

Coding

A Transformative Innovation to Clinical Coding for Better Data and Less Burden

Michael Pine, M.D., M.B.A.¹, Kathy Giannangelo, M.A., R.H.I.A., C.C.S., F.A.H.I.M.A.², Charles Hobson, M.D., M.H.A., Ph.D.³, Christopher Tompkins, Ph.D.⁴

¹ Managing Director, MJP Healthcare Innovations, LLC, ² President, Kathy Giannangelo Consulting, LLC, ³ Chief Medical Officer, HCA Florida Lake City Hospital, ⁴ Research Associate Professor, Brandeis University

Keywords: International Classification of Diseases, Knowledge Bases, Semantics, Morbidity, Mortality, Clinical Coding, Clinical Decision-Making
<https://doi.org/10.63116/FQRG7761>

Advances in Health Information Science and Practice

Vol. 1, Issue 1, 2025

Background

The importance of codes for documenting healthcare concepts has expanded from a historical need to collect statistics to today's wide array of clinical and other secondary uses. The most deeply embedded system for coding diagnoses is the World Health Organization's (WHO) International Classification of Diseases (ICD). WHO's eleventh revision, ICD-11, was evaluated to determine how best to meet the needs of stakeholders concerned about clinical quality, healthcare decision-making, and other secondary use cases.

Methods

This study explored the capability of ICD-10-CM, the standard ICD-11 for Mortality and Morbidity Statistics (ICD-11-MMS), and an enhanced ICD-11-MMS prototype (ICD-11-CCL) to represent clinical patient profiles for two natural clinical language scenarios. Systems were evaluated for granularity, comprehensiveness, accuracy, precision, relevancy, and their capability to provide user-friendly coding support.

Results

Both the standard and enhanced ICD-11-MMS produced code clusters that captured more clinical details and insight than ICD-10-CM's codes. However, ICD-11-MMS was less comprehensive than ICD-11-CCL and introduced occasional new inaccuracies. Unlike ICD-10-CM and ICD-11-MMS, ICD-11-CCL enabled clinicians to designate their clinical perspectives directly. While ICD-11-MMS produced numerous synonymous code clusters, the enhanced syntax of ICD-11-CCL resulted in a unique code cluster for each scenario.

Conclusions

Adoption of an enhanced ICD-11-MMS such as ICD-11-CCL could provide clinically nuanced data to support important secondary data applications more effectively than standard ICD-11-MMS, while retaining ICD-11-MMS's capability to support international comparative analyses such as mortality and morbidity statistics. With automated coding tools, enhanced ICD-11-MMS could also reduce the burden on humans to select, review, memorize or decipher codes. Enhanced ICD-11-MMS represents a substantial improvement over customized modifications of ICD-10 and, with further development and testing, could be a viable option for future adoption.

INTRODUCTION

The importance of clinical codes for documenting healthcare concepts is greatly increasing in the era of electronic health records, integrated data repositories, natural language processing, and artificial and augmented intelligence. Current code sets and coding conventions are becoming limiting factors in realizing the potential of a truly digital healthcare information system in which all relevant

information is captured, organized, and made available securely and appropriately for clinical, administrative, research, and policy-making purposes. New innovations designed to significantly improve the granularity and interoperability of coded patient data, and to vastly reduce the burden on humans to select, review, memorize or decipher codes, are examined in this article.

While clinicians communicate in their natural, native spoken language (eg, English), computers communicate

about healthcare in codes—unique representations consisting of alphanumeric and special symbol combinations. Codes are most useful when they provide concise, standardized, replicable, and unique representations of clinical and social concepts. For example, a longstanding practice has been for clinicians and facilities to summarize patient care on electronic claim forms for reimbursement using codes for diagnoses, medical treatment, procedures, and psychosocial and environmental issues.

Numerous distinct code sets are available for use in national healthcare databases for symptoms and physical findings, for diagnoses, for mental health and substance use disorders, for diagnostic test results, for professional procedures and services, and for other descriptors and determinants of health. The most pervasive and deeply embedded system for coding diagnoses in the US and other countries is the World Health Organization's (WHO) International Classification of Diseases (ICD), 10th Revision (ICD-10). Several countries have created national modifications of ICD-10; one example is the US-specific clinical modification, ICD-10-CM, which extends ICD-10 for tracking of public health conditions, capture of more diagnoses and of disease severity, and capture of inpatient treatment and procedures. ICD-10-CM permeates the US healthcare industry for a host of purposes, including clinical quality measurement, billing and payment, strategic planning, public policy, and research.¹

In 2022, the 11th revision of the ICD (ICD-11) came into effect, resulting in the WHO ending general support for ICD-10. ICD-11 is designed to be continually updated, comprehensive, with a code structure that captures very granular clinical details and uses a knowledge base that optimizes automated computer-based storage and analytics.^{1,2} This continuous update process could negate the need for a successor (ie, an ICD-12).²

The US and other countries now must decide whether to retain unmodified or clinically modified ICD-10 coding systems or to find an optimal approach to implementing ICD-11. With ICD-11 at their doorsteps, how the US and other countries using clinical modifications of ICD-10 will proceed remains unclear. In the US, some stakeholders resisted the transition from ICD-9-CM to ICD-10-CM due to perceived limitations in ICD-10-CM, preferring instead to wait for ICD-11.³ Historically, the transition from ICD-9-CM to ICD-10-CM (which began in 1998 and was not completed until 2015) was both difficult and expensive.^{4,5} Given this past experience transitioning from ICD-9-CM to ICD-10-CM, what implementation approach would be most useful to stakeholders concerned about clinical care or rigorous secondary use cases?

