

Metrics and tools for consistent cohort discovery and financial analyses post-transition to ICD-10-CM

RECEIVED 31 January 2014
 REVISED 27 August 2014
 ACCEPTED 19 October 2014
 PUBLISHED ONLINE FIRST 13 March 2015



OXFORD
 UNIVERSITY PRESS

Andrew D Boyd^{1,2,*}, Jianrong ‘John’ Li^{3,4}, Colleen Kenost^{3,4}, Binoy Joese^{2,5},
 Young Min Yang¹, Olympia A Kalagidis¹, Ilir Zenku², Donald Saner^{4,6}, Neil Bahroos², Yves A Lussier^{2,3,4,5,6,7,8,9,10,S,*}

ABSTRACT

In the United States, International Classification of Disease Clinical Modification (ICD-9-CM, the ninth revision) diagnosis codes are commonly used to identify patient cohorts and to conduct financial analyses related to disease. In October 2015, the healthcare system of the United States will transition to ICD-10-CM (the tenth revision) diagnosis codes. One challenge posed to clinical researchers and other analysts is conducting diagnosis-related queries across datasets containing both coding schemes. Further, healthcare administrators will manage growth, trends, and strategic planning with these dually-coded datasets. The majority of the ICD-9-CM to ICD-10-CM translations are complex and nonreciprocal, creating convoluted representations and meanings. Similarly, mapping back from ICD-10-CM to ICD-9-CM is equally complex, yet different from mapping forward, as relationships are likewise nonreciprocal. Indeed, 10 of the 21 top clinical categories are complex as 78% of their diagnosis codes are labeled as “convoluted” by our analyses. Analysis and research related to external causes of morbidity, injury, and poisoning will face the greatest challenges due to 41 745 (90%) convolutions and a decrease in the number of codes. We created a web portal tool and translation tables to list all ICD-9-CM diagnosis codes related to the specific input of ICD-10-CM diagnosis codes and their level of complexity: “identity” (reciprocal), “class-to-subclass,” “subclass-to-class,” “convoluted,” or “no mapping.” These tools provide guidance on ambiguous and complex translations to reveal where reports or analyses may be challenging to impossible. Web portal: <http://www.lussierlab.org/transition-to-ICD9CM/>

Tables annotated with levels of translation complexity: <http://www.lussierlab.org/publications/ICD10to9>

Key words: ICD-10-CM, medical informatics, network patterns, patient cohort, financial analyses, ICD-9-CM

INTRODUCTION

For the last 30 years, health care managers and clinical staff have relied on health data analytics based on diagnosis codes that are recorded via the International Classification of Diseases, ninth version, Clinical Modification (ICD-9-CM). Most encounters with a healthcare professional generate at least one ICD-9-CM code, which allows for analytical study of the reimbursement practices of insurers, such as Medicaid and Medicare. These diagnosis codes enable an initial identification of a cohort of patients for clinical research. ICD-10-CM (the 10th revision) has been promised as an improvement for analytics with the increased fidelity of the diagnosis.⁶ Our prior work has examined the potential impact of the transition from ICD-9-CM to ICD-10-CM mapping forward with the focus on

the ICD-9-CM diagnosis codes.^{1–5} We have also shown that ICD-9-CM diagnosis codes can be categorized into five increasing levels of translation complexity: identity (simple reciprocal coding), class-to-subclass, subclass-to-class, convoluted (highly complex), and no translation (intractable).^{1–5} However, with the upcoming transition to ICD-10-CM, migration back to the ICD-9-CM system will be required for responsible and comparable financial analyses reporting and patient cohort discovery between classifications. The objective of our study is to reduce the risks associated with the disruption in clinical cohort discovery and financial analyses for clinical data coded in ICD-10-CM and ICD-9-CM. We hypothesize that the new science of networks can provide an effective framework to compare and contrast the two coding schemes.

*Correspondence to University of Arizona, Department of Medicine, 1657 East Helen Street, M/C 2465, Tucson, AZ 85721, USA; Tel: (520) 626-0245;

Fax: (520) 626-4824; Lussier.y@gmail.com

University of Illinois at Chicago, Department of Biomedical and Health Information Science, 1919 W. Taylor St. AHS 227 (MC 530), Chicago, IL, 60612, USA,

Tel: (312) 996-8339; Fax: (312) 996-8342; boyda@uic.edu

©The Author 2015. Published by Oxford University Press on behalf of the American Medical Informatics Association.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use,

please contact journals.permissions@oup.com

For numbered affiliations see end of article.

