Gallagher Re

Can scanning technologies predict claims?

Cyber IQ Report

JUNE 2023

TABLE OF CONTENTS

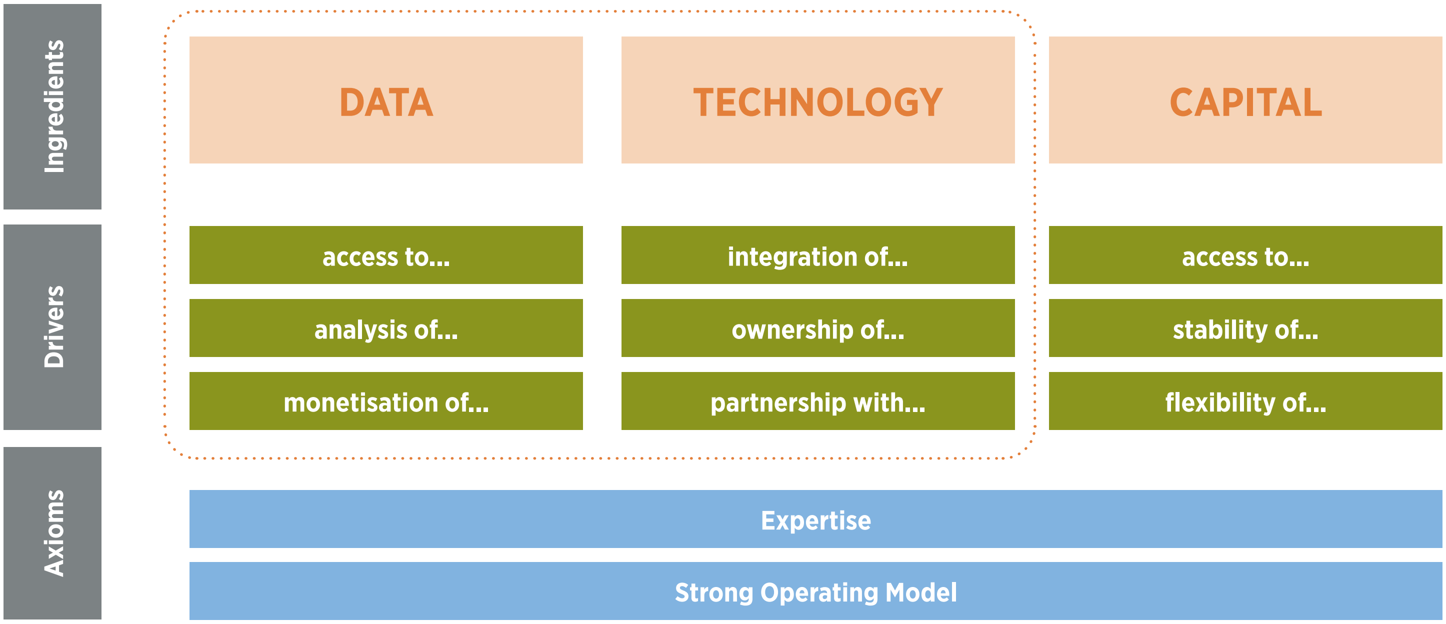
EXECUTIVE SUMMARY

Technology and data in the context of cyber insurance are often used as shorthand to describe the rapid uptake and multifaceted use of external scanning methods by (re)insurers.

However, having access to data alone does not guarantee success. Rather, insurers must ensure they can rigorously translate this access into actionable insights across their business. This includes building infrastructure that works for cyber insurance and overcoming hurdles presented by cumbersome legacy systems.

Figure 1 below outlines the ingredients, drivers and axioms that are critical for (re)insurer success. While this paper focuses on data and technology, our recent whitepaper highlighted the role that capital plays in this fast-growing class.[1]

> External scanning is, and will remain, a key pillar of most successful cyber insurance offerings—although carriers' use of tech and data will broaden in future with the advent of solutions that capture elements of an organisation's internal network.



Ingredients	DATA	TECHNOLOGY	CAPITAL
Drivers	access to...	integration of...	access to...
	analysis of...	ownership of...	stability of...
	monetisation of...	partnership with...	flexibility of...
Axioms	Expertise		
	Strong Operating Model		

Figure 1: The three ingredients we anticipate to be key for securing long-term success in the cyber market.

