

The Path we Have Been On – The 10,000 Foot View of Artificial Intelligence

Regis College President's
October 23, 2024

Adam Landman, MD, MS, MIS, MHS

Outline



Introduction to AI and Generative AI



Healthcare AI Use Cases



Generative AI to Improve Care Team Experience

- Ambient Clinical Documentation
- Drafting In Basket Message Responses



Digital Transformation

Healthcare 2024



Other Industries 2024



Healthcare is at a Crossroads



FINANCIAL PRESSURE



WORKFORCE SHORTAGE



LIMITED ACCESS

Burning Platform for Digital Transformation

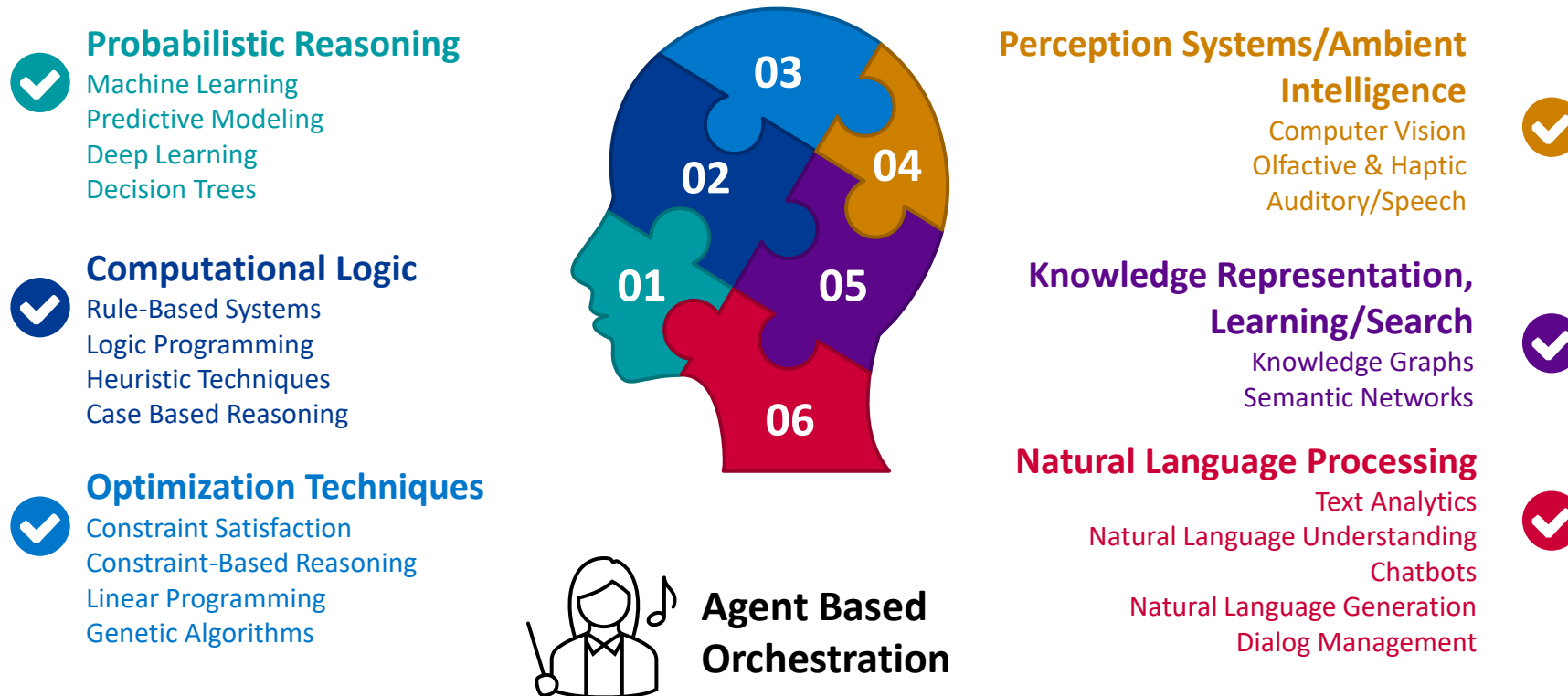


Introduction to AI and Generative AI



What is Artificial Intelligence (AI)?

- “Artificial Intelligence refers to the development of computer algorithms that can perform tasks that typically require human intelligence, such as learning, reasoning, perception, and decision making” (ChatGPT)
- AI is based on machine learning algorithms and other computational techniques:



Two Types of Artificial Intelligence

1. Analytical AI

Analyze datasets to reveal novel insights

- Classification
- Prediction
- Recognition
- Other models

Analytical AI Example: Segmentation of GI Polyps

Use deep learning convolutional neural networks (CNNs) to identify gastrointestinal polyps



Two Types of Artificial Intelligence

2. Generative AI

Generate novel content informed by training dataset

- Summarization
- Generation
- Interaction

Generative AI Examples:

- Text (Bard, ChatGPT)
- Images (DALL-E2)
- Code (Github Copilot)
- Video/Speech (Synthesia)



Generative AI: ChatGPT

What is GPT?

- Large language model (LLM) chatbot developed by OpenAI based on generative pre-trained transformer (GPT)-3.5
- LLMs predict the next word in a series of words in a sentence and the next sentences
- GPT-3 has 175 billion parameters and was trained on ~45 terabytes of text and information from the internet
- Also trained using human feedback (a technique called Reinforcement Learning with Human Feedback) so that the AI learned what humans expected when they asked a question
- Interacts in a text-based conversational way

What is ChatGPT? (As answered by ChatGPT)



What is ChatGPT?



I am ChatGPT, a large language model developed by OpenAI. I have been trained on a massive dataset of text using deep learning techniques, which allows me to understand and generate human-like responses to a wide range of questions and prompts. As a language model, my primary function is to generate natural language responses to user inputs, whether they are in the form of text or speech. I can assist with a variety of tasks, from answering questions and providing information to engaging in conversations and even generating creative writing prompts or stories.



Generative AI: ChatGPT



Early Success

Performed at or near the passing threshold (60%) on the US Medical Licensing Examination (USMLE)

Performed better than other Symptom Checkers on standard clinical vignettes



Limitations

ChatGPT sometimes writes plausible-sounding but incorrect or nonsensical answers

Can exhibit undesirable behavior, including known racial, gender, and religious biases

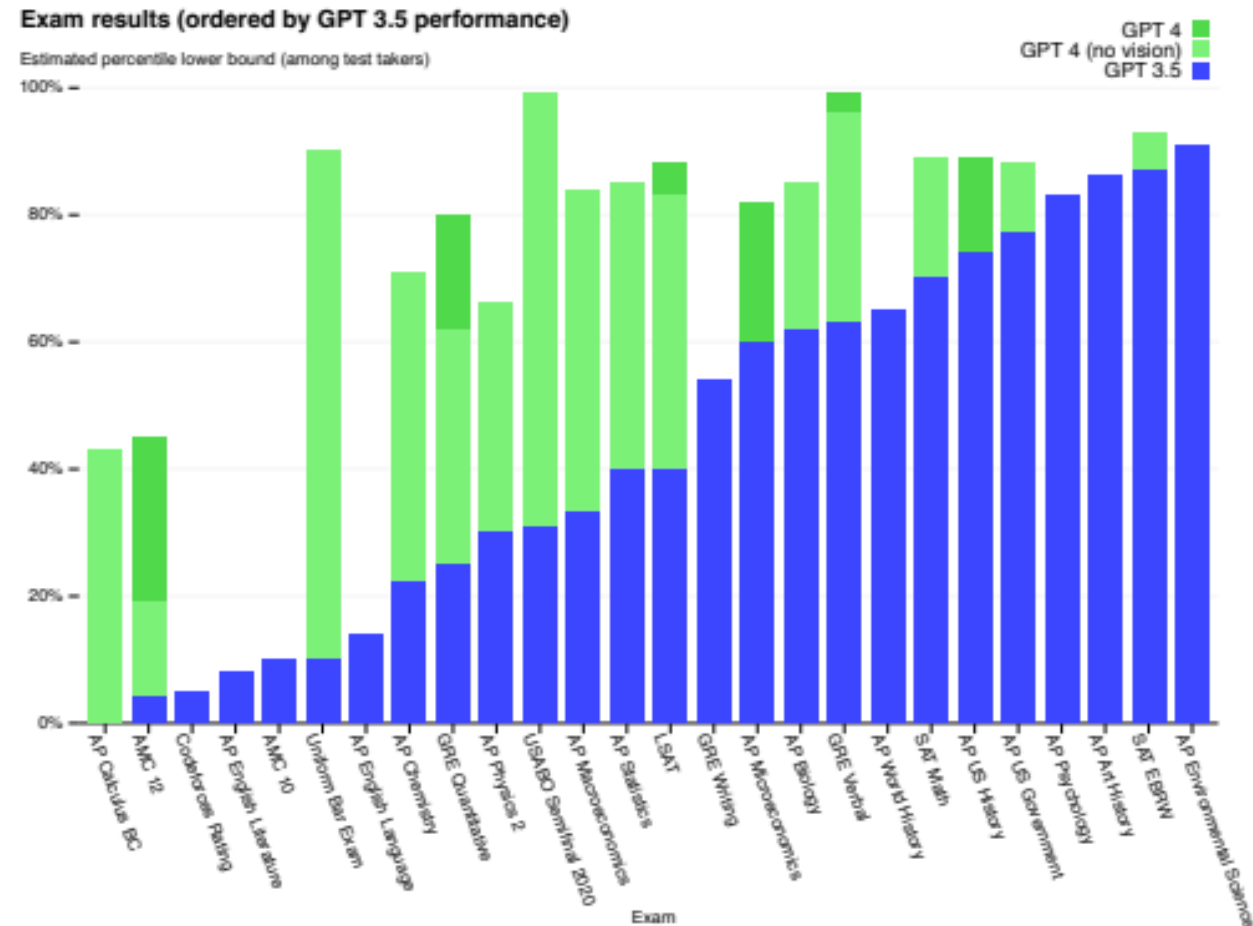
Sensitive to tweaks in the input phrasing or attempting the same prompt multiple times



