

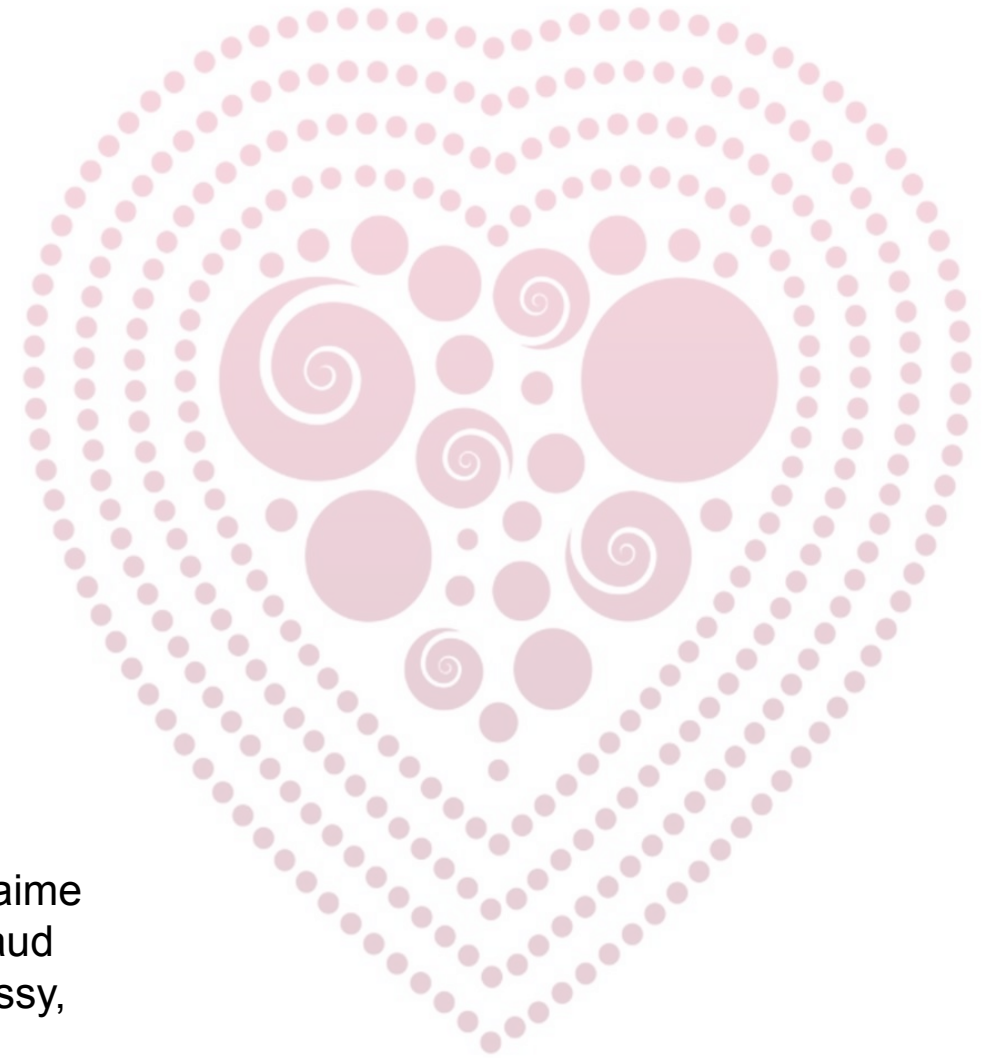
Multilingual Patient Healthy Coach for Cardio-oncology Patients: Leveraging Artificial Intelligence for Cancer Equity in Australia - A Pilot Study

Tu Hao Tran, BMedSci (Hons), MBBS, FRACP

PhD Candidate | Imaging and Phenotyping Lab | University of Sydney
Innovation and Electrophysiology Fellow | Campbelltown Hospital, Sydney
AI and Digital Innovation Fellow for | Australian Centre for Cancer Equity
Ingham Institute of Applied Medical Research (Sydney)

