Evaluation and validation of social and psychological markers in randomised trials of complex interventions in mental health: a methodological research programme

Graham Dunn,[1,2]* Richard Emsley,[1,2] Hanhua Liu,[1] Sabine Landau,[3] Jonathan Green,[4] Ian White[5] and Andrew Pickles[3]

[1]Centre for Biostatistics, Institute of Population Health, University of Manchester and Manchester Academic Health Science Centre, Manchester, UK
[2]Medical Research Council North West Hub for Trials Methodology Research, UK
[3]Department of Biostatistics, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK
[4]Institute of Brain, Behaviour and Mental Health, University of Manchester and Manchester Academic Health Science Centre, Manchester, UK
[5]Medical Research Council Biostatistics Unit, University of Cambridge, Cambridge, UK

*Corresponding author

Declared competing interests of authors: Dr Emsley reports grants from the UK Medical Research Council (MRC) during the conduct of the study. Professor Landau reports grants from the National Institute for Health Research (NIHR) during the conduct of the study. Professor Pickles reports grants from the MRC and from the NIHR during the conduct of the study and royalties from Western Psychological Services outside the submitted work.

Published November 2015
DOI: 10.3310/hta19930

Scientific summary

Social and psychological markers in trials of interventions in mental health

Health Technology Assessment 2015; Vol. 19: No. 93
DOI: 10.3310/hta19930

NIHR Journals Library www.journalslibrary.nihr.ac.uk

Scientific summary

Background

This report describes the development, evaluation and dissemination of statistical and econometric methods for the design of explanatory trials of psychological treatments, and the explanatory analysis of the clinical end points arising from these trials. We are concerned with making valid causal inferences about the mediational mechanisms of treatment-induced change in these clinical outcomes.

Broadly speaking, the research presented in this report aims to answer four questions about complex interventions/treatments:

1. Does it work? Is there a beneficial effect of the treatment compared with some other treatment or treatment as usual?
2. How does it work? What are the underlying mechanisms or targets of the treatment?
3. Who does it work for? Are there subgroups of people who benefit most? Could the treatment be targeted to particular subgroups of the population?
4. What factors make it work better? Is the intervention more effective when delivered as intended or when the alliance with the therapist delivering the intervention is strong?

Objectives

The present project was focused on the use of social and psychological markers to assess treatment effect mediation in the presence of measurement errors in the putative mediator, hidden confounding (selection effects) and missing data. It was also concerned with the evaluation of the role of therapeutic process variables (adherence to treatment protocol, strength of therapeutic alliance, empathy with the therapist, components of therapy) in modifying the efficacy of the therapeutic interventions.

Our aim is to produce a report that is aimed specifically at practitioners (clinical trial statisticians and trial clinicians) and not at the specialist readership comprising causal inference theorists. For this reason we have kept the technical details to a minimum, leaving the reporting of the more theoretical and mathematically demanding results for publication in appropriate specialist journals. We illustrate our methods with applications from psychological and psychosocial intervention trials.

Methods

The programme of work had three integrated components: (1) the extension of instrumental variable (IV) methods to latent growth curve models and growth mixture models for repeated measures data; (2) the development of designs and regression methods for parallel trials; and (3) the evaluation of the sensitivity/robustness of findings to the assumptions necessary for model identifiability. A core feature of the programme was the development of trial designs involving alternative randomisations to different interventions targeted on specified mediators.

A key area of application that has become apparent during the present research programme is the potential role of prognostic and predictive markers in the evaluation of treatment effect mechanisms for personalised therapies (stratified medicine), which links with component 2 as a form of stratified trial design.

We include three of our own methodological developments: (1) new IV methods that are much easier to implement than earlier computationally intensive G-estimation methods (for the case of a binary mediator, using the compliance score to create the instrument, where there are known or suspected unmeasured confounders); (2) extensions of the existing regression methods, adjusting for all known confounders and covering a variety of outcome types (implemented using the new Stata *paramed* command); and (3) extending structural equation modelling methods to account for measurement error in the mediator.

Results

We give a general introduction to the concepts of treatment effects (efficacy), average treatment effects (including the complier-average causal effect) and treatment effect heterogeneity. We discuss the evaluation of mediation by setting the scene and providing motivating examples from our clinical work. We give a critique of the historical analyses of treatment effect mediation and an introduction to the application of modern methods of causal inference to the evaluation of the direct and indirect effects of treatment (mediation). The technical challenge here is to allow for the possibility of hidden confounding (unmeasured common causes, other than treatment, of the mediator and outcome) and errors in the measurements of the mediator.

We introduce the role of post-randomisation sources of treatment effect heterogeneity, known as process measures. As with mediation, the technical challenge is to allow for the possibility of hidden confounding and measurement errors in the process variables. In addition, the process variable clearly cannot be measured in the absence of any treatment (i.e. in the control condition in a randomised trial). Here we illustrate and extend the use of IV methods and principal stratification for this application, the latter paying particular attention to patterns of missing data (both in the process measures themselves and in the clinical outcomes).

