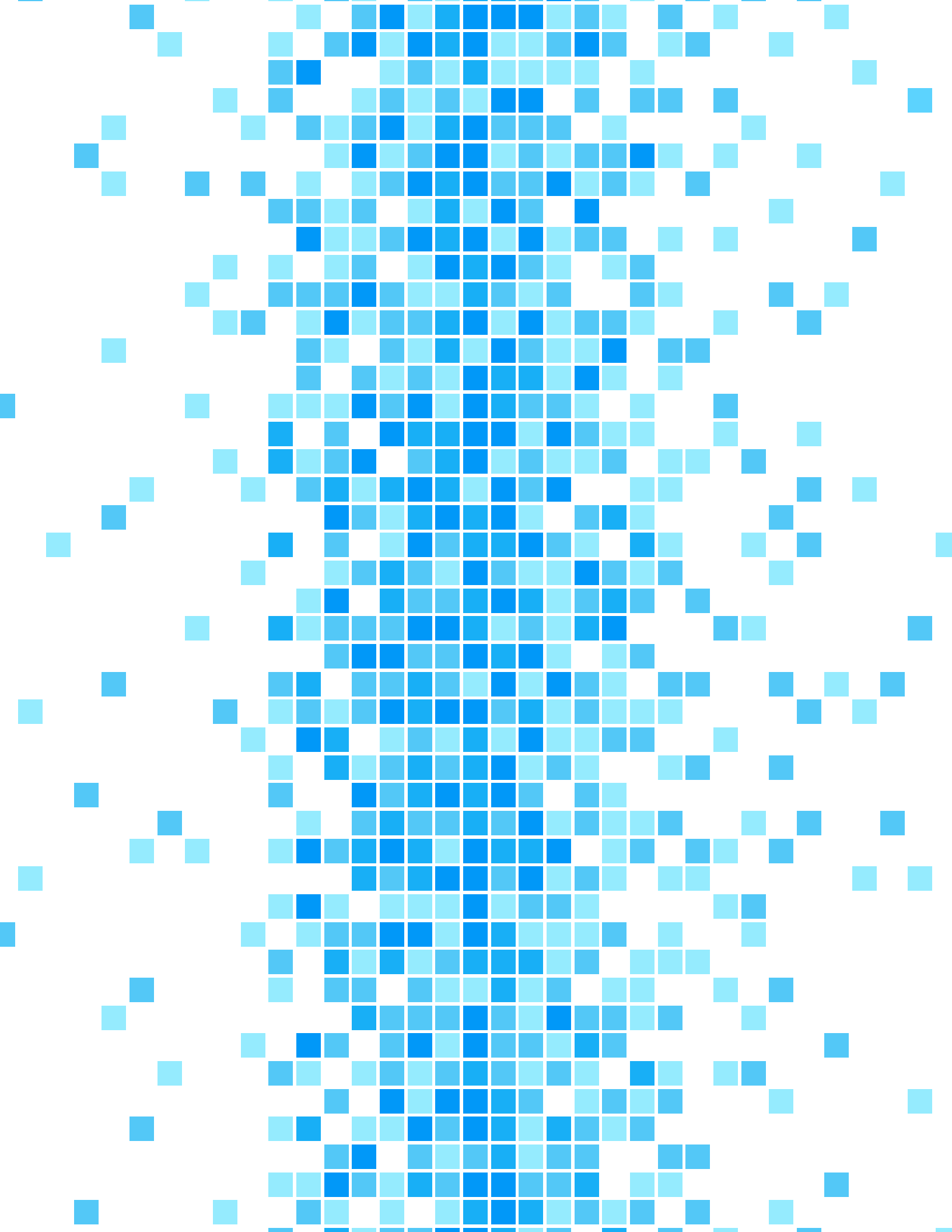
KENNA Security

Cyentia INSTITUTE

PRIORITIZATION TO PREDICTION

Volume 4: Measuring What Matters in Remediation



PRIORITIZATION TO PREDICTION

VOL 4: MEASURING WHAT MATTERS IN REMEDIATION

Introduction & Key Findings 2
Reviewing Remediation Metrics 4
– Coverage and Efficiency 4
– Remediation Velocity. 5
– Remediation Capacity 7
– Overall Remediation Performance 8
Analyzing Factors of Performance. 10
– Maturity and Performance 11
– Assets Under Management. 14
– Organizational Factors 16
– Remediation SLAs 19
– Prioritization Criteria 21
– Process Complexity. 23
– Deployment Methods. 25
Conclusion & Recommendations 28

KENNA Security

This research was commissioned by Kenna Security. Kenna collected and provided the remediation dataset to the Cyentia Institute for independent analysis and drafting of this report.

Kenna Security is a leader in predictive cyber risk. The Kenna Security Platform enables organizations to work cross-functionally to determine and remediate cyber risks. Kenna leverages Cyber Risk Context Technology™ to track and predict real-world exploitations, focusing security teams on what matters most. Headquartered in San Francisco, Kenna counts among its customers many Fortune 100 companies, and serves nearly every major vertical. Find out more at www.kennasecurity.com.

119 Cyentia INSTITUTE

Analysis for this report was provided by the Cyentia Institute. Cyentia seeks to advance cybersecurity knowledge and practice through data-driven research. We curate knowledge for the community, partner with vendors to create analytical reports like this one, and help enterprises gain insight from their data.

Find out more at www.cyentia.com.

Quick Summary of Prior Volumes

Volume 1: Analyzing Vulnerability Remediation Strategies

Proposed a theoretical model to predict which of the thousands of vulnerabilities published each month were most likely to be exploited, and thus deserving of priority remediation.

Volume 2: Getting Real About Remediation

Sought to apply and test the model proposed in Volume 1 in the real world using data extracted from hundreds of production environments inundated with billions of vulnerabilities.

Volume 3: Winning the Remediation Race

Studied remediation data from hundreds of organizations using a technique called survival analysis to measure how many vulnerabilities can be addressed within a given timeframe.

Introduction

> *Never cared for what they say*
> *Never cared for games they play*
> *Never cared for what they do...*
> *Forever trust in who you are*
> *And nothing else matters.*
>
> — *NOTHING ELSE MATTERS*, METALLICA

What matters in vulnerability management (VM)? What enables some programs to achieve higher levels of performance and greater success than others? Is it what they say (e.g., policies) or the products and tools with which they play? Does everything hinge on what they do or how they do it? Is it all about what they know (or don't know)? Or is Metallica right and all that really matters is trusting your team?

These are the kinds of questions we want to answer in this fourth volume of the Prioritization to Prediction (P2P) series. This report represents an analytical first for us and a rarity in the cybersecurity industry. We combine survey and observational data to test how internal VM program factors affect actual remediation performance measures. We hope this combo of "soft" and "hard" analysis yields valuable insights for vulnerability management programs.

If this is your first P2P experience, it's fine to jump in now, and we're glad to have you with us. But we recommend reviewing prior volumes at some point to get the most benefit from this research. In short, our ongoing goal in this series is to analyze tons of data to improve vulnerability management and ultimately reduce risk to organizations. Let's get started!

Key Findings

- **Maturity matters a lot.** Firms that give their VM programs high maturity scores *actually do* perform better overall and across more individual program metrics than any other factor in our study. Metallica was right after all—trust in who you are, and nothing else matters (quite as much).
- **As a rule, tools rule.** Remediation methods definitely matter. The use of patch management tools boosts remediation coverage, efficiency, capacity, and overall performance.
- **More spend, more speed.** Adequate budgets for the VM program correlate with an increased ability to remediate more vulnerabilities at a faster rate.
- **Divide and conquer.** When separate teams are responsible for finding and fixing vulnerabilities, we see increased velocity, capacity, and overall performance.
- **Deadlines shrink timelines.** Defining service-level agreements (SLAs) for vulnerabilities improves remediation velocity and overall performance.
- **Complexity has complex effects.** Complex processes are a mixed bag when it comes to remediation performance. Complexity negatively affects coverage but has some positive effects on velocity. Everything's a trade-off.
- **Get your priorities right.** Firms that heavily leverage Kenna's platform and risk scores to prioritize remediation efforts fix critical flaws a lot faster. Conversely, those basing prioritization decisions primarily on Common Vulnerability Scoring System (CVSS) perform worse than those who ignore it. We'll back those statements up with data, of course.

Key Findings From Previous Volumes

23% of vulnerabilities with published Common Vulnerabilities & Exposure (CVEs) have associated exploit code.

About one-third of published CVEs are actually observed in live enterprise environments.

Only 5% of all CVEs are both observed within organizations and known to be exploited.

40% of vulnerabilities observed in enterprise networks are still open (unremediated) today.

The median time-to-remediation is 100 days. 25% of vulnerabilities remain open over a year.

Any given organization, regardless of size, can address about one out of every 10 vulnerabilities.

Top-performing organizations remediate over twice the number of vulnerabilities at a rate three times faster than the norm.

Reviewing Remediation Metrics

If you've been following the P2P reports, you'll know we introduced and examined several metrics related to vulnerability remediation. If you're new to the series, this section won't make you an expert on everything covered so far, but it will review the basics so newcomers and veterans alike are well-calibrated for what's to come.

While we use consistent metrics from prior volumes, the sample of organizations is not consistent among them. This means we can't compare statistics across volumes to identify what's changed in the interim. In this section, we show remediation metrics for only the sample of ~100 organizations that participated in our recent survey. This preserves sample consistency throughout this report. With that level-setting out of the way, let's get into our first metric duo, coverage and efficiency.

Coverage and Efficiency

The core message of P2P Vol. 1 was that successful vulnerability management balances two simultaneous goals of coverage (fix everything that matters) and efficiency (delay/deprioritize what doesn't matter). The powerful thing about those goals is that they can be objectively measured and compared among different remediation strategies. A quick recap:

- **Coverage:** Measures the completeness of remediation efforts. What percentage of exploited or "high-risk" vulnerabilities were actually remediated?
- **Efficiency:** Measures the precision of remediation efforts. What percentage of vulnerabilities remediated were actually high-risk?

