

# Data Migration: A Thorny Issue in Electronic Health Record Transitions—Case Studies and Review of the Literature

Richard Schreiber<sup>1</sup> Lawrence Garber<sup>2</sup>

<sup>1</sup>Division of Informatics, Department of Medicine, Geisinger Commonwealth School of Medicine, Camp Hill, Pennsylvania, United States

<sup>2</sup>Division of Informatics, Reliant Medical Group, Worcester, Massachusetts, United States

**Address for correspondence** Richard Schreiber, MD, FACP, FAMIA, Division of Informatics, Department of Medicine, Geisinger Commonwealth School of Medicine, Holy Spirit Campus, 431 North 21 Street, Suite 101, Camp Hill, PA 17011, United States (e-mail: rschreiber@geisinger.edu).

ACI Open 2020;4:e48–e58.

## Abstract

**Objective** To review the existing literature regarding data migration during electronic health record (EHR)-to-EHR transitions and add two case studies on this topic.

**Methods** Very few publications exist that detail the processes and potential pitfalls of data migration during EHR transitions. One of the authors participated in a panel discussion at the American Medical Informatics Association symposium in 2015; at the time, only five empiric or experiential research articles on any aspect of EHR transitions were available. Of those, only two mentioned their experiences with data migration or conversions. A detailed PubMed and CINAHL (Cumulative Index to Nursing and Allied Health Literature) search in March 2019 yielded only one more article giving details about data migration.

**Results** The two new case studies contrast starkly: one relied on manual abstraction and data entry, whereas the other leveraged several electronic tools. The literature reflects this diversity of approach: no two sites have reported the same approaches. The authors identify nine domains of potential consequences of the currently available techniques and offer mitigating strategies.

**Discussion** Very little empiric information exists in the peer-reviewed literature regarding data migrations during EHR-to-EHR transitions; yet the case studies reflect that much remains suitable for a prospective study.

**Conclusion** This report adds two new case studies to the six already reported in the literature. There is a wide disparity in techniques of data migration, each with its own set of pros and cons, which sites must consider during an EHR-to-EHR transition. Such transitions would benefit from prospective research on evaluation and knowledge discovery.

## Keywords

- ▶ data migration
- ▶ data conversion
- ▶ information
- ▶ electronic health records
- ▶ EHR transition

## Background and Significance

More and more practices and organizations are transitioning from one electronic health record (EHR) to another due to consolidation and mergers, the need to maintain certified

EHRs for Meaningful Use/Promoting Interoperability, decommissioning of EHR vendor products, and cost considerations (personal communication: Huang, Koppel, McGreevey, Craven, and Schreiber, 2020). One of the most difficult issues is the thorny one of what to do with legacy EHR data. There are many

received  
December 2, 2019  
accepted after revision  
March 12, 2020

DOI <https://doi.org/10.1055/s-0040-1710007>.  
ISSN 2566-9346.

© 2020 Georg Thieme Verlag KG  
Stuttgart · New York

License terms



anecdotes and opinions but little empiric data. Goals regarding data migration range from no migration at all to complete conversions. Strategies for data migration include manually entering abstracted information, scanning images of legacy documents, and/or electronically transferring data from the legacy system to the new EHR. Institutions report many rationales for their strategies, again with a broad variety of cost, time, and personnel considerations; complexities of the process; risk mitigation; perceived needs; and political exigencies.<sup>1-5</sup> This study reports two new case studies of EHR data migrations as part of EHR transitions, reviews the literature on the topic, details the pros and cons of various strategies, offers mitigation strategies for the issues identified, and makes recommendations to assist institutions that are contemplating an EHR transition and data migration.

The latest data in 2015 reveals that 60% of ambulatory clinicians were still on their first EHR, more than 28%

had already transitioned to a second EHR, and more than 10% were on their third or even more EHR product.<sup>6</sup> Another 18% intended to change their EHR vendor product.<sup>7</sup>

Saleem and Herout<sup>8</sup> reviewed 10 transitions from various EHR vendors, 9 of which were to Epic, including one institution which first migrated to a home-grown system but promptly abandoned it in favor of Epic. They described in some detail the data migration strategies of only four of these transitions. →Table 1 includes these sites, another mentioned briefly in Saleem and Herout's review, and another recent report concerning EHR data migration.

These studies agree that there is limited research on EHR transitions in general and on data migration in particular: "few studies addressing the same topic have been conducted"<sup>15</sup>; "Less is known about the complex challenges of transitioning from one EHR to another."<sup>1</sup> In addition, "38% of

