Using a Scripted Data Entry Process to Transfer Legacy Immunization Data While Transitioning Between Electronic Medical Record Systems

J. Michel[1]; A. Hsiao[2]; A. Fenick[2]

[1]Children's Hospital of Philadelphia, Pediatrics, Philadelphia, Pennsylvania, United States; [2]Yale School of Medicine, Pediatrics and Emergency Medicine, New Haven, Connecticut, United States

Keywords

Immunization, electronic health records, organizational efficiency, automatic data processing, electronic data processing

Summary

Background: Transitioning between Electronic Medical Records (EMR) can result in patient data being stranded in legacy systems with subsequent failure to provide appropriate patient care. Manual chart abstraction is labor intensive, error-prone, and difficult to institute for immunizations on a systems level in a timely fashion.

Objectives: We sought to transfer immunization data from two of our health system's soon to be replaced EMRs to the future EMR using a single process instead of separate interfaces for each facility.

Methods: We used scripted data entry, a process where a computer automates manual data entry, to insert data into the future EMR. Using the Center for Disease Control's CVX immunization codes we developed a bridge between immunization identifiers within our system's EMRs. We performed a two-step process evaluation of the data transfer using automated data comparison and manual chart review.

Results: We completed the data migration from two facilities in 16.8 hours with no data loss or corruption. We successfully populated the future EMR with 99.16% of our legacy immunization data – 500,906 records – just prior to our EMR transition date. A subset of immunizations, first recognized during clinical care, had not originally been extracted from the legacy systems. Once identified, this data – 1,695 records – was migrated using the same process with minimal additional effort.

Conclusions: Scripted data entry for immunizations is more accurate than published estimates for manual data entry and we completed our data transfer in 1.2% of the total time we predicted for manual data entry. Performing this process before EMR conversion helped identify obstacles to data migration. Drawing upon this work, we will reuse this process for other healthcare facilities in our health system as they transition to the future EMR.

Correspondence to:
Jeremy J. Michel
Suite 1024
3535 Market Street
Philadelphia, PA 19104
Email: michelj@email.chop.edu

Appl Clin Inform 2014; 5: 284–298
DOI: 10.4338/ACI-2013-11-RA-0096
received: November 12, 2013
accepted: January 27, 2014
published: March 26, 2014
Citation: 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 Inf 2014; 5: 284–298
http://dx.doi.org/10.4338/ACI-2013-11-RA-0096

1. Background

As our institution was transitioning between electronic medical records (EMR), and moving to a single EMR for the health system, we were determined not to lose our accumulated immunization records by abandoning them in inactive clinical systems. Immunization records represent a critical segment of pediatric health data [1] and adult immunizations are becoming increasingly recognized as an effective method for preventing disease [2]. Delivery of appropriate preventative healthcare depends on access to comprehensive immunization records [3]. We sought to institute an efficient, reusable method for preserving legacy immunization data from two different EMRs being retired by our healthcare system, which we planned to reuse for other EMRs being retired by our health system.

We define 'legacy immunization data' as all the immunization records stored in any EMR that we would no longer be using once our health system transitioned to a single EMR, which we refer to as the 'future EMR'. This data included immunizations given within our health system and immunizations given elsewhere then entered as historical data (i.e. abstracted) into a patient's record. We considered the potential effects of not preserving the legacy data and the possible methods for transferring it to the future EMR.

Not carrying forward the legacy immunization data or "starting fresh" could lead to record fragmentation, failure to administer appropriate immunizations (inadequate immunization), and administration of unnecessary immunizations (over-immunization). Record fragmentation could make it difficult to determine the immunization status of our patients from a population health standpoint [4]. The Advisory Committee on Immunization Practices (ACIP) recommends that immunizations be given in most circumstances if records cannot be located [5]. However, in practice, delay of immunization occurs while providers attempt to locate records, potentially resulting in inadequate immunization. Inadequate immunization is a leading cause of preventable infection in infants and children within the United States [6, 7]. The majority of adults in the United States are already inadequately immunized and delaying immunizations would only exacerbate this lack of protection [8]. Increased prevalence of measles [7, 9-11], varicella [12, 13], and pertussis [14, 15], have each been linked with inadequate immunization coverage.

If providers choose to immunize when prior records are unavailable, they run the risk of over-immunizing. Revaccinating already immunized patients would unnecessarily consume medical resources including clinician time and materials, could lead to out-of- pocket costs to patients [16], and could put patients at risk for complications unnecessarily [17, 18]. While immunizations are typically well-tolerated, adverse events, such as anaphylaxis, seizures, and local reactions, can occur [17, 18]. Over-immunization with tetanus, meningitis, and pneumonia vaccines is particularly associated with increased rates of adverse events [19, 20]. For most other immunizations, extra immunizations are as safe as those given on schedule [21]. Sparse evidence correlates the number of immunizations given with increasing risk of adverse events [22].

Based on the quantity of our legacy data and estimating a data entry error rate of 5% [23-25], we determined it would cost between $60,000 to $135,000 in salary for a trained healthcare worker, such as a nurse, to manually abstract and re-enter the data of 45,000 patients. Perhaps more importantly, the time to complete the abstraction was estimated as at least six months given the limited hours available from appropriately skilled staff. Even if six months were available, transferring over this time would have other complications (as six months of additional data would accumulate during data entry).

Transferring data from the legacy system to the future EMR through an interface or through a state registry would be ideal [7]. However, despite evolving EMR standards, such as HL7, electronic transfer of data remains difficult [26-29]. For transferring patient-related data, it is important to confirm that patient identifiers in both systems signify the same patient [30, 31]. Even after the patient is identified, the data from the first system may not be compatible with the second system [32]. Data identifiers are not consistently linked to code standards and need to be mapped between systems prior to data exchange [33]. For our future EMR, this would require interfaces designed for each of our legacy systems, which would be difficult given that our hospital system had six different legacy EMRs.