<https://hai.stanford.edu/news/how-large-language-models-will-transform-science-society-and-ai>
<https://www.usatoday.com/story/news/health/2023/02/26/chatgpt-medical-care-doctors/11253952002/>
<https://www.statnews.com/2023/02/13/chatgpt-assisted-diagnosis/>
<https://journals.plos.org/digitalhealth/article?id=10.1371/journal.pdig.0000198>

GPT-4 Multimodal Model with Improved Performance

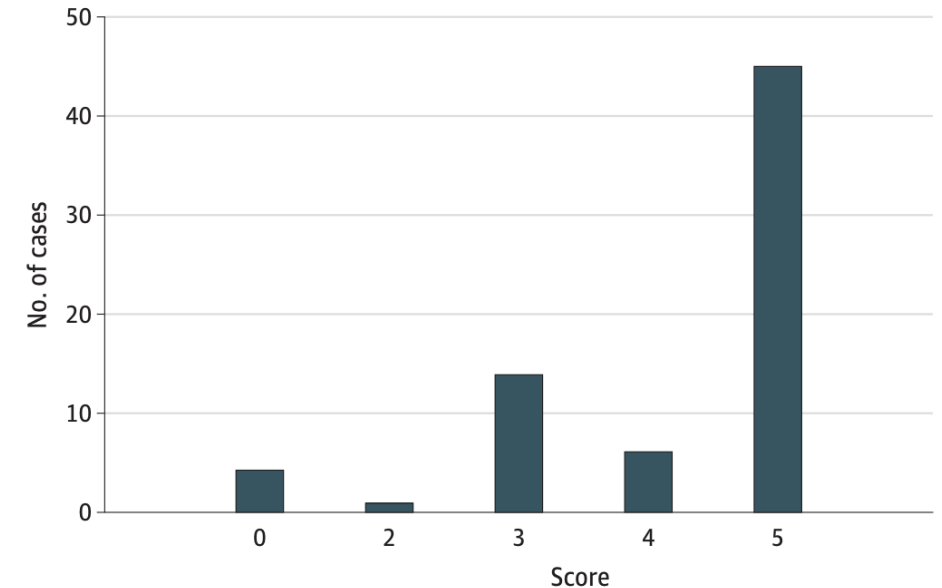
- GPT4 is a large multimodal model that can process image and text inputs and produce text outputs
- Improved ability to understand and generate natural language text, in more complex and nuanced scenarios
- Improved performance on simulated bar exam (top 10% vs GPT-3.5 in bottom 10%)
- Similar limitations to prior GPT models



Accuracy of GPT4 for Complex Diagnostic Challenge

- *New England Journal of Medicine* clinico-pathologic conferences from Jan 2021 – Dec 2022 (only diagnostic dilemmas)
- Outcome: model's top diagnosis matched the final case diagnosis (5-point scale)
- Results:
 - GPT4 top diagnosis agreed with the final diagnosis in 39% (27/70) of cases.
 - In 64% of cases (45/70), the model included the final diagnosis in its differential
- Compares favorably with existing differential diagnosis generators
- Generative AI is a promising adjunct to human cognition in diagnosis

Figure. Performance of Generative Pre-trained Transformer 4 (GPT-4)

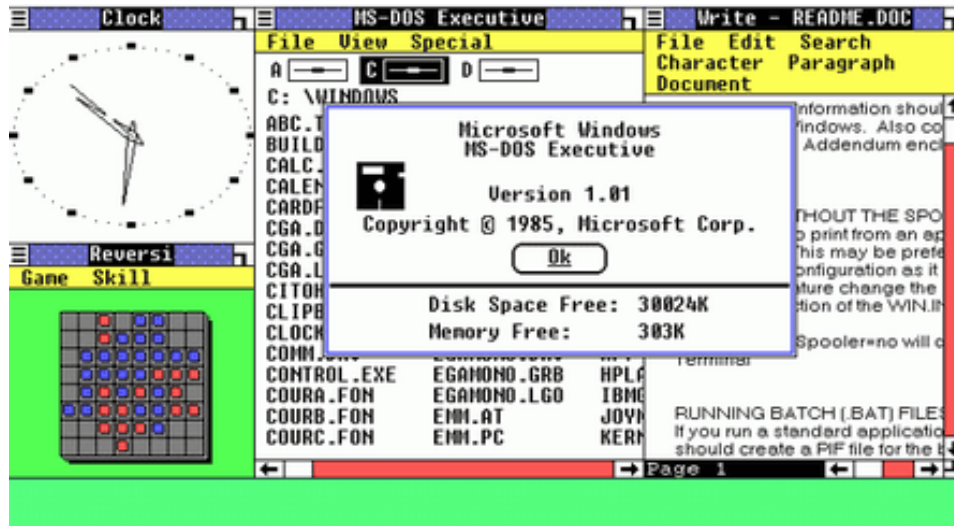


Histogram of GPT-4's performance. Performance scale scores (Bond et al²): 5 = the actual diagnosis was suggested in the differential; 4 = the suggestions included something very close, but not exact; 3 = the suggestions included something closely related that might have been helpful; 2 = the suggestions included something related, but unlikely to be helpful; 0 = no suggestions close to the target diagnosis. (The scale does not contain a score of 1.)

The Age of AI has Begun

“In my lifetime, I’ve seen two demonstrations of technology that struck me as revolutionary.” Bill Gates

Graphical User Interface



ChatGPT

AD What is ChatGPT?



I am ChatGPT, a large language model developed by OpenAI. I have been trained on a massive dataset of text using deep learning techniques, which allows me to understand and generate human-like responses to a wide range of questions and prompts. As a language model, my primary function is to generate natural language responses to user inputs, whether they are in the form of text or speech. I can assist with a variety of tasks, from answering questions and providing information to engaging in conversations and even generating creative writing prompts or stories.



AI is Approaching Human Abilities

Facial Recognition



Google Photos can match infants to children and adults, almost flawlessly

Natural Language

But then, in his haste,
he got into a fight.
He had some emails that he sent
that weren't quite polite.

The SEC said, "Musk,
your tweets are a blight.

GPT-3 generated a Dr. Seuss poem about
Elon Musk upon request

Complex Tasks



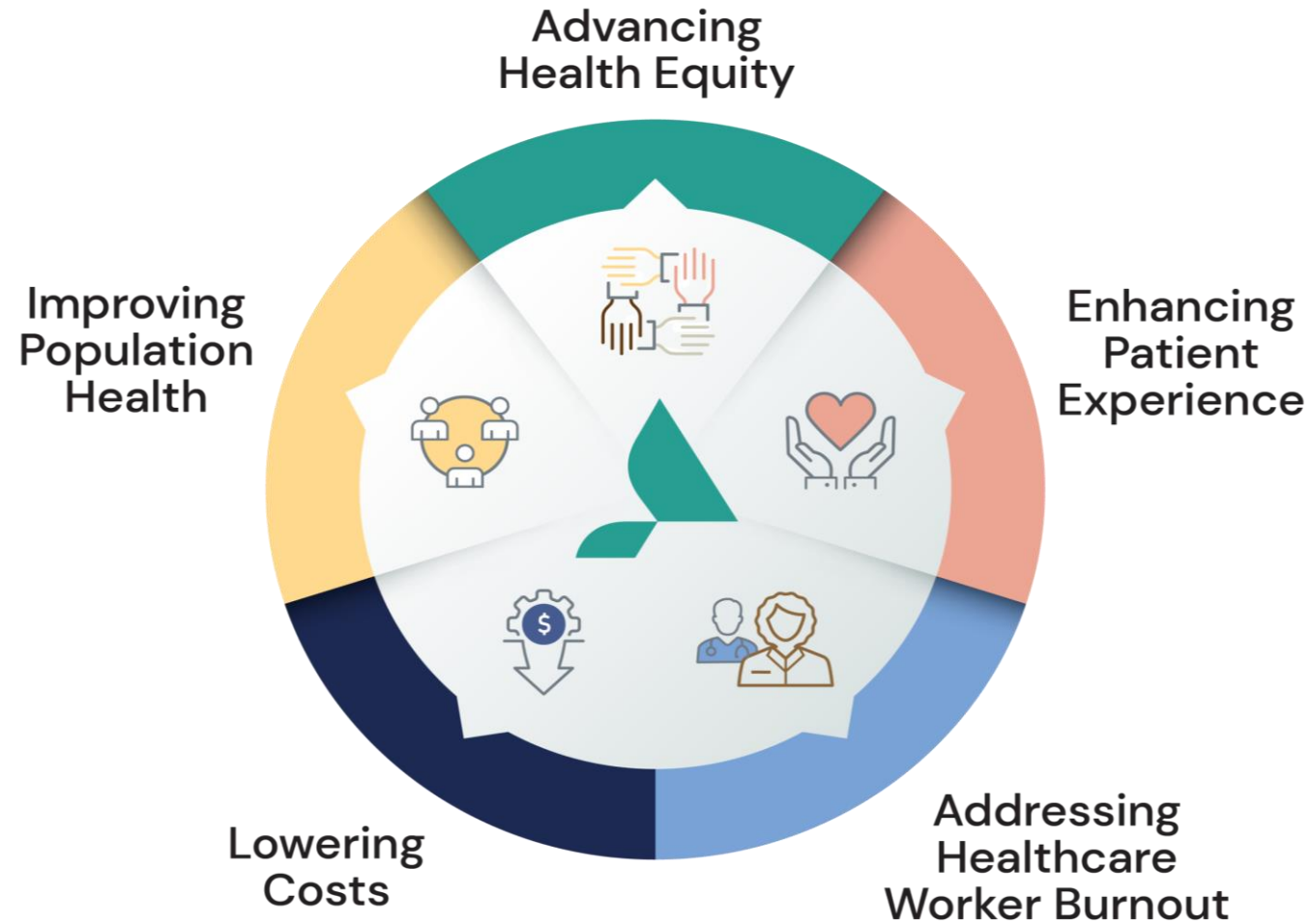
Waymo alone has driven >20M miles
autonomously on public roads



