

Chapter 7

Using EHR to Conduct Outcome and Health Services Research

Laura Myers and Jennifer Stevens

Take Home Messages

- Electronic Health Records have become an essential tool in clinical research, both as a supplement to existing methods, but also in the growing domains of outcomes research and analytics.
- While EHR data is extensive and analytics are powerful, it is essential to fully understand the biases and limitations introduced when used in health services research.

7.1 Introduction

Data from electronic health records (EHR) can be a powerful tool for research. However, researchers must be aware of the fallibility of data collected for clinical purposes and of biases inherent to using EHR data to conduct sound health outcomes and health services research. Innovative methods are currently being developed to improve the quality of data and thus our ability to draw conclusions from studies that use EHR data.

The United States devotes a large share of the Gross Domestic Product (17.6 % in 2009) to health care [1]. With such a huge financial and social investment in healthcare, important questions are fundamental to evaluating this investment:

How do we know what treatment works and for which patients?

How much should health care cost? When is too much to pay? In what type of care should we invest more or less resources?

How does the health system work and how could it function better?

Health services research is a field of research that lives at the intersection of health care policy, management, and clinical care delivery and seeks to answer

these questions. Fundamentally, health services research places the health system under the microscope as the organism of study.

To begin to address these questions, researchers need large volumes of data across multiple patients, across different types of health delivery structures, and across time. The simultaneous growth of this field of research in the past 15 years has coincided with the development of the electronic health record and the increasing number of providers who make use of them in their workspace [2]. The EHR provides large quantities of raw data to fuel this research, both at the granular level of the patient and provider and at the aggregated level of the hospital, state, or nation.

Conducting research with EHR data has many challenges. EHR data are riddled with biases, collected for purposes other than research, inputted by a variety of users for the same patient, and difficult to integrate across health systems [See previous chapter “Confounding by Indication”]. This chapter will focus on the attempts to capitalize on the promise of the EHR for health services research with careful consideration of the challenges researchers must address to derive meaningful and valid conclusions.

7.2 The Rise of EHRs in Health Services Research

7.2.1 *The EHR in Outcomes and Observational Studies*

Observational studies, either retrospective or prospective, attempt to draw inferences about the effects of different exposures. Within health services research, these exposures include both different types of clinical exposures (e.g., does hormone replacement therapy help or hurt patients?) and health care delivery exposures (e.g., does admission to a large hospital for cardiac revascularization improve survival from myocardial infarction over admission to a small hospital). The availability of the extensive health data in electronic health records has fueled this type of research, as data extraction and transcription from paper records has ceased to be a barrier to research. These studies capitalize on the demographic and clinical elements that are routinely recorded as part of an encounter with the health system (e.g., age, sex, race, procedures performed, length of stay, critical care resources used).

We have highlighted a number of examples of this type of research below. Each one is an example of research that has made use of electronic health data, either at the national or hospital level, to draw inferences about health care delivery and care.

Does health care delivery vary? The researchers who compile and examine the Dartmouth Atlas have demonstrated substantial geographic variation in care. In their original article in *Science*, Wennberg and Gittlesohn noted wide variations in the use of health services in Vermont [3]. These authors employed data derived from the use of different types of medical services—home health services, inpatient discharges, etc.—to draw these inferences. Subsequent investigations into national variation in care have been able to capitalize on the availability of such data electronically [4].

Do hospitals with more experience in a particular area perform better? Birkmeyer and colleagues studied the intersection of hospital volume and surgical outcomes with absolute differences in adjusted mortality rates between low volume hospitals and high volume hospitals ranging from 12 % for pancreatic resection to 0.2 % for carotid endarterectomy [5]. Kahn et al. also used data available in over 20,000 patients to demonstrate that mortality associated with mechanical ventilation was 37 % lower in high volume hospitals compared with low volume hospitals [6]. Both of these research groups made use of large volumes of clinical and claims data—Medicare claims data in the case of Birkmeyer and colleagues and the APACHE database from Cerner for Kahn et al.—to ask important questions about where patients should seek different types of care.

How can we identify harm to patients despite usual care? Herzig and colleagues made use of the granular EHR at a single institution and found that the widely-prescribed medications that suppress acid production were associated with an increased risk of pneumonia [7]. Other authors have similarly looked at the EHR found that these types of medications are often continued on discharge from the hospital [8, 9].

