The new normal(ize): A primer on patient data standardization

1. What is data normalization?

Normalization. Standardization. Harmonization. It is a process with many names, but for the purposes of this brief can be understood as follows: Data normalization is the practice of taking information in various formats, gathered from numerous systems, and converting it into a consistently labeled, structured, and unified format. In healthcare, data can be considered normalized when information from multiple sources is tied to the appropriate patient, attributed to the appropriate source(s), and connected to terminology that is standardized, mutually understandable, and fully interoperable.

2. Why is data normalization so important to healthcare?

The volume and diversity of healthcare data has grown exponentially in recent decades due to a variety of regulatory and market factors, like the shift from paper to electronic health records (EHRs), the mandates for a growing number of standardized codes, and the proliferation of new sources of patient data. At the same time, the cost of gathering, standardizing, and sharing this information has quickly and drastically increased.

While there is great potential inherent in this torrent of data, its promise cannot be fully realized without effective, accurate, and scalable normalization – a process that reduces redundancy, improves data integrity, and establishes a common understanding of healthcare information.

3. What are some of the challenges to normalizing patient data?

There are a number of obstacles that can impede the effective normalization of patient data. Chief among them are:

Diverse data sources and systems

Patient data is frequently stored in multiple locations like EHRs and practice management systems, as well as lab, radiology, and pharmacy information systems. Over the years, each vendor – seeking to address provider issues with their
own proprietary solutions – has contributed (albeit inadvertently) to systemic issues with data consistency, exchange, and interoperability. In addition, as data is transferred between systems, details are often lost in the process. The inconsistent adoption of interoperability standards has, unfortunately, only compounded the problem.

Multiple standardized code systems

The codification of essential data elements has taken numerous parallel paths. For example, diagnoses, procedures, and lab results are each represented by several different code systems – none of which connect or “speak” to one another, even within a domain. This leads to siloed information instead of a seamless patient record. In addition, each code system has its own schedule for updates, which necessitates vigilant and meticulous maintenance of the many maps connecting patient data to the codes required for billing and reimbursement, along with other use cases.

Documentation variability

Point of care documentation is subjective. One physician’s heart attack is another’s myocardial infarction and yet another’s MI. As a result, how a patient’s condition is described, the specificity used, and how standard codes are applied can be highly variable.

How does patient data degrade or lose detail?

As patient data is documented, aggregated, stored, and transferred, critical information can easily be lost. Improper use of terminology at the point of care can prevent clinicians from capturing all the pertinent details of an encounter. And, using dictation can also lead to data loss if details converted into free text are not effectively mapped to the appropriate standardized codes.

Even when data is properly stored within the EHR, information is frequently omitted when the data is sent to an aggregator or data warehouse. Often, only partial information transfer occurs when the minimum requirement for a single code – such as a primary ICD-10-CM code – is sent. In such cases, the specificity initially captured in secondary codes or other code systems, such as SNOMED-CT®, is lost. Additionally, in the case of laboratory results, a lab system may send local codes that provide little meaning when aggregated with other non-local data.

Even when the data is properly stored within the EHR, information is frequently omitted when the data is sent to an aggregator or data warehouse.

What is the downstream impact of data loss for things like patient care, analytics, and reporting?

At the individual level, the quality and safety of patient care can be jeopardized by inaccurate or incomplete data. Not having a full picture of a patient’s history – including conditions, procedures, labs, and medications – impedes the ability to make truly informed decisions during care delivery. Pulling back to the population level, the importance of accurate information cannot be overstated. If individual patient data is incomplete or inconsistent, patient cohorts will likely not include the right people, or may miss relevant clinical details. Any subsequent analysis of this incomplete data, perhaps aimed at identifying patterns or trends, may then lead to misleading
in order to understand the value of a terminology partner, one must literally start from the start – the point of care. terminology solutions are frequently integrated into the ehr so they are present during the patient encounter. at their best, these solutions empower and enable clinicians to document with natural, intuitive language. they can also prompt clinicians to document with an appropriate level of specificity, and as a result, help to fully capture clinical intent.

knowing this intent – and having it both documented with granularity and mapped to standardized codes behind the scenes – is the very core of a terminology partner’s inherent value. terminology solutions also help to maintain layers of detail that may otherwise get lost in the multiple steps of data transfer and manual harmonization. they do so by providing the connective tissue that links terms, codes, and other metadata. without a terminology layer, critical information will inevitably be omitted, painting an incomplete picture of a given patient or population. it is the difference between starting a book on page one or beginning five chapters in.

what is the unique value that a terminology partner can bring to those trying to aggregate and normalize vast amounts of data?

in order to understand the value of a terminology partner, one must literally start from the start – the point of care. terminology solutions are frequently integrated into the ehr so they are present during the patient encounter. at their best, these solutions empower and enable clinicians to document with natural, intuitive language. they can also prompt clinicians to document with an appropriate level of specificity, and as a result, help to fully capture clinical intent.

without a terminology layer, critical information will inevitably be omitted, painting an incomplete picture of a given patient or population. it is the difference between starting a book on page one or beginning five chapters in.

what initiatives or processes can be accelerated or improved with normalized data?

a complete and accurate data set is the foundation for a variety of use cases, such as a hospital’s quality improvement initiatives; alerts facilitated by a health information exchange; or the identification of ideal patients for clinical trials. the process of getting to a complete and accurate data set, however, can be a challenge. this crucial step – which is essential to enabling any meaningful data analysis – often takes longer than anticipated due to the extensive manual mapping of standardized codes and terms and the need for human review. some initiatives feel this pain more acutely than others.

for example, the covid-19 crisis has highlighted the importance of normalization related to lab data aggregation and surveillance. at the start of the pandemic, there was (and continues to be) an urgent need to quickly and accurately report...
on test results and identify new cases. However, lab data is often represented in multiple ways; frequently includes only local codes; and can be compromised by other data entry issues. This has led to an increased manual burden at a time when there is already a shortage of helping hands, as well as an analytic burden that has delayed our understanding of the pandemic. Putting technologies in place to quickly normalize and standardize this data would alleviate this type of bottleneck and prepare health systems for similar crises in the future.

To learn about IMO’s unique approach to patient data normalization, visit imohealth.com/imo-precision-normalize.

About Intelligent Medical Objects

At IMO, we are dedicated to powering care as you intended, through a platform that is intelligent, intuitive, and intentional. We offer a portfolio of products that includes terminologies and value sets that are clinically vetted, always current, and maintenance-free. This aligns to provider organizations’ missions, EHR platforms’ inherent power, and the evolving vision of the healthcare industry while ensuring accurate care documentation and administrative codes. So clinicians can get back to being clinicians, health systems can get reimbursed, and patients can more easily engage in their own care. As intended.

© 2021 Intelligent Medical Objects, Inc. All rights reserved.

SNOMED and SNOMED CT® are registered trademarks of SNOMED International.