Beyond Drafting Speed: What Actually Matters in AI for Medical Writing

Barry Drees, PhD
Founder & Advisor, TriloDocs
Former EMWA President
& Editor-in-Chief

Louise Martinsson, PhD
Senior Director & Head of Medical Writing, Hypercare & Enablement R&D, GSK

Disclaimer: My statements (Louise Martinsson) in this article are reflective of my own personal opinions and not those of my employer, GSK.

After more than 40 years in medical writing, one pattern becomes familiar: solutions get described as answers to problems that were never quite the ones people thought they were.

AI in medical writing is at risk of repeating that pattern.

The conversation has been dominated by drafting speed. How fast can a system generate a first draft? How many hours can be saved? Can AI compress a six-week timeline into two? These are reasonable questions to ask, but they are aimed at the wrong target. Drafting speed was never the real bottleneck in regulated medical writing. It never has been.

The bottleneck has always been verification.

> "To shorten overall timelines, i.e. to get to an accurate, internally approved, inspection and submission ready dossier, it takes more than generating a first draft quickly. In my experience it is not the authoring - or content generation - that takes time, but rather the strategic alignment, decision-making and review steps that drives timelines. Process, people, and technology need to come together to truly shorten timelines."
>
> Louise Martinsson, PhD
> Senior Director and Head of Medical Writing, Hypercare & Enablement R&D, GSK

Anyone who has spent serious time on CSRs, integrated summaries, patient narratives, or subgroup analyses understands what I mean immediately. The challenge is not turning a table into a sentence. The challenge is systematically applying clinical judgement across hundreds of pages of structured data - laboratory shifts, adverse event summaries, efficacy analyses - and being able to say with confidence that nothing clinically meaningful has been missed.

That is exhausting work. And exhaustion matters, because humans are not machines. We become fatigued. We start to move faster than we should. We stop questioning things we questioned an hour ago.

The story I keep coming back to

There is a situation that comes to mind often: two experienced, capable medical writers spent three full days reviewing a large laboratory shift output - around 180 pages. They were thorough. They identified several clinically relevant signals. By any normal standard, they did a good job.

But when the same outputs were later reviewed systematically using structured rules, one of the findings they had flagged turned out not to be clinically meaningful. And one signal they had not flagged had been missed entirely.

This is not a criticism of those writers. They were good at their jobs. The point is that they were being asked to do something humans are not well-suited to do at scale: apply the same exhaustive review logic, consistently, repeatedly, across enormous volumes of data, under time pressure, without variation.

> "Utilizing AI to do what humans are not good at, such as reviewing enormous volumes of data, enables us to focus on what humans excel at, such as translating complex data into meaningful communication for the intended audience."
>
> Louise Martinsson, PhD
> Senior Director and Head of Medical Writing, Hypercare & Enablement R&D, GSK

Where the AI conversation goes wrong

When I look at how generative AI is being discussed in medical writing, I see a genuine category error. Systems are being evaluated on how quickly and fluently they can produce text. But fluency is not the problem we needed to solve.

A system that generates impressive prose but still requires a writer to manually verify every number, every comparison, every interpretation, and every omission has not reduced the burden. It has moved it. The work now sits downstream, in QC, at a later stage, often under more pressure.

Speed that creates QC debt is not efficiency. It is risk deferred.

Regulators do not approve prose. They approve scientifically defensible evidence and the logic used to interpret it. That distinction matters more than the industry has yet fully absorbed.

A similar dynamic played out about fifteen or twenty years ago. A new review process was introduced specifically to make regulatory review faster and more manageable - the whole point was to reduce the burden on reviewers and shorten timelines. In practice, it did the opposite. The review process took longer, not less time. It added steps rather than removing them. The intention was sound. The reality was not. That is not unusual with technology solutions in

regulated environments - something can look entirely logical on paper and still not work when it meets the complexity of real workflows. What reviewers actually need is not a process that tells them what to conclude. They need something that makes it easy to see what is actually in the study - clearly, accurately, without having to fight through the document to find it. That is a much more specific problem than most AI tools are currently built to solve.

In conversations with senior medical writing leaders, the shift has been noticeable. The questions are different now. The ask is no longer how fast a system can draft. They are asking: can the output actually be trusted? Can I explain why it said what it said? Will it regenerate consistently if the TLFs are updated? How much checking does it still require?

Those are the right questions. They are questions about trust, not speed.

Generative AI has real value - just not where people think

To be clear: this is not a dismissal of generative AI. Used carefully and within appropriate limits, it has genuine value. Patient-friendly language adaptation, workflow navigation, exploratory analysis support, helping writers manage large information sets - these are real use cases where generative tools can help.

