Journal of Cybersecurity and Information Management

♦ ISSN: 2692-4048 Vol: 18, No. 02, 2026 ♦ PP. 11–25

♦ Article DOI: https://doi.org/10.54216/JCIM.180202



FULL LENGTH ARTICLE

Severity Scores, Listing Latency and Prioritization Efficiency: Evidence from the CISA Known Exploited Vulnerabilities Catalog

Tamara Zhukabayeva[1] **Baumuratova Dilaram**[2]

[1] Eurasian National University, Kazakhstan

[2] Astana International University, Kazakhstan

Emails: Zhukabayeva_tk@enu.kz · Baumuratova.d@gmail.com

Received: August 11, 2026 Revised: September 09, 2026 Accepted: November 05, 2026 ⋆ Corresponding author

ABSTRACT

Vulnerability managers receive more CVE records each year than they can patch, and the Known Exploited Vulnerabilities (KEV) catalog of the U.S. Cybersecurity and Infrastructure Security Agency has become the most widely used public signal of which ones matter. We link the catalog as of 29 July 2026 to every National Vulnerability Database record published from January 2022 to June 2026 and ask which vulnerabilities are listed, how soon, and how well information available at disclosure ranks them. Fewer than one published CVE in two hundred reached the catalog. Listing was front-loaded: among CVEs listed within a year, 29% were listed on or before their publication day and 67% within 30 days. A Critical base score multiplied the listing hazard roughly tenfold, yet a CVSS ranking caught 39% of later listings at a workload of 8.6% of new CVEs; a logistic model adding weakness family and the issuing authority's listing history caught 60% at that workload on a temporal hold-out. Scores supplied by numbering authorities ranked better than NVD's own. Vendor concentration fell after 2023, ransomware labels mostly arrived long after listing, and remediation windows shortened to 14 and then 3 days in 2026.

Keywords: Known exploited vulnerabilities ▪ Vulnerability prioritization ▪ CVSS ▪ Survival analysis ▪ Patch management

1. INTRODUCTION

A security team that subscribes to the National Vulnerability Database (NVD) received roughly 130 new CVE records a day during 2025. Nobody patches at that rate. Every team therefore runs some form of triage, and for most of the past two decades the default sort key has been the base score of the Common Vulnerability Scoring System (CVSS), a 0–10 number that describes how bad a flaw would be if someone used it [1]. Whether anyone will use it is a different question, and the base score was never designed to answer it [2]. Since November 2021 a second signal has existed. The Cybersecurity and Infrastructure Security Agency (CISA) publishes the Known Exploited Vulnerabilities (KEV) catalog, a list of CVEs for which it holds reliable evidence of exploitation in the wild, each with a deadline by which federal civilian agencies must remediate [3]. Private firms, insurers and scanner vendors adopted the list quickly, and by mid-2026 it was the input that two other public systems had been rebuilt around: NIST announced in April 2026 that NVD would enrich KEV entries within one business day while leaving most other records unscored [4], and CISA's Binding Operational Directive 26-04 of 10 June 2026 tied its new three-day and 14-day clocks to KEV status [5].

That dependence makes the catalog's behavior an information-management question in its own right. Timing decides its value. A list that arrives after most exposed systems have been hit serves clean-up and audit, and one that arrives earlier serves prevention. A defender cannot wait for the list, though, so the practical issue is how much of it can be anticipated

at the moment a CVE is published, with the fields that a CVE Numbering Authority (CNA) and NVD actually supply. The exploit-prediction literature has answered a related question with proprietary exploitation feeds [6, 7, 8], and recent comparisons find that CVSS, the Exploit Prediction Scoring System (EPSS) and stakeholder-specific schemes disagree with one another far more than practitioners assume [9]. No published analysis, as far as we can find, builds the same comparison entirely from public, versioned records with an explicit time axis that a reader could rerun against a later snapshot.

We assemble such a record. The KEV catalog at its 29 July 2026 version is joined to every NVD record whose identifier carries a 2022–2026 year and that was published between 1 January 2022 and 30 June 2026, giving a cohort in which each CVE has at least four weeks of follow-up and the oldest have more than four years. Git history of the catalog gives the date on which each ransomware flag changed. Three questions organize the paper: which CVEs get listed and how much of the difference disclosure-time fields explain; how soon after publication listing happens, and whether that timing differs across kinds of weakness; and how much remediation work a CVSS threshold costs per listed CVE caught, set against a ranking that uses the same public fields more fully.

Several answers cut against common practice. Listing is rare (0.44% of the cohort) and front-loaded, so the decision a defender faces is largely made in the first week. CVSS carries real signal, with a Critical score raising the listing hazard about tenfold, yet at the workload a "patch everything rated 9.0 or above" rule implies, a model that adds weakness family and the CNA's listing history catches 60% of later listings where the rule catches 39%. The CNA's own score, which is what a defender sees first, ranks better than the NVD score that replaces it. Ransomware attribution, a field many triage tools treat as a severity multiplier, is applied a median of 287.5 days after listing. The literature and five hypotheses follow in Section 2; Sections 3 and 4 cover data and models, Section 5 the results, and Section 6 what they imply for vulnerability management and for the two 2026 policy changes.

2. RELATED WORK AND HYPOTHESES

2.1 Severity and exploitation

Large-scale studies of vulnerability databases began by asking how disclosure, exploit and patch dates line up [10], and the answer has always been that most published vulnerabilities are never exploited at all. Allodi and Massacci [11] matched NVD records to exploit kits and Symantec attack signatures in a case-control design and found that a high CVSS score was a weak indicator of exploitation: a defender who patched only on CVSS would fix many unexploited flaws and miss a share of the exploited ones. The CVSS specification itself now says that the base score measures severity, and that threat and environmental information must be layered on for risk [1, 12]. Spring et al. [2] argued that the single number invites misuse as a risk ranking, and the Stakeholder-Specific Vulnerability Categorization was proposed as a decision tree that makes exploitation status an explicit input [13]. Scoring practice has its own noise. Ruohonen [14] measured delays between publication and NVD scoring, Anwar et al. [15] documented inconsistent and missing fields across NVD, and the Department of Commerce Inspector General reported in May 2026 that independent evaluators matched NVD's scoring exactly for only 8 of 69 sampled vulnerabilities [16].

2.2 Predicting exploitation

Prediction work treats exploitation as a classification target. Bozorgi et al. [17] trained classifiers on text and metadata to forecast proof-of-concept exploits, and Sabottke et al. [18] added social-media mentions and showed how an attacker who knew the features could game them. Suciu et al. [19] went further and modeled the probability that a functional exploit would be written as a quantity that changes over time after disclosure, so that the same CVE can look harmless in its first week and dangerous in its third month. EPSS, the most widely deployed model, estimates the probability of exploitation activity in the next 30 days from vendor, weakness, reference and threat-feed features [7]; its third version reports large efficiency gains over CVSS thresholds on intrusion-detection telemetry [8], and Jacobs et al. [6] introduced the coverage and efficiency vocabulary used below.

Two features of this literature limit what a reader can check, because the outcome data are usually private and the evaluation is usually cross-sectional, so a model can learn from references and tags that were attached after exploitation became known. Timing matters as much as features. With public exploit archives, Householder et al. [20] found that the share of CVEs with a public exploit keeps growing for years after disclosure, so a label observed at one date can differ from the label observed at another.

2.3 The catalog as a policy instrument

The KEV catalog differs from these targets in being both an outcome and an instrument. Listing requires an assigned CVE, reliable evidence of active exploitation and clear remediation guidance [3], and listing triggers a federal deadline, so the catalog is shaped by CISA's evidence threshold and capacity as well as by attacker behavior. Economic models of patching treat the defender's choice as a trade between the cost of applying fixes and the expected loss from waiting [21], and vendor patch speed responds to disclosure pressure [22]; a public list with deadlines changes both sides of that trade. Ransomware adds a second dimension. CISA marks entries with known use in ransomware campaigns, and work on organizational ransomware incidents finds that the severity of an attack depends heavily on which systems are reached first [23]. Exploit markets and the economics of exploitation suggest that attackers concentrate on widely deployed products where one exploit pays repeatedly [24], which predicts a concentrated vendor profile in the catalog.

2.4 Hypotheses

Five hypotheses follow, stated so that the data can reject them.

H1. Listing probability rises with the CVSS base score. Weakness family and the issuing CNA's prior listing history carry information beyond it.

H2. The listing hazard is front-loaded: most listings that occur within a year of publication occur within the first 30 days.

H3. On a temporal hold-out, and at the same remediation workload, a ranking built from the fields available at disclosure (severity, vector components, weakness family and CNA identity and history) catches more of the CVEs later listed than a ranking by the CVSS base score alone.

H4. Vendor concentration in the catalog declined between the 2021–2023 and 2024–2026 addition periods.

H5. Ransomware attribution is retrospective. With vendor and weakness held fixed, the chance that an entry carries the flag rises with time since listing.

3. DATA

3.1 The KEV catalog

CISA publishes the catalog as JSON and CSV files and mirrors each release to a public Git repository (cisagov/kev-data), and we fixed the analysis at catalog version 2026.07.29, released on 29 July 2026, taking the JSON file from repository commit 564b8c5. One discrepancy turned up. The CSV file at that commit lacked an entry that the JSON contained (CVE-2026-20316, added that day), so the JSON is treated as authoritative throughout. Each of the 1,656 entries records a CVE identifier, vendor and product names, the date added, the remediation due date, a ransomware field taking the values "Known" or "Unknown", and CWE identifiers. We define the remediation window of entry e, in days, as

$$w_e = \text{dueDate}_e - \text{dateAdded}_e, \tag{1}$$

The repository history from 29 January 2025 to the snapshot contains 268 versions of the JSON file. Replaying them in order gives, for each entry, the version in which it first appeared and any version in which its ransomware field changed; 85 such changes occurred, every one from "Unknown" to "Known", and three entries were removed.

3.2 NVD records and score provenance

NVD records came from the community mirror fkie-cad/nvd-json-data-feeds at tag v2026.07.31-000014 (commit 734f9e2, 31 July 2026), which stores one JSON file per CVE as published through the NVD application programming interface. We parsed every record in the CVE-2022 to CVE-2026 folders, 181,947 files in all, keeping for each CVE its publication date, analysis status, CNA source identifier, CVSS v3.x metrics, CWE identifiers, reference tags and vulnerable-product configurations.

A CVE can carry two CVSS v3.x assessments. The "Secondary" assessment comes from the CNA or an authorized data publisher and is normally present at or near publication; the "Primary" assessment is NVD's own, added during enrichment, which since 2024 has lagged by weeks or not happened at all [16]. We built two scores. The disclosure score takes the Secondary assessment when one exists and the Primary otherwise, while the NVD-preferred score reverses that order and so reproduces what most consumers of the NVD feed saw before NIST changed its scoring policy in April 2026. Both use version 3.1 where present and 3.0 otherwise. CVSS v4.0 assessments were too sparse in the 2022–2024 training years to support a parallel analysis [12]. Severity bands follow the specification: Low (0.1–3.9), Medium (4.0–6.9), High (7.0–8.9) and Critical (9.0–10.0).