**Table 1** Data migration strategies, methods, challenges, and solutions

Location	Data migration strategy	Methods	Challenges	Solutions	Reference
Geisinger Holy Spirit (ambulatory and inpatient) case 1	Limited Just in time	Manual	Labor and personnel intensive	Hire extra nontechnical staff Training office staff to sustain ongoing ability	This study
Reliant Medical Group (ambulatory only) case 2	Electronic conversion and data import	Building interfaces Electronic imports CDA parsing	Complex Requires technical skill Different display formats Differing architecture between systems	Fastidious mapping Technically savvy staff Attention to differences	This study
University of Chicago	Full data conversion of RIS and PACS	DICOM	Significant time required for PACS Dissimilar RIS models required extensive labor	Transfer from offline backup media faster than from robotic tape library "Some [RIS] data lost in translation"	Behlen et al <sup>9</sup>
Dartmouth-Hitchcock Medical Center	Data conversion between systems assumed to have similar database structures and content	HL7 interface CCOW	Metadata and configuration considerably different Medications: Free text Mapping difficulties Different display formats	Development of CCOW component Creation of special viewer if data existed in legacy systems Multiple programmers Took 5 months	Gettinger and Csatar <sup>1</sup>
Northern California Kaiser Permanente	Full data migration	Back-loading into a new system	Multiple server farms Slow data transfer Risky for medication data Everybody wanted everything Slow synchronization of data	Restricted to 2 years of radiology data Problem list still daunting with "too much of a good thing"	Bornstein <sup>4</sup>
Three Swedish counties	Linkage from new to legacy database	Manual data entry into the new EHR	Differing architecture between systems	Manual process	Makar <sup>10</sup>
Lucile Packard Children's Hospital, Stanford Children's Health	Assessment of migrated data completeness and accuracy	Statistical sampling for data validation	10 major data categories	Potential to save time and money	Pageler et al <sup>11</sup>
Kelowna, British Columbia, Canada	Full data migration from Vistacan/VistA to OpenEMR	Mirth Connect v 3.x <sup>12-14</sup>	Mapping Establishing the source and destination of data Validation of data migration accuracy	Access to source code for both legacy and new EMR OpenEMR database Use of XML, JavaScript, and writing SQL commands as needed	Lin et al <sup>15</sup>

Abbreviations: CDA, Clinical Document Architecture; CCOW, Clinical Context Object Workgroup; DICOM, Digital Imaging and Communications in Medicine; EHR, electronic health record; PACS, Picture Archiving and Communication System; RIS, radiology information system.

large data migration projects run over budget or are not delivered on time.”<sup>16</sup>

Several other articles<sup>17–21</sup> present recommendations regarding specific data migrations such as imaging databases or are general admonitions for successful management of data migrations but do not include sufficient details of an institution’s strategies and experience with migrations of ambulatory or inpatient records. All-purpose comments advising that poor planning, incomplete data in the legacy system, improper mapping, and insufficient intra- and post-migration testing can thwart the best of intentions.<sup>17</sup>

A common request from providers is to transfer all legacy data to the new EHR, as was the situation in case 2’s Epic-to-Epic conversion. At Dartmouth-Hitchcock, the intent was to “convert the maximum amount of legacy clinical information.”<sup>1</sup> In the Swedish survey and interview study, 11% of users asserted this requirement,<sup>10</sup> whereas 30% advocated for selected categories, 30% noted that access to the legacy system through a portal would suffice, and 21% suggested the use of both systems simultaneously. The remaining 8% felt that a manually written summary of the old information should reside in the new system. Conversely, Epstein et al<sup>22</sup> found that “following conversion to an integrated EHR, providing access to historical anesthesia records by maintaining the legacy [system] is not an effective strategy to promote review of such records.”

## Objective

We report two case studies to add to the existing literature on data migration as part of EHR transitions. Given the lack of guidance regarding best practices or the advantages and disadvantages of various approaches, we offer perspectives on the risks of various strategies and suggestions for risk mitigation.

## Methods

Both the authors helped lead the data migration strategies for their respective organizations and have firsthand experience

with the views of their organization’s leadership and EHR end-users, as well as associated data migration issues. One of the authors (R. S.) chaired a panel discussion on EHR transitions at the American Medical Informatics Association Symposium in 2015. At that time, only five research articles were available on any aspect of EHR transitions. Of those, only two mentioned experiences with data migration or conversions.<sup>1,4</sup> A search on January 13, 2020, in PubMed yielded 7,346 results. **Table 2** outlines further refinement of the discovered literature. This study adds two more case studies to the literature on data migrations in EHR-to-EHR transitions.

## Case Studies

### Case 1

Holy Spirit Hospital is an independent not-for-profit community hospital (270 acute inpatient beds) and outpatient provider in Central Pennsylvania. Until 2015, some of the ambulatory clinics were still using paper, while most were using NextGen, eClinicalWorks, or GEMMS. In 2007, the hospital implemented Eclipsys Sunrise Acute Care v4.5 (Eclipsys, later Allscripts v6.5, Chicago, Illinois, United States) and subsequently earned Most Wired and HIMSS stage 6 certifications.<sup>23</sup> The hospital built out Allscripts with all major features, although the use of electronic progress notes was not universal. None of the clinics was using Eclipsys/Allscripts. In 2014, the organization joined the Geisinger Health System and decided to transition all of its sites to the full suite of the Epic EHR (Epic Systems Inc., Verona, Wisconsin, United States). EHR transitions started in the ambulatory areas. Over 5 months during 2015, the team installed Epic ambulatory in all clinics. The next transition was inpatient, from Allscripts to Epic as a big bang go-live in May 2017. The transition team also deployed a new laboratory information system, upgrades to the billing and ADT (admission/discharge/transfer) systems, a new radiology PACS (Picture Archive and Communication System), and many other modules.

**Table 2** Query criteria and results

PubMed query	Total number of citations retrieved	7,346
((transition* OR “data” OR “data migration” OR “conversion” OR “integration” OR “consolidation”) AND (“electronic health record” OR “EHR” OR “EMR”) AND Humans[Mesh] AND English[lang])	Topics of full articles from query reviewed	Number remaining
	Removal of “data” from query	1,143
	<i>Excluded:</i> after review of article titles and, if necessary, review of abstract	28
	<i>Reasons for exclusion:</i> transitions of care, initial implementations, integration of modules within an EHR, transitions from paper records, multiple other irrelevant topics (e.g., chemical transition states)	
	<i>Excluded:</i> issues of transitions but not data migration	9
	<i>Added:</i> articles from references in above and articles known to the authors	15
	“Gray literature” found through internet search	4
Total applicable references for this manuscript	28	

Abbreviation: EHR, electronic health record.