Making smart choices will require careful evaluation of the options and trade-offs involving relative costs and benefits. Although WHO has emphasized and prepared tools supporting the use case of mortality and morbidity statistics, many countries now rely on their clinically modified ICD-10 systems for use cases requiring more granular information than population statistics. Hence, this paper compares and contrasts ICD-10-CM and two ICD-11 options for their capabilities to capture, represent, and articulate

clinically nuanced information illustrated in a couple of hypothetical patient profiles. We have chosen a demanding use case (ie, communicating comprehensive, nuanced clinical information) to show the capabilities and limitations associated with each option. Less demanding use cases, such as population or cohort statistics for common conditions, are more easily supported by any of the options.

THREE MAJOR OPTIONS

One option is to continue indefinitely with a clinical modification of ICD-10. A second is to implement a currently available WHO coding system, such as ICD-11 for Mortality and Morbidity Statistics (ICD-11-MMS). A third is to adopt an innovation model intended to enhance ICD-11-MMS with content, features, and capabilities that take advantage of the expansiveness of the ICD-11's information network.⁶ This model is designed to support automated coding and review processes to increase reliability and decrease the burden on clinicians and coding professionals.

ICD-10-CM

The coding system currently in use in the US is ICD-10-CM, which is a unique derivative of ICD-10 that took years to develop.⁷ ICD-10-CM reduces nuanced natural clinical language to sets of discrete codes, each of which embodies a clinical concept. Some codes represent single clinical concepts, while others combine (or “pre-coordinate”) two or more clinical concepts into single codes.

Receivers of the coded concepts can interpret transmitted information accurately within the constraints of the classification and its rules. However, because there is no formal mechanism to link discrete codes other than by creating pre-coordinated compound codes or by classifying some codes as secondary diagnoses, some alternatives might each match a clinical concept partially but differently. Consequently, coders might select one or another of the alternative codes, introducing variation in codes representing identical situations. Furthermore, even when coding is consistent, important clinical details may be lost. For example, the presence, severity, or active treatment of one diagnosis might be causing or exacerbating another diagnosed condition. Thus, while pre-coordination can capture complex clinical situations, ICD-10-CM's coding conventions generally fall far short of maximizing the information content from among thousands of available codes. Clinicians and coders must often accept just an approximation of a patient's status to bill for a service or to fulfill a reporting requirement. Much of the disappointment with ICD-10-CM has stemmed from the reality that many of the more complicated and specific codes that were introduced to improve coding and to benefit downstream use cases were rarely used.⁸

ICD-11-MMS

ICD-11 provides a basis for much clearer and more complete communication due to its architecture, and it introduces automated coding tools that foreshadow potentially

lower burden for coders. According to Harrison et al,⁹ “ICD-11’s information framework comprises a semantic knowledge base (the Foundation), a biomedical ontology linked to the Foundation and classifications derived from the Foundation.” ICD-11-MMS is a classification (referred to as a linearization) that is derived from the Foundation. It was created by extracting and organizing Foundation entities that are suitable for ICD-11-MMS’s use case, where the assigned codes create a familiar strict hierarchy that makes ICD-11-MMS appear like ICD-10 and its predecessors. ICD-11-MMS stem codes represent clinical entities (nouns), and extension codes can be used to modify or add additional information (adjectives).

ICD-11 MMS relies much less on the availability of pre-coordinated codes. Instead, clinical specificity typically requires the post-coordination of codes into clusters (phrases) using combinations of stem and extension codes. Thus, a code cluster can describe a condition, its associated manifestation or causality, acuity and severity, etiology, anatomical location, and more using combinations of relevant individual codes and the appropriate coding rules. Post-coordination will be familiar to many because it is used in SNOMED CT.¹⁰ Largely for that reason, data exchange protocols using Fast Healthcare Interoperability Resources (FHIR) also support post-coordination. Nevertheless, post-coordination will be entirely unfamiliar to large segments of the clinician and coding professional communities and would involve substantial training, orientation (and persuasion) for widespread adoption.

We examine three limitations to ICD-11-MMS in this paper. The first two are foundational to mortality and morbidity classification, and the third relates to uses of ICD-11-MMS for other than mortality.

