

**FROM SCAN REPORT TO HARDENED NETWORK: A MITRE ATT&CK-DRIVEN
FRAMEWORK FOR CLOSING THE ATTACK SURFACE REDUCTION GAP****Samuel Bemi**

Independent Researcher, United Kingdom

ORCID iD: 0009-0001-9992-0741

ABSTRACT

Penetration testing is widely practised, yet its findings are rarely converted into durable architectural change. Results are typically ranked by Common Vulnerability Scoring System (CVSS) severity, delivered as static reports and actioned selectively, leaving structural weaknesses such as flat network segmentation and perimeter-centric monitoring largely unresolved between engagements. This study proposes Map, Prioritise, Harden, Validate (MPHV), a four-stage, tools-based framework that maps Nmap, Burp Suite and Metasploit findings onto MITRE ATT&CK techniques and Cyber Kill Chain stages; prioritises remediation using a two-axis Exploit Path Criticality matrix that weighs demonstrated exploitability against lateral blast radius; translates prioritised findings into specific segmentation, firewall and IDS/IPS changes; and closes the loop by re-testing the exact technique each control was intended to stop. The framework was developed through a qualitative, design-oriented methodology that combined a structured review of academic literature and 2025 industry breach data with practitioner experience from enterprise and healthcare engagements, and it is demonstrated through a tool-to-technique mapping, two composite cases and one publicly reported 2025 incident. The analysis indicates that the gap between broad ATT&CK adoption and its shallow operational use is a plausible contributor to the long breach lifecycles and lateral-movement-heavy breach patterns now reported by industry. The paper closes with an explicit account of limitations: ATT&CK coverage is not predictive of novel techniques, prioritisation scoring remains vulnerable to organisational pressure, and continuous validation is not yet affordable for many resource-constrained organisations.

Keywords:

MITRE ATT&CK, Cyber Kill Chain, Network Hardening, Attack Surface Reduction, Penetration Testing, Network Segmentation

INTRODUCTION

Penetration testing has become a routine, and frequently contractual, obligation for enterprise and public sector organisations. Yet the distance between running a test and actually reducing risk has, if anything, widened. The 2025 Verizon Data Breach Investigations Report (DBIR) analysed 12,195 confirmed breaches across 139 countries and found that system intrusion, the breach pattern most closely associated with lateral movement once an attacker is inside a network, grew from 36% to 53% of breaches year on year, while the presence of ransomware rose from 37% to 44% (Verizon, 2025). These are not, in the main, organisations without security programmes. In many cases they are organisations that commissioned a penetration test, received a report and moved on.

At the same time, MITRE ATT&CK has become close to a default reference point across the industry. Survey research reports adoption by more than 80% of enterprises (Basra & Kaushik, 2020), and separate analyst research found that 48% of organisations use the framework extensively for security operations (CSO Online, 2022). The vocabulary of tactics and techniques is now ubiquitous in reports and dashboards. What is far less common is evidence that the framework actually shapes which firewall rule is written or which VLAN is created. That disconnect, between widespread familiarity with a framework and limited operational use of it, is the practical problem this paper addresses.

Problem statement. Penetration test findings are typically delivered as a flat list ranked by CVSS score or a broadly similar severity rating, with little indication of how a given finding fits into a realistic attack sequence or how far an attacker could progress if it were exploited. Remediation teams working from that list gravitate towards fixes that are cheap and fast, such as patching a single host or closing an obviously unnecessary port, while structural issues such as flat network architecture, unrestricted east-west traffic and IDS/IPS sensors placed only at the perimeter remain unresolved, reappearing in the next report almost unchanged.

The cost of this pattern is no longer speculative. The IBM Cost of a Data Breach Report 2025, based on 600 breached organisations, found that the average global breach lifecycle (time to identify plus time to contain) was 241 days, the lowest in nine years but still most of a working year (IBM, 2025). Breaches that took longer than 200 days to contain cost on average USD 1.14 million more than those contained within 200 days. Healthcare breaches averaged USD 7.42 million and 279 days to contain, the costliest of any sector for the fifteenth consecutive year, and breaches involving compromised credentials, which typically depend on unrestricted lateral movement to become material, averaged USD 4.67 million (IBM, 2025). There is also a behavioural dimension: only 49% of breached organisations planned to increase security investment after an incident, down from 63% the previous year (IBM, 2025). If a breach does not consistently move budget, a penetration test report that arrives with no comparable sense of urgency is unlikely to fare better without a fundamentally different way of presenting its findings.

Research objectives. Against this background, the study pursues five objectives: (i) to identify, from current literature and industry data, the specific gaps that prevent penetration test findings from becoming durable hardening; (ii) to design a repeatable framework that links offensive tool output to MITRE ATT&CK techniques and Cyber Kill Chain stages; (iii) to define an evidence-based prioritisation mechanism that replaces severity-only ranking; (iv) to demonstrate how prioritised findings translate into traceable segmentation, firewall and IDS/IPS controls; and (v) to assess critically the limitations of both the proposed framework and the study itself.

Significance and scope. The paper is aimed at practitioners who run or commission offensive testing (penetration testers, security consultants and security operations centre leads) and at the IT managers and chief information security officers who decide what happens to the findings. Its contribution is structural rather than technical: it reuses tools and frameworks practitioners already have, and adds a disciplined cycle that makes every control change traceable to a demonstrated technique and every closure independently re-testable. The scope covers network-layer and application-layer attack surface reduction, network segmentation, firewall and IDS/IPS configuration, and a repeatable prioritisation and validation cycle. It deliberately does not cover cloud security posture management in depth, does not replace a secure software development lifecycle programme or a full red or purple team methodology, and treats ISO/IEC 27001 and NIST mapping as context rather than as a compliance guide.

The remainder of the paper is organised as follows. The literature review examines tooling, ATT&CK research, kill chain models, vulnerability prioritisation, segmentation and continuous validation. The next section consolidates the research gaps and the limitations of current approaches. The methodology section describes the design-oriented research approach. Results and discussion present the MPHV framework, its demonstration and its limitations, followed by future work and conclusions.

