Government
Office for
Science

Blackett Review of High Impact Low Probability Risks

Government Office for Science

Foreword

By the Government Chief Scientific Adviser

As Government Chief Scientific Advisor, I am responsible for the quality of science-based evidence in government decision making. This report from the Government Office for Science is one of many that support that role.

We live in an uncertain world in which citizens are subject to numerous threats and hazards from within and outside the UK. One of the responsibilities of government is to help address these using the very best methodology for identifying, assessing and managing risk.

Cabinet Office and Ministry of Defence commissioned this Blackett Review. The review uses internationally regarded experts from outside government to present up to date and leading edge thinking on the best ways to approach identifying, assessing and managing high impact low probability risk. The report identifies several recommendations for further strengthening UK government's approaches to addressing these types of risk. It will also be of value to the wider Risk Management community.

Professor Sir John Beddington

Blackett panel members

Professor David Spiegelhalter

Trevor Maynard

Proffessor Judith Petts

Dr Willy Aspinall

Professor Patrick Godfrey

Professor Peter Taylor-Gooby

Dr Gordon Woo

Professor Dougal Goodman

Dr Andrew Coburn

Dr David Ball

Professor Philip Bond

Richard Anderson

Dr Scott Steedman

GO-Science Project Team

Dr Chris McFee

Dr Miles Elsden

Richard Meadows

Elizabeth Surkovic

Contents

1. Executive summary

At the request of the Ministry of Defence (MOD) and Cabinet Office (CO) this Blackett Review[1] was established to consider "High Impact Low Probability Risks". Representatives from HM Treasury also participated. The review has considered these issues in the specific context of those Departments and, for the Cabinet Office, from the perspective of the National Risk Assessment (NRA), though recommendations have emerged which it will be useful for all Departments to consider.

The Review has approached the issue with fresh thinking, considering the latest approaches to the risk management cycle. The recommendations build on existing practice, with an emphasis on refreshed thinking in a number of areas. The most notable over-arching factor in these recommendations is the repeated need for the inclusion of external experts and readiness to consider unlikely risks. Additionally, the report makes clear that behavioural matters and the role of social science in risk management needs to be enhanced.

Eleven recommendations have been made as a result of discussion at the Review meetings, seven are focused across all Government Departments and Agencies and four are specifically addressed to the Cabinet Office.

1. Government should make greater use of external experts to inform risk assumptions, judgements and analyses.

2. Government should continue to ensure the optimal and efficient balance of resources is used to address high impact low probability risks versus any other risk.

3. Government departments should enhance their warning systems to better detect early signs of low probability high impact risks as a mitigation measure to avoid strategic surprise. In doing this it should make best use of work and capabilities in government, academia and industry.

4. Government should review the means by which it can assess the effectiveness of its risk mitigation strategies.

5. Government should use probabilistic analysis, where it is available, in support of its risk management process to evaluate defined scenarios and inform decision making about significant individual risks.

---

[1] See Annex 1 for an explanation of the Blackett Review Process

6. Government should strengthen its mechanisms to review risks and include ‘Near Misses’ (where a significant risk almost materialises).

7. Government should work more closely with risk communication experts and behavioural scientists to develop both internal and external communication strategies.

8. Cabinet Office, working with other departments, should strengthen the scrutiny of the NRA by experts drawn from appropriate disciplines in the scientific, analytical and technical fields.

9. Cabinet Office should encourage government departments to develop and maintain a database of appropriate experts for the NRA risks they own, and ensure that it is kept under continual review.

10. Cabinet Office should encourage departmental risk owners to consider using supplementary approaches[2] to inform the likelihood and impact assessments for scenarios within the NRA process.

11. Cabinet Office should work with other government departments and experts to consider potentially linked or compounding[3] risks to inform contingency planning appropriately.

This report has brought together an expert view on this focused aspect of risk assessment. It aims to encapsulate the key issues and particularly highlights contemporary thinking in the field. The central text of this report is of broad interest, has wide applicability, and contains the main considerations that were debated.

---

[2] See pages 8-11 for examples of these approaches
[3] For an explanation of linked and compounding events see pages 10-11

2. Introduction

It can be challenging for the Government to be confident that it has used the best available evidence and expert judgement to identify, assess and prioritise, a representative range of challenging yet plausible risk scenarios to inform decisions on capability planning. This is particularly difficult for high impact low probability risks which, by their very nature, are difficult to identify or occur only infrequently.

Government needs to identify and mitigate risks arising from a wide range of non-malicious hazards and malicious threats. In this context, the overall challenge for Government is "the ongoing review and improvement of risk management[4]".

The National Risk Assessment (NRA) is one example of risk management in government. It aims to inform civil emergency contingency planning at both the national and local level. The NRA identifies, assesses and prioritises a range of representative risk scenarios that are considered challenging yet plausible manifestations of the wider risk they represent. These risks are then characterised on the basis of both likelihood and impact assessments. More information about the NRA is provided in annex 2.

For civil emergencies, the Civil Contingencies Act (CCA)[5] requires local emergency responders to assess risk and maintain Community Risk Registers for the specific areas and remit which they cover.

---

[4] http://www.hm-treasury.gov.uk/d/orange_book.pdf

[5] http://www.legislation.gov.uk/ukpga/2004/36/contents

3. Emerging risk identification

The identification of low probability risks, and the subsequent development of mitigation plans, is complicated by their rare or conjectural nature, and their potential for causing impacts beyond everyday experience.

Planning against potential strategic shocks from a high impact low probability risk often relies on ‘expert judgement’ to identify and provide advice. This is particularly true where quantitative data is sparse. Once a potential risk has been identified, it can be assessed and appropriate mitigation considered. However, it is important that policy makers have confidence in the process used to identify and assess risk, so that appropriate decisions can be made on its management.

It is proposed that there are key challenges that can prevent these risks from being correctly identified, which are identified below. In terms of fully embracing the possibility of those risks, there are several techniques which are widely used within the professional risk community to overcome these challenges and those conclude this section.

All risks can be separated into:

Those which most people would not necessarily identify and characterise, but about which many experts might have a reasonable understanding. For example: a storm surge overtopping the Thames Barrier;

Risks which are identified, but about which little is understood, for example: severe space weather; or

Risks which most, if not all, experts would struggle to identify.

Where the nature of a risk is such that experts can identify it, it may be sufficient that a wide range of experts are consulted[6]. It is preferable that this should be done in a structured way to create a suitably judicious consensus[7] and to limit communication issues that may arise between different communities using different jargon or different values. The majority of high impact low probability risks can be identified, and their likelihoods characterised in this way.

It should be noted that the two strands of a risk characterisation, impact and likelihood, require different types of analysis. For example, it may often be easier to characterise the impact of an event than its likelihood, such as the impact of your wallet being stolen against working out the numerical likelihood of it happening.

There will also be occasions when experts struggle to identify particular risks, for example where a risk falls between (or beyond) traditional expert boundaries.

---

[6] This might include, but not be limited to: social and physical scientists, statisticians, analysts, medical professionals, engineers, lawyers, actuaries etc.

[7] E.g. Aspinall, W. (2010) ‘A route to more tractable expert advice’, Nature, 463, 294-295.

Experts are also often reluctant to consider scenarios or risks which fall outside their ‘comfort zones’. This can lead to the rejection of potential pointers or evidence simply because such indicators challenge current technical or scientific understanding, or lie outside normal situational awareness. This ‘cognitive dissonance’ is partly a cultural problem. People are often unwilling to give credence to improbable notions specifically because their professional or social community consider them too improbable.

In addition, if a problem is thought too complex, there is the danger that organisations will simply ignore it. This may be further complicated by the different emphases and priorities of experts from different backgrounds or disciplinary traditions.

There may also be surprises within fields of expertise – for example, the magnitude of the recent Japanese Earthquake was not thought to be possible on that segment of the fault line.[8]

More generally, there is often a lack of imagination when considering high impact low probability risks. To counter this, tools are needed to enable risk assessors to ‘imagine’ these risks and expand the boundaries of their mental models.

A number of formal methods have been developed that seek to do this. These techniques[9] can be used to incorporate information which is counter to currently held beliefs and assumptions. If a particular piece of ‘surprising information’ is received from a number of different sources, or from more trusted sources, then the risk assessor will need to consider how to judge the plausibility of the evidence and whether or not it should be used. There are a number of questions that could be asked to probe the importance of new or unanticipated evidence:

- How independent is each information source, and how reliable?
- Is this information really ‘surprising’?
- Is it likely that the source could be wrong?
- When considering all of the ‘surprising information’ received, is it contradictory; or does it partially/fully confirm what other sources have been saying?

One technique for capturing low probability risks is to use a very wide range of experts from many different disciplines with different ways of thinking and create an environment where experts are encouraged to express their views candidly. Such a group can then provide the collective ‘imagination’ which challenges established internal thinking.

‘Red teaming’[10] is another approach often used in organisations. This methodology uses the group’s experience and knowledge in problem solving to learn, and discover unidentified risks (collectively known as a heuristic approach). This can help ensure an organization is overall more robust.

