Generative AI and the Primary Care Practice of the Future: A Case Study

Steven Lin MD

Clinical Professor, Medicine

Chief, General Primary Care

Director, Stanford Healthcare AI Applied Research Team

Division of Primary Care and Population Health

Stanford University School of Medicine

Scope and Learning Objectives

Generative AI is a type of artificial intelligence (AI) system capable of generating text, images, or other media in response to prompts

This session will focus largely on applications of large language models like ChatGPT and Gemini

At the end of this session, participants will be able to:

1. Articulate 5 use cases for generative AI in a primary care journey-centered framework

2. Describe current evidence supporting these use cases

Case Study: Part 1

- You are a primary care physician running behind in clinic

- Your next patient is here for a follow-up hospitalization visit

- You recognize his name (Juan), and vaguely recall that he has diabetes and hypertension, but you haven't seen him in 3 years
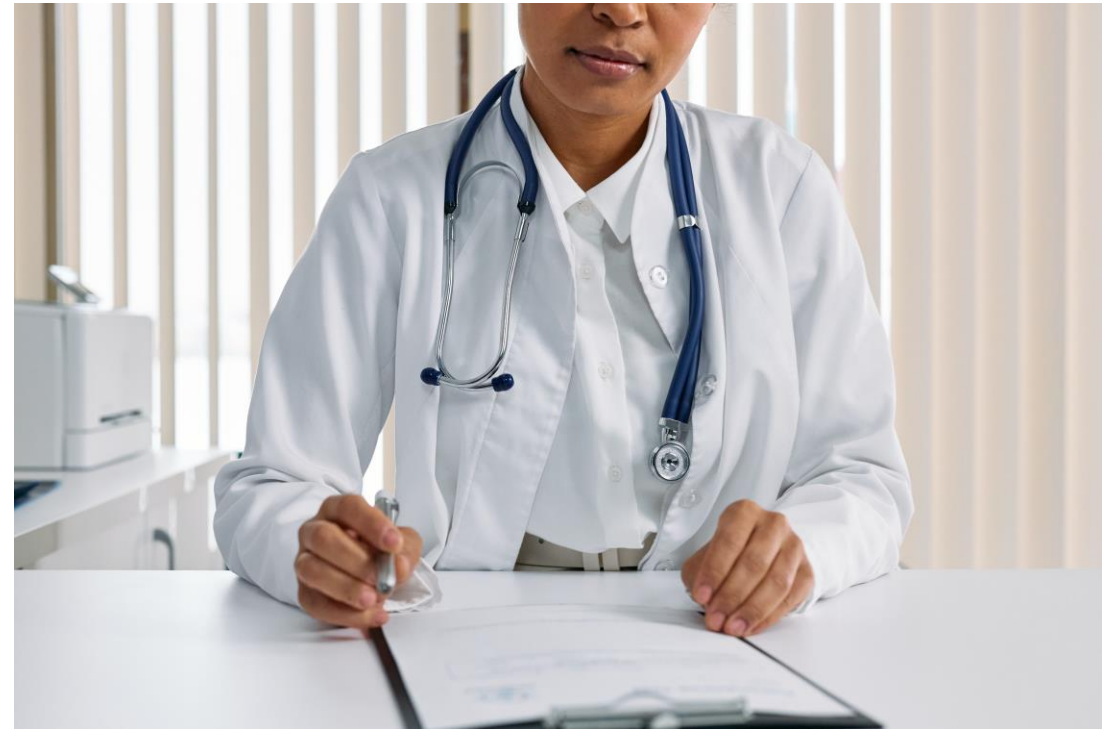
Problem: Data Overload and Fragmentation

- Overabundance and scatter of information within EHRs is a barrier to patient care and key source of physician burnout

- For every hour physicians spend in front of patients, 2 hours are spent in front of the EHR; chart review accounts for the highest proportion of EHR time at 33%

Opportunity: AI Chart Review and Summarization

- AI-based systems that extract, analyze, and summarize patient clinical information from EHRs can reduce physicians' cognitive burden from information chaos

- Study: AI summaries were rated as equivalent (45%) or superior (36%) to summaries from human experts (Nature Medicine, 2024)

nature medicine

Article   https://doi.org/10.1038/s41591-024-02855-5

Adapted large language models can outperform medical experts in clinical text summarization