LITERATURE REVIEW

Penetration Testing Practice and Tooling

Established guidance frames security testing as a structured sequence of planning, discovery, attack and reporting (Scarfone et al., 2008). In practice the offensive tooling landscape is stable: Nmap for reconnaissance and enumeration, Burp Suite for web application assessment, Metasploit for exploitation and post-exploitation, and Wireshark for traffic analysis remain the default toolkit across independent consultancies and in-house red teams. What has changed is less the tooling than the expectation placed on it. Organisations increasingly want assurance that findings correspond to real adversary behaviour rather than to a generic vulnerability count, and the reporting stage, which in most guidance ends with a list of findings and recommendations, has become the weakest link between testing and risk reduction.

MITRE ATT&CK: Adoption, Research and Operational Maturity

MITRE ATT&CK was designed as a curated knowledge base of adversary tactics and techniques derived from real-world observation (Strom et al., 2018; MITRE Corporation, n.d.). A joint study by the UC Berkeley Center for Long-Term Cybersecurity and McAfee found that more than 80% of surveyed enterprises use ATT&CK and that 57% use it specifically to identify gaps in deployed security controls (Basra & Kaushik, 2020). The same study is candid about the limits of that adoption: 45% of respondents cited incompatibility with existing security products as their biggest challenge, 43% found it difficult to map event-specific data onto tactics and techniques, and 61% were not correlating events across cloud, network and endpoint sources at all. Analyst research similarly reported that 48% of organisations use ATT&CK extensively, with a further 41% using it only on a limited basis (CSO Online, 2022). Adoption in name is high; operational maturity is considerably lower.

Academic work reflects the same pattern. Systematic reviews of more than fifty ATT&CK-based contributions show that research concentrates on threat intelligence, detection and modelling, with comparatively little attention to how ATT&CK should drive remediation decisions (Al-Sada et al., 2025; Roy et al., 2023). Xiong et al. (2022) built a threat-modelling language on the Enterprise ATT&CK matrix to simulate attacks against system models, and Ahmed et al. (2022) proposed an ATT&CK-driven cyber risk assessment that links techniques to likelihood and impact. Georgiadou et al. (2021) connected ATT&CK techniques to organisational security culture, showing that human and organisational factors condition technical exposure. Work on automated extraction of tactics and techniques from threat reports addresses the mapping burden on the intelligence side (Rahman et al., 2023), but there is little corresponding work on mapping the output of practitioner testing tools to techniques and then to specific controls.

Kill Chain Models

The Cyber Kill Chain introduced by Hutchins et al. (2011) describes an intrusion as seven sequential phases: reconnaissance, weaponisation, delivery, exploitation, installation, command and control, and actions on objectives. It remains in wide use as a communication tool precisely because it is simple and linear, and it is the model most IT leaders and boards can follow without prior security training. Its simplicity is also its limitation. It foregrounds the perimeter stages and gives little detail about the post-compromise activity in which advanced persistent threats spend most of their time (Alshamrani et al., 2019). The Unified Kill Chain extends the model to eighteen phases to capture pivoting and lateral movement more faithfully (Pols, 2017). This paper retains the original Kill Chain for executive communication alongside ATT&CK for engineering detail, and treats that choice as a limitation rather than an endorsement.

Vulnerability Prioritisation Beyond Severity Scores

CVSS measures the technical severity of a vulnerability rather than its exploitability or business impact in a specific environment; the CVSS v4.0 specification itself distinguishes base severity from threat and environmental metrics that organisations must supply (FIRST, 2023). Spring et al. (2021) argue that CVSS is routinely misused as a risk score and that its outputs are poorly suited to remediation ordering. The Exploit Prediction Scoring System estimates the probability that a vulnerability will be exploited in the wild (Jacobs et al., 2021), an important advance, but it remains a global estimate rather than evidence of what was demonstrated inside a given network. Neither approach captures blast radius: how far an attacker could move laterally once a particular weakness is used.

Segmentation, Zero Trust and Continuous Validation

On the defensive side, network segmentation and zero trust architecture have moved from aspirational best practice to an area of active investment. NIST SP 800-207 defines zero trust as removing implicit trust based on network location (Rose et al., 2020), and surveys document both the promise and the practical difficulty of migrating legacy estates to it (He et al., 2022; Syed et al., 2022). Microsegmentation platforms have grown around the argument that perimeter defence alone cannot contain lateral movement. Breach and attack simulation (BAS) platforms occupy a related space: rather than testing once a year, they validate specific ATT&CK techniques against live defences on an ongoing basis. Both trends support the argument of this paper, but both remain priced and resourced mainly for larger enterprises.

Threat Landscape and Governance Context

The threat side has become harder to ignore for anyone treating hardening as periodic housekeeping. In the 2025 DBIR, third-party involvement in breaches doubled to 30%, exploitation of vulnerabilities as an initial access vector rose by 34%, edge device exploitation grew almost eight-fold, and a human element was present in around 60% of breaches (Verizon, 2025). Healthcare environments are particularly exposed because availability-critical clinical systems are difficult to patch or test (Argaw et al., 2020). Governance frameworks such as ISO/IEC 27001:2022, the NIST Risk Management Framework and the NIST Cybersecurity Framework 2.0 require that risks are identified and treated (ISO/IEC, 2022; Joint Task Force, 2018; NIST, 2024), but their audit evidence typically demonstrates that a test was performed rather than that a specific attack path was closed. Table 1 summarises the literature most relevant to this study.

Source	Focus and contribution	Limitation addressed by this study
Hutchins et al. (2011)	Seven-stage intrusion kill chain for intelligence-driven defence	Linear and perimeter-centred; weak on lateral movement
Strom et al. (2018); MITRE (n.d.)	ATT&CK design philosophy and knowledge base	Describes behaviour; does not prescribe remediation order
Basra & Kaushik (2020)	Enterprise survey of ATT&CK use and challenges	Shows mapping difficulty but offers no operational method
Xiong et al. (2022)	Threat modelling language built on Enterprise ATT&CK	Model-based; not tied to practitioner test evidence
Ahmed et al. (2022)	ATT&CK-driven cyber risk assessment	Risk scoring without closure re-testing
Al-Sada et al. (2025); Roy et al. (2023)	Systematic reviews of ATT&CK research	Identify a gap in remediation-oriented use of ATT&CK
Spring et al. (2021); Jacobs et al. (2021)	Critique of CVSS; exploit prediction scoring	Global estimates; no measure of lateral blast radius
Rose et al. (2020); Syed et al. (2022)	Zero trust architecture and its challenges	Architectural target state; no finding-driven path to it
Verizon (2025); IBM (2025)	Breach patterns, lifecycle and cost data	Quantify the problem; not a method for addressing it

Table 1: Summary of key literature and its relevance to this study

RESEARCH GAP AND LIMITATIONS OF CURRENT APPROACHES

Identified Research and Practice Gaps

Five specific gaps sit behind the problem described above, and each has a different root cause. **First**, penetration test reports are point-in-time and technical, structured around individual findings rather than around the attacker's path through the environment, so remediation teams cannot easily see which fixes close a route to a sensitive asset and which merely tidy up a low-value host. **Second**, ATT&CK adoption is broad but shallow: the framework is referenced in reports and dashboards, yet fewer than half of adopters correlate events across their own environment well enough to use it operationally (Basra & Kaushik, 2020). **Third**, the Cyber Kill Chain used on its own reinforces perimeter-centric thinking, foregrounding the delivery and exploitation stages a firewall is built to stop and giving little attention to what happens after initial compromise, which is precisely where system-intrusion data shows most breaches now play out (Verizon, 2025).

