



H·CUP
HEALTHCARE COST AND UTILIZATION PROJECT

USER GUIDE:

**CLINICAL CLASSIFICATIONS SOFTWARE REFINED
(CCSR)**

VERSION 2019.1

Issued September 2019

Agency for Healthcare Research and Quality
Healthcare Cost and Utilization Project (HCUP)

Phone: (866) 290-HCUP (4287)

Email: hcup@ahrq.gov

Website: www.hcup-us.ahrq.gov

TABLE OF CONTENTS

Introduction	1
Comparison of the CCSR for ICD-10-CM Diagnoses, the Beta Versions of the CCS for ICD-10-CM, and the CCS for ICD-9-CM	2
Description of the CCSR for ICD-10-CM	3
The Structure of the CCSR	3
CCSR Guidelines	4
General Assignment Guidelines	4
Body System-Specific Structure Notes and Exceptions to General Guidelines	6
One-to-One and One-to-Many Mapping of ICD-10-CM Diagnosis Codes to CCSR	8
Why Cross-Classification of Codes are Necessary in the CCSR	8
Prevalence of Cross-Classified Codes in the CCSR	9
Impact on Users	9
Capitalizing on ICD-10-CM Diagnosis Coding Specificity	10
Episode of Care Coding in ICD-10-CM Diagnoses	12
Example of Episode of Care CCSR Categories	13
Trimester Coding in ICD-10-CM Diagnoses	14
Impact of New Clinical CCSR Categories on Users	14
Using the Downloadable CCSR Files	15
System Requirements	15
Downloadable Files	15
Flexible File Structure for Outputting the CCSR for ICD-10-CM Diagnoses	16
Output Option 1, Vertical Output File	17
Output Option 2, Horizontal Array of all CCSR Categories	17
Example of Output from Option 1, Vertical Output File, and Option 2, Horizontal Array of All CCSR Categories	18
Representation of ICD-10-CM Diagnosis Codes	19
Running the SAS Program to Add CCSR Categories to Data	19
Appendix A: Background on the Development of the CCSR	22
Impact on the Clinical Classifications Software	22
Refinement of the Beta Versions of the CCS	22
Refinement Review Process	23
Appendix B: CCSR Categories	25

INDEX OF TABLES AND FIGURE

Table 1. Differences Between the CCSR for ICD-10-CM, CCS for ICD-10-CM, and CCS for ICD-9-CM.....	2
Table 2. Three-Character Abbreviation for ICD-10-CM Body Systems.....	3
Figure 1. Example of One-to-One and One-to-Many CCSR Mappings for Hypertension and Heart Failure	8
Table 3. Type of CCSR Category Assignment by Number of ICD-10-CM Codes, Version 2019.1	9
Table 4. Number of CCSR and CCS (Beta Versions) Categories by Body System	10
Table 5. Example of CCSR Categories for Episode of Care.....	13
Table 6. Example of CCSR Categories by Pregnancy Trimester.....	14
Table 7. Contents of the CCSR for ICD-10-CM Diagnoses Zip File.....	16
Table 8. Sample of Input File Record.....	18
Table 9. Example of the Vertical Output File for the Sample Record.....	18
Table 10. Example of Horizontal Output File.....	19
Table 11. Example of Representation of ICD-10-CM Diagnosis Codes in the CCSR for ICD-10-CM.....	19

ACKNOWLEDGEMENTS

This work was funded by the Agency for Healthcare Research and Quality (AHRQ) under contract HHSA-290-2018-00001-C. AHRQ gratefully acknowledges the contributions of American Health Information Management Association (AHIMA)-certified ICD-10-CM/PCS trainers at the Ohio State University; clinical experts from the University of California, Davis, and the RAND Corporation, the National Institutes of Health, and the technical team at IBM Watson Health. The Healthcare Cost and Utilization Project (HCUP) is a family of health care databases and related software tools and products developed through a Federal-State-Industry partnership and sponsored by AHRQ. HCUP would not be possible without the contributions of the following data collection Partners from across the United States:

Alaska Department of Health and Social Services

Alaska State Hospital and Nursing Home Association

Arizona Department of Health Services

Arkansas Department of Health

California Office of Statewide Health Planning and Development

Colorado Hospital Association

Connecticut Hospital Association

Delaware Division of Public Health

District of Columbia Hospital Association

Florida Agency for Health Care Administration

Georgia Hospital Association

Hawaii Lauima Data Alliance

Illinois Department of Public Health

Indiana Hospital Association

Iowa Hospital Association

Kansas Hospital Association

Kentucky Cabinet for Health and Family Services

Louisiana Department of Health

Maine Health Data Organization

Maryland Health Services Cost Review Commission

Massachusetts Center for Health Information and Analysis

Michigan Health & Hospital Association

Minnesota Hospital Association (provides data for Minnesota and North Dakota)

Mississippi State Department of Health

Missouri Hospital Industry Data Institute

Montana Hospital Association

Nebraska Hospital Association

Nevada Department of Health and Human Services

New Hampshire Department of Health & Human Services

New Jersey Department of Health

New Mexico Department of Health

New York State Department of Health

North Carolina Department of Health and Human Services

North Dakota (data provided by the Minnesota Hospital Association)

Ohio Hospital Association

Oklahoma State Department of Health

Oregon Association of Hospitals and Health Systems

Oregon Office of Health Analytics

Pennsylvania Health Care Cost Containment Council

Rhode Island Department of Health

South Carolina Revenue and Fiscal Affairs Office

South Dakota Association of Healthcare Organizations

Tennessee Hospital Association

Texas Department of State Health Services

Utah Department of Health

Vermont Association of Hospitals and Health Systems

Virginia Health Information

Washington State Department of Health

West Virginia Department of Health and Human Resources, West Virginia Health Care Authority

Wisconsin Department of Health Services

Wyoming Hospital Association

INTRODUCTION

This report provides technical documentation for the Healthcare Cost and Utilization Project (HCUP) Clinical Classifications Software Refined (CCSR) for International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM)-coded diagnoses. Starting in October 1, 2015, diagnoses for hospital inpatient stays and outpatient encounters in the United States are reported using the ICD-10-CM coding system. ICD-10-CM consists of more than 70,000 diagnosis codes. Although it is possible to present descriptive statistics for individual ICD-10-CM diagnosis codes, it is often more useful to aggregate codes into clinically meaningful categories. The CCSR for ICD-10-CM diagnoses provides one method for aggregating codes.

Similar to the original Clinical Classifications Software (CCS) for ICD-9-CM-coded diagnoses, the CCSR for ICD-10-CM:

- Classifies diagnoses into clinically meaningful categories
- Provides a means by which to identify specific clinical conditions using diagnosis codes
- Can be used analytically to examine patterns in healthcare cost, utilization, and outcomes; perform rank utilization by diagnoses; and risk-adjust by clinical condition.

The CCS for ICD-9-CM diagnoses was translated to ICD-10-CM prior to the availability of ICD-10-CM-coded data and released as a beta version. Once ICD-10-CM coded data became available, the beta versions of the CCS was evaluated using the HCUP databases and unexpected discontinuities between the ICD-9-CM and ICD-10-CM versions of the CCS were revealed. In addition, there was interest in taking advantage of the specificity built into ICD-10-CM diagnosis codes.

These findings led to the development of the CCSR for ICD-10-CM diagnoses. The CCSR balances the retention of the clinical concepts included in the CCS categories under ICD-9-CM and capitalizes on the specificity of ICD-10-CM diagnoses by creating new clinical categories. In addition, the CCSR allows ICD-10-CM diagnosis codes to be cross classified into more than one category because individual codes can be used to document multiple conditions or a condition and a common symptom/manifestation. The refinement process was informed by American Health Information Management Association-certified ICD-10-CM trainers and reviewed by a team of clinical experts. The team extensively reviewed the CCSR at each stage of its development using the HCUP State databases for quality control testing. Additional information on the refinement process is available in [Appendix A: Background on the Development of the CCSR](#).

The Version 2019.1 release of the CCSR for ICD-10-CM diagnoses is based on ICD-10-CM diagnosis codes valid from October 1, 2015, through September 30, 2019.¹ The CCSR for ICD-10-CM is updated annually to coincide with fiscal year updates to the ICD-10-CM diagnosis coding system. Files containing the translation of ICD-10-CM diagnosis codes into CCSR for ICD-10-CM categories can be downloaded from the [HCUP User Support \(HCUP-US\)](#) website.²

¹ Version 2020.1 release of the CCSR for ICD-10-CM will be updated for ICD-10-CM diagnosis codes valid through September 30, 2020. These updates are expected to be released in late Fall of 2020.

