Emerging Risk Report – 2015
Innovation Series

UNDERSTANDING RISK

Emerging Liability Risks

Harnessing big data analytics
About Lloyd's

Lloyd's is the world's only specialist insurance and reinsurance market that offers a unique concentration of expertise and talent, backed by strong financial ratings and international licences. It is often the first to insure new, unusual or complex risks, providing innovative insurance solutions for local, cross-border and global risks. Its strength lies in the diversity and expertise of the brokers and managing agents working at Lloyd's, supported by capital from across the world.

In 2015, more than 90 syndicates are underwriting insurance and reinsurance at Lloyd's, covering all lines of business from more than 200 countries and territories worldwide. Lloyd's is regulated by the Prudential Regulatory Authority and Financial Conduct Authority.

Harnessing Big Data Analytics is an Emerging Risk report published by Lloyd's as part of its Innovation Series.

Key contacts

Trevor Maynard
Exposure Management
Trevor.maynard@lloyds.com

Nick Beecroft
Manager, Emerging Risks & Research
nick.beecroft@lloyds.com

For general enquiries about this report and Lloyd’s work on emerging risks, please contact emergingrisks@lloyds.com

Disclaimer

This report has been produced by Lloyd's and Praedicat, Inc. for general information purposes only. While care has been taken in gathering the data and preparing the report, Lloyd's does not make any representations or warranties as to its accuracy or completeness and expressly excludes to the maximum extent permitted by law all those that might otherwise be implied.

Lloyd's accepts no responsibility or liability for any loss or damage of any nature occasioned to any person as a result of acting or refraining from acting as a result of, or in reliance on, any statement, fact, figure or expression of opinion or belief contained in this report. This report does not constitute advice of any kind.

This report addresses the use of publicly available information only. Lloyd's wishes to make it clear (i) that Lloyd's is committed to competition law compliance at all times, and (ii) that Lloyd's does not support or endorse any anti-competitive sharing of confidential or commercially sensitive data, and reserves the right to take appropriate action to prevent such an occurrence in the Lloyd's market.
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Executive summary</td>
<td>03</td>
</tr>
<tr>
<td>Introduction</td>
<td>05</td>
</tr>
<tr>
<td>Big data, catastrophes and the future of casualty</td>
<td>07</td>
</tr>
<tr>
<td>Science, insurance and the human risks of technological progress</td>
<td>07</td>
</tr>
<tr>
<td>The comparison with property catastrophe</td>
<td>09</td>
</tr>
<tr>
<td>Emerging risk oversight</td>
<td>10</td>
</tr>
<tr>
<td>The four challenges of emerging risks</td>
<td>10</td>
</tr>
<tr>
<td>Big data and emerging risks</td>
<td>11</td>
</tr>
<tr>
<td>Liability catastrophe modelling, underwriting and risk management</td>
<td>13</td>
</tr>
<tr>
<td>The example of bisphenol A</td>
<td>14</td>
</tr>
<tr>
<td>Conclusion</td>
<td>18</td>
</tr>
<tr>
<td>References</td>
<td>20</td>
</tr>
</tbody>
</table>
Executive summary

Accumulations of liability risk have the potential to send shockwaves through the insurance industry, and are one of the most complex exposure management challenges faced by insurers.

This report, commissioned by Lloyd’s, presents a new approach to such risk, developed by liability catastrophe modelling company Praedicat Inc., which uses big data analytics to improve understanding in this area.

Liability catastrophes – where the large-scale adverse effects of a product or substance result in an accumulation of claims across portfolios – can generate significant losses for insurers.

The impact of new technologies can be much more complex than anticipated, and it can take several years before the wider consequences are understood. When a liability catastrophe does occur, the products or business practices involved are usually discontinued and the companies that sold them may cease to exist. As a result, each new liability catastrophe is likely to happen in a different industry and in a different way. Claims data therefore typically cannot be used to predict the next liability catastrophe, and this presents a challenge for actuarial modelling.

The field of emerging risks research was established to address the need to identify and investigate early signs of potential liability catastrophes. However, emerging risk research is still a relatively nascent field when compared with actuarial science, and lacks a well-established methodology. The potential remains for liability risks to generate unexpected loss accumulations for insurers.

In this report, Praedicat sets out its new methodology, which uses big data to improve insurers’ understanding of liability risk. Using new technology to search and mine data from scientific research associated with potential liability risks, this approach estimates the probability of a general consensus being reached that exposure to a substance or product causes a particular form of injury. This is the critical threshold at which lawsuits become more likely to succeed; a liability catastrophe could emerge if a successful lawsuit gains traction and triggers mass litigation. This information is then overlaid on an insurer’s portfolio to identify potential accumulations of liability risk. The analysis can be used to develop quantitative estimates of mass litigation, allowing a liability catastrophe model to be built from the bottom up.

This approach helps insurers tackle the four key challenges associated with understanding emerging liability risks. These are:

1. **Identification**: recognising an emerging risk before it manifests as a loss or claim.
2. **Contextualisation**: comparing the size of risks, both relative to other emerging risks and to previous risks.
3. **Projection**: converting an emerging risk from the context in which it is identified, such as scientific literature, into an exposure-relevant context, such as companies and portfolios.
4. **Quantification**: estimating the expected loss from an emerging risk.