Fourth, segmentation is usually scoped as a single large capital project, a full network re-architecture, rather than as an incremental set of changes tied to specific findings. This makes it easy to defer indefinitely and, when partially completed, can create false confidence that is arguably worse than an honestly flat network. **Fifth**, continuous validation tooling exists and works, but is priced and resourced for enterprise budgets; mid-market and public sector organisations, including many healthcare providers most exposed to third-party and credential-based attack patterns, are effectively priced out of the tooling built to address their risk.

Root Causes

These gaps are not simply a matter of insufficient awareness. Technically, most SIEM, firewall and vulnerability management platforms were not built with native ATT&CK technique mapping, so translating a scan finding into a technique identifier is often a manual, spreadsheet-driven exercise that does not survive staff turnover. Economically, organisations budget far more readily for a discrete penetration test than for an ongoing validation capability, because the former fits an annual procurement cycle and the latter does not. Behaviourally, a report handed to an operations team with no shared prioritisation language competes, and usually loses, against feature delivery and other project work with clearer sponsorship. Where a regulatory driver exists, the bar it sets is frequently that a test was performed rather than that demonstrated risk was reduced, which further weakens the incentive to close the loop.

Limitations of Current Approaches

Static, CVSS-ranked reports. A critical-rated finding on an isolated test server and a medium-rated finding on a domain controller with unrestricted network reachability represent very different real-world risk, yet a CVSS-sorted list can present them in the wrong order. Findings also decay unevenly: a CVE is patched and disappears from the next report, while the architectural weakness that made its exploitation dangerous persists unremarked.

ATT&CK used as documentation rather than operational input. Many organisations produce an ATT&CK Navigator layer once, generally after an incident or audit, and treat it as a deliverable rather than a living artefact. A framework used this way documents risk without reducing it.

Perimeter-centric hardening. Firewalls and IDS/IPS are frequently tuned for north-south traffic while east-west traffic between internal segments, the route lateral movement actually takes, goes largely uninspected. A control set optimised for the perimeter is, by construction, optimised against the wrong stage of the kill chain.

One-off segmentation projects. Large segmentation programmes commonly stall midway because of the disruption of re-addressing a live environment or because the budget for a discrete project runs out, leaving an assumption of containment that the architecture does not deliver.

Compliance-driven remediation. Where ISO/IEC 27001 or similar frameworks are the primary driver, the audit trail demonstrates that a policy exists and a test was conducted, not that a specific attack path was closed. Table 2 consolidates these gaps and indicates how the proposed framework responds to each.

Gap	Evidence and root cause	MPHV response
G1: Finding-centric reporting	Point-in-time lists; no attack-path context	Map stage links each finding to a technique and Kill Chain stage
G2: Shallow ATT&CK use	Mapping difficulty and tool incompatibility (Basra & Kaushik, 2020)	Lightweight tool-to-technique mapping usable without SIEM integration
G3: Perimeter-centric thinking	System intrusion now 53% of breaches (Verizon, 2025)	Prioritise stage weights lateral blast radius explicitly
G4: Monolithic segmentation	Projects stall; partial segmentation breeds false confidence	Harden stage segments the highest-value demonstrated boundary first
G5: Unaffordable validation	BAS priced for large enterprises	Validate stage allows manual, targeted closure tests
G6: Compliance as proxy for risk	Audits evidence activity, not outcome	Every backlog item carries a named closure test

Table 2: Research gap matrix and the corresponding MPHV response

METHODOLOGY

Research Design

The study adopts a qualitative, design-oriented research approach in which the primary output is an artefact, a remediation framework, rather than a statistical test of a hypothesis. The design follows five iterative activities shown in Figure 1: problem identification from breach data and literature; definition of objectives and gaps; design of the framework and its prioritisation matrix; demonstration through a tool-to-technique mapping and illustrative cases; and evaluation through structured comparison with the traditional test-and-report cycle and a critical assessment of limitations. Findings from evaluation were fed back into the design, which is why the framework itself takes a cyclical form.

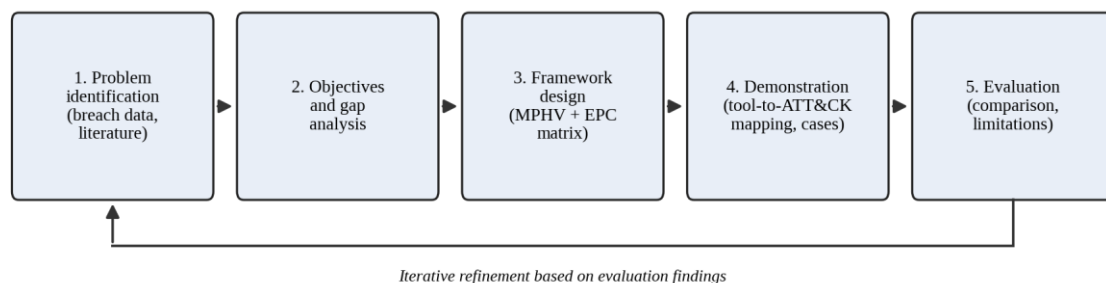


Figure 1: Research design adopted in this study

Data Sources

Three categories of evidence were used. The first is secondary industry data, chiefly the 2025 Verizon DBIR (12,195 confirmed breaches across 139 countries) and the IBM Cost of a Data Breach Report 2025 (600 breached organisations), supplemented by survey research on ATT&CK adoption (Basra & Kaushik, 2020; CSO Online, 2022). The second is the academic and standards literature summarised in Table 1, including the ATT&CK knowledge base, NIST special publications and peer-reviewed work on ATT&CK, kill chains, vulnerability scoring and zero trust. The third is practitioner knowledge accumulated across more than two decades of network engineering, healthcare IT management, enterprise penetration testing and cybersecurity training, including vulnerability assessments of web applications, APIs and backend systems and the design of attack simulations in virtual laboratory environments.

