



BLUE GOAT CYBER
Medical Device Cybersecurity

MEDICAL DEVICE CYBERSECURITY GUIDE

Penetration Testing for Medical Devices: A 2026 Explainer

What medical device penetration testing is, how it supports Section 524B obligations and FDA recommendations, the four expected test categories, and what a credible report contains.

Reviewed 2026-10-22

bluegoatcyber.com

Book a free discovery call
30 minutes with a senior expert
bluegoatcyber.com/discovery-call



IN BRIEF

Medical device penetration testing is hands-on, exploit-driven security testing of a finished or near-finished device, covering firmware, software, wireless interfaces, mobile and cloud companions, and supporting infrastructure, performed by independent testers for an FDA premarket submission. The February 2026 final premarket cybersecurity guidance expects four categories: security requirements testing, threat mitigation testing, vulnerability testing, and penetration testing, documented in a report tying findings to risk and remediation.

What medical device penetration testing is, how it supports Section 524B obligations and FDA recommendations, the four FDA-expected test categories, scope by device archetype, and what a credible deliverable contains.

A plain-language guide to medical device penetration testing - what it is, what the FDA actually requires, how it differs by device archetype, and what a submission-ready deliverable looks like.

Aligned to the [February 2026 final premarket cybersecurity guidance](#) and Section 524B of the FD&C Act.

Reviewed October 22, 2026 against the February 3, 2026 final guidance.

1. What Medical Device Penetration Testing Is (and Is Not)

A medical device penetration test is a **time-boxed, exploit-driven evaluation** of a real device by independent testers. The tester's job is to behave like a motivated attacker: chain weaknesses, bypass controls, and demonstrate impact - not just list missing patches.

It is **not**:

- A vulnerability scan (Nessus, OpenVAS) - automated, signature-based, no exploitation.
- A code review or SAST run - source-level, no runtime proof.
- A threat model - paper exercise that defines what could go wrong; the pen test proves what actually does.
- A red team - broader, goal-oriented adversary simulation, often spanning people, process, and physical access.

A pen test sits between the threat model and the red team: narrower than a red team, more concrete than a threat model, and reproducible enough to land in an eSTAR submission.

2. What the FDA Recommends

Under [Section 524B\(b\)\(2\)](#) the FDA requires manufacturers of "cyber devices" to provide a reasonable assurance of cybersecurity, including evidence of testing. The February 2026 final guidance operationalizes this in **Section V.C - Cybersecurity Testing**, which expects:

1. Independent testers (organizationally separate from the development team).
2. Testing against the final or near-final device build.
3. Coverage of every external interface in the security architecture.
4. A report that the FDA reviewer can map back to your threat model and risk file.

Pen testing also appears - implicitly - in the labeling, vulnerability-management, and postmarket sections. If your submission says you'll handle a class of vulnerability, the pen test is where you prove the control works.

3. The Four FDA-Expected Test Categories

Section V.C of the February 3, 2026 guidance recommends considering, among other testing, four categories. They describe different evidence purposes, not four required standalone files. One report may cover several categories when methods, results, and traceability are explicit.

3.1 Security Requirements Testing

Document evidence that security design input requirements were implemented successfully, including boundary analysis and the rationale for boundary assumptions. Boundary analysis concerns limits of security requirements, not only architecture trust boundaries.

3.2 Threat Mitigation Testing

Document evidence that risk controls identified in the threat model are effective, with traceability from threats to controls and test results.

3.3 Vulnerability Testing

Identify known vulnerabilities and potential weaknesses using appropriate methods. The guidance discusses abuse cases, malformed or unexpected inputs, attack-surface analysis, vulnerability chaining, and closed-box known-vulnerability scanning within this category.

3.4 Penetration Testing

Use independent testers to identify and characterize security issues through adversarial testing. Independence may be provided by independent internal testers or external testers; a third party is not universally required. Document tester independence and expertise, scope, duration, methods, and results.

4. Scope by Device Archetype

The same four test categories apply to every device, but the surface area and tooling differ wildly. Use this table to size an engagement before you scope it.

Archetype	Primary attack surface	Specialist skills needed
Embedded / firmware device	UART, JTAG, SPI flash, secure boot, signed updates, debug interfaces	Hardware reversing, firmware extraction (chip-off, JTAG), Ghidra / IDA, signed-update bypass
SaMD (mobile + cloud)	Mobile app (iOS/Android), backend APIs, IAM, tenant isolation, secrets handling	OWASP MASVS, MobSF, Frida, Burp, cloud auth flow analysis
Implantable + BLE-paired companion	BLE pairing & bonding, GATT services, replay, & MITM, companion app	Sniffers (Ubertooth, Sniffle), BLE fuzzing, ATT/GATT protocol analysis
Imaging modality / DICOM / PACS	DICOM/DIMSE, HL7, network segmentation, vendor service accounts	DICOM toolkits (dcm4che), HL7 fuzzing, hospital-network simulation
Hospital-network-connected device	Wired/wireless onboarding, 802.1X, default credentials, service portals	Network attack tooling, AD/Kerberos basics, segmentation testing
AI/ML SaMD	Model integrity, adversarial inputs, prompt/data poisoning, MLOps pipeline	Adversarial ML, model-signing review, pipeline IAM review

A credible scoping conversation names the archetype, lists the boundaries from the threat model, and assigns hours to each.

