Published in final edited form as:

Epidemiology. 2011 May ; 22(3): 298–301. doi:10.1097/EDE.0b013e318212640c.

Nonexperimental Comparative Effectiveness Research Using Linked Healthcare Databases

Til Stürmer, Michele Jonsson Funk, Charles Poole, and M. Alan Brookhart

Pharmacoepidemiology Program, Department of Epidemiology, UNC Gillings School of Global Public Health, University of North Carolina at Chapel Hill, http://www.sph.unc.edu/epid/pharmacoepidemiology_program_area_7704_10437.html

Comparative Effectiveness Research (CER) has gained a great deal of attention over the past year through the new federal coordinating council,¹ the recent Institute of Medicine (IOM) report,² and the American Recovery & Reinvestment Act (ARRA) stimulus funding.³ CER has a broad scope as defined by the IOM, addressing "...the generation and synthesis of evidence that compares the benefits and harms of alternative methods to prevent, diagnose, treat and monitor a clinical condition or to improve the delivery of care. The purpose of CER is to assist consumers, clinicians, purchasers, and policymakers to make informed decisions that will improve health care at both the individual and population levels."²

So what's new? As pharmacoepidemiologists, we could point out that we have generated evidence on health-relevant drug benefits and harms at the population level for over 25 years (the International Society for Pharmacoepidemiology just had its 26th International conference). And we could point out that we have moved from using untreated as a comparison group when assessing drug effects towards using treated with a realistic clinical alternative as a comparator group well before CER (e.g.,⁴). And we acknowledge that CER is much broader than pharmacoepidemiology. What is really new for us is the implicit acknowledgement of the need for and value of nonexperimental evidence for benefits. Until recently (e.g.,⁵), the Food and Drug Administration has largely dismissed nonexperimental evidence on drug benefits (e.g.,⁶), mainly because of fear about intractable confounding by indication.^{7, 8} The Food and Drug Administration also insists on comparing treatments with untreated (including placebo) and does not accept the idea of a comparator drug for proof of efficacy.⁹ It is interesting to see that other government agencies, including the Agency for Health Care Research and Quality through their network of DEcIDE (Developing Evidence to Inform Decisions about Effectiveness) Centers,¹⁰ see this differently. It is also encouraging to see that some divisions of the Food and Drug Administration, e.g., the Division of Epidemiology in the Office of Surveillance and Biometrics, Center for Devices and Radiological Health, are pioneers in recognizing the value of nonexperimental research.⁵

When we focus on the nonexperimental evaluation of the use and beneficial and harmful effects of drugs in the population (i.e., pharmacoepidemiology), we have to acknowledge

some specific aspects of drugs (Greek: Pharmacon) that we need to take into account. Drugs are the mainstay of contemporary medicine. They are used for the primary, secondary, and tertiary prevention of disease outcomes. A multibillion \$ industry constantly screens chemical compounds for physiologic effects. Drugs are marketed only after experimental proof of efficacy (benefits) in humans. They are therefore likely to affect disease outcomes.

Most prescription drugs are dispensed at pharmacies which then submit claims to payors, such as Medicare, Medicaid or private insurance companies. Likewise, physician visits, hospital stays, laboratory test results, procedures, injections, and other encounters with the healthcare system each generate a paper trail (or the electronic equivalent thereof). By linking these data across various sources (insurance claims, electronic health records, clinical records, billing data, laboratory results, vital statistics), an integrated picture of the patient's health and healthcare emerges. After being stripped of key identifiers, researchers can obtain permission to access these data with appropriate confidentiality safeguards.¹¹

