ELSEVIER

Journal of Clinical Epidemiology 66 (2013) 703−705

Journal of Clinical Epidemiology

COMMENTARIES

Call to RECORD: the need for complete reporting of research using routinely collected health data

Eric I. Benchimol[a,b,c,*], Sinead Langan[d], Astrid Guttmann[a,e,f], on behalf of the RECORD Steering Committee[1]

[a]*Institute for Clinical Evaluative Sciences, Toronto, Ontario, Canada*
[b]*Department of Pediatrics, University of Ottawa, Ottawa, Ontario, Canada*
[c]*Department of Epidemiology and Community Medicine, University of Ottawa, Ottawa, Ontario, Canada*
[d]*Department of Non-communicable Disease Epidemiology, Faculty of Epidemiology & Population Health, London School of Hygiene and Tropical Medicine, University of London, London, UK*
[e]*Department of Paediatrics, University of Toronto, Toronto, Ontario, Canada*
[f]*Department of Health Policy, Management and Evaluation, University of Toronto, Toronto, Ontario, Canada*

Accepted 2 September 2012; Published online 24 November 2012

The rapid entry of health care into the electronic age has led to a proliferation of large data repositories of health-related information. Such data are collected for administrative and clinical purposes, without specific *a priori* research questions. Examples of these ''routinely collected health data'' include health administrative data, data warehouses of electronic medical records, primary care medical record data, and disease registries. These data are increasingly being used for observational, comparative effectiveness and health services research. The expanded use of these data for research has also generated expanded scientific investigation of their specific strengths and weaknesses, with strong research in this field recently published in the *Journal of Clinical Epidemiology* [1−7]. We would therefore strongly advocate for guidelines for the reporting of research using routinely collected health data to improve the transparency of the methods and validity of the results to allow adequate peer review and appropriate application of research evidence by health care providers and policy makers. We are therefore proceeding with the formal guideline development process, the REporting of studies Conducted using Routinely collected Data (RECORD) statement. This statement will be produced as an extension of the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement. Promoting guidelines for reporting health data research reflects a more general global effort to develop and use reporting guidelines to reduce waste in research.

As in any new research field, techniques for analyzing routinely collected health data are rapidly evolving. For example, a growing body of literature has focused on the concept of misclassification bias in the identification and surveillance of patients with chronic disease [8−10]. Misclassification bias in research using health administrative data refers to the concept that errors in classification of patients based on administrative codes could result in incorrect study conclusions such as over- or underestimation of health care costs [9]. Because such data are collected for nonresearch purposes and often poorly quality controlled, the codes and algorithms used to identify patients with disease within the databases may be based on incorrect assumptions and therefore inaccurate, resulting in significant misclassification. Despite the importance of this concept, validation of the algorithms used to identify patients with disease within health administrative data are often poorly reported. Our recent systematic review describes gaps in the reporting of algorithm validation studies, and we have therefore suggested reporting guidelines for future validation studies based on the Standards for the Reporting of Diagnostic accuracy studies (STARD) criteria [11]. Another review article recently identified other sources of potential bias in the methods used to conduct research using health administrative data [12]. These include inaccurate linkage of records across databases and biased estimates owing to selective inclusion of patients. In addition to misclassification bias, these issues are shared by all research using routinely collected health data. With time, researchers will refine methods to reduce the degree of bias and improve validity of research using such data. At

[1] Members of the RECORD Steering Committee: David Moher (Ottawa Hospital Research Institute), Irene Petersen (University College London), Liam Smeeth (London School of Hygiene and Tropical Medicine), Henrik Sørensen (Aarhus University), Fiona Stanley (University of Western Australia), and Erik von Elm (Université de Lausanne).

* Corresponding author. Division of Gastroenterology, Hepatology & Nutrition, Children's Hospital of Eastern Ontario, 401 Smyth Road, Ottawa, Ontario, Canada, K1H 8L1. Tel.: +613-737-2516; fax: +613-738-4854.

E-mail address: ebenchimol@cheo.on.ca (E.I. Benchimol).

http://dx.doi.org/10.1016/j.jclinepi.2012.09.006

present, however, these methodological issues may result in confusion among researchers and consumers of the literature. The confusion is compounded by incomplete or inadequate reporting of research based on routinely collected health data in the literature.