5. The Test Environment Problem

The 2026 guidance expects testing against a **production-equivalent** configuration. In practice this means:

- The exact hardware revision that will ship (not an early prototype with extra debug headers).
- The release-candidate firmware/software build, signed with production keys where possible.
- Surrogate networks, surrogate patient data, and - for hospital-connected devices - a hospital-network simulator (EHR, PACS, DHCP, NTP).
- Documented deviations from production, with a justification for why each deviation doesn't change the security posture.

If the FDA reviewer can't tell from the report what was tested, the test doesn't count. Include build hashes, hardware serials, and network diagrams.

6. Methodologies the FDA Recognizes

Cite the methodology you actually followed. The 2026 guidance and AAMI TIR57 / ANSI/AAMI SW96:2023 explicitly reference:

- **OWASP MASVS / MSTG** - mobile companion apps.
- **OWASP ASVS** - web/cloud backends and clinician portals.
- **OWASP IoT Top 10** - embedded device weakness classes.

- **ISA/IEC 62443-4-2** - component-level cybersecurity requirements.
- **NIST SP 800-115** - overall testing process and reporting structure.
- **AAMI TIR57 / ANSI/AAMI SW96:2023** - security risk management linkage between threats, tests, and risk file.

A report that names its methodology, version, and the specific controls tested is far harder for a reviewer to push back on.

7. What a Credible Pen-Test Deliverable Contains

A practical report structure, not a mandatory FDA nine-item format:

1. **Executive summary** - one page, risk-rated, non-technical.
2. **Scope statement** - device, build, environment, dates, in/out of scope.
3. **Methodology** - which standards were followed, which tools used.
4. **Boundary coverage matrix** - every threat-model boundary mapped to the tests run against it.
5. **Findings** - each with CVSS v3.1, EPSS, exploitability narrative, screenshots/PoCs, remediation guidance, and a CWE.
6. **Chained attack paths** - documented where found; no guaranteed exploit count.
7. **Retest results** - findings that were fixed and re-verified.
8. **Residual risk statement** - aligned to your ISO 14971 risk file.
9. **Tester independence and expertise** - independence, qualifications, hours.

If your current vendor's report is a Nessus export with a cover page, you do not have a submission-ready deliverable. See our companion piece, [12 Critical Findings from Medical Device Penetration Tests](#), for what the body of a real report looks like.

8. Pen Test vs Red Team vs Tabletop vs Threat Model

Activity	Goal	Output	When you need it
Threat model	Identify what could go wrong	STRIDE/attack-tree diagrams, risk register	Pre-design and continuously
Penetration test	Prove which threats are real on the built device	Findings, PoCs, residual risk	Before every FDA submission
Red team	Test detection + response across people/process	Adversary narrative, control gaps	Mature postmarket programs
Tabletop	Stress-test the incident-response plan	Lessons learned, playbook updates	Annually, postmarket

Most manufacturers need a threat model and a pen test for premarket; red teams and tabletops belong to the postmarket cybersecurity program.

One more distinction matters for the submission: a penetration test is adversarial security testing of the built device, while **security requirements testing** is verification against your own design inputs, complete with boundary analysis and the rationale behind those boundary assumptions. Section V.C recommends evidence covering both purposes. One report or several may be used with clear coverage and traceability. See [security requirements testing vs penetration testing](#).

9. Cost and Timeline Ranges

Transparent ranges so you can budget. Actual quotes depend on archetype and number of interfaces:

Device archetype	Typical duration	Typical investment (USD)
SaMD only (mobile + cloud)	2 - 3 weeks	\$25K - \$55K
Single embedded device, no wireless	3 - 4 weeks	\$40K - \$75K
Embedded + BLE companion app	4 - 6 weeks	\$60K - \$110K
Imaging / DICOM / hospital-connected	5 - 8 weeks	\$75K - \$150K
Multi-device platform (gateway + nodes + cloud)	8 - 12 weeks	\$120K - \$250K+

Quotes below these ranges almost always omit one of the four test categories. Quotes above usually include premarket consulting, re-test, and FDA Q-response support.

10. How Blue Goat Cyber Approaches Medical Device Penetration Testing

We run pen tests specifically for FDA-regulated devices - the scope, evidence format, and threat coverage reviewers actually want to see. Here is the model behind every engagement.

- **Threat-model-driven scope.** We ingest your STRIDE model and design documentation first, so the test targets interfaces and data flows tied to declared risks - not a generic OWASP checklist.
- **Full device attack surface.** Firmware, hardware interfaces (JTAG/UART/USB), radios (BLE/Wi-Fi/cellular), cloud APIs, mobile companion apps, and clinician web portals - tested as one connected system.
- **Independent testers.** Our pen-test team does not also engineer your device. FDA reviewers increasingly expect this separation, and we run it as a default.
- **Reviewer-ready deliverables.** Every finding includes CVSS + FDA patient-safety severity, exploit narrative, reproduction steps, and remediation guidance mapped to AAMI TIR57 / ANSI/AAMI SW96:2023 hazards.