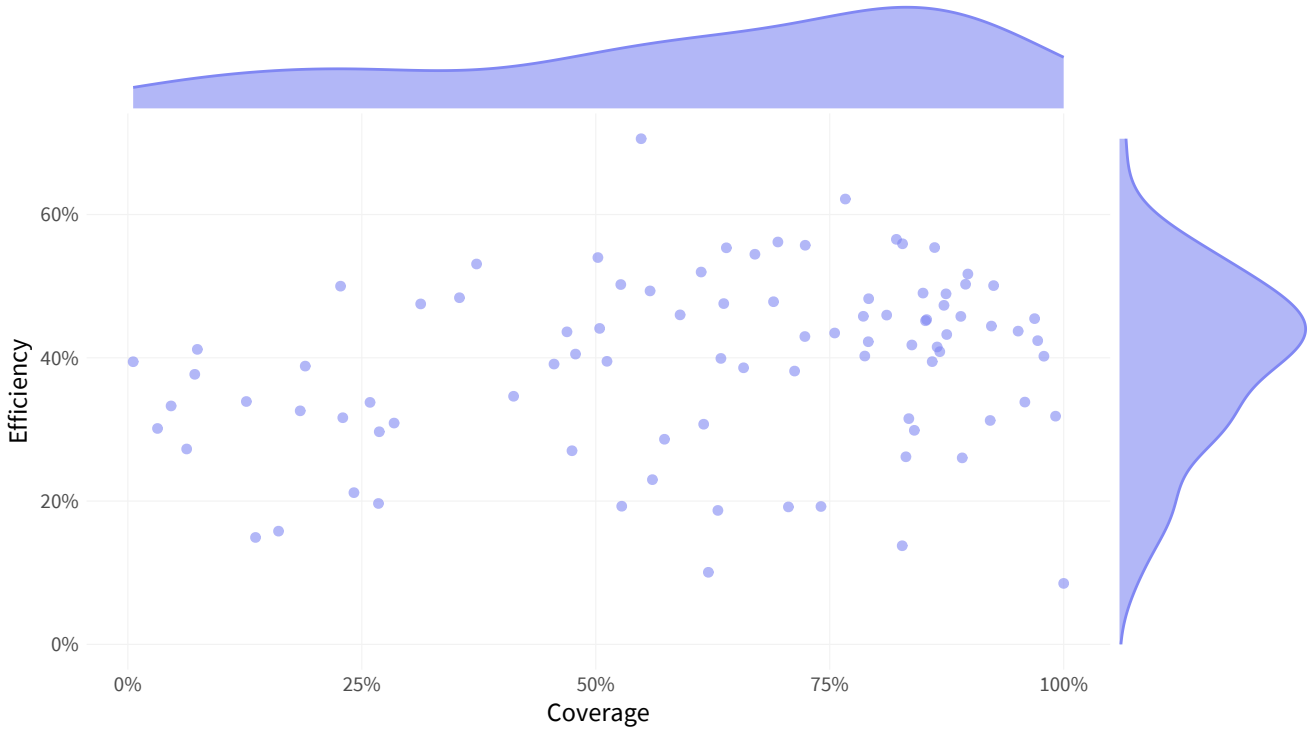
Figure 1 presents coverage and efficiency statistics for the organizations in our sample. It's easy to see why we say "balance two simultaneous goals" because not many do both well. The variation among them reflects different vulnerability remediation strategies, practices, and capabilities.

On the coverage side of Figure 1, we see firms at every point along the scale from 0% to 100%. Those at the extremes represent environments with very few assets under management. The peak density at 82% coverage is promising; most organizations are successfully addressing the large majority of their most risky vulnerabilities.

Things look a bit different over on the efficiency side. Not many organizations cross the 50% line, indicating greater emphasis on coverage than efficiency. That strategy makes sense because the cost of not fixing a vulnerability that winds up being exploited is generally higher than fixing it just in case. But there's another reason as well—patching is inherently inefficient from the perspective of the efficiency algorithm.

Among the hundreds of thousands of patches we studied, about half fix multiple CVEs. Let's say the patch you deploy fixes five CVEs and only one of those is exploited. According to the raw efficiency calculation, you chose "wrong" four out of five times. Your efficiency metric reflects that penalty even though you really didn't explicitly choose to prioritize the other four. The calculations behind Figure 1 attempt to account for this, but keep that in mind as you consider the results.

FIGURE 1: REMEDIATION COVERAGE AND EFFICIENCY METRICS ACROSS FIRMS



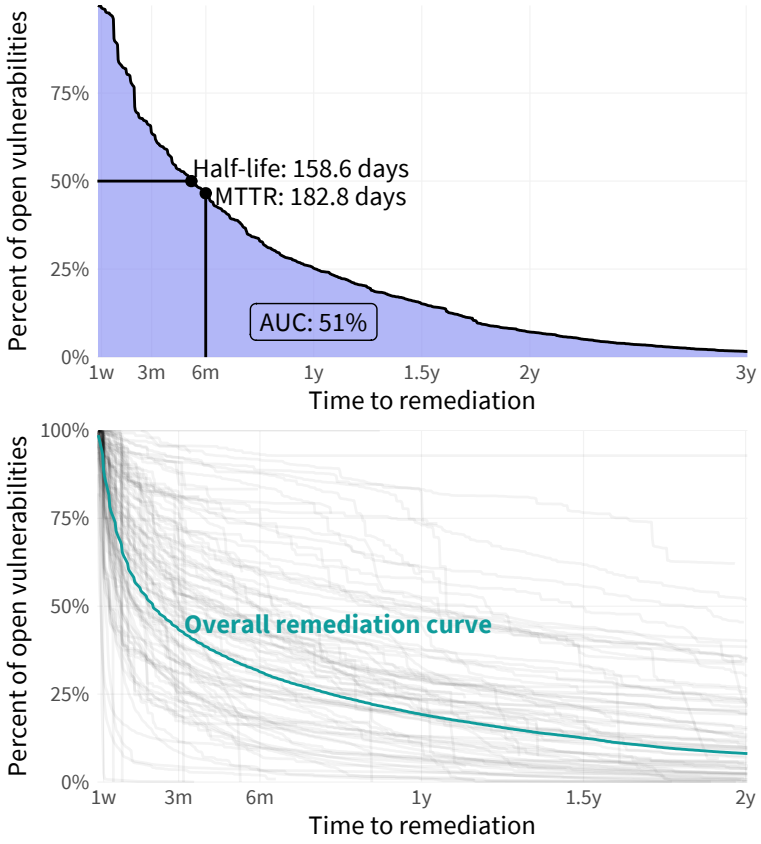
Remediation Velocity

Regardless of where firms end up with respect to coverage and efficiency, they can take dramatically different routes to get there. The second volume in the P2P series introduced the survival analysis technique, which is a set of methods to understand the time duration to an event. For our purposes, the event of interest is remediating a vulnerability, and it's a very useful way to study how long this remediation process takes.

Assume an organization observes 100 open vulnerabilities today (day zero) and manages to fix 10 of them on the same day, leaving 90 to live another day. The survival rate on day zero would be 90% with a 10% remediation rate. As time passes and vulnerabilities continue to be fixed, that proportion will continue to change. Tracking this change across all of the vulnerabilities across all of the firms in our study over time produces a curve like the one shown at the top of Figure 2. We see that the overall half-life of a vulnerability is 159 days. Beyond that, there's clearly a long tail challenge for remediation programs that results in many vulnerabilities remaining open beyond one year.

The bottom of Figure 2 traces the same aggregate vulnerability survival curve (shown in teal), but also includes the (gray) curves for each of the other individual firms. This makes an important point that remediation timeframes vary substantially across firms. This is why we refer to ***velocity***; there's both a directional and a speed aspect to those lines. If you're wondering how things differ across industries, sizes, software, etc., you'll find all that and more in volume three of the P2P series.

FIGURE 2: SURVIVAL ANALYSIS CURVES FOR VULNERABILITY REMEDIATION TIMELINES.



Though it's easy to see a visual difference among the organizations depicted in the top half of Figure 2, it's not easy to quantitatively compare or benchmark them. For that, we can use several metrics from survival analysis, each of which gives a slightly different measure of remediation timelines.

- **Area Under the survival Curve (AUC)**: Measures the area under the survival curve representing "live" (open) vulnerabilities. A lower AUC means higher velocity.
- **Mean Time To Remediation (MTTR)**: The average amount of time it takes to close vulnerabilities.
- **Vulnerability half-life**: The time required to close exactly 50% of open vulnerabilities.

As you review Figure 2, keep in mind that it's not necessarily bad when organizations don't close all vulnerabilities quickly. It would be wasteful, in fact, to close all of them as fast as possible. The more important question is how quickly firms fix the vulnerabilities that really matter. Because of this, firms with higher remediation velocity tend to have lower efficiency, indicating a tradeoff similar to that between coverage and efficiency.

The variation among firms in Figures 1 and 2 reflects different vulnerability remediation strategies, practices, and capabilities.

Remediation Capacity

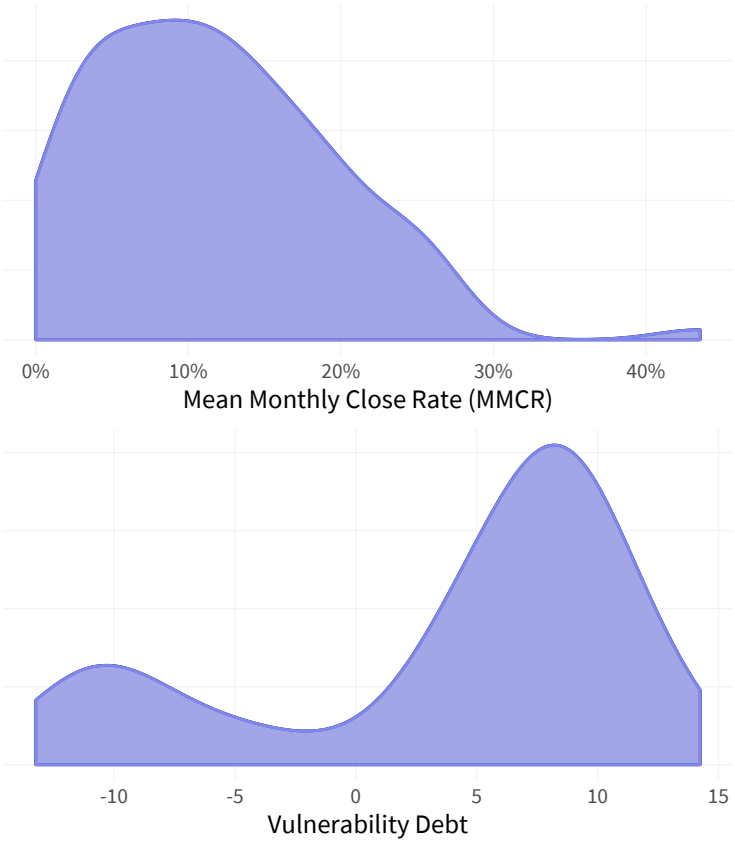
Baseball fans know that the typical closing pitcher throws hard, but only for a couple innings. Starters may not have equivalent velocity but do have a higher capacity to go the distance. Vulnerability management programs have similar distinctions. Some fix fast within a limited scope, while others play the long game.

Remediation capacity in VM is a similar concept with two primary metrics. We quickly define for each and then examine them further in the context of organizational data.

- **Mean Monthly Close Rate (MMCR)**: Measures the proportion of all open vulnerabilities a firm can close within a given timeframe.
- **Vulnerability debt**: Measures the net surplus or deficit of open high-risk vulnerabilities in the environment over time.

Think of MMCR as raw remediation capacity. To derive it, we calculate a ratio for the average number of open and closed vulnerabilities per month for each organization in our sample. Readers of P2P Vol. 3 may remember the astonishing statistic that, on average, organizations remediate about one out of every 10 vulnerabilities in their environment within a given month.

