



9210 Fautleroy WA
Seattle, WA 98136-2618
206-478-8227
www.healthdataconsulting.com

ICD-10

Reconciling Analytics Across the ICD-10 Transition

*Joseph C Nichols MD
Principal*

A Health Data Consulting White Paper

9/22/2012



TABLE OF CONTENTS

THE ICD-10 TRANSITION CHALLENGE	3
MIXED CODING STANDARDS OVER TIME.....	3
MIXED GRANLUARITY	4
CROSSWALKING CHALLENGES.....	5
IMPRECISE DEFINITION OF CURRENT CATEGORIES.....	6
CHANGES IN DEFINITION OF TERMS	7
THE “DATA FOG”	8
UNDERSTANDING THE CURRENT ENVIRONMENT	8
INVENTORY ANALYTIC REQUIREMENTS.....	8
DEFINE CURRENT CATEGORIES OF ANALYSIS.....	8
ESTABLISH A DATA GOVERNANCE STRUCTURE.....	8
IMPLEMENTATION OF A RECONCILED ANALYTIC ENVIRONMENT.....	9
DEFINING THE CATEGORIZATION SCHEME.....	9
DEFINING CATEGORIES.....	11
A CASE EXAMPLE – RECONCILIATION OF THE AHRQ CLINICAL CLASSIFICATION SCHEME.....	11
TESTING THE MODEL.....	12
TEST DATA	12
DEFINING EXPECTED RESULTS	13
KNOWN ISSUES	13
MONITORING POST-IMPLEMENTATION.....	13
SUMMARY	14

The ICD-10 Transition Challenge

The transition to ICD-10 represents one of the largest changes in healthcare information in decades. ICD-10-CM represents the only nationally mandated standard for defining patient’s health conditions. ICD-10-PCS represents the only nationally mandated standard for the definition of inpatient procedures. These codes form a cornerstone of healthcare data and are critical in understanding the nature of the health of the population and burden of illness across healthcare related organizations. These codes are used for a variety of types of analysis including.

- Analysis of disease patterns
- Assessment of risk and severity
- Analysis of patterns of co-morbidities, complications and sequelae
- Evaluation of appropriateness and effectiveness
- Predictive modeling and actuarial analysis
- Disease pattern surveillance
- Quality and Efficiency
- Patient Safety
- And a host of other analytic uses...

As we move from the existing ICD-9 to an ICD-10 code base, there are substantial challenges to assuring the integrity of data and the reliability of information built on that underlying data. This section will explore some of those challenges in preparation for the discussion on how to reconcile data to mitigate the impact to information integrity.

MIXED CODING STANDARDS OVER TIME

Most analytics look at historical data over some period of time to provide stratification, comparisons, trending, prediction and other goals of analysis. As we move from ICD-9 to ICD-10 there are substantial changes in the nature of how these codes are defined. For some time, historical databases will include data defined in ICD-9 and ICD-10.

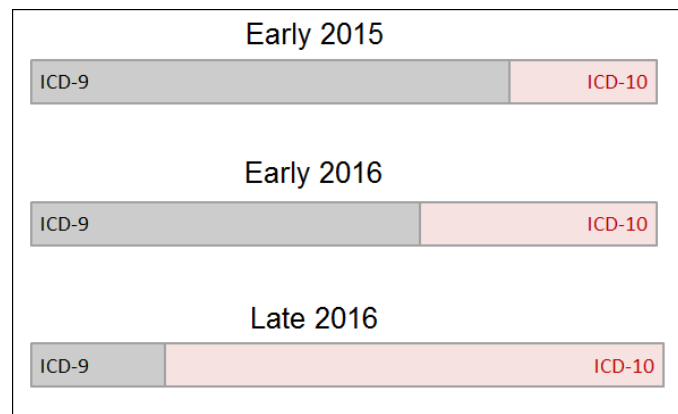


Fig.1

Fig. 1 illustrates the nature of the mixture of ICD-9 and ICD-10 data in a running 3 year historical data set as we move through and past the transition period. As can be seen by this illustration, there will quite a long period of time before there is a clean historical data set composed of ICD-10 codes alone. During this time there will need to be some reconciliation of the categories of data used to support a variety of types of analysis.