To facilitate appropriate modeling and identification of confounders in observational studies, researchers have had to devise methods to extract markers of diagnoses, severity of illness, and patient comorbidities using only the electronic fingerprint. Post et al. [10] developed an algorithm to search for patients who had diuretic-refractory hypertension by querying for patients who had a diagnosis of hypertension despite 6 months treatment with a diuretic. Previously validated methods for reliably measuring the severity of a patient’s illness, such as APACHE or SAPS scores [11, 12], have data elements that are not easily extracted in the absence of manual inputting of data. To meet these challenges, researchers such as Escobar and Elixhauser have proposed alternative, electronically derived methods for both severity of illness measures [13, 14] and identification of comorbidities [14]. Escobar’s work, with a severity of illness measure with an area under the curve of 0.88, makes use of highly granular electronic data including laboratory values; Elixhauser’s comorbidity measure is publically available through the Agency for Healthcare Research and Quality and solely requires billing data [15].

Finally, researchers must develop and employ appropriate mathematical models that can accommodate the short-comings of electronic health data or else they risk drawing inaccurate conclusions. Examples of such modeling techniques are extensive have included propensity scores, causal methods such as marginal structural models and inverse probability weights, and designs from other fields such as instrumental variable analysis [16–19]. The details of these methods are discussed elsewhere in this text.

7.2.2 The EHR as Tool to Facilitate Patient Enrollment in Prospective Trials

Despite the power of the EHR to conduct health services and outcomes research retrospectively, the gold standard in research remains prospective and randomized trials. The EHR has functioned as a valuable tool to screen patients at a large scale

for eligibility. In this instance, research staff uses the data available through the electronic record as a high-volume screening technique to target recruitment efforts to the most appropriate patients. Clinical trials that develop electronic strategies for patient identification and recruitment are at an even greater advantage, although such robust methods have been described as sensitive but not specific, and frequently require coupling screening efforts with manual review of individual records [20]. Embi et al. [21] have proposed using the EHR to simultaneously generate Clinical Trial Alerts, particularly in commercial EHRs such as Epic to leverage the EHR in a point of care strategy. This strategy could expedite enrollment although it must be weighed against the risk of losing patient confidentiality, an ongoing tension between patient care and clinical trial enrollment [22].

7.2.3 The EHR as Tool to Study and Improve Patient Outcomes

Quality can also be tracked and reported through EHRs, either for internal quality improvement or for national benchmarking; the Veterans' Affairs' (VA) healthcare system highlights this. Byrne et al. [23] reported that in the 1990s, the VA spent more money on information technology infrastructure and achieved higher rates of adoption compared to the private sector. Their home-grown EHR, which is called VistA, provided a way to track preventative care processes such as cancer and diabetes screening through electronic pop up messages. Between 2004 and 2007, they found that the VA system achieved better glucose and lipid control for diabetics compared to a Medicare HMO benchmark [23]. While much capital investment was needed during the initial implementation of VistA, it is estimated that adopting this infrastructure saved the VA system \$3.09 billion in the long term. It also continues to be a source of quality improvement as quality metrics evolve over time [23].

7.3 How to Avoid Common Pitfalls When Using EHR to Do Health Services Research

We would propose the following hypothetical research study as a case study to highlight common challenges to conducting health services research with electronic health data:

Proposed research study: Antipsychotic medications (e.g. haloperidol) are prescribed frequently in the intensive care unit to treat patients with active delirium. However, these medications have been associated with their own potential risk of harm [24] that is separate from the overall risk of harm from delirium. The researchers are interested in whether treatment with antipsychotics increases the risk of in-hospital death and increases the cost of care and use of resources in the hospital.

7.3.1 Step 1: Recognize the Fallibility of the EHR

The EHR is rarely complete or correct. Hogan et al. [25] tried to estimate how complete and accurate data are in studies that are conducted on an EHR, finding significant variability in both. Completeness ranged from 31 to 100 % and correctness ranged from 67 to 100 % [25]. Table 7.1 highlights examples of different diagnoses and possible sources of data, which may or may not be present for all patients.