To decrease the expense and time required to transfer legacy immunization data we planned to complete data transfer via scripted data entry, a process by which a computer has been prepared to continually execute the algorithm for manual data entry (the 'script') upon a large quantity of preformatted records [34]. We sought to compare the error rate of our data migration to published manual data entry error rates. Consultants our organization had hired were using this process to transfer administrative data (e.g. appointment schedules) from our legacy systems but not for clinical data. Other health systems have used this approach for transferring administrative data, medication records, and allergies [35-37]. However, there are no published accounts of scripted data entry for transferring legacy immunization data.

2. Methods

This project was conducted at an approximately 1000-bed tertiary academic medical center with spectrum of ambulatory clinics utilizing legacy EMR-1 for over the past ten years. The inpatient EMR system, legacy EMR-2, has been in use since 2006. Both systems were replaced with the shared future EMR in early 2013.

In anticipation of the transition, we worked with consultants to perform the computer programming necessary to adapt their general scripted data entry process to transfer our legacy immunization data. To accomplish this project, we developed a map between immunizations in the legacy systems and the future system and readying the legacy data (system crosswalk & data preparation), leveraged our consultants experience to perform data insertion into the future system (scripted data entry), and assessed the accuracy and efficiency of the data transfer (process evaluation).

2.1 System Crosswalk & Data Preparation

To facilitate immunization data migration from the legacy systems to the future system, we first determined an appropriate code from the *IIS: HL7 Standard Code Set CVX (CVX)* for the legacy systems' and the future EMR's immunization identifiers [38]. CVX is the standard terminology for immunizations established by the CDC [38, 39]. It includes numeric codes for generic and underspecified historical vaccinations, which enhances its utility for encoding legacy immunization data [39]. By mapping both the legacy systems and the future EMR to a comprehensive standard, we avoided biases from only acknowledging the limited immunizations within each system.

All authors independently chose the CVX code thought to best characterize an immunization. Selection of CVX codes was influenced by our experience with immunizations, chart review, consulting the CDC website, and consulting state registry resources [38-40]. Through discussion, we developed consensus on what CVX code best represented each immunization and what immunization were not well represented by any CVX codes. We developed a linkage map by connecting the legacy immunization identifiers to a CVX code, which was in turn connected to the immunization code in the future EMR. We allowed multiple legacy immunizations to be associated with a single CVX code, but strived for a one-to-one relationship between CVX codes and the future EMR identifiers.

We refined our linkage map using the future system's test environment. We created two sets of approximately 50 patients who had at least five immunizations each and as a group used each legacy immunization identifier at least once. Through chart review of the patients in the test sets, we identified immunization identifiers that would require special handling (exceptions, branch points, and exclusions) and incorporated these special relationships into the linkage map. The exceptions were immunization identifiers that could not be mapped between the EMRs using CVX codes. The branch points handled legacy immunizations which required additional information (such as the date of administration or the age of the patient) to determine the appropriate future EMR immunization identifier. The exclusions were immunization identifiers that were mapped through the CVX codes, but ultimately not included in the data migration due to concerns about data validity.

All immunization records were extracted using structured query language (SQL) from two Oracle databases that contained the legacy data. These reports were merged into a single Excel spreadsheet. Records were omitted from the data pull if they contained an invalid medical record number (MRN), belonged to an identifiable test patient, had an invalid immunization identifier, or were an-

notated in the legacy immunization record as 'refused', 'no', 'not done', or 'declined'. Because the MRNs for the legacy system and the future EMR are discrete and independent unique patient identifiers, it was necessary to cross-reference MRNs between the systems using the *master patient index* for the health system.

Our legacy immunization data was divided into 2,448 individual files with all data for a single patient contained in the same file to facilitate the later stages of this project. The segmentation was necessary to support simultaneous data entry while preventing errors that occur with concurrent access of a single patient's immunization record and to limit the potential effects of file corruption on our data transfer.

2.2 Scripted Data Entry

We documented the workflow for manual immunization data entry using the future EMR with screenshots and developed a flow chart for historical immunization entry (▶Figure 1). From this, we determined four error conditions that could arise, which are described in ▶Table 1. We provided this documentation and the system crosswalk to our consultants to encode the manual historical immunization entry workflow as a scripted data entry process using Visual Basic.

Our consultants, from a single computer, accessed 25 server-based virtual machines using Citrix and launched separate instances of the future EMR on each virtual machine. The scripted data entry process was initiated from this computer and our immunization data was entered into the future EMR following the workflow we documented for immunization data entry. Since 25 virtual machines with separate EMR instances were used in parallel, this greatly decreased the actual time for the data migration process. This process continued with remote monitoring by our consultants until all of the segmented files had been processed.

The first complete run occurred three weeks before our facility switched clinical care between the EMRs. The process was repeated according to a pre-arranged schedule twice more two days before the EMR transition and five days after the transition to capture the new data recorded in the legacy system after the first complete run. A final subsequent data migration run was completed two months after the EMR transition to account for immunizations that had not been extracted from the legacy systems because they were not accounted for within our original query parameters.

We used a migration log to capture the disposition of the legacy immunization (transferred vs. not transferred), the total time to complete the data insertion, the machine that was running the script, and the name of the file which contained the immunization record. For records that were not transferred, the migration log also indicated the reason that the scripted data entry was unsuccessful (i.e. which error condition occurred). Using this migration log, we calculated the percent of extracted legacy data appropriately transferred into the future system and the breakdown by error type of the other records.

2.3 Process evaluation

We evaluated the scripted data entry process by using the migration log to evaluate the process efficiency and by using both the migration log and manual chart review to perform data validation. We had previously estimated (while preparing this project) a manual data entry time for immunization records in the future system using 468 historical immunization records from 21 patients, which we used as our comparison value for the scripted data entry process. We identified published estimates for manual data entry as a comparison for data validation.

