Grinrey Conference Proceedings
ICMESR 2026, India, 04-05 May 2026
Volume 01 (2026) 001501
https://doi.org/10.55084/gcp/001501

AI-Enabled Framework for Continuous Cybersecurity Awareness in Healthcare SMEs: A Behavior-Centered Approach

Kalyan Killari[1], Rahul Azmeera[1,*] and Abuzar Hamza[2]

[1]Information Technology, University of the Cumberlands, Williamsburg, USA

[2]7-Eleven Inc, Irving, USA

Emails: rahulazmeera@gmail.com (*Corresponding Author)

Received : 28 Apr 2026; Accepted : 02 Aug 2026; Published : 08 Aug 2026

Abstract. Cybersecurity awareness programs in healthcare small and medium-sized enterprises (SMEs) continue to rely on static, periodic training that offers minimal evaluation of long-term behavioral change. Research indicates that over 90% of security training content is forgotten shortly after one-time sessions, yet existing approaches rarely employ artificial intelligence for real-time assessment or behavioral reinforcement. This paper proposes an AI-enabled framework for continuous evaluation of security awareness programs in healthcare SMEs. Through a systematic review of 20 peer-reviewed studies published between 2018 and 2025, we synthesize how machine learning is applied to monitor employee actions, identify risk behaviors such as phishing susceptibility and policy violations, and deliver real-time feedback through nudges and personalized alerts. Grounded in Kirkpatrick's Four Levels of Evaluation and Nudge Theory, our framework categorizes AI tools by their ability to measure and influence security outcomes across knowledge, behavior, and incident reduction. Findings indicate that AI-based feedback loops significantly enhance training effectiveness by reinforcing lessons in context, identifying vulnerable users, and enabling curriculum adaptation. For resource-constrained healthcare SMEs, the proposed framework provides scalable, behavior-informed interventions that sustain organizational cybersecurity readiness and foster a human-centric security culture.

Keywords: *Cybersecurity awareness; artificial intelligence; machine learning; healthcare SMEs; behavioral nudges.*

1. INTRODUCTION

Small and medium-sized healthcare organizations occupy a uniquely vulnerable position in today's threat landscape. They hold some of the most sensitive personal data imaginable, patient records, billing histories, medication logs yet they often run cybersecurity programs that would not look out of place in a mid-2000s compliance manual. HIPAA and GDPR impose real obligations [1, 2], but meeting those obligations without a dedicated security team is hard. When something goes wrong, it is almost always a person who opened the wrong email or clicked the wrong link [3, 4].

Annual training sessions have been the default response to this problem for decades, and by most measures they do not work. Research puts the forgetting rate above 90% within weeks of a one-time session [5]. Employees who passed the phishing quiz in January still click on convincing lures in March not because they are careless, but because a 45-minute slideshow does not rewire the instincts that social engineers exploit [6]. The knowledge is there. The habit is not. The critical missing part is the feedback loop. An

© 2026.This is an open access article and allows to use the as per CC BY-NC-SA 4.0 license ISBN: 978-81-996627-3-5

employee has minimal knowledge that the training provided was covering all the data types they handle every day unless an incident happens. The organizations were also in a similar black box without human behavior signals and hard to assess the impact of training on security incidents. That gap between training delivery and observable behavior is the central problem this paper tries to address.

AI and machine learning have made enormous inroads in the technical side of cybersecurity threat detection, anomaly flagging, malware analysis but the human side has received far less attention. A tool that can detect the network intrusion within milliseconds has little impact if the employee intentionally sends the unencrypted patient file in email. The field is slow at analytical approaches to detect human behaviors, patterns in real time. What we propose here is a framework that tries to close the gap by AI-driven architecture that monitors users behavior signals and scores the risk level and provides targeted feedback, nudges, alerts and training. The goal is not surveillance but coaching.

Static Versus Continuous Awareness Programs: Ask most healthcare SME employees what cybersecurity training looks like and the answer is predictable: a module that shows up in the LMS once a year, takes about an hour, and ends with a quiz that most people pass by clicking through until they find the right answer. Nobody designed this system to fail; it emerged from the practical reality that security awareness had to fit inside compliance workflows that were never built around learning. The result is an approach that satisfies auditors without particularly changing behavior [3, 7]. For years the research literature argued that continuous feedback with timely prompts connected to the security principles while they are making decisions [9, 10].

AI and Machine Learning in Cybersecurity Behavior Monitoring: Behavioral monitoring in security contexts is not new, but AI has substantially changed what is possible. A few examples: Phishing simulation platforms like KnowBe4 and HoxHunt now analyse behavioral data and build risk profiles [6]. NLP-based analysis has been applied to understand how well the employees understand security concepts 8] [11]. Anomaly detection models flag unusual signals like file download at 2 am, bulk data export to unrecognized drive and lock the accounts [12, 13].

Feedback Loops and Behavioral Nudges: The best moment to teach someone about a security risk is when they are in the middle of taking one. Behavioral nudges [14] operationalize this idea, rather than blocking an action outright which creates frustration, the workaround nudges surface information that helps the user make a better decision on their own. Security alerts by specific behaviors can be mapped to individual risk profiles and provide frequency of intervention [6]. Micro training prompts and a quick question immediately after a violation shown measurable impact on behavior [15]. Systematic review found that employees who received nudges at the precise moment of a risky action were less likely to repeat that behavior at next instance [16]. AI-driven systems make it possible to deliver just-in-time intervention at scale [9, 17, 18].

