Validating AI-Based Detection of I-PASS Elements in Verbal Handoffs: A Proof-of-Concept Study

Joshua Pankin, MD[1] · Vishal Pallerla[2] · Samantha Pendleton, DO[1] · Hannah Miller, MD[1] · Don Woodlock[2] · Jonathan Teich, MD[2] · Qi Li, MD[2] · Christopher Landrigan, MD, MPH[3] · Molly Senn-McNally, MD[1] · Amy Starmer, MD, MPH[1]

[1] Baystate Children's Hospital, Springfield, MA [2] InterSystems [3] Boston Children's Hospital, Boston, MA

Background & Objectives

Handoff communication failures during transitions of care contribute to medical errors and adverse patient outcomes. Implementation of the I-PASS Handoff Program has been associated with 30–47% reductions in preventable adverse events across diverse pediatric settings.[1–4] Despite resident acceptance,[3] adherence to all five I-PASS elements declines when handoffs are not directly observed.[2] At Baystate Children's Hospital, observation is performed by the night hospitalist on weekday nights only with limited opportunity to provide feedback.

OBJECTIVE

We sought to develop and validate an AI tool that automates I-PASS element detection in verbal handoffs, supplementing manual observation while providing residents with consistent, scalable feedback.

Methods

Dataset. 77 pediatric verbal handoff cases varied in quality and speaking style, yielding 385 element assessments. All 77 cases are simulated handoffs containing no protected health information. De-identified transcripts were processed via OpenAI APIs (Whisper for transcription, a GPT-based classifier for element detection).

Validation. Gold-standard criteria established by a 10-reviewer consensus panel (8 physicians, 2 software engineers). A trained reviewer independently assessed all 77 cases before reviewing AI output.

The Five I-PASS Elements

I	**Illness Severity**	Stable, "watcher," unstable
P	**Patient Summary**	Summary statement · events leading to admission · hospital course · ongoing assessment · plan
A	**Action List**	To-do list with timeline and ownership
S	**Situation Awareness & Contingency Planning**	Know what's going on · plan for what might happen
S	**Synthesis by Receiver**	Receiver summarizes what was heard · asks questions · restates key action items

Figure 1. *IPASS Handoff framework and key elements.*

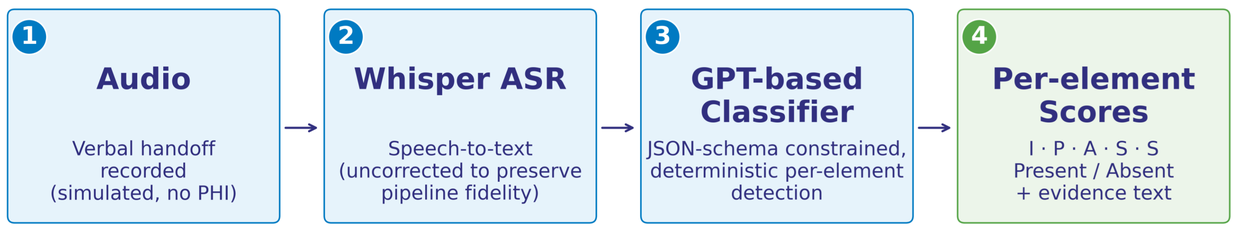

Figure 2. *Detection pipeline. Audio is transcribed uncorrected to preserve real-world fidelity, then classified per-element with strict JSON-schema output.*

Results

94.8% Overall accuracy

77 Simulated pediatric handoffs evaluated free of PHI.

80% of errors were misses, the safer feedback failure mode.

100% **Action** List — the highest per-element accuracy.