This report shows that big data innovations have the potential to create more robust liability risk management for insurers.

Further developments in liability catastrophe modelling using big data could offer insurers a means of managing liability accumulations while also identifying opportunities to increase exposure to certain risks where the accumulation is consistent with their risk appetite.
Introduction

The quantification of potential liability aggregations is one of the most complex exposure management challenges facing insurers. Lloyd's is investigating different methods that aim to reduce uncertainty in this area, and this report presents an approach developed by Praedicat, Inc. which seeks to harness the power of big data analytics to identify emerging liability issues.

The most effective risk transfer is expected to continue to rely on a combination of underwriting expertise and detailed analysis. Emerging technologies offer the promise of creating new tools designed to generate insight and, crucially, foresight by allowing us to test alternative scenarios and to stretch established thinking on plausible events.

This report presents one approach to the challenge of understanding emerging liability risks; Lloyd's hopes that it will inspire further contributions to drive innovation in the insurance industry.
Big data, catastrophes and the future of casualty

Many risks to health and property in modern life emerge from the interaction between people and businesses. On the one hand, there are everyday risks such as when a worker has an accident in the workplace or a customer is injured in a store. On the other hand, there are large-scale catastrophes such as bodily injury from toxic chemical exposures or property damage from accidents during energy production. The large-scale events can involve hundreds of thousands or even millions of injured parties, and tens to hundreds of businesses (see Box 1).

**Box 1: Large-scale liability events**

While not insured under casualty insurance due to longstanding exclusions, the largest settlement of litigation associated with businesses causing bodily injury would be the United States 1998 tobacco settlement, for which the number of injured parties would easily number in the millions. For instance, it is estimated that smoking causes 480,000 deaths a year in the United States. Beyond tobacco, it has been estimated that the number of asbestos claimants in the United States up to 2002 exceeded 700,000 and the number of defendants numbered over 8,000. When global claims are included, and updated to today, the number of injured parties could also exceed 1,000,000. Other large-scale bodily injury or property damage catastrophes with claimants in the tens of thousands include Agent Orange, and pharmaceutical litigation such as fen-phen, Vioxx and Zyprexa. In addition, property damage litigation involving agents such as polychlorinated biphenyls (PCBs), hexavalent chromium and methyl tert-butyl ether (MTBE) has resulted in billions of dollars in settlements and hundreds of defendants. Additional examples of large-scale catastrophes from business activities include the Deepwater Horizon and Exxon Valdez spills.

When adverse events occur, businesses are often liable for the financial consequences. To manage this risk, they often rely upon commercial general liability insurance. It was through the innovation of insurers that everyday risks such as house fires could be spread by pooling them, an idea that matured into actuarial science. This innovation is equally relevant to the everyday risks of liability faced by businesses today. But the large-scale events, which we call ‘liability catastrophes’, have been challenging for actuarial modelling, and this has led to many problems over the years for the insurance industry.

Actuarial modelling is based on projecting historical claims patterns into the future in order to price and manage risks. However, when a liability catastrophe occurs, the products or business practices involved are usually discontinued, and sometimes the companies that sold them cease to exist. The claims data they generated cannot predict the next liability catastrophe. The next catastrophic liability event is likely to happen in a different industry and in a different way. For this reason, in recent years, insurers have begun to invest more heavily in the ‘emerging risks’ function, which includes identification and investigation of early signs that a new liability catastrophe might emerge. Compared with actuarial science, emerging risk is still a relatively new field without a well-established methodology, and insurers continue to search for alternative methods to apply the insights that emerge from the activity.

Big data potentially offers assistance for the problem of liability catastrophes for insurers and a new set of methodologies for the study of emerging risks. While there is no rigorous definition of ‘big data’, the term can be used to refer to the remarkably rapid increase in digital data available and the technologies that are developing to efficiently organise, search and analyse it. The phenomenon is transforming many businesses across a wide spectrum of industries, from retail to healthcare. This report describes how, through the application of big data to peer-reviewed science, the insurance industry is developing innovative ways to identify emerging risks and model business-related, large-scale liability events. These innovations have the potential to create more robust risk management for insurers, and to also create opportunities for growth and innovation.

**Science, insurance and the human risks of technological progress**

Technological progress often involves the invention and commercialisation of technologies that interact with the physical environment in new ways. It is this novel interaction that drives the commercial value, whether it is the development of new technologies’ chemicals that have exciting properties like conductivity, durability, or strength; new methods of generating energy or extracting fuel from the ground; or new ways to store data and rapidly process it. These technologies interact with the world in more ways than previously anticipated and it can take years or even generations before the larger consequences, both positive and negative, of technological progress are understood. These consequences may include bodily injury, property damage or environmental damage.

Commercial general liability insurance is primarily intended to cover the day-to-day and idiosyncratic risks of people interacting with businesses at work and as consumers. As technology progresses, large-scale adverse effects – liability catastrophes – may initially be indistinguishable from these idiosyncratic events. Even when the evidence starts to mount that there is a deeper, more systemic problem, the technological innovators may not want to recognise the problem, and their insurers may not be alerted. Actuarial modelling
using historical claims data may not recognise the potentially systemic effects of technological progress, in which case insurers may not collect enough premium or hold enough capital to pay claims. The insurers and their clients may go out of business and the affected parties may not be compensated.