[1] https://www.ajg.com/gallagherre/news-and-insights/2023/may/the-vital-role-of-capital-in-cyber-reinsurance/

In '*Looking from the Outside-In: Can taking the threat actors' viewpoint help insurers?*', a study produced by Gallagher Re in 2022, the paper argued the value of external scanning was clear: arming insurers with the same data attackers use to select and compromise targets.

There were also likely applications across an insurance lifecycle; offering a complementary view to underwriting questionnaires by focusing on how security controls are deployed in practice (rather than how they are designed) and presenting a potential 'force multiplying' effect for portfolio optimisation.

However, false positives, infrequency of updates and sheer data volume obfuscate the value in underlying data, as well as differences between vendor methodologies, challenging the reliability of scores. This complexity renders it difficult to amplify findings and translate them into digestible insights for insurers.

Arguably, the greatest limitation has not been outside-in scanning itself, but the uncertainty surrounding its ability to predict cyber claims, rendering it difficult for insurers to:

- Evaluate vendors and data objectively;
- Place reliance on technology in an appropriate and proportional way; and
- Gain trust and better terms from capacity providers/strategic partners for the effective use of scanning.

Over the course of 2022, Gallagher Re has built a machine-learning model and combined this with historical claims, to better understand which elements of external scanning data would have been more predictive of claims at the point of underwriting.

This paper presents the key insights from the study into the ability of the data to predict cyber claims, as well as outlining upcoming trends with insurers' uptake of outside-in technology.

TRENDS FOR 2023: INSURANCE'S USE OF EXTERNAL SCANNING IS RAPIDLY EVOLVING

Last year, Gallagher Re remarked on the rapid uptake of external scanning technology by carriers after the cyber market hardened.

In 2023, insurance use of external scanning is better captured by the word 'refinement', with carriers now focused on integrating technology into existing processes (an effort often hampered by cumbersome legacy and cross line of business systems) and licencing vendors that better suit requirements.

Refining how technology is used now will likely deliver a many-fold return on investment in future years.

As this paper will outline, external scanning technology only has the ability to predict cyber claims when the use is focused on a small percentage of its dataset. Existing vendor approaches to consolidating data, and amplifying important findings are (currently) no substitute to internal insurer expertise. Indeed, Gallagher Re has found vendor view of risks (usually articulated by overall scores) to be inconsistent and often weighted in favour of data with little or no value for predicting cyber claims.

Figure 2 below shows the range of ranked scores for select companies in our portfolio, with a taller vertical bar highlighting disagreement between vendors.

> We anticipate 2023 will expose a bifurcation in insurer capability for using external threat scanning to drive loss ratio with the majority of insurers not enjoying these benefits.

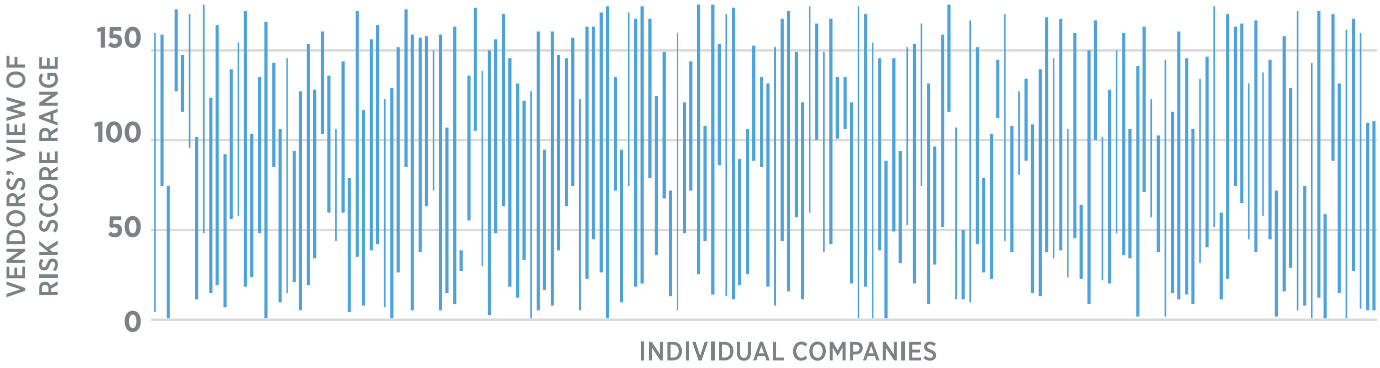


Figure 2: Chart outlining difference between five vendors' view of relative cyber risk presented by companies in a portfolio. The higher the line goes, the more risky the company is considered to be. Each blue line represents one company.

