

Penetration Testing Vendor Evaluation Checklist

30 checks, 12 KPIs, and a scoring rubric for comparing penetration testing firms on who tests, how deep they go, the report, the fixes, and the retest.

By Louis Sanchez, OSCP, OSWA, CISSP, CCSK · Founder, Voke Cyber · References: NIST SP 800-115, PCI DSS v4.0.1 Req. 11.4, PCI SSC Penetration Testing Guidance

What to Look for in a Penetration Testing Firm: The Short Answer

When you evaluate penetration testing firms, look for six things. First, a named, certified tester who also writes the report. Second, a documented manual methodology such as NIST SP 800-115, the OWASP Web Security Testing Guide, or PTES, not a scanner run. Third, written authorization and rules of engagement before testing starts. Fourth, detailed reports that work for both engineers and non-technical stakeholders. Fifth, remediation guidance specific to your technology stack, with a retest included to verify the fixes. Sixth, evidence that your compliance framework (PCI DSS, SOC 2, HIPAA, ISO 27001) will accept. Score each provider on the 30 checks below: a penetration testing vendor that scores 50 of 60 or more, with no hard stops, is a strong choice. The best way to compare pentest companies is to use the same selection criteria and due diligence questions for every one of them, which is what this checklist is for.

Only have 15 minutes per vendor? Start here.

1. **Ask for a sample report.** It shows more about a vendor than any sales call. Check it against Area 4.
2. **Ask who the tester is, then verify them.** Look up their certifications with the issuer and ask to speak with them.
3. **Ask what the retest costs and what it produces.** If the answer is "extra" or "a new report later," factor that into the price.

How to score

- **2 points:** Yes, and the vendor showed evidence (a document, a sample, a name you verified).
- **1 point:** Partially, only on request, or a verbal yes with nothing to back it.
- **0 points:** No, or they could not answer.

50-60 Strong. Shortlist. Put the commitments you scored into the SOW.

40-49 Workable. Negotiate the gaps into the SOW before signing.

Below 40 Keep looking. The risk is that you pay for a scan and call it a pentest.

Hard stops, whatever the score

- No named tester, or certifications you cannot verify.
- No sample report, or a sample that is a scanner export.
- No written authorization or rules of engagement before testing.
- No retest, so you cannot prove the fixes held.

Red flags in a pentest proposal

- The proposal names a firm but no tester.
- The sample report is a scanner export with a cover page.
- Every finding is rated High or Critical, with no CVSS vectors to check.
- Retesting is a paid add-on or is not offered.
- They cannot explain how they test authorization or business logic.
- "Continuous pentesting" pricing that, on inspection, is automated scanning.

The 30-Point Scorecard

Score each check 0, 1, or 2 for up to three vendors (A, B, C). Five checks per area, 10 points per area, 60 points in total.

1 Who Actually Does the Testing

PCI SSC Information Supplement: Penetration Testing Guidance (tester qualifications)

HIGHEST PRIORITY

The single biggest driver of pentest quality is the person at the keyboard. Get a name, verify the credentials, and confirm that person writes the report.

CHECK	A	B	C
1.1 The proposal names the tester(s), and the same person who tests writes the report.			
1.2 Their hands-on certifications (for example OSCP, OSWE, OSPA, GPEN, CREST CRT) are verifiable through the issuer's public lookup.			
1.3 They can describe years of offensive security work specifically, not general IT or compliance experience.			
1.4 The firm states whether any work is subcontracted or offshored, and where your data and evidence will be handled.			
1.5 You can talk to the tester directly before signing, and during testing.			
Area 1 subtotal (max 10)			

2 Methodology and Manual Depth

NIST SP 800-115 · OWASP WSTG / MASTG · PTES · PCI DSS v4.0.1 Req. 11.4.1

DEPTH CHECK

A credible penetration testing firm names its methodology, tests business logic and authorization by hand, and exploits past detection. Scanners find the easy half.

CHECK	A	B	C
2.1 The vendor names the public methodology it follows (for example NIST SP 800-115, OWASP WSTG or MASTG, or PTES) and, if you are in PCI scope, shows how it meets PCI DSS 11.4.1.			
2.2 They can explain how business-logic and authorization flaws are tested by hand, not just which scanners they run.			
2.3 Authenticated testing uses multiple roles, and multiple tenants where your product is multi-tenant.			
2.4 They attempt exploitation and chain findings to show real impact, within agreed rules, rather than stopping at detection.			
2.5 Every scanner result is manually validated before it reaches the report.			
Area 2 subtotal (max 10)			

3 Scope, Rules of Engagement, and Legal

NIST SP 800-115 (planning phase and rules of engagement)

RISK CONTROL

Before a pentest starts, the vendor should provide written authorization, rules of engagement, proof of insurance, and a data-handling policy. That paperwork protects you too.

