

# Maximizing Legacy Data Migration to a New EHR

*An Encore Point of View*

*By Joy Ales, MHA, BSN, RN, Sr. Consultant,  
Encore, A Quintiles Company*

*August 2015*

## AN ENCORE POINT OF VIEW

When organizations change EHR systems, clinicians want all the information that was captured within the legacy system(s) to be present on day one of go-live to support continuity of care. Typically, the process to enter that legacy data in the new EHR system is manual and labor/cost intensive. Many EHR vendors discourage electronic migration of data into their new system because of data quality concerns. Data dictionaries, organizational standards, and pick lists for data entry fields may describe the intent of a particular data field but do not guarantee that the data captured in the source system actually reflects that intent. To maximize legacy data migration, organizations need to profile and evaluate that data to fully understand it.

*When organizations change EHR systems, clinicians want all the information that was captured within the legacy system(s) to be present on day one of go-live to support continuity of care.*

## MAXIMIZING LEGACY DATA MIGRATION TO A NEW EHR

Many healthcare organizations have spent millions of dollars to implement new technology and to use the Electronic Health Record (EHR) systems in a meaningful way. Additionally, the demand for healthcare organizations to understand the quality and cost of the care they provide is increasing. Using data captured in the course of providing care has never been more important.

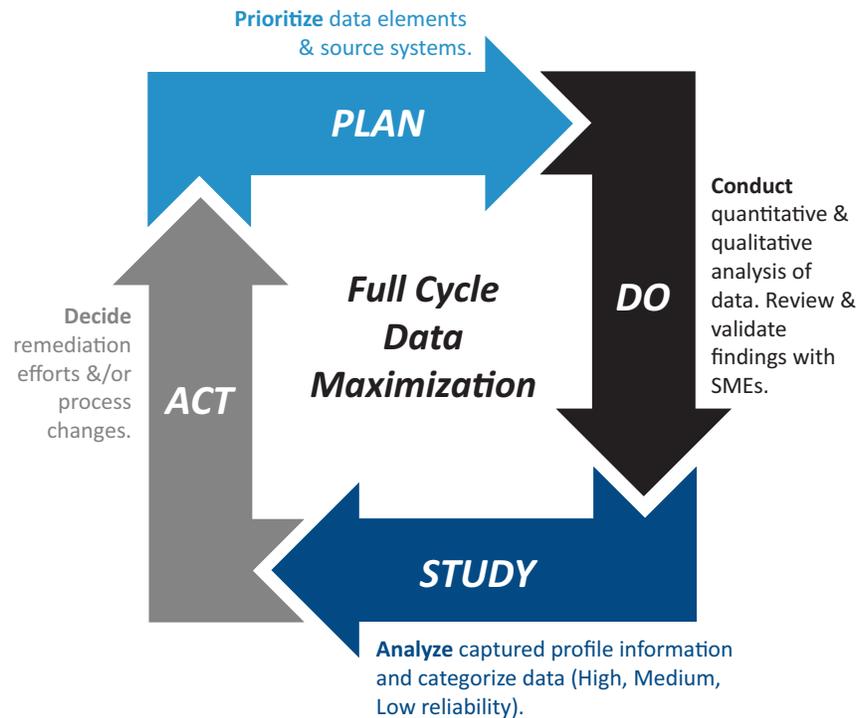
When organizations change EHR systems, clinicians want all the information that was captured within the legacy system(s) to be present on day one of go-live to support continuity of care. Typically, the process to enter that legacy data in the new EHR system is manual and labor/cost intensive. Many EHR vendors discourage electronic migration of data into their new system because of data quality concerns. Data dictionaries, organizational standards, and pick lists for data entry fields may describe the intent of a particular data field but do not guarantee that the data captured in the source system actually reflects that intent. To maximize legacy data migration, organizations need to profile and evaluate that data to fully understand it.