These databases have unique advantages for epidemiologic research. Most are population-based and therefore less prone to the healthy (given the target) selection almost inherent in recruiting and consenting participants for randomized trials or cohort studies.¹² Linked healthcare databases include continuous service dates rather than the interval assessments (e.g., every two years) that are common in epidemiologic cohort studies and many large trials. Continuous assessment of exposure and outcomes allows us to be specific about timing, an important consideration given that the exposure coming before the outcome is arguably the only sine qua non condition for causality.¹³ Linked healthcare databases contain information on almost all drugs prescribed or dispensed in an outpatient setting. And they include codes for outpatient and inpatient diagnoses and procedures that the patient has received.¹¹ They also have major downsides. Without pretending to be exhaustive (this is neither within the scope of this commentary nor necessary for our argument) these include lack of data on important confounders, lack of data on drugs administered during hospitalization or purchased over the counter, lack of mortality data (in some databases), lack of data on the sensitivity and specificity of various algorithms to define outcomes, and lack of data on events not covered by the corresponding insurance plan.¹¹

Let us now step back from linked healthcare databases and highlight several important threats to the validity of nonexperimental research on drug effects in general. In addition to other sources of confounding, there is the potential for confounding by indication. Sicker patients are almost always more likely to be treated and are often more likely to have bad outcomes. If the severity of disease is unknown or measured with error, residual confounding will make the drug look bad. Recent work has focused on another kind of unmeasured confounding, confounding by frailty. Frailty is difficult to measure condition close to death that is not linked to a specific pathology but rather an overall (poor) state of health, and probably easily recognized by trained physicians. Unlike study populations in many other nonexperimental studies and trials, study populations assembled from large linked healthcare databases include very frail patients. Frailty may reduce the likelihood of a particular treatment if physicians focus on a patient's main medical problem and do not initiate useful therapies for secondary conditions.¹⁴ The practitioner may determine that in the presence of competing risks a therapy offers little expected benefit.¹⁵ Because frailty is

hard to measure and a very strong risk factor for poor outcomes (especially mortality), it will also lead to unmeasured and residual confounding. When comparing the treated with the untreated, however, frailty will often tend to make the drug look good. Frailty is a plausible explanation for paradoxical treatment-outcome associations observed in the elderly.^{16–18}

Besides confounding by unmeasured confounders and residual confounding, selection bias over time on treatment is a major problem when assessing longer term effects of drugs. Patients who adhere to treatments over prolonged periods of time tend to be healthier. Conversely, patients stopping treatments tend to be sicker. This leads to increasingly healthy users with increasing duration of treatment.^{19, 20} Similar to confounding by frailty, selection bias from healthy users tends to be again most pronounced for all-cause mortality.²¹ Those not adherent to placebo in randomized trials have been shown to have twice the mortality rates of those adherent to placebo.²² And there is the potential for immortal time bias.²³

None of these threats to validity is specific to pharmacoepidemiology or linked healthcare databases. For example, selection bias is a major issue in occupational epidemiology (healthy worker bias, e.g.,²⁴). Furthermore, only the potential for major confounding is specific to nonexperimental study designs. The lack of an inherent sampling structure compared with ad hoc studies may, however, increase the risk for flawed designs.

Pharmacoepidemiologic methodologists have made substantial progress in recent years addressing the above mentioned threats to validity. Recent developments include the new user design,²⁵ which allows us to focus on treatment initiation decision processes and hypothetical interventions. The new user design allows us to implement propensity scores²⁶ and instrumental variables²⁷, and assess positivity^{28, 29} and treatment contrary to prediction¹⁸. All these can be used to balance cohorts that have some equipoise to initiate all treatments compared with respect to the risk for the disease outcome. The new user design also allows us to address stopping, switching, and augmenting of drug use after baseline separate from the confounding at baseline. We are thus able to apply various methods for censoring, ranging from immediate censoring to ignoring (first treatment carried forward or intention to treat analysis). If we are able to predict adherence, we can apply marginal structural models.^{30–32} Finally, comparing patients initiating a drug to patients who initiate another drug for the same indication reduces the potential for and thus magnitude of most of the biases outlined above, including confounding, selection bias, and immortal time bias. The “comparative” in CER can therefore help us to avoid major biases when making nonexperimental treatment comparisons.