Our data validation involved two steps. The first step was an automated comparison of all the extracted legacy data with the migration log, which provided the opportunity to check for skipped records and data corruption. The second step was a manual chart review using both the legacy systems and the future system. We developed a new set of 50 patients using the Microsoft Excel randomization function, stratified to include 25 patients above 18 years old and 25 patients under 18 years old. We obtained the immunization reports from the future EMR for these patients, who together had 457 immunization entries, and visually and verbally compared these to the data in the legacy systems to determine if there was any data loss or corruption.

3. Results

3.1 System Crosswalk

We successfully developed a map between the legacy immunization identifiers and the future EMR immunization identifiers, predominantly through a CVX intermediate. We achieved consensus on the selection of appropriate CVX codes for immunization identifiers in both the legacy systems and the future system. A breakdown of the immunization identifiers included in the linkage map and used during the migration process is included as ▶Table 2.

Of the 127 immunization identifiers in the outpatient legacy systems, we excluded five legacy immunization identifiers from the linkage map. We excluded an additional five immunization identifiers from the future system. The excluded immunization identifiers and the justification for exclusion are shown in ▶Table 3.

One exception and two branch points were incorporated into this map to handle special situations. The exception was handling of rho (D) immunoglobulin (RhoGAM) for which at that time there was no CVX code. This product was mapped directly between the systems, skipping the CVX code intermediate step. Branch points were developed for the handling of pneumococcal vaccinations (our outpatient legacy system used the same identifier for PCV-7 and PCV-13, we used the date of administration to categorize these products) and tetanus vaccinations (our outpatient legacy system used the same immunization identifier for Td and DT; we used patient age at time of administration to categorize these products). An excerpt from this linkage map (which includes both branch points) is included as ▶Figure 2.

3.2 Scripted Data Entry

We extracted 502,095 immunization data points from our legacy systems. This data included records from 59,538 patients from two legacy systems. Using the scripted data entry process, we transferred 500,906 immunization records from the legacy systems into the future EMR. For those records that our process could not transfer, we determined the root cause that the record was not transferred (▶Table 4). Patients included in the data transfer ranged from 1 day old to 110 years old (average age 32.3 years) and there were on average 8.4 vaccines per patient (range 1–237 vaccines). The manual chart review required approximately four hours by two of the authors working together to compare data in the three systems.

There was an extremely low rate of duplication alerts. The workflow we initially used to handle duplicate immunizations was a natural function of the future EMR and successfully prevented duplicates during testing. However, during the actual data transfer the duplication message was displayed inconsistently by the future EMR. By searching for duplicates within a report of all immunizations contained within the future EMR, we identified 10,171 duplicate immunizations after our migration. Of these, 3,015 should have triggered the duplication alert, because the immunization was already present in the future EMR (due to abstraction or receipt of vaccines at sites that had already transitioned). The remaining 7,155 were immunizations within the legacy system that were faithfully 'reduplicated' to the future EMR.

3.3 Process Evaluation

Because we had 25 machines performing data migration simultaneously, scripted data migration was performed in 16.8 real-time hours. From our experience entering 468 historical immunizations, we calculated that manual data entry in the future EMR took 9.5 seconds-per-record. The scripted data entry process performed this task at an overall rate of 2.47 seconds-per-record. However, because we used 25 virtual machines to perform the data migration, what would have taken six months for a skilled abstractor to accomplish (assuming eight hours a day and no breaks) was accomplished in about two workdays, or 1.2% of the total predicted time. Although we lack precise records, we estimate that it took our project team between 20 and 35 hours to design and test this process, and that the total cost of performing this migration to be at $35,000.

Our side-by-side comparison of the legacy data with the migration logs showed that all the extracted legacy data was accounted for during scripted migration process. There was no data loss or data corruption. There were no errors for the 50 randomly selected patients (457 immunizations) we subjected to manual chart review. There is a 1% chance that we found a sample of 457 vaccinations with no error if our data process had a 1% error rate (the bottom reported rate for manual data entry) [41]. We based our conclusion that the data has been appropriately migrated on our chart review, the migration log containing no unaccounted for data, and reports from our colleagues that the records are accurate.

Some data could not be transferred. The vast majority of data that was not transferred resulted from immunizations with invalid dates. A small portion of the extracted legacy data was not transferred because the immunization identifiers had been excluded from the crosswalk. This was included in our analysis because we did not exclude these records prior to starting the data migration. We elected to abandon this data in the legacy system because we had concerns with data validity. As an unexpected benefit of this process, we identified 52 patients through the patient validation alerts where the birthdate stored in a legacy system did not match the birthdate in the future EMR. This list of patients was forwarded to Health Information Management, which could contact families directly and determine each patient's actual birthdate.

From reports of patient interactions throughout the first week after transitioning, we recognized that two subsets of immunizations were not originally extracted from the legacy systems. The first was from the outpatient legacy system and consisted of historical data entered *after* the first data extraction for immunizations administered *before* the first data extraction. Our SQL queries did not account for these immunizations. The second was the inpatient immunization data since the time of the first extraction, which was not extracted due to a miscommunication between the project team and the data extractors. We extracted these sets of data from the respective legacy systems and prepared them for scripted data entry. Using data exported from the future EMR we eliminated records already existing in the future EMR, and transferred 1,695 missed immunization records with no duplication errors.