Dave Van Veen [1,2], Cara Van Uden[2,3], Louis Blankemeier[1,2], Jean-Benoit Delbrouck[2], Asad Aali[4], Christian Bluethgen [2,5], Anuj Pareek [2,6], Malgorzata Polacin[5], Eduardo Pontes Reis[2,7], Anna Seehofnerová[8,9], Nidhi Rohatgi [8,10], Poonam Hosamani[8], William Collins [8], Neera Ahuja[8], Curtis P. Langlotz [2,8,9,11], Jason Hom[8], Sergios Gatidis[2,9], John Pauly[1] & Akshay S. Chaudhari [2,9,11,12]

Case Study – Part 1 Continued

- The AI tells you that Juan was hospitalized one week ago with a new diagnosis of congestive heart failure; it summarizes the key results of all pertinent labs, imaging, and procedures done

- He was discharged home with carvedilol, lisinopril, atorvastatin and metformin

Case Study: Part 2

- You enter the exam room to see Juan, who appears well; his vital signs, history and physical exam show no signs of fluid overload

- You noted from the AI summary that his hemoglobin A1C = 8.5%

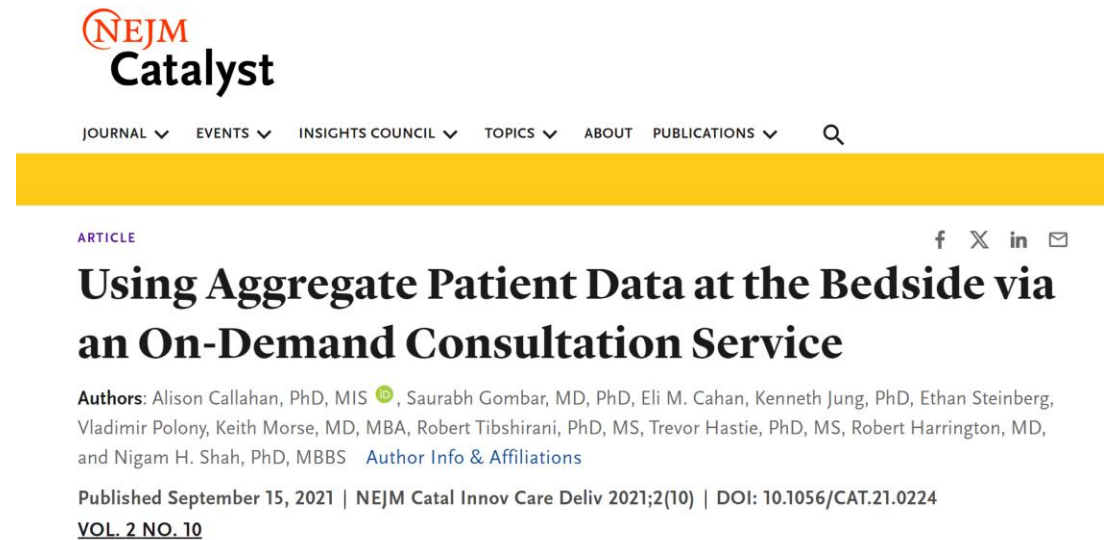- You wonder which medication is the best to add to his regimen?

Problem: Rapidly Changing Evidence & Guidelines

Time To Double Medical Knowledge Is Decreasing
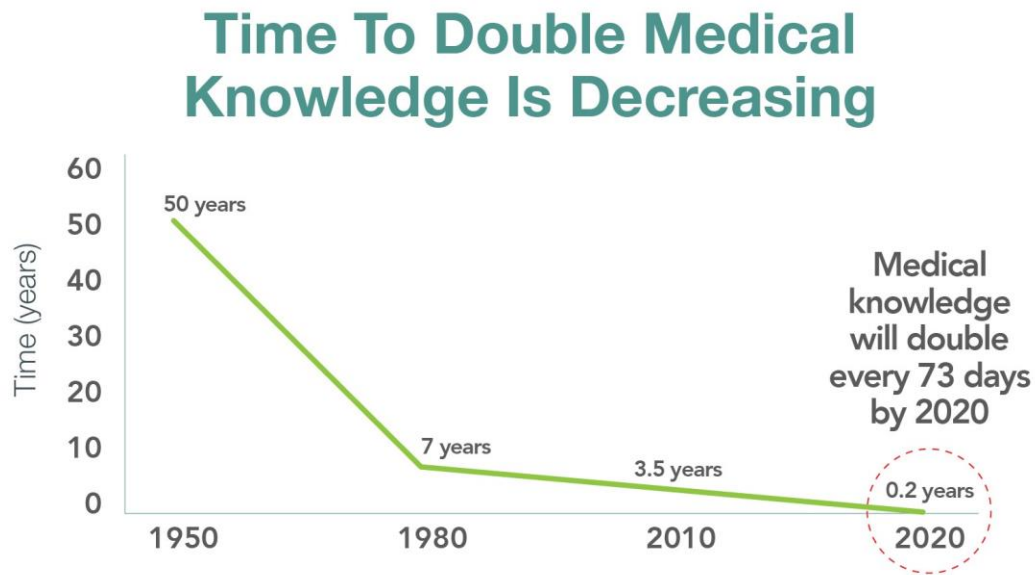
- Medical knowledge doubles every 2.5 months (in 1950, it was 50 years)

