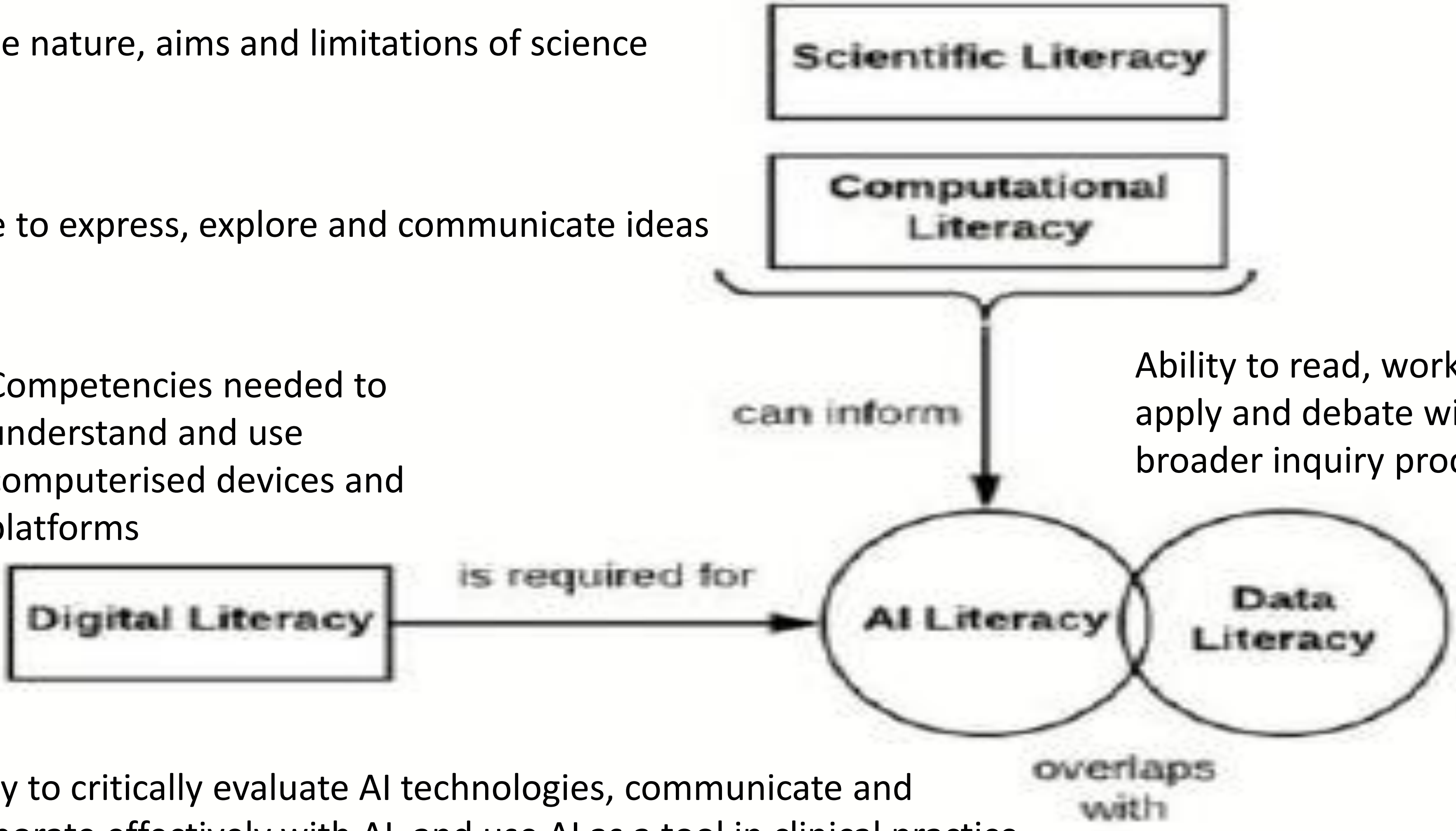
Limited literacy

Appreciation of the nature, aims and limitations of science

Ability to use code to express, explore and communicate ideas

Competencies needed to understand and use computerised devices and platforms

Ability to read, work with, analyse, apply and debate with data as a broader inquiry process



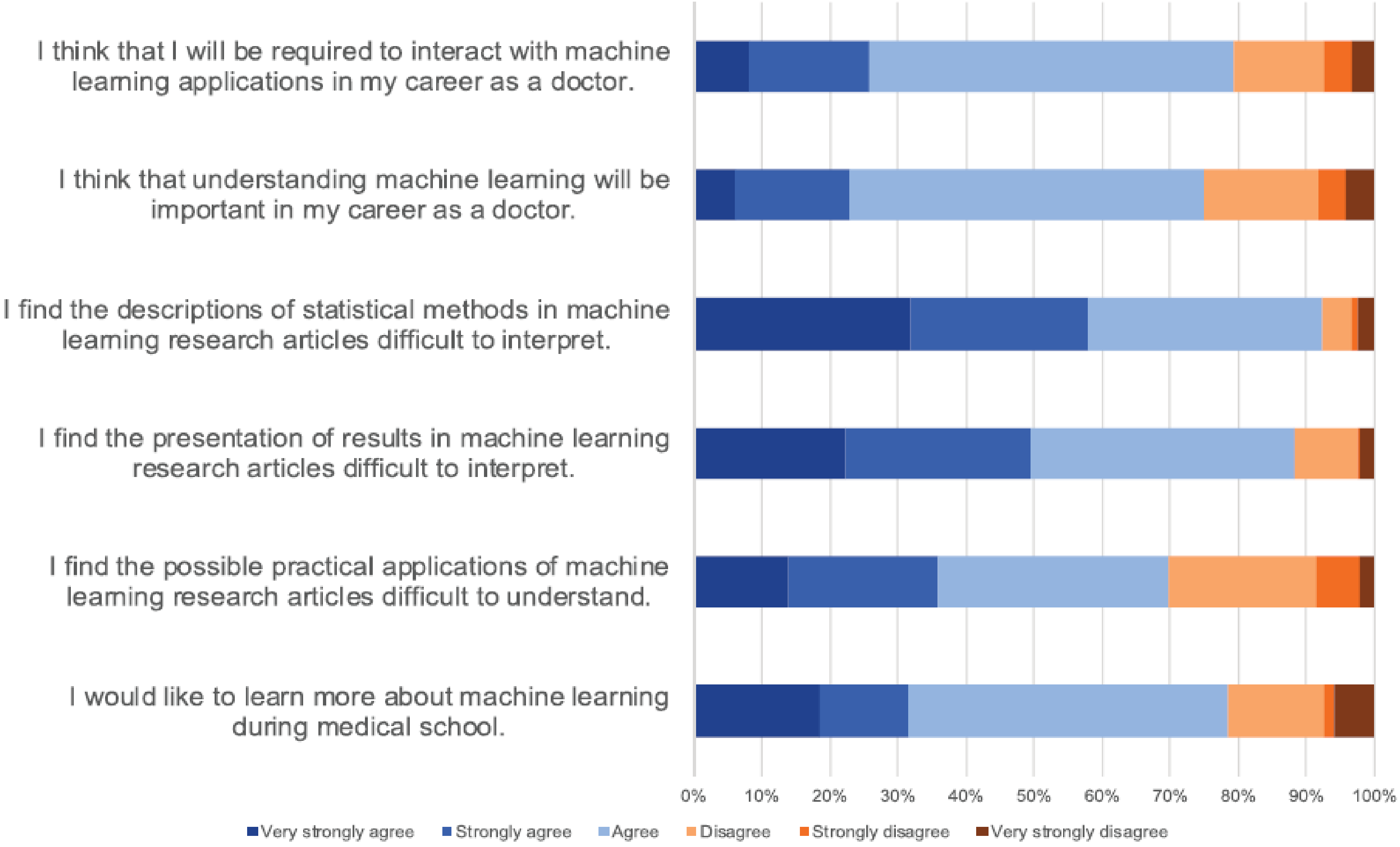
Ability to critically evaluate AI technologies, communicate and collaborate effectively with AI, and use AI as a tool in clinical practice

Adapted from Long & Magerko CHI 2020

Limited literacy

Consistently stated impressions at all levels of training

- AI/ML will change clinical practice
- Faculty of universities and professional colleges do not have expertise and skills in AI
- Need for interdisciplinary training



245 final year medical students (Aust, NZ, US)

Blacketer et al Intern Med J 2021

Strategy: Provide structured education and CPD

- Certified curricula, on-line courses, webinars, conferences, datathons, primers

Annals of Internal Medicine — IDEAS AND OPINIONS

Machine Learning and Evidence-Based Medicine

Ian A. Scott, MBBS, MHA, MEd

Perspectives

Ian A Scott[1,2], David Cook[1], Enrico W Coiera[3], Brent Richards[4]

Machine learning in clinical practice: prospects and pitfalls

EBM analysis: General medicine

Evidence-based medicine and machine learning: a partnership with a common purpose

Ian Scott ,[1,2] David Cook,[3] Enrico Coiera[4]

INTERNAL MEDICINE JOURNAL — RACP Specialists. Together

doi:10.1111/imj.1520

REVIEW

Demystifying machine learning: a primer for physicians

Ian A. Scott [1,2]

[1]Internal Medicine and Clinical Epidemiology, Princess Alexandra Hospital, and [2]School of Clinical Medicine, University of Queensland, Brisbane, Queensland, Australia

Open access — Communication

BMJ Health & Care Informatics

Clinician checklist for assessing suitability of machine learning applications in healthcare

Ian Scott,[1,2] Stacy Carter,[3] Enrico Coiera[4]

MIT MANAGEMENT EXECUTIVE EDUCATION

AI in Health Care
Long course (3 months)
Short course (6 weeks)

HARVARD MEDICAL SCHOOL

Designing and Implementing AI Solutions for Health Care
Short course (2 weeks)

Academy of Medical Royal Colleges

January / 2019

Artificial Intelligence in Healthcare

QUT — The University of Queensland Australia

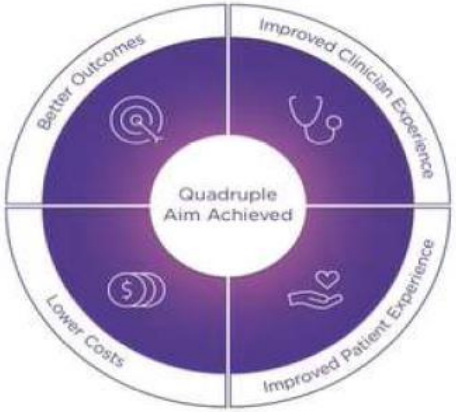
KPMG — intelliHQ

Australian Medical Council Limited

Australian Digital Health Agency and the Australian Medical Council:

Capability Framework in Digital Health in Medicine (DRAFT V10)

Strategy: Educate clinicians to ask key questions of AI/ML applications



- PURPOSE: What will this application do for me and my patient?
- TRUSTWORTHINESS: Can I trust this application?
- INTELLIGIBILITY: Do the predictions of this application make sense to me?
- RESPECT and AUTONOMY: How much respect and choice does this application give me/my patient?
- INTEGRATION: Does this application fit easily into my workflow and operating environment?
- SUSTAINABILITY: What is the life cycle of this application?