Theoretical Frameworks: Kirkpatrick's Four-Level Evaluation Model [19] was originally developed for corporate training assessment, but it maps onto security awareness. Level 1 (Reaction) captures how employees respond to feedback and training. Level 2 (Learning) measures knowledge change. Level 3 (Behavior) whether what was learned is actually applied on the job. Level 4 (Results) did any of this reduce incidents, lower the cost of breaches? Nudge Theory [14], as developed by Thaler and Sunstein, holds that the architecture of a choice environment shapes decisions in predictable ways and that small, well-designed prompts can steer behavior toward better outcomes without taking options away.

Technical cybersecurity AI is mature. You can buy a product today that detects network intrusions, classifies malware, or flags anomalous logins with impressive accuracy. What you cannot easily buy is a system that applies those same capabilities to the human behavioral layer, identifies who is at risk before

an incident happens, and delivers feedback. Particularly within the budget constraints of a small healthcare organization, it does not yet exist as a packaged solution. This paper describes what such a system should look like.

2. METHODOLOGY

We conducted a systematic literature review following PRISMA guidelines, covering AI applications in cybersecurity awareness and behavioral monitoring. Searches ran across Google Scholar, Scopus, ScienceDirect, and MDPI between May and September 2025, using the query: ("AI" OR "machine learning") AND ("cybersecurity awareness") AND ("behavior" OR "feedback") AND ("SMEs" OR "healthcare").

To be included, a study had to meet four criteria: peer-reviewed and published between 2018 and 2025; written in English; focused specifically on AI or ML in cybersecurity awareness or behavioral monitoring (not purely technical intrusion detection); and relevant to SME or healthcare contexts. Pure opinion pieces, white papers, and blog content were excluded regardless of how they showed up in the search results.

Screening proceeded in two passes. Titles and abstracts were reviewed first; anything that clearly did not meet the criteria was set aside. The remaining articles went to full-text review with thematic coding, extracting the AI technique used, the feedback mechanism described, the behavioral outcomes measured, and the organizational context. The process is illustrated in Fig. 1. Twenty studies cleared both passes and formed the final analytical sample. We measured outcomes primarily through behavioral indicators rather than self-reported attitudes: phishing click rate changes, policy violation frequency, and voluntary threat reporting rates all drawn from log data, simulation results, and audit trail analysis. Attitude surveys are useful, but they are a poor proxy for what people actually do.

2.1 Proposed Framework: SENTINEL

2.1.1 Framework Overview

We call the framework SENTINEL Secure, Explainable, Nudge-driven Training and Intelligence Network for Employee Learning. The name is intentional: the system is meant to watch, but also to teach. It operates across four layers that feed into one another: behavioral data collection, AI-powered risk assessment, real-time adaptive feedback, and continuous improvement analytics [16, 17].

Fig. 2 maps those layers and shows how data moves through the system from raw behavioral signals, through risk scoring, out to targeted interventions, and back into a learning loop that updates the models and the curriculum.

2.1.2 Core Components

Component 1: Behavioral Data Collection (Privacy-by-Design)

SENTINEL draws behavioral signals from four sources:

- Email metadata sender, recipient, attachment presence. Content is never captured or stored.
- Web activity logs URL patterns and domain reputation checks
- System access records file downloads, access timestamps, location data
- Training platform interactions quiz scores, completion rates, simulation click patterns

Every collection decision was made with the assumption that employees would eventually see exactly what is being collected about them. That assumption drives a set of hard constraints:

- User identifiers are pseudonymized before any analysis

- Data is aggregated at the department or role level wherever individual attribution is unnecessary
- No keystroke logging. No screen recording. Full stop.
- Staff receive clear documentation of what is monitored and why before deployment begins
- Data is retained on a 90-day rolling window, consistent with HIPAA and GDPR obligations [1, 2]

