2026 APIC MINNESOTA FALL CONFERENCE

Artificial Intelligence in Infection Prevention

Where the technology is today — and how to think about it



Brandon Staunton

Software Developer · Infection Prevention & Control

INTRODUCTION

Where I'm coming from

In the profession — from the software side

I'm an APIC member — I just come at infection prevention from the technology-delivery side

Background

- **11 years at Epic — the last 10 in Infection Prevention**
- Building HAI surveillance and prevention tools
- Hands-on with LLM-assisted chart abstraction and a clinical-support agent

BEFORE WE START

Disclosure



I develop infection-prevention software for a living.

That gives me a builder's view of these tools.

BEFORE WE BEGIN

Learning objectives

1

LEVEL 1 · GENERAL-PURPOSE TOOLS

Identify practical, currently available AI tools that can support routine infection prevention tasks such as policy drafting and documentation.

2

LEVEL 2 · LLM-ASSISTED SURVEILLANCE

Describe how large language models can assist with surveillance activities.

3

LEVEL 3 · AGENTS FOR PREVENTION

Explain how AI agents differ from general-purpose and analytic tools, and recognize the opportunities and limitations of agent-based approaches.

IP Domains: Education & Research · Management & Communication

THE PROBLEM THAT NEVER SLEEPS

1 in 31

hospital patients has a healthcare-associated infection on any given day.

*You know this better than anyone: surveillance is still largely **manual**, largely **retrospective**.*

~30%

of HAIs are urinary tract infections — the single largest category

15–25%

of hospitalized patients get a urinary catheter during their stay

Every day

a device stays in adds measurable, compounding risk

Source: CDC, National HAI Progress Report

THE OTHER HALF OF THE PROBLEM

4 in 5

hospitals are understaffed for infection prevention — and understaffing tracks with higher infection rates.

*The work keeps growing. The workforce **isn't***

40%+

of the IP workforce is over 55 — a retirement wave is already underway

1 in 4

infection-prevention programs has a position sitting vacant

43.6 hrs

the average IP work-week — persistently over 40, across settings

The profession is fighting back: APIC's **IP Staffing Calculator** turns these gaps into a facility-specific, data-driven case for the staff you need.

Sources: APIC IP Staffing Calculator study (AJIC, 2024); APIC MegaSurvey; IP workload survey (AJIC, 2025)

THE FIELD, IN ITS OWN WORDS

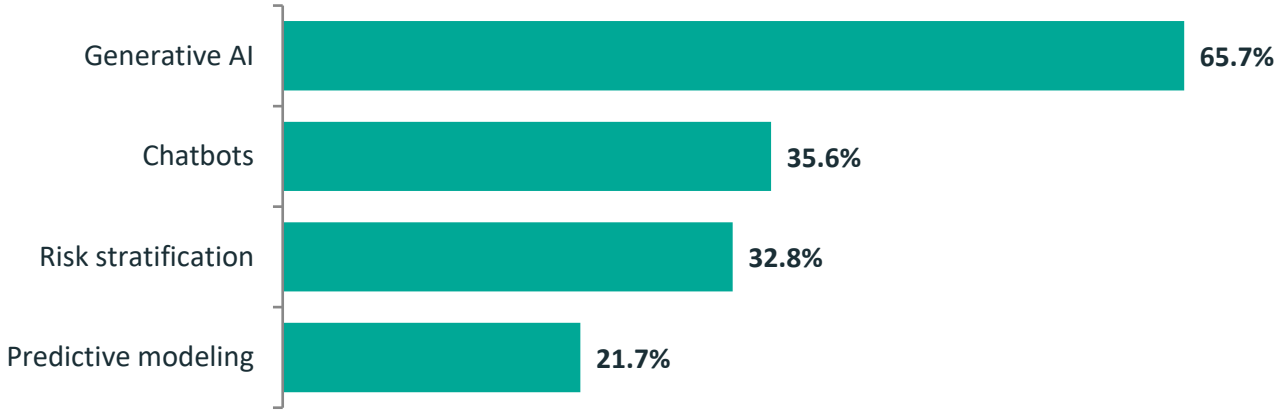
What the profession just told APIC

12.6%
of IPs currently use AI in their work

80%
believe AI will be useful to their practice

75%
want to learn more about it

HOW CURRENT USERS APPLY AI



Generative AI = drafting policies & education, synthesizing research (Level 1). Percentages are of the 12.6% who use AI.