The story of asbestos fits this pattern of technological innovation followed by science describing human harm, and ultimately liability catastrophe. Asbestos was an extraordinarily versatile material resistant to fire and corrosion which could be incorporated into textiles, concrete and insulation, among other things. It was, however, later revealed to be a carcinogen, and the serious and sometimes fatal illnesses caused by its inhalation gave rise to casualty insurance losses in the United States of $85 billion. The problem for insurers was aggravated by the occurrence form as part of the standard general liability contract between an insurer and its client. With the occurrence form, a claim for indemnity is triggered on a particular insurance policy when the bodily injury ‘occurs’. With latent injuries and illnesses, the interpretation of the timing of occurrence is difficult, and many policies have been interpreted by courts to cover claims that emerged years later. As a result, the use of this policy form can cause latent exposures to accumulate for an insurer as an unreserved legacy. In the case of asbestos, when the litigation emerged many insurers were wholly unprepared.

While the world has changed since asbestos litigation emerged, the interplay between science, technology-driven innovation and risk which can drive the accumulation of exposure has not. Three prominent examples include:

- **Endocrine disruptors.** In recent years, scientists have become increasingly concerned about a new kind of bodily injury that has been investigated for its association with exposure to some plastics, pesticides and preservatives, among other commercial products. In laboratory animals, it appears that these exposures can cause alterations to the hormone system, which regulates physiological processes including growth, metabolism and reproduction. Some scientists have claimed that these exposures are the drivers of many modern public health issues, such as obesity and autism.

- **Hydraulic fracturing.** Developments in the technologies for extracting oil and gas from shale deposits, particularly hydraulic fracturing, have dramatically increased energy reserves and transformed global energy markets. At the same time, concerns have been raised that these extraction activities could lead to widespread property damage, including groundwater contamination and increased seismicity, and even to bodily injury.

- **Nanotechnology.** In recent years, the ability to engineer new materials and processes at the atomic or molecular level has led to an explosion of innovation that promises to be the next great technological revolution, transforming medicine, electronics, consumer products and more. At the same time, a growing body of scientific literature has raised concerns about the potential impact of these new materials on human bodies and the environment. For instance, carbon nanotubes appear to be one of the strongest materials ever fabricated, but they have been seen to cause health conditions in laboratory animals that look uncomfortably similar to the impacts of asbestos.

In addition to technology, science itself evolves. The ability to identify chemicals in ever-smaller concentrations, or the ability to demonstrate that a particular gene expression is caused by environmental exposures, can open up new avenues of liability. Indeed, the entire field of endocrine disruption research effectively did not exist 30 years ago, and the field of nanotoxicology, which studies the toxicity of nanomaterials, has only existed for around 20 years.

The range of phenomena described in this section is captured under the broad label of ‘emerging risks’ by

---

* Lloyd's defines an emerging risk as an issue that is perceived to be potentially significant but which may not be fully understood or allowed for in insurance terms and conditions, pricing, reserving or capital setting.

---

* Among the insurers with asbestos-related solvency issues were American Mutual, Home, Ideal Mutual, Integrity, Midland, Mission, Reliance, Texas Employers, Transit Casualty and Union Indemnity.
Emerging liability risks

The insurance industry. Emerging risks possess two key characteristics:

- the potential for large-scale losses for liability insurance, combined with limited or no history to rely upon for pricing and risk management; and
- a significant nexus with the scientific literature such that tracking the science around the risk provides a promising means through which casualty exposures might be identified and managed.

As the commercial landscape that led to asbestos litigation was transformed by this litigation, the insurance industry has been transformed as well. The primary response has been to limit coverage in various ways. For some companies and industries that have been exposed to liability catastrophes in the past, such as large chemical companies, the insurance industry developed alternatives to the occurrence form, such as ‘claims made’ policy forms or the ‘integrated occurrence’ form in Bermuda. In addition, a growing number of risks are excluded from coverage, with some insurance and reinsurance companies listing dozens of chemicals as standard exclusions in general liability coverage. For some other industries, such as large pharmaceutical companies, risk managers will use a small number of specialist insurance companies that understand complex risks and have a wide risk appetite.

Nonetheless, for the insurance industry as a whole, significant exposure to liability catastrophes is expected to remain:

- The United States general liability insurance market, where the risk is perceived to be the highest, continues to be the world’s largest commercial liability insurance market, and most of this risk continues to be written on the occurrence form. Outside the United States, much general liability is also written on the occurrence form – although higher risk exposures are likely to be written on a claims made wording.

- For smaller-scale, single industry events focused on large corporations, casualty insurers may have attempted to limit their exposure, but these adaptations have yet to be tested in the face of a large-scale liability catastrophe that involves a large number of companies simultaneously or multiple industries (such as food production) that have not been involved in this type of litigation in the past on a large scale.

In conclusion, the insurance industry response to liability catastrophes has largely been to limit the emerging risks that are a significant concern – however, despite these efforts, the potential for a large liability risk to reside within casualty emerging risk remains.

