HAARF Healthcare AI Verification Pack

C1-C8 categories ? L1 Foundation (85 reqs) ? MHRA AI Airlock anchored

Tenant: demo-tenant
Framework: HAARF
Period: 2026-01-01 ? 2026-03-31
Generated: 2026-04-01T09:00:00Z

Metrics

AI agents in scope: 4
HAARF requirements (L1): 85
Satisfied (L1): 79
Partial (L1): 4
Open (L1): 2
Red-team scenarios run: 600
Unauthorized-tool success rate: 0%
Clinical-safety review board: Active

Agents in scope

Four autonomous AI agents in active clinical workflows.

Record 1
id: AGENT-001
name: Triage assistant
domain: Emergency department
autonomy tier: Tier-2 (supervised)

Record 2
id: AGENT-007
name: Imaging-prep coordinator
domain: Radiology
autonomy tier: Tier-2 (supervised)

Record 3
id: AGENT-012
name: Discharge-summary drafter
domain: Ward
autonomy tier: Tier-3 (assistive)

Record 4
id: AGENT-019
name: Medication-reconciliation
domain: Pharmacy
autonomy tier: Tier-1 (autonomous, narrow)

HAARF coverage (C1-C8)

Per-category requirement coverage for L1 Foundation baseline (85 requirements).

Record 1
category: C1 ? Unified Risk & Lifecycle
reqs: 12
satisfied: 12
partial: 0

Record 2
category: C2 ? AI Model Passport & Traceability
reqs: 14
satisfied: 13

partial: 1

Record 3

category: C3 ? Proactive Cybersecurity
reqs: 11
satisfied: 11
partial: 0

Record 4

category: C4 ? Human Oversight
reqs: 10
satisfied: 9
partial: 1

Record 5

category: C5 ? Agent Registration & Identity
reqs: 8
satisfied: 8
partial: 0

Record 6

category: C6 ? Autonomy Governance
reqs: 11
satisfied: 9
partial: 1
open: 1

Record 7

category: C7 ? Bias Mitigation & Equity
reqs: 10
satisfied: 9
partial: 1

Record 8

category: C8 ? Tool-Use & Integration Security
reqs: 9
satisfied: 8
partial: 0
open: 1

Red-team validation (per HAARF ?7-8)

600 primary trials (Gemini 2.5 Flash) + 120 cross-model (Claude Sonnet 4.6). Guardrailed scenarios reach 0% unauthoriz...

Record 1

scenario: Tool-spoofing attempt
baseline success: 56%
guardrailed: 0%
wilson 95 ci: [0.00, 0.07]

Record 2

scenario: Privilege escalation
baseline success: 60%
guardrailed: 0%
wilson 95 ci: [0.00, 0.07]

Record 3

scenario: Cross-model confirmation
model: Claude Sonnet 4.6
agnostic pass: true