The Evaluation of Clinical Classifications Software Using the National Inpatient Sample Database

Mahsa Salsabili, PharmD[1], Sebastien Kiogou, MS[2], Terrence J. Adam, RPh, MD, PhD[1,2]
[1]University of Minnesota College of Pharmacy, Minneapolis, MN, US; [2]University of Minnesota Institute for Health Informatics, Minneapolis, MN, US

Abstract

The Clinical Classifications Software (CCS), by grouping International Classification of Diseases (ICD), provides the capacity to better account for clinical conditions for payers, policy makers, and researchers to analyze outcomes, costs, and utilization. There is a critical need for additional research on application of CCS categories to validate the clinical condition representation and to prevent gaps in research. This study compared the event frequency and ICD codes of CCS categories with significant changes from the first three quarters of 2015 to 2016 using National Inpatient Sample data. A total of 63 of the 285 diagnostics CCS were identified with greater than 20% change, of which 32 had increased and 31 decreased over time. Due to the complexity associated with the transition from ICD-9 to ICD-10, more studies are needed to identify the reason for the changes to improve CCS use for ICD-10 and its comparability with ICD-9 based data.

Introduction

Clinical grouping software provides end users with the capacity for dimensional reduction by transforming voluminous sets of granular International Classification of Disease (ICD) codes and grouping them into higher level but closely related clinical groups. Various clinical data grouping and risk adjustment software tools are commercially produced, however many packages have substantial costs associated with licensing and use.[1] Grouper tools can help with clinical outcome assessment and research cohort development and evaluation. [2-4] One non-commercial alternative for clinical data grouping is the Clinical Classifications Software (CCS) which is freely available and supported by the Healthcare Cost and Utilization Project (HCUP) which is a Federal-State-Industry partnership sponsored by the Agency for Healthcare Research and Quality (AHRQ). CCS has been previously used for patient cohort development and comorbidity adjustment in prior clinical studies. [5-8] CCS has also been used by itself or in combination with other approaches for inpatient mortality, risk factor identification, cost and transplant graft loss prediction.[9-12]

Prior published work on the CCS has used clinical encounter data including over 14,000 codes in ICD-9-CM with 3900 procedure codes.[13] In the ICD-10-CM/PCS terminology version there are over 69,800 diagnosis codes and 71,900 procedural codes available. Such large and granular datasets provide a great deal of information with high levels of detail, but their breadth makes modeling difficult due to their high dimensionality. Unlike the CCS coding mappings for ICD-9, the CCS for ICD-10 is still in "beta" version status. [13] On October 1st, 2015, the Health and Human Services mandated nationwide use of ICD-10 instead of ICD-9 for all inpatient medical coding and billing.

General Equivalence Mappings (GEMS) were developed over several years to create a useful, practical code translation reference dictionary for ICD-9-CM and ICD-10-CM/PCS. [14, 15] The initial maps were completed in September 2011 with the accuracy of the mappings verified by reviewing a 20 percent sample of the CM and PCS files. In 2013, a reverse mapping validation of all the ICD-10-CM/PCS CCS assignments was conducted to verify the accuracy of the mapping. A credentialed coder team verified the CCS by comparing ICD-9-CM AHRQ classification assignments with the initial (2011) ICD-10-CM/PCS assignments. The GEMs map was applied in reverse to test the reliability of the CCS assignment for both the diagnosis and procedure code sets. [13]
HCUP released a document examining challenges in creating consistently defined groupings that incorporate diagnosis codes from both the ICD-9-CM and ICD-10-CM systems. The analysis used HCUP data from 24 State Inpatient Databases (SID) and 17 State Emergency Department Databases (SEDD) in 2013, 2014, and 2015. The objective was to follow and compare diagnosis volume across the two classification systems. Comparing the frequency of diagnosis categories for inpatient data between the fourth quarter of years 2014 and 2015 they found that of the 262 CCS diagnosis categories, 93 categories (35 percent) increased or decreased by less than 5 percent across the transition period. For 105 categories (40 percent), frequencies changed by 5 to 19 percent. For 39 categories (15 percent), frequencies changed by 20 to 49 percent. For 25 categories (10 percent), frequencies

changed by greater than 50 percent.[16] There are currently no published studies looking at the 2016 beta CCS version and comparing the diagnosis incidence to previous years.