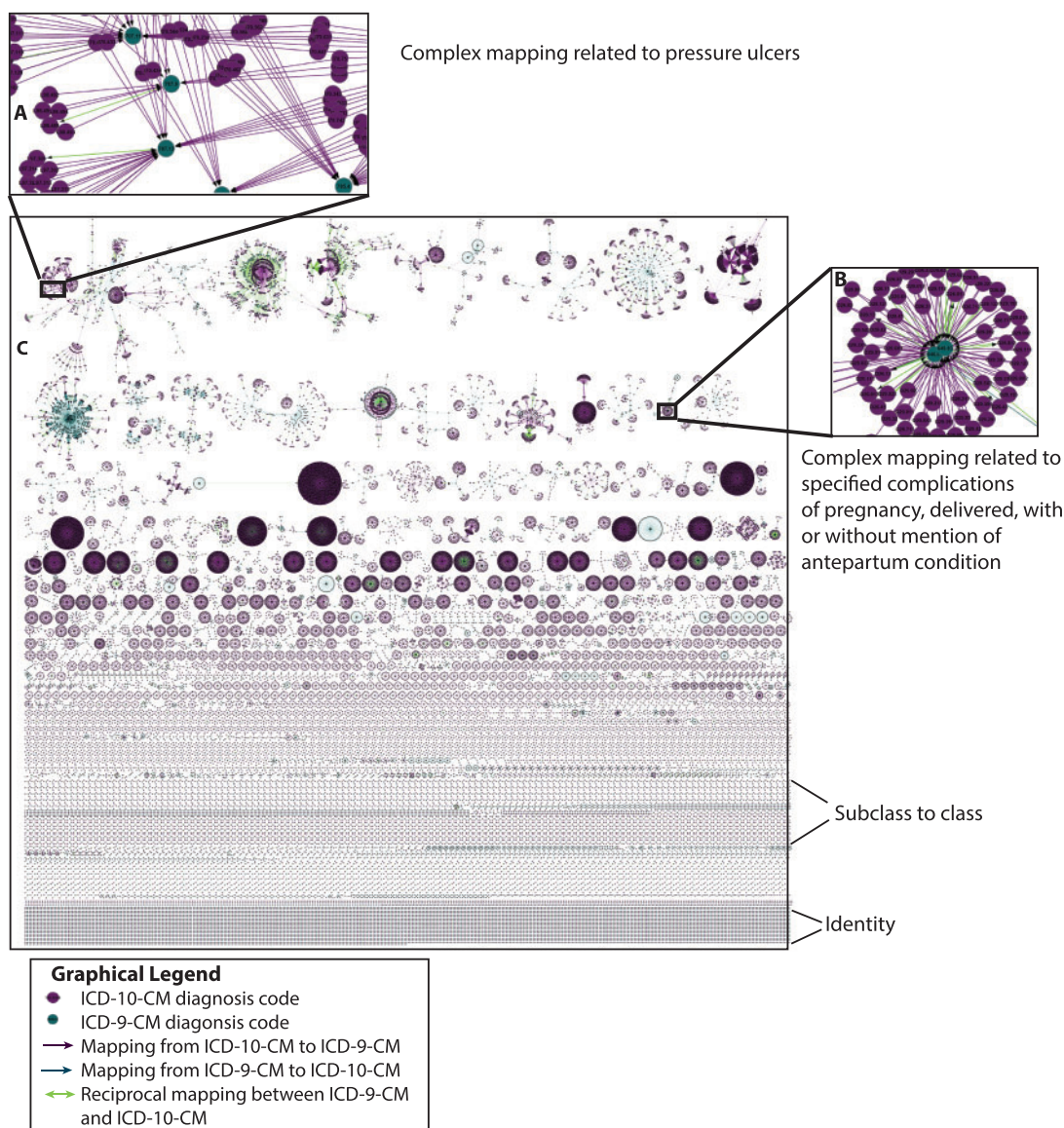
METHODS

Construction of the bidirectional map from GEMs

The US Government has provided bidirectional maps from ICD-9-CM to ICD-10-CM in two different files, called General Equivalence Mappings (GEMs), which are revised annually.⁷

In the present study, we used the 2014 GEMs,⁷ to create a translation large network map of ICD-10-CM to ICD-9-CM using Cytoscape 3.0 with circles representing each ICD-9-CM and ICD-10-CM code and arrows representing the transitions between codes (Figure 1).

Figure 1: ICD-10-CM to ICD-9-CM network demonstrates convoluted terminological translation. The Center for Medicare & Medicaid Services General Equivalence Mappings (GEMs) are used to create the full network initiated (seeded) from ICD-10-CM. The majority of the ICD-10-CM codes do not map straightforwardly to ICD-9-CM codes; 27 distinct types of bilateral relationships (ICD-10-CM to ICD-9-CM motifs, Figure 2) can be observed to deconstruct the network and be used to derive a summary table of complexity according to medical specialty (Figure 3). (A) Detail of the complex and convoluted mapping relation to the ICD-10-CM code of “pressure ulcers.” (B) Detail of the complex ICD-9-CM code related to the ICD-10-CM code of “complications of pregnancy.” (C) A complete representation of the ICD-9-CM to ICD-10-CM transition. Purple and blue circles, respectively, represent ICD-10-CM and ICD-9-CM diagnosis codes. Purple lines indicate a one directional relationship from ICD-10-CM to ICD-9-CM, while blue lines correspond to reverse mapping that are not reciprocal. Green lines represent reciprocal relationships between ICD-9-CM codes and ICD-10-CM codes.