FIGURE 3: DISTRIBUTION OF CAPACITY METRICS ACROSS FIRMS

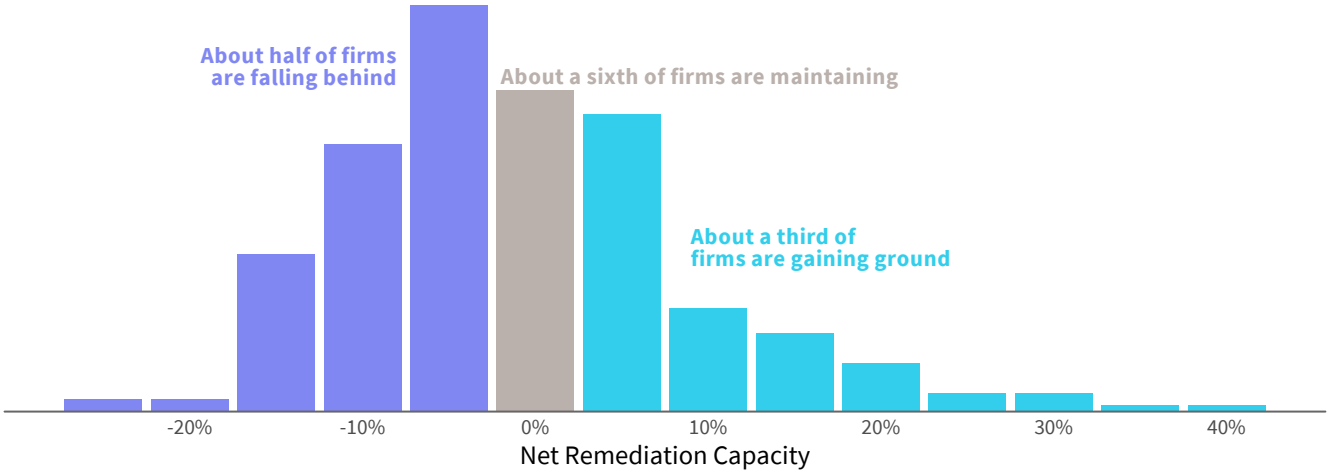


There's variation around this 1-in-10 ratio, of course, and the top chart in Figure 3 showing the distribution of MMCR across organizations demonstrates that fact. Quite a few firms fall below that mark and some exceed it by a large margin. It's tempting to assume those exhibiting higher remediation capacity must have less infrastructure to manage, but the data doesn't support that conclusion. Average capacity remains remarkably consistent, regardless of characteristics such as organization size, number of assets, and total vulnerabilities.

This is all rather depressing for VM programs, but we do see signs of hope in the data. If organizations can't remediate all (or even most) of the vulnerabilities in their environment, can they at least stay ahead of those that represent the highest risk? Enter the concept of vulnerability debt depicted in the bottom half of Figure 3.

Actually, it may be helpful to take a step back before further explanation of vulnerability debt. Figure 4 is a reprint from P2P Vol. 3. To produce it, we calculated the total number of open high-risk vulnerabilities and the number remediated per month for each organization. The resulting ratio identifies which firms are keeping up (closing about as many vulnerabilities as were opened), falling behind (opened > closed), or gaining ground (closed > opened). From that, we see about half of firms falling behind, one in six keeping up, and one in three actually gaining some positive ground. Figure 4 also makes it plain that the degree by which organizations are falling behind or pulling ahead varies widely.

FIGURE 4: COMPARISON OF NET REMEDIATION CAPACITY AMONG FIRMS FOR HIGH-RISK VULNERABILITIES (COPIED FROM P2P VOLUME 3)



We turned this into a metric by tracking the number of open vulnerabilities an organization has over time. We found that most organizations are either building a backlog of open vulnerabilities or slowly chipping away at them. We estimated this change over time by extracting the slope of a linear regression for each organization. Of course, some organizations have many more vulnerabilities than others, so we scaled the slope in order to compare apples to apples. We call this metric vulnerability debt.

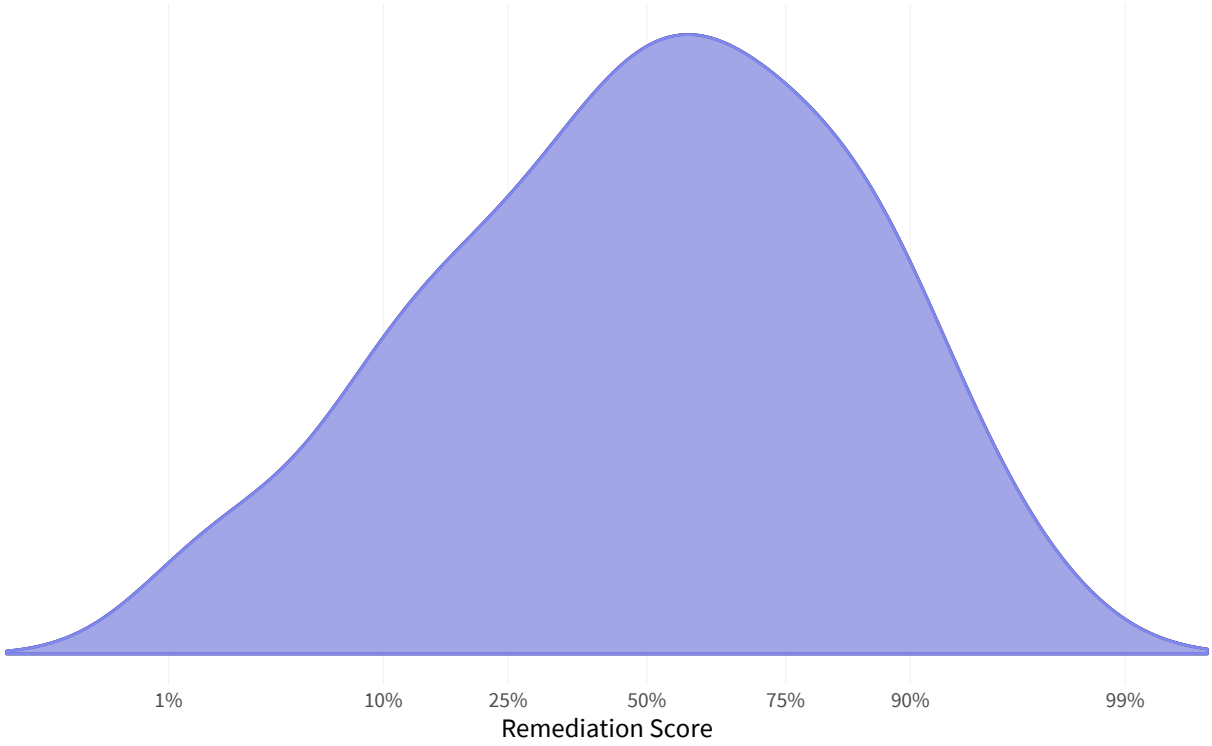
Together, these metrics reveal that remediation capacity, though remarkably consistent across diverse organizational demographics, is not set in stone. What enables some organizations to stay afloat and even pay down their vulnerability debt, while others fall deeper and deeper in the hole? Be patient; we'll explore that soon. But first, let's pull all remediation metrics into a single performance measure.

Overall Remediation Performance

No one metric to rule them all exists within the realms of Middle-Earth, so we'll need to forge one ourselves. Thankfully, a journey into the fires of Mordor isn't required for this; we'll just use the power of math. Our goal is to combine the metrics presented thus far into one credit score-like measure of overall vulnerability remediation performance. From this score, we can compare how any organization compares to others.

To do this we need a mathematical method that can combine all the correlated bits together into a single value. Thankfully, there is a century-old method that does just that: Principal Components Analysis (PCA).[1] All we need is a little prep, a few order-preserving (logarithmic) transformations on high variance variables, an inversion on some metrics to make sure "higher is better," and then we can feed the values into the fires of PCA. We pick out the first component, which in our case accounts for a majority of the correlation in our metrics, and voila—we have one measure to rule them all. We can apply this method to capacity and velocity metrics first to get scores in those categories, and then combine everything to get a single value holistic view of remediation.

FIGURE 5: DISTRIBUTION OF OVERALL REMEDIATION PERFORMANCE SCORES ACROSS FIRMS



The distribution of remediation performance scores in Figure 5 is fairly normal in shape. Measuring vulnerability remediation performance—through particular metrics or an overall score—is a useful exercise for VM programs. Once an organization understands where it is, plotting a course to improve performance requires a deeper understanding of the organizational characteristics and practices that influence performance. We identify these factors in the next section.

HOW DOES THIS DIFFER FROM THE KENNA RISK METER SCORE?

If you're a Kenna Security customer, you may be asking this question. We wanted to address that here to avoid potential confusion. The Risk Meter scores risk across assets in your environment. The remediation performance score, in a nutshell, scores efforts to address that risk by remediating vulnerabilities affecting those assets.

1 *Thanks Lord Sauron… I mean Karl Pearson.*

Survey Methodology

Kenna and the Cyentia Institute collaboratively developed the questions used for this survey and Cyentia implemented it in a web-based survey platform. Inviting participants was accomplished by generating numerous anonymous invitation URLs, which Kenna then distributed to their customers for which we had remediation performance data from the platform (about 300 organizations).

We received 171 responses in total, many of which represented multiple people from the same organization. In such cases, we took the average or consensus answer across those responses for our analysis. Merging answers for multi-respondent organizations left us with completed questionnaires for 103 organizations. This amounts to a response rate of approximately 35% of our target sample frame.

Analyzing Factors of Performance

Up to this point in the P2P series, we've examined remediation data exported from the Kenna platform. This has afforded a unique data-centric perspective on VM performance, but we still lack a view of the various internal factors that influence those outcomes. To explore such factors, we designed and distributed a survey to Kenna customers.

There are two goals for this section: (1) to summarize what we learned from organizations via the survey and then (2) to compare those findings with what we observed in the remediation data.[2] The purpose is to identify factors that correlate, both positively and negatively, with changes in our metrics of remediation performance.

Before we get started, some explanation is in order. We've taken a rigorous and cautious approach to the analysis underlying the findings we present in this section. We won't go into the gritty details but comparing proportions among groups—which is a lot of what we do here—can be very deceiving. That's especially true when variation is as high as it is in this sample. Without proper techniques, seemingly substantial (but statistically insignificant) differences will inevitably prompt all sorts of unsupported conclusions. We bring that up to let you know that we've applied those proper techniques, and anything we include is a statistically significant finding (unless we explicitly state otherwise). We'll also add the proverbial reminder that correlation doesn't imply causation. We focus on interesting correlations and offer insights on what might be behind them, but we can't confirm any causal relationships with certainty. But correlation can certainly support thoughtful consideration about how these findings can help improve VM program performance, and that's our aim in this endeavor.