Use and endorsement of reporting guidelines has been shown to be associated with more complete and transparent research reporting [13,14]. An international collaboration recently produced the STROBE statement to improve reporting of observational research [15]. STROBE has been endorsed and is being used by a wide range of journal editors (including those of the *Journal of Clinical Epidemiology*), researchers, and consumers of research [16]. Because most research using routinely collected health data is observational in design, the STROBE guidelines should apply to these studies. However, at this time, neither STROBE nor other reporting guidelines identified by the Enhancing the QUAlity and Transparency Of health Research (EQUATOR) Network [17] specifically address reporting of research using routinely collected health data. Over the past decade, the International Society for Pharmacoeconomics and Outcomes Research has endeavored to create a series of checklists to aid in the critical appraisal and undertaking of research using retrospective databases [18−20]. The resulting articles represent in-depth analyses of research methods and quality. Although some of the included discussion would apply to reporting, the checklists were not intended as minimum reporting standards for all studies using routinely collected data. Additionally, the checklists often addressed the quality of research, rather than focusing on reporting. Widely endorsed and useful reporting guidelines typically focus on the latter, while avoiding the issue of quality judgment [13]. In the most recently published article by Berger et al. [18], the authors advocated for a standardized approach to the reporting of observational studies using retrospective databases and suggested some modifications to STROBE, which could be applied. These were developed based on the authors' expertise without a formal Delphi process and did not address the intricacies of various internationally available databases.

Formal reporting guidelines for studies using routinely collected health data, with multidisciplinary expert stakeholder involvement, and collaboration with the STROBE group, will help provide transparency of the methods used in this growing field of research. This would ensure adequate peer review and a clear understanding of methods, strengths, and limitations of the research by users. Extensions of the STROBE guidelines have been produced for genetic association studies [21] and molecular epidemiology research [22]. Such extensions address issues specific to those forms of research without overlap with existing STROBE checklist items. The process of reporting guideline development has been described by members of the EQUATOR network [23]. The first step is to assess the need for such guidelines and the possible engagement in guideline development by members of the research community. To gauge the interest and feasibility of producing such guidelines among users of routinely collected health data, a workshop was recently held following the Infectious Diseases Research Network Primary Care Database Symposium 2012 (January 27, 2012 in London, UK). Among the more than 100 participants in the workshop (including conveners of the STROBE committee), strong interest was expressed in the possibility of developing reporting guidelines specific to research using routinely collected health data. There was general agreement that the reporting of research methods was highly variable and that there are sufficient issues specific to studies based on routinely collected health data to warrant an extension of the STROBE statement. Participants at the Primary Care Database Symposium identified multiple issues that may qualify for inclusion as checklist items such as: description of database characteristics, validation of codes and algorithms to identify exposures and outcomes, and record-linkage methodology. These are difficult issues that pose specific challenges with use of routinely collected information and may apply to research using databases or cohorts designed with *a priori* research questions.

The RECORD guidelines will be developed in close conjunction with members of the STROBE group to ensure consistent methods and to make this a useful addition to the original STROBE statement. This process will involve general consultation with stakeholders, a formal Delphi process and the eventual production of useful and widely accepted reporting guidelines. At present, we are engaging stakeholders (such as researchers, data providers, health care providers, and policy makers) to provide their opinions on important items to include in reporting guidelines for research using routinely collected data. To identify further topics for the RECORD statement, a working group is planned and will include expert stakeholders, journal editors, and guideline developers. An international group of experts will then use the information provided by stakeholders to create formal guideline statements and create the final reporting guidelines. If readers of this commentary are interested in participating at the stakeholder level, please visit the website http://record-statement.org.

The availability of specific reporting guidelines will be a resource for researchers and ensure transparency of their research methods, including the strengths, limitations, and biases that may be associated with their data. Furthermore, the guidance, typically in the form of a checklist, could help scientists review the available literature and address areas of methodological concern, thereby improving the quality of research produced based on routinely collected health data. Most importantly, complete and transparent reporting will allow users of such research publications (including journal editors, peer reviewers, scientists, clinicians, and policy makers) to understand what the authors did (methods) and found (results). The checklist will provide substantive guidance for authors reporting research using routinely collected health data. Wide acceptance of

the checklist and engagement with journal editors will ensure that the methods and results of studies using routinely collected health information are clearly reported so that academics, health care providers, and policy makers can evaluate and apply their results confidently.

References

[1] McClish D, Penberthy L, Pugh A. Using Medicare claims to identify second primary cancers and recurrences in order to supplement a cancer registry. J Clin Epidemiol 2003;56:760−7.

[2] Ahmed F, Janes GR, Baron R, Latts LM. Preferred provider organization claims showed high predictive value but missed substantial proportion of adults with high-risk conditions. J Clin Epidemiol 2005;58:624−8.