---

[8] Kerr, Richard A, “New Work Reinforces Megaquake’s Harsh Lessons in Geoscience” Science Vol 332, p 911

9 For example, see Stirling A, ‘Keep it complex’, Nature vol. 468, p 1029-1031

[10] ‘Red teaming’ is a military concept in which a group is formed to try and assess vulnerabilities and limitations of systems or structures. It is used to reveal weaknesses in military readiness.

Section recommendations:

- Government should make greater use of external experts to inform risk assumptions, judgements and analyses.

- Government departments should enhance their warning systems to better detect early signs of low probability high impact risks as a mitigation measure to avoid strategic surprise. In doing this it should make best use of work and capabilities in government, academia and industry.

- Cabinet Office, working with other departments, should strengthen the scrutiny of the NRA by experts drawn from appropriate disciplines in the scientific, analytical and technical fields.

- Cabinet Office should encourage government departments to develop and maintain a database of appropriate experts for the NRA risks they own, and ensure that it is kept under continual review.

4. Assessing and representing risk

4.1 Methods of Risk Assessment

The purpose of any risk assessment is to identify how likely events of concern are to occur and to assess potential losses and impacts. This provides an objective foundation for deciding appropriate prevention or mitigation strategies. There are three main ways of undertaking a risk assessment:

1. a heuristic method to derive a traffic light score (or other non-numeric measure);
2. a deterministic scenario basis; or
3. probabilistically.

Heuristic

The heuristic method refers to a subjective score for a given risk where relevant stakeholders or experts provide their judgement. This method is generally used when there is little measurable information but opinions are still available. Deterministic or probabilistic approaches are usually preferred where possible.

Deterministic

The deterministic approach is the simpler of the two numeric approaches. It requires an evaluation of the likelihood of the impact or losses made from a single specific scenario, quite often chosen to represent some form of defined extreme or 'worst case' scenario. The objective is to represent a range of impacts up to the level of this 'worst case' scenario. For example, the maximum passenger loss from a single-plane air crash could deterministically be set at the capacity of an Airbus A380, the largest aircraft currently in service.

For natural hazards, practical deterministic bounds are hard to establish in an absolute sense. The height of a tsunami is one obvious example that is difficult to predict and, locally, can depend on many factors. Consequently deterministic assessments will almost always rely on expert judgement and specific assumptions, drawing on available evidence and data.

Probabilistic

A probabilistic risk assessment evaluates the relative likelihood of different levels of loss or damage for a number of individual scenarios taken from a range of possible events. Decisions can then be taken whether to act based on a set level of risk tolerance. A decision to take no action may be made if the probability of an event is considered acceptably low or action taken to maximise the probability-weighted expected utility[11].

---

[11] Utility can be defined as the total satisfaction received from consuming a good or service

One key weakness of deterministic assessments is that they are not readily comparable across risks. For example the impact or losses from a 'worst case' flood scenario is difficult to compare to a 'worst case' tornado scenario, as the likelihood and impact will certainly be quite different. Therefore comparisons between deterministic scenarios will not be on a consistent basis as both the impact and likelihood for scenarios will vary. However, in practice, risk managers routinely compare several deterministic scenarios and make decisions on that basis.

Deterministic scenarios are often used to validate the results of probabilistic models, so that undue reliance is not placed on complex mathematical models often requiring complex probabilistic analysis.

For example, in the insurance industry, Lloyd's of London operates a system whereby 'Realistic Disaster Scenarios' are used to stress-test both individual syndicates and the whole market to see how they stand up to chains of accumulated extreme cases. These scenarios include man made threats (such as a terrorist attack on the US) and natural hazards (such as flooding).

4.2 Quantifying uncertainty

Uncertainty in the decision making process

The aversion of individuals to be shown to be 'wrong' can skew any individual risk assessment and lead to inconsistencies. In particular, experts can have an aversion to making decisions when there is incomplete information or where difficult decisions need to be made quickly. There is therefore a danger that some experts will be unwilling to provide information to aid a decision until they consider they have all the facts. This may lead them to defer to colleagues or not contribute to discussions. Problems can then arise when vital binary decisions have to be made, for instance the decision to evacuate a population threatened by hurricane, or to impose a travel ban in a pandemic. Experts may therefore wish to suspend scientific judgment whilst they wait for the situation to develop further and more information accrues before choosing whether or not to act when a decision can have drastic consequences for huge numbers of people. Reluctance by experts to commit can therefore engender 'epistemic risk'[12].

There are advanced, mathematically based methods for pooling scientific advice to minimize this epistemic risk, to quantify uncertainty, and to improve decision making. One such approach is given in annex 2.

Systems based approach

There are situations where one hazard event can affect the likelihood or impact of another, leading to compound or joint risks. *Linked risks* can arise from separate effects due to a common causal event, while *compounding risks* can occur where separate hazard events combine to increase the overall impact on a single area. The Fukushima nuclear incident

---

[12] Epistemic risk is that arising from lack of knowledge or an agreed understanding of an issue.

provides an example of how complex dependencies can escalate the severity of outcome from an initial event.

Linked risks have separate impacts that are connected to the same causal event. For instance, in the summer of 2010, an atmospheric blocking pattern over Russia gave rise to conditions of severe drought and subsequent forest fires in Russia, while simultaneously contributing to extreme rainfall that resulted in unprecedented floods in Pakistan.[13]

Compounding risks occur when there is a direct causal link between a series of hazard events, or where there are entirely separate but coincidental hazard events. In the second case the impact of each event at a particular site or locality can combine with separate events to increase the overall impact. For example, the eruption of Mount Pinatubo in 1991 in the Philippines coincided with a tropical storm. Heavy rain from the storm compounded the effects of ash fall from the volcano, leading to additional deaths from collapsed roofs.

Risk analysis therefore needs to take account of the possibility of linked or compounding and consider a systems level view of multiple hazards and risks. To ensure that key risks are identified, analysis should concentrate on the full range of direct causal events that might produce a series of linked effects and risks, and consider events related to a locality or part of a system with the ability to compound.

Such a systems based approach can be a powerful technique for assessing risk. Substantial amounts of work have been undertaken to identify, assess and understand risk in complex industries using these types of approaches. For example, in the nuclear power industry the concept of ALARP (As Low As Reasonably Possible) is used to define a threshold beyond which available resources could be more effectively used on other risks.

The systems based approach (for example, as outlined by the Royal Academy of Engineering[14]) helps identify links where a minor incident in one sector can have major implications for other sectors. Even when such risks are recognized, the true impact is often underestimated as complex systems can be robust to one type of failure, but vulnerable to a different failure mode. For example, communication systems may in general be very resistant to the random loss of individual parts of that network, but may be vulnerable to the loss of one or two key network nodes.

The conditions under which complex interdependencies can lead to a small impact creating the potential for cascading failures[15] that result in disproportionate consequences are not well understood. This is an area that needs further research[16].

Expressing deeper uncertainties

In many situations the likelihood of an event is uncertain, due to limited evidence, scientific disagreement, or acknowledged ignorance about the underlying processes. It is therefore

---

[13] http://www.earthzine.org/2011/06/01/did-the-russian-heat-wave-trigger-the-pakistan-heavy-rain-in-2010/

[14] http://www.raeng.org.uk/education/vps/pdf/RAE_Systems_Report.pdf

[15] http://mathaware.org/mam/2011/essays/complexsystemsHines.pdf

[16] England J, Blockley, DI & Agarwal, J. 'The Vulnerability of Structures to Unforeseen Events', Computers and Structures, 86, (pp. 1042-1051), 2008. 10.1016/j.compstruc.2007.05.039

important to be able to communicate these deeper uncertainties and not to suggest that the probabilities are 'known'. Some suggestions for describing such uncertainty are summarised below.

'Italian flags'

Complex systems contain interdependencies that are difficult to anticipate. There is a need to recognise, identify and express the extent to which we genuinely do not know something. Methods based on interval probability analysis and evidence theory can be used to express a degree of evidence for a proposition (green), against (red) and don't know (white) as an Italian Flag[17],[18] See annex 4

Quality of evidence – healthcare interventions

Annex 5 outlines the approach adopted by the Cochrane foundation[19]. They have developed a technique to reach a judgment on the overall quality of a collection of risk assessments. All reviews are accompanied by a 'summary of findings' table which presents estimates and gives a rating of confidence in the evidence behind a treatment's benefits and consequences.

Quality of evidence – climate science

Annex 6 summarises the approach promoted by the Intergovernmental Panel for Climate Change (IPCC) in communicating their confidence in the conclusions of their assessment reports. This is a highly controversial area in which substantial effort has been made to harmonise the expression of uncertainties.

The Renn approach

Finally, the Renn approach, described below in section 4.3, uses nine indicators (which include likelihood and impact) to provide a more in-depth representation of the risk, which can then be characterised into six risk classes.

Summary

In summary, effective assessment of emerging risks is for a common problem across both industry and government. There is no simple solution. Common to the effectiveness of the approaches outlined is the following good practice:

- adopting a clear definition of an emerging risk, including medium and longer term issues;

---

[17] Blockley D I, Godfrey P S, (2000), Doing it Differently, Thomas Telford London,

[18] Blockley D and Godfrey P (2007) "Integrating soft and hard risks" Int. J. Risk Assessment and Management Vol. 7, Nos. 6/7, 798- 802

[19] The Cochrane Foundation is a major international organisation that conducts systematic reviews of medical interventions (see www.cochrane.org)

- recognising at senior management levels that emerging risks need explicit management and needs to be considered from a multidisciplinary viewpoint;
- keeping a track of identified risks;
- considering a broad range of scenarios;
- making allowance for behavioural factors; and
- sharing risk identification.

