Microsoft

Transparency report

Examining industry test results, August 2019

Prepared by

Microsoft Defender ATP Research Team

Microsoft

Table of Contents

Microsoft

1 Summary of latest industry test results

This report provides a review of the latest independent industry tests results for Windows Defender Antivirus, the next-generation protection component of Microsoft Defender Advanced Threat Protection (Microsoft Defender ATP), Microsoft's unified endpoint protection platform.

Over the last few years, Microsoft has been improving its performance in industry tests. Today, it consistently achieves high scores in these tests, demonstrating the strength of our protection capabilities and the innovations we continue to make in our security technologies.

While current antivirus tests don't necessarily reflect how attacks operate and how solutions are deployed in the real world, they can influence important business decisions. We are actively working with several leading industry testers to evolve security testing. Meanwhile, we're publishing this report to provide more details, insights, and context on test results. We'd like to be transparent to our customers and to the industry about our wins as well as improvement plans as a result of these tests.

1.1 AV-TEST: Perfect Protection scores (March-April, May-June 2019)

Windows Defender Antivirus achieved perfect scores (6.0/6.0) in the Protection module of AV-TEST's March-April 2019 and May-June 2019 Business User test cycles. The industry-leading antivirus solution has consistently achieved this feat in all AV-TEST cycles in the past 14 months.

In Usability, Windows Defender Antivirus achieved a perfect 6.0/6.0 in March-April but performed slightly lower (5.5/6.0) in May-June.

In the Performance test module, Windows Defender Antivirus improved its scores from 5.5/6.0 in March-April to a perfect 6.0/6.0 in May-June. Learn More >>

1.2 AV-Comparatives: Approved Business Product (March-June 2019)

In July 2019, AV-Comparatives released the Business Security Test 2019 (March-June 2019) report, which combines results from various reports. Windows Defender Antivirus retained the recognition as an Approved Business Product.

Windows Defender Antivirus achieved a protection rate of 99.9% in the Real-World Protection Test (March-June) and 99.5% in Malware Protection Test (March). Learn More >>

1.3 SE Labs: AAA Award (January-March, April-June 2019)

In SE Labs' Enterprise Endpoint Protection test for January-March and April-June 2019, Windows Defender Antivirus won the AAA Award.

Windows Defender Antivirus registered 98% Protection Accuracy rating in January-March and 96% in April-June, combining with 100% Legitimate Accuracy in both test periods for a consistent Total Accuracy rating of 99% on both test periods. Learn More >>

1.4 Industry recognition

This report only covers results from enterprise editions of industry tests. The three testing organizations we include all have consumer versions of these tests, taking into consideration the unique security scenarios for home users. In these consumer tests, Windows Defender Antivirus performs just as good or even better.

For example, in AV-TEST's May-June 2019 Home Users test, Windows Defender Antivirus achieved perfect scores (6.0/6.0) in all three test modules (Protection, Usability, Performance), a feat only three other vendors were able to achieve.

These consumer test results further convey the positive outcomes of our investments to make Windows Defender Antivirus an industry-best solution. The impressive showing of Windows Defender Antivirus solutions also prompted industry publications to run stories with headlines like the following:

- Windows Defender Achieves 'Best Antivirus' Status (PC Mag)
- Windows Defender Gets Perfect Scores in Antivirus Test (Tom's Hardware)
- Microsoft's Windows Defender Is Now One of the Best Antivirus Apps in the World (Softpedia)

Additionally, CNET named Windows Defender Antivirus the best free Windows antivirus in its list of The best antivirus protection of 2019 for Windows 10.

2 Examining AV-TEST results

2.1 Summary of overall AV-TEST scores

The table below summarizes the overall test results for Windows Defender Antivirus in the March-April and May-June 2019 AV-TEST Business User test:

	March-April	May-June
Protection	6.0/6.0 (±0)	6.0/6.0 (±0)
Usability	6.0/6.0 (+0.5)	5.5/6.0 (-0.5)
Performance	5.5/6.0 (±0)	6.0/6.0 (±0)

Table 1. Windows Defender Antivirus' overall antivirus test results in the March-April 2019 and May-June 2019 Business User Test. AV-TEST uses Protection, and Usability, and Performance test modules.

