

# Design and Analysis of Experiments

18 July–21 December 2011

## Report from the Organisers:

R. A. Bailey (Queen Mary, London), S. Biedermann (Southampton), B. Bogacka (Queen Mary, London), H. Dette (Ruhr-Universität Bochum), P. Goos (Antwerp), A. Lynch (Cancer Research UK), H. Maruri-Aguilar (Queen Mary, London)

## Background

The design of experiments is an area of statistical methodology which has applications in many fields where experiments are done to develop new theories or to confirm existing hypotheses. Manufacturing processes, computer experiments to model climate change, trials to improve large-scale primary care or to find effective drugs for cancer patients are only a few examples. Even small improvements in methodology can have huge effects when applied across a range of topics.

However, researchers in the different fields tend to become specialized, to use different vocabulary for the same thing, and to reinvent the same ideas, unaware that they have already been developed in another area. Recent examples include the literature on microarray experiments in genomics, where biologists did not recognize these as row-column designs; and the different meanings attached to technical statistical words such as *block*, *randomize* and *stratum* by different members of the Royal Statistical Society's working party on First-in-Man trials following the dramatic failure of the Te Genero trial at Northwick Park in March 2006.

On the theoretical side, statisticians who use one part of mathematics to further the design of experiments may know little about other parts of mathematics that can also be used. For example, R. A. Fisher effectively used character theory of Abelian groups in his famous first paper on confounding in factorial experiments, but this theory is unknown to many statisticians. Use of symmetry groups can cut down computer searches for designs, but those who know the group theory and those who design search algorithms do not always talk to each other.

There is still considerable scope for increasing the use of both algebra and computer science within design of experiments.

Design of experiments concerns the optimal capture of information to build reliable models in all areas of science, technology and commerce. It involves very careful setting of controllable input variables, or factors, decisions about the best location for sensors, choices of which units to 'treat' and so on. Without this effort there may be huge loss in haphazard decisions about data capture. At its worst, this may mean that the important questions simply cannot be answered: one may just be looking in the wrong place for the answer or not be able to measure the key causal effect.

The short INI programme on *Design of Experiments* in 2008 recognized the fragmentation in the subject. It successfully drew together people from three application areas and three areas of methodology. For 2011 we aimed to expand the coverage to include optimal designs for non-linear models, designs for experiments in healthcare in a wider sense than just clinical trials, adaptive and sequential designs, designs for non-normal data, Bayesian design, computer experiments and their use in industry, choice experiments, multi-stratum experiments, experiments for dynamic processes. We also wanted to strengthen links with other parts of mathematics, such as algebra, combinatorics, optimization.

## Programme Structure

There were six workshops, all in the first half of the programme, to suit the availability of those who had volunteered to organise them. Although this meant

that visitors who were there only during this period found that they had insufficient quiet time for research, it did have the desired effect of introducing people to aspects of the design and analysis of experiments, and application areas, with which they were unfamiliar.

There were two *Open for Business* days, one during the third workshop and one in November.

The *Cambridge Statistics Initiative* organised a one-day meeting in September. J. Stufken was the Rothschild Visiting Professor, in which capacity he gave a general lecture on orthogonal arrays in September. The Fisher Memorial Trust organised a Fisher Memorial lecture, given by A. P. Dawid on causal inference from experimental data.

## Experiments for Processes With Time or Space Dynamics

### Workshop, 18–22 July

Organisers: D. Uciński, A. Curtis.