Encapsulating a year of ‘refinement’ in how external scanning is used by insurers, we’ve identified the following trends for 2023.

‘Attack management’ becomes a standard part of the cyber insurance offering.

This is the practice of determining which insureds are vulnerable to an emerging widespread event and highlighting the potential exposures while offering good practice advice to remediate the risk.

Any insurer with close relationships to their insureds will find it easier to reap the benefits of helping insureds manage emerging events using this technology. Uptake benefits from, but doesn’t require, deep technical integration with existing systems.

Part of the appeal for utilising external scanning for attack management is to avoid ‘alert fatigue’, where insureds (for myriad reasons) do not process and respond to alerts that could prevent/limit the severity of an incident.

Even the most effective deployments of this practice enjoy less than 25% active response, but its speed and ability to directly reduce claim frequency and severity make this a worthy investment.

More insurers will work with multiple external scanning data providers.

Historically, a complex vendor landscape increased the likelihood of insurers working with vendors that do not meet their needs. Vendors have responded to a increased competition by looking to differentiate themselves. Some of this differentiation is by use case, e.g., a vendor focused on incident response use cases may not be a leader for underwriting. Differentiation is also being sought in the granularity of data offered, with some providers looking to help insurers with translating data into actionable insights, whilst others focus on providing a wide array of malleable data for insurers to use in building their own models.

32 of 34
insurers using external scanning data

13 of 34
insurers using multiple scanning vendors

23 of 34
insurers using scanning in underwriting

Figure 3: Uptake of external scanning continues to grow, but insurer focus is now on refinement of how technology is used. While an increasing number of insurers have access to the data, the level of use and reliance varies widely between carriers.

ATTRIBUTES FOR SUCCESS IN USING EXTERNAL THREAT SCANNING

✓ Internal or accessible cybersecurity, data science and actuarial expertise to translate data into insights.

✓ Selecting the right vendor for the right use cases.

✓ Using the right data in a targeted and proportional way.

✓ Use across the insurance lifecycle, from underwriting to portfolio optimisation.

The next generation of AI tools could help us overcome our dependency on URLs.

With external scanning vendors largely unable to offer automated solutions for reliably matching URLs to company name and firmographic data (industry, company size and geography), the limitation in availability of URLs renders it difficult for insurers to reap the rewards of technology use.

This challenge of matching URLs to firmographic data seems a perfect fit for a new generation of AI tools, heralded by the release of ChatGPT in December 2022. However, the hope offered by AI is not a substitute for adjusting processes and systems to capture URLs, and any cyber insurer should prioritise this.

Insurers are more discerning when (re)evaluating vendors.

Akin to the change in how (re)insurers evaluated accumulation modelling providers between 2018–2020, greater rigour and depth are now applied by insurers to selecting the right external scanning vendors. This is driven by myriad factors, including:

- More mature understanding of external scanning technology's strengths and drawbacks.
- Greater access to internal and third-party expertise for evaluating different data sets.
- Automation and API access enabling greater integration of external data with incumbent systems and processes.

While the use in risk selection becomes more targeted, direct application of scanning technology for pricing remains rare.

A growing minority of carriers are moving away from utilising an overall score as a check and balance on questionnaire results and towards working with 'risk factors' and individual data points for targeted and repeatable/automated decision-making. This marks an evolution in how insurers are using scanning data for risk selection.

Other examples include declining a risk due to the presence of exposed Remote Desktop Protocol (RDP) or including sublimits based on a particular risk factor score.

This targeted use of the data lays the foundations for insurers to circumvent insignificant data and focus on only data that adds value.

> Whilst the requirement of URLs to leverage external scanning technology has been clear for many years, to identify the technographic footprint of an organisation, the industry has remained stubbornly resistant to capturing URLs at the point of underwriting.