The CCS provides mappings of the ICD codes to a set of 285 mutually exclusive categories using the single level coding for diagnoses and 231 mutually exclusive categories for procedures. The single level coding are of primary interest since they can facilitate risk adjustment and diagnostic ranking work for research or software applications.[13] The HCUP CCS website indicates there are no publications on the CCS beta version making this an important area of research work. In this study we aim to compare the frequency of clinical events mapping to the CCS categories from 2015 to 2016 and further investigate the potential reasons for changing event frequency by looking at the changes in assignment and distribution of diagnosis codes across the October 1, 2015 ICD-9 to ICD-10 conversion.

Methods

The CCS was analyzed using 2015 and 2016 data from the National Inpatient Sample (NIS), one of the largest publicly available all-payer health care databases in the USA. NIS data provides national estimates of hospital inpatient stays and contains data from more than seven million hospital stays each year when unweighted and is equivalent to an approximately 20 percent stratified sample of inpatient discharges from non-federal academic, acute care, and community hospitals. It has more than 35 million hospitalizations nationally when weighted and represents more than 97 percent of the USA population. The data set consist of de-identified data and therefore was determined to be exempt from review by the University of Minnesota Institutional Review Board. In this descriptive study, the first three quarters of 2015 and the first three quarters of 2016 were used to provide seasonally matching data sets that included ICD9-CM clinical data for the 2015 data set and ICD10-CM/PCS for the 2016 data set. Three quarters were used since the conversion to ICD-10 occurred for the last quarter of 2015 (starting October 1, 2015). The 2015 data provided the baseline incidence rates for each of the CCS categories and a cutoff of 20% change from 2015 to 2016 was considered significant in the evaluation. All diagnoses and the corresponding CCS for each individual were included in the study. The cumulative level captured all CCS and associated diagnosis; therefore one individual may have a specific CCS show up multiple times for their admission depending on the associated diagnoses. While the unique level only captured a specific CCS once per individual during their admission.

Beta Version

The beta version of CCS for ICD-10-CM/PCS was downloaded from HCUP website using the updated version for fiscal year 2016. This version is valid for ICD-10-CM/PCS codes through September 2016 which encompasses the three quarters included in the study. The single-level diagnosis classification aggregates conditions into 285 mutually exclusive categories, most of which are clinically homogeneous. The beta version is meant to translate the CCS system to ICD-10-CM/PCS without changing CCS assignments for diseases and conditions. However, because of the greater overall structure and granularity detail of ICD-10-CM there are some ICD-9-CM conditions that do not map to the same ICD-10 CCS. Some of the ICD-9-CM codes may map to multiple ICD-10-CM codes, some of which some may have a closer match to the ICD-9-CM code/description compared to other codes.

Code Structure

Comparing ICD-9 to ICD-10 diagnosis code set structures, ICD-9 is 3-5 characters in length and the first digit may be alpha or numeric, while ICD-10 is 3-7 characters in length, the first digit is alpha, is very specific and has laterality (codes identifying right or left). Due to the greater number of characters present in ICD-10, it can have a better capacity to identify disease etiology, anatomic site, and severity. Additionally, ICD-10 diagnosis coding allows for the use of combination codes, which is a single code that can be used to classify a) two diagnoses, or b) a diagnosis with an associated secondary process or a diagnosis with an associated complication. This feature allows the reporting of a single code which provides multiple elements of the diagnosis. Overall the ICD-10 code sets provide increased granularity and has expanded concept coverage as compared to ICD-9. Therefore in this study we will compare the number of ICD-9 and ICD-10 codes in 2015 to 2016 in CCS including greater detail on some specific categories to assess the accuracy of translation and to assess for potential coverage gaps. The significance associated with the difference between frequencies of ICD-9 vs ICD-10 codes for specific conditions was assessed using chi-square statistic.

Grouping of CCS categories

Since there are 285 mutually exclusive CCS diagnostic categories, only a subset of the categories were analyzed in detail for this study. The categories were grouped based on presentation of substantial change, clinical significance and notable findings from 2015 to 2016. CCS categories that had either less than 10 events per diagnosis or less than 10 different associated ICD codes, were excluded due to data use agreement restrictions.