2.2 Understanding Protection scores

Below are details on the Protection test scores.

	March-April	May-June
Real World testing	99.6% (276/277)	100% (307/307)
Prevalent Malware testing	100% (6,572/6,572)	100% (2,428/2,428)
Overall malware protection rate (all samples)	99.8% (6,848/6,849)	100% (2,735/2,735)
Overall Protection score >>>	**6.0/6.0** (±0)	**6.0/6.0** (±0)
Overall Protection ranking >>>	1st out of 18 (tied with 14 more)	1st out of 18 (tied with 8 more)

Table 2. Summary of Protection scores for the March-April and May-June 2019 Business User tests.

Windows Defender Antivirus detected 100% of malware files used in the Prevalent Malware testing in both March-April and May-June 2019 cycles. It missed one file in the Real-World testing in March-April, but once again detected 100% of files in May-June. All in all, Windows Defender Antivirus detected all, but one of the 9,583 files used in both tests in the last two cycles.

The diagrams below show Windows Defender Antivirus detection rates in the Prevalent Malware and Real-World tests over a one-year period. Windows Defender AV achieved 100% in 11 out of the 12 monthly Prevalent malware tests and 100% in 10 out of the 12 monthly Real-World tests.

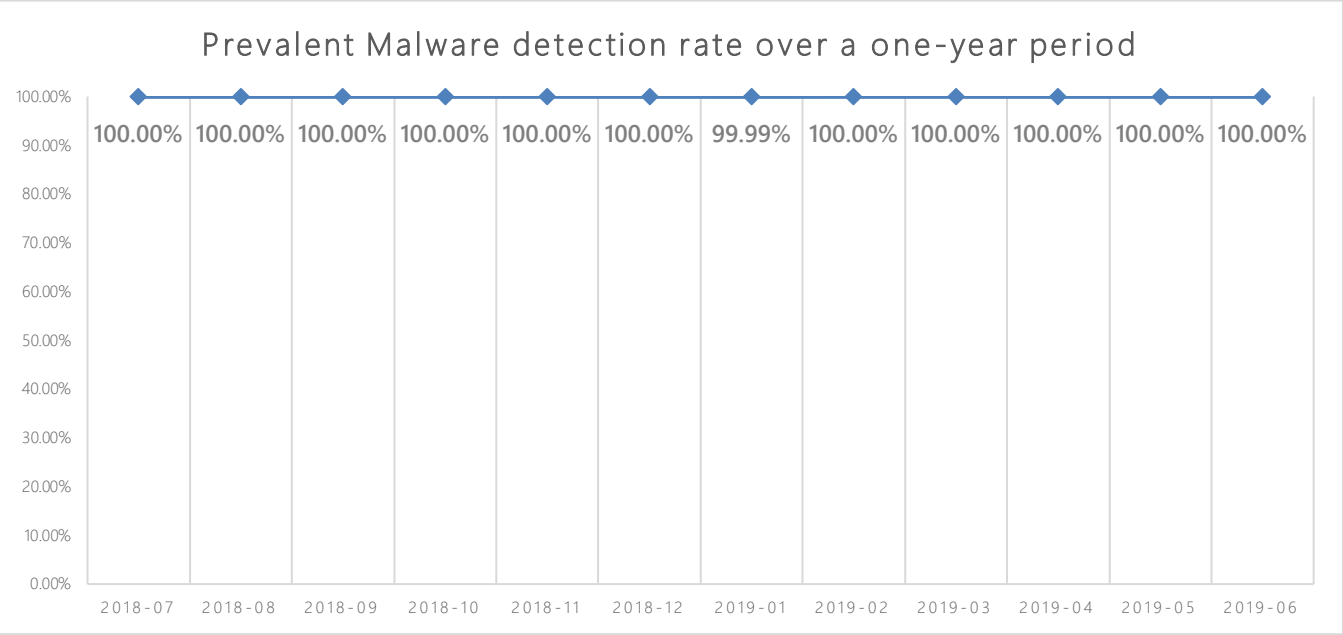

Figure 1. Windows Defender Antivirus detection rates in AV-TEST "Prevalent malware" tests over a one-year period