- It takes an average of 17 years for evidence to change practice

- Example: <25% of heart failure patients are getting guideline-directed medical therapy

Opportunity: AI Assisted Clinical Decision Support

- AI-based systems trained on up-to-date scientific literature can provide on-demand consultation to answer physician's questions

- AI can also generate real-world evidence from medical records to answer questions that can't be answered by studies (NEJM Catalyst, 2021)

NEJM Catalyst

JOURNAL · EVENTS · INSIGHTS COUNCIL · TOPICS · ABOUT · PUBLICATIONS ·

ARTICLE

Using Aggregate Patient Data at the Bedside via an On-Demand Consultation Service

Authors: Alison Callahan, PhD, MIS, Saurabh Gombar, MD, PhD, Eli M. Cahan, Kenneth Jung, PhD, Ethan Steinberg, Vladimir Polony, Keith Morse, MD, MBA, Robert Tibshirani, PhD, MS, Trevor Hastie, PhD, MS, Robert Harrington, MD, and Nigam H. Shah, PhD, MBBS   Author Info & Affiliations

Published September 15, 2021 | NEJM Catal Innov Care Deliv 2021;2(10) | DOI: 10.1056/CAT.21.0224

VOL. 2 NO. 10

Case Study – Part 2 Continued

- The AI tells you that, based on recent trials and both the 2022 AHA/ACC/HFSA and 2023 ADA guidelines, an SGLT2 inhibitor is strongly recommended for heart failure with or without diabetes

- You start Juan on dapagliflozin, and ask him to come back and see you in 3 months

Case Study: Part 3

- You're ready to send Juan home, but he has numerous questions about his medical conditions, his medications, and what foods he should eat or avoid, etc.

- Your generic patient educational materials don't match his exact set of conditions/medications, and they don't come in Spanish

Problem: Lackluster Patient Instructions/Education

- Patients only remember 49% of the decisions made during talks with their doctors

- Written instructions improve recall, but are time-consuming to generate and are often not aligned with patients' language, literacy level, and unique set of medical conditions/medications

Opportunity: AI Written Personalized Instructions

- AI-based systems that retrieve historical instructions, reason related medical knowledge, and refine language to fit patients' needs can create personalized patient instructions/education

- Study: AI instructions were rated as more understandable (81% vs 13%) than human-written ones (JAMA Netw Open, 2024)

JAMA Network | Open

Original Investigation | Health Informatics

Generative Artificial Intelligence to Transform Inpatient Discharge Summaries to Patient-Friendly Language and Format

Jonah Zaretsky, MD; Jeong Min Kim, MD; Samuel Baskharoun, MD; Yunan Zhao, MS; Jonathan Austrian, MD; Yindalon Aphinyanaphongs, MD, PhD; Ravi Gupta, MD; Saul B. Blecker, MD; Jonah Feldman, MD

Abstract

IMPORTANCE By law, patients have immediate access to discharge notes in their medical records. Technical language and abbreviations make notes difficult to read and understand for a typical patient. Large language models (LLMs [eg, GPT-4]) have the potential to transform these notes into patient-friendly language and format.

OBJECTIVE To determine whether an LLM can transform discharge summaries into a format that is more readable and understandable.

DESIGN, SETTING, AND PARTICIPANTS This cross-sectional study evaluated a sample of the discharge summaries of adult patients discharged from the General Internal Medicine service at NYU (New York University) Langone Health from June 1 to 30, 2023. Patients discharged as deceased were excluded. All discharge summaries were processed by the LLM between July 26 and August 5, 2023.

Key Points

Question Can a large language model transform discharge summaries into a format that is more readable and understandable for patients?

Findings In this cross-sectional study of 50 discharge summaries, understandability scores were significantly higher for patient-friendly discharge summaries. Summaries were rated entirely complete in 56 of 100 reviews, but 18 reviews noted safety concerns involving omissions and inaccuracies.