Based on these recent developments, we propose that study design has a larger influence on validity of pharmacoepidemiologic studies than whether we use linked healthcare databases or data from ad hoc studies. For instance, a reanalysis of the data from the Nurses Health Study on the effects of estrogen and progestin therapy on coronary heart disease in postmenopausal women based on a new user design and dealing with selection bias after initiation³³ showed results compatible with the ones from a large randomized trial. Threats to validity in pharmacoepidemiology are by no means specific to linked healthcare databases.

We need timely and trustworthy answers about both drug benefits and harms in the population. Such answers are essential to safeguard public health and they can rarely be obtained by data collection, be it within randomized controlled trials, large simple trials, or cohort studies. Current examples include insulin glargine and angiotensin receptor blockers, which have been implicated with increased risk for malignancies.^{34, 35} Fortunately in both cases we have obvious comparator drugs that will allow us to limit the potential for bias; but we need to be careful to avoid flaws in the study design³⁶.

The answer to the question whether we are brave or foolhardy may depend on what we want to do. We can be brave and study treatments with similar indications (e.g., insulin glargine vs. NPH insulin, angiotensin receptor blockers vs. angiotensin converting enzyme inhibitors), unintended effects, and short term effects. We would probably be foolhardy to study intended long term effects without a comparator, e.g., statins vs. no statin. Note that this list is not exhaustive. There will always be gray zones and it may very well be easier for academics to live with these than for industry and regulatory agencies.³⁷

We conclude that large linked healthcare databases offer major advantages for pharmacoepidemiologic research.³⁸ While certainly not ideal, the population is often close to ideal compared with ad hoc studies because it is unselected (e.g., Medicare, General Practice Research Database, Scandinavia³⁹). And while again not ideal, the information on drug exposure is almost ideal for prescription drugs in the outpatient setting, i.e., for most of the drugs used. We can study clinically relevant outcomes, and given the large size of linked healthcare databases, we can find timely answers to multiple questions without the need to wait for new data to be collected. Among the many downsides is the lack of information on important covariates but we have design options to limit and reduce confounding by unmeasured confounders and selection bias. Large linked healthcare databases allow us to answer some important questions that could otherwise not be answered. The future will allow us to link electronic medical records, cohort studies, and claims data. The ideal database will remain elusive, however.

While we need to increase the use large linked healthcare databases to answer appropriate pharmacoepidemiological and CER questions using state of the art nonexperimental methodology, we also need to acknowledge the need for more head to head, simple, large randomized trials comparing the effect of relevant treatment alternatives on clinically relevant outcomes (e.g., based on linked healthcare databases) in unselected populations.

Acknowledgments

The authors are the core faculty of the Pharmacoepidemiology Program within the Department of Epidemiology at the UNC Gillings School of Global Public Health. The core faculty combines backgrounds from medicine (Til Stürmer), psychology (Michele Jonsson Funk), health administration (Charles Poole), and biostatistics (Alan Brookhart) with strong training and expertise in epidemiology methods into a common interest in the development and assessment of innovative research methods, specifically for the nonexperimental evaluation of drug benefits and harms using large linked healthcare databases. Funding for research and training comes from various sources, including the National Institute on Aging (RO1 AG023178 and K25 AG27400); AHRQ (K02 HS17950); the UNC-GSK Center of excellence in Pharmacoepidemiology and Public Health, an innovative academia-industry collaboration; and unrestricted research grants from the pharmaceutical industry (e.g., Merck, Sanofi-Aventis, Amgen). UNC houses an AHRQ DEcIDE Center.