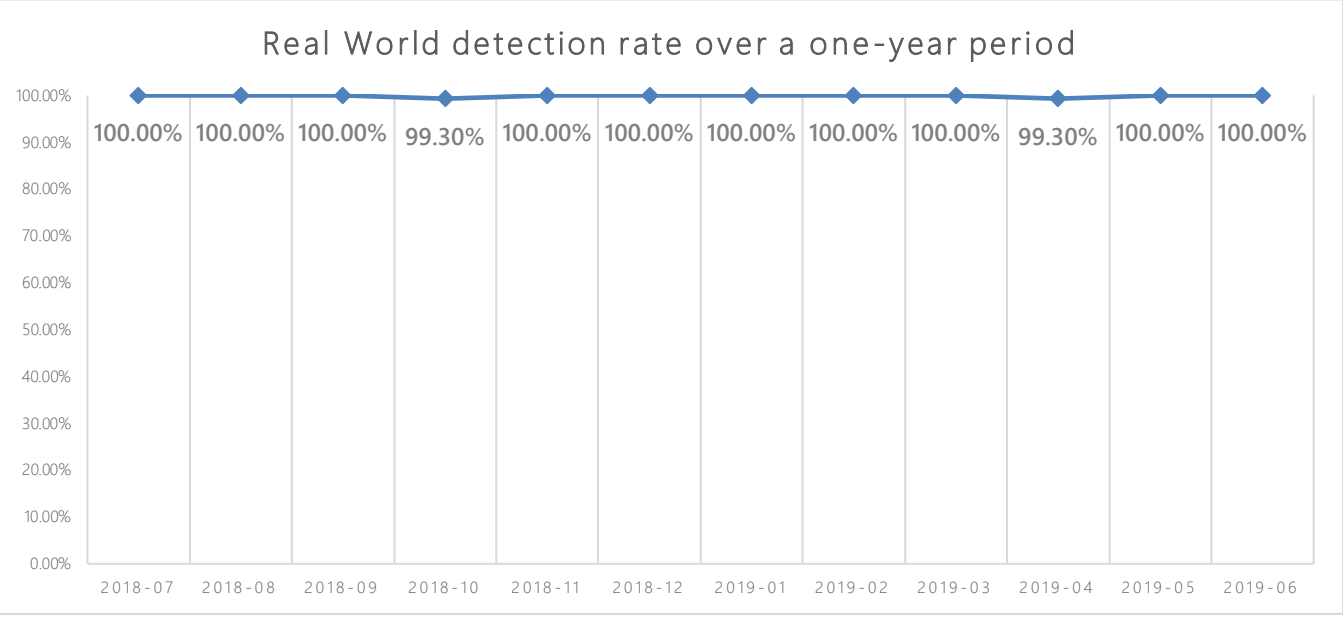

Figure 2. Windows Defender Antivirus detection rates in AV-TEST "Real World" tests over a one-year period

2.3 Understanding Usability scores

In Usability tests, AV-TEST includes clean file samples in the test population and checks whether antivirus products incorrectly classify them as malware (what is known as false positive, or FP). Below is a summary of the results for Windows Defender Antivirus in the Usability test.

	March-April	May-June
Number of misclassified files	1 (out of 1,645,832 samples)	5 (out of 1,516,283)
Overall Usability score >>>	**6.0/6.0** (+0.5)	**5.5/6.0** (-0.5)
Overall Usability ranking >>>	1st out of 18 (tied with 14 more)	1st out of 18 (tied with 3 more)

Table 3. Summary of Usability test scores for the March-April and May-June 2019 Business User test

2.3.1 Analysis: What kinds of files were misclassified?

Our research team analyzed the samples that Windows Defender Antivirus misclassified and assigned proper determination. The team also analyzed the root causes of these misclassifications and worked with different threat research teams to enhance detection accuracy.