- First, ICD-11-MMS references only a subset of the entities in the Foundation and supplies ‘Y codes’ (ie, other specified) to truncate details below a certain point.
- Second, ICD-11-MMS identifies single parent-child relationships and ignores other plausible ancestries to achieve a strict mono-hierarchy. For example, stomach cancer traces back to Neoplasms and to Diseases of the Digestive System in the Foundation, but only to Neoplasms in ICD-11-MMS. It is likely that descriptive statistics is the only use case that benefits from a strict mono-hierarchy, because of the advantages of aggregating elements into larger and larger categories via only one pathway. Many other use cases can benefit from poly-hierarchies because many concepts relate to other concepts in multiple ways, and it can be important to specify such relationships and contexts.
- Third, the rules for selecting and ordering codes are limited, resulting in multiple ways that a unique concept can be coded. Countries are examining their needs for expanding WHO’s guidance on selecting and ordering codes to achieve more accurate and consistent reporting than currently is possible.

ENHANCED ICD-11-MMS (ICD-11-CCL)

Because ICD-11-MMS is built from the comprehensive clinical Foundation, it can be enhanced to more accurately and reliably capture clinical detail. In this paper, we use an enhanced ICD-11-MMS (ICD-11-CCL) described elsewhere⁶ to illustrate this approach. These enhancements could facilitate automation of coding functions, thereby reducing both the overall burden of coding and the dependence on clinicians’ knowledge of the coding process.

METHODS

In this study, we compare the capability of the three coding systems to represent realistic clinical scenarios through individual codes or code clusters for clinical decision-making and healthcare-related analyses using two hypothetical patient profiles. Clinical communication for these purposes requires a solid base of high-quality patient data generated from a multitude of providers over the evolution of patient care.

The patient profiles were first described using relevant ICD-10-CM codes. ICD-11-MMS code clusters were created for the patient profiles using the 2024 WHO coding tool.¹¹ A single, unique ICD-11-CCL code cluster for each patient profile was created using current ICD-11-CCL codes and syntax as detailed in the Supplement. Each profile consists of a natural clinical language description of a single medical condition assessed and summarized by a single physician based on a discrete healthcare encounter. The profiles were translated into codes and code clusters in ICD-10-CM, ICD-11-MMS, and ICD-11-CCL. Each set of codes for each condition then was used as input for translation back into natural clinical language. The US ICD-10-CM Official Guidelines for Coding and Reporting were not applied as there are no corresponding US ICD-11-MMS guidelines. The only coding guidance used were the official conventions and instructions provided with ICD-11-MMS.

The fidelity of these translations to the original description from which they were derived was assessed. This is analogous to an English-speaking individual using a translator to communicate with a Russian-speaking audience. Without knowing the language, the speaker can assess the fidelity of the Russian translation by examining the content after it has been translated back from Russian into English. This arrangement could usher in a new era of coding in which clinicians are no longer required to seek and identify the most suitable pre-coordinated codes or create and interpret detailed coded clusters. Instead, they only will need to provide clinical content in a stylized natural clinical language format that can be interpreted accurately by automated electronic coding systems. Clinicians then can ensure the adequacy of resulting coded clusters by reviewing the reverse translation back into natural clinical language and iteratively editing input when this is appropriate. As such, the widespread burden on clinicians and coders could plummet.

A looming controversy regarding the potential implementation of ICD-11 is how best to facilitate the mapping

of resulting coded clusters cleanly into ICD-10-CM codes in either direction. Of first concern is the question of how ICD-10-CM codes can be replicated in ICD-11. An example is the specification of quality measures, which often are linked directly to specific ICD-10-CM codes. How can those specifications be updated to support ICD-11 and yield seamless longitudinal measurements without creating gaps, discontinuities, or incongruities? Research to date¹² has explored how ICD-11-MMS might be used to mirror ICD-10-CM codes and has uncovered substantial deficiencies. Meanwhile, an even bigger and potentially insurmountable challenge is mapping ICD-11 code clusters back into equivalent ICD-10-CM codes. It seems unlikely that ICD-11-MMS could support one-to-one mapping of code clusters back to ICD-10-CM because ICD-11-MMS's post-coordination features permit unique clinical concepts to be expressed in many different combinations and permutations of codes.

Accordingly, our limited study explored the capability of each coding system to approximate a one-to-one map between the clinical concept embodied in the original description of two scenarios and the content and structure of the transcribed coded data sets. Particular attention was devoted to the accuracy of the coded data, to the correct delineation of relationships among the coded data elements, and to the capability of each system to capture the clinical nuance needed to assess the validity of assigned diagnoses and clinical findings, the evolution of each patient's clinical status, and the appropriateness of clinical interventions.

Finally, the fixed hierarchical structure of ICD-11-MMS does not allow capture of clinically valid alternative perspectives for discrete clinical concepts that are delineated in ICD-11's semantic knowledge base. For example, an amebic hepatic abscess may be viewed by a gastroenterologist as a hepatic abscess caused by an amebic infestation, ie, a disease of the digestive system. However, an infectious disease specialist might view the same condition as amebiasis manifested as a hepatic abscess, ie, a parasitic disease. A coding system that enables a clinician to specify the lens through which each clinical concept is being viewed transcends one limited to a single view.¹³

TWO SCENARIOS

Scenario 1: A patient presents to a pulmonologist for follow-up of an episode of acute salmonella pneumonia.