References

1. [Accessed September 9, 2010] Federal Coordinating Council for Comparative Effectiveness Report to the President and Congress. Jun 30. 2009 <http://www.hhs.gov/recovery/programs/cer/cerannualrpt.pdf>
2. Institute of Medicine. Initial National Priorities for Comparative Effectiveness Research. National Academies Press; Washington DC: 2009.
3. American Recovery and Reinvestment Act. [Accessed September 9, 2010] Comparative Effectiveness Research Funding. <http://www.hhs.gov/recovery/programs/cer/>
4. Brookhart MA, Wang PS, Solomon DH, Schneeweiss S. Evaluating short-term drug effects using a physician-specific prescribing preference as an instrumental variable. *Epidemiology*. 2006; 17:268–275. [PubMed: 16617275]
5. FDA Workshop: Methodologies for Post-Approval Studies of Medical Devices; September 9–10, 2009; <http://www.fda.gov/MedicalDevices/NewsEvents/WorkshopsConferences/ucm177140.htm>
6. Temple R. Problems in the use of large data sets to assess effectiveness. *Int J Technol Assess Health Care*. 1990; 6:211–219. [PubMed: 2203699]
7. Miettinen, Olli S. The need for randomization in the study of intended drug effects. *Stat Med*. 1983; 2:267–271. [PubMed: 6648141]
8. Yusuf, Salim; Collins, Rory; Peto, Richard. Why do we need some large, simple randomized trials? *Stat Med*. 1984; 3:409–420. [PubMed: 6528136]
9. The Social Medicine Portal. [Accessed September 9, 2010] FDA abandons Declaration of Helsinki for international clinical trials. <http://www.socialmedicine.org/2008/06/01/ethics/fda-abandons-declaration-of-helsinki-for-international-clinical-trials/>
10. Agency for Healthcare Research and Quality. [Accessed September 9, 2010] Effective Health Care Program. About the DEcIDE Network. <http://www.effectivehealthcare.ahrq.gov/index.cfm/who-is-involved-in-the-effective-health-care-program/about-the-decide-network/>
11. Strom, BL. *Pharmacoepidemiology*. 4. Chichester, UK: John Wiley & Sons Ltd; 2005.
12. Sesso HD, Gaziano JM, VanDenburgh M, Hennekens CH, Glynn RJ, Buring JE. Comparison of baseline characteristics and mortality experience of participants and nonparticipants in a randomized clinical trial: the Physicians' Health Study. *Control Clin Trials*. 2002; 23:686–702. [PubMed: 12505246]
13. Rothman, KJ.; Greenland, S.; Lash, TL. *Modern Epidemiology*. 3. Philadelphia: Lippincott Williams & Wilkins; 2008.
14. Redelmeier DA, Tan SH, Booth GL. The treatment of unrelated disorders in patients with chronic medical diseases. *N Engl J Med*. 1998; 338:1516–1520. [PubMed: 9593791]
15. Welch HG, Albertsen PC, Nease RF, Bubolz TA, Wasson JH. Estimating treatment benefits for the elderly: the effect of competing risks. *Ann Intern Med*. 1996; 124:577–584. [PubMed: 8597322]
16. Glynn RJ, Knight EL, Levin R, Avorn J. Paradoxical relations of drug treatment with mortality in older persons. *Epidemiology*. 2001; 12:682–689. [PubMed: 11679797]
17. Jackson LA, Jackson ML, Nelson JC, Neuzil KM, Weiss NS. Evidence of bias in estimates of influenza vaccine effectiveness in seniors. *Int J Epidemiol*. 2006; 35:337–344. [PubMed: 16368725]
18. Stürmer T, Rothman KJ, Avorn J, Glynn RJ. Treatment Effects in the Presence of Unmeasured Confounding: Dealing With Observations in the Tails of the Propensity Score Distribution--A Simulation Study. *Am J Epidemiol*. 2010 Aug 17. Epub ahead of print.
19. Brookhart MA, Patrick AR, Dormuth C, Avorn J, Shrank W, Cadarette SM, Solomon DH. Adherence to lipid-lowering therapy and the use of preventive health services: an investigation of the healthy user effect. *Am J Epidemiol*. 2007; 166:348–354. [PubMed: 17504779]
20. Brookhart MA, Stürmer T, Glynn RJ, Rassen J, Schneeweiss S. Confounding control in healthcare database research: challenges and potential approaches. *Med Care*. 2010; 48(6 Suppl):S114–120. [PubMed: 20473199]