STUDY RESULTS: CAN EXTERNAL SCANNING PREDICT CYBER CLAIMS?

Over the course of 2022, Gallagher Re built a machine-learning model and combined this with historical claims, to identify which external scanning factors would have predicted a claim at the point of underwriting. Other public reports on the predictivity of external scanning data have been in partnership with vendors. However, we consider an objective and vendor-agnostic approach to have greater potential for building trust in the targeted and proportional use of technology.

We outline our key findings and our modelling approach below:

Revenue is the greatest claims predictor

It's somewhat comforting that when modelled alongside external scanning (technographic) data, revenue and industry outpace all other data points in their ability to predict claims frequency.

This highlights the importance of doing the basics right and shows that technographic data is no outright substitute for these traditional data points—our model lost some predictive value without their inclusion.

Our next model will consider whether scanning data offers technographic substitutes for this data, e.g., is number of IP addresses observed a better indicator of company size than revenue for cyber insurance?

Point of underwriting scanning offers a limited perspective

Most scanning technologies provide scores and granular data based on point-in-time findings; our modelling finds that evaluating scanning results over 365 days offers material additional value for anticipating claims.

Since much attritional untargeted ransomware is driven by automated/recurring scanning from threat actors, evaluating data over an equivalent policy lifecycle utilises this same method and reveals ad-hoc processes and recurring habits that expose insureds to attackers.

When less is more

When looking at the ability of technographic data individually to predict claims, many hold some value. However, there is a high correlation between many data points. Of the 22 different risk rating factors analysed in our study, a large proportion is highly correlated, meaning features may contain similar information and that a smaller number of scores offer additive value.

Modelling the additive value of risk factors in predicting claims found the majority of technographic data, including data on botnets and unexplained communications add no value to the risk selection for the testing period.

This finding further solidifies the importance of using external scanning data in a targeted way, focusing on specific risk factors regardless of use case.

However, (re)insurers should be cautious of evaluating the technology from the perspective of historical losses alone. Evolutions in the threat landscape mean the risk factors most predictive of claims are likely to be a moving target. For example, assessing claims from 2016 and 2017 might find botnet to be a valuable indicator of claims frequency, due to higher use of botnets to launch damaging cyber attacks in that period. Similarly, with some insurers seeing an increase in business email compromise (BEC) and Fund Transfer Fraud (FTF), we may see the predictivity of email security indicators rise over the coming year.

Methodology

Claims over 18 months drawn from different firmographic groups curated and included.

Policy records included were complete with firmographic data.

Companies technographic data received and analysed.

Security ratings observations provided by BitSight.

Modelling Approach

To determine the predictiveness of factors and understand how to generate better ‘lift’ from the data, a range of explainable models was used, along with more complex ML models. A multi-model approach serves as additional validation, and provides a better understanding of the data, as well as focusing attention on the most relevant areas. For this phase of our analysis, claim frequency was the focus, as this is where the technology is anticipated to be most useful. Testing severity is planned for later in 2023.

The study selected a large number of policies from an 18-month period, across a cross-section of industries, revenue bands and countries. Gathering the associated reported claims from the following period, the policy firmographic data was combined with the outside-in technographic data provided by BitSight to determine the most relevant factors. We focused on analysing the outside-in data at the inception date to reflect an insurance underwriting approach.

Significant time was invested to classify the claims into different subtypes (e.g., ransomware, business email compromise, etc.) and models were fitted to the overall claims frequency and each subtype. This helped determine which factors are most relevant for different claim types. An approach that will also support the continued relevance of this technology as attacker and claims trends evolve.