CHECK	A	B	C
3.1 A written authorization is signed by someone with authority over every in-scope asset, including third-party hosting.			
3.2 Rules of engagement define testing windows, prohibited actions, and an emergency stop contact.			
3.3 The vendor provides certificates of professional liability and cyber insurance on request.			
3.4 A data-handling policy covers how evidence is stored and encrypted and when it is destroyed.			
3.5 Testing source IP addresses are disclosed in advance so your team can separate test traffic from real attacks.			
Area 3 subtotal (max 10)			

4 Report Quality for Two Audiences

NIST SP 800-115 (reporting phase) · PCI SSC Penetration Testing Guidance (reporting)

DELIVERABLE

Leadership needs business risk in plain language. Engineers need evidence and reproduction steps. A good report serves both without a translator.

CHECK	A	B	C
4.1 The executive summary is readable by a non-technical leader and states business impact and priorities.			
4.2 Every finding includes the affected asset, evidence, reproduction steps, and a severity score with its CVSS vector.			
4.3 The report states the scope, test dates, methodology, and limitations, including anything that was not tested.			
4.4 It records attack paths that were attempted and failed, so you can see coverage and not just the failures.			
4.5 The vendor shares a real sample report (redacted or public) before you sign.			
Area 4 subtotal (max 10)			

5 Remediation Guidance and Retesting

PCI DSS v4.0.1 Req. 11.4.4 (correct exploitable findings, repeat testing to verify)

OUTCOME

Pentest findings only matter once they are fixed. Look for fixes specific to your stack, a review call with the tester, and a retest that proves each fix held.

CHECK	A	B	C
5.1 Remediation guidance is specific to your stack and includes a verification step, not just a generic reference link.			
5.2 Findings come with a prioritized fix order that separates quick wins from structural fixes.			
5.3 A findings review call with the tester is included.			
5.4 A retest is included in the price, within a defined window, and produces an updated status for every finding.			
5.5 Critical findings are escalated as soon as they are confirmed, not held until the final report.			
Area 5 subtotal (max 10)			

6 Compliance and Audit Evidence

PCI DSS v4.0.1 Req. 11.4.1–11.4.5 · ISO/IEC 27001:2022 A.8.8

AUDIT READY

If an auditor, customer, or insurer will read the result, confirm the report meets their specific requirements before testing starts, not after.

CHECK	A	B	C
6.1 The vendor can explain your framework's specific pentest requirements (for example PCI DSS 11.4, SOC 2, HIPAA, ISO/IEC 27001 A.8.8).			
6.2 The tester is organizationally independent of the systems tested, as PCI DSS 11.4.1 requires.			
6.3 An attestation letter or summary is available to share with customers and insurers without exposing the full report.			
6.4 Test dates, scope, and methodology are documented in a way that fits your audit period.			
6.5 The vendor can support annual testing and testing after significant changes.			
Area 6 subtotal (max 10)			

Totals

Vendor	Name	Total (max 60)	Any hard stop?	Decision
A				
B				
C				

12 KPIs for Evaluating a Penetration Testing Vendor

Write these into the statement of work so "good" is defined before testing starts.

KPI	WHAT GOOD LOOKS LIKE	RED FLAG
Critical-finding notification	Same business day the finding is confirmed	Criticals first appear in the final report
Named tester	Tester named in the SOW, and the same person writes the report	"Our team" with no names
Manual validation rate	100% of reported findings confirmed by hand	Raw scanner output in the findings list
Reproducibility	100% of findings include evidence and reproduction steps	Findings your team cannot reproduce
False positives	Zero in the final report; any your team disproves are removed and noted	Long lists padded with informational items
Severity scoring	CVSS vector shown for every finding, and the score matches the vector	Scores with no vector, or everything rated High
Remediation specificity	Every finding has a stack-specific fix and a verification step	A generic OWASP link as the only guidance
Executive summary	A non-technical reader can act on it in five minutes	Jargon, CVE lists, or no summary at all
Coverage statement	Scope, dates, methodology, and untested areas are all stated	No way to tell what was actually tested
Retest	Included, within a defined window (ours is 30 days), with an updated status per finding	Retest billed separately, or no retest offered
Tester access	Direct channel to the tester during testing and after the report	All questions routed through an account manager
Time to start	A committed start date once scope and access are agreed	Open-ended scheduling or a rush premium

Score us too.

Voke Cyber is a penetration testing firm in the Charlotte, NC area. We publish our starting prices, put you in direct contact with the tester doing the work, and include a retest within 30 days of every final report. Ask us for a sample report at vokecyber.com/contact, or request a scoped quote at vokecyber.com/get-started.

Online version, with explanations for each area and a FAQ: <https://vokecyber.com/resources/pentest-vendor-evaluation-checklist/>

Detailed Reporting and Remediation Guidance: What Each Stakeholder Should Get

Penetration testing firms and application security services firms are easiest to compare on their reports, because the report is what every stakeholder actually uses. Detailed reporting means one report that works for four audiences: non-technical leaders, engineers, auditors, and the customers or insurers who ask for proof. Remediation guidance means each finding tells the engineer what to change, in their stack, and how to confirm the fix. Use this table to check a sample report from any vendor.

