

Health Services & Outcomes Research

2024 ANNUAL REPORT





OUR VISION

To add years of healthy life to the people of Singapore through excellence in Health Services Research.

OUR MISSION

To improve the value of healthcare by providing best available evidence for decision making and knowledge translation; and build capacity and advance knowledge in Health Services Research.

FOREWORD

The year 2024 marked a pivotal expansion in our health services research capabilities, characterized by the increased integration of Artificial Intelligence (AI) methodologies and a growing expertise in social and behavioural sciences. This interdisciplinary approach enhances our capacity in preventive and population health, ultimately improving health outcomes, optimizing care processes, and elevating service delivery. Working closely with our partners across the National Healthcare Group institutions, we have made meaningful progress towards the cluster's ambitious vision of adding 15 million years of healthy life. The contributions of our partners and our collaborations with them have been instrumental to these achievements.

In this Annual Report, we present our work in five key thematic areas. The first two showcase our department's significant advancements in harnessing generative AI to elevate healthcare analytics, support health services research, and enhance health interventions. The remaining three themes focus on our initiatives in leveraging health services knowledge to understand the factors influencing patient outcomes, analyse population trends, and evaluate the effectiveness of health services.

Building on last year's exploration of generative AI, this report highlights its applications across various domains, from healthcare system dashboards to clinical notes analysis and predictive modelling. The potential of AI in bolstering research and increasing productivity is emphasized through evidence reviews, synthesis, and the transcription and analysis of qualitative study interviews, all of which have demonstrated significant time savings and productivity enhancements. In enhancing healthcare interventions, we showcase the safety and utility of AI-driven screening tools and predictive models, along with novel patient-centric solutions such as a chatbot for lung function tests and AI-assisted diabetic foot screening. These innovations have illustrated the potential for improving clinical decision support, increasing efficiency, and predicting patient outcomes.

Nonetheless, health services research remains our foundational work, providing insights that transcend beyond the capabilities of generative AI. We have deepened our understanding of patients and the factors that drive their health outcomes, which informs both clinical decisions and intervention strategies. Our research includes studies on patients undergoing knee replacements and those facing end-stage organ failure. On a population level, we measured the health span of residents and explored the relationships between educational characteristics, psychosocial factors, and exercise profiles with health outcomes. Additionally, we developed and validated a new tool to assess health resilience among community-dwelling adults, facilitating future efforts to understand its role in our residents' health.

At the health system level, our work has guided resource allocation through specialist manpower planning reviews, as well as evaluations of health interventions and service delivery. The strategic integration of eye care into primary care settings has resulted in high-quality care, while we established the cost-effectiveness of ventilation for obstructive sleep apnoea. The importance of timely identification of care needs has also been underscored in a post-stroke care coordination program and the deployment of a prognostic model for advanced dementia.

As we navigate a rapidly transforming landscape, HSOR is strategically positioned at the intersection of artificial intelligence and health services research. We remain committed to our mission of driving evidence-based decision-making and knowledge translation, continually adapting our toolkits to meet the evolving demands of our healthcare environment. We hope you find the articles in this report both inspiring and insightful for your work.



Prof John Abisheganaden
Director and Senior Consultant, HSOR
National Healthcare Group



Prof Jason Cheah
Deputy Group Chief Executive Officer
(Strategy, Planning and Resourcing)
National Healthcare Group

CONTENTS



Harnessing AI in Healthcare Analytics and Research

- 08** Using a Large Language Model to Analyse Routine Dashboard and Generate Commentary
- 10** A Hybrid Approach Using Math and Large Language Models to Analyse Dashboard Data
- 12** Combining Generative AI and Machine Learning to Improve Prediction of High Healthcare Utilisers
- 14** Harvesting the Wealth of Unstructured Data: CogStack Exploration for Clinical Notes
- 16** Multi-agent AI – An Exploration on Applications
- 18** Using Artificial Intelligence Tools to Support Systematic Reviews – An Overview
- 20** Exploring the Use of ChatGPT for Risk of Bias Assessment in Prognostic Studies
- 22** Exploring Thematic Analysis Processes with Large Language Model
- 24** Transcription in Healthcare: Comparing the Accuracy and Speed of *Whisper* and *Scribe*



Enhancing Health Interventions with AI

- 28** Smart Prep: Enhancing Lung Function Testing with AI
- 30** Enhancing the Patient Journey for Lung Function Tests Using Large Language Model-Powered Chatbots
- 32** AI-assisted Pharmacy Intervention Classification
- 34** Cost-Effectiveness Analysis of AI-enhanced Screening for Diabetic Foot Ulcers
- 36** Predictive Risk Scores for 30-day Readmission



Examining Drivers of Patient Outcomes

- 40** Factors Associated with Longer Length of Stay in Patients Undergoing Primary Total Knee Replacement at a General Hospital in Singapore
- 42** Identifying Groups of Healthcare Expenditure Trajectories in Decedents with End-stage Organ Diseases
- 44** The Impact of Methodologies on Identifying Multimorbidity Clusters in Heart Failure: A Systematic Review
- 46** NHG Knee Arthroscopy Patients: 10-year Trends and Factors Associated with Longer Stay Among Older Patients



Identifying Population Trends and Determinants of Health

- 50** Population Projection for Lung Cancer Screening
- 52** Healthspan of NHG Residents: A Comprehensive Study from 2016 to 2023
- 54** Associations between OGTT Curve Morphology and Physiological, Metabolic, and Exercise Profiles in Healthy Adults: A Cross-sectional Analysis
- 56** Associations of Educational Attainment and Readiness-to-learn with Self-reported Health
- 58** Associations between Psychosocial Factors and Health-Related Quality of Life in Community-dwelling Adults: A 1-year Longitudinal Study
- 60** Validating A Self-report Scale for Assessing Health Resilience in The Community-dwelling Adults: A Cross-sectional Study

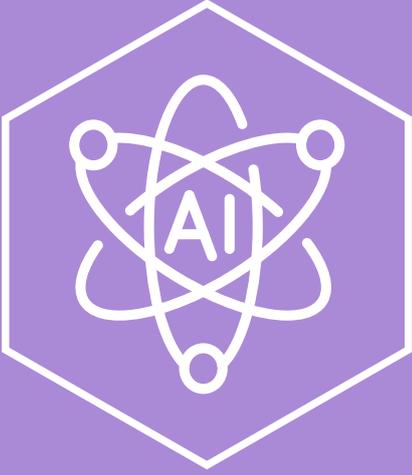


Evaluating Health Systems and Service Delivery

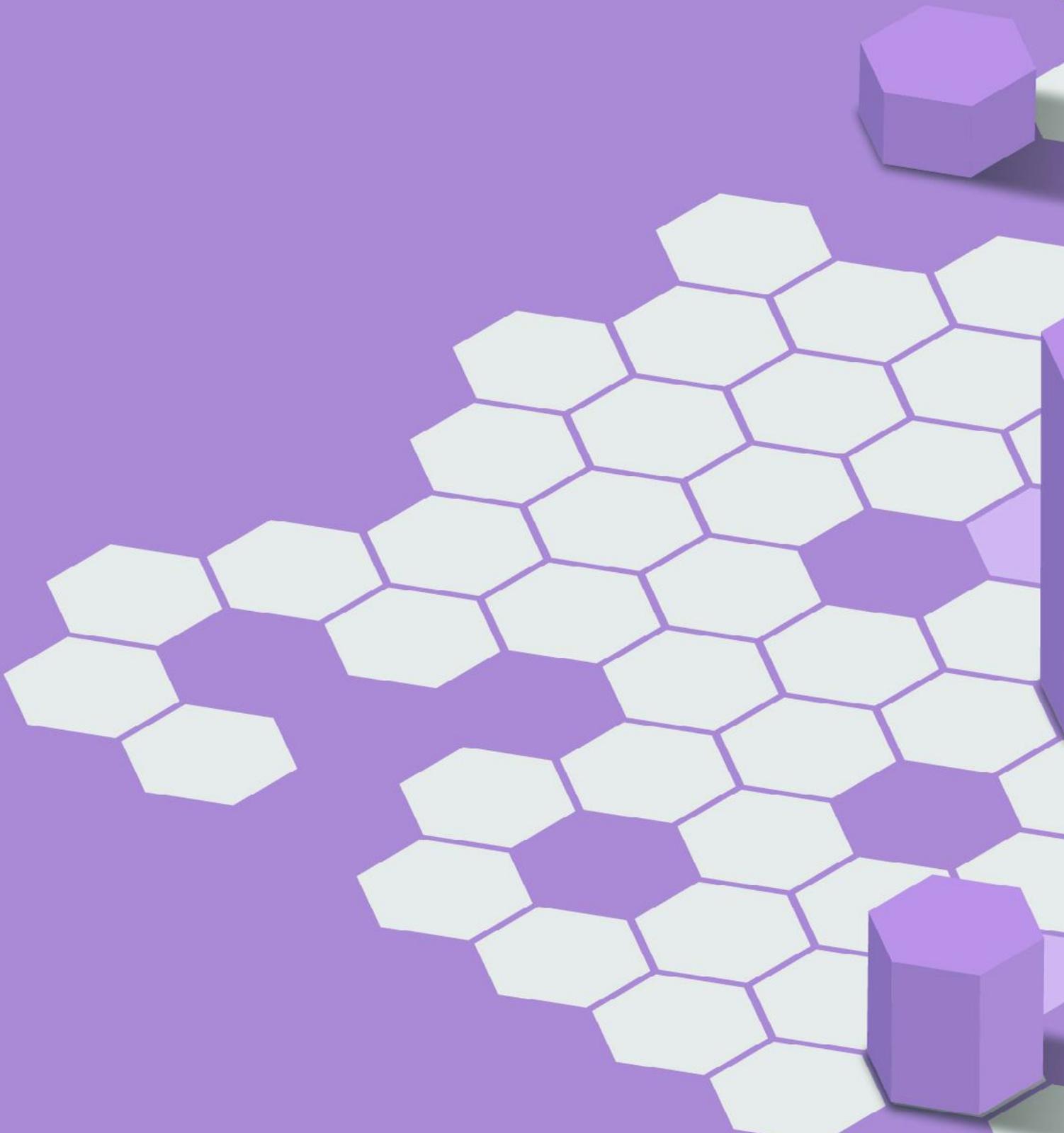
- 64** Specialist Manpower Planning: Projecting Healthcare Workforce Needs
- 66** Impact of Vaccination Strategies on Pneumonia Hospitalisation Rates in Older Adults Aged 65 and Older: A Multi-Variable Big Data Study of Influenza and Pneumococcal Vaccines
- 68** Cost-effectiveness Analysis of Continuous Positive Airway Pressure Therapy for Obstructive Sleep Apnea in Singapore from a Health System Perspective
- 70** Enhancing Post-stroke Recovery and Return-to-Work: Insights from the SLS CHARTER Project
- 72** From Prediction to Palliation: A Case Study of a Prognostic Tool Implementation in a Transmural Model of Early Palliative Care for Advanced Dementia
- 74** Key Drivers of High-quality Care in the Primary Eye Care Model: Insights from a Mixed-methods Study

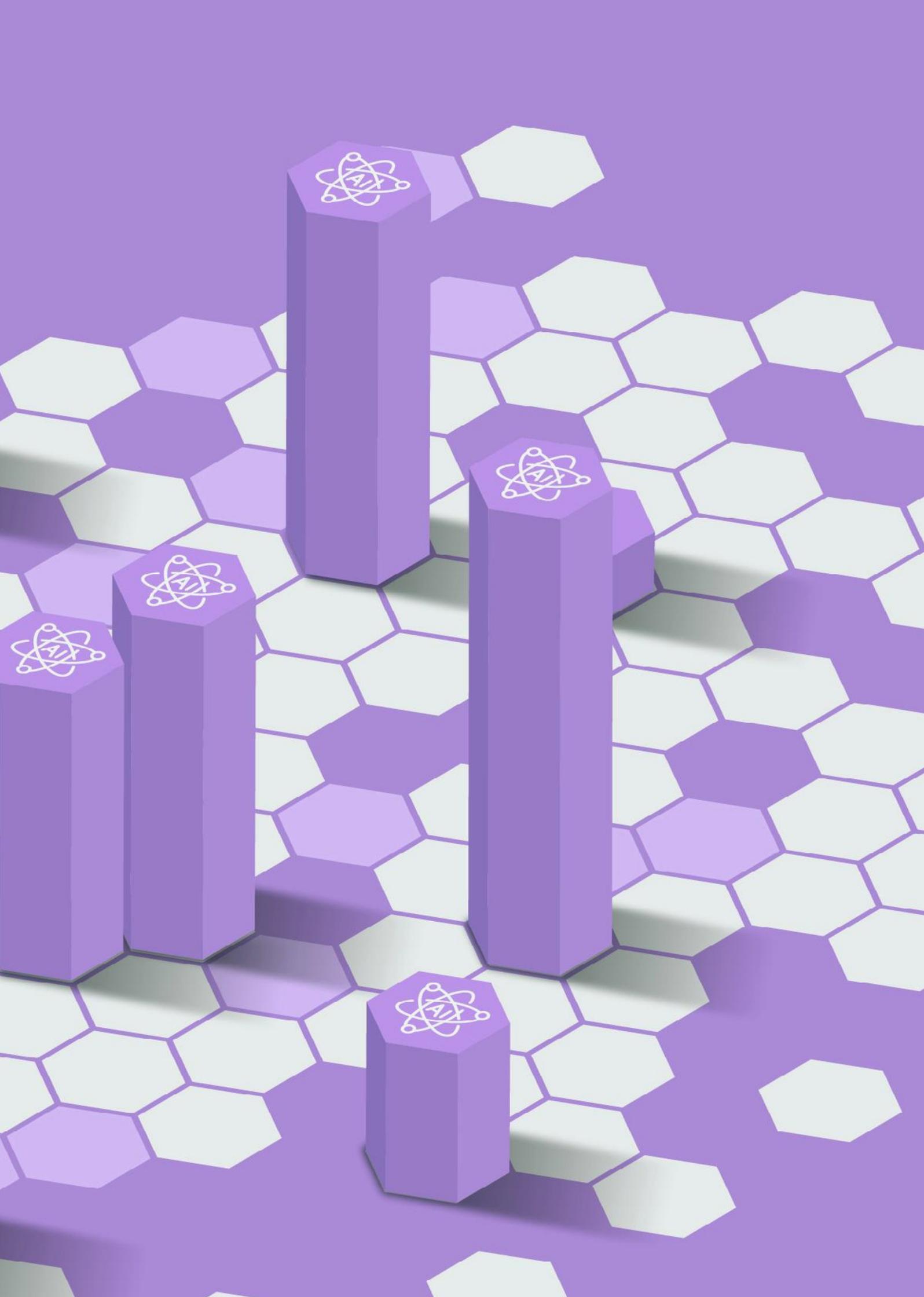


- 76** Publications
- 78** Research Grants and Teaching
- 80** Conference Presentations and Awards
- 87** The Team
- 90** HSOR Activities



HARNESSING AI IN HEALTHCARE ANALYTICS AND RESEARCH







Using Large Language Model to Analyse Routine Dashboard and Generate Commentary

Palvannan R.K.

Highlights

- LLMs can analyse and generate the text-generation aspect of interpreting routine healthcare dashboards.
- We tested this on structured monthly operational data, with domain knowledge encoded as prompts. The commentaries aligned with established rules though variability in style persisted.

Introduction

Senior management in healthcare organizations regularly review dashboards that reflect clinical quality, operations, finance, human resources, and population health. These dashboards rely on senior analysts to interpret data, spot trends, contextualize anomalies, and transform numbers into meaningful insights. This data-to-text chain can be time-intensive, particularly for junior analysts.

Can modern Large Language Models (LLMs) which generate text commentary, also analyse data using its pre-trained knowledge and contextual domain knowledge structured as prompts? This is a throwback to emulating earlier artificial intelligence expert systems with declarative knowledge. If so, junior analysts can be assisted with a LLM co-pilot, leaving senior analysts to focus on complex analysis.

We share our experience of using an LLM to analyse and comment on operational dashboards at the National Healthcare Group (NHG). The aim is twofold: (1) to demonstrate that LLMs can capture and utilize explicit domain knowledge through declarative prompts, (2) to highlight that LLMs can produce contextual, expert-like narratives, thereby assisting and upskilling junior analysts.

Method

Our data source is monthly operational statistics from hospitals from November 2023 to May 2024. They are organized into the care settings: outpatient care, emergency services, inpatient workload, surgical throughput, and community hospital flow. For each domain, we identified three core elements—demand, capacity, and wait times—and supplemented them with relevant statistics. We translated our known domain knowledge into a set of rules and facts (Table 1, column 2). These include known seasonal utilization patterns, statistical thresholds for distinguishing outliers from normal variation, quality benchmarks expressed as mathematical inequalities, target settings, fundamental equations, and relationships, and if-then rules or extended system thinking models, represented as acyclic causal diagrams (see last row to see how two causal links are expressed as triplets), to aid the LLM to explain observed trends. We progressively tested the data set with ChatGPT-4, 4o, o1-preview and o1.

Results

Table 1. A sample of the knowledge base or prompts and the associated LLM commentary

Category	Knowledge Base/Prompt	Result in the commentary by LLM
Seasonal & Utilization Patterns	"Months have a seasonal pattern of utilization, February is a short month, December a vacation period with lower utilization, rebounds are expected in following months."	Total attendances: Month-to-month fluctuations are observed (e.g., from 53,405 in November 2023 down to 45,894 in December 2023, settling around ~55,000 in recent months). Given seasonal patterns (December holidays, shorter February), these variations appear normal. The average volumes after January 2024 seem relatively stable around the mid-50,000 range.
Capacity & Bed Occupancy Rate (BOR) Standards	"A healthy BOR is around 90%..."	BOR: Very high, often above 100%. Values of 102–112% indicate use of temporary beds, which strains resources and leads to longer ED admission waits. This warrants urgent capacity and flow interventions.
Quality Benchmarks	"Door-to-Balloon <60 min"	AMI¹ Door-to-Balloon Time: Mostly below 50 minutes, well within the 60-minute target.
Elective Targets	"Electives are planned care, we target to reach 15% Electives."	Elective Percentage: Generally, 7–9%, stable and below the long-term target of 15%. No strong upward trend toward the target is observed yet.
Medical Home Initiatives	"Medical at Home ... is a key government initiative... should increase in volume."	Medical Home Cases: Increasing from 31 in November 2023 to 69 in May 24, showing a positive trend in utilizing Medical Home services.
Trend vs. Variation	"Broadly speaking those within +/- 10% are statistical and should not be alerted."	Priority 2 Wait Times to see ED² doctor: Fluctuated between 26-34 minutes, variations mostly within ±10%. Appears stable.
Queueing Theory & Wait Times	"Wait times differ by service type..."	The sustained improvement in 95 th percentile wait times is a key positive outcome, indicating that the very long waits are being reduced.
Equations and Ratios	% of ED attendances admitted = Admitted as inpatients/Total ED attendances	% Admitted as Inpatients: Steady around 41-43%, no significant trend. This is stable and likely indicates a consistent acuity mix.
Causal Relationships (CLD)	<Bed wait time, Discharge rate, (+) > <Discharge rate, BOR, (-) >	Infers that higher BOR leads to increased wait times and potentially longer ALOS ³ .

¹ AMI: Acute Myocardial Infarction;
² ED: Emergency Department;
³ ALOS: Average Length of Stay

Table 1 shows sample prompts and commentaries generated by ChatGPT-o1 model. The analysis and commentary are descriptive, correct and generally align with the prompts. LLM does provide its own explanations which may or may not be relevant. The causal loop diagram was partly successful. The limitation is LLM's variability. Traditional expert systems strictly adhere to programmed rules, but LLMs generate stylistically different responses when prompted multiple times. While the core insights remain similar, ensuring consistent format and depth across different runs requires careful prompt design.

Conclusion

In future, healthcare dashboards will integrate built-in LLM engines to generate routine commentary automatically. The next step would be to unify commentaries across multiple functional domains—clinical, operational, financial — since most patient data points connect along the same care continuum. With more advanced "agentic" capabilities, it could drill down into department-level details. To prepare for such a digital analyst, all we have to do now, is to get used to encoding our functional knowledge as prompts.



A Hybrid Approach Using Math and Large Language Models to Analyse Dashboard Data

Ernest Hoang Le Tri Cuong, Palvannan R.K., Teow Kiok Liang, Dr John Abisheganaden, Dr Steven Miller¹

¹ School of Information Systems, Singapore Management University, Singapore

Highlights

- We used a Large Language Model (LLM) to summarise notable trends in the National Healthcare Group (NHG)'s dashboard for monthly operational data of hospitals. However, this did not consistently produce reliable results due to LLMs' inherent limitations at doing mathematical calculations.
- We achieved more reliable results by pre-processing the numeric dashboard data with a mathematical tool before feeding the results to an LLM to create a summary.

Introduction

In a previous report, we evaluated two LLMs (OpenAI's 4o "fast model" and o1 "reasoning model") for generating key observations in NHG's monthly operational data dashboard, which is used by our senior management teams. As expected, the LLM could generate well-written and convincing comments. However, analysing time series data in a dashboard requires mathematical computations and we found that OpenAI's models were not consistently reliable in generating accurate results for this. LLMs are primarily trained for natural language processing, so they lack many mathematical capabilities.

Why are LLMs not reliable at mathematics? A possible explanation lies in how they perceive mathematical inputs. Unlike humans who can see individual numbers and symbols, the smallest unit that LLMs can see is tokens. A token may include multiple characters such as "+2^" or "5>". Moreover, as statistical models, LLMs often provide approximate answers. In the realm of natural language, this approximation would suffice because there are many ways to convey the same idea. In contrast, mathematical calculations demand precision. For example, if we instruct an LLM, which generates tokens of 2 characters, to compare 18 and 9, the ideal sequential output would be "18" and then ">9". However, due to the LLM's statistical nature, it may first produce "18" followed by "<9" instead of ">9". In terms of natural language, this response is still a valid answer to what we asked for, which is comparing "18" and "9", it only fails when we consider the exacting standards of mathematical logic.

To address these limitations, we used a hybrid approach that separates mathematical calculations from natural language processing.

Methods

To systematically describe the raw data programmatically, we referred to prior insights from a senior analyst and identified common commentary themes. Our original dashboard contained monthly metrics, starting from October 2023, which revealed the following types of observations:

1. "Slight decrease in RV:FV ratio" - Trend observation of the ratio of repeat visits to first visits.
2. "New low for hip fracture to surgery time in May" - Significant observation for May.
3. "BOR (Bed Occupancy Rate) higher than 90%" - Concerning data as BOR should be below 90%.

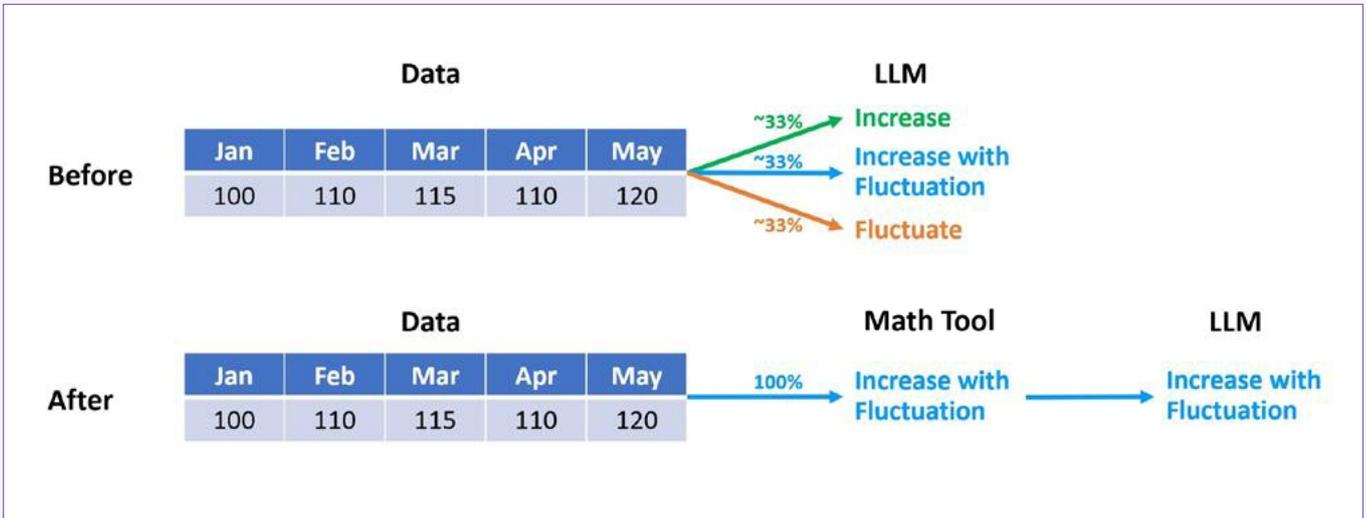
We developed a Python program that scanned the original data for these three types of observations. The program would compute the R-squared, mean, standard deviation, gradient, historical minimum and maximum before generating descriptive comments based on these values.

Once we obtained the observations, we had the option to pass both the original data and the generated observations to an LLM for comment generation. However, to prevent the LLM from executing mathematical calculations, we concealed the original data. To increase the depth of the LLM to assess the pre-processed numerical outputs, we adjusted the LLM prompts to encourage Chain-of-Thought (CoT) reasoning. The LLM was directed to follow these steps: exploring the knowledge base, theorising, and providing evidence. As for which LLM we chose, we went with gpt-4o specifically as it gave us better control over the CoT and was available on the secure GPT platform, Tandem as compared to gpt-o1 which had CoT built-in and was not available on any secured platform. We also supplied the LLM with previous analyses conducted by the senior analyst as learning examples.

Results

The implementation of the mathematical program allowed us to convert the original data into a set of deterministic observations, thereby reducing variability in the LLM’s responses after multiple runs with the same data. See **Figure 1** (below).

Figure 1. Example showing how a math program provides a deterministic outcome.



This approach fostered reliability, accountability, and traceability. This is because the responsibility for calculations lies with the mathematical program, which we can control more effectively. Coupled with CoT prompting, the LLM is better positioned to identify key observations and draw connections.

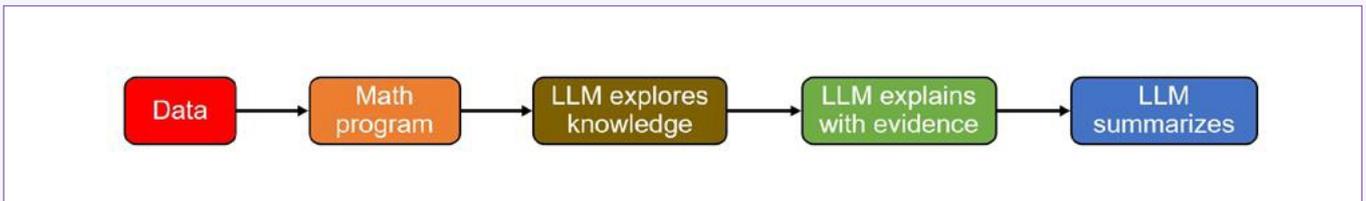
Table 1. Examples of math results and the corresponding LLM answers

Math program results	LLM answers
ED ¹ Attendances, Admissions, Discharges, BOR ² rose by x% in Dec, upward trend for the past 4 months ...	Increased ED attendances and admissions to inpatient services. This higher demand explains the increased discharges and BOR (for inpatient services).
BOR rose to 9x% in Dec. Bed Wait Time rose by x% in Dec.	The inpatient BOR rose to 9x% (above the target of 90%) leading to increased Bed Wait Time.

¹ ED: Emergency Department

² BOR: Bed Occupancy Rate

Figure 2. A pipeline of Math and LLM Agents



Conclusion

Employing a mathematical program to pre-process dashboard data before presenting it to the LLM for analysis effectively abstracts mathematical calculations from the LLM’s processing. This approach ensures accurate observations and limits variability in the LLM’s responses. Moving forward, we plan to segment the CoT prompt into various smaller tasks. By distributing these tasks among specialized agents (as shown in **Figure 2**), we aim to enhance the quality of generated responses.



Combining Generative AI and Machine Learning to Improve Prediction of High Healthcare Utilisers

Dr Yap Chun Wei

Highlights

- Persistent high utilisers can be predicted with PPV of 62.2% using traditional predictors.
- Generative AI could potentially improve PPV to 73.0% through additional vector embeddings.

Introduction

High healthcare utilisers represent only 2% of Central and North residents in Singapore, but their healthcare bills contribute 55% to the total incurred by all these residents. Among high utilisers, 39% are persistent high utilisers, defined as high utilisers for 2 consecutive years or more.

To facilitate identification of high utilisers for appropriate intervention to reduce their utilization in the next one year, the objectives of this study are to:

1. Predict which patients are likely to be persistent high utilizers, and
2. Determine whether generative artificial intelligence (AI) could improve the prediction accuracy of traditional machine learning models.

Methods

High utilisers between 2018 to 2022 living in the Central and North regions were identified using the National Healthcare Group Population Health Data Mart. A total of 10,000 patients who were high utilisers in 2022 were selected through stratified sampling for validation. The remaining 125,359 patients were used for model development.

A total of 309 predictors, comprising of demographics, comorbidities, economic status, risk factors, cost predictors, care predictors, disease severity and Johns Hopkins ACG predictors were calculated for each patient.

A text summary of each patient's characteristics was also generated and provided to generative AI for comments. These comments were converted to vector embeddings and combined with the 309 predictors for modelling. Google gemma-2b-it was used as the generative AI model and BAAI bge-large-en-v1.5 was used as the sentence transformer to convert the responses to vector embeddings.

Results

When trained with 309 predictors, ensemble methods, such as random forest and XGBoost, had higher predictive performance, with XGBoost having the highest predictive performance (Table 1). Table 2 shows the top predictors of persistent high utilisers, which included being a past high utiliser, frail and or of low socio-economic status, measured by use of Medifund. When the generative AI's comments were added as predictors, sensitivity and PPV increased to 75.7% and 73.0%, respectively (Table 3).

Conclusion

In this study, a predictive model for persistent high utilisers was developed. Additionally, we demonstrated how incorporating generative AI can potentially improve the prediction accuracy of a prediction model by using vector embeddings.

Results

Table 1. A sample of the knowledge base or prompts and the associated LLM commentary

Model	Sensitivity	Specificity	PPV	NPV	AUROC	AUROC-5CV
Random ¹	28.3%	71.6%	36.3%	63.6%	0.500	0.499
Always predict positive ²	100.0%	0.0%	36.4%	-	0.500	0.500
Existing true high utiliser ³	47.0%	81.3%	59.0%	72.9%	0.641	0.642
Logistic regression ⁴	66.1%	76.6%	61.7%	79.8%	0.792	0.792
Naïve Bayes ⁴	61.5%	77.7%	61.3%	77.9%	0.775	0.775
K-nearest neighbour ⁴	85.9%	70.3%	62.4%	89.7%	0.872	0.730
Random forest ⁴	80.8%	77.1%	66.8%	87.5%	0.878	0.815
XGBoost⁴	82.2%	82.0%	72.3%	89.0%	0.904	0.824

¹ Randomly predicts a patient to be a high utiliser with a probability of 33.9% (proportion of true high utilisers in training set).

² Always predicts a patient to be high utiliser.

³ Predicts a patient to be high utiliser if the patient is an existing true high utilizer (e.g. a high utiliser patient in 2019 who is also a high utiliser patient in 2018 will be predicted to be a high utiliser for 2020).

⁴ Prediction threshold is set at the probability of true high utilisers in training set to compensate for the lower probability of high utilisers.

Table 2. Top predictors

Predictors	Importance
High utiliser (\geq \$10000) in year - 2	21.6%
eFI – Severe frailty	5.9%
Hospital frailty risk score – high risk	4.9%
Persistent high utilizer (\geq \$10000) in previous two years	3.1%
Use of Medifund	2.8%
No. of acute hospital admissions in previous one year	2.8%
Total gross charge in previous 3 months	2.7%
Total attributable inpatient admissions in previous one year (i.e. expected utilisation based on comorbidities, adjusted based on attributable risk proportion)	2.4%
Total gross charge in year - 2	1.8%
eFI – Fit	1.0%
Total expected SOC gross charge in previous one year (i.e. expected utilisation based on comorbidities)	1.0%

eFI: Electronic Frailty Index

SOC: Specialist Outpatient Clinic

Table 3. XGBoost performance on validation set

XGBoost model	Sensitivity	Specificity	PPV	NPV	AUROC
Predictors only	72.3%	74.8%	62.2%	82.5%	0.819
Vector embeddings only	64.8%	59.6%	64.0%	60.5%	0.677
Predictors + Vector embeddings	75.7%	69.1%	73.0%	71.9%	0.809



Harvesting the Wealth of Unstructured Data: CogStack Exploration for Clinical Notes

Matthias Koh Yong An, Dr Gary Ang Yee, Palvannan R.K.

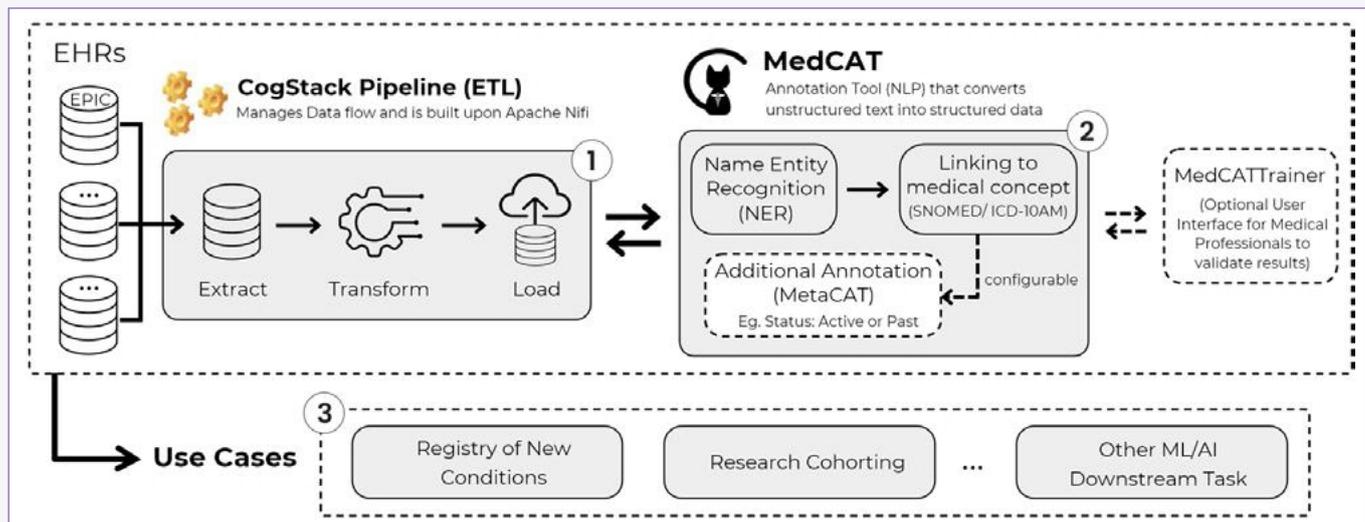
Highlights

- MedCAT, a natural language processing (NLP) annotation tool, demonstrated potential for extracting valuable insights from free-text clinical notes.
- Further testing on larger, localised datasets with licenced MedCAT model and MetaCAT component, is crucial in providing a more definitive evaluation.