**Note:** The reference section lists all 28 citations.

**Table 3** Topics abstracted into all Epic patient charts in case 1

Ambulatory	Inpatient
Advance directives	Active and pending orders
Allergies	Allergies
Health maintenance	Medications
Immunizations	
Medications	
Past family history	
Medical history	
Past social history	
Past surgical history	
Primary care providers	

### Ambulatory Transition

Geisinger had experienced health information employees and hired consultant abstractors to review the legacy EHR or paper records and input data by hand as discrete data elements. ▶Table 3 lists what abstractors entered into all patient charts. Problem lists intentionally were not included in this abstraction as these had not been reviewed, edited, and consolidated by the providers. Depending on the specialty clinic, the abstraction team also entered other data. For example, for vascular surgery, data entry included a list of prior surgeries, and for primary care practices, the abstraction team scanned the most recent mammogram, colonoscopy, EKG (electrocardiogram), and other reports. For patients on warfarin, data entry included at least three recent international normalized ratio (INR) results and future INR orders. In the absence of empiric research, recommendations, or guidelines, the choice of what and how much to abstract was based on Geisinger's prior experience with such transitions.

Abstraction of charts was in a sequence based on the next scheduled appointment after go-live. Starting at least 3 weeks prior to go-live at each ambulatory clinic, the team succeeded in completing charts for appointments scheduled over the next 8 weeks. After go-live, the task devolved to the office staff. They largely completed chart abstractions according to ▶Table 3 over the next several months, and the task is ongoing for patients seen infrequently. Legacy EHRs were and still are accessible. For 2 years after transition, clinicians could log on through a separate icon on the same computer as the new EHR. Now access is through health information management.

Soon, two groups of patients whose charts had not been abstracted in a timely way became apparent: those due for an INR but not yet seen for an appointment and those who were seen close to the time of go-live but not expected back for several months but required a laboratory test or procedure in the interim. These orders still existed in the legacy system, but occasionally they were not in the new system. For patients on warfarin, this meant entering the INR test in the new system and then extracting and entering at least the last three laboratory values and recent warfarin doses. This

resulted in increased wait times and interruption of laboratory and nursing workflows.

### Hospital Transition

In the hospital, cutover—meaning the timing of when to change interface connections—occurred at midnight during a weekend when inpatient census was approximately 180. Inpatient chart abstraction teams consisting of physicians, nurses, pharmacists, and technical analysts familiar with both EHRs reviewed the content in Allscripts and entered active allergies, orders, and medications by hand into Epic, as shown in ▶Table 3, beginning 2 days prior to go-live in the behavioral health unit and during the day and evening just prior to go-live for all others. Abstractors kept logs of each chart and conveyed this information to the laboratory and pharmacy. When clinicians made medication changes after abstraction, pharmacy made the changes in the new system in addition to the legacy system. Laboratory and other ancillary department orders made sure to add new orders into the new system.

There was no attempt to convert or scan progress notes, operative reports, nursing notes or flowsheets, inactive or completed laboratory or other orders, or other documents. These were available for viewing in the legacy system. Aside from convenience, there was no perceived advantage to scan or interface these records. Many physician progress notes were still on paper; therefore, a financial investment would far outweigh the possible benefits.

The emergency department (ED) continued to use the legacy EHR for those patients likely to go home before midnight but used downtime procedures (paper) for those who arrived late afternoon or evening and might still be in the ED after midnight. For those straddling midnight manual chart abstraction occurred ad hoc close to midnight.

There was no attempt to convert data electronically in either the ambulatory or inpatient transitions despite access to a health information exchange, nor did we attempt to interface legacy systems with the new ones or transmit continuity of care documents from old to new systems. There were no known or reported adverse events<sup>24</sup> in our incident reporting system other than complaints about the wait times.

### Geisinger's Experience with Transitions

In both ambulatory and inpatient settings, the key goal was for a rapid transition to minimize disruption across the entire system and to ensure the use of a certified EHR for meaningful use purposes. Geisinger had performed similar manual transfer in seven prior hospital transitions before using the same process at Holy Spirit and its associated outpatient facilities. The legacy hospital system included only inpatients and those having outpatient procedures performed in the hospital. Of the approximately 750,000 unique medical records in the inpatient EHR at the time of the transition, many contained records that were more than 10 years old; approximately 50,000 represented patients who had died; many were of patients with only outpatient tests—some performed as much as 10 years before—and many that did not contain any inpatient data. At the cutover time, only about one-half of patients had inpatient notes, and

these were only in the year or two prior to transition. All other notes such as history and physical examinations, operative notes, discharge summaries, and consultations were dictated and thus available only as word documents and thus not readily searchable. Even if Geisinger Holy Spirit had contemplated electronic or other methods of data migration from the inpatient legacy system to the new EHR, the finding was that the effort and cost would have been disproportional to the need.

Unlike Gettinger and Csatai's experience<sup>1</sup> in which clinicians still accessed the legacy system 10 years after implementation of a new EHR, Geisinger's overall experience is that access to the legacy systems tail off rapidly after the first year. Initial anecdotal reports of a few dissatisfied clinicians who did not see everything they expected in the new EHR quickly tailed off as they learned to review the records in the legacy system and summarize the information in the new one. Nearly 33 months later, access to Allscripts remains very low: access logs show only one clinician logged in in 6 months. Maintaining universal access is very expensive. Only health information management will have access after this 33-month interval. We do not have comparable data for the ambulatory legacy systems.