Below is a list of files that Windows Defender Antivirus misclassified in the two test cycles. Based on our research and on file prevalence data, most of the misclassified samples are not common in enterprise environments.

Sample	File prevalence (30 days)	Description	Digitally signed? (Y/N)
Sample a	184	Gaming application	N

Table 4. Files that Windows Defender antivirus incorrectly classified as malware during March-April 2019 Business User test

Sample	File prevalence (30 days)	Description	Digitally signed? (Y/N)
Sample a	171	Photo editing software	N
Sample b	1,284	Browser installer	Y
Sample c	58	VPN app	Y
Sample d	-	Media converter	Y
Sample e	1,131	JavaScript redirector	N

Table 5. Files that Windows Defender antivirus incorrectly classified as malware during May-June 2019 Business User test

Microsoft encourages software vendors to take steps to raise the level of trust both by security vendors and users alike. These steps include signing software with certificates issued by reputable Certification Authorities.

2.3.2 The synthetic nature of usability tests

Misclassifications in a synthetic test are not necessarily indicative of false positives in real-world scenarios. This is true when the test methodology discounts contextual elements that Windows Defender Antivirus uses for issuing a verdict. For example, when a file is tested, it is not downloaded from the vendor website. Both the original file name and the download site are contextual information that is removed in tests. We've seen many cases where a customer in the real world downloads a clean program from the vendor site without encountering any erroneous detection. However, when a tester gives the file a seemingly random name (e.g., its SHA-256 hash), removes the mark of the web, and doesn't download the file from the vendor website, some of our more aggressive machine learning models issue blocks that don't occur in the real world.

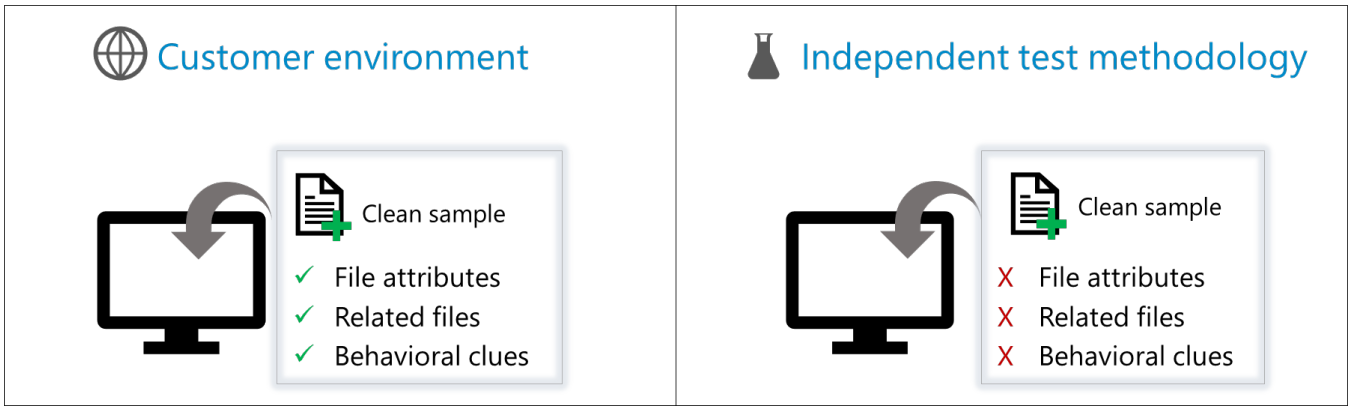

Figure 3. In some cases, samples are incorrectly classified (false positive) in the synthetic test environment but not on customer machines.