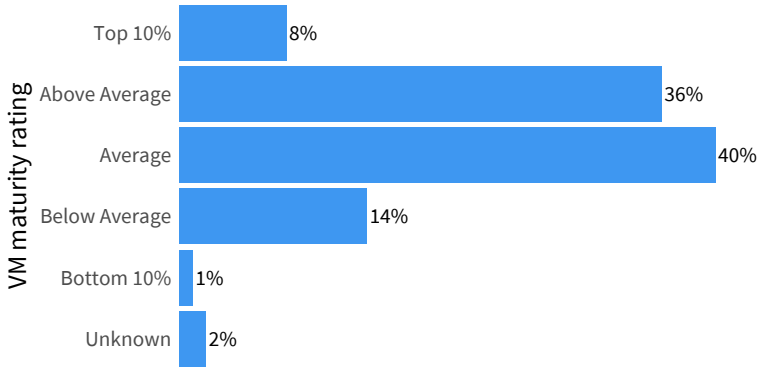
2 *This may appear to contradict our statements above about preserving the anonymity of participants. Unique IDs provided by Kenna allowed Cyentia to match survey responses with remediation data without identifying the organizations represented.*

Maturity and Performance

Let's start with brass tacks. As explained above, we have a remediation performance score for each organization. In the survey, we asked for a self assessment of the overall maturity level of the VM program. Ostensibly, higher maturity should tie to better performance. Does it?

In answering that, we'll first need to know how respondents rated the maturity of their programs. Figure 6 shows the middle "Average" option captured the most votes, but only barely. A near equal proportion felt they were above average, meaning roughly three-quarters of programs fell in the "on par or a bit better" range.

FIGURE 6: SELF-ASSESSMENT OF THE OVERALL MATURITY OF THE VM PROGRAM



It's rather interesting that nearly 10% assessed themselves in the top 10%. That's refreshingly honest compared to many surveys where 80% place themselves in the top 10%. But the real question is whether this maturity self-assessment has any bearing on actual performance.

DOES IT MATTER?

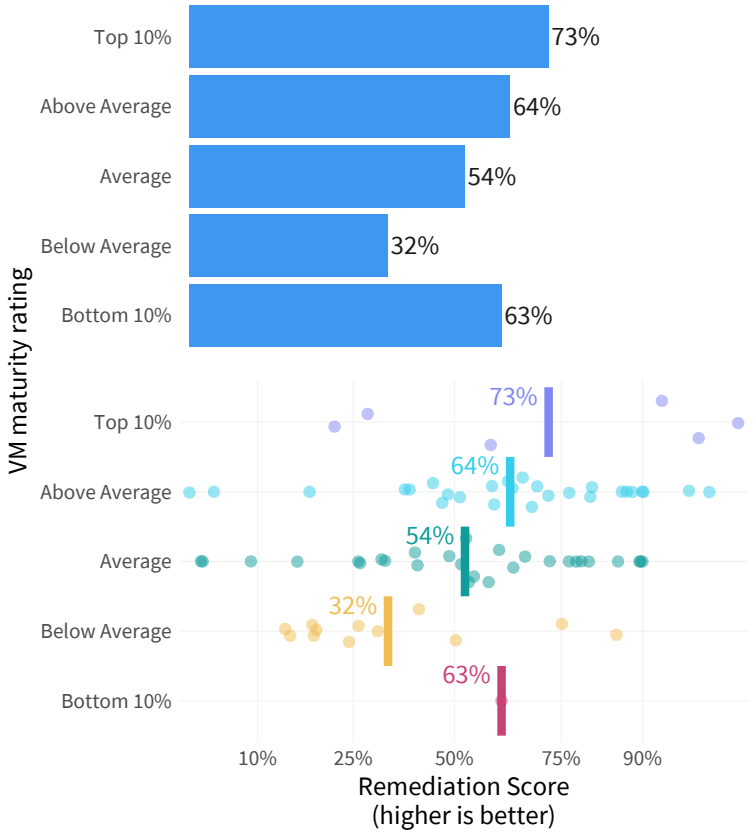
From the top chart in Figure 7, the answer to that question appears to be a fairly solid "yes". Organizations rating their programs toward the upper end of the maturity scale achieve a higher remediation performance score than those that place themselves toward the bottom. But the answer is less clear in an alternate view of the same data given in the bottom chart of Figure 7. Though the dots representing individual firms show a lot of overlap among maturity ratings and performance scores, some lower-maturity organizations perform better than those at the top. So which view is right? Well, both, and the reference lines in Figure 7 corresponding to the group means hopefully reinforce that point.

Correlation isn't Causation, but...

...it's something to consider and investigate. We've taken a rigorous and cautious approach to the analysis underlying the findings we present in this section. Comparing proportions among groups—which is a lot of what we do here—can be very deceiving. Know that anything we include in this section is a statistically significant finding, unless we explicitly state otherwise.

> "Immaturity in managing assets and the speed of vulnerabilities coming out are key challenges for us."
>
> FinTech respondent

FIGURE 7: TEST COMPARING VM MATURITY RATING WITH REMEDIATION PERFORMANCE SCORE.



You will see many of the "swarm plots" from the bottom of Figure 7 in the pages to follow. We chose them because they're the most candid view of the data. Each dot represents a particular respondent. If you find the dots distracting, just focus on the vertical lines/labels that mark the means for each group. Those fall where the bars would end in a standard bar chart, which you can see in Figure 7 by visually tracing the corresponding bars and lines.

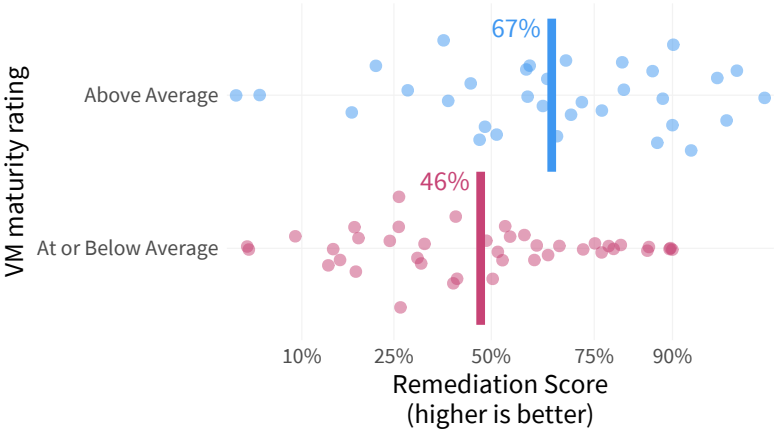
Chart choices aside, what do we make of the performance overlap among maturity tiers? That outcome partially stems from the survey instrument. Is it reasonable to expect organizations to know where they stand in relation to others? It also comes down to statistics; the score differences among maturity ratings in Figure 7 just aren't as significant as they seem. Another part of the issue relates to the metrics themselves. While our combined score is supposed to represent overall performance, it should be recognized that organizations may intentionally pursue strategies that focus on subcomponents of that score (e.g., maximize coverage) rather than balance everything equally. So a mediocre score doesn't necessarily mean the firm isn't meeting its goals. It could even be posited that the highest maturity programs have the confidence and capability to let one metric slip in order to maximize another, thereby acceptably lowering their overall performance score.

This outcome foreshadows what you'll see a lot in this section. Things don't always fall perfectly into place or follow predictable patterns. Welcome to vulnerability management, amirite?

> "Vulnerability management is typically seen as technical debt unless it's a critical finding on an assets that is publicly facing."
>
> IT Services respondent

Rather than taking the easy way out by flashing a bunch of simple charts that tell a misleading story, we adjusted our analytical approach to better coax the subtle and stubborn (yet statistically valid) findings from the data. As an example of that approach, let's see what we can glean from the simplified view of maturity and performance found in Figure 8.

FIGURE 8: SIMPLIFIED TEST COMPARING VM MATURITY RATING WITH REMEDIATION PERFORMANCE SCORE.



Dots representing individual firms show a lot of overlap among maturity ratings and performance scores.

Figure 8 splits organizations into two groups—those with above-average maturity (combines "Above Average" and "Top 10%" answers) and those rating themselves at or below average. Results are now a bit cleaner and clearer. **On average**, firms with higher maturity ratings have higher performance scores, even though individual firms don't all follow that pattern. What's more, maturity ratings led all other factors in our survey in terms of the number of individual correlating metrics.[3] All that suggests an important link between confidence—if not actual maturity—and performance.

We emphasize "On average" here because that's the proper way to think about results in this section. We will try to consistently remind you of that with phrases like "Overall, " and "typically" but caveating every statement gets redundant. So assume it's universally implied unless we state otherwise.

To be completely honest, we didn't expect to find strong correlation between the subjective maturity ratings from respondents and the objective remediation data from the Kenna platform. We should have listened more closely to Metallica: "*Forever trust in who you are, and nothing else matters.*"

But is that actually true? Does nothing else really matter? Thankfully, Metallica didn't get that last part right. We found lots of things that matter for different aspects of vulnerability remediation performance and explore them with you in the pages that follow!

3 The Conclusion sections contains a chart showing correlations between survey factors and performance metrics. The complete list of metrics associated with maturity ratings can be found there.

“I think vuln management is a struggle for any org that isn’t in a highly regulated environment with teeth to enforce schedules.”

IT Services respondent

“Vulnerabilities not on endpoint environments are a nightmare.”

Mass Media respondent

Assets Under Management

The first two questions in the survey sought a baseline understanding of the infrastructure under management. Namely, what proportion of the organization’s assets were a) reflected in Kenna’s platform and b) hosted in the cloud or other off-premise environment. The first is useful because it sets the scope of infrastructure for which we have remediation data. The second is interesting to us as a potential influencing factor in remediation performance.

FIGURE 9: PROPORTION OF ASSETS MANAGED WITH KENNA

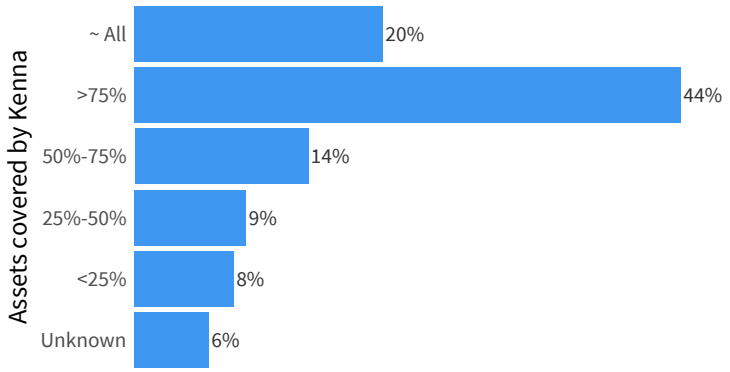
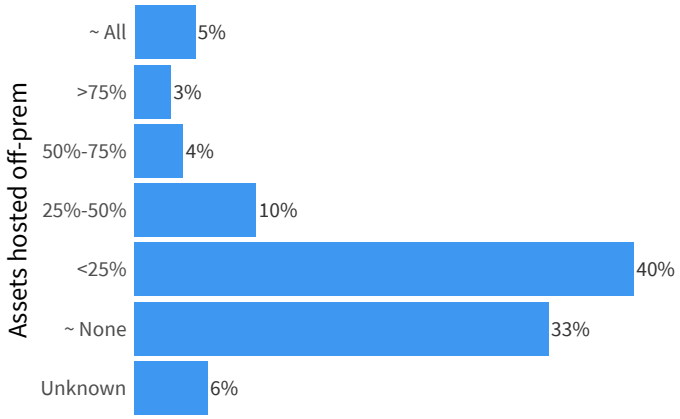


FIGURE 10: PROPORTION OF ASSETS HOSTED OFF-PREMISES

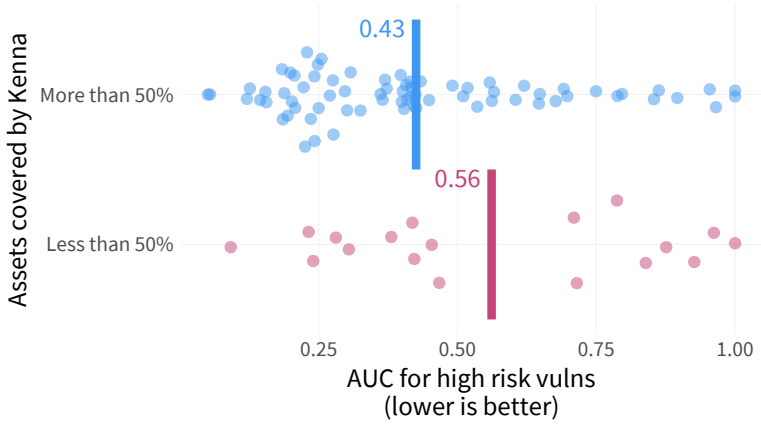


Responses to the two infrastructure questions are found in Figures 9 and 10. Let us explain…no, there is too much…let us sum up: most respondents (who are all Kenna customers) manage the majority of their assets in the Kenna platform, and a minority of those assets are hosted in the cloud. Now on to what inquiring minds want to know—does this affect performance?