4. Discussion

Using our system crosswalk and a scripted data entry process we appropriately transferred 99.16% of the immunization records extracted from two legacy systems into the future EMR with no data loss or corruption. We accomplished this data transfer in 16.8 hours and completed the migration of immunization data shortly after the transition between EMRs. Because we were continuing to provide immunizations throughout our transitioning process, it would be complicated to start manual chart abstraction before transitioning. It would then be necessary to abstract each patient's immunization record in preparation for visits (days to weeks before anticipated visits or immediately for unanticipated visits). This is burdensome to the clinical staff, who would be involved in immunization data entry [42]. As most patients see their primary doctor yearly, we could potentially be entering legacy immunization data for a year or more after transitioning. While the actual cost of this project is difficult to separate from the total costs of our EMR installation, the fact that we can reuse this process for the other legacy systems in our health system speaks to its cost effectiveness.

We expected that the data entry would occur significantly faster than the 2.47 seconds-per-record and were initially surprised with this result. Some of this can be explained by a few records with prolonged times but the majority of this time was likely related to using a scripted process, which needed to wait for the EMR to refresh between steps. The data entry script paused while the EMR moved between screens but time measurement continued. We feel the improved time of this process is more representative of the number of virtual machines available than the effect on record entry time. While one could always find more people to perform manual data entry, the expense of finding and training these people may end up being cost prohibitive [42]. Virtual computers are relatively inexpensive and can be repurposed once data migration is complete.

We had anticipated many duplication alerts because the new EMR had been active at other sites in our health system for months. However, there were far fewer alerts than we expected due to the future EMR not consistently displaying the duplicate immunization alert. We could not identify the

reason why the duplication alert was triggered in some instances and not others despite attempts to investigate. We did not go back and remove the duplicate vaccines from the system. After discussions with clinicians and leadership, we ultimately decided that A) the re-duplicated vaccines were an accurate reflection of the legacy record, B) the new duplicates entered during this process would not adversely affect patient care, and C) the effort to remove these duplicates as part of the migration process would not justify the expense. As we had modified the process to prevent this in the future (and for our other legacy systems) we were satisfied with this solution.

We compared our results to manual chart abstraction because manual chart abstraction and data entry has been the traditional method of preserving immunization data. However, manual chart abstraction is time consuming and error prone [4, 23, 35, 36]. Estimates for clinical data entry error rates range from 1.0–26.9% with one article citing an immunization transcription error rate of 10.2% [4, 23-25, 43-47]. Using chart abstraction and manual data entry would require training the person/people performing the abstraction and developing a cost-effective data validation method. Additionally, for this method, we would still need an algorithm to map the historical immunizations to the future EMR because the data entry personnel would not be familiar with the future EMR. Developing this method and training the personnel would have taken at minimum the time and effort as it took for us to develop the map between the legacy systems and the future EMR using the CVX code bridge.

A better comparison point would have been an externally managed immunization repository (e.g. a state immunization registry). While our state has an immunization registry it is not comprehensive for childhood vaccines, did not yet support EMR reporting, and does not include adult vaccines [48, 49]. For states where the immunization registry is a patient's 'official' immunization record, this type of external comparison may be possible. Immunization interfaces could have been developed and would have served as a lasting solution if we were planning to retain the use of our legacy systems. However, these take at least six months to build and were not needed for ongoing patient care.

From the initiation of this project, we were committed to using the CVX to bridge the immunization identifiers between systems. This linkage map may be an unnecessary step in systems that use the CVX standard inherently. However, our future EMR uses immunization identifiers linked to the medication orders and not to the CVX. In the end, using the CVX helped us to identify an immunization identifier that was missing from the future system. Additionally, because we are transitioning as a health system to one record, for legacy immunization data from other sites within our health system we will reuse this portion of the linkage map. To repeat this process on our other legacy systems we will

1. link the legacy identifiers to the CVX,
2. extract immunization data from the legacy system,
3. format for insertion,
4. remove duplicates by comparing against a future system data extraction, and
5. insert the data using the already functional scripted data entry process.

4.1 Limitations

Our process of mapping immunization data and automating transfer to a new EMR from a legacy EMR was developed and evaluated at a single hospital system. This process may work less well in environments where:

1. legacy systems lack specificity about the actual vaccine products that were administered;
2. have more frequent errors in immunization dates; or
3. lack a master patient index.

We selected as a comparison point the time it would take humans to perform this data entry. In our comparison we related the time to perform data entry to the predicted time it would take manual data entry personnel to perform data entry. This does not take into consideration the time it took for the project team to develop and test this process. We would favorably equate the time it took to develop and test this process with the time it would take to locate, retain, and train manual data entry personnel and with the time to develop an immunization interface.

We did not perform a formal cost analysis, which could have helped future organizations interested in reproducing this work. As this project was incorporated into the EMR migration, we did not have an itemized budget. The best we can report is that the approximated cost of this data transfer was significantly less than our prediction for a skilled healthcare worker to perform manual data entry.

We compared our results to published estimates for large-scale data entry, which may not accurately reflect the error rate were data entry to be performed during each patient's visit. Comparing our process to other methods of automated immunization transfer between systems would have been interesting, but we could find no error rates for electronic transfer of immunization records to permit this analysis.

Finally, we calculated the rate of appropriate transfer using the immunizations extracted from the legacy system. Because our legacy systems did not intrinsically differentiate between immunizations and other data, it is impossible to be certain that we accounted for every immunization identifier. Our legacy system also did not differentiate test patients from actual patients and therefore we cannot be certain that none of the patients excluded with invalid MRNs represented real patients. Since the first week after EMR transition, we have not identified any new missing data nor heard of missing immunizations from other physicians within the health system. We will continue to investigate for erroneous entries and missing data with each facility we transition to our future EMR.

5. Conclusion

Transfer of immunization records from two of our legacy systems to our future EMR was completed via scripted data entry. This process was more efficient and more accurate than would be expected of manual chart abstraction and data entry and did not require a separate interface for each legacy system. We were able to adjust our process, account for duplicate vaccines, and enter residual data with minimal additional effort. Additionally, coordinating immunizations between systems helped to identify one immunization not accounted for within the future EMR and to identify patient birthdate discrepancies between our legacy systems and our future EMR.