2.3.3 Criteria for evaluating files may vary across vendors and testers

The criteria for classification can vary between antivirus vendors and testers depending on their policies. Some files identified as clean by some vendors could be files that Windows Defender Antivirus identifies as a potentially unwanted application (PUA) and thus would be blocked. Microsoft's policy aims to protect customers against malicious software while minimizing the restrictions on developers. The diagram below demonstrates the high-level evaluation criteria Microsoft uses for classifying samples:

- Malicious software: Performs malicious actions on a computer.
- Unwanted software: Exhibits the behavior of adware, browser modifier, misleading, monitoring tool, or software bundler
- Potentially unwanted application (PUA): Exhibits behaviors that degrade the Windows experience
- Clean: We trust that the file is not malicious, is not inappropriate for an enterprise environment, and does not degrade the Windows experience

Figure 4. Microsoft's high-level sample classification criteria

2.4 Understanding Performance scores

Performance tests measure the effectiveness of certain user actions, which are executed as part of the test, on system speed. In the May-June test, Windows Defender Antivirus achieved a perfect performance score, demonstrating the emphasis we have been putting in improving the impact of our antivirus solution to user actions. The table below summarizes Performance test results.

	March-April	May-June
Overall Performance test score >>>	**5.5/6.0 (+0)**	**6.0/6.0 (+0.5)**
Performance ranking >>>	1st out of 18 (tied with 6 more)	1st out of 18 (tied with 11 more)

Table 6. Summary of Performance test scores for the March-April and May-June 2019 Business User test

The tables below present Windows Defender Antivirus' performance test results compared to industry averages during March-April and May-June 2019 test cycle. Performance is measured by the average impact of the product on computer speed; therefore, a smaller number is favorable. Green boxes indicate areas where Windows Defender Antivirus performed better than or the same as the industry average; red boxes indicate performance lower than the industry average.

Action	Standard PC	*Industry average*	High-End PC	*Industry average*
Launching popular websites	5%	*16%*	7%	*12%*
Downloading frequently used applications*	0%	*1%*	0%	*1%*
Launching standard software applications	8%	*8%*	10%	*8%*
Installation of frequently used applications	44%	*27%*	39%	*24%*
Copying of files (locally and in a network)	1%	*3%*	0%	*4%*

Table 7. The average impact of the product on computer speed in daily usage during March-April 2019

*The description for these operations is given by AV-TEST and might not be aligned with what Microsoft's data indicates as realistic.

Action	Standard PC	*Industry average*	High-End PC	*Industry average*
Launching popular websites	10%	*15%*	13%	*13%*
Downloading frequently used applications*	1%	*1%*	0%	*1%*
Launching standard software applications	8%	*10%*	9%	*9%*
Installation of frequently used applications	27%	*22%*	23%	*19%*
Copying of files (locally and in a network)	0%	*2%*	1%	*3%*

Table 8. The average impact of the product on computer speed in daily usage during May-June 2019

2.4.1 Areas that matter the most to customers

Windows Defender Antivirus performed better than the industry average in most areas and had significant shortcoming in the area that AV-TEST labels as "*Installation of frequently-used applications*". There are several factors to consider for driving the right conclusion out of these test results:

- **Consider the frequency of the action**
  Most users in enterprise environments are information workers whose common user activities include:
  - Browsing the web
  - Using email clients
  - Processing documents
  - Accessing network resources

  Users spend substantially less time installing new applications compared to the activities listed above. This is true for all user segments, but especially for enterprises, where software installation is usually governed by usage policies. Windows Defender Antivirus is optimized for delivering high levels of performance during high-frequency actions. Performance is a priority area for the Windows Defender Antivirus team, and we're working to improve it even further.

- **Consider the level of risk**
  Windows Defender Antivirus is designed to perform thorough scanning during the software installation process. This could have a performance cost. One reason for this is that software installation is a relatively complex operation that touches different areas of the operating system. A thorough inspection is necessary to reduce the risk of introducing malicious software on the system.

