DOI: 10.13140/RG.2.2.20914.71363

Is decision-making using historical data alone more dangerous than lawnmowers?

Norman Fenton and Martin Neil

Norman Fenton and Martin Neil are both Professors at Queen Mary University of London, and Directors of Agena Ltd agena.co.uk

11 January 2018

In our recent article[1] we discussed the Royal Statistical Society's 'International Statistic of the Year': The statistic was: "69: the annual number of Americans killed, on average, by lawnmowers - compared to two Americans killed annually, on average, by immigrant Jihadist terrorists", and the CEO of the RSS tweeted[2] about it:

> *"Americans 34 times more likely to be killed by lawnmowers than foreign-born jihadis".*

Nassim Nicholas Taleb summed up the problems in comparing the two numbers by stating[3]: "Your lawnmower is not trying to kill you". In addition to concerns raised by Taleb about both the choice of the statistic and the conclusions drawn from it (such as those of the CEO above), we explained the need to consider causal and explanatory factors, rather than just counting deaths, when assessing risk.

An especially concerning aspect of the RSS citation was the implication that the relatively low number of terrorist deaths implied that new measures to counter terrorism were unnecessary because of the 'low risk'. But then, equally, we might conclude that the relatively high number of deaths from lawnmowers requires us either to spend money educating people on lawnmower safety or perhaps simply ban lawnmowers. In fact, and rather obviously, to make such decision reasoning explicit and rational, we would have to perform a cost-benefit and trade-off analysis (see Figure 1 for the kind of model required for the terrorist case). As implied by the RSS, imposing new measures to counter terrorist threats involve both a financial cost and a human rights cost. But they also involve potential *benefits* - not just in terms of lives saved but also in reduction of other existing (secondary) security costs and resulting improved quality of life. The implication from the RSS was that the costs were greater than the benefits.

[1] Fenton, N.E.,& Neil, M. (2018). "Are lawnmowers a greater risk than terrorists?", Open Access Report DOI:10.13140/RG.2.2.34461.00486/1
[2] @HetanShah
[3] via twitter @nntaleb

This is an extended type of Bayesian network, called an ***influence diagram***. The 2018 release of AgenaRisk (version 10.0) enables users to build and compute influence diagrams like this that include both discrete and continuous nodes.

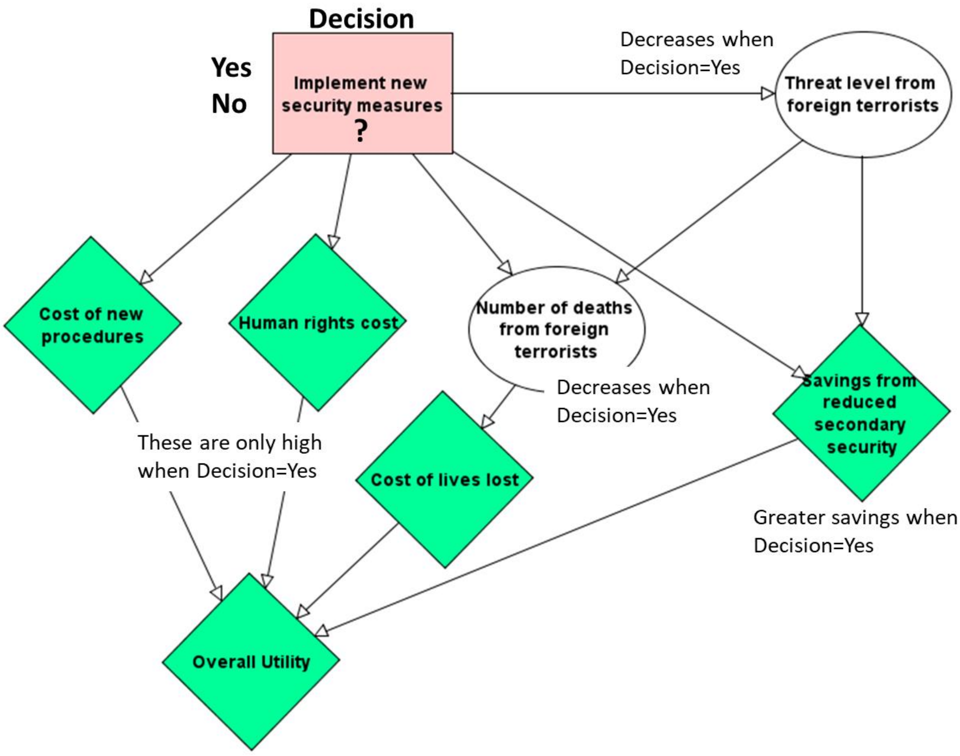

Figure 1 The kind of cost-benefit trade-off analysis required for informed decision-making[4]

But even if this trade-off analysis had been made explicit (which would involve putting actual numbers to all the costs and benefits, as well as the number of expected deaths from jihadis who would otherwise not have entered the USA) ***there is a fundamental flaw in relying only on historical fatality data***.