WHAT HOLDS EVERYONE ELSE BACK

- Lack of organizational readiness
- Limited access to tools
- Concerns about accuracy
- Data privacy
- Fear of job displacement

Source: Kon et al., 2025 APIC MegaSurvey. Am J Infect Control (2026); n = 3,220 responding to the AI section.

FRAMING

Three levels of AI as I see them



LEVEL 1

General-purpose tools

Chat assistants anyone can use today — draft, summarize, compare, explain. You steer; it responds.



LEVEL 2

LLM-assisted surveillance

Models read unstructured charts and propose answers for surveillance. It suggests; you adjudicate.



LEVEL 3

Agents for prevention

Systems that take steps inside a workflow to act during care, not just report after. It acts; you supervise.

more capable → more autonomy → more oversight required

FRAMING

AI isn't new to healthcare — access to it is

Rule-based logic

OPAs, NHSN rule engines, reporting queries



Machine learning

risk scores, semi-automated surveillance



TODAY'S FOCUS

LLMs & agents

read plain language, act in a workflow



For decades — deterministic: same input, same answer.

The new slice — non-deterministic, like us.

FRAMING

Why this is happening now

Language models got good at the exact thing infection prevention drowns in:

reading unstructured clinical text.

Notes, not fields

Much of what defines an HAI lives in free-text notes, not tidy structured data.

No code required

The interface is plain language. An IP can use these tools without a data team in the loop.

Fast enough to act

Responses arrive in seconds — fast enough to fit inside a clinical workflow.

FRAMING

What it's good at — and what it isn't

Strong at

- Reading and summarizing long text
- Drafting a first version of almost anything
- Comparing documents and flagging differences
- First-pass triage of a large pile of records

Not a source of truth

- Can state something false with full confidence
- Doesn't “know” your definitions unless given them
- Isn't the one making the final surveillance call — today
- Needs PHI handled under your own policies

The rule for the whole hour: the human stays in the loop. AI proposes — a person decides.



LEVEL ONE

General-purpose tools

1

The everyday AI that any infection preventionist can use on Monday morning

A tangible example: make it readable

The prompt: *“Rewrite this handout for a patient at a 6th-grade reading level — keep the medical facts.”*

BEFORE

~GRADE 11

“You have been placed on contact precautions due to Clostridioides difficile colonization. Personnel will don gowns and gloves, and meticulous hand hygiene with soap and water is required given the sporicidal limitations of alcohol-based antiseptics.”

AFTER — ONE PROMPT LATER

~GRADE 6

“You have a germ called C. diff, so we're taking extra steps to keep it from spreading. Staff will wear a gown and gloves in your room. Everyone should clean their hands with soap and water — for this germ, hand gel doesn't work as well.”

In a JMIR study, LLMs simplified 60 patient handouts from ~10th- to 6th–8th-grade reading level with accuracy preserved (2025). You confirm the clinical facts; the tool does the rewriting.

LEVEL 1 · GENERAL-PURPOSE TOOLS

Three things you could try this week

1

Precautions & exposure questions

The isolation and exposure-follow-up questions that flood your inbox — ask a general assistant for a fast first pass, then confirm against your own policy.

2

Draft a teaching scenario

Turn a real risk into a quick quiz or competency check for a nurse or physician huddle

3

Literature fast-scan

Drop in a handful of abstracts on a prevention question. Ask for a plain-language synthesis and the points of disagreement

Illustrative and reproducible with any general-purpose assistant. Peer-reviewed work found commercial LLMs answered real IPC inquiries at ~99% accuracy (GPT-4; Abosi et al., ICHE 2025) — still a first pass that needs expert review.

LEVEL 1 · GENERAL-PURPOSE TOOLS

AI for education & communication

Gamify a policy

Adults retain more from active recall and scenarios than from reading a policy

— and an LLM is the fast way to build it.

Turn a policy or competency into a quiz, a “spot the breach” scenario, or a Jeopardy-style board in minutes. You refine; it drafts.

Make the visual

Draft a hand-hygiene infographic, a unit poster, or an illustration for a handout or huddle slide

— a first version in seconds, not a design ticket.

The caveat: check every fact and label, and follow your organization's rules on AI-generated images.