- **What impactful areas are not being tested?**
  There are several areas that are not being tested for performance by AV-TEST that are critical to user experience. Examples include:

  - Shutdown and startup
  - Universal Windows app launch
  - Battery consumption

3 Examining AV-Comparatives results

The table below summarizes overall test results for Windows Defender Antivirus in the March-June 2019 antivirus testing by AV-Comparatives:

	Real-World	Malware
Overall scores for this cycle >>>	**99.9%**	**99.5%**

Table 9. Windows Defender Antivirus' overall antivirus test results in the March-June 2019 AV-Comparatives Business Security test. AV-Comparatives uses Real-world protection, and Malware protection, test modules.

3.1 Understanding Real-world protection test scores

The below table displays more details on the results of the Real-World Protection test. The results are based on a test set consisting of 732 test cases (such as malicious URLs) tested from the beginning of March through the end of June 2019.

	March-June
Blocked	730
User dependent	2
Compromised	-
Overall Real-world protection rate** (all samples)	99.9% (732/732)
Overall Real-world protection score >>>>	**99.9%**
False positives	24

Table 10. Summary of Real world protection scores for the March-June 2019 Business security test

**[Blocked % + (User dependent % / 2)]

The table below shows Windows Defender Antivirus detection rates in Real-World protection tests consistently improving over a one-year period.

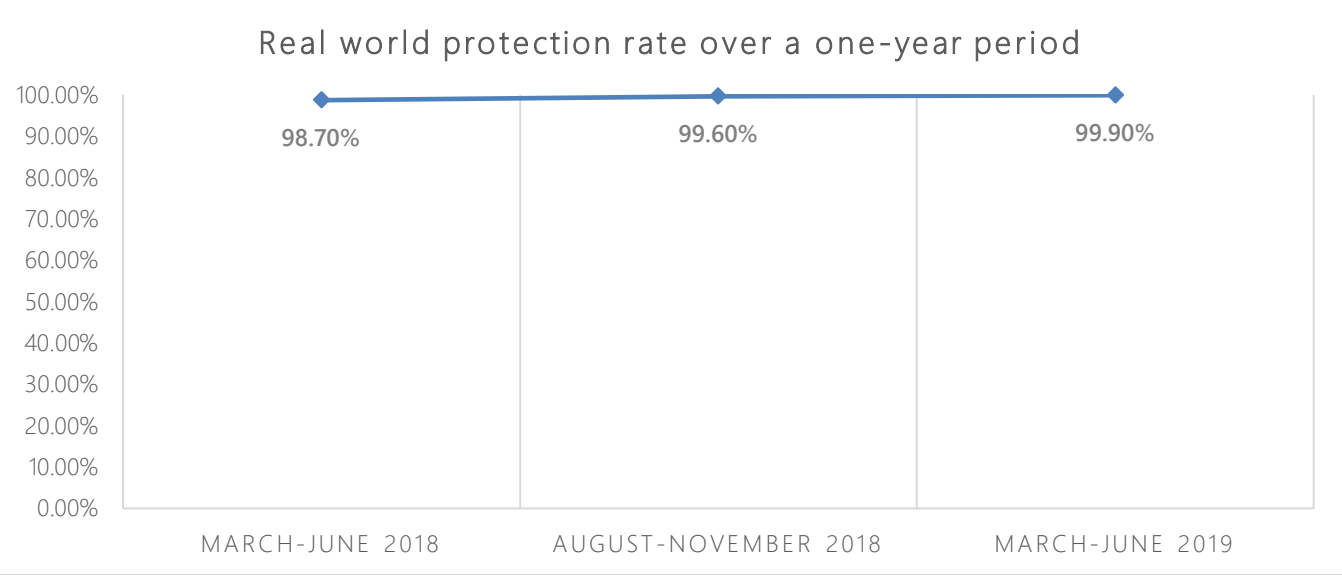

Figure 5. Windows Defender Antivirus detection rates in AV-Comparatives Real-World protection tests

3.2 Understanding Malware protection test scores

The below table gives a brief overview of the results of the Business Malware Protection test run in March 2019. The results are based on a test set consisting of 1,311 recent malware samples used during March 2019. Below are details on the Malware Protection test scores.

