

# **Harnessing Data to Make Better-informed Decisions**

---

Professor Stefan Steiner

# HARNESSING DATA TO MAKE BETTER-INFORMED DECISIONS

There are many situations where large volumes of data are collected over time, and processes can be greatly improved by gleaning insights from that data. For example, hospitals and healthcare authorities collect data on patient outcomes following treatment or surgery. By better analysing such data, patterns can be revealed and process changes can be implemented to improve patient outcomes. **Professor Stefan Steiner** and his colleagues at the University of Waterloo develop new models and statistical methods that can obtain such insights across a wide array of sectors, from improving healthcare to reducing road accidents.

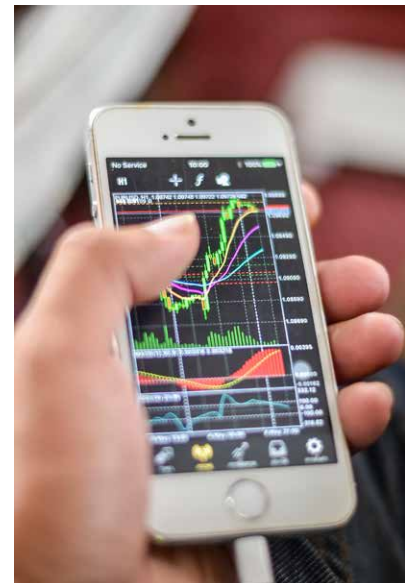
## Big Data

Data generation has increased at a rapid pace over the last few decades. In 2016 alone, about 16 trillion gigabytes of data was generated, and this number is estimated to be 10 times higher by 2025. With the availability of such large volumes of data, researchers are developing better ways to extract new information that can improve our lives and businesses.

They do so by developing statistical models, which can analyse data to gain a clearer picture of problems and successes within a system, and the factors that may be behind them. For example, long-term traffic data collected on busy roads can help to identify bottlenecks, so that plans can be developed to divert traffic in appropriate ways. In hospitals, surgical outcomes monitored over time can be analysed to make better-informed decisions about patient care. In popular culture, the movie 'Moneyball', starring Brad Pitt, recounts how a baseball team used statistical analysis to build a competitive team on a low budget.

The accuracy of such statistical models depends on the source of the data and the variables that are taken into account. Therefore, there is a great need for new models and other statistical methods that can improve the accuracy of our current approaches. Towards this aim, Professor Stefan Steiner from the University of Waterloo and his colleagues develop innovative statistical methods to improve efficiency and decision making in numerous areas – from healthcare to road safety. His team has leveraged decades of experience in modelling, process monitoring and improvement, and experimental design to support decision-making based on large data sources.

'In this era with a rapidly increasing volume of collected data there is a need for new methodology to extract useful information to provide insights that at the same time acknowledges the limitations imposed by the source of the data,' explains Professor Steiner. 'My research develops innovative statistical methods for process improvement and decision making. The general goals include improving efficiency, reducing costs and better decision making and



allocation of resources in business, industrial and medical applications.'

## Using Telematics Data to Assess Risky Driving

One focus of Professor Steiner's research involves developing models that analyse data collected by telematics devices in vehicles. Fitted in most of today's new cars, telematics devices include cameras, GPS navigation devices, accelerometers



and other sensors, which continuously collect and transmit data on driving conditions and driver behaviour. These data include vehicle location, length and time of the current trip, and incidents of acceleration, braking, and cornering.

Telematics data can be used to allow insurance companies to provide fairer premiums for safe drivers. In contrast, conventional automotive insurance policies rely on vehicle characteristics and driver demographics such as age, profession, place of residence and income. These policies often mean that young, low-income drivers need to pay expensive premiums, despite being safe drivers – increasing social inequality.

Professor Steiner and his team accessed automotive telematics data from over 28 million trips, and incorporated this data into their statistical model. They used four telematics measurements in the study: penalties recorded from acceleration, braking, speeding, and cornering. To investigate which driver behaviours were correlated with a higher likelihood of having an accident, the researchers used a case-control technique. In a case-control study, a driver that was involved in a crash is compared with other drivers who were not involved in a crash but are otherwise similar – for example in age, and gender.

By analysing the data, Professor Steiner and his team showed that speeding was the most significant driver behaviour linked to the risk of having an accident.

‘Our models would be useful in better pricing insurance products by stratifying drivers based on their observed driving behaviours and not just on static demographic data, such as age, sex and home location,’ says Professor Steiner. ‘A follow-on second goal is to develop driver monitoring protocols that quickly identify risky actions or important changes in driving behaviour that avoid overreacting to natural variation. Such monitoring procedures could be helpful, for instance, in training new drivers or monitoring elderly drivers where rapid detection of improvement or deterioration could improve decision making.’

#### **Estimating Process Performance Over Time**

Many important processes can be improved by monitoring them over time and changing the process inputs as needed to compensate for performance drifts or upsets. However, most methods estimate current process performance either by only using present-time (streaming) data or by including all historical data.

Both of these types of methods can give an inaccurate picture of process performance. By giving too much weight to historical data, one cannot achieve an accurate picture of process performance in the present. On the other hand, focusing only on very recent data means that decision makers may need to rely on only small amounts of data to make decisions which increases the uncertainty in the conclusions.