Decomposing the mapping network from the perspective of ICD-10-CM

As the network of Figure 1 is too complex to understand in its entirety, we identified smaller network patterns (Figure 2, motifs) that are meaningful to coders, administrators, and clinicians. As shown in Figure 2, we systematically associated each of the 36 network patterns of ICD-10-CM translation to ICD-9-CM with a level of complexity: “identity” (reciprocal), “class-to-subclass,” “subclass-to-class,” “convoluted,” or “no mapping” (Figure 2, color coding of the cells and legend). We utilized an algorithm that we previously described for the converse mapping of ICD-9-CM to ICD-10-CM.⁴ Specifically, one ICD-10-CM code is used as an input (Figure 2, rows in Greek letters) and follows its translation to ICD-9-CM using the GEMs tables (Figure 2, columns in Roman numerals). As the relationships in GEMs are not always reciprocal, this ICD-9-CM code may paradoxically code back to different ICD-10-CM codes, leading to a complex pattern that we termed “convoluted” (Figure 2).⁴

Clinical classes

The ICD-10-CM diagnosis codes in each of the 21 clinical classes in ICD-10-CM were analyzed together to calculate the percentage of codes in each category (Figure 3B). The categories were derived from the hierarchy within ICD-10-CM. Additionally, each ICD-10-CM clinical class was analyzed for all of the ICD-9-CM diagnosis codes mapping backward to calculate the percentage decrease in the number of codes (Table 1).

Web portal and supplement tables

We created a web portal tool and translation tables to list all ICD-9-CM diagnosis codes related to the specific input of ICD-10-CM diagnosis codes and their level of complexity: “identity” (reciprocal), “class-to-subclass,” “subclass-to-class,” “convoluted,” or “no mapping.” These tools provide guidance on ambiguous and complex translations to reveal where reports or analyses may be challenging to impossible.

Web portal: <http://www.lussierlab.org/transition-to-ICD9CM/>

Tables annotated with levels of translation complexity: <http://www.lussierlab.org/publications/ICD10to9>

RESULTS

Figure 1 provides an overview of a bidirectional network map of Centers for Medicare & Medicaid Services (CMS) approved translations between ICD-9-CM and ICD-10-CM diagnosis codes, using GEMs mappings. We simplified the components of this network into smaller patterns described in Figure 2 (each cells of the table). We further classify these 36 network patterns according to five levels of translation complexity and illustrate each level with one example (Figure 3A). We identified 4127 ICD-10-CM codes with the simple reciprocal translations to ICD-9-CM translations (Figures 2 and 3, identity). Unsurprisingly, as ICD-10-CM is more comprehensive than ICD-9-CM, an additional 536 class-to-subclass relationships were found. Similarly, 7478 subclass-to-class translations were identified. Importantly, a substantial number of

relationships were convoluted (57 013) or had no mapping to ICD-9-CM (669).

