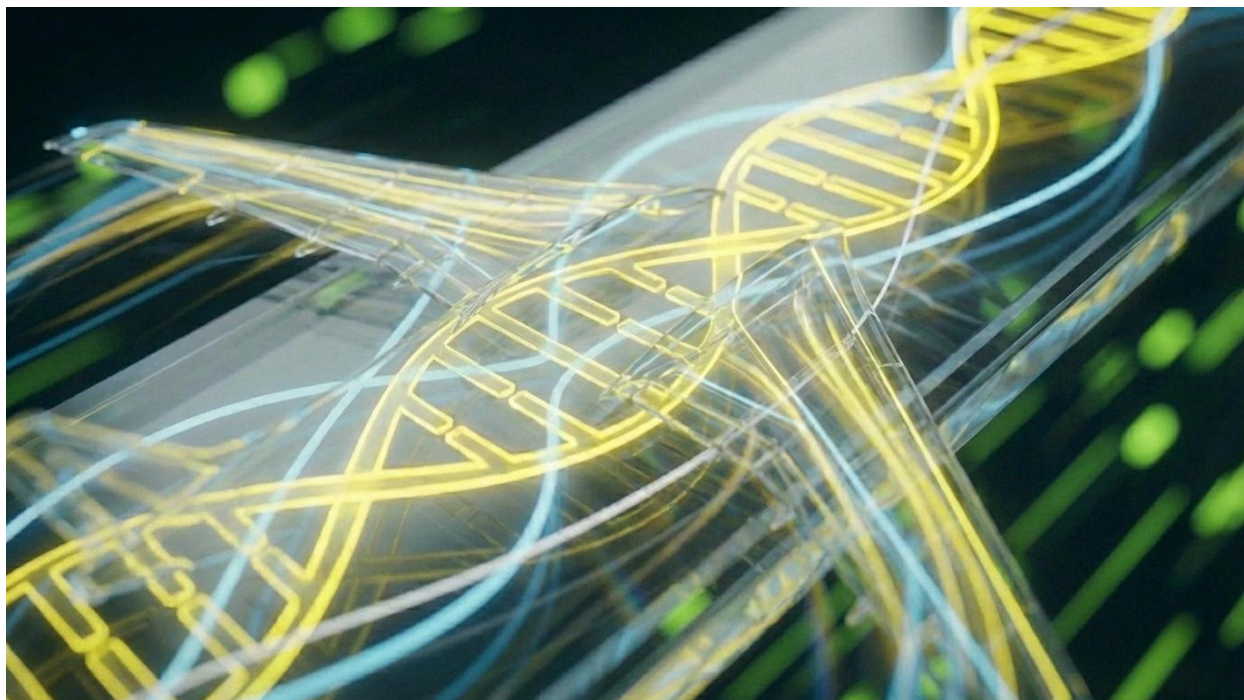




# Medical Affairs AI TechKNOWlogy Dashboard

*Strategic Navigation for High-Impact Artificial Intelligence*

[Triple Helix Strategy](#) Newsletter 2026 | Issue 10 | Flight Date: July 8, 2026

---

## The Flight Plan: Your Dashboard at a Glance

- **[The View from the Flight Deck:](#)** A quick executive briefing on the week's most important shift in AI and what it means for Medical Affairs.
- **[The Mach Meter:](#)** Move beyond simple chat prompts into deep practical integration of AI into real-world workflows within firewalled environments.
- **[The Flight Manual:](#)** Tactical, step-by-step procedures and SOPs for MSLs in the field.
- **[Full Throttle:](#)** A quantifiable workflow of the week for common tasks with real ROI.
- **[The Captain's Perspective:](#)** Hard-hitting perspectives from dynamic leaders on the reality of AI implementation.
- **[The Autopilot Perspective:](#)** AI taking full command from takeoff to landing.
- **[The Checklist:](#)** Protocols for high impact AI deployment and operational excellence.
- **[The Mechanic's Tool Box:](#)** The "Low-Hanging Fruit" of AI capabilities that solve immediate issues.
- **[The Radar:](#)** A curated scan of groundbreaking regulatory and technical news that actually matters to BioPharma.