4.3 Representing risk

The National Risk Register (NRR) (the public-facing version of the National Risk Assessment (NRA)), categorises risks scenarios into types of risk that are represented relative to each other on an impact/likelihood matrix. This matrix purposefully excludes numbers or scales to avoid a false sense of accuracy. The relative positioning of these risk types is defined by the underpinning NRA assessment.

The register illustrates the types of emergency that can happen in the country as a whole. These are designed to help readers identify and visualise the risks. The National Risk Register includes risks such as an outbreak of a pandemic human disease; coastal flooding; an attack on crowded places or on transport; a major industrial accident, or an outbreak of an animal disease.

An alternative way to characterise and represent risk was developed by Ortwin Renn using nine indicators (which include likelihood and impact) such as 'reversibility' or 'persistence' to provide a more in-depth representation of the risk. For example, environmental radiation contamination has a long term impact that is harder to remediate than the damage from a conventional explosion. Renn's system suggests these risks should be treated differently, both in evaluation of the impact and in management strategies.

Renn distilled these nine criteria into six genuine risk classes, and assigned them names from Greek mythology. Each class has a different type of risk associated with it, for instance a "Medusa" risk is low probability and low damage but nonetheless a cause of considerable concern for people as they are not fully understood – mobile phone usage is one example. These are discussed in more detail in annex 7 and an example matrix is below at Figure 2.

Each risk class entails different mitigation strategies. For example: science based; precautionary; or a strong focus on risk communication. It is claimed that this descriptive approach can provide more information and better demonstrate the uncertainty associated with a particular risk.

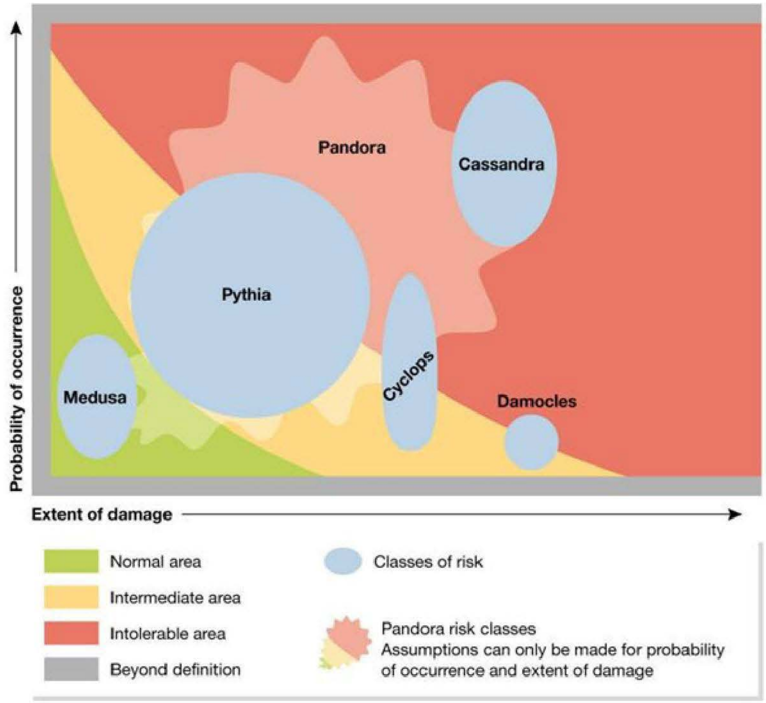


Figure 2, a descriptive risk approach (see annex 7)

Section Recommendations:

- Government should use probabilistic analysis, where it is available, in support of its risk management process to evaluate defined scenarios and inform decision making about significant individual risks.

- Cabinet Office should encourage departmental risk owners to consider using supplementary approaches to inform the likelihood and impact assessments for scenarios within the NRA process.

- Cabinet Office should work with other government departments and experts to consider potentially linked or compounding risks to inform contingency planning appropriately.

5. Managing Risk

5.1 Moving from assessment to management

It is accepted within government that the assessment and representation of risk is a separate process to that of managing risk. Managing risk focuses on how best to respond to the risk. This may include the 'do nothing' option, but more likely involves measures to mitigate the likelihood or impact (or both) of the risk, and to manage any residual risk.

This distinction between assessing and managing risk requires a different type of expert advice. For instance, a statistician and a meteorologist might be suitable experts to contribute to a risk assessment of a storm surge. However, development of mitigation measures would need to reflect the wider context, which might include ethical, social and political sensitivities of the various options, such as of the need for enforced relocation of homes. Pre-event mitigation decisions also need to take account of wider issues such as the economic or ecological impact of implementation, such as a sea defences.

This distinction between assessment and management can often create conflicting priorities. Those responsible for providing the risk assessment will have a tendency to want all the facts before reaching a decision (as discussed in section 3), whereas a decision-maker will want to have mitigation measures in place as soon as possible. This tension needs to be managed effectively by ensuring risk assessment is timely and based on the latest evidence. This will give the optimal time to build, develop and maintain mitigation capabilities.

In some circumstances, for instance where there is insufficient evidence to support a robust assessment of a risk in a timely manner, risk owners need to directly commission research to address key knowledge gaps.

One common approach used in industry to share risk liability is the concept of risk pooling. The cost of sizeable claims, that may unsupportable by a single risk owner, is distributed as a number of smaller liabilities across multiple entities. Specific risk pools generally cover claims in the same category, such as fire or flood, and in a specific geographic area. They can be between a group of companies or as public/private risk pools where some of the risk is underwritten by Government.

Risk transference is another option where the burden of the risk is transferred to a third party though, for example, legislation, regulation or insurance (including issuing catastrophe bonds).

Neither risk pooling nor risk transference are considered further here as explicit tools for dealing with high impact low probability government risks, though they may form part of the mitigation strategy for individual risks.

Care must be taken when managing risk to ensure scarce resources are targeted for maximum effect. Before judging what (if any) resources are needed to manage a specific risk, senior decision makers need to be confident that the risk assessment, both in terms of its impact and likelihood, is as reliable as is feasible.

Risk assessment of potential strategic shocks is used to drive planning to reduce the likelihood, or to mitigate the impact of such risks. From a government perspective, planning to mitigate these risks must be conscious of the time frame considered by the National Risk Register (the next five years), but also needs to consider longer term adaptability over the next twenty to thirty years. For example, to mitigate against climate change.

5.2 Risk Mitigation

Measures aimed at minimizing the impact of high impact risks may be expensive and, as with all risks, may never be utilized if the risk does not occur. Decisions on mitigation strategies including those for high impact low probability risks, need to be made on as robust a basis as practicable. This should include a cost benefit analysis to inform decisions on the level of risk acceptance and allocation of resources.

There may be circumstances where the cost of installing mitigation measures outweighs the cost associated with any damage caused. This is shown graphically in figure 3 below.

The red line represents the cost of repairing the damage as the size/impact of the event increases, the blue line the cost of installing mitigation measures. The point at which the red and blue lines cross is the point at which it no longer becomes economically reasonable to install mitigation measures.

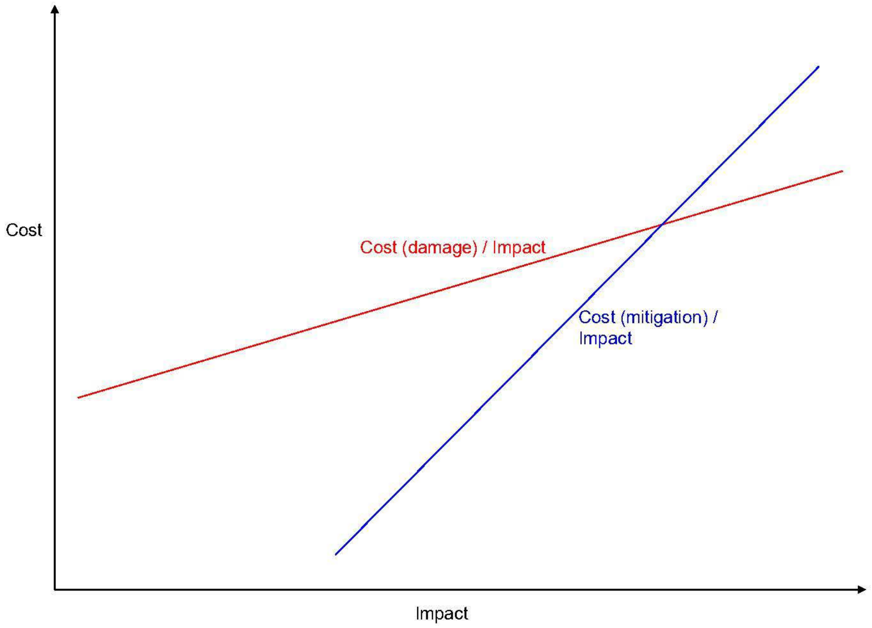


Figure 3 – damage cost versus mitigation cost

It may also be necessary to take into account other economic factors such as:

- the mitigation measure may not completely mitigate the risk and therefore it would be necessary to account for residual damage costs;

- mitigation measures may eliminate damage from smaller events up to its design threshold but may also reduce the damage from a larger event.