Healthcare AI Use Cases



The Opportunity for Technology in Healthcare



Health Affairs 27, no.3 (2008):759-769

<http://www.ihl.org/Engage/Initiatives/TripleAim/pages/default.aspx>

<https://www.ahah.net/who-we-are/tripleaim.png>

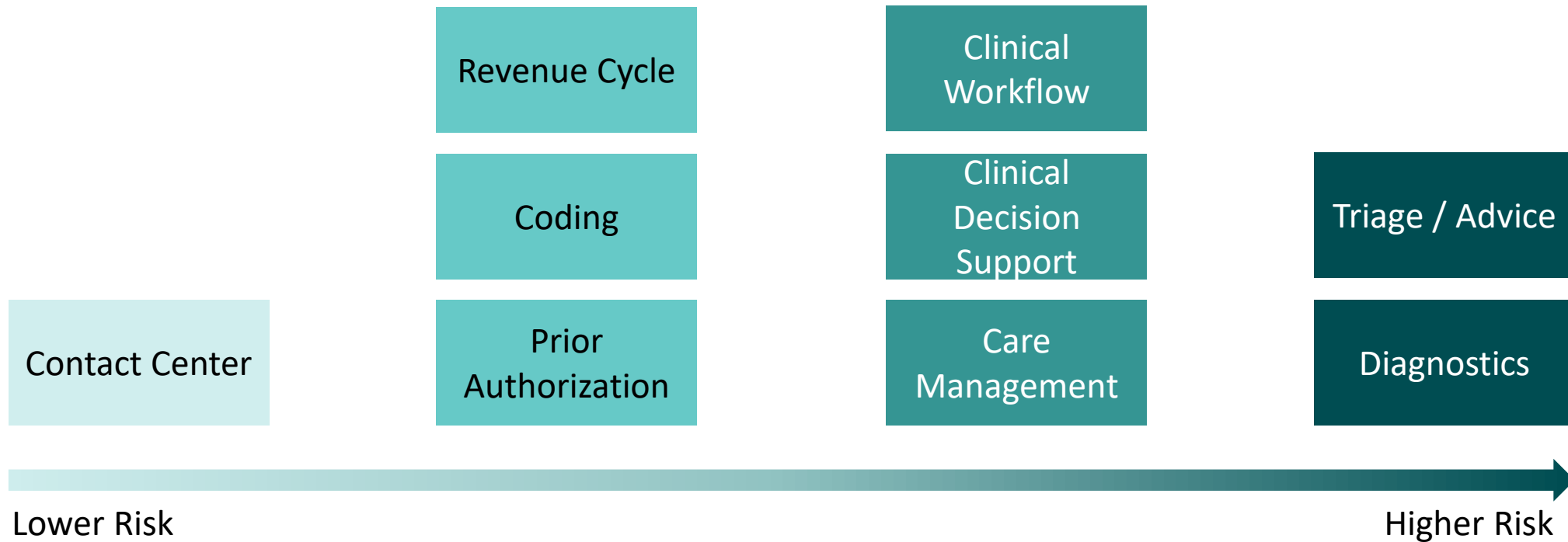
Ann Fam Med. 2014 Nov; 12(6): 573–576.

Photo Credit: Accolade: <https://www.accolade.com/health-equity/>

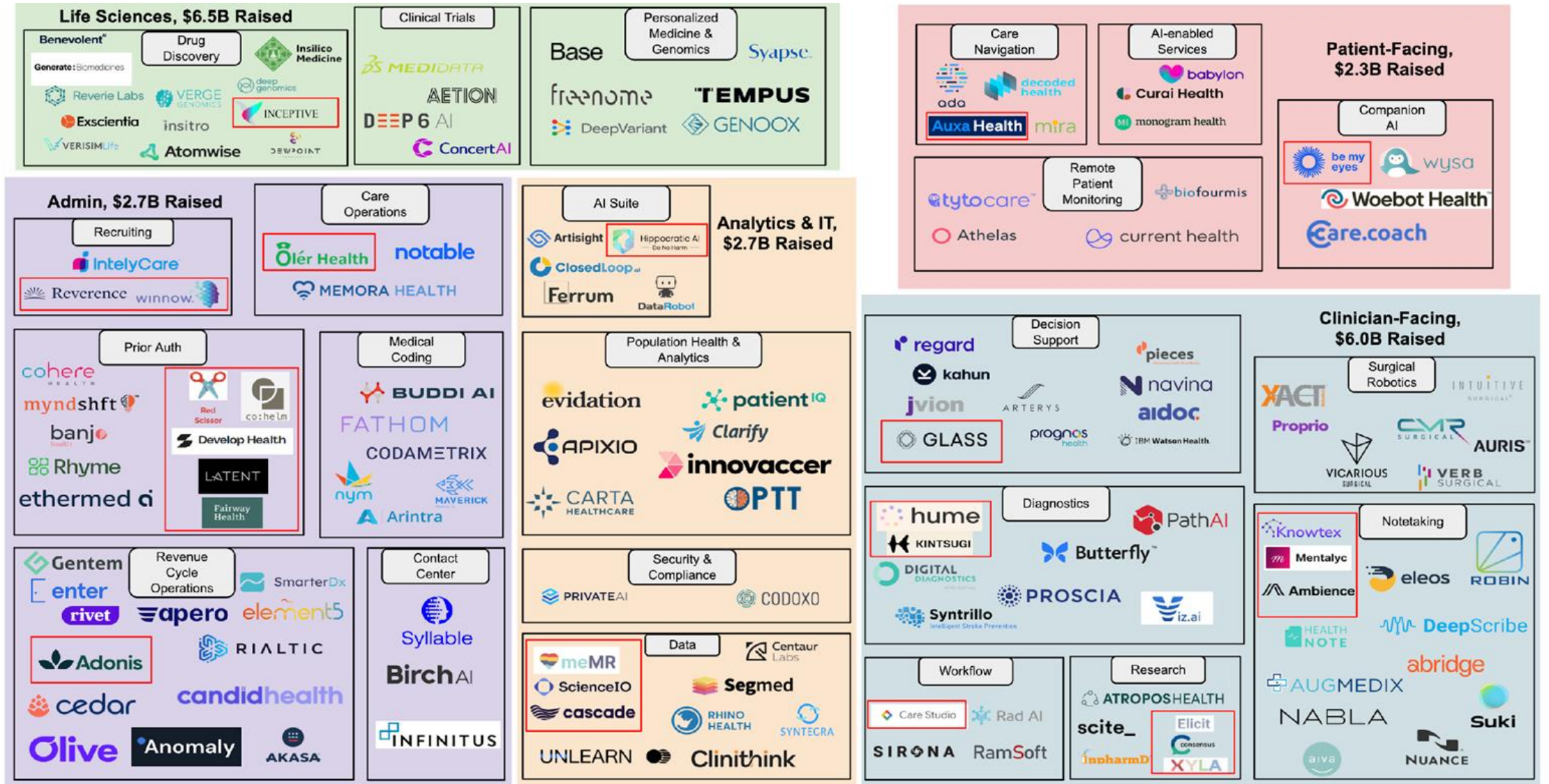
The Opportunity for AI in Healthcare

We are scratching the surface of applying AI to address the Quintuple Aim and MGB Strategy