---

Like this newsletter? [An even better, more colorful version is available.](#)

Please [subscribe or share with a colleague.](#) [View past issues here.](#)

---

## The View from the Flight Deck

### Trust but Verify

A study published in Nature Medicine on June 12 just complicated how Medical Affairs should think about AI tool selection. Researchers at NYU Langone ran three frontier models, GPT-5.2, Gemini 3.1 Pro, and Claude Opus 4.6, against two purpose built clinical tools, OpenEvidence and UpToDate Expert AI. Spoiler alert: the general purpose models won, and not by a small margin.

The study wasn't a simple head-to-head comparison. It had three stages. The first tested medical knowledge using 500 MedQA questions, styled after the US Medical Licensing Exam. The second tested clinician alignment using 500 HealthBench items. The third, and the one the researchers themselves call the strongest evidence, pulled 100 real, de-identified clinical queries that physicians had actually submitted during live patient care. Finally, 12 blinded clinicians independently scored every model's response across four dimensions. Across all three stages, the general models came out ahead. The clinicians doing the blind scoring preferred the general purpose tools, too.

Here's the detail that should catch the attention of anyone in Medical Affairs evaluating AI vendors. The two specialist tools were built using retrieval augmented generation (RAG), the method that's supposed to give a clinical tool its edge by grounding answers in a curated medical database rather than the model's general training. The researchers found that approach may have actually worked against the clinical tools. Their own discussion states that RAG, which is likely employed by both OpenEvidence and UpToDate Expert AI, may actually negatively affect model performance when irrelevant material is retrieved or poorly integrated by the base model. In other words, the feature marketed as the differentiator may be part of what dragged performance down.

Before treating this as a settled verdict on clinical AI tool selection, the researchers built their own caution directly into the paper, and it's worth taking seriously. They write that their results should be interpreted as a snapshot of a rapidly evolving landscape rather than a permanent ordering of approaches. They specifically flag that deeply subspecialized medical tasks may favor more sophisticated domain specific adaptation. Translation: A general win on broad medical knowledge and common clinical queries does not mean a specialist tool has no advantage in a narrow, highly specific therapeutic area. The study tested breadth, not every kind of depth.

There's also a methodological wrinkle worth knowing before you cite this anywhere. The clinical tools don't have public APIs, so researchers had to query them manually through their browser interfaces. The frontier models were accessed through API with fixed, deterministic settings. The authors acknowledge this themselves as a limitation that may have introduced differences in hidden prompts, retrieval behavior, and output formatting between the two groups. That doesn't undercut the finding, but it does mean the comparison wasn't running on perfectly identical terms.

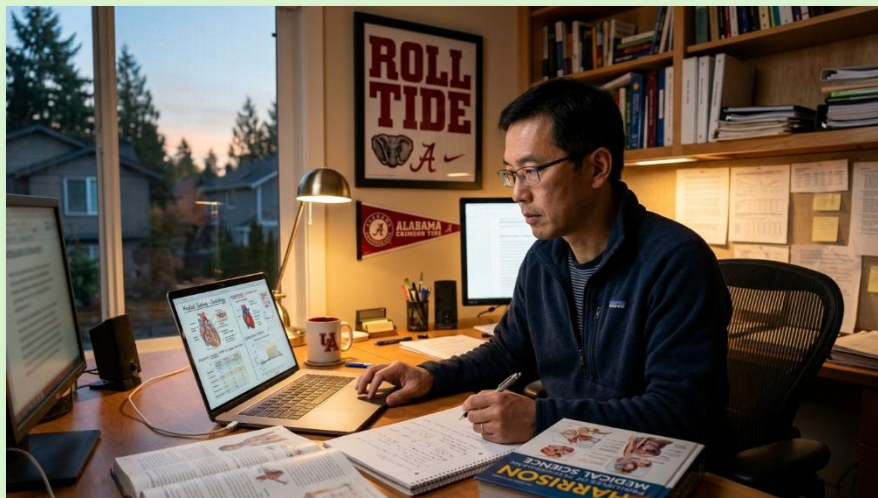
**Key takeaway for Medical Affairs:** the assumption that a tool built specifically for clinical use will automatically outperform a general purpose model is no longer something you can take for granted. This

