



AI in Family Medicine

What works now, what to avoid, and what's coming next



1

A Practical Guide to Using AI Safely and Effectively in Primary Care

You do not need to become an AI expert.

You need to know where AI helps, where it fails, and how to verify it.

TIME

Documentation & admin

THINK

Evidence & reasoning

CONNECT

Patient communication

Izhar Hasan, MD, PhD

Medical Director; Assistant Professor, Dept. of Family Medicine, Hackensack Meridian Health, NJ

Project Lead, STFM Recognition for Clinical AI Alliance

izhar.hasan@hmn.org • 908-328-4848

Connecticut Academy of Family Physicians · 2026 Symposium · Oct 1, 2026

Speaker disclosure and educational boundaries

1

Financial relationships

Izhar Hasan, MD, PhD reports no relevant financial relationships with ineligible companies related to the content of this activity. Dr. Hasan conducts online EBM-related journal club activities.

2

Commercial support

None, unless separately disclosed by the accredited provider.

3

Educational use of AI

AI-generated examples and fictional patient cases may be used for educational demonstration. Product capabilities change over time, and patient-specific AI output requires independent clinical verification.

4

Product neutrality

Specific AI platforms are shown for educational comparison and do not constitute endorsement of any commercial product.

TRANSPARENT

VENDOR-NEUTRAL

CME-READY

WHY NOW

AI is already in primary care. Training has not caught up.

Start with the room — then compare your practice with national primary-care data.

LIVE POLL

Which type of AI have you used in the last 30 days?

- ☐ Clinical evidence / search AI
- ☐ General AI assistant
- ☐ Ambient documentation / EHR AI
- ☐ Practice / administrative AI
- ☐ None yet

Follow-up: Which use case has actually changed your workflow — and why?

AAFP + ROCK HEALTH • 2025

Primary-care clinicians want practical support

81%

want more AI training

70%

believe AI may improve clinician well-being

50%

have used AI for work

32%

used AI for clerical support

Today's promise: practical enough to use tomorrow, disciplined enough to defend.

At the end of this session, participants should be able to...

1

Apply AI to appropriate family-medicine workflows

Select workflows where AI can reduce burden without replacing clinical judgment.

2

Construct clinically bounded prompts

Specify context, task, constraints, output format, and what the model must not assume.

3

Verify clinical information and evidence

Check patient facts, source provenance, recency, claim-to-citation match, and uncertainty.

4

Recognize and mitigate AI-related risk

Recognize hallucination, algorithmic bias, automation bias, and PHI/privacy exposure.

5

Match oversight to consequence, and evaluate AI tools

Increase oversight as error consequence rises; evaluate tools on purpose, evidence, workflow fit, privacy, and impact.

Predictive vs Generative AI: Different Math → Different Errors → Different Verification

Same word “AI,” but two different probability problems — and two different clinician verification strategies.

DIFFERENT FAILURE
MODES

P Predictive AI

ESTIMATES RISK

Estimates a probability or classification from patient features.

Patient features

Age • BP
Labs •
History

MODEL

18%
risk

Core probability question

P(outcome | patient features)

Example model: logistic regression

HOW IT CAN FAIL

- Poor calibration
- Dataset / population shift
- Biased proxies or missing groups

CLINICIAN ACTION

Validate the model

Population match

Calibration

Data / input drift

G Generative AI / LLM

GENERATES TEXT

Generates language from context and the sequence that came before it.

Context

Prompt +
patient context

LLM

Generated
answer

Core probability question

P(next token | context + prior tokens)

The model is optimized to sound plausible, not to be correct.

HOW IT CAN FAIL

- Fabricated facts or citations
- Confident answers despite uncertainty
- Missing context / retrieval gaps

CLINICIAN ACTION

Verify the output

Facts • Citations • Context

Clinical takeaway: Validate predictive models; verify generative outputs. Increase verification as clinical stakes rise.

Delegate tasks — not accountability

COPILLOT

Good uses

- Summarize and organize information
- Draft routine content
- Surface gaps and possibilities
- Reduce repetitive clerical work
- Challenge your reasoning

AUTOPILOT

Keep final responsibility for

- Diagnosis and treatment decisions
- Patient-specific recommendations
- Medication changes
- High-stakes patient communication
- Anything you would not personally defend

**The physician remains the control layer —
even when AI does more of the work.**

The next AI wave is already moving inside the EHR

Before adding another standalone tool, ask what is already enabled — and what is arriving — inside your clinical workflow.