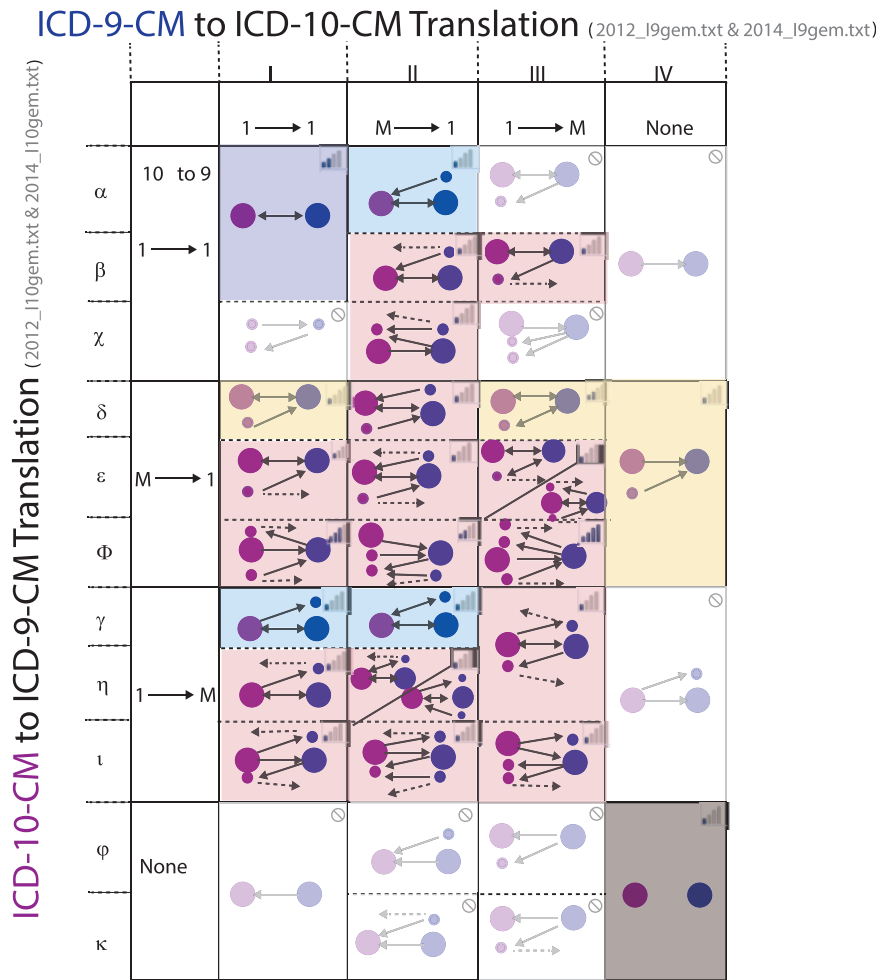
Finally, diagnoses were organized according to the top ICD-10-CM classification codes (Figure 3B, clinical classes). The count of these diagnoses according to their respective levels of translation complexity per clinical classes are provided as a bar graph (Figure 3B). Diseases of the blood and conditions that occur during the perinatal period had the smallest percentage of convoluted diagnosis codes (Figure 3B). External causes of morbidity and injury, poisoning, and certain other consequences of external causes all had the largest percentage of convolution diagnosis codes (Figure 3B). Each clinical class also had a decrease in the number of diagnosis codes from ICD-10-CM diagnosis codes being mapped back to ICD-9-CM diagnosis codes. The clinical class with the largest decrease in the number of concepts is injury, poisoning, and certain other consequences of external causes, with a 95% decrease (Table 1).

DISCUSSION/CONCLUSION

The challenges of the transition from ICD-10-CM to ICD-9-CM are complex and it will be important to rely on queries involving diagnosis codes to assess the financial impact of diseases on clinicians and health care systems. Examining the network graphs of individual ICD-10-CM diagnosis codes from the online tool can provide a quick view of the challenges facing administrators evaluating high-cost diagnoses. The more complex and convoluted the code(s), the more time and resources needed to conduct queries and analytics across the coding divide. We are providing a web portal and annotated tables to help administrators, clinicians, and coders quantitatively and qualitatively evaluate the financial and compliance risks associated to querying and analyzing datasets coded historically in ICD-9-CM and thereafter in ICD-10-CM (“Methods,” Table 2). Indeed, consulting firms and specialty organizations have even recommended dual coding during a few months of the transition period, which very few organizations can afford.⁸

The upcoming implementation of ICD-10-CM diagnosis codes, which will occur in October 2015, will focus attention on the stylistic differences between healthcare facilities. Indeed, with a threefold increase in the number of codes as compared to ICD-9-CM, ICD-10-CM creates ample opportunities for larger variations in coding alternatives between coders, coding agencies, departments, and institutions. In the computer sciences, for example, the standards for comments, variable capitalization, and naming of variables have become normalized to increase readability and understandability across computer programs, as well as to create consistency between software programs and programmers. While standards and guidelines are taught to professional medical coders⁹ who attempt to normalize the stylistic differences, many clinics and physicians create a punch sheet, or a list of codes, that will likely introduce biases in the use of ICD-10-CM codes. In reality, some physicians and coders are becoming “artists,” similar to the original computer programmers, as the guidelines and the misunderstanding of style issues will lead to a wide