The comparison with property catastrophe

After a period of several decades of relatively limited hurricane or earthquake experience in the United States, the property insurance industry was unprepared for hurricanes Hugo and Andrew, or the Loma Prieta and Northridge earthquakes. Hurricane Andrew, a Category 5 hurricane which struck southern Florida in 1992, resulted in $27.5 billion in losses and led to the insolvencies of eight insurers.

In the years after Hurricane Andrew, industry practices were transformed by the rise of a new kind of risk modelling: property catastrophe (or ‘cat’) modelling. As a result, when Hurricane Katrina struck New Orleans in 2005, despite a record $62.2 billion insurance payout, the losses were spread across the global insurance and reinsurance market without significant industry dislocation.

The use of cat modelling is today part of the fabric of property insurance. Insurers and reinsurers manage accumulations of natural catastrophe risk by controlling exposures to geographical areas that are probabilistically exposed to common storm, earthquake and other natural hazard risks. Portfolios are balanced across geographies based on modelled risks. Reinsurance is priced using cat models, and rating agencies rely on the reporting of modelled output to evaluate capital adequacy. More recently, the growth of the insurance-linked securities market has been heavily dependent on catastrophe risk modelling.

Property catastrophe has some similarities to liability catastrophe. Like liability catastrophe, natural disasters are low probability, high consequence events that can threaten the insurance industry’s capital. In addition, since natural disasters often occur in places where they have not occurred in recent history, claims data is typically not helpful and other data must then be brought to bear. Despite these similarities, there are several reasons why property catastrophe modelling was developed earlier than liability catastrophe modelling:

- There are advanced sciences available for characterising the underlying risk. Property catastrophe modelling depends upon meteorology, seismology and other developed sciences to build the event set and model the risk. In contrast, liability catastrophes are complex social and economic phenomena and have, to date, lacked an obvious foundation for the event set.

- Liability catastrophes have two tails instead of one. Property cat modellers talk about ‘tail risk,’ which refers to the size of the loss at low probability. Managing the tail risk is the essence of property.
catastrophe risk management. Liability catastrophe has both this tail and another tail: the long tail of the occurrence form. Liability catastrophe involves a convergence of two types of tail risk, which requires modelling the size and timing of a complex multi-year event.

- **There is currently no way to scalably characterise the coordinates of prospective mass litigation risk in the currency of liability insurance, namely companies.** The development of geocoding facilitated the growth of property insurance as properties could be associated with latitude and longitude – the event set would describe the path of storms or the activation of faults in terms of this latitude and longitude and the associated hazard conditions. This development made the science actionable. In contrast, there currently exists no way to scalably map emerging liability risks to policyholders beyond the largely unscalable due diligence of traditional underwriting. Emerging risk groups within insurance companies routinely report struggling with the translation of information on emerging risks into portfolio exposures.

In the years since Hurricane Andrew, property insurance has seen significantly higher premium growth than liability insurance, and the rise of risk-linked insurance has seen significantly higher premium growth in the years since Hurricane Andrew, property insurance has seen significantly higher premium growth than liability insurance, and the rise of risk-linked insurance has seen significantly higher premium growth. Given the divergent paths of these two adjacent markets over the past two decades, it is reasonable to conclude that liability should instead strive to follow the path of property, and that liability catastrophe modelling could encourage better exposure management.

**Emerging Risk Oversight**

Awareness of the prospective nature of systemic risk in casualty has led some insurers to allocate responsibility for identifying and managing emerging risks to a dedicated Emerging Risk Group (ERG), while other insurers share this responsibility among different functions of the business. Today, the oversight of emerging risks is an integral piece of many insurers’ risk management strategy, and the creation of an ERG or equivalent function is sometimes required by rating agencies. In many ways, the activities of these individuals and teams are the foundation for liability catastrophe modelling.

The individuals and teams responsible for the oversight of emerging risks typically have the following responsibilities:

- **Identify emerging risks.** This is typically done by monitoring a diverse range of possible sources, such as claims, published science, blogs, websites, etc. ERG members will also participate in related industry activities.

- **Develop an emerging risk list.** An ERG, or others carrying out the emerging risk function, will often have the responsibility of maintaining a list of emerging risks. This activity includes identifying items to add, prioritising items for further action and removing items from the list.

- **Develop and disseminate underwriting strategies.** A limited number of risks are selected for the development of underwriting strategies. Those responsible for managing emerging risks will review scientific and trade literature about the risk, evaluate the insurer’s exposure, and then make recommendations for underwriting strategies.

While the resources applied to emerging risk research vary across insurers, many that consider it a competitive advantage will devote significant resources to an ERG or similar dedicated emerging risk function.

**The four challenges of emerging risks**

Individuals and teams undertaking emerging risk research must overcome four challenges for the resulting information to be considered sufficient to address the liability catastrophe problem. The four problems are identification, contextualisation, quantification and projection.