Contact: hao@grievelab.com

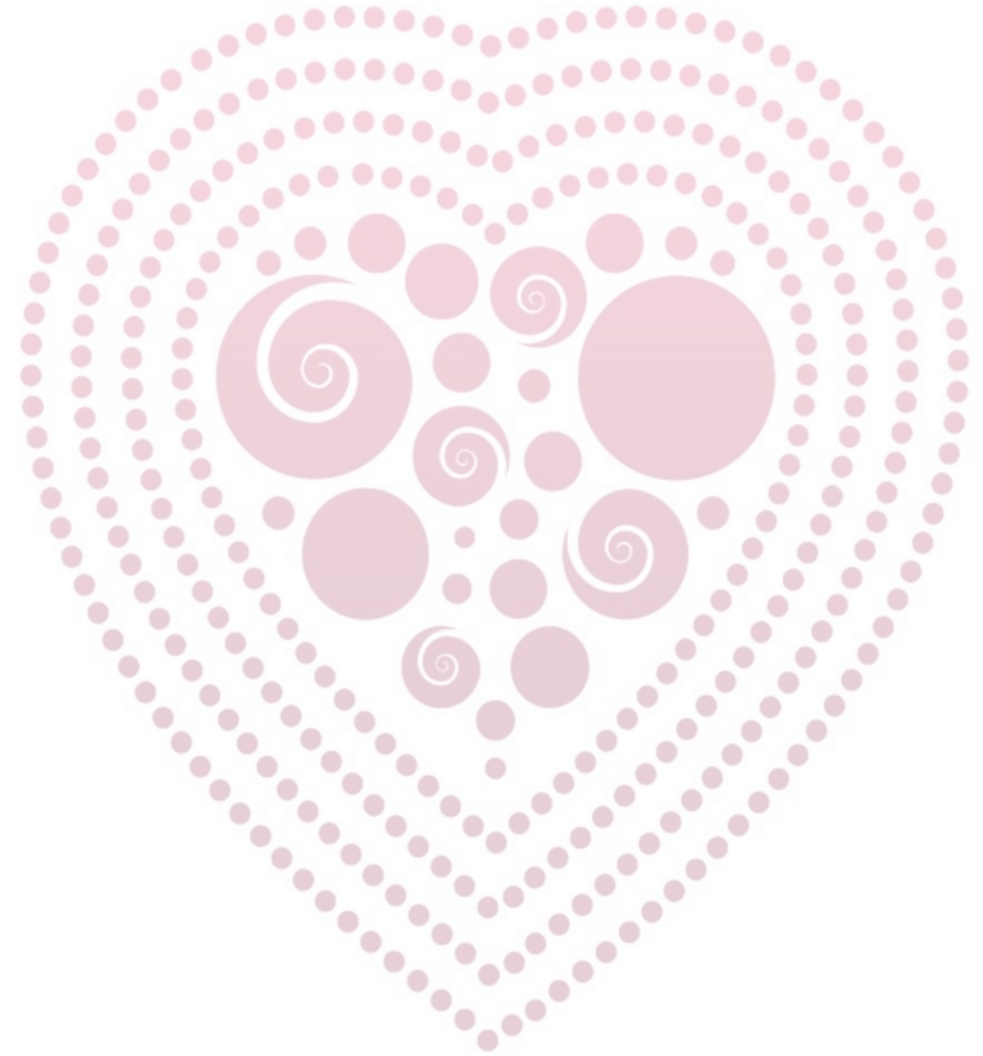
Co-authors: Kazuaki Negishi, Wei Chua, Bernadette Brady, Viren Joseph, Jaime Baquaran, Upul Premawardhana, Nagendraprasad Sungala, Sidney Lo, Arnaud Mazier, Avanti Shrikumar, Tuan Nguyen, Zaidon Al-Falahi, Annemarie Hennessy, Gemma Figtree, **Stuart M Grieve**