The NVD fields need one warning before use. All 732 listed CVEs in the cohort had the status "Analyzed" at the snapshot, against 36% of the cohort as a whole, because NVD enriches KEV entries first [4]. Fields that NVD adds during analysis (product configurations, reference tags such as "Exploit", and the Primary score itself) are therefore partly consequences of listing, so we use them only in clearly marked sensitivity analyses and build the main models from the fields that a defender could have seen on the day a CVE was published.

3.3 Cohort, outcome and covariates

The cohort contains every NVD record with a 2022-or-later identifier year and a status other than "Rejected" published from 2 January 2022 (the first publication in the folders) to 30 June 2026: 167,579 CVEs, of which 732 were in the catalog on 29 July 2026. Follow-up therefore runs from 29 days for the newest records to 1,669 days for the oldest. For CVE i with publication date t_i and listing date A_i the observed time is

$$T_i = \begin{cases} \max(A_i - t_i, 0) + 0.5, & \text{if listed by 29 July 2026,} \\ 29\ \text{July 2026} - t_i + 0.5, & \text{otherwise,} \end{cases} \tag{2}$$

with event indicator δ_i equal to one for listed CVEs. The half-day offset places same-day listings strictly after time zero. Fifty-six listed CVEs (7.7%) were added to the catalog before NVD published them, the usual case being a vendor advisory and CISA alert issued while the record was still reserved; we set their lag to zero for survival models and report them separately.

Disclosure-time covariates are the disclosure-score severity band; three CVSS vector components (attack vector network, no privileges required, no user interaction); a weakness family built from CWE identifiers (memory safety, injection and code execution, authentication and access control, path traversal with file upload and server-side request forgery, SQL injection, cross-site scripting, information exposure, other, and unspecified); the CNA identity, with 18 large vendor and third-party CNAs kept as separate categories; and the CNA's listing history. The last of these is the number of earlier catalog additions among CVEs issued by the same CNA,

$$P_i = \#\{j : \text{CNA}_j = \text{CNA}_i,\ \delta_j = 1,\ A_j < t_i\}, \tag{3}$$

which uses only listings dated before the CVE's own publication and so cannot see the future. Because the cohort starts in 2022, P_i counts listings of 2022-and-later CVEs only. Table 1 lists the sources and the role each plays.

The cohort is skewed in ways that matter later. Three content-management-ecosystem CNAs (Patchstack, Wordfence and WPScan) issued 29,806 cohort CVEs, nearly all in WordPress plug-ins, and none of them was listed, while VulDB, a third-party CNA that accepts reports on almost any product, issued 14,154 of which four were listed. The Linux kernel CNA, which began assigning identifiers to most kernel fixes in 2024, issued 11,777 with seven listings. Rates computed over "all CVEs" are thus partly rates over these streams. Section 5.2 separates them.

Table 1. Data sources, snapshot identifiers and use in the analysis

Source	Snapshot	Fields used	Role
CISA KEV catalog (JSON)	Version 2026.07.29; commit 564b8c5	CVE, vendor, dateAdded, dueDate, ransomware field, CWE	Outcome; catalog analyses
KEV Git history	268 versions, 29 Jan 2025 – 29 Jul 2026	Ransomware field by version; removals	Flag timing
NVD JSON mirror	Tag v2026.07.31-000014; commit 734f9e2	Publication date, status, CNA, CVSS v3.x Primary and Secondary, CWE	Cohort; disclosure covariates
NVD enrichment fields	Same	Reference tags, CPE vendor	Sensitivity analyses only

4. METHODS

4.1 Time to listing

We estimate the cumulative probability of listing by day t with the Kaplan–Meier product-limit estimator [25],

$$\widehat{S}(t) = \prod_{j:t_j \le t} \left(1 - \frac{d_j}{n_j}\right), \qquad \widehat{F}(t) = 1 - \widehat{S}(t), \tag{4}$$

where d_j is the number of listings at day t_j and n_j the number of CVEs still unlisted and under observation. Withdrawal of a CVE from the catalog is not a competing event here (three removals in 18 months), so F reads directly as a cumulative listing probability. Covariate effects come from the Cox proportional hazards model [26],

$$h(t \mid \mathbf{x}_i) = h_0(t) \exp\left(\mathbf{x}_i^\top \beta\right), \tag{5}$$

fitted with Efron's handling of ties in the lifelines library [27] and summarized by Harrell's concordance index [28], with a publication-year term to absorb the changes in CVE volume and CNA mix that separate one cohort from the next. Proportionality is tested with the rank-time Schoenfeld residual test [29]. Because several covariates failed that test, we also fit a piecewise model that lets every coefficient differ before and after day 30,

$$h(t \mid \mathbf{x}_i) = h_0(t) \exp\left(\mathbf{x}_i^\top \beta_1 \mathbf{1}[t \le 30] + \mathbf{x}_i^\top \beta_2 \mathbf{1}[t > 30]\right), \tag{6}$$

estimated as two Cox models: one with follow-up truncated at day 30, and one on CVEs still unlisted after day 30 with time measured from that point.

4.2 Ranking and prioritization

The prioritization analysis uses a fixed 365-day horizon. Let $y_i = 1$ if CVE i was listed within 365 days of publication. The training set holds CVEs published from January 2022 to June 2024 and the test set those published from July 2024 to June 2025; every test CVE has a full year of follow-up by 30 June 2026, so all CVEs are judged over the same window, and the split is strictly forward in time, so no test-period information reaches the model. We fit a ridge-penalized logistic regression,

$$\Pr(y_i = 1 \mid \mathbf{z}_i) = \frac{1}{1 + \exp\left[-\left(\alpha + \mathbf{z}_i^\top \gamma\right)\right]}, \tag{7}$$

on standardized disclosure-time features z_i (the severity band and base score, all eight CVSS vector components, weakness family, CNA indicators, $log1 + P_i$ and the log length of the description), with penalty weight $C = 0.5$ in scikit-learn [30]. A gradient-boosted tree ensemble [31] fitted to the same features checks whether the linear form leaves signal unused, and for interpretation we also report unpenalized odds ratios from a training-set logit on the core covariates that the Cox model uses.

A ranking s is judged the way a remediation team experiences it. If the team fixes the k highest-ranked CVEs out of n published, its workload is k/n, and

$$\text{Coverage}(k) = \frac{\sum_{i \in T_k} y_i}{\sum_{i=1}^{n} y_i}, \qquad \text{Efficiency}(k) = \frac{1}{k} \sum_{i \in T_k} y_i, \tag{8}$$

where T_k is the top-k set [6]. Coverage is recall at a fixed workload; efficiency is precision. Two summaries cover all workloads at once. The area under the receiver operating characteristic curve,

$$\text{AUC} = \Pr(s_i > s_j \mid y_i = 1,\ y_j = 0), \tag{9}$$

measures ranking quality irrespective of prevalence, while average precision,

$$\text{AP} = \sum_k (R_k - R_{k-1})\ \text{Efficiency}(k), \tag{10}$$

with R_k the coverage at the k-th threshold, weights the top of the list and is the more informative summary at a prevalence below 1% [32]. Confidence intervals come from 1,000 bootstrap resamples of the test set, and differences between rankings from paired resamples [33]. Thresholds are handled separately. We evaluate the CVSS rules at 7.0 and 9.0 as fixed policies and then cut each model's ranking at exactly the workload the corresponding rule implies, so that the comparison holds the remediation budget constant.

4.3 Catalog dynamics

Monthly additions N_m from December 2021 to July 2026, which exclude the founding batch of 3 November 2021, are overdispersed, so we fit a negative binomial regression [34],

$$\mathbb{E}[N_m] = \mu_m = \exp(\beta_0 + \beta_1 \tau_m), \qquad \text{Var}(N_m) = \mu_m + \alpha \mu_m^2, \tag{11}$$

with τ_m in years. A single change in mean is located by maximizing the Poisson likelihood-ratio statistic over admissible break months,

$$\begin{aligned} \Lambda_\kappa &= 2\left[\ell(N_{1:\kappa}) + \ell(N_{\kappa+1:M}) - \ell(N_{1:M})\right], \\ \widehat{\kappa} &= \operatorname*{arg\,max}_{6 \le \kappa \le M-6} \Lambda_\kappa, \end{aligned} \tag{12}$$

where l is the maximized Poisson log-likelihood of a segment; its significance comes from 2,000 permutations of the

monthly series, a simpler relative of the penalized searches in Killick et al. [35]. Concentration gets two measures. They are the Gini coefficient of entry counts across vendors and the Herfindahl–Hirschman index,

$$G = \frac{\sum_{i=1}^{V}(2i - V - 1)x_{(i)}}{V\sum_{i=1}^{V}x_{(i)}}, \qquad \mathrm{HHI} = \sum_{v=1}^{V} s_v^2, \tag{13}$$

where x_i are vendor counts in ascending order and s_v vendor shares [36]. The ransomware field is modeled with a logistic regression on vendor class, catalog weakness family, whether the CVE was three or more years old when listed, and years since listing; proportions carry Wilson score intervals,

$$\frac{\widehat{p} + \frac{z^2}{2n} \pm z\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n} + \frac{z^2}{4n^2}}}{1 + \frac{z^2}{n}}, \tag{14}$$

with $z = 1.96$.

5. RESULTS

5.1 How the catalog grew and how its deadlines changed

The catalog opened on 3 November 2021 with 287 entries and held 1,656 by 29 July 2026. Its growth had two regimes (Figure 1, panel a). Through mid-2022 CISA was clearing a backlog of older, already-exploited CVEs, including 555 additions in 2022 alone with a median CVE age of four years at listing (Table 2). The likelihood-ratio search places the break in July 2022, when mean monthly additions fell from 70.6 to 17.8, with $\Lambda = 500.3$ and none of 2,000 permuted series reaching that value. Monthly counts are heavily overdispersed (variance 905 against a mean of 24.4; Pearson dispersion of a Poisson fit 24.7), which is why we use Equation 11, and over the whole period its trend is downward, with an incidence-rate ratio of 0.85 per year (95% CI 0.76–0.95). That slope is the backlog draining. After the break the sign flips. Additions grew by 17% a year (IRR 1.17, 95% CI 1.08–1.27, 49 months), reaching 245 entries in 2025 and 172 in the first seven months of 2026.