Divergent stakeholder perceptions and expectations

What do clinicians think of AI?

Positive perceptions	Negative perceptions
Improved diagnostic accuracy; fewer errors (n=2)	Liability for AI-mediated errors (n=1)
More efficient work flows (n=4)	Insufficient training and continuing professional development in AI (n=4)
Less time spent on administrative and other mundane tasks (n=2)	Reputational loss and reduced demand for specialist opinion ('DIY medicine') (n=2)
Synthesis of clinical information (n=2)	Potential erosion of empathetic communication with patients (n=2)
Updating of clinical records (n=1)	Risk of privacy breaches and loss of confidentiality of patient information (n=1)
More time spent with patients (n=1)	Lack of proof of efficacy of AI applications in clinical settings (n=2)
Improved access to care (n=1)	Limited explainability (n=1)

Scott et al BMJ Health Care Inform 2021 (in press)

Divergent stakeholder perceptions and expectations

What do patients/consumers think of AI?

Positive perceptions	Negative perceptions
Second opinions to clinicians yielding better decisions (n=3) **Improved access to care (n=1)**	Dehumanisation of the clinician-patient relationship (n=2) Threat to shared decision-making involving patients (n=1) Low trustworthiness of AI advice (n=3) Insufficient clinician and regulatory oversight (n=1) Uncertainty around fairness and equity in treatment allocation (n=1)

Scott et al BMJ Health Care Inform 2021 (in press)

Divergent stakeholder perceptions and expectations

What do healthcare executives think of AI?

Positive perceptions	Negative perceptions
Improved operational efficiency, cybersecurity, analytic capacity, cost savings (n=1)	Uncertainty around patient satisfaction, access to care, improved patient outcomes (n=1)

What do industry professionals think of AI?

Positive perceptions	Negative perceptions
Most of the positive perceptions already listed (n=3)	Limited access to high quality data for model development (n=1) Unresolved legal liability question (n=1) Lack of explicit and robust regulatory frameworks (n=1) Low levels of funding for independent, investigator-led research in AI (n=1)

Scott et al BMJ Health Care Inform 2021 (in press)

Strategy: AI/ML charter that reconciles stakeholder expectations

Expectation	Dependency
Ensure accuracy, freedom from bias, trustworthiness	Model development and testing must involve domain experts, use high quality data sets, minimise bias, and demonstrate accurate results in the populations for which they are to be used
Improve efficiency and reduced administrative burden	Applications must be fitted to, and complement, routine clinical workflows and, where possible, self-populate the required data with minimal clinician input.
Improve clinical decision-making and outcomes	Applications must be as or more effective in improving clinical decision-making and patient outcomes than current care, and be accompanied with clinician oversight.
Augment clinician-patient interaction	Applications should not distract from, or degrade, human to human interaction and shared decision-making.
Ensure explainability and transparency	Applications should aim to provide explainability and transparency in regards to their inner workings, while acknowledging limits to the extent this can be achieved.
Preserve and enhance professional roles	Applications must be sensitive to potential loss of jobs or professional reputation, remove tedium, improve job satisfaction, provide new skills, meet training needs.
Obtain regulatory approval	Applications should be subject to regulatory standards that are robust, transparent and responsive to updates of existing applications.
Determine liability for error	Applications should be associated with clear lines of responsibility regarding liability for error, including no-fault provisions.
Ensure privacy, confidentiality and security	Application developers must ensure they adhere to legal and community expectations regarding data privacy, confidentiality and security for health and medical data.

Scott et al BMJ Health Care Inform 2021 (in press)

Fragmented AI/ML ecosystems

Foundations for AI/ML in healthcare

- Workforce capacity and expertise
  - Clinicians, data scientists, computer scientists, clinical informaticians, statisticians, methodologists, engineers, ethicists, economists, psychologists, sociologists, researchers, vendors
    - AI/ML is a sociotechnical ecosystem
- Data sources, data storage, data analytics
- Computational capacity
- Reliable long term funding
- Established ethical, legislative, regulatory policy frameworks
- Patient and public involvement

Strategy: Establish statewide clinical AI/ML R & D collaborative

- Metro South Clinical AI Working Group
- QH Sepsis Clinical AI Working Group
- QH Deterioration Analytics Risk Tool Working Group

- Digital Health CRC
  - Prof Tim Shaw (University of Sydney)
  - Prof Steven McPhail (QUT)
- CSIRO e-Health Research Centre/Data61
  - Prof David Hansen (Queensland)
- Queensland AI Hub

- Other HHS-specific initiatives
  - Various consultancies
- Other academic discipline-specific groups

Queensland AI4H Collaborative

- Australian Institute of Machine Learning (University of Adelaide)
- Australian Artificial Intelligence Institute (University of Technology Sydney)
- NHMRC CRE in Digital Health (Macquarie University)

- QH/UQ Digital Health Research Group
  - UQ SMART project
- eHealth Qld Clinical & Business Intelligence

Co-design
Co-produce
Co-evaluate

with Clinicians
and Consumers

Australian Alliance for AI in Healthcare
Prof Enrico Coiera
(Macquarie University)

AAAiH
Australian Alliance for Artificial Intelligence in Healthcare

Strategy: Establish multi-level governance processes

Data governance

- **Data access**: how, from what source and from whom will data be obtained?
- **Data protection:** how will patient consent and privacy be guaranteed?
- **Data transparency**: what are the processes for sharing data?

Technical infrastructure governance

- **System reliability:** how stable are the software platforms
- **System interoperability**: how seamless and efficient is the flow of data across interfaces?
- **System responsiveness and adaptability**: is the system capable of keeping up with rapidly evolving technology?

The Power of One

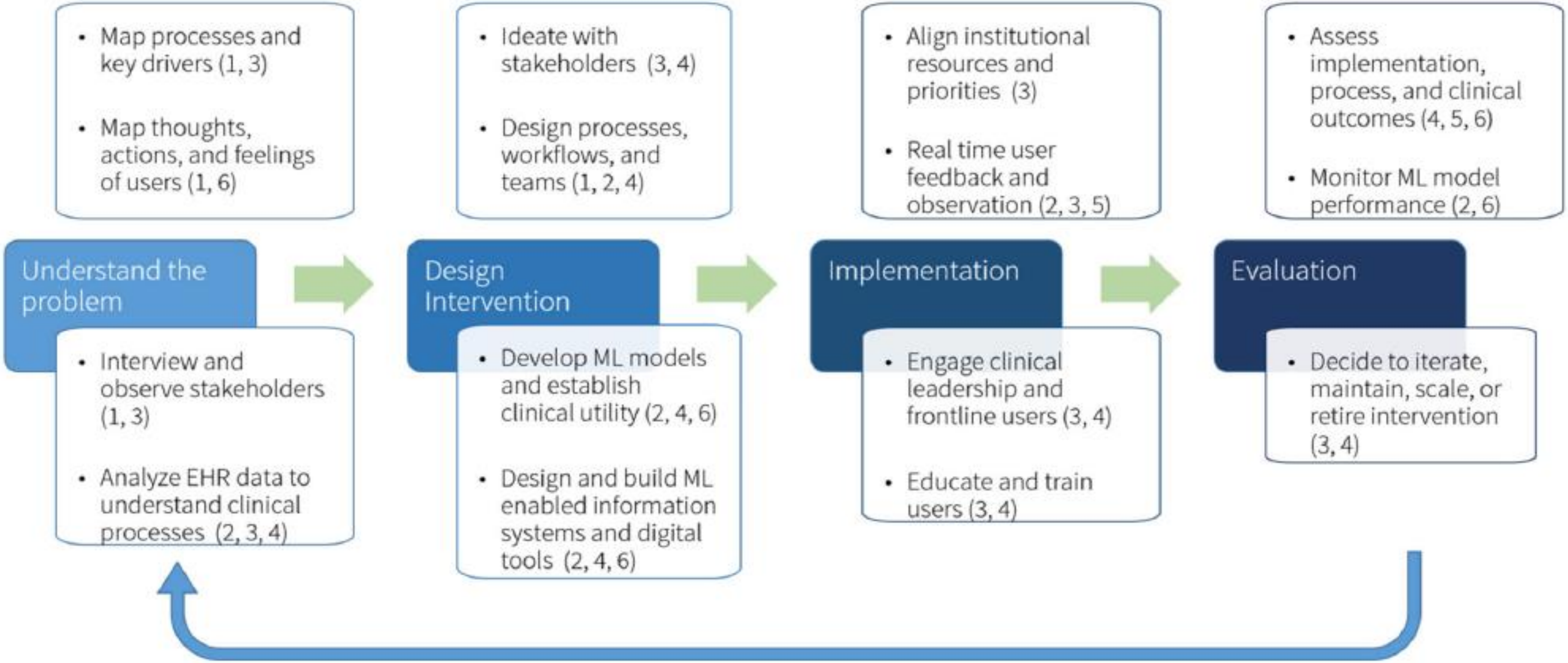
Clinical governance

- **Value proposition**: how will AI/ML projects be selected for investment?
- **Problem definition**: what is the AI application supposed to do
- **Accountability**: how and by whom is optimal performance to be decided and assessed
- **Clinical safety**: what risk mitigation will be employed and who is responsible?