MIXED GRANULARITY

Granularity refers to the level of detail or “grain” of codes. The term is a bit confusing since many describe ICD-10-CM as “more granular” but in reality ICD-10-CM is generally made up of more combinations of concepts than ICD-9. For example in ICD-9, there is a single code for “non-union” but in ICD-10 the concept of non-union is included in nearly 3,000 codes as part of a combination of concepts buried in each code.

Example: If we look at an ICD-10-CM code - **S52031M** - (*Displaced fracture of olecranon process with intra-articular extension of right ulna, subsequent encounter for open fracture type I or II with nonunion.*) It can be seen that this one code includes many different concepts related to the patient’s health condition.

- Fracture
- Displaced
- Olecranon process
- Intra-articular
- Right
- Ulna
- Subsequent Encounter
- Open Fracture
- Type I or II
- Nonunion

Each of the concepts is important and may be part of some analysis of similar types of conditions, but the concepts are buried in a single code. Getting at this information requires the ability to “get inside” the code by mapping concepts across a variety of codes that might be otherwise categorized in areas that are not readily apparent.

Because of the use of these combination codes, categorization of codes can be more challenging since there may be a dilemma in where to “bucket” the code in an exclusive hierarchal taxonomy or classification scheme.

Example: If you were to search for all codes related to the concept of “Pneumonia”, you might logically search in the official tabular index for ICD-10-CM codes under the section “Influenza and Pneumonia”. There are 38 codes in this section. However with further searching for the concept of pneumonia you can find 42 other codes that include the concept of Pneumonia in 18 other categories. The reason for this is



that codes for “Staphylococcal Pneumonia” are in the section related to Staphylococcal diseases. Other forms of Pneumonia may be in other section paired with other concepts as part of a combination code.

As a result of this “mixed granularity”, simply cross walking one code to another will result in the loss of key concepts or the assumption of concepts that may not be true in over 95% of the cases based on an analysis of the reported “approximate matches” in the GEM (General Equivalency Mapping) file¹.

CROSSWALKING CHALLENGES

There is a temptation to suggest that reconciliation of mixed ICD-9 and ICD-10 can be accomplished by simply mapping the ICD-10 codes in a dataset back to ICD-9 and use existing categories. It is assumed in this approach that ICD-9 is “less granular”, and therefore any mapping back from ICD-10 will result in comparable categories at the higher level currently mapped for ICD-9 categories. Unfortunately because of the mixed granularity between the two code sets and the change in the combination nature of the codes in both ICD-10-CM as well as in ICD-10-PCS, mapping in this manner can result in a number of impacts to data integrity and reliability². CMS supported by 3M has developed the GEM (General Equivalency Mapping) files to assist in the mapping between ICD-9 and ICD-10 codes. These are bi-directional files, meaning that they are intended to be used in the desired direction of mapping; forward mapping (ICD-9 to ICD-10) or backward mapping (ICD-10 to ICD-9). These files were never created to redefine the intent of an aggregation of codes in ICD-9 to codes in ICD-10. The files can be a valuable research tool to help identify codes in ICD-10 that may belong to some category that has been defined in ICD-9. To do this bi-directional mapping is required. Even using this approach however, there are often codes in that category that may be identified that don’t belong or more frequently a large number of codes left out that should be included.

Example: Consider the case below (fig.2) where the category of injuries to median nerve is defined by a set of ICD-9 codes, then mapped Bi-directionally using GEM to a set of ICD-10 codes vs. redefining the original intent (median nerve injury) in ICD-10 natively. If just the bidirectional mapping from GEMs were used to redefine this category, over 50% of the codes would be left out of the category redefinition that should have been included given the original definition of the intent.

¹ <http://www.cdc.gov/nchs/icd/icd10cm.htm#10update>

² <http://www.cms.gov/Medicare/Coding/ICD10/Downloads/MsdrConversion.zip>

Aggregation of codes that represent Median nerve injuries

- **Native ICD-9 definition = [3] Codes**
 - ✓ 1 code related specifically to Median nerve injury
 - ✓ 2 codes for review related to potential injury
- **GEM Bidirectional = [15]**
- **Native ICD-10 definition = [33] Codes**
 - ✓ 27 codes related specifically to median nerve injury
 - ✓ 6 codes related to potential injury (Carpal Tunnel/Median nerve lesion)

Fig 2.