² The HCUP User Support website can be found at www.hcup-us.ahrq.gov/.

COMPARISON OF THE CCSR FOR ICD-10-CM DIAGNOSES, THE BETA VERSIONS OF THE CCS FOR ICD-10-CM, AND THE CCS FOR ICD-9-CM

There are several key differences to note across the HCUP ICD-based diagnosis classification software: the CCSR for ICD-10-CM diagnoses, the beta versions of the CCS for ICD-10-CM diagnoses, and the CCS for ICD-9-CM diagnoses. Table 1 summarizes these differences. Additional details are provided throughout this User Guide and on the [HCUP-US website](#).

Table 1. Differences Between the CCSR for ICD-10-CM, CCS for ICD-10-CM, and CCS for ICD-9-CM

CCSR for ICD-10-CM Diagnoses	Beta Versions of the CCS for ICD-10-CM Diagnoses	CCS for ICD-9-CM Diagnoses
More than 70,000 ICD-10-CM codes are categorized into 538 categories	More than 70,000 ICD-10-CM codes are categorized into 283 categories	More than 14,000 ICD-9-CM codes are categorized into 283 categories
Capitalizes on the specificity of ICD-10-CM coding by creating new clinical categories that did not previously exist	Uses the same categories as previously available under CCS for ICD-9-CM diagnoses	283 CCS diagnosis and external cause of injury categories beginning in fiscal year 2008 after the CCS for Mental Health and Substance Use categories were permanently integrated into the CCS tool
Categories are organized into 21 body systems, generally following the ICD-10-CM diagnosis codebook chapters	Uses the same categories and order as previously available under CCS for ICD-9-CM diagnoses	Categories are ordered to generally follow the ICD-9-CM codebook across 17 chapters plus two supplementary classifications
Each ICD-10-CM code maps to one or more CCSR categories	Each ICD-10-CM code maps to one CCS category	Each ICD-9-CM code maps to one CCS category
Codes mapped to multiple CCSR categories do not have an embedded hierarchy for easily ranking principal diagnosis categories	Mutually exclusive categories facilitate ranking principal diagnoses categories	Mutually exclusive categories facilitate ranking principal diagnoses categories
No multi-level system with additional diagnostic specificity has been developed	Multi-level system with additional diagnostic specificity available for up to two levels	Multi-level system with additional diagnostic specificity available for up to four levels
SAS® programming code available. Comma separated values (CSV) mapping file available for use with other programming languages	SAS programming code available. CSV mapping file available for use with other programming languages	SAS and Stata® programming code available. CSV mapping file available for use with other programming languages
Allows users the flexibility to choose between output files structured horizontally or vertically. The vertical file structure improves storage efficiency.	Horizontal output only, one CCS data element for every diagnosis code.	Horizontal output only, one CCS data element for every diagnosis code.

Abbreviations: CCS, Clinical Classifications Software; CCSR, Clinical Classification Software Refined; CSV, comma separated values; ICD-9-CM, International Classification of Diseases, Ninth Revision, Clinical Modification; ICD-10-CM, International Classification of Diseases, Tenth Revision, Clinical Modification.

DESCRIPTION OF THE CCSR FOR ICD-10-CM

The Structure of the CCSR

The CCSR for ICD-10-CM diagnoses aggregates more than 70,000 ICD-10-CM diagnosis codes into a manageable number of clinically meaningful categories across 21 body systems, which generally follow the structure of the ICD-10-CM diagnosis chapters. Additional information on specific differences between the CCSR body systems and the ICD-10-CM diagnosis chapters is detailed in the next section ([CCSR Guidelines](#)). CCSR categories are organized by body system.³ Each body system is abbreviated using a three-character scheme as shown in Table 2. Individual CCSR categories are numbered sequentially with the numbering scheme starting at “001” within each body system (i.e., there is a CCSR 001 for each body system). A complete listing of all CCSR categories and their associated descriptions can be found in [Appendix B: CCSR Categories](#), available on the [CCSR](#) page of the HCUP User Support website.