The aim of this workshop was to bring together researchers from a number of scientific areas that use optimum experimental design as a mathematical tool to deal with large-scale and highly complex systems where time and/or space are inevitable components. Accordingly, it focused on the applications, rather than on the theory for its own sake, and met the urgent need for cross-fertilization between the engineering areas, applied mathematics and DOE experts. To this end, the workshop invited several speakers from outside the usual statistical community to help guide researchers towards potential areas of application for experimental design. Thus, it included talks on design in chemical engineering by Macchietto and Bezzo; in biochemical engineering by Van Impe; in control engineering by Hjalmarsson and Jaubertie, and in geophysics (Curtis, Wilkinson, Gibson, Winterfors). They were accompanied by talks outlining recent advances in application-driven design theory for mixed-effect models by Schwabe and Mielke; on stochastic dynamic models by Pagendam and Gibson; on adaptive design by Pronzato; on percolation and random-graph models by Bejan; on statistical learning by Wynn and Skubalska-Rafajlowicz. Other talks gave approaches to design accounting for correlations in time (López Fidalgo, Zhigljavsky, Harman) and space (Müller, Stehlík), design for processes modelled by partial differential equations (Uciński, Patan, Carraro), and design for ill-posed problems (Biedermann,

Bardow). Finally, some hard computational issues were addressed by Körkel and Melas.

Overall, an overwhelming majority of participants rated the scientific content as extremely stimulating for both theoreticians and practitioners, while highlighting most interesting links between theoretical and applied developments. The participation of many young researchers ranging from statisticians to engineers, which was evidenced by a high-level poster session among other things, and the contacts established between them were an unquestionable added value of the workshop. Because of the different research areas, most of the participants had never met before: therefore, the interdisciplinary links which were established between them during the meeting can hardly be overestimated.

## Optimum Design for Mixed Effects Non-Linear and Generalized Linear Models

### Workshop, 9–12 August

Organisers: B. Bogacka, S. Leonov.

As one of the respondents to the workshop questionnaire commented, it was “an interesting ‘first’, i.e., a conference bringing together a new set of topics”. Indeed, we are not aware of other meetings on optimum design for mixed-effects models apart from the yearly Population Optimum Design of Experiments (PODE) workshops, which are mostly focused on applications in drug development.

The three lectures (Bates, Demidenko and Fedorov) gave very good theoretical background embedding design issues in a more general setting of modelling and estimation when random parameters are present in the model. Other talks covered more specific issues of optimum design when population variability is represented by such parameters.

Accuracy of estimation of the variance components (Donev and Loeza-Serrano) and the trade-off between estimation and prediction (Schwabe) were discussed for the linear-model set-up. Gilmour, Woods and Waite presented their views on optimum design in generalized linear mixed models. Other talks were mostly related to non-linear models. Interesting Bayesian approaches were presented by Rosner and Mueller with applications to cancer research. Other applications were shown by Ueckert (Alzheimer’s disease) and Mentré (HIV). The problem of sensor loca-

tion for monitoring networks was discussed by Patan, while Latif talked about design for drug-drug interaction experiments when covariates play an important role in the choice of experimental treatments.

The last day of the workshop was devoted to PODE. Leonov, Mielke and Nyberg discussed various approximations of the Fisher information matrix, Waterhouse presented his experience in designing experiments for various phases of clinical trials in Eli Lilly, while Duffull talked about his method of sampling windows. Finally, Mentré compared several computer packages specifically devoted to finding optimum designs of experiments for population PK/PD models.

There were also four posters presented by Dumont, McGree, Nguyen and Satagopan.

## Design of Experiments in Healthcare

### Workshop, 15–19 August

Organisers: S. Biedermann, V. Dragalin, S. Eldridge, H. Großmann, M. Krams, P. Mueller

The workshop was organised around several focus themes, with separate sessions serving as almost self-contained mini-workshops for each theme. There was also an opening session with prominent speakers across all areas. The dose-ranging sessions were organised as an Open-for-Business day, with additional industry participation.

Overall, the workshop was a great success. The organisation around focus themes made it possible to get some of the leading researchers in the respective areas to present and participate. The condensed attention on one topic at a time created the right atmosphere to generate discussion and reflection on challenges and open problems that we sought to achieve. It also facilitated interactions and exploration of new ideas by exposing participants to cutting-edge research in a very focused form that often helped to isolate some of the most important research questions.

