LETTER FROM THE PRESIDENT

ENBIS and the future of applied statistics

Ron Kenett, the new president of ENBIS, has a vision of the future of applied statistics and how it relates to ENBIS

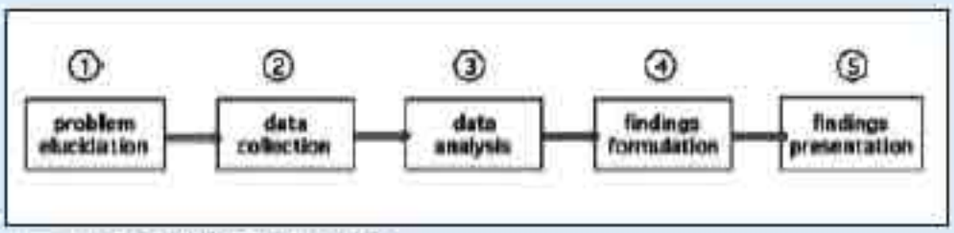

Figure 1: **The statistical consulting cycle.**

What do we need to develop applied statistics? How do these needs apply to ENBIS?

In 1922, Ronald Fisher, one of the founders of modern statistics, wrote his fundamental paper in the *Philosophical Transactions of the Royal Society*[1]. In that paper he states that 'the object of statistical method is the reduction of data'. He then identifies 'three problems which arise in the reduction of data'. These are:
Specification – choose the right mathematical model for a population;
Estimation – methods to calculate, from a sample, estimates of the parameters of the hypothetical population; and
Distribution – properties of statistics derived from samples.

Later, Colin Mallows, a distinguished researcher at AT&T Bell Laboratories, added a 'zeroth problem' – considering the relevance of the observed data, and other data that might be observed, to the substantive problem[2].

The statistical consulting cycle to solve real problems, has five steps: 1) Elucidating the problem, 2) Collecting the data, 3) Analysing the data, 4) Formulating the findings, and 5) Presenting the findings (see Figure 1). Academia typically focuses on step three: the data analysis step[3].

The application of statistics requires the full cycle. The three problems addressed by Fisher focus on steps two and three. The zeroth problem, identified by Mallows, relates to steps one and two.

To expand the view of statistical methods, Kenett *et al* suggested evaluating statistical work using a formula of Practical Statistical Efficiency (PSE). The PSE formula accounts for eight components and is computed as follows[4]:

PSE = E{R} x T {I} x P {I} x P {S} x V {PS} x V {P} x V {M} x V {D}

Where:
V{D} – value of the data actually collected.
V{M} – value of the statistical method employed.
V{P} – value of the problem to be solved.
V{PS} – value of the problem actually solved.
P{S} – probability that the problem actually gets solved.
P{I} – probability the solution is actually implemented.
T{I} – time the solution stays implemented.
E{R} – expected number of replications.

Applied statistics, in my view, is about data analysis, efficient problem solving and the full statistical consulting cycle.

Enhancements and future developments in applied statistics therefore require advances in a variety of areas such as problem elicitation techniques (see p30-31), data management and data quality techniques (see p32), data analysis methodology, graphics and visual analysis, deduction and reconstruction theory and presentation techniques.

ENBIS is a unique society with unique individuals. As a young society with our first official annual conference held in 2001 in Oslo and following annual conferences in Rimini, Barcelona, Copenhagen, Newcastle, and Wroclaw, we have proved that applied statistics can generate strong interest in skilful individuals from a variety of countries and industries. Conference participants come from academic institutions as well as banks, health care organisations, electronic, chemical, software and many other industries.

As Søren Bisgaard, one of the founders of ENBIS wrote: 'ENBIS should be the meeting ground for theory and application, a network for exchange between users, researchers and educators of business and industrial statistics. ENBIS should consider applied statistics as a wide enterprise where there should be room for such things as consulting, statistical computing and graphics, that the traditional statistical societies don't care for very much.' This is indeed the ENBIS differentiating factor.

As president of ENBIS, I look forward to an exciting year where we all work together to promote the application of statistics within and beyond the statistical community. The publication of the ENBIS Magazine within SCW is indeed one approach to reach these communities. I am determined that we shall find ways to sustain and expand the energy needed for this mission.

References

1. Fisher, R.A. (1922) On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society A*, 222, pp. 309-368)

2. The American Statistician, 52, 1998, pp 1-9.

3. Kenett and Thyregod, 'Aspects of statistical consulting not taught by academia', *Statistica Neerlandica*, special issue on Industrial Statistics, 60, 3, pp. 396-412, August 2006.