## Case 2

Reliant Medical Group, a 500-provider multispecialty group practice in Central and MetroWest Massachusetts, took a different approach to data conversions. During the initial HIMSS Nicholas E. Davies's award-winning implementation<sup>25</sup> of their EHR (EpicCare Ambulatory from Epic Systems Inc.) as well as during acquisition of three new practices, the goal was to transfer electronically as much historical data as possible. In 2007, Reliant migrated from a hybrid paper and homegrown electronic (QuickChart) record to the Epic EHR, as outlined in [Table 4](#).

To transfer these data electronically, Dr. Garber meticulously mapped approximately 100,000 terms (mostly procedure codes, result components, diagnoses, medications, and allergies) and then, using an internally developed interface engine, transformed and loaded more than 100 million clinical data records spanning 15 years into the Epic EHR. Standard Epic inbound interfaces for registration, results, transcriptions, clinical observations, pharmacy, and scheduling supported these data loads.

Medical records staff manually entered the remaining approximately 5% of relevant data from the paper medical record that was not amenable to electronic conversion, including some discrete data, as shown in [Table 4](#). A team of physicians and nurses identified specific criteria for what to scan, including the type, age, and quantity of documents.

As a result of this extensive and methodical conversion and abstraction, physicians no longer needed the paper medical record. QuickChart and the paper medical record continued to be available for a year for use mostly by physicians and staff who learned that they could trust the data in the Epic EHR. QuickChart's audit trails were archived before completely decommissioning in case of future medicolegal needs.

In 2015, Reliant acquired a two-provider practice on the Aprima EHR and a four-provider practice on the athenahealth EHR, as well as merged with a 74-provider practice on another instance of the Epic EHR. As with the initial Epic EHR implementation, the goal was to electronically convert as much of the legacy EHR data as possible. Most of the data were transformed electronically using standard interfaces, except for athenahealth EHR, which required Reliant to use a home-grown Clinical Document Architecture (CDA) parser to extract discrete data ([Table 4](#)).

The general sequence for electronic data conversion was the same for all EHRs, starting with loading patients and their demographic information. This was followed by loading historical encounters for each patient. These encounters become the shell to which notes, orders, and results were attached. Finally, other patient-level data types (e.g., Problem List, Medication List, Allergies, Immunizations, etc.; [Table 4](#)) and recent encounters, notes, orders, and results were loaded so that the new EHR had the latest clinical data.

[Table 5](#) lists the possible complications of large-volume data conversions into a live EHR, and strategies Reliant used to mitigate these potential issues. While some of these were obvious, others did not become apparent until testing the data conversion. This reinforces the importance of preemptive planning as well as having data validated by both clinical and technical staff who each might recognize different types of data conversion issues.

Depending on the legacy EHR, there was huge variability in terms of how many types of data could be converted electronically versus manually versus not at all ([Table 4](#)). For instance, the athenahealth conversion involved 19 different types of data going back 3 years, and of those, all notes and most test results had to be manually scanned or abstracted into the Epic EHR. The Aprima conversion involved 24 data types going back 5 years, and most were converted electronically while notes had to be manually scanned ([Table 4](#)). In contrast, the Epic conversion involved 65 different data types going back more than 30 years, and 100% of them were converted electronically. The Epic-to-Epic conversion moved 42 million data records from 11 million encounters on 163,000 patients. While there was a query-based health information exchange that could have been used to transfer CDA documents for the Epic-to-Epic conversion, it would not have supported almost 50 of the data types needed for conversion and thus was felt to offer little added value.

One added benefit of doing the comprehensive Epic-to-Epic data conversion is that access to the legacy EHR was terminated after 1 year, having served its purpose of giving end-users confidence in the data conversion. This is in contrast to the other conversions that required continued access to the legacy EHR for 2 to 3 years.

Reliant's three EHR-to-EHR data conversions also varied significantly in their total data conversion costs (electronic + scanning + manual abstracting). [Fig. 1A-C](#) displays total costs graphed against each potential predictor of cost: number of providers, years of data, and number of data types (refer to [Table 4](#) for a full list of data types). The relationship between the number of data types and the total cost is the closest to linear

**Table 4** Reliant's data types and electronic strategies for data conversion

Type of data migrated to the new EHR	Strategy per source
Patient-level data types	
Patients/demographics/nickname/aliases	Registration interface (PAEQ)
Emergency contacts	Registration interface (PAEQ)
Insurance information	Registration interface (PAEQ)
PCP/treatment team	Registration interface (PAEQ)
Comments (phone/permanent/family/specialty)	Electronic import (E)
Patient photo	Electronic import (E)
Document list/patient-level scans	Transcription interface (E)
Patient lists	Electronic import (E)
Communication preferences and FYIs	Electronic import (E)
Code status	Electronic import (E)
Problems/care coordination note	Problems interface (PAE) after CDA parsing (A) Manual (Q)
Goals	Electronic import (E)
Allergies	Registration interface (PAE) after CDA parsing (A) Manual (Q)
Preferred pharmacy	Electronic import (E)
Medications	Pharmacy interface (PAEQ) after CDA parsing (A)
Immunizations	Immunization interface (PAEQ) after CDA parsing (A)
Medical and surgical history	Observation interface (PAEQ) after CDA parsing (A)
Family history	Observation interface (PAE) after CDA parsing (A) Manual (Q)
Social history (tobacco, alcohol, sex, drugs, birth, ADLs, socioeconomic, obstetric)	Observation interface (PAE) after CDA parsing (A) (to- bacco-only for P and A)
Health maintenance modifiers	Observation interface (E)
Encounter-level data types	
Encounters (chief complaint, admit/DC dates, LOS, diagnoses, disposition/follow-up)	Registration interface (PAEQ)
Notes (progress, nursing, patient instructions, hospital, telephone, letters, patient portal)	Transcription interface (PEQ) Manual scan (A)
Encounter-level scans	Transcription interface (E) Manual scan (PAQ)
Vitals/growth chart exclusion flag	Observation interface (PAE) after CDA parsing (A) Manual (Q)
Flowsheets	Flowsheet interface (E)
Encounter-level data elements	Observation interface (E)
Future appointments	Scheduling interface (PAE)
Order-level data types	
Pharmacy Fills	Pharmacy interface (EQ)
Labs	Results interface (PAEQ) after CDA parsing (A)
Imaging	Results interface (PAEQ) after CDA parsing (A)
EKGs	Results interface (EQ)
Encounter procedures	Transcription interface (PEQ) Manual scan (A)
Other tests/procedures/order-level scans	Transcription interface (E) Manual scan (PAQ)
Referrals	Electronic import (E)
Future/open orders	Results interface (EQ)