Introduction

Despite the large volumes of clinical data stored in unstructured formats like free-text clinical notes, much are underutilised. Harnessing this information will enrich the structured data, potentially improving decision-making processes and optimising patient outcomes. CogStack, a comprehensive technological stack that was developed by Bio and Health Informatics at King’s College London, was explored. Its key components comprise the (i) ETL Pipeline that manages data flow; (ii) MedCAT, a natural language processing (NLP) annotation tool for Named Entity Recognition and Linking (NER+L); (iii) MetaCAT for contextualizing annotations; and (iv) MedCATTrainer user interface (UI), which facilitates human reinforcement learning by medical professionals (Figure 1). This study focused primarily on MedCAT, which performs NER+L to medical ontologies. The aim was to extract insights from unstructured clinical data.

Figure 1: Abstracted view of CogStack’s architecture



Methods

An open-source MedCAT model pre-trained on the Medical Information Mart for Intensive Care III (MIMIC-III) dataset was evaluated for its ability to extract and link medical entities to Systematized Nomenclature of Medicine – Clinical Terms (SNOMED CT) concepts. Following a literature review, a custom Python pipeline was developed for this evaluation. 30 clinical notes were randomly selected from CogStack’s open-source Patient Notes dataset, covering 9 clinical departments, including General Medicine, Radiology and Psychiatry. MedCAT’s performance was assessed for its ability to (1) accurately identify all possible entities in the clinical notes and (2) correctly link them to the appropriate SNOMED CT concepts within the “disorders” category, which contained approximately 350 entities. A physician reviewed the annotations generated by MedCAT to assess its accuracy. Annotations were categorised into three groups namely, “Correct NER+L”, “Wrong Link to SNOMED”, and “Entity Not Detected”. Metrics were calculated for each clinical note separately. The metrics included recall, which measured the proportion of correctly identified and linked annotations (Correct NER+L), and error rates that were classified as ‘Wrong Link to SNOMED’ and ‘Entity Not Detected’. The results of each metric were then averaged across all clinical notes to obtain the macro scores.

Results

The open-source MedCAT achieved a macro recall of 86.2% in identifying and linking entities to SNOMED CT, with error rates of 2.2% of 'Wrong Link to SNOMED' and 11.7% of 'Entity Not Detected'. This suggested that MedCAT could potentially be effective in supporting automated data enrichment in clinical workflows. However, entities extracted in isolation lack sufficient context to establish the relevance of disorders to the patient. For example, MedCAT has limited abilities to distinguish between current and past conditions (e.g. "has fever" vs. "had fever") or determine whether a condition pertains to the patient or their family. Ensuring that only patient-relevant disorders are retrieved requires understanding the MedCAT-extracted entities with their surrounding context, which is beyond MedCAT's standalone capabilities. While not available for this evaluation, MetaCAT has been described in literature to address this gap to conduct further contextualisation with its ability to interpret contextual nuances. Apart from this, some disorders require inference from contextual cues like lab tests or prescribed medications, when not explicitly stated. For example, medications like Diovan and Zocor as shown in the case notes in **Figure 2**, could suggest hypertension and dyslipidemia, respectively. These inferences fall outside MedCAT's design, necessitating the integration of other models to complement its functionality.

Figure 2: Example of case notes and MedCAT annotations

Patient Case Notes:	MedCAT Annotations	
Category: General Medicine	Types: Disorder ('Findings' type are not included eg hemoptysis)	
<p>REASON FOR CONSULTATION: Hemoptysis.HISTORY OF PRESENT ILLNESS: The patient is an 80-year-old African-American male, very well known to my service, with a past medical history significant for asbestos exposure. The patient also has a very extensive cardiac history that would be outlined below. He is being admitted with worsening shortness of breath and constipation. He is also complaining of cough and blood mixed with sputum production, but there is no fever.PAST MEDICAL HISTORY:1. Benign prostatic hypertrophy.2. Peptic ulcer disease.3. Atrial fibrillation.4. Coronary artery disease.5. Aortic valve replacement in 1991. St. Jude mechanical valve #23.6. ICD implantation.7. Peripheral vascular disease.8. CABG in 1991 and 1998.9. Congestive heart failure. EF 40%.10. Asbestos exposure.</p> <p>MEDICATIONS. 1. Coumadin 6 mg alternating with 9 mg.2. Prevacid 30 mg once a day.3. Diovan 160 mg every day.4. Flomax 0.4 mg every day.5. Coreg 25 mg in the morning and 12.5 mg at night.6. Aldactone 25 mg a day.7. Lasix 20 mg a day.8. Zocor 40 mg every day..ALLERGIES:1. DARVOCET.2. CLONIDINE.PHYSICAL EXAMINATION,GENERAL: The patient is an elderly male; awake, alert, and oriented. In no acute distress. VITAL SIGNS: Blood pressure is 136/80, pulse is 70, respiratory rate is 20, temperature 99.3, pulse oximetry 96% on 2 L nasal cannula.HEENT: Significant for peripheral cyanosis,NECK: Supple. ,LUNGS: Bibasilar crackles with decreased breath sounds in the left base,CARDIOVASCULAR: Regular rate and rhythm with murmur and metallic click.ABDOMEN: Soft and benign,EXTREMITIES: 1+ cyanosis. No clubbing. No edema.LABORATORY DATA: Shows a white count of 6.9, hemoglobin 10.6, hematocrit 31.2, and platelet count 160,000. CK 266, PTT 37, PT 34, and INR 3.7. Sodium 141, potassium 4.2, chloride 111, CO2 23, BUN 18, creatinine 1.7, glucose 91, calcium 8.6, total protein 6.1, albumin 3.3, total bilirubin 1.4, alkaline phosphatase 56, and troponin I 0.085 and 0.074.DIAGNOSTIC STUDIES: Chest x-ray shows previous sternotomy with ICD implantation and aortic valve mechanical implant with left-sided opacification of the diaphragm worrisome for pleural effusion.ASSSESSMENT:1. Hemoptysis.2. Acute bronchitis.3. Coagulopathy.4. Asbestos exposure.5. Left pleural effusion.RECOMMENDATIONS:1. Antibiotics.</p>	<p>1 'pretty_name': 'Benign prostatic hyperplasia', 'cur': '266569009', 'source_value': 'Benign prostatic hypertrophy', 'detected_name': 'benign-prostatic-hypertrophy'</p>	<p>6 'pretty_name': 'Congestive heart failure', 'cur': '42343007', 'source_value': 'Congestive heart failure', 'detected_name': 'congestive-heart-failure'</p>
	<p>2 'pretty_name': 'Peptic ulcer', 'cur': '13200003', 'source_value': 'Peptic ulcer disease', 'detected_name': 'peptic-ulcer-disease'</p>	<p>7 'pretty_name': 'Pleural effusion', 'cur': '60046008', 'source_value': 'pleural effusion', 'detected_name': 'pleural-effusion'</p>
	<p>3 'pretty_name': 'Atrial fibrillation', 'cur': '49436004', 'source_value': 'Atrial fibrillation', 'detected_name': 'atrial-fibrillation'</p>	<p>8 'pretty_name': 'Acute bronchitis', 'cur': '10509002', 'source_value': 'Acute bronchitis', 'detected_name': 'acute-bronchitis'</p>
	<p>4 'pretty_name': 'Coronary arteriosclerosis', 'cur': '53741008', 'source_value': 'Coronary artery disease', 'detected_name': 'coronary-artery-disease'</p>	<p>9 'pretty_name': 'Blood coagulation disorder', 'cur': '64779008', 'source_value': 'Coagulopathy', 'detected_name': 'coagulopathy'</p>
	<p>5 'pretty_name': 'Peripheral vascular disease', 'cur': '400047006', 'source_value': 'Peripheral vascular disease', 'detected_name': 'peripheral-vascular-disease'</p>	<p>10 'pretty_name': 'Pleural effusion', 'cur': '60046008', 'source_value': 'pleural effusion', 'detected_name': 'pleural-effusion'</p>

Conclusion

This exploration highlights MedCAT as a lightweight tool capable of performing NER+L tasks. However, MedCAT's design to only perform NER+L task underscores the need to utilise it in conjunction with contextual understanding tools like MetaCAT. Further testing with state-of-the-art (SOTA) licensed MedCAT models with MetaCAT, on larger and localised datasets is recommended to assess its performance in handling local sentence structures, acronyms and embedded context. To experiment on local contexts and datasets, a simple Python package with an intuitive Jupyter Notebook interface containing MedCAT has been developed and deployed across National Healthcare Group (NHG) institutions. These efforts will enable scalable and precise data enrichment across the NHG cluster, driving better healthcare outcomes.



Multi-agent AI – An Exploration on Applications

Dr Cheng San Chye

Highlights

- Adoption of a multi-agent artificial intelligence (AI) framework can help users to accomplish an overarching goal comprising a series of specific tasks.
- An example of a potential use would be the simulation of interactions between a health coach and individuals with varying levels of motivation, facilitating testing and evaluation of the quality of these interactions at scale.
- Prior to adoption, users of the frameworks must consider the suitability of the goal they wish to accomplish, as well as ethics, safety and governance processes.

Introduction

In a multi-agent AI framework, multiple specialised AI agents can help users to accomplish an overarching goal comprising a series of specific tasks, with or without a human in the loop. Specific tasks are distributed among specialised agents. Each agent can be powered by its own large language model (LLM), is capable of autonomous decision-making, and can be assigned tools to carry out its task.

Various frameworks (e.g., CrewAI, Microsoft AutoGen) and examples of their use have been developed (e.g., create job postings;¹ code generation, execution, and debugging²). Explorations of the utility of a multi-agent AI set-up in NHG are underway. A potential use could be the development and implementation process of a generative AI chatbot prototype. As part of development, iterations of rapid testing and user feedback are performed. However, acquiring a large number of user interactions for chatbot testing can be time-consuming. Alongside manual tests and users feedback, to accelerate and/or add rigour to this aspect, multiple AI agents can be used to:

- (a) simulate conversations in which AI agents role-play different profiles of the target population that the chatbot is intended for.
- (b) apply scoring rubrics and potentially assist in developing rubrics to classify interactions and evaluate the quality of conversations.

Using AI agents to complement manual tests and users feedback can form part of the examination on whether a generative AI chatbot is performing its intended functions or not. Hence, not only are manual tests and users feedback conducted, larger scale automated tests and evaluations can also be performed.

Description of an application of a multi-agent AI framework

For this article, I applied a multi-agent AI framework called Microsoft AutoGen to simulate various autonomous AI agents role-playing different roles to converse with one another in the context of health coaching of overweight adults to adopt healthy lifestyles.

To simulate conversations between the various AI agents, I used Ollama to run LLMs (e.g., Llama3.1, Phi-4) locally to power the AI agents. Prompts were provided for an AI agent to role-play a health coach, experienced in conducting motivational interviews to help overweight adults with Type 2 diabetes self-manage their weight. Prompts were also provided for AI agents to role-play two broad profiles of overweight adults with: (a) little motivation to manage their weight; and (b) some motivation to manage weight. These prompts were developed with reference to a commonly used behaviour change model, COM-B,³ where Capability (C), opportunity (O), and motivation (M) are three necessary components for any behaviour (B) to occur.³ Each simulated conversation involved two AI agents: one role-playing a health coach and the other role-playing one of the broad profiles of overweight adults. **Exhibit 1** shows excerpts from various automated conversations between AI agents (powered by Phi-4), reflecting their different roles. For example, the reply from the AI agent role-playing an adult with little motivation showed limited confidence to engage in behaviour change to lose weight.

Exhibit 1. Excerpts from various conversations between different role-playing AI agents

"I'm feeling a bit overwhelmed, honestly... My schedule and environment don't really support making big changes right now, and I often feel discouraged when things don't go as planned. Maybe we can talk about small steps that are manageable for me? It might help if we focus on one thing at a time so it doesn't seem too daunting."

- an AI agent role-playing an adult with little motivation

"It sounds like you're facing some real challenges with motivation and your environment, which is completely understandable. Recognizing the importance of weight management for your diabetes is a great starting point. Let's explore small steps together that can fit into your schedule and feel manageable. What types of activities or changes do you think might be achievable right now? Maybe we could focus on one area at a time to make it less overwhelming."

- an AI agent role-playing a health coach

"... I know losing weight is important for managing my diabetes, but it's a bit overwhelming at times. I do want to make some changes, though, Let's start small with something manageable. Maybe we could focus on incorporating more physical activity into my daily routine? Just simple things like taking short walks or doing light exercises that fit into my schedule could be a good beginning. What do you think about starting there?"

- an AI agent role-playing an adult with some motivation

Several factors should be considered before deciding whether a multi-agent AI framework should be applied to accomplish the goal that users aim to accomplish. These include the nature of the overarching goal, exact tasks/stages to be performed, how each task/stage is linked, the collaborative processes involved, tools and choice of LLMs required. Also, ethics, as well as safety and governance processes for AI agents⁴, will be continual considerations.

Conclusion

Creators of various multi-agent AI frameworks have featured examples of their use cases to illustrate the value of applying such frameworks. I have adopted such a framework to simulate interactions between a health coach and overweight adults with limited motivation to manage their weight. Such simulations can be in the context of testing and evaluating generative AI chatbot prototypes at scale, covering a wide range of targeted user profiles, on top of manual tests and user feedback. Prior to adoption, users of frameworks must consider the suitability of the goal they wish to accomplish, as well as ethics, safety and governance processes.

References

- ¹ CrewAI. Examples. <https://docs.crewai.com/examples/example>
- ² Microsoft AutoGen. Examples. <https://microsoft.github.io/autogen/0.2/docs/Examples/>
- ³ Michie, S., van Stralen, M.M. & West, R. The behaviour change wheel: A new method for characterising and designing behaviour change interventions. *Implementation Sci* 6, 42 (2011). <https://doi.org/10.1186/1748-5908-6-42>
- ⁴ Hodjat, B. (2025, 14 January). How to ensure the safety of modern AI agents and multi-agent systems. <https://www.weforum.org/stories/2025/01/ai-agents-multi-agent-systems-safety/>



Using Artificial Intelligence Tools to Support Systematic Reviews – An Overview

Ge Lixia, Dr Rupesh Agrawal¹, Dr Maxwell Singer², Palvannan R. K., Teow Kiok Liang, Dr Joseph Antonio D. Molina, Dr Yap Chun Wei, Dr John Abisheganaden

¹ National Healthcare Group Eye Institute, Tan Tock Seng Hospital, Singapore

² Department of Ophthalmology and Visual Science, Yale University School of Medicine, USA

Highlights

- Artificial Intelligence (AI) tools enhance efficiency in systematic reviews mainly by automating literature search, screening, and data extraction processes.
- While existing AI tools reduce workload and enhance efficiency, human validation is still required.
- AI integration in systematic reviews is still evolving, with the automation of data extraction being less advanced than screening methods but progressing rapidly.

Introduction

Systematic reviews (SRs) are essential for synthesising research evidence in healthcare. However, the traditional manual process is time-consuming and resource intensive. Artificial Intelligence (AI) technologies, including machine learning, natural language processing, deep learning, and large language models (LLMs) have been increasingly integrated to enhance efficiency in conducting SRs. AI tools can support almost all the key stages of SRs, including planning, systematic literature searching, appraisal & synthesis, and reporting findings. We aim to provide an overview of the existing AI tools which have been used to support SRs.

Methods

A review of existing AI tools for SRs was conducted based on literature and platform analyses. AI tools were categorised according to their function in different SR processes involved in the four stages. The effectiveness of AI tools was assessed based on their ability to reduce screening burden, improve accuracy, and streamline workflow.

Results

Several AI tools have been applied in various processes of SRs. **Figure 1** illustrates the commonly used AI tools in different SR processes.

At the stage of study appraisal and synthesis, AI-based methods can reduce title and abstract screening burden up to 60%¹ and represent time savings of more than 80 hours². Several tools, like Rayyan, Distiller SR, Abstrackr, and EPPI-Reviewer are popular and validated in supporting paper selection in SRs in the medical domain. The performance of these tools is influenced by several factors, including the effectiveness of the underlying machine learning paradigms and algorithms, the quality and quantity of data used for training, and the degree of human involvement.

While methods for automating / semi-automating data extraction have been explored, it remains less mature than automated screening technologies for practical use³. In recent years, the use of LLM has made automating data extraction possible, although it generally has neither been integrated in existing SR platforms nor validated in published papers. Despite these advancements, human oversight remains critical in ensuring accuracy and reliability throughout the whole SR processes. AI techniques or tools should complement rather than replace human judgment.

Figure 1. The commonly used AI tools in different systematic review processes

Stages	Processes	Examples of AI tools	Remarks
Planning	Search for existing reviews/protocols		
	Formulate review question		Prompting ChatGPT to "Define a research question for a systematic review comparing the efficacy of two drugs"
	Develop a review protocol	-	-
Systematic literature searching	Search strategy		Prompt ChatGPT to "Define a search strategy for this research question".
	Systematically search databases/grey literature	Python or R to interact with database APIs	
	Download citations to a reference manager		
Appraisal and synthesis	Select studies	 	
	Extract data		
	Synthesise data		
Reporting findings	Writing the review		Generating ideas on how to structure the discussion section, to synthesize data points, or to summarize key ideas; Provide citation suggestion.
	Publishing the review	-	-

Conclusion

AI technologies offer significant potential to optimise SR processes by reducing workload, improving efficiency, and enhancing reproducibility. However, the need for extensive human validation persists, particularly in complex tasks such as data extraction and bias assessment. The integration of LLMs and interactive AI assistants into SR platforms is a promising direction for future development. AI should be viewed as a complementary tool rather than a replacement for human judgment in evidence synthesis.

References

¹ Tsou AY, Treadwell JR, Erinoff E, Schoelles K (2020) Machine learning for screening prioritization in systematic reviews: comparative performance of Abstrackr and EPPI-Reviewer. *Syst Rev* 9(1):73

² Gates A, Gates M, Sebastianski M, Guitard S, Elliott SA, Hartling L (2020) The semi-automation of title and abstract screening: a retrospective exploration of ways to leverage Abstrackr’s relevance predictions in systematic and rapid reviews. *BMC Med Res Methodol* 20(1):139

³ Marshall, I.J., Wallace, B.C. Toward systematic review automation: a practical guide to using machine learning tools in research synthesis. *Syst Rev* 8, 163 (2019). <https://doi.org/10.1186/s13643-019-1074-9>



Exploring the Use of ChatGPT for Risk of Bias Assessment in Prognostic Studies

Sheryl Ng Hui Xian, Dr Yip Wan Fen

Highlights

- Having a Q&A with ChatGPT before asking it to perform an assessment showed promise in replicating a human reviewer’s ratings but will require additional encoding to comprehensively replicate a reviewer’s thought process.
- Reproducibility of ChatGPT’s appraisal was limited due to its inherent ‘creativity’ encoded in the underlying model.

Introduction

Systematic reviews of literature require rigorous appraisal of studies for their risk of bias (ROB) in design, conduct, analysis and reporting. Traditionally, two or more reviewers would appraise a study independently, before comparing and reconciling their assessment of the ROB. For reviews with a large volume of studies, this would be a laborious and manual process as both reviewers would have to read, understand and assess each study before comparing their ratings for consensus.

The emergence of large language models, such as the widely adopted ChatGPT, presents opportunities to streamline this appraisal process, given their ability to read volumes of text and retrieve information quickly. If effective, ChatGPT could reduce the total appraisal workload by replacing one reviewer. Hence, this study aimed to establish the accuracy and reliability of ChatGPT in ROB assessment.

Methods

We identified three studies included in a scoping review of prognostic studies in chronic lung disease and used the ROB assessments by the human reviewers as the gold standard. We included one study each rated to be at low, moderate and high risk of bias, to test the adaptability of ChatGPT to a diversity of studies. The ROB tool used in the scoping review was the Quality in Prognostic Studies (QUIPS) tool (**Table 1**).

Table 1. Description of the QUIPS tool by Hayden et al.

Domain	Characteristics of unbiased study
Study Participation	The study sample adequately represents the population of interest
Study Attrition	The study data available (i.e., participants not lost to follow-up) adequately represent the study sample
Prognostic Factor Measurement	The PF is measured in a similar way for all participants
Outcome Measurement	The outcome of interest is measured in a similar way for all participants
Study Confounding	Important potential confounding factors are appropriately accounted for
Statistical Analysis and Reporting	The statistical analysis is appropriate, and all primary outcomes are reported

To have ChatGPT perform the appraisal, we developed four types of prompts. The first, a direct prompt, requested for ChatGPT to rate the three papers using the QUIPS risk of bias tool, using only the information provided in the paper and without making any assumptions. If there was information unavailable or unclear, it was to provide the response “This information is unavailable or unclear” and rate the domain as ‘Not rated’. Three variants of this prompt were then developed in attempts to improve the performance of the prompting (**Table 2**).

Table 2. Additional prompts developed

Prompt type	Objective
Prompt with context	Provide background information to ChatGPT in terms of the objective of a prognostic study, the QUIPS tools and the prompting items and considerations within QUIPS, before asking it to perform the assessment
Chained prompt	Query ChatGPT on the objective of a risk of bias assessment and the process of doing so, before asking it to perform the assessment
ChatGPT-generated prompt	Have ChatGPT optimise its own prompt to perform the appraisal

Each prompt was tested on each paper, over five days. The performance of ChatGPT was assessed by the percentage of ratings that matched the human reviewers'. Inter-day reliability of ChatGPT's assessments were measured by level of agreement through the Cohen's weighted kappa.

Results

The chained prompt had the highest accuracy in ROB ratings across all studies and domains (69%), while the ChatGPT-generated prompt had the lowest (53%). Examining the results by study (**Table 3**), prompt chaining achieved an accuracy of 77% for the low and high ROB studies, but only 53% for the moderate ROB study. Notably, lower accuracy was observed across all prompt types for this study.

Table 3. Accuracy by study and prompt type

	Direct prompt	Prompt with context	Chained prompt	ChatGPT-generated prompt
Low ROB	70%	80%	77%	70%
Moderate ROB	50%	57%	53%	43%
High ROB	47%	57%	77%	47%

Of the six QUIPS domains, the chained prompt replicated the human reviewers' rating in all instances for the domain of outcome measurement. However, it consistently had the lowest accuracy in the domain of statistical analysis and reporting (range: 20%-40%). Across five different days of testing, agreement in the rating was only moderate (Kappa=0.41).

Conclusion

ChatGPT demonstrated potential in appraising prognostic studies using a ROB framework, especially with the aid of chained prompts. However, it had low accuracy in rating a domain requiring technical knowledge. In addition, the variation in responses across days of testing raised concerns on reliability and reproducibility as an objective assessment tool. To ensure its effectiveness to support or replace a reviewer in this process, prompts must be encoded with the entirety of the human reviewer's knowledge and judgement. Further, as the ChatGPT interface has an inherent 'creativity', or variation, encoded to allow for adapting to context and flexibility in problem-solving, for reproducibility, the underlying large language models without a preset variability (e.g. GPT-4) should be adopted for this task.

Reference

Hayden JA, van der Windt DA, Cartwright JL, Côté P, Bombardier C. Assessing bias in studies of prognostic factors. *Ann Intern Med.* 2013;158(4):280-286. doi:10.7326/0003-4819-158-4-201302190-00009



Exploring Thematic Analysis Processes with Large Language Model

Dr Charmaine Krishnasamy, Dr Mary Lee Ching Ling, Dr Li Ruijie

Highlights

- The use of a large language model (LLM) to assist in coding and categorisation of themes can facilitate the qualitative data analysis process (De Paoli, 2024). However, it is necessary for the researcher to assess the accuracy of the analysis and themes suggested by the LLM.
- A human-in-the-loop hybrid approach for thematic analytic tasks is preferred in which the LLM performs data coding and initial analysis, while researchers conceptualise research questions, analytical tasks and prompts, and appraise the findings.

Introduction

Thematic analysis of interview transcript data requires a substantial amount of time and effort to code, interpret and produce useful themes. In recent years, with advancements in large language models (LLMs), these platforms were explored for their potential to assist in the analyses of text and language structures. We used ChatGPT 4o on 21 February 2025 to explore the strengths and limitations of using LLMs to analyse qualitative data and generate codes and themes.

Methods

Three publicly available transcripts from an oral history repository (Matthews & Lloyd, 2016) were read and analysed for facilitators and barriers to self-management of Type 2 diabetes. Specific prompts were constructed to guide the LLM to do likewise, and the relevance of responses generated were compared. The prompts were also refined and used iteratively in the following steps:

1. Uploaded three interviews into the LLM. Inserted the definition of what initial coding is (Braun & Clarke, 2006).
2. Prompted the LLM with: "Please generate initial codes and illustrative quotes from the interviews, to address the research question, 'What are the barriers and facilitators to self-management for people with type 2 diabetes?'. Store one meaningful quote for 10 of the most salient initial codes."
3. After obtaining the output of initial codes and quotes for all three interviews, we prompted it with: "Please generate initial themes and illustrative quotes from your list of initial codes for the three interviews. Each code should contain up to a maximum of 10 words. Store one meaningful quote for 10 of the most salient focussed codes."
4. After obtaining the output of focussed codes, we sent the definitions of themes and sub-themes (Braun & Clarke, 2006) together with the prompt: "Please identify 3-5 themes and subthemes, and their descriptions by sorting and grouping these initial codes according to the respective themes and subthemes. Each theme and subtheme should be illustrated by one meaningful quote."
5. We obtained 3 themes with two subthemes for each and their descriptions, all of which were supported by illustrative quotes.
6. We reviewed them for validity, i.e., whether the themes applied to the entire dataset. We also selected the final set of themes and subthemes.
7. Our final prompt was: "Please connect the themes and subthemes into one description of the challenges faced by people managing their type 2 diabetes."
8. To which we got a list of barriers and facilitators to self-management for people with type 2 diabetes.

Results

Following a series of prompts comprising adequate provision for conceptual understanding, and specific task descriptions including the expected output format, the LLM generated 10 initial codes with accurate illustrative quotes, as prompted. It also generated an overview of themes and subthemes:

- Theme 1: Challenges in diabetes self-management
(Subthemes: Difficulties in monitoring & treatment, Struggles with diet & nutrition)
- Theme 2: Psychosocial and emotional aspects
(Subthemes: Psychological barriers to lifestyle changes, Emotional resilience & coping)
- Theme 3: Support systems and resources
(Subthemes: Family & social influence, Healthcare system & availability)

The subthemes were accompanied by corresponding descriptions of their content and an illustrative quote. However, the illustrative quotes were summarised versions and not accurate quotations from the data provided. Below is an extract of one such theme:

Theme 1: Challenges in diabetes self-management

- Subtheme 1.1: Difficulties in monitoring & treatment
Challenges related to monitoring blood sugar levels and insulin use.
Illustrative Quote: "The machine wasn't successful, I couldn't see the blood drop."
- Subtheme 1.2: Struggles with diet & nutrition
Struggles in maintaining a balanced diet and weight control.
Illustrative Quote: "Even with small portions, I couldn't lose weight on insulin."

Conclusion

The LLM was able to identify initial codes and group them into themes and subthemes, quickly and accurately. Its thematic output included descriptors of each iteration of analysis, as well as narrative summaries of the results as requested. However, the themes it suggested read more like topic summaries. While this was acceptable for a quick and rough grasp of the content of the transcript data, the codes and themes the LLM generated did not make for a meaning-based interpretive story. Good prompt engineering enables the LLM to generate initial codes and themes. Further prompts and active and iterative engagement are needed if researchers are seeking meaning-based interpretations of their qualitative data. We noted that some of the outputs such as illustrative quotes were summarisations where they should be in-vivo evidence. We recommend using a hybrid model of AI and human analysis for thematic analytic tasks, in which LLM performs data coding and initial analysis, while researchers conceptualise research questions, analytical tasks and prompts, and appraise the findings.

References

- ¹ Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101.
- ² De Paoli, S. (2024). Performing an inductive thematic analysis of semi-structured interviews with a large language model: An exploration and provocation on the limits of the approach. *Social Science Computer Review*, 42(4), 997-1019. <https://doi.org/10.1177/08944393231220483>
- ³ Matthews, D., & Lloyd, H. (2016). *Diabetes Stories*. <https://diabetesmemories.com>



Transcription in Healthcare: Comparing the Accuracy and Speed of *Whisper* and *Scribe*

Estelle Foo Jie Wei, Dr Joseph Antonio D. Molina, Dr Yip Wan Fen

Highlights

- *Whisper* outperformed *Scribe* in transcription accuracy. Both tools transcribed significantly faster than human transcription.
- *Whisper* and *Scribe* address healthcare data privacy through distinct methods. *Whisper* stores data locally for enhanced security, while *Scribe* integrates with Care360, streaming data without local retention.

Introduction

Transcription plays a key role in healthcare, supporting documentation of patient-physician interactions and qualitative research methods. However, human transcription is costly and time-consuming. Consequently, clinicians are increasingly tapping on automated speech recognition (ASR) technologies to transcribe patient-physician interactions. Critically, there is a lack of frameworks to evaluate ASR's effectiveness. Establishing ASR's reliability is critical to facilitate broader adoption across healthcare settings. *Whisper* stores data locally, ensuring enhanced control over sensitive information, while *Scribe* integrates with Care360, streaming data from personal devices without retaining it locally. These privacy-focused designs address critical concerns in healthcare, where safeguarding patient information is paramount. *Scribe* is currently being used by medical social workers (MSWs) to help transcribe patient conversations. The aim was to compare *Whisper's* (Large-v3-turbo, quantised) and *Scribe's* (*Whisper-large-v3*) transcription accuracy and speed.

Methods

Three audio recordings, each averaging 20 seconds, from an evaluation study were used for preliminary analysis. Each audio file was transcribed using *Whisper* and *Scribe*, with transcription time recorded. The outputs from *Whisper* and *Scribe* were compared to human transcription for accuracy and speed.

Transcription accuracy was assessed using two metrics:

1. Overall word error rate (WER): The percentage of insertions, deletions, substitutions, and spelling errors.
2. Major word class (MWC) WER: The percentage of errors in critical word classes (nouns, verbs, adjectives, and adverbs) that misconstrued the meaning.

While overall WER provides a general measure of accuracy, MWC WER offers a more meaningful assessment by focusing on key word classes essential for preserving meaning in clinical contexts.

To account for varying transcript lengths, total weighted overall WER and total weighted MWC WER were calculated by multiplying each WER by the transcript's word count, summing these values across all transcripts, and dividing by the total word count.

Results

Whisper outperformed *Scribe* in both accuracy metrics. *Whisper's* overall WER (ranging from 2.5%-18.2%) was lower compared to *Scribe's* overall WER (ranging from 2.5%-30.3%) (**Table 1**). *Whisper's* MWC WER (ranging from 0%-10.0%) was also lower than *Scribe's* MWC WER (ranging from 5.3%-25.0%). The total weighted overall WER for *Whisper* was lower than that for *Scribe* (8.2% vs 15.4%) and total weighted MWC WER for *Whisper* was also lower than that of *Scribe* (5.5% vs 14.6%).

In terms of speed, the average time taken for transcription by *Whisper* (57 seconds) and *Scribe* (4 seconds) were significantly shorter than the human transcriber (135 seconds; $p = 0.0061$ and 0.0019 , respectively). *Whisper* took a significantly longer time than *Scribe* to transcribe ($p = 0.0001$).

Table 1. Comparison of transcription times and word error rates.

	Length (sec)	Transcription done by	Time taken (sec)	Total word count	MWC word count	Overall WER (%)	Weighted overall WER (%)	MWC WER (%)	Weighted MWC WER (%)
Audio 1	23	Human	150	33	16				
		<i>Whisper</i>	57			18.2	6.0	6.3	1.0
		<i>Scribe</i>	5			30.3	10.0	25.0	4.0
Audio 2	20	Human	120	37	20				
		<i>Whisper</i>	57			5.4	2.0	10.0	2.0
		<i>Scribe</i>	3			16.2	6.0	15.0	3.0
Audio 3	20	Human	135	40	19				
		<i>Whisper</i>	57			2.5	1.0	0.0	0.0
		<i>Scribe</i>	3			2.5	1.0	5.3	1.0

MWC: Major Word Class
 WER: Word Error Rate

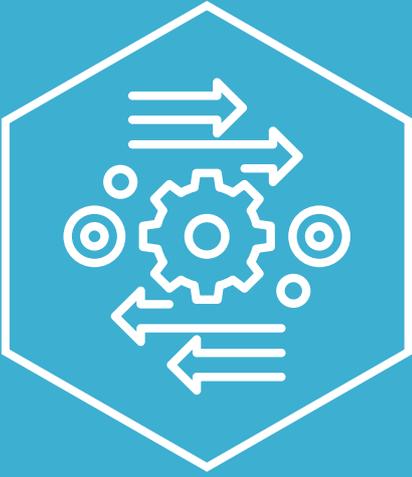
Conclusion

Both *Whisper* and *Scribe* transcribed significantly faster than human transcription, with *Whisper* transcribing slower than *Scribe*. This slower speed is likely because we did not run the model on a graphics processing unit (GPU). In terms of accuracy, *Whisper* seemed to outperform *Scribe* across all audio files. However, we acknowledge that the audio files used were short and the sample size was small. To fully validate the accuracy of both tools, longer and more audio files are required.