Faculty disclosure

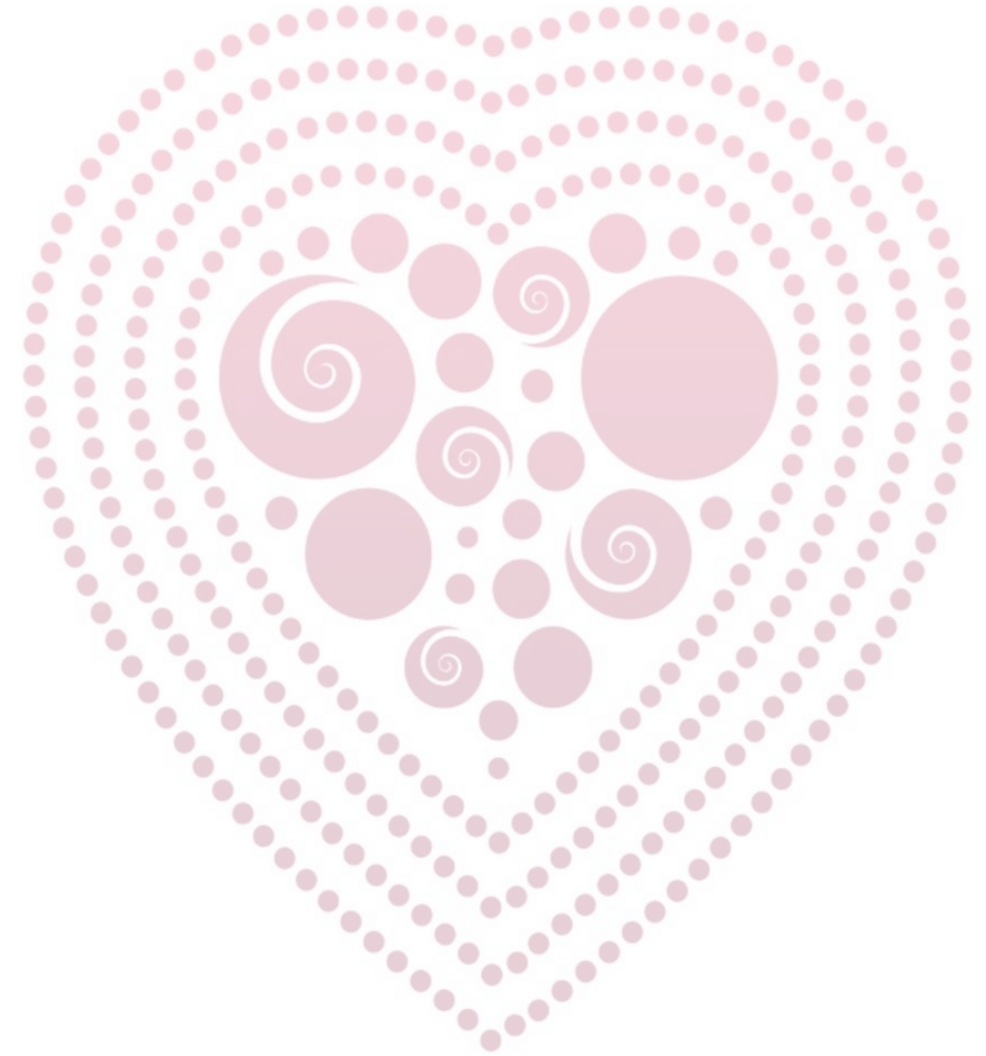
No relevant disclosures.

The presenter has advised that the following presentation will NOT include discussion on any commercial products or service and that there are NO financial interests or relationships with any of the Commercial Supporters of this years ASM.



Background

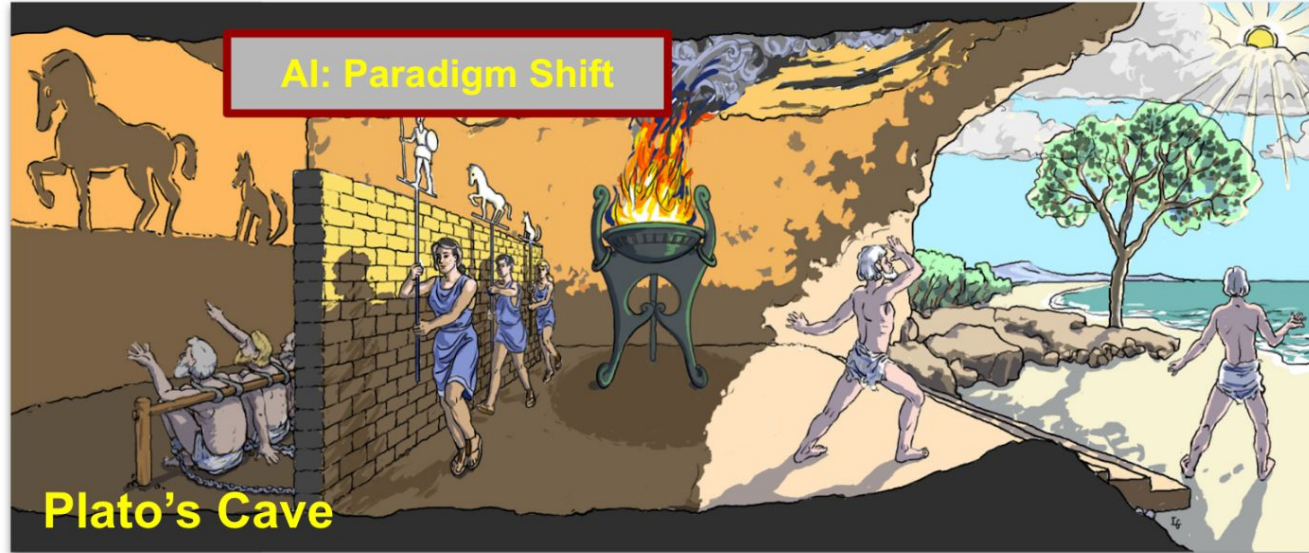
- Cardio-oncology (CO) patients benefit from multidisciplinary care.
- Linguistic barriers affect health access equity.
- Artificial intelligence (AI), specifically Large Language Models (LLM), may assist patients with translation and education.
- The veracity of LLMs for medical translation needs careful human/physician adjudication.



