THE ATTENDING

47

Documentation Failure Modes of Clinical AI

A classified reference for clinicians who already use AI and want a defensible way to do it.

Clinician-built · Safety-first
Designed around human review and data minimization

Part of **The Attending Standard**

Version 1.0 · August 2026

How to use this document

This is not a list of things AI gets wrong. It is a classification of the specific ways a clinical document can be damaged between your dictation and your signature so that you can name the failure, know which protocol catches it, and stop looking for the rest.

Each entry gives you five things: what the failure looks like, a synthetic example, which of **The Four Boxes** the task belongs in, an Attending Risk Level, and the protocol or test that catches it.

The six families

F	**Fabrication** - the model asserts what was never there	12
O	**Omission** - what was there silently disappears	8
D	**Distortion** - what was there is changed in transit	9
A	**Attribution** - who said or decided it becomes unclear	6
P	**Privacy** - identifiable information leaves your environment	6
G	**Governance** - the failure is in the workflow, not the text	6

The reading order that actually helps

Do not read all 47 today. Read family **F** before your next documentation session and family **P** before you next paste anything into a model. Those two cover the failures you are most likely to meet this week. The rest will still be here.

Who wrote this

The Attending is built by a group of international physicians and designed for clinical responsibility. The taxonomy below was assembled from documentation review, payer-side appeal review, and structured critique by practicing clinicians across emergency medicine, internal medicine, primary care, cardiology and psychiatry. It carries no individual byline by design: a protocol should be judged on whether it holds, not on who signed it.

> **One thing this document will not do.** Nothing you read, download or buy makes an AI tool HIPAA-compliant, GDPR-compliant, or approved by your institution. Compliance is a property of your environment, your vendor contract and your local policy - not of a file. Anyone telling you otherwise is telling you something untrue.

F

FABRICATION

The model asserts something that was never in your source.

The dominant failure class, and the hardest to see. Fabrications are fluent, specialty-appropriate, and shaped exactly like true statements.

12 failure modes

F01 The Completed Exam

CHECK BLOCK AMBER

You dictate three findings; the model returns a full systems examination. Nothing false was stated - everything absent was invented.

```
Source: "Chest clear. Abdomen soft."
Output: "...HEENT normal. Neck supple, no JVD.
 Heart RRR, no murmurs. Extremities without edema."
```

CAUGHT BY AS-06 check 1 - invented finding

F02 The Confident Dose

90-SEC AUDIT RED

You said "start an NSAID." The model produces a specific agent, dose, route and frequency you never specified.

```
Source: "start an NSAID"
Output: "ibuprofen 600 mg PO TID with food"
```

CAUGHT BY AS-06 check 2 - invented number

F03 The Phantom Negative

CHECK BLOCK RED

The model asserts a pertinent negative that was never elicited. Pertinent negatives carry medicolegal weight precisely because they imply you asked.

```
Source: (no mention of chest pain)
Output: "Denies chest pain, dyspnea, or palpitations."
```

CAUGHT BY AS-06 check 1 - invented finding

F04 The Invented Timeline

90-SEC AUDIT RED

Times, durations and intervals appear that you never supplied. A fabricated onset time changes eligibility for time-dependent therapy.

```
Source: "symptoms started this morning"
Output: "symptom onset at approximately 07:30, roughly 4 hours prior to arrival"
```

CAUGHT BY AS-06 check 2 - invented number

F05 The Borrowed Vital

90-SEC AUDIT RED

Plausible vital signs appear in a note where none were dictated. They look like they came from the chart. They came from the training distribution.

```
Source: "vitals stable"
Output: "BP 128/76, HR 82, RR 16, SpO2 98% on RA, T 36.8C"
```

CAUGHT BY AS-06 check 2 - invented number

F06 The Textbook Completion

CHECK BLOCK RED

One element of a classic triad is present; the model supplies the other two because they usually travel together.

```
Source: "fever"
Output: "fever, neck stiffness, and photophobia"
```