Example: In some cases, mapping results in the exclusion of a dramatic number of codes that should have been included.

Aggregation of codes that represent “Open Fractures of the Femur”

- **Native ICD-9 definition = [15] Codes**
 - **GEM Bidirectional map of the ICD-9 codes = [270]**
- **Native ICD-10 definition = [1530] Codes**

Fig 3.

There are many examples of code mapping that prove the fact that you can't crosswalk your way to new categories. The process of creating equivalent categories requires the redefinition of the intent of the original category in ICD-10 natively. GEM is a great tool to help, but there is no magical mapping process or tool that redefines these categories directly from current ICD-9 category groups without significant informed human intervention.

IMPRECISE DEFINITION OF CURRENT CATEGORIES

In many cases, current categories of analysis have been in place with little documentation as to the definition of the categories. A lack of precise category definitions can impede the ability to make sure that the intended purpose of the category is met regardless of the code standard.

Example: If a category refers to “Secondary Diabetes” there will need to be a more explicit definition of the concept of secondary diabetes to assure that the proper codes are aggregated to represent the intent of the category. In ICD-9, codes



related to secondary diabetes use the term “secondary diabetes” in the description. In ICD-10 however the term “secondary diabetes” is not used, but rather, an explicit description of a variety of types of secondary diabetes. In determining what to include in this category we need to further define the intent of this category. Does secondary diabetes include or exclude:

- Drug induced diabetes
- Chemical induced diabetes
- Diabetes related to Cushing’s disease
- Neoplasm related diabetes
- Diabetes associated with malnutrition
- Post procedural (i.e. pancreatectomy) Diabetes
- Gestational diabetes
- Diabetes associated with kidney disease
- Diabetes associated with other related conditions...?

Most other categories of health related conditions could be interpreted to include or exclude a number of condition related concepts. It is critically important that categories are clearly define or analysts and other users will not understand the anticipated intent of the category and may make erroneous conclusions related to stratification, trending, comparisons, predictive analytics or any other uses of the data.

CHANGES IN DEFINITION OF TERMS

Changes in definition of key terms and concepts between ICD-9 and ICD-10 particularly in the area of the institutional procedure codes also pose potential issues with comparable categorization.

Example: A bunionectomy is a common procedure today and there are specific ICD-9 procedure codes related to a bunionectomy. In ICD-10-PCS the concept of bunionectomy is not used and terms like “metatarsal resection” are used instead. Unfortunately “metatarsal” resection can be used for a number of procedures that are not bunionectomies. Reconciling categorization of “bunionectomies” becomes somewhat problematic.

Though less common, there are also changes in definitions between ICD-9 diagnosis and ICD-10-CM codes.

Example: Acute Myocardial Infarction (AMI) is defined as occurring within 8 weeks of the initial cardiac event in ICD-9, but within 4 weeks of the initial cardiac event in ICD-10. The concept of initial vs. subsequent encounter is a part of the ICD-9 code, but there is no indication of initial vs. subsequent encounter in ICD-10. In ICD-10 the concept of a “subsequent MI” within 4 weeks of the acute MI is included in the codes, but no such designation of “subsequent MI” exists in ICD-9. Despite the most



accurate coding and mapping, in this example, comparability will be problematic without additional data outside of the diagnosis codes.

THE “DATA FOG”

Even given a full and accurate reconciliation of all code categories and definitions there remains the challenge of the “learning” curve. Clinicians may be less than familiar with the new requirements for documentation and coders may be less than fully competent in proper coding under the ICD-10 model. This fact plus the challenges already noted lead some to claim that we may have a significant period of time when there is a “data fog” or less than reliable cohort of data for analytic purposes. It could take 1-3 years before we can claim that the fog has lifted.

Understanding the Current Environment

It should be apparent that there is a substantial risk to the integrity of information across the transition period from ICD-9 to ICD-10 as well as for some time after the transition. It is critical to develop a strategy that will minimize the risk to information integrity and provide the greatest level of reliability in data analysis. Additionally it is important to understand those known issues that are problematic and cannot be fully mitigated to avoid making incorrect assessments.