1. **Identification: is the ability to recognise an emerging risk prior to its manifestation as a loss or a claim.** This problem arises because of the prospective nature of emerging risks and the insurance industry’s historical reliance on retrospective information. Identification requires the selection of data sources that capture risks at an earlier stage in their evolution and a framework for understanding the evolution of risk. When writing casualty risks on an occurrence form, several years of advance warning are required to avoid accumulation prior to identification. The key question is whether the identification mechanism can be regarded as comprehensive. For the purpose of the development of a liability catastrophe modelling framework, a comprehensive identification mechanism is necessary.
2. Contextualisation: is the ability to compare the magnitude or size of risks, both relative to other emerging risks and to previous risks. Without contextualisation, human biases in risk assessment emerge that can undermine the emerging risk function. For instance, a common bias in emerging risk management is the recency effect, which is that the most recently identified emerging risk is the one that leads to the most alarm or concern. In general, without contextualisation, human biases tend to convert the emerging risk function into a risk avoidance activity. This can lead to the emerging risk function relying on exclusion as its primary underwriting strategy. The most common approach to contextualisation is the Realistic Disaster Scenario, which presents emerging risks as scenarios.

3. Projection: is the ability to convert an emerging risk from the context in which it is identified, such as scientific literature, into an exposure-relevant context, such as companies and portfolios. The problem of projection can lead to an overestimation of risk when, for instance, all different types of exposure associated with an emerging risk are assumed to lead to risk of bodily injury or property damage even when certain exposures may be safe. It can also lead to underestimation when only upstream activities such as production are identified as sources of risk, even when bodily injury and litigation are possible from downstream exposures. Projection is required for the efforts of those responsible for emerging risks oversight to be actionable.

4. Quantification: is the ability to estimate the expected loss from an emerging risk. It includes the ability to estimate the scale of the risk and the probability of its occurrence. Quantification is the ultimate goal of any emerging risk function as it potentially informs a wide range of critical insurance functions – from pricing to reserving to risk transfer – though it is a high bar and frequently controversial, particularly in the estimation of probabilities.

The greatest effort of research into emerging risks has been around identification, which often results in the creation of an emerging risk list. The creation of this list is often the core activity of ERGs. Various data sources are relied upon and it is typically a challenge to judge whether the results are comprehensive. Some progress has also been made in contextualisation, typically through the creation of a limited number of scenarios that serve as a proxy for the larger range of potential scenarios on the list. Projection is required for the development of underwriting strategies from the emerging risk list, and requires significant resources since exposure to an emerging risk is not a standard part of any submission. Typically, those with responsibility for emerging risks will limit projection efforts to a small number of risks each year, which are selected after periodically prioritising the emerging risk list. Quantification efforts have perhaps proven the most challenging. Overall, with an ad hoc and limited identification technology feeding a limited number of underwriting strategies and a limited number of risks profiled in depth, it is unlikely that these individuals and teams alone can solve the liability catastrophe problem.

The rise of emerging risk research in recent years has arguably been a tremendously positive development for the insurance industry in that it provides the foundation for a prospective approach to evaluating and managing risk, which is necessary to solve the problem of insuring technological progress. For this foundation to reach its potential, the field would benefit from tools to optimise the flow of data on emerging risks, convert the resulting information into quantified risk data, and transform today’s emerging risk avoidance strategy into a portfolio management strategy that is as likely to identify opportunities as risks, and which uses capital efficiently.

Big data and emerging risks

Big data is arguably changing the face of business in the 21st century. The most prominent set of applications of big data has been in profiling customers and improving customer experience as interactions with customers are increasingly tracked, stored and analysed. Through these big data applications, businesses are increasingly optimising their procedures for interacting with customers and targeting their marketing efforts. The second most prominent set of applications of big data has been risk and threat analysis (see Box 2).

Box 2: Top big data use cases

1. Customer Analytics (churn, segmentation etc.): 48%
2. Customer Experience Analytics: 45%
3. Risk Analysis: 37%
4. Threat Analysis: 30%
5. Regulatory Compliance Analysis: 26%
6. Campaign Optimisation: 26%
7. Location-based Targeting: 25%
8. Fraud Analysis: 22%
9. Brand Sentiment Analysis: 16%
10. Product Placement Optimisation: 16%
11. Other: 9%
12. Drug Discovery: 1%
The overwhelming source of data for most big data applications has been internally generated data, such as the clicks of customers interacting with a company website. The use of large amounts of internal data by an insurer, namely those relating to claims, is simply called actuarial science. Actuaries are arguably the original data scientists, even if their data has only recently been digital. However, recent advances in data science are allowing property and casualty insurers to use new data mining tools and visualisations more efficiently to extract information from their claims data – though, given the long history, these developments are arguably more incremental than transformative. In addition, for all but the largest insurers, the data being tapped into is not on the scale of big data and new methods of analysis may not be critical to success.

Where big data offers a transformative opportunity for commercial casualty insurers is in the mining of external data sources for emerging risk. At one level, this data is the source that the ERGs, underwriters and others responsible for emerging risks routinely rely upon today as reference sources. They are publicly available news feeds, scientific journals, trade journals, regulatory data sources, legal data sources and the sprawling web. However, it may be possible to move beyond using these as reference sources and convert them into data sources. Using these publicly available external data sources, it is then possible to begin to solve the four problems of emerging risk described previously: identification, contextualisation, quantification and projection.

First, big data offers the possibility of comprehensive identification through the appropriate selection of data sources that describe a class of emerging risks. For instance, if the class of emerging risks is latent bodily injury, such as with asbestos, and if some published science is necessary to support causation in a case of alleged latent bodily injury, as in the United States, then a corpus of biomedical science can serve as the basis of a database with comprehensive coverage of potential emerging risks in this class.