But there is a meaningful difference between assisting a writer and constructing regulatory evidence. Generative systems are probabilistic by design. That creates important challenges around reproducibility, lifecycle stability, and governed regeneration in regulated environments. That is not the same as producing accurate, traceable, regulatorily defensible text where the reasoning can be inspected and the output reproduced identically on request.

As long as the underlying processing remains probabilistic in nature, consistency becomes a distribution rather than a guarantee. Guardrails and monitoring can constrain or redistribute that variability, but they do not eliminate it at an architectural level.

Experienced medical writers do something that is far harder than summarising information. They interpret the decision of the clinical team as to what is clinically meaningful and what is not. They recognise when evidence is insufficient to draw a conclusion. They avoid over-interpretation. They understand what regulators are likely to scrutinise. And perhaps most importantly, they know what not to say.

> “Gen AI has huge potential, but it’s not a universal solution. In the context of regulatory writing, we need to be careful to not end up with lots of content just because content generation at a click of a button has become very easy.”
>
> Louise Martinsson, PhD
>
> Senior Director and Head of Medical Writing, Hypercare & Enablement R&D, GSK

That level of judgement requires governance, traceability, and reproducibility. It requires that identical inputs produce identical outputs. It requires that when data updates, regeneration is stable and explainable - not a new document that has to be reviewed from scratch.

What deterministic systems actually change

This is why structured, deterministic approaches will become increasingly important for high-stakes regulatory documentation - not as a replacement for human expertise, but as the right infrastructure to support it.

When a system is deterministic, reviewers are no longer spending their attention questioning whether numbers or interpretations may have shifted unexpectedly. They can focus on what they are actually there to do: scientific interpretation, contextual judgement, and regulatory thinking. That is where experienced medical writers create the most value. And it is where AI should be freeing them up to operate, not generating new checking work for them downstream.

Medical writers need to be in the room

AI systems in regulated medical writing are too often built externally and presented to writing teams afterward. That approach misses something important.

Experienced medical writers understand things that are very difficult to capture through software engineering alone. Where the review burden actually sits. How scientific interpretation really happens in practice. Why over-interpretation creates regulatory risk. Why consistency matters more than impressive prose. How a subtle shift in wording can materially change what a regulator understands.

If AI is going to genuinely help this profession, medical writers cannot simply be the recipients of tools built without their input. They need to help define how those tools are designed, validated, and governed. That has to be part of the conversation from the beginning.

The consequences of ignoring that are no longer hypothetical.

The regulatory stakes of getting this wrong are becoming clearer. In April 2026, the FDA issued its first-ever warning letter explicitly citing AI misuse as a compliance violation - addressed to Purolea Cosmetics Lab, a drug manufacturer in Michigan. The company had used AI agents to generate drug specifications, manufacturing procedures, and production records, without adequate human review of those outputs. Process validation requirements were missed entirely. When inspectors asked why, the company’s response was essentially that the AI had not flagged it as required. The FDA’s warning letter included a dedicated section titled “Inappropriate Use of Artificial Intelligence in Pharmaceutical Manufacturing” - a first for the agency. The consequences for the organisation were significant. This is not an isolated case. It is a signal. AI that operates without sufficient human oversight and transparency does not just create quality risk. In a regulated environment, it creates an existential risk for businesses. What that means for how we design and govern these systems is not a technical question. It is a governance question - and one the industry can no longer defer.

What Louise and I will be discussing at the EMWA hosted discussion

The future of AI in regulated documentation is unlikely to be autonomous drafting. It is far more likely to be governed evidence orchestration - systems that help humans construct, interrogate, regenerate, and defend regulatory evidence with transparency and consistency. The question is not whether AI will be part of regulated medical writing. It is whether the systems being built are designed to meet the actual standards that regulated evidence requires.

On 11 June, this will be the focus of a hosted discussion with Louise Martinsson with the European Medical Writers Association. We will be talking about exactly these questions - whether the focus on drafting speed has been solving the wrong problem, what trust actually requires in a regulated writing context, and where the real distinction between generative and structured AI lies in practice.

If you work in regulatory or medical writing and are thinking seriously about AI implementation, it will be a useful conversation. Not because all the answers are there, but because the right questions are finally being asked.

Join the conversation - Thursday 11th June 2026, 4PM CET
EMWA HOSTED DISCUSSION
The Role of AI in Medical Writing: Beyond Drafting Speed, Accuracy, QC Burden and What Actually Works

Click here to register