Research governance

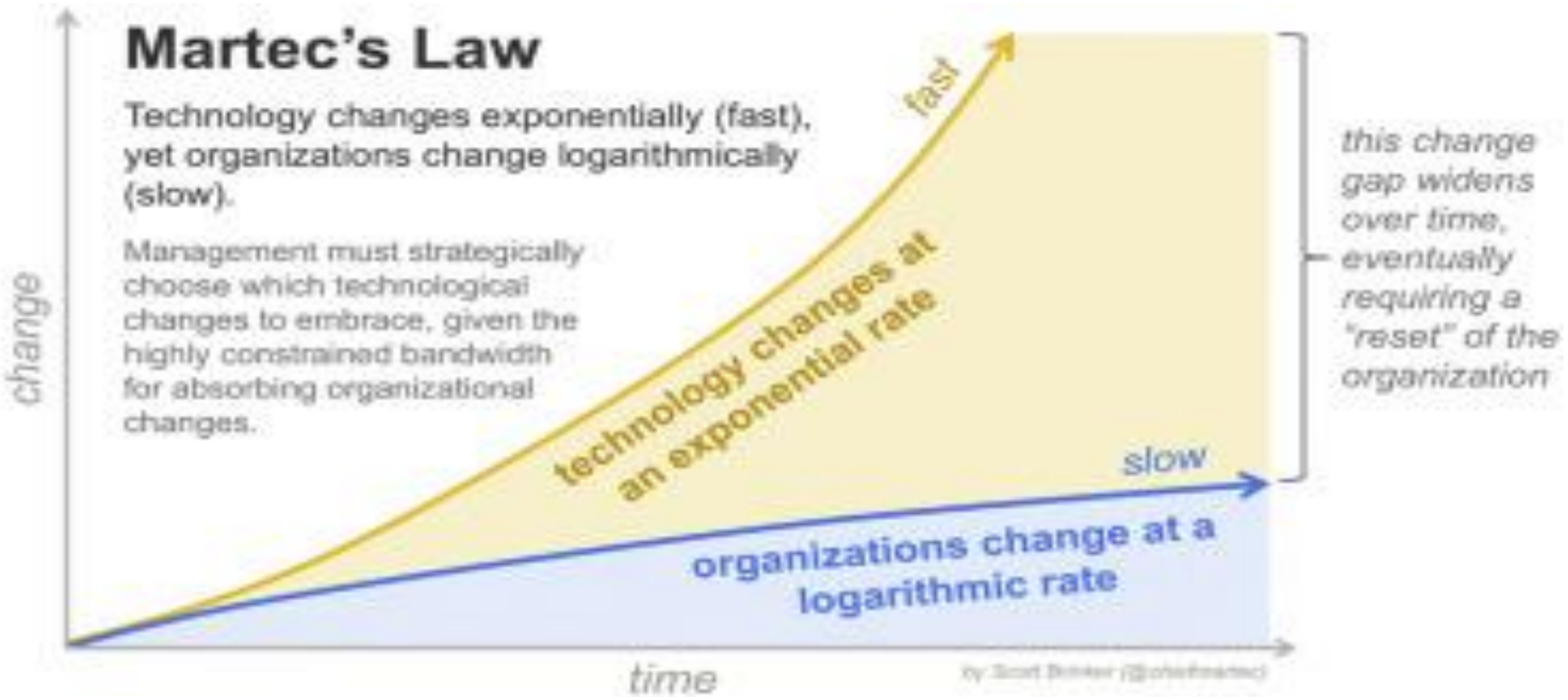
- **Development and validation:** how will applications be developed and validated?
- **Implementation**: what are the implementation and usability testing strategies?
- **Effectiveness**: what are the metrics of real-world effectiveness and fidelity and how will they be measured?
- **Ethics**: how will fairness, equity and freedom from vested commercial interest be assured?

Strategy: Prioritise end to end projects



Li et al NPJ Digital Med 2020

Lagging organisational readiness and capacity

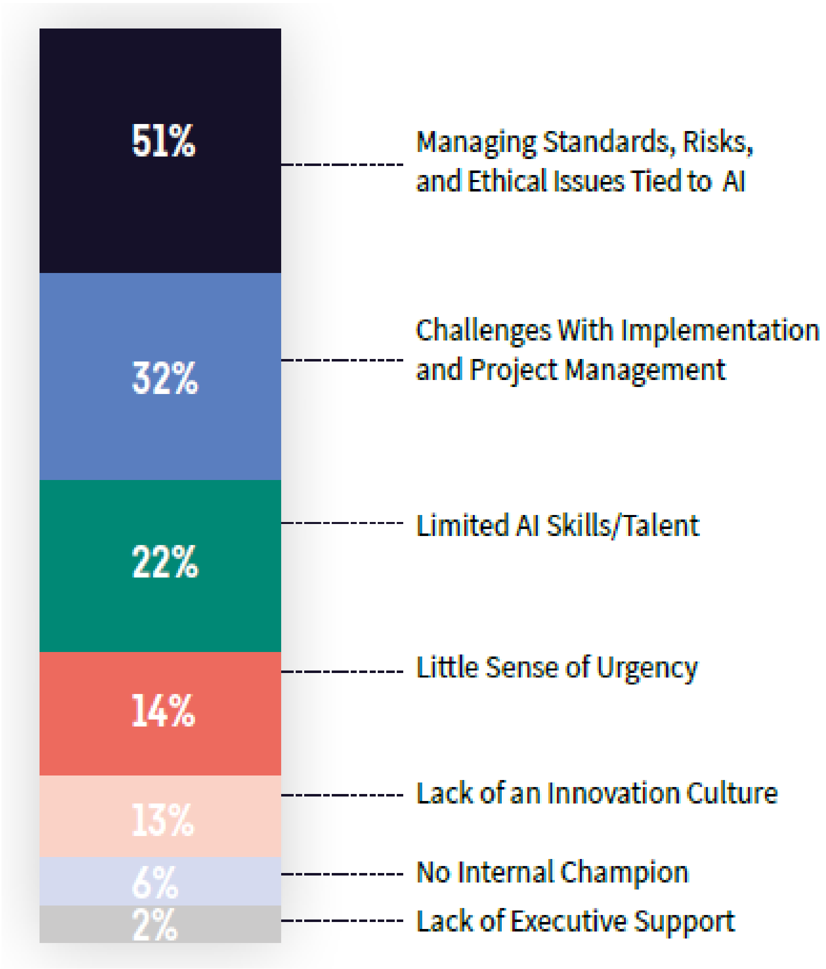


Brinker S. (2016). Martec's law: technology changes exponentially, organizations change logarithmically. Najdeno.https://chiefmartec.com/2016/11/martecs-law-great-management-challenge-21st-century/

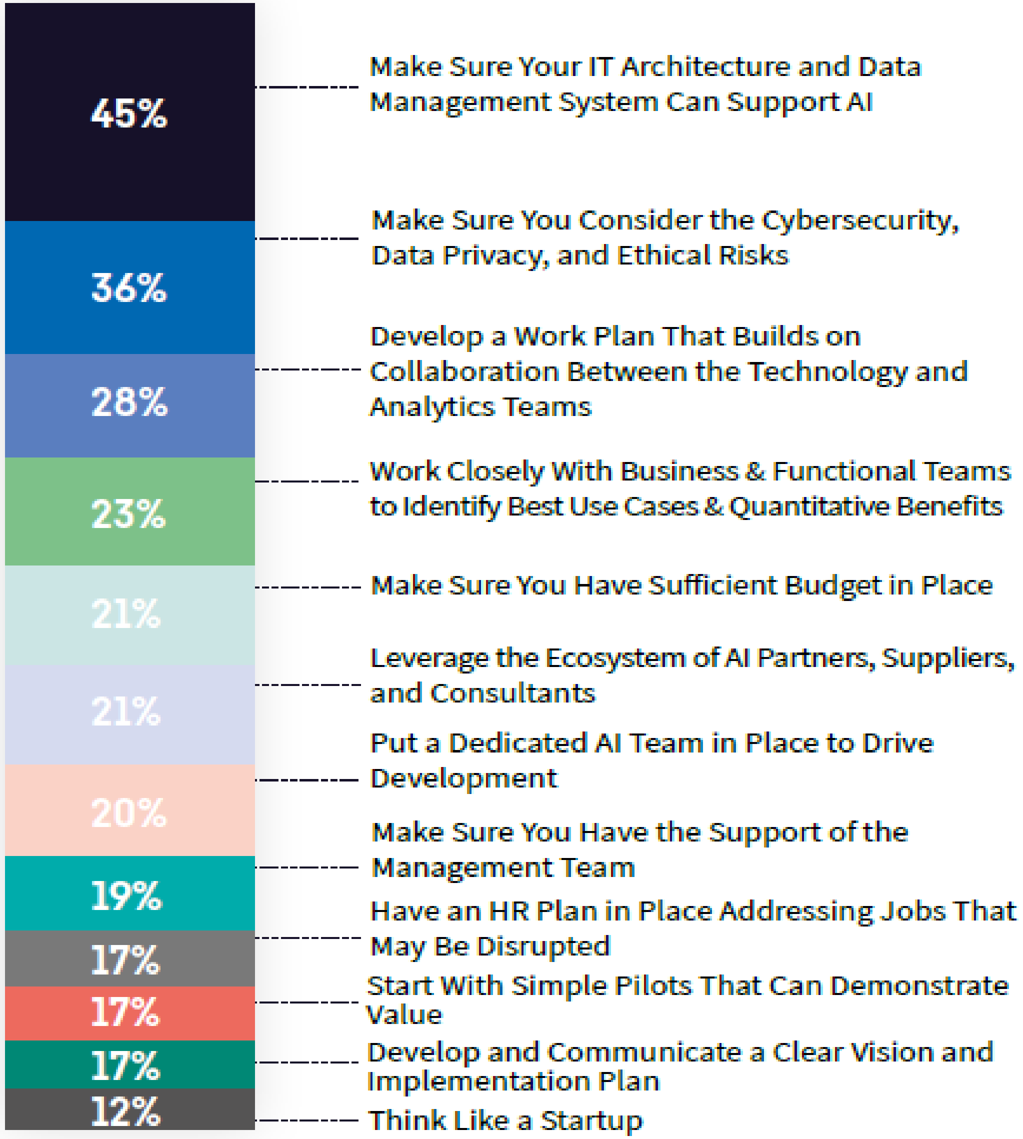
Lagging organisational readiness and capacity

1200 data executives - Dataiku survey Sept 2020

What are the biggest challenges
in adopting AI in your healthcare organisation?



Which of the following do you believe have been the most important lessons learned in implementing applied AI within your healthcare organization?



Strategy: Embed AI/ML aspirations into health service strategic plans

Digital Health 2031

A digital vision for Queensland's health system

Queensland Government

Horizon 3

Embedding digital

Within 10 years

Augment and embed digital models of care and ways of working that leverage technology and innovation and realising the full benefits.

- Consumer managed mobile healthcare
- Reliable connectivity and access
- Digital hospital (full)
- Insight-driven and intelligent care
- Precision medicine and public health

3. Insight-enabled Clinicians

Provide clinicians meaningful insights and tools to enable smarter, safer and higher quality care delivery. Key to this will be leveraging Artificial Intelligence (AI), analytics, decision support, and integrated health records to support prevention, treatment and ongoing wellness using comprehensive health data and evidence-based care.

- The use of proven AI models is embedded that provide decision support and insights at the point of care
- The digital ecosystem is continuously enhanced to leverage new technologies, while maintaining the ongoing security and safety of digital assets

Adam Shepherd – eHealth Qld

Strategy: Present AI/ML as essential to LHS

Making the move to a learning healthcare system:
has the pandemic brought us one step closer?