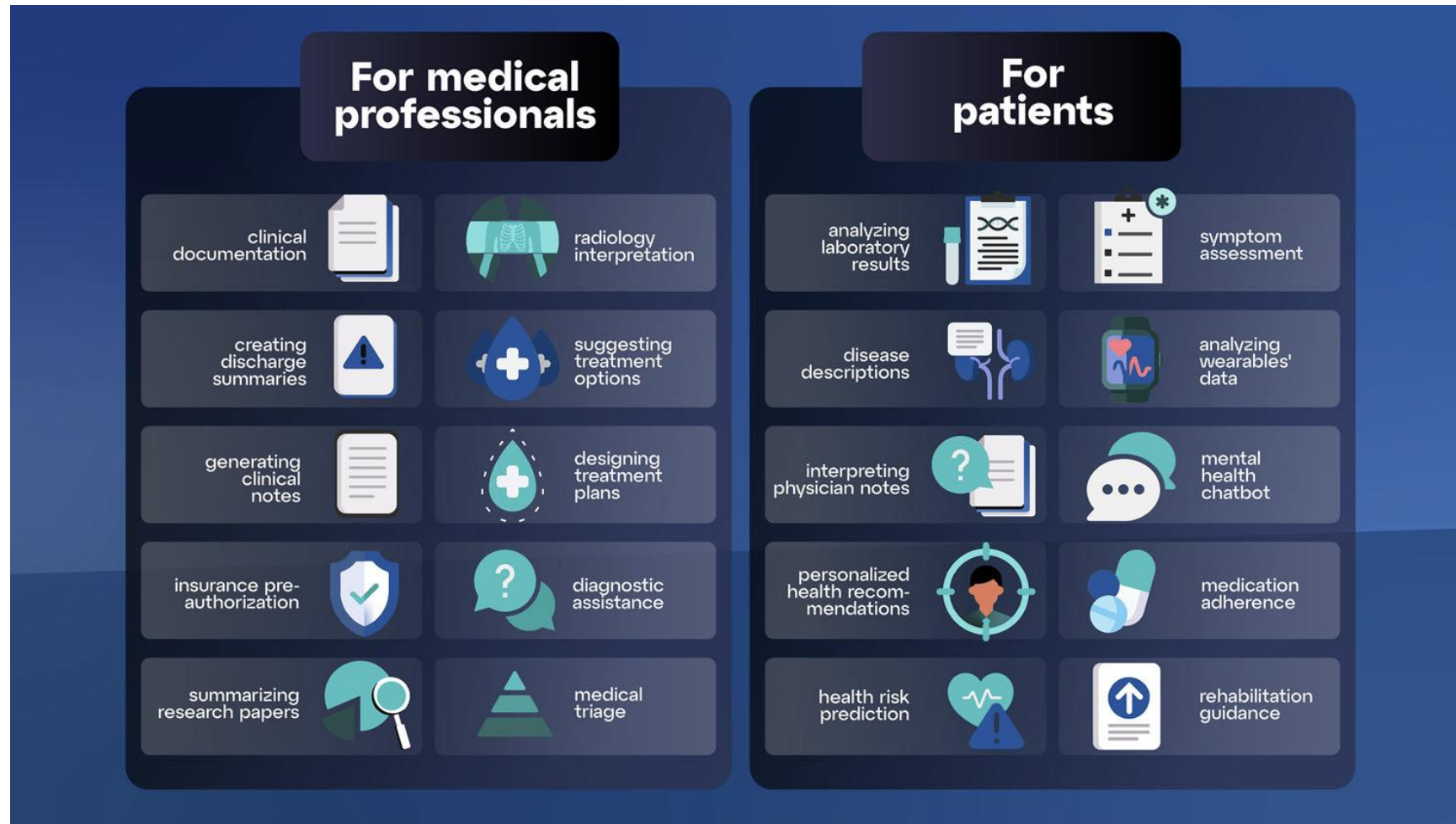
Sample AI Use Cases by Risk



The Healthcare AI Landscape



Opportunities for LLMs in Healthcare



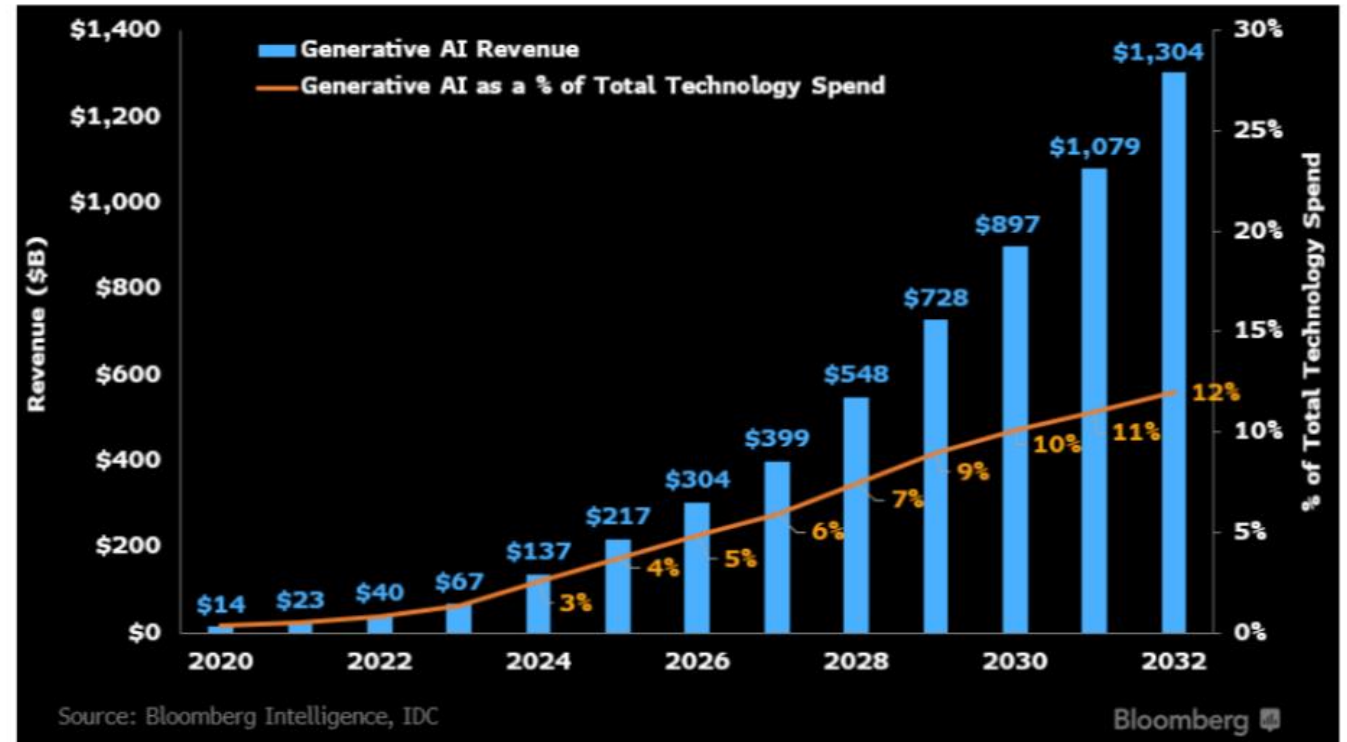
The Opportunity for Generative AI in Healthcare

Market **may** hit \$800-900 billion by 2030

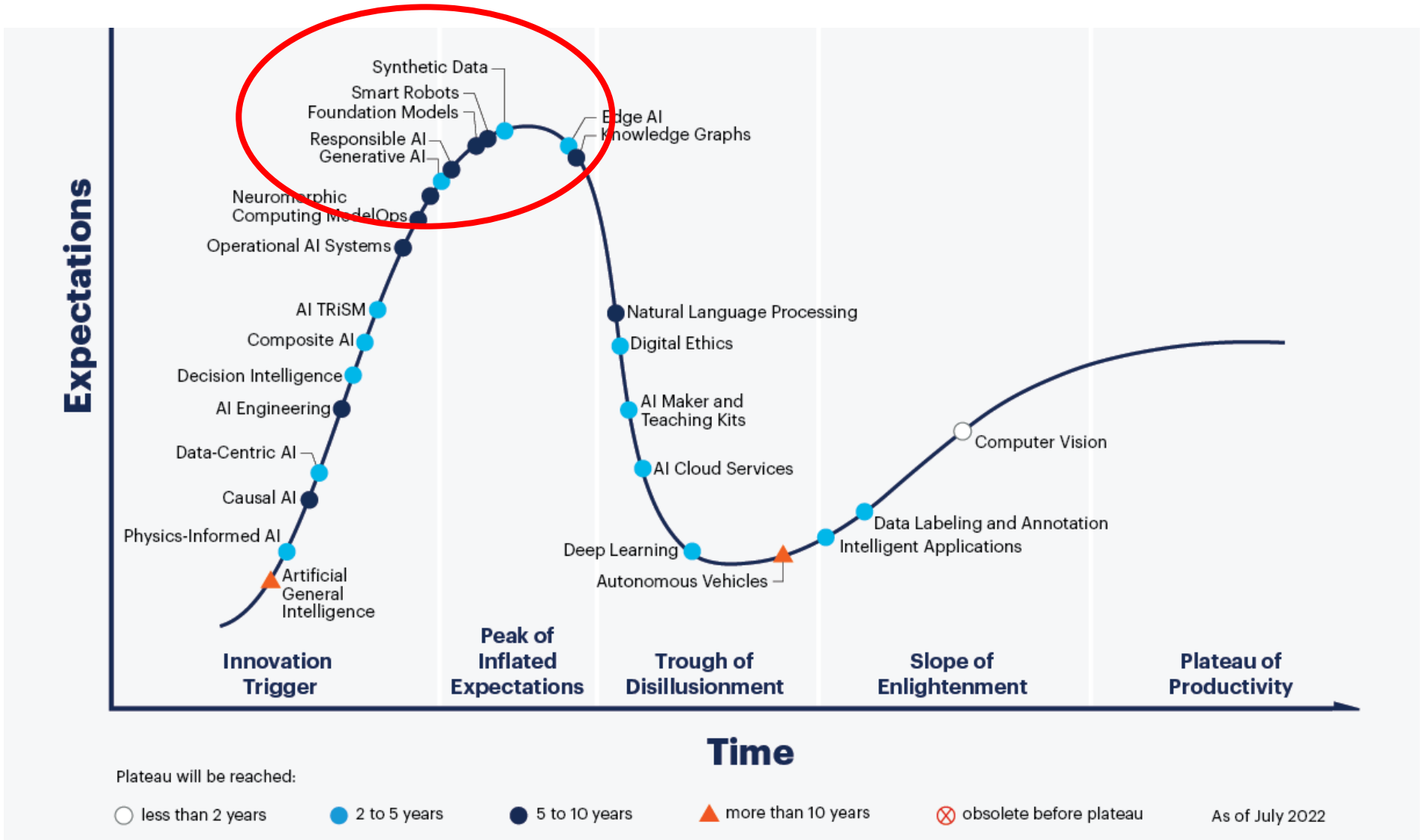
Healthcare use-cases are expected to make up 25% of the generative AI market by 2023

Top projected medical cases:

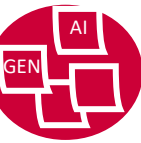
- healthcare analytics
- documentation
- customer interaction
- medical simulation/education
- decision support
- clinical trials
- drug development



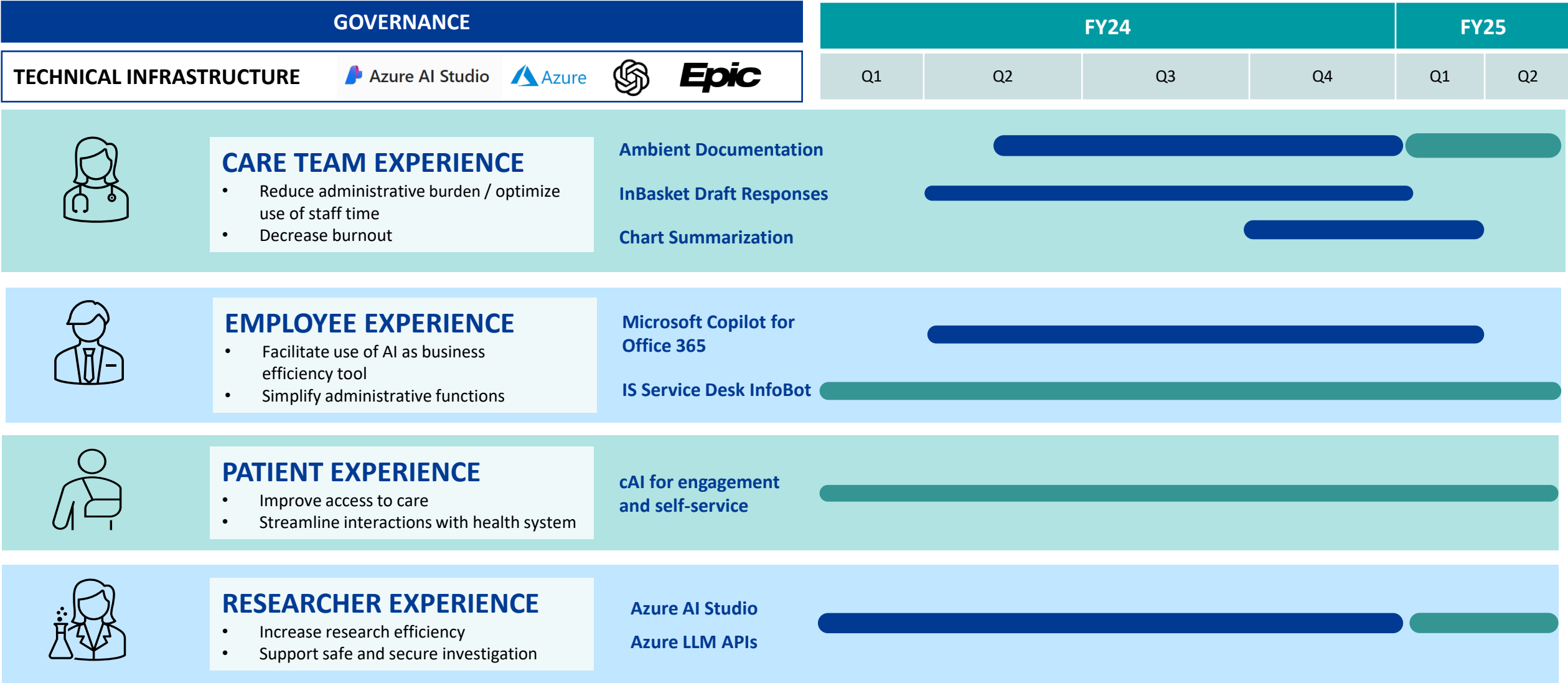
Gen AI is largely at the Gartner Peak of Inflated Expectations