It is useful to note that a simple analysis of one potential scenario will not be universally applicable – different risks will require different analyses, and indeed this approach may not apply to some risks. It is also worth noting that it may be difficult to quantify, in advance, the costs of the damage caused by an event and of implementing the mitigation measure.

Of course there are more factors at play than pure material cost; social, ethical and political considerations will need to be taken into account when making the decision to implement mitigation measures. Economics will form only one part of this decision, especially when considering the allocation of scarce resources. There may even be secondary benefits to implementing mitigation measure such as the longevity of civil engineering structures or infrastructure as a result of mitigation against natural hazards such as high winds or earthquakes.

5.3 Resource Allocation

'Reverse stress testing' can be a useful tool in addressing the issues of where to allocate scarce resources by identifying the point at which the current mitigation begins to fail. Senior decision-makers can then make a judgment on whether it is prudent to increase resources allocated to those risks, or whether other risks have a more pressing need.

By definition, low likelihood risks rarely occur. There is therefore the danger that once an initial full risk assessment has been undertaken and mitigation planned, an organisation considers the problem solved and it can be 'surprised' when the risk occurs some years later. Pre-warning or new evidence that such a risk could be about to occur would clearly be extremely valuable to reduce any potential 'surprise'. For some risks, such as a volcanic eruption, it may be possible to identify a number of small events, (such as tremors in the vicinity of the volcano) that signify that a major incident might be about to happen.

However, for many risks it must be recognised that no prior warning will be available. For example, no indicators have yet been found which could provide <u>reliable</u> warning of an impending earthquake. Nevertheless, in situations where meaningful 'precursors' or cues do exist, there should be robust procedures in place to ensure these indicators are monitored and acted upon.

Similarly, because there is little opportunity to see how effectively plans perform under the pressure of a real incident, it is important that appropriate testing and exercising is regularly undertaken.

A 'Near Miss' (whereby a significant risk almost materialises) can also be used to assess how current plans would have coped if the risk had actually occurred. Similarly events, which share some of the impacts of the original risk, can also be used to assess existing plans.

Section Recommendations:

- Government should continue to ensure the optimal balance of resources is used to address high impact low probability risks versus any other risk.
- Government should strengthen its mechanisms to review risks and include ‘Near Misses’, (whereby a significant risk almost materialises).

6. Communicating the risk

As with all communication, credibility is key when communicating complex ideas such as risk.

The communication of risks to senior decision makers in a timely manner is important to ensure risk can be considered within the context of overall resource availability and wider priorities. Discussion of risks should also be framed so that senior decision makers can feel comfortable in being able to challenge ‘expert’ assessments.

Effective communication of risk is an active area of work that government could make better use of, drawing on the experience gained in large private sector organisations. It is very important to use the correct ‘vocabulary’ when explaining risk and to use the language and context which the organisation normally works. ‘Learning by doing’ (such as working through scenarios or exercises) can also be very effective for increasing the risk awareness of senior individuals.

It is important to try and identify ways in which a lack of knowledge can be articulated alongside any uncertainty. For example, the lack of a particular piece of evidence will add to the overall uncertainty within a risk assessment. Absence of this specific piece of evidence can often be straightforward to communicate, but getting across the implications for the uncertainty in the risk itself may be more challenging.

A large number of behavioural and psychological studies have been undertaken in academia to understand how people approach and understand risk. Many people have difficulty in assessing the longer term implications of a high impact risk (for example, the geopolitical impact of 9/11) and have a tendency to focus on lower impact, more probable risks. This means the public can sometimes under-estimate risks. People also tend to focus on short term outcomes; often demanding immediate action from authorities following and event without full consideration of the longer term implications.

In trying to identify effective communication methods, it is important to recognise that the public ‘users’ of the risk assessments and risk owners do not usually think in the same way as expert risk assessors. This doesn’t mean that people do not understand, or are ignorant of risks, but that they simply use a different language to express them.

However, while the public often understands the risks, they may not always apply the logical cost-benefit analysis which risk analysts use when dealing with them. For many high impact risks we do not understand what the public actually expects in a situation, or how tolerant they may be of ‘abnormal’ risks during a crisis. Exploration of these issues can be used to help inform how these types of risks are communicated to the public.

One of the key drivers in a major crisis is to avoid collateral impacts and restore ‘normality’ as quickly as possible. Managing public expectations of the speed of recovery and what a return to normality will look like, is extremely important, for example impatience with prolonged power cuts following extreme weather. This may be compounded by a lack of understanding within government of the public expectations in a crisis. The use of feedback mechanisms, such as public surveys on the issues of public concern around specific risks, are key to ensuring that risk communication is a two-way process and that the communication matches public concerns.

The Chernobyl explosion in 1986 was a useful illustration of the need to account for public understanding and reaction in managing a crisis. Media coverage amplified fears about radiation and this is thought to have contributed to a sense of hopelessness in the population in the Ukraine and neighbouring countries. This in turn has been linked to high rates of alcoholism and smoking — factors that have had a hugely bigger health impact than the radiation itself.[20]

Similarly, the public has a much stronger reaction to a group of deaths resulting from a single event than for a similar number of people dying in a number of unrelated events. Consider the impact on the public of the M40 Minibus Crash in 1993 (13 dead – including 12 children) which eventually lead to calls for compulsory seatbelts in minibuses and coaches,[21] compared to the much larger annual number of road death, for example 5,217 deaths on British roads in 1990 (a rate of 14.3 per day).

Finally, trust is vital. Existing government guidance[22] rightly stresses the importance of: providing a trusted source of consistent information; being transparent about what is known and unknown; clearly distinguishing 'worst case scenarios' and what is expected; providing regular updates; and giving clear guidance as to suitable action for people to take.

---

[20] Peplow, M. 'Chernobyl's legacy', Nature 471, 562-565 (2011) | doi:10.1038/471562a

[21] http://www.rospa.com/news/releases/detail/default.aspx?id=62

[22] For example Treasury Orange Book Annex http://www.hm-treasury.gov.uk/d/managingrisks_appraisal220705.pdf

Section Recommendation:

- Government should work more closely with risk communication experts and behavioural scientists to develop both internal and external communication strategies.

7. Conclusions

Both industry and government face the same difficult challenge in assessing and understanding high impact low probability risks. During this Blackett Review, a number of experts in risk identification, analysis and communication, from both industry and academic, have had the opportunity to hear about and discuss high impact low probability risks from a Government perspective.

The Blackett panel have provided a peer review role, challenging existing practices in a number of areas. They have identified eleven recommendations where Government should consider using complementary techniques or approaches to strengthen existing risk identification and management processes. Government will need to assess each recommendation in more detail, but consideration the recommendations of this Blackett Review will help bring additional robustness to the Government's approach to high impact low probability risks

Annexes

Annex I

The Blackett Review Process

The Government Chief Scientific Advisor (GCSA), Sir John Beddington, has established a process for government to engage with academia and industry to answer specific scientific and/or technical questions primarily in the security domain. These Blackett Reviews provide fresh, multi-disciplinary thinking in a specific area. In each review, a small panel of 10-12 experts are tasked with answering a well defined question or set of questions of relevance to a challenging technical problem.

In autumn 2010 the GCSA convened a group to address the question *"How can we ensure that we minimise strategic surprises from high impact low probability risks"*. The panel considered how Government could best identified, assess, communicate and quantify the inherent uncertainty in these types of risk.

The panel met three times over a nine month period and provided an independent scientific review of the current risk management approach of a number of departments, primarily the MOD and the Cabinet Office. It is written by the Government Office for Science based on the discussions at the meetings and additional information provided by panel members.

Annex 2 – The National Risk Assessment

For civil emergencies, the Civil Contingencies Act (CCA) requires local emergency responders to assess risk and maintain Community Risk Registers for the specific areas which they cover. The National Risk Assessment (NRA) process is designed to help inform and assist local risk assessments. The NRA is also designed to supplement local emergency preparedness by acting as a prioritisation tool which can inform the Government's decisions about capability planning. For these reasons, the NRA is not intended to be an exhaustive register of all civil emergencies. Instead it aims to identify and assess a representative range of risks which can be communicated to local emergency planners to inform and assist local assessments and used to enable the Government to optimally spend its resources on preparing for civil emergencies.

The NRA process

Each year, Government departments identify any new potential risks, assess these and reassess existing ones to identify potentially challenging yet plausible manifestations of risks (reasonable worst case scenarios) for inclusion in the NRA prioritisation process. It is expected that they will consult a range of experts in doing this, including their Chief Scientific Advisor and Scientific Advisory Committees. For risk scenarios to be included in the NRA process they must meet the:

a)Civil Contingencies Act's criteria for an emergency (e.g. threaten or seriously damage human welfare, the environment or the security of the UK);
b)pre-defined likelihood thresholds (scenarios deemed to be extremely unlikely are excluded);
c)pre-defined impact thresholds (scenarios that can be managed by local responders or within existing resources are excluded); and
d)the criteria that, adding it will provide: i) new insights on the common consequences that government should plan for or ii) identify high impact risks where a risk-based decision on whether specific planning is required is needed.