Proposed research study: The researchers will need to extract which patients were exposed to antipsychotics and which were not. However, there is unlikely to be one single place where this information is stored. Should they use pharmacy dispensing data? Nursing administration data? Should they look at which patients were charged for the medications? What if they need these data from multiple hospitals with different electronic health records?

Additionally, even with a robust data extraction strategy, the fidelity of different types of data is variable [26–33]. For example, many EHR systems have the option of entering free text for a medical condition, which may be spelled wrong or be worded unconventionally. As another example, the relative reimbursement of a particular billing code may influence the incidence of that code in the electronic health record so billing may not reflect the true incidence and prevalence of the disease [34, 35].

7.3.2 Step 2: Understand Confounding, Bias, and Missing Data When Using the EHR for Research

We would highlight the following methodological issues inherent in conducting research with electronic health records: selection bias, confounding, and missing data. These are explored in greater depth in other chapters of this text.

Table 7.1 Examples of the range of data elements that may be used to identify patients with either ischemic heart disease or acute lung injury through the electronic health record

Disease state	Data source	Example
Ischemic heart disease	Billing data	ICD-9 code 410 [48]
	Laboratory data	Positive troponin during admission
	Physician documentation	In the discharge summary: “the patient was noted to have ST elevations on ECG and was taken to the cath lab”...
Acute lung injury	Billing data	ICD-9 code 518.5 and 518.82 with the procedural codes 96.70, 96.71 and 96.72 for mechanical ventilation [49]
	Radiology data	“Bilateral” and “infiltrates” on chest x-ray reads [50]
	Laboratory data	PaO ₂ /FiO ₂ < 300 mmHg

Selection bias, or the failure of the population of study to represent the generalizable population, can occur if all the patients, including controls, are already seeking medical care within an EHR-based system. For example, in EHR-based studies comparing medical versus surgical approaches to the same condition may not be comparing equivalent patients in each group; patients seeking a surgical correction may fundamentally differ from those seeking a more conservative approach. Hripcsak et al. [36] used a large clinical data set from a tertiary center in 2007 to compare mortality from pneumonia to a hand-collected data set that had been published previously; the different search criteria altered the patient population and the subsequent risk of death. While it is not eliminated entirely, selection bias is reduced when prospective randomization takes place [37].

Confounding bias represents the failure to appropriately account for an additional variable that influences both the dependent and independent variable. In research with electronic health records, confounding represents a particular challenge, as identification of all possible confounding variables is nearly impossible.

Proposed research study: The researchers in this study are interested in the patient-level outcomes of what happens to those patients exposed to antipsychotics during their stay. But patients who are actively delirious while in the ICU are likely to be sicker than those who are not actively delirious and sicker patients require more hospital resources. As a result, antipsychotics will appear to be associated with a higher risk of in-hospital mortality and use of hospital resources not due to the independent effect of the drug but rather as a result of confounding by indication.

Missing data or unevenly sampled data collected as part of the EHR creates its own complex set of challenges for health services research. For example, restricting the analysis to patients with only a complete set of data may yield very different (and poorly generalizable) inferences. The multidimensionality of this problem often goes unexamined and underestimated. Nearly all conventional analytic software presumes completeness of the matrix of data, leading many researchers to fail to fully address these issues. For example, data can be misaligned due to lack of sampling, missing data, or simple misalignment. In other words, the data could not be measured during a period of time for an intentional reason (e.g., a patient was extubated and therefore no values for mechanical ventilation were documented) and should not be imputed or the data was measured but was unintentionally not recorded and therefore can be imputed. Rusanov et al. studied 10,000 outpatients at a tertiary center who underwent general anesthesia for elective procedures. Patients with a higher risk of adverse outcome going into surgery had more data points including laboratory values, medication orders and possibly admission orders compared to less sick patients [38], making the missing data for less sick patients intentional. Methods for handling missing data have included omitting cases are not complete, pairwise deletion, mean substitution, regression substitution, or using modeling techniques for maximum likelihood and multiple imputation [39].