Abbreviations: A, athenahealth; ADL, activity of daily living; CDA, **Clinical Document Architecture**; DC, discharge; EKG, electrocardiogram; E, Epic; EHR, electronic health record; FYI, for your information; LOS, length of stay; P, Aprima; Q, QuickChart; PCP, primary care provider.

**Table 5** Mitigating solutions for possible complications of data conversions

Issue	Mitigating strategy
Quantity of data impacts the duration of conversion	Increase time allowed for conversion
Large data loads may interfere with real-time data and cause a lag between live and backup system	Intentional slowing of loading during daytime hours Limit the total number of records loaded per hour
Data values change between the start and end of data load (e.g., finalized notes, corrected results)	Maintain unique legacy encounter numbers, order numbers, and documents IDs to enable updating Just-in-time incremental load at go-live
Certain symbols in free-text notes interfere with HL7 2.x interface messages	Pipe [   ] replaced with slash [ / ] Backslash [ \ ] replaced with slash [ / ] Caret [ ^ ] replaced with dash [ - ] Tilde [ ~ ] replaced with dash [ - ]
RxNorm codes not specific enough for certain medications (e.g., various methylphenidate formulations)	Careful attention to extended-release medications
Varied use of units of measure (e.g., English vs. metric, pounds vs. ounces)	Careful data quality checking and electronic conversion when needed
Legacy data corruption	Filtering, e.g., skipping data with future dates or prior to a patient's date of birth
Risk of duplicating data for patients in the old system who are already active in the new EHR	Avoid load of similar immunizations or history that differ by 2 d or less in the new EHR Allow Problem list entries with comments to create duplicates to limit the risk of critical data loss
Mapping errors	Load text description into the comment field visible to end-users to aid identification of error Load name of source system with each data element

Abbreviation: EHR, electronic health record.

and thus the best predictor of the total cost. Indeed, with the heavy emphasis on electronic data conversion, it was far less important how many patients' charts were converted or how many years of data, but rather how much effort was involved in setting up the imports or interfaces and doing the mapping. These findings may not be translatable to organizations relying heavily on manual abstraction, as in the first case study.

## Discussion

Our two case studies add significant details to the scanty literature on methodologies and consequences of data migration. Of the six previously published reports, five involve some form of electronic format,<sup>1,4,9,11,15</sup> and we are aware of other conversions that successfully used varying degrees and types of manual and electronic abstraction.

### Methods of Data Conversion

Manual extraction has several advantages: it is relatively inexpensive, requires no sophisticated software or interface development thus there is no associated testing, is generally fast for any given chart, and allows for selectivity of what documents to scan or abstract. There are also clear disadvantages: there is no practical way to abstract a complete chart, much less one with several years of data, does require a second set of eyes at the time of abstraction to ensure accuracy of manually entered data, and is time- and personnel-intensive with a fairly fixed cost per provider, which can become particularly high for very large hospitals and ambulatory clinics, especially if a deeper level of history is desired.