LEVEL 1 · GENERAL-PURPOSE TOOLS

A few simple guardrails

Protect PHI

Public consumer chatbots are not a place for identifiable patient data. Use only tools cleared under your organization's own data-use and BAA policies.

Assume it can be wrong

Treat every output as a confident draft, not a fact. If a claim matters, verify it against the primary source before you act.

You own the output

The tool narrows and drafts; a qualified person makes the call and owns it



LEVEL TWO

LLM-assisted surveillance

2

Where the field is heading fast: models that read the chart and propose the answer, so the IP shifts from data entry to adjudication.

The problem: manual chart abstraction

Surveillance case-finding means reading charts — often many, often long, often on a deadline. Three costs come with that:

Time

AHRQ: at ~20 min/chart and ~1% positivity, roughly 33 hours of review for every SSI found.

Variability

Two reviewers, two readings — consistency is hard on nuanced definitions.

Backlog

When volume spikes, review falls behind exactly when it matters most.

5h 8m

one community-hospital IP's measured daily time on HAI data collection and reporting alone.

Sources: AHRQ SSI surveillance analysis; RWJ/APIC 2015 case report

LEVEL 2 · LLM-ASSISTED SURVEILLANCE

How it works — conceptually

Model reads the chart

The LLM ingests the unstructured notes and pulls out what's relevant to the criteria.

It proposes a value

It returns a suggested answer with the supporting text it relied on.

A person reviews & decides

The IP confirms, corrects, or rejects. Their judgment is key

And it closes the loop: a thumbs-up or thumbs-down with a note feeds back, so the tool **improves over time**.

In an AJIC pilot, an LLM correctly applied NHSN CLABSI/CAUTI definitions on training scenarios given clear prompts (Wiemken & Carrico, 2024).

LEVEL 2 · ESCALATING EXAMPLES

Start simple: one field

IN THE EXAM NOTE

"…febrile to 38.6°C. Abdomen soft, mildly tender suprapubically. Back: ***positive CVA tenderness on the right****; left CVA nontender."*

→

THE FIELD IT FILLS

CVA pain / tenderness

Yes — right

evidence: *"positive CVA tenderness on the right"*

proposed for your review — the model never closes the case

One field, one sentence of evidence — you confirm in seconds. Now scale it up.

LEVEL 2 · ESCALATING EXAMPLES

Now the whole determination: an SSI

Yes	Within the 30-day surveillance window	evidence: *post-op day 12*
Yes	Purulent drainage from the incision	evidence: *“purulent drainage expressed”*
Yes	Organism from a wound culture	evidence: *wound culture: S. aureus*
No	Fever	evidence: *“afebrile”*
Yes	Surgeon's diagnosis of SSI	evidence: *A/P: “surgical site infection”*

Proposed: **meets Superficial Incisional SSI** — you adjudicate. The model assembles the criteria; the call stays yours.

In a Stanford evaluation, an LLM read 5,299 surgical charts against NHSN SSI definitions — catching nearly all true SSIs, with IPs still the adjudicators (ICHE, 2026).

LEVEL 2 · WHY IT WORKS

How it reads: meaning, not words

Keyword: hit

“No fever.”

Negated — the finding is absent

Keyword: hit

“Call provider if signs of fever develop.”

Conditional — an instruction, not a finding

Keyword: hit

“Fever to 39°C overnight.”

Present — the only real one

Keyword search flags all three the same. An LLM reads them the way you would — negated, conditional, present — **which is what makes the field- and SSI-level reads possible.**

LEVEL 2 · LLM-ASSISTED SURVEILLANCE

What it changes for the IP's day

BEFORE

- Read every candidate chart end to end
- Hunt for the criteria buried in the notes
- Most of the hour spent on retrieval
- Consistency depends on who's reviewing

AFTER

- Start from a proposed answer + its evidence
- Confirm, correct, or reject — quickly
- Most of the hour spent on judgment
- A consistent first read across every case

≈75 min → ≈5 min: a first-pass LLM screen vs. manual CLABSI review (Stanford + Univ. of Maryland, ICHE 2025). The IP still decides.

The shift isn't fewer people — it's the same people spending their expertise where it matters most.

LEVEL 2 · THE EVIDENCE

A closer look

75 → 5 min

per case — IP review vs. LLM first pass

80%

sensitivity — caught 16 of 20 CLABSIs

35% → 80%

specificity — once it had the full chart

Rodriguez-Nava et al. — secure GPT-4 on 40 real CLABSI cases (Stanford + Univ. of Maryland, ICHE 2025).