4. Kenett R.S., Coleman, S. and Stewardson, D., 'Statistical Efficiency: The Practical Perspective' *Quality and Reliability Engineering International*, 19, pp. 265-272, July-August 2003.

ENBIS

www.enbis.org

The tale of the

Antje Christensen, Karin Lund Nielsen and Peter Rasn Brinck tell us how they had to obtain and explain data for a Go – No Go decision in a pharmaceutical development project

Starting a project to develop new, better and cheaper medicines is an uplifting experience. But it is quite a different story if you must close down a project that cannot live up to this promise. The project team needs to understand, fast and reliably, what is going on, and to make the information accessible to management, so they can take the right decision. The wounded rat was the beginning of the end of such a project.

Background to the project

Biopharmaceutical medicines are not the most easy-going acquaintances a chemist can have. Production processes are often complicated, and the molecules can behave in a plethora of unwanted ways. That's the trade-off for a molecule that works in the human body in the same way as its natural counterpart. Glucagon is no exception. It is a naturally occurring hormone that raises blood sugar levels. In its pharmaceutical form, it is used as part of the treatment of diabetes, where it is injected in emergency cases of very low blood sugar to counteract insulin. It is on the market as a freeze-dried formulation. It comes in a small vial as a white powder. During the manufacturing process, there is a stage where the hormone is dissolved in water. This liquid can, under certain circumstances, turn into a jelly. Once the medicine is freeze-dried, this cannot happen, hence it is not a quality problem experienced by users. But in the production process, it can be quite a bit of a hassle. Therefore, we considered introducing a stabiliser into the formulation, some substance that prevented the liquid from turning into a jelly. The stabiliser we chose had not been earlier used in medicinal products. This did not concern us much, as the substance occurred naturally in the human body. Nevertheless, safety data had to be generated to test and document that it was safe to introduce the substance from the outside together with glucagon in the specific formulation at hand. Such tests are routinely done in animals. According to the rules of pharmaceutical development, we set up a comprehensive programme involving rats, rabbits, dogs, and pigs.

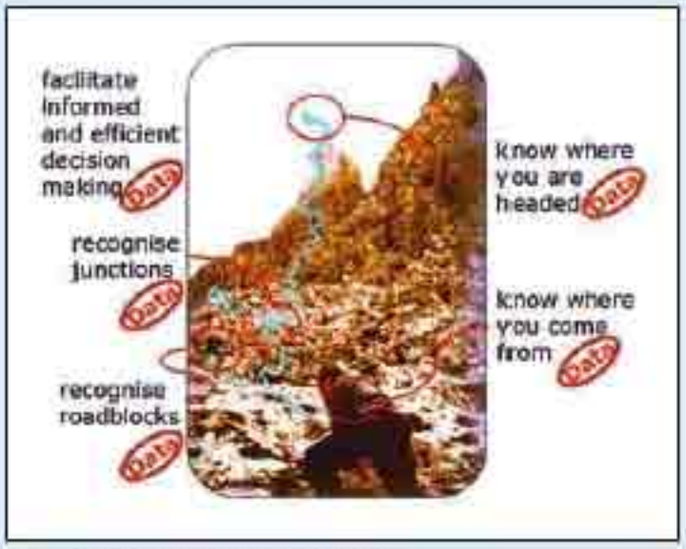

Figure 2: **Project management is path finding.**

The course of events

Onto the stage comes the wounded rat. One of the first animal studies was done in rats. It was designed to show us systemic effects of the stabiliser. We did not see any such effects – but the rats developed wounds around the injection site that really got us concerned. A study in rabbits had been started in parallel, which was designed to look out for localised skin reactions. The rabbits were better off than the rats, but they, too, showed reactions. This could be the end of the story for our project. We needed to find out two things: first, was this a rodent-specific problem, or should we expect problems in our patients as well if we went on? Second, could we change the formulation in a way that would prevent the problem and yet stabilise the solution for us? We needed to find out fast before spending more time and money on the development programme.

'We need to find the right path, hopefully the one that leads us to the summit, but sometimes also the one that leads us back to base camp'

We set up a study in pigs, the animals whose skin most resembles human skin. The pigs were given a number of different formulations: We tested the formulation that we had expected to move on with (Figure 1, [4]), against the current formulation without the stabiliser (Figure 1, [3]). We had two formulations which contained lower amounts of the stabiliser ([5] and [6]). We had a formulation without the hormone, which just contained the stabiliser and the other excipients ([7]). And we had two controls, a placebo with the current excipients ([2]), and physiological salt solution ([1]). The latter would enable us to separate the effect of

wounded rat

the injection from the effect of the injected molecules. The design of the experiment was a standard set-up for this kind of test – statistics at work below the surface.