LEARNING
HEALTH SYSTEM

Research

Care delivery

Building a Learning Health System for Canadians

Report of the Artificial Intelligence
for Health Task Force

July 2020

Rapid improvement & decision making that is:

- Timely
- Centered on member needs
- Informed by science
- Driven by our own data
- Designed to be sustainable

Allen et al Learn Health Syst 2021

Strategy: Assess organisational readiness and capacity

	1	2	3	4	5
Problem relevance	Organisation feels AI not relevant and cannot identify any need	Organisation feels AI of little relevance and cannot identify current need	Organisation feels AI relevant to meet some current needs. AI pilot project(s) planned	Organisation feels AI is very relevant and has identified pressing need. AI project(s) ongoing with defined metrics	Organisation feels AI very relevant to meet urgent need. AI mainstream technology
Organisational leadership	No/little interest from executive leadership	Some (<50%) interest from executive leadership	Major (>50%) interest from executive leadership	3+ commitment from executive leadership	AI playing a role in executive leadership decisions
Organisational culture	No/little interest/knowledge from staff	Few staff (<5%) engaged in AI education/training	Some staff (5–10%) engaged in AI education/training	More staff (11–20%)+ ongoing education/training program	4+ regional/national meetings
Technology infrastructure	Poor	Below average	Average	Above average	Excellent[A]
Data quality	PACS and/or CPOE only	1+ physician documentation	Complete EMR	3+ HIE and patient portal	4+ enterprise data warehouse
Data analytics	Basic analytics (reporting)	Advanced analytics (forecasting)	Descriptive analytics (machine learning)	Predictive analytics (machine learning)	Real-time, time series (reinforcement learning)
Cybersecurity	Poor	Below average	Average	Above average	Excellent[B]
Development life cycle	One-off design, development and implementation	Occasional post-implementation user surveys and safety checks	Set time interval for review using pre-defined performance indicators	Periodic review and recalibration as required, according to new evidence, user feedback	Continuous surveillance and optimisation
Team competencies	One full-time data scientist or data champion	More than one data scientist (at least one at PhD level)	2+ more scientists but no team leader	Team and designated senior leader in data and/or AI	Dedicated AI entity
Financial resources	No financial support or business plan	Ad hoc support for small projects (≤$25 000) but no business plan	Support for larger projects (>$25 000) and business plan	Part of IT budget designated for AI projects	Separate designated budget for AI projects/team

Scott et al Aust Health Rev 2021

Strategy: Provide a roadmap for AI/ML realisation

Conceptualise use case
- Identify clinical need with stakeholders
- Undertake literature review for existing AI solutions suitable for local customisation
- Propose idea for new AI solution
- Obtain organisational approval and funding
- Formalise stakeholder collaboration
- Obtain ethics and governance approvals

Develop and test AI application
- Define data elements; acquire, clean and curate required data
- Formulate target performance metrics
- Build and test prototype application under research conditions
- Demonstrate pre-specified performance metrics
- Refine and validate the application using external datasets

Implement and evaluate the application in clinical trials conducted real-world settings
- Design a hybrid implementation-effectiveness trial
- Obtain ethical approval and funding
- Conduct the trial under best practice standards and assess application usability, impacts on workflows and care processes, patient outcomes, costs
- Gather end-user feedback on ease of use and acceptance of advice

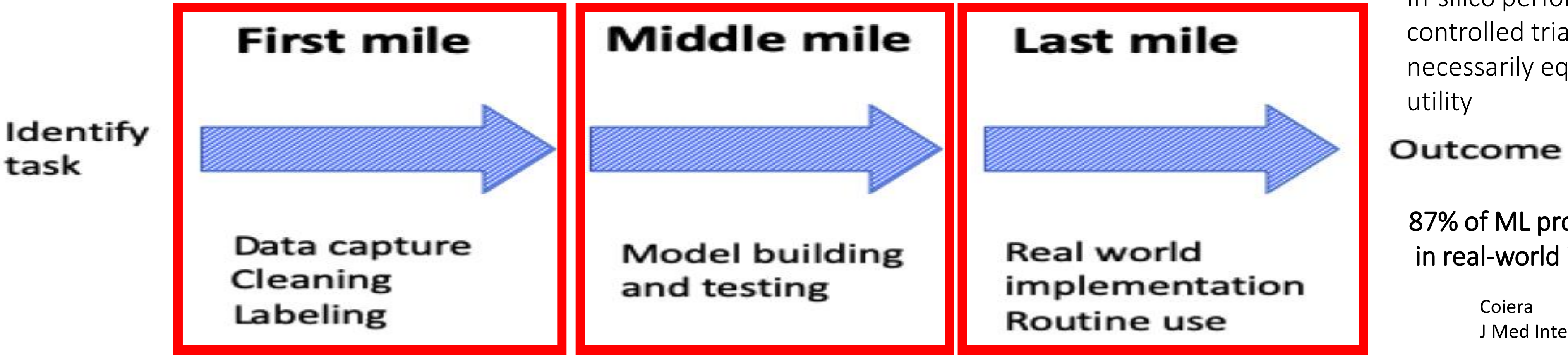
Optimise the application
- Refine the application in response to trial results and user feedback
- Recalibrate the application using additional data
- Modify interface design to optimise ease of use
- Re-assess application performance in an additional testing cycle

Seek peer review and regulatory approval
- Publish findings in peer reviewed journals
- Apply TGA for licensing as software as a medical device
- Negotiate patents and commercialisation contracts
- Address logistics of dissemination, training and education, life-cycle monitoring and updates
- Develop business case templates for organisations

Seek professional endorsement
- Incorporate recommendations for using the application in clinical practice guidelines
- Maintain registry of endorsing organisations
- Obtain feedback on tool performance from post-marketing surveys

Scott et al Aust Health Rev 2021

Constraints in the AI/ML application pipeline



In-silico performance or controlled trials do not necessarily equate with clinical utility

87% of ML projects fail in real-world implementation

Coiera
J Med Internet Res 2019

Development and validation phase (in-silico)

Production phase (real world)

Poor use case
- Not well defined or justified
- Insufficient cognizant of context
  - Workflows, resource, cost

Limited data quality, quantity, access
- In silos, not standardised/structured, not interoperable
- Incorrect, biased, missing data
- Insufficient labelled data, incorrect or inappropriate labels
- Difficulties in accessing the data
- Inadequate data storage capacity

Problems in model development and optimisation
- Hit and miss feature selection and engineering
- Time consuming recurrent testing and iteration of different models
- Subjective choice of best models
- Limited computational processing capacity

Barriers to model deployment
- Inability to convert the model into a usable and implementable application at point of care
- Inability to embed applications into electronic health record
- Absence of checks and balances to ensure application accuracy and safety
- No responsibility for life cycle management - monitoring, sustaining, updating the application
- Lack of clarity as to who is in charge and taking responsibility
- User rejection

Constraints in the AI/ML application pipeline

Defining the task

Use case

Data collection, processing, curating, labelling (‘Data Janitorial’)

HARD ENOUGH (but does it need to be?)

First mile

Data scientists spend around 50%-80% of their time on just preparing the data for processing and 76% of data scientists view data preparation as the least enjoyable part of their work’

BUT once done on large datasets, can be used for multiple tasks

EASY (and getting easier)

Feature selection and engineering

Model development and refinement

Model validation (internal)

Model validation (external, multi-site)

Second mile

REALLY HARD (and very few have done it)

System interface
Feasibility
/usability testing

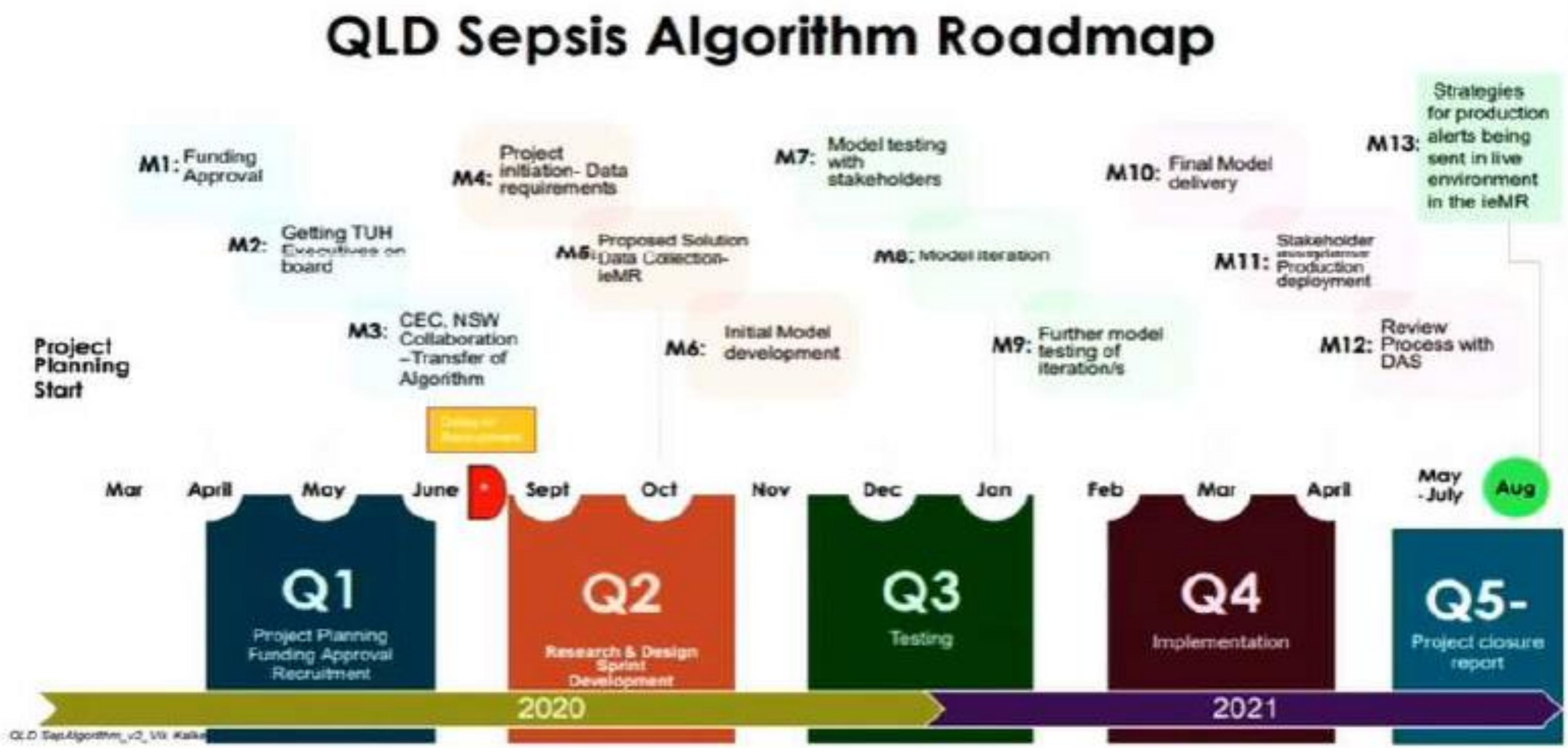
Implementation
with live data
(silent mode)

Effectiveness
(active mode)
Before-after, ITS, RCT

Prospective evaluation

Third mile

Constraints in the AI/ML application pipeline



Courtesy of Paul Lane, Rudolph Schletner, Vikrant Kalke

Constraints in the AI/ML application pipeline

Courtesy of Paul Lane, Rudolph Schletner, Vikrant Kalke

WASTE of TIME, EFFORT, RESOURCES

'A significant data access barrier relates to the complex, multilayered, opaque, process of obtaining appropriate authorisation for use.' Narelle Doss eHealth Qld 2019