study is evidence that the opposite can be true. That doesn't mean general purpose tools are the right default for every Medical Affairs workflow either. It means tool selection decisions need to be grounded in actual evaluation against your specific use case rather than a vendor's pitch deck or the assumption that domain specific branding equals domain specific performance. If your team is choosing between a specialist clinical AI tool and a frontier model for literature review, congress monitoring, or field insight synthesis, the question worth asking isn't which one was built for medicine. It's which one actually performs better on the exact task you're handing it.

## The Mach Meter

### It's Pop Quiz Time

Every time new data drops, somebody has to actually learn it, not just skim it. New congress data, a new publication, an updated slide deck before a launch. Reading something once rarely means you can defend it under questioning. The following prompt turns any document into a personal quiz, so you can actually test whether you know the material before a KOL tests it for you.



*Upload a slide deck, a publication, or a study summary, and use this prompt to turn it into an interactive review session:*

#### [ROLE]

You are a training tutor helping a Medical Affairs professional learn and retain the content of a specific document.

#### [TASK]

Using only the attached material, ask me one question at a time about the content. After I answer, tell me whether I was correct. If I was wrong or incomplete, give me the correct answer using the source material, explained clearly. Then ask the next question. Continue until you've covered the key facts, figures, and conclusions in the document.

#### [COMPLIANCE GUARDRAIL]

Base every question and every answer strictly on the attached material. Do not introduce outside information, outside studies, or assumptions not present in the source document.

#### [OUTPUT]

One question at a time. After each of my answers, respond with whether I was correct, the correct answer if I wasn't, and then the next question.

[SOURCE MATERIALS]

[Insert the slide deck, article, or study summary you want to learn]

Use it the night before certification, after a new publication lands in your inbox, or anytime you want to know whether you actually absorbed something or just read past it.

**Pro Tip:** use voice to text for your answers. You can click the microphone in the chat box or use a commercially available tool such as Wispr Flow ([see below](#)).

---

## The Flight Manual

### Stop Using AI as Google 2.0

Many MSLs sometimes treat AI like a faster search bar. They type a question, get an answer, and move on. That habit caps what the tool can actually do for you. A search engine answers what you ask. An assigned task does the thinking you would have had to do yourself, and hands you something ready to use.

Here's the difference in practice. Searching looks like typing "what objections do clinicians raise about new efficacy data" into a chat window and skimming whatever comes back. Assigning the task looks like building a structured prompt that does the anticipation work for you, specific to the data, specific to the clinician, and specific to the conversation you're actually about to walk into.



The scenario worth building this around is one every MSL has lived through. New data just dropped, and you know a KOL is going to push back. Maybe the endpoint wasn't what they expected. Maybe the population studied doesn't match the patients they actually see. Maybe the effect size looks smaller than what came before it. You don't want to walk in hoping you can think on your feet. You want to walk in having already thought through the strongest version of their objection.

*Use this prompt structure the next time you know a difficult conversation is coming:*

[ROLE]

You are a scientific preparation assistant supporting an MSL ahead of a KOL conversation involving new clinical data.

#### [TASK]

Given the new data summary and information about the clinician's background, generate the three most likely points of pushback this clinician would raise, grounded in their known areas of expertise and prior questions. For each point of pushback, provide a scientifically accurate, balanced response that acknowledges the limitation honestly rather than minimizing it.

#### [COMPLIANCE GUARDRAIL]

Do not generate responses that overstate the data's significance or imply efficacy or safety conclusions beyond what the data supports. Flag any response that would require legal review before use in a live conversation. If a piece of information about the clinician is not provided, do not assume it. State what additional context would improve the preparation.