‘Current methods for these applications are not efficient and result in ineffective decisions being taken,’ explains Professor Steiner. ‘Good estimates and realistic measures of uncertainty are critical to inform important decisions, such as which hospitals are identified as underperforming, or how customer satisfaction is changing over time.’

In a paper published in 2017, Professor Steiner and his colleagues proposed a new Weighted Estimating Equation (WEE) method that more accurately monitors process performance by using equations that include both streaming data and historical data, but give less weight to the historical data. ‘By developing methods that “borrow strength” across time periods in a sensible way, we can greatly improve decision making and drive increased quality and reduced cost,’ says Professor Steiner.

The WEE approach offers a trade-off between the estimation of a performance parameter using streaming data or historical data. This trade-off is especially important if the sample size is small, because the parameter under investigation can drift slowly in an unpredictable way over time.

### **Improving Patient Care Through Risk-adjusted Monitoring**

Many hospitals and other healthcare providers are becoming increasingly interested in monitoring the outcomes of their patients after treatment. They can do this by analysing a variety of different factors, such as the length of hospital stays and 30-day mortality following surgery.

By monitoring such outcomes over time, serious problems can be rapidly detected and the performance of different healthcare providers can be compared. Such insights can allow providers to improve their processes and offer better care for their patients.

‘Effective monitoring of healthcare processes requires a statistical method



for two main reasons,’ states Professor Steiner. ‘First, decisions should incorporate a measure of uncertainty to avoid overreacting to expected outcome variability while still promptly detecting important problems. Second, patients (unlike manufactured parts) are not expected to be homogenous and importantly can have dramatically different risks prior to treatment, due for instance to underlying health.’

Therefore, Professor Steiner recommends using risk-adjusted statistical models for monitoring healthcare processes. Such models consider patient-level risk factors, in order to obtain fair and accurate comparisons between different healthcare providers.

To monitor healthcare processes and outcomes, Professor Steiner proposed the Risk Adjusted Cumulative Sum Chart, which monitors the observed patient deaths in a hospital and uses an adapted likelihood ratio test to compare this data to a predicted number of deaths.

### **Using Baseline Data to Improve Analysis of Experiments**

In industrial environments, scientists and engineers conduct experiments to understand how processes might be improved, in order to optimise efficiency and resource allocation. However, designing and implementing an experiment to identify process improvements comes at a heavy cost, in terms of planning, resources, and interrupting the routine workflow.

Professor Steiner and his team propose a low-cost method using baseline data to improve how such experiments are conducted on existing processes.

Baseline data is a measure of the current conditions as a particular process is running. This data can be later used for comparisons and calibrations – providing a historical reference point for project monitoring and assessment. Past research efforts have scarcely considered the use of baseline data to improve process-improvement experiments. ‘For process improvement projects large amounts of baseline or other prior data are often “freely” available. By exploiting this baseline information, we can reduce the cost and improve the efficiency of experiments, thereby substantially enhancing process improvement,’ states Professor Steiner.

### **Improving Decision Making Across Diverse Applications**

Professor Steiner and his team have developed numerous statistical methods and recommendations for assessing how processes can be improved across diverse applications – from surgery and road safety, to manufacturing. Their work shows how new methods that combine data in a sensible way can greatly improve decision making. The team’s risk-adjusted method for monitoring surgical outcomes has been well received and is now widely used in practice, where it is undoubtedly improving patient outcomes and saving lives.



# Meet the researcher

**Professor Stefan Steiner**

Department of Statistics and Actuarial Sciences

University of Waterloo

Waterloo, Ontario

Canada

Professor Stefan Steiner obtained his PhD in Business Administration from McMaster University in 1994, for a thesis on quality control and improvement based on grouped data. He then joined the University of Waterloo in 1995 as Assistant Professor in the Department of Statistics and Actuarial Sciences. After his promotion to Full Professor in 2010, Professor Steiner was appointed Chair of the Department in 2014. Professor Steiner's research interests cover the broad area of business and industrial statistics focusing on process improvement. The overall goal of his research is the development of innovative ways to use process data and statistical methods to drive process improvement and variation reduction. Professor Steiner has won numerous awards during his distinguished career for his research contributions in the field of statistics.

## CONTACT

**E:** [shsteiner@uwaterloo.ca](mailto:shsteiner@uwaterloo.ca)

**W:** <https://uwaterloo.ca/statistics-and-actuarial-science/people-profiles/stefan-steiner>

## KEY COLLABORATORS

Professor Nathaniel Stevens, University of Waterloo  
Professor Jock MacKay, University of Waterloo

## FUNDING

Natural Science and Engineering Research Council of Canada (NSERC)

## FURTHER READING

M Winlaw, SH Steiner, R Jock MacKay, AR Hilal, Using telematics data to find risky driver behaviour, *Accident Analysis & Prevention*, 2019, 131, 131.

PL Cooper Barfoot, SH Steiner, R Jock MacKay, Bias/Variance Trade-Off in Estimates of a Process Parameter Based on Temporal Data, *Journal of Quality Technology*, 2017, 49, 301.

SH Steiner, Risk-adjusted Monitoring of Outcomes in Health Care, In book: *Statistics in Action*, 2014, 225.

SH Steiner, RJ Cook, VT Farewell, T Treasure, Monitoring surgical performance using risk-adjusted cumulative sum charts, *Biostatistics*, 2000, 1, 441.



**UNIVERSITY OF  
WATERLOO**