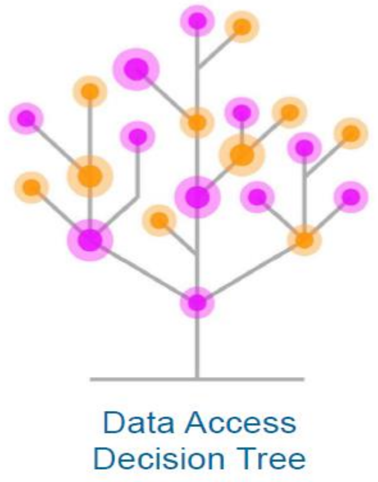
Strategy: Accelerate the approval processes for multi-site projects

- Ethics – Data custodians -- PHA --- SSA ---------
  - Single QH ethics approval process
  - Single QH digital data custodian and authoriser
  - Single data repository and sharing platform with real-time accessibility

Potential Future State

Will deliver a decision tree algorithm that will provide the following benefits:

- Reduced complexity and time taken to facilitate data access authorisation
- Enhanced fluidity of data for appropriate secondary uses
- Improved standardisation of processes regarding data access and governance
- Reduction of inappropriate data blocking
- Identification of areas for streamlining of the process

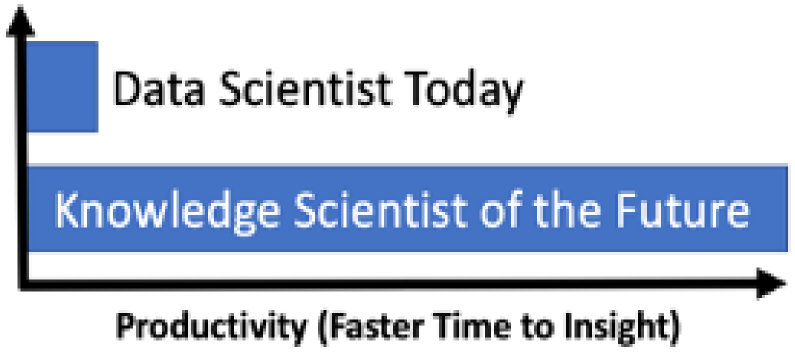


  - Central AI/ML approval co-ordination centre/service

- Opt-out patient consent
  - Inform public of how and why their data will be used
    - Strict privacy provisions; de-identified, anonymised datasets
  - Remove personal health information from meta-data
  - No ownership of data by commercial third parties
  - Any AI application will be 'common good' and not monetised

Strategy: Accelerate data collection and aggregation

- Optimise 'structuredness' of data sources (esp EMR)
    - Data definitions and data dictionaries
    - Templates, power plans, coding nomenclatures (e.g. SNOMED-CT)



- Use NLP wherever possible

- Employ unified data formats (using HL7-Fast Healthcare Interoperability Resources [FHIR]) for aggregating data from different sources

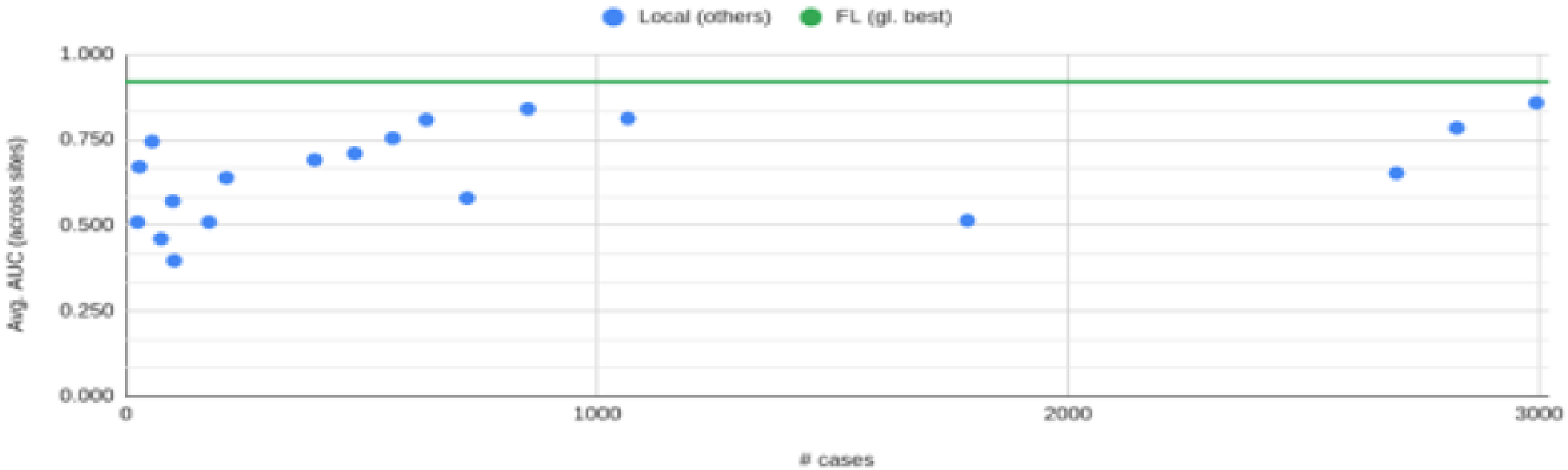
- Consider dynamic graphic databases rather than relational databases and flat files
    - To optimally preserve relationships between variables within large datasets and create high quality meaningful knowledge graphs

- Interrogate datasets for quality with new scoring tools (e.g. MarkLogic) and reject those of poor quality

- Optimise variable selection (and required data) by choosing predictive features identified by clinical domain experts and prior regression models as being relevant, and accessible when required

Strategy: Use open source libraries, transfer learning, federated learning

- Open sources libraries
  - Large ML-ready annotated training datasets
  - ML-ready model codes
- Transfer learning
  - Model developed for one task using a specific training dataset, available as open source code is used as a base model to be retrained on data for a similar or related task.
- Federated learning
  - Model sited in one central server shared with other client sites which is then run on local data without the need to share data across sites.
  - Each client site learns locally and shares model weight updates with central server that aggregates contributions using secure encryption and communication protocols.
  - Central server and client sites iterate back and forth minimising global loss function and synchronising learned site weights until the model converges.



FL (best) AUROC 0.920

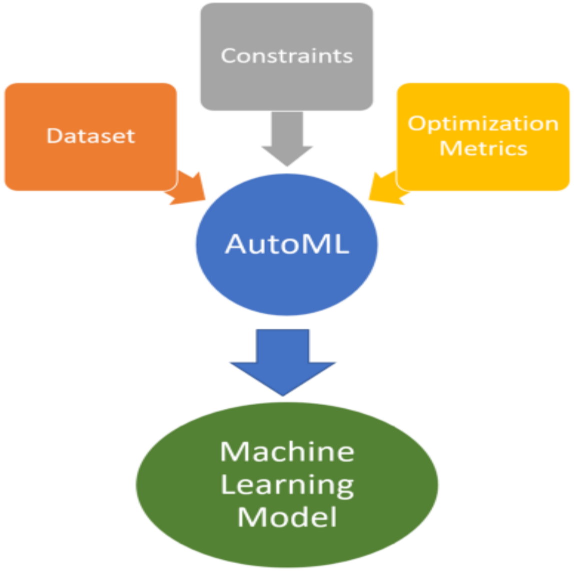
Av local AUROC 0.795 (n=19)

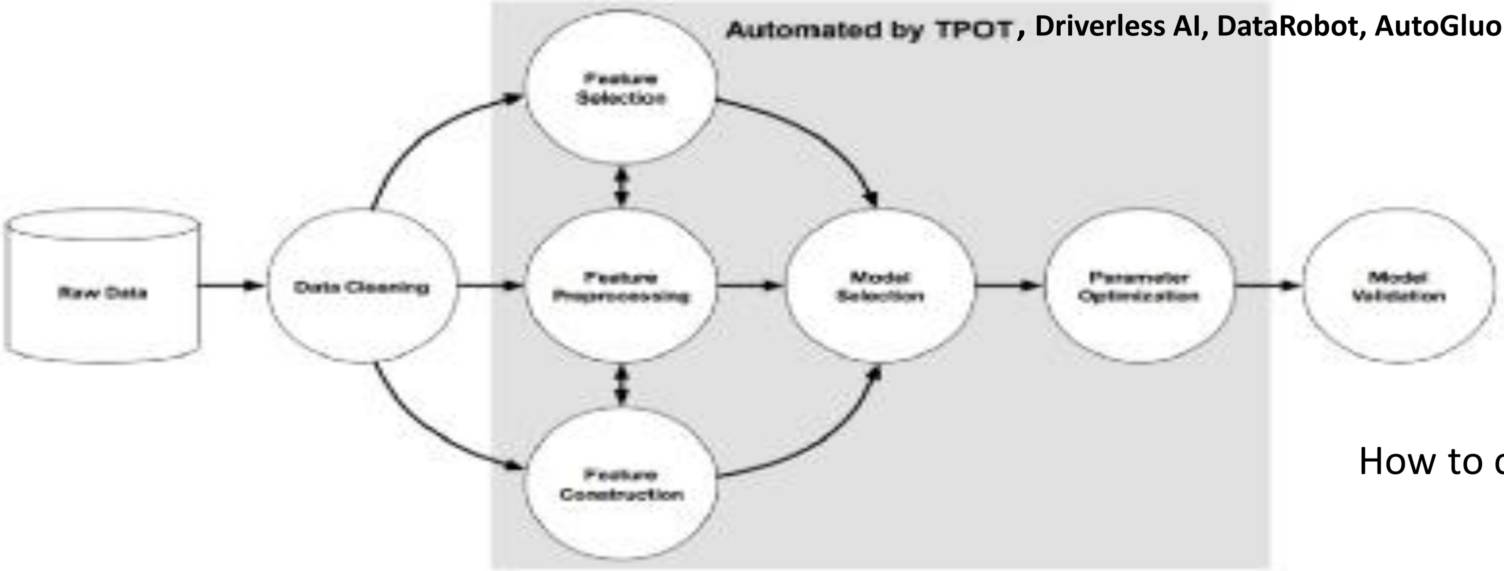
Flores et al Res Square 2021

Predicting need for respiratory support at ED presentation in COVID patients

Strategy: Automate model development where possible

- AutoML attempts to totally or partially automate the ML workflow
  - feature preprocessing, feature selection, feature construction, model selection and (hyper)parameter optimization





, Driverless AI, DataRobot, AutoGluon, MS AzureML

How to choose?

