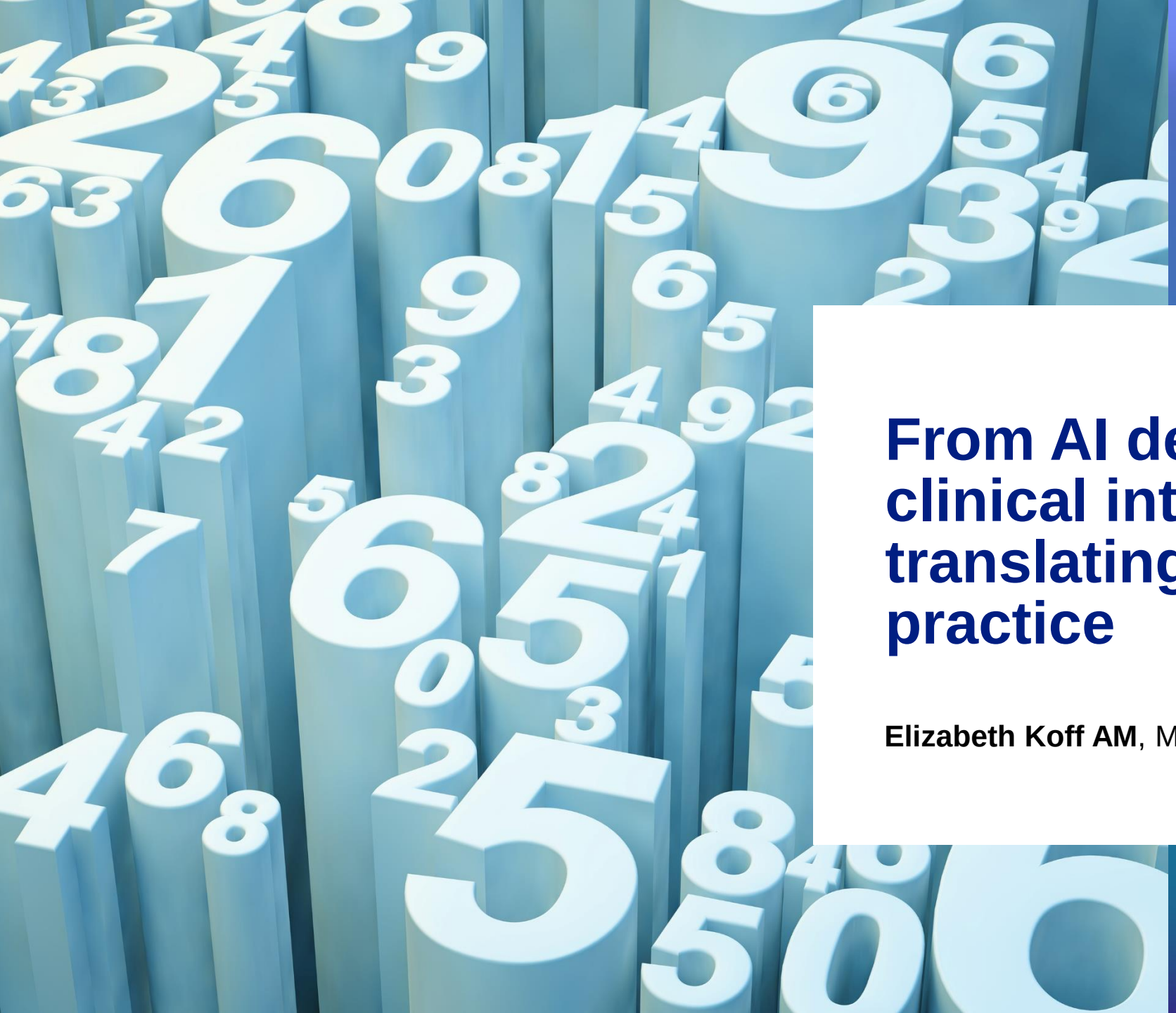
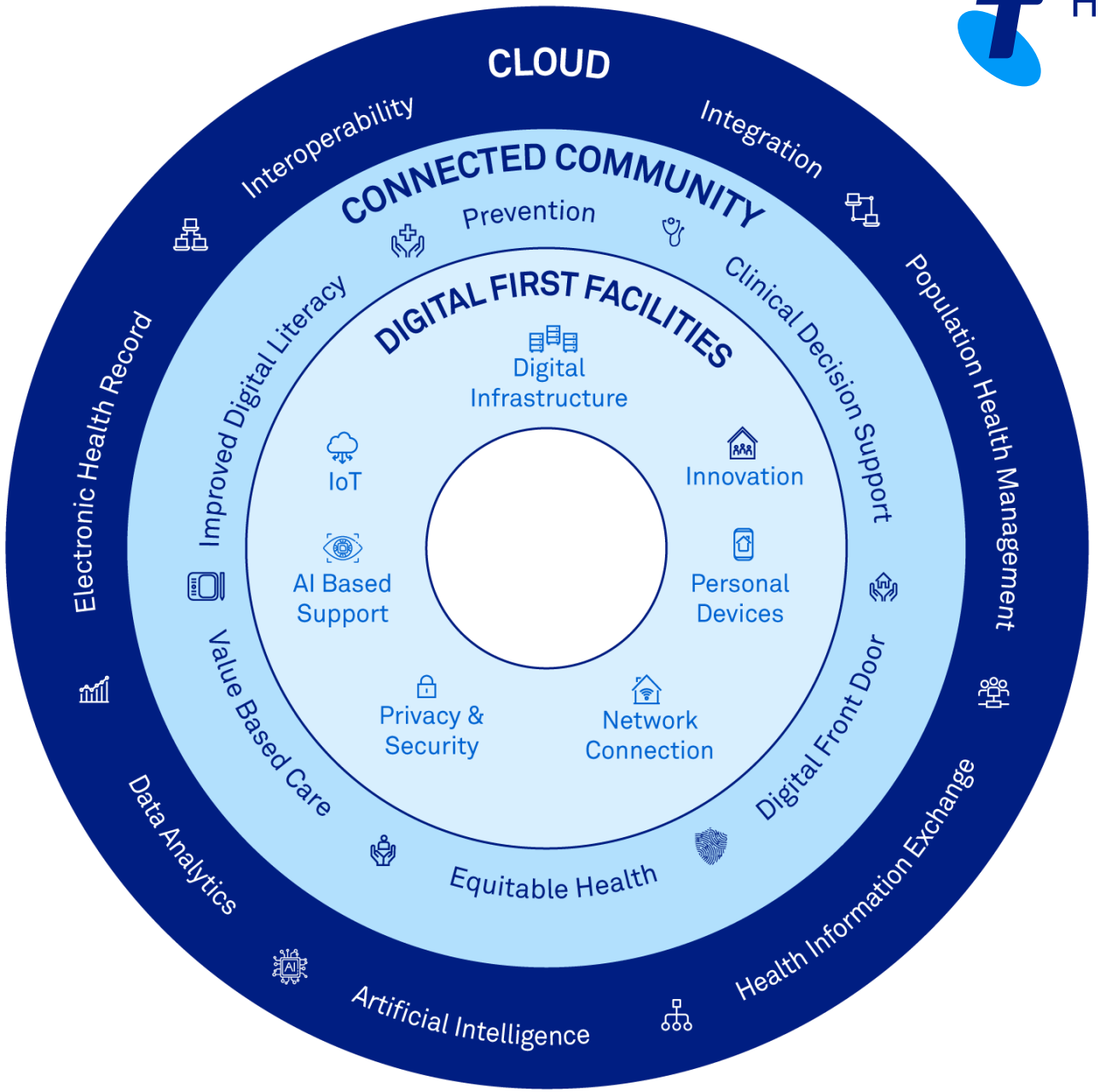