Once such a database is identified, how that data changes over time can be expected to carry signals about how the risks are evolving. It is therefore critical to have a model for the process by which the data source evolves. This could be a fully parameterised structural model, like a property catastrophe model. However, it could also be a framework for understanding the motivations of the actors that populate the database. For instance, in the case of the corpus of biomedical science, individual scientists identify possible sources of bodily injury and conduct research studies which are then published. Other scientists respond to the original studies and literature develops. Over time, as more studies accumulate around a specific bodily injury and a specific hypothesis that a particular exposure caused that bodily injury, the literature may evolve to the point where there is general acceptance of the hypothesis in the scientific community.

General acceptance of a bodily injury causation hypothesis is a critical threshold for a defendant and its insurer. At that point, the scientific evidence can be strong enough that a lawsuit against defendants accused of causing the bodily injury could be considered more likely to be successful (see Box 3 on causation and toxic tort litigation). The causation threshold provides an opportunity for the insurance industry data scientist to develop statistics that track hypotheses through the biomedical literature as they evolve towards general causation. In effect, scientists are motivated to find the same things as an ERG – namely the emerging risk that is causing the largest harm to the largest number of people – but start earlier and, in aggregate, bring millions of hours of work to the task. Furthermore, scientists are aware of each other’s activities and build on the work of their colleagues, driving the literature towards a conclusion on risk hypotheses. To have staff with responsibility for managing emerging risks reading this literature is operationally unviable, as it can involve sifting through hundreds of articles on thousands of hypotheses annually. A data scientist’s ability to scan, extract and aggregate large amounts of text-based data would be of significant value here.

Box 3: Scientific evidence in the US courts: implications for data science

For biomedical science to provide a predictive signal for liability insurers, it is necessary that the courts treat scientific evidence in a predictable fashion. This is true in the United States as a result of three US Supreme Court cases, which transformed American jurisprudence around scientific evidence in the 1990s. Daubert v. Merrell Dow Pharmaceuticals (509 US 579, 1993) established that the judge is to be the gatekeeper for scientific evidence and that the evidence must be reliable. The Supreme Court further clarified its intent in Joiner v. General Electric Co. and Kunho Tire Co. v. Carmichael, and the instructions were ultimately codified in Rule 702 of the Federal Rules of Evidence, which dictate evidentiary procedure in the US federal courts and many states. The change in science-driven tort litigation was ‘revolutionary’.

1 General causation is a necessary condition for success in latent bodily injury tort litigation, but not sufficient. Plaintiffs must also prove that their actual bodily injury was caused by the specific defendant, which is referred to as ‘specific causation’. Therefore, tracking science provides a probabilistic perspective on risk, but not a deterministic one.

1 An analyst using big data methodologies and analysing big data is sometimes referred to as a ‘data scientist’. 
The data scientist’s task is to statistically uncover the signal generated by the research activities of tens of thousands of scientists investigating injury hypotheses. Statistics can be generated by indexing and electronically monitoring biomedical research over time. These statistics would seek to measure quantitatively how new scientific hypotheses emerge and evolve. With this early identification, new science can be monitored before it matures to the point that it can support litigation. Metadata around the relative maturity of the literature could assist with contextualisation. With additional effort to describe the commercial context of the risks described by the science, the metadata can further support quantification.

The final piece is using big data to project information gained onto companies and portfolios. In the context of property catastrophe modelling, this has proved to be a considerable challenge. Many years elapsed before detailed addresses or latitude and longitude data was collected on property insurance submissions and then geocoded by cat modellers; this slowed the growth of the modelling. In the context of casualty, because policies are written to cover companies rather than a physical location, risk is characterised by companies rather than latitudes and longitudes. The challenge is therefore to apply emerging risk information to companies.

For external big data to be used effectively by liability insurers, it is likely that it will be necessary to manage internal policy data more effectively. Information about specific companies which have been underwritten should be captured in such a way as to permit linking to external data sources. More critically, liability insurance contract information, such as limits, attachment points and triggers, should be captured in insurer databases in a manner that permits internal exposure modelling. It is possible that a standard for liability insurance exposure data will evolve over time that will permit better communication with reinsurers about exposure and easier linkage to external data sources. With improved data management, the combination of internal and external data is likely to create an enormously valuable big data asset for insurers.

While there are challenges in organising the exposure data and linking it to external data, the quantity of external data on companies is significant. In recent years, an explosion of information about corporations has become publicly accessible. Much of this is made available by the corporations themselves, for example as marketing material or corporate responsibility reports, and these are made electronically available as documents on websites. These documents represent unstructured data that is potentially available to insurers’ data scientists for analysis, just as social media posts are analysed by companies for marketing. Additional information is collected, published and made searchable as digital data by a wide range of third parties, including governments and non-governmental organisations.

With a focused description of commercial exposures for emerging risks, it is now possible to mine digital data for information linking companies to these commercial exposures and therefore to the emerging risks themselves. The resulting data on companies linked to emerging risks can then be assembled into portfolio information and analysed for aggregation and clash. This data science strategy, if successful, effectively solves the projection problem.