Case Study – Part 3 Continued

- The AI generates personalized after-visit instructions for Juan: clear, Spanish-language, patient-friendly education for his unique set of conditions, medications, a culturally sensitive diet plan, and a heart failure action plan

- Juan goes home feeling more confident and empowered

Case Study: Part 4

- You are finally finished with your clinic for the day

- You have not written any of your progress notes

- It's your daughter's birthday and you had promised her that you would be home for dinner
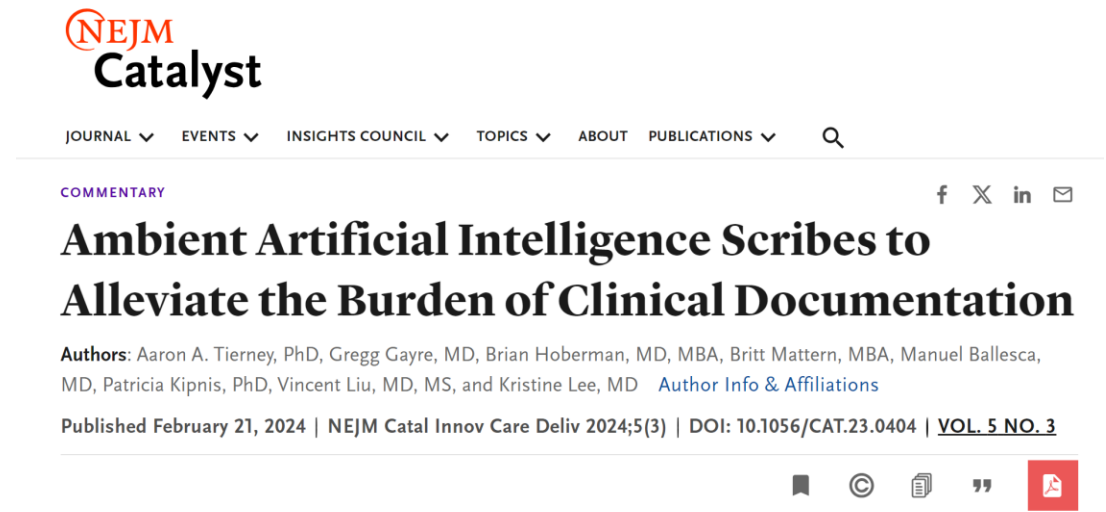
Problem: Burden of Clinical Documentation

- Burden of documentation within EHRs is a key source of physician burnout and lost productivity

- For every hour physicians spend in front of patients, 2 hours are spent in front of the EHR; clinical documentation accounts for the second highest proportion of EHR time

Opportunity: Ambient AI Scribes

- AI systems that capture audio from conversations, convert audio to text, and summarize EHR data can draft notes for providers to review and edit

- Study: AI lowers charting time by 10 to 25%, and is associated with high provider/patient satisfaction (NEJM Catalyst, 2024)

NEJM Catalyst

JOURNAL ⌄   EVENTS ⌄   INSIGHTS COUNCIL ⌄   TOPICS ⌄   ABOUT   PUBLICATIONS ⌄   🔍

COMMENTARY

Ambient Artificial Intelligence Scribes to Alleviate the Burden of Clinical Documentation

Authors: Aaron A. Tierney, PhD, Gregg Gayre, MD, Brian Hoberman, MD, MBA, Britt Mattern, MBA, Manuel Ballesca, MD, Patricia Kipnis, PhD, Vincent Liu, MD, MS, and Kristine Lee, MD   Author Info & Affiliations

Published February 21, 2024 | NEJM Catal Innov Care Deliv 2024;5(3) | DOI: 10.1056/CAT.23.0404 | VOL. 5 NO. 3

Case Study – Part 4 Continued

- The AI scribe has drafted all of your progress notes

- You review the drafts, edit them, and sign your notes in minutes

- You get home in time for dinner, much to your daughter's delight

Case Study: Part 5

- One month later, Juan feels unwell after a weekend of partying and eating BBQ

- He's not sure if he needs to see you, so he messages you using the portal on a Thursday night: "Hey Doc, it's hard to catch my breath and my feet are swollen. Should I be worried?"