A 67-year-old woman with a history of chronic obstructive lung disease and a one-second forced expiratory volume of 82 percent returns to clinic with acute, rapid onset, active nosocomial pneumonia. Her sputum culture was positive for salmonella bongori. Her chest X-ray revealed a patchy left lower lobe pulmonary infiltrate. Arterial oximetry reveals moderate hypoxemia secondary to her pneumonia. She is currently being treated with cefalexin. Diagnosis: pneumonia of moderate severity caused by salmonella bongori.

Scenario 2: A patient with acute enterococcal endocarditis is evaluated by a cardiovascular specialist.

A 76-year-old man is seen for the first time with a history of permanent atrial fibrillation currently treated with propranolol and apixaban, and a history of chronic NYHA Class II systolic left ventricular failure with an ejection fraction of 55% documented by echocardiogram. He presents with moderate acute bacterial endocarditis of the aortic valve caused by enterococcus faecalis, documented by echocardiogram and blood culture, currently being treated with ampicillin. He has complicating moderate acute aortic regurgitation, currently being treated with benazepril, mild acute left ventricular dilation, and an ejection fraction of 45%, all documented by echocardiogram. A chest X-ray reveals mild pulmonary vascular congestion. Diagnosis: moderate acute bacterial endocarditis caused by enterococcus faecalis complicated by moderate acute aortic regurgitation and moderate acute left ventricular failure.

PERFORMANCE CRITERIA

The American Health Information Management Association's (AHIMA) characteristics of data quality¹⁴ were adapted to assess the capabilities of alternative coding systems to capture high quality information about the health of individuals. The four AHIMA characteristics used in this preliminary assessment of ICD-11-MMS and ICD-11-CCL relative to ICD-10-CM were: (1) granularity, ie, the level of detail required to capture important clinical nuance; (2) comprehensiveness, ie, all elements required for evaluations of clinical performance, healthcare decision-making, and other secondary use cases; (3) accuracy, ie, freedom from errors; and (4) relevancy, ie, capability of coded data to support accurate assessments of the appropriateness of diagnostic and therapeutic interventions which then can serve as the basis for important secondary healthcare data applications.¹⁴ Because multiple representations of the same concept can compromise the creation of inclusive analytic subgroups, unique representation (ie, one-to-one correspondence of each code or cluster to a single clinical concept) was added to the list of performance criteria. Finally, because the burden on clinicians collecting data is an important consideration in the adoption of any health information system, these systems' potential to support the application of user-friendly tools to acquire properly processed coded data also were compared.

RESULTS

SCENARIO 1: SALMONELLA PNEUMONIA

Translation into ICD-10-CM resulted in codes for salmonella pneumonia, chronic obstructive pulmonary disease with acute lower respiratory infection, and hypoxemia. (See the Supplement for these ICD-10-CM codes.) It is clear from the text that salmonella pneumonia is the primary, (first) diagnosis in the set. The ordering of the other two codes was arbitrary.

For ICD-10-CM, a reasonable translation into natural clinical language would be:

A patient presents with acute salmonella pneumonia, hypoxemia, and chronic obstructive pulmonary disease with acute lower respiratory infection.

Translation into ICD-11-MMS resulted in three distinct clusters. The first cluster focused on salmonella pneumonia and contained 16 codes, the second contained two codes for chronic obstructive pulmonary disease, and the third contained three codes for respiratory compromise. (See the Supplement for details of each cluster.) It is clear from the text that the first cluster was primary. The ordering of the other two clusters was arbitrary.

For ICD-11-MMS, a reasonable translation into natural clinical language would be:

A 65-to-80-year-old female is seen for follow-up of a salmonella bongori infection manifested as pneumonia of the left lower lobe of the lung. The diagnosis was confirmed by culture and imaging. The pneumonia was nosocomial, with an acute, rapid onset. It currently is active, of moderate severity, and is being treated with cefalexin. Related conditions include mild chronic obstructive pulmonary disease (GOLD 1) with an FEV1 \geq 80%, and asphyxia [sic] confirmed by laboratory examination.

Translation into ICD-11-CCL based on ICD-11-CCL's syntactical rules and explicit coding hierarchy resulted in a single uniquely ordered cluster consisting of 28 codes. (See the Supplement for details of this cluster.) Also, although ICD-11-MMS classifies salmonella pneumonia as an infectious disease, ICD-11-CCL explicitly documents that the pulmonologist has selected the alternative view in which salmonella pneumonia was classified as a respiratory disease.