Health

From AI development to clinical integration: translating AI into clinical practice

Elizabeth Koff AM, Managing Director, *Telstra Health*

The digital ecosystem of healthcare



AI continuum

Health

Diagnostic AI

Why did this happen?

- Analyse
- Scenario
- Segment

Predictive AI

What might happen in the future?

- Pattern
- Forecast
- Model

What should we do next?

- Simulate
- Optimise
- Recommend

Generative AI

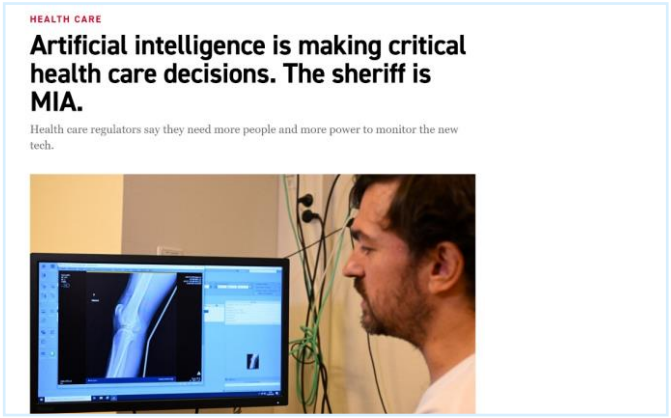
What can AI do?

- Advise
- Create
- Code
- Automate
- Protect

Current State of Play: AI in the Fourth Industrial Revolution

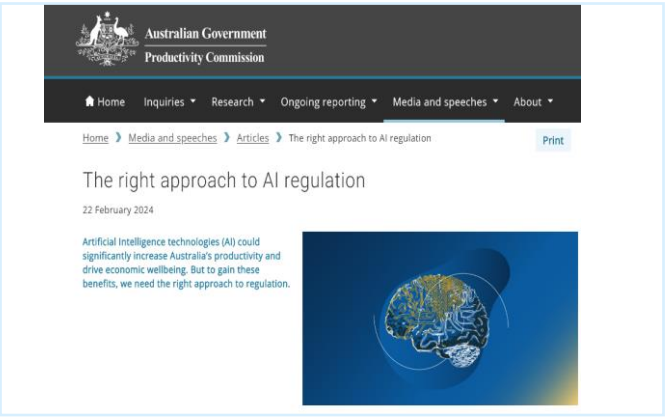
"Our AI Act will make a substantial contribution to the development of **global rules and principles for human-centric AI**. By **focusing regulation on identifiable risks**, today's agreement will foster responsible innovation in Europe. By guaranteeing the safety and fundamental rights of people and businesses, **it will support the development, deployment and take-up of trustworthy AI in the EU**."

Ursula von der Leyen, President of the European Commission
Dec 2023



"Unlike medical devices or drugs, AI software changes. Rather than issuing a one-time approval, **FDA wants to monitor artificial intelligence products over time**, something it's never done proactively."

