Clinical AI Checklist

10 Questions Every Physician Should Ask Before Using an AI Medical Scribe

A vendor-neutral evaluation tool from ZayedMD — By Ahmed Zayed, MBBCh, Physician-Engineer

---

How to use this checklist. AI medical scribes succeed or fail based on how well they fit your clinic, your evidence threshold, and your tolerance for review burden. Run through these questions in any vendor demo or trial — before any commercial conversation. Tick each item as the vendor answers it. The questions that remain unticked at the end are the questions that should remain open at decision time. This checklist is educational. It is not legal advice, medical advice, or a purchasing recommendation.

The 10 Questions

- [ ] **1. Auditability — can I trace a sentence in the note back to the source audio?**

  > A note that cannot be traced back to source audio is a generated artifact you cannot verify. For medico-legal exposure, peer review, and your own confidence in the record, auditability is a category-defining feature, not a luxury.

- [ ] **2. EHR integration — does it work on my specific EHR, in my configuration?**

  > Ask for a live demonstration on the EHR your clinic actually runs. The last mile is often a copy-paste or browser extension — not what "one-click integration" implies. You should see the integration before you agree to anything.

- [ ] **3. Setting-specific evidence — what data do you have from environments like mine?**

  > Ambient capture degrades in environments with overlapping speakers, monitor alarms, and ambient noise. Generalist accuracy figures from a quiet outpatient demo do not generalize to a pediatric room or an emergency department.

- [ ] **4. Published evidence — what peer-reviewed studies support your claims?**

  > The independent evidence base for AI scribes is small but growing. White papers and press releases do not count. A vendor whose only evidence is their own materials has not been independently studied at a level that should drive a purchasing decision.

- [ ] **5. Audio and data handling — where does the audio go, for how long, and under whose control?**

  > Ask explicitly: where is audio stored, in which jurisdiction, for how long, who has access, and is any of it used to train models? You need these answers before you can have an informed consent conversation with a patient about recording.

- [ ] **6. Patient disclosure and consent — what template does the vendor provide?**

  > Patient comfort with AI scribe recording holds when patients are clearly informed. Template language from the vendor is a starting point, not a substitute for your own legal review, but its presence or absence signals how seriously the vendor takes consent.

- [ ] **7. Error reporting and hallucination escalation — what happens when the note is wrong?**

  > Any tool built on a large language model carries a non-zero risk of hallucination. A vendor that cannot describe a hallucination-reporting pathway has not thought carefully enough about the failure mode that matters most.

- [ ] **8. Review workflow — what does end-of-day actually look like?**

  > Time saved is a function of review burden. A short, high-quality draft that needs minor edits is genuinely time-saving. A long, over-complete draft that buries the clinical reasoning costs you twice — once in reading and once in editing. Evaluate signal density, not completeness.

- [ ] **9. Reliability and fallback — what happens when your service goes down?**

  > Almost every AI scribe depends on stable, high-bandwidth connectivity. Rural deployments, mobile clinics, and any setting where continuity of care cannot tolerate a network outage need to understand the single-point-of-failure risk.

- [ ] **10. Twelve-month real cost — what is the total cost, including review time?**

  > ROI is review burden minus time saved, not the vendor's headline figure. If the vendor cannot describe twelve-month cost as a range that includes your editing time, they have not measured ROI in a way that should change your purchasing decision.

Red Flags — When to Walk Away

A single instance is a question. Two or more in the same conversation is a signal.

- ☐ **"Zero hallucinations."** No LLM-based tool has zero error rate.
- ☐ **"One-click integration"** with no live demo on your actual EHR.
- ☐ **No published, peer-reviewed evidence** beyond white papers or press releases.
- ☐ **Vague answers on audio storage, jurisdiction, or training use.**
- ☐ **No template patient disclosure language**, or language not reviewed by counsel.
- ☐ **No documented hallucination-reporting pathway** for clinically significant errors.
- ☐ **Fixed-figure ROI claims** not bounded by review time.
- ☐ **No published failure-mode or fallback documentation.**
- ☐ **Refusal to demonstrate auditability** on a real recording.
- ☐ **Urgency or exclusivity pressure** during the demo.

Physician Authentication — The Non-Negotiable

Every note is a draft until you have read it, corrected it, and authenticated it.

You are the author of the medical record. No AI scribe product on the market in 2026 removes the requirement for full physician review. The moment your scribe leaves your control, the conversation shifts to consent, privacy, and medico-legal exposure.

This checklist accompanies the ZayedMD article *AI Medical Scribes for Doctors: What They Actually Do, What to Check, and Where the Risks Are.* The 10 questions distill the four-question diagnostic framework from that article into a single sheet you can take into a vendor meeting. For the broader landscape — the three market categories, the four vendor archetypes, and the structural risks — read the full article.

Read the full article at ZayedMD.com [URL to be inserted after publication]

This checklist is educational. It is not legal, medical, or purchasing advice. Vendor names are intentionally absent. Free to print, share with colleagues, and use in your own evaluation process. Do not republish, white-label, or repackage without attribution.