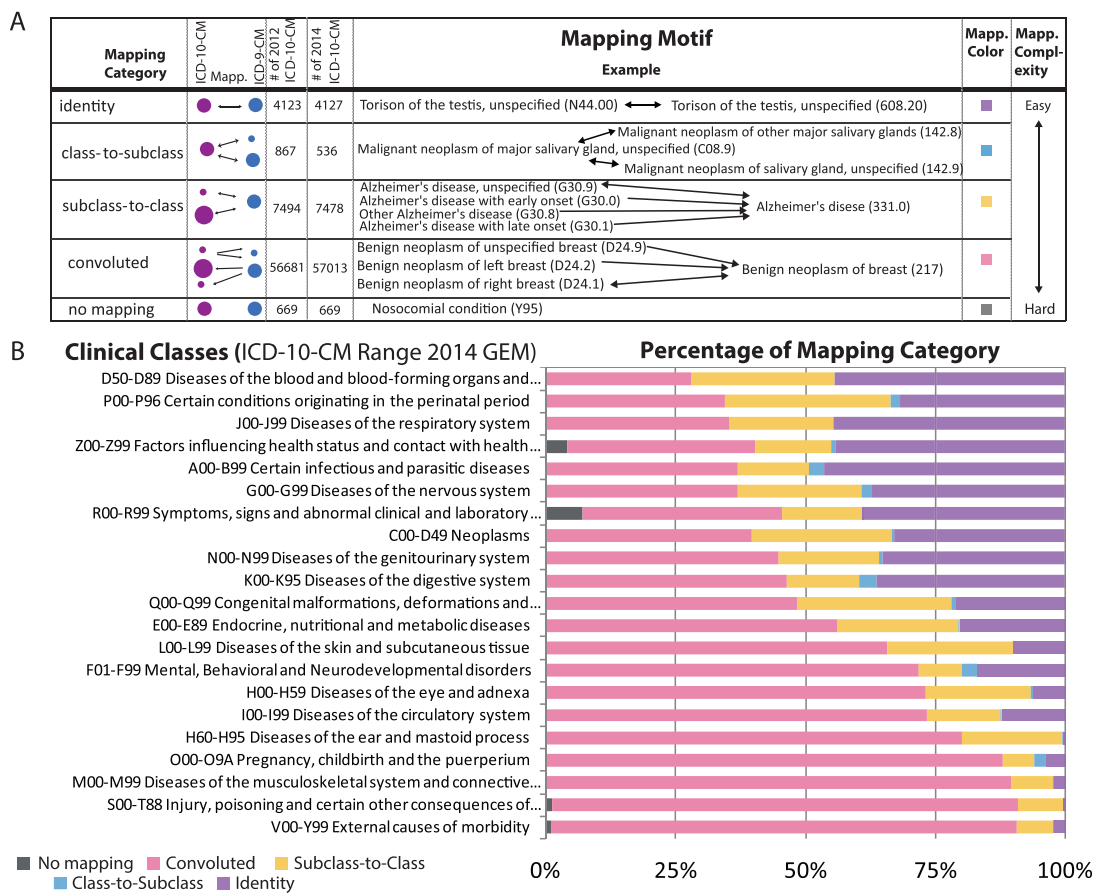
Figure 2: ICD-10-CM to ICD-9-CM motifs enabling comprehension of ICD-10-CM transition to ICD-9-CM. From previously published methodology,³ the complete network from Figure 1 has been converted into individual elementary network motifs, seeded from the ICD-10-CM diagnosis codes. The y-axis is the grouping of ICD-10-CM codes by their relationship to ICD-9-CM codes with fundamental relationships of one-to-one, one-to-many, many-to-one, and none. The x-axis is the grouping of individual motifs with respect to the ICD-9-CM to ICD-10-CM codes after the initial seeding from ICD-10-CM; 75% of all ICD-10-CM seeded diagnosis codes are represented by five network motifs. The most frequently encountered motif is Φ -III, where a nonreciprocal relationship from ICD-10-CM to ICD-9-CM causes convolution. The categorization of individual motifs is: identity = purple; class-to-subclass = blue; subclass-to-class = yellow; pink = convoluted; and gray = no mapping. Blurred matrix cells contain no ICD-10-CM codes in the specified motifs. The color scheme of the categories (blue, light blue, yellow, pink, and gray) link the specific motifs from Figure 2 to the results in Figure 3. In each motif, the four bars in the top right corner represent which quartile the motif is assigned by the number of diagnosis codes.



LEGEND:

	DESCRIPTIVE		ANALYTIC	
	Primary code	Secondary code	Background color: mapping categories	
ICD-10-CM				identity
ICD-9-CM				class-to-subclass
				subclass-to-class
				convoluted
				no mapping
	Arrow direction: mapping		Motif quartiles of code counts (n=72139)	
	ICD10CM to ICD9CM	ICD9CM to ICD10CM		First (n = 25379)
		Reciprocal mapping		Second (n = 12512)
bounded mapping (in the motif)				Third (n = 20502)
unbounded mapping (out of the motif)				Fourth (n = 17869)
				None (n = 0)