Suggestions that do not make these criteria or risks for which there is insufficient evidence are kept on a "risks under review list" which is reviewed annually. Defining whether risks meet the criteria outlined above, is usually an iterative process.

Selected representative risks, are compared using a likelihood and impact matrix which uses defined and logarithmic scales to ensure consistency, avoid a false sense of accuracy and enable relative comparisons of very different risks. The annual assessment of the relative positioning of risks on the NRA matrix is used to identify those risks that are:

a)high impact and high probability and require specific planning;

b) high impact and low probability and may require a more proportionate approach using existing resources and mechanisms (Ministers are asked to make a risk-based judgement on these risks);
c) “moderate” risks with lower impacts and probabilities – the common consequences from which can be used to inform planning; and
d) “low risks” with lower impacts that common consequence planning will cover.

Annex 3 - Epistemic risk

The empirical sciences have always placed great emphasis on rigour and engendered a culture of restraint when it comes to inference and inductive logic. The philosopher David Hume advised: “there is a degree of doubt, caution, and modesty, which, in all kinds of scrutiny and decision, ought for ever to accompany a just reasoner”. Most scientists dread being proved

wrong in any scientific assertion they make, in claiming support for a hypothesis that turns out to be incorrect. Thus, in drawing an inferential conclusion or accepting a hypothesis as valid, a scientist accepts a certain level of ‘epistemic risk’—the risk of being wrong. As a consequence, the majority of academic scientists endeavour to minimize their exposure to epistemic risk, usually by resisting temptation or blandishment to reach conclusions that go beyond the bounds of observational data (however, this may be less true for some scientists in the pay of certain industries, see below).

But, at the same time, scientific research is pursued with the goal of advancing knowledge and furthering its utility in all areas of life – and this must include consideration of the type of low probability events characterized by Taleb[23] as ‘Black Swans’. There is, therefore, a tension between the goal of advancing scientific knowledge and understanding of nature, and the desire to avoid epistemic risk. Research scientists have developed mechanisms and rules for managing epistemic risk, but these rules tend to enforce an acute aversion to that risk so that, in some circumstances, its avoidance can act counter to the broader, exploratory aims of the scientific endeavour; moreover, strategies for managing epistemic risk can have unforeseen or unexpressed moral or societal consequences[24].

When dichotomous yes/no decisions must be made - say, about whether to evacuate a population threatened by a volcano or a hurricane, or to ban commuting and travel if a pandemic strikes - choosing to act or not act may have drastic consequences for huge numbers of people. Academic scientists may wish for the comfort of being able to suspend scientific decision on the issue, reducing their epistemic risk exposure by waiting to see the eventual outcome of the situation, but this strategy is untenable for many decisions in the ‘real world’, including policy addressing societal risk. Concern for epistemic risk reduction in relation to decision-making on any course of action that could have life and death, health or even human rights implications raises the stakes in areas where policy is informed by science – and, increasingly, scientific reluctance to engage in such matters is being exacerbated by growing litigiousness.

Techniques for managing epistemic risk and for quantifying degrees of uncertainty and evidential support have evolved, but not uniformly in and across all branches of science. Statisticians have proposed rules of thumb for managing epistemic risk, but in the main these rules focus on avoiding Type I errors, and neglecting the consequences of Type II errors -respectively, the error of rejecting the null hypothesis when it is true, and failing to

---

[23] Taleb, N.N. (2010) *The Black Swan*. Second Edition, Penguin.

[24] Parascandola M. (2010) Epistemic risk: empirical science and the fear of being wrong. Law, Probability and Risk9, 201−214. doi:10.1093/lpr/mgq005

reject it when it is false. When applied universally, Parascandola[25] argues this approach can lead sometimes to unacceptable consequences counter to the aims of science to explain and predict natural phenomena.

In the biomedical sciences, for instance, attention is paid, almost exclusively, to the *p* value, adopting a threshold for statistical significance (usually $p < 0.05$) with the aim of avoiding Type I errors. However, dependence on such *p* values comes at a cost: the usual *p* value threshold reduces the risk of Type I errors at the expense of Type II errors – in other words, false positives are implicitly regarded as somehow less acceptable than false negatives. This is an application of one form of epistemic risk management that involves an implicit element of increased epistemic risk taking on the Type II error side of the balance sheet[26].

Thus, a challenge exists for mainstream science: how to measure and establish a threshold for epistemic risk that does not engender undesirable or unintended consequences. In decision support, the situation is made more difficult by the fact that, inevitably in any pressing issue, available scientific evidence is always incomplete, imprecise and insufficient. Rudolph Carnap introduced the Principle of Total Evidence to explain how beliefs can be supported by incomplete evidence, the principle requiring accounting for all evidence available to the scientist, and that the scientist does not ignore accessible relevant information (Carnap uses as a counter-example the case of a scientist who publishes results of experiments that support his theory, but omits those results that do not support it).

It should follow from Carnap's Principle that the more evidence there is, the better the epistemic position. Indeed, I. J. Good went on to show through an elaborate mathematical proof that introducing more evidence always lead to higher maximum 'expected utilities' in decision making. In other words, putting aside factors of cost, it is always better to make an observation or perform an experiment than to not do so. New evidence, on this view, must move understanding closer to the truth. In some cases, the value of additional evidence may be very low (i.e. the experiment is weak or poorly designed, or a single observation adds little to the totality of evidence), but it is never negative according to Good's estimation. However, scientific evidence, because it is always limited in some way, can be misleading[27],

---

[25] Ibid.

[26] In performing an experiment, say, there is no guarantee that the epistemic benefits will outweigh the costs - the epistemic utility of an experiment depends significantly on how risk averse is the person conducting the investigation. A person very averse to epistemic risk may find that the prospect of suffering loss to his or her reputation, because they could be found to have furnished a misleading inference, outweighs the potential utility of uncovering a valid result. Individual scientists are likely to take differing attitudes towards epistemic risk, and on when to suspend judgement. In contrast, others may be so anxious to supply an 'answer' that they do not suspend judgment at all: in other words, rather than the avoidance of error these individuals set more store by the removal of doubt - or the elimination of any appearance of doubt; it is easy to see how such personalities, in the face of intense political pressure, can stray in the direction of providing definitive advice, even though their basis is weakly supported by available evidence.

[27] In the 1960s, public health scientists predicted that the risks of cigarette smoking would be reduced by getting smokers to switch to cigarettes with lower tar and nicotine content, as long as those smokers did not compensate by smoking more cigarettes. Evidence collected from surveys and epidemiological studies suggested that smokers who switched did not increase the number of cigarettes smoked, but the public health scientists did not anticipate that tobacco manufacturers would redesign cigarettes using ventilation holes to increase the elasticity of the smokers' intake. Within the context of limited scientific understanding of nicotine

and Good later acknowledged[28] that the value of research evidence can appear negative in the eyes of someone else who has other information and also knows the knowledge state of the investigator, such that the investigator could be misled. Thus, while it may seem counterintuitive, gathering new evidence can actually introduce additional risk of error. Prior to performing the experiment it can be argued, moreover, that the p value is a poor measure of epistemic risk, and that broad application of a universal “standard” threshold of statistical significance $p < 0.05$ can be inimical to the fundamental aims of scientific inquiry. As scientific research methods become more and more precise, the risk of error may actually increase if this is the criterion relied on.

While all scientists ultimately are (or should be) interested in both eliminating doubt and avoiding error, there is no universal rule for determining where the balance should be struck between the two, especially in the context of providing decision support advice. Individual scientists may differ in the degree of caution they apply in different circumstances.

As noted above, the academic empirical scientific tradition has generally taken a stance of strong aversion to epistemic risk, and research is needed into how this impinges on science-informed and risk-informed decision-making, especially in the context of low probability, high consequence ‘black swan’ events.

Some suggestions

The rigour of practice is different from the rigour of science and the pursuit of truth[29]. Rigour is strict enforcement of rules to an end. Logical rigour in mathematics and science require an exactness and adherence to logical rules that derive from the pursuit of strict truth. These same rules are necessary but not sufficient for practice. There is a larger purpose in practice which is to deliver a system that is valued in a rich variety of ways. Practical rigour requires creative foresight and analysis of hindsight. The possible unintended consequences of human decisions and actions are legion and the rigour of practice is about anticipating and managing them. Scientific rigour requires selective inattention to the difficulties it cannot yet address. Practice does not have that luxury - it must include everything that is relevant even if only understood qualitatively. Any idea that, faced with a decision about a very complex matter, a decision maker can simply state "OK we'll sort this out if it happens" is not practically rigorous - every possibility must be thought through, including the inconceivable. For example, *a priori* the height of the tsunami at Fukushima might have been regarded as inconceivable. However if our lack of knowledge about such rare events had been admitted then in view of the consequences there should have been an incentive to ensure that the design was robust to the lack of knowledge.

---

dependence and smoking behaviour, the available evidence led policy in a failed direction: a tobacco company scientist with an understanding of cigarette design and how smokers could, and would, subtly change their behaviour (by covering the ventilation holes and taking more frequent puffs) could see how the survey evidence collected about number of cigarettes smoked might be misleading to other scientists, less informed about the propensities of smokers.