LEVEL 2 · LLM-ASSISTED SURVEILLANCE

How to make this class of tool prove itself

Measured against a gold standard

Agreement with expert human abstraction on YOUR cases — not a demo dataset.

Shows its work

Every proposed answer should point to the chart text it used.

Tested for the failure that matters

For surveillance, a missed case is worse than a false flag.

Re-checked over time

Notes and documentation drift. Performance measured once is not performance guaranteed.

LEVEL 2 · LLM-ASSISTED SURVEILLANCE

What to watch for

Only as good as the note

If the documentation is thin, wrong, or missing, the model inherits the gap.

Edge cases are hard

The thorniest, judgment-heavy cases — exactly the ones you'd most want help with — are where it's least reliable.

It can sound certain when it's wrong

Fluent, confident phrasing hides errors. A wrong answer reads exactly like a right one.

It never removes accountability

The facility still owns every number it reports. The tool changes the workload, not the responsibility.



LEVEL THREE

Agents for prevention

3

Think of a small army of very literal interns who never sleep - each with one job, a cost, and a standard to meet.

Understanding agents

A HUMAN INTERN		AN AI AGENT
Job description	→	The prompt — its written instructions
Salary + benefits	→	Token cost — cents, not a salary
You check their work at first	→	Human-in-the-loop review
Works one shift	→	Never sleeps — nights, weekends, holidays

The differences are the point — near-zero marginal cost, tireless, and it scales. But it still needs the job description, the oversight, and your judgment.

LEVEL 3 · AGENTS FOR PREVENTION

From counting harm to preventing it

Admin & surveillance (Levels 1–2)

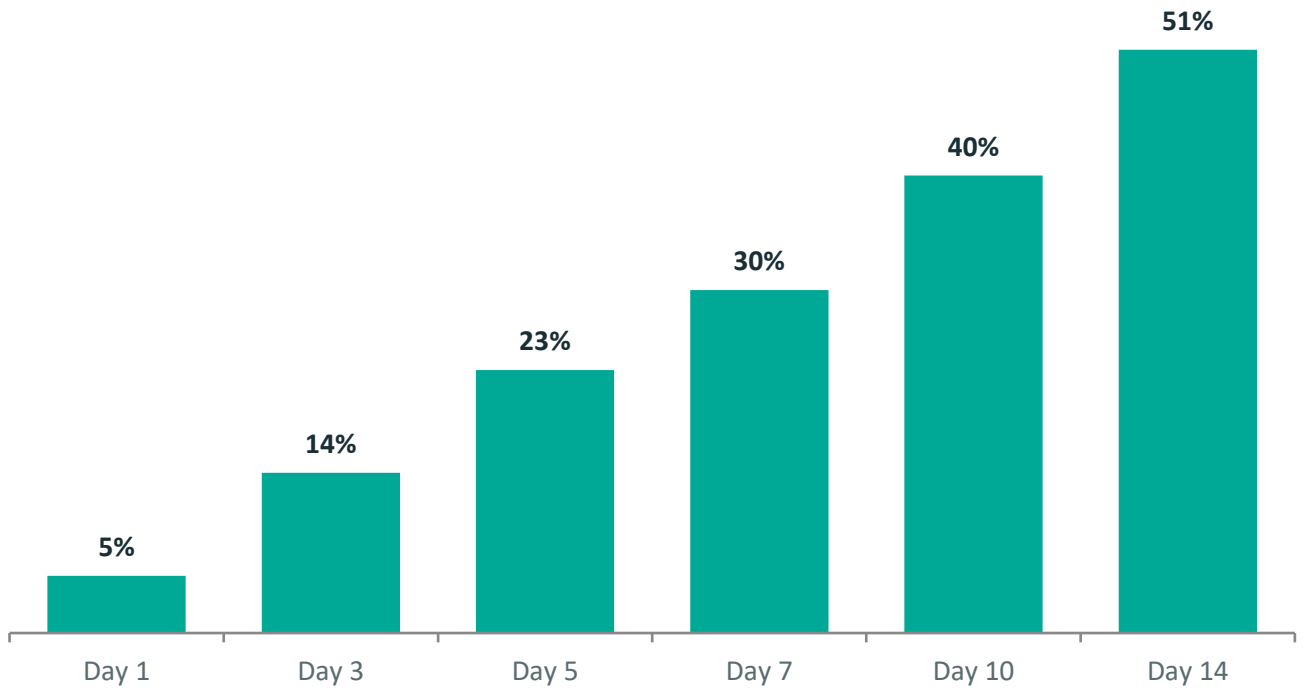
Helps with the work around care — drafting, education, reading charts — mostly after the fact.