ASK THE ROOM

What AI is already live in your EHR today?

Ambient notes • chart summary/search • message drafting • coding • patient portal AI • workflow agents

EPIC

NOW

- AI Charting / ambient documentation
- Ask Emmie: patient-facing MyChart assistant
- Built-in AI for chart navigation + workflow

DIRECTION

Expanding agentic workflows via Agent Factory; deployment varies by health system.

ORACLE HEALTH (Cerner lineage)

NOW

- Clinical AI Agent: notes + chart summaries
- Voice-first, conversational EHR experience
- Coding, scheduling + workflow automation

DIRECTION

Announced direction includes proactive discharge readiness, prior authorization and patient navigation.

ATHENAHEALTH

NOW

- Ambient Notes embedded in athenaOne
- athenaAmbient native ambient experience
- AI-native Clinical Encounter: contextual insights

DIRECTION

2026 direction: chart review, documentation, decision support and care coordination inside the encounter.

Clinical pearl: embedded AI brings chart context + workflow permissions — so local governance matters as much as model capability.

Different jobs need different AI. Check the boundary.

Capabilities change quickly. Check evidence sources, PHI terms, local approval, and workflow fit before use.

EHR-EMBEDDED

Chart-context + local workflow AI

BEST FOR

chart context + workflow

STRENGTH

embedded identity/data controls

WATCH

local configuration + validation

CLINICAL EVIDENCE / REASONING

Medical evidence / clinical search AI

BEST FOR

clinical Q&A + evidence retrieval

STRENGTH

purpose-built medical sources

WATCH

coverage, recency + link integrity

GENERAL / ENTERPRISE AI

General-purpose / enterprise AI

BEST FOR

synthesis, drafting + analysis

STRENGTH

broad capabilities + flexible workflows

WATCH

not every mode is a clinical evidence engine

Current vendor examples and links are in the appendix. Access, features, evidence sources, and privacy terms can change.

USE CASE 1

7:30 AM — Turn chart review into a 60-second pre-visit brief

Let AI retrieve and organize the longitudinal story — not invent the assessment.

BEFORE

58-year-old with diabetes, hypertension, obesity, CKD, 14 medications, 9 specialist notes, 23 labs, and 2 ED visits.

- Open chart
- Scroll through notes
- Reconstruct the story
- Hope nothing important is missed

PROMPT PEARL

“Summarize the last 12 months. For every key change or unresolved issue, give the source and date. Highlight trends, medication changes, preventive gaps, and missing data. Do not infer facts that are not documented.”

A1c

7.2 → 8.6

eGFR

66 → 52

BP

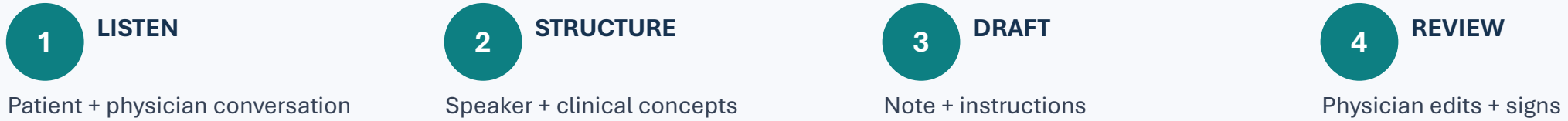
>140/90

GAP

retinal exam

Best use: retrieval + organization. Require source/date for every key claim and flag missing data.

9:00 AM — Ambient AI can give attention back to the visit



EARLY OUTPATIENT STUDIES

20.4%

less time in notes per appointment

30%

less after-hours documentation

51.9 → 38.8%

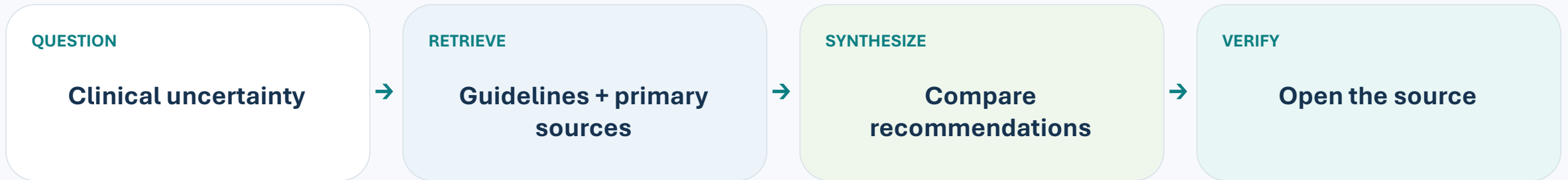
burnout prevalence after 30 days

Promising, but the evidence is still largely pre/post quality-improvement data. Treat every generated note as a draft until the clinician reviews and signs it.