CAUGHT BY AS-06 check 1 - invented finding

F07 The Confabulated Citation

90-SEC AUDIT RED

A guideline, year, society, or recommendation class is invented - or a real guideline is cited with the wrong content.

```
Output: "per the 2023 ACC/AHA guideline, Class IIa"
(the recommendation exists; the class and year do not)
```

CAUGHT BY AS-06 check 1 - invented finding

F08 The Manufactured Consent

FOUR BOXES RED

The output asserts that risks, benefits and alternatives were discussed and that the patient agreed. This is a legal claim about an event.

```
Output: "Risks, benefits and alternatives were discussed;
 the patient verbalized understanding and consented."
```

CAUGHT BY AS-02 - DON'T box (consent)

F09 The Invented Follow-Up

90-SEC AUDIT AMBER

A follow-up interval appears that no one decided. The patient reads it. The next clinician inherits it.

```
Source: "arrange follow-up"
Output: "Follow up in clinic in 2 weeks."
```

CAUGHT BY AS-06 check 2 - invented number

F10 The Ghost Consult

CHECK BLOCK RED

The output states that a service was contacted, or that advice was received, when neither happened.

```
Source: "may need cardiology"
Output: "Cardiology was consulted and advised outpatient follow-up."
```

CAUGHT BY AS-06 check 1 - invented finding

F11 The Fabricated Failed Therapy

SECOND SIGNATURE RED

In a prior-authorization appeal, the model invents prior treatments, doses or durations to satisfy a step-therapy criterion. This is the highest-consequence fabrication in the entire set.

```
Output: "Patient failed agent A at therapeutic dose for 10 weeks
 and agent B with documented intolerance."
```

CAUGHT BY AS-01 - Second Signature Rule

F12 The Synthetic Reassurance

90-SEC AUDIT RED

Patient-facing text acquires a comforting sentence that overstates certainty or promises an outcome.

```
Output: "This is nothing to worry about and will resolve on its own."
```

CAUGHT BY AS-06 check 3 - hardened hedge

0

OMISSION

Something that was in your source silently disappears.

Fabrication hides inside completeness; omission hides inside fluency. No CHECK block catches an omission - only you do, and only if you look for the gap rather than reading the text.

001 The Silent Drop

90-SEC AUDIT AMBER

A finding you stated does not appear in the output. The note reads perfectly and is missing a fact.

```
Source: "...and a 2 cm laceration to the left forearm"
Output: (no mention of the laceration)
```

CAUGHT BY AS-06 check 5 - would this read the same for the next patient

002 The Vanished Uncertainty

90-SEC AUDIT RED

Hedges are removed in summarization. "Possibly consistent with" becomes "consistent with."

```
Source: "appearances may represent early changes"
Output: "imaging shows early changes"
```

CAUGHT BY AS-06 check 3 - hardened hedge

003 The Missing Red Flag

90-SEC AUDIT RED

The single worrying detail is compressed away because it was one clause in a long narrative.

```
Source: "...pain woke him from sleep..."
Output: (nocturnal onset absent)
```

CAUGHT BY AS-06 check 5 - cloned-note test

004 The Lost Pending

SECOND SIGNATURE RED

Pending investigations vanish from a handoff. This is the omission most likely to reach a patient.

```
Source: "CT pending, blood cultures sent"
Output: "Workup underway."
```

CAUGHT BY AS-01 - Second Signature Rule

005 The Truncated Medication List

90-SEC AUDIT RED

Long lists are silently shortened. The model does not tell you it stopped early.

```
Source: 14 medications
Output: 9 medications, no ellipsis, no note
```

CAUGHT BY AS-06 check 2 - invented number

006 The Dropped Allergy

SECOND SIGNATURE RED

An allergy or adverse reaction does not carry through into a referral, discharge summary, or patient letter.

```
Source: "penicillin - rash"
Output: (allergy section absent entirely)
```