Prevention (Level 3)

Acts inside the encounter, while the patient is still at risk, so the harm can be avoided instead of tallied.

The whole point of Level 3: move the moment of usefulness earlier — from the report after discharge to the decision at the bedside.

Why the catheter is a natural fit for an agent



Illustrative: cumulative risk at ~5%/day, the midpoint of the published 3–7% daily range.

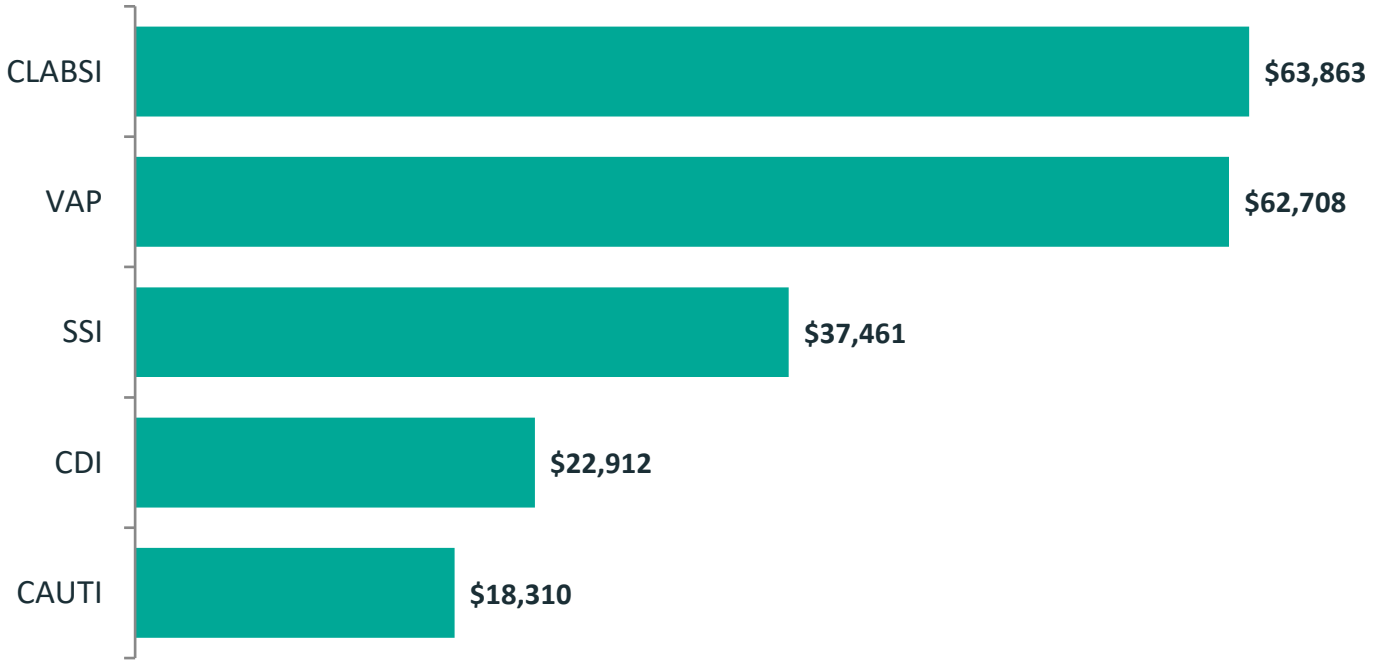
The gap isn't knowing — it's watching every day.

Risk climbs **3–7%** per catheter-day, so you pull it early. Nothing new to an IP. But day 5 might be a Saturday when we don't staff IPs.

$18,310 added cost per CAUTI (AHRQ) — every catheter-day saved is safety and dollars.

Sources: CDC / IDSA CAUTI guidance; AHRQ HAC cost estimates

The stakes, in dollars



Source: AHRQ, additional inpatient cost per hospital-acquired condition, inflation-adjusted.

Prevention is where patient safety and the bottom line point the same way.