Feb 2024



"The landscape for AI in Australia is **still developing**. **Australia will often be an international 'regulation taker'**. Other jurisdictions, such as the EU, are designing specific AI-regulations. **Developers, including those based in Australia, will need to meet these new rules if they want to access some of the world's biggest markets**. This means that **AI-technologies released in Australia**, even if developed locally, **will often be designed to meet overseas standards.**

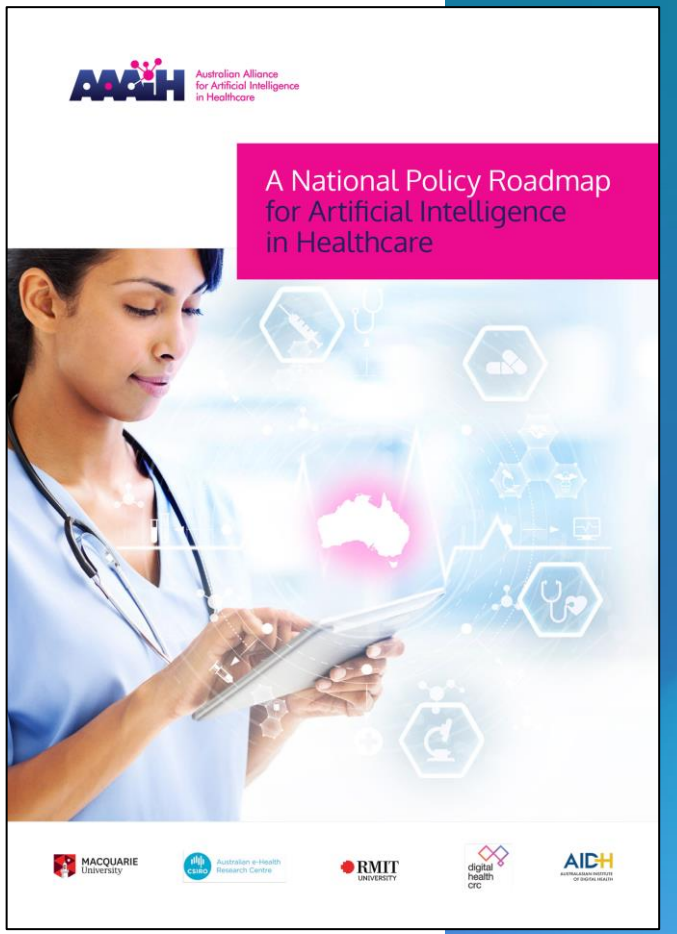
The **starting point** for any regulation is the **real-world use of AI technology**. Governments should identify how a specific AI-based technology is being used, or likely to be used in the immediate future. This could be **based on,** for instance, stated **intended uses or overseas experience**."

Feb 2024

A National Policy Roadmap for Artificial Intelligence in Healthcare in Australia

Vision -An AI enabled healthcare system delivering personalised healthcare safely, ethically and sustainably

1. AI ethics and safety
2. Workforce
3. Consumers
4. Industry
5. Research



Adoption rates of AI in health care

- Lack of clinician trust in unexplainable models
- Data privacy concerns from consumers
- Health inequity concerns due to underlying data biases
- Underdeveloped or absent government regulation

Additional anecdotal reflections

  - Digital maturity of health systems
  - Absence of standards for information exchange

Why clinical artificial intelligence is (almost) non-existent in Australian hospitals and how to fix it

Anton van der Vegt[1]
Victoria Campbell[2]
Guido Zuccon[3]

1 Centre for Health Services Research, University of Queensland, Brisbane, QLD.

In-hospital clinical artificial intelligence (AI) encompasses learning algorithms that use real-time electronic medical record (EMR) data to support clinicians in making treatment, prognostic or diagnostic decisions. In the United States, the implementation of hospital-based clinical AI, such as sepsis or deterioration prediction, has accelerated over the past five years,[1] while in Australia, outside of digital imaging-based AI products, nearly all hospitals remain clinical AI-free zones. Some would argue this is a good thing, both prudent and sensibly cautious given the wide ranging ethical, privacy and safety concerns;[2,3] others contend our consumers are missing out on important interventions that save lives and improve care.[4,5] In this perspective article, we argue that in-hospital clinical AI excluding imaging-based products (herein referred to as "clinical AI") can improve care and we examine what is preventing clinical AI uptake in Australia and how to start to remedy it.

Most clinicians know about the failures of the Epic Sepsis Model in the US.[6,7] The AI model missed 67% of septic patients (prevalence 7%) and generated alerts for 18% of all hospitalised patients and, when compared across 15 hospitals (six without the prediction tool), there were no improvements in antibiotic treatment rates or patient outcomes. Other similar high profile failures litter both journals and general news outlets, for example:

- Google's Verily Health Science product to detect diagnostic retinopathy in Thailand, where 21% of images were rejected by the system because of different lighting conditions on site and different patient preparation procedures;[8]
- Stanford's evaluation of skin cancer detection products revealing significant effectiveness drops between light and dark skin patients;[9] and
- IBM's abject failure to get Dr Watson off the ground after spending US$5 billion, with clinicians "wrestling with the technology rather than caring for patients".[10]

Bad news makes for memorable news, but what is missed in the burgeoning number of journal articles on clinical AI[11] is the myriad success stories. In a recent systematic review of implemented sepsis prediction models worldwide, eight systems were installed across more than 40 hospitals; five systems prediction models for stroke, hypertension, venous thromboembolism and appendicitis.[1] Of these, 82% were implemented in the US and none in Australia.