Problem: Burden of Electronic Patient Messages

- The COVID pandemic hastened the adoption of virtual care, resulting in a 1.6-fold increase in electronic patient messages

- Current approaches like limiting number of messages patients can send, billing for responses, or delegating responses to less trained staff limit access to care

Opportunity: AI Triage and Response to Messages

- Large language models can draft responses to patient messages and triage or provide medical advice with human-quality text

- Study: Use of AI-generated replies to patient messages was associated with improvements in physician burden and burnout (JAMA Netw Open, 2024)

JAMA Network | Open™

Original Investigation | Health Informatics
Artificial Intelligence–Generated Draft Replies to Patient Inbox Messages

Patricia Garcia, MD; Stephen P. Ma, MD, PhD; Shreya Shah, MD; Margaret Smith, MBA; Yejin Jeong, BA; Anna Devon-Sand, MPH; Ming Tai-Seale, PhD, MPH; Kevin Takazawa, BBA; Danyelle Clutter, MBA; Kyle Vogt, BA; Carlene Lugtu, MCiM; Matthew Rojo, MS; Steven Lin, MD; Tait Shanafelt, MD; Michael A. Pfeffer, MD; Christopher Sharp, MD

Abstract

IMPORTANCE The emergence and promise of generative artificial intelligence (AI) represent a turning point for health care. Rigorous evaluation of generative AI deployment in clinical practice is needed to inform strategic decision-making.

OBJECTIVE To evaluate the implementation of a large language model used to draft responses to patient messages in the electronic inbox.

DESIGN, SETTING, AND PARTICIPANTS A 5-week, prospective, single-group quality improvement study was conducted from July 10 through August 13, 2023, at a single academic medical center (Stanford Health Care). All attending physicians, advanced practice practitioners, clinic nurses, and clinical pharmacists from the Divisions of Primary Care and Gastroenterology and Hepatology were enrolled in the pilot.

Key Points

Question What is the adoption of and clinician experience with clinical practice deployment of a large language model used to draft responses to patient inbox messages?

Findings In this 5-week, single-group, quality improvement study of 162 clinicians, the mean draft utilization rate was 20%, there were statistically significant reductions in burden and burnout score derivatives, and there was no change in time.

Case Study – Part 5 Continued
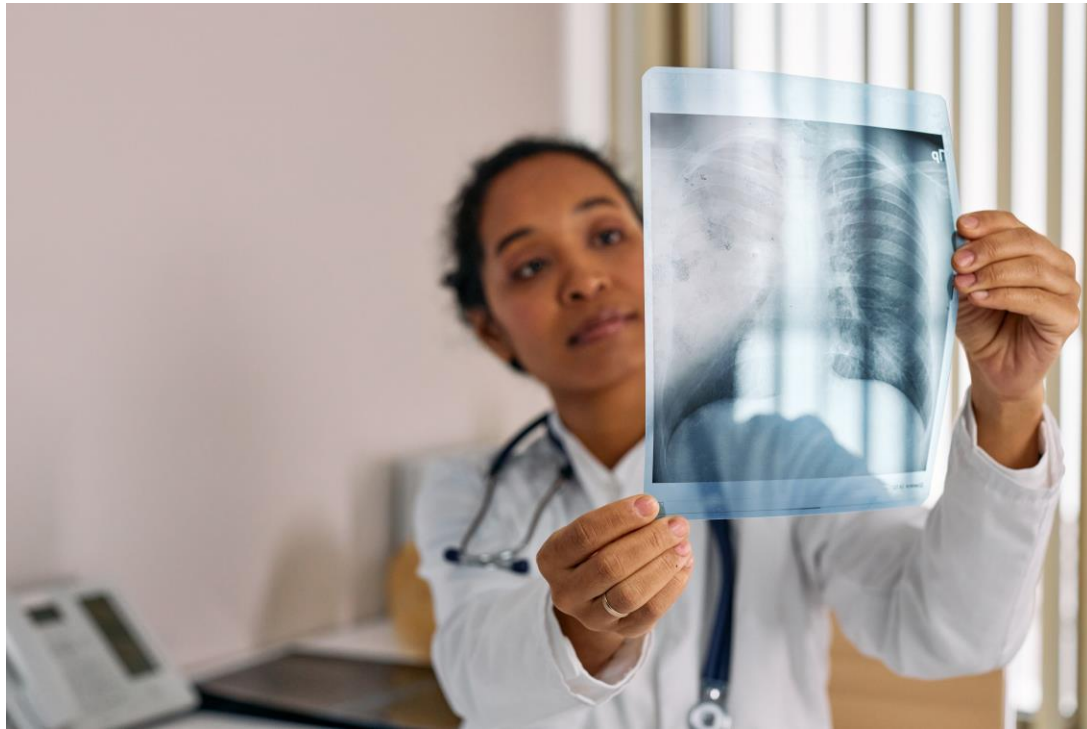
- The AI triages Juan to the on-call nurse, who finds out that he has gained 5 lbs in the last week and has trouble laying flat to sleep; the nurse schedules him a next-day visit with you