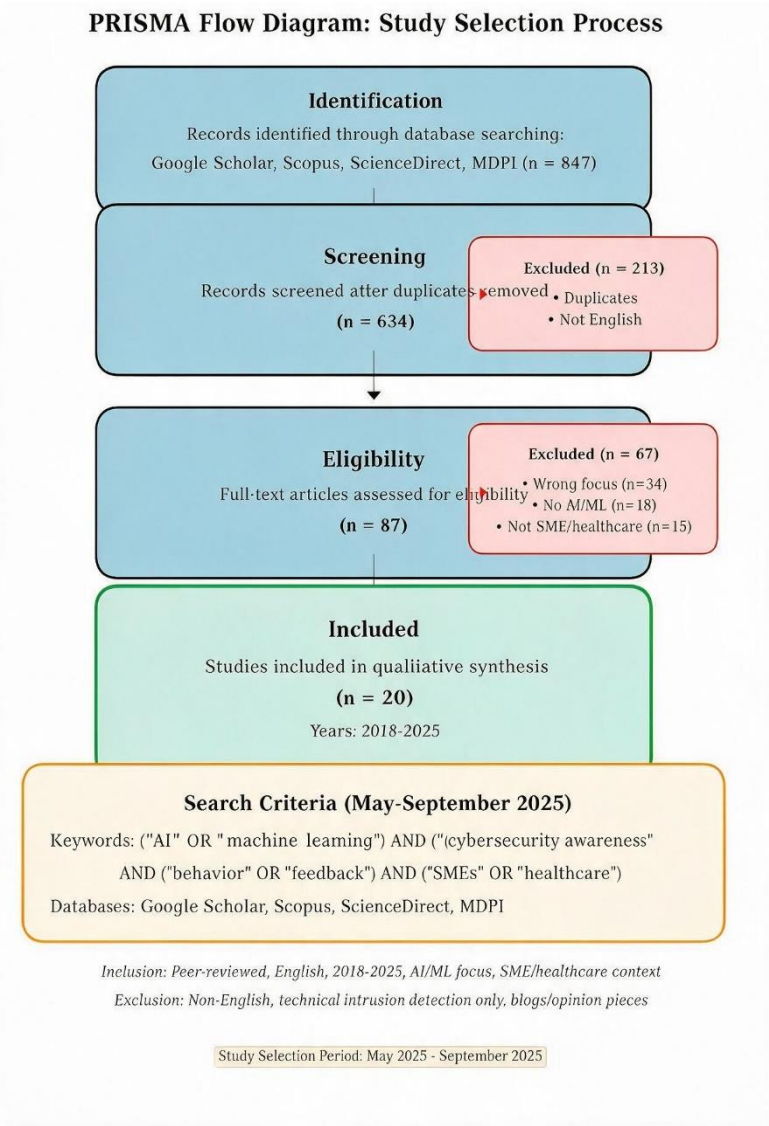

Fig. 1. PRISMA flow diagram showing the systematic literature review selection process. Initial database searches across Google Scholar, Scopus, ScienceDirect, and MDPI identified 147 records; after deduplication and sequential screening, 20 studies met the final inclusion criteria.

Component 2: AI-Powered Risk Assessment Engine

Three model types work in parallel inside the risk assessment layer:

- **Phishing Susceptibility Predictor:** A Random Forest classifier trained on simulation history, click patterns, and voluntary report behaviors. Output is a 0–100 risk score per user, updated continuously as new behavioral data arrives.
- **Policy Violation Detector:** An Isolation Forest anomaly model that ingests access logs, file transfer records, and encryption status signals. Unusual patterns surface as real-time alerts rather than end-of-week reports.

- **Behavioral Risk Profiler:** A K-means clustering model that segments the user population into Low, Medium, and High-risk groups based on aggregated behavioral metrics. This segmentation drives how interventions are calibrated across the organization.

All three models retrain on a weekly cycle using fresh behavioral data. Decision thresholds are not fixed they adjust automatically based on observed false-positive rates and are validated continuously against actual incident records. A threshold that generates too many false alarms gets loosened; one that misses real events gets tightened.