For ICD-11-CCL, a reasonable translation into natural clinical language would be:

A 65-to-80-year-old female is seen for follow-up of pneumonia caused by salmonella bongori affecting the left lower lobe of the lung. The diagnosis was confirmed by culture and imaging. The pneumonia was nosocomial, with an acute, rapid onset. It currently is active, of moderate severity, and is being treated with cefalexin. The patient has complicating moderate hypoxia confirmed by laboratory examination. Related conditions include mild chronic obstructive pulmonary disease (GOLD 1) with an FEV1 \geq 80%.

SCENARIO 2: ENTEROCOCCAL ENDOCARDITIS

Translation into ICD-10-CM resulted in codes for acute and subacute infective endocarditis, enterococcus as the cause of diseases classified elsewhere, nonrheumatic aortic (valve) insufficiency, acute on chronic systolic (congestive) heart failure, and permanent atrial fibrillation. (See the Supplement for these ICD-10-CM codes.) It is clear from the text that endocarditis is the primary (first) diagnosis in the set. The ordering of the other four codes was arbitrary.

For ICD-10-CM, a reasonable translation into natural clinical language would be:

A patient presents with enterococcal acute and subacute infective endocarditis with coexisting non-rheumatic aortic valve insufficiency, acute on chronic systolic congestive heart failure, and permanent atrial fibrillation.

Translation into ICD-11-MMS resulted in five distinct clusters. The first cluster contained 11 codes, the second contained eight codes, the third contained five codes, the fourth contained four codes, the fifth contained three codes. (See the Supplement for details of each cluster.) It is clear from the text that the first cluster was primary. The ordering of the other four clusters was arbitrary.

For ICD-11-MMS, a reasonable translation into natural clinical language would be:

A 65-to-80-year-old male is seen for the first time with moderate acute bacterial endocarditis of the aortic valve caused by enterococcus faecalis, diagnosed by culture and imaging and treated with ampicillin. He has complicating moderate acute left ventricular failure with a midrange ejection fraction diagnosed by imaging secondary to moderate acute aortic valve insufficiency diagnosed by imaging that is being treated with benazepril. Related conditions include a mild acute specified acquired left ventricular abnormality diagnosed by imaging, permanent atrial fibrillation currently treated with apixaban and propranolol, and chronic NYHA Class II left ventricular failure with preserved ejection fraction diagnosed by imaging.

Translation into ICD-11-CCL based on ICD-11-CCL's syntactical rules and explicit coding hierarchy resulted in a single uniquely ordered cluster consisting of 40 codes. (See the Supplement for details of this cluster.)

For ICD-11-CCL, a reasonable translation into natural clinical language would be:

A 65-to-80-year-old male is seen for the first time with moderate acute bacterial endocarditis of the aortic valve caused by enterococcus faecalis which was documented by culture and imaging and is currently being treated with ampicillin. Complications include moderate acute aortic valve insufficiency currently being treated with benazepril, mild acute left ventricular dilation, and moderate acute left ventricular failure with midrange ejection fraction, all of which were diagnosed by imaging. Related conditions include permanent atrial fibrillation currently treated with apixaban and propranolol and chronic NYHA Class II left ventricular failure with a preserved ejection fraction diagnosed by imaging.

SUMMARY OF PRELIMINARY ASSESSMENT

[Table 1](#) summarizes the preliminary assessment of the three coding systems based on the six performance characteristics described above. The rationale for assigned scores is contained in the Supplement.

Table 1. Preliminary Assessment of ICD-11-MMS and ICD-11-CCL Compared to ICD-10-CM as the Reference Standard

Performance Criteria	MMS*	CCL*
Granularity (clinical nuance)	+	++
Comprehensiveness	-	+
Accuracy	-	+
Relevancy (clinical context)	0	+
Unique representation	--	0
User-friendly coding support	+	++

* 0 = similar; plus = better; minus = worse

DISCUSSION

The three coding systems differed significantly in the level of detail conveyed, the interrelationships among the clinical elements, and the general ability to provide a comprehensive and clinically nuanced representation of a patient's condition according to the clinician and its impact on the clinician's choice of interventions.

ICD-10-CM

The eight ICD-10-CM standalone codes provided an accurate high-level overview of the clinical information contained in the two patient profiles. This is at least in part an artifact of our choice of relatively straightforward scenarios as illustrative examples for which precise ICD-10-CM codes were likely to be available, and our plausible but not reliable assumption that a coder would identify and utilize the most accurate available codes. However, ICD-10-CM failed to capture some potentially important clinical nuances in the original scenario, including severity of illness, specific location of the disease, specific causative organism, and specific treatments being undertaken. Furthermore, the code lists did not indicate that the first patient's hypoxemia was a complication of her pneumonia and not a manifestation of her chronic lung disease. Similarly, it was unclear from the coded data whether the second patient's aortic regurgitation was entirely due to his endocarditis or whether aortic regurgitation was present prior to the acute onset of the infection.