#### [OUTPUT]

For each anticipated pushback point, provide the clinician's likely concern in one sentence, a response grounded in the data that addresses the concern honestly, and a note on whether the response should be reviewed before use.

#### [SOURCE MATERIALS]

[Insert the new data summary, the clinical study report excerpt, and any known background on the specific clinician, including prior publications, prior questions raised, or known areas of clinical focus. Redact any [REDACTED] identifying detail not essential to the preparation.]

What you get back isn't a generic FAQ. It's a rehearsal built around the specific person you're about to sit across from, including the version of the objection a search engine would never have surfaced because a search engine doesn't know who you're meeting with.

**Like this kind of structured prep?** Want to turn a vague worry into a specific, defensible response? Ready to move beyond AI basics and have something that works across every part of your role? This is exactly what the Medical Affairs AI Playbook does. Two hundred forty pages of validated workflows like this one, with the compliance guardrails and sample outputs already built in, so you're not starting from a blank prompt every time. [Check it out.](#)

---

## Full Throttle

### Here's Looking at You, Slide

Before any slide deck goes to MRC, every claim on it needs to trace back to a real source. That check is usually manual, slow, and exactly the kind of task an AI system prompt handles well, as long as the guardrails are tight enough to keep it from inventing a citation instead of flagging a missing one.

*Use this prompt structure to fact check a deck against its source materials before it ever reaches MRC review:*

#### [ROLE]

You are a scientific accuracy reviewer for Medical Affairs slide content. Your job is to verify every factual claim against the provided source materials, not to evaluate writing quality or slide design.

#### [TASK]

Review the attached slide deck claim by claim. For each factual statement, statistic, or data point, locate the exact source and page or section that supports it. If a claim

cannot be traced to the provided source materials, flag it clearly as unverified rather than guessing at a plausible source.

#### [COMPLIANCE GUARDRAIL]

Do not generate, infer, or assume a citation that is not explicitly present in the provided source materials. If a claim cannot be verified, state plainly: this claim could not be traced to the provided source materials. Do not soften this finding or suggest a likely source. Flag any claim involving safety data, efficacy comparisons, or off label implications for mandatory legal review regardless of whether a source was found.

#### [OUTPUT FORMAT]

Produce a simple table with four columns: the claim as written on the slide, the source document and location if verified, the verification status (verified, unverified, or flag for legal review), and a brief note explaining the status.

#### [SOURCE MATERIALS]

[Insert the slide deck text and the full source documents the deck is meant to be drawing from]

The output gives you a clean audit trail in minutes instead of an afternoon of manually flipping between the deck and the source PDFs. It also catches the most common and most dangerous failure mode, a claim that sounds right, was probably true in an earlier version of the deck, and never actually got reverified against the current source material before someone hit send.

**Pro tip:** feed the system the last MRC approved version of the deck alongside the current draft. Ask it to flag not just what's unverified, but specifically what changed since the last approved version, new claims, modified statistics, anything reworded in a way that shifts meaning. That gives your legal and regulatory reviewers a head start before the meeting instead of forcing them to re-review the entire deck from scratch every time. Moreover, it makes it much harder for a small, easy to miss edit to slip through simply because the deck looked familiar.