Now a time of anxious anticipation started. We heard from the stable that the pigs looked just fine. But the final answer had to wait until the animals were sacrificed, and the histologist had had a good look at the injection sites through the microscope. Her report left not much hope: the pigs had a serious problem as well. So it was time to bring the bad news before the management body that had to decide on our project's future.

The pivotal presentation

We wanted management to see the data that had told us the bad news. So we turned the protocol of the pig study, and the 14 page result report from the histologist, into the central slide shown in figure 1.

The upper part of the figure shows the setup of the study. The light blue bars represent the excipients. They are present in all formulations except for the salt solution. The dark blue bars represent the hormone, glucagon. The red bars represent the stabiliser. Their sizes show the relative amount of the stabiliser between the formulations.

The lower part of the figure shows the results. Each formulation has been injected eight times at different sites in different animals. Four of the injections took place two days before the animals were sacrificed, while the other four took place five days before sacrifice. The histologist had examined all injection sites and classified them according to a grading system. Green is class 0, which is normal tissue. Classes A-D cover various degrees of effect, which are more or less acceptable depending on circumstances. They are all shown in grey here to avoid information overload. Class E stands for marked inflammation and cell death – a clearly unacceptable effect. It is marked in red in the graph. For all formulations containing the stabiliser, the majority of injection sites has progressed to this stage after five days.

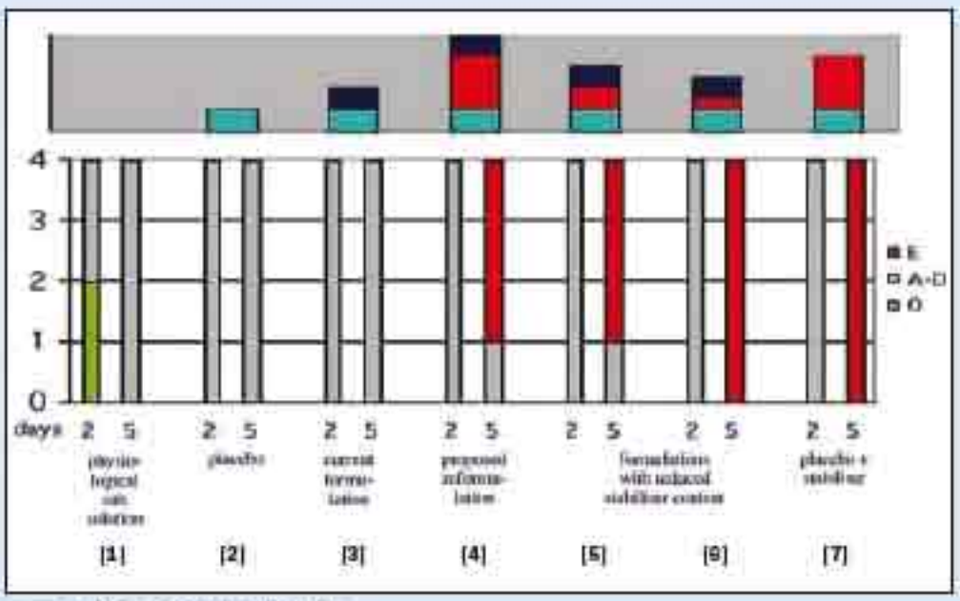

Figure 1: **Pig study setup and results.**

Not only did we have a serious problem – there was no hint as to how we could solve it, other than by keeping the stabiliser out of the product.

So that's what we proposed: Keep the stabiliser out, stick to the current formulation, close the project. After a short discussion, the management group endorsed this recommendation. We went back to finalising the reports, paying the last bills, and sharing our disappointment over a glass of beer – the traditional funeral get-together in Denmark.

The moral

How we made our data accessible at a glance
The management group that governed our project consists of representatives from such diverse areas as production, marketing, and regulatory affairs. They are not presented with preclinical findings on a daily basis. But they appreciated the depth of the information they received, and our efforts to make it as easily accessible as possible.

The study design was much easier to overview in the semi quantitative graph at the top of figure 1 than in a description. Colour coding worked in the right direction: Red for the toxic substance and red for the most serious reactions set the focus where we needed it. We changed the order of the formulations from the protocol for the graph. By doing so, we reduced the seven formulations to four principles: new versus old – lower amounts of stabilisers – no glucagon – controls. The human working memory, no matter whether it belongs to a manager or a scientist, can hold four items, but not seven.

Project management is path finding
The tale of the wounded rat exemplifies a principle that is underlying all project management activities: we need to find the right path, hopefully the one that leads us to the summit, but sometimes also the one that brings us back to base camp – see figure 2. To find the right path, we need to know where we come from and where we are headed. We need to recognise roadblocks, as the rat turned out to be, and junctions where a decision has to be made. The team needs to facilitate informed and efficient decision-making, both for decisions taken by the team itself and for decisions taken by a governing body, like the one to close the reformulation project. For all of these, we need to obtain, understand, and speak with data.