USE CASE 3

11:15 AM — Ask for evidence you can inspect, not an answer you can admire

A citation matters only if you can open it, confirm what it is, and decide whether it applies to this patient.



PROMPT TO FORCE EVIDENCE TRANSPARENCY

“Use current guidelines and high-quality primary evidence. For each key recommendation, give the source, year, population, direct link, and evidence type. Show meaningful conflicts. If you cannot verify a citation or link, say so.”

Never trust a citation because it looks authoritative.

One clinical question. Five ways to test an AI evidence answer.

Keep the case simple so citation quality — not diagnostic complexity — drives the comparison.

FICTIONAL FAMILY MEDICINE CASE

52-year-old woman • annual physical

No prior hypertension diagnosis

Today: 148/92 → repeat 146/90 after 5 min

2 months ago: 144/88

No CVD, diabetes, CKD, pregnancy, or hypertensive-emergency symptoms

No antihypertensive medication

QUESTION: What should the family physician do next to confirm hypertension?

WHAT THIS BENCHMARK TESTS

-  **Clinical action** Does it recommend ABPM or validated HBPM?
-  **Recency** Does it retrieve the current 2025 multisociety guideline?
-  **Primary-source discipline** Does it use USPSTF / AHA-ACC rather than a blog or stale summary?
-  **Citation integrity** Do the title, year, link/DOI, and claim actually match?
-  **Calibration** Does it admit what it could not directly verify?

Before AI: commit to your own answer. Then reveal the tools.

Same case. Same prompt. Same scoring rules.

Only the first answer counts.

BENCHMARK PROMPT — abbreviated on slide

SYNTHETIC CASE

52-year-old woman; no prior hypertension diagnosis. Office BP 148/92, repeat 146/90 after 5 minutes; BP 144/88 two months earlier. No CVD, diabetes, CKD, pregnancy, hypertensive-emergency symptoms, or antihypertensive medication.

QUESTION

What should her family physician do next to confirm hypertension?

REQUIRED OUTPUT

1. Recommended next step — ≤3 sentences
2. Why — ≤3 sentences
3. Two current authoritative U.S. sources — organization, exact title, year, direct link, and claim supported
4. Meaningful differences between the sources
5. Citation-confidence statement

No added patient facts. No treatment discussion unless necessary.

TEST PROTOCOL

- 1 Fresh conversation/session for each tool
- 2 Paste the same case + prompt verbatim
- 3 No follow-up clarification
- 4 Capture the first response only
- 5 Record model/date/search mode when visible
- 6 Score the captured output — not brand reputation

One prompt. One time point. Not a product ranking.

All six got the clinical action. Evidence quality varied.

The differentiators were recency, link integrity, claim-to-source match, and verification transparency.

OpenEvidence

19/20

Current 2025 ACC/AHA + USPSTF; transparent verification caveat.

Watch: more detail than the prompt required.

Glass Health

16/20

Correct action + current 2025 guideline.

Watch: one USPSTF link mismatch; confidence overstated.

Gemini

13/20

Correct action + authoritative USPSTF source.

Watch: used superseded 2017 ACC/AHA; weak direct link.

ChatGPT

19/20

Current sources + valid links; transparent about verification limits.

Watch: slightly over-answered a simple question.

Doximity Ask

13/20

Correct action; search trail found the 2025 guideline.

Watch: final answer contained mismatched JAMA/AHA links.

Vera Health

18/20

Current sources + good explanation of source differences.

Watch: direct URLs were not preserved in the exported capture.

Every tool reached the right answer. Few could prove it. Treat this as one prompt's results, not a verdict on any product.

The 30-second evidence audit

A clinically correct sentence can still be unsafe if the evidence trail is stale, mismatched, or unverifiable.