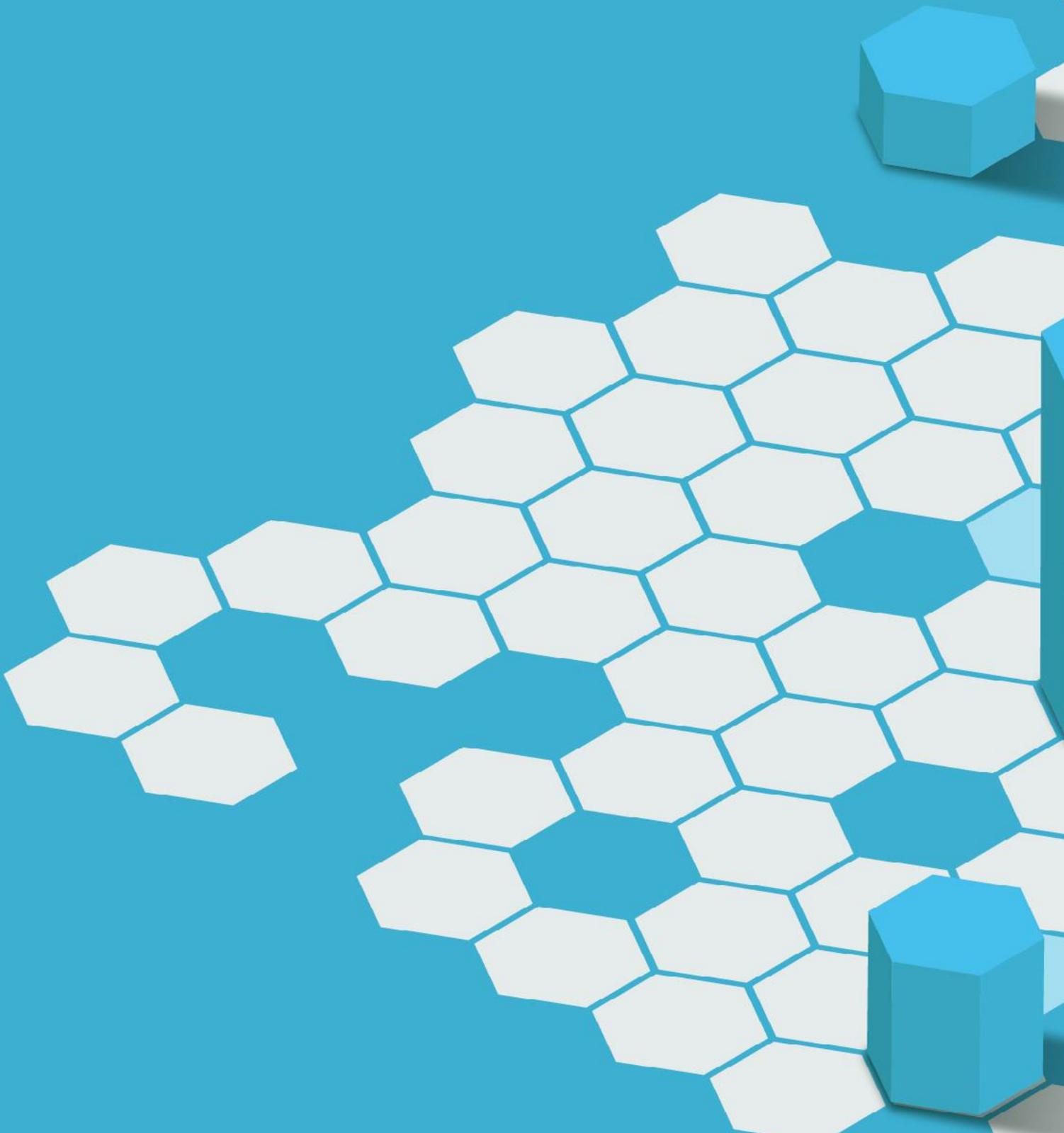
Notably, both *Whisper* (large-v3-turbo) and *Scribe* (*Whisper*-large-v3) do not transcribe disfluencies. *Whisper* (base model), on the other hand, can do so, and hence may be a more suitable choice when capturing disfluencies is critical.

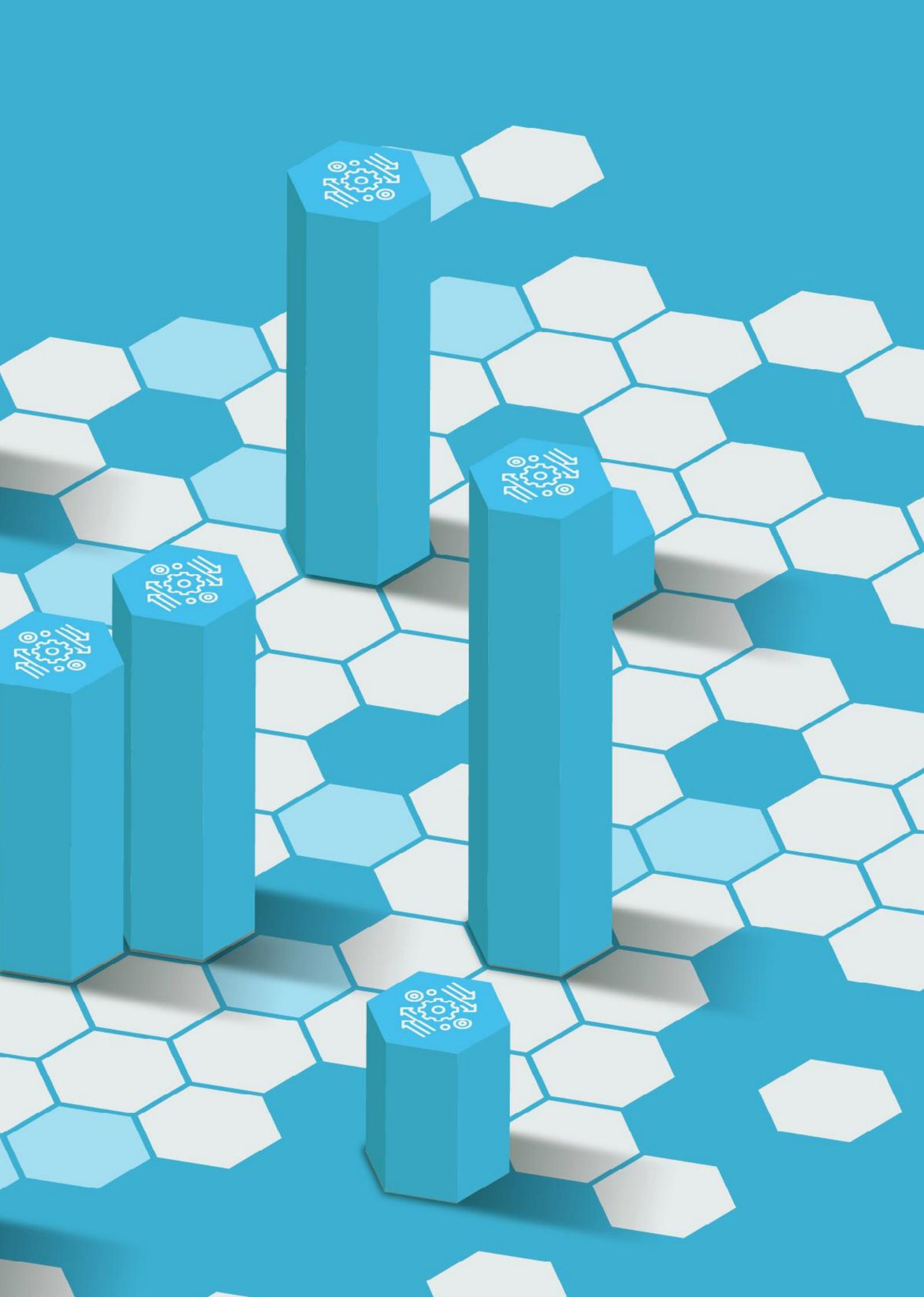
While overall WER provides a general measure of accuracy, MWC WER offers a more meaningful assessment by focusing on key word classes, such as nouns and verbs, which are vital for preserving meaning in clinical applications. It is hence useful to look at both overall WER and MWC WER when evaluating the accuracy of a transcription tool.

Further studies are needed to evaluate the clinical effectiveness of these transcription tools across diverse healthcare settings and establish accuracy thresholds to guide clinicians in selecting the most suitable ASR model for their needs.



ENHANCING HEALTH INTERVENTIONS WITH AI







Smart Prep: Enhancing Lung Function Testing with AI

Dr Ang Yee Gary, Ernest Hoang Le Tri Cuong, Dr Tan Geak Poh¹, James Nguyen², Kane Nguyen², Wang Li Jia², Dr John Abisheganaden

¹ Respiratory and Critical Care Medicine, Tan Tock Seng Hospital, Singapore

² Synapxe Pte Ltd, Singapore

Highlights

- We developed an AI-powered chatbot to enhance patient preparation for lung function tests, with the aim of improving patient understanding, reducing anxiety, and decreasing repeated tests.
- Its empathy-driven design was based on healthcare provider interviews and patient observation, and the prototype was among the Top 10 posters at the Ministry of Health Value Based Healthcare Conference 2024.

Introduction

Lung function tests play a crucial role in diagnosing and managing respiratory conditions such as asthma and Chronic Obstructive Pulmonary Disease (COPD). However, staff conducting these tests face significant challenges, including a high percentage of repeated tests due to poor patient preparation and anxiety. Many patients often ignore printed pre-test instructions, leading to suboptimal test results and inefficiencies in healthcare delivery.

To address these issues, a team of clinical and technical experts from the National Healthcare Group, Tan Tock Seng Hospital, and Synapxe have developed an innovative artificial intelligence (AI)-powered chatbot. This project, named “Smart Prep,” aims to enhance patient preparation and participation in lung function tests through the use of advanced AI technologies.

The primary objectives of the Smart Prep project are to:

1. Improve patient understanding of test procedures
2. Reduce anxiety associated with lung function testing
3. Decrease the rate of repeated tests due to poor preparation

By leveraging AI and cloud technologies, the project seeks to provide patients with personalised, interactive guidance that can significantly improve the efficiency and effectiveness of lung function testing.

Methods

The development of the chatbot followed an empathy-driven design approach, combining technical innovation with a deep understanding of patient needs and healthcare processes. The methodology included:

Stakeholder Interviews: The team conducted interviews with healthcare providers and staff to gain insights into the challenges faced during lung function testing.

Patient Journey Observation: Firsthand observation of the patient journey helped identify key pain points, including missed appointments, uncertainty about test preparation, and performance anxiety.

Technical Architecture Creation Set-up: The chatbot was built using the Claude 3.5 Sonnet large language model (LLM) for processing of inputs, with a scalable service-oriented design implemented on the Amazon Bedrock platform, within the nation-wide Healthcare Commercial Cloud adopted by all public healthcare institutions.

Knowledge Base Development: The team created a comprehensive knowledge base using text embedding (Cohere Model in Amazon Bedrock) for information retrieval and the Claude LLM for text generation. Sources included NHG Lung Function Test guidelines and general respiratory knowledge from reputable sources.

Iterative Development: The chatbot underwent continuous refinement based on feedback from both technical and clinical team members.

Results

The Smart Prep project has achieved several significant milestones:

Working Prototype: A functional chatbot prototype has been developed and is ready for testing (**Figure 1**). This prototype offers step-by-step preparation guidance and real-time query resolution for patients undergoing lung function tests.

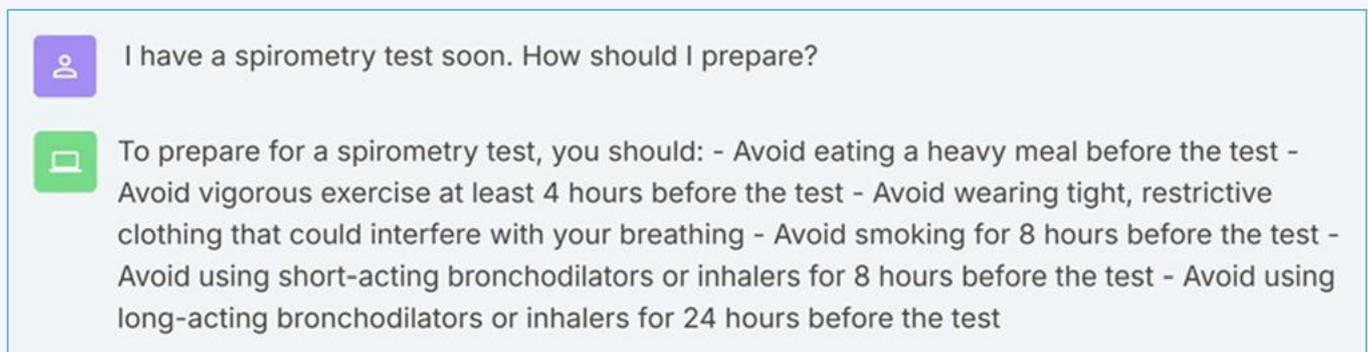
Cloud Integration: The chatbot is the first in the NHG to be deployed in the Healthcare Commercial Cloud, overcoming multiple hurdles to access LLMs in this secure environment.

Knowledge Base Integration: The chatbot successfully incorporates NHG-specific guidelines, ensuring that users receive information aligned with institutional protocols.

Recognition: The project won a Top 10 poster award at the Ministry of Health Value Based Healthcare Conference 2024, demonstrating its potential impact and innovation in healthcare delivery.

Additional Funding: Based on the prototype's promise, the project secured an additional \$7,000 in funding to further develop the chatbot.

Figure 1. Screenshot of the Smart Prep chatbot



Conclusion

The Smart Prep AI-powered chatbot represents a significant advancement in preparing patients for lung function tests. By leveraging cutting-edge AI technologies and cloud-based solutions, the project has the potential to greatly improve patient understanding, reduce anxiety, and decrease the rate of repeated tests.

Future directions for the project include: (i) applying for exemption from ethics review for testing with de-identified data; (ii) conducting acceptability testing and design improvements; (iii) refining the chatbot's expression style to be less robotic and more empathetic; (iv) improving the knowledge retrieval and chunking strategies to address information gaps; and (v) integrating additional features such as scheduling aid and reminder flows.



Enhancing the Patient Journey for Lung Function Tests Using Large Language Model-powered Chatbots

Ernest Hoang Le Tri Cuong, Dr Ang Yee Gary, Dr Tan Geak Poh¹, James Nguyen², Kane Nguyen², Wang Li Jia², Dr John Abisheganaden

¹ Respiratory and Critical Care Medicine, Tan Tock Seng Hospital, Singapore

² Synapxe Pte Ltd, Singapore

Highlights

- We developed a lung function test chatbot prototype using Retrieval Augmented Generation (RAG) on the Healthcare Commercial Cloud.
- RAG performance is heavily reliant on a robust Knowledge Base.

Introduction

Lung function tests require meticulous pre-test and post-test preparations by patients. A chatbot designed to educate patients about these tests and address their enquiries could be invaluable. Current chatbots often utilize Large Language Models (LLMs), which possess advanced conversational abilities but are prone to hallucinations and have limited contextual memory.

In our chatbot development, we faced two primary challenges. First, it was essential that our chatbot provides answers based on guidelines from the National Healthcare Group (NHG), rather than generate responses from its memory. Second, it was impractical to include all necessary information in a single prompt. Overcoming these challenges was crucial for enhancing the chatbot's performance.

To address the hallucination issue and the limitations of LLMs, we implemented a well-regarded framework known as Retrieval Augmented Generation (RAG). This system allows the storage of NHG-specific guidelines, enabling the LLM to retrieve only relevant information to formulate its responses. By doing so, we aimed to ensure efficient knowledge retrieval and minimize the risk of generating incorrect answers.

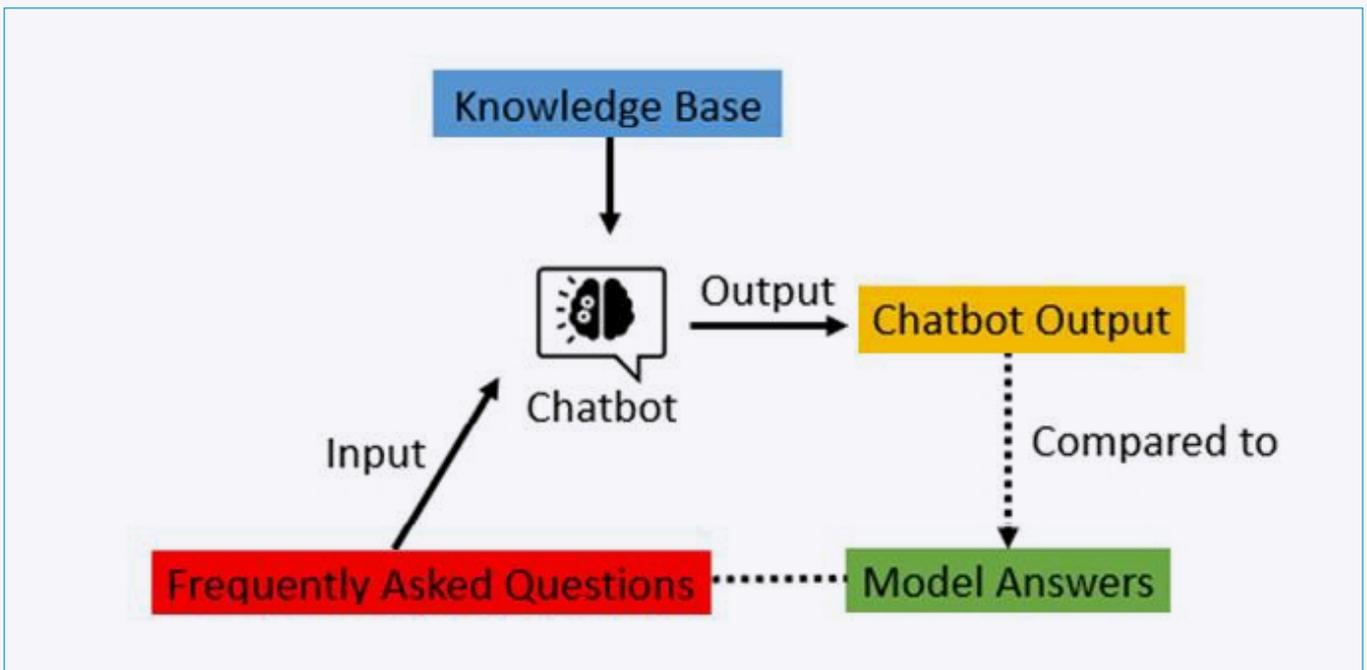
Our chatbot was deployed on Amazon Web Services (AWS) using the Healthcare Commercial Cloud (HCC), which offered both scalability and access to advanced LLMs available through the AWS Bedrock platform.

Methods

An RAG system usually has 2 components as shown in **Figure 1**: a Knowledge Base (KB) and an LLM (in this case a Chatbot).

The KB serves as the information repository, retrieving data based on natural language queries, much like a search engine. Our KB was distinct from the LLM and housed NHG documents related to lung function tests, primarily in PDF or Word formats. Only administrative users could update the KB, eliminating the need for patients to upload any information.

Uploaded documents were segmented into smaller text chunks, each converted into embeddings. An embedding functions as a vector-based summary of a text chunk; similar embeddings indicate related content. We utilized Cohere EN to create these embeddings, and AWS Bedrock served as our KB engine.

Figure 1. Schematic Representation of the RAG Architecture

Each embedding was paired with its corresponding text chunk when stored in the KB, where the embedding acted as the key and the original text chunk as the value. When a patient poses a question, the KB converts the inquiry into an embedding to query the database. The closest matching embeddings, along with their associated text chunks, are retrieved and provided to the LLM, which then formulates an answer. For our LLM, we employed Anthropic's Claude-3.5-Sonnet.

We populated the KB with publicly available NHG documents. To evaluate the chatbot's effectiveness, we consulted a respiratory physician for frequently asked questions and obtained model answers for benchmarking.

Results

Initial testing revealed several issues with our chatbot, including:

1. Inability to answer questions about asthma despite relevant documentation.
2. Incorrect responses regarding spirometry when asked about lung diffusion tests.
3. Missing information in preparation instructions.

Upon further investigation, we traced these problems back to our methods for uploading, chunking, and embedding information. Occasionally, our software struggled with PDF file extraction, prompting us to incorporate Optical Character Recognition for improved PDF-to-text conversion. Additionally, we began storing not just embedding-text pairs but also metadata, such as topic names and document titles, to enhance retrieval accuracy. These adjustments successfully addressed the identified issues.

We also assessed the chatbot's responses to malicious and irrelevant questions, confirming that it appropriately refused to answer, a feature inherent to AWS Bedrock. Nevertheless, further testing within a structured framework is necessary to comprehensively evaluate the chatbot's performance.

Conclusion

While the concept of RAG appears straightforward, its implementation can be complex. Securing permission for certain essential cloud services on the HCC proved challenging, and we had to find workarounds for some access limitations. Moving forward, we will benchmark our chatbot against key metrics, including accuracy, relevance, bias, and clarity. The results will guide our iterative improvement process. Future enhancements may involve further refining retrieval performance through data augmentation techniques and integrating our chatbot with existing front-end platforms like WhatsApp, providing a more familiar interface for patients and enhancing usability.



AI-assisted Pharmacy Intervention Classification

Dr Yap Chun Wei, Sia Beng Yi¹, Tan Wei Qin¹, Lim Hong Yee¹, Crisann Moh Xin Yun¹, Audrey Chong Wei Kuan¹, Dr John Abisheganaden

¹ Department of Pharmacy, Tan Tock Seng Hospital, Singapore

Highlights

- AI models can potentially assist pharmacists in the classification of inpatient and outpatient pharmacy interventions.
- Fine-tuned LLM has the best accuracy among the AI methods for this task.

Introduction

Pharmacists initiate interventions to prevent prescribing errors. These are documented and analyzed periodically to create educational materials and workflow changes.

At Tan Tock Seng Hospital, pharmacists document interventions based on six different drug related problems: 1. Drug Selection related (dosage form, strength, better or safer choice); 2. Dosage Regimen related (frequency, timing, rate, quantity, duration, add "once order"); 3. Route, Site, Diluent, Container, Dilution-related; 4. Monitoring; 5. Adherence and Education; 6. Operational. These are further subclassified into their respective subtypes, different types of Action Taken, and then into different types of Process Related Causes, with a total of 84 different categories. To standardise and harmonise intervention classifications, pharmacists spend over 65 hours monthly to check and re-classify these interventions accurately.

This study aims to develop an artificial intelligence (AI) model to assist pharmacists in the classification of these interventions.

Methods

A total of 150,763 inpatient and outpatient pharmacy interventions from August 2022 to June 2024 were used in this study. These were split into training (121,283 interventions) and testing sets (29,480 interventions) each running in chronological order.

Four types of AI methods were explored:

1. Global Vectors for Word Representation (GloVe) word embedding with Long-Short Term Memory (LSTM)
2. Sentence embedding with single hidden layer neural network (NN)
3. Retrieval-augmented generation (RAG) pipeline
4. Fine-tuning of Large Language Models (LLMs)

Results

The testing set accuracies of the best models from the four AI methods are shown in **Table 1** below. Fine-tuned LLM had the best predictive performance with an accuracy of 72.0%.

Table 1. Testing set accuracies of best model for each AI method

AI methods	Testing accuracy
GloVe + LSTM	62.6%
Sentence embedding + NN	58.1%
RAG pipeline	43.5%
Fine-tuned LLM	72.0%

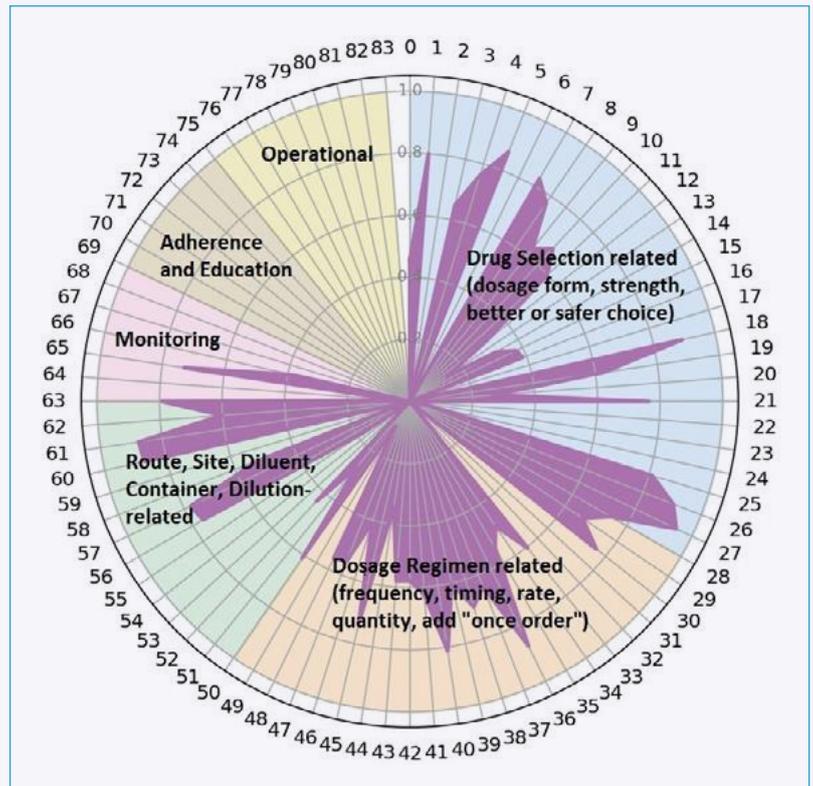
Table 2 presents fine-tuning results for LLMs. Similarly sized LLMs showed comparable performance. Models below 7 billion parameters underperformed, while those above 10 billion showed no significant improvements. Fine-tuned foundation and chat models performed similarly, suggesting an optimal parameter range of 7-10 billion for effective fine-tuning.

Table 2. Testing set accuracies of fine-tuned LLMs

LLM	No. of parameters	Testing accuracy
Similarly sized fine-tuned LLMs		
mistralai/Mistral-7B-Instruct-v0.3	7B	71.9%
meta-llama/Meta-Llama-3.1-8B-Instruct	8B	71.1%
google/gemma-2-9b-it	9B	72.0%
Different sized fine-tuned LLMs		
microsoft/Phi-3-mini-4k-instruct	3.8B	70.7%
microsoft/Phi-3-medium-4k-instruct	14B	70.7%
google/gemma-2-2b-it	2B	70.6%
google/gemma-2-9b-it	9B	72.0%
google/gemma-2-27b-it	27B	72.0%
Fine-tuned foundation vs instruct LLMs		
mistralai/Mistral-7B-v0.3	7B	71.7%
mistralai/Mistral-7B-Instruct-v0.3	7B	71.9%
meta-llama/Meta-Llama-3.1-8B	8B	71.4%
meta-llama/Meta-Llama-3.1-8B-Instruct	8B	71.7%

For the 84 intervention categories, the best fine-tuned LLM (gemma-2-9b-it) had testing accuracies and precision ranging from 0% to 94.5% and from 0% to 100.0%, respectively (**Figure 1**). It achieved testing accuracies of 73.4% for inpatient and 67.6% for outpatient interventions.

Figure 1 (right).
Testing accuracies for each of the 84 intervention categories



Conclusion

Among the four different types of AI methods, fine-tuning of LLMs had the best predictive performance for classification of pharmacy interventions. However, due to the wide variation in accuracies and PPV (positive predictive value) for the different categories, the AI model can only assist in the classification of half of the interventions, assuming that the model is used only for those categories where its precision is 75% or higher.



Cost-effectiveness Analysis of AI-enhanced Screening for Diabetic Foot Ulcers

*Dr Sun Yan, Ge Lixia, Dr Gary Ang Yee, Dr Joseph Lo Zhiwen¹, Dr Liew Huiling²,
Dr Donna Tan M.L.³, Dr Daniel Chew^{2,4}, Dr John Abisheganaden*

¹ Vascular Surgery Service, Department of Surgery, Woodlands Health, Singapore

² Department of Endocrinology, Tan Tock Seng Hospital, Singapore

³ National Healthcare Group Polyclinics, Singapore

⁴ Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore

Highlights

- AI-enhanced screening estimated savings of SGD657.5 million over 40 years for half-million patients with diabetes.
- AI-enhanced screening offers a viable strategy to optimise resource allocation in diabetic foot ulcer management.

Introduction

Diabetic foot ulcers (DFUs) are a serious complication of diabetes mellitus, with a lifetime risk estimated to be between 19% and 34%. Without timely prevention and management, DFUs can lead to lower extremity amputations (LEAs) and premature death.¹ DFUs also impose significant healthcare and societal costs. Regular foot screenings are essential for preventing these complications.

The guidelines set by the 2023 International Working Group on the Diabetic Foot² recommend foot screening frequency based on their risk stratification system: annually for very low risk, once every 6–12 months for low risk, once every 3–6 months for moderate risk, and once every 1–3 months for high risk. However, adherence to these guidelines reveals significant challenges and variability in practice due to barriers such as resource limitations and staff shortages. In Singapore, the Agency for Care Effectiveness recommends differentiated screening intervals: annually for low-risk patients, every six months for moderate-risk patients, and every 3–4 months for high-risk patients.³ Low-risk patients are defined as those without any of the following risk factors: calluses, deformities, peripheral artery disease, neuropathy, previous foot ulcers or amputations. Evidence suggests that the risk of ulceration in low-risk patients remains stable over time. Based on data from National Healthcare Group's Chronic Disease Management Datamart (CDMD), we estimated that 60–80% patients with diabetes can be classified as low risk, and their annual DFU incidence risk is lower than 1%, raising concerns about over-screening and inefficient resource allocation.

Artificial intelligence (AI) offers a potential solution by enabling personalised risk-tailored screening. This study evaluated the cost-effectiveness and clinical outcomes of AI-enhanced screening versus routine annual screening for low-risk patients. Using an XGBoost-based predictive model, the AI-enhanced screening approach involves screening based on individual risk: patients flagged as positive are screened annually, while negative low-risk patients are screened every three years. The predictive model achieved an area under the receiver operating characteristics curve of 0.81, with a sensitivity of 0.7 and a specificity of 0.8. A cohort of 500,000 low-risk patients with diabetes in Singapore, with an average age of 50, was simulated over a lifetime to assess the long-term impact of the AI-enhanced screening.

Methods

A Markov state-transition model was constructed to simulate disease progression through five health states: diabetes, DFU, minor LEA, major LEA and death. Model parameters, including both base values and distributions of disease transition probabilities and costs, were derived from CDMD and validated against international sources.⁴ Direct medical cost for foot screening, examination, consultation, DFU treatment and management, LEA procedures, and related services were included in the cost analysis. Effectiveness was measured in quality-adjusted life years (QALYs), based on EQ-5D data from a local study. A 3% discount rate was applied for both costs and effectiveness. The incremental cost-effectiveness ratio (ICER) was then assessed to determine the most cost-effective screening strategy using gross domestic product (GDP) per capita as the willingness-to-pay (WTP) threshold as commonly used in Singapore. Other outcomes measured in the simulation included the number of patients screened, DFU detections, minor and major LEAs performed, and deaths. Monte Carlo micro-simulations with 1,000 samples, randomly drawn from parameter distributions, were performed for probability sensitivity analysis (PSA) to account for uncertainties in model parameters. Results were summarised in cost-effectiveness acceptability curves (CEAC) and expected loss curves (ELC) to determine the optimal screening strategy under uncertainty. The CEAC shows the chance of each strategy being the most cost-effective and picks the one with the highest chance. The ELC shows the potential cost of choosing the wrong strategy and selects the one with the lowest risk of loss.

Results

Based on our modeling and simulation, AI-enhanced screening resulted in an average lifelong cost of SGD54,272 per patient, compared to SGD55,587 for annual screening, saving SGD1,315 per patient. The lifelong effectiveness for AI-enhanced screening was 23.150 QALYs per patient, slightly lower than 23.154 QALYs for annual screening, with a minimal difference of 0.004 QALYs per patient. AI-enhanced screening required 4,372,523 DFU screenings, compared to 11,178,861 screenings under annual screening. LEA rates were similar between strategies, with 52,720 minor and 68,646 major LEAs for AI-enhanced screening versus 52,689 minor and 68,710 major LEAs for routine screening. AI-enhanced screening led to 26 additional deaths. The AI-enhanced screening for Singapore's diabetic population reduced 6,806,338 unnecessary DFU screenings and saved SGD657.5 million over 40 years compared to routine annual screening, with minimal impact on quality-adjusted life years (QALYs). The ICER was SGD292,181 per QALY gained, well above Singapore's GDP per capita in 2023 (SGD110,000), indicating that AI-enhanced screening was more cost-effective than annual screening.

PSA showed that AI-enhanced screening had an average cost of SGD52,178 and effectiveness of 22.866 QALYs, compared to SGD53,993 and 22.876 QALYs for routine screening. The ICER from PSA was SGD174,572 (standard deviation: SGD13,296) per QALY gained. Both CEAC and ELC indicated that AI-enhanced screening was optimal when WTP was below SGD180,000 per QALY.

Conclusion

This study demonstrates that AI-enhanced DFU screening can offer significant cost savings while maintaining high standards of care. By tailoring screening intervals based on individual risk, healthcare systems can improve resource efficiency while maintaining high standards of care. Future research should focus on validating these findings in diverse populations and addressing ethical concerns to ensure that AI-driven healthcare innovations are accessible to all.

References

- ¹ Armstrong DG, Boulton AJM, Bus SA. Diabetic Foot Ulcers and Their Recurrence. *N Engl J Med* 2017;376:2367-75.
- ² Schaper NC, van Netten JJ, Apelqvist J, et al. Practical guidelines on the prevention and management of diabetes-related foot disease: IWGDF 2023 update. The International Working Group on the Diabetic Foot 2023. <https://iwgdfguidelines.org/wp-content/uploads/2023/07/IWGDF-Guidelines-2023.pdf>. Accessed 17 September 2024.
- ³ Agency for Care Effectiveness. Foot assessment in patients with diabetes mellitus, 6 June 2019. Updated 8 August 2024. [https://www.ace-hta.gov.sg/docs/default-source/acgs/foot-assessment-in-patients-with-diabetes-mellitus-\(aug-2024\).pdf?sfvrsn=7ee3b22c_5](https://www.ace-hta.gov.sg/docs/default-source/acgs/foot-assessment-in-patients-with-diabetes-mellitus-(aug-2024).pdf?sfvrsn=7ee3b22c_5). Accessed 17 September 2024.
- ⁴ Lo ZJ, Surendra NK, Saxena A, et al. Clinical and economic burden of diabetic foot ulcers: A 5-year longitudinal multi-ethnic cohort study from the tropics. *Int Wound J* 2021;18:375-86

The content of this write-up has been published in part and licensed under a Creative Commons BY-NC-SA 4.0 License. Citation: Sun, Y., Ge, L., Ang, Y. G., Lo, Z. J., Liew, H., Tan, D. M., ... & Abisheganaden, J. A. (2024). Cost-effectiveness and clinical outcomes of artificial intelligence-enhanced screening for diabetic foot ulcers: A simulation study. *Annals of the Academy of Medicine, Singapore*, 53(10), 638-640.



Predictive Risk Scores for 30-day Readmission

*Dr John Abisheganaden, Robin Tan Kah Heng¹, Dr Tay Seow Yian¹,
Dr Terence Tang², Dr Yap Chun Wei*

¹ Group Digital Health, National Healthcare Group, Singapore

² Group Health Informatics, National Healthcare Group, Singapore

Highlights

- Validation of the National University Health System 30-day readmission risk score in Tan Tock Seng Hospital and Khoo Teck Puat Hospital reported AUROC values of 0.74 and 0.73 respectively, comparable to the estimate of 0.78 in NUHS.
- The model exhibits varying effectiveness across different medical departments, highlighting the need for tailored approaches in risk assessment.