It's your data – look after it!

John Logsdon urges you to care for your data. You may need it 30 years later.

Data are the foundation of analysis, the evidence of science. Without data we can only imagine a result. While there is much to be said for pre-Socratic thought experiments to build theories, in today's data-centric world we need numbers or at least something like them. We must care for our data just as much as we treasure formulae and algorithms.

Much work in industry looks for solutions to problems and applies them immediately. You may be interested in improving the quality of a product or decreasing the cost. In these cases I hope you immediately think of using a designed experiment. But sometimes experiments are run over very long time-scales and data need to be stored before analysis.

Data needs integrity. By this I mean not only that they are right – numbers are correctly measured and recorded – but also that the measurements are self-consistent. We can all envisage problems in our own field. It is all too easy to take a reading of one parameter today but another parameter after the long weekend because the process is believed to be stable. Temperatures are recorded without realising that the thermocouples have different cold junctions. And when you come to look at these data after many years, the immediate memory will have faded. Maintaining the integrity of the data is of paramount importance.

Unfortunately it is also one of the most boring activities. We have become very lazy and imagine that by storing petabytes of information in some haphazard way we will always be able to recover and interpret the past. But the way that the data are stored and maintained affects how well we can conduct retrospective investigations – or even if they are possible.

Let me give a small example. Some years ago I was involved in a long-term corrosion experiment. This was a big problem at the time and required quantification. Many very clever people were involved. And me. All had their own ideas. Many meetings were held to discuss the issues, postulate solutions and report results.

My role was initially to store individual observations on a computer – the sort of machine now eclipsed by your wrist watch. There were other laboratories doing much the same thing but while they shared the analyses, there was no standardised storage format. There were no unique identifiers and other parameters in these collaborative experiments. My laboratory did have a consistent coding system but it was not universally adopted. This is the Identification problem.

My laboratory covered wide conditions and was probably close to an optimum design but the other laboratories ran exploratory tests. While there were often good reasons for these, a few examples do not always produce sensible results as there is insufficient data to estimate the uncertainties. These tests were often run at only a couple of conditions. This is the Design problem.

Analysis depended on data from all laboratories. A query about an individual result could always be referred to the operator but some laboratories were cagey about the actual measurements, preferring to give only processed numbers. Algorithms used to fit the data and produce intermediate results were different – some were simple least squares and others used search algorithms to identify the 'best' subset of data. These are examples of the Recording problem.

After many years I am looking at the data again. Statistical methodology has moved a long way forward in the last 30 years and while the original analysis has been proved accurate, a more sophisticated analysis is needed.

More information has come to light, there have been more experiments and they are all in slightly different formats. We now find that data have not only been recorded in different ways but often the same results are recorded in different databases. Or we have results from one specimen and then later results from the same specimen. Or the identifier has been mis-written somewhere and that two supposedly different observations are the same. These are examples of problems in Identification, Design and Recording.

These are all instances of what I call *data entropy* – the increasing disorder of data. If it is stored on a computer surely nothing can go wrong. They take backups, don't they? Well that doesn't help when a system is decommissioned without reference to the data owners so the numbers have to be re-entered. It doesn't help when people leave, companies are merged, split or taken over, responsibility has been 'outsourced' so that on paper the data is there but there is no 'ownership' of the problem. The institution ceases to care. This is the Archiving problem.

The result is that I am spending an enormous amount of time just sorting the data, checking the results, removing duplicates and validating the data before any analysis.

So before you even think of a designed experiment, look after your data, archive it, have a mandatory protocol and make one person responsible on pain of death – and payment of a lot of money – for the integrity of your data. You never know when you will need it.

Possible problems are:

Problem	Issue
Identification	Ensure that naming conventions are consistent, particularly with big experiments and diffused collaborative research.
Design	Ideally a designed experiment should be used to ensure that all variables and possible interactions are covered. At least ensure that a balance is maintained in the data that experts consider is important. Sometimes this is not practicable or even necessary: no-one would start a test for failure at a condition where failure is impossible.
Recording	We need to agree what data are recorded and that all data are consistent. Record the raw data if at all possible so that they can be re-examined. Organise the data to ensure consistency. Don't forget the silly things like date, time, operator, production line. Resist petty arguments of confidentiality if at all possible in the interests of integrity.
Archiving	Ensure that all data are stored in more than one location but also that there is no divergence between the data. This is the most difficult discipline to achieve.