[28] Good, I. J. (1974) A little learning can be dangerous. British Journal for the Philosophy of Science 25: 340–342.

[29] Blockley, Godfrey 2000 Doing it Differently, Thomas Telford, London Page 113

What can be said about this issue? Of course, Bayesian philosophers of science have long argued that, in order to prioritize which trials or experiments should be conducted with limited resources, scientists must rank alternative hypotheses [30]. In this sense, the Bayesian paradigm offers an obvious rational framework for evaluating scientific evidence for decision support.

For instance, in assessing volcanic risks Bayes' rule was invoked in a retrospective analysis[31] of an unfortunate episode, the 1993 eruption of Galeras volcano, Colombia. In this case, several scientists and a few accompanying tourists were killed by an unexpected explosion following the appearance of a few rare and poorly understood seismic signals, called tornillos. This analysis exemplified, using Bayes' rule, how perceptions of the level of risk at the time might have differed between a scientist visiting the crater - who was assumed averse to epistemic risk (but only mildly so) - and a safety-conscious risk-averse person.

The comparison proceeded by assuming that *a priori* each would have attributed their own different but equally plausible weights to the significance of the tornillos and to the counter evidence – absence of other usual precursory signs of eruption. What seems clear is that some, but not all, of the unlucky scientists may have judged (validly by their own values) that the potential benefits of obtaining new data by visiting the crater outweighed the perceived risk to their own lives and limbs - i.e. they were, tacitly at least, accepting the epistemic risk that the hypothesis of no imminent eruption was wrong. This said, there is no question the physical and epistemic risk levels were too high to justify the tourists being there, at the time.

Nowadays, these principles for weighing multiple strands of (scientific) evidence are readily applied with graphical Bayesian Belief Network implementations[32], and the results can be directly linked into decision support through operational forecasting concepts[33]. However, on the basis of this (very limited) experience in volcanology, such formalized probabilistic approaches are yet to gain much traction with scientists, including those senior colleagues who can, or could, provide advice to political decision-makers. Given the intrinsic stochastic nature of black swans, it is difficult to see how their occurrence and impacts can be considered and treated, in the present context of the national risk register, other than by probabilistic methods. How best to do this, and which experts' views should be accessed, are the key issues.

---

[30] Howson, C. and P. Urbach (2005) Scientific Reasoning: the Bayesian Approach. Open Court Publishing Company, 3rd ed.: ISBN 978-0812695786

[31] Aspinall, W. P., G. Woo, P.J. Baxter and B. Voight (2003). Evidence-based volcanology; application to eruption crises. Journal of Volcanology and Geothermal Research: Putting volcano seismology in a physical context: in memory of Bruno Martinelli 128: 273-285.

[32] 8 Cowell, R.G., A.P. Dawid, S.L. Lauritzen and D.J. Spiegelhalter (1999) Probabilistic Networks and Expert Systems. Springer: ISBN 978-0-387-98767-5

[33] Woo, G. (2010). Operational earthquake forecasting and risk management. Seismological Research Letters 81:778-782. doi:10.1785/gssrl.81.5.778.

Where data are available, one possible avenue is to apply extreme value statistical analysis to search for conditions that might constitute black swan events. Recent advances in this area offer hope: some datasets evince heavy-tailed behaviour, allowing the prospect of estimating, by extrapolation, the probability of an event intensity that is beyond previous experience[34].

Social scientists such as Turner and Pidgeon (1998)[35] have argued that major disasters do not occur 'out of the blue' but incubate over a period of time with potentially identifiable patterns of attributes. The need is for methods of identifying those preconditions with sufficient dependability to enable decision makers to make such politically difficult and potentially expensive decisions to avoid the even greater costs and consequences of a disaster.

Where data are sparse or uninformative, recourse to expert judgment may be necessary in order to quantify low probability event likelihoods and to scale the intensities of such events. Formalized procedures for eliciting expert judgment for such problems have been developed, and are being applied increasingly to a widening range of medical, technical and scientific risk estimation challenges[36], including multifactorial diseases and other processes with significant complexity. Applying this structured approach offers an avenue for reasoning about uncertainties, and the possibility of putting sensible numbers to otherwise intractable probability estimates. However, it needs to be recognized that the process of eliciting judgments from a group of knowledgeable subject matter experts is a non-trivial exercise - one of the main challenges for a problem owner, seeking inputs from an expert elicitation, is the selection of his or her expert panel. Guidance can be given on this.

For the issue at hand, that is, consideration of low probability, high consequence black swan events for the national risk register, other recent advances in probabilistic methods may have roles to play, too (e.g. discrete choice probabilistic inversion[37] in relation to ranking some financial and other, more qualitative risks); a review of options in the field would be beneficial.

---

[34] For instance, unpublished work done on the flight operations database of a major airline, using the Generalized Pareto Distribution, suggests that the patterns of certain undesirable event occurrences are heavy-tailed and that new 'record' peak values could be encountered in time, albeit with a low probability in each case. One example is the measure of 'deep landings' at Heathrow: the implication of these data is that, one day, a jumbo jet could land so far down the runway that it overshoots the end of the runway and end up on the M25. Normally, deep landings are manageable, but an extreme incident might occur as the result of combination of a deep landing with one or more other low probability factors – a deep landing followed by brake failure in very wet conditions. Each is a low probability event, and manageable on its own. But, while a 'wrong' combination may be some orders of magnitude lower probability, the chance of all three happening is still non-zero. With the count of landings at Heathrow numbering over 230,000 per year, this hypothetical scenario is being tested rather frequently.

[35] Turner B A, Pidgeon N F (1998) Man Made Disasters 2nd ed) Butterworth-Heinemann, Oxford

[36] Cooke, R. M. and L. L. H. J. Goossens (2008) TU Delft expert judgment data base. Reliability Engineering & System Safety - Expert Judgement 93: 657-674.

[37] Kurowicka, D., C. Bucura, et al. (2010). Probabilistic Inversion in Priority Setting of Emerging Zoonoses. Risk Analysis 30: 715-723.

And, harking back to what was said earlier in this note, if opinions on the probabilities of very rare or unprecedented events are sought from experts, then these should be obtained from the ‘right’ type of expert - i.e. those not afflicted by acute epistemic risk aversion. Whether this is an appropriate precept to adopt in the current context, and how to determine whether a person ‘suffers’ from this trait are relevant open questions.

Annex 4

"We need the wisdom to know what we do not know"[38]

With sufficient knowledge and informed judgement uncertainty can be characterised statistically. It follows that strategic surprises arise from lack of knowledge or the inability to perceive the consequences of what is known. An evidence based process to elicit what we do not know can be very helpful. The 'italian flag'[39] [40] is a simple evidence based tool for that purpose.

The method invites consideration of the evidence for failure 'Red' and evidence for success 'Green' and what is not known 'White'. Its intention is to elicit what we do not know and by making it explicit invite strategic approaches to mitigation. It is a means of encouraging transparency, focusing on evidence and away from the fear of being wrong.