DOES IT MATTER?

Tests reveal that, on average, organizations managing more than half of their assets in Kenna's platform exhibited a significantly better remediation velocity for high-risk vulnerabilities (13% reduction in AUC). That may seem suspiciously self-serving, but keep in mind this result comes from the Cyentia side of the partnership rather than from Kenna.

FIGURE 11: TEST COMPARING SCOPE OF ASSETS MANAGED IN KENNA'S PLATFORM WITH REMEDIATION VELOCITY



Appearance of self-service aside, this actually makes sense. One of Kenna's main value propositions is helping organizations focus remediation efforts on the vulnerabilities that matter most. Extending that guidance to a larger scope of assets would logically result in addressing critical issues more quickly. It would be interesting to see whether remediation velocity improved even more between Kenna and non-Kenna customers, but we don't have the data to support that comparison.

Surprisingly, the proportion of assets hosted in the cloud did not correlate with any of the performance variables we examined. Well, that's not entirely true; there was one minor measure under the velocity category that showed a weak positive correlation with cloud adoption, but nothing that should influence moving to/from the cloud for the sake of remediation performance. And that's the main point we wanted to make in bringing up this non-finding for the cloud. There are enough factors to consider already, and the cloud doesn't appear to be one of them.

> "Organizing our assets with Kenna Risk Meters helps us a lot. This gives us an idea of where to start when we fix vulnerabilities and saves time from searching CVEs that are common among our assets."
>
> Banking respondent

Surprisingly, the proportion of assets hosted in the cloud did not correlate with any of the performance variables we examined.

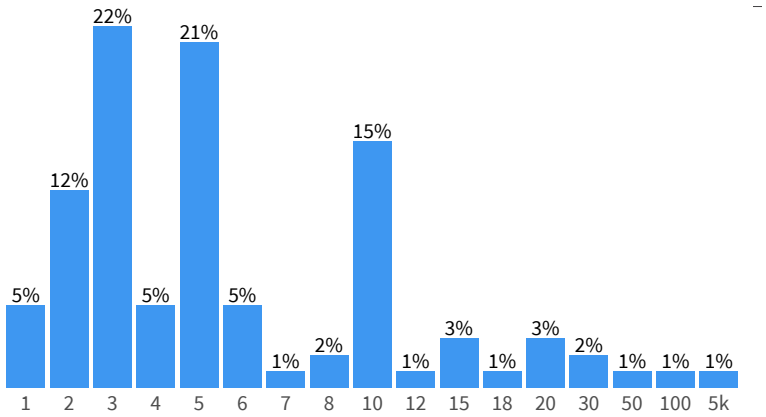
> We have weekly standup meetings to address outstanding vulnerabilities with security, compliance, and asset owners."
>
> Healthcare respondent

> The lack of resources on our team is the key factor and influences in the speed and efficacy for remediation"
>
> Financial Services respondent

Organizational Factors

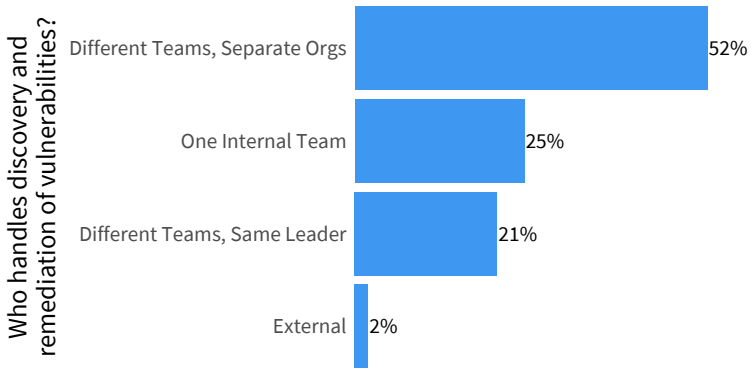
We then asked a few questions about the VM program, starting with how many people participate in the remediation process. A glance at Figure 12 shows it's a handful or less for the majority of firms, with at least one outlier respondent who apparently takes the stance that "VM is everyone's job."

FIGURE 12: PARTICIPANTS IN REMEDIATION



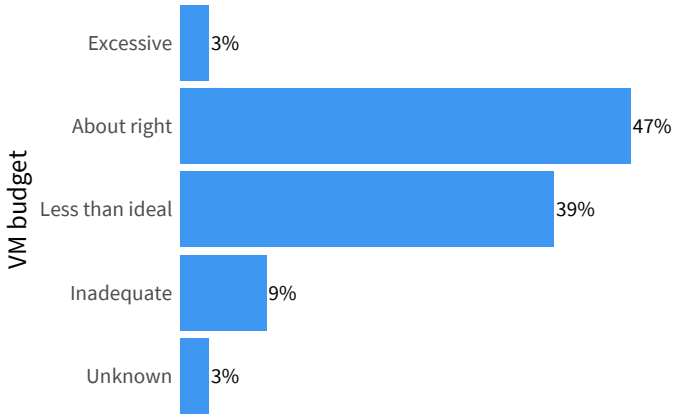
In terms of VM team structure, the main question we wanted to answer was whether those who identify vulnerabilities were also responsible for remediating them. According to Figure 13, finders and fixers are typically either in completely separate organizations or separate teams reporting up through a common chain of command. Although in the minority, the number of "same team or person" responses is actually more than expected. It's possible that respondents misinterpreted the distinctions we made in the question.

FIGURE 13: DISCOVERY AND REMEDIATION DUTIES



Another data point we gathered focused on the adequacy of budgets backing vulnerability management programs. It's not hard to imagine how this could make or break remediation performance. Most felt budgets were about right or a little less than ideal (see Figure 14), which is almost certainly by design. Both frivolity and famine come with unexpected costs and complications.

FIGURE 14: VULNERABILITY MANAGEMENT BUDGET



DOES IT MATTER?

People obviously matter a great deal in VM. But from a strictly statistical perspective, the number of people participating in the remediation process did not have much bearing on performance metrics. We suspect several reasons for that, but the main one is that team size is intercorrelated with organization size and other such factors. We'd need to compare firms with abnormally large or small VM teams relative to overall organization size while controlling for other variables, and we just don't have that data at present. Our apologies to all those who were hoping to use this to justify a hiring spree.

But this next one may help loosen the purse strings a bit. We found some evidence that budgets help things get done faster. Per Figure 15, organizations reporting adequate (or better) VM budgets cleared their environment of vulnerabilities significantly faster (.13 off overall AUC) than their budget-strapped peers. Perhaps money ***can*** buy time after all.

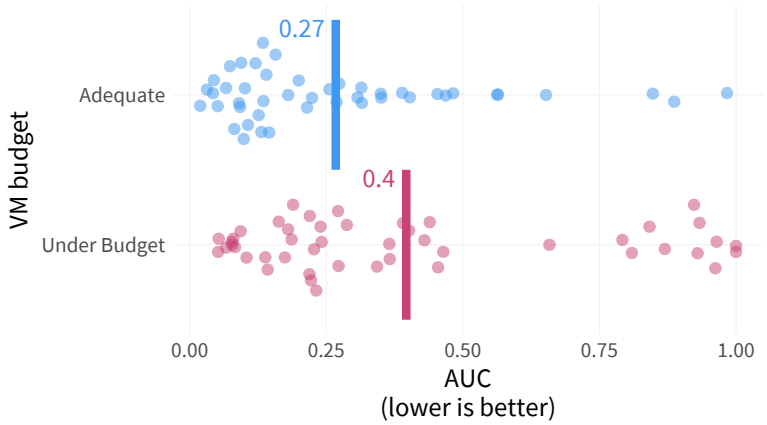
> We found some evidence that budgets help things get done faster (see Figure 15). Perhaps money can buy time after all.

> "Conflicting business requirements are a major issue. The same team members are tasked with deployment of new software code for new products and functionality as well as vulnerability remediation."
>
> Healthcare respondent

> "We need people available for remediation, funding to upgrade older systems, and down time for remediation"
>
> Manufacturing respondent

> The level of engagement of teams outside security matters a lot."
>
> Healthcare respondent

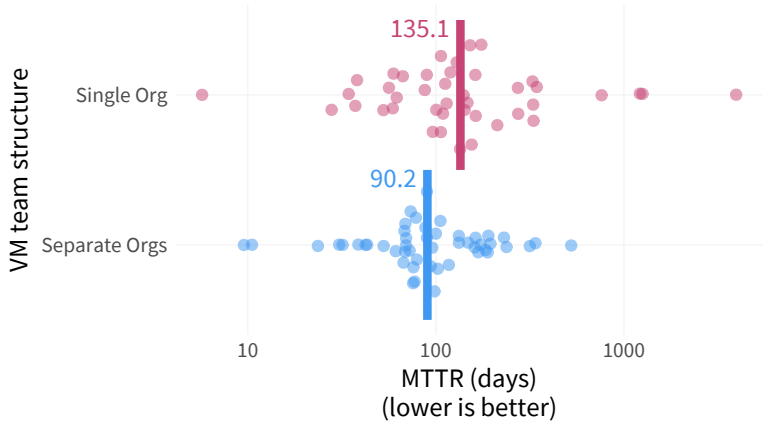
FIGURE 15: TEST COMPARING VM PROGRAM BUDGET WITH REMEDIATION VELOCITY



Regardless of budget or team size, the structure of VM staff appears to make a difference. It's one of the few factors we tested that correlates with the overall remediation performance score. In terms of specific metrics, analysis indicates that MTTR is about a month and a half shorter among firms that house find and fix responsibilities in separate organizations (Figure 16). And we found similar time savings when reducing the scope to just high-risk vulnerabilities. Separating duties also corresponds to higher remediation capacity (Figure 17), meaning firms are less likely to be in that "falling behind" zone from Figure 4.