From 2023 onward the catalog became a list of new vulnerabilities, with between 76% and 78% of each year's additions carrying a CVE identifier from that year or the year before, against 26% in 2022, when the backlog of older CVEs was still being cleared. The remediation windows changed more abruptly (Figure 1, panel b). Between 2023 and 2025, 92% to 95% of entries received a 21-day window, while windows of a week or less were reserved for a few emergencies (4% to 8%). That pattern ended in 2026. In March the first standard 14-day windows appeared (14 of 26 entries), from April no entry received 21 days, and after BOD 26-04 came into force on 10 June windows of three days or fewer became the majority: 17 of 23 entries in June and 22 of 26 in July. Across 2026, 39.5% of additions carried a window of a week or less.

Note. In 2021 and 2022 the remaining entries carried mostly six-month windows for older CVEs, as BOD 22-01 prescribed. Ransomware shares are as recorded on 29 July 2026; later years have had less time to be flagged (Section 5.6).

5.2 Which CVEs are listed

Of the 167,579 cohort CVEs, 732 were listed by the snapshot date. That is 0.44%. The rate fell by publication year from 0.60% for 2022 and 2023 CVEs to 0.41% for 2024 and 2025 CVEs; 2026 CVEs, observed for at most seven months, stood at 0.28%. Part of that decline is denominator growth. Annual CVE publication more than doubled between 2022 and 2025, much of it from CNAs whose CVEs are almost never listed, while the number of listed CVEs per publication year moved only from 119 to 194.

Severity matters a great deal (Figure 2, Table 3). The listing rate climbs monotonically from 0.06% in the 4.0–4.9 bin to 0.46% at 7.0–7.9, 0.81% at 8.0–8.9 and 1.67% at 9.0–10.0, roughly doubling with each point above 7, and the Wilson intervals of adjacent bins above 6.0 do not overlap. Critical CVEs make up 10.2% of the cohort and 39.1% of listings, and High CVEs 35.7% and 47.0%, so the two upper bands together hold 86.1% of listings in 45.9% of the cohort. Two facts limit what this buys a defender. Even in the top bin, 98.3% of CVEs were never listed. And 13.9% of listings came from CVEs scored Medium or Low at disclosure, which a "High and above" rule would never reach.

Weakness family separates the cohort about as sharply as severity does (Figure 3). Injection and code-execution weaknesses (OS command injection, code injection, deserialization and expression-language injection) were listed at 1.24%, twice the memory-safety rate of 0.60%. They supplied 20.4% of listings from 7.2% of the cohort. At the other end, cross-site scripting made up 14.9% of the cohort and 2.3% of listings (a rate of 0.07%), and SQL injection, a perennial entry on weakness top-ten lists, was listed at 0.12%. Authentication and access-control flaws (0.54%) and the path, upload and request-forgery group (0.55%) sat close to memory safety, so that four families with very different mechanics, from use-after-free bugs in browsers to missing authorization checks in management consoles, were listed at almost the same rate.

The issuing CNA is the strongest single marker in the data (Figure 4). Vendor CNAs for network-edge and platform products had listing rates 5 to 12 times the cohort mean. Palo Alto Networks led at 5.0% (12 of 238), ahead of Fortinet at 3.1%, Microsoft and Cisco at 2.5% each (130 of 5,101 for Microsoft) and Apple at 2.0%. The high-volume third-party and ecosystem CNAs sat at or near zero, Patchstack, Wordfence and WPScan with none in 29,806 and VulDB with 4 in 14,154, while MITRE, which assigns identifiers as CNA of last resort and therefore covers a mixture of everything, came close to the cohort mean at 0.43%. CNAs differ in what software they cover; attackers concentrate on enterprise products exposed to the internet. The record is public. A CNA's history of listings can be read off the catalog at the moment a new CVE appears, and 83.5% of listed cohort CVEs came from CNAs with at least one earlier listing, against 56.9% of all cohort CVEs.

5.3 How soon listing happens

Listing is fast when it happens (Figure 5). Of the 732 listed CVEs, 27.0% were listed on or before their NVD publication day, including 7.7% listed before it; half were listed within 8 days and 64.1% within 30 days, while 6.7% were listed more than a year after publication. Same-day listing is a vendor

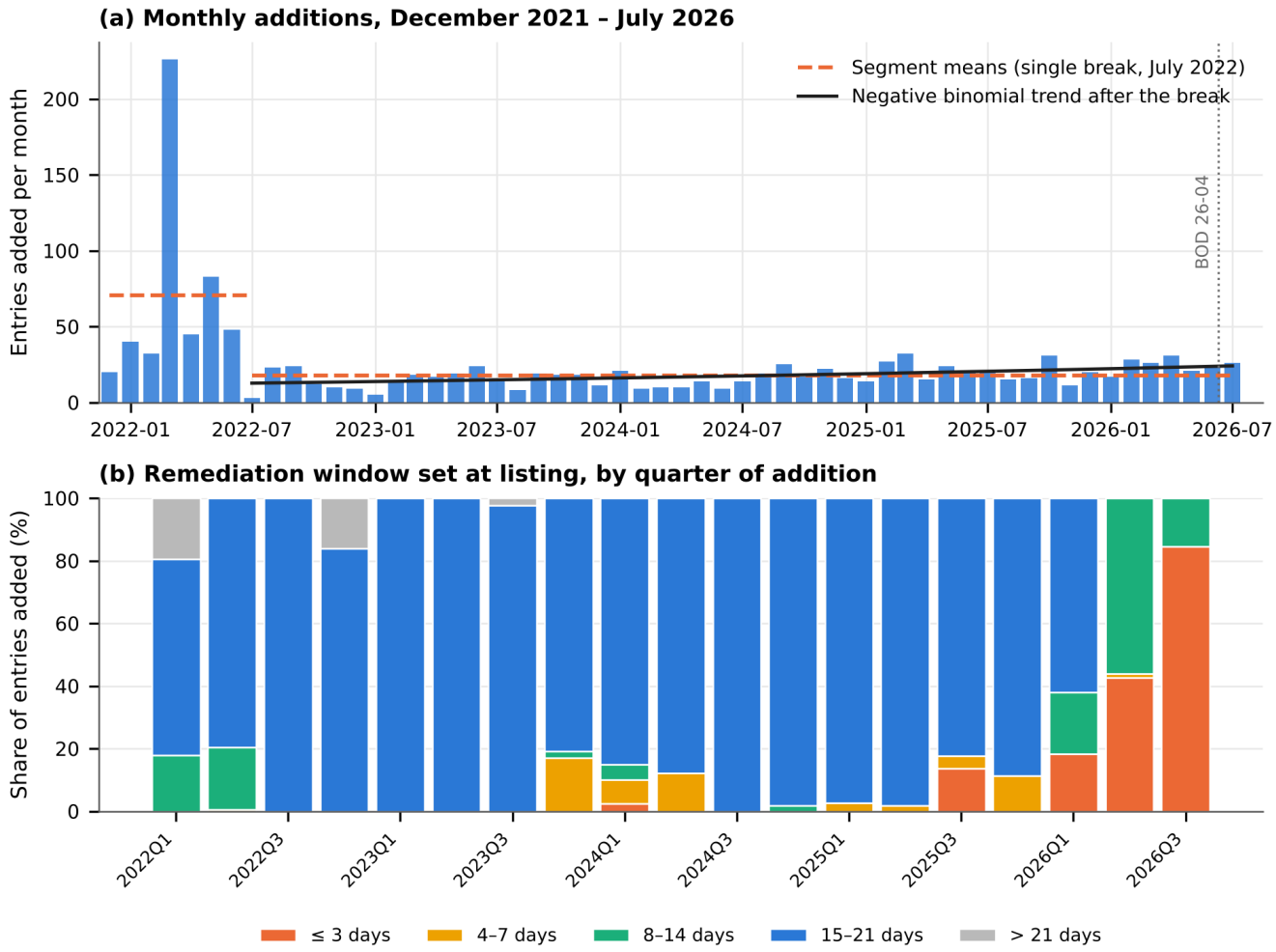


Figure 1. KEV catalog additions and remediation windows. (a) Entries added per month, December 2021 to July 2026, excluding the founding batch of 3 November 2021, with the segment means on either side of the estimated break and the negative binomial trend fitted after it. (b) Distribution of the remediation window (due date minus date added) by quarter of addition; 2026 Q3 contains July only.

Table 2. Catalog additions by calendar year of addition

Year added	Entries	Median CVE age at listing (years)	CVE from same or previous year (%)	21-day window (%)	14-day window (%)	Window $\leq$ 7 days (%)	Known ransomware use (%)
2021	311	1	69.5	0.0	35.7	0.0	24.8
2022	555	4	25.8	72.6	15.9	0.2	22.5
2023	187	0	77.5	94.7	0.0	4.3	23.0
2024	186	0	76.3	94.1	0.0	4.3	22.6
2025	245	0	77.6	92.2	0.0	7.8	10.6
2026 (to 29 Jul)	172	0	77.9	25.6	34.9	39.5	11.1

Table 3. Cohort composition and KEV listing by disclosure severity band

Severity at disclosure	CVEs	Share of cohort (%)	Listed	Listing rate (%)	Share of listings (%)
Critical (9.0–10.0)	17,135	10.2	286	1.67	39.1
High (7.0–8.9)	59,862	35.7	344	0.57	47.0
Medium (4.0–6.9)	78,572	46.9	98	0.12	13.4
Low (0.1–3.9)	7,266	4.3	4	0.06	0.5
No v3.x score	4,744	2.8	0	0.00	0.0
All	167,579	100.0	732	0.44	100.0

phenomenon. It applied to 64% of Microsoft's listed CVEs, many disclosed on Patch Tuesday as already exploited, to 48% of Apple's and 43% of Fortinet's, and to none of the CVEs issued by Apache, Oracle or the Linux kernel CNA.

Raw lag distributions mislead when follow-up differs by cohort, because a 2026 CVE can only appear in the data if it was listed quickly, and the restriction used in Table 4 handles this for the 365-day horizon by keeping only CVEs with a full year of follow-up. Among the 517 CVEs published by 29 July 2025 and listed within a year of publication, 29.0% were listed by day 0, 53.0% by day 7, 66.5% by day 30, 79.3% by day 90 and 89.6% by day 180. So much for the first month. A defender who waits a month for confirmation has, by then, let two-thirds of the year's eventual listings pass.

Note. Lag columns describe listed CVEs only. Later cohorts are right-censored, so their medians and day-30 shares are biased toward faster listing.