Go backwards to go forwards...

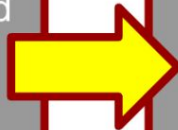


AI: Aspiration / Dance



LLM

- **Probabilistic** (probability distribution of tokens): training and token prediction from data
- **Stochastic** (random sampling of probability distribution): for text generation during execution



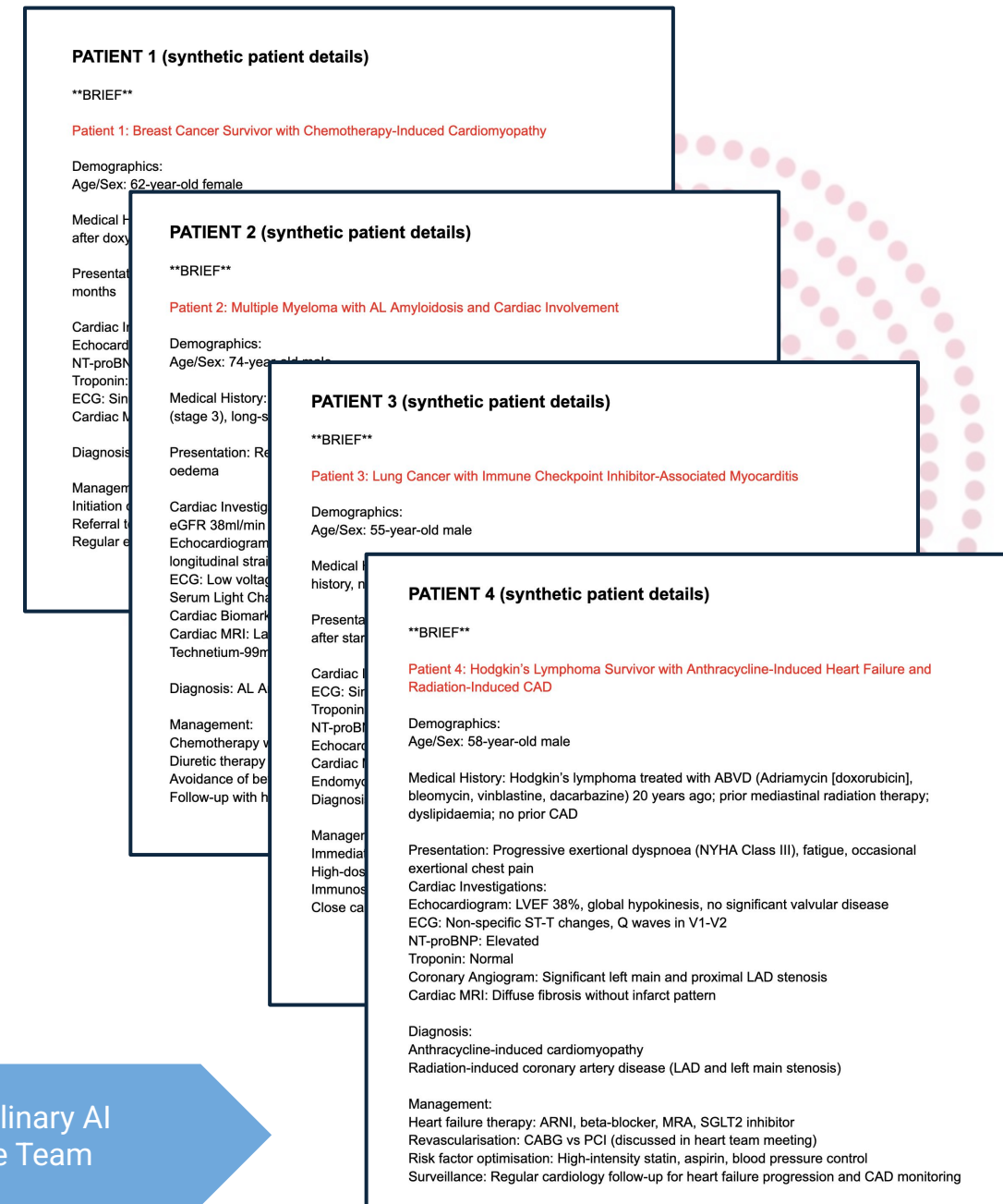
Deterministic Tools

- **Parroting Models** (multi-head self attention, multi-head latent attention) to **Reasoning Models** (chain of thought, long-range context coherence, multi-agent systems)
- **Control of Stochastic Randomness:** prompt engineering, temperature, Top-k sampling, Top-p (nucleus) sampling

Method

4 '**synthetic**' cardio-oncology patient archetypes were generated by human & LLMs:

1. 62F with **HER-2 positive breast cancer** and **trastuzumab-induced cardiomyopathy** (LVEF 40%, diffuse fibrosis) on GDMT
2. 74M with **IgG lambda multiple myeloma** and **AL amyloidosis** (lambda, CKD3) with **cardiac involvement** (LVH, apical sparing, LGE, positive PYP) on daratumumab-based chemotherapy