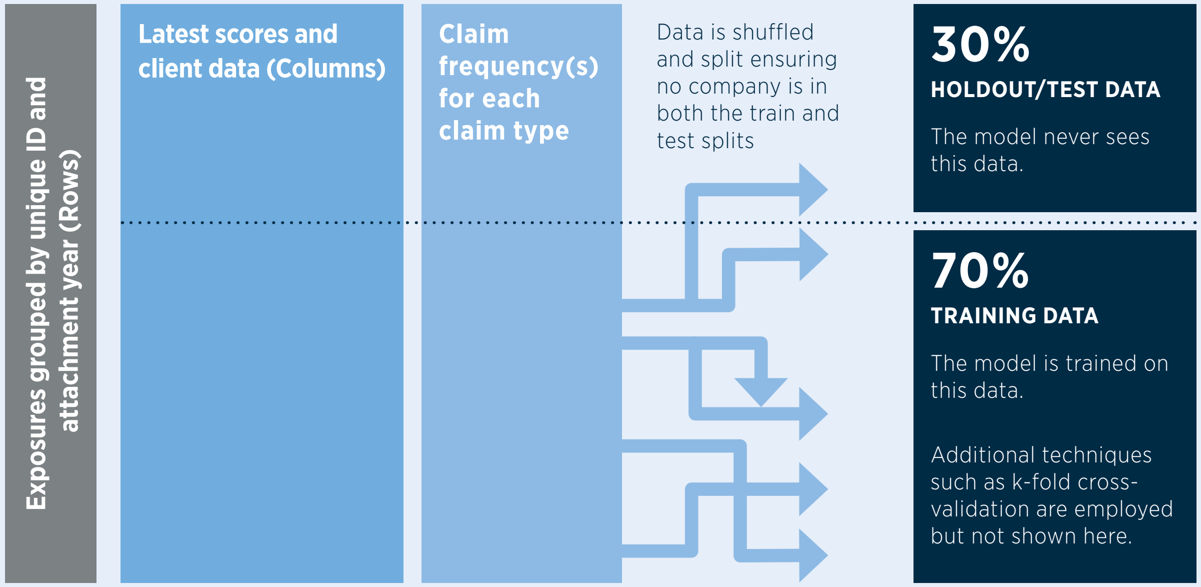
Simplistic one-factor linear regression models (GLMs) were developed first, to get an understanding of which factors seemed most correlated to distinguishing an elevated risk of claims. These simple models are easily communicable and accepted in the insurance world, but they also provide validation for the results of more complex and ‘black-box’ ML models.

A Gradient Boosted Model (GBM) was then trained on the dataset using all firmographic and technographic information available, allowing the model to determine the most important features and the weights to put on these.

The results of our study, and factors we found to be most predictive, are discussed below. More advanced methodology and model refinement techniques will be deployed in 2023 to better optimise the outcome and improve predictive lift.

Training and Testing Strategy

Splitting data into a training and holdout set enabled better understanding of the real-world predictive performance of the model.


Exposures grouped by unique ID and attachment year (Rows)
Latest scores and client data (Columns)
Claim frequency(s) for each claim type

Data is shuffled and split ensuring no company is in both the train and test splits

30%

HOLDOUT/TEST DATA

The model never sees this data.

70%

TRAINING DATA

The model is trained on this data.

Additional techniques such as k-fold cross-validation are employed but not shown here.

Patching cadence is the strongest technographic predictive indicator

Patching cadence refers to the speed by which organisations apply patches to critical external facing vulnerabilities.

Whilst point-in-time patching cadence showed predictive potential, our own scoring feature—to understand how an organisation patched key vulnerabilities over 365 days—showed materially greater predictive capability.

This predictiveness was especially pronounced for malware claims, highlighting the importance for organisations to maintain a rigorous approach to vulnerability identification and patching.

With an increased percentage of ransomware attacks resulting from the exploitation of external facing software vulnerabilities (Coveware 2022),[2] we can anticipate this to remain a strong indicator of claims frequency for the foreseeable future.

Port security is still a clear driver of claims

Our own scoring feature to understand how ports are exposed over 365 days found port security to be predictive of claims, behind only patching cadence and firmographic data.

The past 18 months has seen a steady move away from RDP being a primary attack vector for threat actors, largely due to security posture improvements.

However, companies with particular exposed ports will appear enticing to passing attackers.

Whilst negative port security findings might be used to decline business in the SME segment, individual port security findings can be utilised to drive more targeted conversations in underwriting meetings for large companies.

[2] https://www.corvusinsurance.com/blog/data-science-insight-how-vpn-vulnerabilities-affect-ransomware-risk?utm_campaign=Impact%20of%20Blog%20CTAs&utm_source=Blog%20CTA

Web security is a material driver of claims

Web security aims to represent the security posture of an organisation's external facing web presence, highlighting use of outdated protocols, poor certificate management practices and exposure to web-based attacks.