- You see Juan the next day, treat him with diuretics, and save him from another hospitalization
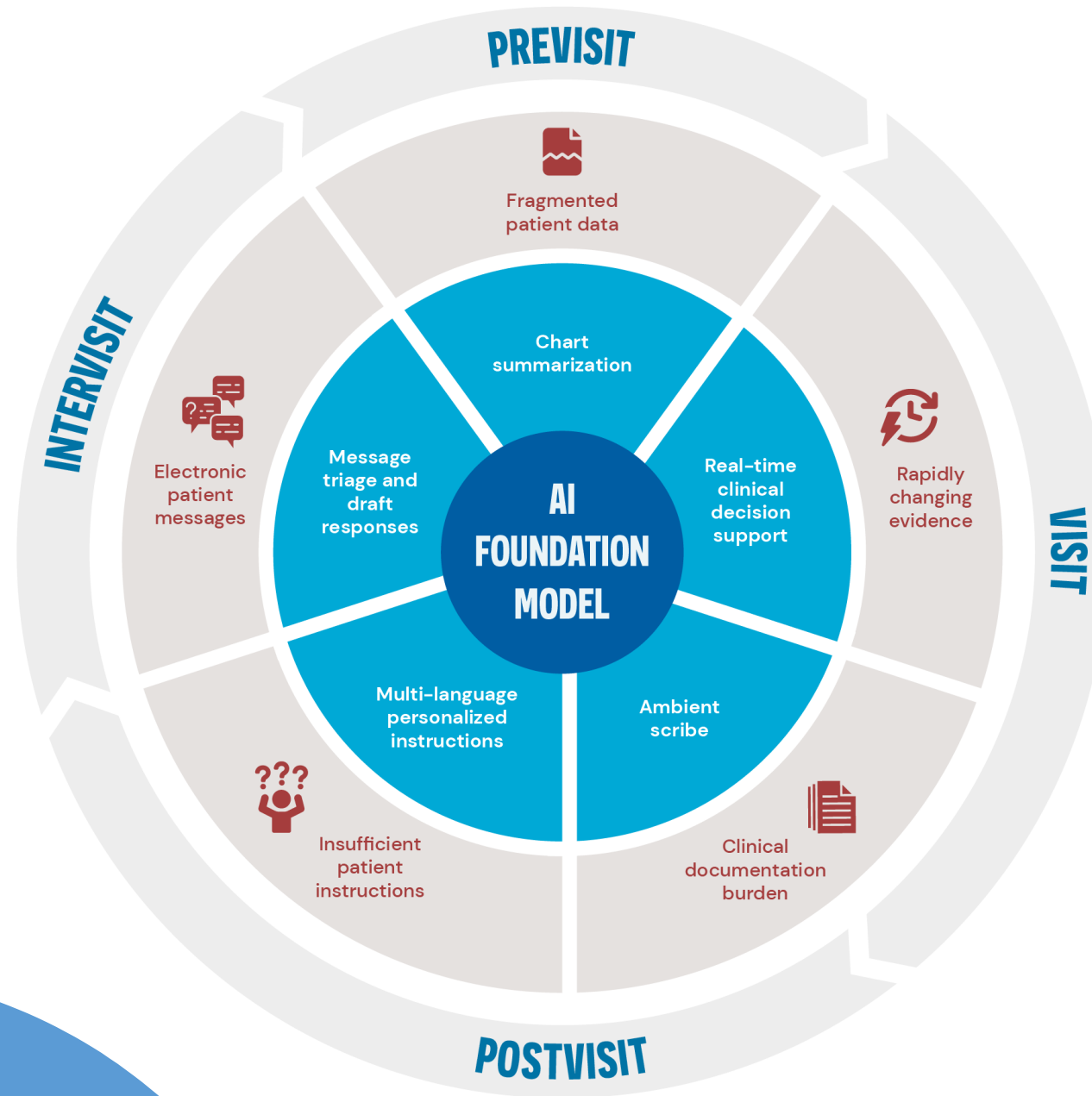
Thank you! Questions?

Email

stevenlin@stanford.edu

LinkedIn

www.linkedin.com/in/steven-lin-md

Scan Me
To visit our website