Source: 2022 Hype Cycle of AI, Gartner



High Impact Gen AI Use Cases at MGB



 = Scaling

 = Pilot & Evaluation

Clinical trial informed approach to implementing AI (after model statistical evaluation and responsible use of AI assessment)



Phase I: Safety

Evaluate safety
Design workflows
Engaged stakeholders



Phase II: Efficacy

Refine workflows
Assess impact:

- Quality (incl equity)
- Efficiency
- Financial



Phase III: Effectiveness

Scale
Compare to standards
Design best practice workflows / implementation guides
Monitor safety, workflow, impact



Phase IV: Monitor

Monitor safety, workflow, impact
Disseminate / share outcomes, best practices
Ongoing technology evolution

Combating Provider Burnout

InBasket Optimization
Ambient Documentation

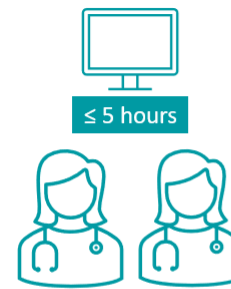


Burnout is one of the biggest challenges in healthcare: highly prevalent and multi-factorial

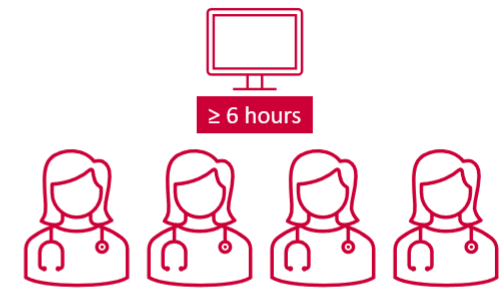
Physician burnout is one of our most significant challenges: 60+% of physicians



Reduced patient access, patient safety risks, and morbidity of individual physicians

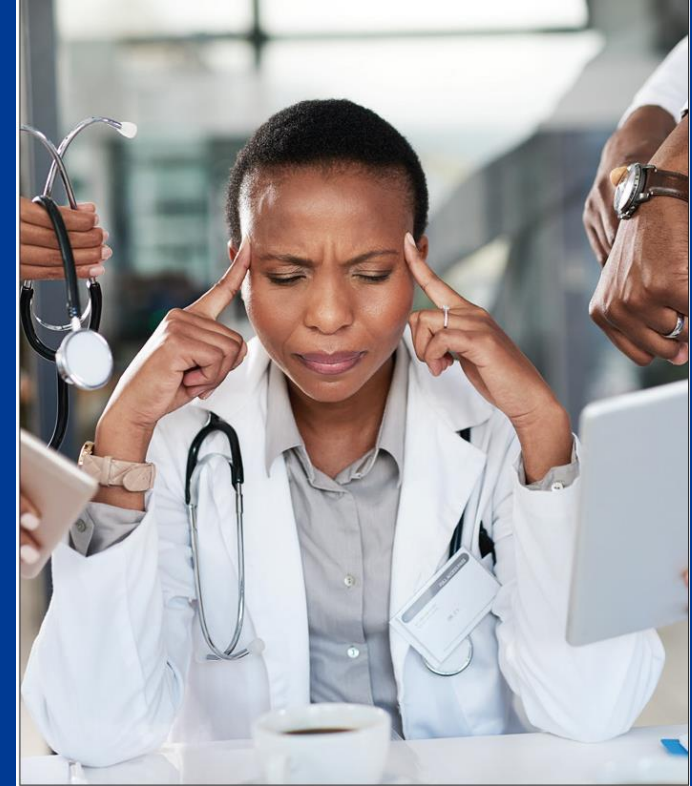


<
Less burnout



Physicians with ≤ 5 hours of after-hours charting per week were 2x as likely to report lower burnout than those with ≥ 6 hours

Can we use
technology to put
humanism back in
healthcare?



Generative AI for In Basket Draft Responses

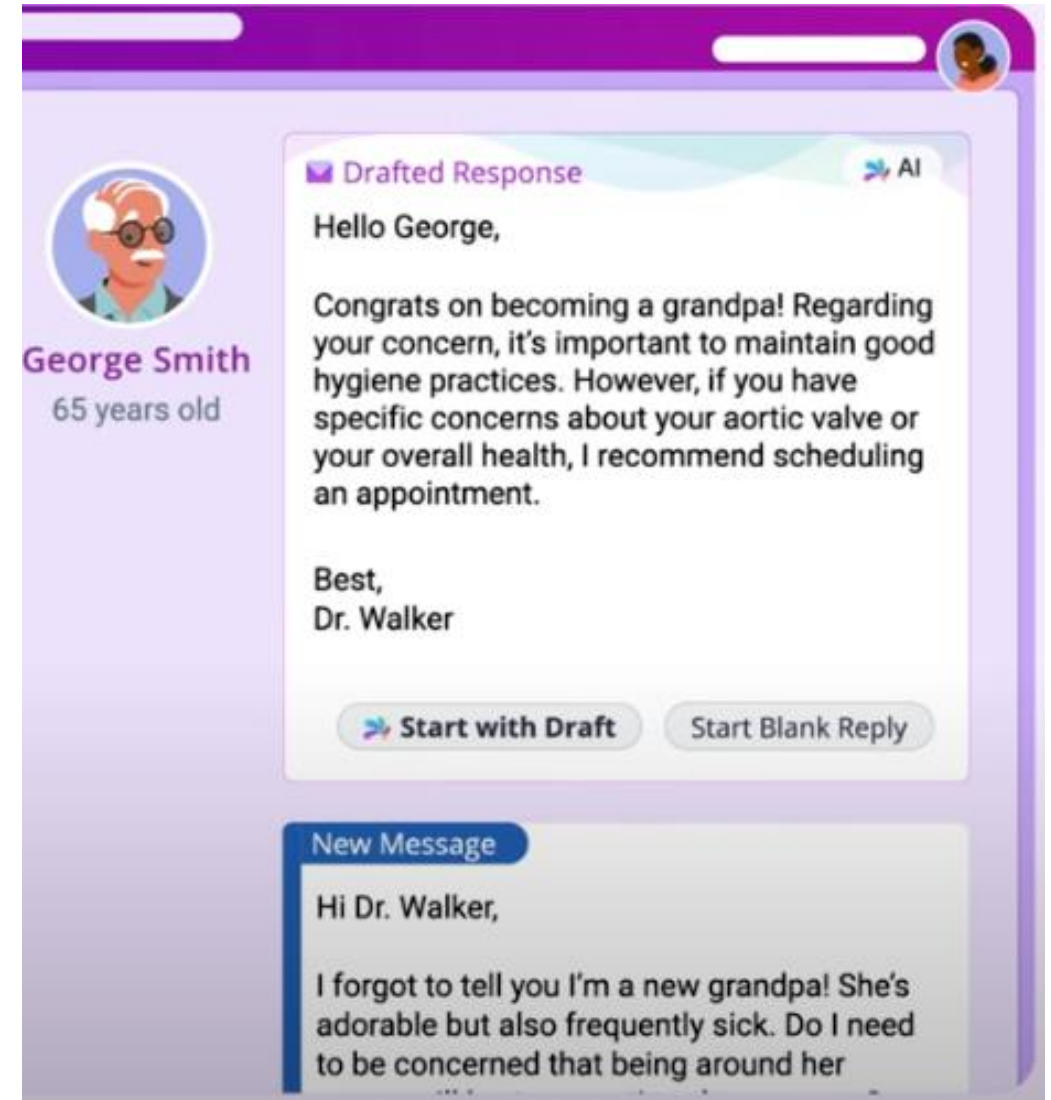
Addressing In Basket Burden

- ~2.4 million patients on MyChart (August 2023)
- ~3.3 million MyChart messages/month

GPT-4 driven In Basket **draft replies** to patients based on chart data.

Specific prompts have been engineered to handle generating draft responses for:

- Medication refills
- Letter requests
- Test result responses
- General messages



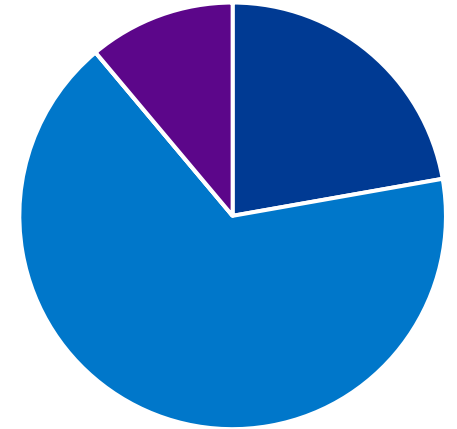
Early Results & Feedback

- Piloted for 3 months with 5 MGB primary care sites
- Initial high utilization and excitement has dwindled over time (currently only used for ~10% of responses)
 - Average draft rating of 4.7/10
 - Consideration of best user- content may be more applicable to nurses and front desk staff
 - ‘Voice of clinician’ not captured; spending more time on editing
 - Reverted to starting with blank draft since does not fit with workflow of message response (cultural & usability)



Have the drafted replies to patient messages impacted your workload?