[3] Schneeweiss S, Avorn J. A review of uses of health care utilization databases for epidemiologic research on therapeutics. J Clin Epidemiol 2005;58:323−37.

[4] Hall S, Schulze K, Groome P, Mackillop W, Holowaty E. Using cancer registry data for survival studies: the example of the Ontario Cancer Registry. J Clin Epidemiol 2006;59:67−76.

[5] Bottle A, Aylin P. Comorbidity scores for administrative data benefited from adaptation to local coding and diagnostic practices. J Clin Epidemiol 2011;64:1426−33.

[6] Smeets HM, de Wit NJ, Hoes AW. Routine health insurance data for scientific research: potential and limitations of the Agis Health Database. J Clin Epidemiol 2011;64:424−30.

[7] van Walraven C, Bennett C, Forster AJ. Administrative database research infrequently used validated diagnostic or procedural codes. J Clin Epidemiol 2011;64:1054−9.

[8] Kopec JA, Fines P, Manuel DG, Buckeridge DL, Flanagan WM, Oderkirk J, et al. Validation of population-based disease simulation models: a review of concepts and methods. BMC Public Health 2010;10:710.

[9] Manuel DG, Rosella LC, Stukel TA. Importance of accurately identifying disease in studies using electronic health records. BMJ 2010;341:c4226.

[10] van Walraven C, Austin PC, Manuel D, Knoll G, Jennings A, Forster AJ. The usefulness of administrative databases for identifying disease cohorts is increased with a multivariate model. J Clin Epidemiol 2010;63:1332−41.

[11] Benchimol EI, Manuel DG, To T, Griffiths AM, Rabeneck L, Guttmann A. Development and use of reporting guidelines for assessing the quality of validation studies of health administrative data. J Clin Epidemiol 2011;64:821−9.

[12] van Walraven C, Austin P. Administrative database research has unique characteristics that can risk biased results. J Clin Epidemiol 2012;65:126−31.

[13] Moher D, Weeks L, Ocampo M, Seely D, Sampson M, Altman DG, et al. Describing reporting guidelines for health research: a systematic review. J Clin Epidemiol 2011;64:718−42.

[14] Plint AC, Moher D, Morrison A, Schulz K, Altman DG, Hill C, et al. Does the CONSORT checklist improve the quality of reports of randomised controlled trials? A systematic review. Med J Aust 2006; 185:263−7.

[15] von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. J Clin Epidemiol 2008;61:344−9.

[16] Altman DG, Egger M, Gotzsche PC, Pocock SJ, Vandenbroucke JP, von Elm E. STROBE statement: endorsements. Bern, Switzerland: ISPM—University of Bern; 2009. Available at http://www.strobe-statement.org/index.php?id=strobe-endorsement. Accessed February 14, 2012.

[17] Equator Network. Enhancing the quality and transparency of health research. Oxford, UK: Minervation, Ltd.; 2011. Available at http://www.equator-network.org. Accessed January 28, 2012.

[18] Berger ML, Mamdani M, Atkins D, Johnson ML. Good research practices for comparative effectiveness research: defining, reporting and interpreting nonrandomized studies of treatment effects using secondary data sources: the ISPOR Good Research Practices for Retrospective Database Analysis Task Force Report—Part I. Value Health 2009;12(8):1044−52.

[19] Motheral B, Brooks J, Clark MA, Crown WH, Davey P, Hutchins D, et al. A checklist for retrospective database studies—report of the ISPOR Task Force on Retrospective Databases. Value Health 2003; 6(2):90−7.

[20] Peterson AM, Nau DP, Cramer JA, Benner J, Gwadry-Sridhar F, Nichol M. A checklist for medication compliance and persistence studies using retrospective databases. Value Health 2007;10(1):3−12.

[21] Little J, Higgins JP, Ioannidis JP, Moher D, Gagnon F, von Elm E, et al. STrengthening the REporting of Genetic Association Studies (STREGA): an extension of the STROBE statement. PLoS Med 2009;6(2):e22.

[22] Gallo V, Egger M, McCormack V, Farmer PB, Ioannidis JP, Kirsch-Volders M, et al. STrengthening the Reporting of OBservational studies in Epidemiology—molecular Epidemiology (STROBE-ME): an extension of the STROBE Statement. PLoS Med 2011;8(10):e1001117.

[23] Moher D, Schulz KF, Simera I, Altman DG. Guidance for developers of health research reporting guidelines. PLoS Med 2010;7(2): e1000217.