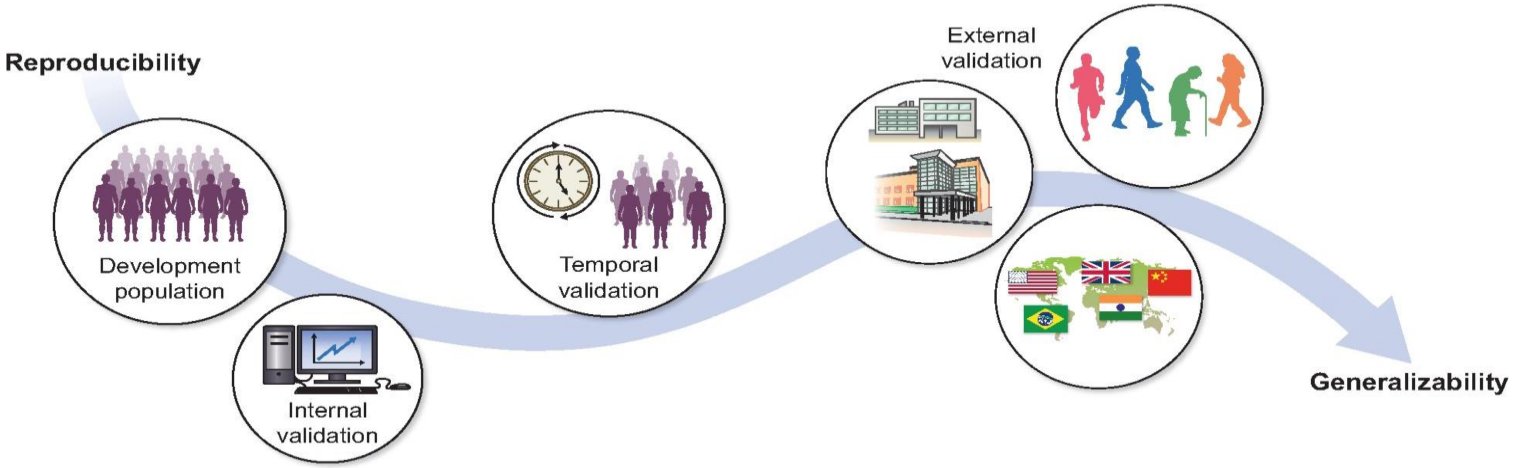
Traditional AI process

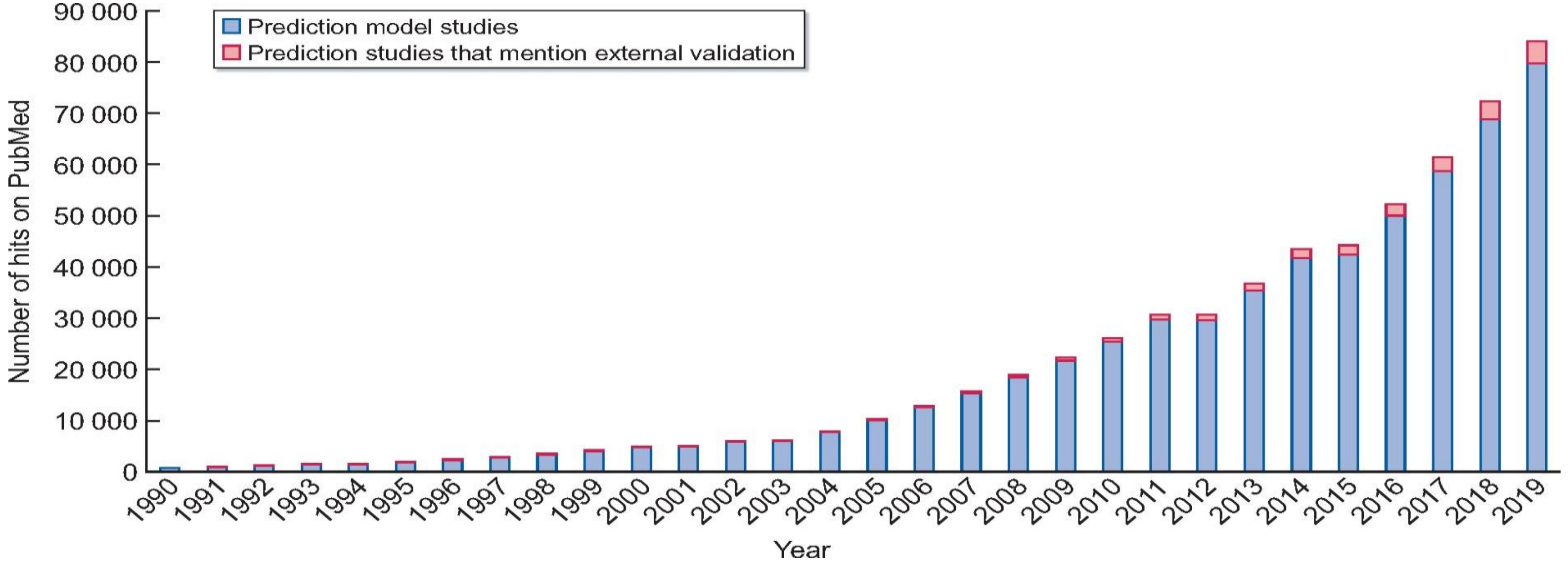
No-code AI process **Maker Pad, Zeroqode, NoCode**

Intersection of no-code, SaaS, and AI

May not have to choose two-out-of-three between fast/cheap/good

Strategy: Ensure external validation





Ramspek et al
Clin Kind J 2021

The Epic Sepsis Model Falls Short—The Importance of External Validation

Anand R. Habib, MD, MPhil; Anthony L. Lin, MD; Richard W. Grant, MD, MPH

Strategy: Use add-on apps rather than reconfigure EMR platform

- Cerner's FHIR API Platform allows an in-house team to easily iterate a best-fit application for end-users
- 'The SMART® (Substitutable Medical Apps and Reusable Technology) platform defines a specification for an EMR to safely and securely open other applications with context
- SMART applications commonly web applications but may also be native mobile applications that use HL7® FHIR® standard to read and write data from the EHR
- Cerner can embed a SMART app in the EMR and Cerner will also support FHIR access through mobile SMART applications (through their Ignite platform), as those specifications emerge from the SMART web site.
- Cerner expects majority of SMART apps to be provided to clients in a SaaS (software as a service) model.
  - SMART app is hosted or managed by the SMART developer or provider.
  - Clients do not need to install any code or package in order to implement a particular SMART app.'

Data taken from Cerner website [SMART (cerner.com)]

Strategy: Formulate and implement robust implementation plan

Reddy et al J Am Inform Assoc 2020

Development Check

Pre-Development Phase	Objective	Dataset Source and Integrity
Pre-Development Phase	Use Case	Transparency
Development Phase	Performance Metrics	Internal Validity

Deployment Check

Pre-Deployment Phase	Generalizability and Contextualization	External Validity
Deployment Phase	Technical Integration	Privacy
Post-Deployment Phase	Safety and Quality	

Discernment Check

Short Term Phase	Non-Maleficence	
Mid Term Phase	Use in a Healthcare Setting	
Long Term Phase	Number of Services	Alignment with Domain

Reddy et al. BMJ Health Care Inform 2021
Translational Evaluation of Healthcare Artificial Intelligence (TEHAI)

Strategy: Secure adequate end to end funding

- Identify potential funding sources
  - MRFF, NHMRC, Government (state/federal), Health insurance funds, etc
  - Prioritise long term funding (preferably 3 years)

- Formulate business case
  - Clear AI/Ml need and suitability
  - Cost to build, deploy and maintain the model
    - human effort and expertise needed to acquire necessary data, ensure data quality, and establish the infrastructure and data pipelines
    - ongoing model maintenance and support needs to detect and allay decline in model performance or model shifts
  - Estimate the benefits and ROI in dollar terms

- Secure binding contract with quarantined budget
  - Not in an HHS/facility operational cost centre

- Transition to permanent embedment

Under-developed regulatory guidance

Australian Government
Department of Health
Therapeutic Goods Administration

Regulatory changes fo
based medical devices

Version 1.2, August 2021

TGA Health Safety Regulation

Australian Government
Department of Health
Therapeutic Goods Administration

Clinical decision suppo

Scope and examples

Version 1.1, October 2021

TGA Health Safety Regulation

Australian Government
Department of Health
Therapeutic Goods Administration

Examples of regulated and unregulated software (excluded) software based medical devices

Version 1.3, October 2021

TGA Health Safety Regulation

Strategy: Consider and adapt policies of other countries

European Commission Artificial Intelligence Act April 2021

- Four levels of AI risk: unacceptable risk, high risk, limited risk, minimal risk.
- Healthcare AI applications would generally fall into the high-risk category
- Will need to fulfill the following criteria to achieve regulatory approval:
  - Adequate risk assessment and mitigation systems
  - High quality of the datasets feeding the system to reduce risks and discriminatory outcomes
  - Logging of activity to ensure traceability of results
  - Detailed documentation providing all information necessary on the system and its purpose
  - Clear and adequate information to the user
  - Appropriate human oversight measures to reduce risk
  - High level of robustness, security and accuracy

FDA AI/ML-based SaMD Action Plan January 2021

- Five areas of focus, each with actions the FDA intends to take:
  - Issuance of draft guidance on a predetermined change control plan (for software's learning over time)
  - Supporting the development of good ML practices to evaluate and improve ML algorithms
  - Fostering a patient-centred approach, including device transparency to users
  - Developing methods to evaluate and improve ML algorithms
  - Advancing real-world performance monitoring pilots

Collaborate with TGA in working up case exemplars

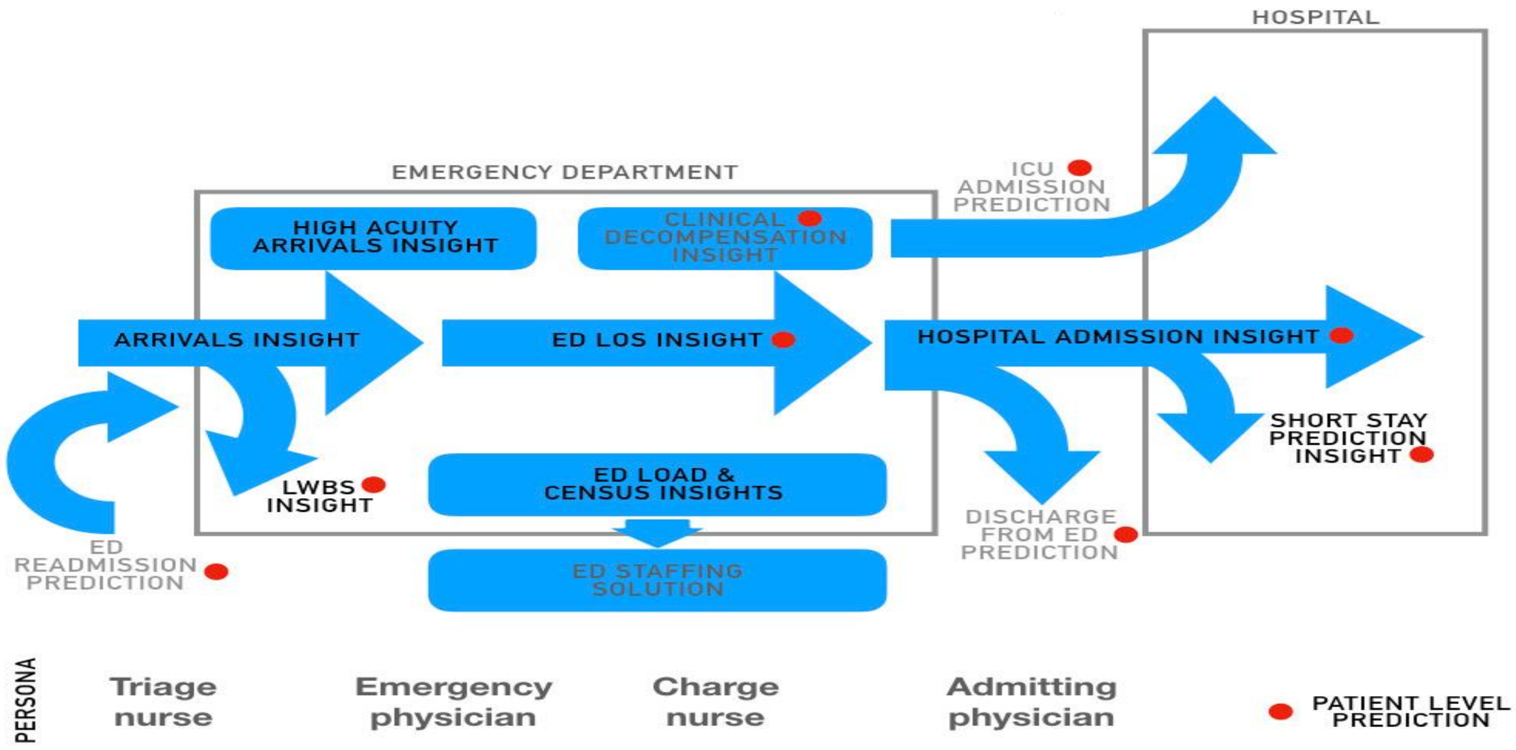
Use cases in progress

Predicting ED disposition and LOS

Problem: ED congestion due to impaired patient flow leads to ambulance ramping and longer ED-LOS – associated with increased patient mortality and safety events; need to accelerate patient flow through ED

