Journal of Clinical Epidemiology 89 (2017) 1−3

Journal of Clinical Epidemiology

EDITORIAL

Big-5 Quasi-Experimental designs

Quasi-Experimental study designs

JCE readers are invited to send letters or commentaries on whether **clinical epidemiologists and guideline groups should distinguish Quasi-Experimental study designs from other Non-Randomised designs.** Advances in health care are increasingly expensive; this increases the imperative for making policy and health care choices based on evidence. In high income countries such as Canada these compete against other basic human needs for limited resources, threatening other national priorities such as education, social programs, housing, roads etc. In LMICs global initiatives such as the Sustainable Development Goals and Universal Health Coverage also need a strong evidence-base to make the compelling arguments for their adoption if they are to succeed.

This issue offers a rich trove of papers in a series of 13 articles on Quasi-Experimental designs, led by guest-editors Barnighausen and Rockers. This series arose out of a meeting at Harvard School of Public Health to review the uses and utility of Quasi-Experimental designs in evaluating the effectiveness [benefits and harms] of healthcare practice, programs and policy. The series covers the methods of both primary research Quasi-Experimental studies and systematic reviews that include Quasi-Experimental designs in evidence synthesis. The term Quasi-Experimental was popularised by Campbell and Cook in their many papers and texts [1]. The definition of Quasi-Experimental used by them and other authors varies but the figure in the article by Waddington et al. demonstrates the two key central principles that supports them being called 'as if randomised' − namely (1) that the determination of allocation to the intervention versus no intervention [or to different interventions] being evaluated is 'exogenous' − i.e. is truly independent of the outcome of interest. (2) That this independence applies to both known confounders ['observables'] and those unknown confounders that cannot be measured ['unobservables']. These series' authors nominate the 'Big 5' study designs and their main variants as meeting these criteria: (1) Instrumental Variables, (2) Regression Discontinuity, (3) Interrupted Time Series, (4) Difference-in-Differences, (5) Fixed Effects Designs. Good examples to illustrate the key features of the sometimes complex concepts are abundant throughout the series. The authors argue that these provide substantively stronger causal evidence for benefits and harms than the traditional non-randomised designs [cohort, case control] that are taught in introductory epidemiology 101 courses. Experimental designs are still accepted as the design with the strongest internal validity but as the authors emphasize, RCTs are often not the best design for assessing external validity [including both generalisability and applicability] given the need to ensure the experiment is valid that necessitates additional artificial activities [informed consent, monitoring, additional measurements etc.] that the subject knows is for research, not their healthcare. Quasi-Experimental designs have much stronger external validity since they do not require individuals to volunteer to participate and reflect real-life practice; they are also much cheaper and quicker since they use existing data. As Barnighausen states [cf intro in this series] RCTs can be thought of as the 'first translation' between laboratory discoveries to humans to show demonstrate internal validity efficacy in ideal circumstances; then Quasi-Experimental can be thought of as the 'second translation' for the later parts of the research value chain to demonstrate the external validity impact upon clinical practice and community effectiveness. Interestingly Quasi-Experimental designs are much more commonly used in agriculture, education and social protection [2].

Several authors in the series make the case that this is not an adversarial situation of 'randophiles' protecting RCTs and beating back Quasi-Experimental advocates − both types of studies are needed for optimal evidence based policy and practice − Frenk and Gomes Dantes, and Pascal Geldsetzera and Wafaie Fawzi describe the impressive examples of Conditional Cash Transfer and HIV programs being evaluated by complementary RCT and Quasi-Experimental designs that have had major influences on practice and policy.

Twelve recommendations from the various papers are worth highlighting here as they lay out a potential roadmap for redressing the current paucity of Quasi-Experimental studies in health and healthcare:

(1) Support Capacity Development by

(a) Establishing a global working group with stakeholders (Doers [authors, library scientist, statistician, content/subject expert, methods expert, synthesis

http://dx.doi.org/10.1016/j.jclinepi.2017.09.010
0895-4356/© 2017 Elsevier Inc. All rights reserved.

organisations), Users (program managers, policy makers, practitioners, patients and public, private industry, payors, press)-charged with developing approaches to strengthen the global capacity for Quasi-Experimental studies and increase the uptake of RCT and Quasi-Experimental combinations in policy and practice (Rockers et al.).