The Kaplan–Meier curves in Figure 6 put these timings on the scale of the whole cohort, where the cumulative listing

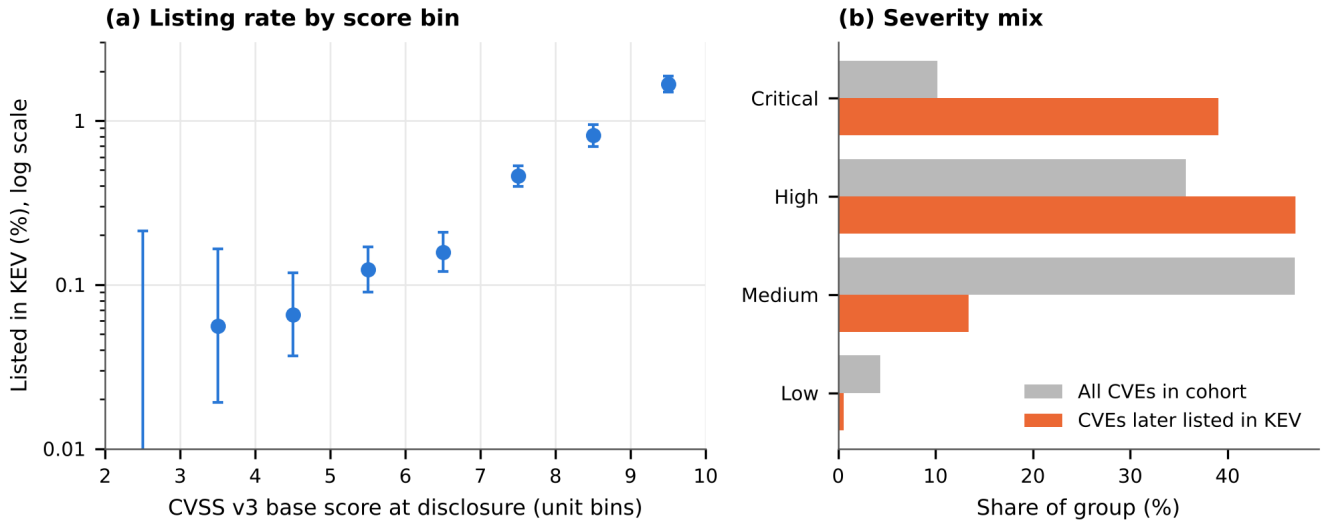


Figure 2. Severity at disclosure and KEV listing. (a) Share of CVEs listed by 29 July 2026, by unit bin of the disclosure CVSS v3 base score, with 95% Wilson intervals (log scale; the 2.0–2.9 bin had no listings). (b) Severity band shares among all cohort CVEs and among listed CVEs.

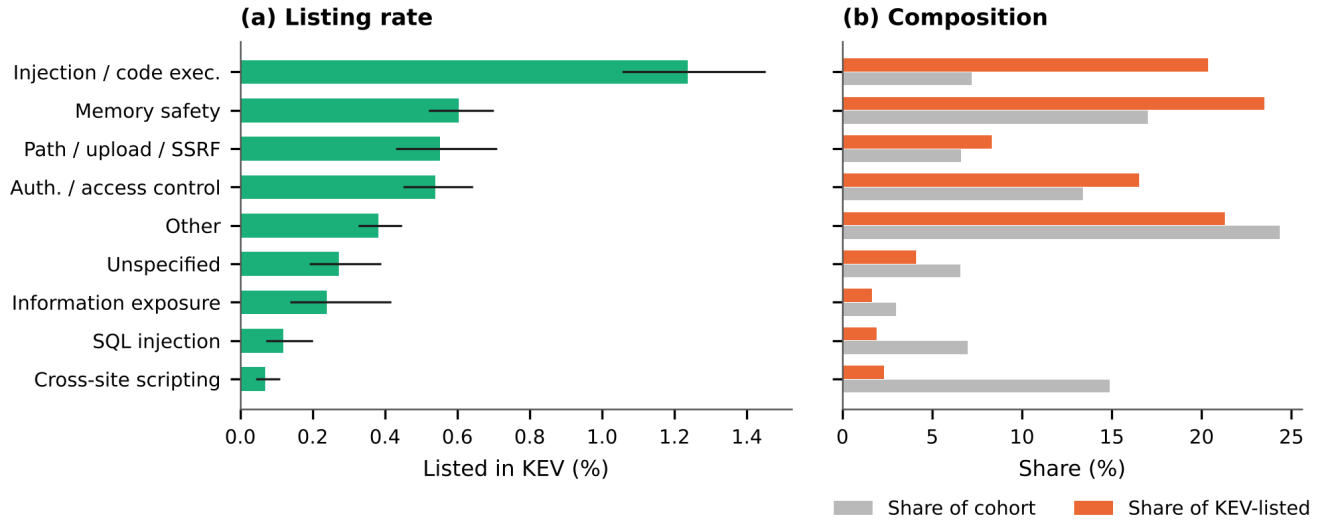


Figure 3. Weakness family and KEV listing. (a) Listing rate with 95% Wilson intervals. (b) Share of cohort CVEs and of listed CVEs in each family.

Table 4. Listing outcomes and lags by publication year of the CVE

Published	CVEs	Listed	Listing rate (%)	Median lag (days)	Listed by day 0 (%)	Listed by day 30 (%)
2022	19,795	119	0.60	21.0	31.1	53.8
2023	27,194	163	0.60	6.0	27.6	65.0
2024	38,709	160	0.41	11.0	25.0	59.4
2025	47,587	194	0.41	7.5	25.3	64.9
2026 (Jan–Jun)	34,294	96	0.28	5.0	28.1	81.3

probability reached 0.12% by day 1, 0.28% by day 30, 0.43% by day 365 and 0.50% by day 1,460, so that more than half of the four-year total had accumulated within a month of publication. For Critical CVEs the corresponding values were 0.25%, 1.03%, 1.68% and 1.83%; for High, 0.21%, 0.38%, 0.54% and 0.68%; for Medium, 0.04%, 0.08%, 0.13% and 0.14%. High CVEs come close to Critical ones on the first day (0.21% against 0.25%) and fall behind later, a sign that same-day vendor disclosures of exploited flaws are often scored High.

The Cox model (Table 5) separates these influences. Holding weakness family, CNA history and publication year fixed, a Critical score multiplied the listing hazard by 10.5 (95% CI 8.0–13.8) and a High score by 4.1 (3.2–5.1) relative to Medium and Low. Injection and code-execution weaknesses raised the hazard by a factor of 1.87 and authentication flaws by 1.45; SQL injection (0.19) and cross-site scripting (0.31) lowered it. CNA history mattered too. Each doubling of $1 + P_i$ raised the hazard by about 21% (HR 1.32 per log unit). Network attack vector and absence of required privileges, the two vector fields most often cited as signs of exploitability, had no detectable effect once severity was in the model; their information is already in the base score. The model ranks pairs well (concordance 0.82), and its publication-year term of 0.86 per year matches the dilution of the denominator by low-yield CNAs that Section 5.2 described.

Note. Reference categories are Medium or Low severity and the "other" and "unspecified" weakness families. Concor-

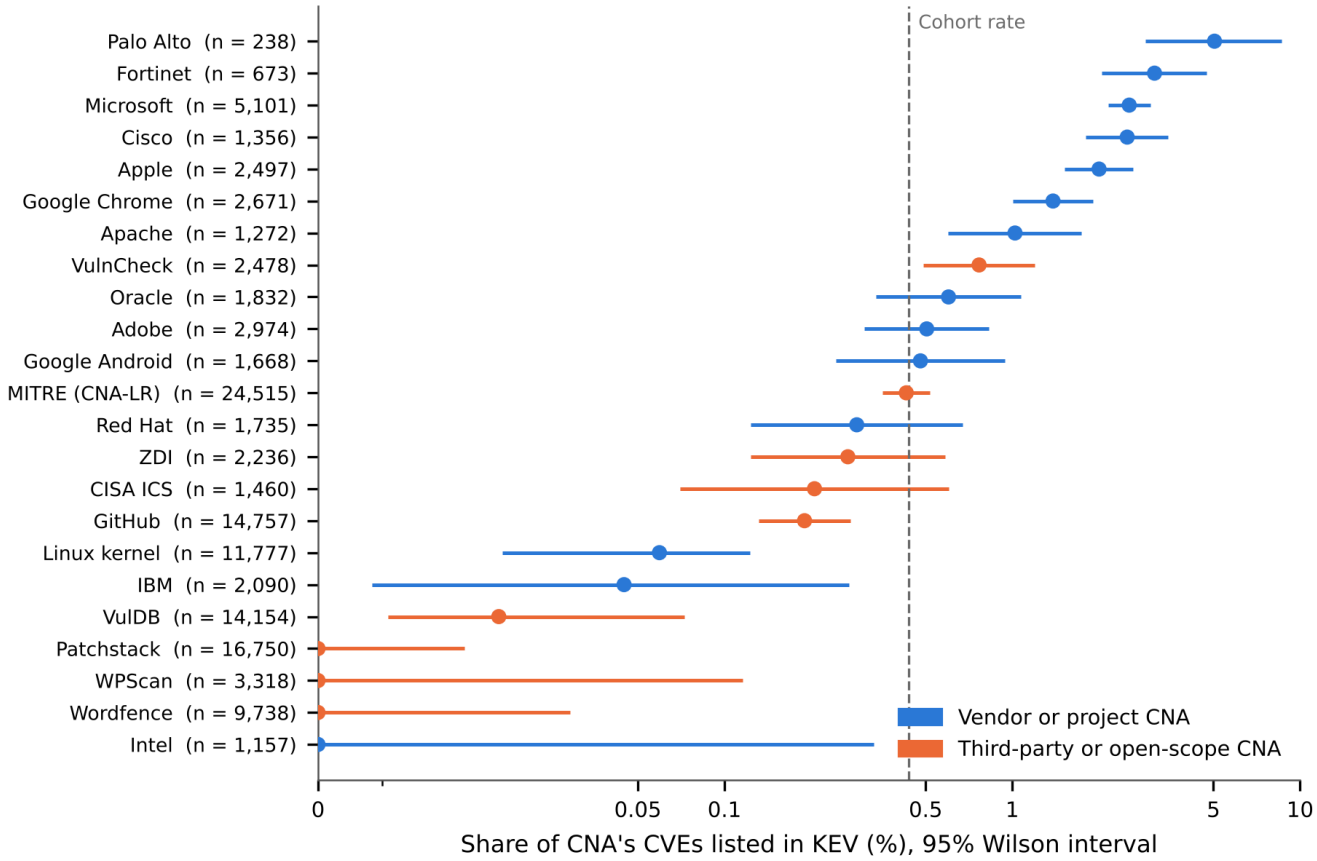


Figure 4. KEV listing rate by issuing CNA, for the 23 named CNAs in the cohort, with 95% Wilson intervals on a symmetric-log scale. The dashed line marks the cohort rate of 0.44%.