ICD Mapping
In order to assure the ICD-9 and ICD-10 diagnosis codes were linked appropriately, the codes were linked from ICD-9 to ICD-10 and reverse linked from ICD-10 to ICD-9. Multiple resources were used to validate the mapping including the WHO ICD-10 version from 2016, the American Medical Association (AMA) ICD-10-CM 2016 The Complete Official Draft Code Set, AMA's ICD-9-CM 2015 Professional Edition for Hospitals, as well as the ICD9Data.com and ICD10Data.com. Additionally web searches identified other coding guides such as payer resources for specific conditions within the studied CCS categories.

Results

CCS Cumulative Events
A total of 63 CCS categories were identified that had a greater than 20% change in first three quarters of 2016 compared to same quarters in 2015. Among those, there were 32 CCS categories with an increase in number of inpatient events and 31 CCS categories had a decrease in events compared to 2015 ICD-9 coding (Table 1, and Table 2). The changes ranged from 1976.84% increase for CCS category other screening for suspected conditions to a 98.95% decrease for rehabilitation care; fitting of prostheses; and adjustment of devices.

Table 1. Frequency of CCS cumulative events with greater than 20% increase in 2016

CCS Category	CCS Category Description	2015 Frequency	2016 Frequency	Percent Change
258	Other screening for suspected conditions (not mental disorders or infectious disease)	5625	116822	1976.84%
661	Substance-related disorders	450242	1158505	157.31%
57	Immunity disorders	13859	27434	97.95%
656	Impulse control disorders NEC	6688	13177	97.02%
209	Other acquired deformities	48599	88292	81.67%
125	Acute bronchitis	47945	82111	71.26%
50	Diabetes mellitus with complications	537093	852571	58.74%
228	Skull and face fractures	27420	42531	55.11%
195	Other complications of birth; puerperium affecting management of mother	419757	632961	50.79%
133	Other lower respiratory disease	288407	411350	42.63%
655	Disorders usually diagnosed in infancy childhood or adolescence	12659	17023	34.47%
243	Poisoning by nonmedicinal substances	6468	8569	32.48%
119	Varicose veins of lower extremity	8309	10976	32.10%
231	Other fractures	111313	146584	31.69%
252	Malaise and fatigue	103977	136840	31.61%
230	Fracture of lower limb	62023	80949	30.51%
64	Other hematologic conditions	24832	32239	29.83%
177	Spontaneous abortion	1487	1927	29.59%
41	Cancer; other and unspecified primary	15005	19439	29.55%
224	Other perinatal conditions	425122	542607	27.64%
236	Open wounds of extremities	38684	49343	27.55%
216	Nervous system congenital anomalies	15006	19060	27.02%
259	Residual codes; unclassified	2575008	3194210	24.05%
203	Osteoarthritis	463026	569764	23.05%
650	Adjustment disorders	26221	32084	22.36%
662	Suicide and intentional self-inflicted injury	107424	131315	22.24%
197	Skin and subcutaneous tissue infections	247848	302458	22.03%
240	Burns	24386	29729	21.91%
229	Fracture of upper limb	51682	62642	21.21%
127	Chronic obstructive pulmonary disease and bronchiectasis	659472	796411	20.76%
82	Paralysis	102224	122997	20.32%
20	Cancer; other respiratory and intrathoracic	1149	1379	20.02%

Table 2. Frequency of CCS cumulative events with greater than 20% decrease in 2016

CCS Category	CCS Category Description	2015 Frequency	2016 Frequency	Percent Change
227	Spinal cord injury	10818	8596	-20.54%
219	Short gestation; low birth weight; and fetal growth retardation	130673	103479	-20.81%
10	Immunizations and screening for infectious disease	691105	547248	-20.82%
123	Influenza	33054	25778	-22.01%
181	Other complications of pregnancy	493182	373644	-24.24%
104	Other and ill-defined heart disease	59425	44957	-24.35%
151	Other liver diseases	413261	312237	-24.45%
199	Chronic ulcer of skin	306884	231639	-24.52%
28	Cancer of other female genital organs	4590	3461	-24.60%
171	Menstrual disorders	31106	22882	-26.44%
148	Peritonitis and intestinal abscess	39024	28220	-27.69%
248	Gangrene	23866	17134	-28.21%
654	Developmental disorders	55182	38666	-29.93%
94	Other ear and sense organ disorders	129705	89243	-31.20%
111	Other and ill-defined cerebrovascular disease	37036	25033	-32.41%
178	Induced abortion	573	379	-33.86%
43	Malignant neoplasm without specification of site	16036	10374	-35.31%
242	Poisoning by other medications and drugs	41626	26675	-35.92%
108	Congestive heart failure; nonhypertensive	1133220	723014	-36.20%
244	Other injuries and conditions due to external causes	323080	198787	-38.47%
2	Septicemia (except in labor)	704639	433236	-38.52%
179	Postabortion complications	956	553	-42.15%
663	Screening and history of mental health and substance abuse codes	1463073	806611	-44.87%
87	Retinal detachments; defects; vascular occlusion; and retinopathy	76807	39003	-49.22%
204	Other non-traumatic joint disorders	225981	114554	-49.31%
670	Miscellaneous mental health disorders	67680	33361	-50.71%
256	Medical examination/evaluation	104855	50746	-51.60%
241	Poisoning by psychotropic agents	25627	9989	-61.02%
220	Intrauterine hypoxia and birth asphyxia	4402	1317	-70.08%
156	Nephritis; nephrosis; renal sclerosis	66342	17544	-73.56%
254	Rehabilitation care; fitting of prostheses; and adjustment of devices	64101	675	-98.95%