Clinical Relevance Statement

This work helped us to develop a process for transferring data, which we plan to use with the remaining legacy systems within our healthcare system. Coordinating immunization identifiers using the CVX helped create a potentially reusable and sharable process. This data migration process could be employed by other health systems seeking to preserve legacy immunization records when transitioning between EMRs.

Conflicts of Interest

The authors declare that they have no conflicts of interest in the research. In particular, the authors are not associated with nor do they derive any income from the consulting group who was hired to assist in this project.

Human Subject Research Approval

No humans or animal subjects were included in this study.

Acknowledgements

This work was supported in part by the National Library of Medicine through grant 5T15 LM007056. Special thanks to the data extractors who obtained data from each of our systems EMRs: Grace Vitkauskas, Kevin Early, and Rebecca Tylutki.

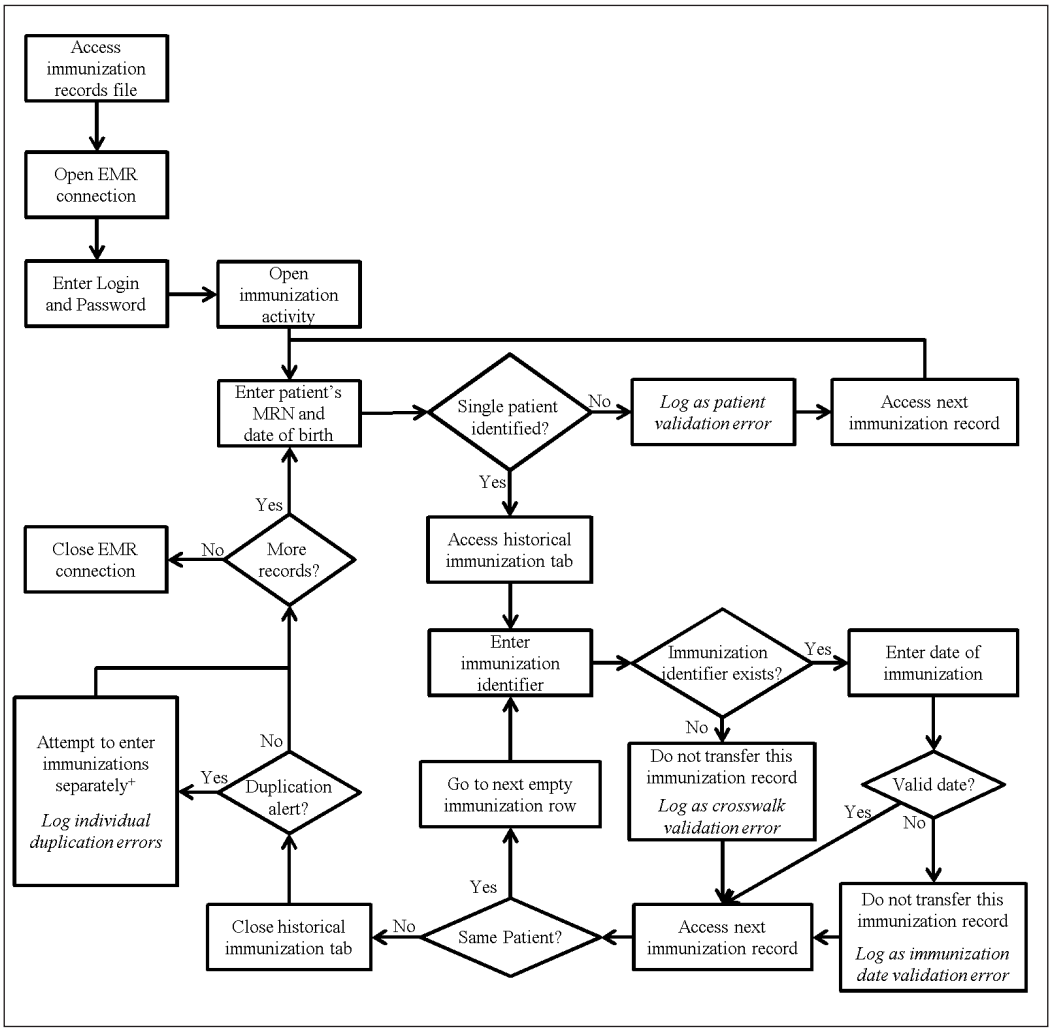

Fig. 1 Flow diagram for historical immunization data entry into the future EMR. Each file contains multiple patients. Data insertion errors are indicated in the flow chart with italics. Immunizations triggering errors are logged but not transferred. + The future EMR's duplication alert indicates one of (potentially many) immunizations entered was already present. However, other immunizations being entered may not be duplicates. To account for this, when a duplication error occurs, the immunization data for that patient are entered one-by-one (closing the record between insertions) to remove actual duplicates and appropriately handle other immunizations.