Table 2. Three-Character Abbreviation for ICD-10-CM Body Systems³

ICD-10-CM Body System (largely aligned with ICD-10-CM chapters)	Three Character Abbreviation
Certain infectious and parasitic diseases	INF
Neoplasms	NEO
Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism	BLD
Endocrine, nutritional and metabolic diseases	END
Mental, behavioral and neurodevelopmental disorders	MBD
Diseases of the nervous system	NVS
Diseases of the eye and adnexa	EYE
Diseases of the ear and mastoid process	EAR
Diseases of the circulatory system	CIR
Diseases of the respiratory system	RSP
Diseases of the digestive system	DIG
Diseases of the skin and subcutaneous tissue	SKN
Diseases of the musculoskeletal system and connective tissue	MUS
Diseases of the genitourinary system	GEN
Pregnancy, childbirth and the puerperium	PRG
Certain conditions originating in the perinatal period	PNL
Congenital malformations, deformations and chromosomal abnormalities	MAL
Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified	SYM
Injury, poisoning and certain other consequences of external causes	INJ
External causes of morbidity	EXT

³ The term “body system” is used to describe the organization of conditions within the CCSR tool. When referring specifically to the organization of codes within the ICD-10-CM codebook, we use the term “chapter.”

ICD-10-CM Body System (largely aligned with ICD-10-CM chapters)	Three Character Abbreviation
Factors influencing health status and contact with health services	FAC

CCSR Guidelines

During the development of the CCSR, several guidelines were adopted to address common questions that the team encountered. The following sections describe the general guidelines that apply to all body systems as well as guidelines that are specific to an individual body system.

General Assignment Guidelines

1. The overarching goal of the CCSR is to categorize codes into a manageable number of clinically meaningful categories. The categories themselves should capture a clinical concept, and the codes within a specific category should maintain the clinical intent of that category. Although there was an attempt to keep the clinically meaningful categories from previous versions, there was no attempt to match the trends or utilization counts of the CCS for ICD-9-CM or the beta versions of the CCS for ICD-10-CM, even in cases in which the category title remains the same. Instead, the clinical concept guided the assignment of codes.
2. The number of catch-all categories (those that start with “Other specified”) are minimized, because they are of limited value if the ICD-10-CM codes within the category are highly heterogeneous. When a catch-all category is necessary, the number of codes is minimized by separating out clinically relevant groups. This often is a trade-off with the number of overall categories created in the tool.
3. Different CCSR categories may include an overlapping set of codes, but no CCSR should be nested completely within another CCSR.
4. Every ICD-10-CM code is assigned to at least one CCSR category but can be assigned to multiple categories. Individual ICD-10-CM diagnosis codes can be used to document multiple conditions or a condition and a common symptom/manifestation. Assigning only one CCSR category to these codes would require prioritizing the assignment. Clinical coding guidelines were reviewed and considered when CCSR category assignments were made. Each code is mapped to no more than five categories and an attempt was made to minimize the number of assigned categories. In addition, when possible, a code is mapped to no more than one catch-all category.
5. Cross-classified codes (i.e., codes mapped to more than one CCSR category) are not designed to be hierarchical because the preferred hierarchy of conditions might be different depending on the research question or intended application of the tool. There is no implication or intent to rank one condition as more important than another. If a code is assigned to more than one CCSR category, the user must decide how to best use that information in his or her own research study⁴.

⁴ In the future, AHRQ will conduct analyses to create a multi-level system using the CCSR as a foundation. A hierarchical approach will be developed as part of the creation of the multi-level system.

