The uses of statistics
Michael Wood

I am writing this just after the death at the age of 92 of the British epidemiologist, Sir Richard Doll. He was one of the first scientists to establish the link between smoking and lung cancer in the 1950s. This was based a survey of patients with a possible diagnosis of lung cancer to try to find out which aspects of their lifestyle might be linked to the incidence of lung cancer. The results showed that 647 out of 649 patients diagnosed with lung cancer were smokers.

Once this link was accepted, and the dangers of smoking appreciated, the number of people smoking and the number of deaths from lung cancer, and other smoking-related diseases, dropped dramatically. Richard Doll's obituaries credited him with saving millions of lives over the last half century. This is possibly a little over-generous in the sense that if Richard Doll had not done this research, it would probably have been done by someone else—but the fact remains that this statistical research has made a massive contribution to the management of public health.

But you may think that the statistics just confirmed the obvious. Wasn't it just a formal proof of what everyone knew already? No it was not! Smoking was not viewed as a serious health hazard, and the main cause of the epidemic of lung cancer was unknown. The approach Doll used was essentially detective work—thinking of possible explanations and then searching for evidence to check them. Doll's initial guess was that the problem had something to do with the motorcar. Despite the strength of the statistical connection that was eventually found, the links with smoking were not appreciated at all at the time. The statistics were necessary both to find and to confirm the connection.

The basic approach of statistics is simply to take a largish number of examples of whatever we are interested in, and then work out what happens *on average*. Doll's statistical survey demonstrated that, on average, smokers are more likely to develop to develop lung cancer than non-smokers.

Very similar principles apply to the management of business processes. Customer satisfaction surveys can help you understand how satisfied customers are, and which factors are related to higher levels of satisfaction. Perhaps females are more satisfied—on average—than males; perhaps there is a tendency for one aspect of your service to yield a greater level of satisfaction than other aspects; perhaps satisfaction has been declining gradually over time? The right data, analysed using simple statistical techniques, can help to pick up clues like this so that the service can be improved. Similarly, monitoring the output from a manufacturing process may reveal that the quality of the output is too variable and is deteriorating over time. Taking a snapshot at a particular point in time cannot provide this information; it is necessary to take systematic samples over a period of time so that the trends and the average amount of variability can be assessed.

The danger with *not* taking a systematic statistical approach is that you may be misled by individual events, or by cases which are not typical of the wider picture. If you have one

vociferous customer who complains loudly it may be tempting to overhaul your whole system to try to satisfy this customer. However, if this customer turns out to be atypical, this strategy may be counterproductive. It is obviously essential to try to get the whole picture—and it is here that statistics can help.

For these reasons it has been widely recognised for a long time that statistical methods are an essential tool for quality management. Two of the early advocates of the importance of statistics in this context were Walter Shewhart and W Edwards Deming who wrote in 1939:

> "The long-range contribution of statistics depends not so much upon getting a lot of highly trained statisticians into industry as it does on creating a statistically minded generation of physicists, chemists, engineers and others who will in any way have a hand in developing and directing the production processes of tomorrow."[1]

This was written in 1939, and refers specifically to production processes. Nowadays, this is just as true of production processes, but it is also equally true of service processes.

---

[1] Quoted by Oakland, J. S. and Followell, R. F., *Statistical process control, a practical guide (2nd edition)*, Oxford, Butterworth-Heinemann, 1990, p. 11.