Australian health care seems impervious to the alure of AI. We searched each Australian state and territory's public health care websites and found just two clinical AI stories. Across a network of clinicians in a national AI working group, only one hospital was known to have an AI trial underway. As far as we are aware, there is no clinical AI implemented across Queensland Health despite having Australia's largest centralised EMR system, which could make large-scale AI feasible. In stark contrast to the number of implemented AI systems, AI research abounds, with nearly 10 000 journal articles published each year across the world.[11] Why is clinical AI not translating to clinical practice in Australia?

There are many common reasons cited for a lack of AI uptake within health care, including lack of clinician trust in unexplainable models,[2] data privacy concerns from consumers,[13] health inequity concerns due to underlying data biases,[2,13] and underdeveloped or absent government regulation.[14] But these reasons are similarly applicable worldwide and yet do not prevent other countries from translating AI from the laboratory to their clinical practice. Perhaps the World Health Organization's recent caution can provide a clue, "Precipitous adoption of untested systems could lead to errors by health-care workers, cause harm to patients, erode trust in AI".[15]

What does an "untested system" mean, or, more importantly, what constitutes a tested AI system, that is, one that health authorities would be willing to implement within their hospitals? As far as we are aware, no such framework for the safe introduction of AI into clinical practice exists in Australia. However, it does overseas. To explore this question further, we conducted a scoping review of clinical AI implementation guidelines, standards and frameworks and identified 20 published articles since 2019 from seven countries.[16] We found there were common stages to AI implementation to ensure the safe, effective and equitable introduction of AI into clinical practice. Although these stages vary, they generally always include a stage for problem definition to check that AI is needed and possible (stage I); retrospective (in silico or laboratory)

Areas of impact for AI in healthcare

- **Improving population-health management**
- **Improving operations**
- **Strengthening innovation**

AI enabled patient journey

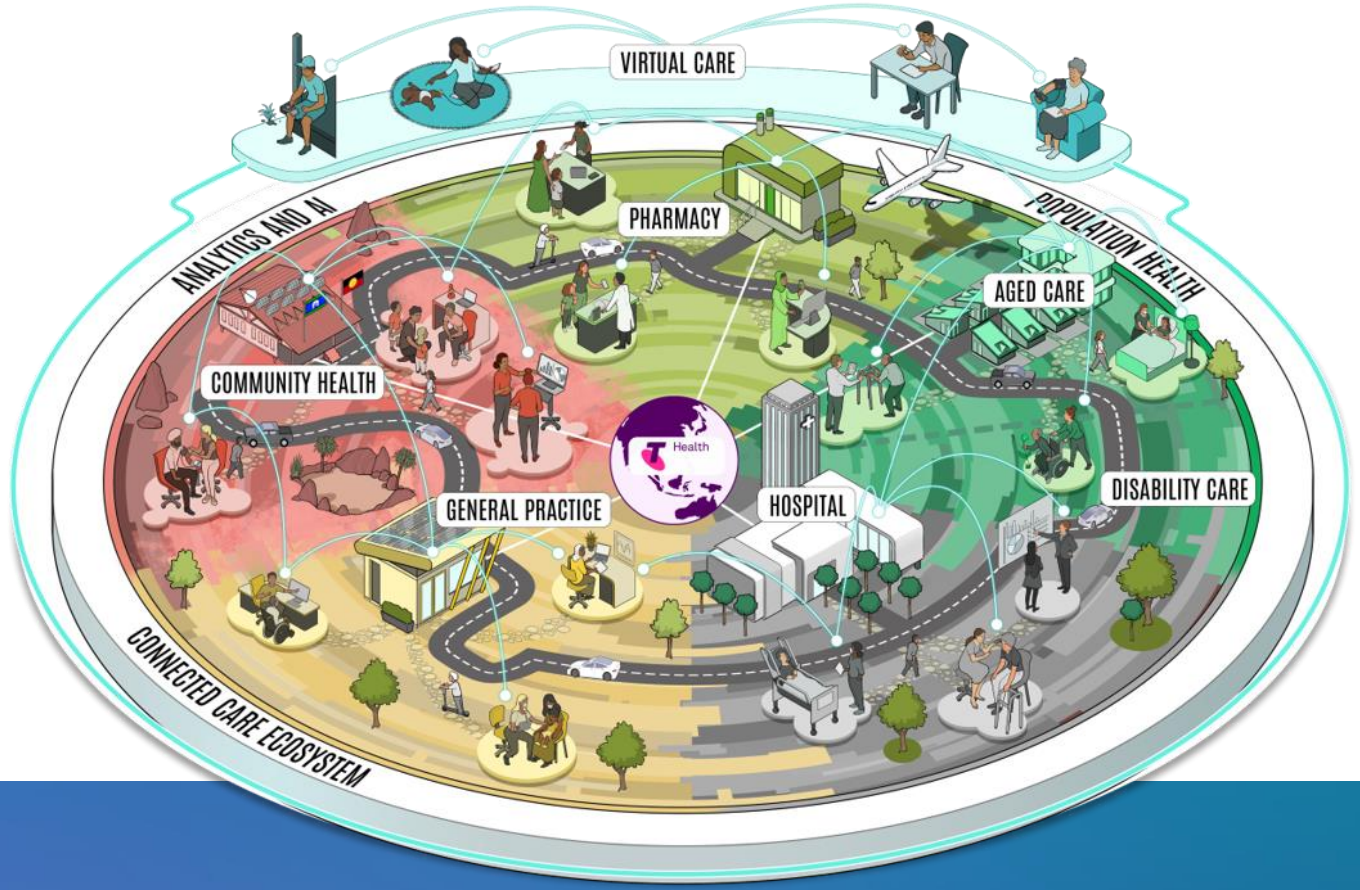
Health

01 AI assisted self-triage and self-management

02 AI assisted pre-hospital care

03 AI assisted acute / hospital care

04 AI assisted rehabilitation / monitoring / community care



Patient Journey 1
AI-assisted self-triage & self-management

Hope

Redesigning Primary Care: The Emergence of Artificial-Intelligence-Driven Symptom Diagnostic Tools

by Christian J. Wiedermann [1,2,*], Angelika Mahlknecht [1], Giuliano Piccoliori [1] and Adolf Engl [1]