ICD-11-MMS

The eight ICD-11-MMS clusters with a total of 52 codes provided a similarly accurate overview with a couple of notable exceptions. Because ICD-11-MMS does not have a specific code for hypoxemia, the ICD-11-MMS coding tool search for hypoxemia returned the code for asphyxia, which is defined as "a life-threatening condition in which oxygen is prevented from reaching the tissues by obstruction or damage to any part of the respiratory system."¹¹ This substitution vastly overstated the severity of the moderate hypoxemia described in the natural clinical language abstract. Another issue was the translation of the ICD-11-MMS code for mild acute left ventricular dilation back into a clinically

uninformative description (ie, "a mild acute specified acquired left ventricular abnormality diagnosed by imaging"), which failed to provide insight into the nature of this abnormality (left ventricular dilation).

Finally, the ICD-11-MMS clusters captured substantial important clinical nuances that ICD-10-CM did not, including information about causality, acuity, severity, anatomical location, objective support for diagnoses, and current medications.

ICD-11-CCL

ICD-11-CCL summarized each scenario in a single cluster with 28 codes for the first patient and 40 codes for the second patient. Because it can access all entities in the ICD-11 Foundation, ICD-11-CCL achieved high-level accuracy by utilizing the ICD-11-CCL code for hypoxia instead of the default MMS code for asphyxia, and the ICD-11-CCL code for acquired left ventricular dilation for the default MMS code for other specified acquired left ventricular abnormality.

In addition, because of its syntactical precision, ICD-11-CCL enables specialists to capture their clinical perspectives rather than be limited to the default taxonomy of ICD-11-MMS. For example, in the first scenario, ICD-11-CCL assigned each stem code to the ICD-11 chapter that captured a pulmonary specialist's viewpoint rather than an infectious disease specialist's, and it related secondary stem codes explicitly to their initial stem code as manifestations or related comorbidities of a respiratory condition.

Furthermore, while both ICD-11-MMS and ICD-11-CCL made use of all available extension codes to capture important clinical nuance, ICD-11-CCL's codes and syntactical rules ordered these extension codes so that each clinical concept, no matter how complex, had a unique representation. In contrast, in these two scenarios the ICD-11-MMS clusters and the codes they contain could be arranged in numerous different ways. Hence, this additional feature of ICD-11-CCL can support rapid identification of individual concepts without the need to evaluate numerous possible variations to determine which are simple permutations of one another and which differ in clinically important ways. Illustrating this capability with input as complex as paragraphs in full clinical notes implies similar capability to provide unique representations of progressively smaller or simpler scenarios. In other words, ICD-11-CCL alone might be expected to support one-to-one matching with the vast majority of ICD-10-CM codes or other clinical concepts of interest and the creation of seamless longitudinal data sets across the transition from ICD-10-CM to ICD-11.

COMPARISON OF APPROPRIATENESS OF CLINICAL DECISION-MAKING ASSESSED USING ICD-10-CM AND ICD-11-CCL

Clinical communication and healthcare-related analyses require computer-friendly highly nuanced clinical data that provide an accurate representation of the evolution of each patient's health status. The appropriateness of diagnostic

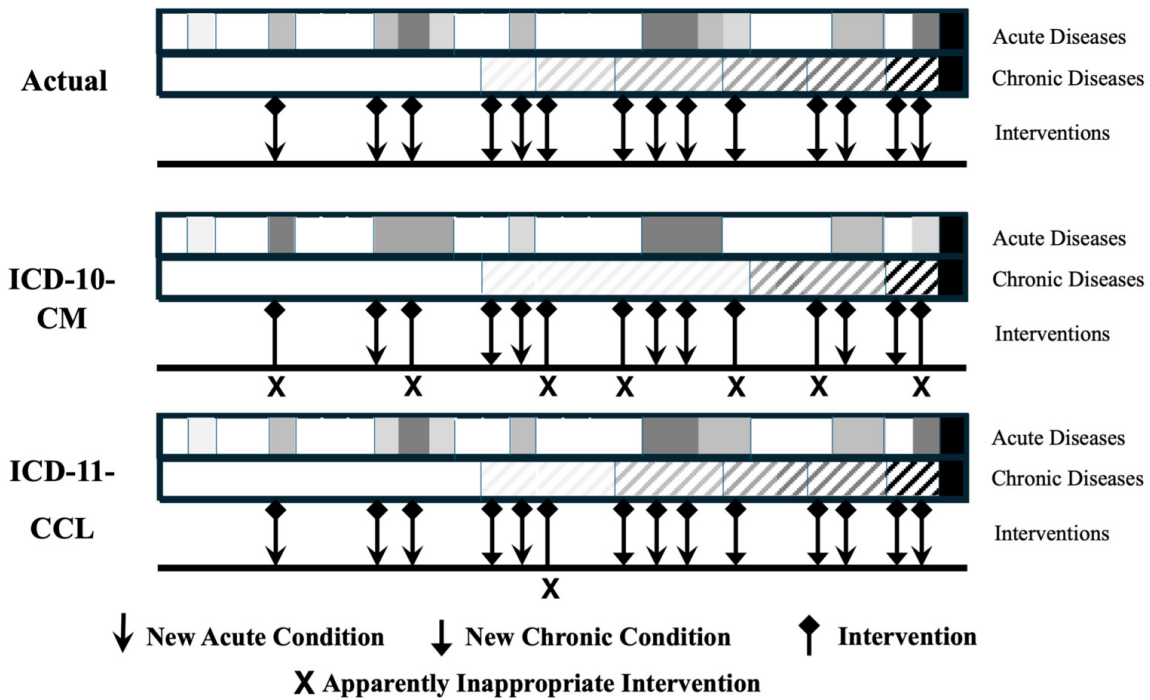


Figure 1. Evolution of Patient Health Status and Appropriateness of Clinical Interventional Responses As Evaluated Based on Actual Health Status, on ICD-10-CM Reported Health Status, and on ICD-11-CCL Reported Health Status