Prior research: ML models outperform best rolling average estimates of LOS by over 20% and reduce the number of patients with large underpredicted waiting times by 42%. (Pak et al Int J Med Inform 2021)

Tasks: Early identification of patients who need inpatient admission so that bed allocation and patient transfers (eg direct to ward admissions) can be expedited
Early identification of patients who are likely to have LOS >4 hours to enable targeted assessment and intervention



Working Group
Andrew Staib
Stephen Canaris
Jasmine Wilkersen – KenSci
Will Barnett – KenSci
Asish Singh - KenSci
Ian Scott

Predicting ED disposition and LOS

ED disposition

Binary classifier (ED discharge vs admit)

AUC = 0.97

Model	AUC	Accuracy	Precision	Recall	F1 Score
Best Historical Model	0.688	0.693	0.679	0.693	0.678
Best Realtime Model	0.965	0.918	0.918	0.918	0.918
Best Historical + Realtime Model	**0.971**	**0.925**	**0.925**	**0.925**	**0.925**

Historical+ Realtime Model

ACTUAL \ PREDICTED	ED Discharge	EDIP
ED Discharge	99	5
EDIP	7	49

ED LOS

Binary classifier (< 4 hrs vs > 4 hs)

AUC = 0.848

Model	AUC	Accuracy	Precision	Recall	F1 Score
Best Historical Model	0.639	0.633	0.629	0.633	0.598
Best Realtime Model	**0.848**	**0.778**	**0.784**	**0.778**	**0.771**
Best Historical + Realtime Model	0.832	0.759	0.767	0.759	0.749

Realtime Model

ACTUAL \ PREDICTED	LOS < 4 hours	Overstay
LOS < 4 hours	84	9
Overstay	26	40

- Predictions generated within 30 minutes of ED presentation
- Use MS data Nov 2015 to Jan 2020

Early identification of patient developing sepsis

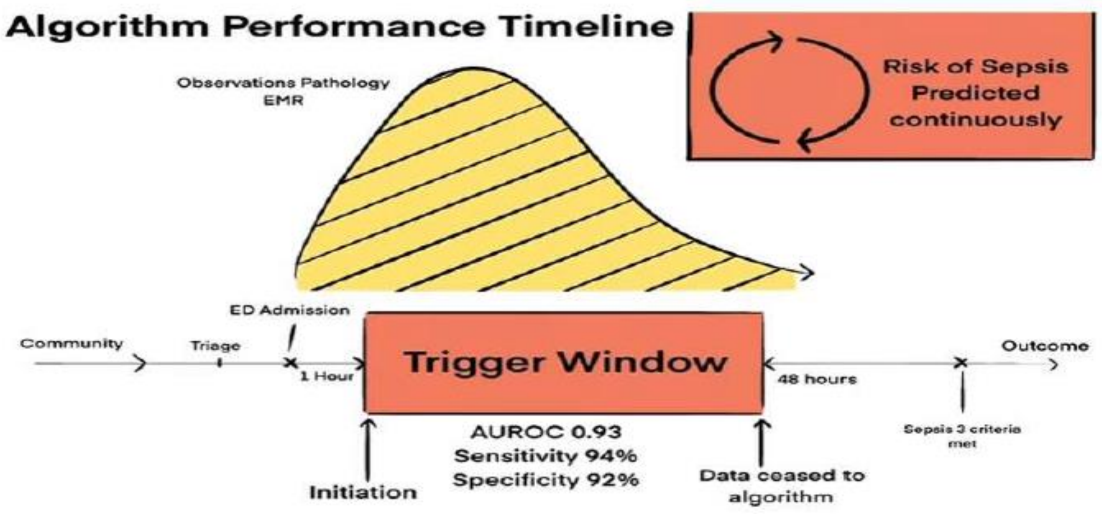
Problem: Hospital-acquired sepsis affects up to 8% of inpatients; mortality up to 20%; ICU up to 50%
At least 30% septic episodes deemed preventable if septic-incipient patients identified early & care bundles instituted.
Clinicians recognise no more than 70% of incipient sepsis (Barker et al unpublished data 2021)

Prior research: ML models show high accuracy (AUROC >0.90) in identifying sepsis patients resulting in decreased mortality and LOS (Shimabukuro et al BMJ Open Resp Res 2017; McCoy et al. BMJ Open Quality 2017). None have undergone validation and feasibility testing in Australian settings.
Paul Lane and colleagues at Townsville HHS and NSW CEC: model using 4 years data from 10 QH digital hospitals (1.13 m encounters; 26,753 sepsis cases – Sepsis 3; mortality rate 10%) that identifies incipient sepsis up to 48 hours before clinical recognition (Schletner et al unpublished data 2021).

Task: Optimize the model further, place it into silent production mode, and verify accuracy in real time as well as user and feasibility testing prior to prospective before-after studies.

QH Sepsis AI Working Group

Paul Lane
Stephen Canaris
Ahmed Abdel-Hafez
Luke Lawton
Oscar Bonilla
Ian Scott
Rudolph Schnetler
David Cook
Amith Shetty
Aldo Saavedra
Vikrant Kalke
Lyndell Redpath



'Katharine'

Predicting IV heparin bolus and maintenance dosing

Problem: Weight based dosing nomograms achieve therapeutic aPTT in only 22% of patients; inappropriate dosing results in bleeding in 6-8% patients, and recurrent thrombotic events in up to 10%; frequent dose adjustments by medical/nursing staff; more blood tests for patients

Prior research: ML models for estimating heparin dosing are very few, restricted to specific populations (ICU, dialysis) and lacking external validation (single centre)   (Falconer et al Br J Clin Pharmacol 2021)

Tasks: Develop and validate a model for estimating aPTT and its classification (sub-, supra-, therapeutic) based on bolus and maintenance dosing; optimise this model to select optimal dosing rates in individual patients

Heparin Dosing Working Group
Nazanin Falconer
Michael Barras
Ahmed Abdel-Hafez
Aaron van Garderen
Sven Marxen
Ian Scott
Stephen Canaris
Oscar Bonilla

Tool	Model	Accuracy	Macro Precision	Macro Recall	Macro F1-Score	Macro AUC
H2O DAI	FINAL ENSEMBLE MODEL	**0.599**	**0.554**	**0.686**	**0.613**	**0.735**
SKlearn	LogisticRegression	0.562	0.51	0.56	0.52	0.691
SKlearn	LogisticRegression with RFE*	0.557	0.49	0.56	0.5	0.687
SKlearn	SVM* – Linear SVC*	0.535	0.51	0.54	0.517	0.679
SKlearn	SVM – Polynomial SVC	0.451	0.46	0.45	0.457	0.614

		Predicted			
		aPTT < 70	100 ≥ aPTT ≥ 70	aPTT > 100	Total
Actual	aPTT < 70	943	25	104	1072
	100 ≥ aPTT ≥ 70	295	49	154	498
	aPTT > 100	251	36	301	588
	Total	1489	110	559	2158
	Accuracy	88%	51%	10%	

Abdel-Hafez et al JMIR Med Inform (pre-print 2021)

Class imbalance due to under-representation of this class in dataset

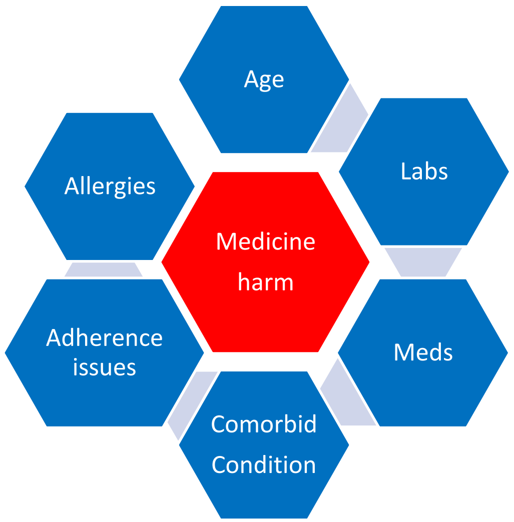
Early identification of patients at risk of ADE

Problem: Adverse drug events affect 15% to 30% of hospitalised older patients with at least half deemed preventable by virtue of review and adjustment or deprescribing of risk-inducing medications (Paradissis et al. J Pharm Pract Res 2017; Scott et al. JAMA Intern Med 2015)

Prior research: Systematic review of existing risk prediction rules (Falconer et al Br J Clin Pharmacol 2018); logistic regression model (AIME) for predicting ADE risk developed and validated using data extracted from EMR of 1982 patients admitted PAH over 6 mo (2017). AUROC = 0.70 (Falconer et al. Br J Clin Pharmacol 2020); recent addition of CFS (AIME-FRAIL) validated in 3948 patients (2020) improved AUROC to 0.79 (Falconer et al. unpublished data). Tool provides a score (8 point scale; score >5 high risk) inserted into EMR that clinical pharmacists use to identify and intervene in at risk patients shortly after admission; to be evaluated in a controlled trial.

Task: Optimize the model further using ML methods applied to larger patient sample using labelled data; repeat controlled trial with new tool as an app embedded in EMR, place it into silent production mode, and verify accuracy in real time as well as user and feasibility testing prior to prospective before-after studies.