In the past, healthcare systems have focused on adoption of EHRs, paying little attention to consistency or governance of the end-user documentation screens, customizing the screens to fit an individual department and/or office practice workflow. We just want our clinicians to use the system! As regulatory requirements and financial incentives have emerged, many organizations created additional documentation fields to capture information for specific purposes. In many instances, these fields are duplicative of similar content or subject areas. We've even implemented mandatory fields to ensure that the information is captured. Front-end users are incredibly clever and inventive at finding ways to circumvent the data field requirements to expedite completing their task (e.g., registering a patient). Consequently, between the concerns with the quality of legacy system data in the EHR and the lack of consistency and governance around new system data capture, clinicians know that certain data within the EHR system is inaccurate, and therefore, not useful.

## THE ROLE OF DATA PROFILING

Maximizing legacy data migration involves data profiling – the process of assessing and diagnosing the health of the data within the source systems. This process helps an organization understand its data from a technical perspective, and clarifies the data's context and significance.

Maximizing legacy data migration involves data profiling – the process of assessing and diagnosing the health of the data within the source systems.



By knowing what your data looks like and what kind of shape it's in (i.e., a diagnosis), you can decide on a treatment plan to improve its usefulness.

There are two aspects of data profiling – quantitative and qualitative. The first can be done with any number of data profiling tools – or even spreadsheets – but the second requires subject matter expertise specific to the type of data (e.g., a clinician to evaluate clinical data).

## QUANTITATIVE DATA ASSESSMENT

This assessment involves an objective evaluation of the data, comparing the data facts with the standards and rules for that data. We also seek to understand the expected or the data specifications. These are the standards, the business rules, or reference data. These specifications tell us what the data should be.

In medical terms, these specifications would equate to the physical assessment findings and test results compared to what is normal. A quantitative assessment evaluates the integrity of the data as it exists in the source system. This quantitative assessment results in a set of statistics about the data, with examples below.

*A quantitative assessment evaluates the integrity of the data as it exists in the source system while a qualitative assessment looks at data accuracy, and evaluates data consistency and synchronization.*

- The existence or non-existence (NULL) of data and how many NULL values are present as a percentage
- The type of data (e.g., text or numeric)
- Any data value ranges (e.g., 4–10 or 37.6 – 540)
- The minimum and maximum values we find
- Any patterns or format of the data (e.g., nn.n or mm/dd/yyyy)
- Reference table matches (e.g., discharge disposition codes)

## QUALITATIVE DATA ASSESSMENT

Qualitative aspects of data profiling involve a subject matter expert manually examining the data values for rational values, i.e., do the objective findings make sense? A qualitative assessment looks at data accuracy, and evaluates data consistency and synchronization. Accuracy has two parts: is the data valid and is the data right? Validity tells you if the facts of the data meet the data standards (quantitative assessment). For example, a field in an EHR labeled “temperature” may contain data in the correct type and format (nn.n – nnn.n), but the range of values may exceed what is logical from a clinical perspective, such as a patient temperature exceeding 500 degrees Fahrenheit. Another example is a zip code. A data standard for a zip code is either five or nine numeric digits. Facts of the data may show numeric digits, either five or nine character lengths. However, when looking at the data values, if the zip codes are all “9s,” then the data is suspect.

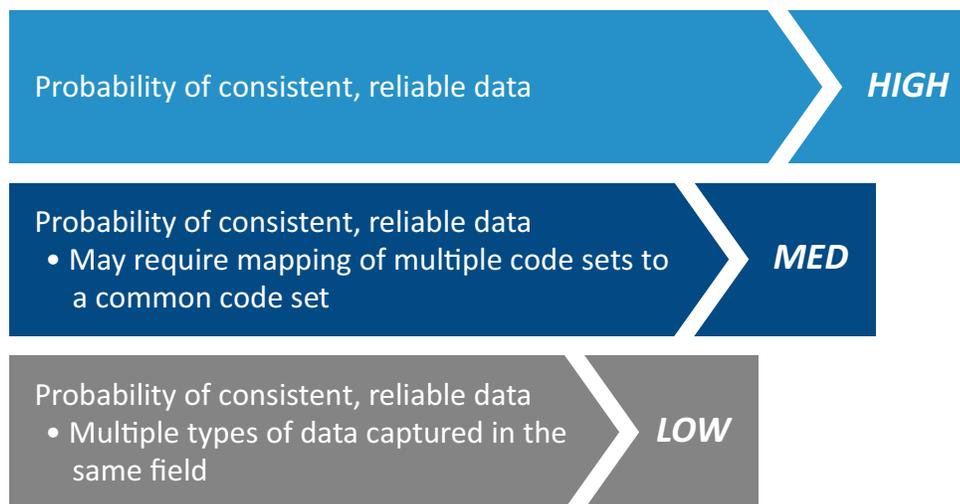
The final part of the qualitative assessment involves evaluating consistency and synchronization, comparing information and determining which data are equivalent and which may need to be made equivalent. For example, when looking at different her systems, a patient’s gender information may be captured differently. One system may use the letters “F” and “M”, while another system may use the words “female” or “male.” The data are equivalent, but the data values differ. As another example, a field labeled “Tobacco Type” may have a pick list for the type of tobacco used (e.g., cigarettes, oral, cigar, pipe). The same or different EHR system may contain individual fields that indicate a yes/no response for field names “cigar use,” “oral tobacco use,” “pipe use.” A subject-matter expert would identify these duplicate or equivalent fields during a qualitative assessment.