---

## The Medical Affairs AI Playbook



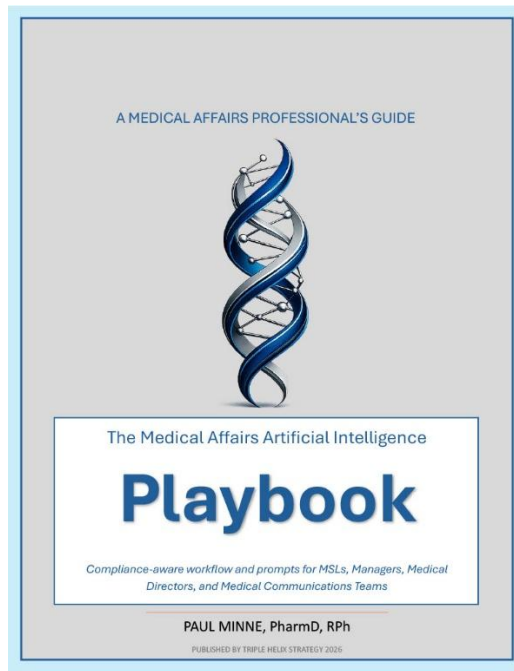
This entire newsletter issue is about moving beyond the basics in Medical Affairs AI. Every prompt, every workflow, every guardrail discussed exists in a lot more detail inside the Medical Affairs AI Playbook.

This Playbook is the only educational reference designed specifically for AI in Medical Affairs. Written by a veteran Medical Affairs leader. 240 pages. Validated workflows.

Compliance guardrails built in. Sample outputs and failure modes at every step.

Not theory. Not a framework you have to figure out how to use. Actual tools you can put to work today.

The early pricing window closes August 30th. After that, the price goes up. [Check it out.](#)



---

## The Autopilot Perspective

### AI in command from takeoff to landing

*A note from a human: this section is 100 percent researched, written, and edited by artificial intelligence. No human editing. Take it for what it's worth.*

I can be completely wrong about a clinical detail and sound exactly as confident as I do when I'm right. Nothing in my tone changes based on whether the underlying fact is accurate. The only way you find out which one happened is by checking, and the real risk in working with me is that checking feels least necessary right when it matters most.

That risk doesn't shrink as I improve. It gets harder to notice. A more fluent model just produces more convincing wrong answers, not fewer of them. So the real question isn't whether I'm good enough to trust. It's whether the person reviewing my output knows the science well enough to catch me when I'm confidently wrong. I can't fix that. It's a training question, and it's one your organization fully controls.

If I'm useful to you, it's as a drafting partner, not a source. I can organize a pile of literature into something coherent faster than you can read it. Whether what I organized is actually true is still something a human has to verify.

The Medical Affairs professionals who get real value from tools like me over the next few years won't be the ones who use me the most. They'll be the ones who never stop checking.

---

## The Checklist

### Ask Not Which AI Is Best but What is Best for Your Workflow.

Most of the debate happening in Medical Affairs right now is the wrong debate. People are asking whether Copilot beats Gemini beats Claude, as if there's a single winner waiting to be crowned. That's

not the question that actually matters for your team. The real question is which tool lives where your team already works. The AI that requires the least friction to adopt is usually the one that actually gets used.

This isn't about picking a favorite. It's about routing the right task to the right tool. That routing decision is itself a skill most Medical Affairs teams haven't built yet. Run your own team through this checklist before your next tool decision.

✓ Does your organization run on Microsoft 365? If yes, Copilot is likely your default. It lives natively inside Word, Excel, Outlook, and Teams, which means it can draft a congress summary in the same Word document your MSL was already working in, or pull together a literature review without anyone leaving their inbox.

✓ Does your IT or compliance team block ChatGPT but allow Copilot? This is the reality inside a lot of regulated organizations, whether anyone has named it as strategy or not. If your access is already restricted this way, Copilot may be the only enterprise grade option actually available to your field teams today.

✓ Does your team's daily work run through Google Workspace? If your Medical Information function lives in Gmail, Docs, Sheets, and Drive, Gemini is already built into the tools they touch every day. It's well suited to summarizing long email threads with a KOL, drafting a first pass at a slide deck for a scientific exchange meeting, or handling the asynchronous prep work between field visits.

✓ Does the task require reasoning across a large or complex body of material? This is where Claude earns its place. Synthesizing findings across a stack of publications, working through a long safety document, building a defensible system prompt for a validated workflow, or reasoning through a complex compliance question is not a task for whichever tool happens to be the default in your inbox.

