Reference Data Management: Overcoming Challenges, Taking Hold of Opportunities

Save to myBoK

By Cheryl Mason, MHSI

The scope of healthcare's Big Data imperative is growing by the day. For instance, one study reveals that the doubling time of medical data in 1950 was 50 years; in 1980, seven years; and in 2010, three and a half years. In 2020 it is projected to be 0.2 years—just 73 days.[1] In truth, the industry has made notable progress in its ability to collect and store vital information that can be used to drive improvements in clinical outcomes, costs, and research. But that's just the first step.

Without harmonization, data aggregation and understanding is difficult at best. Central to overcoming this challenge is the industry's ability to fully leverage an important asset: reference data. Comprised of industry terminology standards, like ICD-10, and other proprietary content, reference data provide not only the building blocks for analytics efforts and better understanding of our patient population, but the foundation of interoperability needed to support the free flow of information between systems.

Reference data management (RDM) is a strategy that employs people, processes, and technology to organize and understand its data. RDM encompasses oversight of important data assets to ensure that all stakeholders across an enterprise are drawing from a single source of truth for analytics and reporting. A component of an organization's master data management framework, RDM is increasingly central to any forward-thinking data governance strategy that supports mission-critical activities.

Consider two proposed rules: the Centers for Medicare and Medicaid Services' Interoperability and Patient Access Proposed Rule and the Office of the National Coordinator for Health IT's Proposed Rule for Improving the Interoperability of Health Information. Together, these rules aim to advance information exchange to allow patients greater access to their health data and call for the implementation of the US Core Data for Interoperability (USCDI). Among other important changes, they require that electronic health information (EHI) be made available via common Fast Healthcare Interoperability Resources (FHIR) application program interfaces (APIs).

If passed in its current form, this new legislation will impact every sector in healthcare. Health IT vendors will have new requirements to certify their software; providers will be required to share additional types of data found in free text notes such as laboratory and imaging narratives, consultation and discharge summaries, and progress notes; and payers will be expected to make claims, clinical, and cost data readily available to their members. Without a complete RDM strategy, healthcare organizations will struggle to comply.

As the industry continues to up the ante on information exchange, the ability to maintain the integrity of the patient story as it moves through electronic systems is increasingly important. HIM professionals have always been responsible for managing that critical part of care delivery. It's important that HIM leaders have a seat at the data governance table and a keen understanding of the challenges and opportunities of managing all an organization's data assets.

Big Data Challenges: A Deeper Look

The industry is sitting on a wealth of information that has potential to transform care delivery. Consider the various forms of data needed to drive performance improvement:

- Structured data such as claims information derived from CPT, ICD-10, MS-DRG, and HCPCS codes
- Semi-structured data such as labs, medications, and social determinants of health found in electronic health record (EHR) drop-down menus
- Unstructured data such as free text found in clinician notes and PDF documents

Healthcare organizations are increasingly trying to draw insights from all this information, but much of it remains locked in silos due to communication barriers between systems. To achieve a foundation of accurate analytics, all this data must be represented by a standard terminology—content that makes up an organization's reference data—to establish a single source of truth that allows information to flow easily.

While this equation may seem simple on the surface, the reality is that many organizations struggle because they lack an effective RDM strategy. This state of affairs leads to poor data quality and notable downstream consequences including cost duplications, negative impacts to reimbursement, and ineffective approaches to care and disease management.

The RDM Opportunity

RDM plays an important role in an organization's data governance strategy. When healthcare organizations centralize management of their data assets through RDM, they can simplify complex data governance processes, establish a single source of terminology truth to optimize analytics, and reduce overhead. The key is having systems in place that ensure reference data is current and accurate and used consistently across an enterprise. Otherwise, inconsistencies can diminish the opportunity.

A comprehensive RDM strategy addresses five components: Governance, Acquisition and Promotion, Content Authoring, List or Value Set Management, and Integration and Distribution.

Governance. Healthcare organizations must ensure alignment of a RDM strategy across people, processes, and technology. Some questions to ask include: How will my team govern terminology definition, intended use, versioning, and implementation across data domains? How do I align my enterprise around a single source of terminology truth?

Acquisition and Promotion. Identifying all code sets used across an enterprise and defining what it takes to maintain them in an optimal way is the next step in designing an RDM strategy. For instance, CPT is a commonly used code set throughout healthcare today. It is found in many disparate systems such as EHRs, admission and registration, billing, and financial systems. Each of these systems is frequently managed on the department level, requiring CPT to be acquired departmentally. By adopting a centrally managed data acquisition process, healthcare organizations can reduce cost duplications related to the acquisition, updating, and maintenance of CPT while also ensuring that all systems are operating on the same version of the code set. Some questions organizations can ask include: Are we sourcing the same thing from different vendors? Are we paying duplicate fees? How do we keep up to date with changes in our reference data as new versions are released?

Content Authoring. Once sourcing is addressed, an RDM strategy should consider an organization's needs around enriching and authoring data. As such, the next step requires defining what that information model looks like and supporting it with toolsets and services. Some questions to consider are: How can custom content be authored in a consistent manner? How do we notify downstream users of any changes to both standards and local data?

List or Value Set Management. Moving beyond sourcing and authoring data, the next step is the management of groups or lists of codes. Code groups or lists are often used as building blocks for business rules that help inform such initiatives as population health and quality measures. For example, how does your organization identify patient cohorts for research, create preference lists for provider workflows, and know which patients need additional intervention for their chronic conditions? It is important to not only curate these lists in a logical and data-driven way, but to ensure that once created, lists are kept up to date as the underlying standards update. A good RDM platform can keep these processes automated, accurate, and available throughout the enterprise.

Integration and Distribution. Finally, RDM must address the distribution of content and how an organization will integrate data into its infrastructure either manually or automatically via APIs. This part of RDM should be supported by processes that handle communication and change management across an organization. Often, this begins with a question: How do I ensure systems are receiving the updates they need?

Implementing the above five steps lays the groundwork for effective RDM. Many resource-strapped healthcare organizations find that the business case for leveraging a framework of automation to do the heavy lifting is an easy one to make. The good news is that technological infrastructures exist that can help healthcare organizations extract the greatest value from reference data.

The best strategies draw on the right data, software, and services to advance RDM. First, healthcare organizations should consider single sourcing reference data to reduce costs and overhead that can result from managing multiple suppliers.

Once content is in place, advanced infrastructures can be deployed to overcome the burden of managing reference data on spreadsheets. Healthcare organizations can consider applications that provide tools for modeling, grouping, and searching data, as well as automating the distribution of updates.

Even with an advanced technological infrastructure in place, many organizations are still challenged to allocate resources to RDM initiatives. In these cases, organizations can lean on third-party informaticists, clinicians, and coders with intimate knowledge of reference data to help augment staff as necessary.

Note

1. Densen, Peter. "Challenges and Opportunities Facing Medical Education." *Transactions of the American Clinical and Climatological Association* 122 (2011): 48-58. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3116346/.

Cheryl Mason (cheryl.mason@wolterskluwer.com) is director of clinical informatics consulting at Wolters Kluwer, Health Language.

Article citation:
Mason, Cheryl. "Reference Data Management: Overcoming Challenges, Taking Hold of Opportunities." *Journal of AHIMA*, no. 7 (Jul-Aug): 38-39.

Driving the Power of Knowledge