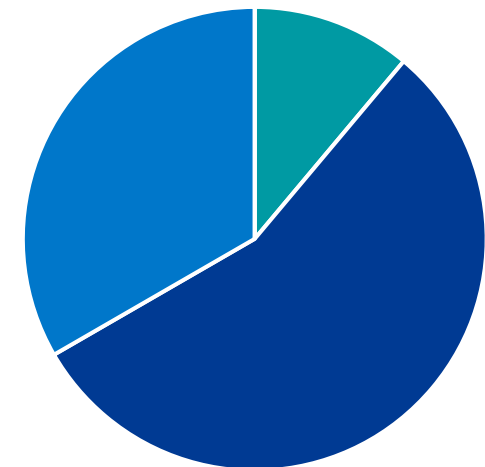
- Significantly decreased workload
- Slightly decreased workload
- No significant impact
- Slight increased workload
- Significantly increased workload



N = 9

Would you recommend the AI drafts to other colleagues?

- Strongly recommend
- Somewhat recommend
- Somewhat do not recommend
- Strongly do not recommend



Artificial Intelligence–Generated Draft Replies

Pilot Results from Stanford

- 5-week, prospective, single-group quality improvement study (pre and post)
- 197 clinicians enrolled in pilot; 162 included in study analysis
- Mean AI-generated draft reply utilization across clinicians was 20%
- No change in reply action time, write time, or read time
- Statistically significant reductions in the 4-item physician task load score and work exhaustion scores

Table 2. Draft Utilization per Clinician Stratified by Specialty and Role

Specialty and role	Mean (SD)			
	Reply action count	Reply action count with draft available	Draft used count	Draft utilization rate
Overall	79.3 (95.5)	59.4 (72.6)	8.6 (16.9)	0.203 (0.268)
Primary care	98.5 (84.4)	74.1 (62.9)	9.3 (11.3)	0.176 (0.212)
Physician and APP	102.0 (75.5)	78.5 (61.0)	9.9 (11.9)	0.153 (0.185)
Nurse	164.8 (215.0)	97.0 (109.0)	5.0 (6.8)	0.111 (0.136)
Clinical pharmacist	29.5 (26.0)	17.4 (15.9)	5.1 (3.8)	0.444 (0.317)
Gastroenterology and hepatology	52.8 (103.9)	39.1 (80.3)	7.6 (22.6)	0.250 (0.342)
Physician and APP	19.3 (33.2)	12.9 (20.6)	1.1 (1.8)	0.240 (0.365)
Nurse	246.5 (156.3)	191.1 (123.5)	45.0 (44.2)	0.293 (0.219)



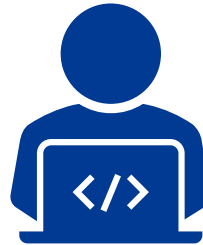
Ambient Documentation



Technology to Improve Provider Documentation Burden



Voice Recognition



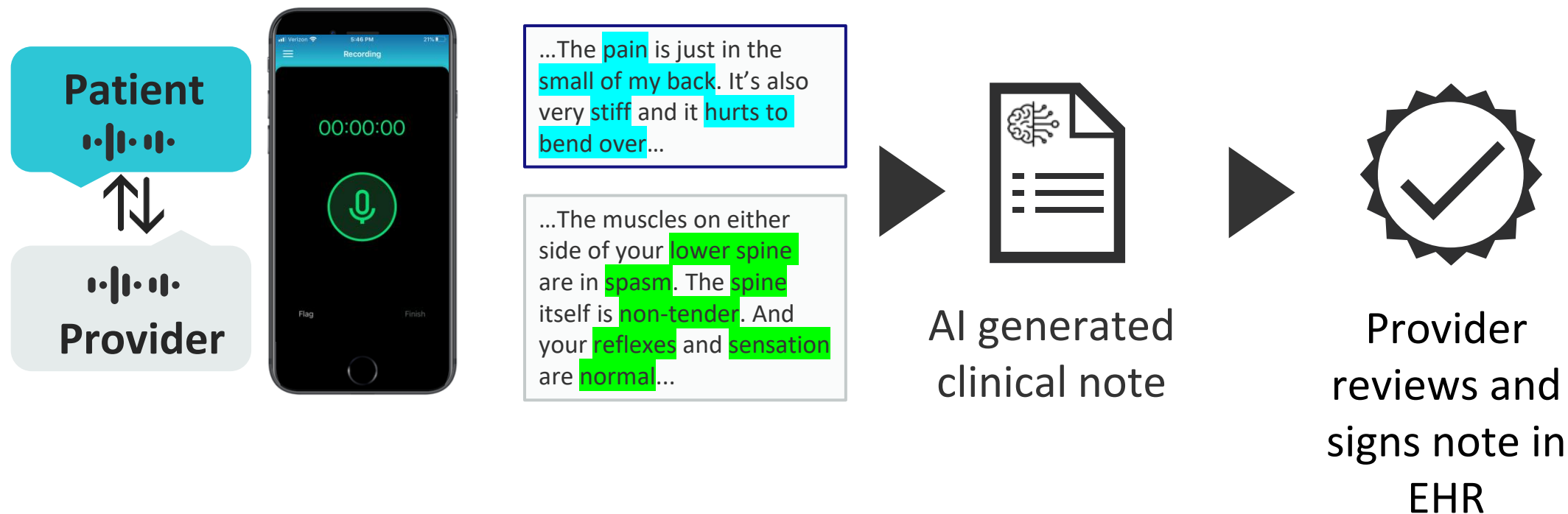
Scribes
(in-person or virtual)



Ambient Documentation

Using Generative AI to Summarize Patient Visits into Clinical Documentation

Ambient documentation (sans scribe review) is rapidly improving due to advanced machine learning, including new generative AI/large language models



Initial Pilot Outcomes

Proof of Concept with 18 Mass General Brigham physicians

- Recorded 400+ sessions
- Average time to draft note ready = 1.5 minutes
- 91% of drafted text is retained by the user

Feedback: Epic integration crucial for workflow efficiency

Pilot of sufficient scale needed to assess for satisfaction, efficiency, ROI

History of Present Illness

- The patient is a 53 year old male with a past medical history of atrial fibrillation, diabetes, hypertension. The patient presented with shortness of breath and chest tightness, also associated with dizziness. The symptoms occur during strenuous activity like walking up stairs or going on a hike, and with lifting heavy objects. He first noticed the symptoms about a week ago, with symptoms now occurring on a near daily basis.
- They denied palpitations, swelling in the ankles, lightheadedness, or trouble lying flat. They also denied fevers, chills.

> Add a new line

Assessment & Plan

- Shortness of breath and chest tightness: Patient reported shortness of breath and tightness in chest during strenuous activity like walking up stairs or going on a hike, and with lifting heavy objects. No swelling in ankles, no nausea or vomiting, and some dizziness reported. Plan to perform EKG and echocardiogram this afternoon to assess heart function.
- Hypertension: Patient will start taking lisinopril 10mg a day to protect heart and lower blood pressure.
- Diabetes: Last hemoglobin A1c elevated at 8. Patient has not been able to afford metformin. Social worker consulted to help with medication supplies for metformin 500mg twice daily.

> Add a new line

Transcript

Hello, Patrick. What's brought you in today?

So, I'm having some issues with breathing and chest tightness. It seems to happen mostly when I do things like walk upstairs or go for a hike, and when I lift heavy things. Like, I was helping my partner with gardening a few days ago and after lifting the bags of soil I was just so short of breath.

I see. Can you tell me when this first started?

Well, I first noticed it a week ago, and it's been happening almost every day since.

Got it. Do you have any palpitations, swelling in the ankles?

No, not that I've noticed.

How about lightheadedness? Any trouble lying flat on your back?

No, nothing like that. Well, I do get a little dizzy, I guess.

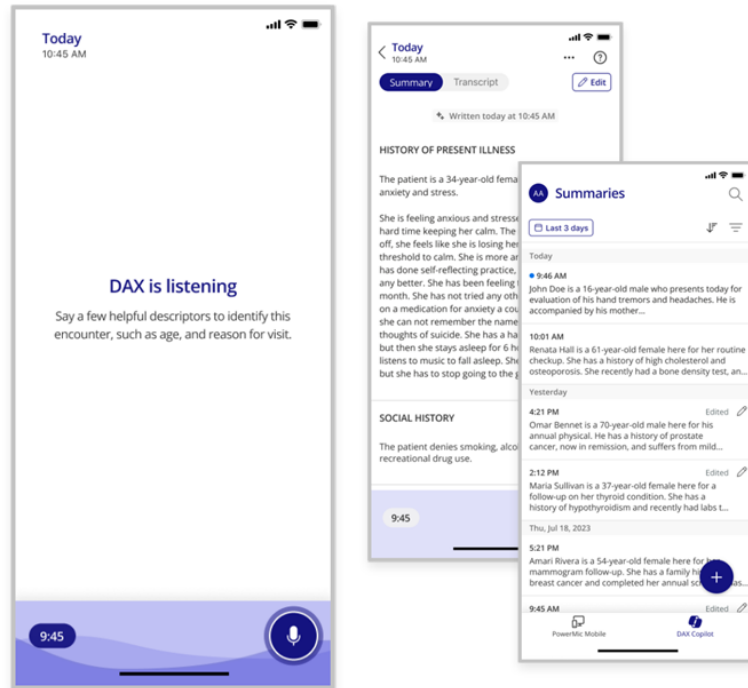
Alright. Well, given your past medical history I think it's probably wise to check things out thoroughly. Let's get you an EKG and an echocardiogram to get a better idea of how your heart is functioning.

