


Randomisation in single-case experimental designs

Response-guided intervention in phase designs

There is a strong tradition of response-guided intervention in single-case rehabilitation studies. This is perhaps most noticeable in phase (eg, AB, ABA etc.) designs, where the common recommendation is to continue with baseline observations until the baseline stabilises; that is, several observations in a row show little variation.^{1,2} Regardless of whether visual or statistical analysis of the data follows, this response-guided determination of the point at which the treatment intervention is introduced creates a bias in favour of finding a treatment effect where none exists. This is illustrated with some simulated data pertaining to an AB design.

Figure 1 displays simulated data representing a linear trend, such as might be associated with a practice effect, but with random variations around the linear trend. The equation used to generate the linear trend was: score = 5 + 0.2(obs. no.). Thus the regression line meets the score axis at 5 and for every increase of 1 observation, the score increases by 0.2 of a score unit. The vertical arrows indicate ran-

Author



John Todman is Professor of Applied Cognitive Psychology in the Department of Psychology at the University of Dundee. He has worked in the Department for 26 years and, apart from his work on single-case experimental methodology, he engages in research on the development and evaluation of conversation aids for people without speech.

treatment effect was 33.0 for those with response guided intervention points and 11.4 for those with randomly determined intervention points (related $t(11) = 9.84$; $p < 0.001$; 2-tailed). Furthermore, the method of assignment of the intervention point accounted for 90% of the variance.

Randomisation tests for phase designs

The bias that has been described above exists whether visual or statistical analysis is used. However, introduction of a random procedure into the design, opens the way for a valid statistical test of the treatment effect. In view of the established difficulty of evaluating causal hypotheses by visual inspection of the data,⁴ a statistical test may well be worthwhile, if only to confirm the conclusion from a visual analysis.

Randomisation tests work by considering all possible recombinations of the data, given the randomisation procedure that was used in the study. In the AB example, if it had been decided that there were to be 30 observations in all and that there were to be at least 5 in

the baseline and treatment conditions respectively, the randomly determined intervention point could have occurred at any observation from 6 to 26. That is, there were 21 potential intervention points. The difference between actual baseline and treatment means is calculated, as is the mean difference for every split between baseline and treatment observations that could have occurred if a different random intervention point had been selected. If the actual difference is greater than the difference for any other potential intervention point, the probability of this happening by chance is $1/21 = 0.048$. In general, the obtained difference between means will be statistically significant at the 5% level if the obtained difference falls in the 5% most extreme differences in the (real) distribution of possible recombinations of the data. Obviously, there must be at least 20 potential intervention points to make possible a significant effect at the 5% level. Thus, many practical applications of the AB design will have quite low power to find a real effect. Power can be increased, however, by adding phases to the design (as in an ABA reversal design) or by introducing multiple baselines.

If a stable baseline is considered necessary for interpretation of an apparent treatment effect, it is possible to have the best of both methodologies by waiting until a “pre-experimental baseline” has stabilised before proceeding to a baseline-treatment experiment with a randomly selected intervention point.

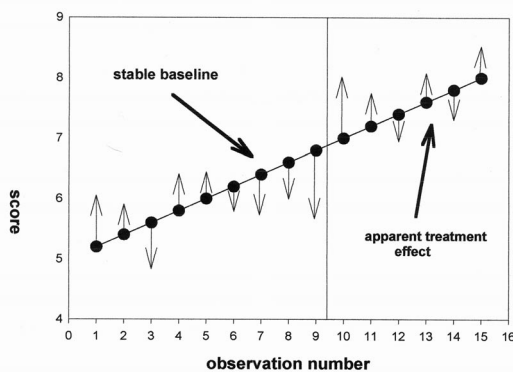
Randomisation tests for other designs

The principles underlying randomisation tests generalise to numerous other possible single-case designs, provided some form of randomisation procedure has been introduced into the design. For example, randomisation procedures (and, hence, randomisation tests) are readily applicable to alternating designs, which are the other major category of single-case designs alongside phase designs. Consider a rehabilitation study designed to compare the efficacy of two prostheses. If we have eight observation periods available, we could make four observations with each prosthesis, with prosthesis type randomly allocated to the observation periods. If the mean score difference between prosthesis types is greater than the mean difference for any other possible split of the data into two piles of four, what is the probability that this was a chance effect? It is $1/70 = 0.014$ (i.e., $p < 0.05$). For those who are interested, this result can be obtained using the formula for the number of different combinations possible for N things with n_1 and n_2 of each of two different kinds:

$$N! / (n_1! n_2!) = 8! / (4! 4!) = 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1 / (4 \times 3 \times 2 \times 1 \times 4 \times 3 \times 2 \times 1) = 70.$$

As the number of observations increases, the number of possible combinations increases rapidly. For example, with double the number of observations with each prosthesis, there are about 600 million possible combinations! This has two implications. First, alternating

Figure 1. Data generated by equation: $y = 5 + 0.2x$ with random variation added



dom fluctuations of the score introduced at each observation point (ie, random normal deviates set at mean = 0 and SD = 1.5). The important point is that we know there is no treatment effect; just a continuous linear trend with random variations. However, a couple of upward random variations (observations 4 and 5) followed by several downward random variations (observations 6, 7, 8 and 9) may easily be interpreted as stabilisation of the baseline. If the treatment intervention is introduced at this fortuitous point, the expected continuation of the upward trend is likely to appear to be a discontinuity coinciding with the commencement of the intervention. If this led to the inference of a causal relationship between the treatment and the observed discontinuity at the intervention point, it would be an erroneous conclusion.

It is reasonable to ask how frequently mistaken inferences of treatment efficacy might arise in this way. In a study carried out to assess the risk,³ 80 graphs were generated using the same equation and random deviate as in Figure 1. In 50 of these graphs a “stabilisation of the baseline” could be discerned and two copies of each of these graphs were prepared. One of each pair had an intervention point (a vertical line) drawn in immediately following the roughly horizontal sequence of observations. The other member of each pair had the intervention point randomly assigned, with the constraint that there must be at least 4 observations in each of the baseline and treatment phases. Twelve psychology students, used to interpreting graphs, were each asked to sort a randomly ordered pile of the 100 graphs into those that did and those that did not appear to show a treatment effect. The mean number of graphs judged to be consistent with a

designs, where they are practical, can have very high power. Second, powerful computers are required to make randomisation tests routinely available. Programs that allow researchers to carry out randomisation tests for a range of single-case designs within SPSS, Minitab or Excel are now available on CD ROM supplied with a recent book on randomisation tests by Todman and Dugard.⁵

To summarise; just as true experiments in multi-participant designs require random allocation of treatments to participants (and to observation times), so too is some form of random allocation of treatments to observation times essential for internal validity in single case designs. The principle of random allocation is paramount for the establishment of causal effects in single-case as in group designs, whether or not statistical analysis is carried out. The availability of valid statistical tests made possible by the random allocation procedures is “icing on the cake”.

Conference Preview

13th European Congress of Physical and Rehabilitation Medicine

28-31 May 2002, Brighton, UK

Those involved in the speciality of Physical & Rehabilitation Medicine (P&RM) – represented in the UK as Rehabilitation Medicine – at a European level know that the discipline is a lot more healthy in some countries than in Britain and that a greater range of clinical activity exists. There are whole areas of work undertaken in continental Europe that are not considered as part of the speciality’s area of expertise here. These include cardiac, pulmonary and cancer rehabilitation and span the whole age range. The lack of interest by British doctors in the activities of our colleagues abroad is also surprising and the view was formulated at the European Congress in Rome in 1997 (where there was only a handful of British doctors and not one trainee), that the only way to stimulate some interest was to hold the Congress in the UK and bring Europe to British clinicians.