INVENTORY ANALYTIC REQUIREMENTS

To understand what will be needed, it is first necessary to inventory current analytic requirements today. For a number of organizations there are hundreds of existing reports, but on further evaluation, only a small number of reports are frequently if every used. An inventory of reports and other forms of analysis and the usage of that information will be important in prioritizing efforts.

DEFINE CURRENT CATEGORIES OF ANALYSIS

There may be a number of categories that have been used for analysis, but the historical knowledge of the intent of those categories may be difficult to identify. It is important to make sure that any categories of analysis have been defined and the definition has been review by stakeholders before moving to reconciliation of standards.

ESTABLISH A DATA GOVERNANCE STRUCTURE

It is important that some data governance structure is established prior to beginning the reconciliation process. This governance structure should have executive support and sponsorship and should include information technology, business operations, quality, compliance, clinical, financial and other key stakeholders to assure that decisions can be effectively made and definitions are commonly shared by all information stakeholders.



This governance structures must be empowered and resourced to do the job to insure success during the transition period and beyond.

Implementation of a Reconciled Analytic Environment

From an analytic perspective, ICD-10 will require a re-definition of all categories of analysis that require ICD-9 codes today. To mitigate the risk to information integrity, organizations will need to prepare their analytic models to support an environment where health conditions and inpatient procedures are coded in both ICD-9 and ICD-10 in the historical data set. This preparation requires a data warehouse; code set repository and data aggregation schema design (usually established in data marts) that will handle both code set standards.

DEFINING THE CATEGORIZATION SCHEME

Based on an analysis of existing analytic needs the organization will need to identify the type of categorization scheme that will support those needs. There are two basic types of categorization schemes that will be needed by most organizations:

EXCLUSIVE HIERARCHAL TAXONOMY

This scheme involves the categorization of all conditions and/or inpatient procedures into unique hierarchal categories that are defined in a parent-child relationship. Each categorization level should be at essentially the same level of granularity. There may be one or many levels of categorization depending on the level of detail needed. In this type of model codes are aggregated to one and only one categorization at the lowest level and all code aggregations are rolled up from the lower level to the higher level. This assures that stratification counts data only once and provides a consistent comparison of associated data across the same level of categorization. This type of categorization is common across a variety of reports where the counts, associated dollars or other values must balance across the entire reporting scheme. For example, AHRQ (the Agency for Healthcare Research and Quality) has designed a four level hierarchal taxonomy (Clinical Classification Software) that has been used for the analysis of public data for many years.

As previously noted ICD-10 brings some additional challenges in this type of taxonomy since the use of combination codes may create a dilemma as to where to categorize the code when only one category applies. There is however a level of categorization that can be reconciled at some level to reasonably assure that “buckets” are comparable across ICD-9 and ICD-10 data.

Example: ICD-10 breaks down the definition of open fractures into the “Gustilo” classifications. This is very valuable since there is a dramatic difference in risk, severity, treatment complexity, cost, sequelae and other important parameters of this type on injury between a Class I and a Class IIIc type injury. For comparison of care, and efficiency of treatment for these types of injuries, the Gustilo classification system provides a refined level of risk segmentation and severity

adjustment. Unfortunately in ICD-9, fractures are reported as either open or closed and the classification of open injuries is not supported at the level of ICD-10. In order to compare data that contains both ICD-9 and ICD-10 codes, categories can only be reconciled at the level of “open fractures” and not the level of “Gustilo classification”. At some point in the future when the entire historical data set being used is now only reported with ICD-10 codes, then analysis will be able to leverage these important differences.

NON-EXCLUSIVE CLASSIFICATION (ONTOLOGY)

Unlike an exclusive hierarchal taxonomy, ontologies link key conceptual attributes to many codes. There is no attempt to confine a code to one concept. It could belong to many concepts based on its defined relationship to these concepts.