- **Retest included.** Fixing findings without a documented retest is a common deficiency trigger. Retest is part of the engagement, not an add-on.
- **Timeline aligned to submission.** We plan pen testing after design freeze with enough runway to remediate and retest before eSTAR lock.

Our [Medical Device Penetration Testing](#) engagement is the entry point.

11. Where Automation Helps, and Where It Doesn't

Automated tooling does a defined job well. Port and service discovery, protocol fuzzing, static analysis, SBOM generation, and triage of known CVEs against the component inventory are all faster and more consistent when a tool runs them, and every credible test plan uses them.

What tooling cannot do is reason about purpose. A scanner does not know that an infusion rate field should never accept that value, that a service technician role should not be able to reach patient data through a legitimate maintenance workflow, or that three individually low findings chain into a path that ends in patient harm. Business logic flaws, authorization bypasses, clinical workflow abuse, and multi-step exploit chains are found by people, not by signatures.

That is why the majority of a credible medical device engagement is manual penetration testing, with tool output treated as an input rather than as the report. It is also why reviewers react badly to a test report that is visibly a scan export: it demonstrates coverage of the known and silence on everything else. If you are weighing automated vs manual security testing, plan for tool-assisted breadth and manual depth, and say in the test plan which findings came from which.

12. Frequently Asked Questions

Does the FDA actually require a pen test? Section 524B(b)(2) requires cybersecurity processes providing reasonable assurance, not a separately named penetration-test document. Section V.C of the February 3, 2026 final guidance recommends penetration testing among its testing categories. Determine applicable evidence for the device and submission pathway.

Is your testing automated or manual? Both, weighted toward manual. Tooling handles scanning, fuzzing, SBOM generation, and known-CVE triage; the majority of each engagement is manual testing by a senior tester, because automation cannot find business logic flaws, authorization bypasses, clinical workflow abuse, or chained exploits.

Can our internal security team do it? Yes, if the internal testers are independent of development and have appropriate expertise. Section V.C allows independent internal or external testers; third parties may be needed in some cases.

When in the development cycle should we test? Against the release-candidate build, with enough runway to remediate critical findings and retest before submission. Most manufacturers schedule the pen test 8 - 12 weeks before their target FDA filing date.

Does a pen test cover SBOM / vulnerability management? Partly. The pen test will exercise components in your SBOM and may surface exploitable CVEs, but the ongoing CVE-matching loop is a separate program. See our [SBOM vulnerability management guide](#).

Do we need a new pen test for every software update? Material design changes - new interfaces, new wireless, new cloud integration, new AI/ML model - trigger a re-test. Routine maintenance releases do not, provided your postmarket vulnerability-management plan is operating.

How does this differ for Class III / PMA devices? Higher scrutiny, more required artifacts, longer engagements. See [FDA PMA cybersecurity requirements](#) and [FDA pathway cybersecurity differences](#).

13. Where to Go Next

- **Service overview:** [Medical Device Penetration Testing](#)
- **Requirements deep-dive:** [Medical Device Pen Test Requirements](#)
- **Field findings:** [12 Critical Findings from Medical Device Penetration Tests](#)
- **Premarket roadmap:** [FDA Premarket Cybersecurity Submission Checklist](#)
- **Standards context:** [The MedTech Cybersecurity Standards Decoder](#)
- **Verification side:** [FDA Security Requirements Testing](#) - the design-controls evidence a pen test does not replace.
- [FDA Penetration Testing Scope for 510\(k\)](#) - how to bound scope so the report satisfies review.

Ready to scope a pen test against the February 2026 guidance? [Contact Blue Goat Cyber](#) for a fixed-scope, fixed-price proposal.

References

Primary sources cited or relied on in this guide:

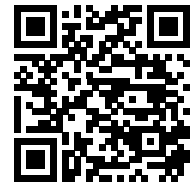
- [Cybersecurity in Medical Devices: Quality Management System Considerations and Content of Premarket Submissions \(Feb 2026 final\)](#), U.S. FDA
- [ANSI/AAMI SW96:2023, Standard for Medical Device Security](#), AAMI/ANSI
- [IEC 81001-5-1:2021, Health software and health IT systems safety, effectiveness, and security](#), IEC/ISO
- [OWASP Application Security Verification Standard \(ASVS\)](#), OWASP
- [Section 524B of the FD&C Act \(Consolidated Appropriations Act, 2023, §3305\)](#), U.S. Congress

Current web guide: bluegoatcyber.com/guides/penetration-testing-for-medical-devices



Talk to a senior medical device cybersecurity expert

Book a free 30-minute discovery call. Scan the code or visit bluegoatcyber.com/discovery-call. You get a written quote within 24 hours.



Informational material. Confirm requirements for your device and market. The manufacturer owns risk decisions, QMS approvals, and regulatory submissions. No regulatory outcome is promised.