Examples of this include Cross-Site Scripting (XSS), where an attacker can impersonate victims through the injection of malicious client-side web code; or Man-in-the-Middle (MITM) attack, which involve a malicious actor intercepting traffic between a client and a web server, enabling the capture of login details and other sensitive information.

The risk factors we analysed broke web security into three components, SSL/TLS configuration, web certificates and Hypertext Transfer Protocol (HTTP) headers.

When these three factors are combined, they present a greater predictive potential for claims than either patching cadence or port security.

Mobile application security can't be ignored

The shift to remote working for many at the start of the pandemic mirrored a fundamental shift in the attack surface of many companies globally.

Identifying mobile devices connected to an organisation running outdated or vulnerable operating systems and/or applications is well-placed to capture that risk and added material value to predicting historical claims in the study.

An evaluation into the services an organisation uses to secure remote devices may further reveal exposure, as remote working technology increasingly becomes a target of attacks.

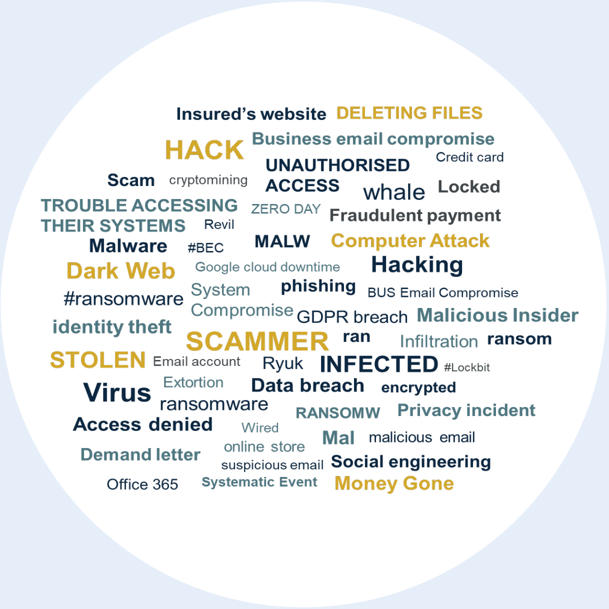
Cyber insurer Corvus (2022) have reported that organisations using a high-risk Virtual Private Network (VPN) is three times more likely to have a security incident than those without a VPN and five times more likely than with a low-risk provider.[3]

Case study: Claim type classification

The way claims are captured and classified can limit insurer capability to derive value from external scanning data and other technologies. Claims data is littered with inaccurate and misleading terminology which significantly hinders useful analysis. To combat this and further this study, Gallagher Re developed a taxonomy for claims.

Additional standardisation for cyber claims could unlock improvements in the ability of (re)insurers to improve accuracy and focus of risk selection and exposure management, whilst catalysing anticipation of/response to changes in the threat landscape.

For example, insurers with a limited approach to claims classification will likely be both underreporting and conflating incidents of FTF and BEC.



[3] https://www.corvusinsurance.com/blog/data-science-insight-how-vpn-vulnerabilities-affect-ransomware-risk?utm_campaign=Impact%20of%20Blog%20CTAs&utm_source=Blog%20CTA

External scanning excels at identifying the worst 20% of risks

One output from our model was a classification of a policy’s likelihood of suffering a claim, placing the portfolio into five buckets, A–E. In the training portfolio, each of these buckets is equally sized with 20% of policies placed in each category. Figure 5 below outlines that external scanning finds it hard to distinguish between the security posture of the strongest 80% of organisations—it is most effective when identifying organisations most at risk of suffering a claim. This is consistent with how many carriers are using external scanning data for risk selection.