3. 55M with **stage 4 non-small cell lung cancer** and **pembrolizumab-induced myocarditis** (↑troponin, LVEF 50%, MRI confirmed) treated with steroids and immunosuppression.
4. 58M with **Hodgkin's lymphoma** and **anthracycline-induced cardiomyopathy** (LVEF 38%, diffuse fibrosis) and **radiation-induced CAD** (LM/LAD stenosis) managed with GDMT and for revascularization (CABG vs PCI)



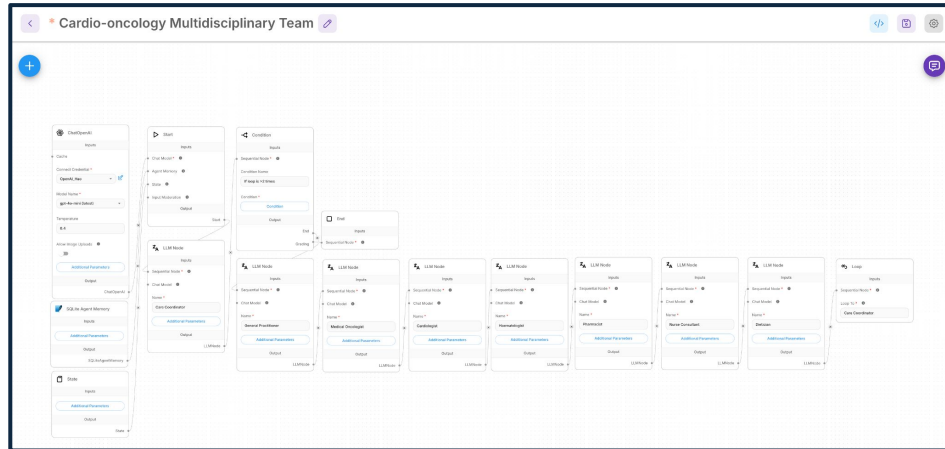
Patient Medical Information

Multi-agent Models generate
AI recommendations

Multi-disciplinary AI
Healthcare Team

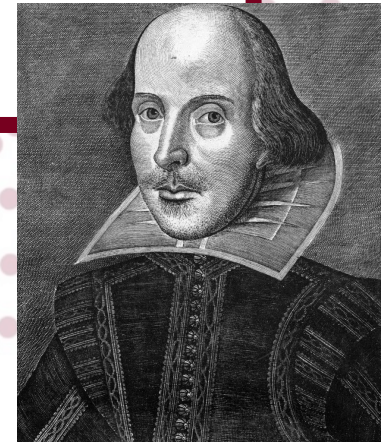
Method

A **multi-agent AI architecture** was created to emulate a **Cardio-oncology Multidisciplinary Team** which generated **advice** for the patient archetypes.



8 personifications:

- Care Coordinator
- General Practitioner
- Medical Oncologist
- Haematologist
- Cardiologist
- Pharmacist
- Nurse Consultant
- Dietician



Patient Medical Information

Multi-agent Models generate AI recommendations

Multi-disciplinary AI Healthcare Team

Method

AI Patient Health Coach chatbots were created for each archetype with prompt engineering.