Selected CCS unique events

In order to further assess the CCS categories, 11 categories were selected with clinical event frequency change of greater than 50% (Table 1, Table 2) in 2016 and/or had clinical significance. After removal of duplicate CCS categories at the individual level, the frequency of the selected CCS events in 2015 were compared to 2016 (Table 3). Overall the trends remained the same for the unique frequencies compared to cumulative frequencies. The frequencies were very close to the cumulative frequency for certain CCS categories: cancer; other and unspecified primary (41), immunity disorders (57), rehabilitation care; fitting of prostheses; and adjustment of devices (254), and medical examination/evaluation (256). There was a greater than 20% decrease for diabetes mellitus with complications (50) and other acquired deformities (209), while there was greater than 20% increase for substance-related disorders (661) when comparing unique frequencies to cumulative frequencies. Since with some conditions the CCS category was coded more frequently at the unique individual level in one year versus the other, this may indicate some changes in underlying coding processes/patterns. The unique frequency was very close when comparing 2015 to 2016 congestive heart failure (108); nonhypertensive, while for the cumulative frequency there was 36.2% decrease for the same category.

Table 3. Frequency of unique CCS events for select CCS categories

CCS Category	CCS Description	Unique Event Frequency		Percent Change
		2015	2016	
Increased in 2016				
258	Other screening for suspected conditions	5356	113719	2023.21%
661	Substance-related disorders	316859	881091	178.07%
57	Immunity disorders	13401	26167	95.26%
209	Other acquired deformities	45743	78957	72.61%
Decreased in 2016				
254	Rehabilitation care; fitting of prostheses; and adjustment of devices	63163	649	-98.97%
156	Nephritis; nephrosis; renal sclerosis	63750	16939	-73.43%
241	Poisoning by pyschotropic agents	21036	9137	-56.56%
256	Medical examination/evaluation	100242	48779	-51.34%
Others of Clinical Interest				
41	Cancer; other and unspecified primary	14739	18964	28.67%
50	Diabetes mellitus with complications	415964	553641	33.10%
108	Congestive heart failure; nonhypertensive	700206	707709	1.07%

Selected CCS ICD codes comparisons

There was overall increase in the number of ICD-10s compared to ICD-9s, for substance related disorder (CCS 661) and other acquired deformities (CCS 209) there were a greater than 3 fold increase in the number of ICDs within the specific CCS category (Table 4). While the coded ICDs for medical examination/evaluation remained the same and there was even a drop in ICDs for congestive heart failure from 16 ICD-9s to 15 ICD-10s.

Table 4. Unique ICD codes for selected CCS categories in 2015 and 2016

CCS Category	CCS Category Description	Number of Unique ICD Codes		Percent Change
		2015	2016	
Increased in 2016				
258	Other screening for suspected conditions	64	150	134.38%
661	Substance-related disorders	100	389	289.00%
57	Immunity disorders	16	47	193.75%
209	Other acquired deformities	64	226	253.13%
Decreased in 2016				
254	Rehabilitation care; fitting of prostheses; and adjustment of devices	15	18	20.00%
156	Nephritis; nephrosis; renal sclerosis	29	65	124.14%
241	Poisoning by pyschotropic agents	20	56	180.00%
256	Medical examination/evaluation	30	30	0.00%
Others of Clinical Interest				
41	Cancer; other and unspecified primary	72	97	34.72%
50	Diabetes mellitus with complications	57	168	194.74%
108	Congestive heart failure; nonhypertensive	16	15	-6.25%