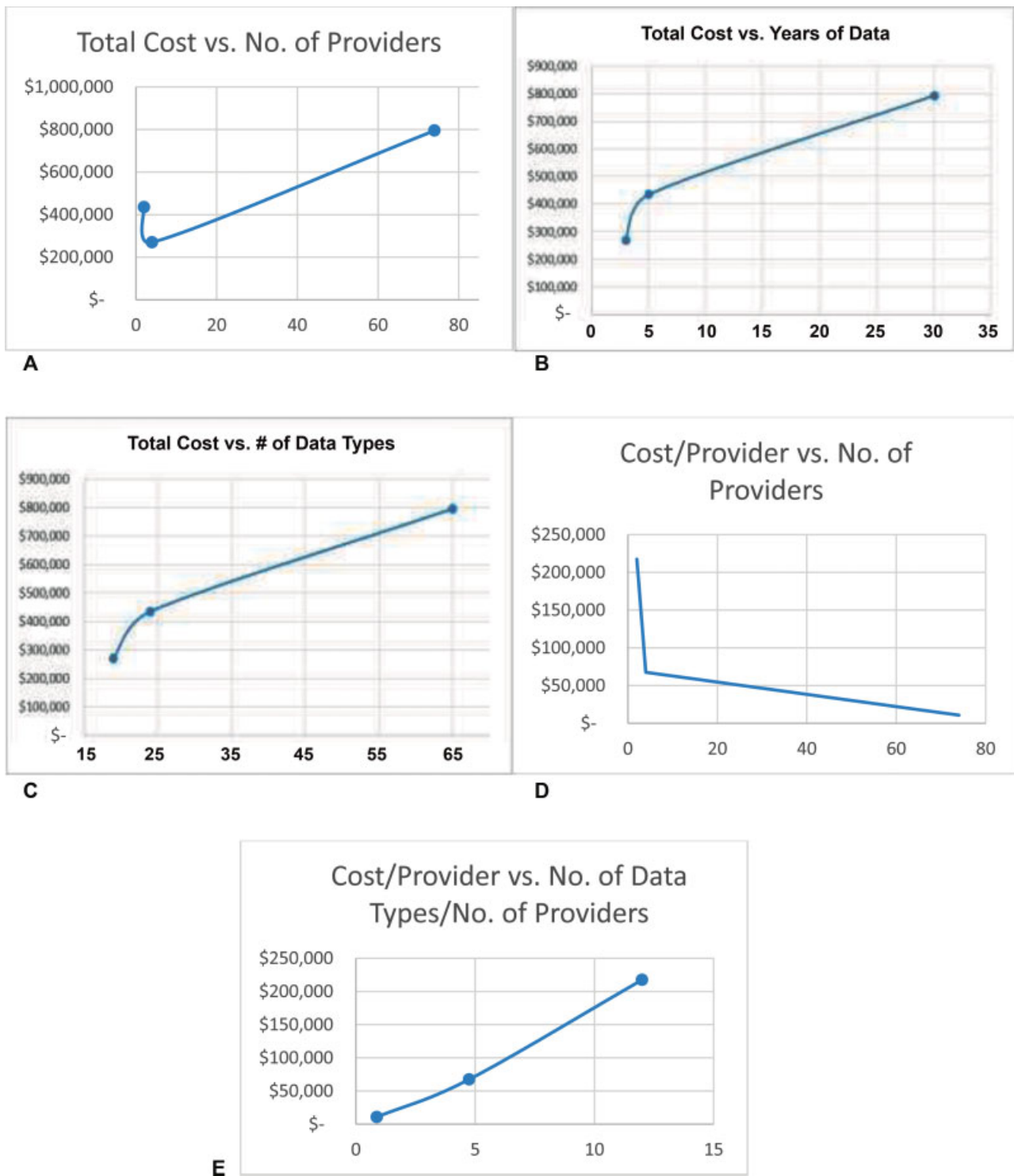
Also, failing to abstract some charts in a timely way is inevitable.

As the authors of the cited papers and case 2 point out, electronic transfer is potentially expensive, requires programming skills, extensive mapping, and thorough testing, and is dependent on the capabilities of the EHRs. Procedures and allergy data<sup>26</sup> are particularly difficult to transfer electronically due to the limited use of standards and mapping incompatibilities. Pageler et al<sup>11</sup> and Lin et al<sup>15</sup> suggest alternative testing strategies such as migration of samples of mapped records and validation of that sample using statistically determined randomized sample size for different data elements or using open source technologies that may save considerable time, money, and personnel commitments, but these studies require replication.

### Cost

Both institutions in the case studies in this report accomplished their data migrations on time and within budget based on their perceived needs<sup>17</sup>; however, there were significant variations in cost. The cost per provider in case 2 had a 20-fold variation between the smallest practice (2 providers: \$217,500 each) and the largest practice (74 providers: \$10,743 each). This suggests that decisions regarding whether to perform electronic data conversion may be determined in part by the number of providers having records converted, although the relationship is not linear, as shown in **Fig. 1D**.

The most accurate predictor of cost per provider for electronic conversion for case 2 comes from comparing it to the number of data types converted, divided by the number of



**Fig. 1** Case 2: total data conversion costs (electronic + scanning + manual abstracting) including salaries (with fringe benefits, both for electronic and manual data conversion), consultants, and electronic health record (EHR) vendor fees (data extraction, interfaces, consultation). (A) Nonlinear relationship between the number of providers having their EHR records converted and the total cost of the data conversion. (B) Nonlinear relationship between the number of years of data converted and the total cost of the data conversion. (C) Almost linear relationship between the number of data types converted and the total cost of the data conversion. (D) Nonlinear relationship between cost per provider's records converted and the total number of providers converted. (E) Linear relationship between cost per provider's records converted and the number of data types converted divided by the total number of providers converted.

providers converted, as shown in **Fig. 1E**. Electronically converting small practices is only affordable if the number of data types being converted is limited, whereas larger practices can afford to electronically convert more types of data.

It appears that using the CDA document may mitigate the cost of electronic data conversion for small practices. Much of the cost in case 2 for electronically converting the four-person athenahealth practice involved developing the CDA



parser, which extracts discrete clinical data out of the CDA document. Future conversions would not incur that cost and thus would lower the per-provider cost closer to that of electronically converting large practices. Indeed, there are several third-party vendors that offer CDA parsing. Emerging Fast Healthcare Interoperability Resources (FHIR) standards will also enable querying and importing of data from one EHR to another at a lower cost in the future.

### How Much Data to Convert

There is a difference in the needs of data migration between ambulatory and inpatient records. Certainly, for longitudinal care, a comprehensive collection or at least a summary of prior data is necessary. Then the question arises as to what constitutes “comprehensive.” Aside from medicolegal and research considerations, it is unlikely that all normal mammograms, chest X-rays, and similar reports are necessary for ongoing care. In contrast, the last result for expensive or invasive tests or an initial report of an abnormality in a study may be significant. On the other hand, it is the experience of one author (L. G.) that being readily able to show patients similar problems, medications, or results from 20 years earlier that they forgot about is reassuring and helps provide confidence in treatment and trust in their physician. But even with a full data conversion, short-term access to the legacy EHR is necessary to allow validation when data quality questions arise and to give clinicians confidence in the conversion.