Framework Construction and Analytical Procedure

The framework was constructed by working backwards from the outcome that current approaches fail to deliver, namely a control change whose effectiveness against a named technique has been demonstrated. Four analytical procedures were defined. In the **mapping procedure**, each finding produced by Nmap, Burp Suite, Metasploit or Wireshark is assigned an ATT&CK tactic, a technique identifier and a Cyber Kill Chain stage, using the public ATT&CK knowledge base as the reference vocabulary. In the **scoring procedure**, each mapped technique is rated on two ordinal axes: likelihood, recorded as demonstrated (exploited during authorised testing) or theoretical (plausible but not demonstrated), and blast radius, recorded as large or contained according to the lateral reach and sensitivity of assets exposed by successful exploitation. In the **translation procedure**, prioritised techniques are converted into specific segmentation, firewall or IDS/IPS changes. In the **validation procedure**, each change is associated with a closure test that re-attempts the same technique.

Demonstration and Evaluation Approach

Because the study did not have access to before-and-after data from a controlled deployment, the framework is demonstrated rather than statistically tested. Demonstration uses three instruments: a mapping of common practitioner tool outputs to ATT&CK techniques (Table 5); two illustrative composite cases built from patterns typical of healthcare and enterprise engagements; and one publicly reported 2025 incident used as an external reference point. Evaluation compares MPHV with the traditional cycle across six practical dimensions (Table 6) and assesses limitations of both the framework and the study (Table 7). The composites are deliberately presented as composites rather than sourced case studies, and no claim of quantified effect is made where none can be supported.

Ethical Considerations

All offensive activity referred to in this paper is assumed to take place within clearly authorised scope and rules of engagement. Exploitation against availability-critical systems, such as clinical platforms, is assumed to occur against staging mirrors rather than live production. The composite cases contain no client-identifiable data, and all industry statistics are drawn from published sources.

RESULTS AND DISCUSSION**The Map, Prioritise, Harden, Validate (MPHV) Framework**

The principal result of the study is the MPHV framework (Figure 2, Table 3). MPHV is not a new offensive methodology; it deliberately reuses established tools and frameworks rather than proposing new ones. Its contribution is structural: a repeatable four-stage cycle that forces every finding to be mapped to a technique, every technique to be weighed by realistic impact rather than raw severity, every prioritised item to be converted into a specific control change, and every control change to be re-tested rather than assumed effective. Because Validate feeds back into Map, the model is continuous rather than a one-off project.

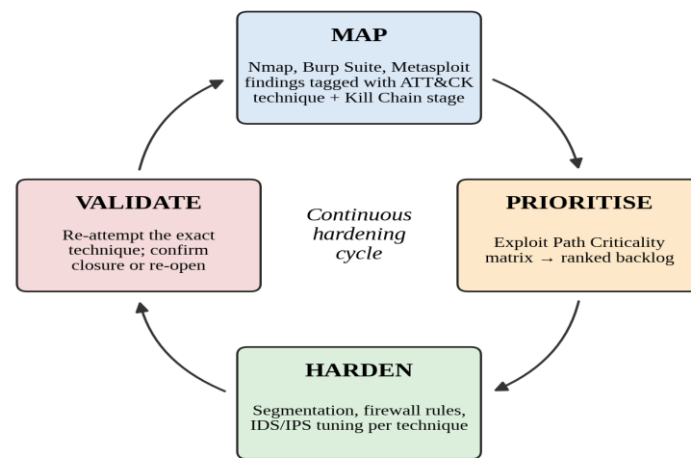


Figure 2: The MPHV continuous hardening cycle

Stage	Core activity	Primary tools / inputs	Output
Map	Enumerate the environment, demonstrate real attack paths, and tag every finding with an ATT&CK technique ID and Kill Chain stage	Nmap, Burp Suite, Metasploit, ATT&CK Navigator	Technique-mapped finding set (Navigator heatmap)
Prioritise	Score each mapped technique by demonstrated ease of exploitation and lateral blast radius, not CVSS alone	Exploit Path Criticality matrix	Ranked remediation backlog with owners
Harden	Translate top-ranked items into specific segmentation, firewall and IDS/IPS changes tied to named techniques	Segmentation design, firewall rule sets, IDS/IPS tuning	Deployed controls, each traceable to a technique ID
Validate	Re-attempt the specific technique the control was meant to stop and feed the result back into Map	Targeted re-test, purple team session or BAS platform	Confirmed closure or re-opened backlog item

Table 3: Stages of the MPHV framework

The Exploit Path Criticality Matrix

Rather than sorting findings by CVSS alone, each mapped technique is scored on two axes: demonstrated likelihood (was it actually exploited in testing, or only theoretically possible?) and blast radius (how much lateral movement or asset exposure would follow from successful exploitation?). Plotting findings on this two-by-two matrix produces four practical categories (Figure 3, Table 4).

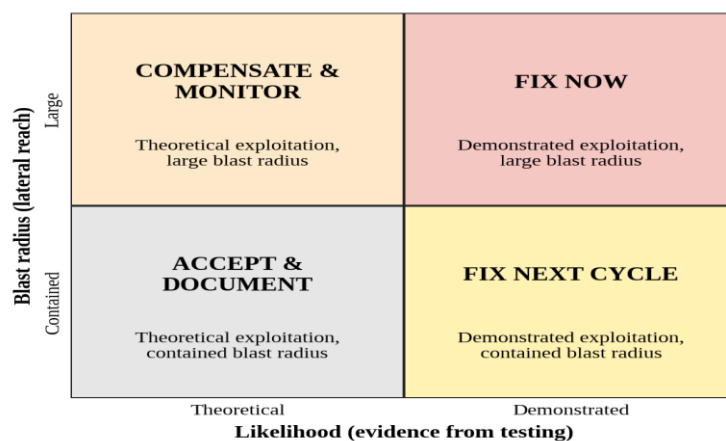


Figure 3: The Exploit Path Criticality (EPC) matrix

Quadrant	Definition	Typical action
Fix Now	Demonstrated exploitation, large blast radius (e.g. unrestricted path to a domain controller or clinical data store)	Immediate remediation, tracked at executive level
Fix Next Cycle	Demonstrated exploitation, contained blast radius	Scheduled into the next remediation sprint or quarter
Compensate & Monitor	Theoretical exploitation, large blast radius, no clean fix available immediately	Interim IDS/IPS signature or segmentation control, re-assessed at next Validate cycle
Accept & Document	Theoretical exploitation, contained blast radius	Formally risk-accepted with a named owner and review date, not silently dropped