**Liability catastrophe modelling, underwriting and risk management**

The full range of science-based risks that a commercial general liability insurer is exposed to is extensive and varied, including such potential risks as mobile telephones causing brain tumours, hydraulic fracturing causing earthquakes and nanomaterials causing lung damage. All of these risks represent sources of potential correlation and accumulation in an insurer’s portfolio. All of them also have a body of scientific and regulatory literature that can be mined using big data methods to characterise the science and quantitatively evaluate the direction of future research.

As each body of literature is mined and the exposures are overlaid on an insurer’s portfolio, the resulting analysis could be used to transform and scale the emerging risk function into liability catastrophe modelling. A number of assumptions will need to be made to transition from big data analysis to full probabilistic modelling, and the precise details of the process followed by Praedicat is proprietary information that cannot be disclosed in this report. The framework of the process is as follows:

1. **Identification:** Using algorithms to identify bodily injury and environmental property damage hypotheses in the scientific literature, the full range of risks can be identified, catalogued and tracked.

2. **Contextualisation:** While the risks are diverse, the means through which scientists establish causation are common to all of them. Each new hypothesis published in the scientific literature sits somewhere along the road to establishing causation, and it is possible to estimate whether causation will ever be established based on the current pace at which the research is progressing. Since science-based emerging risks are typically associated with a body of literature, this approach provides a means of contextualising them.

3. **Projection:** Using scientists as an ERG, the mining of scientific literature can provide a map of the relevant exposure settings for every science-based emerging
risk. With knowledge of the full range of commercial activities of companies in an insurer’s portfolio – data which can increasingly be obtained from public sources – it is possible to place companies on a ‘map’ of emerging risks. Some companies will be located on none of the emerging risks, and some will be located on many of them – and the map itself will vary according to what is on each insurer’s list of emerging risks. Assembling companies into a portfolio and then adding up the risks across companies provides a way to measure portfolio accumulation.

4. Quantification: The final step requires models in the same way that property catastrophe management requires models. Since science plays a critical role in the ability of a given lawsuit to gain traction and become a mass litigation, the scientific data used for contextualisation can be used to develop quantitative estimates of mass litigation, which can then be allocated to portfolios using the data used in projection.

The resulting liability catastrophe model is built from the underlying exposures up to the portfolio accumulations, and could provide a means of managing liability accumulations while also identifying opportunities to increase exposure to certain risks where the accumulation is consistent with the insurer’s risk appetite.

**The example of bisphenol A**

Bisphenol A (BPA) is a high-production chemical used to make hard plastics for a wide range of applications, from baby bottles to the canopies of fighter jets, and is used in the lining of cans and on thermal paper. It is also controversial because it is purportedly an endocrine disruptor. Its use in baby bottles, now largely discontinued, has been alleged to expose infants to a synthetic oestrogen, which is hypothesised to cause various bodily injuries ranging from breast cancer to obesity.46–49 Outside of this most prominent exposure to BPA, the scientific and regulatory literature has expressed concern about exposure in food supply, among retail workers handling sales receipts, and in a wide range of occupational settings.50, 51 However, the biomedical literature is overwhelmingly comprised of laboratory and animal studies, and most global regulatory agencies have largely concluded that the science is not compelling enough to warrant a ban.

While BPA has not resulted in bodily injury mass litigation to date, a prospective analysis of the risk based on the current body of science could conclude that there is cause for caution by insurers, particularly when writing on the occurrence form. If this fear is realised, BPA could expose many industries to potential future litigation, as illustrated in the accompanying Box 4.

Based on this information, insurers should consider their options for how best to handle this risk.

Should the insurance industry avoid BPA? Given the wide range of commercial exposures, from food and manufacturing to chemicals and retail, avoiding BPA would require avoiding a significant amount of commercial activity, which may not make sense for an emerging risk that has not yet resulted in significant litigation and may never do so.

Should the insurance industry exclude BPA? Many of the industry’s customers seek insurance for precisely this type of problem, namely the low likelihood of bodily injury litigation about their products. Excluding BPA makes the general liability product less valuable to the customer.

If avoiding or excluding the risk is undesirable, how is an insurer to manage the low probability of a large BPA exposure? Quantification of the risk using the data science approach outlined in the previous section can help address this question.

**Box 4: Some industries hypothesised by scientists as potentially exposing workers, customers or the public to bisphenol A**

- Bisphenol A manufacturer
- Epoxy resin manufacturer
- Polycarbonate resin manufacturer
- Thermal paper manufacturer
- Thermoplastic resin manufacturer
- Dental polymers manufacturer
- Household appliance manufacturer
- Plastic food packaging manufacturer
- Plastic bottle manufacturer
- Beverage can/canned beverage manufacturer
- Food can/canned food manufacturer
- Infant formula manufacturer
- Meat packing plant
- Retail store
Figure 1 represents a ground-up exceedance probability curve for BPA (all losses before the application of insurance coverage) generated based upon big data, using Praedicat’s liability catastrophe model. The top line represents estimated losses including defence costs, while the bottom line represents indemnity-only losses.