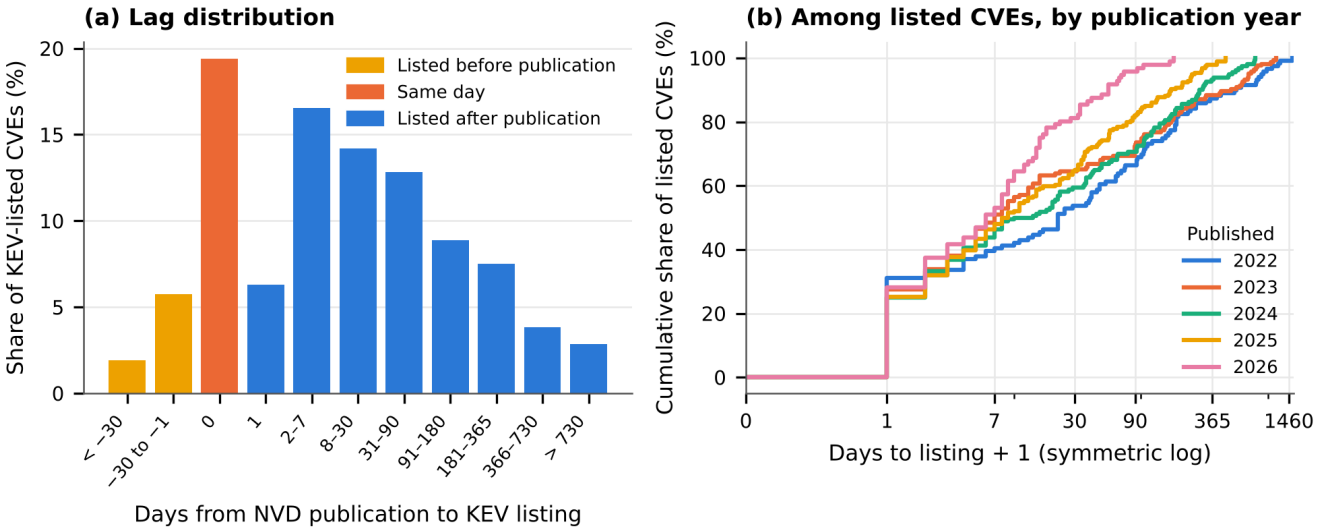


Figure 5. Days from NVD publication to KEV listing. (a) Distribution among all listed cohort CVEs. (b) Cumulative distribution among listed CVEs by publication year; curves for 2025 and especially 2026 are truncated by shorter follow-up and overstate speed.

dance of the full model is 0.82.

The proportional-hazards assumption fails for seven of the thirteen covariates (Schoenfeld tests at the 5% level), and the failures have a pattern that the piecewise model in Equation 6 makes visible (Figure 7). Memory-safety flaws carry an elevated hazard in the first 30 days (HR 1.30) and a reduced one afterwards (0.60), which fits the browser and operating-system bugs that vendors disclose and patch as zero-days: if such a flaw is listed at all, it is listed at once. Injection, authentication and path-traversal flaws run the other way. After day 30 their hazard ratios rise to 2.63, 1.68 and 1.58, a profile consistent with appliance and web-facing enterprise software that is weaponized after disclosure, once attackers have reverse-engineered the patch. Cross-site scripting is nearly excluded from early listing (0.12) and indistinguishable from the reference after day 30. A rule that treats "no user interaction" as an exploitability signal works only in the first month (0.59 early, 1.51 later), when most early listings are client-side bugs triggered by opening a page or file.

5.4 Ranking new CVEs for remediation

The prioritization test asks whether a team that ranked CVEs as they were published between July 2024 and June 2025

Table 5. Cox models of time from NVD publication to KEV listing: hazard ratios with 95% confidence intervals

Covariate	Full follow-up	Days 0–30	After day 30
Critical (vs Medium/Low)	10.49 (7.96–13.81)	11.52 (8.13–16.33)	8.91 (5.68–13.98)
High (vs Medium/Low)	4.08 (3.24–5.13)	4.15 (3.11–5.54)	3.86 (2.64–5.65)
Attack vector: network	0.90 (0.73–1.09)	0.88 (0.69–1.13)	0.90 (0.63–1.28)
No privileges required	1.08 (0.89–1.30)	1.06 (0.84–1.35)	1.09 (0.80–1.48)
No user interaction	0.76 (0.62–0.94)	0.59 (0.46–0.75)	1.51 (0.99–2.30)
CWE: memory safety	1.08 (0.88–1.34)	1.30 (1.02–1.66)	0.60 (0.39–0.93)
CWE: injection / code execution	1.87 (1.49–2.34)	1.48 (1.10–1.98)	2.63 (1.84–3.75)
CWE: authentication / access control	1.45 (1.15–1.83)	1.32 (0.98–1.77)	1.68 (1.15–2.44)
CWE: path / upload / SSRF	1.13 (0.85–1.52)	0.89 (0.60–1.32)	1.58 (1.02–2.46)
CWE: SQL injection	0.19 (0.11–0.34)	0.20 (0.10–0.40)	0.18 (0.07–0.46)
CWE: cross-site scripting	0.31 (0.19–0.53)	0.12 (0.05–0.31)	0.91 (0.46–1.81)
log(1 + CNA prior listings)	1.32 (1.26–1.38)	1.38 (1.29–1.46)	1.22 (1.13–1.32)
Publication year (per year)	0.86 (0.81–0.91)	0.84 (0.78–0.90)	0.90 (0.81–1.00)
Listings	732	469	263

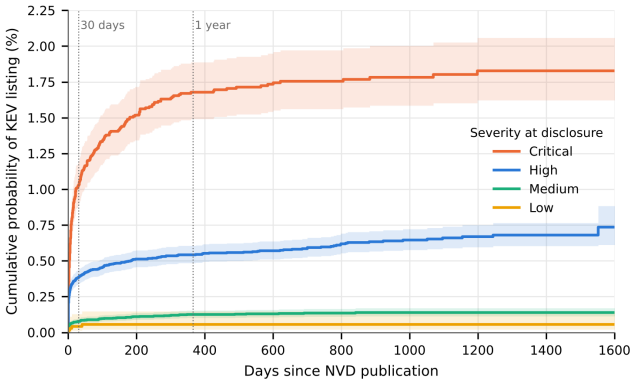


Figure 6. Cumulative probability of KEV listing by days since NVD publication and disclosure severity band (one minus the Kaplan–Meier estimate), with 95% pointwise intervals.

would have caught the 193 that were listed within a year, using a model trained on January 2022 to June 2024, when 311 CVEs met the same definition. Nothing from the test year enters the fit. Table 6 reports the interpretable training-set model, and its odds ratios repeat the Cox pattern: 8.85 for Critical and 3.74 for High, 2.00 for injection and 0.13 for SQL injection, and 1.24 per log unit of CNA history.

Note. McFadden pseudo-R^2 is 0.094. Prevalence in training is 0.47%.

Note. Workload is the share of test-period CVEs remediated, taken from the top of each ranking. The last row uses NVD fields that are partly added after listing and is an upper bound, not a usable ranking at disclosure.

Most organizations run threshold rules. Table 8 evaluates them at their own workloads. Patching every CVE scored 9.0 or above meant remediating 8.6% of new CVEs and caught 39.4% of later listings, with 2.1% of the patched CVEs turning out to be listed, whereas the logistic model cut at the same 8.6% caught 60.1% and boosted trees 64.2%, roughly 40 and 48 more listed CVEs in a year for an identical amount of patching. The 7.0 rule caught 82.4% of listings by remediating 43.4% of CVEs; at that workload the models caught 92.7% and 94.8%. Restricting the 9.0 rule to network-reachable flaws needing no privileges shrank the workload to 7.5% and coverage to 36.3%, which means that the vector filter removed listed and unlisted CVEs in nearly the same proportion. Efficiency barely moved (2.2%).

Which features do the work? Table 9 removes groups one at a time. Entered as separate components, the CVSS vector already beats the base score on AUC (0.806 against 0.767), though not on average precision, because the components carry information that the scoring formula compresses away into a single number. Adding weakness family moves cov-erage at 5% workload from 22.3% to 26.4%. Adding CNA identity and history is the big step, to an AUC of 0.879 and 46.6% coverage, and CNA information alone, with no sever-ity at all, ranks better than the CVSS score (AUC 0.833). Description length adds nothing measurable.

Calibration matters if the probabilities themselves are to feed a cost model of the kind proposed by Dey et al. [21], in which the expected loss from leaving a flaw open is weighed against the cost of an emergency patch. By decile of predicted probability (Figure 9), the logistic model is well calibrated in the two deciles that hold almost all listings, where the top decile predicted 3.30% and observed 3.04% and the ninth predicted 0.56% and observed 0.63%. In the seven lowest deciles the observed rates are noisy, from zero to 0.21%, and the model under-predicts in four of them; every rate there is below 0.25%. The sensitivity model with NVD reference tags reached an average precision of 0.212, nearly four times the disclosure model. Most of that gap is leakage. The “Exploit” tag is frequently attached to references added after exploitation was reported, and NVD analyzes listed CVEs first.

5.5 Whose score: the CNA’s or NVD’s

Whose score is used changes the ranking (Figure 10). A Secondary score from the CNA existed for 81.7% of cohort CVEs and for 92.9% of listed ones, while NVD’s Primary score existed for 54.7% and 64.3%, a gap that follows the enrichment backlog NVD accumulated from February 2024 onward [16]. Where both existed (65,794 CVEs), NVD’s score was higher by 0.58 points on average, the two differed by a full point or more in 47.1% of cases, and they fell in the same severity band only 62.3% of the time. Disagreement was asymmetric: 73% of CVEs scored Low by the CNA were scored Medium by NVD, and 25% of CNA-Medium CVEs were scored High. Listed CVEs behaved differently. Among the 419 with both scores the average gap was 0.05 points, with NVD higher in 15.3% and lower in 15.0%. NVD’s upward revisions thus fell mostly on CVEs that were never exploited, which explains the ranking result in Table 7: the NVD-preferred score had an AUC of 0.739 and an average precision of 0.015, both below the disclosure score’s 0.767 and 0.025.