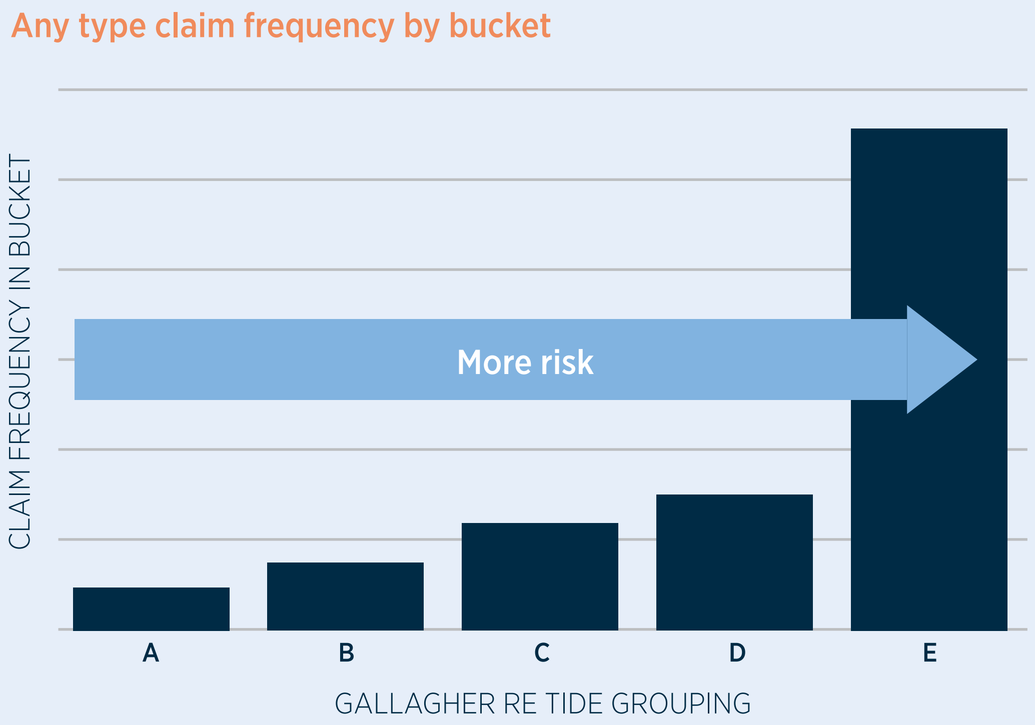


Figure 5: The actual claim frequency of policies in each grouping of Gallagher Re’s scoring methodology’s (TIDE) assessment of risk for the modelled portfolio. There is a clear increase in claim likelihood for the lowest-scoring policies.

The divide was particularly pronounced when looking specifically at malware claims. This is likely explained by the ability of external scanning to present the ‘attacker’s view’ in highlighting the specific vulnerabilities and exposures, e.g., RDP that drive untargeted and attritional ransomware.

Organisations without externally exposed popular compromise vectors for largely automated attacks are likely insulated from most untargeted malware events.

When looking specifically at BEC events, we found the worst 60% of organisations presented an elevated likelihood of a claim. This perhaps indicates that a greater proportion of organisations are susceptible to these events and whilst external scanning captures elements of email security, many BEC events have an element of human security failure.

EXTERNAL SCANNING: A COMPLEMENTARY TOOL

The landscape of how (re)insurers are utilising external scanning has shifted profoundly over the past year—uptake of technology continues to rocket, whilst the application of it becomes more varied and nuanced.

Insurers have become more discerning around evaluating vendors and embedding the technology in post-bind 'attack management', but many are still hamstrung in optimising uptake by limitations, including legacy or incompatible systems and processes; the availability of data; and uncertainty over which data points to incorporate in underwriting.

Our study combined cybersecurity ratings with firmographic and claims data, using machine learning algorithms. The research concluded that some technographic data holds the ability to predict cyber claims and the potential to provide a powerful resource for an insurer's arsenal when used in an appropriate and proportionate manner.

Our findings indicate that external scanning data is a particularly valuable tool for identifying the worst 20% of risks, which appear to be materially more likely to suffer a claim.

Therefore, despite only a small percentage of external scanning data being predictive of claims, it appears to be an invaluable tool for underwriters, capturing complementing aspects of a company's cybersecurity posture to questionnaires and evidencing whether insureds are applying their security policies in practice. Web security, patching cadence and port security were the risk factors we found to be most predictive of claims.

Beyond this, one of the most exciting prospects of external scanning technology is its ability to provide real-time feedback on a company's cybersecurity posture. In the context of emerging events, this can be used to alert companies and insurers to new risks as they emerge, enabling them to assess their portfolio's exposure and to act proactively, supporting insureds in mitigating risks before they turn into claims.