Table 4: Exploit Path Criticality quadrants and typical actions

The matrix is not a novel scoring theory; it borrows directly from long-established likelihood-times-impact risk matrices of the kind embedded in ISO/IEC 27001 and the NIST Risk Management Framework. What differs is that both axes are tied explicitly to what was demonstrated during testing rather than to abstract asset value. This keeps the matrix grounded in evidence rather than opinion, which is where informal risk registers most often lose credibility with operations teams. It also complements rather than replaces CVSS and EPSS: those scores remain useful inputs when deciding whether an undemonstrated technique belongs in the Compensate & Monitor or Accept & Document quadrant.

Mapping Practitioner Tool Outputs to ATT&CK

Table 5 presents the mapping produced by applying the mapping procedure to the outputs that the standard toolkit typically generates. Nmap-based service and version enumeration maps most directly onto the Reconnaissance and Discovery tactics; Network Service Discovery (T1046) and Active Scanning (T1595) describe almost exactly what an authorised scan is doing. The value of the mapping is practical: every exposed service found during enumeration, such as an unpatched SMBv1 share, an exposed RDP endpoint or a legacy Telnet listener, is recorded not merely as a finding but as a named technique that adversaries are documented to use, which makes the case for closing it far easier to make to a non-specialist audience than a bare port number.

Tool	Typical finding	ATT&CK tactic / technique	Kill Chain stage
Nmap	Exposed services, versions and open ports	Reconnaissance: Active Scanning (T1595); Discovery: Network Service Discovery (T1046)	Reconnaissance
Burp Suite	Injection or access-control flaw in an internet-facing application	Initial Access: Exploit Public-Facing Application (T1190)	Delivery / Exploitation
Burp Suite	Weak authentication or session handling	Credential Access: Brute Force (T1110); Steal Web Session Cookie (T1539)	Exploitation
Burp Suite	IDOR exposing other users' records	Collection: Data from Information Repositories (T1213)	Actions on Objectives
Metasploit	Local privilege escalation on a foothold host	Privilege Escalation: Exploitation for Privilege Escalation (T1068)	Installation
Metasploit	Credential harvesting from memory or hashes	Credential Access: OS Credential Dumping (T1003)	Installation
Metasploit	Pivot to a second host with harvested credentials	Lateral Movement: Remote Services (T1021); Valid Accounts (T1078)	Command and Control / Actions on Objectives
Wireshark	Cleartext protocols on internal segments	Credential Access: Network Sniffing (T1040)	Actions on Objectives

Table 5: Mapping of practitioner tool outputs to ATT&CK techniques and Kill Chain stages

Burp Suite-driven testing of authentication, session management, injection points and access control typically surfaces findings that map onto Initial Access, particularly Exploit Public-Facing Application (T1190), and, where credential or session weaknesses are found, onto Credential Access techniques. An insecure direct object reference that exposes another user's records is not just an OWASP Top 10 item; mapped correctly, it demonstrates a specific, low-effort route to Collection that a defender can explicitly test detection coverage against.

Metasploit, used within authorised scope and, wherever service availability is a concern, within a staging mirror, demonstrates the tactics that matter most for hardening decisions: Execution, Persistence, Privilege Escalation and, critically, Lateral Movement. A successful chain from an initial foothold to a second host using harvested credentials is the single most persuasive piece of evidence a practitioner can bring to a segmentation discussion, because it replaces an abstract recommendation to segment the network with a concrete statement of the exact path used to move from, for example, the guest VLAN to the finance server.

Navigator Heatmap and Kill Chain Overlay

Every technique observed or judged plausible during an engagement is logged into an ATT&CK Navigator layer and colour-coded by whether an existing control fully mitigates it, partially mitigates it or leaves it uncovered. This single artefact serves two audiences: it is detailed enough for a security engineer to action and visual enough for an executive audience to understand without a walkthrough. Overlaying the same techniques onto the seven Kill Chain stages then shows where coverage clusters and where it thins out. In the composite engagements on which this framework draws, coverage was consistently strongest at Delivery, where perimeter email and web filtering sit, and weakest at Actions on Objectives, where data movement and exfiltration monitoring are required. This is a qualitative pattern from practitioner experience rather than a measured distribution, and it should be checked in any environment rather than assumed.

From Heatmap to Architecture: Segmentation, Firewalls and IDS/IPS

Segmentation decisions are drawn directly from the Lateral Movement and Discovery techniques recorded on the heatmap rather than from generic best-practice diagrams. If testing demonstrated an unrestricted path from a general user VLAN to a clinical or financial system, that specific boundary is prioritised for a segmentation control ahead of theoretically desirable but untested boundaries elsewhere. Firewall rule changes and IDS/IPS signature tuning follow the same logic. A signature set chosen to detect the specific lateral movement technique demonstrated in testing, for example SMB-based remote service abuse, will catch an attacker following the same path more reliably than a generic vendor default pack, and its effectiveness can be tested directly in the Validate stage by re-attempting exactly that technique. This incremental, evidence-led approach is consistent with the zero trust principle of removing implicit trust between network zones (Rose et al., 2020) while avoiding the all-or-nothing re-architecture that causes many segmentation projects to stall.

Remediation Roadmap and Closure Tests

The final methodological step converts the prioritised, technique-mapped findings into a backlog with three mandatory fields beyond the usual description and severity: an owner, a target remediation window and a named closure test, which is the specific technique that will be re-attempted to prove the fix worked. Without the third field, remediation tends to be marked complete once a ticket is closed rather than once the underlying attack path has stopped working, which is precisely the gap this framework is built to close.

Case Evidence

The two scenarios below are illustrative composites built from patterns typical of enterprise and healthcare engagements rather than from any single named client, and they are presented as such rather than as sourced fact. The third example is a real, publicly reported 2025 incident.