STAKEHOLDER	WHAT A DETAILED PENTEST REPORT GIVES THEM	QUESTION TO ASK THE VENDOR
Executives and non-technical stakeholders	A one- or two-page executive summary in plain language: overall risk, what an attacker could actually do, and the top three priorities. No CVE lists or tool output.	"Show me the executive summary from a recent report. Could our CEO act on it in five minutes?"
Engineers and developers	Every finding with the affected asset, evidence (requests, responses, screenshots), reproduction steps, a CVSS vector, and stack-specific remediation guidance with a verification step.	"Can my team reproduce a finding from the report alone, without a call?"
Auditors and compliance	Scope, test dates, methodology, tester independence, limitations, and retest results showing each finding's status at closure.	"Will this report satisfy PCI DSS 11.4, SOC 2, HIPAA, or ISO 27001 evidence requests as is?"
Customers, partners, and insurers	An attestation letter or summary that confirms testing was done and findings were addressed, without exposing exploit details.	"Do you provide a shareable attestation, and what does it include?"

The Pentest Findings Remediation Process, Step by Step

A good penetration testing vendor supports the whole remediation process, not just the report. These are the six steps for taking pentest findings from report to verified closure, and what the vendor should contribute at each step.

1. **Triage the findings with the tester.** On the findings review call, confirm each finding, check its severity against the CVSS vector, and add business context the tester could not see.
2. **Assign an owner and a due date to every finding.** How long remediation should take after a pentest depends on severity, so set deadlines by severity in your vulnerability management policy before the test starts. PCI DSS v4.0.1 Requirement 6.3.3, for example, requires critical and high-severity security patches within one month of release.
3. **Fix the root cause, not just the instance.** If one endpoint is missing an authorization check, review the shared code path so the same flaw does not survive on the endpoints the tester did not reach.
4. **Verify the fix internally.** Use the report's reproduction steps to confirm the finding no longer reproduces before you request the retest.
5. **Retest with the vendor.** The tester repeats the attack and records an updated status for every finding. PCI DSS v4.0.1 Requirement 11.4.4 requires this repeat testing for exploitable vulnerabilities.
6. **Document closure.** Attach the retest results to your risk register and audit evidence, and formally accept and document any risk you decide not to fix.

Manual Penetration Testing vs. Continuous Vulnerability Management for Compliance

Continuous compliance programs run vulnerability management all year: automated scanning, patching, and tracking. A manual penetration test is a separate requirement. Scanning finds known weaknesses; a penetration tester proves which ones can be exploited and finds the authorization and business-logic flaws scanners miss. Where a compliance framework names penetration testing (PCI DSS, NY DFS Part 500, FedRAMP), scanning does not substitute for it: both are required, on different cadences. When you evaluate penetration testing services for compliance, confirm the vendor knows which of these frameworks applies to you.

FRAMEWORK	MANUAL PENETRATION TESTING	VULNERABILITY SCANNING	CITATION
PCI DSS v4.0.1	Required: internal and external, at least annually and after significant changes; segmentation testing every 12 months (every 6 months for service providers)	Quarterly internal scans and quarterly external ASV scans	Req. 11.3, 11.4
SOC 2	Expected by auditors: penetration testing is named in the CC4.1 points of focus; annual is the de facto standard	Commonly used to satisfy CC7.1 monitoring	CC4.1, CC7.1
HIPAA	Not named in the current Security Rule, but risk analysis and evaluation drive it; annual is the industry standard. A proposed Security Rule update would require it every 12 months	Proposed update would require scanning every 6 months	45 CFR 164.308
ISO/IEC 27001:2022	Not named, but auditors expect it as evidence for technical vulnerability management; annual	Part of A.8.8 technical vulnerability management	A.8.8, A.5.36
NY DFS 23 NYCRR 500	Required: annually, from inside and outside the system boundaries, by a qualified internal or external party	Risk-based automated scans plus manual review (since May 1, 2025)	Section 500.5
FedRAMP	Required: annually, by an independent third-party assessor (3PAO), following FedRAMP penetration test guidance	Monthly, with 30-day remediation for high-risk findings	FedRAMP Pentest Guidance

Sources

The checks cite only what these primary sources say. KPI targets and the scoring rubric are Voke Cyber's own recommendations, not requirements of any standard.

- [NIST SP 800-115, Technical Guide to Information Security Testing and Assessment](#): four-phase penetration testing methodology (planning, discovery, attack, reporting) and rules of engagement.
- [PCI DSS v4.0.1 and the PCI SSC Penetration Testing Guidance information supplement](#): Requirements 6.3.3, 11.3, and 11.4.1 through 11.4.6; tester qualifications and report content.
- [OWASP Web Security Testing Guide \(WSTG\)](#): web application testing methodology.
- [OWASP Mobile Application Security Testing Guide \(MASTG\)](#): mobile application testing methodology.
- [FIRST Common Vulnerability Scoring System \(CVSS\)](#): severity scores and vector strings.
- [FedRAMP](#): annual third-party penetration testing and monthly scanning for cloud service providers.