# Taking decisions, not risks


Risks are an unavoidable part of modern life, but mathematicians and statisticians have developed a variety of methods to help mitigate its effects. These techniques enable hospitals, banks and other organisations to make better decisions, based on evidence and facts.

Are some potentially life-saving drugs too expensive to be prescribed by the NHS? Should banks lend money to people with poor credit histories? Is it ever safe to live near a volcano? Answering questions such as these requires a careful analysis of the potential risks and consequences involved, and thanks to the latest advances in statistics we're able to make better decisions than ever before.

The consideration of risk isn't a purely mathematical process, but a solid mathematical foundation is vital when making evidence-based decisions. In the past this meant using simple formulas to perform a basic statistical analysis, but the wealth of data and vast computing resources now available let statisticians develop new statistical methods, uncovering trends that were previously impossible to see.

For example, David Spiegelhalter and colleagues at the Medical Research Council Biostatistics Unit in Cambridge have developed a statistical software package that can model a variety of real-world problems, such as comparing the cost-effectiveness of the different drugs and treatments available for a particular medical condition. Statisticians can use this software to calculate a value for each treatment in terms of the expected gain in quality-adjusted life years (QALYs), where one QALY indicates that the treatment will enable a patient to live for the equivalent of one year of perfect health.

This standard measure means the cost of each treatment can easily be compared, which is why the National Institute for Health and Clinical Excellence (NICE) use Spiegelhalter's program and similar statistical packages to evaluate the risks and benefits of competing medicines. NICE makes recommendations to the NHS about the most cost-effective drugs for a particular medical condition, and it is recognised as a world leader in setting healthcare standards. This international reputation, founded on strong statistics, leads countries such as China, Thailand and Georgia to seek expert healthcare advice from NICE.

> "The consideration of risk isn't a purely mathematical process, but a solid mathematical foundation is vital when making evidence-based decisions"

In addition to his work at the MRC Biostatistics Unit, Spiegelhalter investigates risk in a number of settings as the Winton Professor of the Public Understanding of Risk. For example, estimating how much longer a new drug will enable a patient to live turns out to be mathematically similar to predicting when a volcano will erupt - the main difference is the time scale involved. This mathematical technique, known as survival analysis, is also used by manufacturers to estimate how long machine components will last.

Risk analysis is also very important in the financial sector, especially during uncertain economic times. Retail banks use statistical techniques to model risk in a variety of situations, such as deciding whether to approve loans, credit cards and other financial product, or monitoring their customers' behaviour to ensure prompt repayments. Banks also use risk analysis to sort through transactions and look for any anomalies that could indicate financial fraud.

Statisticians such as David Hand at Imperial College London help banks to make risky decisions by analysing their extensive data records. A credit card supplier might process a billion transactions each year, each with 80 items of information, and this large amount of data allows for the construction of very sophisticated models. These models must also be able to produce answers quickly, as a system that can flawlessly detect 100% of fraud is no use if it takes three months to run, so the banks must take advantage of the latest statistical techniques to make the best decisions they can in the available time.

Institute of mathematics & its applications
EPSRC Pioneering research and skills

Card transactions in particular require rapid assessment, as it is easy for a fraudster to quickly run up a large bill. This kind of credit and debit card fraud costs the UK around £500 million per year, so identifying fraudulent transactions is a top priority. One method is to create a statistical model of a customer's card use over time, and automatically flag any transactions that don't meet their usual pattern. For example, if a customer who normally spends relatively small amounts in supermarkets and clothing shops suddenly purchases £5000 worth of jewellery, it is possible that someone else is using their card. When the model identifies this kind of deviation, the bank can get in touch with the customer and verify whether the transaction is genuine.

This method only works if the bank can accurately identify their customers' spending patterns, since mistakenly flagging legitimate transactions as fraudulent is both a waste of resources and likely to annoy customers. The complexity of an average customer's spending means that the statistical analysis must be quite advanced – for example, a simple model that compares transactions on a month-to-month basis will likely fail when customers splash out at Christmas, marking their unusual but legitimate spending as fraudulent.

Hand and his colleagues have created more accurate methods for fraud detection that can deal with these situations. In an approach called "peer group analysis", customers are grouped with others who have behaved similarly in the past. If all of the customers in one group show an increase in transactions around Christmas then the bank can safely assume these purchases are genuine, but if a single customer deviates from the general group behaviour, the bank can investigate further.