7.4 Future Directions for the EHR and Health Services Research

7.4.1 Ensuring Adequate Patient Privacy Protection

It is controversial whether using EHR for research goes against our national privacy standard. In large cohorts, many patients may be present with the same health information, thereby rendering the data sufficiently deidentified. Further, Ingelfinger et al. acknowledge that countries with healthcare registries such as Scandinavia have a distinct research advantage [40]. However, health information is a protected class of information under the Health Insurance Portability and Accountability Act, so there is significant awareness among U.S. healthcare professionals and researchers about its proper storage and dissemination. Some argue that patients should be consented (versus just notified) that their information could be used for research purposes in the future. Ingelfinger et al. [40] recommends IRB approval of registries and a rigorous deidentification process.

Public perception on the secondary use of EHR may not be as prohibitive as policymakers may have believed. In a survey of 3300 people, they were more willing to have their information used for research by university hospitals, compared to public health departments or for quality improvement purposes [41]. They were much less willing to contribute to marketing efforts or have the information used by pharmaceutical companies [41].

With the growing amount of information being entered into EHRs across the country, the American Medical Informatics Association convened a panel to make recommendations for how best to use EHR securely for purposes other than direct patient care. In 2006, the panel called for a national standard to deal with the issue of privacy. They described complex situations where there were security breaches due to problems with deidentification or data was being sold by physicians for profit [42]. While the panel demanded that the national framework be transparent, comprehensive and publicly accepted, they did not propose a particular standard at that time [42]. Other groups such as the Patient-Centered Outcomes Research Institute have since addressed the same conflict in a national forum in 2012. Similarly, while visions were discussed, no explicit recommendation was set forth [PCORI]. Controversy continues in this area.

7.5 Multidimensional Collaborations

Going forward, the true power of integrated data can only be harnessed by forming more collaborations, both within institutions and between them. Research on a national scale in the U.S. has been shown to be feasible. The FDA implemented a pilot program in 2009 called the Mini-Sentinel program. It brought together 31 academic and private organizations to monitor for safety events related to

medications and devices currently on the market [43]. Admittedly, merging databases may require significant financial resources, especially if the datasets need to be coded and/or validated, but researchers like Bradley et al. [44] believe this is a cost-effective use of grant money because of the vast potential to make advances in the way we deliver care. Fundamental to the feasibility of multidimensional collaborations is the ability to ensure accuracy of large-scale data and integrate it across multiple health record technologies and platforms. Efforts to ensure data quality and accessibility must be promoted alongside patient privacy.

7.6 Conclusion

Researchers continue to ask fundamental questions of our health system, making use of the deluge of data generated by EHRs. Unfortunately, that deluge is messy and problematic. As the field of health services research with EHRs continues to evolve, we must hold researchers to rigorous standards [45] and encourage more investment in research-friendly clinical databases as well as cross-institutional collaborations. Only then will the discoveries in health outcomes and health services research be one click away [46, 47]. It is time for healthcare to reap the same reward from a rich data source that is already in existence.

Open Access This chapter is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, duplication, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, a link is provided to the Creative Commons license and any changes made are indicated.

The images or other third party material in this chapter are included in the work's Creative Commons license, unless indicated otherwise in the credit line; if such material is not included in the work's Creative Commons license and the respective action is not permitted by statutory regulation, users will need to obtain permission from the license holder to duplicate, adapt or reproduce the material.

References

1. Center for Medicare and Medicaid Services (2015) National health expenditure data fact sheet
2. Jha AK, DesRoches CM, Campbell EG, Donelan K, Rao SR et al (2009) Use of electronic health records in U.S. hospitals. *N Engl J Med* 360:1628–1638
3. Wennberg J, Gittelsohn (1973) Small area variations in health care delivery. *Science* 182:1102–1108
4. Stevens JP, Nyweide D, Maresh S, Zaslavsky A, Shrank W et al (2015) Variation in inpatient consultation among older adults in the United States. *J Gen Intern Med* 30:992–999
5. Birkmeyer JD (2000) Relation of surgical volume to outcome. *Ann Surg* 232:724–725
6. Kahn JM, Goss CH, Heagerty PJ, Kramer AA, O'Brien CR et al (2006) Hospital volume and the outcomes of mechanical ventilation. *N Engl J Med* 355:41–50