Cardio-oncology Patient Health Coach 1 (Lien Nguyen)



HER-2 positive breast cancer
Trastuzumab-induced cardiomyopathy

Cardio-oncology Patient Health Coach 2 (Omar Khaled Al-Farsi)



IgG lambda multiple myeloma
Cardiac amyloidosis

Cardio-oncology Patient Health Coach 3 (John Matthew Anderson)



Stage 4 non-small cell lung cancer
Pembrolizumab-induced myocarditis

Cardio-oncology Patient Health Coach 4 (Wei Zhang)



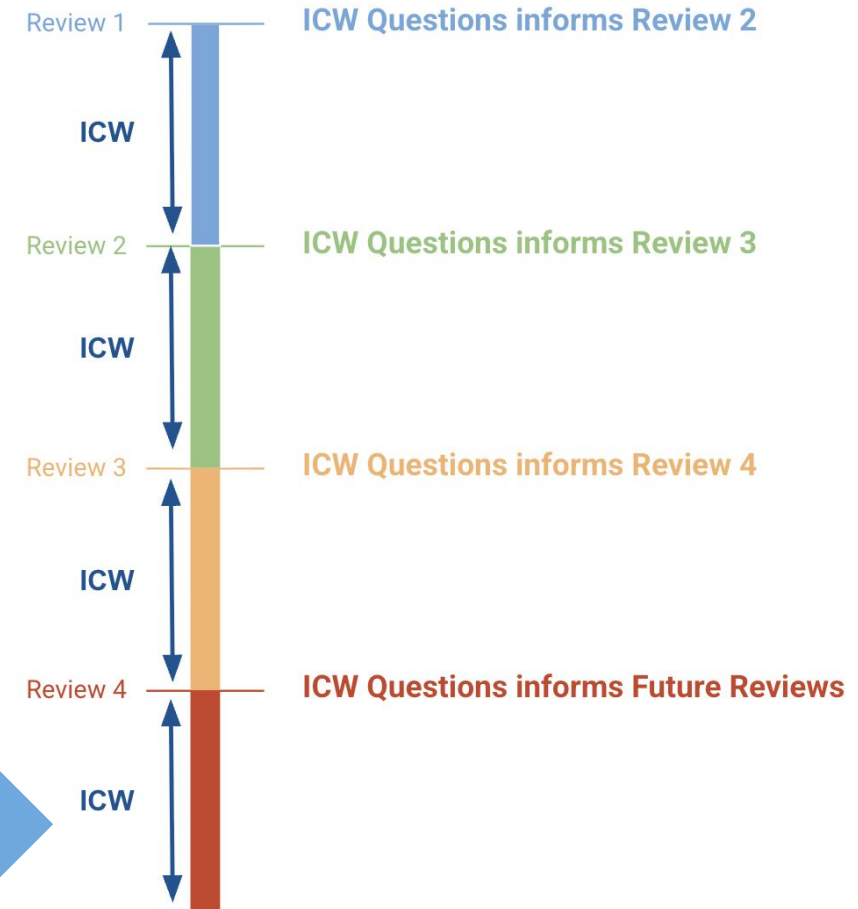
Hodgkin's lymphoma
Anthracycline cardiomyopathy
Radiation-induced CAD

Patient Medical Information

Multi-agent Models generate
AI recommendations

Multi-disciplinary AI
Healthcare Team

Inter-Clinical Window (ICW)



Method

- AI Patient Health Coach chatbots were asked 3 standardised questions:
 - Tell me about my condition
 - Explain my medications to me
 - Is my condition serious, should I be worried?
- Multilingual outputs were assessed by 7 physicians (one for each language, language-congruent)
- Kurskal-Wallis test for significant differences within archetypes, language and score categories (small sample size)

Lien Nguyen - PHC Chatbot OUTPUT in your language

Q1: Tell me about my condition

හලේගේ ලියන! ඔබට ඇති තත්වය යනු කාන්තා රෝගීන් සඳහා ප්රතිකාරයක් ලෙස ලබා දෙන රසායනික ප්රතිකාරයකින් පසු ඇතිවන හෘද රෝගයක් වන "රසායනික ප්රතිකාරය හේතුවෙන් ඇතිවන හෘද පීඩාව" (chemotherapy-induced cardiomyopathy) වේ. මෙය වෛෂයෙන් ඇත්නුසයිකලීන් සහ වරාසුසුමැබ් (trastuzumab) වැනි ප්රතිකාර ලබා දීමෙන් පසු සිදුවන හෘද රෝගයක් වේ.