Example: Looking at the prior code example; **S52031M** - (*Displaced fracture of olecranon process with intra-articular extension of right ulna, subsequent encounter for open fracture type I or II with nonunion.*) The following concepts are related to this code:

- Fracture
- Displaced
- Olecranon process
- Intra-articular
- Right
- Ulna
- Subsequent Encounter
- Open Fracture
- Type I or II

In an ontology, the concept of ‘fracture’ is related to all codes that are appropriate to this concept. Similarly the concepts of ‘ulna’, ‘right’, ‘intra-articular’, ‘displaced’ and other concepts are also mapped to all relevant codes that are appropriate to these concepts regardless of how those codes might be categorized in a hierarchal taxonomy. The ontology does not attempt to put a code into one category; it simply maps concepts to all relevant codes. While it might be difficult to create a fully stratified balanced report with an ontology, it does provide the freedom to ask the question; “How many cases of fractures of the ulna that were displaced and went into the joint, occurred in a care environment in the past year, and what was the average cost of treatment?” Any set of concepts can be combined to support a variety of analytic questions on the fly based on concept mapping in a very flexible way.

Almost all health care organizations use (sometimes unknowingly) some combination of hierarchical taxonomies or ontologies to support analysis at different levels and for different purposes. In preparing for reconciliation of data across this transition it is



important to understand how these models work in this mixed coding environment in order to preserve information integrity.

DEFINING CATEGORIES

As previously mentioned the most common cause of confusion in data reporting is the lack of well-defined categories where there is little ambiguity as to which codes should be included or excluded. The data governance structure should establish the lead in making sure that all categories have been rigorously defined, documented and maintained and that there is a shared understanding and buy-in by all analytic stakeholders. Each definition should state not only the intent of the category, but what should be included and/or excluded as needed to assure clarity. Validation of code sets should be consistent with these definitions. Occasionally the process of validating these code sets will illustrate the need for a re-look at the clarity of the original definition.

Wherever there are areas of definition that remain open to interpretation, the rationale for choosing one approach or another should also be documented.

A Case Example – Reconciliation of the AHRQ Clinical Classification Scheme

AHRQ maintains a taxonomy (aggregation scheme) presented in their Clinical Classification Software (CCS) that is currently mapped from ICD-9 to four levels of clinical categories³. This standard taxonomy has been used for large national research projects such as HCUP for many years and has been updated on an annual basis. Currently there is an AHRQ categorization for ICD-10, but it is only at one level and has not been reconciled with the ICD-9 classification. Health Data Consulting (HDC) has completed an exercise in reconciliation of ICD-10 codes to a slightly modified ICD-9 version of the AHRQ classification scheme. Several lessons were learned in this process.

- Most of the current ICD-9 AHRQ categories were compatible with ICD-10 mapping with only minor tweaks to the ICD-9 CCS structure
- The first step in the mapping process leveraged bi-directional GEM files to create the first level of categorization of ICD-10 to the ICD-9 categories based on existing mapped ICD-9 codes.
- The second step in the process was to reconcile several thousand ICD-10 codes that mapped to duplicate AHRQ categories when mapped bi-directionally with the GEM files. Most of this reconciliation could be done programmatically based on mapping patterns, but reconciliation of several hundred codes required further manual mapping based on clinical judgment.

³ <http://www.hcup-us.ahrq.gov/toolssoftware/ccs/ccs.jsp>

- The third step in the process required the reconciliation of over 10,000 codes that did not map to any AHRQ category based on GEM mapping to existing ICD-9 codes. Most of these codes could also be mapped based on readily apparent patterns on reviewing unmapped codes. Some codes however required manual review and clinical judgment to assign to the most logical categories.
- The fourth and most difficult step involved QA of the entire completed mapping. Based on this review several programmatic and manual steps were used to resolve the inappropriate assignment of GEM mapped codes that were not addressed in the steps to resolve duplicates and missing mappings.
- When complete it was felt by this author that categorization of both ICD-9 and ICD-10 to common AHRQ based categories could be done down to 3 levels without compromising the integrity of the categorization. Of course this will require further outside validation at some point.
- The entire exercise took approximately 3 weeks of heads down analysis by the author.