Introduction

National University Health System (NUHS) has developed a risk score for real-time prediction of 30-day readmission risk for hospitalized patients using machine-learning models, and deployed it in the EPIC Next Generation Electronic Medical Record (NGEMR) system in 2017. The model had an Area under Receiver Operating Characteristic (AUROC) value of 0.78. Using a cut-off score of ≥ 30 to identify patients at high risk of readmissions within 30 days, the model had sensitivity of 42%, specificity of 88%, positive predictive value (PPV) of 39%, and negative predictive value (NPV) of 90%.

Hence, we aimed to: i) validate NUHS' model on National Healthcare Group (NHG) data; ii) compare its accuracy with both the original and modified LACE models in NHG; and iii) determine the accuracy of the risk score for sub-specialty departments and conditions.

Methods

Patients studied were discharged from Tan Tock Seng Hospital (TTSH) between 30 Jul 2022 to 1 Dec 2023, and from Khoo Teck Puat Hospital (KTPH) between 25 Feb 2023 to 1 Dec 2023.

The 30-day readmission risk scores were extracted from our EPIC NGEMR. The LACE model incorporates Length of stay (L), Acuity of admission (A), Comorbidities (C), and recent Emergency department use (E) as predictors of readmission or death within 30 days. LACE scores were computed using the original LACE algorithm, and a modified LACE were extracted from our Population Health Data Mart (PHDM) for the same group of patients. Episodes with 30-day readmission from all-causes were identified using the PHDM.

AUROC were calculated based on the raw risk scores. Sensitivity, specificity, PPV and NPV were determined using the same cut-off score of ≥ 30 for the NUHS model, a cut-off score ≥ 10 for the original LACE model, and a cut-off score of the top 15% of patients for the modified LACE model.

Results

The readmission risk model developed by NUHS had a comparable AUROC in patients discharged from TTSH (0.739) and KTPH (0.725), and were similar to that of the modified LACE model (0.738 and 0.722, **Table 1**). In contrast, the original LACE model had a much poorer performance (AUROC: 0.590 and 0.679).

Among the Medical departments, the model had varying performance. Based on the PPV, we found the risk score suitable for use by Gastroenterology and Renal departments in both TTSH and KTPH, and for Hematology and Neurology patients in TTSH. However, the model had lower AUROC for surgical departments. Among the conditions of higher risk of readmissions, it had the highest AUROC and PPV for asthma and chronic obstructive pulmonary disease (**Table 2**).

Table 1. Comparison of performance of the 30-day readmission risk score developed by NUHS

Hospital	Model	AUROC	Sensitivity	Specificity	PPV	NPV
NTEFGH	NUHS	~0.783	~42.2%	~88.4%	~39.2%	~89.7%
TTSH	NUHS	0.739	42.3%	87.1%	40.1%	88.1%
	Modified LACE	0.738	39.4%	89.1%	42.4%	87.8%
	Original LACE	0.590	70.0%	39.7%	19.2%	86.6%
KTPH	NUHS	0.725	38.7%	87.5%	35.4%	89.0%
	Modified LACE	0.722	32.5%	91.0%	39.0%	88.4%
	Original LACE	0.679	79.3%	43.0%	19.8%	92.2%

AUROC: Area Under the Receiver Operating Characteristic

KTPH: Khoo Teck Puat Hospital

NPV: Negative Predictive Value

NTEFGH: Ng Teng Fong General Hospital

NUHS: National University Health System

PPV: Positive Predictive Value

TTSH: Tan Tock Seng Hospital

Table 2. Comparison of model performance among primary discharge diagnoses

Primary discharge diagnosis	TTSH		KTPH	
	AUROC	PPV	Specificity	PPV
AMI	0.640	32.2%	0.621	27.9%
Asthma	0.776	47.1%	0.749	47.8%
COPD	0.756	41.7%	0.759	50.0%
Diabetes	0.705	38.6%	0.689	36.8%
Heart failure	0.670	34.1%	0.658	36.9%
Pneumonia	0.672	30.0%	0.681	31.8%
Stroke	0.639	21.5%	0.615	23.1%

AMI: Acute Myocardial Infarction

AUROC: Area Under the Receiver Operating Characteristic

COPD: Chronic Obstructive Pulmonary Disease

KTPH: Khoo Teck Puat Hospital

PPV: Positive Predictive Value

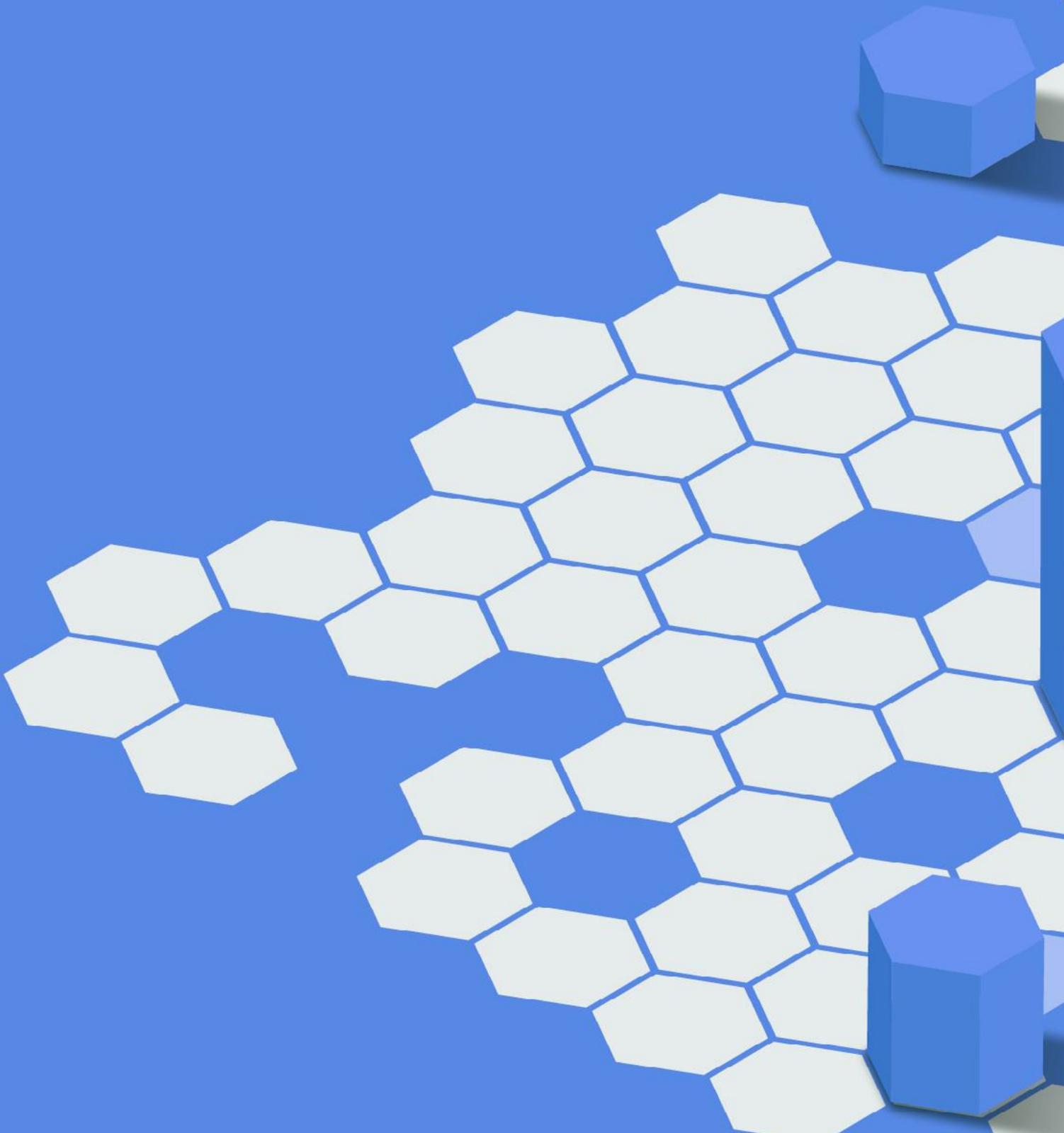
TTSH: Tan Tock Seng Hospital

Conclusion

Unplanned readmissions to a hospital add a significant strain on healthcare systems and aggravates the bed crunch faced by hospitals in Singapore. Development and adoption of risk prediction models may guide clinicians and help alleviate this. Validation of the NUHS 30-day predictive risk score in NHG showed comparable performance with NUHS and the modified LACE model. We also found that it may be useful to predict better 30-day readmission risk for certain medical departments or medical conditions.



EXAMINING DRIVERS OF PATIENT OUTCOMES







Factors Associated with Longer Length of Stay in Patients Undergoing Primary Total Knee Replacement at a General Hospital in Singapore

Ng Su Yee¹, Dr Shiek Abdullah Bin Ismail¹, Christopher Ng Thong Lian², Mohd Khalid Bin Mohd Anuar^{3,4}, Ivan Tan Peng Choon⁵, Dr Michelle Jessica Pereira

¹ Department of Rehabilitation Service, Khoo Teck Puat Hospital, Singapore

² Department of Physiotherapy, Tan Tock Seng Hospital, Singapore

³ Group Allied Health, National Healthcare Group, Singapore

⁴ Allied Health, Tan Tock Seng Hospital, Singapore

⁵ Value Based Care, National Healthcare Group, Singapore

Highlights

- Factors associated with longer length of stay were older age, female gender, non-Chinese ethnicity, unmarried status, and surgery requiring augmentation among studied patients with Total Knee Replacement at KTPH.
- Recognition of these factors can enhance recovery protocols and better meet the needs of these patients.

Introduction

Total knee replacement (TKR) is a widely performed surgical procedure aimed at alleviating pain and enhancing mobility in patients with severe knee osteoarthritis. Monitoring the length of stay (LOS) for TKR is crucial for controlling healthcare costs, optimising care delivery, and ensuring the efficient use of hospital resources.¹ This study seeks to identify and examine the socio-demographic and surgical factors associated with varying LOS among patients undergoing primary TKR at Khoo Teck Puat Hospital (KTPH).

Methods

A retrospective analysis was conducted on patients who underwent primary TKR at KTPH across a 10-year period (2014 to 2023) using secondary, de-identified data from administrative databases. Patients with primary TKR were identified with standardised procedure codes [SB716K – Primary TKR (Unilateral) with implant/bone augmentation; SB810K – Primary TKR (Unilateral)] using institutional surgical procedures data. Patients with bilateral knee surgeries, revision TKR, emergency and ambulatory surgery cases were excluded from the analysis. Unique patient identifiers were then used to link to other institutional databases for socio-demographic (age, gender, ethnicity, marital status, housing type, referral source) factors and surgical episode LOS, which were extracted and analysed. We used a Poisson model to examine LOS as the dependent variable and investigated its associations with variables such as socio-demographic factors and surgical types.

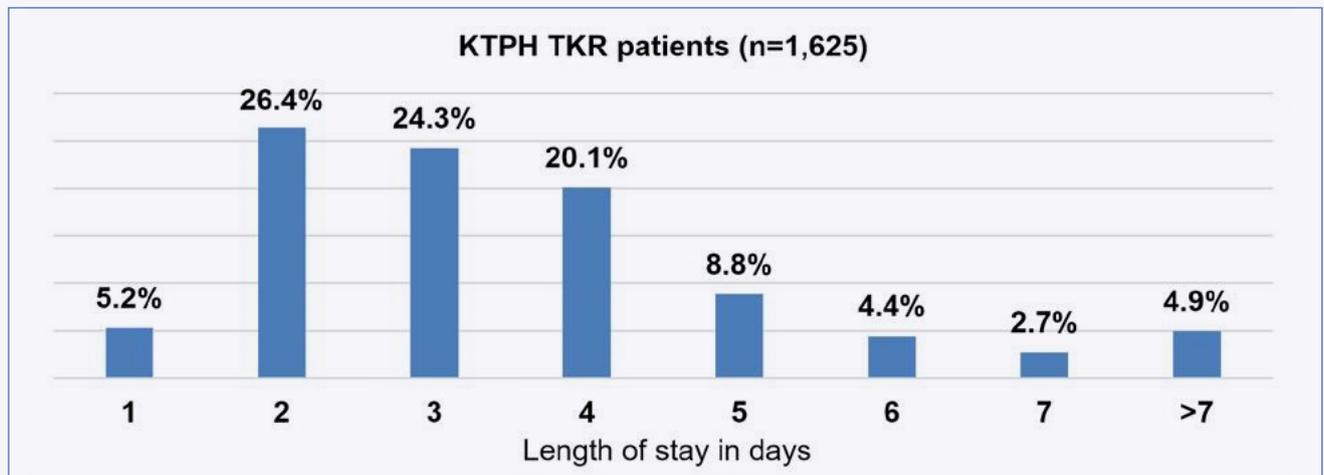
Results

A total of 1,625 patients with a mean age of 66.0 years [standard deviation (SD) = 7.8], and largely females (n=1,100; 67.7%), were included in the analysis. Other socio-demographic factors are shown in **Table 1**. Unilateral TKR (SB810K) constituted the main surgical procedure type (n=1,553; 95.6%), and most patients received one TKR procedure during the timeframe of the study (n=1,280; 78.8%). The mean LOS was 3.7 days (SD = 2.5). A breakdown of LOS categories is shown in **Figure 1**.

In our main model examining the socio-demographic and surgical factors associated with varying LOS, we found that factors significantly associated with longer LOS in patients with TKR included older age [Incident Risk Ratio (IRR) = 1.006; 95% Confidence Interval (CI): 1.003 – 1.010; p=0.001], female gender (IRR=1.119; 95% CI: 1.054 – 1.188; p<0.001), non-Chinese ethnicity (IRR=1.064; 95% CI: 1.002 – 1.130; p=0.042), unmarried status (IRR=1.118; 95% CI: 1.028 – 1.216; p=0.010), and primary TKR with augmentation (IRR=1.288; 95% CI: 1.142 – 1.453; p<0.000).

Table 1. Socio-demographic characteristics of study patients

Characteristic	Study population: n=1,625
Age in years [mean (standard deviation)]	66.0 (7.8)
Gender – Female [n (%)]	1,100 (67.7%)
Ethnicity [n (%)]	
<i>Chinese</i>	1,080 (66.5%)
<i>Malay</i>	287 (17.7%)
<i>Indian</i>	145 (8.9%)
<i>Others</i>	113 (7.0%)
House type [n (%)]	
1-/2-room flat	30 (1.9%)
3-/4-room flat	902 (55.5%)
5-Room/Executive/Multi-generational flat	378 (23.3%)
Private Housing	315 (19.4%)
Marital status [n (%)]	
Married	1,272 (78.3%)
Non-married	158 (9.7%)
Missing	195 (12.0%)
Referring source for surgery [n (%)]	
Specialist Outpatient Clinics (Subsidised)	1,486 (91.5%)
Specialist Outpatient Clinics (Private)	131 (8.1%)
Others	8 (0.5%)

Figure 1. Breakdown of length of stay, across duration categories

Conclusion

This study identified key socio-demographic and surgical factors associated with longer LOS in patients undergoing primary TKR at KPTH. Older age, female gender, non-Chinese ethnicity, unmarried status, and the need for surgical augmentation were all significantly associated with longer LOS. Understanding these factors can aid clinicians and healthcare administrators in developing targeted strategies to minimise LOS and enhance postoperative recovery. This may be achieved by optimising preoperative conditions for older adults, providing robust social support for unmarried individuals, offering culturally sensitive care for individuals of non-Chinese ethnicity, and conducting preoperative planning for patients requiring augmentation.

Reference

¹ Molloy, I. B., Martin, B. I., Moschetti, W. E., & Jevsevar, D. S. (2017). Effects of the Length of Stay on the Cost of Total Knee and Total Hip Arthroplasty from 2002 to 2013. *The Journal of bone and joint surgery. American volume*, 99(5), 402–407. <https://doi.org/10.2106/JBJS.16.00019>.



Identifying Groups of Healthcare Expenditure Trajectories in Decedents with End-stage Organ Diseases

Sheryl Ng Hui Xian, Palvinder Kaur, Dr Laurence Tan Lean Chin¹,
Dr Mervyn Koh Yong Hwang², Dr Andy Ho^{3,4,5}, Dr Allyn Hum^{5,6}, Dr Tan Woan Shin⁷

¹ Department of Geriatric Medicine, Khoo Teck Puat Hospital, Singapore

² Dover Park Hospice, Singapore

³ Department of Psychology, School of Social Sciences, Nanyang Technological University, Singapore

⁴ Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore

⁵ The Palliative Care Centre for Excellence in Research and Education, Singapore

⁶ Department of Palliative Medicine, Tan Tock Seng Hospital, Singapore

⁷ Geriatric Education & Research Institute, Singapore

Highlights

- Decedents with end-stage organ diseases had one of three trajectories of healthcare expenditure over their last five years of life: consistently low, moderately high near death and escalating near death.
- High acute care use in the final months of life highlights the pertinence of timely access to needs assessment and care planning at the end of life to mitigate aggressive care use.

Introduction

Patients with end-stage organ disease (ESOD) live with a disease trajectory of intermittent exacerbations and recovery before death, often resulting in high healthcare costs in the final year of life. Studies on end-of-life (EOL) healthcare expenditure (HCE) trajectories over one to five years before death have identified distinct cost trajectories, e.g. persistently high, or moderately high cost. However, few studies have characterized resource use of patients with ESOD prior to their final year. A detailed understanding of cost trajectories among patients with ESOD will be useful to inform the review of current care processes and targeted intervention. Hence, we aimed to identify distinct HCE trajectories among patients with ESODs over five years prior to their death.

Methods

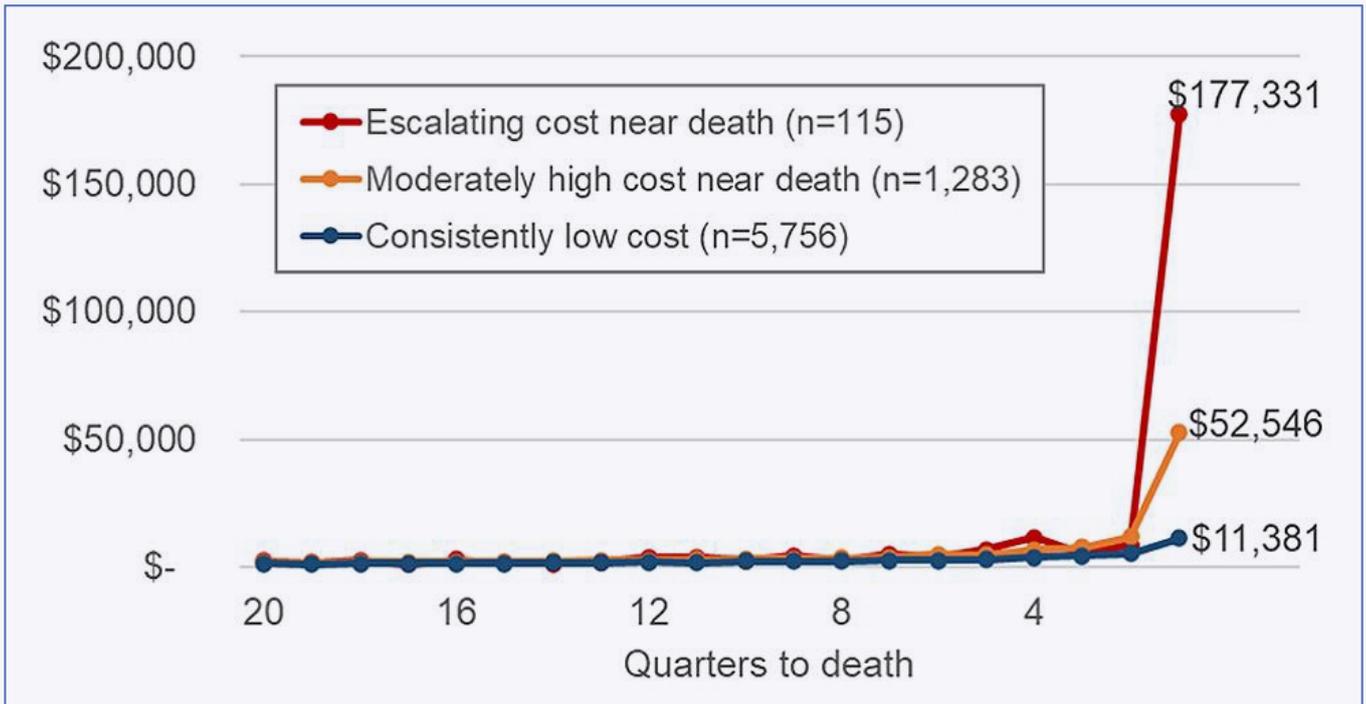
We conducted a retrospective cohort study of decedents who died between 2017 and 2019, and were diagnosed with either advanced dementia, severe liver disease, or any of heart, renal, or respiratory failure in the last five years of their lives. HCE was defined as the gross charge to the patient prior to governmental subsidy for each hospital encounter, across inpatient, specialist outpatient, emergency and day procedure settings in Tan Tock Seng Hospital and Khoo Teck Puat Hospital. This was then aggregated by quarter, or 90 days, for the entire five-year period prior to the patient's death.

We applied longitudinal k-means estimation to cluster patients into groups based on the similarity of their HCE trajectories. We tested the clustering of patients into two to seven groups and identified the number of groups which yielded the best performance across multiple criterion scores. We then profiled the patients in each group in terms of age, ESOD prevalence and EOL acute care use.

Results

The 7,154 decedents on average incurred \$70,608 in HCE over the last five years of life. They were classified into three groups based on their cost trajectories: consistently low cost (LC), moderately high cost near death (MC), and escalating cost near death (EC) (**Figure 1**). The group with LC was the largest, followed by the groups with MC and EC (LC: n=5,756, 80%; MC: n=1,283, 18%; EC: n=115, 2%). In the group with LC, most patients were aged ≥ 75 years (n=2,142, 74%), and four in 10 had a diagnosis of advanced dementia. In the group with MC, half of patients were aged ≥ 75 years (n=701, 55%), or had multiple ESODs (n=637, 50%). In the group with EC, most patients were aged < 75 years (n=79, 69%), or had respiratory failure (n=65, 57%). Across all groups, most patients had at least one emergency attendance in the final three months of life and had late or no access to palliative care (**Table 1**). Moreover, half of patients in the group with MC and almost all patients in EC had at least one admission requiring intensive or high dependency care.

Figure 1. Average quarterly healthcare expenditure over the last five years of life, by group



Dollar amounts reported in the figure reflect the average HCE incurred in the final three months of life among patients in each group.

Table 1. Descriptives by group

	LC (n=5,756, 80%) n (%)	MC (n=1,283, 18%) n (%)	EC (n=115, 2%) n (%)
End-stage organ disease			
Advanced dementia	2,576 (44.8%)	381 (29.7%)	13 (11.3%)
End-stage renal failure	1,587 (27.6%)	437 (34.1%)	50 (43.5%)
End-stage heart failure	1,810 (31.4%)	579 (45.1%)	46 (40.0%)
Respiratory failure	1,457 (25.3%)	553 (43.1%)	65 (56.5%)
Severe liver disease	610 (10.6%)	202 (15.7%)	16 (13.9%)
Acute care use in final 3 months of life			
≥ 1 emergency attendance	4,758 (82.7%)	1,185 (92.4%)	85 (73.9%)
≥ 1 intensive / high dependency care admission	705 (12.2%)	654 (51.0%)	103 (89.6%)
Underwent endotracheal intubation	285 (5.0%)	226 (17.6%)	59 (51.3%)
Access to palliative care			
Ever referred	1,219 (21.2%)	482 (37.6%)	57 (49.6%)
Days from first known referral to death, median (Q1-Q3)	15.0 (3.0-114.5)	9.0 (3.0-40.8)	12.0 (4.0-101.0)

EC: escalating cost near death; MC: moderately high cost near death; LC: consistently low cost
 Q1: 25th percentile; Q3: 75th percentile

Conclusion

Decedents with ESODs had one of three HCE trajectories, differentiated by the extent of increase in HCE in the final months of life. Acute care use was high in the final months, with low access to palliative care. Facilitating the timely referral of patients with ESODs for needs assessment and care planning will be essential to ensure they receive care appropriate to their needs and preferences in the final months and mitigate aggressive or unnecessary treatment at the EOL.



The Impact of Methodologies on Identifying Multimorbidity Clusters in Heart Failure: A Systematic Review

Palvinder Kaur, Joey Ha Wei Yee, Natalie Raye¹, Dr Wouter Ouwerkerk^{2,3},
Dr Bart J. van Essen⁴, Dr Laurence Tan⁵, Dr Tan Chong Keat⁶, Dr Allyn Hum^{7,8},
Dr Alex R Cook^{1,9}, Dr Jasper Tromp^{1,9}

¹ Saw Swee Hock School of Public Health, National University of Singapore and National University Health System, Singapore

² National Heart Research Institute Singapore, National Heart Centre Singapore, Singapore

³ Department of Dermatology, University Medical Center Amsterdam, Netherlands

⁴ Department of Cardiology, University of Groningen, University Medical Centre Groningen, Groningen, The Netherlands

⁵ Geriatric Medicine, Khoo Teck Puat Hospital, Singapore

⁶ Department of Cardiology, Tan Tock Seng Hospital, Singapore

⁷ Palliative Care Centre for Excellence in Research and Education, Singapore

⁸ Department of Palliative Medicine, Tan Tock Seng Hospital, Singapore

⁹ Duke-NUS Medical School, Singapore

Highlights

- We identified significant gaps in analysis quality and reporting among studies that performed clustering of patients with heart failure. Implementing a methodological framework to develop, validate, and report clustering analyses can improve the clinical utility and reproducibility of clustering outcomes.

Introduction

Heart failure (HF) is a complex condition, and clustering algorithms offer a valuable tool for identifying distinct patient subgroups to inform the development of phenotype-guided therapy. However, the choice of algorithms, modelling process, and input variables can impact clustering outcomes. Therefore, we reviewed analytical methods and variables used in studies that performed clustering in patients with HF.

Methods

We systematically searched CINAHL, COCHRANE, EMBASE, OVID Medline, and Web of Science for eligible articles between inception and April 2023. We included primary studies that identified distinct HF multimorbid subgroups and were appraised for risk-of-bias and against methodological recommendations for cluster analysis. A narrative synthesis was performed.

Results

Our analysis included 43 studies, mostly following a cohort design (n=34) and conducted primarily in Europe (n=15) and North America (n=13). Model-based (n=23), centre-based (n=10), and hierarchical class clustering (n=9) were the most frequently employed algorithms, identifying a range of 2-10 multimorbid clusters. As input variables for clustering, most studies used a combination of multiple parameters (i.e., socio-demographics, biochemistry, clinical characteristics, comorbidities and risk factors, cardiac imaging, and biomarkers) (n=27). The phenogroups identified differed by input variables and algorithms used. We identified notable methodological weaknesses in the development of clustering models, variable selection, reporting of the modeling process, and model validation (**Figure 1**).

Conclusion

Cluster analysis results differed based on the clustering algorithms used and input variables. This review found significant gaps in analysis quality and reporting. Implementing a methodological framework to develop, validate, and report clustering analysis, can improve the clinical utility and reproducibility of clustering outcomes.

Figure 1. Framework to assess quality of clustering methodology used

Model types	Centre-based models	Hierarchical clustering	Model-based	Network
Description of clustering processes	Ahmad 2018			
	Horiuchi 2018			
	Tromp 2018b			
	Zikos 2019			
	Harada 2020			
	Steinen 2020*			
	Perry 2021			
	deLange 2022			
	Urban 2022			
	Nakano 2023			
	Ahmad 2014			
	Bose 2018			
	Schrub 2020			
	Casebeer 2021			
	Gevaert 2021			
	Gu 2021			
	Woolley 2021			
	Fayol 2022			
	Murray 2022*			
	Lee 2014			
Kao 2015				
Shah 2015				
Ferreira 2018				
Tromp 2018a				
Wiley 2018				
Gimeno-M 2019				
Cohen 2020				
Hedman 2020				
Segar 2020				
Gulea 2021				
Nguyen 2021				
Sison 2021				
Ujil 2021				
Choy 2022				
Sotomi 2022				
Thompson 2022				
Zheng 2022*				
Blum 2023				
Meijs 2023				
Nakamaru 2023				
Carmona-P 2021				
Reza 2022				
Uszko-L 2022				
1. Establishing research question				
1.1 Presented rationale for clustering				
2. Data pre-processing				
2.1 Described representativeness of the data				
2.2 Provided clear rationale for candidate variable selection				
2.3 Described nature of the candidate variables				
2.4 Described feature selection methods				
2.5 Performed missing data analysis				
3. Pre-clustering testing				
3.1 Evaluated data for multimodality and randomness				
4. Clustering				
4.1 Described modeling techniques used				
4.2 Reported model parameters implemented				
4.3 Transparency in modeling processes				
5. Evaluation of model, visual exploration and explanation of the results				
5.1 Presented evaluation criteria to select optimal model				
5.2 Visual exploration of the results				
5.3 Presented final model				
5.4 Interpretation of final model				
5.5 Performed internal validation				
6. External validation				
6.1 Type of external validation performed				
6.2 Compared characteristics of derivative and validation cohort				
6.3 Reported external fit statistics				

Requirements met
 Requirements partially met
 Requirements not met
 * Used >1 methods



NHG Knee Arthroscopy Patients: 10-year Trends and Factors Associated with Longer Stay Among Older Patients

Chen Wei Ling¹, Estelle Foo Jie Wei, Christopher Ng Thong Lian²,
Mohd Khalid Bin Mohd Anuar^{3,4}, Dr Shiek Abdullah Bin Ismail⁵, Ivan Tan Peng Choon⁶,
Dr Michelle Jessica Pereira

¹ Saw Swee Hock School of Public Health, National University of Singapore, Singapore

² Department of Physiotherapy, Tan Tock Seng Hospital, Singapore

³ Allied Health, Tan Tock Seng Hospital, Singapore.

⁴ Group Allied Health, National Healthcare Group, Singapore

⁵ Rehabilitation Service, Khoo Teck Puat Hospital, Singapore

⁶ Value Based Care, National Healthcare Group, Singapore

Highlights

- Knee arthroscopy demand and length of stay in NHG hospitals have generally reduced in the past decade, but patients have appeared younger. However, the procedural charges for day surgery and inpatient patients had year-on-year mean increases.
- Among patients aged 50 and above, those who were unmarried and in private-paying wards had a longer length of stay.

Introduction

Knee arthroscopies (KAs) are minimally invasive procedures for diagnoses or repair. Our main aim was to profile demographic, care path-related and utilisation factors of NHG KA patients across 10 years of data. Additionally, as older patients often have higher comorbidity burdens, slower recovery, and greater post-operative needs leading to prolonged hospitalisation compared to younger patients, we studied patients aged 50 and above more closely, delving deeper into similar non-clinical factors associated with different length of stay (LOS).

Methods

We conducted a retrospective cohort study using de-identified, administrative data. Patients seen in Tan Tock Seng Hospital (TTSH) and Khoo Teck Puat Hospital (KTPH) from 2014–2023 (SB700H: therapeutic KA+/-synovectomy; SB700K: KA with ligament reconstruction; SB808K: arthroscopic meniscectomy) were identified. We extracted demographic (age, gender, ethnicity), care path-related (referral source, discharge disposition) and utilisation factors {length of stay [LOS], costs [MOH healthcare Consumer Price Indices (CPI) 2019-adjusted levels], admission type [day surgery (DS), inpatient (IP)]}. We analysed yearly trends and descriptive statistics of KA patients' data by procedure year and admission type.

For our subgroup analysis, we identified KA patients who were aged 50 and above (excluding emergency admissions). We then split these patients into short stay (LOS = 1 day) and long stay (LOS > 1 day) groups, before profiling them based on similar non-clinical factors of socio-demographic and care path-related measures. We performed a logistic regression to study the associations of demographic and care path-related factors with different LOS (short versus long stay) among these patients.

Results

For our main analysis, among 1,196 patients (TTSH=856, KTPH=340) with a mean age of 41.0 years (SD=15.7 years), most were male (66.2%) and Chinese (56.0%). Majority (93.5%) were referred from specialist outpatient clinics (SOCs) and discharged home +/- SOC follow-ups (97.6%). KA-associated all-cause mortality was low (in-institution=0.3%, 30-day=0.2%). KA patients were younger over the years of the study (year-on-year mean age decrease: 0.2 years).

Therapeutic-type (SB700H) KAs (62.0%) were the majority (SB700K=20.2%, SB808K=17.8%). KA numbers decreased 11.1% annually on average (**Figure 1**). KAs from 2014–2022 were mainly IP procedures (mean-annual%=78.3%), but in 2023, 69.3% were DS. IP-KAs LOS decreased over the study period (year-on-year mean LOS decrease: 0.09 days) except 2023. IP and DS KAs' charges had year-on-year mean increases of 19.9% and 13.7%, respectively.

Figure 1. 2014 – 2023 NHG Knee Arthroscopy volume, length of stay and charges.