Figure 3: Summary of ICD-10-CM motifs and implication in clinical specialty. (A) Twenty-five distinct patterns of mapping motifs (Figure 1, background color) are observed and classified into five mapping categories organized by increasing complexity (A, first column). Each category has a specific color scheme (a fifth column) utilized in the background (Figure 1 and the bar graph of B). The abbreviation, Mapp., refers to mapping. Each mapping category is illustrated with an example. The examples of the two last categories demonstrate the difficulties that may arise from interpreting data collected in ICD-9-CM or ICD-10-CM, which may affect cohort discovery. For example, benign neoplasm of unspecified breast (D24.9) is convoluted since the ICD-9-CM code benign neoplasm of breast maps forward to only the right breast (D24.1). (B) Challenge in patient cohort by clinical specialty. Furthermore, clinical class is unequally impacted, as shown with the percentage of ICD-10-CM codes per mapping category (color coding of the bars from A, column 5). Ten of the clinical classes have >50% convoluted codes.



variation in ICD-10-CM coding and coding practices across healthcare facilities. In practice, a replacement coder or a change in coding agency may significantly affect the accountability of specific groups of codes, often misleading business intelligence, healthcare agencies, or researchers querying these codes over time. The evolution of ICD-10-CM coding will occur as well as new individuals joining teams in the future will learn from their colleagues and predecessors.

All of the above challenges will also impact patient cohorts. One use of patient cohorts is for the evaluation of residencies, fellowships, group practices, and physicians. Knowing the patient cohort of clinicians, for example, facilitates additional training when needed and evaluation of the breadth of patients

for training programs. In identifying patient cohorts, one drawback of using a single GEMs file for a patient cohort is the published limitations in the documentation from the CMS.² The CMS GEMs file mapping diagnosis codes from ICD-9-CM to ICD-10-CM is designed to be comprehensive of all ICD-9-CM codes; 24% of all ICD-10-CM diagnosis codes are included in the CMS GEMs file that maps from ICD-9-CM to ICD-10-CM.⁵ The CMS GEMs file for ICD-10-CM is designed to be comprehensive for all ICD-10-CM but maps backward to a limited subset of ICD-9-CM codes; 70% of all ICD-9-CM diagnosis codes are included in the file that maps from ICD-10-CM to ICD-9-CM.⁵ Due to the way the GEMs files are designed, a researcher

Table 1: ICD-10-CM clinical category loss of fidelity in transition to ICD-9-CM

Clinical categories	ICD-10-CM	ICD-9-CM	% Decrease
S00–T88 Injury, poisoning, and certain other consequences of external causes	39869	1995	95
V00–Y99 External causes of morbidity	6812	593	91
M00–M99 Diseases of the musculoskeletal system and connective tissue	6339	836	87
L00–L99 Diseases of the skin and subcutaneous tissue	769	210	73
H60–H95 Diseases of the ear and mastoid process	642	190	70
H00–H59 Diseases of the eye and adnexa	2432	727	70
I00–I99 Diseases of the circulatory system	1254	435	65
O00–O9A Pregnancy, childbirth, and the puerperium	2155	771	64
E00–E89 Endocrine, nutritional, and metabolic diseases	676	295	56
F01–F99 Mental, behavioral, and neurodevelopmental disorders	723	363	50
Q00–Q99 Congenital malformations, deformations, and chromosomal abnormalities	790	404	49
C00–D49 Neoplasms	1620	935	42
P00–P96 Certain conditions originating in the perinatal period	417	249	40
R00–R99 Symptoms, signs, and abnormal clinical and laboratory findings, NEC	639	382	40
D50–D89 Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism	238	153	36
N00–N99 Diseases of the genitourinary system	591	384	35
K00–K95 Diseases of the digestive system	706	474	33
G00–G99 Diseases of the nervous system	591	416	30
Z00–Z99 Factors influencing health status and contact with health services	1178	855	27
J00–J99 Diseases of the respiratory system	336	261	22
A00–B99 Certain infectious and parasitic diseases	1056	852	19

The first column is a list of all of the ICD-10-CM clinical categories as outlined by CMS when creating the 21 higher order classes. The second column is the number of ICD-10-CM diagnosis codes in each category. The third column is all of the ICD-9-CM diagnosis codes that are mapped through GEMs. The fourth column is the percentage decreased is $(\text{ICD-10-CM codes} - \text{ICD-9-CM codes}) / \text{ICD-10-CM codes}$. Injury, poisoning, and certain other consequences of external causes have the greatest decrease in the number of concepts at 95%.

or evaluator will miss 30% of the ICD-9-CM diagnosis codes and potentially miss patients, as well, if only the ICD-10-CM GEMs files are used.