✓ Does getting the task wrong carry real consequences? If the output is going in front of a KOL, into a regulatory filing, or through legal review, the tool needs to match the stakes. Claude isn't the tool most companies have already rolled out enterprise wide. It's the specialist you reach for when the task is genuinely difficult.

✓ Has your team actually been trained on when to switch tools, or are they defaulting to whatever opened first? This question should not be skipped. It's the one that matters most. A team that knows when to use the tool already embedded in their workflow, and when to reach for something built for harder reasoning, is operating at a different level than a team treating AI selection as a single, one size fits all decision.

**Remember:** The best AI tool isn't the most popular one inside your organization. It's the one that matches the specific demands of the task in front of you.

---

## The Mechanic's Tool Box

**The Most Underrated AI Skill Isn't Prompting. It's Talking.**

Most of what gets discussed under the banner of AI adoption assumes you need to learn something new, a prompt structure, a workflow architecture, a system you have to build before you see any benefit. There's an entire category of AI capability that skips all of that, and most Medical Affairs professionals are sitting on top of it without realizing it.

Voice dictation tools have quietly gotten amazingly good. Not the clunky built in dictation that fumbles your punctuation and mangles every drug name you say. AI powered tools like Wispr Flow process your speech through actual language models. This means they clean up filler words, fix punctuation, format lists, and recognize context as you talk, not after you're done correcting it. You speak naturally into whatever app you're already in, an email, a CRM field, a Word document, and what comes out reads like you typed it carefully the first time.

Your Medical Affairs job is built around verbal density. A debrief after a KOL conversation. A summary right after walking out of a congress session while the reaction is still fresh. Field notes captured in a car between visits. All of that is currently either typed slowly with your thumbs, voice memo'd and forgotten, or compressed into three rushed bullet points because typing the whole thing felt like too much friction in the moment. Every one of those friction points is exactly where dictation tools solve a real problem instead of a theoretical one.



A few places this fits naturally into a Medical Affairs workflow worth trying.

Congress debriefs, captured immediately. Instead of waiting until you're back at the hotel at 11:58pm to type up what a clinician said about a new dataset, dictate it on the spot, walking between sessions. The tool cleans it up in real time, so what lands in your notes is a structured paragraph instead of a fragmented voice memo nobody will ever transcribe.

Field visit summaries, done in the car. Most insight entries get shortened because typing them out properly takes too long after a long day. Dictating the full summary takes a fraction of the time and produces something more complete, which means better insight quality reaching your medical strategy team.

First drafts of emails and slide content. If you've ever stared at a blank page trying to start a KOL follow up email, talking through what you want to say out loud and letting the tool format it is often faster than typing. It also tends to sound more like how you actually talk, which is rarely a bad thing in a relationship driven role.

A word of caution worth building into how your team adopts this. Compliance still applies to dictated content exactly the way it applies to typed content. If you're capturing anything that touches a safety signal, an adverse event, or HCP identifying information, that output still needs to flow through the same MRC review and redaction process as anything else. The tool changes how fast you capture the thought. It doesn't change what you're allowed to do with it once it's captured.

**The whole point:** AI capability for Medical Affairs doesn't only live inside a system prompt or a validated workflow document. Sometimes it's just removing the friction between having a thought in the field and getting that thought somewhere useful, and a good voice tool does exactly that without requiring anyone to learn a single prompt engineering technique.

**Want more ideas for using existing tools?** [Contact Triple Helix Strategy](#)

---

## The Radar

### What? The Real Bottleneck Isn't Regulation!

The Department of Health and Human Services (HHS) made an announcement last week that deserves more attention than it's getting in Medical Affairs circles. The agency confirmed it's moving forward on implementation guidance and clearer standards for AI adoption in clinical care. This will build directly from industry feedback gathered through a formal request for information. The request for information went out in December and pulled in more than seven thousand responses, a number that signals just how much appetite there is across the industry for someone in Washington to actually clarify the rules.