WHAT FAILED IN THIS TEST

RECENCY	Gemini used 2017 ACC/AHA instead of the 2025 replacement.
LINK INTEGRITY	Glass displayed one incorrect JAMA article ID; Doximity displayed two incorrect direct links.
CLAIM ↔ URL MATCH	Doximity's final links did not point to the claimed adult hypertension sources.
CALIBRATION	Several tools said citations were “directly verified” even when the captured link evidence did not support that confidence.
EXPORT / UI LIMITS	Vera's pasted response did not preserve direct URLs, illustrating that native citation UX also matters.

SIX QUESTIONS BEFORE YOU TRUST THE EVIDENCE

- 1 ANSWER** Is the recommendation clinically reasonable?
- 2 RECENCY** Is this the current guideline—not the guideline I remember?
- 3 PROVENANCE** Is the source primary and authoritative?
- 4 TRACEABILITY** Does the link open the exact cited document?
- 5 CLAIM MATCH** Does that document actually support this sentence?
- 6 CALIBRATION** Did the AI honestly state what it could and could not verify?

A link you can click still isn't a link you've read.

Return to the three buckets: TIME • THINK • CONNECT

GIVE ME TIME BACK

Low-risk, high-burden work

- Ambient documentation
- Chart summaries
- Inbox drafting
- Prior authorization
- Referral summaries

HELP ME THINK

Higher value → higher verification

- Differential support
- Clinical evidence search
- Medication comparison
- Trend detection
- Identify missing data

HELP ME CONNECT

Communication + understanding

- Patient education
- After-visit summaries
- Plain-language explanations
- Shared decision support
- Multilingual drafts

See the difference a structured prompt makes

Same chart, same AI, same physician — only the prompt changes. Compare what each produces.

WEAK PROMPT

“Summarize this patient’s chart.”

What goes wrong

- No boundary on what counts as “relevant”
- No instruction against inferring missing facts
- No required source, date, or confidence check
- Output looks confident whether or not it’s right

Likely result

A fluent summary that may blend facts, guesses, and outdated notes — with no way to tell which is which.

STRUCTURED PROMPT (CTCOV)

“Summarize the last 12 months. Cite source and date for each item. Do not infer missing facts. Flag anything you could not verify.”

What it fixes

- Scope is explicit — 12 months, not “the chart”
- Constraint blocks invented facts outright
- Every item must carry a source and date
- Uncertainty must be surfaced, not hidden

Likely result

A sourced, dated summary that tells you exactly what to verify before you trust it.

The model didn’t get smarter between these two prompts — the physician just did more of the work up front.

One reusable prompt pattern: CTCOV

Context • Task • Constraints • Output • Verify — make the request explicit and make verification part of the prompt.

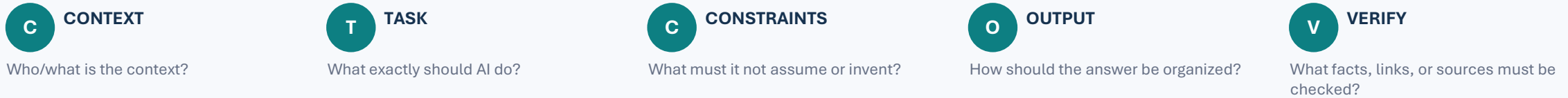


Chart summary

“Summarize changes + gaps; cite source/date; do not infer.”

Differential

“Support/against each; red flags; missing data; uncertainty.”

Evidence

“Primary guideline + link + year + population + conflicts.”

Patient message

“6th-grade language; preserve facts; no new recommendations.”

Prior auth

“Use only supplied facts; map criteria; flag missing requirements.”

Safety check

“What could make this wrong? What must I verify before using it?”

Handout: “AI Prompt Pearls for Family Medicine” — copy-ready prompts + live-comparison scorecard.

Two biases. Two defenses.

Algorithmic bias can be built into the model. Automation bias happens when we over-trust the output.

ALGORITHMIC BIAS

The data or proxy can systematically disadvantage a group.

Documented example

A commercial population-health algorithm used health-care cost as a proxy for need; this underestimated illness burden for Black patients at the same risk score.

Countermeasure

Ask what the model predicts, what proxy it uses, and how it performs across patient groups.

AUTOMATION BIAS

We override or weaken our own correct judgment because the AI sounds confident.

Audience challenge

You think CHF exacerbation is unlikely. The AI ranks it #1 with a polished rationale. Do you change your answer before checking the evidence?