The methods that were used to generate the probabilities and magnitudes shown in this curve are experimental and should be regarded as illustrative of the potential for the methods rather than as a reliable indication of the risk of BPA. However, these experimental estimates indicate that BPA is ‘probably not’ the next asbestos, because the probability of mass litigation is estimated to be less than 25% and the probability of significant mass litigation at a size that could threaten the capital of major insurers is estimated to be low (depending on the litigation’s spread through the industry). Nonetheless, the risk is comparable with that of a major hurricane hitting some cities along the eastern seaboard of the United States in the next year.

The shape of the curve is driven by the difference in the science today across the various contexts in which exposure may occur and the different bodily injuries that may arise, which leads to a different likelihood of litigation emerging. For example, there is generally less science supporting a hypothesis of subsequent harm for adult exposures to BPA than for early childhood or prenatal exposures. Using the data analysis methods outlined in the previous section, it can be quantitatively shown that the likelihood that the body of science will come to support hypotheses that adult BPA exposure causes cancer, for example, is significantly lower than the likelihood that science will ultimately support the hypothesis that BPA exposure in utero causes developmental disabilities. In more general terms, using the big data approach one can sort the various exposure settings and harm hypotheses associated with an emerging risk, generate predictions for the trajectory of the science, and estimate the risk of future litigation in each setting. Using this methodology, a data scientist can identify particular associations of BPA exposure to bodily harm that have a low probability of ever being supported by science. By distinguishing across different types of exposure, an insurer can begin to develop a more tailored underwriting strategy for BPA.

Figure 1 represents a fully quantitative approach, which involves estimating a model of potential future litigation using predictions of future scientific progress towards understanding the risk of BPA. With the model output a liability insurer could begin to manage BPA risk in a similar way to how a property insurer manages exposure to hurricanes and earthquakes.
The use cases for a science-based liability catastrophe model underpinned by big data technologies include the following:

- **Underwriting and risk selection.** In this framework, a company is a collection of emerging risks which together could contribute to an accumulation of exposure to the insurer. An underwriter can use the quantitative information for risk selection. Given the probability and size of unlikely liability catastrophe events, an excess casualty underwriter can evaluate attachment points and develop an attachment point strategy.

- **Underwriting strategies.** With liability catastrophe information around emerging risks, underwriting and risk management teams can define risk appetites and identify industries and companies that contribute the most premium while providing the least exposure to emerging risks.

- **Accumulation management.** Insurers can estimate exposures to dozens, or even hundreds, of emerging risks, and develop targeted risk transfer strategies. Insurers will be able to stress test their portfolios against unlikely events of the ‘next asbestos’ type. In addition, insurers may be able to provide much more granular accumulation risk information to reinsurers, facilitating more competitive treaty reinsurance pricing, terms and conditions.

- **Reinsurer product development.** Liability catastrophe modelling could provide opportunities for the development of new reinsurance products, from named peril to clash. It may also provide opportunities for the development of insurance-linked securities.

More generally, with big data and technology emerging risk becomes a portfolio problem. Instead of the handful of issues able to be monitored by individuals responsible for emerging risk oversight – which is particularly constrained when this role exists alongside other responsibilities – insurers can use technology to track hundreds of issues, set risk appetites and track accumulations based on granular information about exposures. In turn, reinsurers can offer tailored products to address these accumulations, and both sides of a reinsurance transaction can have greater transparency about the risk. As a result, the risk of liability catastrophe could be spread prudently across the global insurance and reinsurance community.
Conclusion

This report has described a long-standing problem of insuring commercial liability in the face of technological and scientific change. In particular, science-based risk can lead to liability catastrophe, as seen with asbestos. The insurance industry, to date, has not had the ability to manage this risk. However, the explosion of external data describing bodily injury and property damage risks and about corporations in recent years could help solve this problem. While this report has focused on the opportunities that can arise out of mining biomedical literature, big data innovation is very likely to lead to opportunities relating to sources and methods beyond those described here.

Historically, there have been two major ways in which science has been incorporated into decision making. The first is 'scientific scepticism', which counsels that no action should be taken until scientific consensus is reached. This approach has been criticised on the grounds that the point at which consensus is reached can be too late for many injured parties. The competing approach is the 'precautionary principle', which counsels that the overall risk should be considered and steps to manage the risk should be taken before consensus is reached. When the steps taken involve bans and restrictions of technologies or products based on immature science, the precautionary principle can undermine promising new technologies that may in fact be safe.

There has been an irony in recent years: insurers have available to them a continuous range of methods to manage science-based risk, but have lacked the quantitative, forward-looking information about liability risks needed to use these methods effectively. In addition, they have frequently relied on scientific scepticism with respect to new technologies, waiting until scientists have reached consensus about a risk, and therefore exposed themselves to the risk of accumulating large latent exposures as a result. In contrast, regulators are increasingly relying on the precautionary principle as justification for banning or restricting various products and technologies.

With science-based liability cat modelling, insurers may be able to profitably incorporate the precautionary principle into their businesses. Big data makes identification, contextualisation, projection and quantification of casualty emerging risks possible; with this information at their disposal, casualty insurers can create the incentives needed for their clients to take appropriate precautionary actions. As a result, bans and exclusions could be employed more sparingly by regulators, and only when the science warrants. With big data about science informing insurer decision making, liability insurers can help make the world cleaner, healthier and safer for the public and their investors alike.
Emerging liability risks