7. Herzig SJ, Howell MD, Ngo LH, Marcantonio ER (2009) Acid-suppressive medication use and the risk for hospital-acquired pneumonia. *JAMA* 301:2120–2128
8. Murphy CE, Stevens AM, Ferrentino N, Crookes BA, Hebert JC et al (2008) Frequency of inappropriate continuation of acid suppressive therapy after discharge in patients who began therapy in the surgical intensive care unit. *Pharmacotherapy* 28:968–976
9. Zink DA, Pohlman M, Barnes M, Cannon ME (2005) Long-term use of acid suppression started inappropriately during hospitalization. *Aliment Pharmacol Ther* 21:1203–1209
10. Post AR, Kurc T, Cholleti S, Gao J, Lin X et al (2013) The analytic information warehouse (AIW): a platform for analytics using electronic health record data. *J Biomed Inform* 46:410–424
11. Zimmerman JE, Kramer AA, McNair DS, Malila FM (2006) Acute physiology and chronic health evaluation (APACHE) IV: hospital mortality assessment for today's critically ill patients. *Crit Care Med* 34:1297–1310
12. Moreno RP, Metnitz PG, Almeida E, Jordan B, Bauer P et al (2005) SAPS 3—from evaluation of the patient to evaluation of the intensive care unit. Part 2: development of a prognostic model for hospital mortality at ICU admission. *Intensive Care Med* 31:1345–1355
13. Escobar GJ, Greene JD, Scheirer P, Gardner MN, Draper D et al (2008) Risk-adjusting hospital inpatient mortality using automated inpatient, outpatient, and laboratory databases. *Med Care* 46:232–239
14. Elixhauser A, Steiner C, Harris DR, Coffey RM (1998) Comorbidity measures for use with administrative data. *Med Care* 36:8–27
15. Project HCaU (2015) Comorbidity software, Version 3.7
16. Rubin DB, Thomas N (1996) Matching using estimated propensity scores: relating theory to practice. *Biometrics* 52:249–264
17. Rubin DB (1997) Estimating causal effects from large data sets using propensity scores. *Ann Intern Med* 127:757–763
18. Howell MD, Novack V, Grgurich P, Soulliard D, Novack L et al (2010) Iatrogenic gastric acid suppression and the risk of nosocomial *Clostridium difficile* infection. *Arch Intern Med* 170:784–790
19. Hernan MA, Brumback B, Robins JM (2000) Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology* 11:561–570
20. Thadani SR, Weng C, Bigger JT, Ennever JF, Wajngurt D (2009) Electronic screening improves efficiency in clinical trial recruitment. *J Am Med Inform Assoc* 16:869–873
21. Embi PJ, Jain A, Clark J, Harris CM (2005) Development of an electronic health record-based clinical trial alert system to enhance recruitment at the point of care. *AMIA Annu Symp Proc*, 231–235
22. PCORnet (2015) Rethinking clinical trials: a living textbook of pragmatic clinical trials
23. Byrne CM, Mercincavage LM, Pan EC, Vincent AG, Johnston DS et al (2010) The value from investments in health information technology at the U.S. Department of Veterans Affairs. *Health Aff (Millwood)* 29:629–638
24. Ray WA, Chung CP, Murray KT, Hall K, Stein CM (2009) Atypical antipsychotic drugs and the risk of sudden cardiac death. *N Engl J Med* 360:225–235
25. Hogan WR, Wagner MM (1997) Accuracy of data in computer-based patient records. *J Am Med Inform Assoc* 4:342–355
26. Lee DS, Donovan L, Austin PC, Gong Y, Liu PP et al (2005) Comparison of coding of heart failure and comorbidities in administrative and clinical data for use in outcomes research. *Med Care* 43:182–188
27. Iwashyna TJ, Odden A, Rohde J, Bonham C, Kuhn L et al (2014) Identifying patients with severe sepsis using administrative claims: patient-level validation of the angus implementation of the international consensus conference definition of severe sepsis. *Med Care* 52:e39–e43
28. Jones G, Taright N, Boelle PY, Marty J, Lalande V et al (2012) Accuracy of ICD-10 codes for surveillance of *clostridium difficile* infections, France. *Emerg Infect Dis* 18:979–981
29. Kramer JR, Davila JA, Miller ED, Richardson P, Giordano TP et al (2008) The validity of viral hepatitis and chronic liver disease diagnoses in Veterans Affairs Administrative databases. *Aliment Pharmacol Ther* 27:274–282