First, the actual evolution of the patient’s health status is illustrated both for acute and for chronic conditions, shown left to right on a horizontal axis of time. The occurrence of acute conditions is illustrated by gray rectangles that become darker as severity increases and black when the patient dies. The progressively increasing burden of chronic diseases is illustrated below in a similar manner. Diamonds indicate 14 clinical interventions based on the accurate assessments of acute and chronic conditions, designated by arrows. Next, the evolution of the same patient’s health status based on ICD-10-CM codes is illustrated in the middle section below the actual status. Loss of clinical nuance is consistent with prior² and current empirical findings. In this illustration, seven of the 14 appropriate interventions would no longer be considered appropriate (marked by X). For example, the chronologically first intervention appears to be too intensive because the ICD-10-CM code for an acute condition suggested greater than actual severity (darker shaded box), and the sixth chronological service appears to be inappropriate because ICD-10-CM coding missed the progression of severity in a chronic condition. Finally, the evolution of the same patient’s health status is illustrated based on ICD-11-CCL clusters. Loss of clinical nuance again is consistent with prior⁶ and current empirical findings. In this illustration, only one of the 14 appropriate interventions would no longer be considered appropriate. As above for ICD-10-CM, the sixth chronological service appears to be inappropriate because, in this illustration, ICD-11-CCL coding still missed the progression of severity in one chronic condition.

and therapeutic interventions can be assessed accurately only if the clinical indications for these interventions are apparent from analyzable data. High-level overviews based solely on ICD-10-CM standalone diagnostic and descriptive codes often fail to justify appropriate interventions, while more nuanced representations of the evolution of a patient’s health status such as those obtained using ICD-11-CCL can better clarify the rationale and timing underlying appropriate clinical care.

The improved capability of an enhanced ICD-11-MMS such as ICD-11-CCL to support important analytic use cases is illustrated in [Figure 1](#). This figure illustrates the potential accuracy of assessments of the appropriateness of clinical interventions when a patient’s health status is reported using ICD-10-CM and using ICD-11-CCL. It was prepared to illustrate the advantages of ICD-11 compared to ICD-10 using ICD-10-CM and ICD-11-CCL. Future research could evaluate ICD-11-MMS using criteria similar to those employed in this illustration.

STUDY LIMITATIONS AND RECOMMENDED NEXT STEPS

The current study is a limited initial exploration of two scenarios for evaluating potential benefits and burdens associ-

ated with the implementation of an enhanced ICD-11-MMS (eg, ICD-11-CCL), as a successor to ICD-10-CM or other comparable clinical modifications used in countries other than the US. It provides evidence that this transition is potentially beneficial in its effect on the quality, transmission, and use of coded clinical information. Additional research, development, and planning will be required to realize the full potential of this important innovation in the most expeditious and cost-effective manner.

CONCLUSIONS

This study explored the capability of three coding systems to represent realistic clinical scenarios through individual codes or code clusters for two natural language scenarios. It illustrated the limitations of each coding system in capturing, representing, and articulating clinically nuanced information by first coding and then translating the coded data back into natural clinical language. In addition, it investigated how several enhancements to ICD-11-MMS, as embodied in ICD-11-CCL, could facilitate comprehensive improvements in clinical data capture by translating not just isolated concepts, but whole “paragraphs” (code clusters) of information representing patients’ needs, services, and by

extension, the trajectory over time of each patient's clinical status in relation to healthcare services received.

Within limits established by ICD-11's Foundation, enhanced ICD-11-MMS could facilitate data collection, organization, and universal translation of digital data transmissions supporting a variety of use cases such as quality measurement, accurate comparative assessments of the cost and value of healthcare services, and clinical and health services research. Some use cases might benefit from incremental improvements in enhanced ICD-11-MMS data, while others might not require all available details. In either case, the adoption of an enhanced ICD-11-MMS data system would provide clinically nuanced data required to discern clinical context well enough to effectively support important secondary data applications and broad, industry-wide programs. These include inferences about patients' needs, clinical appropriateness and gaps in services, efficient resource use, the incidence and severity of untoward events, and the relation of clinical interventions to patient outcomes generally.