As Figure 1 indicates, the number of fatalities depends (among other things) on the security measures that are put in place. To make the point stronger consider the following analogous example:

> *In several decades prior to 1974 the number of deaths in London due to the River Thames flooding was zero.*

Based on this data (and applying the RSS reasoning) what possible justification could there have been for the British Government to decide to build the **Thames Barrier** - a flood barrier which cost £650 million pounds before its completion in May 1984 - rather than do nothing?

The decision to build the Thames Barrier was made because steadily rising water levels were already causing expensive (but non-fatal flooding) in other coastal areas and reliable models predicted

[4] www.agenarisk.com

catastrophic flooding within 50 years if no barrier was built. In this case the simplistic counts of past number of fatalities were clearly insufficient for rational risk assessment and decision-making.

While the Thames Barrier decision still made use of historical data (namely monthly water levels and cost of flood damage etc.), the key point is that we need to go beyond the simplistic data and consider contextual and situational factors. Moreover, in many risk scenarios 'triggers' and 'threats' that are analogous to rising water levels in this example might require expert judgments and models *in addition to* data. Without an explanatory or causal model, the data alone would be meaningless from an inferential or decision-making perspective.

Similarly, despite zero verifiable fatalities from man-made climate change, the RSS has strongly supported extensive measures to combat it. This contrasts to their conclusions drawn from low terrorist fatality numbers.

Completely novel risks (such as crashing civilian planes into skyscrapers prior to 9/11) can only be quantified using expert imagination and judgement. Indeed, the 9/11 scenario had previously been considered seriously by security experts (and movie scriptwriters), and terrorist 'chatter' suggested the threat was increasing. However, the probability of such an event was considered sufficiently low not to merit additional security measures that could have been put in place to avert it. Had security measures - which are now routine at all airports in the world - been put in place before 9/11, there would have been no mass fatalities on 9/11. Yet we find the same flawed 'data centric' reasoning being applied yet again. Witness the response to the partial immigration ban proposed by President Trump in 2017. There has been very strong opposing arguments made claiming that the proposed measures were unnecessary because everyday risks from, for example, lawnmowers is greater than the risk from jihadis. This demonstrates again the problems of relying solely on historical fatality data.

One of our own areas of research – predicting system reliability– is especially prone to these kinds of misconceptions about over-simplistic past data. A 'system' could be a software program, a physical device (phone, TV, computer, microprocessor, or component thereof) or even a process (such as a method for manufacturing steel). It is standard to measure a system's reliability in terms of the frequency with which it fails – both before 'releasing' the system for general use and also after release. But using this data alone is problematic. Why? For instance, consider a system where after two years there are very few or zero reports of system failures. At first glance this might suggest that the system is very reliable. But there is another possible explanation - the cause of the low number of failures may well be that the system was so bad that it was rarely or never used. So, here we have competing causal explanations that are very different but give rise to the same observable data.

In many areas of life past data is a good indicator of future behaviour and may be sufficient for good decision-making. Based on average temperatures in previous years we can be pretty confident that if we are going to Cairo in June we will not need a fur coat to keep warm. You don't even need the past data to be 'constant'. A company that has seen a steady year-on-year increase in sales of widgets can be confident of next year's sales based on simple regression models. The same is true in many industries. In both of these examples we do not use the data alone. We use it with a (often implied) model to interpret and make inferences, either using other relevant circumstances connected to weather or customer demand. But as soon as there are novel circumstances and factors this type of model for decision-making is likely to be poor.

Our book describes how to build models that incorporate expert subjective judgement with data in order to provide fully quantified risk assessment. The new edition (out summer 2018) has extensive new material on influence diagrams that enable us to automatically compute optimal decisions based on maximising utility:

> Fenton, N.E. and M. Neil, *Risk Assessment and Decision Analysis with Bayesian Networks*. CRC Press
>
> www.bayesianrisk.com

We acknowledge the financial support by the European Research Council (ERC) under research project, ERC-2013-AdG339182-BAYES_KNOWLEDGE. See project website: http://bayes-knowledge.com/