[1] Institute of General Practice and Public Health, Claudiana—College of Health Professions, 39100 Bolzano, Italy

[2] Department of Public Health, Medical Decision Making and HTA, University of Health Sciences, Medical Informatics and Technology-Tyrol, 6060 Hall, Austria

* Author to whom correspondence should be addressed.

J. Pers. Med. **2023**, *13*(9), 1379; https://doi.org/10.3390/jpm13091379

Submission received: 4 September 2023 / Revised: 13 September 2023 / Accepted: 14 September 2023 / Published: 15 September 2023

Intended Role & Potential Benefits

- Uses chatbots to engage users in conversations regarding their symptoms.
- Uses AI algorithms to offer potential diagnoses, triage advice, and/or personalised healthcare guidance.
- Assists individuals in healthcare decision-making and navigation, including self-management options
- AI-assisted triage helps prioritise caseloads based on urgency and clinical relevance
- App collects preliminary symptom & demographic data, reducing admin burden for clinicians

Factors Affecting Success:

- Quality of algorithms
- Accuracy of user-input data
- Complexity of medical conditions assessed

Patient Journey 1
AI-assisted self-triage & self-management

Hard reality

Limitations and Challenges:

Evidence indicates that the accuracy of digital symptom checkers for both triage and diagnosis is generally low:

2022 Systematic review of 10 studies involving 48 symptom trackers found that:

- primacy diagnostic accuracy, or listing the correct diagnosis first, ranged from 19% to 38%
- triage accuracy, or sending users to the right level of care, was higher, ranging from 49% to 90%

"As a whole, these issues raise multiple concerns about the use of symptom checkers as patient-facing tools, especially given their increasingly endorsed role within health systems as triage services that direct patients towards appropriate treatment pathways."

Wallace et al 2022

2017: Babylon launches GP at Hand app to address NHS waitlists, offering virtual GP consultations and automating patient inquiries.

Same year, concerns arise over app's diagnostic capabilities (e.g. failure to recognise common symptoms of heart attack), triggering regulatory investigation in the UK.

UK regulator notes concerns but app remains available, highlighting gaps in medical device regulations for eHealth apps.

2018: App launches in Rwanda despite excluding common diseases like tuberculosis or malaria relevant to the region.

2019: UK MPs call for inquiry on Babylon model's impact after £26M funding deficit announced by Hammersmith and Fulham Clinical Commissioning Group caused by excessive 'out of catchment' enrolments

2019: App expands to Canada.

2020: Data breach exposes private video consultation recordings.

2021: **GP at Hand (Babylon UK)** criticised for discouraging registration of older and complex-needs patients

2021: Canadian regulator finds app non-compliant with privacy regulations: identity verification (use of facial recognition tech in tandem with Govt ID), excessive data collection (i.e. recording of video consultations deemed 'beyond what is essential to provide a health service') and inadequate disclosure of data collection & reuse in privacy policy

2023: Company collapses due to financial losses ($284M), regulatory restrictions and reputational damage

"The collapse of Babylon Health is a cautionary tale about government officials, regulators, investors and journalists believing the AI hype in healthcare" ***Jeremy Hsu, Tech Reporter, New Scientist***

Patient Journey 2
AI-assisted pre-hospital care

Health

In the wild:

- Smartwatches equipped with sensors and AI
- emerging positive evidence that AI linked to consumer-grade smartwatches can recognise biometrics

At dispatch:

- AI systems analyse emergency call center content, extracting patient information and automatically transcribing calls.
- AI also being used to translate emergency calls to text in real-time, aiding operators when language barriers exist
- AI technology is being implemented in Europe 4 and Australia[5] to analyse incoming emergency calls, identifying keywords, language, and sound patterns indicative of cardiac arrest, immediately alerting operators for high-priority ambulance dispatch

En route:

- AI offers real-time traffic analysis, efficient route planning to the emergency scene and nearest medical facility.
- AI analyses patient data, suggesting optimal management strategies, alleviating guesswork for paramedics in stressful situations.
- Generative AI systems assist emergency responders in accessing relevant medical protocols, drug treatments, and procedures during field care.

Rewards

10-19-23

AI bots are helping 911 dispatchers with their workload

AI is attempting to revolutionize non-emergency calls in several 911 dispa the country.

Campus Safety — News School

News

How Data-Driven Video Technology is Elevating Air Ambulance Safety

A combination of safety measures and technologies, such as cameras and sensors, can reduce the likelihood that your helicopter ambulances will have an accident.

Risks

The New York Times

Driverless Taxis Blocked Ambulance in Fatal Accident, San Francisco Fire Dept. Says

Two Cruise taxis delayed an ambulance carrying a car accident victim to a hospital, a department report said. The company said it was not at fault.