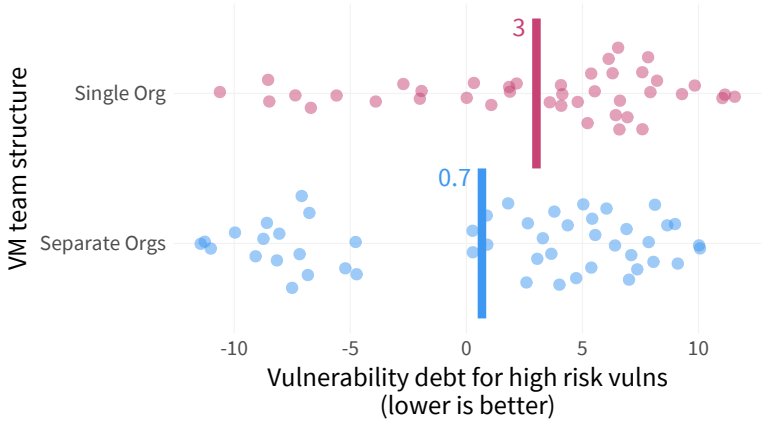
> Upper management support, better team synergy, and correct prioritization are key success factors."
>
> Financial Services respondent

FIGURE 16: TEST COMPARING VM TEAM STRUCTURE WITH REMEDIATION VELOCITY



The reasons such performance improvements are tied to organizational structure isn't exactly clear. But we suspect having separate teams handling identification and remediation of vulnerabilities is a sign of more resources and maturity. It may also indicate that security has its own culture and mission rather than just being one of the many duties falling under broader IT operations.

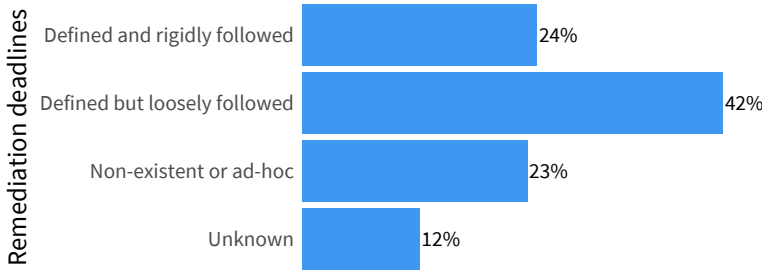
FIGURE 17: TEST COMPARING VM TEAM STRUCTURE WITH REMEDIATION CAPACITY



Remediation SLAs

Having a better understanding of organizational factors from respondents, we next wanted to learn about the policies governing their VM programs. Specifically, we were interested in any service-level agreements (SLAs) or deadlines established for vulnerability remediation timeframes. As seen in Figure 18, about two-thirds of organizations have defined deadlines and follow them in some capacity. The remainder either handle things on a case-by-case basis or aren't sure if SLAs exist (which would seem to suggest they do not).

FIGURE 18: ARE REMEDIATION DEADLINES DEFINED?



We then asked respondents with defined deadlines how many days their firms allow for fixing vulnerabilities of varying priority or risk levels. Not surprisingly, remediation windows vary both within and across priority levels. The peaks of the distributions in Figure 19 point to a week for the highest priority vulnerabilities, a month for high priorities, and three months for moderates.

> "Transparency and communication seem to have the biggest impact."
>
> Manufacturing respondent

> "External service providers adhering to SLAs is a key factor in our success or failure."
>
> Financial Services respondent

“

App owners failing to provide maintenance windows hinders our success.”

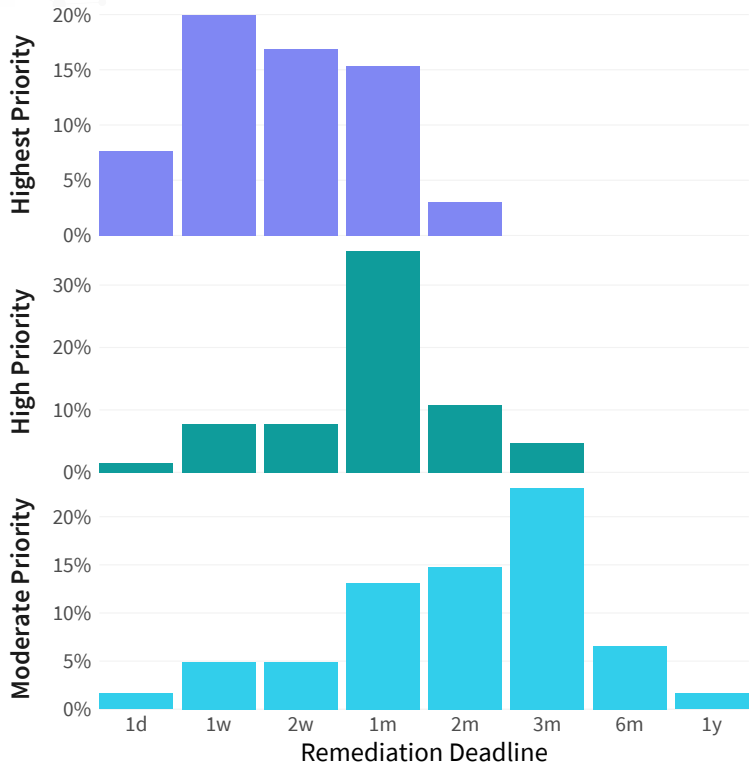
Education respondent

“

As with most things the urgent outweighs the important.”

Banking respondent

FIGURE 19: REMEDIATION DEADLINES BY PRIORITY LEVEL



DOES IT MATTER?

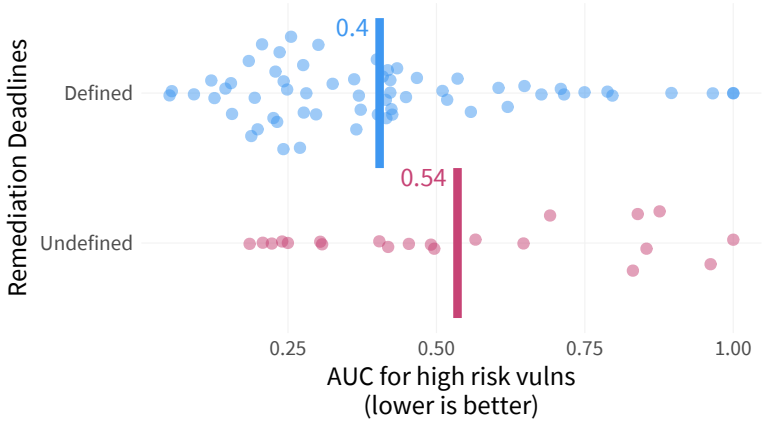
Every single one of us has crammed for a test or burned the midnight oil to get a project across the finish line, so we know the…*ahem*… motivating power of deadlines. Does that power extend to vulnerability remediation too? It appears so. Our analysis suggests that firms with defined SLAs address vulnerabilities more quickly than those with non-existent or ad-hoc schedules. In terms of metrics, that’s observed as a 15% improvement in AUC for high-risk vulns and 10% better AUC across all vulns (Figure 20). Sure, deadlines are missed from time to time, but not having them allows tardiness to become a habit.

In addition to boosting velocity, it seems that priority-driven SLAs correlate with expanded overall remediation capacity as well. Adding a few percentage points to the proportion of vulnerabilities you can close in a given month may not seem like much, but a point here and a point there may be the difference between running a surplus vs. adding to the deficit.

We also compared remediation velocity among organizations with shorter and longer SLAs within each priority level but found no major differences. In other words, firms on the left side of the distribution in Figure 19 didn’t fix vulnerabilities significantly more quickly than those towards the right.

That's a bit counterintuitive to be sure. Our takeaway is that establishing SLAs is more important to performance than pushing for overly-aggressive deadlines. From a practical perspective, that may simply translate to "set realistic goals."

FIGURE 20: TESTS COMPARING DEFINED AND UNDEFINED SLAS ON REMEDIATION VELOCITY



Prioritization Criteria

So SLAs are helpful. But how are these priority or risk levels determined in the first place? We asked firms about their top three influencers for prioritizing vulnerabilities for remediation and consistently heard a mix of Kenna Risk Scores (obviously), asset value/criticality/sensitivity, and Common Vulnerability Scoring System (CVSS). Beyond those, some organizations mentioned compliance requirements and their own internal analysis as the basis for vulnerability prioritization.

Rather than launch into commentary on prioritization strategies, we will instead refer you to the first volume in the P2P series. It's chock-full of examples, measurements, and comparisons relevant to that topic. What we want to know here is whether the source or the method of prioritization influences performance.

DOES IT MATTER?

Every organization in our sample is a Kenna customer, but those that said Kenna Risk Meter scores were a primary influencer of their remediation efforts reduced the half-life of vulnerabilities by an average of 46 days (Figure 21). They also measure a 10% to 15% reduction in AUC for high-risk vulnerabilities. We've demonstrated several times in the P2P series that it's impossible to fix every flaw, and certainly not with expediency. Guidance on where to expend precious remediation resources and when to double down to get it done ASAP would logically result in fewer distractions and faster response.

> "Prioritization is key for us. Once we have that down, mitigation/remediation is not that difficult as a process for us."
>
> Utilities respondent

> "Knee-jerk reactions to vulns that "make the news" without taking into account the viability of the threat are a big problem."
>
> Retail respondent

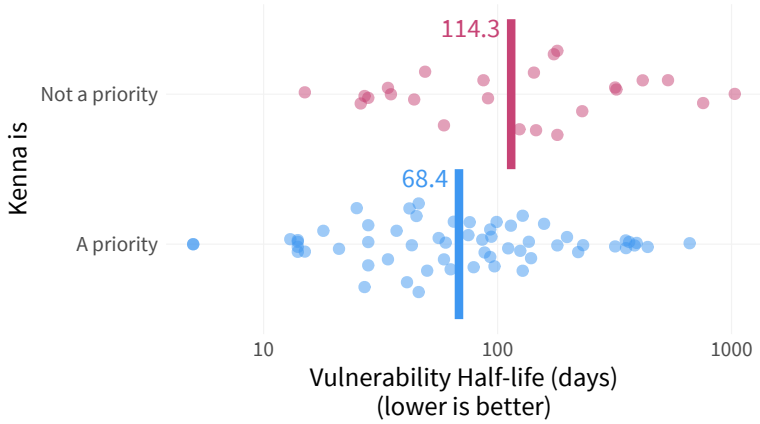
“Kenna is a key technology in influencing executives.”

Banking respondent

“Media coverage of a specific vulnerability sometimes has an overwhelming influence, even when it goes against any advice on security forums.”