Talks and discussions in *Treatment individualization* and *Covariate-adaptive designs* highlighted the exciting opportunities that are in principle available for an adaptive model-based approach that allows investigators to learn about important subpopulations and match patients with treatments in an optimal fashion. The presentation by Don Berry showcased the potential and high expectations of the research community and other stakeholders in a currently ongoing high-profile trial. Related presentations by Jack

Lee and Kyle Wathen explored more of the technical details in such trials. Yuan Ji explored in more detail the challenges and still unresolved problems in planning such designs. Problems and challenges were further highlighted by an excellent discussion by Bhramar Mukherjee. Several participants commented informally that they thought that the discussion was one of the highlights of the meeting.

The sessions *Early-phase model-based design* and *Sequential and other Bayesian designs* explored specific novel clinical-trial designs, with an emphasis on Bayesian designs. Among the themes that emerged as interesting challenges are the appropriate use of PK data in the design of early-phase studies and the use of constraints and more prior information in Bayesian designs.

The issue of individual versus collective ethics in clinical trials was highlighted in the session on *Clinical trial design*. In many situations, such as typical phase-two dose-finding studies, inference and patient-gain objectives contradict each other. Compound criteria to balance between relative ethical and inferential gain were discussed, and adaptive randomization rules were explored with respect to their convergence to an ethical target allocation. With many statisticians from the pharmaceutical industry present, it was hoped that these ideas would soon have impact on clinical-trial design in practice.

The sessions on *Cluster-randomized trials* and *Stepped wedge designs* gave participants an insight into the use of trial design outside the field of drug development, and highlighted issues around calculation of sample size, avoidance of bias, ethics, and the practical and logistic issues in design that are often considerable in these trials.

Three sessions were devoted to *Designs for choice experiments*, with an emphasis on applications in health economics. The first gave an introduction to the general area and a review of current practice in health economics by leading international experts. Subsequent sessions presented technical advances as well as some specific applications. The sessions had a multidisciplinary flavour reflecting the broad range of backgrounds of people working on choice experiments: they were very successful in bringing together and facilitating an exchange of ideas between experts from different disciplines.

### Open-for-Business Pharmaceutical Day

This focused on the design of experiments for *Dose-*

*ranging studies*. Presentations describing current attempts to use design of experiments to overcome the recognized inefficiencies of traditional drug development as well as new challenges in implementing DoE in clinical trials were the basis for discussion and exchange of information among researchers from pharmaceutical companies and academia. We had three talks on methodology (two speakers from academia and one from industry). Industry challenges were presented by representatives of both the pharmaceutical industry and the regulatory authorities, with an additional three talks on real applications. The day closed with a panel discussion on general issues of optimal design in drug development.

## **DEMA 2011: Designed Experiments: Recent Advances in Methods and Applications**

### **Workshop, 30 August–2 September**

Organisers: S. Biedermann, S. G. Gilmour, H. Großmann, S. Lewis, B. Torsney, D. Woods.

This workshop was the third in a series and attracted a large number of high-profile speakers from around the world. A full programme was packed into the four days of the workshop, starting with Jeff Wu's keynote talk on post-Fisherian experimentation and covering a range of methodological topics, including high-dimensional responses and screening in the presence of model uncertainty, and application areas such as clinical trials and computer experiments.

With no parallel sessions, not all speakers could be accommodated, but this led to lively and high-quality poster sessions. The programme participants were very positive about the scientific programme in their questionnaire responses; many also appreciated the visit to Rothamsted Experimental Station, the birthplace of modern design of experiments.

## **Accelerating Industrial Productivity via Deterministic Computer Experiments and Stochastic Simulation Experiments**

### **Workshop, 5–9 September**

Organisers: D. Bingham, A. M. Dean, T. Santner, B. Ankenman, B. Nelson

The goal of this workshop was to bring together researchers from the two communities concerned with deterministic computer experiments and with stochastic simulation experiments, in order to share advances and to discuss the diverse approaches used. It thus provided the basis for enhancing the design and analysis of experiments for both groups. The programme featured speakers from the US, UK, France, Germany, Israel, Singapore, Netherlands, Canada, and attracted over 55 participants.