0:03:32 0:16:24 1x Complete & Send

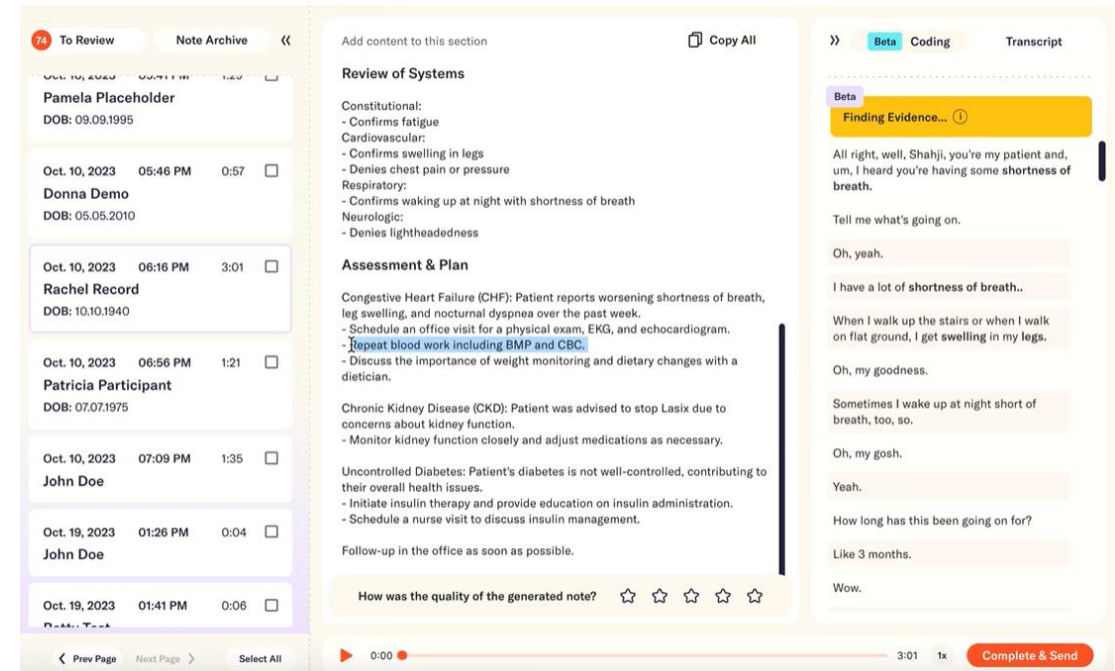
"When using Abridge during a visit, I have noticed that I am able to have more face-to-face time with my patients, rather than staring at my computer screen"
- Physician POC participant



Pilot: Two platforms across ~800 Physicians and APPs



iOS Haiku integrated
Ambulatory visits only



iOS or Android not integrated (planned June 2024)
Any care setting
Spanish concordant visits



Demo



Ambient Clinical Documentation

WORKLIST

OCT 17, 2024

10:46 AM

5:54

Mit course example

ABRIDGE



Mit course example

CLINICAL NOTE PATIENT SUMMARY (PVS)

History of Present Illness

The patient, with a history of recurrent knee injuries, presents with right knee pain that began a few weeks ago after a run. They deny any specific trauma or fall, but noticed the pain upon completion of their run. The pain has persisted since then, and is particularly noticeable when running or walking at certain angles. They have been managing the pain with Advil and previous physical therapy exercises, but report no improvement. The patient has a history of annual knee injuries, each affecting a different part of the knee. They have had multiple x-rays, but no MRI or CT scans. They have seen an orthopedist annually, and have been recommended cortisone injections, but have not received any. After each injury, they have undergone physical therapy, which typically resolves the issue after a few months. They deny any other medical issues or medications, and work a sedentary job in consulting.

Physical Exam

VITALS: T- 98.3, P- 68, BP- 118/68, RR- 20, SaO2- 98%
MUSCULOSKELETAL: Pain upon palpation in the middle of the knee, directly on the bone. Mild pain with knee flexion. Normal gait observed. Anterior draw test negative. Collateral ligaments intact. Profusion and pulses 2+ in the extremity.

Results

Click to add content to this section

Assessment & Plan

Right Knee Pain
Pain for several weeks after running, no trauma. Pain on extension and palpation of the knee. History of recurrent knee injuries. No prior imaging beyond X-rays.
-Order MRI of the right knee to evaluate for possible strain or partial tear.
-Refer to orthopedist for further evaluation.

TRANSCRIPT

Um, what brings you in today?

Uh, my knees hurt.

Your knees hurt?

Yeah.

Um, both of your knees?

Uh, well, my **right knee** hurts.

Your **right knee**?

Yeah.

THEY/THEM

COPY ALL

MARK AS DONE



0:00

5:54

1x

Ambient Pilot Volunteers Span Sites and Departments weighted to medicine (60%) and AMCs (60%)

	Medicine	Med Specialties	Surgical/ Anesth	Psych	Pedi Specialty	ED/ Urgent	Rad Onc	Pedi Primary	Neuro	Home Hospital	Total
BWH	36	49	39	7		16	11	1	1		160 (20%)
Comm. Hospitals	91	80	33	9	4	10		1	4	1	233 (29%)
MEE		8									8
MGB MG	41	12	5			14		1	3		76
MGH	64	95	54	15	12	49	1	7	25		322 (40%)
SRN		10		1							11
McLean		1		4							5
Grand Total	232 (28%)	255 (31%)	131 (16%)	36	16	89 (11%)	12	10	33	1	815



(Ambient) is **truly life changing** for me. I conduct over **50% of my visits in Spanish** and documentation has always been a struggle for me. There is a **huge weight off** my shoulders at the end of the day. - *Pediatrician*

I'm **realizing how much I'm doing** for my patients." - *Neurologist*

I **look at the patient's eyes and that is magical.** I'm addicted. I want it to be a part of every encounter I have. - *Neurologist*

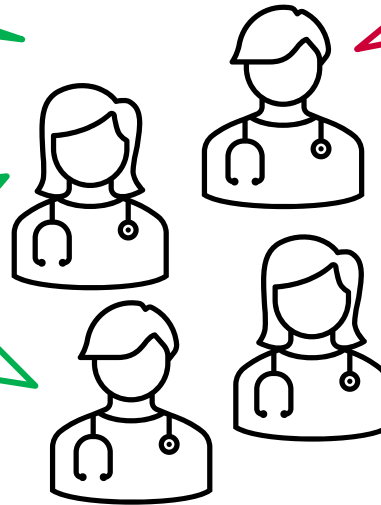
In short this is the **greatest change to my practice of medicine in the 26 years** I have been working. I can **never go back** and again **am enjoying my job** rather than dreading being behind on notes. - *Urologist*

I am in love with (ambient). Already I **am much less likely to change careers** in the next year.
- *Primary Care (after 2 days of use)*

I'm still typing because **what if there's a bug** that it doesn't work. Then I've looked at the (ambient) note and it's perfect, so I delete everything I typed. - *Primary Care*

I spent **a lot of time explaining to my patients** why they weren't having an emergency...and very little of that made it into my note. **I would like to see more of that in my note.** - *Emergency Medicine*

(Ambient) is a **complete game changer.** The first time in 5 years can I **actually finish my notes on the day of clinic and it allows me to focus on the patient conversation** rather than on documentation. - *Orthopedic Surgeon*



Generative AI Pilot Assessment Approach



Validity / safety

- User survey, focus groups
- Epic data on clinician marking draft as useful or not
- Epic data on usage of Draft Reply without further editing



Efficiency and satisfaction

- User EHR efficiency: pajama time, notes completed same day, total documentation time, total EHR time, IB message latency time, total IB time
- Satisfaction survey

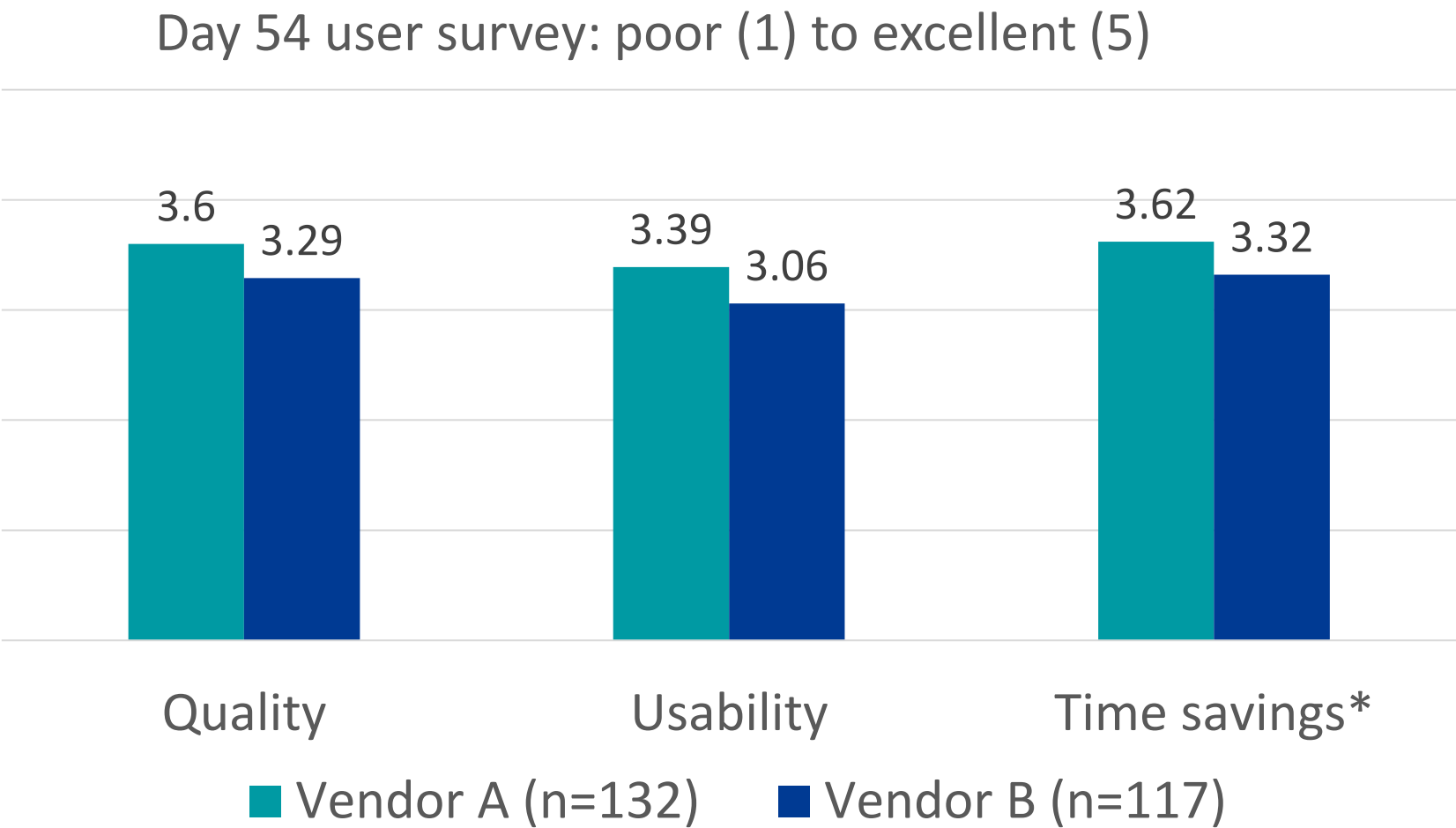
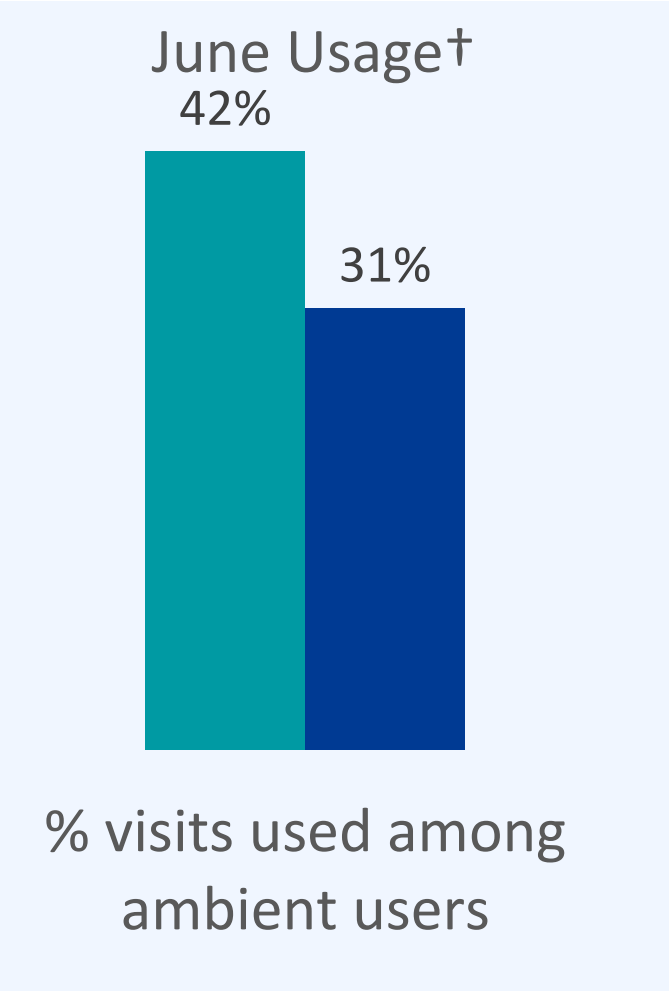