Composite A: healthcare provider network. A composite typical of a multi-site healthcare environment supporting more than 100 users illustrates a recurring pattern: a flat clinical network in which a single compromised administrative workstation, reached via phishing, could route directly to imaging and patient record systems with no intermediate control. Testing against a staging mirror using Nmap and Metasploit demonstrated the lateral path in full. Mapping that path to Remote Services (T1021) and Valid Accounts (T1078), placing it in the Fix Now quadrant and overlaying it on the Kill Chain gave the case for VLAN segmentation of clinical systems a concreteness that a generic recommendation had previously failed to achieve, and the segmentation work was approved and scheduled within the same budget cycle.

Composite B: enterprise web application. Burp Suite testing of an internally facing enterprise web application surfaced an insecure direct object reference and weak session management, mapped to Initial Access, Collection and Credential Access techniques. Once prioritised through the Exploit Path Criticality matrix as Fix Now, because exploitation was demonstrated and the application offered a direct path to a shared API gateway, the fix was actioned ahead of eight-month-old, lower-impact web application firewall tuning tickets that had been sitting in the same backlog with no clear prioritisation basis.

Real-world reference: 2025 UK retail attacks. In April 2025, Marks & Spencer and the Co-operative Group suffered disruptive cyber attacks, with Harrods also targeted around the same time. The Cyber Monitoring Centre

classified the M&S and Co-op incidents as a single combined Category 2 systemic event with an estimated financial impact of GBP 270 million to 440 million, and reporting indicates that initial access relied on social engineering of IT help desks, with the Scattered Spider group believed to be responsible (Cyber Monitoring Centre, 2025; Lakshmanan, 2025). The case is instructive precisely because it sits largely outside the scope of a purely technical hardening framework. It shows why the early Kill Chain stages must include social engineering and third-party access paths, and why the 30% third-party involvement reported in the 2025 DBIR is a segmentation and access-review problem as much as a technology problem.

Comparative Evaluation and Value Proposition

Table 6 compares the traditional penetration-test-and-report cycle with MPHV across six practical dimensions. The central claim of MPHV is traceability: every control change can be traced to a demonstrated technique, and every technique's closure can be independently re-tested rather than assumed. That traceability has a plausible, though not precisely quantifiable, relationship to the cost data reviewed earlier. The USD 1.14 million difference between breaches contained within 200 days and those that take longer (IBM, 2025) suggests that keeping segmentation and detection focused on demonstrated lateral movement paths is a reasonable route to faster containment; however, the saving for any given organisation cannot be asserted without its own before-and-after data and should be treated as illustrative rather than guaranteed.

Dimension	Traditional pentest-and-report cycle	MPHV framework
Prioritisation basis	CVSS severity score	Demonstrated exploitability and lateral blast radius
Output format	Static PDF report	Living ATT&CK Navigator heatmap and backlog
Traceability	Finding to fix, informally	Finding → technique → control → named closure test
Revalidation	Next annual engagement	Continuous or quarterly, tied to backlog closure
Executive communication	Count of findings by severity	Kill Chain coverage view with plain-language gaps
Cost profile	One discrete engagement fee	Base engagement plus lower-cost incremental validation

Table 6: Comparison of the traditional cycle and the MPHV framework

Beyond cost, the more defensible benefit is organisational. The Navigator heatmap serves simultaneously as an engineering backlog and a board-level risk narrative, removing the usual translation step in which a technical report is separately rewritten into an executive summary that then drifts from what engineers are actually working on.

Implementation Considerations

Several conditions influence successful adoption. **Scope and rules of engagement** must define explicitly where exploitation is permitted against live production systems and where staging mirrors are required, particularly for clinical or financial platforms, with change windows agreed in advance. **ATT&CK competence, not just awareness**, is needed: because mapping data to techniques is the most commonly cited adoption difficulty (Basra & Kaushik, 2020), organisations should budget for training and a Navigator-literate owner. **Tooling should start light**: a spreadsheet or the free Navigator tool is sufficient to begin the Map stage, and smaller organisations should not wait for SIEM-native integration. **Segmentation should be incremental**, ring-fencing first the single highest-value boundary identified in testing. **Governance** requires a named backlog owner, a fixed Validate cadence (quarterly is a reasonable default for mid-market organisations) and a standing steering-group agenda item that reviews the heatmap rather than a raw finding count. Finally, **the budget case** is more persuasive when built on containment-time cost data than on abstract risk language.

Risks, Challenges and Limitations of the Proposed Framework

No framework of this kind is complete, and it is important to be direct about where MPHV falls short. **ATT&CK is not predictive**. It is a knowledge base of documented behaviour; a fully coloured heatmap demonstrates coverage of known techniques and says nothing about a genuinely novel or zero-day technique. Treating a complete heatmap as equivalent to being safe is a real failure mode of adopting this framework uncritically.

Scoring depends on honesty. Asset owners have an incentive to argue their systems into the Accept & Document quadrant, and the matrix provides no structural defence against that pressure beyond requiring a named owner and review date, the same discipline that many risk registers fail to sustain. The binary axes are also coarse, and two techniques in the same quadrant may differ materially in risk.

Exploitation carries operational risk. Simulated exploitation against legacy clinical or industrial systems can cause disruption. Staging mirrors reduce that risk but add cost and time, and organisations without a representative staging environment, common in under-resourced healthcare and public sector settings, often face a choice between testing live and not testing that path at all.

Continuous validation remains aspirational for resource-constrained organisations. BAS platforms are priced for enterprise budgets, and manual re-testing depends on retaining skilled staff in a profession with a well-documented skills shortage (ISC2, 2024). The framework makes the resourcing gap visible; it does not solve it.

Asset visibility is a prerequisite. Organisations cannot segment what they have not mapped, and asset discovery in legacy, operational technology and clinical estates is frequently incomplete, in which case the Map stage must begin with a basic inventory exercise. **The Kill Chain is dated.** It maps poorly onto cloud-native and identity-centric attack paths with no clear perimeter, and more granular alternatives such as the Unified Kill Chain exist (Pols, 2017); retaining the original model is a pragmatic concession to audience familiarity. **The human element is out of scope.** With a human element present in around 60% of breaches (Verizon, 2025), a perfectly segmented, ATT&CK-mapped network can still be compromised by a convincing call to a help desk, as the 2025 retail incidents illustrate.