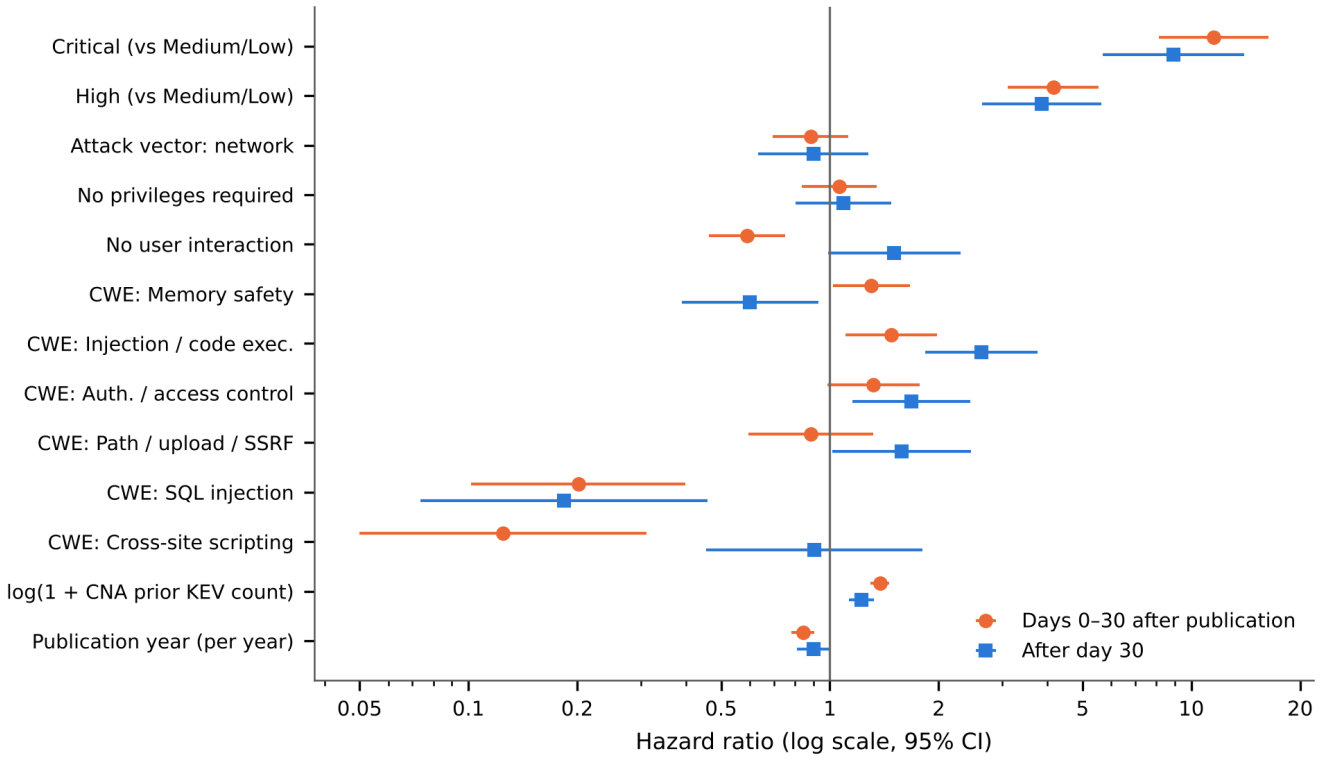


Figure 7. Hazard ratios for KEV listing in the first 30 days after publication and after day 30, from the piecewise Cox model, with 95% confidence intervals on a log scale.

Table 6. Logistic model of listing within 365 days, training period January 2022 – June 2024: odds ratios with 95% confidence intervals

Covariate	Odds ratio	95% CI	p
Critical (vs Medium/Low)	8.85	5.75–13.62	< 0.001
High (vs Medium/Low)	3.74	2.63–5.31	< 0.001
Attack vector: network	0.97	0.72–1.31	0.856
No privileges required	1.07	0.79–1.44	0.679
No user interaction	0.59	0.43–0.80	0.001
CWE: memory safety	1.31	0.96–1.78	0.087
CWE: injection / code execution	2.00	1.40–2.86	< 0.001
CWE: authentication / access control	1.66	1.15–2.40	0.007
CWE: path / upload / SSRF	0.91	0.55–1.49	0.704
CWE: SQL injection	0.13	0.05–0.36	< 0.001
CWE: cross-site scripting	0.20	0.09–0.48	< 0.001
log(1 + CNA prior listings)	1.24	1.15–1.34	< 0.001

Table 7. Ranking new CVEs for KEV listing within 365 days, test period July 2024 – June 2025

Ranking	AUC (95% CI)	Average precision (95% CI)	Coverage, 1% workload (%)	Coverage, 5% workload (%)	Coverage, 10% workload (%)	Efficiency, 5% workload (%)
CVSS, disclosure score	0.767 (0.732–0.799)	0.025 (0.015–0.043)	8.3	22.3	42.5	2.0
CVSS, NVD-preferred score	0.739 (0.705–0.770)	0.015 (0.011–0.024)	6.7	17.6	30.6	1.6
Logistic, disclosure features	0.881 (0.857–0.902)	0.058 (0.038–0.091)	16.6	43.0	67.4	3.9
Boosted trees, disclosure features	0.891 (0.872–0.911)	0.075 (0.048–0.116)	15.5	53.4	68.9	4.8
Logistic + NVD reference tags (sensitivity)	0.922 (0.902–0.941)	0.212 (0.157–0.271)	34.2	63.2	77.7	5.7

Table 8. Threshold rules against model rankings cut at the same workload, test period July 2024 – June 2025

Policy	Workload (%)	Coverage (%)	Efficiency (%)
CVSS $\geq$ 9.0	8.6	39.4	2.1
Logistic model, same workload	8.6	60.1	3.2
Boosted trees, same workload	8.6	64.2	3.4
CVSS $\geq$ 9.0 with network vector and no privileges	7.5	36.3	2.2
CVSS $\geq$ 7.0	43.4	82.4	0.9
Logistic model, same workload	43.4	92.7	1.0
Boosted trees, same workload	43.4	94.8	1.0

5.6 Vendor concentration and ransomware labels

The catalog spans 276 vendor names, 144 of which appear once, and Microsoft alone accounts for 382 entries (23.1%). Concentration fell between the two addition periods (Table 10, Figure 11). The Gini coefficient dropped from 0.738 for 2021–2023 additions to 0.599 for 2024–2026, the Herfind-

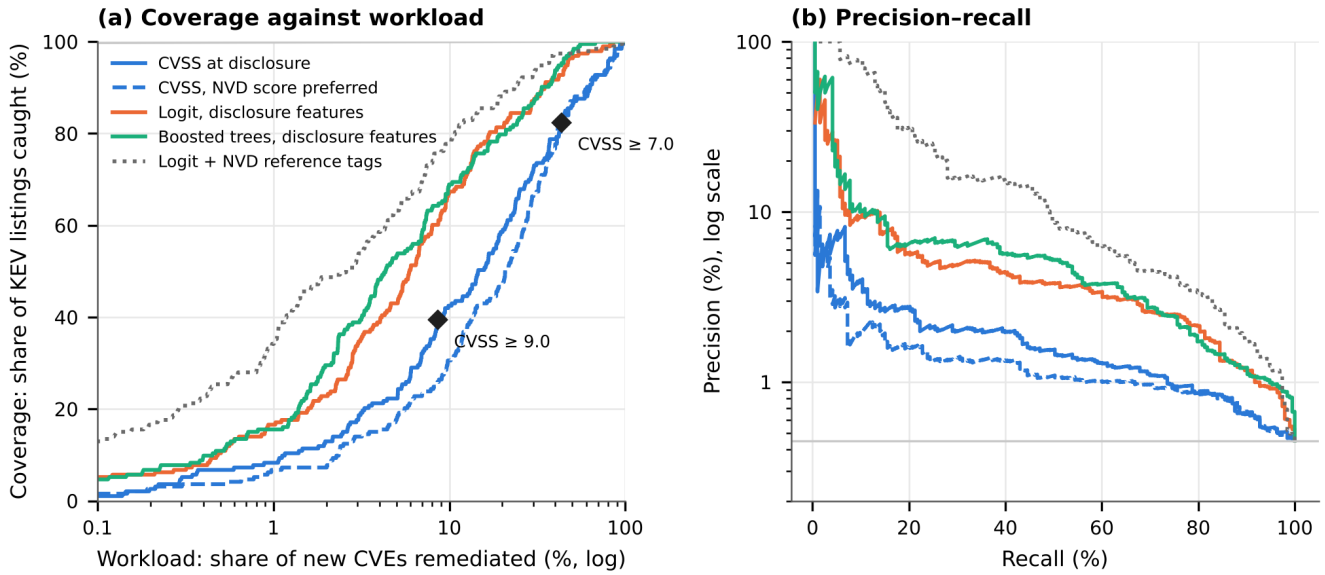


Figure 8. Ranking performance on the test year. (a) Coverage of KEV listings within 365 days against workload, on a log workload axis; diamonds mark the CVSS 7.0 and 9.0 threshold rules. (b) Precision–recall curves on a log precision axis; the grey line marks prevalence.

Table 9. Feature-group ablation for the logistic ranking, test period July 2024 – June 2025

Features	AUC	Average precision	Coverage at 5% workload (%)
CVSS base score only (reference)	0.767	0.025	22.3
CVSS vector components	0.806	0.019	22.3
CVSS vector + weakness family	0.813	0.025	26.4
CVSS vector + weakness family + CNA	0.879	0.049	46.6
CNA identity and history only	0.833	0.021	35.2
All disclosure features	0.880	0.055	45.1

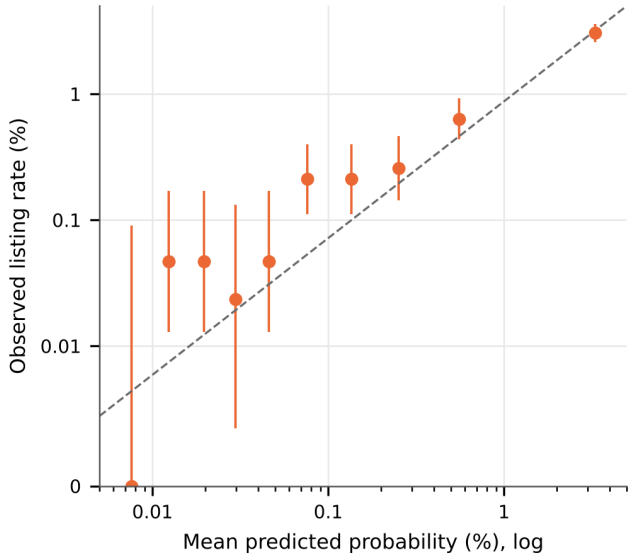


Figure 9. Calibration of the disclosure-feature logistic model on the test year: observed listing rate against mean predicted probability by decile, with 95% Wilson intervals. The dashed line is perfect calibration.

ahl–Hirschman index halved from 0.087 to 0.044, the top-ten share fell from 60.6% to 45.1%, and Microsoft's share fell from 26.1% to 17.7% while Ivanti, Fortinet and Linux grew, which is the decline H4 predicted. The catalog is still far narrower than the vulnerabilities themselves, since NVD product configurations name 13,501 vendors in the cohort (Gini 0.825 but an HHI of only 0.013) and the catalog draws a few hundred of them from that long tail. Repeat vendors dominate. Of 790 entries added since January 2023, 82.8% named a vendor already present in the catalog.

Note. The NVD column counts the first vendor in each CVE's product configuration and excludes CVEs that NVD had not configured.