This exercise clearly demonstrated the value of GEM files as a tool in assisting the category mapping process, but illustrated that GEM alone could not fully reconcile the two code sets to appropriate category designations. The work needed to reconcile reporting was doable within a reasonable time frame, but incorporation of this process into an overall project at any healthcare enterprise with the appropriate governance structure in place will likely take months of effort.

Testing the Model

Assuming that there has been validation of the defined code sets; reports, data marts and other analytic processes will require testing to assure that analysis provides the intended result to stakeholders. Most organization know how to build test plans, test cases and test data for their analytic environments, but the ICD-10 transition poses some unique challenges in testing that should be considered.

TEST DATA

Since we have no historical data that has been coded in ICD-10, identifying sufficient data to test the analytic environment is a unique challenge. Absent any historical ICD-10 data, the only option is to create ICD-10 data by crosswalking current ICD-9 codes to “equivalent” codes in ICD-10 programmatically across a large enough set of exiting data to support testing of analytic functions. Unfortunately this type of data does not represent reality for several reasons:

- Since crosswalking will lose some key medical concepts and assume other concepts in over 95% of the cases, the data at the outset will not be fully comparable



- Crosswalking is a poor substitute for coded data. There is no way to know what the actual coding pattern will be in ICD-10 as compared to ICD-9.
- There are many instances where there is a choice of codes to select from in the GEM files. Selection of one code as compared to another may be an arbitrary process.
- Depending on the type of analysis, there will need to be a sufficient level of data in the various categories being tested. The identification of these categories in ICD-9 and ICD-10 may be a bit of a challenge since we are testing the process that defines these categories

Ideally, as some organizations are planning now, test data could be derived from parallel coding of the same experience in both ICD-9 and ICD-10 for a year prior to implementation. This provides the ability to train and refine coding prior to the implementation date, and also provides advance data for prediction of impacts and for testing prior to “go live”.

DEFINING EXPECTED RESULTS

As previously noted, we have no experience with ICD-10. While we might expect to see similar results in analysis to what we see today, based on the discussion so far, even with the best design and implementation, results will be somewhat unpredictable. Most organizations are using a broad definition of “revenue neutrality” as a target goal for the initial results of implementation; this is hardly a model for defining expected results in testing. There will be considerable judgment as to whether the results of testing are “in the ballpark”. While some levels of testing can provide fairly definitive expected results, much of the analytic testing will require a healthy dose of fuzzy logic to assess success.

KNOWN ISSUES

All testing will result at some level in a list of “known issues”, where an issue has been identified, but it has been agreed that there is no ability to fully address the issue at the time of release. For ICD-10 this list may be longer than anticipated. Documentation of these known issues will be critical to guide interpretation of analytics going forward. The governance structure will need to provide direction in how to resolve or “live with” these issues as part of the implementation.

MONITORING POST-IMPLEMENTATION

Because many of the challenges mentioned so far, there will be an unfortunately high level of “testing in production” in many cases. Early detection of issues or anomalies in analytics is important to make sure that decisions are not based on flawed information. Understanding known issues and monitoring the impacts of these issues post –



implementation is critical to define the limitations of analysis during a period of “data fog”

There is however light at the end of the tunnel. At some point we will have a homogenous set of data for analysis that hopefully will be more accurate and reliable. That won't happen passively however. At that point in time, the analytic structure will need to rebase its design and retest based on actual historical data in ICD-10

Summary

ICD-10 represents one of the most significant changes to health data in decades. Since these codes represent the cornerstones of healthcare information around the patient's health state and the institutional procedures done to maintain or improve that health state, they are critical factors in analytic activities of all types.

There are substantial challenges to re-defining the categorization structure that will support ICD-10 across the transition period and into a new cohort of data built on ICD-10. Not all challenges will be fully resolved and much of the effort initially is to mitigate the impact of what may be an inevitable data fog.

Testing will be a very different experience under this transition and will “test the testers”. Scenarios and simulations will require creativity. Defining and documenting known issues will be critical to identify limitations of analysis. Monitoring post-implementation will be a necessary extension of the “testing in production” to rapidly identify, quantify and remediate issues.

There is light at the end of the tunnel for those who are aware of the challenges, mitigate the impacts and create the roadmap to leveraging the substantial advantages of analysis offered by ICD-10 once we move through this transition.