ඔබගේ හෘදය වෛෂික ලෙස කාර්ය සාධනයට බලපා ඇති බව පෙන්වන ලද පරිදි, ඔබගේ වමතික වාතනු (left ventricular ejection fraction - LVEF) 40% ක් වන අතර, මෙය සාමාන්ය පරාසයට වඩා අඩු වේ. මෙය ඔබට කුඩා කාර්ය සාධනයකදී (dyspnoea on exertion) හෝ නිදා සිටියදී (orthopnoea) හෝ අනුපූර්ණ වූ විට (oedema) වැනි ලක්ෂණ ඇති විය හැක.

ඔබගේ සෞඛ්‍ය කණ්ඩායම ඔබගේ තත්වය නිරීක්ෂණය කරමින් සිටි සහ ඔබට හෘද පීඩාව කළමනාකරණය කිරීමට උපකාරී වන කිපයක්ම මාෂඩ් ලබා දී ඇත. මෙහිදී ARNI (sacubitril/valsartan), beta-blocker, mineralocorticoid receptor antagonist (MRA), සහ SGLT2 inhibitor ඇතුළුව වේ. ඔවුන් ඔබගේ හෘද කාර්ය සාධනය නිරීක්ෂණය කිරීමට සාමාන්යයෙන් එකතෝකාර්ඩියෝග්‍රෑම් සහ උදර පරික්ෂණ සිදු කරනු ඇත.

ඔබගේ ප්රතිකාර සැලැස්ම අනුගමනය කිරීම සහ අඩු සෞඛ්යමි ආහාරයක් සහ ඔබගේ හෘදය සඳහා සුදුසු සාමාන්ය ක්රියාකාරකම් කර ගැනීම වඩාත් වැදගත් වේ. ඔබට ඔබගේ තත්වය පිළිබඳ කිසියම් වෛෂික ප්රශ්න හෝ සැකයන් ඇත්නම්, කුරුණාගර අසන්න!

ආරක්ෂිතව සිටින්න!
AI Cardiologist

Sinhalese

Method

- Language-congruent Physician adjudication of LLM output:
 - 7 Languages (South-West Sydney): English, Sinhalese, Japanese, Chinese, Vietnamese, Hindi, Arabic
 - 6 categories: Accuracy, Completeness, Empathy, Cultural Sensitivity, Clarity, Relevance
 - 5 grades:
 - 1 = Poor
 - 2 = Below Acceptable
 - 3 = Acceptance
 - 4 = Acceptable
 - 5 = Excellent

