Efficient design and sample size calculation for trials with clustered data

SMMR
STATISTICAL METHODS IN MEDICAL RESEARCH

Editorial

Statistical Methods in Medical Research
2015, Vol. 24(5) 491–493

DOI: 10.1177/0962280215608718
smm.sagepub.com
SAGE

Efficient design and sample size calculation for trials with clustered data

Gerard JP van Breukelen and Math JJM Candel

This special issue on 'Efficient design and sample size calculation for trials with clustered data' is a post-hoc issue. Over the last years, six papers on this topic from our department have been accepted by *Statistical Methods in Medical Research* and have become available online. Some months ago, it occurred to us that they might be more easily found by combining them into a single issue. We are grateful to the editor, Professor Brian Everitt, for his positive response to this idea. We apologise to all statisticians who are, or once were, active in the same area as we, and whose work is not included into this issue for the reason given above: Anthony Atkinson, Valerii Fedorov, Stephen Raudenbush, Xiaofeng Liu, Allan Donner, Neil Klar, David Murray, Larry Hedges, Sonja McKinlay, Henry Feldman, Don Hedeker, Robert Gibbons, Scott Maxwell, John Overall, Weng Kee Wong, Andrew Forbes, Chul Ahn, Moonseong Heo, Kaifeng Lu, Richard Hayes, Lawrence Moulton, Michael Campbell, Chris and Stephen Roberts, Rebecca Turner, Donna Spiegelman, Xavier Basagaña, Sandra Eldridge, Karla Hemming, Bruno Giraudeau, Peter Goos, Mirjam Moerbeek, Steven Teerenstra, Joop Hox, Cora Maas, Tom Snijders, Peter Van De Ven, and our own colleagues Hubert Schouten, Bjorn Winkens and Valeria Limapassos. There are many more and the whole editorial might perhaps be filled by their names alone. We are sorry not to be able to do so.

Of the six papers in this issue, five are about cluster randomised and multicentre trials, and one is about longitudinal studies with a repeatedly measured binary outcome. We give a summary of each paper in turn. The first paper by Lemme et al. discusses the optimal treatment allocation in cluster randomised and multicentre trials with a two-by-two factorial design and a quantitative outcome, to be analysed with mixed linear regression, using as fixed part either the classical ANOVA model or Helmert contrasts. The optimality criterion is maximum precision of treatment effect estimation, as measured by the determinant of the covariance matrix of the fixed effect estimators (D-criterion), or its trace (A-criterion), or the variance of one specific fixed effect estimator. For each criterion and model, the authors derive the optimal allocation of the clusters to the four treatment arms in a cluster randomised trial and compare it with the popular balanced allocation. The theory is extended to the optimal allocation of individuals in a multicentre trial and is illustrated on a cluster randomised trial of smoking prevention and a multicentre trial on lifestyle.

Maastricht University, the Netherlands

Corresponding author:
Gerard JP van Breukelen, Department of Methodology and Statistics, School for Public Health and Primary Care CAPHRI, Maastricht University, Maastricht 6200 MD, the Netherlands.
Email: gerard.vbreukelen@maastrichtuniversity.nl

Manju et al. derive optimal and maximin sample sizes in multicentre cost-effectiveness trials with two treatment arms. Such trials have two outcomes, treatment effectiveness and treatment costs, that can be combined into a single one, called net monetary benefit (NMB), which expresses the treatment effectiveness on the same scale as the costs with a so-called willingness-to-pay parameter. Assuming a mixed model for this NMB and using as optimality criterion the precision of the treatment effect estimator, Manju et al. derive the optimal sample size (how many centres, how many patients per centre) as a function of all model parameters and the willingness-to-pay parameter, under a budget constraint and given the trial cost per centre and per patient. Since this optimal sample size depends on covariance parameters that are unknown in the design stage, the authors derive maximin sample sizes, that is the optimal sample sizes for the worst case scenario of covariance parameter values that give the minimum precision. This maximises the minimum precision and guarantees a user-specified precision level over the whole range of covariance parameter values at the lowest costs. The theory is applied to a cost-effectiveness trial comparing laparoscopic-assisted hysterectomy with standard hysterectomy, with power plots for the various designs and for various parameter values.