(b) Enlist large and small synthesis organisations to sign up to implementation of the above (Rockers et al.)

(c) Increase the supports for infrastructure [e.g. forums where policymakers can meet with stakeholders and researchers to hold deliberative dialogues that include Quasi-Experimental options (Rockers et al.)

(d) Incubator funding to support methods development to improve Quasi-Experimental studies and reviews (Rockers et al.)

(e) Quasi-Experimental courses should be integrated into graduate health science programs (Rockers et al.)

(f) An integrated approach to knowledge translation and exchange (Lavis et al.)

(2) Develop new funding opportunities for Quasi-Experimental studies; investments should be made to develop new innovations in Quasi-Experimental methods (Rockers et al.)

(3) Make available administrative data and clinical registers from health programs from all publically funded health systems for Quasi-Experimental studies of primary research and evidence synthesis (Rockers et al.)

(4) Develop and validate uniform Quasi-Experimental guidelines using the Equator methods for conducting and reporting primary Quasi-Experimental research studies (Rockers et al.)

(5) Standard Quasi-Experimental guidelines should be developed for conducting synthesis of evidence on effectiveness that includes RCTs and Quasi-Experimental designs; when RCTs and Quasi-Experimental studies are combined, the results should also be disaggregated to show the contribution of each (Rockers et al.; Becker et al.).

(6) Establish a registry [or expand a current one] to register titles and protocols of Quasi-Experimental primary studies, rapid reviews and full systematic reviews (Lavis et al.).

(7) Support a research synthesis library that identifies both RCTs and the Big Five Quasi-Experimental designs (Lavis et al.)

(8) Extend validated Risk of Bias instrument[s] to include the five Quasi-Experimental designs in studies and evidence synthesis systematic reviews (Waddington et al.)

(9) Extend validated best practice Quality Assessment instrument[s] for use by GRADE and other guideline groups for the five Quasi-Experimental study designs, distinguished from other non-randomised designs (Geldsetzer et al.)

(10) Justify any QE design 'label': Study design labels work for the study designs that clinical epidemiologists use frequently, e.g., RCTs, Cohort studies, and Case Control studies, but Quasi-Experimental study labels should be always justified by details of the component features such as with the proposed checklist developed by Reeves et al.

(11) Establish search strategies for Quasi-Experimental designs. For Quasi-Experimental studies evidence synthesis to be trusted all robust Quasi-Experimental studies addressing the question of interest need to be sought out; the types of search filters that have been successfully developed for RCTs are missing for Quasi-Experimental studies (Reeves et al.)

(12) Standardise data collection of true confounding variables given the importance of adjusting effect sizes for true estimates of effects (Aloe et al.)

Patient-Reported Outcomes [PROs]

In this issue is also a second series, with 10 articles, entitled the Montreal Accord on Patient-Reported Outcomes Use Series (arising from a meeting in Montreal in 2013). Patient Reported Outcomes [PROs] are everywhere − they number over a 1000 now.

For this series guest editors, Bartlett and Ahmed have commissioned a series of articles to address the challenge of how to bring order to this chaos. These authors argue persuasively that this dream of appropriate universal use will only be realised if (a) this is at least loosely coordinated through national networks linked to other national and global networks; (b) that there is consensus on the taxonomy, methods and metrics.

This can be tracked back to 1948 when the World Health Organisation defined health as a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity'. This led to many groups around the world committing substantial resources to quantify and measure physical, emotional and social wellbeing.

The US Institute of Medicine and equivalent authorities in many other countries have adopted the patient-centred model to ensure that priority is given to the health of individuals and communities rather than economic efficiency alone. Research agencies such as Canadian Institutes of Health Research, UK MRC, Australian NH&MRC, US NIH now mandate that the 'patient voice ' be included in demonstrating efficacy. Originally developed for use in research, PROs have evolved into an essential component for monitoring individuals and populations, managing care, evaluating services and providers, and informing policy.

However there has been confusion and some understandable resistance to this morass of instruments, many of which are of dubious validity. As emphasised in the publications of COMET [3] and OMERACT [4] this is similar to other types of clinical trial outcomes where researchers

have been guilty of concentrating heavily [perhaps obsessively] on churning out new measures [Schmidt 2010]; all too often these are fatally flawed in that (a) they were developed without input from the patients/public [despite the fact that they are the source of PRO data]; and (b) the criteria for success have usually been based on statistical significance instead of meaningful substantive importance.