New techniques like this are vital to the banks, as they are caught in an ever-escalating fight with the fraudsters. Anti-fraud methods are only successful for a short amount of time, as demonstrated by the chip and PIN system rolled out in 2006. Although the added security reduced crimes in which the perpetrator is physically present, the number of fraudulent purchases via the phone or internet has actually increased.

Risk is unavoidable in finance, health and many other aspects of modern life, but we can use statistics to mitigate its effects. While mathematical techniques developed by Spiegelhalter and Hand can't tell us exactly what will happen in the future, they enable hospitals, banks and other organisations to collate information from different sources and make decisions based on the best possible evidence.

# TECHNICAL SUPPLEMENT

## Quality-adjusted life years

A quality-adjusted life year (QALY) is calculated by multiplying the years a treatment adds to a patient's life with a rating of the quality of life they will lead. Quality of life is measured on a scale ranging from 1 (best possible health state) through 0 (death) to negative values that indicate the worst possible health states, such as being confined to bed with the inability to wash or dress oneself.

For example, if treatment A extends a patient's life by 1 year with a quality of life of 0.4, while treatment B adds 1.25 years with a quality of 0.6, then the QALYs gained are 0.4 and 0.75 respectively, a difference of 0.35. Dividing the treatment cost by the QALYs gained provides the cost per QALY. In this case, if treatment B is £7,000 more expensive than A, the cost per QALY is £20,000. NICE typically consider costs of more than £20-30,000 per QALY to not be cost effective.

Although this example is simple, in practice the number of life years gained must be estimated using advanced statistical techniques such as those developed by David Spiegelhalter. His WinBUGS (Bayesian inference Using Gibbs Sampling) software employs Markov chain Monte Carlo methods, simulating the treatment of a group of patients over time.

## Peer-group analysis

Basic statistical models of credit and debit card transactions flag sudden large purchases as potentially fraudulent, but these assessments aren't always correct. David Hand's peer group analysis technique improves the accuracy of these models by comparing the transactions of people with a similar history of purchases, meaning sudden events that impact all accounts, such as Christmas shopping, will not trigger false positives.

This involves identifying similar spending patterns from data that is asynchronous – different people don't go to the supermarket at exactly the same time. To get around this, Hand's method creates a summary statistic from the transactions in each account over a certain time frame, then measures the difference in this statistic between all of the accounts, splitting them into related peer groups. Fraudulent transactions are identified by watching for any account that exceeds a certain distance from the others, which indicates that a customer has become less similar to their peer group and something is unusual about their spending behaviour.

### References


Bolton, R.J. & Hand D.J. (2002) Statistical fraud detection: a review. *Statistical Science*, 17(3), 235-255. DOI: 10.1214/ss/1042727940

Hand, D. J. & Weston D. J. (2008) Statistical Techniques for Fraud Detection, Prevention and Assessment. In: Fogelman-Soulié, F., Perrotta, D., Pisorski, J. and Steinberger, R. (eds.) *Mining Massive Data Sets for Security*, IOS Press. pp 257-270.

National Institute for Health and Clinical Excellence (2010) *Measuring effectiveness and cost effectiveness: the QALY*. URL: http://www.nice.org.uk/newsroom/features measuringeffectivenessandcosteffectivenesstheqaly.jsp

Spiegelhalter, D. J. & Best, N. G. (2003) Bayesian approaches to multiple sources of evidence and uncertainty in complex cost-effectiveness modelling. *Statistics in Medicine*, 22(23), 3687-3709. DOI: 10.1002/sim.1586

Weston, D., Hand, D., Adams, N., Whitrow, C. & Juszczak, P. (2008) Plastic card fraud detection using peer group analysis. *Advances in Data Analysis and Classification*, 2(1), 45-62. DOI: 10.1007/s11634-008-0021-8


### EPSRC Grants


Reference: EP/D505380/1
Title: Risk Management in the Personal Financial Services Sector

Reference: EP/C532589/1
Title: Statistical and machine learning tools for plastic card and other personal banking fraud detection

*The IMA would like to thank Professor David Spiegelhalter, University of Cambridge and Professor David Hand, Imperial College London for their help in the preparation of this document*