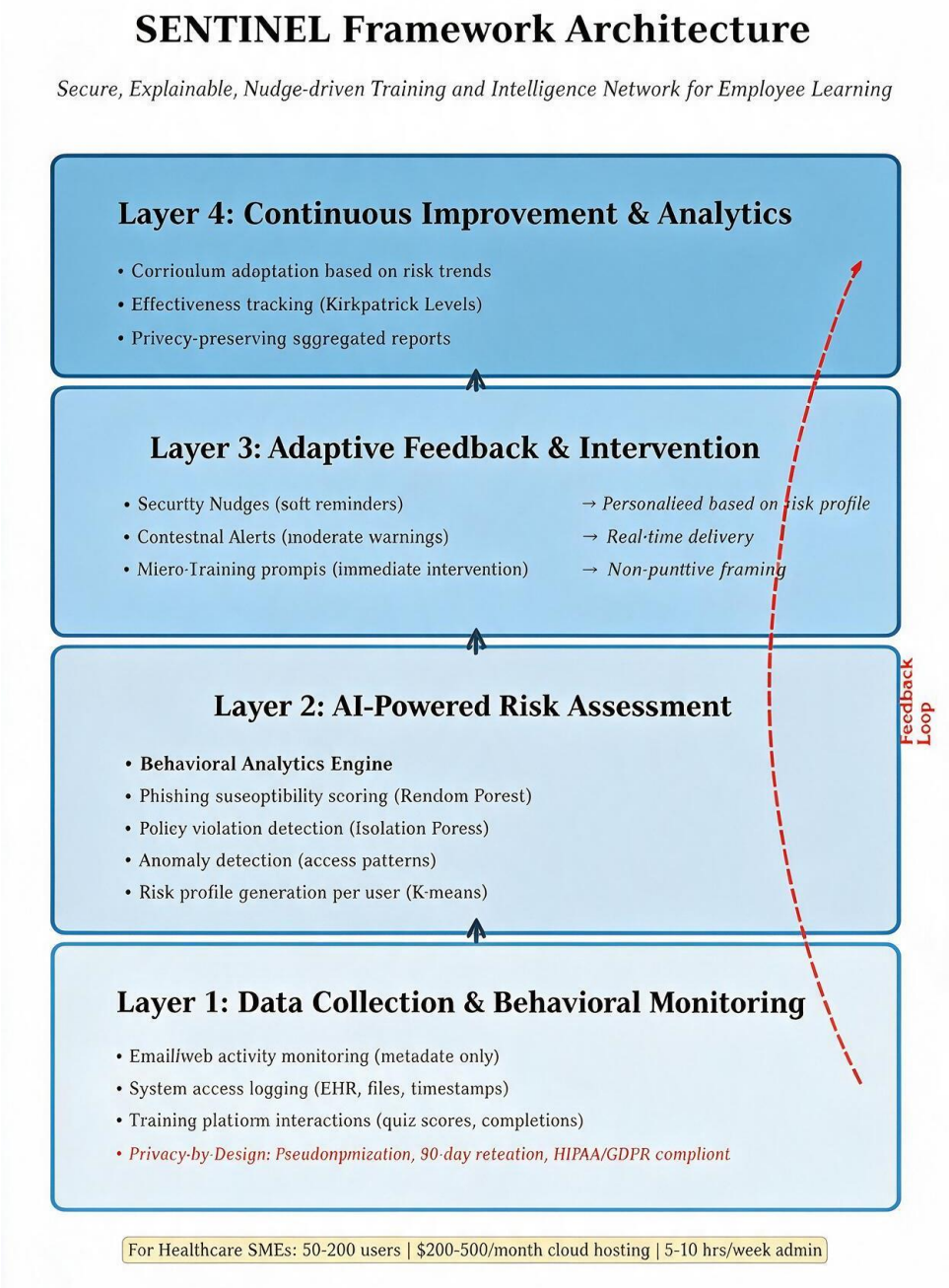

Fig. 2. SENTINEL Framework Architecture showing the four interconnected operational layers and the data flow between them.

Component 3: Real-Time Feedback Mechanisms

Feedback runs on three escalating tiers, grounded in Nudge Theory [14]:

- **Level 1 – Soft Nudges (non-intrusive):** Low-stakes reminders tied to routine risk events. Password expiring in three days: one passive banner, easy to dismiss, no workflow interruption. The bar for triggering a Level 1 nudge is deliberately low.
- **Level 2 – Contextual Alerts (moderate interruption):** Triggered when a medium-risk action is detected attempting to send an unencrypted file with patient data, for instance. A pop-up or sidebar warning appears with a clear, non-accusatory explanation of what the risk is and a one-click path to the safer option.
- **Level 3 – Immediate Intervention with Micro-Training (high interruption):** Reserved for genuinely high-risk events: credential entry on a flagged domain, suspected malware download. The action is temporarily blocked. A 60-second training video explains the specific risk. A short quiz must be completed before the user can proceed. The IT administrator is notified.

A personalization engine decides which tier applies to which user. Chronic high-risk individuals get more frequent, more detailed nudges. Low-risk users with strong behavioral records experience minimal interruption. The messaging tone is calibrated too gentler for first-time mistakes, more direct when the same behavior keeps appearing [18].

Component 4: Continuous Improvement Loop

The framework evaluates itself against Kirkpatrick's four levels [19], rather than treating the evaluation as something done externally once a year:

- Level 1 (Reaction): Post-feedback satisfaction surveys, plus sentiment analysis on IT support tickets
- Level 2 (Learning): Quiz performance trends tracked at 30, 60, and 90 days post-training
- Level 3 (Behavior): Phishing click rates, policy violation counts, voluntary threat reporting frequency
- Level 4 (Results): Incident frequency, near-miss events logged, regulatory audit outcomes

The curriculum updates automatically when risk patterns shift. If credential phishing attempts spike across the user base, SENTINEL flags the trend, surfaces targeted training modules for high-risk groups, and pushes the topic up in the queue for lower-risk users. Management receives quarterly reports built from aggregated, privacy-preserving data enough to make decisions, not enough to surveil individuals.

2.2 Implementation Workflow

Deployment follows four phases designed around the reality that most healthcare SMEs cannot shut down operations for a technology rollout:

- **Phase 1 – Setup (Weeks 1–2):** Lightweight monitoring agents go on endpoints. Integrations with the email gateway and EHR access logs are configured. Thirty days of baseline behavioral data are collected before any models are trained or any alerts are fired.
- **Phase 2 – Model Training (Weeks 3–4):** Initial risk models train on the baseline. Accuracy is validated against any historical incident data the organization holds. Feedback thresholds are set conservatively alert fatigue is a real risk, especially in clinical environments where staff are already stretched.
- **Phase 3 – Soft Launch (Weeks 5–8):** Nudges and alerts go live in shadow mode: the system logs what it would have triggered, but does not interrupt users yet. A pilot group of 10–20% of staff provides feedback on message clarity and timing. Adjustments are made before the wider rollout.
- **Phase 4 – Full Deployment (Week 9+):** All feedback mechanisms activate organization-wide. Weekly model retraining begins. Monthly reviews with stakeholder's track whether the metrics are moving in the right direction.

On the resource side: a cloud-hosted platform for 50–200 users typically run $200–$500 per month. A part-time IT administrator, five to ten hours per week, can handle ongoing operations. Staff need a two-hour orientation before launch enough to explain what the system monitors, why it exists, and how to read the feedback they will receive.

2.3 Framework Validation Approach

Analytical validation was performed as the SENTINEL is developed by systematic literature review. The four layers were directly deduced by evidence from the literature, behavioral data collection [7, 12, 13], AI based risk assessment [7,11, 12, 18] real time adaptive feedback 14, 15, 16, 17], continuous improvement evaluation [9, 10, 19]. Internal consistency was validated by nudge theory [14] and Kirkpatrick's model [[19]. Empirical validation through pilot study is planned for future study.

3. RESULTS

3.1 AI Techniques in Awareness Evaluation

The 20 studies used a wider range of AI approaches than we initially expected. Fig. 3 shows the breakdown. Behavioral analytics led by a wide margin nine of the twenty studies, 45% covering click pattern tracking, access violation detection, and continuous risk scoring. This dominance reflects less a consensus that behavioral analytics is the best approach and more the fact that it is the most practically accessible one; the tools exist, the data pipelines exist, and the interpretability is good enough for non-specialist users to act on. Predictive modeling accounted for another five studies (25%), mostly focused on forecasting individual susceptibility to phishing and social engineering. Multimodal AI combining biometrics, access logs, and behavioral signals appeared in three studies (15%). NLP and sentiment analysis, used to probe the depth of security understanding from quiz responses and communication data, turned up in two (10%). Adaptive feedback engines that adjust training pathways dynamically arguably the most promising category appeared in only one study (5%), pointing to how much room the field still has to grow.