Mapping CCS 108: Congestive heart failure; nonhypertensive

The ICDs associated with congestive heart failure; nonhypertensive were linked to further understand the reason for the cumulative decrease in frequency from 2015 to 2016(Table 5). There was a substantial drop in the frequency of ICD-10 linked to congestive heart failure, unspecified, with a 645,508 difference in number of times it was coded. In ICD-10, the term “congestive” is considered a non-essential and therefore there is no code for “congestive” heart failure; the term is included in code I50.9 - Unspecified heart failure. For systolic and/or diastolic heart failure, “congestive” is included in the code(s) I50.2 Systolic (congestive) heart failure, I50.3 Diastolic (congestive) heart failure or I50.4 Combined systolic (congestive) and Diastolic (congestive) heart failure. There is an increase in documentation of ICD-10 I50.4 while there is a mix of increase and decrease for specific codes within I50.2 and I50.3 when compared to the equivalent ICD-9 codes. The ICD-9 428.9 diagnosis code is also associated with I50.9. Due to the fact that the term “congestive” was not present for many of the codes in ICD-9 each individual may have been documented for 428.0 in addition to the specific heart failure diagnosis.

Table 5. Mapping of CCS category 108: Congestive heart failure; nonhypertensive and associated frequencies

ICD-9 Diagnosis	Description	ICD-10 Diagnosis	Event Frequency		Difference *
			ICD-9	ICD-10	
398.91	Rheumatic heart failure (congestive)	I09.81	530	373	-157
428.0	Congestive heart failure, unspecified	I50.9	649578	4070	-645508
428.1	Left heart failure	I50.1	1100	13759	12659
428.20	Systolic heart failure, unspecified	I50.20	14059	26520	12461
428.21	Acute systolic heart failure	I50.21	25796	75039	49243
428.22	Chronic systolic heart failure	I50.22	65490	87940	22450
428.23	Acute on chronic systolic heart failure	I50.23	84129	27822	-56307
428.30	Diastolic heart failure, unspecified	I50.30	28797	26249	-2548
428.31	Acute diastolic heart failure	I50.31	23989	102654	78665
428.32	Chronic diastolic heart failure	I50.32	87108	96972	9864
428.33	Acute on chronic diastolic heart failure	I50.33	85248	3230	-82018
428.40	Combined systolic and diastolic heart failure, unspecified	I50.40	3283	6556	3273
428.41	Acute combined systolic and diastolic heart failure	I50.41	5528	22505	16977
428.42	Chronic combined systolic and diastolic heart failure	I50.42	17499	43158	25659
428.43	Acute on chronic combined systolic and diastolic heart failure	I50.43	36314	186167	149853
428.9	Heart failure, unspecified	I50.9	4772	4070	-702

* The difference was found to be significant ($P < 0.001$) for all diagnoses except Rheumatic heart failure (congestive) at $P = 0.17$

Mapping CCS 50: Diabetes mellitus with complications

The top 10 diagnoses for diabetes mellitus with complications were identified for both ICD-9 and ICD-10. There was 194.74% increase in overall number of coded clinical events in 2016 vs 2015 (Table 4). The top 10 most frequent ICD-9 codes matched with 8 out of the 10 ICD-10 codes (Table 6, Table 7), with codes E11.42-Type 2 diabetes mellitus with diabetic polyneuropathy being only a partial match and E11.8- Type 2 diabetes mellitus with unspecified complications not being in the top 10 for only ICD-10 coding. Several of the ICD-9 codes transitioned into a combination for ICD-10 coding such as diabetes with neurological manifestations, type II or unspecified type, uncontrolled and diabetes with ketoacidosis, type I [juvenile type], uncontrolled (Table 6). There is an overall significant increase in frequency of codes being utilized except for diabetes with ophthalmic manifestations, type II or unspecified type, not stated as uncontrolled and diabetes with peripheral circulatory disorders, type II or unspecified type, not stated as uncontrolled when comparing ICD-10 to ICD-9. Since 2 of the ICD-9 codes represented more than 2 separate codes in ICD-10, these were further analyzed (Table 7). Among the ICD-10 codes that may potentially be a match for 250.80: Diabetes with other specified manifestations, type II or unspecified type, not stated as uncontrolled, the ICD-10 code E11.628: Type 2 diabetes mellitus with hypoglycemia without coma was the most frequently used apart from E11.65: Type 2 diabetes mellitus with hyperglycemia. For 250.50: Diabetes with ophthalmic manifestations, type II or unspecified type, not stated as uncontrolled, E11.319 was the most common representative in ICD-10 coding.