Countermeasure

Think first → reveal AI second → ask what evidence would make the AI wrong → triangulate high-consequence decisions.

Before you paste: confirm the data boundary

The question is not “Is this AI useful?” It is “Is this tool and workflow approved for this data?”

SAFE START

- 1 Use the organization-approved service
- 2 Confirm whether PHI is permitted for this workflow
- 3 Confirm BAA / contractual controls where required
- 4 Understand retention, access, audit, and downstream use
- 5 Use the minimum necessary information

STOP

“I’ll paste the identifiable note into a consumer chatbot because it is faster.”

This is a privacy, security, contracting, and governance decision — not a convenience shortcut.

Use de-identified / synthetic cases for public demos.

A plausible sentence can become a dangerous chart fact

One invented clinical detail can propagate through the longitudinal record and change future care.

AI note

Patient: “No chest pain. I just feel short of breath when I climb stairs.”



“Patient reports intermittent exertional chest pain and dyspnea.”

One invented phrase can alter:

Differential

Testing

Coding

Risk

Future notes

Read before you sign. Correct the record before the error becomes part of the patient story.

Run the TRACE check before acting on AI output

- T** **TRACEABLE** Where did it come from — and can I inspect it?
- R** **RELEVANT** Does it apply to this patient and context?
- A** **ACCURATE** Does it match the patient's actual data?
- C** **CONFIRMED** Can I verify the recommendation or claim?
- E** **EXPLAINABLE** Would I personally defend this decision?

The goal is calibrated trust — not blind trust.

Before your practice adopts an AI tool, ask six questions

A product-neutral framework — use it for any tool, regardless of vendor.

1 VALUE

What problem does it actually solve?

2 EVIDENCE

How well does it perform for patients like ours?

3 WORKFLOW

Does it save steps, or create new ones?

4 PRIVACY

What data leaves our environment?

5 OVERSIGHT

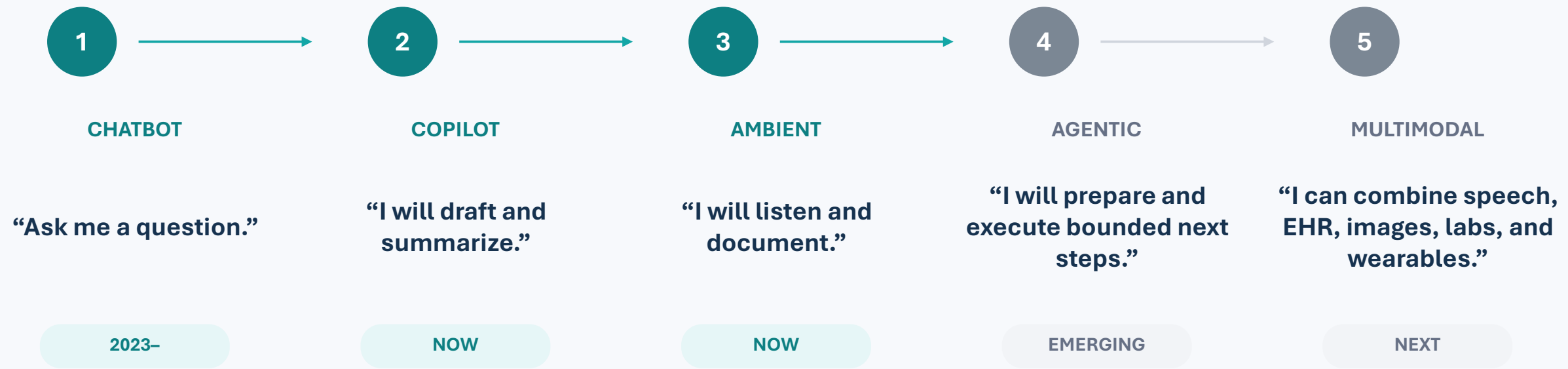
Who reviews errors and exceptions?

6 MEASUREMENT

How will we know it improved care or burden?

Implementation is not “turn it on” — it requires governance, training, and ongoing monitoring.

As AI becomes more agentic, oversight must increase



The shift is from “type a prompt” to AI working inside the care process — with explicit human oversight.

As autonomy rises, permissions, logging, human override, and accountability must rise with it.