CAUGHT BY AS-01 - Second Signature Rule

007 The Absent Denominator

90-SEC AUDIT AMBER

A number appears without the base it was drawn from, making it uninterpretable.

```
Output: "3 episodes" (of how many days? how many attempts?)
```

CAUGHT BY AS-06 check 2 - invented number

008 The Unstated Gap

CHECK BLOCK AMBER

Nothing records that an element was not assessed. The reader cannot distinguish "normal" from "not examined."

```
Correct: "Cranial nerves: NOT DOCUMENTED"
Failure: silence
```

CAUGHT BY AS-06 check 4 - every wildcard has real content

D

DISTORTION

A fact that was present is changed in transit.

Distortions are the most dangerous class per unit of text, because the output remains superficially faithful to your input while inverting its meaning.

9 failure modes

D01 The Smoothed Chronology

90-SEC AUDIT RED

Events given out of order, with gaps, are returned as a clean linear sequence. The gaps were clinically meaningful.

```
Source: fragmented history with two unwitnessed periods
Output: continuous narrative, no gaps marked
```

CAUGHT BY AS-06 check 5 - cloned-note test

D02 The Upgraded Certainty

90-SEC AUDIT RED

Diagnostic language is promoted one level. "Query" becomes "suspected"; "suspected" becomes "consistent with."

```
Source: "query pneumonia"
Output: "findings consistent with pneumonia"
```

CAUGHT BY AS-06 check 3 - hardened hedge

D03 The Laterality Flip

90-SEC AUDIT RED

Left and right swap. This survives review because the sentence is otherwise perfect.

```
Source: "left knee"
Output: "right knee"
```

CAUGHT BY AS-06 check 5 - cloned-note test

D04 The Unit Slide

90-SEC AUDIT RED

Units change between source and output - mg and mcg, kg and lb, mmol/L and mg/dL.

```
Source: "levothyroxine 50 mcg"
Output: "levothyroxine 50 mg"
```

CAUGHT BY AS-06 check 2 - invented number

D05 The Rounded Value

90-SEC AUDIT RED

A value is rounded across a clinically meaningful threshold.

```
Source: "potassium 5.4"
Output: "potassium 5"
```

CAUGHT BY AS-06 check 2 - invented number

D06 The Severity Drift

90-SEC AUDIT AMBER

Severity descriptors move a notch in either direction, changing triage, coding, and payer response.

```
Source: "moderate"
Output: "severe"
```

CAUGHT BY AS-06 check 3 - hardened hedge

D07 The Negation Collapse

90-SEC AUDIT RED

A double negative or a hedged exclusion is flattened into an affirmative exclusion.

```
Source: "cannot be excluded"
Output: "excluded"
```

CAUGHT BY AS-06 check 3 - hardened hedge

D08 The Paraphrased Criterion

SECOND SIGNATURE AMBER

In an appeal, the payer's criterion is restated in your own words. The reviewer is matching against their document, not your summary - this is why clinically correct appeals lose.

```
Correct: paste the criterion verbatim, answer beneath it
Failure: "their policy requires prior therapy"
```

CAUGHT BY AS-01 - quote the criterion verbatim

D09 The Reading-Level Overshoot

RISK LEVEL AMBER

A patient handout is generated above the stated literacy level, or reverts to clinical register midway.

```
Requested: 6th-grade reading level
Output: "myocardial ischemia secondary to..."
```

CAUGHT BY AS-05 - AMBER: verify before it goes anywhere

A

ATTRIBUTION

Who said, decided, or recommended something becomes unclear.

Attribution failures rarely change a clinical fact. They change who is accountable for it - which is the only question that matters when a record is examined later.