For inpatients, a comprehensive history and physical examination that summarize prior records, with a complete, accurate, and up-to-date problem list, immunizations,<sup>27</sup> and allergies, may suffice. Leaving aside that conversions for current inpatients differ from all other situations due to active orders including laboratory and rad orders, a minimalist approach to conversion is reasonable, especially since some data change frequently (e.g., active medications) and many patients may never be at the hospital again. Setting realistic expectations for clinicians has a major impact on the success of the data migration in particular and for the EHR transition in general.<sup>2,3</sup> High expectations prior to transition may be unre-

alized and lower satisfaction after the transition. We ascribe in part the success of the EHR transition to setting in advance realistic expectations of what would be in the new EHR and what would not, although we acknowledge a lack of empiric data such as a survey to support this statement.

### Data Conversion Risks, Risk Mitigation Strategies, and Barriers

While the clinical benefits of having data migrated from the legacy EHR to the new EHR are universal regardless of approach, the risks involved in manual abstraction versus electronic conversion differ significantly by approach. **Table 6** summarizes the risks inherent in the data migration strategies.

Given the high volume of data conversion in both manual and electronic data migrations, even with the best planning there is a high likelihood of some conversion errors, whether they are mapping to the wrong data type, wrong date, or wrong patient. It is critical to plan for these inevitable events, have policies on how to identify and correct them, and plan strategies to make sure that correction is always possible regardless of how extensive the problem may be.<sup>11,12,17</sup>

Another consequence of a high-volume electronic data conversion is the prolonged time (weeks to months) to load legacy data into a new EHR. As case 2 shows, when to load the data—either prior to turning on a new EHR versus having to load some later into an already live system—requires different strategies to manage updates to data, redundant data, and system performance. Trial runs to measure data bandwidth and live system performance impact are a critical part of the planning. Similarly, loading more recent data prior to older data increases the likelihood that the most relevant data will be available in the new EHR for patient appointments after transition.

The quality and quantity of personnel can significantly impact the quality and overall conversion time for both approaches to data migration. This impact is greater for manual abstraction of large practices where it may be difficult to hire, train, and have work space for a large team of temporary staff. On the other hand, electronic conversion of many

**Table 6** Risks inherent with various data migration strategies

Risk domain	Possible risks for manual extraction	Possible risks for electronic data conversions
Chart content	May miss clinically relevant content	Mapping may be difficult
Timeliness	May not be available for first patient encounter in the new EHR	May not be available if conversion strategy not properly planned
Accuracy and validation of content	Dependent on human process; therefore, errors are random and hard to detect	Errors tend to be system-wide and therefore easier to detect after go-live Requires extensive testing to detect before go-live
Personnel	Personnel-intensive	Requires highly skilled programmers and content experts
Time	Labor-intensive and dependent on staffing levels	Programming/mapping and data transfer can take a long time
EHR capabilities	Limited risk	Highly dependent on the capabilities of both the legacy and new EHR to export and import data, respectively
Cost	Expensive for large practice conversions and more historical data	Expensive for small practice conversions and more data types

Abbreviation: EHR, electronic health record.

data element types may make it difficult to find enough content experts to work for a limited period of time, or programmers with sufficient skills to migrate the data.

The capabilities of the EHR are another discriminating factor between the two data migration approaches. Most EHRs provide minimal barriers to manual abstraction. However, electronic data conversion is highly dependent on the legacy EHR's ability to export the desired data in a useful format, and the new EHR's ability to import the desired data into the desired fields. The authors' experience with the Epic EHR provided no limitations with either requirement, but some other EHRs were less capable of fully supporting electronic data conversion either due to database constraints, willingness to support data extraction, or the ability to generate encounter-specific CDA documents in bulk.

## Conclusion

This report adds two cases to the body of work on data migration, an area of informatics research where there is only scant literature. Case 1 was a minimalist approach for both ambulatory and inpatient records given that the legacy record was readily available. Case 2 electronically converted massive amounts of data for ambulatory EHRs, suggesting that with robust technical tools at scale it may also be reasonable to convert considerably more information into inpatient charts than case 1 attempted. Review of the available knowledge corpus shows that there are numerous techniques, strategies, and approaches, as summarized in [–Tables 1, 3, and 4](#), but no agreed-upon set of best practices. There appear to be starkly different approaches to conversions of ambulatory versus hospital charts, some of which are reasonable due to the special circumstances of inpatient records. Currently, the cost per ambulatory provider to electronically convert data is proportional to the number of data types converted divided by the number of providers' records converted. The authors believe that newer and more sophisticated technical tools may offer opportunities for more robust conversions at a lower cost, greater efficiency, and increased completeness. Those contemplating such conversions must be aware of the inherent risks and mitigation strategies of each technique, be prepared to test and validate the approach vigorously, and be watchful for omissions, incorrect mapping, and corrupted data.

The two case studies represent highly contrasting strategies. We suggest that those contemplating data migration consider all the methods described in this review and find the best balance between manual abstraction and electronic data conversion, depending on the size of the organization, the number of data types converted, and the capabilities of the EHRs. Conversion by large organizations can justify the cost of electronic data conversion to minimize the use of manual abstraction, bringing the per-provider cost to a reasonable level and allowing for a more comprehensive data conversion if desired. For smaller practice conversions, manual abstraction is currently much more cost-effective, and its cost is controllable by a thoughtful choice of which data to abstract. In the case of multiple small practice

conversions, particularly with CDA documents and CDA parsers becoming more readily available, electronic data conversion will become an increasingly used option.

We offer summary information for those considering data migrations as part of EHR transitions including issues identification, mitigation strategies, and explicit recommendations in [–Tables 4, 5, and 6](#).