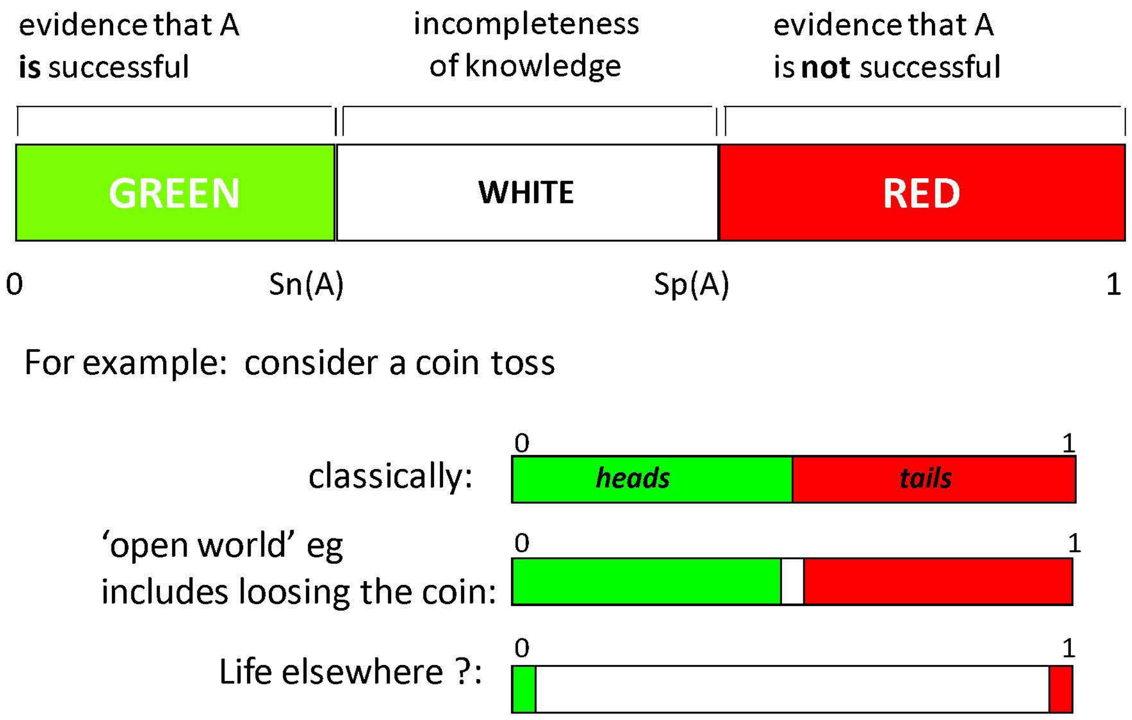


Figure 1 The Italian flag method of articulating incomplete knowledge

---

[38] Theaetetus of Plato

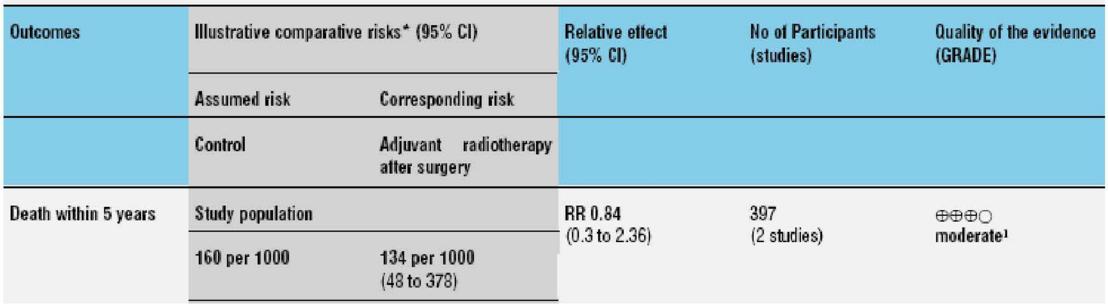
[39] Blockley D and Godfrey P (2007) "Integrating soft and hard risks" Int. J. Risk Assessment and Management Vol. 7, Nos. 6/7, 798 - 802

[40] Hall J Blockley and Davis (1998) "Uncertain inference using interval probability theory" International Journal of Approximate Reasoning 19 (1998) 247-264

Once it is accepted that for low probability events in complex systems there will always be incomplete knowledge, then it should be good practice to consider what we do not know and develop strategies that are insensitive to the lack of knowledge. For example: in the meltdown at Fukushima Daiichi Nuclear Plant, the height and damage from a tsunami wave is uncertain. If the cooling systems for the fuel rods had failed to a passively safe mode of operation such as convective heat transfer to the adjacent Pacific then the system could have been much less vulnerable to what occurred. Alternatively if sufficient time exists between the evidence of existence of the hazard and ability to act, a rigorous observational risk management approach can be used as in geotechnical engineering

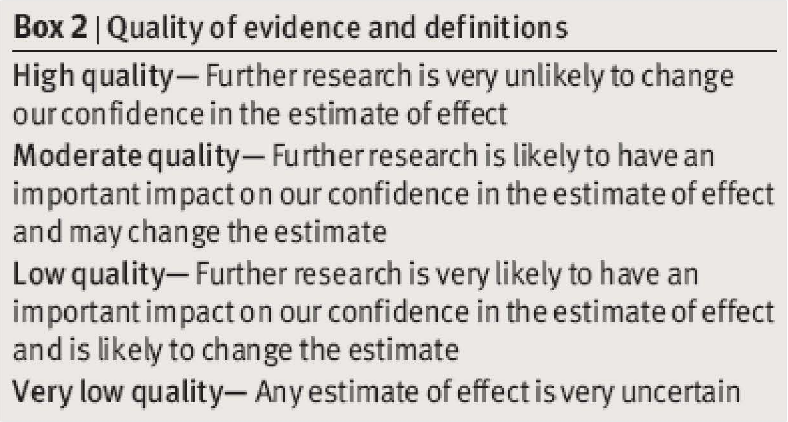
Annex 5 – Grading the quality of evidence - the Cochrane Collaboration

The Cochrane Collaboration is a major international organisation that conducts systematic reviews of medical interventions. All the 20,000+ reviews are accompanied by a ‘Summary of Findings’ table that presents estimates and confidence intervals for the treatment benefits and side effects. Below is part of the table for adjuvant chemotherapy after surgery for cervical cancer.



Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
	Control	Adjuvant radiotherapy after surgery			
Death within 5 years	Study population		RR 0.84 (0.3 to 2.36)	397 (2 studies)	⊕⊕⊕○ moderate[1]
	160 per 1000	134 per 1000 (48 to 378)			

Each conclusion has an associated ‘GRADE’ score which expresses the quality of the underlying evidence. There are 4 levels of quality.



Box 2 | Quality of evidence and definitions

High quality— Further research is very unlikely to change our confidence in the estimate of effect

Moderate quality— Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate

Low quality— Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate

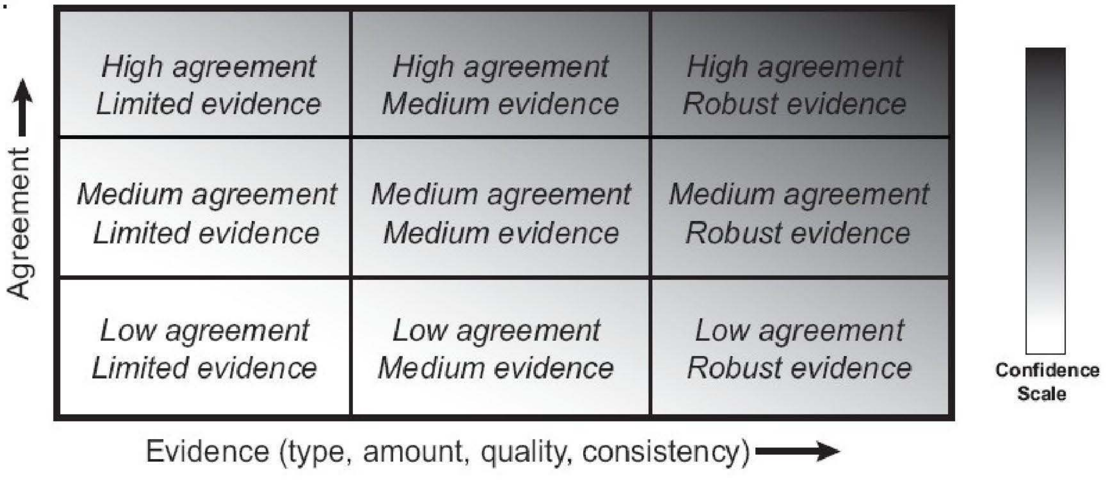
Very low quality— Any estimate of effect is very uncertain

This method is now used by many other international organisations, and appears to be adaptable to any circumstances where a quantitative risk assessment needs to be accompanied by a qualitative judgment of its reliability.

Annex 6 Communicating quality of evidence in climate change

The IPCC Guidance Notes for the Fifth Assessment report has followed, to some extent, the critical report of the Inter-Academy Council in recommending that author teams (a) summarise the quality of evidence and level of agreement as in Table 1 below, (b) if there is high agreement and/or robust evidence, express qualitative confidence in findings on a scale ‘very low’, ‘low’, ‘medium’, ‘high’ and ‘very high’, (c) when appropriate, assign a quantitative probability. When expressing uncertainty about key unknown quantities, a 6-level scale of ‘calibrated language’ is also provided, ranging from situations of ambiguity and ignorance where confidence or likelihood should not be assigned, through only specifying an order of magnitude, to full specification of a probability distribution. Authors are urged to provide traceable account for their uncertainty assessment.

We can therefore identify two somewhat different approaches to expressing uncertainty when we recognise indeterminacy through disagreement or poor quality evidence. The GRADE approach encourages a provisional numerical expression of uncertainty, but qualified by an additional judgement about the robustness of the analysis. In contrast, the revised IPCC guidance recommends that if conclusions cannot be given with high confidence then quantities should not be given probability distributions. However the difference between these approaches should not be exaggerated: the vital element is that each allows an expression of caution and humility concerning a precise quantification of uncertainty.



High agreement Limited evidence	*High agreement Medium evidence*	*High agreement Robust evidence*
Medium agreement Limited evidence	*Medium agreement Medium evidence*	*Medium agreement Robust evidence*
Low agreement Limited evidence	*Low agreement Medium evidence*	*Low agreement Robust evidence*

Figure 1: A depiction of evidence and agreement statements and their relationship to confidence. Confidence increases towards the top-right corner as suggested by the increasing strength of shading. Generally, evidence is most robust when there are multiple, consistent independent lines of high-quality evidence.

Reference

IPCC Cross-Working Group Meeting on Consistent Treatment of Uncertainties 2010 Guidance Note for Lead Authors of the IPCC Fifth Assessment Report on Consistent Treatment of Uncertainties Available from: http://www.ipcc-wg2.gov/meetings/CGCs/Uncertainties-GN_IPCCbrochure_lo.pdf

Annex 7 - A descriptive approach to risk assessment

Renn (2008) has developed a new system of risk categorisation and evaluation based on the 1999 annual report of the German Advisory Council on Global Change (WBGU, 2000). His belief was that modern societies need a more holistic concept of risk that goes beyond the two classic components of risk assessment: extent of damage and probability of occurrence. In order to achieve a balanced and reasonable judgment on the acceptability of risks, a more comprehensive set of attributes is needed that reflects public concerns and acknowledges the uncertainty and assumptions inherent in the task.