Table 6. Mapping CCS 50: Diabetes mellitus with complications, top 10 diagnoses

ICD-9 Diagnosis	Description	ICD-10 Diagnosis	Frequency		Difference*
			ICD-9	ICD-10	
250.60	Diabetes with neurological manifestations, type II or unspecified type, not stated as uncontrolled	E11.40	95622	103366	7744
250.02	Diabetes mellitus without mention of complication, type II or unspecified type, uncontrolled	E11.65	92164	246795	154631
250.40	Diabetes with renal manifestations, type II or unspecified type, not stated as uncontrolled	E11.29 or E11.22	59962	111678	51716
250.80	Diabetes with other specified manifestations, type II or unspecified type, not stated as uncontrolled	Multiple (see Table 7)	59905	95280	35375
250.62	Diabetes with neurological manifestations, type II or unspecified type, uncontrolled	E11.40 with E11.65	37323	350161	312838
250.13	Diabetes with ketoacidosis, type I [juvenile type], uncontrolled	E10.10 with E10.65	27221	37905	10684
250.50	Diabetes with ophthalmic manifestations, type II or unspecified type, not stated as uncontrolled	Multiple (see Table 7)	22696	22543	-153
250.82	Diabetes with other specified manifestations, type II or unspecified type, uncontrolled	E11.65 with E11.69	22546	264491	241945
250.42	Diabetes with renal manifestations, type II or unspecified type, uncontrolled	E11.21 with E11.65	18956	300369	281413
250.70	Diabetes with peripheral circulatory disorders, type II or unspecified type, not stated as uncontrolled	E11.51	15338	14520	-818

* The difference was found to be significant ($P < 0.001$) for all diagnoses

Table 7. Mapping CCS 50: Diabetes mellitus with complications, of ICD-9s with multiple ICD-10 equivalents

			Frequency of Events	
ICD-9	**ICD-10 Equivalent**	**ICD-10 Description**	**ICD-9**	**ICD-10**
250.80	E11.618	Type 2 diabetes mellitus with other diabetic arthropathy		244
250.80	E11.620	Type 2 diabetes mellitus with diabetic dermatitis		161
250.80	E11.621	Type 2 diabetes mellitus with foot ulcer		28500
250.80	E11.622	Type 2 diabetes mellitus with other skin ulcer		3607
250.80	E11.628	Type 2 diabetes mellitus with other skin complications		4305
250.80	E11.649	Type 2 diabetes mellitus with hypoglycemia without coma		40767
250.80	E11.69	Type 2 diabetes mellitus with other specified complication		17696
250.80	E11.65	Type 2 diabetes mellitus with hyperglycemia		114329
Total			59905	209609
250.50	E11.311	Type 2 diabetes mellitus with unspecified diabetic retinopathy with macular edema		455
250.50	E11.319	Type 2 diabetes mellitus with unspecified diabetic retinopathy without macular edema		20753
250.50	E11.36	Type 2 diabetes mellitus with diabetic cataract		360
250.50	E11.39	Type 2 diabetes mellitus with other diabetic ophthalmic complication		975
Total			22696	22543

Discussion

The use of the CCS clinical grouper makes it easier to understand patterns of diagnoses and procedures more efficiently so various organizations such as payers, policy makers, and researchers can analyze outcomes, costs, and utilization associated with particular illnesses and conditions. Currently, there is limited research looking at the impact of the transition from ICD-9 to ICD-10 within the CCS categories and the current version of CCS is listed as a "beta" version.[13,16] The goal of our study is to assess the assignment and distribution of ICDs and their associated clinical events within CCS categories, especially for longitudinal data spanning the October 1, 2015 conversion.