### Protection of Human and Animal Subjects

This study does not contain any patient information and was not research, and thus it did not require Institutional Review Board review.

### Funding

None.

### Conflicts of Interest

None declared.

### Acknowledgments

The authors thank Edie A. Asbury for professional librarian assistance with the literature query.

## References

- 1 Gettinger A, Csatori A. Transitioning from a legacy EHR to a commercial, vendor-supplied, EHR: one academic health system's experience. *Appl Clin Inform* 2012;3(04):367–376
- 2 Abramson EL, Patel V, Malhotra S, et al. Physician experiences transitioning between an older versus newer electronic health record for electronic prescribing. *Int J Med Inform* 2012;81(08):539–548
- 3 Zandieh SO, Abramson EL, Pfoh ER, Yoon-Flannery K, Edwards A, Kaushal R. Transitioning between ambulatory EHRs: a study of practitioners' perspectives. *J Am Med Inform Assoc* 2012;19(03):401–406
- 4 Bornstein S. An integrated EHR at Northern California Kaiser Permanente: pitfalls, challenges, and benefits experienced in transitioning. *Appl Clin Inform* 2012;3(03):318–325
- 5 Pfoh ER, Abramson E, Zandieh S, Edwards A, Kaushal R. Satisfaction after the transition between electronic health record systems at six ambulatory practices. *J Eval Clin Pract* 2012;18(06):1133–1139
- 6 MJH Life Sciences. 2015 Technology Survey Results. Available at: <https://www.physicianspractice.com/technology-survey/2015-technology-survey-results>. Accessed January 14, 2020
- 7 Walsh B. Survey finds Allscripts top-ranked EHR, satisfaction increasing. *Black Book Rankings 2015 Survey Results*. 19 August 2015
- 8 Saleem JJ, Herout J. Transitioning from one electronic health record (EHR) to another: a narrative literature review. *Proc HFES* 2018;62(01):489–493
- 9 Behlen FM, Sayre RE, Weldy JB, Michael JS. "Permanent" records: experience with data migration in radiology information system and picture archiving and communication system replacement. *J Digit Imaging* 2000;13(02, Suppl 1):171–174
- 10 Makar M. Dealing with Existing Data in Legacy Systems when Transitioning between Electronic Health Records in Three Swedish Counties [Master's thesis]. Stockholm, Sweden: Karolinska Institutet, Stockholm University; 2014
- 11 Pageler NM, Grazier G'Sell MJ, Chandler W, Mailes E, Yang C, Longhurst CA. A rational approach to legacy data validation when transitioning between electronic health record systems. *J Am Med Inform Assoc* 2016;23(05):991–994

- 12 Camacho Rodriguez JC, Stäubert S, Löbe M. Automated import of clinical data from HL7 messages into OpenClinica and transSMART using Mirth Connect. *Stud Health Technol Inform* 2016; 228:317–321
- 13 Bortis G. Experience with Mirth: an open source health care integration engine. Presented at the 30th International Conference on Software Engineering, Leipzig, Germany, May 10–18, 2008
- 14 Yang YL, Zhang YH. Application of Mirth Connect interface integration engine in medical data transmission. *Industrial Control Computer*. 2016;29:112–113
- 15 Lin J, Ranslam K, Shi F, et al. Data migration from operating EMRs to OpenEMR with Mirth Connect. *Stud Health Tech Inform* 2019; 257:288–292
- 16 Howard P. Data Migration Customer Survey. Available at: <https://www.bloorresearch.com/research/data-migration-customer-survey/>. Accessed January 14, 2020
- 17 Noreen N. Data migration lessons learned: one facility relates hard-learned best practices after migrating to a new EHR. *J AHIMA* 2013;84(09):40–42
- 18 West S. Need versus cost: understanding EHR data migration options. *J Med Pract Manage* 2013;29(03):181–183
- 19 van Ooijen PM, Aryanto KY, Broekema A, Horii S. DICOM data migration for PACS transition: procedure and pitfalls. *Int J CARS* 2015;10(07):1055–1064
- 20 Agarwal TK, Sanjeev. Vendor neutral archive in PACS. *Indian J Radiol Imaging* 2012;22(04):242–245
- 21 Berkowitz SJ, Wei JL, Halabi S. Migrating to the modern PACS: challenges and opportunities. *Radiographics* 2018;38(06): 1761–1772
- 22 Epstein RH, Dexter F, Schwenk ES. Provider access to legacy electronic anesthesia records following implementation of an electronic health record system. *J Med Syst* 2019;43(05):105
- 23 HIMSS Analytics. Electronic Medical Record Adoption Model. Available at: <https://www.himssanalytics.org/emram>. Accessed January 14, 2020
- 24 Barnett ML, Mehrotra A, Jena AB. Adverse inpatient outcomes during the transition to a new electronic health record system: observational study. *BMJ* 2016;354:i3835
- 25 Garber L, Nicholas E. Davies Award of Excellence – Ambulatory Care Application – 2011. Available at: [www.himss.org/reliant-medical-group-formerly-fallon-clinic-davies-ambulatory-award](http://www.himss.org/reliant-medical-group-formerly-fallon-clinic-davies-ambulatory-award). Accessed June 13, 2019
- 26 Graham J, Levick D, Schreiber R. AMDIS case conference: intrusive medication safety alerts. *Appl Clin Inform* 2010;1(01):68–78
- 27 Michel J, Hsiao A, Fenick A. Using a scripted data entry process to transfer legacy immunization data while transitioning between electronic medical record systems. *Appl Clin Inform* 2014;5(01): 284–298