3.2 Feedback Mechanisms and Effectiveness

Table. 1 maps the feedback types we found across the sample onto the Kirkpatrick levels they most directly address. Fig. 4 shows the distribution.

A few things stand out on the Table. 1. Security nudges and contextual alerts both cluster at Level 3 they are tools for changing what people do, not just what they know. Micro-training touches both Level 2 and Level 3, which is part of what makes it valuable: it delivers knowledge and behavioral practice in the same moment. Chatbot feedback occupies the Level 1–2 range, which may explain why it remains a minority approach; reactions and learning are necessary but insufficient if behavior does not follow. A prior literature finding worth highlighting is that context-aware interventions delivered immediately after a risky action reduced repeat errors [16]. That is a large effect for a relatively low-cost intervention.

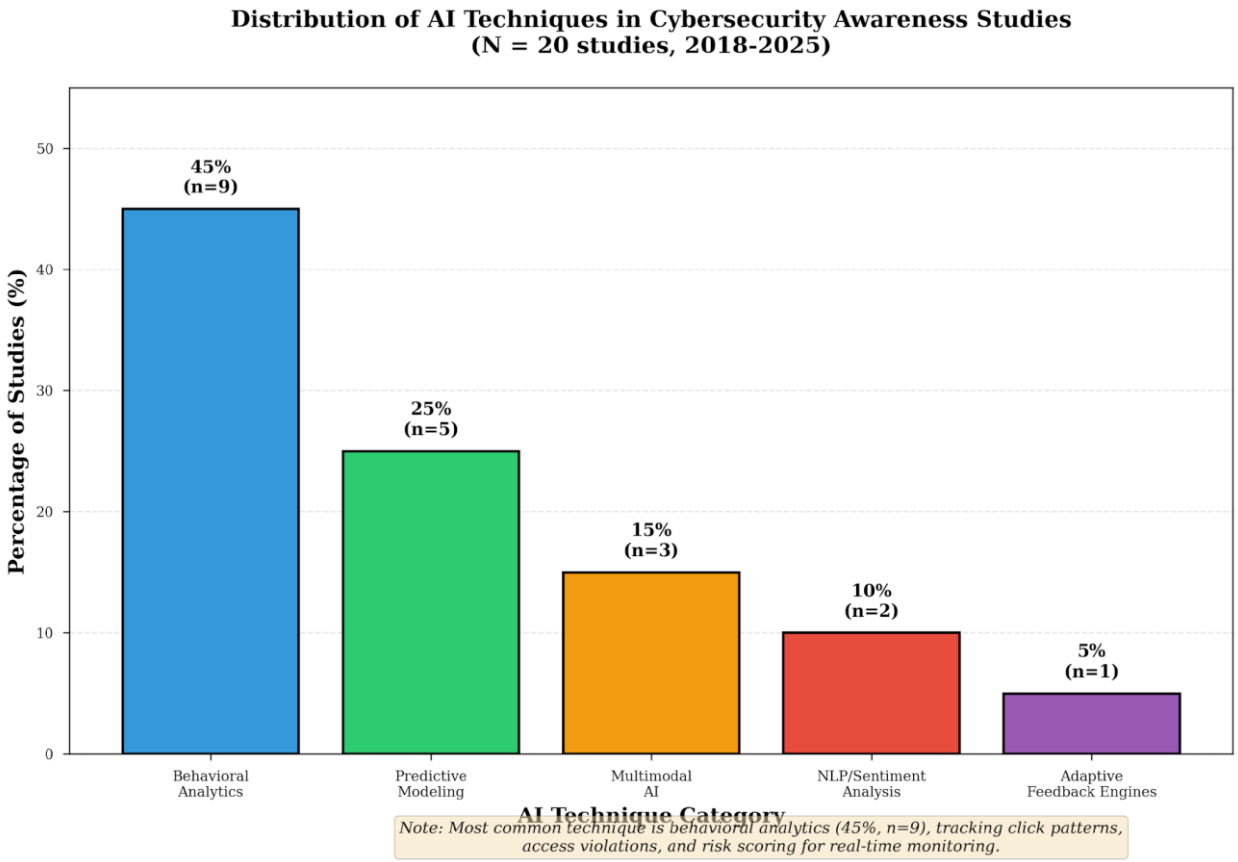

Fig. 3. Distribution of AI and ML techniques across the 20 included studies. Behavioral analytics was the dominant approach, reflecting its practical maturity for real-time monitoring environments.