Ransomware labels behave differently. On 29 July 2026, 332 of 1,656 entries (20.0%) carried "Known" ransomware use. Rates differed by vendor class (Figure 12), at 27.2%for Microsoft, 25.5% for edge and enterprise-infrastructure vendors (firewalls, VPN gateways, file-transfer and virtual-ization products), 19.6% for other vendors and one entry in 231 (0.4%) for mobile and browser vendors. Path-traversal, upload and request-forgery flaws were flagged most often (33.8%), memory-safety flaws least (9.7%). In Table 11 these differences survive adjustment for time since listing, which it-self carries an odds ratio of 1.32 per year (95% CI 1.21–1.45), so an entry listed three years ago is more than twice as likely to carry the flag as an otherwise similar entry listed this year. All 85 flag changes between January 2025 and July 2026 went from "Unknown" to "Known", at a median of 287.5 days after listing (interquartile range 20 to 818 days, maxi-mum 1,695). The git history explains the time effect. Of 403 entries added after 29 January 2025, 39 were flagged by the snapshot, but only 3 of these carried the flag on the day they were added, so H5 is supported, and the lower ransomware shares for 2025 and 2026 in Table 2 are, at least in part, labels not yet applied.

Note. All 1,656 entries. McFadden pseudo-R^2 is 0.112; AUC 0.715.

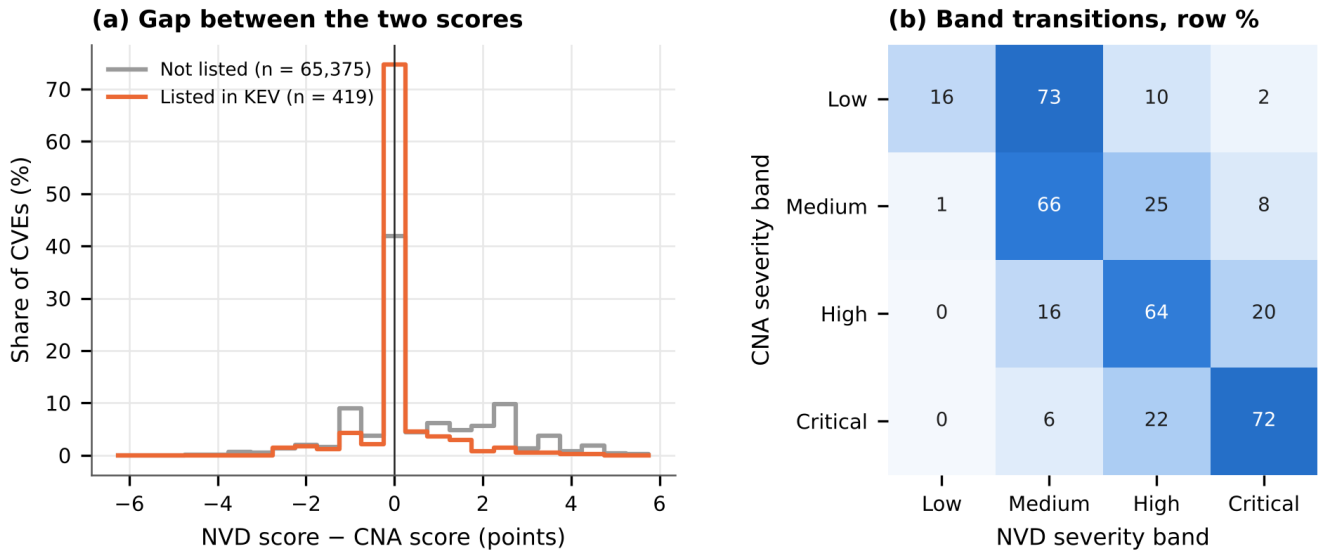


Figure 10. CNA and NVD CVSS v3 scores for CVEs carrying both. (a) Distribution of the difference (NVD minus CNA) for listed and unlisted CVEs. (b) Severity band transitions from the CNA score (rows) to the NVD score (columns), as row percentages.

Table 10. Vendor concentration in the KEV catalog by addition period, with the NVD cohort for comparison

Measure	KEV, added 2021–2023	KEV, added 2024–2026	NVD cohort (CPE vendor)
Entries	1,053	603	129,500
Distinct vendors	167	173	13,501
Gini coefficient	0.738	0.599	0.825
Herfindahl–Hirschman index	0.087	0.044	0.013
Top-ten vendor share (%)	60.6	45.1	—
Microsoft share (%)	26.1	17.7	—

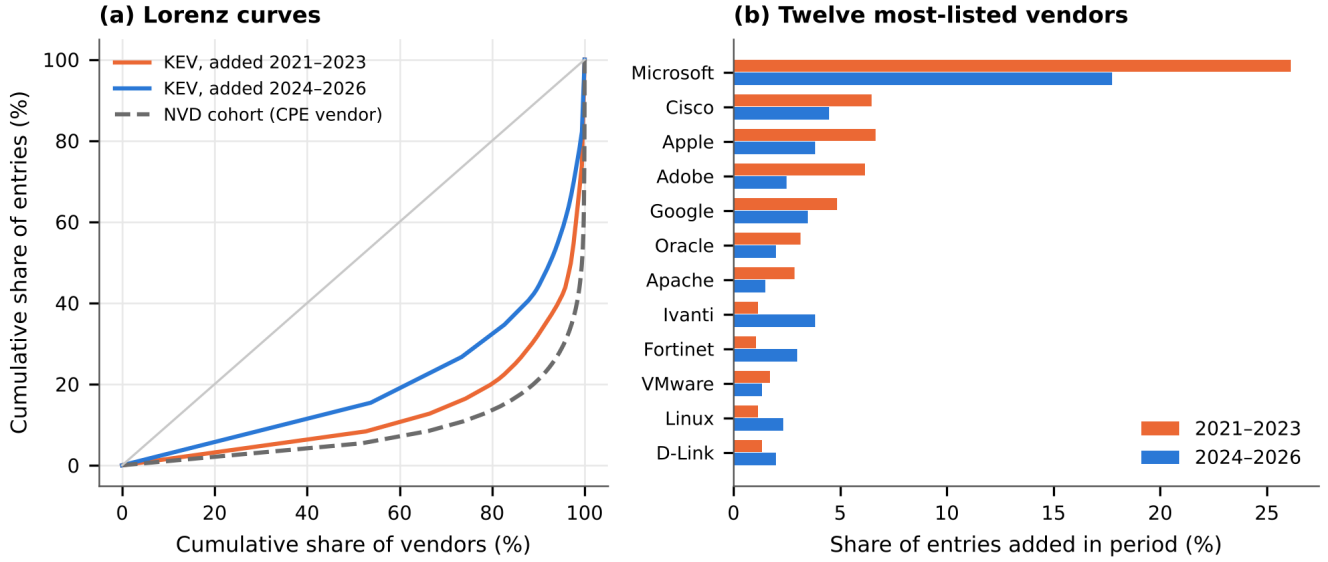


Figure 11. Vendor concentration. (a) Lorenz curves of entries across vendors for the two KEV addition periods and for the NVD cohort. (b) Shares of the twelve most-listed vendors in each KEV addition period.

Table 11. Logistic model of "Known" ransomware use among KEV entries: odds ratios with 95% confidence intervals

Covariate	Odds ratio	95% CI	p
Vendor class: Microsoft (vs other vendors)	1.69	1.23–2.34	0.001
Vendor class: edge / enterprise infrastructure	1.38	1.01–1.88	0.046
Vendor class: mobile / browser	0.02	0.003–0.16	< 0.001
CWE: authentication / access control (vs other)	1.79	1.19–2.69	0.005
CWE: injection / code execution	0.87	0.59–1.27	0.465
CWE: memory safety	0.57	0.37–0.87	0.010
CWE: path / upload / SSRF	2.05	1.31–3.22	0.002
CWE: unspecified	1.02	0.65–1.60	0.929
CVE three or more years old when listed	0.66	0.50–0.88	0.004
Years since listing (per year)	1.32	1.21–1.45	< 0.001

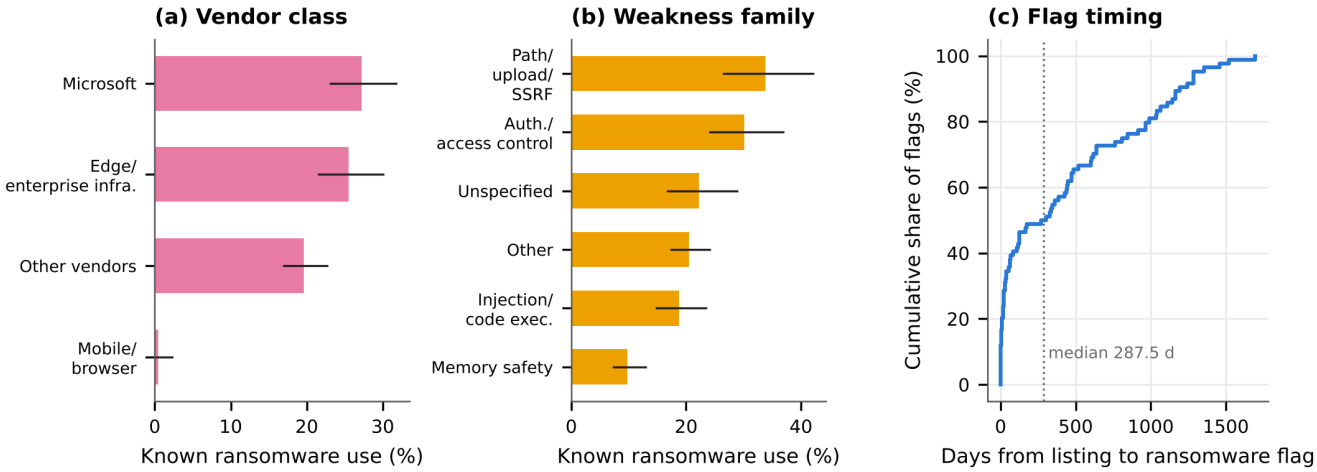


Figure 12. Ransomware labels in the KEV catalog. (a) Share of entries with "Known" ransomware use by vendor class and (b) by weakness family, with 95% Wilson intervals. (c) Cumulative distribution of the days between listing and the change of the ransomware field to "Known", for the 84 changes between January 2025 and July 2026 matched to current entries.

6. DISCUSSION

6.1 What the evidence says about the hypotheses

6.2 Implications for vulnerability management

Start with the CVSS threshold. The score carries real signal (the Critical hazard ratio is 10.5), yet a single threshold spends most of its effort on CVEs from streams that are almost never listed. A team that patches everything at 9.0 or above in a typical year spends about 2% of that effort on CVEs that are later listed; ranking with the same public fields would have raised that to over 3% and caught half again as many listings. None of the extra features requires a threat feed. CNA identity, CWE and the CNA's own listing history are published with the CVE, and the history term can be computed from the catalog itself. EPSS reports a similar pattern with richer, partly private inputs [8], which fits the disagreement among scoring systems documented by Koscinski et al. [9]; what this paper adds is a measure of how much of the gap a fully public model, trained only on information dated before each prediction, can close.