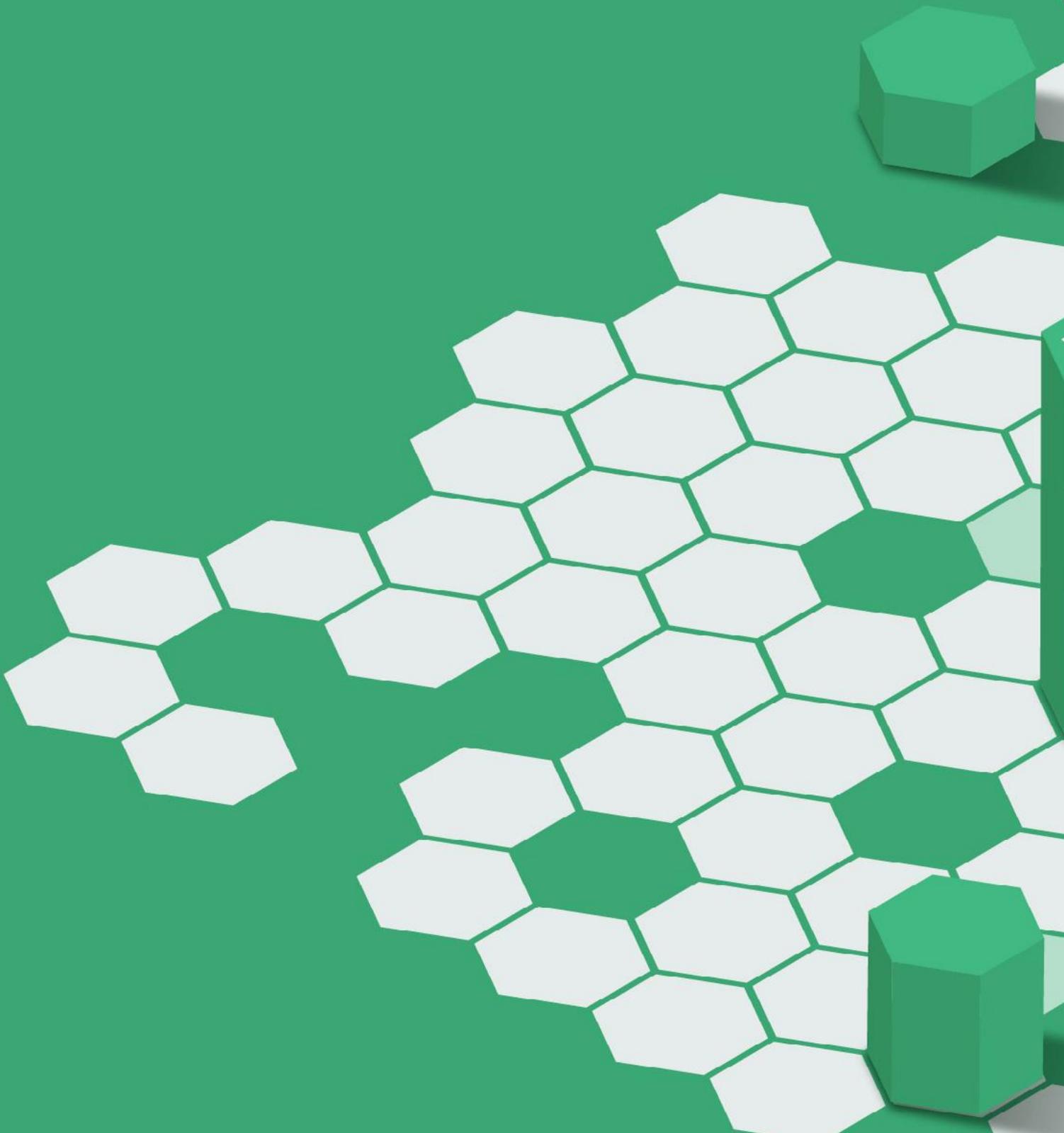
For the subgroup analysis, among 358 patients aged 50 and over, most had a short stay (n=312; 87.2%). We were unable to include education level, employment status and discharge disposition due to low data availability. We found those who were unmarried were more likely to have a long stay compared to those that were married (Odds ratio {OR} = 3.48, 95% CI = 1.45-8.34, p-value = 0.005) and those that were in private-paying wards were more likely to have a long stay than those in subsidised wards (OR = 7.10, 95% CI = 2.22-22.62, p-value = 0.001).

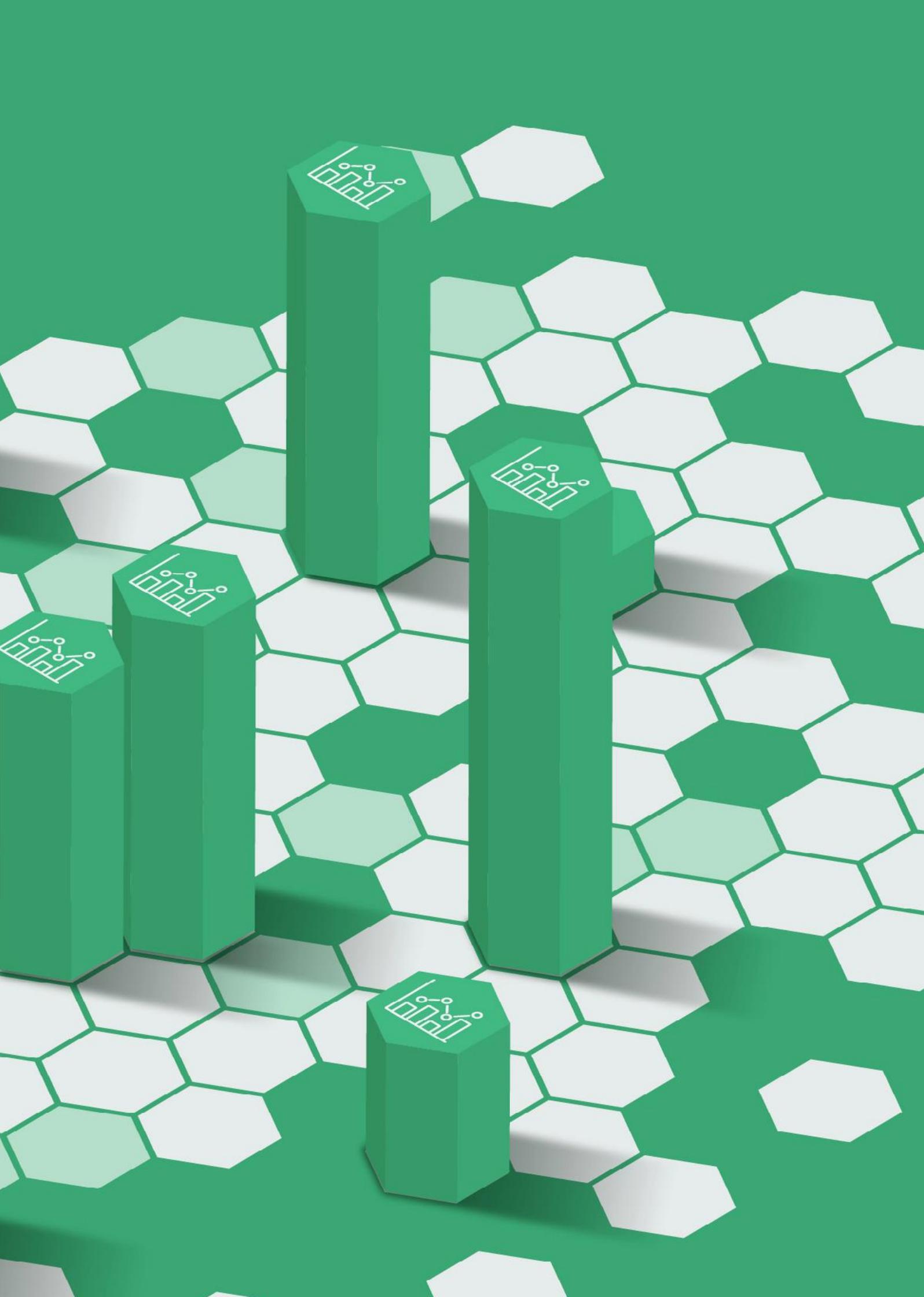
Conclusion

NHG KA demand and length of stay has generally reduced in the past decade, potentially due to changes in surgical practices, but procedural charges have increased. To better examine age-related LOS, we investigated NHG KA patients aged 50 and above, and found that those who were unmarried or in private-paying wards stayed longer, potentially due to the lack of a strong social support system that can facilitate faster recovery and discharge or preference for longer stays when self-funding their procedures. Further studies can be done to examine drivers of increased procedural charges.



IDENTIFYING POPULATION TRENDS AND DETERMINANTS OF HEALTH







Population Projection for Lung Cancer Screening

Dr John Abisheganaden, Dr Yap Chun Wei

Highlights

- New lung cancer screening recommendations and guidelines were released in 2024 for Singapore.
- Using data from the National Population Health Survey (NPHS), we estimate that the number of CT scans performed annually will range from 40,000 to 100,000.

Introduction

In 2024, the Academy of Medicine, Singapore, released national recommendations for lung cancer screening. The key recommendations are:

- Annual low-dose CT scans of the thorax may be offered to individuals aged 50 to 80 years who have a smoking history of at least 20 pack-years and are either current smokers or have quit smoking within the last 15 years.
- Screening should not be offered to non-smokers.
- Annual screening should be discontinued for individuals who have quit smoking for more than 15 years, who become medically unfit, or who are unwilling to undergo invasive evaluations or anti-cancer treatments.

This study aims to estimate the number of CT scans that will be performed for lung cancer screening in Singapore.

Methods

We analysed data from all Singaporeans and permanent residents alive as of January 1, 2022, using the National Population Health Staple Database. We excluded residents diagnosed with lung cancer (C33, C34) in 2021 or earlier. The primary outcome measure was the number of newly diagnosed lung cancer cases (C33, C34) in 2022. Covariates included the last known status of residents prior to their lung cancer diagnosis in 2022, if applicable, or their last known status for that year. The final study population consisted of 4,405,958 residents, with 1,828 newly diagnosed cases of lung cancer identified in 2022. Detailed information on risk factors for lung cancer, along with the number and percentage of residents with newly diagnosed lung cancer, is presented in Tables 1-2.

Results

Using the NPHS 2023 figures, we examined the current take-up rates for chronic disease screenings: approximately 63% for general chronic disease screening, 35% for breast cancer screening, 45% for cervical cancer screening, and 42% for colorectal cancer screening. Assuming a take-up rate for lung cancer screening between 30% to 60%, we estimate that the annual number of CT scans performed will range from 40,000 to 100,000.

Table 1. Risk factors for lung cancer and adjusted odds ratio

Risk factors	Adjusted odds ratio (95% Confidence Interval)	P value
Age		
0 to 49	Reference	
50 to 59	7.1 (5.6 – 9.0)	<0.001
60 to 74	15.0 (12.1 – 18.5)	<0.001
75 to 80	19.8 (15.7 – 25.1)	<0.001
81 and above	14.1 (11.1 – 18.0)	<0.001
Gender		
Female	Reference	<0.001
Male	1.4 (1.3 – 1.5)	
Race		
Chinese	Reference	
Malay	0.9 (0.7 – 1.1)	0.184
Indian	0.7 (0.5 – 0.9)	0.001
Other	0.7 (0.5 – 0.8)	0.001
Smoker	2.0 (1.7 – 2.2)	<0.001
Chronic Obstructive Pulmonary Disease	3.6 (3.0 – 4.2)	<0.001
History/Presence of other cancers	16.7 (15.2 – 18.4)	<0.001

COPD: Chronic Obstructive Pulmonary Disease

Table 2. Number and percentage of residents with newly diagnosed lung cancer in 2022

	No. of residents	No. (%) with newly diagnosed lung cancer in 2022
All residents	4,403,958	1,828 (0.04%)
Age 50 to 80	1,555,646	1,484 (0.10%)
Age 50 to 80, Smoker	106,519	260 (0.24%)
Age 60 to 80, Smoker	72,686	237 (0.33%)
Age 50 to 80, Smoker, COPD	7,597	67 (0.88%)
Age 60 to 80, Smoker, COPD	6,422	64 (1.00%)
Age 50 to 80, Smoker, History/Presence of other cancers	6,870	127 (1.85%)

COPD: Chronic Obstructive Pulmonary Disease

Conclusion

Lung cancer screening is now recommended both internationally and locally. The projected number of scans performed annually is substantial. The healthcare system must assess the resources required to support these recommendations and evaluate the cost-effectiveness of the screening services.



Healthspan of NHG Residents: A Comprehensive Study from 2016 to 2023

Dr Zhu Zhecheng, Dr Jonty Heaversedge¹, Palvannan R. K., Dr Ang Yee Gary, Dr Aidan Lyanzhiang Tan, Dr John Abisheganaden

¹Population Health Management, National Healthcare Group, Singapore

Highlights

- Healthspan measures the number of years spent in good health of a target population.
- Addressing precursors to prevalent and early-onset chronic diseases is pertinent to improve Healthspan estimates.

Introduction

Healthspan is a population health measure proposed by Outcomes Based Healthcare UK. It measures the number of years spent in good health, and the age at which people first enter a period of significant ill health. This study applied the Healthspan concept to National Healthcare Group (NHG) residents to explore its performance as a descriptive indicator of population health.

Methods

The Healthspan estimate of the NHG resident population was calculated as the median onset age of the first chronic disease across all residents. Chronic diseases were defined as those under the Chronic Disease Management Programme maintained by the Ministry of Health, as well as cancer. The age of disease onset was retrieved from NHG's Chronic Disease Management System. The Lifespan estimate was derived from Singapore's life expectancy data. All estimates were reported for the period of 2016 to 2023, and by gender and ethnic groups. Conditions commonly diagnosed as a first chronic disease were also highlighted.

Results

Figure 1 showed the Healthspan, Lifespan and Healthspan to Lifespan ratios of NHG residents from 2016-2023. It revealed a stable Healthspan and Healthspan to Lifespan ratio over the observed period.

Comparing subgroups, female residents demonstrated a longer Healthspan compared to males but had a lower Healthspan to Lifespan ratio due to their longer overall lifespan. Among ethnic groups, Chinese residents had the longest Healthspan, followed by Indians, with Malays exhibiting the shortest Healthspan, and a decreasing trend from 2016 to 2023. Central zones also showed longer Healthspan than northern zones.

Figure 2 illustrated that dyslipidaemia, hypertension, diabetes, allergic rhinitis, and cancer were the most common initial diagnoses leading towards the end of Healthspan in 2023. Younger residents (aged 0-40) were more likely to be diagnosed with allergic rhinitis and mental health issues, while older age groups (aged 41 and above) predominantly faced dyslipidaemia, hypertension, diabetes and cancer. The inclusion of diseases with high prevalence and incidence and early onset age, such as allergic rhinitis, asthma, and depression, significantly impacted Healthspan estimates.

Figure 1. Healthspan estimates of NHG residents from 2016 - 2023

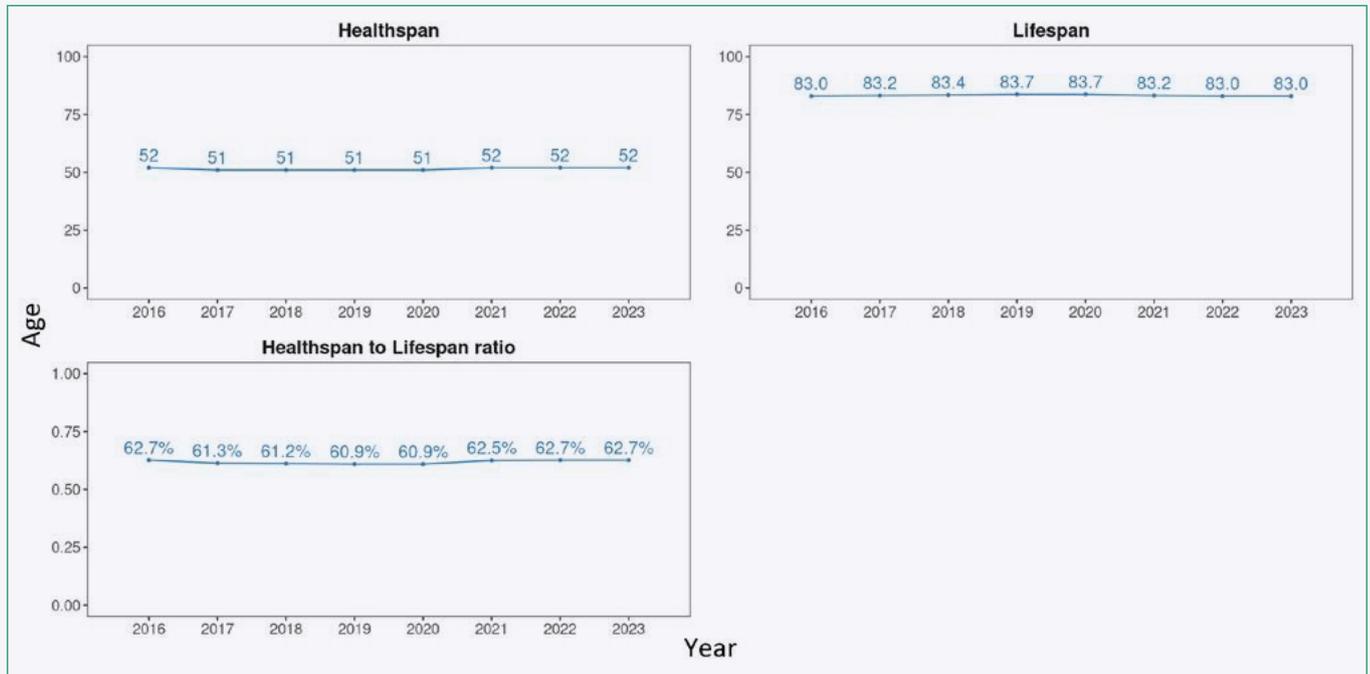
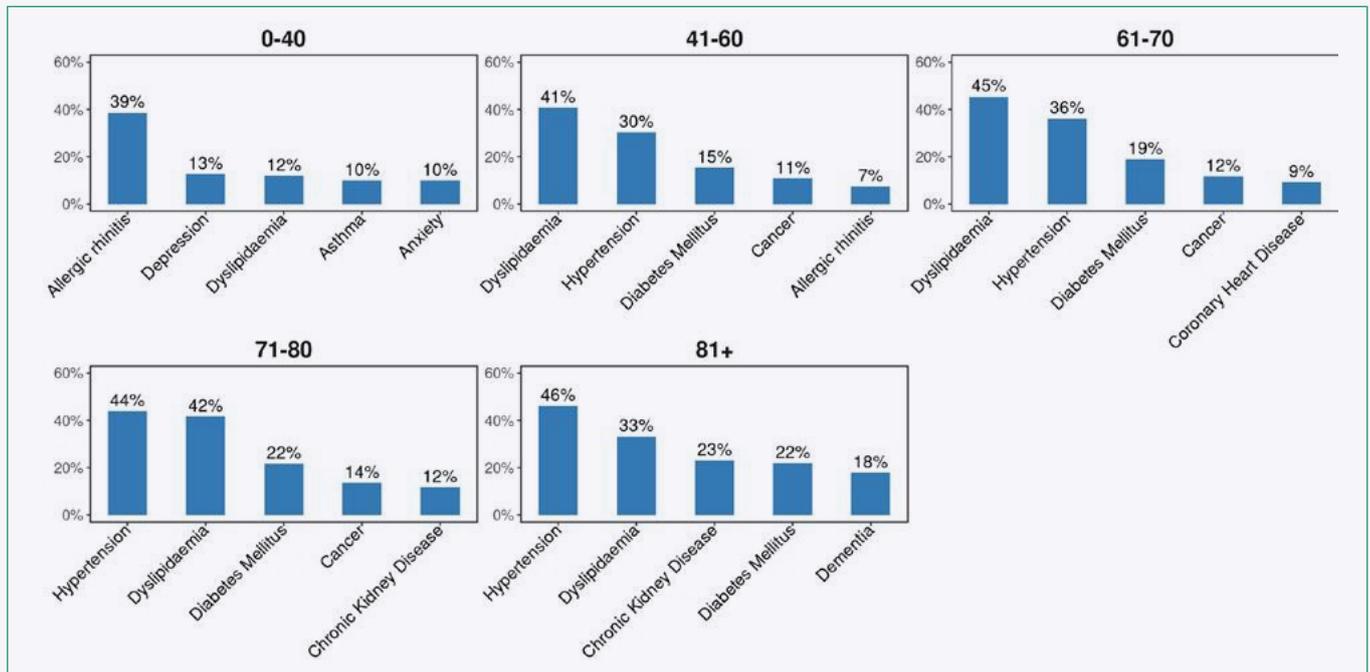


Figure 2. Five most frequent first diagnoses by age group in 2023



Conclusion

This study demonstrated the importance of addressing precursors to prevalent and early-onset chronic diseases, to improve Healthspan of the NHG population and achieve the goal of adding 15 million years of healthy life.

Acknowledgements

We would like to extend our gratitude to Outcomes Based Healthcare for their innovative concept and invaluable works of Healthspan, which is the foundation of our piece of work.



Associations between OGTT curve Morphology and Physiological, Metabolic, and Exercise Profiles in Healthy Adults: A Cross-sectional Analysis

Dr Meng Fanwen, Dr Ang Yee Gary, Palvannan R.K., Margaret Mei Chan Yap¹, Dr Shuen Yee Lee¹, Dr Melvin Khee Shing Leow^{1,2,3}, Dr Chin Leong Lim¹

¹ Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore

² Department of Endocrinology, Tan Tock Seng Hospital, Singapore

³ Cardiovascular and Metabolic Disorders Programme, Duke-NUS Medical School, Singapore

Highlights

- Significant differences found in biomarkers between monophasic and biphasic glucose curves, sedentary and active, females and males in healthy individuals.
- Healthy subjects with biphasic responses had lower insulin resistance and better beta cell function.
- Sedentary individuals had higher insulin resistance than active individuals.
- More active subjects exhibited biphasic curves, suggesting a lower risk of type 2 diabetes.

Introduction

Type 2 diabetes mellitus (T2DM) is a chronic metabolic disease characterized by impaired insulin secretion and insulin resistance. Commonly used measures of insulin resistance include homeostatic model assessment for insulin resistance (HOMA-IR), Matsuda index (MI), and triglycerides/high-density lipoprotein (TG/HDL) ratio, while insulin secretion is assessed using homeostatic model assessment for beta cell function (HOMA-B). However, these indices are mostly used in research rather than in clinical practice. The oral glucose tolerance test (OGTT) is a gold standard used by doctors to diagnose diabetes and impaired glucose tolerance. Recent studies indicate the shape of OGTT curves not only correlates with beta cell function and insulin resistance, but also differentiates diabetes risk, where individuals with biphasic curve less likely to develop T2DM than those with a monophasic curve. This study aims to unravel an association between the shape of glucose response curves during OGTT and indices of insulin resistance, beta cell function, metabolic phenotype, exercise profile, gender, and age.

Methods

We analysed data of active and sedentary adults clinically assessed to be healthy as detailed below.

Inclusion criteria	<ul style="list-style-type: none">• more than 18 years of age• not undergoing treatment for chronic disease, e.g., metabolic, cardiovascular, and pulmonary disease• not taking oral contraceptives or undergoing hormone replacement therapy• received medical clearance from appointed service providers• non-smokers
Exclusion criteria	<ul style="list-style-type: none">• the contraindications for the above

None were obese or in a high-risk category of BMI (body mass index) for Asians. Their physical activity profiles were based on a self-declared five-year history. Physical activity was defined as engaging in three or more sessions per week of moderate to vigorous-intensity exercises, for at least 45 minutes per session. Sedentary behaviour was defined as engaging in two or less sessions per week of light-intensity exercises, lasting no more than 30 minutes per session. Study subjects took the two-hour OGTT after an overnight fast, involving five time-points of blood samples at 0, 30, 60, 90, 120 minutes after ingestion of the glucose solution. Glucose response curve morphology was categorized based on stationary points of fitted smoothing curves. Insulin resistance and beta-cell function were assessed using HOMA-IR, MI, TG/HDL ratio, HOMA-B calculated based on published formulae. For categorical variables, frequencies and proportions were presented and Pearson's χ^2 test was applied to compare between specific groups. For continuous variables, means (standard deviations) were presented and unpaired Student's t test (non-parametric Wilcoxon test for the sample size being less than 30) was used. Statistical significance level was taken as $p < 0.05$.

Results

Among 157 healthy adults (50.3% females, 82.8% Chinese) with 86 being physically active (54.8%) and 71 sedentary (mean age 47.4 and 44.7, respectively), monophasic (76.4%) and biphasic (21.7%) glucose response curves predominated (**Figure 1**), which is in alignment with the findings in literature. 3 adults (who had either triphasic or unclassified response) were excluded from analysis. A higher proportion (24.4%) of active subjects had biphasic curves, compared to sedentary subjects (18.3%). Odds of a biphasic response were 1.5 times higher in the active group compared to the sedentary group. Biphasic responders exhibited lower insulin resistance, better beta-cell function, and significant differences in TC/HDL, LDL/HDL, and SBP compared to monophasic responders. Among younger individuals (age < 50) with monophasic responses, sedentary and active groups differed significantly in MI, TG/HDL, and HOMA-B. Furthermore, significant differences in fasting glucose, BMI, HDL, TC/HDL, LDL/HDL, TG, TG/HDL, and SBP were observed between genders (**Table 1**).

Figure 1. Monophasic and biphasic glucose curves and their distribution during oral glucose tolerance test

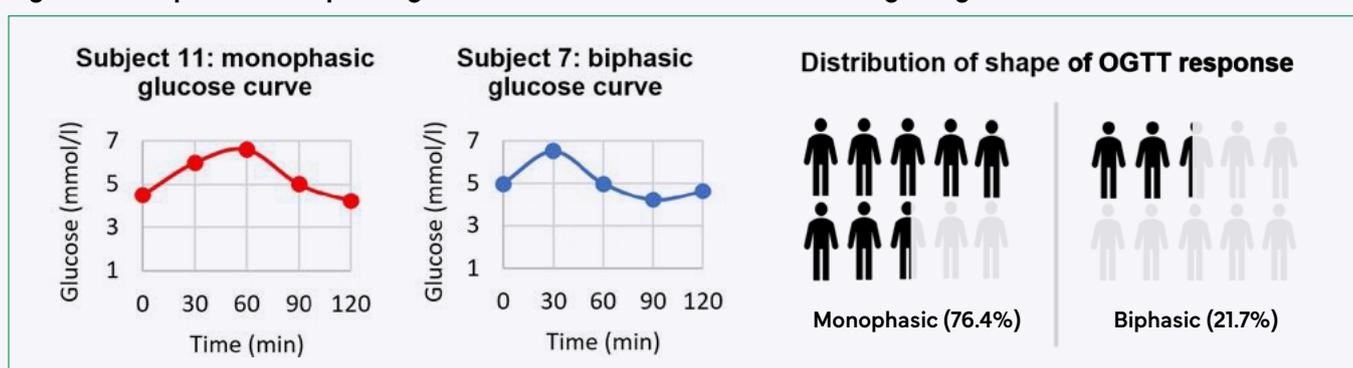


Table 1. Univariate comparisons of the study cohort by gender

Variables	Total (N=154)	Female (N=77)	Male (N=77)	p-value
Fasting glucose (mmol/l) mean(sd)	5.01 (0.41)	4.93 (0.41)	5.09 (0.41)	0.015
BMI (kg/m ²) mean(sd)	22.16 (2.82)	21.56 (2.39)	22.77 (3.08)	0.007
TC (mmol/l) mean(sd)	5.31 (0.98)	5.39 (0.86)	5.23 (1.09)	0.323
HDL (mmol/l) mean(sd)	1.75 (0.49)	1.9 (0.48)	1.59 (0.44)	< 0.001
LDL (mmol/l) mean(sd)	3.08 (0.78)	3.05 (0.65)	3.1 (0.89)	0.672
TC/HDL mean(sd)	3.2 (0.86)	2.96 (0.69)	3.45 (0.95)	< 0.001
LDL/HDL mean(sd)	1.86 (0.65)	1.69 (0.57)	2.02 (0.69)	0.002
TG (mmol/l) mean(sd)	1.06 (0.56)	0.96 (0.43)	1.17 (0.65)	0.018
TG/HDL mean(sd)	0.71 (0.58)	0.56 (0.36)	0.86 (0.71)	0.001
SBP (mmHg) mean(sd)	114.35 (10.93)	111.88 (10.58)	116.82 (10.79)	0.005
DBP (mmHg) mean(sd)	69.64 (8.34)	68.95 (8.06)	70.32 (8.61)	0.307
HOMA-IR mean(sd)	1.28 (0.81)	1.27 (0.84)	1.29 (0.79)	0.889
Matsuda index mean(sd)	9.31 (5.91)	8.95 (5.35)	9.67 (6.44)	0.456
HOMA-B mean(sd)	78.76 (50.14)	84.56 (54.77)	72.96 (44.65)	0.152

Conclusion

In this study, our results revealed significant differences in biomarkers between monophasic and biphasic glucose curves. The higher proportion of biphasic curves in active subjects suggests that biphasic OGTT curve morphology may confer a lower risk of developing type 2 diabetes.



Associations of Educational Attainment and Readiness-to-learn with Self-reported Health

Dr Cheng San Chye, Loo May Eng

Highlights

- The associations between educational attainment and readiness-to-learn with self-reported health persist, even after accounting for a range of key information-processing skills, psychological, and social factors.

Introduction

Few studies have examined whether there is an association between readiness-to-learn and health, and whether the association of educational attainment with health persists after taking into account various key information-processing skills, psychological and social factors, including separating the independent effects of key information processing skills from educational attainment. Our study explores whether there is such a relationship, considering key information-processing skills (literacy, numeracy, problem-solving in technology-rich environment) and various psychological and social factors.

Specifically, we examine the following research questions:

- i) How do adults with different educational attainment differ in their self-reported health outcomes?
- ii) How do adults with different readiness-to-learn scores differ in their self-reported health outcomes?
- iii) To what extent are educational attainment and readiness-to-learn associated with self-reported health, if key information-processing skills, psychological and social factors are taken into account?

Methods

Our study uses data from Singapore's participation in Round 2 (2014-2015) of the Organization for Economic Cooperation and Development Programme for the International Assessment of Adult Competences¹ (PIAAC) 1st Cycle.

Our analyses focused on employed adults who participated in PIAAC (n=3989). Key measures included: (a) proficiency in literacy, numeracy, and problem-solving in a technology-rich environment (scores from 0 to 500, categorised into two proficiency levels); (b) Frequency of skill use at work and in everyday life (Likert scale of 1 to 5, coded as Never to Every day); (c) Participation in adult learning during the last 12 months (yes/no); (d) Readiness-to-learn (Likert scale of 1 to 5); (e) Educational attainment; (f) Self-reported health status (Likert scale of 1-5, coded as Excellent to Poor). For self-reported health status, "Excellent", "Very good", and "Good" were considered as favourable self-reported health status while "Fair" and "Poor" were considered as non-favourable.

Besides examining descriptive statistics, we also modelled the likelihood of non-favourable health status with logistic regression, with educational attainment and readiness-to-learn as key predictors, with and without accounting for various factors (e.g., key information-processing skills).

Results

The demographic characteristics of our study sample by self-reported health status are presented in **Table 1**. A lower percentage of adults who completed Bachelor's, Master's or Doctorate degrees, and adults with readiness-to-learn scores above the threshold reported non-favourable health status.

Table 1. Characteristics of study sample of employed adults by self-reported health status

Characteristics	Self-reported health status	
	Favourable (weighted, row %)	Non-favourable (weighted, row %)
Age group (years)		
16-24	83.5	16.5
25-34	81.3	18.7
35-44	78.6	21.4
45-54	70.9	29.1
55-65	68.1	31.9
Gender		
Male	75.9	24.1
Female	75.9	24.1
Educational attainment		
Lower secondary or less	61.4	38.6
Upper secondary	74.0	26.0
Post-secondary non-tertiary	73.9	26.1
Diploma	77.3	22.7
Bachelor's, Master's or Doctorate degree	84.0	16.0
Readiness-to-learn*		
<median score	70.1	29.9
≥median score	80.7	19.3

*Based on average score across all 6 items. The actual average score was used in the logistic regressions.

From logistic regression, we found that readiness-to-learn and educational attainment have significant and independent associations with self-reported health. After accounting for demographic characteristics and a range of key information-processing skills, social and psychological factors (e.g., proficiency in literacy, numeracy, and ability to solve problems in technology-rich environment, skill use at work/in everyday life, participation in adult learning, readiness-to-learn), educational attainment was still positively associated with self-reported health status. Adults who completed Bachelor's, Master's or Doctorate degrees were less likely to report having poor health than adults who completed Diplomas.

Conclusion

Readiness-to-learn and educational attainment have independent and statistically significant associations with self-reported health. The association between educational attainment and health persists, even after accounting for a range of key information-processing skills, psychological and social factors. Further research could include and examine health behaviors as part of understanding the effect of educational attainment on health.

Reference

¹ Organisation for Economic Cooperation and Development (OECD) (2016). Programme for the International Assessment of Adult Competencies (PIAAC), Singapore Public Use File (prgsgpp1.csv). Paris: OECD Publishing.



Associations between Psychosocial Factors and Health-related Quality of Life in Community-dwelling Adults: A 1-year Longitudinal Study

Joey Ha Wei Yee, Dr Michelle Jessica Pereira, Ge Lixia, Dr Aidan Tan Lyanzhiang, Dr Yap Chun Wei, Dr Joseph Antonio D. Molina

Highlights

- In a cohort study of community-dwelling adults residing in central Singapore, we found that 3% had worsened depressive symptoms, 4% became lonely, and 26% had worsened social isolation at the end of a year.
- Individuals with worsened severity of depressive symptoms and those who became lonely had decreased EQ-5D scores.

Introduction

Health-related quality of life (HRQoL) is a multidimensional measure for wellbeing. Understanding the dynamic relationship between psychosocial determinants of health and HRQoL is valuable in informing interventions. We studied the association between HRQoL and changes in depressive symptoms, loneliness, and social isolation, respectively.

Methods

We carried out a 1-year cohort study using retrospective anonymized data from a longitudinal population health survey in Central Singapore. The dependent variable was changes in HRQoL (EuroQoL five-dimensions – 5-level (EQ-5D)) at 1-year. The independent variables were 1-year changes in depressive symptom severity (Patient Health Questionnaire-9), loneliness status (UCLA Loneliness Scale – 3-item), and risk category of social isolation (Lubben Social Network Scale-6). Multivariate analyses were conducted using linear regressions, adjusting for chronic conditions (diabetes, hypertension and dyslipidemia); socio-demographics (age, gender, ethnicity, living arrangement, and marital status); and baseline EQ-5D levels.

Results

Data from 1,218 adults (≥ 21 years) were analysed. 56.7% were females with a mean age of 53.2 years. Baseline distributions of depressive symptoms, social isolation and loneliness are shown in **Figure 1**.

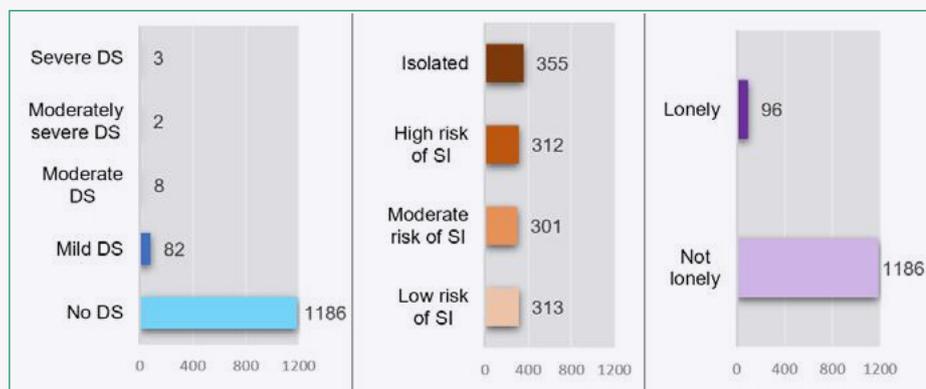


Figure 1. Baseline depressive symptoms (DS), social isolation (SI) and loneliness categories

At 1-year from baseline, 90.7%, 6.3% and 3.0% had maintained, improved, and worsened severity of depressive symptoms respectively; 90.4% remained lonely, 5.6% were less lonely, and 4.0% became lonely; and 44.7%, 29.2% and 26.1% had maintained, improved, and worsened social isolation respectively (**Figure 2**).