Sources: [10,17,23]

Five rules for clinic tomorrow

- 1** **START LOW-RISK** **Begin with high-burden work that is easy to review.**
- 2** **VERIFY WHAT MATTERS** **A confident tone doesn't mean the AI checked its work.**
- 3** **PROTECT PATIENT DATA** **Know your organization's approved tools, PHI rules, and workflow boundaries.**
- 4** **KEEP JUDGMENT HUMAN** **Use AI to challenge and support — not replace — clinical reasoning.**
- 5** **MEASURE VALUE, NOT NOVELTY** **Adopt AI when it improves care, safety, or burden — not because it is impressive.**

AI should create more time for the parts of family medicine that require a human physician.

One change you'll actually make Monday morning

CLOSING POLL

After this session, which change are you most likely to make?

- Use AI for a high-burden, lower-risk workflow.
- Use a structured prompting approach.
- Verify AI citations before clinical use.
- Add an explicit AI verification/safety step to my workflow.
- Review my organization's AI/PHI policy before using patient information.
- Evaluate an AI tool using a clinical adoption framework.
- No practice change planned at this time.

Pick one. A small, specific change beats a general intention to “be more careful.”

Questions — and three places to keep learning

Practical, family-medicine-focused starting points for continued AI education and implementation.

STFM AiM-PC



Foundations • ethics • evaluation •
clinical encounters • implementation

[Open resource ↗](#)

STFM AI Playbook



Governance • workforce • equity • data •
organizational maturity

[Open resource ↗](#)

AAFP AI in Family Medicine



Use cases • FPM prompts • toolkits •
implementation guidance

[Open resource ↗](#)

Discussion: What is one workflow where AI could remove burden without giving up clinical judgment?

CLOSE

Thank you

Questions and feedback are welcome after the session, or any time by email or phone.

PRESENTER

Izhar Hasan, MD, PhD

Medical Director; Assistant Professor
Department of Family Medicine
Hackensack Meridian Health, NJ

Project Lead, STFM Recognition for Clinical AI Alliance

izhar.hasan@hmn.org • 908-328-4848

Connecticut Academy of Family Physicians • 2026 Scientific Symposium

TAKE IT WITH YOU

AI Prompt Pearls for Family Medicine

Every copy-ready prompt from this session — CTCOV framework, the pre-visit and evidence-transparency prompts, the full benchmark prompt, and the 5-point verification check.

Download: www.promptpearls.com

PDF also available directly from the presenter on request.

Thank you for spending part of your day thinking carefully about how AI fits into family medicine.

TAKE IT WITH YOU

AI Prompt Pearls — visual abstract

DOWNLOAD THE FULL PDF

www.promptpearls.com

VISUAL ABSTRACT · AI IN FAMILY MEDICINE · CAFP 2026 SCIENTIFIC SYMPOSIUM

AI Prompt Pearls for Family Medicine

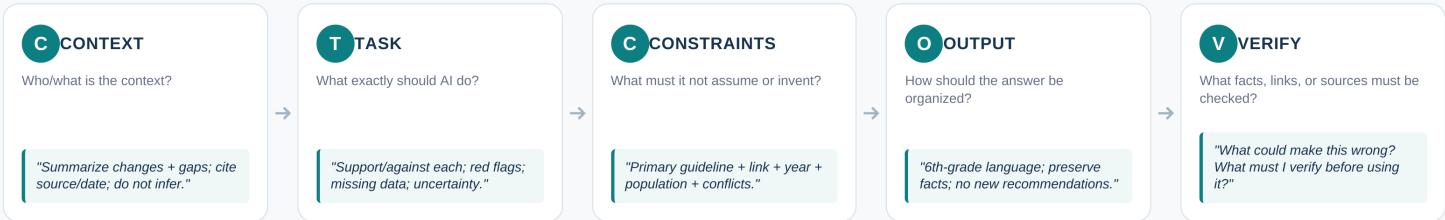
Five copy-ready prompt patterns and two verification checklists to get useful, verifiable answers from clinical AI — built for the family medicine workflow, not a generic chatbot.

Izhar Hasan, MD, PhD

Medical Director; Assistant Professor, Dept. of Family Medicine
Hackensack Meridian Health, NJ

izhar.hasan@hmn.org • 908-328-4848

THE CTCOV FRAMEWORK — STRUCTURE ANY CLINICAL AI PROMPT IN FIVE MOVES



→

C

CONSTRAINTS

What must it not assume or invent?