The first day focused on applications and features of computer experiments and stochastic simulation experiments, described by speakers from academia, industry, and UK and US national research centres. Topics included applications in the atmospheric sciences, oceanography, medical research and engineering. The second day was devoted to the construction of designs for use in deterministic computer experiments, using both algorithmic and combinatorial techniques. The following two days featured stochastic simulation techniques and, in particular, issues in dealing with input uncertainty and model uncertainty. On the final day, issues for computer simulators were discussed, such as interpolation, calibration and discrepancy of simulator results from physical experiment data. The workshop included a poster session as well as several panel discussions dealing with issues such as robustness, the interface of physical experiments and computer models, and future challenges for incorporating multiple modes of experimentation.

90% of questionnaire respondents rated the scientific content of the workshop as excellent. Written comments stated that the programme was “very strong and balanced” and that this was an “excellent workshop with many of the top people from both sides of the Atlantic”. The workshop provided a unique chance for the computer experiment and stochastic simulation communities to come together to share their ideas and to discuss problems of common interest.

## **Algebraic Method in Experimental Design**

### **Workshop, 26–27 October**

Organisers: H. Maruri-Aguilar, H. P. Wynn

Hugo Maruri-Aguilar and Henry Wynn had re-

quested that they be allowed to organise a special workshop on algebraic methods in experimental design, taking into account the rapid acceleration in the field and the fact that a number of participants in the main programme were either active or interested in the topic. They took advice particularly from Thomas Kahle and Giovanni Pistone. The organisers were very pleased to obtain the go-ahead and to be allowed to run the workshop in a more informal style. The respondents to questionnaires were appreciative of the organisation, including this informal approach (the new table arrangement was good but INI may consider improving this).

The central aim of the workshop was to apply the new algebraic method in experimental design to classical combinatorial designs, which were well represented at the main programme. The workshop was fortunate to have an introductory lecture on graphs and block designs from Rosemary Bailey to help in this process. On the algebraic side, there were sessions on monomial ideals and on boundary exponential models.

Real progress was made in a number of areas. Some extensive computations had shown strong correlations between different measures of connectivity for block designs, some graph-theoretic and some from recent work by Thomas Kahle on support polytopes. It is likely that the workshop will lead to further collaborations, and there is some feeling that the group should meet again to check on progress on the ideas.

## Cambridge Statistics Initiative

### 26 September

This special one-day meeting of the Cambridge Statistics Initiative (CSI) (the third such meeting) was held in the INI. These days are designed to allow networking of the statistical community within Cambridge and to expand discussions beyond the two hubs of the Statistical Laboratory and MRC Biostatistics Unit. As the DAE programme represented a considerable expansion of the Cambridge statistical community in the latter half of 2011 it was natural that the INI should play a part in the event. The academic content was organized by Ioanna Cosmo and Silvia Chiappa of the Statistical Laboratory and Andy Lynch from the DAE programme, with administrative support from Julia Blackwell of the Statistical Laboratory and the team at the INI. As the INI

provided the venue and the CSI arranged catering, attendance at the meeting was free for all participants.

Delegates were treated to 37 presentations (23 talks and 14 posters) covering topics such as genomics, statistical theory, the presentation and communication of statistical ideas, collaboration with scientists from a range of different areas, and—of course—the design and analysis of experiments. There were over 100 participants, with approximately 20 coming from the DAE programme (although it proved difficult to persuade visiting fellows that they should register for an event held within the programme, so the true number may have been higher). More importantly, there were opportunities for participants to mingle and identify areas of shared interest with near-neighbours whom they otherwise might not have met.

## Open-for-Business Industry Day

### 30 November

Organiser: D. Woods

Well-designed experiments are increasingly being recognised as key components in industrial competitiveness and scientific innovation in a variety of areas. In addition, challenges from industry and science continue to provide stimulus for new research directions in the field. This one-day meeting brought together academic and industrial researchers and practitioners for the interchange of ideas on the design and analysis of experiments. In addition to the DAE programme participants, there were a number of other academic participants, including a very encouraging number of early-career researchers and PhD students, and participants from UK industries ranging from software development to speciality chemicals.