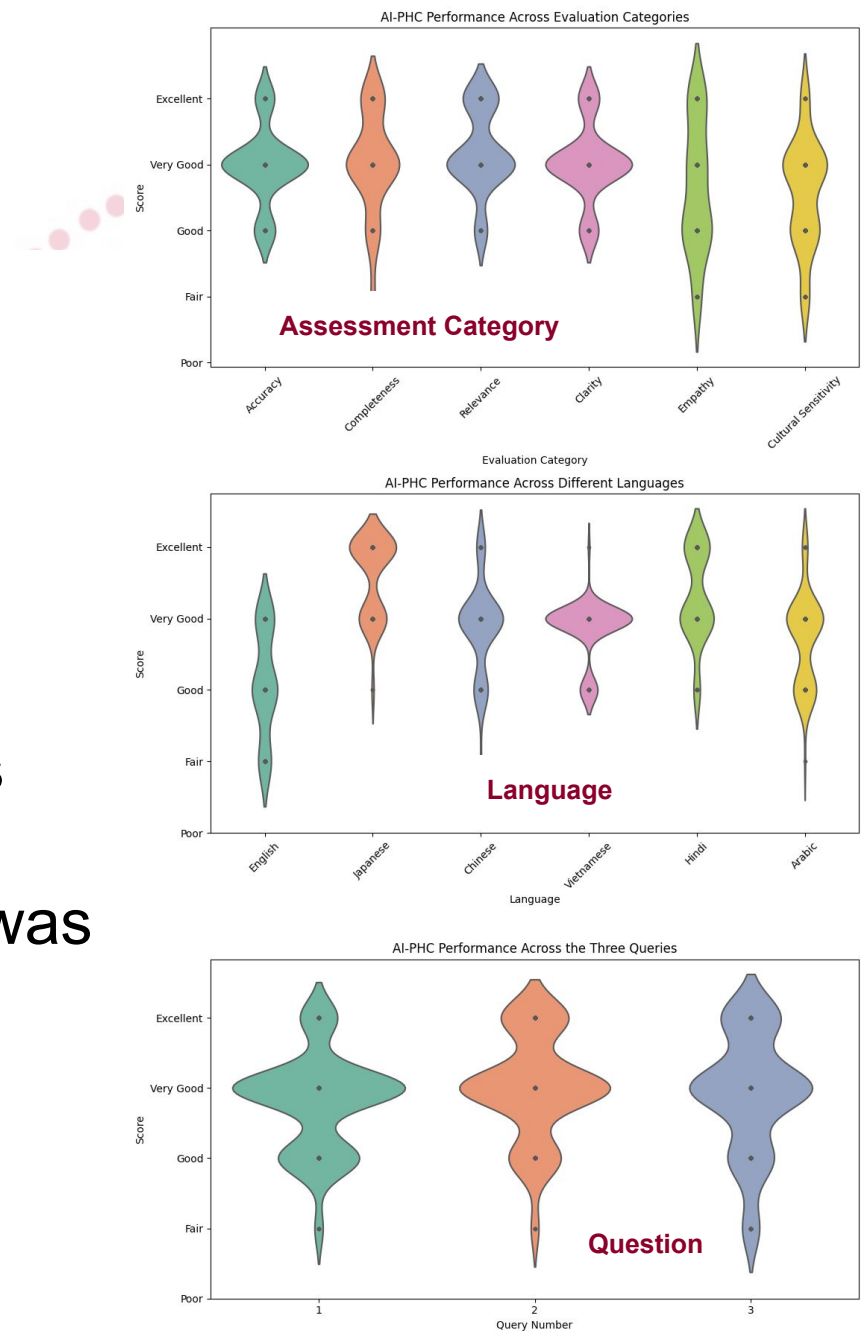
Lien Nguyen Q1 Score Card

Category of Grade	1 Poor	2 Below Acceptable	3 Acceptable	4 Acceptable	5 Excellent	Insert Grade Here
Accuracy	Major inaccuracies	Several errors	Mostly correct	Nearly flawless	Fully accurate	
Completeness	Critical info missing	Many gaps	Mostly complete	Nearly all aspects covered	All necessary details provided	
Empathy	Dismissive or insensitive	Minimally empathetic	Moderately empathetic	Very empathetic	Compassionate and supportive	
Cultural Sensitivity	Offensive or culturally inappropriate	Some cultural missteps	Generally appropriate	Respectful	Fully adapted to audience	
Clarity	Very confusing	Hard to understand	Some structure issues	Well-organized	Logical and easy to read	
Relevance	Off-topic; does not address the inquiry	Largely irrelevant	Partially relevant	Mostly relevant	Fully addresses the inquiry	

Sinhalese

Results

- High variability between Languages and Evaluation Categories and greater consistency in the Queries (Query 3 has highest variance)
- Mean output score was 3.7 +/- 0.8 (out of 5)
- Significant variance was observed in scores across language and output score category ($p < 0.001$ Kruskal-Wallis), whilst the patient query domain was more consistent ($p = 0.043$ Kruskal-Wallis).
- Sinhalese language was omitted as chatbot output was 'incomprehensible', phonetic translation only.



Conclusion

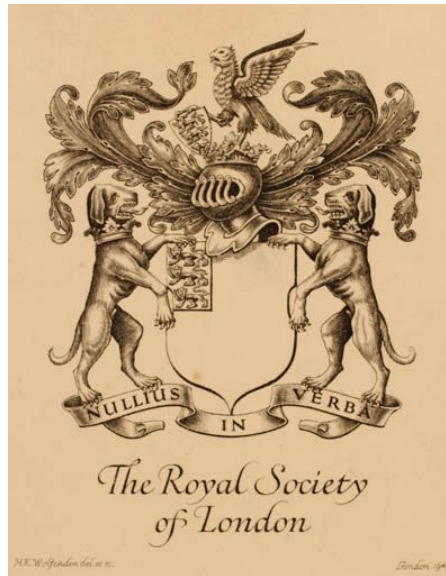
- Overall performance was satisfactory (3.7 of 5)
- Significant heterogeneity exists between languages and score categories.
- Differences in language may be confounded by differences in the scorer, need larger number of physician for each language to elucidate.
- There was no bias for English in performance.
- The relative performance consistency with queries was slightly less robust when conveying a more subjective message of disease severity.
- Sinhalese appears to be not represented in the LLM.
- The use of LLMs for patient use requires further evaluation for safety and utility - proof of concept only, need a larger study.

Old is new and new is old...

Primum Non Nocere

Sidere Mens Eadem Mutato

Nullius In Verba



Thank you.