We consider extensions of the methods described above to more realistic longitudinal data structures (with the possibility of repeated measures of the mediators or therapeutic process indicators as well as of the clinical outcomes). We introduce and illustrate the use of latent growth curve and latent mixture models.

We focus on the design of Efficacy and Mechanism Evaluation (EME) trials, in particular the application of design to create convincing IVs, bearing in mind assumptions needed to attain model identifiability (i.e. the ability to obtain unique estimates of the causal effects of interest). We consider the role of targeted therapies and multiarm trials and the use of parallel trials to help elucidate the evaluation of mediators working in parallel. We give particular attention to the role of stratification (based on treatment effect moderators or predictive markers) in the evaluation of treatment effect mechanisms motivating the development of personalised therapies. We describe our new trial design, the biomarker-stratified EME trial that fully integrates marker information in trials designed to evaluate treatment effect mechanisms underlying the development of personalised therapies. We report the results of Monte Carlo simulation studies to evaluate the performance of the design and how it is affected by misclassification errors in the stratifying factor (treatment effect moderator or predictive marker).

Conclusions

First, we provide the following recommendations for EME triallists. In order to demonstrate both efficacy and mechanism, it is necessary to:

1. demonstrate a treatment effect on the primary (clinical) outcome
2. demonstrate a treatment effect on the putative mediator (mechanism)
3. demonstrate a causal effect from the mediator to the outcome.

The first two steps are necessary but not sufficient to demonstrate a causal pathway from treatment to mediator to outcome. The final step requires careful justification of the assumption that there are no other common causes of the mediator and outcome other than the treatment. These common causes may be characteristics of the trial participant prior to treatment (i.e. potential covariates or prognostic markers that could, in principle, be measured prior to randomisation). There could also be common causes that arise after the onset of treatment (such as comorbidity, life events, etc.); these are much more difficult to handle. Appropriate regression models should be applied for step 3, or alternative IV procedures, which account for unmeasured confounding provided that a valid instrument can be identified. The key to finding useful and convincing instruments appears to be treatment effect heterogeneity; in our examples we have access to treatment effect moderators, the effects of which can be observed in terms of their influence of treatment effects on both the proposed mediator and the final outcome. However, in addition to this, we need an additional assumption that the moderation of the treatment effect on outcome is wholly explained by the moderation of the treatment effect on the mediator (treatment effect mechanism). Sensitivity analysis should be conducted to assess these key assumptions.

Second, we give the following recommendations regarding the role of EME in the development of personalised therapies (stratified medicine):

1. Personalised therapy (stratified medicine) and treatment effect mechanism evaluation are inextricably linked.
2. Stratification without corresponding mechanisms evaluation lacks credibility.
3. In the almost certain presence of mediator–outcome confounding, mechanisms evaluation is dependent on stratification for its validity.
4. Both stratification and treatment effect mediation can be evaluated using a biomarker-stratified trial design together with detailed baseline measurement of all known prognostic biomarkers and other prognostic covariates.
5. Direct and indirect (mediated) effects should be estimated through the use of IV methods (the IV being the predictive marker by treatment interaction) together with adjustments for all known prognostic markers (confounders), these adjustments contributing to increased precision (as in a conventional analysis of treatment effects) rather than bias reduction.

Recommendations for future research

We conclude by giving some brief recommendations for future research in five key areas:

1. linking efficacy and mechanism evaluation explicitly
2. design of trials for EME and implications for sample size
3. measurement of mediators (reliability and measurement error)
4. other forms of outcome variable
5. sensitivity analysis.

Funding

The project presents independent research funded under the MRC–NIHR Methodology Research Programme (grant reference G0900678).

Health Technology Assessment

HTA MRP–MRC

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 5.116

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, The Cochrane Library and the ISI Science Citation Index.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: nihredit@southampton.ac.uk

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme or, commissioned/managed through the Methodology research programme (MRP), and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search, appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

The HTA programme, part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the effectiveness, costs and broader impact of health technologies for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined as all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

For more information about the HTA programme please visit the website: http://www.nets.nihr.ac.uk/programmes/hta

This report

This issue of the Health Technology Assessment journal series contains a project commissioned/managed by the Methodology research programme (MRP). The Medical Research Council (MRC) is working with NIHR to deliver the single joint health strategy and the MRP was launched in 2008 as part of the delivery model. MRC is lead funding partner for MRP and part of this programme is the joint MRC–NIHR funding panel 'The Methodology Research Programme Panel'.

To strengthen the evidence base for health research, the MRP oversees and implements the evolving strategy for high-quality methodological research. In addition to the MRC and NIHR funding partners, the MRP takes into account the needs of other stakeholders including the devolved administrations, industry R&D, and regulatory/advisory agencies and other public bodies. The MRP funds investigator-led and needs-led research proposals from across the UK. In addition to the standard MRC and RCUK terms and conditions, projects commissioned/managed by the MRP are expected to provide a detailed report on the research findings and may publish the findings in the HTA journal, if supported by NIHR funds.

The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded under a MRC–NIHR partnership. The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, the MRC, NETSCC, the HTA programme or the Department of Health. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, the MRC, NETSCC, the HTA programme or the Department of Health.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).