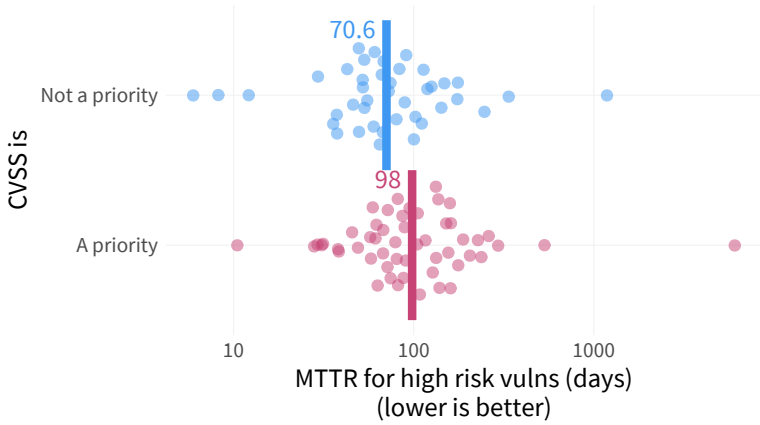
Airline respondent

FIGURE 21: TEST COMPARING KENNA RISK METER PRIORITIZATION WITH REMEDIATION VELOCITY



You know what prioritization source doesn’t seem to improve remediation velocity? CVSS scores. In fact, Figure 22 shows that those who identify them as a primary influencer of which vulnerabilities get priority add just under a month to the time it takes to address high-risk vulnerabilities. If that seems like it can’t be true, we invite you to review our prior reports. Remediation strategies based primarily on CVSS underperform in just about every metric we’ve ever measured.

FIGURE 22: TEST COMPARING CVSS-BASED PRIORITIZATION WITH REMEDIATION VELOCITY

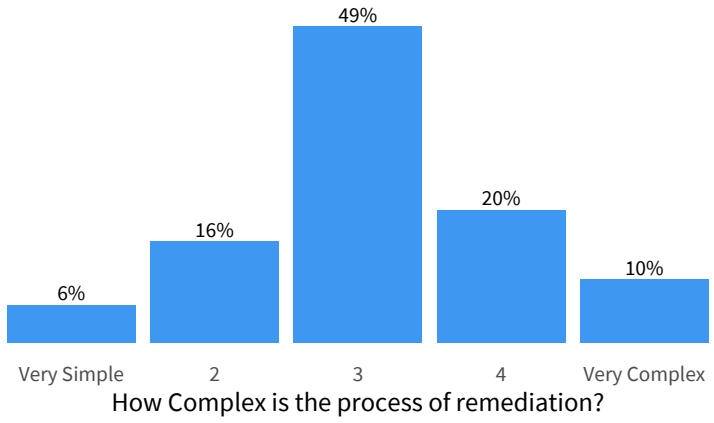


Compliance-driven remediation also misses the mark. The effects weren’t strong, but we did find a correlation between using compliance requirements as a primary driver in prioritizing vulnerabilities and lower coverage rates.

Process Complexity

Once priorities are set and the decision to remediate is made, how complex is the process of actually getting it done? Just over half of respondents pegged it midway on a scale between "very simple" and "very complex," reminding us why we generally don't like offering the neutral option in surveys. It invites the middle bar salute in return, which Figure 23 is clearly sending us.

FIGURE 23: RATING OF REMEDIATION PROCESS COMPLEXITY



Interestingly, the number of people involved in the process (refer back to Figure 12) does not appear to correlate positively or negatively with the complexity. We had both "many hands make light work" and "too many cooks spoil the broth" quotes ready to go and are a tad peeved that neither fits. But we just used both anyway, so there.

DOES IT MATTER?

Process complexity is...well...complex. Our tests point to complexity being a double-edged sword for remediation performance. One edge of that sword looks like a coverage killer. Organizations claiming to have simpler remediation processes averaged nearly 20% higher coverage than those with comparably complex processes. Not every firm can adopt the KISS method for addressing vulnerabilities, but avoiding unnecessary complication appears to be a good general strategy for comprehensive risk remediation.

Our tests point to complexity being a double-edged sword for remediation performance (see Figures 24 & 25).

> "Determining relevance of patches has to be the most costly time factor when remediating vulnerable systems.."
>
> Banking respondent

> "Having detailed processes enables our success."
>
> Pharmaceutical respondent

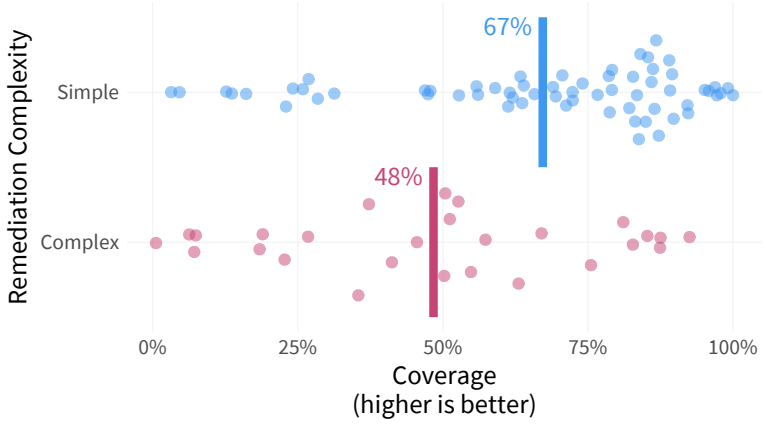
“Application and Middleware vulnerabilities are our key pain areas.”

Pharmaceutical respondent

“Dealing with telecom, IoT, and medical devices is a new level of complexity given those devices are developed for performance in very specific conditions.”

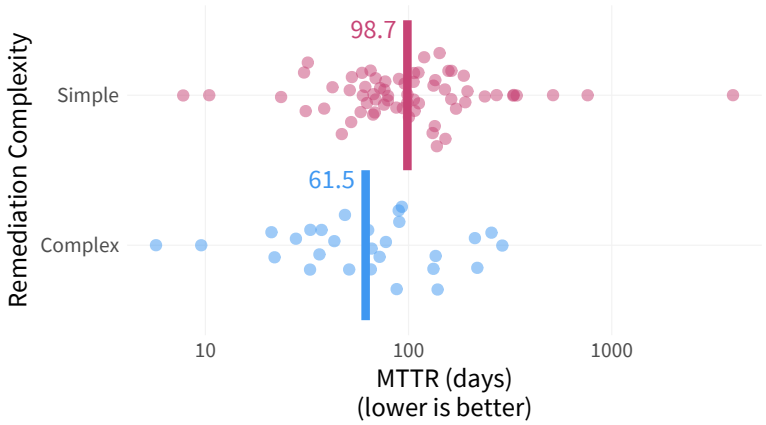
Telecommunications respondent

FIGURE 24: THE EFFECT OF COMPLEXITY ON REMEDIATION COVERAGE



The other edge of the complexity sword, curiously, seems to cut the average time required to close vulnerabilities. According to Figure 25, organizations with simpler processes post an MTTR that’s more than a month longer than those describing remediation as complex.

FIGURE 25: THE EFFECT OF COMPLEXITY ON REMEDIATION VELOCITY

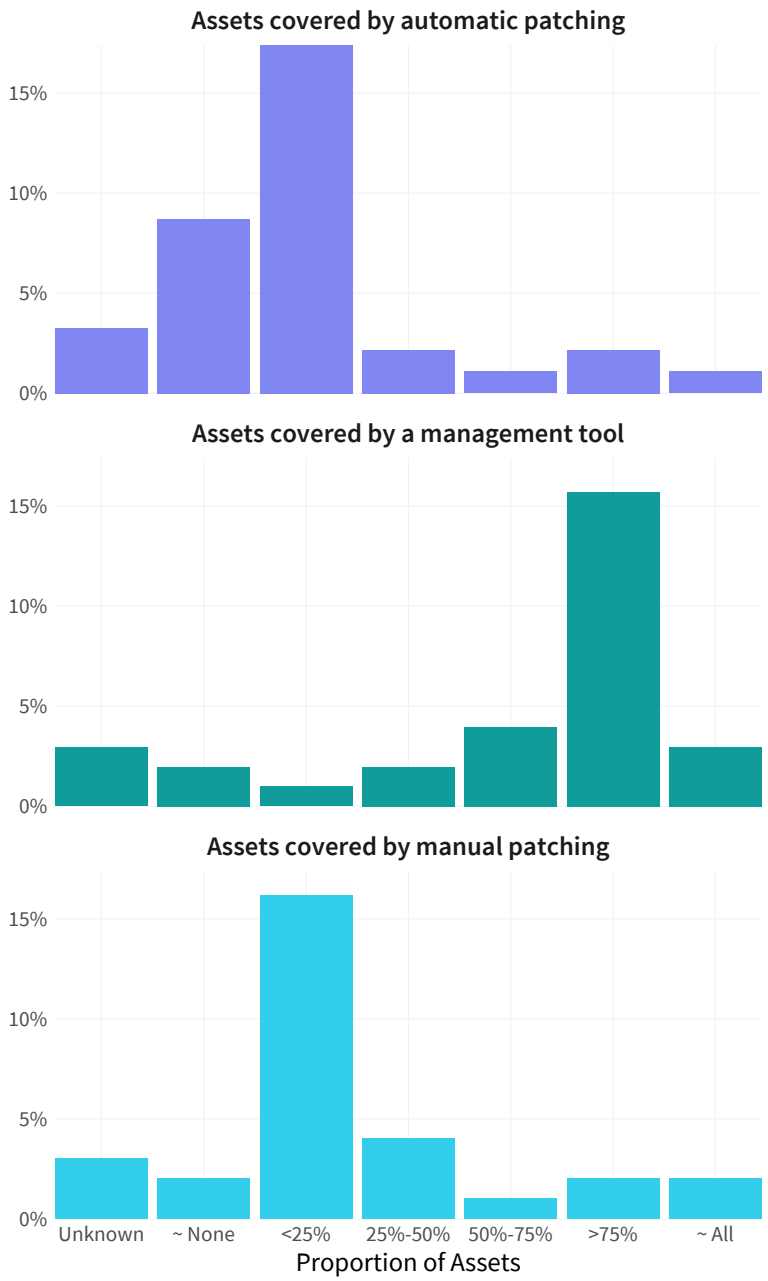


It’s difficult to speculate on what’s happening here without follow-up conversations to better understand what participating firms consider “complex” about their remediation efforts. While we tend to think of process complexity as adding unnecessary barriers that waste precious time, it’s possible for a complex system to actually be more efficient. Consider a situation where the VM program leverages threat intelligence, asset inventories, patch management tools, and automation to streamline remediation. All that increases complexity, but results suggest potential increases to velocity as well.

Deployment Methods

The method of deploying patches is a critical part of the remediation process and ostensibly related to the overall complexity. There's a vast set of tools and services to help with this, but we asked respondents about their use of three broad options for updating systems: 1) automatic updates direct from the vendor, 2) patch management tools (WSUS, SCCM, Jamf, LanGuard, BigFix, etc.), and 3) manual or decentralized deployment. Figure 26 compares the usage of these options across organizations.