This will now hopefully happen when, after five years of planning, the 13th European Congress of P&RM comes to Brighton. The European Federation of P&RM granted the Congress to be jointly hosted by the British Society of Rehabilitation Medicine and the multidisciplinary Society for Research in Rehabilitation, which is a new departure for these events and reflects changing practice in Europe. It will also hopefully see the inauguration of the new European Society of P&RM, which will succeed the European Federation. Some of the traditions of recent European Congresses will continue, but Brighton will break new ground in putting on “state of the art” plenary and concurrent sessions, along with practical teaching to meet the demands for continuing professional development and to allow trainees and junior members of staff to acquire new skills. This Congress should interest a wide variety of professionals and trainees in all professions will be particularly welcome.

The event will open on the evening Tuesday 28th May and the three-day scientific programme on the Wednesday, Thursday and Friday, 29th – 31st will be built across three main areas:

- Neurological rehabilitation,
- Musculoskeletal rehabilitation and
- Prosthetics and the technical aspects of rehabilitation.

The Congress will open with a look at Clinical Governance with a guest lecture by Professor Aidan Halligan, who heads up the NHS Clinical Governance Support Team. The themes will change each day and will cover the following through the daily morning plenary sessions:

- Day 1: Clinical standards,
- Day 2: Measurement of outcome
- Day 3: Evidence for effectiveness

References

1. Kazdin, A.E. (1982). *Single-case research designs: Methods for clinical and applied settings*. New York: OUP.
2. Shaughnessy, J.J., & Zechmeister, E.B. (1994). *Research methods in psychology (3rd ed.)*. New York: McGraw Hill.
3. Todman, J., & Dugard, P. (1999). *Accessible randomization tests for single-case and small-n experimental designs in AAC research*. *Augmentative and Alternative Communication*, 15, 69-82.
4. Franklin, R.D., Gorman, B.S., Beasley, T.M., & Allison, D.B. (1966). *Graphical display and visual analysis*. In R.D. Franklin, D.B. Allison, & B.S. Gorman (Eds.), *Design and analysis of single-case research* (pp. 119-158). Mahwah, NJ: Lawrence Erlbaum.
5. Todman, J. and Dugard, P. (2001). *Single-case and small-n experimental designs: A practical guide to randomization tests*. Mahwah, NJ: Lawrence Erlbaum.

Correspondence Address

Professor John Todman, Dept of Psychology, University of Dundee, Dundee, DD1 4HN. Tel. 01382 344615, E-Mail. j.todman@dundee.ac.uk

There will then be three concurrent sessions throughout the rest of the day covering a range of topics relevant to the speciality’s work and it is here that I hope a wider view of specialist rehabilitation will be apparent. There will be three parallel sessions in each 90-minute section across each area of rehabilitation and the aim is to allow suitable discussion during each. There will thus only be two speakers and it is hoped that the experts in the audience will have a real chance to interact with the experts on the podium. Of course, the scientific committee, headed by Professor Lynne Turner-Stokes has encouraged scientific papers and the submission date has now passed. A creditable 240 have been received and scoring has been carried out using the SRR’s well-established system. Submitted papers fall into the categories of research, audit and quality of clinical care, clinical cases, technological developments and education. These have been placed into number of topic areas (diagnosis and assessment, management, technological advances, rehabilitation approaches, audit and education) to help delegates categorise them. Poster presentation is the preferred method and there are only a few slots for oral presentations where the point of the paper would be lost in a poster. Posters are thus held in the same esteem as oral papers and this is of particular importance where the authors’ first language is not English. In addition, there are a number of workshops, seminars, updates and practical demonstrations covering many aspects of clinical work and the whole experience will be valuable to both established clinicians and to those in training.

The Congress takes over the usual summer meetings of both the BSRM and SRR and an international meeting like this allows greater choice in topics, as well as the chance to “network”. There are thus a number of social events to allow delegates to mix and, while the scientific sessions will be in English only, it will be fun to try the odd words of French, German, Dutch, Swedish, Italian or Greek! We in Britain are proud and fortunate to host this exciting event and I hope as many people interested in rehabilitation make the effort to come. The Congress’ web site has all the necessary information and can be found by clicking on to:

www.ecprpm2002.co.uk

So go on – treat yourself – come to Brighton, acquire new knowledge and skills, renew old acquaintances and make new friends.

*Dr Anthony B Ward
President, Organising Committee*