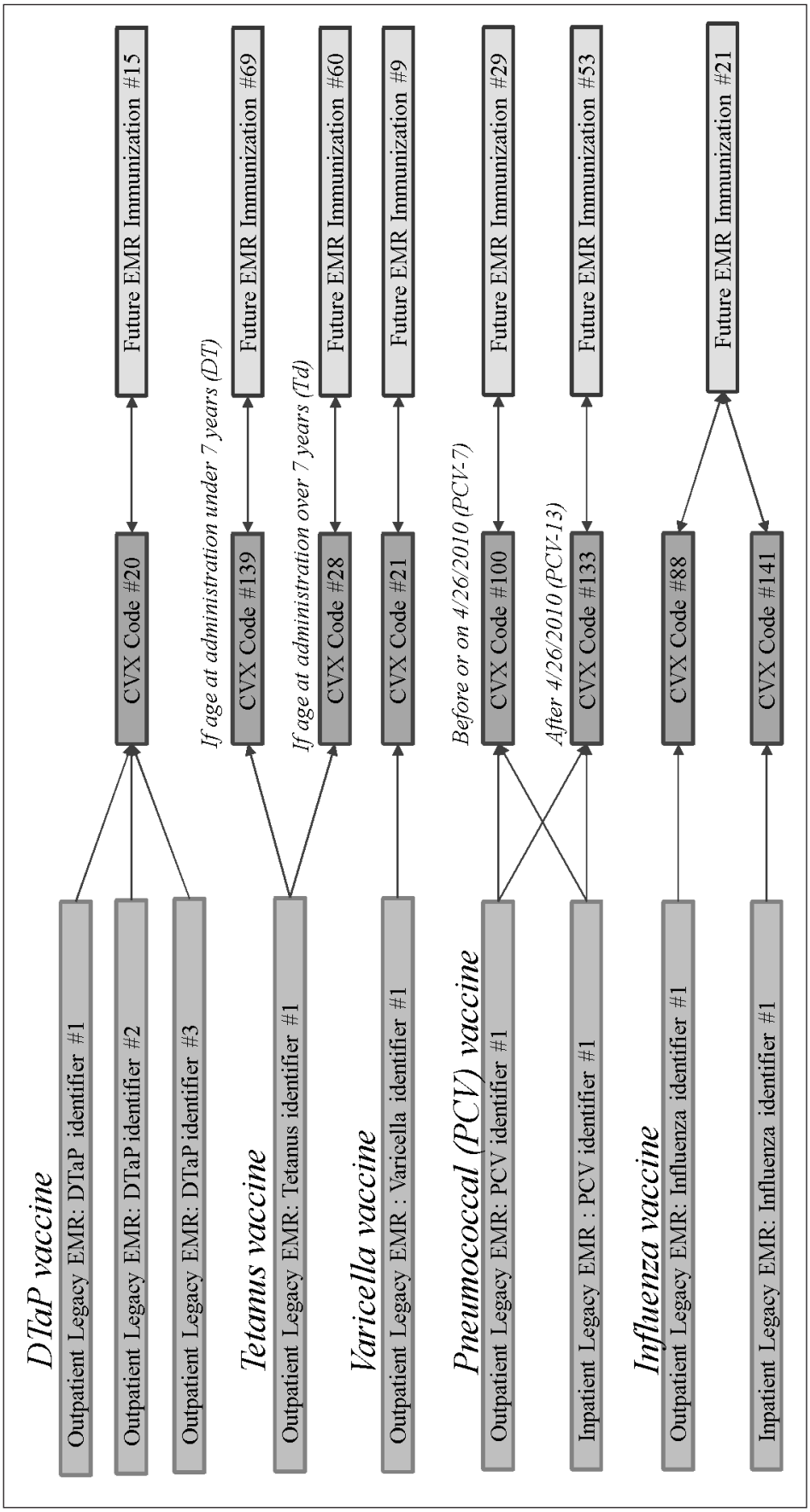

Fig. 2 Excerpt of the linkage map. The excerpt includes DTaP, tetanus (either Td or DT), varicella, pneumococcal, and influenza immunizations. The relationship between legacy system immunizations, the CVX intermediate and the future EMR is denoted by arrows. The branch points for tetanus and pneumococcal are demonstrated at the CVX intermediate level.

Table 1 Description of system alerts for the immunization migration process

Error	Description
Patient validation	No patient in the future EMR with a matching MRN AND matching date of birth
Crosswalk validation	No match for a legacy system immunization identifier in the future EMR
Date validation (2 types)	
• Recorded *before birth*	The immunization administration date in the legacy system is before the patient's birth date in the future system
• Recorded *in future*	The immunization administration date in the legacy system is after the date of the data transfer
Duplicate record	An immunization record already exists within the future EMR for the same administration date **AND** type of immunization as a legacy system immunization.

Table 2 Characteristics of systems involved in the migration

System	Use	Immunization Identifiers	Included in Linkage Map	Used in migration
Legacy System 1	Outpatient	127*	122*	117*
Legacy System 2	Inpatient	3	3	3
Future EMR	Inpatient/outpatient	68+*	63+*	44+*

*Includes RhoGAM, which had no CVX code
+Includes the 'H. influenza type B (3 dose series)', which we added to the future EMR to accurately reflect our legacy data

Table 3 Immunization identifiers excluded from the crosswalk

Immunization	Legacy System	Future EMR	Justification for exclusion
Purified protein derivative (PPD)	2	0	Not a vaccine: The PPD is not an immunization, although it is sometimes treated as one [50]. The legacy system only provided the PPD administration date and PPD evaluation date, not the actual PPD result.*
TdaP, cocoon	2	0	Data validity concern: TdaP cocooning immunizations are given to new mothers (in some research protocols other caregivers), to protect newborns from pertussis [51]. We have 4800 births per year in our hospital system, yet only 19 'TdaP, cocoon' instances, of which eight are recorded for men or women who were not of childbearing age.
Typhoid, generic	1	0	Data validity concern: We could not determine if the immunizations were oral typhoid or injectable typhoid and the follow-up for these immunizations is different [52]. Medical use concern: All patients with 'typhoid' immunizations recorded in our legacy system would need a booster before travelling again [52].
Encephalitis, generic	0	1	Crosswalk validity concern: No CVX code for generic encephalitis.
Influenza, specific formulations	0	4	Legacy system drawback: The legacy systems immunization identifier for influenza did not differentiate preservative vs. preservative-free or whole vs. split product influenza. To decrease the complexity of data entry for seasonal influenza within the future EMR we used: • 1 identifier for injectable influenza • 1 identifier for intranasal influenza • 1 identifier for high-dose influenza

*Because PPD records were excluded before data extraction this data is not included in our analysis of the system crosswalk.