	March
Blocked	1,304/1,311
User dependent	0
Compromised	0.5%
Overall Malware protection rate (all samples)	99.5% (1,304/1,311)
Overall Malware protection score >>>	**99.5%**
False positives	0

Table 11. Summary of Malware protection scores for the March 2019 Business User test

The table below shows Windows Defender Antivirus detection rates in Malware protection tests over a one-year period. This test is conducted once every six months.

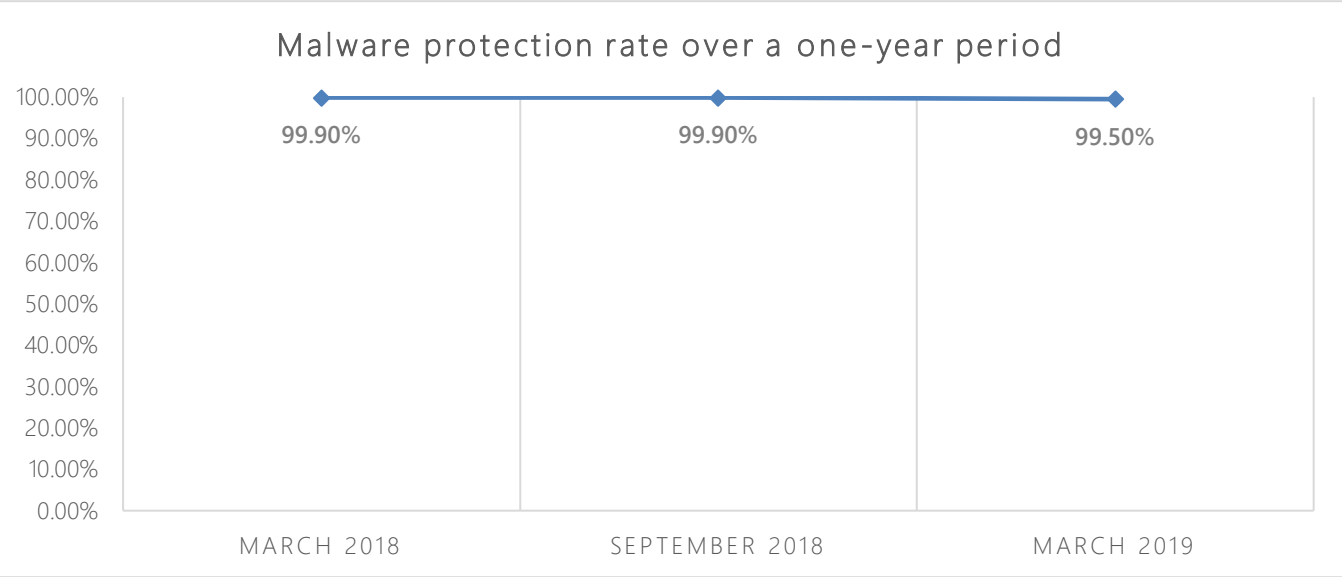

Figure 6. Windows Defender Antivirus Malware Protection rates in AV-Comparatives Malware protection tests

3.3 Analyzing false positives

In the Real-world protection test, Microsoft Defender ATP misclassified 24 files. Below is a list of the files that Windows Defender Antivirus misclassified. As we do for all test results, we analyzed these false positives.

Based on global prevalence data, these files are not common in enterprise environments. All misclassified files are not digitally signed. Microsoft encourages software vendors to help minimize false positives by taking steps to raise the level of trust both by security vendors and users.