It is encouraging to see that some progress has been made on international consensus amongst different stakeholders. In 2001 the PRO Harmonization Group was established by four organisations [European Regulatory Issues on Quality of Life Assessment Group, International Society for Quality of Life Research, International Society for Pharmacoeconomics and Outcomes Research, and the Health Outcomes Committee of PhRMA], with observers from other agencies such as EMEA and FDA. The objectives were to highlight problematic topics/issues; harmonize recommendations proposed by each organization; and work on solutions to achieve consensus on problematic issues/topics. In 2004 the PROMIS initiative was funded by the NIH to establish a common resource of PROs with a common set of readily accessible item banks (currently for over 60 adult conditions and 40 pediatric conditions). PROMIS International, set up in 2013, is a volunteer consortium involving a number of countries (including Canada, France, the Netherlands, Spain, Sweden) for the use of PROMIS for addressing cultural relevance and coordinate translation for PROs in research, clinical care and population monitoring worldwide. As detailed on paper 6 (Bartlett et al.), the Montreal Accord recommendation is to establish a collaborative PRO Network in Canada with the main objectives being to (a) identify a harmonised set of relevant 'universal' [generic] and disease-specific measures; (b) provide guidance on selecting other PROs; (c) facilitate access to existing and new item banks with documentation of their clinimetric and psychometric performance; (d) standardize methods for content development, validation and calibration; (e) provide resources for the clinimetric and psychometric approaches; (f) provide training in outcomes science and the skills needed for integrated knowledge translation, dissemination and facilitation of uptake; (g) conduct research to increase acceptance and use of PROs. Several of these objectives are shared by other outcome networks/organisations such as COMET [Core Outcome Measures in Effectiveness Trials: http://www.comet-initiative.org/about/aimsandobjectives], COSMIN [COnsensus-based Standards for the selection of health Measurement INstruments: http://www.cosmin.nl/], HTAi [Health Technology Assessment international: https://www.htai.org/htai/about-htai/], ICHOM [International Consortium for Health Outcomes Measurement: www.ichom.org/], OMERACT [Outcome Measures in Rheumatology: www.omeract.org], SONG [Standardised Outcomes on NephroloGy: http://songinitiative.org/] − for example the importance of integrated knowledge translation involving all stakeholders from the very beginning as reflected in the concept of Integrated Knowledge Translation [5], so perhaps a global network is needed to advocate for the best PROs by harmonizing terminology, agreeing on minimal standards and speaking with a unified voice.

The series contains state of the art articles on the taxonomy; PROs in shared decision making; clinical decision making including screening, referral, diagnosis, prognosis, assessment, setting goals and monitoring therapy; use in population and public health; the experience and lessons learned from the PROMIS example; value implications; and use in electronic health records.

We look forward to clinical epidemiologists contributing to resolving these challenges.

Peter Tugwell
J. André Knottnerus
Jessie McGowan
Andrea Tricco
E-mail address: Laura.Tugwell@uottawa.ca (P. Tugwell)

References

[1] Campbell DT, Stanley JC. Experimental and Quasi-Experimental designs for research. Chicago: Rand McNally; 1966.
[2] Cameron DB, Mishra A, Brown AN. The growth of impact evaluation for international development: how much have we learned? J Dev Effect 2016;8(1):1−21.
[3] Prinsen CA, Vohra S, Rose MR, Boers M, Tugwell P, Clarke M, et al. How to select outcome measurement instruments for outcomes included in a "Core Outcome Set" - a practical guideline. Trials 2016;17(1):449.
[4] Boers M, Kirwan JR, Wells G, Beaton D, Gossec L, d'Agostino MA, et al. Developing core outcome measurement sets for clinical trials: OMERACT filter 2.0. J Clin Epidemiol 2014;67:745−53.
[5] Tunis SR, Maxwell LJ, Graham ID, Shea BJ, Beaton DE, Bingham CO 3rd, et al. Engaging stakeholders and promoting uptake of OMERACT core outcome instrument sets. J Rheumatol 2017;44(10):1551−9.