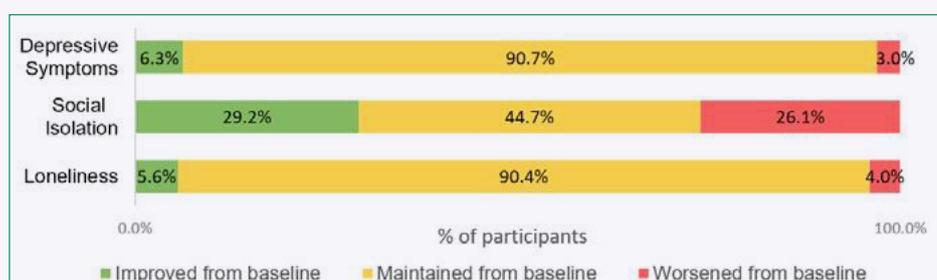
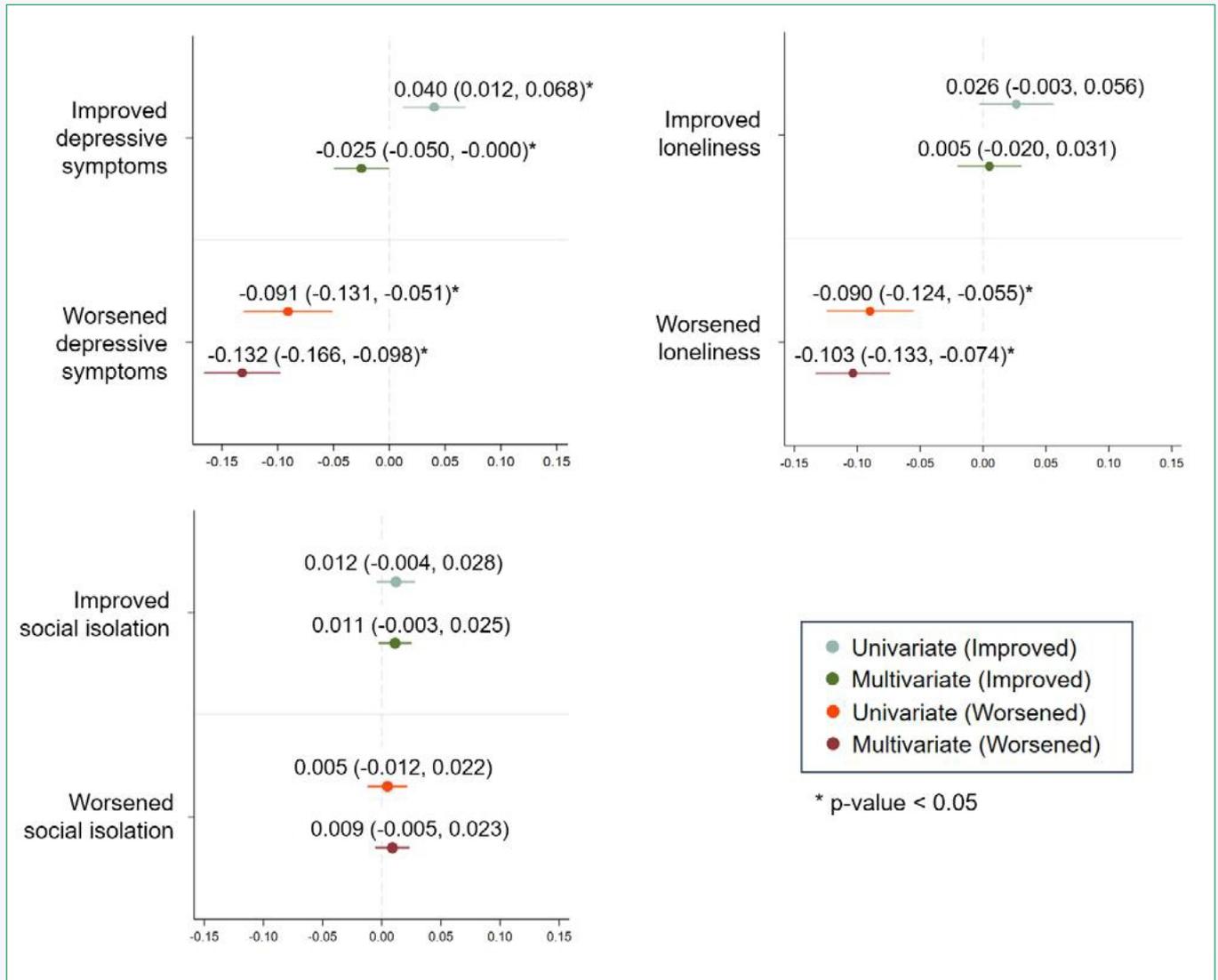


Figure 2. Changes in depressive symptoms, loneliness and social isolation categories

Individuals with worsened severity of depressive symptoms and those who became lonely had decreased EQ-5D scores (Worsened depressive symptoms: -0.13, 95% Confidence Interval: -0.17, -0.10; Became lonely: -0.10, 95% Confidence Interval: 0.13, -0.07), as presented in **Figure 3**.

Figure 3. Results of univariate and multivariate EQ-5D levels (95% Confidence Intervals shown)



Conclusion

We observed a decline in HRQoL among those with worsened depressive symptoms and those who became lonely. Worsening social determinants of health is negatively associated with HRQoL, and interventions targeting psychosocial factors could potentially contribute to better quality of life. To better assess the longer-term relationship between psychosocial factors and quality of life outcomes, modelling methodologies such as the Markov model may be explored in the future.



Validating a Self-report Scale for Assessing Health Resilience in Community-dwelling Adults: A Cross-sectional Study

Ge Lixia, Dr Yip Wan Fen, Dr Li Ruijie, Eric Chua Siang Seng, Dr Ringo Ho¹, Dr Andy Ho^{1,2}, Evon Chua³, Dolly Cheng⁴, Dr Ian Leong⁵, Chieh Pann Pei, Dr Tan Woan Shin

¹ School of Social Sciences, Nanyang Technological University, Singapore

² Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore

³ Population Health & Community Transformation, Yishun Health, Singapore

⁴ Office of Community Integration, Woodlands Health, Singapore

⁵ Continuing and Community Care, Tan Tock Seng Hospital, Singapore

Highlights

- The 22-item multi-dimensional Health Resilience Scale demonstrated good reliability and validity, making it an effective tool for assessing health resilience and guiding initiatives aimed to enhance well-being among community members.

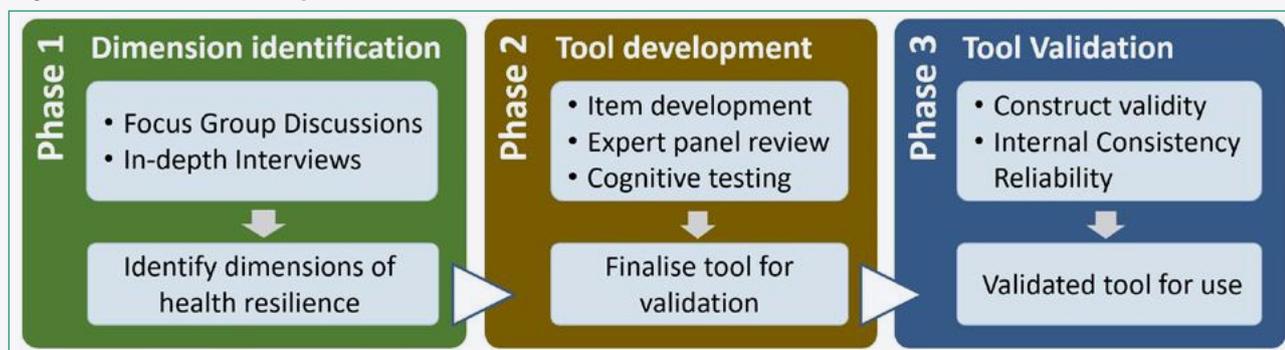
Introduction

Strengthening resilience is key to transforming Singapore's healthcare system into a sustainable care ecosystem. To effectively assess the collective efforts, a culturally relevant tool is essential for measuring health resilience of the population. Existing resilience measures lack specificity in health contexts and fail to comprehensively encompass the multi-level factors influencing individuals' ability to overcome health challenges. To address this gap, we developed and validated a multi-dimensional Health Resilience Scale (HRS) specifically tailored for community-dwelling adults in Singapore.

Methods

A multi-phase mixed-method approach was taken to develop and validate the HRS (**Figure 1**). The 35-item HRS refined at the end of Phase 2 was administered to 650 eligible community-dwelling adults in a cross-sectional survey, along with various validation measures to assess construct validity (including factorial, concurrent, convergent, and divergent validity) and internal consistency reliability. Exploratory factor analysis (EFA) was conducted to identify the underlying structure using 300 randomly selected responses through item reduction and factor extraction. Confirmatory factor analysis (CFA) was then performed to test dimensionality and confirm the factor structure using the remaining 350 responses. Concurrent, convergency and divergent validity of the HRS was examined by testing its correlation with validation measures. Hypothesis-testing validity was evaluated by assessing its relationship with self-reported physical and mental health status. Internal consistency reliability was measured using Cronbach's alpha.

Figure 1. Phases of development and validation of the HRS



Results

Approximately 47% of the participants were aged 51 years and above, with the majority being women (62.9%) and 72.8% being of Chinese ethnicity. Following an iterative process, a total of 22 items were retained, loading onto five distinct factors which accounted for 64.8% of the total variance. Based on the nature and scope of the items, the five factors were renamed to “Adaptive adjustment”, “Health mindset”, “Relational support”, “Perceived health access”, and “Social resourcefulness”. All 22 remaining items exhibited moderate to strong factor loading (>0.50) on their respective factors (**Table 1**). CFA confirmed the five-factor structure of the 22-item HRS with good model fit (RMSEA=0.05, CFI=0.94, TLI=0.93, SRMR=0.05).

Table 1. Component loadings and internal consistency reliability for five-factor solution

	<i>Adaptive adjustment</i>	<i>Perceived health access</i>	<i>Health mindset</i>	<i>Relational support</i>	<i>Social resourcefulness</i>
HRS item number	30, 32-35	12, 13, 15-17	1-5	24-26	10, 21-23
Loading (range)	0.62 - 0.85	0.57 - 0.88	0.56 - 0.78	0.88 - 0.92	0.80 - 0.89
Cronbach’s alpha	0.82	0.83	0.76	0.90	0.87

HRS: Health Resilience Scale

The HRS showed moderate correlation with the Connor–Davidson Resilience Scale, indicating good concurrent validity. It also exhibited moderate correlation with measures of hope and self-efficacy measured by the Herth Hope Index and General Self-Efficacy respectively, supporting its convergent validity. Conversely, the HRS demonstrated weak or non-significant correlations with maladaptive coping strategies and measures of depression and anxiety, providing evidence for its divergent validity (**Table 2**). Individuals reporting good/better physical and mental health status generally scored higher in HRS dimensions ($p<0.05$).

Table 2. Spearman correlation between HRS dimensions and validation measures

<i>HRS dimensions and total</i>	<i>Concurrent validity</i>	<i>Convergent validity</i>		<i>Divergent validity</i>			
	<i>CD_RISC10</i>	<i>HHI</i>	<i>GSE</i>	<i>Brief-COPE Emotion-focused coping</i>	<i>Brief-COPE Avoidant coping</i>	<i>PHQ-9</i>	<i>GAD-7</i>
Health mindset	0.40*	0.46*	0.39*	-0.04	-0.11*	-0.18*	-0.20*
Health access	0.35*	0.26*	0.41*	0.03	-0.09*	-0.20*	-0.13*
Social resourcefulness	0.08*	-0.01	0.09*	0.15*	0.03	-0.13*	-0.04
Relational support	0.21*	0.47*	0.05	-0.13*	-0.17*	-0.15*	-0.29*
Adaptive adjustment	0.40*	0.37*	0.30*	-0.05	-0.14	-0.11*	-0.24*
HRS total	0.44*	0.42*	0.37*	-0.001	-0.11*	-0.21*	-0.25*

Brief-COPE: Brief-Coping Orientation to Problems Experienced Inventory

CD-RISC10: Connor-Davidson Resilience Scale

GAD-7: General Anxiety Disorder

GSE: General Self-Efficacy

HHI: Herth Hope Index

HRS: Health Resilience Scale

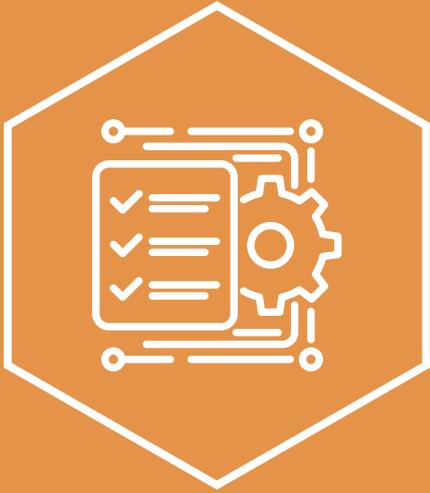
PHQ-9: Patient Health Questionnaire

* p -value<0.05

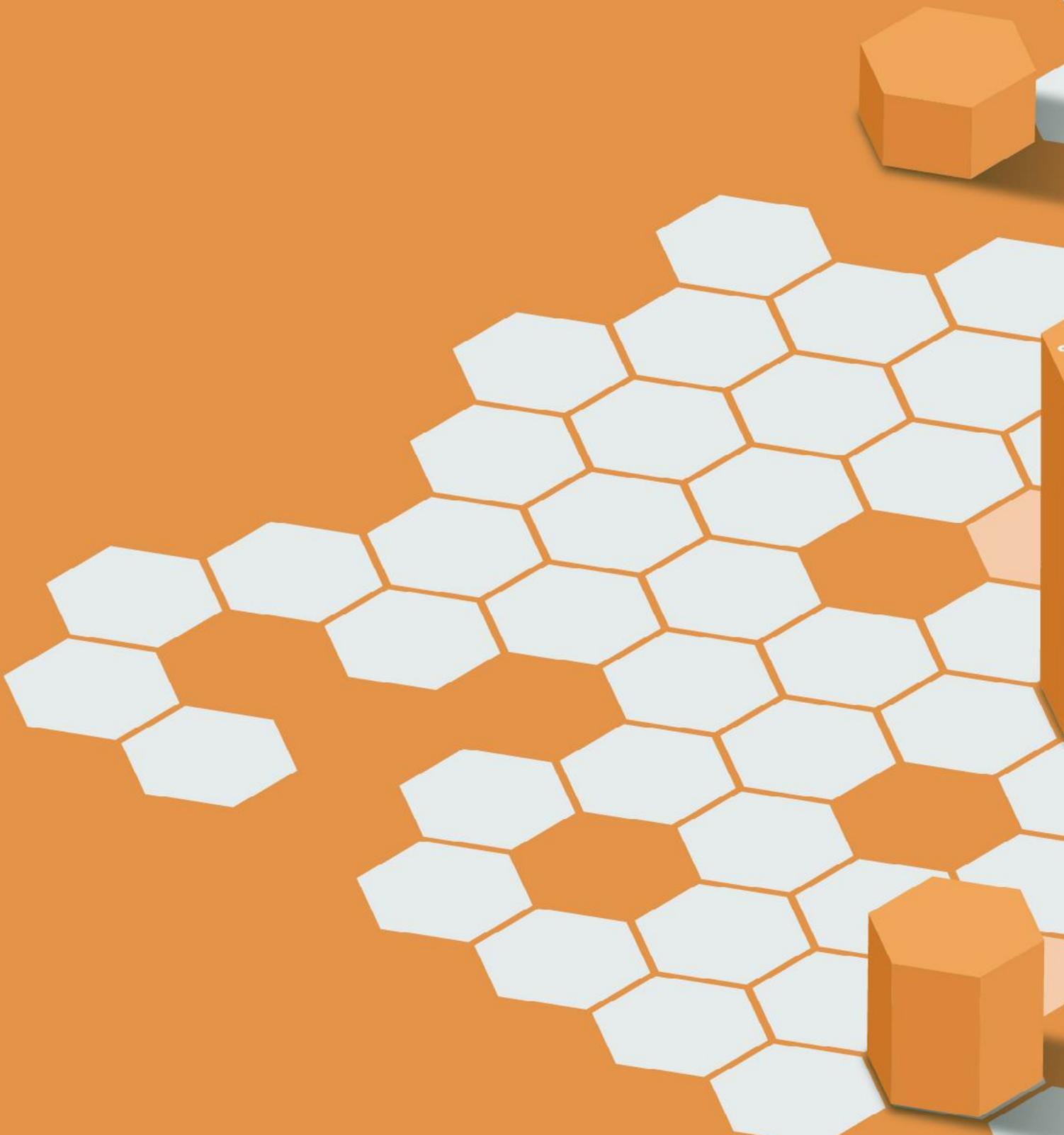
Conclusion

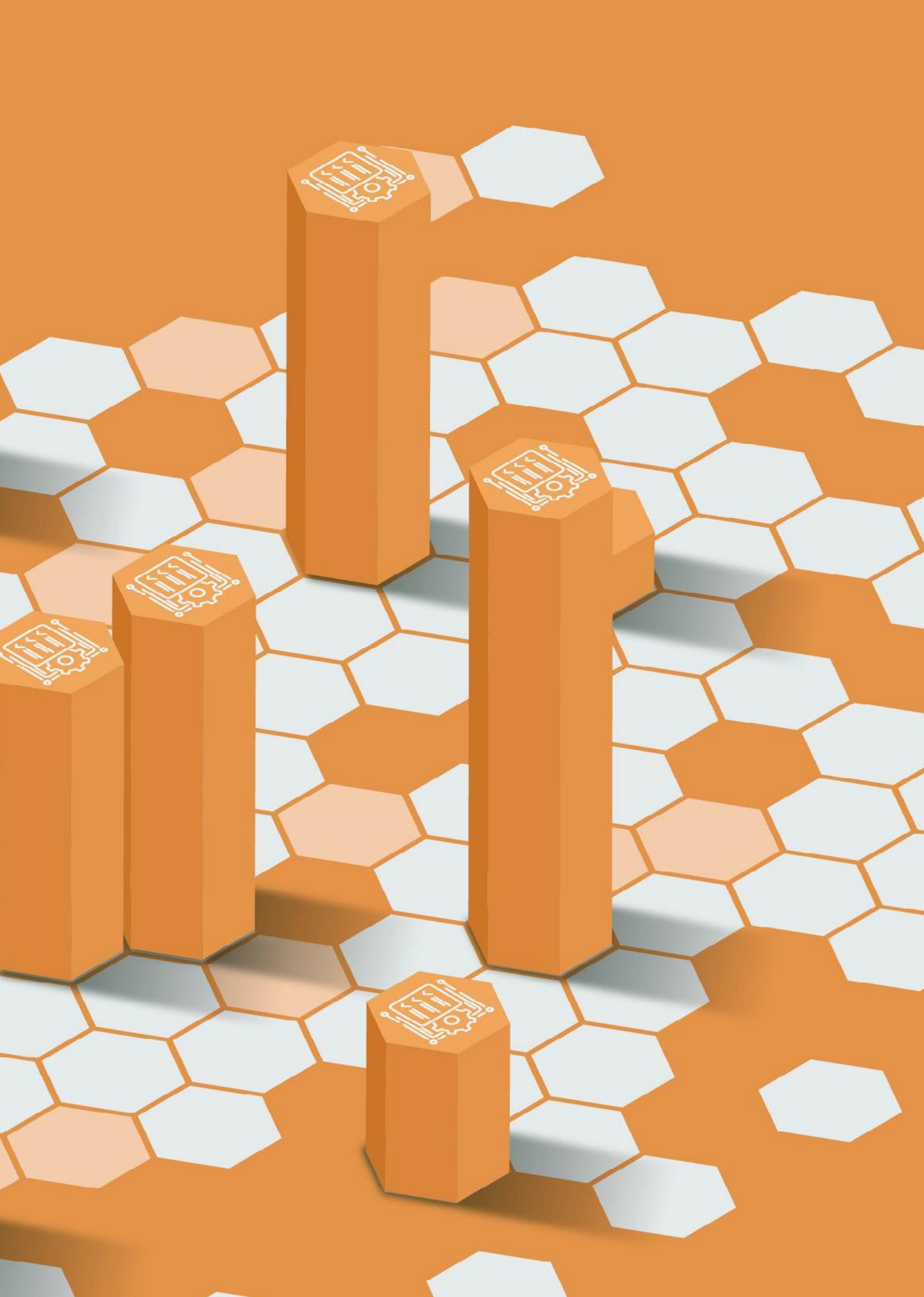
The HRS exhibited good validity and reliability, making it a valuable tool for assessing resilience in adults. It can be instrumental for health education and care professionals in identifying and promoting health resilience among adults, ultimately contributing to better health outcomes.

The content of this write-up has been published in part and licensed under a Creative Commons BY 4.0 License. Citation: Ge, L., Yip, W. F., Li, R., Chua Siang Seng, E., Ho, M. H. R., Ho, A. H. Y., ... & Tan, W. S. Development and validation of the multi-dimensional Health Resilience Scale for community-dwelling adults. *Frontiers in Public Health*, 13, 1452738.



EVALUATING HEALTH SYSTEMS AND SERVICE DELIVERY







Specialist Manpower Planning: Projecting Healthcare Workforce Needs

Teow Kiok Liang, Dr Nicholas Chew Wuen Ming¹, Dr Michelle Jong², Dr Ng Shu Li³, Dr Howard Foo Shing Liang⁴, Aver Lim⁵

¹ Medical Board, Woodlands Health, Singapore

² Group Clinical Education, National Healthcare Group, Singapore

³ Clinical Services, Woodlands Health, Singapore

⁴ Human Resource, Woodlands Health, Singapore

⁵ Group Human Resource, National Healthcare Group, Singapore

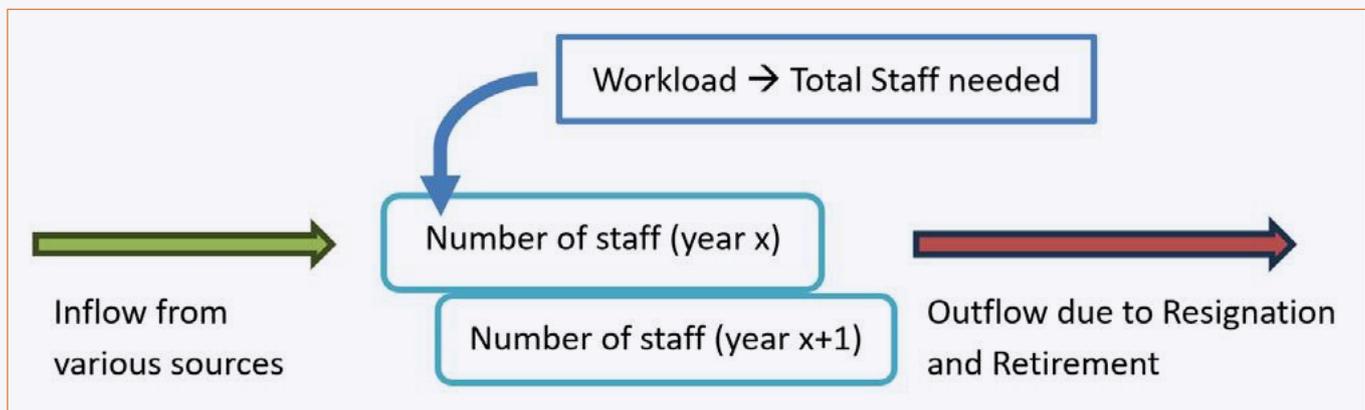
Highlights

- Medium-to-long term manpower planning is required to balance ensuring an adequate number of specialists to meet the future workload and over-production of new staff.
- We constructed a holistic manpower projection model incorporating information on workload, staffing levels and residency volume.

Introduction

Training medical specialists requires extensive time and resources. To ensure healthcare systems maintain adequate specialist coverage, long-term workforce planning is essential. We developed a Stock-and-Flow model to analyse current and projected pipeline specialist numbers, project future needs, and identify potential workforce leakages and gaps before they impact patient care (**Figure 1**).

Figure 1. Stock-and-Flow diagram of Manpower



Methods

When considering manpower, the term “demand” can be ambiguous. It may refer to the “Total number of staff needed” for the organisation / department, or the “Number of new staff to be recruited” in the coming year. We use the Stock-and-Flow framework to articulate and define the issue more clearly: Stock is the pool of existing staff, and Inflow is the new staffing joining the pool.

Our analysis framework separates three key components:

1. **Demand Projection:** We calculated total specialist staffing needs based on hospital workload metrics (both inpatient and outpatient volumes). Rather than attempting to determine an ideal workload-to-specialist ratio, we used current staffing levels as our baseline and projected future needs based on expected workload increases. These increases were modelled primarily using population growth and demographic aging trends; and does not take into account potential savings which may be achieved through initiatives such as AI adoption or job redesign.
2. **Workforce Outflow:** We analysed historical data on specialist departures, stratified by age, to project future attrition through both retirement and other causes. This allowed us to create age-specific attrition rates and forecast future staffing gaps.
3. **New Specialist Pipeline:** We incorporated residency program data, including typical completion rates, typical attraction rates; to project the number of new specialists entering the workforce annually.

This multi-source approach required extensive collaboration between clinical departments, human resources, and analytical teams to ensure data accuracy and practical applicability.

Results

In our analysis, we have noted that our current workforce is relatively young, thus limiting our near-term retirement impact. Our historical specialist attrition rates range from 2-5% annually. We have projected annual workload growth of approximately 3%, driven by population expansion and aging demographics. In parallel, we have assumed that current specialist training slots remain relatively constant.

With this as a background, our model projected a growing disconnect between increasing healthcare demands and the static supply of new specialists. This imbalance will require proactive intervention, possibly through enhanced productivity and efficiency initiatives, strategic management of specialist recruitment, talent retention, as well as implementation of demand management strategies through workforce transformation initiatives and a shift in the model of care delivery.

Conclusions

This study integrates various information together, and with the use of Stock-and-Flow model, provides insight to future medical specialists’ demand-supply situation.

Future refinements could include discipline-specific modelling for targeted interventions, assessment of current staffing adequacy across specialties, and incorporation of known healthcare initiatives that may impact healthcare demand beyond demographic changes.



Impact of Vaccination Strategies on Pneumonia Hospitalisation Rates in Older Adults Aged 65 and Older: A Multi-variable Big Data Study of Influenza and Pneumococcal Vaccines

Matthias Koh Yong An, Dr Yap Chun Wei, Dr John Abisheganaden

Highlights

- Up-to-date influenza vaccinations are a key defence against pneumonia inpatient hospitalisation among older adults aged 65 and above.
- Potential value in public health interventions to encourage up-to-date vaccinations.

Introduction

Pneumonia poses a significant health risk in older adults. As we age, our immunity weakens, and our chances of contracting pneumonia are higher, leading to severe pneumonia, increased morbidity and mortality. In addition to good hygiene, sufficient physical activity and a healthy diet, vaccinations are key public health interventions to reduce pneumonia hospitalisation rates. Utilising statistical artificial intelligence (AI), this study aimed to evaluate the effect of influenza and pneumococcal vaccinations – either influenza only, pneumococcal only, or both, on the hospitalisation rate for pneumonia among individuals aged 65 and older.

Methods

We conducted a one-year retrospective cohort study using a large healthcare dataset ($n = 180,511$) on Singapore Residents aged 65 and older within the National Healthcare Group (NHG) catchment area between 2021 and 2022. The intervention group consisted of individuals vaccinated with either the influenza vaccine, pneumococcal vaccine or both. For patients who had been influenza vaccinated in 2021, their vaccination date was regarded as the study entry date. As pneumococcal vaccination provides lifetime protection, we considered the latest visit date to any NHG institution in 2021 as the study entry date for those who had only been vaccinated with pneumococcal. The control group comprised unvaccinated individuals of which their study entry date was their latest visit date to any NHG institution in 2021. We then studied if they had a primary pneumonia hospital admission one-year from their study entry date.

We employed logistic regression to model the effect of vaccination status on pneumonia hospitalisation risk, with adjustments made for age, gender, race, and comorbidities (chronic obstructive pulmonary disease (COPD), asthma and diabetes). Adjusted odds ratios were calculated to compare the effectiveness of the different vaccination regimens. We used the machine learning algorithm, Decision Tree to cross-validate the variable importance output produced by the logistic regression.

Results

Patients who were vaccinated with influenza only or both influenza and pneumococcal vaccines had around 40% reduced risk of pneumonia hospitalisation. However, patients who were vaccinated with pneumococcal vaccination had a 56% higher risk of pneumonia hospitalisation (**Table 1**). Additionally, feature importance calculated from Decision Tree corroborated with the highest coefficients of the odds ratios from the logistic regression (**Table 2**).

Table 1. Adjusted odds ratios of the variables

Variables	Adjusted Odds Ratio (95% CI)	Times more likely Reference Group is admitted to Hospital for Pneumonia
Vaccination		
Not Vaccinated	Reference	-
Influenza only	0.59 (0.42 – 0.81) *** ¹	x1.69
Pneumococcal only	1.56 (1.42 – 1.70) ***	x0.64
Both	0.58 (0.50 – 0.67) ***	x1.72
Gender		
Male	Reference	-
Female	0.70 (0.65 – 0.75) ***	x1.43
Race		
Chinese	Reference	-
Malay	1.54 (1.34 – 1.76) ***	x0.65
Indian	1.28 (1.10 – 1.48) ***	x0.78
Others	1.46 (1.22 – 1.76) ***	x0.68
Age		
65 – 74	Reference	-
75 – 84	2.61 (2.37 – 2.88) ***	x0.38
≥ 85	7.50 (6.75 – 8.25) ***	x0.13
COPD²		
No	Reference	-
Yes	3.39 (3.04 – 3.78) ***	x0.29
Diabetes		
No	Reference	-
Yes	1.41 (1.31 – 1.53) ***	x0.71
Asthma		
No	Reference	-
Yes	0.99 (0.85 – 1.15)	x1.01

¹ p-value<0.05; ² COPD: Chronic Obstructive Pulmonary Disease

Table 2: Top 5 rank of importance of variables

	Logistic Regression	Decision Tree
	Age	Age
	COPD	COPD
	Pneumococcal Vaccinated	Asthma
	Race	Diabetes
	Diabetes	Pneumococcal Vaccinated

Conclusions

Recent influenza vaccination is associated with better prevention against hospital admissions for pneumonia amongst older adults. Hence, there is potential value in public health interventions to encourage a greater uptake of up-to-date vaccinations among older adults. Future research could compare vaccination impacts over multiple years and incorporate variables that measure frailty. To specifically assess pneumococcal vaccination effectiveness, studies with matched controls could provide deeper insights.



Cost-effectiveness Analysis of Continuous Positive Airway Pressure Therapy for Obstructive Sleep Apnea in Singapore from a Health System Perspective

Dr Meng Fanwen, Dr Ang Yee Gary, Dr Ruth Chang Rui Ying¹, Dr Lee Chuen Peng², Dr Kelvin Bryan Tan^{3,4,5,6}, Dr John Abisheganaden

¹ Department of Respiratory Medicine, Woodlands Health, Singapore

² Department of Respiratory and Critical Care Medicine, Tan Tock Seng Hospital, Singapore

³ Chief Health Economist Office, Ministry of Health, Singapore

⁴ Saw Swee Hock School of Public Health, National University of Singapore, Singapore

⁵ Centre of Regulatory Excellence, Duke-NUS Medical School, Singapore

⁶ Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore

Highlights

- CPAP treatment was highly cost-effective with an ICER of US\$13,822 per DALY averted.
- The total annual cost for patients diagnosed with the home sleep test can be reduced by 27%, compared to those diagnosed with the inpatient sleep study.

Introduction

Obstructive sleep apnea (OSA) is a prevalent sleep disorder, involving episodes of upper airway narrowing or complete closure associated with physiological disturbances during sleep. This common disorder poses significant health and economic challenges, linked to cardiovascular issues, daytime sleepiness, accidents, healthcare costs, and reduced quality of life. Continuous positive airway pressure (CPAP) therapy is the gold standard for OSA treatment. While assessment of its cost-effectiveness is important to public health decision-making, there is a scarcity of publications in Singapore. This study aimed to assess the cost-effectiveness of CPAP treatment in Singapore from a health system perspective.

Methods

The analysis evaluated a 5-year care pathway using a Markov model by considering per-patient costs of treatment, health system cost savings of OSA and attributed conditions. Effectiveness was measured in disability-adjusted life years (DALYs) with a discount rate of 3%, and a weighted 5-year CPAP adherence of 74.1% from local studies to estimate the incremental cost-effectiveness ratio (ICER). Per-patient costs of treatment were from a large public hospital in Singapore. Efficacy of CPAP treatment, health system costs, DALYs were obtained from the literature. Costs were in US dollars, and we used the willingness-to-pay (WTP) threshold of US\$50,000 per DALY used in the United States. We also compared the total annual costs for patients diagnosed with the inpatient sleep study and with those from the home sleep test (HST). Probabilistic and one-way sensitivity analyses were conducted to assess impacts of model parameters on ICER. A what-if analysis was conducted to derive the minimum adherence required for cost-effectiveness.

Results

Table 1 provides a summary of the estimated ICERs associated with adherence rates of two local studies in Singapore together with the weighted adherence, which were US\$17,939, US\$13,822, US\$13,258 per DALY averted, respectively, indicating that CPAP therapy was highly cost-effective. In addition, the average total cost for patients diagnosed by the inpatient sleep study was US\$856 per year per case, while the corresponding cost was \$625 for patients diagnosed by the HST, resulting in a remarkable 27% reduction in the total yearly costs. According to the cost-effectiveness acceptability curve, at the WTP threshold of US\$50,000, CPAP therapy was cost effective with 100% probability. **Figure 1** indicated that costs, effectiveness, and adherence to treatment had higher impact on the cost-effectiveness of CPAP therapy. What-if analysis suggested that for CPAP treatment to be cost-effective, adherence rate should be at least 16.1%, as shown in **Figure 2**.

Table 1. Cost-effectiveness analysis of CPAP treatment for patients with OSA with different adherence rates.