The comprehensive approach of our tool will allow physicians, training programs, researchers, administrators, health systems, and others to compare diseases across this transition from ICD-9-CM to ICD-10-CM, since the tool for cohort discovery includes all but 1% of the ICD-9-CM codes and 1% of ICD-10-CM codes. The output of the online program creates a single diagnosis code delineated in the complete graph. In our prior paper examining the transition from ICD-9-CM to ICD-10-CM,⁴ the five hardest clinical classes to classify were (1) obstetrics and gynecology; (2) mental disorders; (3) injury and poisoning; (4) external cause of injury; and (5) infectious diseases. In our current analysis for the transition to ICD-10-CM

from ICD-9-CM, the top five clinical classes are (1) external causes of morbidity; (2) injury; (3) diseases of the musculoskeletal system; (4) pregnancy and childbirth; and (5) diseases of the ear and mastoid process. The increase in convolution of mapping back to ICD-9-CM codes of the ear and musculoskeletal system leads one to consider mapping forward to ICD-10-CM to reduce the convolution in the analyses or patient cohort discovery. The convoluted classification includes the ICD-10-CM codes with challenging transitions, which results in 10 of the clinical classes having a convoluted classification of >50% (Figure 2B). This further demonstrates the challenges of mapping from ICD-10-CM back to ICD-9-CM.

The CMS GEMs files do have limitations. A retrospective cohort for asthma using the ICD-10-CM (J45.XX) coding demonstrates this limitation; the complete mapping misses three

Table 2: ICD-10-CM resources for transition to ICD-9-CM

Resource sharing work product	Use case or targeted audience	Description or content
Comprehensive network in high resolution	Within the complex entire network, identify specific ICD-10-CM and ICD-9-CM codes searchable in PDF format. Audience: clinical informaticians and analysts.	http://www.lussierlab.net/publications/ICD10to9/2014Update.pdf
Tables of mapping motifs and categories (.xls format)	Rapid reuse in software developed by health information technologists and informaticians.	http://www.lussierlab.net/publications/ICD10to9/2014categories-motifs.xlsx
SQL database of mapping motifs and categories	Lookup of sql queries and specific results by health system analysts to strategically improve health system operations and plan transition to ICD-9-CM.	http://www.lussierlab.net/publications/ICD10to9/ICD10ToICD9.sql.gz
Web portal	Administrator, clinicians, and other users studying a practice pattern in ICD-10-CM. By typing or pasting in the ICD-10-CM codes the motifs will be generated	http://www.lussierlab.org/transition-to-ICD9CM/ Input: Insert multiple ICD-10-CMs codes of interest Output: Visualization of ICD-10-CM, ICD-9-CM, relationships and associated mapping categories in two formats: dynamic network figure or tabular.

All resources used to develop the motif analysis tool for additional patient cohort discovery and additional analytics.

asthma codes related to chronic obstructive asthma. We did not evaluate the clinical correctness of the GEMs files as a number of studies have already examined the clinical correctness of GEMs, with the error being a small percentage.^{10–13} A major challenge with using GEMs files is the overall complexity of the medical field. For example, in one study, about 20% of the clinicians disagreed about whether or not the GEMs mapping was clinically correct.¹⁴ Also, CMS does not intend to maintain GEMs files indefinitely; in the future, new ICD-10-CM codes may not have any backward mapping due to lack of GEMs mappings.

Applying our novel, web-based tool (<http://www.lussierlab.org/transition-to-ICD9CM/>), financial analysts, administrators, health systems, and clinical researchers can maintain easy access to the ICD-9-CM data sets. The tool can be used to identify possible ambiguities and redundancies of definitions when mapping backward from ICD-10-CM to ICD-9-CM. Some clinical domains of retrospective trials will never be equivalent, such as external causes of morbidity and injury, poisoning, and certain other consequences of external causes, due to significant losses in codes and a high percentage of diagnosis codes labeled as convoluted.

Future studies will need to look for inpatient consistencies across healthcare facilities, as well as variations in procedures and diagnoses. Coding styles will also need to be evaluated for consistency. One possible outcome is an individual who switches companies and brings along new specific ICD-10-CM codes, which could lead to financial enrichment or losses due to the application of different coding styles. This individual could bring either a level of precision or a level of ambiguity to the new ICD-10-CM diagnosis codes, which could lead to an

overall increase/decrease in reimbursements. While analyzing the bidirectional mapping of diagnosis codes is challenging, our innovative, web-based tool helps ensure that individuals with diseases of interest are included in retrospective trials. When researchers insert ICD-10-CM codes of interest into our system, it generates a complete mapping of related diagnosis codes. We have created a quick, user-friendly tool enabling researchers to evaluate diagnosis codes. Our free software program facilitates fast evaluation of queries using ICD-10-CM and ICD-9-CM codes, while illuminating the convolution.

AUTHOR CONTRIBUTIONS

Jianrong Li assisted with the coding and development of the ICD-10-CM tool, reviewed and revised the manuscript, and approved the final submitted manuscript. Binoy Joese, Young Min Yang, Olympia A. Kalagidis, Ilir Zenku, Neil Bahroos, Colleen Kenost, and Donald Saner assisted in the review of the tools, interpretation of the findings, and implications of the findings for the ICD-10-CM tool. They also reviewed and revised the manuscript, and approved the final submitted manuscript. Drs. Boyd and Lussier had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

CONFLICT OF INTEREST DISCLOSURES

Authors did not report any disclosures.