Medication Harm Working Group
Nazanin Falconer
Michael Barras
Ahmed Abdel-Hafez
Aaron van Garderen
Sam Radburn
Neil Cottrell
Ian Scott



Accelerating radiology image reporting

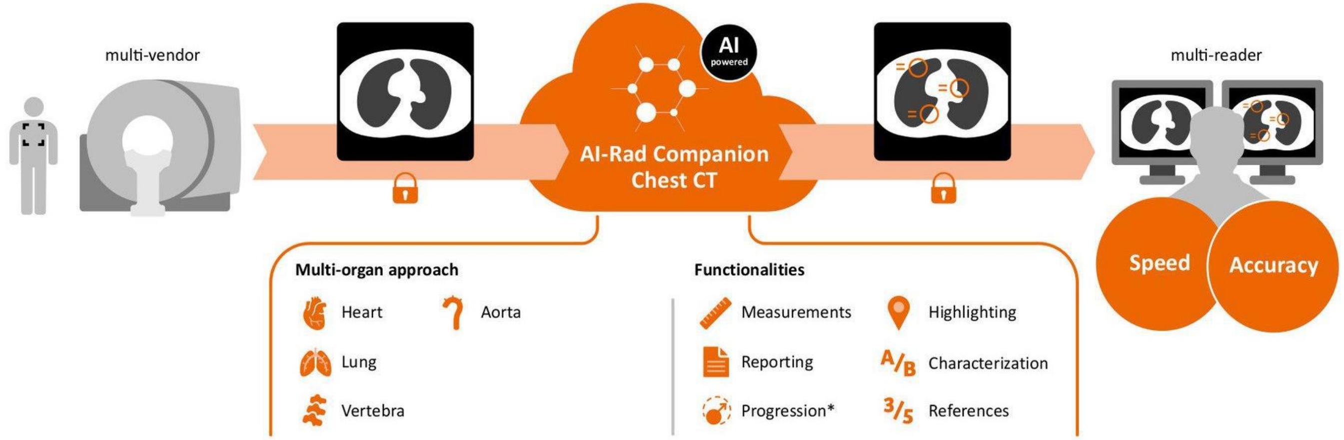
Problem: Backlogs of routine reporting of XRs and CT scans results in delayed decision-making and poses risk to patient safety; taking radiologists off-line to clear backlogs prevents attendance at MDT meetings and interferes with other duties (eg interventional radiology); incurs costs in overtime; may predispose to interpretive errors

Prior research:
For CXRs - well validated, Australian DL model improved accuracy of radiologists for 127 clinical CXR findings: AUC model 0·957 (0·954–0·959) vs AUC radiologists 0·713 (0·645–0·785); reduced interpretation time from 122 secs to 107 secs (Seah et al Lancet Digit Health 2021)
For CT/MRI scans - algorithms that facilitate lesion detection, localisation and segmentation/measurement to expedite reporting, and help with currently manual tasks (eg NLP tools for querying reports and other text-based records, generating x-ray report drafts from other AI model outputs) (Law et al Med J Aust 2021)

Tasks: Implement DL tool (AI-RAD) into CT scanner software for image segmentation and lesion identification for chest CTs; assess user acceptability, reduction in reporting times, detection of previously missed lesions.

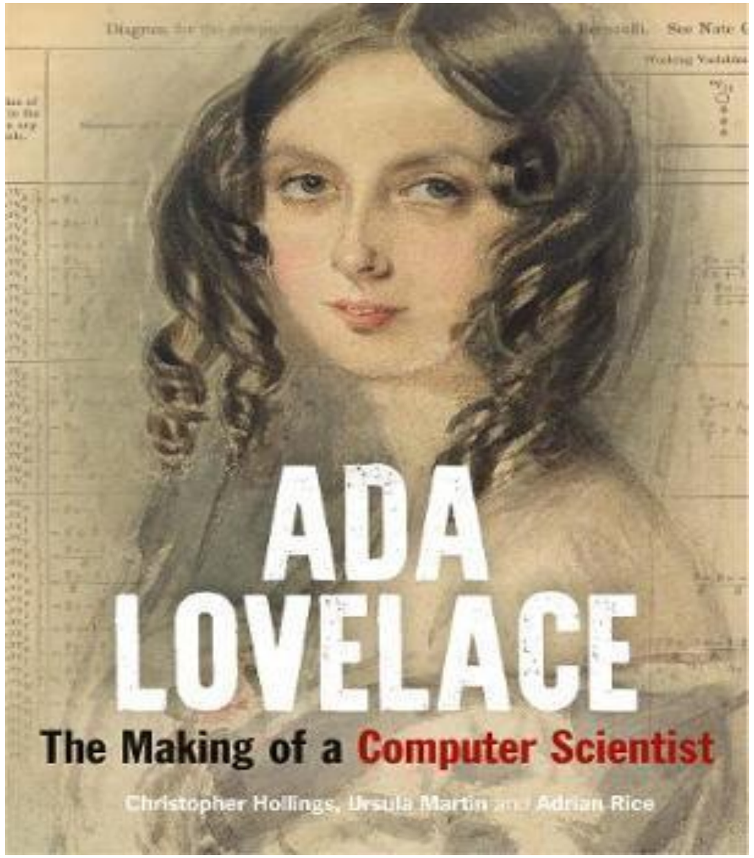
Radiology Working Group

Susanne Jeavons
Paul Schmidt
Bruce Shah
Gorane Santamaria
Kate McLean
Navine Saad
Amit Chacko
Rachele Quested
Steven McPhail/QUT
Ian Scott



Future directions

Lovelace 5-5-5 AI/ML program

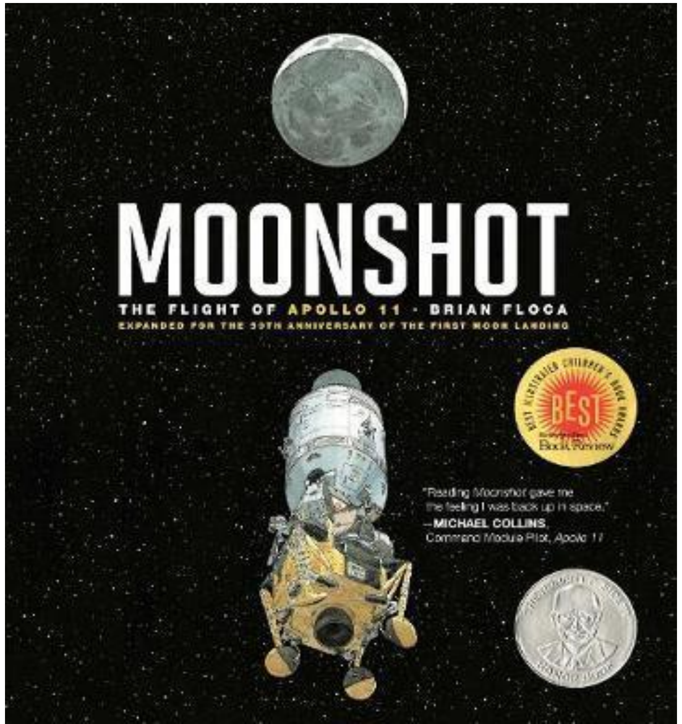


Ada, Countess of Lovelace (1815-1852), daughter of romantic poet Lord Byron and his highly educated wife, Anne Isabella, is sometimes called the world's first computer programmer. Despite no access to formal school or university education, a correspondence course with the eminent mathematician Augustus De Morgan helped her to develop into a gifted and perceptive mathematician. She became very interested in Charles Babbage's 'Analytical Engine' for which she published a table of mathematical formulae - the 'first programme.' Her paper provided clear explanations of the principles of computing, and broader ideas on computer music and AI.

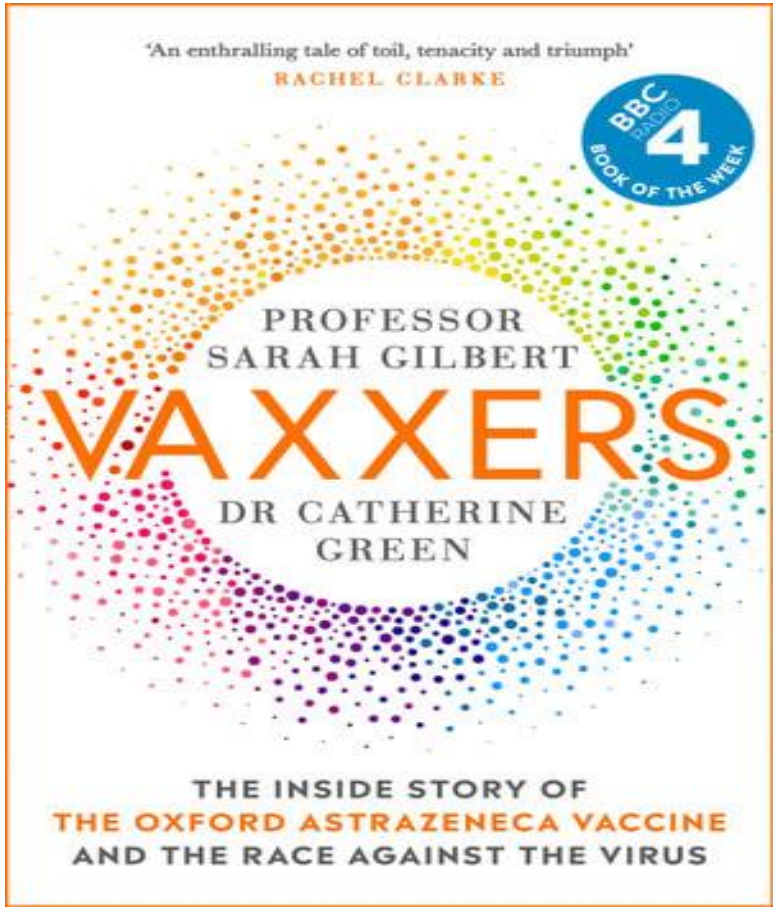
In **5 years**

We will have **5 AI/ML applications** operational

In at least **5 digital hospitals**



It can be done



- Early recognition of an urgent need
- Faith in their vaccinology
- Ability to form and galvanise coalitions
  - Collaboration not competition
  - Infrastructure for trials, production, distribution, delivery
- Parallel processing – not sequential
  - Lab-to-jab in 12 months
- Rapid efficacy testing in clinical trials
- Adequate and secure funding
  - Although not initially
- Responsive, co-operative regulatory agencies
- Accepting the risk of failure
  - 'We would ask for forgiveness, not permission'
- Never giving up

Concluding comments

Key messages

- Clinician engagement and literacy in AI/ML
- Stakeholder charter for AI/ML
- Statewide AI/ML research and development collaborative
- Streamlined data access and sharing processes and platforms
- Integrated governance structures that cover four quadrants: data, clinical, technical, research
- Organisational responsiveness to rapidly evolving AI/ML technologies
- Acceleration of the AI/ML application pipeline

Issues for consideration

- In-house only vs Partnership with commercial vendors
  - Legal contracts, commercialisation, IP
- Queensland only vs Multi-state collaboration
- SaaS/IaaS/PaaS vs On-premise software-based SaMD
- Project phase vs Life cycle management
- Explainability vs Trust
- Current care challenges vs Future care challenges

Questions???