Item per case	Adherence: 52.6% ¹	Adherence: 74.1%	Adherence: 78.5% ²
Cost of treatment (US\$)	586	654	668
Health system cost savings (US\$)	55	77	82
Net cost (US\$)	531	577	586
DALYs averted	0.0296	0.0417	0.0442
ICER (US\$ per DALY)	17,939	13,822	13,258

Figure 1. Tornado diagram of one-way sensitivity analysis.

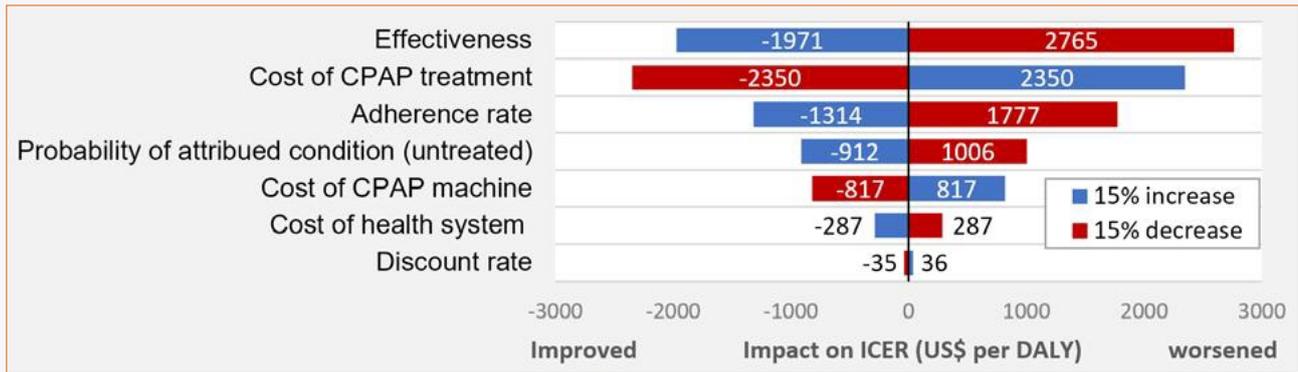
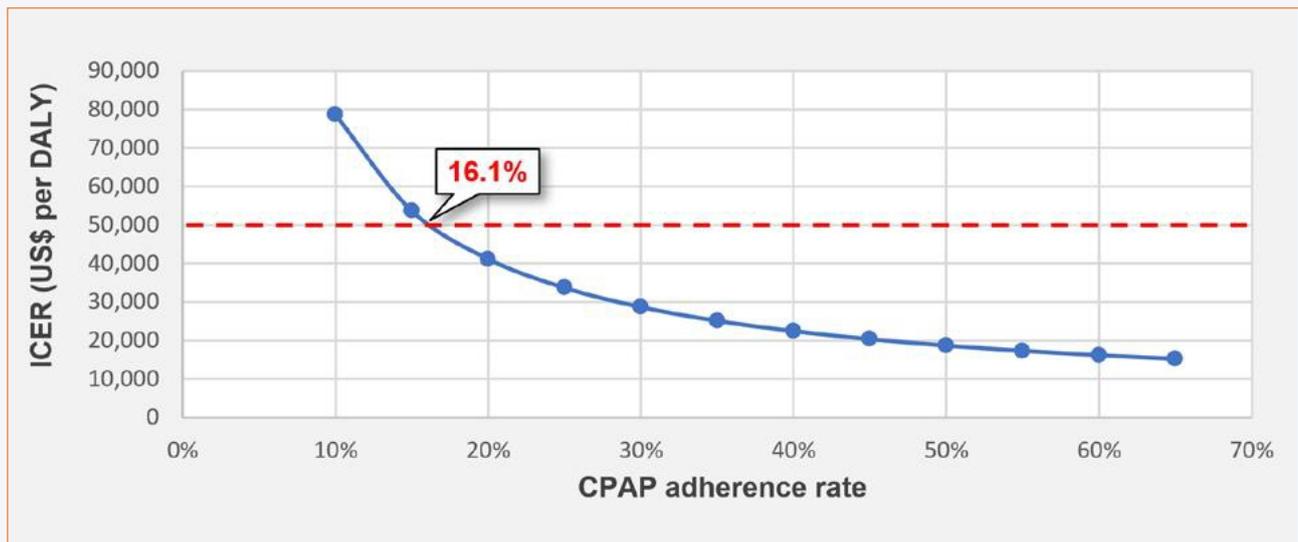


Figure 2. ICERs associated with different CPAP adherence rates.



Conclusions

CPAP treatment was demonstrated to be highly cost-effective. The results of this study provide valuable insights for policymakers to make informed decisions on funding diagnosis and CPAP therapy within Singapore’s healthcare system.

References

¹ Lee, C. H. K., Leow, L. C., Song, P. R., Li, H., & Ong, T. H. (2017). Acceptance and Adherence to Continuous Positive Airway Pressure Therapy in patients with Obstructive Sleep Apnea (OSA) in a Southeast Asian privately funded healthcare system. *Sleep Science*. 10(2), 57-63.
² Tan, B., Tan, A., Chan, Y. H., Mok, Y., Wong, H. S., & Hsu, P. P. (2018). Adherence to continuous positive airway pressure therapy in Singaporean patients with obstructive sleep apnea. *American Journal of Otolaryngology*. 39(5), 501-506.

The content of this write-up has been published in part: Meng F, Ang G. Y., Chang R. R. Y., Lee C. P., Tan K. B., & Abisheganaden J. A. (2024). Cost-effectiveness analysis of continuous positive airway pressure treatment for obstructive sleep apnea in Singapore from a health system perspective. *Journal of Sleep Research*, e14326.



Enhancing Post-stroke Recovery and Return-to-work: Insights from the SLS CHARTER Project

Estelle Foo Jie Wei, Dr Wan Fen Yip, Joey Ha Wei Yee, Ryan Cheok Chi Min, Dr Nuez Odessa Setiota¹, Dr Deshan Kumar¹, Dr Carol Tham¹, Lim Pang Hung¹, Teo Li Huan¹, Nur Nazaria Baharudin¹, Sng Yee Leng¹, Dr Loh Yong Joo¹, Dr Michelle Jessica Pereira

¹ Tan Tock Seng Hospital, Singapore

Highlights

- Stroke Life Support Central Health Alliance, Registry & Technology Enabled caRe (SLS CHARTER) Project increased Post-Stroke Checklist (PSC) usage from 50.5% to 86.9% over six months, demonstrating the feasibility of using the PSC to identify post-stroke needs to facilitate appropriate referrals.
- Return-to-work (RTW) status was influenced by occupation, education, and income, highlighting the need for targeted support for vulnerable groups like blue-collar workers and lower-income individuals.

Introduction

Post-stroke needs often go unmet due to insufficient follow-up and inadequate information, leading to poor reintegration into society and delayed return-to-work (RTW). These challenges contribute to significant economic burdens on individuals and society, highlighting the need for structured care coordination and patient empowerment. The Stroke Life Support Central Health Alliance, Registry & Technology Enabled caRe (SLS CHARTER) Project addresses these issues by educating stroke patients and caregivers on use the Post Stroke Checklist (PSC) and providing care coordination through the Central Health Alliance in Stroke Enablement (CHASE) coordinators. This study aimed to 1) describe the frequency of PSC usage among SLS participants, and 2) examine the association of sociodemographic factors and intervention status with RTW status.

Methods

Patients admitted to the Acute Stroke Unit at Tan Tock Seng Hospital between January and June 2023, with a diagnosis of either an ischemic or haemorrhagic stroke, were recruited for this study. A total of 206 SLS and 55 control patients were included for this analysis. Data on PSC usage, including the number of patients who utilised the PSC and the mean number of PSCs used per patient at one- and six-months post-stroke, were collected.

RTW outcomes were assessed as the proportion of patients who returned to work by six months. Logistic regression was used to examine associations between socio-demographic factors, intervention status (SLS participants versus control participants), and RTW outcomes, adjusting for age, gender, race, stroke type, and baseline Modified Barthel Index (MBI) severity category.

Results

PSC usage among SLS participants increased from 50.5% (n = 104) at one month to 86.9% (n = 179) at six months post-stroke (**Figure 1**). The mean number of PSCs completed per patient also increased over the six-month period (1.76 ± 1.09 to 4.25 ± 1.09).

Among patients employed at baseline, 58.3% (n = 67) of SLS participants (n = 115) and 54.3% (n = 19) of control patients (n = 35) returned to work (**Figure 2**). After adjusting for age, gender, ethnicity, stroke type, and baseline MBI severity category, no significant difference in RTW status was observed between SLS participants and non-SLS participants.

Figure 1: PSC usage of SLS patients over 6 months

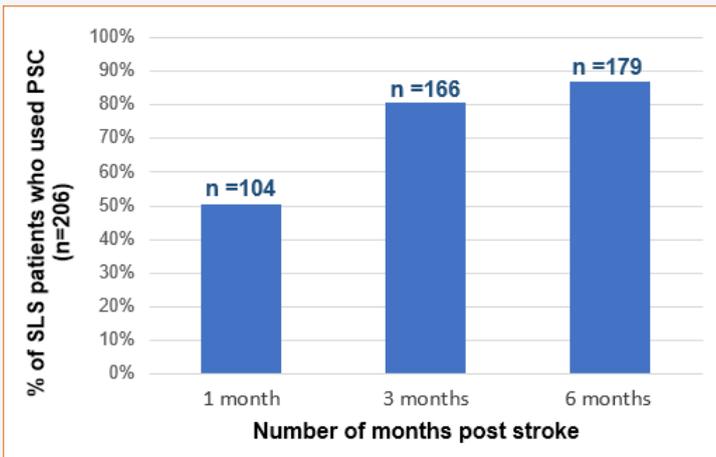
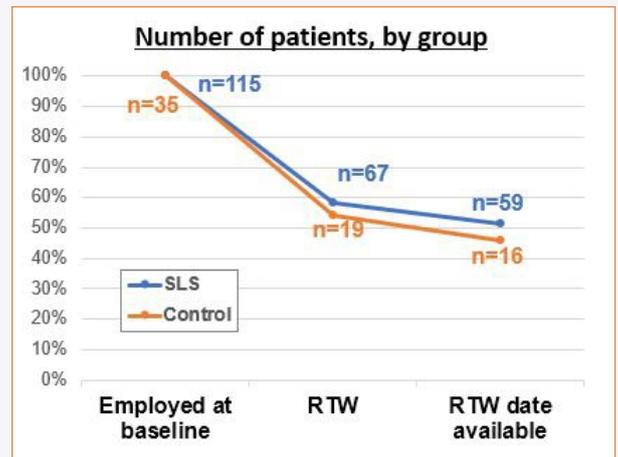
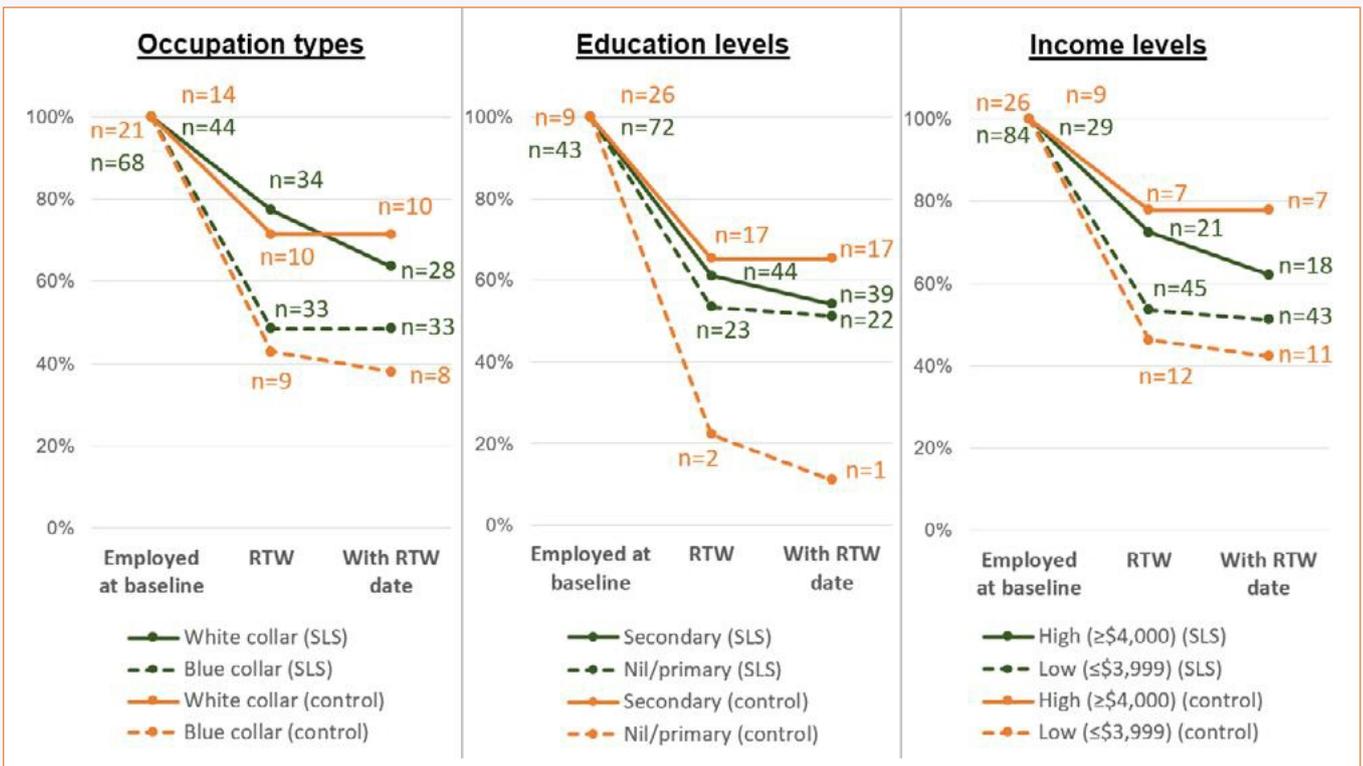


Figure 2: RTW status of SLS and control patients



Sociodemographic factors such as occupation, education, and income strongly influenced RTW status (Figure 3). White-collar workers were four times more likely to RTW compared to blue-collar workers (95% Confidence Interval (CI): 1.6–10.1). Workers with secondary education or higher were 2.5 times more likely to RTW compared to those with primary education or lower (95% CI: 1.0–6.0). Workers with a monthly income \geq \$4000 were 6.5 times more likely to RTW compared to those earning $<$ \$4000 (95% CI: 1.9–22.6).

Figure 3. Associations between sociodemographic factors and RTW outcomes for the SLS and control group



Conclusions

The SLS CHARTER project demonstrated the feasibility of using the PSC to identify post-stroke needs and facilitate appropriate referrals. Targeted interventions may be required to improve RTW outcomes among patients with lower income, lower education levels, and blue-collar workers.



From Prediction to Palliation: A Case Study of A Prognostic Tool Implementation in a Transmural Model of Early Palliative Care for Advanced Dementia

Palvinder Kaur, Low Zhi Jun¹, Michelle Wong¹, Araviinthansai S/O Subramaniam², Chang Chui Rong², Wang Qing², Dr Ding Yew Yoong^{2,3}, Dr Ong Chin Ee¹, Dr Jackie Tan³, Dr Nihar Pandit⁴, Dr Robyn Tan⁵, Dr Tan Woan Shin², Dr Allyn Hum^{1,6}

¹ Department of Palliative Medicine, Tan Tock Seng Hospital, Singapore

² Geriatric Education and Research Institute, Singapore

³ Department of Geriatric Medicine, Tan Tock Seng Hospital, Singapore

⁴ Department of General Medicine, Tan Tock Seng Hospital, Singapore

⁵ Institute of Policy Studies, National University of Singapore, Singapore

⁶ The Palliative Care Centre for Excellence in Research and Education, Singapore

Highlights

- PRO-MADE is integrated into a transmural early palliative care model, stratifying patients for tailored palliative care interventions.
- An ongoing implementation evaluation and prospective validation of the model will evaluate the feasibility and clinical impact of this integrated approach.

Introduction

Accurate prognostication is crucial for timely palliative care, yet few prognostic tools exist for patients diagnosed with advanced dementia (AD). This study implements a transmural early palliative care model that utilises the PROgnostic Model for Advanced Dementia (PRO-MADE) predictive algorithm to identify high-risk AD patients for tailored interventions (**Figure 1**).

Implementing PRO-MADE in TTSH

PRO-MADE, a multivariable logistic model predicting one-year mortality, was developed, and validated according to the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD) guidelines¹. It incorporates readily available clinical parameters: age, sex, pneumonia diagnosis, pressure ulcers, Charlson's Comorbidity Index, functional dependency, dysphagia, urea, and albumin levels². The implementation of PRO-MADE at Tan Tock Seng Hospital (TTSH) was phased. (**Figure 2**). First, we mapped the variables to corresponding data in TTSH's EPIC system. Upon admission, the primary care team is prompted to stage dementia patients using the Functional Assessment Staging Tool (FAST). If lab tests (urea, and albumin) are missing for AD patients, the primary care team receives prompts to order them. The PRO-MADE model is triggered daily in the Next Generation Electronic Medical Record (NGEMR) system to compute risk scores for AD patients staged as FAST 7C and above. Before discharge, high-risk patients (probability ≥ 0.4) are flagged to the primary care team. High-risk AD patients or their caregivers receive further support, including education on disease trajectory, advance care planning (ACP), palliative interventions (e.g., pain assessment using PAINAD), and referral to community services. Low-risk patients receive enhanced usual care, including structured plans of care discussions. In Phase 1, Research Assistants (RAs) manually prompted the primary care team through tiger texts/email, transitioning to digital prompts in Phase 2.

Conclusions

Integrating PRO-MADE into a transmural palliative care model enables risk-stratified tailored interventions and continuity of care across clinical settings. Implementation evaluation and prospective validation is ongoing in the departments of Geriatric and General Medicine at TTSH. We aim to assess acceptability, adoption, feasibility, appropriateness, and sustainability of the model. Primary outcomes include concordance with ACP care preferences, place of death, healthcare costs and utilisation in the last year of life.

References

¹ Kaur P et al 2021. "Predicting mortality in patients diagnosed with advanced dementia presenting at an acute care hospital: the PROgnostic Model for Advanced Dementia (PRO-MADE)." *BMC Geriatrics* vol. 23,1 255. 28 Apr. 2023.

² Kaur P et al 2021. Deaths in dementia: a scoping review of prognostic variables *BMJ Supportive & Palliative Care* 2021;11:242-252.

Figure 1. Model implementation workflow and palliative care interventions

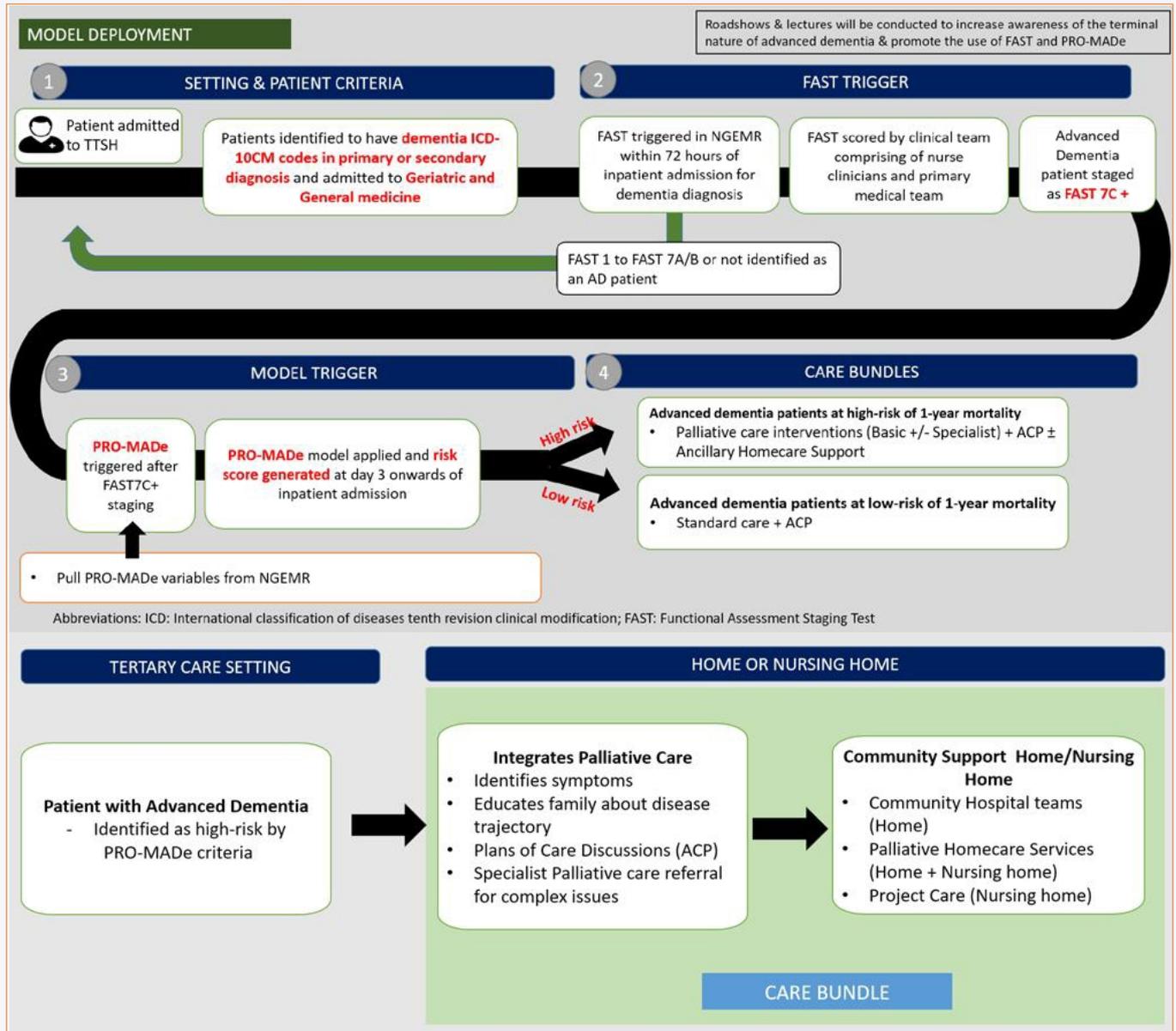
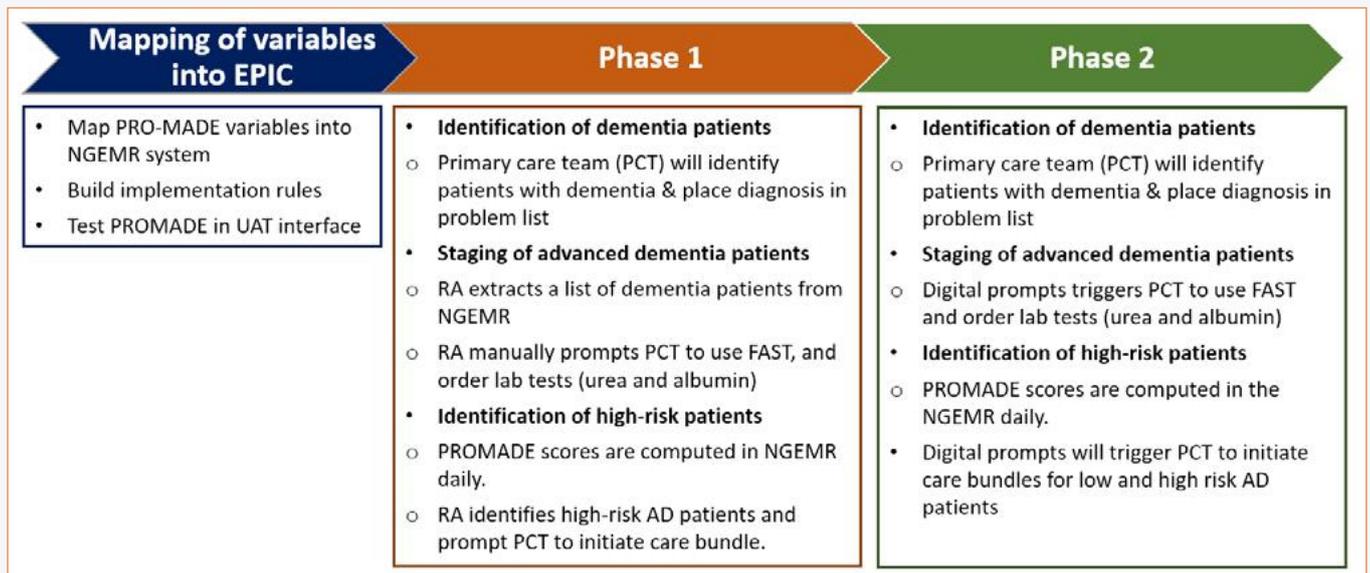


Figure 2. Implementation of PRO-MAde and prompts





Key Drivers of High-quality Care in the Primary Eye Care Model: Insights from a Mixed-methods Study

Dr Yip Wan Fen, Celeste Gue Ci Ying, Dr Michelle Jessica Pereira, Dr Joseph Antonio D. Molina, Dr Ding Si Yan¹, Dr Jonathan Foo¹, Dr Vivien Yip Cherng Hui², Dr Vernon Yong Khet Yau², Dr Wong Hon Tym², Dr Lim Tock Han²

¹ National Healthcare Group Polyclinics, Singapore

² National Healthcare Group Eye Institute, Tan Tock Seng Hospital, Singapore

Highlights

- **Demonstrating Excellence in an Optometrist-Led Care Model:** This mixed-methods study highlighted the Primary Eye Care model's high quality of care and identified factors such as leadership support, comprehensive training, and interprofessional collaboration as critical to maintaining care quality.
- **Bridging Gaps in Awareness:** Limited interactions among healthcare professionals influenced trust and referrals to the Primary Eye Care model, emphasising the need for continuous engagement, trust-building, and improved interprofessional communication to enhance scalability and sustainability.

Introduction

Primary eye care (PEC), established by the National Healthcare Group Eye Institute (NHGEI) is an optometrist-led care model that facilitates the right-siting of patients with stable and non-complex eye conditions in the community. However, there is a lack of studies examining the quality of care provided through PEC. Additionally, there is a lack of understanding on key healthcare professionals' perceptions towards PEC's care quality. The aim of this study was to evaluate PEC's quality of care and to explore stakeholders' perceptions of its quality, as well as the key influences on PEC's level of care.

Methods

A mixed-methods sequential explanatory design, guided by the Practical, Robust, Implementation and Sustainability Model (PRISM) framework, was employed.

Quantitative Analysis

Medical records (between 2018–2022) were retrospectively analysed. Ophthalmologists assessed concordance between optometrists' and ophthalmologists' management plans (proportion of cases where plans matched). Concordance of plan components, such as follow-up location and duration, was measured using Cohen's kappa.

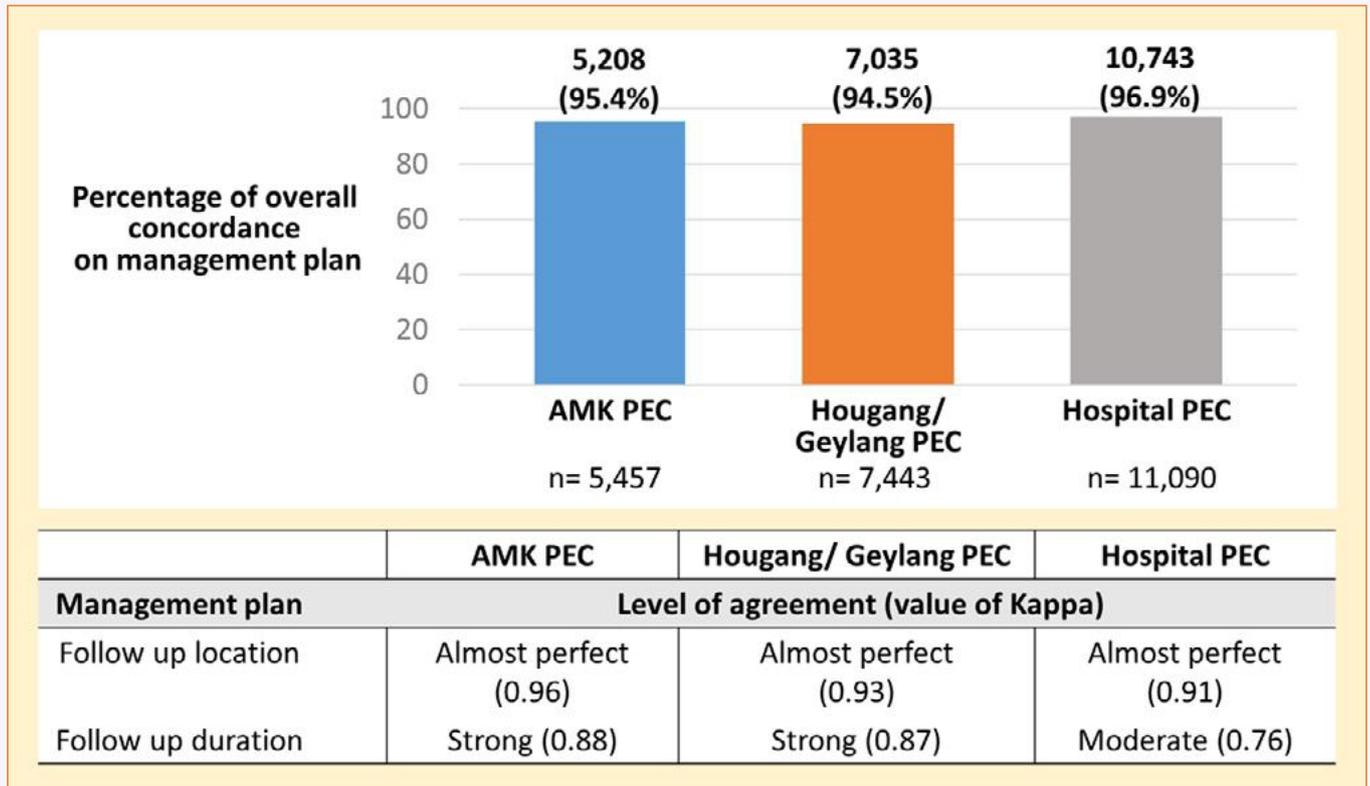
Qualitative Analysis

Stakeholder perceptions of PEC's quality of care were explored through focus group discussions (FGDs) with 7 optometrists, 3 ophthalmologists, and 16 polyclinic healthcare professionals (doctors and nurses) and in-depth interviews (IDIs) with 4 programme leaders and polyclinic directors. Themes were identified and corroborated with the quantitative findings.

Results

Among 23,990 cases, high concordance for management plans (94.5%–96.9%) and strong agreement for plan components (Cohen's kappa: 0.76–0.96) were observed (**Figure 1**). Ophthalmologists perceived PEC's quality of care as high, fostering trust in optometrists and a willingness to refer patients. Conversely, limited interactions between polyclinic healthcare professionals and optometrists influenced the polyclinic healthcare professionals' awareness of PEC's care quality and services as well as referral motivation.

Figure 1. PEC provided consistent high-quality care across all locations



Key factors influencing care perceptions included: 1) existing inter-professional relationships, 2) awareness of and trust in optometrists’ training and PEC’s services, and 3) opportunities for inter-professional interaction.

Five factors supported PEC’s high care quality: 1) hospital leadership, programme leaders’ and ophthalmologists’ support for optometrists-led care; 2) comprehensive training programme for optometrists 3) PEC’s clear clinical guidelines and real-time ophthalmologist helpline; 4) open communication channels among healthcare professionals; and 5) optometrists’ motivation for continuous learning. However, guidelines were also perceived as a barrier as it may not align with ophthalmologists’ and optometrists’ management plan recommendations. Furthermore, variations in management plan recommendations from ophthalmologists in different sub-specialties also affected concordance.

Conclusions

PEC provides high-quality eye care, but variations in perceptions and awareness among healthcare professionals influenced referral motivation. Sustaining PEC high care quality requires continuous interprofessional engagement, fostering trust, and improving awareness. Key elements for sustained high concordance included stakeholders’ support, comprehensive training programme, clear guidelines, real-time helpline, open communication, and optometrists’ motivation. These findings underscore the importance of mixed-methods research in identifying strategies for maintaining high-quality care while scaling allied health-led care models.