FUNDING/SUPPORT

Drs. Boyd and Lussier are supported in part by Center for Clinical and Translational Sciences of the University of Illinois

(NIH 1UL1RR029879-01, NIH/NCATS UL1TR000050), the Institute for Translational Health Informatics of the University of Illinois at Chicago, The Office of the Vice-President for Health Affairs of the University of Illinois Hospital and Health Science System, the Office of the Vice-President for Health Sciences of the University of Arizona, The Arizona Health Sciences Center, The BIO5 Institute, The National Library of Medicine (K22 LM008308-04), and The University of Arizona Cancer Center (P30CA023074).

ROLE OF THE SPONSOR

None of the funding sources had a role in the design or conduct of the study; in the collection, analysis, and interpretation of the data; or in the preparation, review, or approval of the manuscript.

REFERENCES

- Venepalli NK, Qamruzzaman Y, Li J, et al. Identifying clinically disruptive ICD-10-CM conversions to mitigate financial costs using an online tool. *J Oncol Pract*. 2014;10(2):97–103.
- Caskey R, Zaman J, Nam H, et al. The transition to ICD-10-CM: potential challenges for pediatric practice. *Pediatrics* 2014;134(1):31–36.
- Boyd AD, Yang YM, Li J, et al. Challenges and remediation for Patient Safety Indicators in the transition to ICD-10-CM. *J Am Med Inform Assoc*. 2015;22(1):19–28.
- Boyd AD, Li JJ, Burton MD, et al. The discriminatory cost of ICD-10-CM transition between clinical specialties: metrics, case study and mitigating tools. *J Am Med Inform Assoc*. 2013;20:708–717.
- Boyd AD, Lussier YA. Mitigating Tools for ICD-10-CM Transition Through Convolution and Entanglement. Healthcare Information and Management Systems Society Annual Meeting; February 26, 2014. Orlando, FL.
- CMS. Overview ICD-10. 2011. <https://www.cms.gov/ICD10/>. Accessed February 29, 2012.
- CMS. 2014 ICD-10-CM and GEMs. 2014. <https://www.cms.gov/Medicare/Coding/ICD10/2014-ICD-10-CM-and-GEMs.html>. Accessed March 1, 2014.
- HIMSS G7 ADVISORY REPORT. ICD-10 Transformation: Five Critical Risk-Mitigation Strategies. 2011. http://www.himss.org/files/HIMSSorg/content/files/icd10/G7AdvisoryReport_ICD10%20Version12.pdf. Accessed April 11, 2014.
- AHIMA. ICD-10 Training that's right for you. 2014. <http://www.ahima.org/education/onlineed/Programs/ICD10>. Accessed April 11, 2014.
- Riordan R. PS3-11: Beyond General Equivalency Mappings (GEMs): Understanding the Implications of ICD 10 in Research. *Clin Med Res*. 2013;11:177.
- Ross-Davis SV. Preparing for ICD-10-CM/PCS: one payer's experience with general equivalence mappings (GEMs). *Perspect Health Inf Manag*. 2012;9:1e.
- Butler R. The ICD-10 General Equivalence Mappings. Bridging the translation gap from ICD-9. *J AHIMA*. 2007;78: 84–85.
- Venepalli NK, Qamruzzaman Y, Li J, et al. Identifying clinically disruptive international classification of diseases 10th revision clinical modification conversions to mitigate financial costs using an online tool. *J Oncol Pract*. 2014;10(2):97–103.
- Caskey R, Zaman JS, Chae S, et al. The Transition to ICD-10-CM: Potential Challenges for Pediatric Practice. *Pediatrics* 2014, published online June 2, 2014 doi: 10.1542/peds.2013-4147

AUTHOR AFFILIATIONS

¹Department of Biomedical and Health Information Sciences, University of Illinois at Chicago, Chicago, IL, USA

²University of Illinois Hospital and Health Science System, Chicago, IL, USA

³Department of Medicine, University of Arizona, Tucson, AZ, USA

⁴The University of Arizona Health Sciences Center, Tucson, AZ, USA

⁵Department of Medicine, University of Illinois at Chicago, Chicago, IL, USA

⁶Biomedical Informatics Service Group, Arizona Health Science Center, University of Arizona, Tucson, AZ, USA

⁷Department of Pharmaceutical Sciences, University of Illinois at Chicago, Chicago, IL, USA

⁸Department of Bioengineering, University of Illinois at Chicago, Chicago, IL, USA

⁹BIO5 Institute, University of Arizona, Tucson, AZ, USA

¹⁰University of Arizona Cancer Center, Tucson, AZ, USA

[§]The work was completed in part at The University of Illinois