Table 4 Disposition of legacy immunization data

Disposition of legacy immunization data	Immunization Count (%)	Patients affected
Records transferred appropriately	**500,906 (99.16%)**	**59,486**
• Unique immunization records	490,735 (97.7%)	59,486
• Legacy duplicate immunizations	7155 (1.43%)+	3723
Records transferred inappropriately	**3015 (0.6%)**	**797**
• Novel duplicate immunizations	3015 (0.6%)+	797
Records not transferred	**1189 (0.24%)**	**556**
• Patient validation error	334 (0.07%)	52*
• Crosswalk validation error	73 (0.01%)	66
• Date validation error (before birth)	716 (0.14%)	404
• Date validation error (in future)	6 (0.00%)	3
• Duplication error	60 (0.01%)+	31
All records	**502,095 (100%)**	**59,538**

* No immunization records were transferred for these patients. Lack of matching birthdates between the systems was the primary reason.
+ The duplication error alert did not occur during data migration for the vast majority of duplicate immunizations

References

1. Lehmann CU, Kim GR, Johnson KB. Pediatric Informatics: Computer Applications in Child Health: Springer; 2009.
2. Hinman AR, Orenstein WA. Adult Immunization: What Can We Learn from the Childhood Immunization Program? Clinical Infectious Diseases 2007; 44(12): 1532–1535.
3. Stokley S, Rodewald LE, Maes EF. The impact of record scattering on the measurement of immunization coverage. Pediatrics 2001; 107(1): 91–96.
4. Wilton R, Pennisi AJ. Evaluating the Accuracy of Transcribed Computer-Stored Immunization Data. Pediatrics 1994; 94(6): 902–906.
5. Center for Disease Control and Prevention. General recommendations on immunization --- recommendations of the Advisory Committee on Immunization Practices (ACIP). MMWR Recomm Rep 2011; 60(2): 1–64.
6. National Vaccine Advisory Committee. Protecting the Public's Health: Critical Functions of the Section 317 Immunization Program — A Report of the National Vaccine Advisory Committee. In: Services US-DoHH, editor. 2012.
7. Orenstein WA, Hinman AR. The immunization system in the United States — The role of school immunization laws. Vaccine 1999; 17 Supplement 3(0): S19-S24.
8. Center for Disease Control and Prevention. Noninfluenza vaccination coverage among adults – United States, 2011. MMWR Morb Mortal Wkly Rep 2013; 62(4): 66–72.
9. De Serres G, Markowski F, Toth E, Landry M, Auger D, Mercier M, Belanger P, Turmel B, Arruda H, Boulianne N, Ward B, Skowronski D. Largest measles epidemic in North America in a decade - Quebec, Canada, 2011: contribution of susceptibility, serendipity, and superspreading events. J Infect Dis 2013; 207(6): 990–998.
10. Center for Disease Control and Prevention. Outbreak of measles--San Diego, California, January-February 2008. MMWR Morb Mortal Wkly Rep 2008; 57(8): 203–206.
11. Sugerman D, Barskey A, Delea M, Ortega-Sanchez I, Bi D, Ralston K, Rota P, Waters-Montijo K, Lebaron C. Measles outbreak in a highly vaccinated population, San Diego, 2008: role of the intentionally undervaccinated. Pediatrics 2010; 125(4): 747–755.
12. Nguyen MD, Perella D, Watson B, Marin M, Renwick M, Spain CV. Incremental effectiveness of second dose varicella vaccination for outbreak control at an elementary school in Philadelphia, pennsylvania, 2006. Pediatr Infect Dis J 2010; 29(8): 685–689.
13. Lu L, Suo L, Li J, Zhai L, Zheng Q, Pang X, Bialek SR, Wang C. A varicella outbreak in a school with high one-dose vaccination coverage, Beijing, China. Vaccine 2012; 30(34): 5094–5098.
14. Chiappini E, Stival A, Galli L, de Martino M. Pertussis re-emergence in the post-vaccination era. BMC Infect Dis 2013; 13: 151.
15. Glanz JM, McClure DL, Magid DJ, Daley MF, France EK, Salmon DA, Hambidge SJ. Parental Refusal of Pertussis Vaccination Is Associated With an Increased Risk of Pertussis Infection in Children. Pediatrics 2009; 123(6): 1446–1451.
16. Molinari N-AM, Kolasa M, Messonnier ML, Schieber RA. Out-of-Pocket Costs of Childhood Immunizations: A Comparison by Type of Insurance Plan. Pediatrics 2007; 120(5): e1148-e1156.
17. Loughlin AM, Marchant CD, Adams W, Barnett E, Baxter R, Black S, Casey C, Dekker C, Edwards KM, Klein J, Klein NP, LaRussa P, Sparks R, Jakob K. Causality assessment of adverse events reported to the Vaccine Adverse Event Reporting System (VAERS). Vaccine 2012; 30(50): 7253–7259.
18. Center for Disease Control and Prevention. Possible Side-effects from Vaccines. Center for Disease Control and Prevention,2012; Available from: http://www.cdc.gov/vaccines/vac-gen/side-effects.htm.
19. Darden PM, Gustafson KK, Nietert PJ, Jacobson RM. Extra-immunization as a clinical indicator for fragmentation of care. Public Health Rep 2011; 126 (Suppl 2): 48–59.
20. Borrow R, Joseph H, Andrews N, Acuna M, Longworth E, Martin S, Peake N, Rahim R, Richmond P, Kaczmarski E, Miller E. Reduced antibody response to revaccination with meningococcal serogroup A polysaccharide vaccine in adults. Vaccine 2000; 19(9–10): 1129–1132.
21. Abedi GR, Mutuc JD, Lawler J, Leroy ZC, Hudson JM, Blog DS, Schulte CR, Rausch-Phung E, Ogbuanu IU, Gallagher K, Kutty PK. Adverse events following a third dose of measles, mumps, and rubella vaccine in a mumps outbreak. Vaccine 2012; 30(49): 7052–7058.
22. Goldman GS, Miller NZ. Relative trends in hospitalizations and mortality among infants by the number of vaccine doses and age, based on the Vaccine Adverse Event Reporting System (VAERS), 1990–2010. Hum Exp Toxicol 2012; 31(10): 1012–1021.
23. Atkinson I. Accuracy of data transfer: double data entry and estimating levels of error. J Clin Nurs 2012; 21(19–20): 2730–2735.