6 failure modes

A01 The Orphan Recommendation

SECOND SIGNATURE RED

A suggestion generated by the model appears in the record as the clinician's own reasoned decision, with no decision behind it.

```
Output: "Plan: outpatient management."
(no clinician ever weighed admission)
```

CAUGHT BY AS-01 - Second Signature Rule

A02 The Assumed Author

SECOND SIGNATURE AMBER

A draft is signed without the disclosure the institution's policy requires for AI-assisted documentation.

```
Policy: AI-assisted drafts must be flagged
Practice: nothing recorded
```

CAUGHT BY AS-01 - Second Signature Rule

A03 The Merged Voice

SECOND SIGNATURE AMBER

The patient's own words and the clinician's interpretation become indistinguishable in the same paragraph.

```
Output: "The patient reports the pain is cardiac in origin."
(interpretation attributed to the patient)
```

CAUGHT BY AS-01 - Second Signature Rule

A04 The Silent Context Bleed

PHI FIREWALL RED

In one conversation covering several cases, detail from an earlier case appears in a later output.

```
Case 1 mentions a pacemaker.
Case 3's output mentions a pacemaker. Case 3 has none.
```

CAUGHT BY AS-03 - PHI Firewall

A05 The Uncredited Guideline

90-SEC AUDIT AMBER

A recommendation is stated as settled fact with no source, so no one can check it later.

```
Output: "Standard of care is 7 days of therapy."
```

CAUGHT BY AS-06 check 1 - invented finding

A06 The False Consensus

90-SEC AUDIT AMBER

"Guidelines recommend" is asserted where guidelines conflict or where the evidence is genuinely contested.

```
Output: "Guidelines recommend screening from age X."
(two major bodies disagree)
```

CAUGHT BY AS-06 check 3 - hardened hedge

P

PRIVACY

Identifiable information leaves the controlled environment.

None of these are cured by a purchase. They are cured by a protocol you run before you type. De-identification reduces one specific risk; it does not make any tool compliant.

6 failure modes

P01 The Mosaic Leak

PHI FIREWALL RED

No single identifier is present, but the combination singles out one person in a catchment area.

```
"34-year-old, Fontan circulation, admitted to our
 cardiology service last week"
```

CAUGHT BY AS-03 - the mosaic test

P02 The Rare Diagnosis Beacon

PHI FIREWALL RED

Rarity itself is an identifier - and the more interesting the case, the more you want a second opinion on it.

```
Any presentation your department would recognize
from the description alone
```

CAUGHT BY AS-03 - the mosaic test

P03 The Habitual Shorthand

PHI FIREWALL RED

Internal shorthand carries identity: bed numbers, MRN fragments, "the gentleman from bay four on Tuesday."

```
"bed 12" · "MRN ending 4471" · "Tuesday's transfer"
```

CAUGHT BY AS-03 - de-identification protocol

P04 The Date Residue

PHI FIREWALL RED

Dates more specific than the year survive de-identification because they do not feel like identifiers.

```
"admitted 3 March, procedure on the 5th"
```

CAUGHT BY AS-03 - Safe Harbor identifier 3

P05 The Round-Trip Re-Identification

PHI FIREWALL RED

The case is properly de-identified, then the key is pasted back into the same conversation to restore names.

```
Prompt 1: de-identified case
Prompt 2: "the patient's name is..."
```

CAUGHT BY AS-03 - keep the map offline

P06 The Persistent Context

PHI FIREWALL RED

Identifiable text entered earlier in a long session remains in context and contaminates later "clean" prompts.

```
Turn 2 contained PHI.
Turn 40 is a new case in the same window.
```

CAUGHT BY AS-03 - tool-tier classification

G

GOVERNANCE

The failure is in the workflow, not in the text.

These are the failure modes that produce a bad outcome even when every individual output was correct. They are also the ones a practice can eliminate in a single meeting.