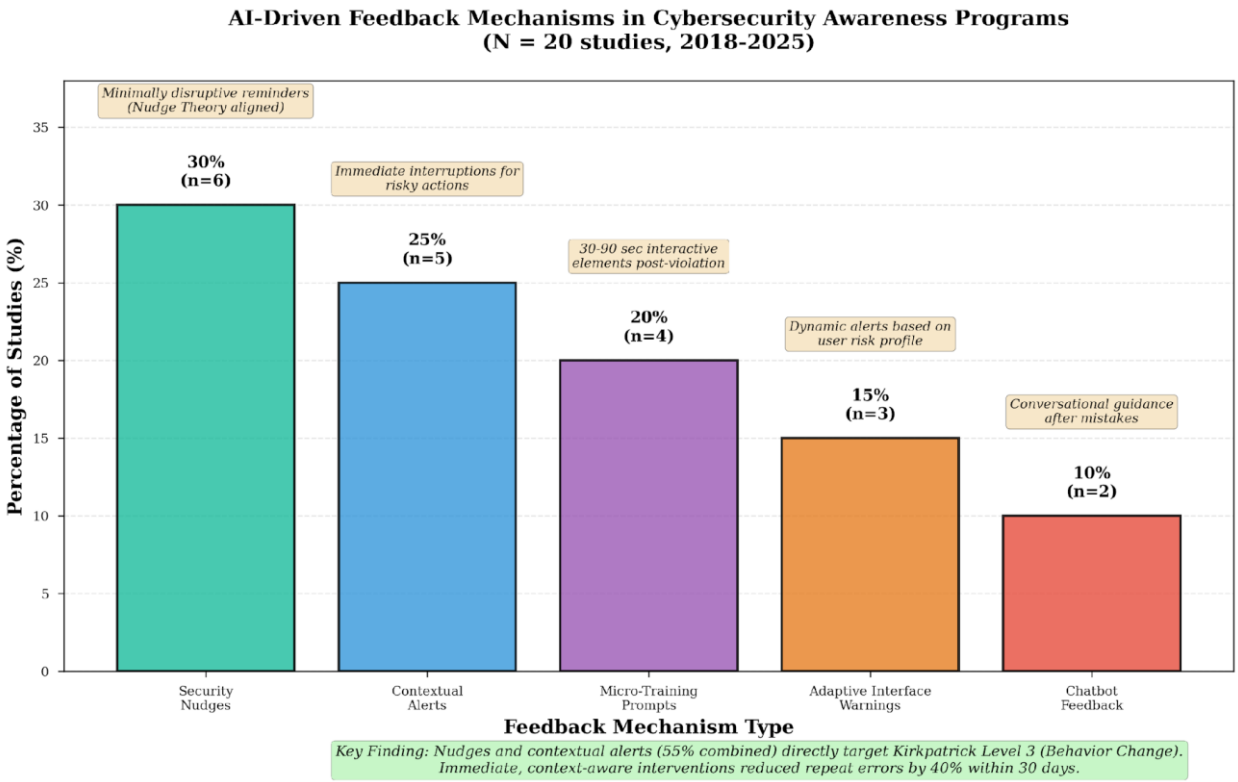

Fig. 4. Distribution of AI-driven feedback mechanism types across the 20 included studies.

Table 1. Mapping of Feedback Mechanism Types to Kirkpatrick Evaluation Levels

Feedback Type	Kirkpatrick Level	Studies (n)
Security Nudges	Level 3 (Behavior)	6
Contextual Alerts	Level 3 (Behavior)	5
Micro-Training Prompts	Level 2–3 (Learning/Behavior)	4
Adaptive Interface Warnings	Level 2 (Learning)	3
Chatbot Feedback	Level 1–2 (Reaction/Learning)	2

3.3 Alignment with Theoretical Frameworks

Table. 2 and Fig. 5 show how the included studies mapped across Kirkpatrick's four levels.

Table 2. Distribution of Kirkpatrick Evaluation Levels Across the 20 Included Studies

Kirkpatrick Level	Studies (n)	Primary Outcomes Measured
Level 3 (Behavior)	12 (60%)	Phishing click rate reduction, increased threat reporting
Level 2 (Learning)	5 (25%)	Quiz performance, knowledge retention
Level 4 (Results)	2 (10%)	Incident reduction, compliance improvement
Level 1 (Reaction)	1 (5%)	User satisfaction surveys

