Electronic Health Record–Enabled Research in Children Using the Electronic Health Record for Clinical Discovery

Scott M. Sutherland, MD[a,b,*], David C. Kaelber, MD, PhD, MPH[c],
N. Lance Downing, MD[b], Veena V. Goel, MD[a,b],
Christopher A. Longhurst, MD, MS[d]

KEYWORDS

- EHR • EMR • Electronic health record • Electronic medical record • Research
- Clinical discovery • Children

KEY POINTS

- The electronic health record (EHR) contains a massive amount of discrete patient data that are generated through the routine provision of patient care.
- EHR data can be so-called big data based on volume (total number of patients/data points), velocity (the rate at which it is generated), and/or variety.
- Data validation is imperative because many of the data were collected for clinical, rather than research, purposes.
- EHR data can be used to build large patient cohorts and/or identify patients with rare conditions, allowing pediatric researchers to overcome small samples sizes.
- The EHR can be used for interventional studies and prospective trials at the point of care.

Disclosure: None of the authors have anything to disclose.
[a] Department of Pediatrics, Stanford University School of Medicine, 300 Pasteur Drive, Room G-306, Stanford, CA 94304, USA; [b] Department of Clinical Informatics, Stanford Children's Health, 1265 Welch Road, MSOB XIC65A, Stanford, CA 94305, USA; [c] Departments of Information Services, Internal Medicine, Pediatrics, Epidemiology and Biostatistics, Center for Clinical Informatics Research and Education, The MetroHealth System, Case Western Reserve University, 2500 MetroHeatlh Drive, Cleveland, OH 44109, USA; [d] Department of Biomedical Informatics, UC San Diego School of Medicine, 9560 Towne Centre Drive, San Diego, CA 92121, USA
* Corresponding author. 300 Pasteur Drive, Room G-306, Stanford, CA 94304.
E-mail address: suthersm@stanford.edu

INTRODUCTION

Although electronic health records (EHRs) were first described more than 50 years ago, it is only recently that EHRs have become pervasive.[1] Notably, the Unites States saw EHR adoption triple between 2009 and 2013; adoption among children's hospitals increased from 21% in 2008 to 59% in 2011.[2,3] The growth seen over the last 5 years is likely to be progressive and sustainable as, with time, countries such as Norway, the Netherlands, New Zealand, and the United Kingdom have achieved near-universal EHR adoption.[4] As experience with EHRs has increased, clinicians and researchers have begun to see the EHR in a different light; as a colossal database and interventional tool at the intersection of physicians, patients, and care delivery.[5,6] In the past, research databases, tools, and records have been separate from clinical databases, tools, and records; EHRs blur these distinctions and merge these silos.

In parallel with increased EHR adoption, there has been the evolution of clinical informatics, the "scientific and medical field that concerns itself with the cognitive, information processing and communication tasks of medical practice, education and research, including the information science and the technology to support these tasks."[7] More specifically, clinical research informatics, which generates the tools and techniques to use the EHR for clinical research, has been developed within the larger field of clinical informatics.[8]

The combination of increased EHR implementation, the exponential growth of digital data available within the EHR, and the development of clinical research informatics tools and techniques present a unique, previously unavailable opportunity to enhance clinical discovery and improve patient care. This phenomenon has particular relevance in children's health care. Pediatric research often involves small samples sizes[9]; by exploiting the EHR data sets across 1 or more institutions, investigators have access to a vast number of patients and can mitigate concerns about statistical power and significance. In addition, when faced with clinical decisions, pediatricians are often forced to rely on studies and trials that have been performed in adults; trials are often not repeated in children because of cost, safety, and redundancy issues.[10] EHR data can be used to retrospectively assess the impact of clinical decisions, making it possible to determine the efficacy of a certain intervention or the safety of a particular medication. In addition, when pediatric trials are implemented, they are often hampered by slow recruitment and erratic referrals.[11] The EHR can be used to facilitate recruitment for trials across institutions and improve the efficacy of trial procedures.[4,12] It is within this context that this article presents an overview of EHR-enabled clinical discovery, focusing on concepts and studies relevant to pediatrics.

THE ELECTRONIC HEALTH RECORD DATA SET

The EHR data set is immense; the scale is on par with many of the big data disciplines such as genomics and proteomics. Across an entire children's hospital, clinical care generates hundreds of thousands of data points per day and tens of millions of data points annually; data generated from ambulatory care and the narrative data contained within clinical notes add substantially more information. However, although the volume of data is alluring, some elements are easier to extract, some have higher fidelity, and some require validation.[13] Thus, an understanding of the types and quality of data available is essential to using EHRs for clinical discovery.

Electronic Health Record Data Elements

The EHR contains clinical data generated through the routine provision of care: physician orders, test results, vital signs, demographics, progress notes, medication