21. Andersen M, Brookhart MA, Glynn RJ, Stovring H, Stürmer T. Practical issues in measuring cessation and re-initiation of drug use in databases [Abstract]. *Pharmacoepidemiol Drug Saf.* 2008; 17(suppl 1):S27.
22. Simpson SH, Eurich DT, Majumdar SR, Padwal RS, Tsuyuki RT, Varney J, Johnson JA. A meta-analysis of the association between adherence to drug therapy and mortality. *BMJ.* 2006; 333:15. [PubMed: 16790458]
23. Suissa S. Immortal time bias in observational studies of drug effects. *Pharmacoepidemiol Drug Saf.* 2007; 16:241–249. [PubMed: 17252614]
24. Pearce N, Checkoway H, Kriebel D. Bias in occupational epidemiology studies. *Occup Environ Med.* 2007; 64:562–568. [PubMed: 17053019]
25. Ray, Wayne A. Evaluating Medication Effects Outside of Clinical Trials: New-User Designs. *Am J Epidemiol.* 2003; 158:915–920. [PubMed: 14585769]
26. Glynn RJ, Schneeweiss S, Stürmer T. Indications for propensity scores and review of their use in pharmacoepidemiology. *Basic Clin Pharmacol Toxicol.* 2006; 98:253–259. [PubMed: 16611199]
27. Brookhart MA, Wang PS, Solomon DH, Schneeweiss S. Evaluating short-term drug effects using a physician-specific prescribing preference as an instrumental variable. *Epidemiology.* 2006; 17:268–275. [PubMed: 16617275]
28. Messer LC, Oakes JM, Mason S. Effects of socioeconomic and racial residential segregation on preterm birth: a cautionary tale of structural confounding. *Am J Epidemiol.* 2010; 171:664–673. [PubMed: 20139129]
29. Westreich D, Cole SR. Invited commentary: positivity in practice [Commentary]. *Am J Epidemiol.* 2010; 171:674–677. [PubMed: 20139125]
30. Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology.* 2000; 11:550–560. [PubMed: 10955408]
31. Hernán MA, Brumback B, Robins JM. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology.* 2000; 11:561–570. [PubMed: 10955409]
32. Cole SR, Hernán MA. Constructing inverse probability weights for marginal structural models. *Am J Epidemiol.* 2008; 168:656–664. [PubMed: 18682488]
33. Hernán MA, Alonso A, Logan R, Grodstein F, Michels KB, Willett WC, Manson JE, Robins JM. Observational studies analyzed like randomized experiments: an application to postmenopausal hormone therapy and coronary heart disease. *Epidemiology.* 2008; 19:766–779. [PubMed: 18854702]
34. Hemkens LG, Grouven U, Bender R, Günster C, Gutschmidt S, Selke GW, Sawicki PT. Risk of malignancies in patients with diabetes treated with human insulin or insulin analogues: a cohort study. *Diabetologia.* 2009; 52:1732–1744. [PubMed: 19565214]
35. Sipahi I, Debanne SM, Rowland DY, Simon DI, Fang JC. Angiotensin-receptor blockade and risk of cancer: meta-analysis of randomised controlled trials. *Lancet Oncol.* 2010; 11:627–636. [PubMed: 20542468]
36. Pocock SJ, Smeeth L. Insulin glargine and malignancy: an unwarranted alarm [Letter]. *The Lancet.* 2009; 374:511–513.
37. Jan P, Vandenbroucke. Observational research, randomized trials, and two views of medical science. *PLoS Medicine.* 2008; 5:e67. [PubMed: 18336067]
38. ISPE. Guidelines for good pharmacoepidemiology practices (GPP). *Pharmacoepidemiol Drug Saf.* 2008; 17:200–208. [PubMed: 17868186]
39. Frank L. Epidemiology: When an entire country is a cohort. *Science.* 2000; 287:2398–2399. [PubMed: 10766613]