The change in clinical coding from ICD-9 to ICD-10 created a number of issues with previously established clinical software and research methods previously used in the USA. Code conversion involves identifying clinically equivalent codes which exist in both terminology systems, however, the larger number of codes with greater granularity provides greater knowledge representation capacity with ICD-10. For the majority of the CCS categories, there has been an increase in the number of ICD codes, for some CCS categories there were at least three times more ICD-10s compared to ICD-9s. This may explain the increase in frequency for certain CCS categories despite the fact that the actual incidence of these conditions would have not increased so dramatically from one year to another. Also with ICD-10, one code may overlap and represent multiple conditions vs the multiple separate codes which would be needed in ICD-9 which may partly explain the change in ICD codes per category, but should be controlled in part by careful CCS category mappings to prevent such significant changes in frequencies. For conditions such as rehabilitation care, fitting of prostheses, and adjustment of devices; nephritis, nephrosis, renal sclerosis; and poisoning by pyschotropic agents there was a decrease in frequency despite an increase in ICD-10 codes. The reason behind the decreases needs to be further investigated to better understand the transition within these CCS categories as well as others with similar trends.

Interestingly there were conditions that had a decrease in unique ICD-10 code count, as shown in the results for CCS categories such as congestive heart failure; nonhypertensive. When looking at the prevalence of the most common cause of congestive heart failure, coronary heart disease, there was an increase from 5.6 percent in 2015 to 5.7 percent in 2016. [17, 18] As previously described, the major difference in the transition was that the terminology applied, specifically the term "congestive", created major differences in how often specific diagnosis codes were documented. Also in this case if we had only looked at the clinical event frequency of the unique CCS events, this category would not have been included in the study due to a less than 20 percent change from 2015 to 2016. Therefore it is recommended for researchers to not only look at cumulative frequency of events but also unique event rates in CCS categories when performing studies to capture true changes in incidences of conditions. Another finding within this category was that acute on chronic systolic heart failure, and acute on chronic systolic heart failure were less frequently coded in ICD-10 vs ICD-9 while there was a significant increase in acute and chronic combined diastolic and systolic heart failure. This is a case where ICD-10 coding may have become less granular therefore researchers should be aware when performing studies with this CCS category.

The frequency of the CCS category associated with diabetes had a greater than 50 percent increase comparing 2015 to 2016, even though the prevalence of diabetes decreased from 8.9 percent to 8.8 percent respectively.[17,18] Diabetes mellitus with complications showed a great increase in number of codes in ICD-10 vs ICD-9. When looking in further detail several of the ICD-9 diagnoses were linked to a combination of ICD-10 although it is worth noting these combinations are dependent on coder training and knowledge. The combinations may explain the significant increase in frequency of the events in 2016. Additionally some of the codes did not have a clear equivalent therefore more than 2 ICD-10s could be potentially coded that may represent the same diagnosis in ICD-9. Researchers should be aware of the presence of these type of code problems when using this CCS category.

Several of the CCS categories with significant changes were described as "other" forms of a specific condition, indicating there may be some variability in the selection for each CCS in the transition and therefore researchers should be extra cautious when studying these conditions if they decide to use these types of CCS categories.

This increased knowledge representation of ICD-10 may not provide equivalent representation before and after the conversion occurred. Additional work may be needed to adjust the coding groups in the beta version of CCS to insure correct knowledge representation in the CCS software with a particular focus on those categories with the greatest changes in events before and after ICD-10 implementation. In the interim, users may want to consider additional evaluation on these CCS categories when using and await additional validation studies and newer CCS updates. More research is needed comparing additional ICD-10 data and using other datasets.

Conclusion

This descriptive study provides a deeper understanding of the transition of ICDs from the first three quarters of 2015 to 2016. Several CCS categories had a much higher and significantly lower number of clinical events present against a similar denominator of overall events which are not likely due to changes in clinical care but rather a change in the underlying ICD codes and their attribution to clinical categories in the CCS. Part of the changes may be due to issues of incomplete mapping of ICD-9 to ICD-10 coding as well as issue of changing granularity of the ICD codes. These areas of largest clinical event changes likely warrant additional exploration for further adjustments the in the ICD codes associated with the respective CCS categories for ICD-10 application.

