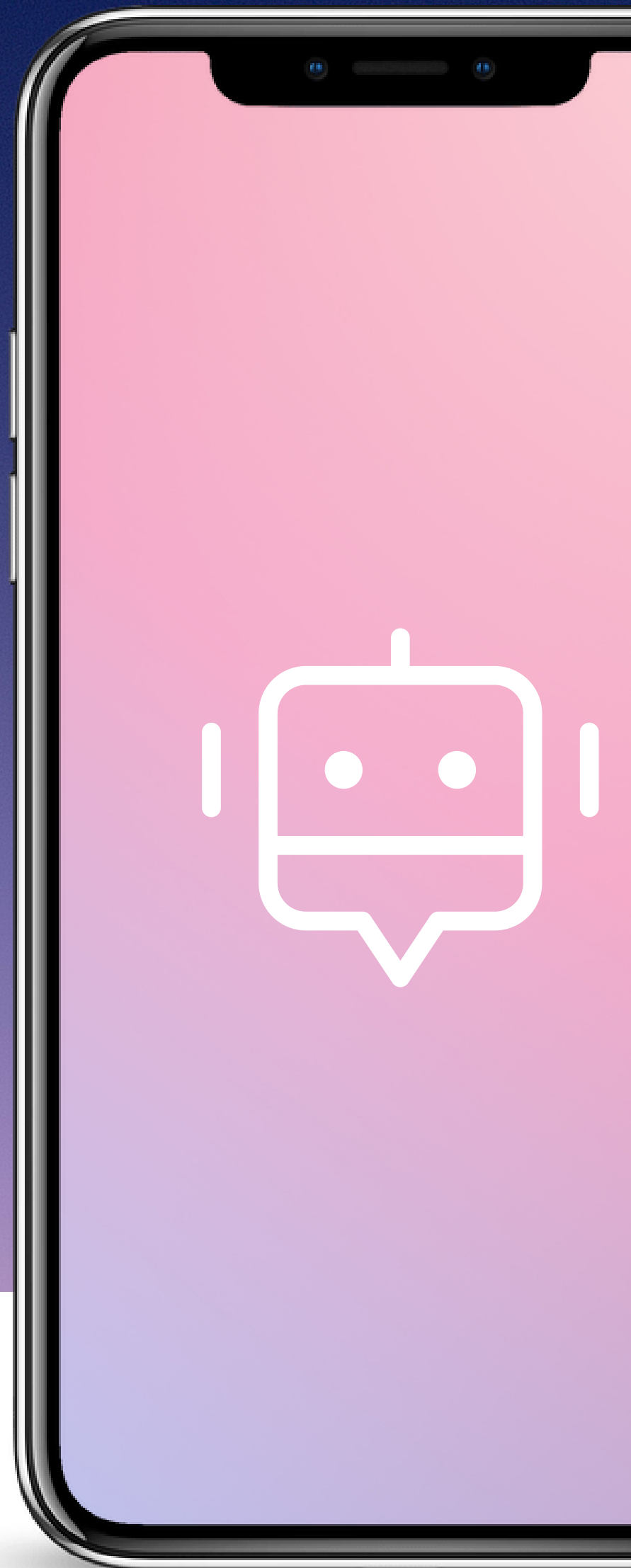


✦ Continuous HTA for
a Learning Health
System

Ambient AI Scribes: From Innovation to Evidence

A rapid scan of emerging
solutions, clinical evidence
and implementation
considerations



Health Technology
Assessment Unit (TAU)
MUHC

What does the evidence say?

We evaluated:

9 AI scribes:

DAX Copilot · Nabla · Ambience ·
Scribeberry · NoteDx · Heidi ·
Coeurway · Tali AI · Plume AI

3 key dimensions:

- Core functionalities
- Clinical benefit
- Implementation considerations

Outcome	What the evidence says*	Overall Signal
 Documentation burden	Studies suggest reduced documentation time, but savings are generally modest & don't consistently translate into increased productivity/well-being	Some benefit
 Accuracy & note quality	Errors, omissions and hallucinations have been reported	Mixed
 Physician well-being	Evidence on workload and burnout is limited & inconsistent	Mixed
 Patient experience	Potential for greater clinician attention, but recording raises consent/privacy concerns	Evidence gap

* Published peer-reviewed evidence, last updated February 2025

Ambient AI scribes: promising, but the evidence is still evolving

What differentiates the vendors?*

Scribe	EHR integration	Evidence base	Canadian / Québec fit	Customization & workflow
DAX Copilot	★★★	★★★	★	★★
Nabla	★★★	★★★	★	★★
Ambience	★★★	★	★	★★
Canadian / Québec solutions	★	★	★★★	★★
Heidi	★★	★	★★	★★

* Vendor characteristics were identified from publicly available information (data last accessed: Feb 2025). These were not independently verified by TAU and may change as products evolve.

5 THINGS MATTER MOST

- EHR integration** Can impact workflow and editing burden
- Note accuracy** Performance during complex encounters requires evaluation
- Customization** Specialty-specific templates can influence adoption
- Local fit** Privacy & data architecture/ local language fit
- Evidence maturity** Vendor-specific evidence concentrated on only 2 vendors

What does this mean for the MUHC?

TAU recommendation:

- Evaluate via a limited, prospective pilot before scaling
- Maintain clinician accountability
- Develop continuous monitoring

Pilot measure	What should we assess?
Efficiency	Documentation time, after-hours work, editing time
Quality	Accuracy, completeness, hallucinations, clinically meaningful errors
Clinician experience	Satisfaction, usability, perceived usefulness, workload & well-being
Patient experience	Acceptance, experience, consent/privacy concerns
Implementation	EHR integration, privacy/security, cost