It's also the case an IP can take to the C-suite — a Level-3 assistant that prevents infections is measurable in dollars, not just SIRs.

Inside a prevention agent

Inputs

- Clinical notes
- Line & device data
- Flowsheets — I&O, temps
- Orders

the same places you'd look

The agent

Reads the inputs against one clear rule, on a schedule. Bounded scope — nothing open-ended.

"Catheter still in, with no documented indication?"

Narrative + recommendation

"Mr. R in Bed 12 has had a urinary catheter for 5 days. There's no documented indication since admission, so it looks ready to come out — worth raising for removal on rounds today."

A chatbot waits, answers, and stops. An agent watches, evaluates, and hands a person the decision **— it proposes, you decide.**

What we tell the intern

Your job	Each morning, review every patient with an indwelling urinary catheter.
What to check	Is there a documented indication today — surgery, strict output monitoring, retention?
When to speak up	No indication documented, and the catheter has been in two or more days.
How to respond	One plain sentence to the nurse — patient, catheter-day, “consider removal” — and cite the note you used.
What you never do	Never guess. Every recommendation cites the exact note it's based on.

The “brains” is mostly a clear, bounded set of instructions — not mystery.

LEVEL 3 · INSIDE THE AGENT

It shows its work

THE RECOMMENDATION

Mr. R in Bed 12 has had a urinary catheter for 5 days[1]. No indication is documented in the last 48 hours[2]. Consider removal — worth raising on rounds.

1

ORDER · the device data / order

Indwelling urinary catheter — placed 5 days ago.

2

NOTE · a nursing note

"…no documented indication for continued catheterization."

Every claim links back to the chart — the order and the note it came from. Review takes seconds, not a re-read.

LEVEL 3 · INSIDE THE AGENT

One action of many

Tools to gather

- Read the notes
- Check orders & device data
- Pull flowsheets & labs
- Look across every patient at once

Tools to act

✓ **Message the nurse**

— what it chose to do

- Page the provider
- Draft an order change for sign-off
- Discontinue the order (gated — human decides)

The hard part is workflow, not the model

Fitting into real care

- Right moment: surface the prompt when the nurse can act on it, not buried in an inbox
- Right person: to the one who can make the change
- Low friction: acting must be easier than ignoring
- No alarm fatigue: fire rarely, and be right when it does

The business case, in the abstract

Earlier removal prompts

↓

Fewer unnecessary catheter-days

↓

Lower cumulative CAUTI risk

↓

Fewer infections, lower cost & LOS

LEVEL 3 · AGENTS FOR PREVENTION

Safety & oversight of a prevention agent



Human-in-the-loop, today

The agent recommends; a clinician acts and owns the decision. Autonomy only widens as validation and trust do.



Bounded scope

One narrow job with clear criteria — not open-ended “decide what's best.”



Auditable

Every prompt it fires can be traced, reviewed, and switched off.

BRINGING IT HOME

Getting started

1

Start at Level 1

Pick one repetitive writing or reading task. Use an approved tool. Learn the feel of it on low-stakes work first.

2

Find your friction, then press your vendor

Where do you lose the most hours? Take that to your EHR/surveillance vendor with your evaluation questions in hand

3

Match caution to risk

Surveillance help isn't in the clinical workflow and you're a ready-made expert review cohort — adopt broadly. Save the narrow pilots for clinical-facing agents.

BRINGING IT HOME

Where this is heading

Surveillance gets easier before it gets automated

Expect manual abstraction to keep shrinking — more proposing, more assisting — well before anything runs unattended.

More prevention agents, narrowly scoped

The catheter case is a template. Similar bounded assistants are emerging for central-line removal, C. diff and antibiotic triggers, outbreak exposure tracking, and proactive syndromic alerts.

The role moves up, not out

As tools take the first pass, the IP's judgment, validation, and governance become more of the value — not less.

IF YOU REMEMBER THREE THINGS

1



AI is already useful today

The everyday, general-purpose tools are here now — drafting, education, answering IP questions — with the right guardrails.

2



AI proposes, a person decides

At every level, the human stays in the loop. The tool changes the workload, never the accountability.

3



Your judgment is the safeguard

Validate on your own cases, govern from day one, and keep the decision yours — your expertise is what makes this safe.

Thank you

Questions?

Brandon Staunton

Software Developer · Infection Prevention & Control