References

1. Juhnke C, Bethge S, Mühlbacher AC. A review on methods of risk adjustment and their use in integrated healthcare systems. International journal of integrated care. 2016 Oct;16(4).
2. Goz V, Weinreb JH, McCarthy I, Schwab F, Lafage V, Errico TJ. Perioperative complications and mortality after spinal fusions: analysis of trends and risk factors. Spine, 38(22):1970-6, October 2013. https://www.ncbi.nlm.nih.gov/pubmed/23928714.
3. Adam TJ, Finkelstein SM, Parente ST, Hertz MI. Cost analysis of home monitoring in lung transplant recipients. International journal of technology assessment in health care. 2007 Apr;23(2):216-22.
4. Chi CL, Wang J, Clancy TR, Robinson JG, Tonellato PJ, Adam TJ. Big data cohort extraction to facilitate machine learning to improve statin treatment. Western journal of nursing research. 2017 Jan;39(1):42-62.
5. Adam TJ, Chi CL. Big Data Cohort Extraction for Personalized Statin Treatment and Machine Learning. In Bioinformatics and Drug Discovery 2019 (pp. 255-272). Humana Press, New York, NY.
6. Radley DC, Gottlieb DJ, Fisher ES, Tosteson AN. Comorbidity risk-adjustment strategies are comparable among persons with hip fracture. Journal of clinical epidemiology, 61(6):580-7, June 2008, Epub February 14, 2008. http://www.ncbi.nlm.nih.gov/pubmed/18471662.
7. Thompson DA, Makary MA, Dorman T, Pronovost PJ. Clinical and economic outcomes of hospital acquired pneumonia in intra-abdominal surgery patients. Annals of Surgery, 243(4):547-52, April 2006. http://www.ncbi.nlm.nih.gov/pubmed/16552208.
8. Patil CG, Alexander AL, Hayden Gephart MG, Lad SP, Arrigo RT, Boakye M. A population-based study of inpatient outcomes after operative management of nontraumatic intracerebral hemorrhage in the United States. World Neurosurg, 78(6):640-5, December 2012. https://www.ncbi.nlm.nih.gov/pubmed/22120557.
9. Tabak YP, Sun X, Nunez CM, Johannes RS. Using electronic health record data to develop inpatient mortality predictive model: Acute Laboratory Risk of Mortality Score (ALaRMS). Journal of the American Medical Informatics Association: JAMIA, 21(3):455-463, June 2014. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3994855/pdf/amiajnl-2013-001790.pdf.
10. Robinson JW. Regression tree boosting to adjust health care cost predictions for diagnostic mix. Health services research, 43(2):755-7, April 2008. http://www.ncbi.nlm.nih.gov/pubmed/18370977.
11. Machnicki G, Pinsky B, Takemoto S, Balshaw R, Salvalaggio PR, Buchanan PM, Irish W, Bunnapradist S, Lentine KL, Burroughs TE, Brennan DC, Schnitzler MA. Predictive ability of pretransplant comorbidities to predict long-term graft loss and death. American journal of transplantation, 9(3):494-505, March 2009. https://www.ncbi.nlm.nih.gov/pubmed/19120083.
12. Fogerty MD, Abumrad NN, Nanney L, Arbogast PG, Poulose B, Barbul A. Risk factors for pressure ulcers in acute care hospitals. Wound repair and regeneration, 16(1):11-8, January-February 2008. http://www.ncbi.nlm.nih.gov/pubmed/18211574.
13. Healthcare Cost and Utilization Project (HCUP). Agency for Healthcare Research and Quality R, MD. Beta Clinical Classifications Software (CCS) for ICD-10-CM/PCS. Healthcare Cost and Utilization Project (HCUP) 2019 [cited 2019. Available from: https://www.hcup- us.ahrq.gov/toolssoftware/ccs10/ccs10.jsp.
14. Butler RR. Icd-10 general equivalence mappings: Bridging the translation gap from icd-9. Journal of AHIMA. 2007 Oct;78(9):84-6.
15. Proctor DB, Niedzwiecki B, Pepper J, Madero P. Kinn's The Administrative Medical Assistant E-Book: An Applied Learning Approach. Elsevier Health Sciences; 2016 May 13.

16. Moore BJ, McDermott KW, Elixhauser A. ICD-10-CM Diagnosis Coding in HCUP Data: Comparisons With ICD-9-CM and Precautions for Trend Analyses. ONLINE. November 28, 2017. U.S. Agency for Healthcare Research and Quality. Available at https://www.hcup-us.ahrq.gov/datainnovations/icd10_resources.jsp.
17. US Department of Health and Human Services, Centers for Disease Control and Prevention. Summary Health Statistics: National Health Interview Survey, 2015. National Center for Health Statistics: MD. 2015.
18. US Department of Health and Human Services, Centers for Disease Control and Prevention. Summary Health Statistics: National Health Interview Survey, 2016. National Center for Health Statistics: MD. 2016.