ROI: clinician well-being, revenue

- User survey
- Pajama time
- Visit volume
- CPT billing
- Risk adjustment
- Future: Patient Experience

If first gate of validity and safety are passed, we will seek to expand the pilot to 50-100 users to assess for efficiency, satisfaction, and ROI

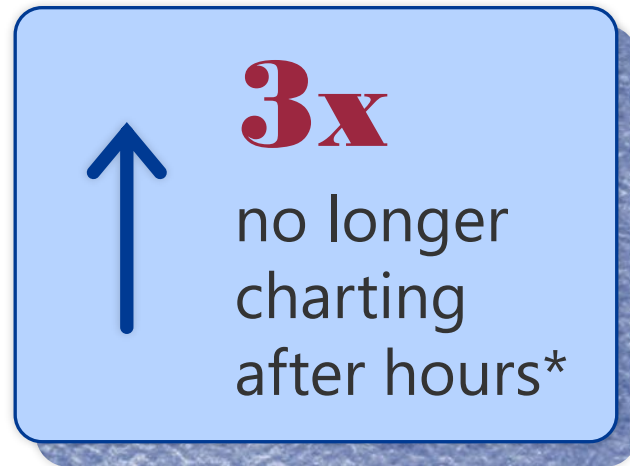
Initial feedback: significant usage and positive experience



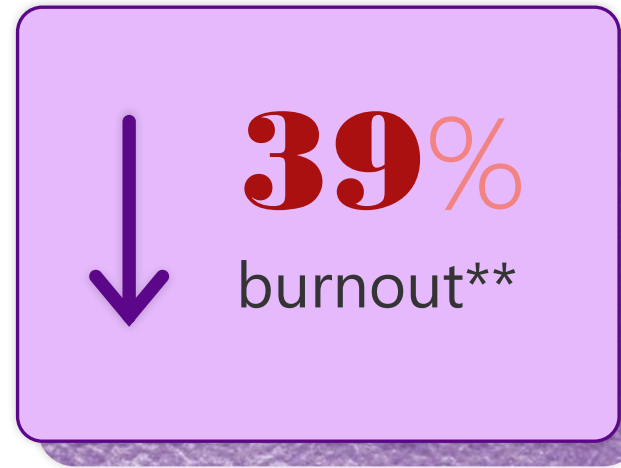
†% of ambulatory visits among ambient users where ambient documentation was used; based on vendor usage data for providers with denominator available (Vendor A n=86, Vendor B n=85)

*Objective data available September 2024

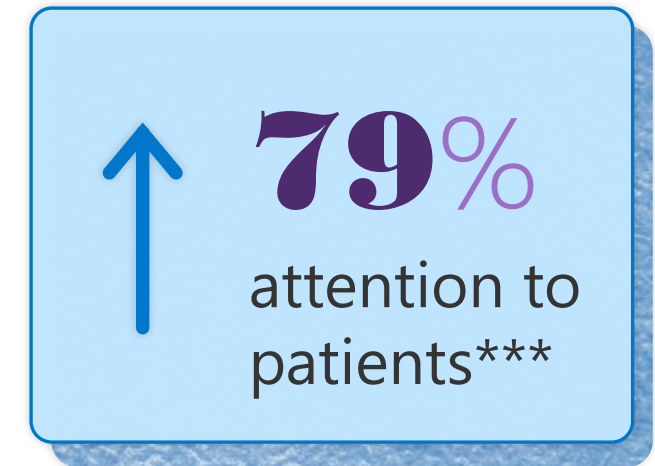
Ambient documentation has significant impact on documentation burden, rapport, and burnout



Usage: 38% of
Ambulatory visits†



Intent to stay: 62%
report more likely to
extend length of career



Note quality: 3.5/5

**Reported by providers who completed both pre-pilot and post-pilot surveys (n=124)*

***% reporting burnout by PFI cut-off at 6 weeks (n = 223), p < 0.0001*

****Reported by providers in the 12-week post-pilot survey (n=152)*

†During June, for active providers who have used Ambient for at least 5 encounters



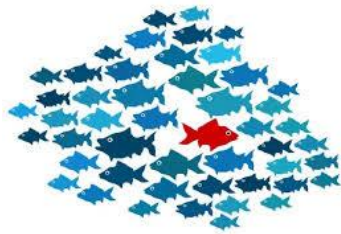
Ambient Documentation for Nursing

Flowsheet Row (Structured Documentation) Needs to be Addressed



Fisher, Amy 74 F 04/23/1945 ADM IN 95 954-1											
Full Code 1.55m 150lb 6.349oz BSA: 1.74m² BMI: 28.4kg/m² Allergy/Adv: calcium carbonate											
EB0000032984 IA00006606 100007280											
Launch Oblix Monitor											
Rate Precision											
Close Save											
Acknowledge Orders											
Next Task											
Bedside Bulletin											
Items	6 Hr	Tue Mar 3 03:00	Tue Mar 3 04:00	Tue Mar 3 05:00	Tue Mar 3 06:00	Tue Mar 3 07:00	Tue Mar 3 08:00	Tue Mar 3 09:00	Tue Mar 3 10:00	Tue Mar 3 11:00	Tue Mar 3 12:00
Waveform											
Vital Signs (Complete) (1h) Never Captured BID											
Temperature											
Pulse											
Pulse Rate (60-90 b...)	95 H	95 H	92 H	90	85	96 H	95 H	88	93 H	95 H	
Location											
Rhythm											
Respirations											
Respiratory Rate (12...)	25 H	20	21	25 H	26 H	27 H	25 H	24	23	25 H	
Pulse Oximetry (95...)											
Oxygen Delivery ...											
Flow Rate (L/min)											
Blood Pressure											
Mechanical Ventilator (1h) Never Captured											
Vent Settings											
Ventilator Mode	VC-CMV	VC-C...	VC-C...	VC-C...	VC-C...	VC-C...	VC-C...	VC-C...	VC-C...	VC-C...	
Respiratory Rate Set...	14	14	14	14	14	14	14	14	14	14	
Tidal Volume Setting...	600	600	600	600	600	600	600	600	600	600	
FiO2 Setting (%)	50	50	50	50	50	50	50	50	50	50	
PEEP Setting (cm H2...)	5	5	5	5	5	5	5	5	5	5	
CPAP Setting (cm H2...)											
Pressure Support Se...											
Measured Patient Data											
Calculated Patient Data											
Static Lung Compla...											
Dynamic Lung Comp...	62.22	62.22	62.22	62.22	62.22	62.22	62.22	62.22	62.22	62.22	
Vent Checks											
Ventilator Checks											
Intake & Output (C)											
Infusions/ Titrations											
Propofol 100 ml @ 5 MCG/KG/MIN 2.047 mls/hr IV TITR PRN											
Current Dose: 40 mcg/kg/min, 16.373 mls/hr											
000026658 M(P) (11) com											
Assessment											
Patient Descr...											

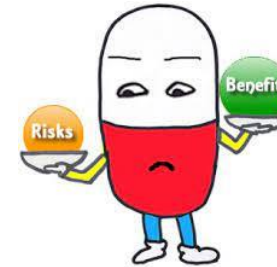
What is Needed to Advance Healthcare AI?



Reliability



Safety



Benefit/ROI



FDA Approval



Bias



Privacy/Security



Transparency



Informed Consent

Discussion



Questions for the Group

Ethical Guidelines and Governance Frameworks

- How are your organizations governing AI?
- What additional frameworks or tools are needed to advance AI in healthcare?

Potential Impact on Healthcare Systems and Society

- What do you think the impact of AI will be on Health care and Health in the short and long-term?

Future Directions for AI in Healthcare

- We all talked about current applications of AI in healthcare, what will future applications of AI in health look like?





Mass General Brigham