Sample	Global file prevalence (30 days)	Description	Digitally signed? (Y/N)
Sample a	0	Compiler	N
Sample b	3	Video encoder library	N
Sample c	0	Image rename software	N
Sample d	16	Finance tool	N
Sample e	89	Backup service tool	N
Sample f	0	Driver update tool	N
Sample g	2	Lock screen image tool	N

Sample	Global file prevalence (30 days)	Description	Digitally signed? (Y/N)
Sample h	2	Image-embedding software	N
Sample i	1	Media converter	N
Sample j	1	Text editing tool	N
Sample k	0	Development and testing tool	N
Sample l	0	Event logger	N
Sample m	0	Web proxy application	N
Sample n	1	Text editing tool	N
Sample o	2	Keylogger tool	N
Sample p	2	File compressor	N
Sample q	9	Application organizer	N
Sample r	1	Image application plugins	N
Sample s	43	Video encoder	N
Sample t	26	Social media and games blocker	N
Sample u	0	Audio production software	N
Sample v	3	Mail migration tool	N
Sample w	30	Email client tool	N
Sample x	5	Game setup tool	N

Table 12. Files that Windows Defender Antivirus incorrectly classified as malware

As part of the Malware protection test, AV-Comparatives also ran a false positive test with common business software. Windows Defender Antivirus, like all other enterprise security solutions included in the test, had zero false positives. This is consistent with our observation about the files that Microsoft Defender Antivirus misclassifies on some tests. Revisit section 2.3.3 for more insights and commentary on false positives.

4 Examining the SE Labs results

4.1 Summary of overall results

The table below summarizes the overall test results for Windows Defender Antivirus in the January-March and April-June 2019 testing by SE Labs:

Test category	January-March	April-June
Protection accuracy	98%	96%
Web downloads score	75/75	74/75
Targeted attacks score	25/25	25/25
Legitimate software accuracy	100%	100%
Total accuracy rating	**99%**	**99%**

Table 13. Overall Windows Defender Antivirus test results in the SE Labs test.

4.2 Understanding Protection Accuracy test scores

SE Labs determines the Protection accuracy scores based on the combined outcome of two tests:

1. Web downloads (75 test cases)
2. Targeted attacks (25 test cases)

SE Labs goes beyond the binary rating (i.e., blocked vs. compromised) in rating protection effectiveness. Instead, SE Labs considers the nuances of the interaction between the product and the threat. For example, it issues a different rating for *Blocked (+2 points)* from what is given for *Complete remediation (+1 points)* or *Compromised system (-5 points)*. The other ratings used by SE Labs for both Web downloads and Targeted attacks tests are: *Detected (+1)*, *Neutralized (+1)*, *Persistent neutralization (-2)*. A rating is assigned to each product-threat interaction operation and a combined score is calculated for each product.

Windows Defender Antivirus achieved the following combined score for Web downloads and the Targeted attacks tests.

	Jan-Mar	Apr-June
Detected	100	99
Blocked	96	99
Neutralized	4	0
Compromised	0	1
Protected	100	99

Table 14. Summary of Windows Defender Antivirus scores in the Protection accuracy test

In the Apr-June test, Windows Defender Antivirus missed 1 of the 75 samples used in the Web downloads test.

When it comes to the Targeted attacks test, the protection score considers the extent of protection demonstrated by the product (i.e., the attack stage in which the product was able to block the threat). Points are deducted for *Access (-1)*, *Action (-1)*, *Escalation (-2)*, and *Post-escalation action (-1)*. Windows Defender Antivirus detected or blocked all the targeted attacks in the test.

4.3 Understanding Legitimate Software Accuracy test scores

SE Labs Legitimate Software Accuracy test measures the endpoint product's ability to correctly classify legitimate applications. SE Labs assigns ratings based on how the product classifies an object (safe, unknown, not classified, suspicious, unwanted, or malicious) and the level of interaction required of the user (e.g., click, or no interaction required).

SE Labs also takes into consideration the prevalence of the legitimate application to account for the breadth of the business impact of incorrectly blocking. This prevalence factor is expressed as a modifier and is multiplied by the interaction rating to determine the product score.

Windows Defender Antivirus correctly classified 100% of legitimate applications as safe in both Jan-Mar and Apr-June 2019 test cycle.