Limitations of the Study

The study itself has limitations that bound the strength of its conclusions. These are summarised in Table 7 together with the steps taken to mitigate them and the further work needed to remove them.

Limitation	Implication	Mitigation / further work
No controlled empirical evaluation	Effect of MPHV on dwell time, cost or breach likelihood is not measured	Claims stated as plausible, not proven; longitudinal multi-site study proposed
Composite rather than sourced cases	Cases illustrate but cannot verify the framework	Composites labelled explicitly; one public incident used as an external reference
Reliance on industry reports	Vendor-produced data may carry sampling and commercial bias	Two independent primary reports used; figures cited as reported
Single-practitioner perspective	Design choices may reflect one practitioner's experience	Grounded in published frameworks; expert panel validation proposed
Coarse binary scoring axes	Limited discrimination within quadrants	Future ordinal or weighted scales with inter-rater reliability testing
Scope excludes cloud and human factors	Generalisability to cloud-native and social-engineering paths is limited	Extension to ATT&CK cloud matrices and identity attack paths proposed

Table 7: Limitations of the study and mitigation

In particular, the absence of quantitative before-and-after data means that the relationships drawn between MPHV and reduced breach lifecycles are argued from secondary evidence rather than observed directly. Generalisability is also constrained: the framework was conceived for on-premises and hybrid enterprise and healthcare networks of moderate size, and its fit for very large, cloud-first or operational technology environments has not been examined.

FUTURE WORK

Several developments over the next three to five years are likely to shape how this framework must evolve, although some, particularly regulatory trajectory, remain unsettled. AI-assisted defence will increasingly automate parts of the Map and Validate stages; organisations using AI and automation extensively already report breach lifecycles shortened by roughly 80 days, yet 63% of organisations lack a formal AI governance policy (IBM, 2025), which means the Map stage must treat AI and machine learning endpoints and agentic tooling as first-class assets. Microsegmentation and zero trust platforms are likely to lower the cost of the Harden stage, at the price of new vendor integration and lock-in considerations not evaluated here.

ATT&CK itself continues to evolve, with expanding cloud matrices and new stewardship arrangements including the ATT&CK Advisory Council (MITRE, 2026), so a single mapping exercise is obsolete almost as soon as it is finished and periodic re-mapping must be built into governance. Regulatory pressure in the United Kingdom and European Union appears to be moving gradually towards treating continuous testing and segmentation as baseline expectations for essential services, and the cybersecurity skills shortage shows no sign of resolving (ISC2, 2024), increasing the importance of managed or hybrid validation services.

For research, four directions follow directly from the limitations above: a longitudinal, multi-organisation evaluation measuring dwell time and closure rates before and after MPHV adoption; validation of the Exploit Path Criticality matrix by an expert panel with inter-rater reliability testing; automation of tool-to-technique mapping using structured output parsing and language models; and extension of the framework to cloud, identity-centric and social-engineering attack paths using the Unified Kill Chain or attack-flow representations.

CONCLUSION

This paper has argued that penetration testing findings reduce risk only when they are mapped to a shared behavioural model, prioritised by demonstrated impact rather than abstract severity, converted into specific and traceable control changes, and re-tested to confirm that each change actually works. MITRE ATT&CK and the Cyber Kill Chain provide the vocabulary; Nmap, Burp Suite and Metasploit already generate the evidence base most organisations need. What has generally been missing is the disciplined, continuous cycle connecting the two.

The study's contributions are threefold: a consolidated analysis of the gaps between ATT&CK adoption and operational hardening; the MPHV framework with its Exploit Path Criticality matrix and closure-test discipline; and a practical tool-to-technique mapping that allows organisations to begin without expensive integration. Its limitations are equally clear. The framework has been demonstrated rather than empirically evaluated, its cases are composites, it inherits the predictive limits of ATT&CK and the linearity of the Kill Chain, and it does not address the human and third-party attack paths that featured prominently in recent incidents. MPHV is therefore offered as a starting structure for continuous hardening, not as a finished or universally applicable solution.

For practitioners holding an existing penetration test report, the practical next step is not to commission another test. It is to take the last report, map its findings to ATT&CK techniques and Kill Chain stages, and ask of each one: if this were exploited today, what specific control would stop it, and has that control ever been tested against this specific technique? Where the honest answer is no, that is the backlog to start with.

NOMENCLATURE

Abbreviation	Meaning
ATT&CK	Adversarial Tactics, Techniques and Common Knowledge; MITRE's knowledge base of documented adversary behaviour
BAS	Breach and Attack Simulation; platforms that continuously test defences against specific techniques
CVSS	Common Vulnerability Scoring System; an industry-standard severity score for individual vulnerabilities
EPC	Exploit Path Criticality; the two-axis prioritisation matrix proposed in this paper
EPSS	Exploit Prediction Scoring System; a probability estimate of exploitation in the wild
IDS/IPS	Intrusion Detection System / Intrusion Prevention System
IDOR	Insecure Direct Object Reference; an access-control flaw exposing other users' data
MPHV	Map, Prioritise, Harden, Validate; the framework proposed in this paper
TTP	Tactics, Techniques and Procedures
Dwell time	Period between an attacker's initial access and detection
Lateral movement	Movement from an initially compromised host to other systems within a network

Table 8: Abbreviations used in this paper

CONFLICT OF INTEREST

The author declares no conflict of interest.

FUNDING

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

DATA AVAILABILITY STATEMENT

No new datasets were generated in this study. The industry statistics analysed are available in the published sources cited in the reference list, and the composite cases are illustrative constructions that contain no client data.