“Quality of Care and Stakeholders’ Perceptions of PEC Implementation in Singapore: A Mixed-Methods Study Protocol”



PUBLICATIONS



1. **Abisheganaden J**, Hin MCC, **Yap CW**, Jao-yiu JS. Impact of the COVID-19 pandemic on Emergency Department attendances and hospitalisation for Asthma and Chronic Obstructive Pulmonary Diseases (COPD) in Singapore. *Pulm Crit Care Med*. 2024 Apr 22; 8(1-6): doi: 10.15761/PCCM.1000174
2. Ng GW, Gan KF, Liew H, **Ge L, Ang G, Molina J, Sun Y**, Prakash PS, Harish KB, Lo ZJ. A Systematic Review and Classification of Factors Influencing Diabetic Foot Ulcer Treatment Adherence, in Accordance With the WHO Dimensions of Adherence to Long-Term Therapies. *Int J Low Extrem Wounds*. 2024 Feb 20;15347346241233962. doi: 10.1177/15347346241233962.
3. **Ge L**, Zhao J, Tan M, Tan E, Liew H, Yong E, Hoe J, Shi C, Chan DYS, **Ang GY, Molina JA, Sun Y**, Hoi WH, Chandraskear S, Lo ZJ. Multi-disciplinary diabetic limb salvage programme in octogenarians with diabetic foot ulcers is not futile: An observational study with historical controls. *Int Wound J*. 2024 Mar;21(3):e14801. doi: 10.1111/iwj.14801
4. Lin JHX, **Ge L**, Liew H, Tan E, Hoe J, Yong E, Loh Z, Chew T, Farhan MFM, Li Z, Tan D, Loh YB, Zhu J, Ang P, Shi C, Seng Chan DY, Law C, Raman N, **Molina J, Ang G**, Koo HY, Low KQ, Choo J, Tan CY, Lim JA, Siow J, Chan SW, Chandraskear S, Chew D, Lo ZJ. Rapid Access From Primary Care to a Multidisciplinary Clinic at Tertiary Care Improves Clinical Outcomes for Patients With Diabetic Foot Ulcers: Results From Diabetic Foot in Primary and Tertiary (DEFINITE) Care's Lower Extremity Amputation Prevention Program (LEAPP) Clinic. *Int J Low Extrem Wounds*. 2024 May 15;15347346241252200. doi: 10.1177/15347346241252200.
5. **Ge L**. Associations between Socioeconomic Status and Social Isolation across Age Groups: A Cross-sectional Study in Singapore. *Journal of Social Health*. 2024 Aug 15; 6(1-16). doi: 10.61072/j.osh.2024.041.
6. **Ge L, Molina J**, Karthigayan R, Foo HT, Tang M, Chua R, Ong CF. Development and validation of the Health Activation Scale for Children (HAS-C): an important intermediate outcome measure for health promotion initiatives. *BMC Health Serv Res*. 2024 Sep 27;24(1):1120. doi: 10.1186/s12913-024-11526-7.
7. **Ge L**, Agrawal R, Singer M, **Kannapiran P, De Castro Molina JA, Teow KL, Yap CW, Abisheganaden JA**. Leveraging artificial intelligence to enhance systematic reviews in health research: advanced tools and challenges. *Syst Rev*. 2024 Oct 25;13(1):269. doi: 10.1186/s13643-024-02682-2.
8. **Ge L, Ang YG, Molina J, Sun Y**, Tan E, Liew H, Hoe J, Hui Xian Lin J, Koo HY, Low KQ, Choo J, Yap T, Bte Azmi NAS, Tan D, Loh YB, Zhu X, Yong E, Hong Q, Chew T, Chan DYS, Shi C, Law C, Hoi WH, Siow J, Lim JA, **Abisheganaden JA**, Chew D, Lo ZJ. Investigating Nonadherence in an Integrated Diabetic Limb Salvage Programme: Reasons, Associated Factors, and Impacts on Care Outcomes. *Int J Low Extrem Wounds*. 2024 Nov 5;15347346241294178. doi: 10.1177/15347346241294178.
9. **Ha JWY, Pereira MJ**, Chong EKC, **Molina JADC**, Lim WS. Comparing patient- and proxy-reported health-related quality of life of frail older patients at the emergency department. *Discov Public Health*. 2024 Jul 30;21(34). doi:10.1186/s12982-024-00165-2
10. **Ha JWY, Yip W, Ge L, Yap CW, Pereira MJ**. Association of social isolation and depressive symptoms with workplace productivity loss in a multi-ethnic Asian study. *Sci Rep*. 2024 Sep 27;14(1):22145. doi: 10.1038/s41598-024-73272-4.
11. **Kaur P**, George PP, **Ng SHX, Yip WF, Seng ECS**, Tay RY, Tan J, Chu J, Low ZJ, Tey LH, Hoon V, Tan CK, Tan L, Aw CH, Tan WS, Hum A. Risk Factors for All-Cause Mortality in Patients Diagnosed with Advanced Heart Failure: A Scoping Review. *J Palliat Med*. 2024 Jul 31. doi: 10.1089/jpm.2024.0067.
12. **Lee M**, Khoo HS, **Krishnasamy C, Loo ME**, Wong SKW, **Cheng SC**, Bei E, Tan SY, Xie P, Lee ES, Tang WE. Experiences of living with overweight/obesity and early type 2 diabetes in Singapore—a qualitative interview study. *BMJ Open*. 2024 May 7;14(5):e079082. doi: 10.1136/bmjopen-2023-079082.
13. Han S, Goh J, **Meng F**, Leow MK, Rubin DB. Contrast-specific propensity scores for causal inference with multiple interventions. *Stat Methods Med Res*. 2024 May;33(5):825-837. doi: 10.1177/09622802241236952.
14. **Meng F, Ang GY**, Chang RRY, Lee CP, Tan KB, **Abisheganaden JA**. Cost-effectiveness analysis of continuous positive airway pressure treatment for obstructive sleep apnea in Singapore from a health system perspective. *J Sleep Res*. 2024 Sep 3:e14326. doi: 10.1111/jsr.14326.
15. **Ng SHX**, Chiam ZY, Chai GT, **Kaur P, Yip WF**, Low ZJ, Chu J, Tey LH, Neo HY, Tan WS, Hum A. The PROgnostic Model for chronic lung disease (PRO-MEL): development and temporal validation. *BMC Pulm Med*. 2024 Aug 30;24(1):429. doi: 10.1186/s12890-024-03233-0.
16. Nah MFK, **Pereira MJ**, Hemaavathi M, Wong SW, Lim CJ, Tan BY. Study on proximal humerus evaluation of effective treatment (SPHEER) - what is the effect of rehabilitation compliance on clinical outcomes of proximal humerus fractures. *BMC Musculoskelet Disord*. 2023 Oct 2;24(1):778. doi: 10.1186/s12891-023-06894-w.
17. Ong PW, Lim CJ, **Pereira MJ**, Kwek EBK, Tan BY. Achieving satisfactory functional outcomes in conservatively treated proximal humerus fractures: relationship between shoulder range of motion and patient-reported clinical outcome scores. *JSEInt*. 2024 Feb 23;8(3):440-445. doi: 10.1016/j.jseint.2024.02.003.
18. Tan BY, Tan CY, Tan TL, Yang SY, Chew GLS, Tan SI, Chua YC, Yan YW, Soh DBQ, Goh TH, Ng PJ, Ng YT, Kuan SB, Teo BS, Kong KH, **Pereira MJ**, Ng HP. Heat and Acupuncture to Manage Osteoarthritis of the Knee (HARMOKnee): Protocol for an Effectiveness-Implementation Hybrid Randomized Controlled Trial. *JMIR Res Protoc*. 2024 Apr 3;13:e54352. doi: 10.2196/54352.
19. Tan BY, Yang SY, **Pereira MJ**, Tan CY, Lim CJ, Ng JP, Lee KT, Pua YH, Briggs AM, Hunter DJ, Skou ST, Thumboo J, Car J. Collaborative model of care between orthopaedics and allied healthcare professionals (CONNACT) in knee osteoarthritis: Effectiveness-implementation hybrid randomized controlled trial of a community-based, multidisciplinary, stratified intervention. *Osteoarthritis and Cartilage*. 2024 Aug;32(8):972-981. doi: 10.1016/j.joca.2024.04.018.
20. Ng H, Chong S, Li Y, Goh T, Pang K, **Pereira M**, Huang C. Objective Analysis of Traditional Chinese Medicine Syndrome Differentiation of Patients With Diabetes and Prediabetes: Protocol for a Nonrandomized, Exploratory, Observational Case-Control Study Using Digitalized Traditional Chinese Medicine Diagnostic Tools. *JMIR Res Protoc* 2024;13:e56024. DOI: 10.2196/56024
21. Kon MHA, **Pereira MJ, Molina JAC**, Yip VCH, **Abisheganaden JA, Yip WF**. Unravelling ChatGPT's potential in summarising qualitative in-depth interviews. *Eye (Lond)*. 2024 Nov 5. doi: 10.1038/s41433-024-03419-0.
22. Gue CCY, Rahim NDA, Rojas-Carabali W, Agrawal R, Rk P, **Abisheganaden J, Yip WF**. Evaluating the OpenAI's GPT-3.5 Turbo's performance in extracting information from scientific articles on diabetic retinopathy. *Syst Rev*. 2024 May 16;13(1):135. doi: 10.1186/s13643-024-02523-2.
23. Surian NU, Batagov A, Wu A, Lai WB, **Sun Y**, Bee YM, Dalan R. A digital twin model incorporating generalized metabolic fluxes to identify and predict chronic kidney disease in type 2 diabetes mellitus. *NPJ Digit Med*. 2024 May 24;7(1):140. doi: 10.1038/s41746-024-01108-6.
24. **Sun Y**, Ge L, Ang YG, Lo ZJ, Liew H, Tan DM, Chew D, **Abisheganaden JA**. Cost-effectiveness and clinical outcomes of artificial intelligence-enhanced screening for diabetic foot ulcers: A simulation study. *Ann Acad Med Singap*. 2024 Oct 30;53(10):638-640. doi: 10.47102/annals-acadmedsg.2024220.
25. Lim ETH, **Tan AL, D Molina JA, Abisheganaden J**. Healthcare utilisation patterns by persons with newly diagnosed Chronic obstructive pulmonary disease (COPD) in Singapore. *Pulm Crit Care Med*. 2024 June 10;8: doi: 10.15761/PCCM.1000176
26. Rojas-Carabali W, Agrawal R, Gutierrez-Sinisterra L, Baxter SL, Cifuentes-González C, **Yap CW, Abisheganaden J, Kannapiran P**, Wong S, Lee B, de-la-Torre A, Agrawal R. Natural Language Processing in medicine and ophthalmology: A review for the 21st-century clinician. *Asia Pac J Ophthalmol (Phila)*. 2024 Jul-Aug;13(4):100084. doi: 10.1016/j.apjo.2024.100084. Epub 2024 Jul 25. PMID: 39059557.
27. **Yip WF, Ng SHX, Kaur P**, George PP, Guan JHC, Lee G, Koh TJK, Tan WS, Hum AYM. Risk factors for short-term all-cause mortality in patients with end stage renal disease: a scoping review. *BMC Nephrol*. 2024 Feb 27;25(1):71. doi: 10.1186/s12882-024-03503-3.



RESEARCH GRANTS AND TEACHING



Research Grants Awarded

National Healthcare Group (NHG)

NHG Translational Research Programme Funding

Strategic Research in Metabolic Health

- Prof Lim Su Chi (PI)
- Dr Mary Lee Ching Ling (Site-PI)
- Dr Charmaine Krishnasamy (Co-I)

Grant Amount: **\$3 million**

- Dr Cheng San Chye (Co-I)
- Ge Lixia (Co-I)
- Joey Ha Wei Yee (Co-I)

Synapxe

Improving the Patient Journey for Lung Function Tests using Artificial Intelligence

- Dr Ang Yee Gary (Clinical)
- Dr Tan Geak Poh (Respiratory specialist)
- Prof John Abisheganaden (Respiratory specialist)
- James Nguyen (Technical Project Manager)

Grant Amount: **\$7,544**

- Kane Nguyen (Technical)
- Wang Lijia (Technical)
- Ernest Cuong (Technical)

Tan Tock Seng Hospital Community Fund

Ng Teng Fong Healthcare Innovation Programme (NTF HIP)

Improve patient monitoring with predictive analytics and process redesign

- Sister Chen Li (PI)
- Dr Sun Yan (Co-I)

Grant Amount: **\$78,268**

External Teaching Engagements

Dr Ang Yee Gary – Sessional Lecturer

Newcastle Australia Institute of Higher Education

Master of Health Economics, Management and Policy

Dr Ang Yee Gary – Invited Lecturer

Singapore University of Social Sciences

AI and Public Administration

ESSEC Business School

MSc in Management in Health Industries

National University of Singapore

Masters in Business Analytics

Dr Aidan LyanZhiang Tan – Invited Lecturer

National University Health System

Preventive Medicine Residency Programme



CONFERENCE PRESENTATIONS AND AWARDS



Health Hack 2024 (Singapore)

Medi-Bot-Charbot for patients with Type 2 Diabetes Mellitus

Oral Presentation

Dr Ang Yee Gary, Eugene Wang, Goh Woon Xuan, Justine Lim

Singapore Sleep Conference 2024 (Singapore)

Shortlisted for Best Poster Award

Cost-effectiveness Analysis of Diagnosis and Treatment with Continuous Positive Airway Pressure for Patients with Obstructive Sleep Apnea in Singapore from Health System Perspective

Oral Presentation

Dr Meng Fanwen, Dr Ang Yee Gary, Dr Ruth Chang Rui Ying, Dr Lee Chuen Peng, Dr John Abisheganaden

Demand Projection of Obstructive Sleep Apnea (OSA) in the Region of National Healthcare Group in Singapore from 2023 to 2035

Oral Presentation

Dr Meng Fanwen, Dr Ang Yee Gary, Dr Ruth Chang Rui Ying, Dr Lee Chuen Peng, Dr John Abisheganaden

Projection of Total FTEs and Number of Sleep Physicians in Singapore based on Obstructive Sleep Apnea Demand during 2025 to 2035

Oral Presentation

Dr Meng Fanwen, Dr Ang Yee Gary, **Dr Ruth Chang Rui Ying**, Dr Lee Chuen Peng, Dr John Abisheganaden

International Conference on Integrated Care (Singapore)

Assessing the Effectiveness of an Integrated Primary and Tertiary Multidisciplinary Diabetic Foot Care Programme: A Historical Control Study

Oral Presentation

Lixia Ge, Dr Ang Yee Gary, Dr Joseph Antonio D. Molina, Dr Sun Yan, Dr Elaine Tan, Dr Liew Huiling, Dr Jeremy Hoe, Dr Jaime Lin, Dr Joseph Lo Zhiwen

Value-based Healthcare Conference 2024 (Singapore)

Best Poster Award

Cost-effectiveness Analysis of Diabetic Foot Screening for Ulcerations: AI-Enhanced Risk Tailored Screening vs. Routine Annual Screening

Poster Presentation

Dr Sun Yan, Lixia Ge, Dr Ang Yee Gary, Dr Joseph Lo Zhiwen, Dr Liew Huiling, Donna Tan, Danial Chew, Dr John Abisheganaden

Top 10 Best Poster

Impact of podiatry services in an integrated diabetic limb salvage programme (DEFINITE) on health outcomes and healthcare utilisation

Poster Presentation

Lixia Ge, Dr Ang Yee Gary, Dr Sun Yan, Dr Joseph Antonio D. Molina, Tiffany Chew, Chelsea Law, Jo Ann Lim, Dr Joseph Lo Zhiwen

Shortlisted for Best Poster Award

Cost-effectiveness Analysis of Continuous Positive Airway Pressure Treatment for Obstructive Sleep Apnea in Singapore from a Health System Perspective

Poster Presentation

Dr Meng Fanwen, Dr Ang Yee Gary, Dr Ruth Chang Rui Ying, Dr Lee Chuen Peng, Dr Kelvin Bryan Tan, Dr John Abisheganaden

Improving the patient journey Improving the Patient Journey for Lung Function Tests using Artificial Intelligence

Poster Presentation

Dr Ang Yee Gary, James Nguyen, Dr Tan Geak Poh, Kane Nguyen, Wang Li Jia, Kelly Lee

18th Singapore Public Health & Occupational Medicine Conference
(Singapore)

1st prize for Best Poster Award

Development and validation of the Health Activation Scale for Children (HAS-C): an important intermediate outcome measure for health promotion initiatives

Poster Presentation

Lixia Ge, Dr Joseph Antonio D. Molina, Foo Hui Ting, Marcus Tang, Rochelle Chua Ming Jie, Dr Ong Chin Fung

Developing and validating a self-report scale for assessing health resilience in the community-dwelling adults: a cross-sectional study

Oral Presentation

Lixia Ge, Dr Yip Wan Fen, Dr Li Ruijie, Eric Chua Siang Seng, Dr Ho Moon-Ho Ringo, Dr Andy Hau Yan Ho, Evon Chua Yiwen, Dolly Cheng S.C., Dr Ian Leong Yi Onn, Chieh Pann Pei, Dr Tan Woan Shin.

Exploring the associations between health confidence, self-efficacy, health activation and healthy behaviours in primary school children

Poster Presentation

Lixia Ge, Dr Joseph Antonio D. Molina, Foo Hui Ting, Rochelle Chua Ming Jie, Dr Ong Chin Fung

10th International Society for Physical Activity and Health Congress (ISPAH)
(Paris, France)

A geriatric inpatient mobility-focused intervention improves function while requiring attention on its feasibility and sustainability

Poster Presentation

Dr Lau Lay Khoon, Dr Michelle Jessica Pereira, Dr Tou Nien Xiang, Chua Shi Min, Russell Yoong Kuok Leong, Dr Lim Seok Mei, Chan Poh Choo, Dr Lydia Au Shu Yi, Dr Ding Yew Yoong, Dr Melanie Tan Mei Ying

Korea Academy of Tuberculosis and Respiratory Diseases International Conference 2024 (KATRDIC 2024)
(Korea)

Forecasting the Demand of Obstructive Sleep Apnea in the National Healthcare Group Region of Singapore (2025 – 2035)

Poster Presentation

Dr Meng Fanwen, Dr Ang Yee Gary, Dr Ruth Chang Rui Ying, Dr Lee Chuen Peng, Dr Kelvin Bryan Tan, Dr John Abisheganaden

Cost-effectiveness Analysis of Continuous Positive Airway Pressure Therapy for Obstructive Sleep Apnea in Singapore from a Health System Perspective

Poster Presentation

Dr Meng Fanwen, Dr Ang Yee Gary, Dr Chang Ruth Rui Ying, Dr Lee Chuen Peng, Dr Kelvin Bryan Tan, Dr John Abisheganaden

Projection of Full-Time Equivalent and Number of Sleep Physicians in Singapore based on Obstructive Sleep Apnea Demand during 2025 to 2035

Poster Presentation

Dr Meng Fanwen, Dr Ang Yee Gary, Dr Ruth Chang Rui Ying, Dr Lee Chuen Peng, Dr Kelvin Bryan Tan, Dr John Abisheganaden

Singapore Health & Biomedical Congress 2024
(Singapore)

Best Poster Award (Health Services Research): Gold

Impact of podiatry services in an integrated diabetic limb salvage programme (DEFINITE) on health outcomes and healthcare utilisation

Poster Presentation

Lixia Ge, Dr Ang Yee Gary, Dr Sun Yan, Dr Joseph Antonio D. Molina, Tiffany Chew, Chelsea Law, Jo Ann Lim, Dr Joseph Lo Zhiwen

Best Poster Award (Nursing): Gold

Development and validation of the Health Activation Scale for Children (HAS-C): an important intermediate outcome measure for health promotion initiatives

Poster Presentation

Lixia Ge, Dr Joseph Antonio D. Molina, Foo Hui Ting, Marcus Tang, Rochelle Chua Ming Jie, Dr Ong Chin Fung

Best Poster Award (Health Services Research): Silver

Different 'strokes' for different folks (who return-to-work): Secondary analysis of employment outcomes from an evaluation of a post-stroke care-coordination programme

Poster Presentation

Dr Michelle Jessica Pereira, Joey Ha Wei Yee, Ryan Cheok Chi Min, Nuez Odessa Setiota, Deshan Kumar, Carol Tham, Lim Pang Hung, Teo Li Huan, Nur Nazaria Baharudin, Sng Yee Leng, Loh Yong Joo, Dr Yip Wan Fen

Best Poster Award (Health Services Research): Merit

Association of capability, motivation and opportunity with physical activity and diet behaviours in patients with Type 2 diabetes

Poster Presentation

Loo May Eng, Dr Charmaine Krishnasamy, Dr Cheng San Chye, Eileen Bei, Dr Khoo Hwee Sing, Dr Mary Lee, Dr Tan Shuyun, Pauline Xie, Dr Sabrina Wong, Dr Lee Eng Sing, Dr Sabrina Lee, Shannon Goh, Dr Tang Wern Ee

Singapore Young Investigator Award (Health Services Research): Bronze

Improving palliative outcomes and honouring patient's preference: A prospective analysis of the Integrated Management and Palliative Care for Terminally ill Non-cancer Patients (IMPACT) programme

Oral Presentation

Palvinder Kaur, Michelle Jessica Pereira, Joseph Antonio D. Molina, Darryl Kok, Wu Huei Yaw

Singapore Young Investigator Award (Health Services Research): Merit

Identifying archetypes of healthcare expenditure trajectories in decedents with end-stage organ diseases

Oral Presentation

Sheryl Ng Hui Xian, Palvinder Kaur, Tay Ri Yin, Dr Laurence Tan Lean Chin, Dr Mervyn Koh Yong Hwang, Dr Andy Ho Hau Yan, Dr Allyn Hum, Dr Tan Woan Shin

SHBC Medical Humanities 2024 Token of Appreciation

What does Health Means to me

Poster Presentation

Dr Ang Yee Gary

Invited to the inaugural research showcase "The next big idea" for Medical & House Officers (Junior doctors)/Residents/Students

Improved short-term knee function and reduced pain interference: Preliminary findings of the potential benefit from acupuncture and heat therapy for patients with knee osteoarthritis

Poster Presentation

Jessica Widyawati, Dr Michelle Jessica Pereira, Tan Chun Yue, Dr Tan Tong Leng, Dr Yang Su-Yin, Gillian Chew Long Szu, Tan Siang Ing, Chua Yu Chun, Yan Yew Wai, Derrick Soh Bing Quan, Goh Tong Hwee, Ng Pu Jue, Ng Yee Teck, Kuan Swee Boey, Teo Bo Siang, Dr Kong Keng He, Dr Bryan Tan Yijia

Remaining and passing on at home: What a specialist-led homecare programme can help end-stage organ failure patients with

Poster Presentation

Dr Michelle Jessica Pereira, Palvinder Kaur, Dr Joseph Antonio D. Molina, Darryl Kok, Dr Wu Huei Yaw

Living with Diabetes: A Survey of Patient Experiences and Influencing Factors On Diet and Physical Activity

Poster Presentation

Dr Charmaine Krishnasamy, Loo May Eng, Dr Cheng San Chye, Eileen Bei, Dr Khoo Hwee Sing, Dr Mary Lee, Dr Tan Shuyun, Pauline Xie, Dr Sabrina Wong, Dr Lee Eng Sing, Dr Sabrina Lee, Shannon Goh, Dr Tang Wern Ee

Evaluating the economic impact and cost-effectiveness of Workplace Health Promotion Interventions: Insights for Healthier Sg Strategy

Poster Presentation

Jiang Wenli, Carmen Lim, **Nguyen Que Chi**, Do Diem Quynh, Nguyen Truong Giang George, Dr Ang Yee Gary

Singapore Health & Biomedical Congress 2024

(Singapore)

Cost-effectiveness Analysis of Diagnosis and Treatment with Continuous Positive Airway Pressure for Patients with Obstructive Sleep Apnea in Singapore from a Health System Perspective

Poster Presentation

Dr Meng Fanwen, Dr Ang Yee Gary, Dr Ruth Rui Ying Chang, Dr Lee Chuen Peng, Dr Kelvin Bryan Tan, Dr John Abisheganaden

Exploring the associations between health confidence, self-efficacy, health activation and healthy behaviours in primary school children

Poster Presentation

Lixia Ge, Dr Joseph Antonio D. Molina, Foo Hui Ting, Rochelle Chua Ming Jie, Dr Ong Chin Fung

Self identification of Rehabilitation Needs for Stroke Patients

Poster Presentation

Summer Chow Shi Qi, Dr Yip Wan Fen , Joey Ha Wei Yee, Ryan Cheok Chi Min, Celeste Ci Ying Gue, Nuez Odessa Setiota, Deshan Kumar, Carol Tham, Lim PangHung, Teo Li Huan, Nur Nazaria Baharudin, Sng Yee Leng, Loh Yong Joo, Dr Michelle Jessica Pereira

Scoping Out Knee Arthroscopies: 10-year trends from 2014 to 2023 at NHG

Poster Presentation

Chen Wei Ling, Christopher Ng Thong Lian, Mohd Khalid Bin Mohd Anuar, Shiek Abdullah Bin Ismail, Ivan Tan Peng Choon, Dr Michelle Jessica Pereira

Critical influences that contributed to high quality care at NHG's primary eye care model

Poster Presentation

Dr Yip Wan Fen, Celeste Ci Ying Gue , Dr Michelle Jessica Pereira , Dr Joseph Antonio D. Molina , Robyn Tan , Dr Ding Si Yan

Forecasting Demand of Obstructive Sleep Apnea in the Region of National Healthcare Group of Singapore during 2025 to 2035

Poster Presentation

Dr Meng Fanwen, Dr Ang Yee Gary, Dr Ruth Rui Ying Chang, Dr Lee Chuen Peng, Dr Kelvin Bryan Tan, Dr John Abisheganaden

Worsening psychosocial factors are associated with poorer health-related quality of life: A 1-year longitudinal study

Poster Presentation

Joey Ha Wei Yee, Dr Michelle Jessica Pereira, Ge Lixia, Dr Aidan Tan Lyanzhiang, Dr Yap Chun Wei, Dr Joseph Antonio D. Molina

Combining generative AI and machine learning to improve prediction of high healthcare utilizers

Poster Presentation

Dr Yap Chun Wei

Teamworking in a multi-disciplinary diabetic foot clinic: Podiatrists' bridging role

Poster Presentation

Dr Mary Lee, Dr Winnie Teo, Dr Yong Enming, Dr Hong Qian Tai, Dr Zhang Li, Dr Loh Zhiwen, Dr Brenda Lim, Dr Chong Rhan Chaen, Dr Glenn Tan, Tiffany Chew, Faezah Sani, Dr Muhd Farhan bin Mohd Fadil, Dr Liew Huiling

**1st Asia Pacific
Population Health
Conference & 2nd Asia
Pacific Social Prescribing
Conference 2024**
(Singapore)

Development and Psychometric Evaluation of the Health Activation Scale for Children (HAS-C)
ePoster Presentation
Lixia Ge, Dr Joseph Antonio D. Molina, Foo Hui Ting, Karthigayan Ramakrishnan, Marcus Tang, Rochelle Chua Ming Jie, Dr Ong Chin Fung

**Asian Pacific Society
of Respirology (APSR)
2024**
(Hong Kong)

Longitudinal Study on Quality of life of Patients and Caregivers Receiving Home Mechanical Ventilation (HMV) and Respiratory Support
Poster Presentation
Dr Tan Geak Poh, Chieh Pann Pei, Dr Michelle Jessica Pereira, Joey Wei Yee Ha, Dr Joseph Antonio D. Molina, Lydia Soon Hse Yin, Sun Tao, Dr Adrian Tan Kok Heng, Dr Kor Ai Ching, Dr Chan Yeow

IMPACT conference 2024
(Switzerland)

High-Quality Care at Singapore's Primary Eye Care Model: A Mixed Methods Study
Poster Presentation
Dr Yip Wan Fen, Celeste Ci Ying Gue, Dr Michelle Jessica Pereira, Dr Joseph Antonio D. Molina, Dr Ding Si Yan, Dr Jonathan Foo, Dr Vivien Yip Cherng Hui, Dr Vernon Yong Khet Yau, Dr Wong Hon Tym, Dr Lim Tock Han

**International Allied
Health Conference 2024**
(Singapore)

Assessing Care Quality at The Virtual Ophthalmology Service (TVOS): A Study of Optometrist-Provided Care in the Community
Oral Presentation
Denise Chloe Chua Wenxin, Dr Yip Wan Fen, Dr Michelle Jessica Pereira, Dr Joseph Antonio D. Molina, Dr Vivien Yip Cherng Hui, Quek Zuoling, Aw Ling Li, Dr Vernon Yong Khet Yau, Dr Wong Hon Tym, Dr Lim Tock Han

**IUBMB-FEBS-ENABLE
Conference 2024**
(Singapore)

Identifying archetypes of healthcare expenditure trajectories in decedents with end-stage organ diseases
Poster Presentation
Sheryl Ng Hui Xian, Palvinder Kaur, Tay Ri Yin, Dr Laurence Tan Lean Chin, Dr Mervyn Koh Yong Hwang, Dr Andy Ho Hau Yan, Dr Allyn Hum, Dr Tan Woan Shin

Associations between OGTT Curve Morphology and Physiological, Metabolic, and Exercise Profiles in Healthy Adults: A Cross-Sectional Analysis
Poster Presentation
Dr Meng Fanwen, Dr Ang Yee Gary, Palvannan R.K., Margaret Yap Mei Chan, Dr Lee Shuen Yee, Dr Melvin Leow Khee Shing, Dr Lim Chin Leong

Leveraging Artificial Intelligence to Enhance Systematic Reviews in Health Research: Advanced Tools and Challenges
Poster Presentation
Lixia Ge, Dr Rupesh Agrawal, Dr Maxwell Singer, Palvanna Kannapiran, Dr Joseph Antonio D. Molina, Kiok Liang Teow, Dr Yap Chun Wei, Dr John Abisheganaden

AI-assisted pharmacy intervention classification
Poster Presentation
Dr Yap Chun Wei, Sia Beng Yi, Tan Wei Qin, Lim Hong Yee, Dr John Abisheganaden

Predictive risk score for 30-day readmission
Poster Presentation
Dr John Abisheganaden, Robin Tan, Dr Tay Seow Yian, Dr Terence Tang, Dr Yap Chun Wei

Exploring the use of ChatGPT 4o to aid thematic analysis
Poster Presentation
Dr Mary Lee, Dr Charmaine Krishnasamy, Dr Li Ruijie

Cost-Effectiveness Analysis of the DEFINITE Care Programme for Patients with

Poster Presentation

Dr Yan Sun, Lixia Ge, Dr Gary Ang, Dr Elaine Tan, Dr Huiling Liew, Jeremy Hoe Kian Ming, Jaime Lin, Joseph Antonio De Castro Molina, John Arputhan Abisheganaden, Dr Zhiwen Joseph Lo

Evaluating the impact of Vaccination Strategies on Pneumonia Hospitalisation Rates in Elderly Aged 65 and Older in 2022: A Multi-Variable Big Data Study of Influenza and Pneumococcal Vaccines

Poster Presentation

Matthias Koh Yong An, Dr Yap Chun Wei, Dr John Abisheganaden

Enhancing Lung Function Tests Using Biodesign Principles and AI Chatbots

Oral Presentation

Dr Ang Yee, Gary, Ernest Cuong, James Nguyen, Dr Tan Geak Poh, Kane Ngyuen, Wang Lijia, Prof John Abisheganaden

Enhancing Lung Function Tests Using AI Chatbot

Oral Presentation

Ernest Hoang, Dr Ang Yee Gary, Dr John Arputhan Abisheganaden, Dr Tan Geak Poh, James Nguyen, Kane Nguyen, Wang Lijia

Identification of disease progression and recommendation of interventionable timepoint: Using ChatGPT to analyse Timeline Visualisation of Home Ventilation Patients' Disease Trajectory

Oral Presentation

Joey Ha Wei Yee, Dr Michelle Jessica Pereira, Dr Joseph Antonio D. Molina, Dr Tan Geak Poh, Dr Adrian Tan Kok Heng, Dr Kor Ai Ching, Lydia Soon Hse Yin, Sun Tao, Dr Chan Yeow



THE TEAM



The Team



Professor John Abisheganaden
MBBS, MRCP, M.Med, FAMS, FRCP
Director and Senior Consultant



Christina Chieh Pann Pei
BSc (Biotechnology)
Senior Executive



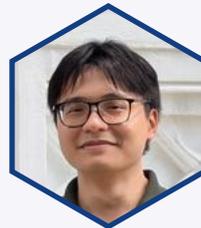
Dr Cheng San Chye
BSc (Hons), EdM, MSc, EdD (Doctor of Education)
Head of Social & Behavioural Sciences Outcomes Research



Eric Chua Siang Seng
BSc (Hons) (Business Management),
Dip (Electrical & Electronics Engineering)
Senior Executive



Dr Joseph Antonio D. Molina
MD, MSc (Public Health)
Head of Evaluation Science and Outcomes Research



Ernest Hoang Le Tri Cuong
BComp (Hons) (Computer Science)
Executive



Palvannan R. K.
BEng, MEng (Industrial Engineering)
Head of Data Science & Operations Research



Ge Lixia
BMed (Hons)(Nursing), MSc (Physiology), MPH
Principal Research Analyst



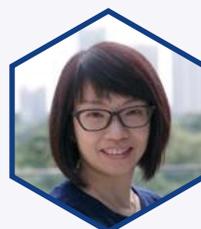
Dr Aidan Lyanzhiang Tan
MBBS, MPH, Dip (Family Med), GDOM
Consultant (Public Health)



Gloria Ho
BA (Psychology), MPH
Senior Research Analyst



Dr Ang Yee Gary
MBBS, MPH, Dip (Family Med), GDMH,
Dip (Family Practice Dermatology)
Consultant (Public Health)



Jasmine Tan Hwee Ling
Adv Dip B. S. (Business Studies)
Executive



Dr Charmaine Krishnasamy
Dip (Occupational Therapy),
Bachelor of Occupational Therapy (Hons), PhD
Principal Research Analyst



Joey Ha Wei Yee
BSocSci (Hons) (Economics)
Research Assistant



Dr Li Ruijie

Dip (Occupational Therapy),
MSc (Occupational Therapy), PhD (Biostatistics)
Principal Research Analyst



Sheryl Ng Hui Xian

BSc (Hons) (Statistics), MPH
Senior Research Analyst



Dr Mary Lee Ching Ling

BA, MA (English Language),
PhD (Communications)
Principal Research Analyst



Dr Sun Yan

MSc (Data Mining), PhD (Medical Informatics)
Senior Principal Research Analyst



Dr Meng Fanwen

MSc (Operations Research),
PhD (Operations Research)
Principal Research Analyst



Teow Kiok Liang

BEng (Electrical Engineering),
MSc (Industrial & Systems Engineering)
Senior Principal Research Analyst



Dr Michelle Jessica Pereira

BPhty (Hons I), MPhty (Sports Physiotherapy),
PhD (Public Health and Health Services,
Applied Health Economics)
Principal Research Analyst



Toh Chai Hoon

Bachelor of Information Technology
Assistant Manager



Ng Siok Leng

BBA (Hons), MEd, MSc (Statistics)
Assistant Director



Dr Yap Chun Wei

BSc (Hons) (Pharm), PhD
Senior Principal Research Analyst



Palvinder Kaur

BSc (Biomedical), MSc (Public Health)
Principal Research Analyst



Dr Yip Wan Fen

BSc (Hons) (Optometry), PhD (Ophthalmology)
Principal Research Analyst



Samuel Wang Xiande

BSc, MSc (Physics)
Clinical Informatics Analyst

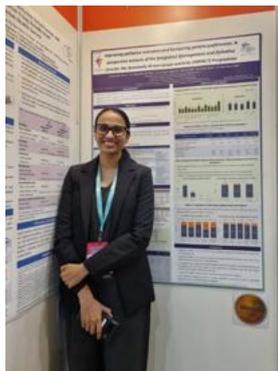
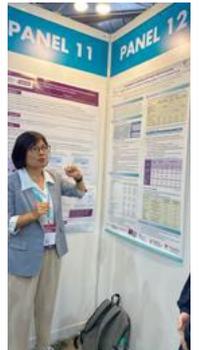
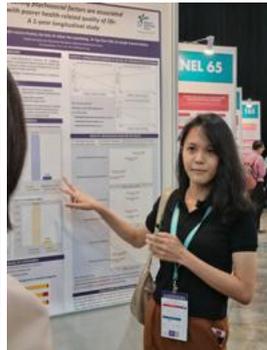
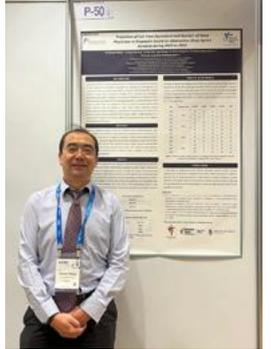


Dr Zhu Zhecheng

MSc (Information Engineering),
PhD (Industrial & Systems Engineering)
Senior Principal Research Analyst

HSOR Activities







Health Services and Outcomes Research (HSOR)
National Healthcare Group

E-mail: hsor@nhg.com.sg

Website: <https://corp.nhg.com.sg/HSOR>