30. van de Garde EM, Oosterheert JJ, Bonten M, Kaplan RC, Leufkens HG (2007) International classification of diseases codes showed modest sensitivity for detecting community-acquired pneumonia. *J Clin Epidemiol* 60:834–838
31. Movig KL, Leufkens HG, Lenderink AW, Egberts AC (2003) Validity of hospital discharge International classification of diseases (ICD) codes for identifying patients with hyponatremia. *J Clin Epidemiol* 56:530–535
32. Sickbert-Bennett EE, Weber DJ, Poole C, MacDonald PD, Maillard JM (2010) Utility of international classification of diseases, ninth revision, clinical modification codes for communicable disease surveillance. *Am J Epidemiol* 172:1299–1305
33. Jung MA, Banerjee SN (2009) Administrative coding data and health care-associated infections. *Clin Infect Dis* 49:949–955
34. O'Malley KJ, Cook KF, Price MD, Wildes KR, Hurdle JF et al (2005) Measuring diagnoses: ICD code accuracy. *Health Serv Res* 40:1620–1639
35. Richesson RL, Rusincovitch SA, Wixted D, Batch BC, Feinglos MN et al (2013) A comparison of phenotype definitions for diabetes mellitus. *J Am Med Inform Assoc* 20:e319–e326
36. Hripcsak G, Knirsch C, Zhou L, Wilcox A, Melton G (2011) Bias associated with mining electronic health records. *J Biomed Discov Collab* 6:48–52
37. Hernan MA, Alonso A, Logan R, Grodstein F, Michels KB et al (2008) Observational studies analyzed like randomized experiments: an application to postmenopausal hormone therapy and coronary heart disease. *Epidemiology* 19:766–779
38. Rusanov A, Weiskopf NG, Wang S, Weng C (2014) Hidden in plain sight: bias towards sick patients when sampling patients with sufficient electronic health record data for research. *BMC Med Inform Decis Mak* 14:51
39. Allison PD (2001) Missing data. Sage Publishers, Thousand Oaks
40. Ingelfinger JR, Drazen JM (2004) Registry research and medical privacy. *N Engl J Med* 350:1452–1453
41. Grande D, Mitra N, Shah A, Wan F, Asch DA (2013) Public preferences about secondary uses of electronic health information. *JAMA Intern Med* 173:1798–1806
42. Safran C, Bloomrosen M, Hammond WE, Labkoff S, Markel-Fox S et al (2007) Toward a national framework for the secondary use of health data: an American medical informatics association white paper. *J Am Med Inform Assoc* 14:1–9
43. Platt R, Carnahan RM, Brown JS, Chrischilles E, Curtis LH et al (2012) The U.S. food and drug administration's mini-sentinel program: status and direction. *Pharmacoepidemiol Drug Saf* 21(Suppl 1):1–8
44. Bradley CJ, Penberthy L, Devers KJ, Holden DJ (2010) Health services research and data linkages: issues, methods, and directions for the future. *Health Serv Res* 45:1468–1488
45. Kahn MG, Raebel MA, Glanz JM, Riedlinger K, Steiner JF (2012) A pragmatic framework for single-site and multisite data quality assessment in electronic health record-based clinical research. *Med Care* 50(Suppl):S21–S29
46. Weber GM, Mandl KD, Kohane IS (2014) Finding the missing link for big biomedical data. *JAMA* 311:2479–2480
47. Murdoch TB, Detsky AS (2013) The inevitable application of big data to health care. *JAMA* 309:1351–1352
48. Birman-Deych E, Waterman AD, Yan Y, Nilasena DS, Radford MJ et al (2005) Accuracy of ICD-9-CM codes for identifying cardiovascular and stroke risk factors. *Med Care* 43:480–485
49. Reynolds HN, McCunn M, Borg U, Habashi N, Cottingham C et al (1998) Acute respiratory distress syndrome: estimated incidence and mortality rate in a 5 million-person population base. *Crit Care* 2:29–34
50. Herasevich V, Tsapenko M, Kojicic M, Ahmed A, Kashyap R et al (2011) Limiting ventilator-induced lung injury through individual electronic medical record surveillance. *Crit Care Med* 39:34–39