The keynote lecture was given by Tim Davis, a former Henry Ford Technical Fellow and Quality Director for the Ford Motor Company, and past Vice-President of the Royal Statistical Society. His talk demonstrated the importance and utility of considering scientific theory and knowledge when designing and analysing industrial experiments. This was followed by four further presentations from scientists and statisticians in the automotive, pharmaceutical, consumer and oil industries, demonstrating the application of modern and novel design methods to industrial problems and presenting some new challenging problems. There was ample opportunity for both informal and formal discussions, including some very

successful break-out sessions and the panel discussion that concluded the event.

## Outcome and Achievements

The main programme had over 80 visiting fellows and other participants, from over 20 countries. A further 200 people attended one or more of the workshops. In addition to university departments of mathematics and statistics, participants came from medical departments, environmental and agricultural science, software firms, pharmaceutical companies, the aerospace industry and utility companies.

Long-term participants included both established and early-career researchers. Thanks to extra funds obtained by INI staff, some PhD students whose supervisors were visiting fellows were able to attend for short periods and gain from interactions and collaborations with people whom they would not otherwise have met.

The major achievement of the programme was the new research generated by face-to-face discussion in the wake of the new ideas introduced in the workshops. A person with a problem would hear a talk by a person with a method (or vice versa); then the two came together and were able to make progress.

The next edition of Rosenberger's book on randomization will incorporate material on cluster-randomized trials from the third workshop. Santer's forthcoming book on the design of computer experiments has been influenced by talks in the fifth workshop. Uciński is editing a book based on the first workshop. Six further books on DAE by participants will contain new material developed during the programme.

Many new collaborations were formed; old collaborators began new work together; others completed work begun before the programme. Everyone commented on how the facilities at the INI helped them to concentrate on research: the availability of books and journals; the proximity of other researchers; the layout of the building and the helpfulness of the staff.

Outside the workshops, most weeks had a seminar about work in progress. There were also many informal discussions of 1–3 hours with about 10–20 people: topics included systematic versus randomized designs; factorial designs; multi-stratum designs; algebraic and graph-theoretical approaches to block designs; numerical algorithms for finding opti-

mal designs; computer experiments and space-filling designs; designs for neighbour effects.

Non-UK participants made fruitful visits to other UK institutions, to give talks or to develop new research on DAE with people unable to attend INI.

During November, Gilmour and Trinca presented a read paper to the Royal Statistical Society on work which they had begun at the 2008 DOE programme. About 20 participants travelled to London to contribute to the discussion at the meeting; these contributions will be published with the paper.

As the workshops demonstrated, different application areas favour different optimality criteria. Designs for categorical treatments have different desiderata from those for continuous treatments. Constraints in chemical engineering are quite different from those in clinical trials. New work was done in and between all these areas.

Examples include: deployment of mobile sensors to observe spatio-temporal processes (Müller and Uciński); optimal block designs for categorical treatments in the presence of continuous covariates (Jones and Morgan); conference matrices (Cameron and Lin); designs for competing objectives (Flournoy, Haines and Rosenberger); dose selection (Bogacka and Patan); best-intention designs (Federov and Flournoy); adaptive designs (Atkinson, Biswas and Pronzato; Coad and May; Giovagnoli et al.); sample mixups in cancer studies (Lynch et al.); crossover designs (Kunert and Stufken; Bailey and Druilhet); designs to fill a spatial region that is not known *a priori* (Challenor, Dean and Santner); design polytopes (Kahle, Maruri-Aguilar and Wynn); optimality under various correlation structures (Dette, Pepelyshev and Zhigljavsky; Filipiak and Markiewicz; Müller); order of items in choice experiments (Bush, Dean and Großmann). There are already 30 preprints on the INI website.

Computation was another strong theme: the provision of software to enable both statisticians and others to design and analyse factorial designs with many strata (Großmann; Monod; Tsai); or to find designs optimizing various criteria (Goos and Jones; Haines; Harman; Torsney; Wong; Zhigljavsky).

Most participants established at least two new collaborations and worked on at least three new papers during their time at INI. Many said that this was the most effective research period in their scientific life, and that they left with plenty of excellent new ideas for future research.