In addition, having a stylized natural clinical language format that can be interpreted accurately by automated electronic coding systems would reduce the dependence on

clinicians' knowledge of the details of the coding process. Hence, although a transition from ICD-10-CM to an enhanced ICD-11-MMS such as ICD-11-CCL will entail education on its structure, once this is accomplished, automating the coding function, which ICD-11-CCL supports, would substantially reduce the overall burdens associated with reporting, deciphering, communication, and analyses from those associated with ICD-10-CM.

.....

DISCLOSURES

The authors have nothing to disclose.

FUNDING

The authors received no funding for this research.

Submitted: November 21, 2024 EDT. Accepted: March 15, 2025 EDT.

REFERENCES

1. Feinstein JA, Gill PJ, Anderson BR. Preparing for the International Classification of Diseases, 11th revision (ICD-11) in the US health care system. *JAMA Health Forum*. 2023;4(7):e232253. doi:[10.1001/jamahealthforum.2023.2253](https://doi.org/10.1001/jamahealthforum.2023.2253). PMID:37505488
2. Fung KW, Xu J, Bodenreider O. The new International Classification of Diseases 11th edition: a comparative analysis with ICD-10 and ICD-10-CM. *Journal of the American Medical Informatics Association*. 2020;27(5):738-746. doi:[10.1093/jamia/ocaa030](https://doi.org/10.1093/jamia/ocaa030). PMID:32364236
3. Chute CG, Huff SM, Ferguson JA, Walker JM, Halamka JD. There are important reasons for delaying implementation of the new ICD-10 coding system. *Health Affairs (Millwood)*. 2012;31(4):836-842. doi:[10.1377/hlthaff.2011.1258](https://doi.org/10.1377/hlthaff.2011.1258)
4. Topaz M, Shafran-Topaz L, Bowles KH. ICD-9 to ICD-10: evolution, revolution, and current debates in the United States. *Perspectives in Health Information Management*. 2013;10(Spring):1d. <https://pubmed.ncbi.nlm.nih.gov/23805064/>
5. Monestime JP, Mayer RW, Blackwood A. Analyzing the ICD-10-CM transition and post-implementation Stages: a public health institution case study. *Perspectives in Health Information Management*. 2019;16(Spring):1a. <https://pubmed.ncbi.nlm.nih.gov/31019430/>
6. Giannangelo KL, Pine MB, Tompkins CP. Leveraging an innovation model to facilitate ICD-11 implementation. *Perspectives in Health Information Management*. 2024;21(1):1d. <https://pmc.ncbi.nlm.nih.gov/articles/PMC11102059/>
7. NCVHS (National Committee on Vital and Health Statistics). ICD-10-CM Major Milestones. June 26, 2019. <https://ncvhs.hhs.gov/wp-content/uploads/2019/12/Timelines-for-Adoption-and-Implementation-of-ICD-10.pdf>
8. Zhu JR, Galivanche AR, Gouzoulis MJ, et al. International Classification of Diseases coding for inflammatory arthritides. *JAMA Network Open*. 2024;7(4):e246544. doi:[10.1001/jamanetworkopen.2024.6544](https://doi.org/10.1001/jamanetworkopen.2024.6544)
9. Harrison JE, Weber S, Jakob R, Chute CG. ICD-11: an international classification of diseases for the twenty-first century. *BMC Medical Informatics and Decision Making*. 2021;21(Suppl 6):206. doi:[10.1186/s12911-021-01534-6](https://doi.org/10.1186/s12911-021-01534-6)
10. Miñarro-Giménez JA, Martínez-Costa C, López-García P, Schulz S. Building SNOMED CT post-coordinated expressions from annotation groups. *Studies in Health Technology and Informatics*. 2017;235:446-450. <https://ebooks.iospress.nl/publication/46380>
11. World Health Organization. ICD-11 Coding Tool: Mortality and Morbidity Statistics (MMS) 2024-01. 2024. Accessed November 19, 2024. https://icd.who.int/ct/icd11_mms/en/release
12. Fung KW, Xu J, McConnell-Lampsey S, Pickett D, Bodenreider O. Feasibility of replacing the ICD-10-CM with the ICD-11 for morbidity coding: A content analysis. *Journal of the American Medical Informatics Association*. 2021;28(11):2404-2411. doi:[10.1093/jamia/ocab156](https://doi.org/10.1093/jamia/ocab156)
13. Cimino JJ. Desiderata for controlled medical vocabularies in the twenty-first century. *Methods of Information in Medicine*. 1998;37(4-5):394-403. doi:[10.1055/s-0038-1634558](https://doi.org/10.1055/s-0038-1634558)
14. AHIMA. Healthcare Data Governance. 2022. Accessed November 19, 2024. <https://www.ahima.org/media/pmcb0fr5/healthcare-data-governance-practice-brief-final.pdf>

SUPPLEMENTARY MATERIALS

Supplementary Material

Download: https://ahisp.scholasticahq.com/article/133904-a-transformative-innovation-to-clinical-coding-for-better-data-and-less-burden/attachment/275087.docx?auth_token=0NpIBLV34X013N1wz9i0