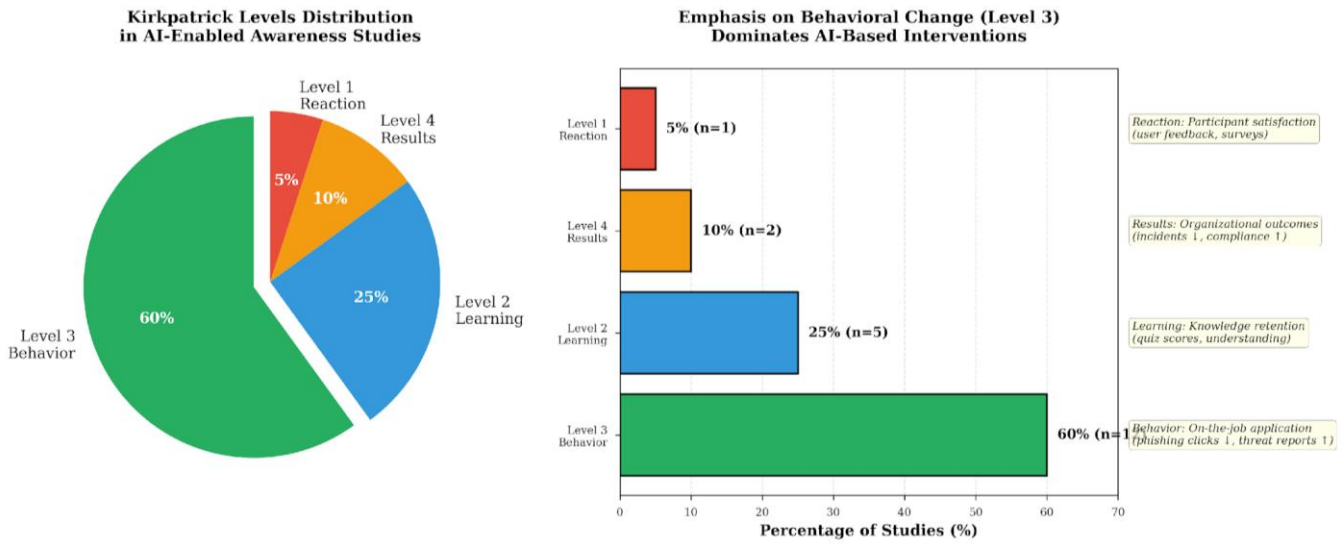

Fig. 5. Distribution of Kirkpatrick evaluation levels addressed across the 20 included studies. The concentration at Level 3 reflects the field's focus on behavioral change over knowledge measurement alone.

From Table. 2 level 3 concentration twelve of twenty studies tell us something important about where the field thinks the problem actually lives. It is not that employees do not know phishing is bad. It is that knowing does not reliably translate into not clicking. The AI tools reviewed in this literature have largely accepted that framing and are trying to solve the behavioral problem directly, rather than doubling down on knowledge transfer. The Nudge Theory alignment is consistent with this: 55% of studies (n=11) used behaviorally informed interventions designed to shape decisions at the moment they occur, without relying on employees to recall and apply training they received months earlier.

4. DISCUSSION

4.1 Addressing Research Questions

RQ1: How is AI used to assess cybersecurity behavior in real time?

The short answer is: by watching what people actually do rather than what they say they would do. Behavioral analytics and predictive modelling track phishing responses, access patterns, and policy compliance continuously. Individual risk scoring makes it possible to target interventions where they are actually needed, rather than applying the same training to everyone regardless of their actual behavioral profile.

RQ2: How do feedback loops influence secure behavior?

The literature is fairly consistent here. Feedback works best when it is immediate, specific, and tied to an action the person just took or is about to take. Nudges, contextual alerts, and micro-training all share that property. Generic reminders sent on a calendar schedule do not, which is why the field is moving away from them.

RQ3: What frameworks support AI integration into continuous improvement?

Kirkpatrick-based evaluation gives the clearest structure for tracking whether an AI-enabled awareness program is actually working. Combining that with privacy-by-design data practices and positioning the AI as a coaching tool rather than a monitoring one seems to be the approach most likely to achieve both technical effectiveness and user acceptance.

4.2 Implications for Healthcare SMEs

Four practical advantages of AI-enabled awareness programs stand out for smaller healthcare organizations specifically. First, real-time monitoring can replace manual observation that simply is not feasible without a dedicated security team. The AI does the watching that a CISO would do in a larger organization. Second, automated nudges reduce the problem of training saturation: instead of overloading staff with quarterly refreshers, the system delivers targeted reminders when they are relevant. Third, behavioral data allows curriculum updates that reflect what is actually happening in the organization rather than generic threat categories. Fourth and perhaps most importantly the system keeps improving as it collects more data, rather than becoming stale between update cycles.

None of that happens automatically, though, and the ethical design choices matter as much as the technical ones. A system that employees experience as surveillance will be undermined, worked around, and eventually abandoned. Transparency about what is monitored, anonymization of sensitive behavioral data, non-punitive framing of interventions, and genuine opt-in processes are not optional extras they are the preconditions for the system to function [8, 20].

4.3 Limitations and Future Directions

Three gaps in the existing literature are worth naming clearly. Almost none of the studies we reviewed followed employees for more than six months, which means we have limited evidence about whether behavioral improvements persist over time or fade once the novelty of the intervention wears off. The privacy and ethical dimensions of continuous behavioral monitoring are discussed in principle in many papers but examined in practice in very few. And the specific constraints of low-resource healthcare settings where even a $300/month platform might represent a real budget decision are largely absent from the literature. Another limitation is that the SENTINEL framework is a theoretical conceptual framework which is not empirically validated in real audience settings; this was considered for future studies.

Future research should treat pilot deployments in operational healthcare SMEs as a priority, not just a recommendation. The framework needs real-world validation: what does alert fatigue actually look like in a ten-person clinic? What privacy concerns do employees raise when they understand the system? What does the cost-benefit calculation look like for a small practice that had one breach five years ago? These are answerable questions, but only through implementation research.

5. CONCLUSION

The argument this paper makes is fairly direct: static, compliance-driven cybersecurity training has a structural problem that more of the same training cannot fix. Healthcare SMEs are caught in that problem of acutely high risk, limited resources, and training programs that satisfy auditors without reliably changing behavior. The 20 studies we reviewed collectively suggest that AI-driven behavioral monitoring and real-time feedback can do something that annual modules cannot: reach employees at the decision point, before the mistake happens, with information that is specific to what they are about to do. The numbers from those studies are modest but meaningful. Behavioral analytics dominated (45% of applications). Nudge-based interventions appeared in more than half the studies. Kirkpatrick Level 3 outcomes actual behavior change, not just knowledge gain was what most of the research was trying to measure and improve. The SENTINEL framework is our attempt to organize the tools that healthcare SME could actually deploy. The architecture is conceptual at this stage, but it is grounded in what the literature shows works. The next step is empirical. For resource constrained SME organizations, the appeal of AI as a workforce multiplier is viable. AI based continuous coaching systems help employees stay secure.