6 failure modes

G01 The Tier Crossing

FOUR BOXES RED

A clinician who has safely drafted referral letters for months asks the model to "help think through" a capacity assessment. Same tool, same interface, different category of decision - and nothing signals the crossing.

```
Month 1: referral letters (DRAFT)
Month 6: capacity assessment (DON'T)
```

CAUGHT BY AS-02 - run the sort before you type

G02 The Skim Signature

SECOND SIGNATURE RED

Reading degrades to skimming on the exact shifts where output volume is highest. The Second Signature Rule fails silently, without any decision to abandon it.

```
Chart 3 of 11: read
Chart 9 of 11: signed
```

CAUGHT BY AS-01 - read every word

G03 The Undocumented Tool

RISK LEVEL RED

An approved workflow is executed in an unapproved tool, because the approved one was slower that day.

```
Approved: institutional environment
Used: personal account, personal device
```

CAUGHT BY AS-05 - approved-tool policy

G04 The Silent Policy Drift

RISK LEVEL AMBER

Institutional or vendor policy changes; the workflow built against the old policy continues unchanged.

```
Policy updated in March.
Workflow last reviewed in October.
```

CAUGHT BY AS-05 - scheduled review

G05 The Unlogged Incident

SECOND SIGNATURE RED

An AI-assisted error reaches a record and is corrected quietly, with no incident record. "We had no process" is a materially worse position than "we had a process and here is what we did."

```
Error found. Note amended.
Nothing recorded anywhere.
```

CAUGHT BY AS-01 - incident procedure

G06 The Orphaned Template

RISK LEVEL AMBER

A phrase file or prompt outlives the clinician who validated it, and is inherited by someone who never checked its assumptions.

```
Written for one specialty and one EHR.
In use across three.
```

CAUGHT BY AS-05 - versioned, named owner

The six protocols that catch them

Every failure mode in this document is governed by one of six named protocols, cited by clause number. All six are published free and in full at **theattending.net/free**.

AS-01

The Second Signature Rule

No AI-generated content enters a record, reaches a patient, or goes to a payer until a licensed clinician has read every word and taken responsibility for it. Not skimmed. Read.

AS-02

The Four Boxes

Before anything is typed, the task goes in exactly one box. **DON'T** - diagnosis, treatment, consent, capacity. **DOUBLE-CHECK** - coding, guidelines, anything with a number. **DRAFT** - notes, letters, appeals. **DELEGATE** - reformatting, reading level, admin.

AS-03

The PHI Firewall

A three-tier classification of tools by the data they may receive, a de-identification protocol, and the mosaic test - because a rare diagnosis in a small town identifies someone no matter how many fields you strip.

AS-04

The CHECK Block

Every prompt ends by forcing the model to list what it added that you did not supply. You are not asking it to be honest - you are requiring a diff. Trust is not a prompt technique.

AS-05

The Attending Risk Level

Every task carries one of three marks. **GREEN** use as written · **AMBER** verify the output against source before it goes anywhere · **RED** never delegate this. A library without triage is a pile.

AS-06

The Ninety-Second Audit

Five checks between the draft and your signature: **1.** did it invent a finding **2.** did it invent a number **3.** did it harden a hedge into certainty **4.** does every wildcard have real content **5.** would this read the same for the next patient. If yes to the last one, it is a cloned note.

Notices

> **Not clinical advice.** This document is an educational reference about documentation workflow and information handling. It is not clinical advice, is not a substitute for clinical judgment, and must not be used to drive clinical decisions. It is not a medical device and makes no diagnostic or treatment recommendation.

All examples are synthetic

Every example in this document was written for illustration. None describes a real patient, and none is derived from, based on, or adapted from a real encounter.

Compliance

Nothing in this document, and nothing sold by The Attending, makes any AI tool compliant with HIPAA, UK GDPR, EU GDPR, or any other framework. Compliance depends on your environment, your vendor agreements, and your institution's policy. Verify locally before you use any AI tool clinically. Where a Business Associate Agreement is required, a purchase does not create one.

License

Corrections

If you find an error, or a failure mode that belongs in this taxonomy and is not here, write to **support@theattending.net**. Corrections are published with the version number they were fixed in. A reference that teaches verification has to be verifiable itself.

- The Attending

Built by a group of international physicians. Designed for clinical responsibility.

The Attending Standard · Version 1.0 · August 2026 · theattending.net