Patient Journey 3
AI-assisted clinical care: hospital settings

Reducing errors

Limited but emerging positive evidence that AI can support drug decisions and dispensing processes[1] and:

- reduce prescription errors[2]
- identify harmful drug interactions[3,4] identify harmful drug-food interactions[5]
- support drug dosing decisions for high-risk drugs[6]

Diagnosis

- decrease their false positive rates and reduce requested biopsies[7]
- expedite patient triage, radiology reporting, and managing incidental finding follow-ups
- improve radiology report generation and interpretation, overcoming issues like fatigue-induced errors and inconsistencies in expertise inherent in traditional reporting.
- Reduce burnout through automation of routine radiology tasks (eg. image segmentation, measurement and quality evaluation)[8]

Identification of patients at risk

- Falls[7] (and can lead to associated cost savings[9])
- Healthcare associated infections[10]
- A range of adverse events (all-cause mortality, cardiac arrest, transfer to intensive care, and evaluation by rapid response teams)[11]

AI-assisted Clinical Monitoring

Development and Validation:

- Algorithm developed on data from 405,000 patient encounters across three hospital systems from 2013- 2015, internally validated by EPIC.
- Limited public information available on performance due to proprietary nature, with no independent validations published[12]
- Initially achieved high AUC of 0.73 for sepsis detection BUT it was later discovered that this performance data was based on comparing ESM alerts to billing codes for sepsis, rather than the clinical criteria typically utilised by medical researchers[13]

Temporal performance

- Algorithm performance deteriorated over time, with the AUC dropping below 0.60 by the end of 2015, due to changes in input variables, such as updates from ICD-9 to ICD-10 coding system adding new codes affecting sepsis detection, and acquisition of new hospitals like Lahey Health in 2019, leading to demographic shifts compromising accuracy.[14]

External Validation Study at Michigan Medical (2021)[15]

- 67% of patients with sepsis were not identified by the ESM algorithm but were identified by the clinical staff and appropriately treated
- ESM algorithm detected only 7% of sepsis cases missed by clinicians
- False positives by the algorithm meant that eight patients had to be evaluated to identify one patient that truly had sepsis. False positives create needless "busywork" and alarm fatigue
- The study concluded that ESM had poor discrimination and calibration in predicting sepsis onset.

Implications and Recommendations:

- Widespread adoption of ESM without independent validation raised concerns about AI-assisted clinical monitoring of sepsis nationally
- Identified the need for improved clinical governance of AI algorithms both before and after real-world implementation
- Recognition of need for periodic 'calibration' and 'retraining" of AI predictive tools in clinical settings

"The increase and growth in deployment of proprietary models has led to an underbelly of confidential, non-peer-reviewed model performance documents that may not accurately reflect real-world model performance." ***Wong et al***

Patient Journey 3
AI for hospital settings

Analysis and automation of patient flow

Limited but emerging positive evidence that **AI can predict and improve patient flow**:

- for inpatient settings including improving bed capacity, and achievement of discharge targets[1,2]
- for outpatient settings[3]

International Journal of Environmental Research and Public Health

MDPI

Int J Environ Res Public Health. 2022 Aug; 19(15): 9667.
Published online 2022 Aug 5. doi: 10.3390/ijerph19159667

PMCID: PMC9368666
PMID: 35955022

Use of Artificial Intelligence to Manage Patient Flow in Emergency Department during the COVID-19 Pandemic: A Prospective, Single-Center Study

Emilien Arnaud,[1,2] Mahmoud Elbattah,[2,3] Christine Ammirati,[1,4] Gilles Dequen,[2] and Daniel Aiham Ghazali[2,5,*]

Paul B. Tchounwou, Academic Editor

Author information · Article notes · Copyright and License information PMC Disclaimer

Journal Article

AI analysis of system resource allocation

Limited but emerging positive evidence that AI can suggest improvements in system organisation and resourcing:

- during major emergencies:
- allocating emergency services[4]
- determining the optimum plan for resource allocation across multiple hospitals[5]

TECHNOLOGY > AI AND AUTOMATION | February 27, 2023 | updated 26 Jun 2023 4:07pm

NHS turns to AI to solve its 'missed appointments' problem

Patient Journey 3
AI for hospital settings

Clinical Documentation

Generative AI can automate tasks and free up clinician time:

- Generative AI can transcribe patient consults then summarise these into clinical notes
  - medical staff in hospitals where this software has been piloted report spending up to 66% less time on report writing.2
  - roughly 44% of administrative tasks carried out by staff in general practice are ‘mostly’ or ‘completely’ automatable using currently available technology.3
  - Currently no comprehensive standards for LLM-generated clinical summaries beyond the general criteria that summaries should be consistently **accurate** and **concise**
  - **Risks**
  - **confirmation bias** that could increase **diagnostic error**
  - Some prompts (eg, “summarise my patient’s history relevant to risk of heart failure”) cause LLMs to effectively function as medical device

Automation of administrative tasks

Rostering

Result in much faster creation of staff rosters which are fairer, include multiple combinations for redundancy and reduce staff burnout.4

Billing and RCM

Optimise revenue cycle costs by retrieving, matching, standardizing, federating, and transforming billing and collections data from disparate systems,

