Feasibility of Integrating Research Data Collection into Routine Clinical Practice Using the Electronic Health Record

David J. Chalmers, Sara J. Deakyne, Marisa L. Payan, Michelle R. Torok, Michael G. Kahn and Vijaya M. Vemulakonda*

From the Division of Urology, Department of Surgery (DJC, VMV), Children's Outcomes Research Program (MLP, MRT) and Department of Pediatrics, Section of Pediatric Epidemiology (MGK), University of Colorado Anschutz Medical Campus, and Department of Pediatric Urology (DJC, VMV) and Department of Research Informatics (SJD), Children's Hospital Colorado, Aurora, Colorado

Purpose: The electronic health record is becoming central to routine medical practice and has the potential to facilitate large scale clinical research. We evaluated the completeness and accuracy of data collection using designated research fields integrated into a semistructured clinical note. We hypothesized that prospective research data collection as part of routine clinical charting is feasible, with a high rate of utilization (greater than 80%) and accuracy (kappa greater than 0.80).

Materials and Methods: Infants with congenital hydronephrosis were followed prospectively at a single institution. Existing functionality in the electronic health record was used for data collection by creation of 28 different data elements captured from a hydronephrosis note or phrase template. Completeness (percent utilization) was calculated and accuracy was assessed by comparing the structured data to manual chart review. Comparisons were conducted using the chi-square test, with 2-tailed p values <0.05 considered statistically significant.

Results: A total of 80 patients were eligible for manual chart review. Data were recorded through template use in 64 patients for an overall completeness of 80.0%. Of 28 elements 17 (60%) demonstrated "almost perfect" agreement (kappa greater than 0.80), and all variables reached at least "moderate" agreement (greater than 0.40).

Conclusions: Integrating research fields into routine clinical practice is feasible by using semistructured clinical templates within an electronic health record. High completion and accuracy rates were captured from a variety of fields within a hydronephrosis template.

Key Words: data collection, electronic health records, research

Abbreviations and Acronyms

EHR = electronic health record

Accepted for publication April 21, 2014.

Study received institutional review board approval (Colorado Multiple Institutional Review Board No. 11-1748).

Supported by a pilot grant from Children's Hospital Colorado Research Institute.

* Correspondence: Division of Urology, Department of Surgery, University of Colorado Anschutz Medical Campus, 13123 E. 16th Ave., Aurora, Colorado 80045 (telephone: 720-777-4052; FAX: 720-777-7370; e-mail: vijaya.vemulakonda@childrenscolorado.org).

THE electronic health record has become an integral part of medical practice. The Institute of Medicine identifies implementation of a national health care information technology infrastructure as essential to improving American health care.[1–5] Furthermore, federal legislation has incentivized "meaningful use" of the electronic health record to improve quality of care, case management and medical research.[6,7] Given the federal focus on comparative effectiveness research, the electronic health record has been identified as a tool to facilitate large scale clinical research.[8,9]

0022-5347/14/1924-1215/0
THE JOURNAL OF UROLOGY®

http://dx.doi.org/10.1016/j.juro.2014.04.091
Vol. 192, 1215-1220, October 2014
Printed in U.S.A.

Despite expanding EHR adoption, its potential remains largely untapped. One barrier to the use of EHRs in research has been the reliance on free text documentation in the course of clinical care.[10] Recent studies suggest that the use of structured documentation, including templated notes, may improve data extraction for research purposes.[10–12] However, to date no study has evaluated this approach in the pediatric surgical or pediatric urology ambulatory care setting. We evaluated the feasibility, completeness and accuracy of data collection using designated research data fields integrated into a semistructured clinical note. We hypothesized that integration of prospective research data collection into routine clinical charting is feasible, with high utilization (greater than 80%) and accuracy (kappa greater than 0.80), similar to previous studies assessing EHR based data capture.[10,13,14]

METHODS

After institutional review board approval was obtained we retrospectively reviewed data on a monthly basis from an ongoing prospective cohort study of infants with congenital hydronephrosis at Children's Hospital Colorado. Patients were included in the study if they were younger than 1 year, seen between January and June 2013, and had a diagnosis of hydronephrosis based on clinical documentation or ICD-9 code (591, 753.29) in the encounter or problem list. Patients were excluded if 1) there was associated urological pathology requiring early surgical intervention, 2) they had other syndromes associated with urinary tract pathology requiring intervention or 3) there was a history of urological fetal intervention.

Clinic Note Templates, Phrase Templates and Research Fields

Existing functionality in the Epic® EHR was used to facilitate data collection by creation of data elements to capture discrete information from notes that would otherwise only be recorded in a free text form. The Epic EHR was implemented at Children's Hospital Colorado in 2004 with semistructured notes and phrase templates. Existing hydronephrosis note templates and phrase templates were modified to include discrete data elements. Lists appearing as a dropdown menu were added to each data element within the study templates. Each list contained an exhaustive selection of choices and allowed for multiple selections when applicable. Each choice was stored as a discrete variable for subsequent data extraction. Four elements were recorded from the face sheet as controls, and 28 elements were recorded from the body of the templated note. Data elements collected through these lists included age at diagnosis, antibiotic prophylaxis, imaging time and results, and followup plan. Five providers used the standardized hydronephrosis note and 1 used phrase templates added to the end of a free text note.

Templates were created by a single author and reviewed by all providers (4 physicians and 2 mid level practitioners) before adoption. An informal educational session regarding the study was conducted within the pediatric urology department, emphasizing the importance of template use. The templates were then implemented by the clinical informatics department in October 2012 with designated research fields "concept enabled" to allow for data extraction. Data were queried on a monthly basis and feedback regarding template use was sent to providers.

The report function in the EHR was used to obtain a list of patients eligible for the study for purposes of manual chart review. To confirm identification of all patients with a diagnosis of hydronephrosis in the problem list or progress note, patients were also required to have a provider assigned encounter diagnosis for ICD9-CM billing codes associated with hydronephrosis (753.20, 753.21, 753.29, 591.0, 593.3, 593.4, 593.70, 593.5, 753.3, 593.89, 753.23, 753.22, 593.7, 753.15 or 753.9). The report criteria included at least 1 visit with a specified provider between January 1, 2013 and June 30, 2013, and age younger than 12 months at the encounter.

Data Collection and Analysis

Data were extracted by the research informatics department. In addition to the study variables embedded at the end of the semistructured note, demographics were extracted from the face sheet. Data elements were extracted from fields embedded at the end of the note. The same elements were manually extracted from the body of the note (Appendixes 1 and 2, http://jurology.com/) and entered by a trained research assistant into REDCap™, a National Institutes of Health sponsored electronic data capture tool.[15] All data were extracted as discrete categorical elements.

Completeness (percent utilization) was calculated by dividing the number of visits identified from template use and ICD9-CM code by the number of encounters identified by ICD9-CM code alone. Accuracy was assessed by comparing data extracted from concept enabled data fields to manual review using the kappa statistic (fig. 1).[16] Evaluation of completeness through time and by template type was conducted using chi-square testing. Missing records were also compared to captured records by chi-square analysis to determine potential reasons for incomplete data capture. Two-tailed p values less than 0.05 were considered statistically significant. All analyses were conducted using SAS®, version 9.3.

RESULTS

A total of 80 patients met study eligibility requirements (fig. 2). Data were recorded through template use for 64 of these patients, for an overall completeness of 80.0%. Agreement between structured data collection and manual sources is shown in table 1. The kappa statistic revealed agreement ranging from 0.45 to 1.00. All variables from the face sheet demonstrated "almost perfect" agreement. Of