REFERENCES

- 1) Ahmed, M., Panda, S., Xenakis, C., & Panaousis, E. (2022). MITRE ATT&CK-driven cyber risk assessment. In *Proceedings of the 17th International Conference on Availability, Reliability and Security (ARES '22)*. ACM. <https://doi.org/10.1145/3538969.3544420>
- 2) Al-Sada, B., Sadighian, A., & Oligeri, G. (2025). MITRE ATT&CK: State of the art and way forward. *ACM Computing Surveys*, 57(1), Article 12. <https://doi.org/10.1145/3687300>
- 3) Alshamrani, A., Myneni, S., Chowdhary, A., & Huang, D. (2019). A survey on advanced persistent threats: Techniques, solutions, challenges, and research opportunities. *IEEE Communications Surveys & Tutorials*, 21(2), 1851–1877. <https://doi.org/10.1109/COMST.2019.2891891>
- 4) Argaw, S. T., Troncoso-Pastoriza, J. R., Lacey, D., Florin, M.-V., Calcavecchia, F., Anderson, D., Burleson, W., Vogel, J.-M., O'Leary, C., Eshaya-Chauvin, B., & Flahault, A. (2020). Cybersecurity of hospitals: Discussing the challenges and working towards mitigating the risks. *BMC Medical Informatics and Decision Making*, 20, Article 146. <https://doi.org/10.1186/s12911-020-01161-7>
- 5) Basra, J., & Kaushik, T. (2020). MITRE ATT&CK as a framework for cloud threat investigation. Center for Long-Term Cybersecurity, UC Berkeley. <https://cltc.berkeley.edu/publication/mitre-attck/>
- 6) CSO Online. (2022). *The changing role of the MITRE ATT&CK framework*. <https://www.csoonline.com/>
- 7) Cyber Monitoring Centre. (2025, June 20). *Cyber Monitoring Centre statement on ransomware incidents in the retail sector – June 2025*. <https://cybermonitoringcentre.com/2025/06/20/cyber-monitoring-centre-statement-on-ransomware-incidents-in-the-retail-sector-june-2025/>
- 8) FIRST. (2023). *Common Vulnerability Scoring System version 4.0: Specification document*. Forum of Incident Response and Security Teams. <https://www.first.org/cvss/v4.0/specification-document>
- 9) Georgiadou, A., Mouzakitis, S., & Askounis, D. (2021). Assessing MITRE ATT&CK risk using a cybersecurity culture framework. *Sensors*, 21(9), 3267. <https://doi.org/10.3390/s21093267>
- 10) He, Y., Huang, D., Chen, L., Ni, Y., & Ma, X. (2022). A survey on zero trust architecture: Challenges and future trends. *Wireless Communications and Mobile Computing*, 2022, Article 6476274. <https://doi.org/10.1155/2022/6476274>
- 11) Hutchins, E. M., Cloppert, M. J., & Amin, R. M. (2011). Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains. *Leading Issues in Information Warfare & Security Research*, 1(1), 80–106.
- 12) IBM. (2025). *Cost of a data breach report 2025*. IBM Security and Ponemon Institute. <https://www.ibm.com/reports/data-breach>
- 13) ISC2. (2024). *ISC2 cybersecurity workforce study 2024*. <https://www.isc2.org/research>
- 14) ISO/IEC. (2022). *ISO/IEC 27001:2022 Information security, cybersecurity and privacy protection – Information security management systems – Requirements*. International Organization for Standardization.
- 15) Jacobs, J., Romanosky, S., Edwards, B., Adjerid, I., & Roytman, M. (2021). Exploit prediction scoring system (EPSS). *Digital Threats: Research and Practice*, 2(3), Article 20. <https://doi.org/10.1145/3436242>
- 16) Joint Task Force. (2018). *Risk management framework for information systems and organizations: A system life cycle approach for security and privacy* (NIST Special Publication 800-37, Rev. 2). National Institute of Standards and Technology. <https://doi.org/10.6028/NIST.SP.800-37r2>
- 17) Lakshmanan, R. (2025, June 21). Scattered Spider behind cyberattacks on M&S and Co-op, causing up to \$592M in damages. *The Hacker News*. <https://thehackernews.com/2025/06/scattered-spider-behind-cyberattacks-on.html>

- 18) MITRE. (2026). *MITRE announces formation of the ATT&CK Advisory Council*. <https://www.mitre.org/news-insights>
- 19) MITRE Corporation. (n.d.). *MITRE ATT&CK knowledge base*. Retrieved September 2026, from <https://attack.mitre.org/>
- 20) NIST. (2024). *The NIST cybersecurity framework (CSF) 2.0* (NIST CSWP 29). National Institute of Standards and Technology. <https://doi.org/10.6028/NIST.CSWP.29>
- 21) Pols, P. (2017). *The unified kill chain* [Master's thesis, Cyber Security Academy]. <https://www.unifiedkillchain.com/>
- 22) Rahman, M. R., Mahdavi-Hezaveh, R., & Williams, L. (2023). What are the attackers doing now? Automating cyberthreat intelligence extraction from text on pace with the changing threat landscape: A survey. *ACM Computing Surveys*, 55(12), Article 241. <https://doi.org/10.1145/3571726>
- 23) Rose, S., Borchert, O., Mitchell, S., & Connelly, S. (2020). *Zero trust architecture* (NIST Special Publication 800-207). National Institute of Standards and Technology. <https://doi.org/10.6028/NIST.SP.800-207>
- 24) Roy, S., Panaousis, E., Noakes, C., Laszka, A., Panda, S., & Loukas, G. (2023). *SoK: The MITRE ATT&CK framework in research and practice* (arXiv:2304.07411). arXiv. <https://doi.org/10.48550/arXiv.2304.07411>
- 25) Scarfone, K., Souppaya, M., Cody, A., & Orebaugh, A. (2008). *Technical guide to information security testing and assessment* (NIST Special Publication 800-115). National Institute of Standards and Technology. <https://doi.org/10.6028/NIST.SP.800-115>
- 26) Spring, J., Hatleback, E., Householder, A., Manion, A., & Shick, D. (2021). Time to change the CVSS? *IEEE Security & Privacy*, 19(2), 74–78. <https://doi.org/10.1109/MSEC.2020.3044475>
- 27) Strom, B. E., Applebaum, A., Miller, D. P., Nickels, K. C., Pennington, A. G., & Thomas, C. B. (2018). *MITRE ATT&CK: Design and philosophy* (Technical Report MP180360). The MITRE Corporation.
- 28) Syed, N. F., Shah, S. W., Shaghaghi, A., Anwar, A., Baig, Z., & Doss, R. (2022). Zero trust architecture (ZTA): A comprehensive survey. *IEEE Access*, 10, 57143–57179. <https://doi.org/10.1109/ACCESS.2022.3174679>
- 29) Verizon. (2025). *2025 data breach investigations report*. Verizon Business. <https://www.verizon.com/business/resources/reports/dbir/>
- 30) Xiong, W., Legrand, E., Åberg, O., & Lagerström, R. (2022). Cyber security threat modeling based on the MITRE Enterprise ATT&CK Matrix. *Software and Systems Modeling*, 21, 157–177. <https://doi.org/10.1007/s10270-021-00898-7>