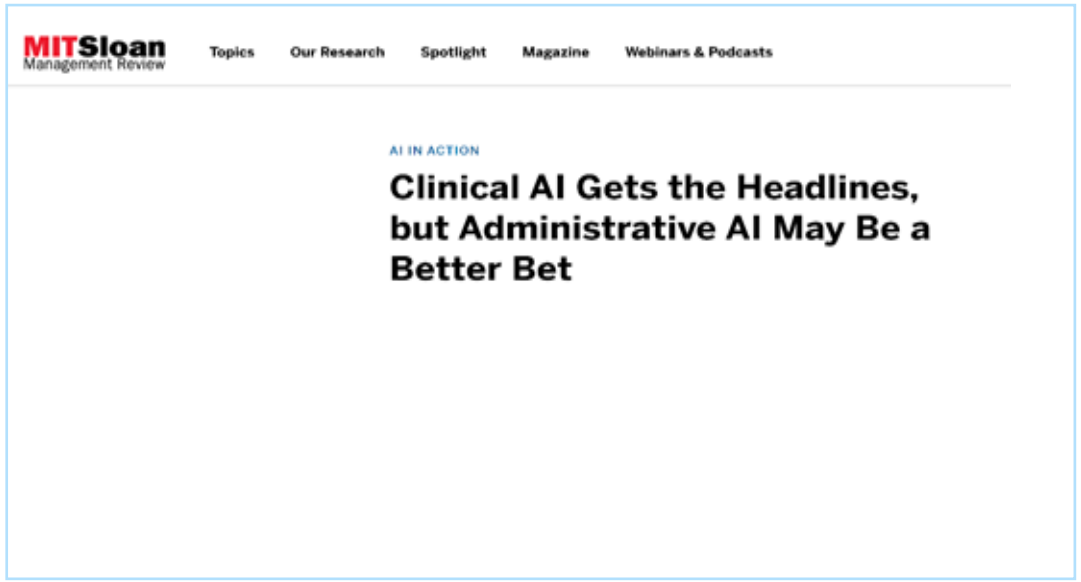
Coding

enhance manual coding accuracy, quality, and efficiency BUT challenges remain due to limited expert-labelled Electronic Health Record synthetic datasets needed for model training and under-coding issues in widely used MIMIC-III dataset, impacting model accuracy and real-world generalization6.

“

doi:10.1001/jama.2024.0555

Clinical AI Vs Administrative AI


MITSloan Management Review

Topics Our Research Spotlight Magazine Webinars & Podcasts

AI IN ACTION

Clinical AI Gets the Headlines, but Administrative AI May Be a Better Bet

"No one will win the Nobel Prize in medicine for applying AI to health care administration. Accolades and much of the media and public attention will go toward clinical applications of the technology [but]…the challenges and cycle times for developing and implementing those advances mean that many organizations will want to strongly consider administrative AI as well. If that type of AI can substantially reduce the cost of care, it could be as useful to the health care system overall – and many patients individually – as any clinical breakthrough"

Davenport et al. 2022

AI and Population Health: Breast Screening

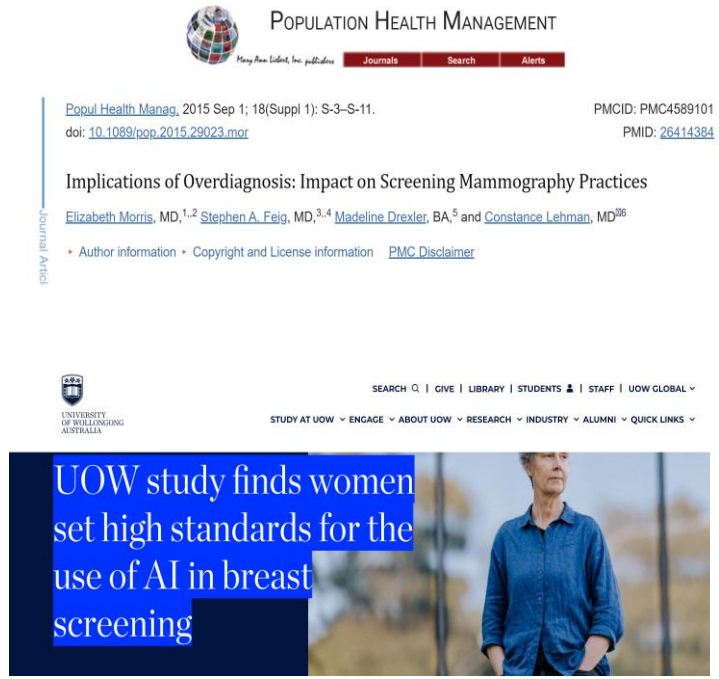
Promising AI innovations

Enhance predictive risk models by combining traditional risk models with imaging features, increasing accuracy.

- Yala et al.'s study showed that a hybrid DL model, integrating traditional risk factors with mammogram analysis, outperformed the IBIS/TC model alone.
  - It identified 31% of patients in the top risk category compared to 18% by the IBIS/TC model and identifying new features associated with long-term risk beyond early disease detection.[6]

Consumer expectations

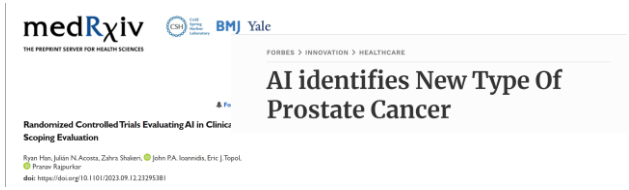
- University of Wollongong researchers conducted eight online discussion groups with 50 women engaged in breast screening to gauge their attitudes towards AI in mammogram interpretation.[7]
- Despite overall positivity and excitement about AI, trust was contingent on two factors: the integration of human oversight and the performance of the AI.
- Women stressed the necessity for AI systems to exhibit excellent performance with an **expectation that AI systems perform better than the status quo**, especially in management of a high-risk condition
- Ongoing monitoring and quality assurance were deemed essential to maintain trust and ensure AI effectiveness in breast screening



Population Health: Potential of AI in Clinical Trials

Medical data is expected to **double up to 5 times per year**. A typical Phase III trial already generates around **3.56 million data points**[1]

2023 systematic review RCTs involving AI found that 82.1% (69 out of 84) of the trials documented favorable outcomes[2]



"This data breach challenges how we think about privacy, data security and corporate accountability in the information economy"

Other areas of impact:

Regulatory Documents: AI algorithms may expedite the Investigational New Drug application process, accelerating the introduction of new drugs and therapies.