The WBGU report outlines nine criteria that were finally chosen to represent most of the experts' and public concerns as the result of a long exercise of deliberation and investigations. These were extent of damage, probability of occurrence, incertitude, ubiquity, persistency, reversibility, delay effect, violation of equity and potential of mobilization.

Renn distilled these nine criteria into six genuine risk classes, and assigned them names from Greek mythology to signify the complex issues associated with the new self-awareness of creating manageable risks, rather than just being exposed to fate:

1. **Damocles**. Risk sources that have a very high potential for damage but a very low probability of occurrence. e.g. technological risks such as nuclear energy and large-scale chemical facilities.

2. **Cyclops**. Events where the probability of occurrence is largely uncertain, but the maximum damage can be estimated. e.g. natural events, such as floods and earthquakes.

3. **Pythia**. Highly uncertain risks, where the probability of occurrence, the extent of damage and the way in which the damage manifests itself is unknown due to high complexity. e.g. human interventions in ecosystems and the greenhouse effect.

4. **Pandora**. Characterised by both uncertainty in probability of occurrence and the extent of damage, and high persistency, hence the large area that is demarcated in the diagram for this risk type. e.g. organic pollutants and endocrine disruptors.

5. **Cassandra**. Paradoxical in that probability of occurrence and extent of damage are known, but there is no imminent societal concern because damage will only occur in the future. There is a high degree of delay between the initial event and the impact of the damage. e.g. anthropogenic climate change.

6. **Medusa**. Low probability and low damage events, which due to specific characteristics nonetheless cause considerable concern for people. Often a large number of people are affected by these risks, but harmful results cannot be proven scientifically. e.g. mobile phone usage and electromagnetic fields.

Each of these risk classes is indicative of a different configuration of what Renn cites as the main challenges in risk management: complexity, uncertainty and ambiguity. Renn contends that the inclusion of such variables in a risk prioritisation system would enable us to deal with risks more realistically, effectively and in a targeted way, and in doing so would ultimately lead to more rational risk policymaking.

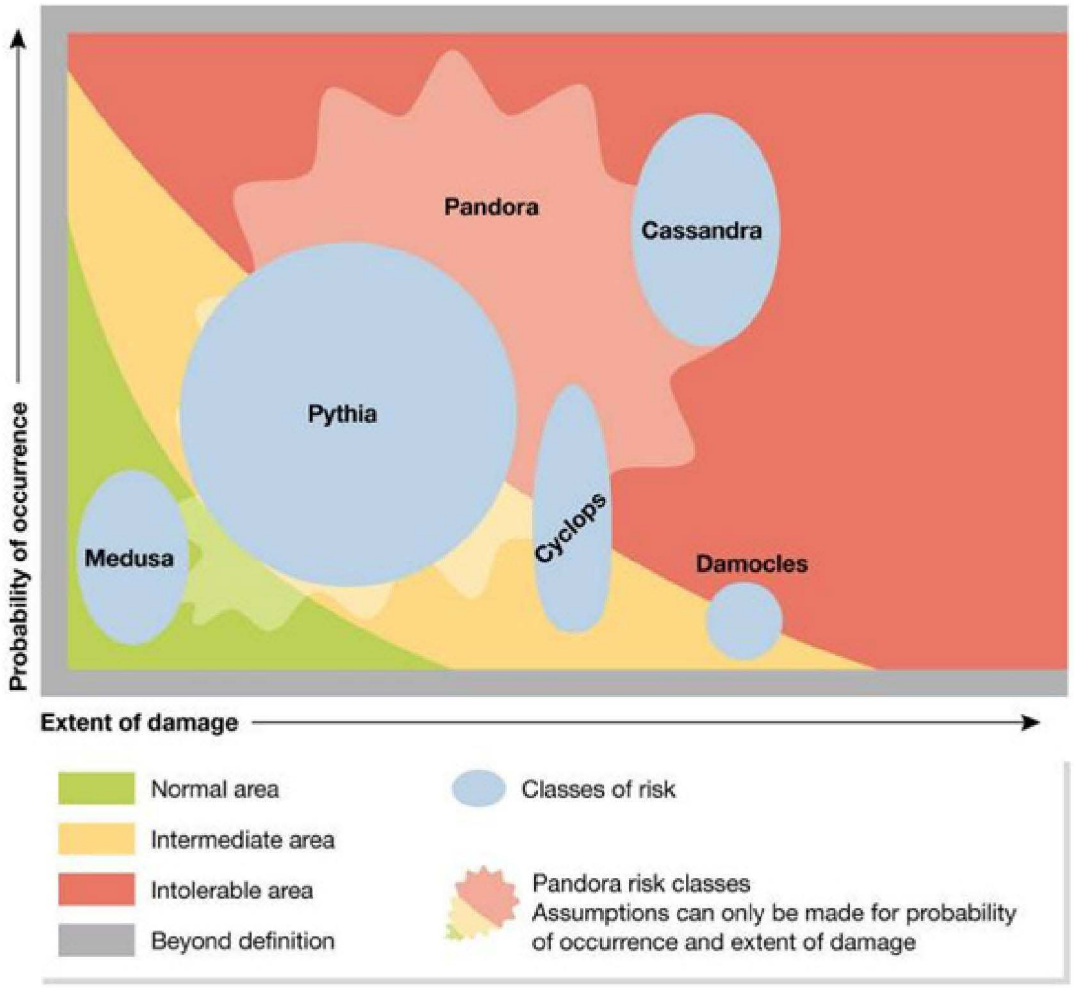
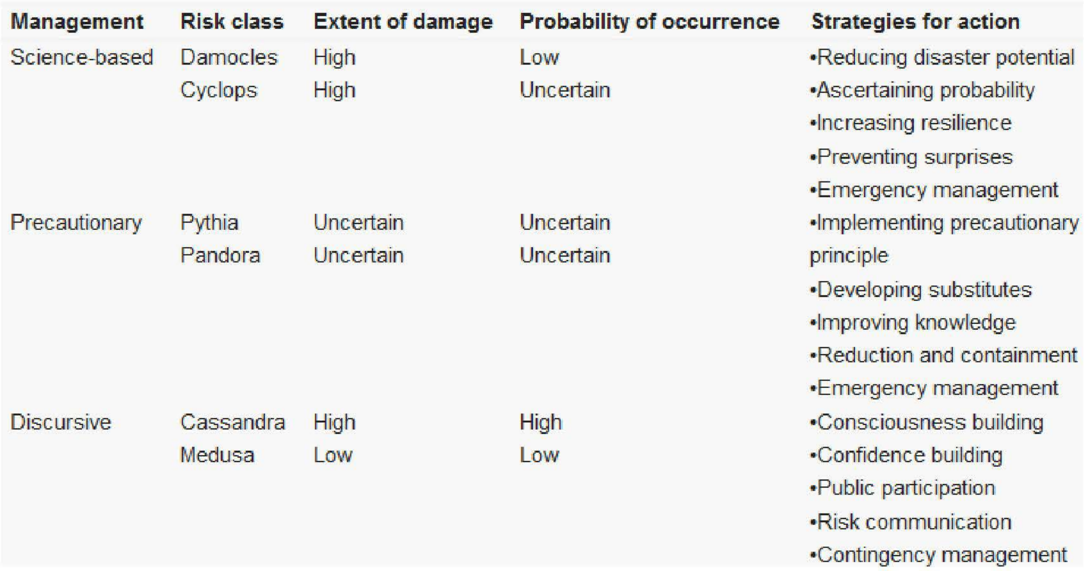


Figure 1- Risk Classes

This characterisation of risk provides a knowledge base that decision makers can use to select a management strategy to deal with each risk class. Renn distinguishes three main categories of risk management: science-based, precautionary and discursive. Damocles and Cyclops require mainly science-based management strategies, Pythia and Pandora demand the application of the precautionary principle, and the risk classes Cassandra and Medusa require discursive strategies for building consciousness, trust and credibility.



Management	Risk class	Extent of damage	Probability of occurrence	Strategies for action
Science-based	Damocles Cyclops	High High	Low Uncertain	•Reducing disaster potential •Ascertaining probability •Increasing resilience •Preventing surprises •Emergency management
Precautionary	Pythia Pandora	Uncertain Uncertain	Uncertain Uncertain	•Implementing precautionary principle •Developing substitutes •Improving knowledge •Reduction and containment •Emergency management
Discursive	Cassandra Medusa	High Low	High Low	•Consciousness building •Confidence building •Public participation •Risk communication •Contingency management

Figure 2 – Management strategies

References

Renn, O. (2008) Risk Governance: Coping with Uncertainty in a Complex World. London: Earthscan.

WBGU (2000) World in Transition. Strategies for Managing Global Environmental Risks. Annual Report 1998. Berlin: Springer.

This publication is also available on our website at http://www.bis.gov.uk/foresight

Any enquiries regarding this publication should be sent to:

Department for Business, Innovation and Skills
1 Victoria Street
London SW1H 0ET
Tel: 020 7215 5000

If you require this publication in an alternative format, email enquiries@bis.gsi.gov.uk, or call 020 7215 5000.

URN 12/519