Van Breukelen and Candel address the problem that the optimal design of a cluster randomised trial depends on the intraclass correlation (ICC), the ratio of outcome variance between clusters to total outcome variance, which is unknown in the design stage. The larger the ICC, the larger the optimal number of clusters and the smaller the optimal sample size per cluster, at least under a fixed budget constraint. The authors derive two different maximin designs to handle this dependency or 'local optimality' problem. The first design maximises the minimum precision of the treatment effect estimator across a user-specified ICC range, thereby guaranteeing a user-specified level of precision at the price of a possibly too large sample size. The second design maximises the minimum relative efficiency compared to the locally optimal design and thus stays closer to the optimal design across the whole ICC range, but no longer guarantees the user-specified level of precision across that range. The two maximin designs are compared with each other and with two simple alternatives, and all four designs are computed for a wide range of cluster-to-person cost ratios and ICC ranges. Finally, the theory is extended to multicentre trials and to binary outcomes.

Candel and Van Breukelen derive optimal and maximin numbers of clusters for trials where the outcome variance parameters and the study costs may differ between treatment arms, and the cluster size is fixed (but not necessarily the same in each treatment arm) by practical constraints. An example is the comparison between two types of group psychotherapy, where the cluster size depends on the ideal group size for that therapy. Optimal and maximin sample sizes are derived first under a budget constraint and then under a power constraint. A large scale numerical evaluation shows how these sample sizes have to be adjusted for the loss of degrees of freedom due to unknown and possibly heterogeneous variance parameters when testing treatment effects. For a range of variance parameters, numbers of clusters, cluster sizes and effect sizes, tables with corrections for the numbers of clusters are established, both for tests with 5% and 1% type I error rates. The application in planning a study's sample size is illustrated for a trial comparing two group-administered treatments for students with body image concerns, also showing the potential savings in research costs or gains in power obtained by optimal instead of equal allocation of clusters to treatments.

The fifth paper, again by Lemme et al., extends the results of the first Lemme paper to the case of heterogeneous (treatment-dependent) outcome variance in a cluster randomised trial with a two-by-two factorial design. Allowing the outcome variance to be treatment-dependent at the cluster level, or at the individual level, or at both levels, the authors derive the optimal sample size (number of clusters and number of individuals per cluster) for each treatment arm under a budget constraint.

Since this optimal design again depends on the unknown outcome variances, of which there can now be eight, the authors evaluate the relative efficiency of the balanced design compared with the optimal design, as a function of the amount of heterogeneity of variance, for various heterogeneity scenarios, cluster-to-person cost ratios and ICCs. Finally, they show how to compute the sample size for a cluster randomised trial with a 2×2 balanced design, and how to adjust that sample size for heterogeneity of variance with an application to a published trial.

The last paper, by Abebe et al., is about the best number and timing of repeated measurements of a binary outcome in a longitudinal study with a fixed follow-up time, under cost constraints. The data are assumed to be analysed with mixed logistic regression, with as fixed part linear or quadratic growth, and a random intercept, random linear slope, and first-order autocorrelation. The optimality criterion is the determinant of the (approximate) information matrix of fixed effects estimators, which is to be maximised as a function of the number and timing of repeated measures (the number of individuals then follows from the budget constraint). Since the optimal design depends on the unknown fixed effects in logistic regression, the authors use Bayesian optimal design. By numerical search, they maximise the expected log-determinant of the information matrix, using a multivariate normal prior for the unknown fixed effects with various hyperparameter values and various covariance and autocorrelation values. The Bayesian optimal design hardly depends on the hyperparameters, or the covariance parameters, or the autocorrelation. The optimal number of measurements, however, depends on the person-to-measurement cost ratio and their optimal timing is approximately equidistant. The theory is applied to a study on respiratory infections in Indonesian children.

There is much more to say about the best sample sizes for trials with clustered data. One issue is which constraint to use: a cost constraint, a power constraint, or a practical constraint on the number of clusters, or on the number of individuals per cluster, or on the number of repeated measures. All methods in this paper can be adapted to a change of constraint, however. Another issue is how to handle the local optimality problem, that is the problem that the optimal design depends on parameters that are unknown in the design stage. Maximin and Bayesian design are two approaches. Sequential and adaptive design, especially popular in drug trials, are alternatives. Different approaches can be combined, however. For instance, a maximin or Bayesian design may be used as initial design, which can then be adapted as data come in.

In all research on optimal designs for trials with clustered data one should keep in mind the practical usefulness of results and keep things as simple as can be. In this respect, the high efficiencies of the balanced design in cluster randomised and multicentre trials, and of equidistant timing of repeated measurements, under a variety of scenarios, are welcome results.