With rates beginning to stabilise this year, insurers need to ensure they have a clear understanding of each policyholder's risk profile if they are to price policies accurately and sustainably. A new phase in the uptake of external scanning technology is anticipated, with companies using the technology in a more targeted way amid greater integration into existing processes.

These three reflections will guide Gallagher Re's continued analysis of external scanning over the coming year.

1 **Modelling:** Creating separate models for small and medium enterprises in consideration of larger risks.

2 **Financial Impact:** Investigating and estimating the financial impact of re-underwriting based on findings.

3 **Market Engagement:** Improving market engagement and feedback in relation to the implementation of Organoid intelligence (OI) technology.

Authors

Ed Pocock
Senior Cyber Security Consultant
Ed_Pocock@GallagherRe.com

Michael Georgiou
Senior Pricing Actuary
Michael_Georgiou@GallagherRe.com

Contributors

Justyna Pikinska
Head of Cyber Analytics
Justyna_Pikinska@GallagherRe.com

James Poynter
Lead Data Scientist
James_Poynter@GallagherRe.com

James Rayner
Cyber Security Consultant
James_Rayner@GallagherRe.com

It’s the *way* we do it.

Drawing on our network of reinsurance and market specialists worldwide, and as part of the wider Gallagher company, Gallagher Re offers the benefits of a top-tier reinsurance broker, one that has comprehensive analytics and transactional capabilities, with an on-the-ground presence and local understanding. Whether your operations are global, national or local, Gallagher Re can help you make better reinsurance and capital decisions, access worldwide markets, negotiate optimum terms and boost your business performance.

For more information, visit **GallagherRe.com**.

Gallagher Re

© Copyright 2023 Arthur J. Gallagher & Co. and subsidiaries. All rights reserved: No part of this publication may be reproduced, disseminated, distributed, stored in a retrieval system, transmitted or otherwise transferred in any form or by any means, whether electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of Arthur J. Gallagher & Co. Gallagher Re is a business unit that includes a number of subsidiaries and affiliates of Arthur J. Gallagher & Co. which are engaged in the reinsurance intermediary and advisory business. All references to Gallagher Re below, to the extent relevant, include the parent and applicable affiliate companies of Gallagher Re. Some information contained in this document may be compiled from third party sources and Gallagher Re does not guarantee and is not responsible for the accuracy of such. This document is for general information only and is not intended to be relied upon. Any action based on or in connection with anything contained herein should be taken only after obtaining specific advice from independent professional advisors of your choice. The views expressed in this document are not necessarily those of Gallagher Re. Gallagher Re is not responsible for the accuracy or completeness of the contents herein and expressly disclaims any responsibility or liability, based on any legal theory, for damages in any form or amount, based upon, arising from or in connection with for the reader’s application of any of the contents herein to any analysis or other matter, or for any results or conclusions based upon, arising from or in connection with the contents herein, nor do the contents herein guarantee, and should not be construed to guarantee, any particular result or outcome. Gallagher Re accepts no responsibility for the content or quality of any third-party websites that are referenced.

The contents herein are provided for informational purposes only and do not constitute and should not be construed as professional advice. Any and all examples used herein are for illustrative purposes only, are purely hypothetical in nature, and offered merely to describe concepts or ideas. They are not offered as solutions for actual issues or to produce specific results and are not to be relied upon. The reader is cautioned to consult independent professional advisors of his/her choice and formulate independent conclusions and opinions regarding the subject matter discussed herein. Gallagher Re is not responsible for the accuracy or completeness of the contents herein and expressly disclaims any responsibility or liability based on any legal theory or in any form or amount, based upon, arising from or in connection with for the reader’s application of any of the contents herein to any analysis or other matter, nor do the contents herein guarantee, and should not be construed to guarantee any particular result or outcome. Gallagher Re is a trading name of Arthur J. Gallagher (UK) Limited, which is authorised and regulated by the Financial Conduct Authority. Registered Office: The Walbrook Building, 25 Walbrook, London EC4N 8AW. Registered in England and Wales. Company Number: 1193013. AJG.com/uk