Protocol Generation: Generative AI language programs can rapidly create first drafts of clinical protocols using inputs from published literature, previous trials and multiple medical sources.

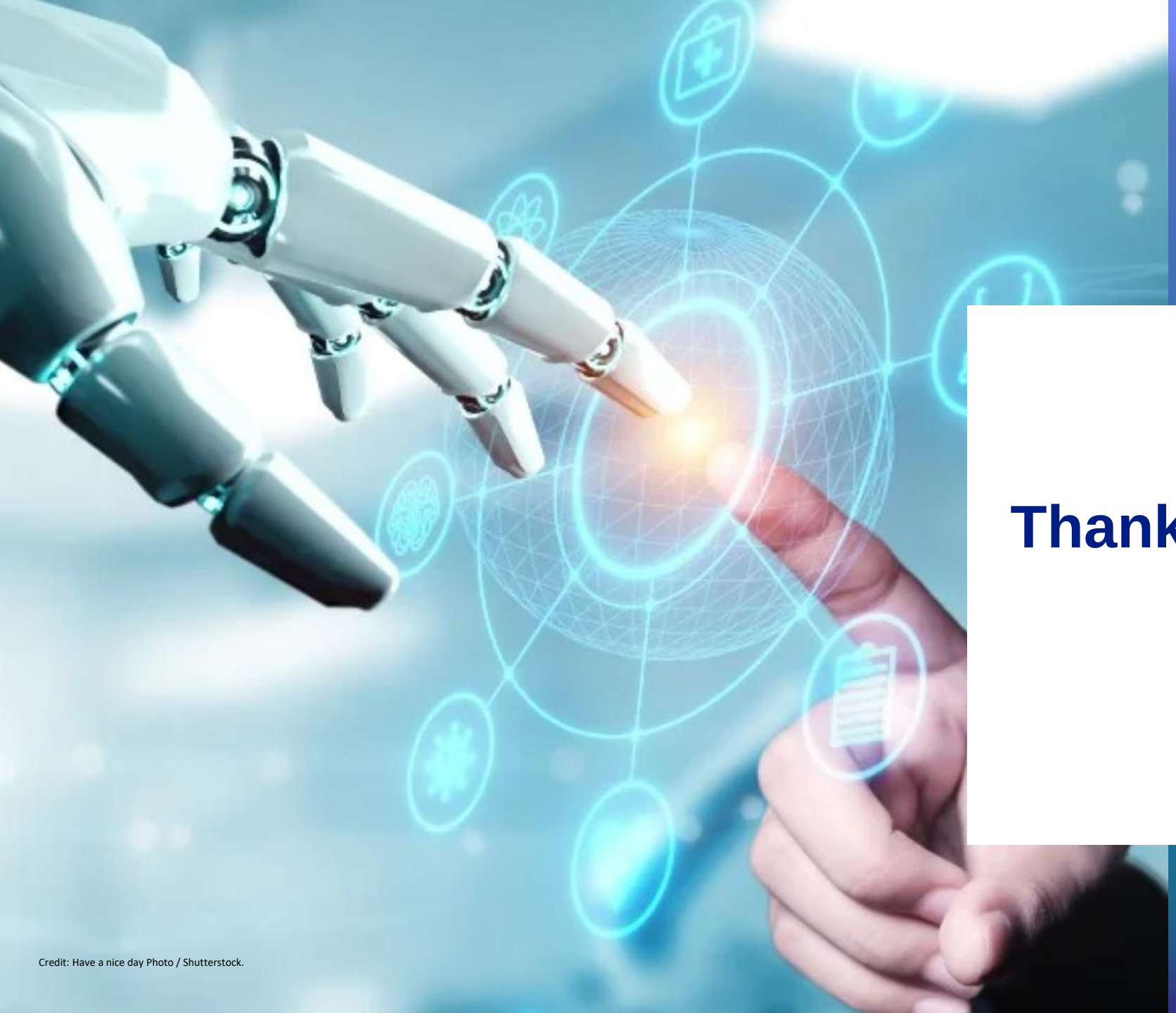
Patient/Site Selection & Matching: AI can match patients to clinical trials more effectively, improving recruitment and trial success rates.

Safety Signal Prediction: Predictive AI models can anticipate safety concerns, potentially averting adverse events before they occur.

Digital Twins: Virtual images of patients, created and monitored ay AI, can provide real-time insights into individual health, facilitating personalised treatments. There are applications in both clinical trials as well as drug manufacturing to help predict biological responses based on biomarkers[1]

Reflections on Use of AI in healthcare

- **Genuinely does have the ability to transform healthcare**
- **3 phases implementation**
  - Administrative tasks
  - Shift from hospital based to home-based sensors and monitoring for a larger range of conditions
  - AI in clinical practice
- **Ethical Considerations should be paramount**
  - Data privacy and ethical use of patient information are critical.
  - Balancing innovation with patient safety and privacy is essential
  - Trust is the key



Credit: Have a nice day Photo / Shutterstock.

Health

Thank you