FIGURE 26: METHODS FOR DEPLOYING SYSTEM UPDATES



> "Assets owners need to approve changes and down time."
>
> Financial industry respondent

> "We're impacted by management and team resistance to requests with minimal repercussions."
>
> Insurance respondent

> Microsoft Windows and Office patches have the most mature remediation process through SCCM. We struggle getting straggling teams to update their software at times and have old versions of Adobe on some servers. Desktops are much more uniform than the servers."
>
> Legal services respondent

> Our success factors? Better tools implemented and more people."
>
> Business Services respondent

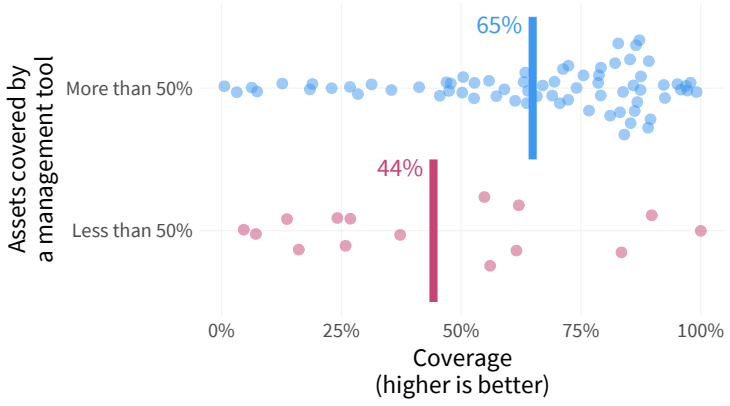
Bear in mind that these methods aren't mutually exclusive; many respondents report using all three in their organizations. It's also worth noting that results in Figure 26 likely indicate relevance to assets under management more than the popularity or utility of a method. For instance, automatic updates from the vendor often aren't an option outside of some desktop operating systems and browsers.

DOES IT MATTER?

Aside from overall VM program maturity, methods used to keep systems updated showed some of the broadest and strongest effects of all factors we tested. To our surprise, however, one that does not seem to help or harm performance is the proportion of systems that receive automatic updates from the vendor.

We say that's surprising because P2P volume 3 found that software like Microsoft Windows and Google Chrome, which auto update frequently, exhibit remediation timelines that lead the pack by a wide margin. Figure 26 above might hint at what's going on here. Most organizations enable automatic updating only on a small minority of systems in their environment. Because of the skewed distribution, we tried altering the groups for statistical tests (e.g., >50% vs. <50% and >25% vs. <25%). But the outcome didn't change much, other than hinting at a possible benefit to remediation capacity. To do this justice, we'd need to make our survey question more granular and probably collect those estimates by asset type. For now, let's move on to the other deployment methods, where we do see some interesting findings.

FIGURE 27: THE EFFECT OF PATCH MANAGEMENT TOOLS ON REMEDIATION COVERAGE



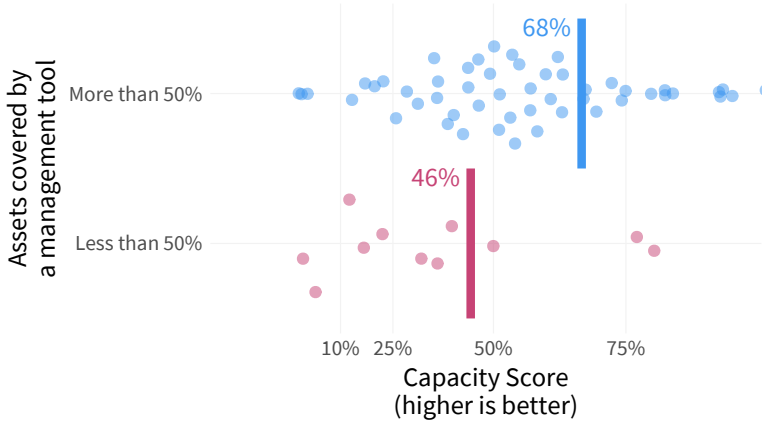
Organizations that make extensive use of centralized patch management tools posted better overall remediation performance and improved multiple component metrics. Firms using these tools to deploy patches

across the majority of assets in their environment successfully closed over 20% more of their high-risk vulnerabilities on average (see Figure 27). This is not unexpected and very much aligns with the reason such tools exist. But it's always nice to see evidence that products do what they claim.

Conversely, a heavy reliance on manual patching dropped average coverage rates by more than 20%. The need to deploy patches in a non-automated, decentralized manner likely indicates challenges such as legacy systems, critical infrastructure, fractured systems, etc. And if that's the case, organizations may have little control over this. But given the choice, these results give good support for investing in more efficient methods of patch deployment as a means of better addressing risk.

In addition to improving coverage, broad use of patch management tools may aid efficiency as well. Firms making heavy use of them show 10% higher accuracy in targeting high-risk vulnerabilities, but the statistical significance of that effect is lower than that of other metrics. These tools grant fine-grained control over which fixes get deployed where and when. By that logic, one might assume deploying patches manually might benefit remediation precision, but that is not the case. Higher rates of manual patching correlate with lower rates of efficiency.

FIGURE 28: THE EFFECT OF PATCH MANAGEMENT TOOL ON REMEDIATION CAPACITY



As if you need another reason to love and leverage patch management tools, we also found evidence suggesting they increase remediation capacity. Firms using them extensively closed more vulnerabilities (overall and high-risk) per month than lighter users. And that makes patch tools the Triple Crown winner (coverage, efficiency, capacity) in the remediation race among deployment methods.

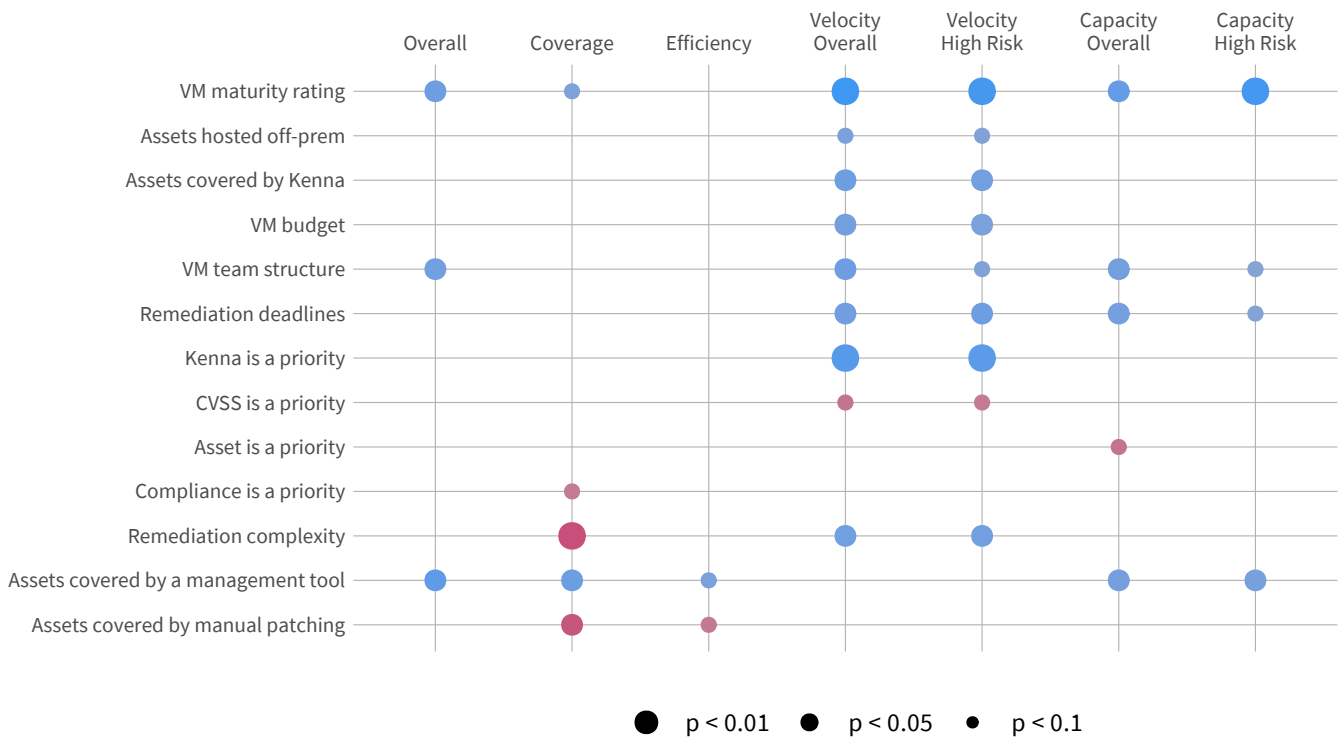
> "If Big Fix pushes a patch, we have to make sure that app and middleware teams aren't duplicating effort by working on them manually."
>
> Financial Services respondent

> "We rely on automation in identifying and patching critical and high risk vulnerabilities in critical systems."
>
> Business Services respondent

Conclusion & Recommendations

We threw an awful lot of words, factors, metrics, charts, and statistics at you in this report. Please don't feel bad if you are struggling to get your head around it all. We certainly did. And we lost count of how many times we asked something like "I forget—how was such and so related to this and that?" Let's just say we're grateful for the interpretive guardrails laid down by statistical analysis.

FIGURE 29: SUMMARY OF CORRELATIONS BETWEEN VM PROGRAM FACTORS AND PERFORMANCE METRICS



To save you some of that confusion, we thought the best way to conclude this report was to sum it all up in a single chart for easy reference. Figure 29 captures several key aspects of our analysis. First, it indicates correlations between survey factors and performance measures with a dot. Second, the dot size corresponds to the level of significance for each pair of correlating variables (bigger dots mean a stronger relationship). Third, the color reveals the nature of that relationship. Blue means the factor has a positive or beneficial correlation with the associated metric, while red signals negative effects.

In addition to a handy concluding summary of everything we've covered, Figure 29 also doubles as a set of data-driven recommendations. Or, at least, options to consider for improving various aspects of your VM program. Many won't find new ideas or practices in our list of remediation factors, but we hope having evidence that some things show promise for improving performance and where you're likely to see the associated benefit is useful information.

And Nothing Else Matters?

Of course other things matter in remediation performance beyond what we tested and found significant in this report. Do not toss out processes or tools that seem to be working for your program just because they don't have a blue dot above. This was an initial attempt to see if a sample of internal program factors we thought might impact remediation performance actually did. And in that respect, we're pleased with the results. But we're not stopping here. We plan to continue this research over the next few years and would love to hear ideas on what you've found helpful thus far and where you'd like to see us go in the future. Thanks for reading.

Send your comments or questions to research@kennasecurity.com or research@cyentia.com.

> Figure 29 captures several key aspects of our analysis. First, it indicates correlations between survey factors and performance measures with a dot. Second, the dot size corresponds to the level of significance for each pair of correlating variables. Third, color reveals the nature of that relationship. Blue means the factor has a positive or beneficial correlation with the associated metric, while red signals negative effects.

PRIORITIZATION TO PREDICTION

VOL 4: MEASURING WHAT MATTERS IN REMEDIATION

KENNA Security

119 Cyentia INSTITUTE

> "We hope this combo of 'soft' survey and 'hard' data analysis yields valuable insights for vulnerability management programs."

© 2019 Cyentia Institute, LLC. All Rights Reserved. The Cyentia Institute and Kenna Security names and logos are the property of their respective owners.