Other industries may not have the same need for the qualitative aspects of data profiling as healthcare. The variety of both source systems and the data captured in these systems is quite broad in our industry. Other factors have contributed to the variability of actual data values compared to expected data values, particularly when it comes to clinical systems or any system when time is scarce and the pressures to encourage adoption or complete tasks is great.

Although most clinicians will say they want 'all' the data migrated, care should be taken to prioritize what data is truly needed.

## MAXIMIZING MIGRATION OF LEGACY DATA

In the case of EHR systems, adoption frequently trumps standardization of data capture. Often, as long as the providers can find the data they need to understand prior clinical decisions and interventions, standardizing where and how data is captured is not the priority; at least not to them. When data is migrated to a new system, understanding data in the legacy system(s) becomes paramount. Following completing of data profiling, data quality can be classified in three ways to support decision making for migration:



- High reliability – the data can be migrated “as is.”
- Medium reliability – Some effort will be required to transform or map the data to a standard.
- Low reliability – Data is not a candidate for migration into a discrete field. Consider alternative ways of displaying data within the EHR system (i.e., pdf).

Although most clinicians will say they want 'all' the data migrated, care should be taken to prioritize what data is truly needed. Much of the data captured within an EHR is transactional and only relevant at the time of the recording. For instance, vital signs recorded at the time of providing care may be obsolete the next time the vital signs are recorded. Consider the frequency of recording such data points when choosing which data to migrate. For example:

- Do you really want to migrate 3 years of vital signs data from your inpatient system?
- Do you need all of the medication details about vaccinations (e.g., route, site administered, dose), or do you just need the name of the vaccination and date administered?
- What is the minimally necessary information needed to safely and efficiently care for patients upon go-live?

*Knowledge is power.*

*Knowing the data and the quality of that data provides the understanding needed to formulate a plan to populate the new EHR system.*

An EHR vendor will most likely have a list of recommended subject areas/data that should be present on go-live. This is a minimum recommended list. Data maximization can help with analyzing and confidently migrating additional subject areas that may enhance end-user productivity upon go-live.

To successfully determine which data should be migrated, organizations should engage end-users to understand their current workflows and information that will make their jobs easier. For example, does the legacy system contain embedded patient primary pharmacy information that aids in refilling prescriptions? What about scheduled appointments? Organizations should also consider evaluating the data within specialty modules (i.e., Operating Room, Labor and Delivery, Emergency Department, Supply Chain, etc.). Understanding the quality of data can help with expanding the data availability beyond the recommended minimum to that which is necessary for optimum productivity while avoiding including unnecessary data.

## CONCLUSION

Knowledge is power. Knowing the data and the quality of that data provides the understanding needed to formulate a plan to populate the new EHR system. It supports decision making and provides confidence for migrating the right data the right way by applying any necessary coding rules or data clean up prior to migration. Knowing the data enables organizations to populate the new EHR system with clean and trusted data to support clinicians and patients. Maximizing legacy data migration ensures satisfied patients (they do not have to repeat information already provided) satisfied clinicians (they trust the data and do not have to look in multiple systems) and helps ensure a successful new EHR system implementation.