Timing comes next. Two-thirds of first-year listings arrive within 30 days and more than a quarter by the publication day, so most of the decision falls in the window when only CNA data exist. Waiting for NVD enrichment, or for the KEV listing itself, turns prioritization into clean-up. The piecewise results point to a two-track policy: client-side and memory-safety flaws in widely deployed platforms deserve immediate action or nothing, while injection, authentication and traversal flaws in edge appliances keep accumulating risk for months and justify sustained monitoring, especially once a patch is public and can be diffed [19].

Then there is the question of which score to trust. NVD's April 2026 decision to defer to CNA scores [4] was made on capacity grounds. Our results give it an accuracy argument as well, at least for ranking by exploitation: on the 2024–2025 test year the CNA score ranked better, and NVD's upward revisions of Low and Medium scores fell mostly on CVEs that were never listed. Whether the same holds after April 2026, when fewer NVD scores exist to compare, needs a later snapshot to settle.

6.3 Implications for the catalog as an instrument

Under BOD 26-04 the KEV status of a CVE now determines whether federal agencies have three days or fourteen [5], and the shift is already visible in the 2026 data. Windows of three days or fewer applied to 22 of 26 additions in July 2026. Shorter clocks raise the value of anticipation. If 29% of first-year listings arrive on the publication day, an agency that starts work only at listing loses nothing on those; for the remaining 71% the listing comes after days or weeks during which a ranked queue could have had the patch underway. The data cannot say whether the new deadlines are met, because the catalog records deadlines and not completions.

The ransomware field needs a different reading than it usually receives. Tools that boost a vulnerability's priority when KEV says "Known" are using a label that, for 36 of the 39 recently added entries that carry it, appeared after listing and, at the median, the better part of a year later. For anticipating ransomware use, the flag's correlates do better: Microsoft and edge-device vendors, and traversal and authentication flaws, which the logistic model in Table 11 identifies even after holding time since listing fixed.

6.4 Limitations

Five limits bound these conclusions. First, KEV is a lower bound on exploitation: it requires public evidence that CISA judges reliable and a clear remediation path, so exploited CVEs without vendor fixes or with only private evidence are missing, and some listed CVEs may have been exploited long before listing. Our outcome is "listed by CISA", which is what federal deadlines and many commercial tools key on, and the results should not be read as rates of exploitation in the wild.

Second, the NVD snapshot is a single July 2026 state. CNA scores can be revised after publication, and we cannot recover the score a defender saw on publication day for every CVE; the near-zero average gap between CNA and NVD scores for listed CVEs suggests that revision after listing is not driving the provenance result, but a series of daily snapshots would settle it. Third, the CNA history term is computed from 2022-and-later CVEs only and so understates the history of long-standing vendor CNAs. Fourth, the test year is one

Table 12. Hypotheses and verdicts

Hypothesis	Evidence	Verdict
H1: CVSS raises listing; weakness and CNA history add information	HR 10.5 Critical; injection 1.87; CNA history 1.32 per log unit (Table 5)	Supported
H2: hazard front-loaded	66.5% of first-year listings by day 30	Supported, weakness-dependent
H3: disclosure ranking beats CVSS at equal workload	Coverage 60.1% vs 39.4% at 8.6% workload; AUC gain 0.115	Supported
H4: vendor concentration declined	Gini 0.738 to 0.599; HHI 0.087 to 0.044	Supported
H5: ransomware flag is retrospective	OR 1.32 per year listed; 3 of 39 recent flags present at addition	Supported

year. A forward evaluation on the July 2025 to June 2026 cohort becomes possible in mid-2027. Fifth, we did not benchmark against EPSS. Its daily score archives sit outside the two repositories used here, so the paper sets a public-data baseline; a matched-date comparison with EPSS is the obvious next step.

7. CONCLUSION

Joined to four and a half years of NVD records, the July 2026 KEV catalog tells a consistent story: listing is rare, fast and predictable from public data to a degree the CVSS threshold leaves unused. Fewer than 0.5% of CVEs reach the catalog; two-thirds of those listed within a year are listed within a month; and a ranking built from severity, weakness family and the CNA's listing history catches 60% of later listings at the workload where a 9.0 threshold catches 39%. The CNA score ranks better than NVD's, which matters now that NVD has stepped back from scoring. Ransomware labels mostly arrive after the fact, and the deadlines attached to listing shortened sharply in 2026. The next test of these results is the 2025–2026 publication cohort, observable in full from mid-2027, which will also be the first scored largely by CNAs alone.

REFERENCES

[1] Forum of Incident Response and Security Teams. Common Vulnerability Scoring System version 3.1: Specification document, 2019.

[2] J. Spring, E. Hatleback, A. Householder, A. Manion, and D. Shick. Time to change the CVSS? IEEE Security & Privacy, 19(2):74–78, 2021.

[3] Cybersecurity and Infrastructure Security Agency. Binding Operational Directive 22-01: Reducing the significant risk of known exploited vulnerabilities, 2021.

[4] National Institute of Standards and Technology. NIST updates NVD operations to address record CVE growth. News release, April 15, 2026.

[5] Cybersecurity and Infrastructure Security Agency. Binding Operational Directive 26-04: Prioritizing security updates based on risk, 2026.

[6] J. Jacobs, S. Romanosky, I. Adjerid, and W. Baker. Improving vulnerability remediation through better exploit prediction. Journal of Cybersecurity, 6(1):tyaa015, 2020.

[7] J. Jacobs, S. Romanosky, B. Edwards, I. Adjerid, and M. Roytman. Exploit Prediction Scoring System (EPSS). Digital Threats: Research and Practice, 2(3):Article 20, 2021.

[8] J. Jacobs, S. Romanosky, O. Suciu, B. Edwards, and A. Sarabi. Enhancing vulnerability prioritization: Data-driven exploit predictions with community-driven insights. In 2023 IEEE European Symposium on Security and Privacy Workshops, pages 194–206, 2023.

[9] V. Koscinski, M. Nelson, A. Okutan, R. Falso, and M. Mirakhorli. Conflicting scores, confusing signals: An empirical study of vulnerability scoring systems. In Proceedings of the 2025 ACM SIGSAC Conference on Computer and Communications Security, 2025.

[10] S. Frei, M. May, U. Fiedler, and B. Plattner. Large-scale vulnerability analysis. In Proceedings of the 2006 SIGCOMM Workshop on Large-Scale Attack Defense, pages 131–138, 2006.

[11] L. Allodi and F. Massacci. Comparing vulnerability severity and exploits using case-control studies. ACM Transactions on Information and System Security, 17(1):Article 1, 2014.

[12] Forum of Incident Response and Security Teams. Common Vulnerability Scoring System version 4.0: Specification document, 2023.

[13] J. M. Spring, E. Hatleback, A. Householder, A. Manion, and D. Shick. Prioritizing vulnerability response: A stakeholder-specific vulnerability categorization. Software Engineering Institute, Carnegie Mellon University, 2019.

[14] J. Ruohonen. A look at the time delays in CVSS vulnerability scoring. Applied Computing and Informatics, 15(2):129–135, 2019.

[15] A. Anwar, A. Abusnaina, S. Chen, F. Li, and D. Mohaisen. Cleaning the NVD: Comprehensive quality assessment, improvements, and analyses. IEEE Transactions on Dependable and Secure Computing, 19(6):4255–4269, 2022.

[16] U.S. Department of Commerce Office of Inspector General. Evaluation of NIST's management of the National Vulnerability Database. Report No. OIG-26-020-I, 2026.

[17] M. Bozorgi, L. K. Saul, S. Savage, and G. M. Voelker. Beyond heuristics: Learning to classify vulnerabilities and predict exploits. In Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pages 105–114, 2010.

[18] C. Sabottke, O. Suciu, and T. Dumitraș. Vulnerability disclosure in the age of social media: Exploiting Twitter for predicting real-world exploits. In Proceedings of the 24th USENIX Security Symposium, pages 1041–1056, 2015.

[19] O. Suciu, C. Nelson, Z. Lyu, T. Bao, and T. Dumitraș. Expected exploitability: Predicting the development of functional vulnerability exploits. In Proceedings of the 31st USENIX Security Symposium, pages 377–394, 2022.

[20] A. D. Householder, J. Chrabaszcz, T. Novelly, D. Warren, and J. M. Spring. Historical analysis of exploit availability timelines. In 13th USENIX Workshop on Cyber Security Experimentation and Test (CSET 20), 2020.

[21] D. Dey, A. Lahiri, and G. Zhang. Optimal policies for security patch management. INFORMS Journal on Computing, 27(3):462–477, 2015.

[22] A. Arora, R. Krishnan, R. Telang, and Y. Yang. An empirical analysis of software vendors' patch release behavior: Impact of vulnerability disclosure. Information Systems Research, 21(1):115–132, 2010.

[23] L. Y. Connolly, D. S. Wall, M. Lang, and B. Oddson. An empirical study of ransomware attacks on organizations: An assessment of severity and salient factors affecting vulnerability. Journal of Cybersecurity, 6(1):tyaa023, 2020.

[24] L. Allodi. Economic factors of vulnerability trade and exploitation. In Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security, pages 1483–1499, 2017.

[25] E. L. Kaplan and P. Meier. Nonparametric estimation from incomplete observations. Journal of the American Statistical Association, 53(282):457–481, 1958.

[26] D. R. Cox. Regression models and life-tables. Journal of the Royal Statistical Society: Series B (Methodological), 34(2):187–220, 1972.

[27] C. Davidson-Pilon. lifelines: Survival analysis in Python. Journal of Open Source Software, 4(40):1317, 2019.

[28] F. E. Harrell, K. L. Lee, and D. B. Mark. Multivariable prognostic models: Issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Statistics in Medicine, 15(4):361–387, 1996.

[29] P. M. Grambsch and T. M. Therneau. Proportional hazards tests and diagnostics based on weighted residuals. Biometrika, 81(3):515–526, 1994.

[30] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, G. Grisel, O. Blondel, M. Prettenhofer, R. Weiss, V. Dubourg, A. Vanderplas, J. Passos, D. Cournapeau, M. Brucher, M. Perrot, and É. Duchesnay. Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12:2825–2830, 2011.

[31] J. H. Friedman. Greedy function approximation: A gradient boosting machine. The Annals of Statistics, 29(5):1189–1232, 2001.

[32] T. Saito and M. Rehmsmeier. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. PLOS ONE, 10(3):e0118432, 2015.

[33] B. Efron and R. J. Tibshirani. An introduction to the bootstrap. Chapman & Hall, 1993.

[34] A. C. Cameron and P. K. Trivedi. Regression analysis of count data, 2nd ed. Cambridge University Press, 2013.

[35] R. Killick, P. Fearnhead, and I. Eckley. Optimal detection of changepoints with a linear computational cost. Journal of the American Statistical Association, 107(500):1590–1598, 2012.

[36] S. A. Rhoades. The Herfindahl–Hirschman index. Federal Reserve Bulletin, 79(3):188–189, 1993.