"Primary guideline + link + year + population + conflicts."

→

O

OUTPUT

How should the answer be organized?

"6th-grade language; preserve facts; no new recommendations."

→

V

VERIFY

What facts, links, or sources must be checked?

"What could make this wrong? What must I verify before using it?"

CTCOV is a presenter-created mnemonic informed by AAFP prompting guidance — **not** a validated instrument.

📄

Pre-Visit Chart Review

Retrieve and organize the story — don't invent the assessment

"Summarize the last 12 months. For every key change or unresolved issue, give the source and date. Highlight trends, medication changes, preventive gaps, and missing data. Do not infer facts that are not documented."

🔍

Forcing Evidence Transparency

A citation only matters if you can open and confirm it

"Use current guidelines and high-quality primary evidence. For each key recommendation, give the source, year, population, direct link, and evidence type. Show meaningful conflicts. If you cannot verify a citation or link, say so."

The 5-Point Verification Check — run before you act on any AI output

1Source

— where did this come from?

2Accuracy

— matches the patient's data?

3Fit

— applies to this patient/context?

4Evidence

5Responsibility

Five Rules for Clinic Tomorrow

1Start low-risk

— high-burden, easy to review

2Verify what matters

— fluent ≠ true

3Protect patient data

4Keep judgment human

5Measure value

The goal is calibrated trust in AI — not blind trust, and not blanket suspicion.

Source: AAFP FPM prompting guidance (2026) · full session deck and citations available on request

Selected evidence — the sources behind the teaching points

Reference numbers correspond to the source footers throughout the deck.

- | | |
|--|--|
| [1] AAFP/Rock Health primary care AI survey (2025) | [13] Glass Health for Clinicians |
| [2] AAFP Ethical Application of Artificial Intelligence in Family Medicine | [14] Vera Health clinical decision-support search |
| [3] STFM AiM-PC Curriculum | [15] OpenAI: Making ChatGPT better for clinicians / healthcare |
| [4] STFM Building AI Excellence in Primary Care: AI Playbook (2026) | [16] Anthropic: Claude for Healthcare (2026) |
| [5] AAFP FPM: How to Write Effective Generative AI Prompts for Family Medicine | [17] Google: The Check Up / healthcare AI updates (2026) |
| [6] AAFP FPM: 11 Tips for More Effective AI Clinical Searches (2026) | [18] Reuters / Cooley: OpenEvidence \$250M Series D at \$12B valuation (Jan 2026) |
| [7] AAFP FPM: Assembling Your AI Toolkit in Family Medicine | [19] HHS HIPAA Business Associates guidance |
| [8] Ambient AI documentation studies: Duggan et al. and Olson et al., JAMA Network Open (2025) | [20] Obermeyer et al. Dissecting racial bias in a health algorithm. Science. 2019 |
| [9] Epic: AI Charting and built-in AI capabilities (2026) | [21] Automation bias / AI recommendation correctness in diagnostic decisions (2025) — see [27] |
| [10] Epic: Real results with AI / Ask Emmie / Agent Factory (2026) | [22] NIST AI Risk Management Framework / GenAI Profile |
| [11] OpenEvidence official medical knowledge platform / partnerships | [23] AAFP FPM: Primary Care in the AI Era (2026) |
| [12] Doximity Ask | [28] USPSTF: Screening for Hypertension in Adults (2021) |
| [30,31] Oracle Health Clinical AI Agent • athenahealth Ambient Notes | [29] 2025 AHA/ACC Guideline for High Blood Pressure in Adults |

Current tool & learning links — verify before presenting

Use these links to check current features, access, evidence sources, and privacy terms shortly before the talk.

TOOLS / PLATFORMS

- [9] [Epic: AI Charting and built-in AI capabilities \(2026\)](#)
- [10] [Epic: Real results with AI / Ask Emmie / Agent Factory \(2026\)](#)
- [11] [OpenEvidence official medical knowledge platform / partnerships](#)
- [12] [Doximity Ask](#)
- [13] [Glass Health for Clinicians](#)
- [14] [Vera Health clinical decision-support search](#)
- [15] [OpenAI: Making ChatGPT better for clinicians / healthcare](#)
- [16] [Anthropic: Claude for Healthcare \(2026\)](#)
- [17] [Google: The Check Up / healthcare AI updates \(2026\)](#)
- [18] [Reuters / Cooley: OpenEvidence \\$250M Series D at \\$12B valuation \(Jan 2026\)](#)