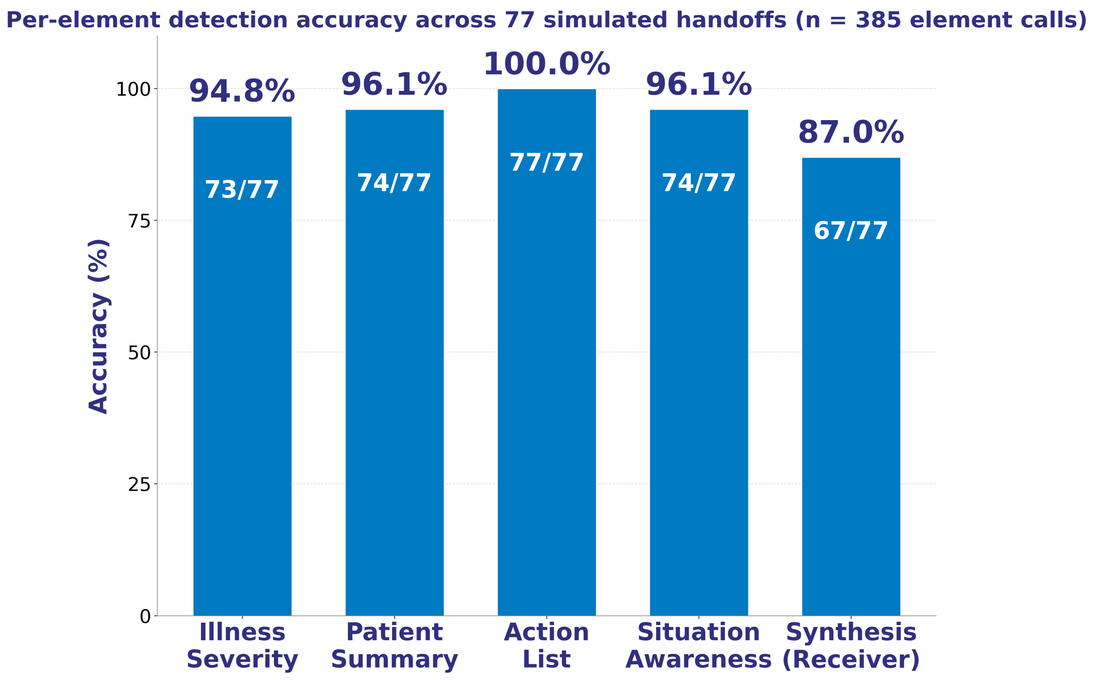

Figure 3. *Per-element accuracy. Action List detection was perfect; Synthesis (Receiver) was the lowest-performing element (87.0%).*

Patient 11 — 10-day-old · hyperbilirubinemia on phototherapy

Handoff #025 · simulated, no PHI

I ✗ Absent · P ✓ Present · A ✓ Present · S ✗ Absent · S ✓ Present

So, we have a 10 day old patient who was born at 373 spontaneous vaginal delivery who was admitted for hyperbilly. He was treated with phototherapy in the nursery and presented to the PCP with jaundice and what was sent to the ED. His levels on admission was 20.3 and he was started on lights and will continue throughout the night and we will just need to follow up a repeat serum, Billy Rubin. We're also trying to investigate.. Wait, so is he on lights now? He is. Okay. Mm. And the plan is for him to continue. And we have also sent some labs to figure out what's going on. There is known ABO and compatibility, and we've also sent a G6PD And Should I I follow those up overnight? Or are those, Do you think they'll come back overnight?, not sure. Probably, but I don't think you need to worry about it overnight, no. Okay. Yeah, and that's pretty much it just making sure that baby stays on phototherapy overnight and follow up the repeat Billy level. All right, so, baby on a belly bad. or with lights and is we'll get a belly level in the morning. Sounds good. All right.

Highlighted phrases = gold-standard reviewer's element-relevant evidence. Gaps signal what feedback should target.

Figure 4. *Annotated handoff excerpt. Highlighted spans triggered element detection; gaps reveal where resident feedback should target.*

Did the AI detect each I-PASS element across 77 handoffs?

		AI detected: Yes	AI detected: No
I	In handoff	26 TP *caught*	4 FN *missed*
	Not in handoff	0 FP *false alarm*	47 TN *correctly skipped*
P	In handoff	72 TP *caught*	0 FN *missed*
	Not in handoff	3 FP *false alarm*	2 TN *correctly skipped*
A	In handoff	74 TP *caught*	0 FN *missed*
	Not in handoff	0 FP *false alarm*	3 TN *correctly skipped*
S	In handoff	53 TP *caught*	3 FN *missed*
	Not in handoff	0 FP *false alarm*	21 TN *correctly skipped*
S	In handoff	30 TP *caught*	9 FN *missed*
	Not in handoff	1 FP *false alarm*	37 TN *correctly skipped*

Figure 5. *Per-element confusion matrices across 77 handoffs. Errors cluster in Synthesis (Receiver), where receiver readback is brief or interleaved with sender speech.*

Conclusions

Automated I-PASS element detection is feasible at 94.8% accuracy across 77 simulated handoffs. Errors strongly favored false negatives (80%) — the safer failure mode for educational feedback, since the tool defers rather than falsely confirms adherence.

FUTURE DIRECTIONS

- Live resident-handoff pilot at Baystate
- Iterative PDSA cycles to test feedback delivery
- Local model deployment within Baystate's environment to enable use with live, PHI-containing handoffs

Limitations

Single-institution pediatric setting. Simulated rather than live handoffs. Pilot did not use independent per-rater scoring, so inter-rater reliability (Cohen's κ) was not computed.

ASR errors and overlapping speech disproportionately affect Synthesis (Receiver) detection — the lowest-performing element.

Contact: Josh.Pankin@BaystateHealth.org

References & Acknowledgments

1. Starmer AJ et al. N Engl J Med. 2014;371(19):1803–1812.
2. Hughes HK et al. J Grad Med Educ. 2019;11(3):301–306.
3. Coffey M et al. J Grad Med Educ. 2017;9(3):313–320.
4. Starmer AJ et al. J Hosp Med. 2023;18(1):5–14.

Acknowledgments: Baystate Children's Hospital pediatric hospitalist team · Boston Children's Hospital · InterSystems Corporation · Baystate Pediatric Residency Program.

Disclosures: V. Pallerla, D. Woodlock, J. Teich, and Q. Li are employees of InterSystems Corporation. The remaining authors report no financial conflicts. No financial support was provided to any author for this work.