24. Goldberg SI, Niemierko A, Turchin A. Analysis of data errors in clinical research databases. AMIA Annu Symp Proc 2008: 242–246.
25. Wahi MM, Parks DV, Skeate RC, Goldin SB. Reducing errors from the electronic transcription of data collected on paper forms: a research data case study. J Am Med Inform Assoc 2008; 15(3): 386–389.
26. Mead CN. Data interchange standards in healthcare IT--computable semantic interoperability: now possible but still difficult, do we really need a better mousetrap? J Healthc Inf Manag 2006; 20(1): 71–78.
27. Namli T, Aluc G, Dogac A. An interoperability test framework for HL7-based systems. IEEE Trans Inf Technol Biomed 2009; 13(3): 389–399.
28. Scott P, Worden R. Semantic mapping to simplify deployment of HL7 v3 Clinical Document Architecture. J Biomed Inform 2012; 45(4): 697–702.
29. Dombkowski KJ, Cowan AE, Harrington LB, Allred NJ, Hudson E, Clark SJ. Feasibility of initiating and sustaining registry-based immunization recall in private practices. Acad Pediatr 2012; 12(2): 104–109.
30. McCoy AB, Wright A, Kahn MG, Shapiro JS, Bernstam EV, Sittig DF. Matching identifiers in electronic health records: implications for duplicate records and patient safety. BMJ Qual Saf 2013; 22(3): 219–224.
31. Kijsanayotin B, Speedie SM, Connelly DP. Linking patients' records across organizations while maintaining anonymity. AMIA Annu Symp Proc 2007; 11 (1008).
32. King G, O'Donnell C, Boddy D, Smith F, Heaney D, Mair FS. Boundaries and e-health implementation in health and social care. BMC Med Inform Decis Mak 2012; 12(100): 1472–6947.
33. Giannangelo K, Fenton SH. SNOMED CT Survey: An Assessment of Implementation in EMR/EHR Applications. Perspectives in Health Information Management 2008; 5(7): 1–13.
34. Augustin W. Examples for Open Office Automation with Scripting Languages. 2005.
35. Case Study: Automating Lab Results. 2013.
36. Case Study: CPOE Implementation During System Migration. 2013.
37. Automation Improves the Continuum of Care: Riverside Health System. Newport News, VA2011. p. 1–2.
38. HL7. IIS: HL7 Standard Code Set CVX - Vaccines Administered. HL7 Table 0292: CDC's National Center of Immunization and Respiratory Diseases; 2013.
39. Center for Disease Control and Prevention. Understanding the Rules for Creating CVX and MVX Codes. In: Control. CfD, editor. 2010.
40. Washington State Department of Health. Complete List of Vaccine Names and CPT/CVX Codes. 2013.
41. Lowry R. VassarStats: Website for Statistical Computation. Poughkeepsie, NY2013.
42. Kushinka S. Chart Abstraction: EHR Deployment Techniques. California HealthCare Foundation, 2010.
43. Shelby-James TM, Abernethy AP, McAlindon A, Currow DC. Handheld computers for data entry: high tech has its problems too. Trials 2007; 8: 5.
44. Hills RA, Revere D, Altamore R, Abernethy NF, Lober WB. Timeliness and data element completeness of immunization data in Washington State in 2010: a comparison of data exchange methods. AMIA Annu Symp Proc 2012; 2012: 340–349.
45. Smith LB, Banner L, Lozano D, Olney CM, Friedman B. Connected care: reducing errors through automated vital signs data upload. Comput Inform Nurs 2009; 27(5): 318–323.
46. Heidebrecht CL, Quach S, Pereira JA, Quan SD, Kolbe F, Finkelstein M, Buckeridge DL, Kwong JC. Incorporating scannable forms into immunization data collection processes: a mixed-methods study. PLoS One 2012; 7(12): e49627.
47. Cole D. The Real Cost of Manual Asset Management. No Limits Software 2011.
48. State of CT DoPH. DPH: Connecticut Immunization Registry and Tracking System (CIRTS). State of CT, Department of Public Health; 2013; Available from: http://www.ct.gov/dph/cwp/view.asp?a=3136&q=388268.
49. State of CT DoPH. DPH: CIRTS – FAQs. State of CT, Department of Public Health; 2013; Available from: http://www.ct.gov/dph/cwp/view.asp?a=3136&q=388268.
50. Christ A. Tuberculosis, Tuberculin Skin Test, and BCG Vaccine. Military Vaccine Agency; 2007.
51. Wiley KE, Zuo Y, Macartney KK, McIntyre PB. Sources of pertussis infection in young infants: a review of key evidence informing targeting of the cocoon strategy. Vaccine 2013; 31(4): 618–625.
52. Center for Disease Control and Prevention. CDC Health Information for International Travel 2012 2011. Available from: http://wwwnc.cdc.gov/travel/yellowbook/2012/chapter-3-infectious-diseases-related-to-travel/typhoid-and-paratyphoid-fever.htm.