REFERENCES

[1] European Union, "General Data Protection Regulation (GDPR): Regulation (EU) 2016/679." [Online]. Available: https://gdpr-info.eu/. Accessed: Sep. 15, 2025. [CrossRef] [Publisher]

[2] U.S. Dept. Health and Human Services, "Summary of the HIPAA Privacy Rule." [Online]. Available: /. Accessed: Sep. 15, 2025. [Publisher]

[3] D. D. Caputo, S. L. Pfleeger, J. D. Freeman, and M. E. Johnson, "Going spear phishing: Exploring embedded training and awareness," *IEEE Secur. Privacy*, vol. 12, no. 1, pp. 28–38, Jan. 2014. [CrossRef] [Publisher]

[4] M. A. Sasse, S. Brostoff, and D. Weirich, "Transforming the 'weakest link': A human/computer interaction approach to usable and effective security," *BT Technol. J.*, vol. 19, no. 3, pp. 122–131, Jul. 2001. [CrossRef] [Publisher]

[5] J. M. J. Murre and J. Dros, "Replication and analysis of Ebbinghaus' forgetting curve," *PLOS ONE*, vol. 10, no. 7, p. e0120644, Jul. 2015. [CrossRef] [Publisher]

[6] M. Alshaikh, S. B. Maynard, A. Ahmad, and S. Chang, "An exploratory study of current information security training and awareness practices in organizations," in *Proc. 51st Hawaii Int. Conf. System Sciences (HICSS)*, Waikoloa, HI, USA, Jan. 2018, vol. 9, pp. 5085–5094. [CrossRef] [Publisher]

[7] T. Sutter, A. S. Bozkir, B. Gehring, and P. Berlich, "Avoiding the hook: Influential factors of phishing awareness training on click-rates and a data-driven approach to predict email difficulty perception," *IEEE Access*, vol. 10, pp. 100540–100565, 2022. [CrossRef] [Publisher]

[8] M. Alshaikh, "Developing cybersecurity culture to influence employee behavior: A practice perspective," *Comput. Secur.*, vol. 98, p. 102003, Nov. 2020. [CrossRef] [Publisher]

[9] M. Bada, A. M. Sasse, and J. R. C. Nurse, "Cyber security awareness campaigns: Why do they fail to change behaviour?". [CrossRef] [Publisher]

[10] D. Lain, K. Kostiainen, and S. Čapkun, "Phishing in organizations: Findings from a large-scale and long-term study," in *Proc. 43rd IEEE Symp. Security and Privacy (SP)*, San Francisco, CA, USA, May 2022, pp. 842–859 [CrossRef] [Publisher]

[11] S. Silvestri, S. Islam, S. Papastergiou, C. Tzagkarakis, and M. Ciampi, "A machine learning approach for the NLP-based analysis of cyber threats and vulnerabilities of the healthcare ecosystem," *Sensors*, vol. 23, no. 2, p. 651, Jan. 2023. [CrossRef] [Publisher]

[12] S. Yuan and X. Wu, "Deep learning for insider threat detection: Review, challenges and opportunities," *Comput. Secur.*, vol. 104, p. 102221, May 2021. [CrossRef] [Publisher]

[13] L. Coventry and D. Branley, "Cybersecurity in healthcare: A narrative review of trends, threats and ways forward," *Maturitas*, vol. 113, pp. 48–52, Jul. 2018. [CrossRef] [Publisher]

[14] R. H. Thaler and C. R. Sunstein, *Nudge: Improving Decisions About Health, Wealth, and Happiness*. New Haven, CT, USA: Yale Univ. Press, 2008. [CrossRef] [Publisher]

[15] W. J. Gordon, A. Wright, R. J. Glynn, J. Kadakia, C. Mazzone, E. Leinbach, and A. Landman, "Evaluation of a mandatory phishing training program for high-risk employees at a US healthcare system," *J. Am. Med. Inform. Assoc.*, vol. 26, no. 6, pp. 547–552, Jun. 2019. [CrossRef] [Publisher]

[16] V. Zimmermann and K. Renaud, "The nudge puzzle: Matching nudge interventions to cybersecurity decisions," *ACM Trans. Comput.-Hum. Interact.*, vol. 28, no. 1, pp. 1–45, Jan. 2021. [CrossRef] [Publisher]

[17] K. Renaud and V. Zimmermann, "Ethical guidelines for nudging in information security & privacy," *Int. J. Hum.-Comput. Stud.*, vol. 120, pp. 22–35, Dec. 2018. [CrossRef] [Publisher]

[18] W. J. Gordon, A. Wright, R. Aiyagari, L. Corbo, R. J. Glynn, J. Kadakia, J. Kufahl, C. Mazzone, J. Noga, M. Parkulo, B. Sanford, P. Scheib, and A. B. Landman, "Assessment of employee susceptibility to phishing attacks at US health care institutions," *JAMA Netw. Open*, vol. 2, no. 3, p. e190393, Mar. 2019. [CrossRef] [Publisher]

[19] D. L. Kirkpatrick and J. D. Kirkpatrick, *Evaluating Training Programs: The Four Levels*, 3rd ed. San Francisco, CA, USA: Berrett-Koehler Publishers, 2006. [CrossRef] [Publisher]

[20] K. Parsons, D. Calic, M. Pattinson, M. Butavicius, A. McCormac, and T. Zwaans, "The human aspects of information security questionnaire (HAIS-Q): Two further validation studies," *Comput. Secur.*, vol. 66, pp. 40–51, May 2017. [CrossRef] [Publisher]