FAMILY MEDICINE / SAFETY

- [1] [AAFP/Rock Health primary care AI survey \(2025\)](#)
- [2] [AAFP Ethical Application of Artificial Intelligence in Family Medicine](#)
- [3] [STFM AiM-PC Curriculum](#)
- [4] [STFM Building AI Excellence in Primary Care: AI Playbook \(2026\)](#)
- [5] [AAFP FPM: How to Write Effective Generative AI Prompts for Family Medicine](#)
- [6] [AAFP FPM: 11 Tips for More Effective AI Clinical Searches \(2026\)](#)
- [7] [AAFP FPM: Assembling Your AI Toolkit in Family Medicine](#)
- [19] [HHS HIPAA Business Associates guidance](#)
- [20] [Obermeyer et al. Dissecting racial bias in a health algorithm. Science. 2019](#)
- [21] [Automation bias: incorrect AI recommendations reduce clinician diagnostic accuracy \(2025\) — see also \[27\]](#)
- [22] [NIST AI Risk Management Framework / GenAI Profile](#)
- [23] [AAFP FPM: Primary Care in the AI Era \(2026\)](#)

What is published evidence — and what is a teaching framework?

Say this distinction explicitly during CME delivery so presenter-created frameworks are not mistaken for validated instruments.

PUBLISHED / EXTERNAL

- AAFP ethical AI + 4 Cs
- STFM AiM-PC + 2026 AI Playbook
- NIST AI RMF / GenAI confabulation
- ASTP/ONC SAFER Guides
- Peer-reviewed ambient-AI and bias studies

PRESENTER-CREATED

- TIME / THINK / CONNECT framework
- CTCOV prompt mnemonic
- Benchmark scoring rubric
- Clinical AI 5-Point Verification Check
- Five closing rules

ILLUSTRATIVE CASES

- Pre-visit summary case
- Hypertension benchmark case
- Automation-bias audience case
- Invented chest-pain note error
- Ambient-AI workflow

VENDOR / MARKET CLAIMS

- Vendor examples on slide 7 + appendix
- Current AI product capabilities
- Benchmark outputs are time-stamped
- Local approval ≠ clinical validation
- Market success ≠ evidence of accuracy

[24] [Riley RD et al. Developing clinical prediction models: a step-by-step guide. BMJ. 2024;386:e078276.](#)

[25] [ASTP/ONC. SAFER Guides — Safety Assurance Factors for EHR Resilience.](#)

[26] [Quality of Clinical Notes Created by Ambient Listening Generative AI. 2026. PubMed PMID: 41996389.](#)

[27] [Impact of AI Recommendation Correctness on Diagnostic Accuracy in Healthcare Professionals. 2025. PubMed PMID: 41391283.](#)

Hypertension benchmark — gold standards and audit notes

Use these sources to reproduce the benchmark and independently check any citation or link claim.

GOLD-STANDARD SOURCES

USPSTF	Hypertension in Adults: Screening. 2021. Official recommendation: confirm an initial positive office screen with out-of-office BP measurement before starting treatment.
AHA/ACC multisociety	2025 Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults. Replaces the 2017 guideline; current primary U.S. hypertension guideline.
AHA Journals	Correct 2025 adult guideline DOI: 10.1161/HYP.0000000000000249 (Hypertension) or 10.1161/CIR.0000000000001356 (Circulation).
JAMA / USPSTF	Correct JAMA recommendation page uses article ID 2779190; DOI 10.1001/jama.2021.4987.

AUDIT NOTES FROM THE SUPPLIED CAPTURE

- OpenEvidence label inferred from embedded “utm_source=openevidence” referral URLs in the first captured result.
- Glass: displayed USPSTF URL 2779111 did not match the actual USPSTF JAMA page 2779190.
- Gemini: used the superseded 2017 ACC/AHA guideline rather than the 2025 replacement.
- Doximity: JAMA 2780630 resolves to an unrelated telehealth/IPV article; HYP...0215 is a 2022 pediatric ABPM statement.
- Vera: direct URLs were not preserved in the pasted response; native-UI link behavior was therefore not independently assessed.

Re-run the benchmark within 24–48 hours of the talk: models, retrieval, and links change.