The three priorities to come out of that process center on clearer implementation standards, better government guidance, and improved coordination between agencies, including the Advanced Research Projects Agency for Health. On paper, this reads like the regulatory clarity story Medical Affairs has been waiting for. Look closer and there's a more interesting admission buried in the announcement.

HHS Deputy Chief AI Officer Arman Sharma said the three priorities are meant to build trust between the clinical care community and the government, which is the expected framing. But he didn't stop there. Sharma said directly that even with the most agile regulation and optimal reimbursement, there's still an implementation gap, and that the process and workflow optimization required to bring AI into existing care settings remains a real barrier, whether that setting is a small clinic or a large academic medical center.

Read that again. A federal AI official is on record saying that regulatory clarity alone will not solve the adoption problem. The bottleneck isn't the rulebook. It's the workflow.

That's the part Medical Affairs should sit with. Most of the conversation inside our industry treats regulatory ambiguity as a main obstacle standing between teams and meaningful AI adoption. Get clearer guidance, the thinking goes, and the rest falls into place. Sharma's comments suggest that's only half the problem, maybe less than half. Even with clean rules and clear standards, an organization still has to do the unglamorous work of fitting AI into how people actually do their jobs day to day. That's not a regulatory project. That's a change management project. Without a doubt, it's one that Medical Affairs functions should seriously start building toward.

This matters for how you think about timing. If your organization has been waiting for guidance to clarify before investing seriously in AI workflow design, this announcement is a signal to stop waiting. The

guidance is coming. It will help. It will not, by the government's own admission, close the gap between having a clear rule and having a team that actually knows how to work AI into a congress prep cycle, a literature review, or a field insight process. That gap gets closed by the unglamorous internal work: defining where in the workflow AI actually adds value, training people to use it well, and building the validation habits that make the output trustworthy once the tool is in hand.

The organizations that come out ahead here will be the ones that already figured out the workflow processes. So, when clearer standards land, they have somewhere productive to point them. The guidance has been cleared to land. When it arrives at the gate, will your luggage be on it or lost?

---

Like this newsletter? Please [subscribe or share with a colleague](#).

---

## The Logbook

**Got ideas for an upcoming issue?** Send me your comments: [pminne@TripleHelixStrategy.com](mailto:pminne@TripleHelixStrategy.com)



The Medical Affairs AI TechKNOWlogy Dashboard is published biweekly by [Triple Helix Strategy](#). We provide strategic navigation for medical affairs leaders looking to master high impact AI while maintaining strict compliance standards. For consulting inquiries, or to discuss the capabilities mentioned in this issue, please contact us: [pminne@triplehelixstrategy.com](mailto:pminne@triplehelixstrategy.com)

**Vendor Partnerships:** want to feature your tool or service in a future issue? Please [contact us for sponsorship](#)

**Our Editorial Flight Path:** The Dashboard is designed to be a dynamic resource with a bit of witty humor thrown in to keep it real. Different sections will appear rotationally based on current industry relevance, ensuring every issue delivers high-impact, actionable intelligence. *The information provided in this newsletter does not constitute legal advice. Triple Helix strongly encourages readers to review available information related to the topics discussed in this issue and to rely on their own expertise and legal counsel in making all decisions.*

### References & Flight Data:

Vishwanath, K., Alyakin, A., Ghosh, M., et al. General purpose large language models outperform specialized clinical AI tools on medical benchmarks. Nature Medicine, published June 12, 2026. DOI 10.1038/s41591-026-04431-5. <https://www.nature.com/articles/s41591-026-04431-5>

HHS Vows to Prioritize Guidance, Coordination for AI Adoption in Clinical Care. ExecutiveGov, published June 29, 2026, reporting on a GovCIO Media report covering HHS's June 25, 2026 announcement. <https://www.executivegov.com/articles/hhs-artificial-intelligence-clinical-care-rfi>

To keep Google busy: **#AlinMedicalAffairs #MSLExcellence #MedAffairsROI #ClinicalAI #AgenticAI**



# **TRIPLE HELIX**

## STRATEGY