6. Diagnosis codes are mostly retained in the chapters assigned by the World Health Organization⁵ unless the clinical review team had a specific rationale for grouping the code only in another body system. For example, an ICD-10-CM code that is located in the Diseases and Disorders of the Circulatory System chapter of the codebook is assigned to at least one CCSR category in the *Diseases of the circulatory system* body system (i.e., the category begins with the characters “CIR”). There are a few notable exceptions⁶ to this rule:
- ICD-10-CM codes for Periprosthetic fracture around internal prosthetic joint (i.e., codes starting with characters M97) are included in the *Injury, poisoning and certain other consequences of external causes* body system and not in the *Diseases of the musculoskeletal system and connective tissue* body system. This decision was made after consultation with injury experts at the Centers for Disease Control and Prevention (CDC).
 - Several Mental, behavioral, and neurodevelopmental disorders codes (i.e., within the range of ICD-10-CM codes starting with characters F01–F99) related to the clinical concepts of neurocognitive disorders and sleep wake disorders are only mapped to categories in the *Diseases of the nervous system* body system so that they are categorized with similar clinical concepts.
 - Several ICD-10-CM codes from the chapter on Symptoms, Signs and Abnormal Clinical and Laboratory Findings, Not Elsewhere Classified (i.e., ICD-10-CM codes starting with characters R01–R99) are included only under other body system categories because they aligned with specific conditions or disorders. For example, several R codes are exclusively assigned to a CCSR category within the *Diseases of the circulatory* body system. In contrast, codes that do not definitively indicate the presence of a condition or disorder (e.g., codes R78.0–R78.5 Finding of substance in blood) are retained within the body system of *Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified*.
 - Several ICD-10-CM codes from the chapter on Factors Influencing Health Status and Contact with Health Services (i.e., ICD-10-CM codes starting with characters Z01–Z99) are included only under other body system categories because they aligned with specific conditions or disorders. For example, several Z codes are exclusively assigned to a CCSR category with codes under *Pregnancy, childbirth, and puerperium* or within *Certain conditions originating during the perinatal period* body systems. In contrast, codes that do not definitively indicate the presence of a condition or disorder (e.g., code Z79.891 long term (current)

⁵ Please see “The ICD-10-CM Tabular List of Diseases and Injuries” for details on the chapter assignments for ICD-10-CM codes (available at www.cdc.gov/nchs/icd/icd10cm.htm or www.cms.gov/Medicare/Coding/ICD10/).

⁶ There are rare exceptions to this rule not covered by these explanations. Each exception code was identified by the American Health Information Management Association (AHIMA)-certified ICD-10-CM trainers and reviewed by the team of clinical experts before a final decision was made. In each case, the team agreed the best assignment for the ICD-10-CM code in question was only to a CCSR category in a different body system than the codebook chapter.

use of opiate analgesic and Z72.0 tobacco use) are retained within the body system of *Factors influencing health status and contact with health services*.

Body System-Specific Structure Notes and Exceptions to General Guidelines

1. The *Neoplasms* body system follows the National Cancer Institute (NCI)⁷ organization of cancer types by body system and location whenever possible. In addition, the team sought guidance from cancer researchers on beneficial subcategories that go beyond the NCI structure, such as more specific leukemia categories.
2. The *Mental, behavioral, and neurodevelopmental disorder* body system follows the organization of the American Psychiatric Association (APA) Diagnostic and Statistical Manual of Mental Disorders (DSM-5),⁸ unless there was a specific reason to do otherwise. When questions arose regarding assignment of codes and categories, the ICD-11 classification system was consulted. Two exceptions are that the Neurocognitive disorders category (NVS011) and Sleep wake disorders category (NVS016)—which each contained codes from the Mental, Behavioral, and Neurodevelopmental Disorders chapter as well as the Diseases of the Nervous System chapter—were placed under the *Diseases of the nervous system* body system even though they are DSM-5 categories.
3. Codes in the *Congenital malformations, deformations and chromosomal abnormalities*, *Certain conditions originating during the perinatal period*, and *Pregnancy, childbirth, and puerperium* body systems are frequently, but not extensively, cross-listed to other body systems. The clinical review team believed that most researchers would be interested in either this specific set of congenital, perinatal, or pregnancy codes or the more general codes in the specific body system categories, but not both together. The conditions often are clinically and epidemiologically different, with unique treatment plans. There are exceptions to this decision, such as infectious diseases or critical chronic conditions such as hypertension and diabetes, in which users will observe codes cross-listed across these body systems.
4. The *Injury, poisoning and certain other consequences of external causes* and *External causes of morbidity* body systems follow the CDC Injury and External Cause of Morbidity matrix definitions for ICD-10-CM.⁹ Additional categories were created to capture concepts not explicitly included in the CDC injury matrices or codes from other body systems that are not included in CDC definitions.
5. Separate Injury categories were created to maintain the difference between poisoning, adverse effects, and toxic effects with respect to the occurrence of drug toxicity. The guidelines for ICD-10-CM coding¹⁰ define these terms as follows:
 - “Adverse effect” is used when a drug has been properly prescribed and properly administered
 - “Poisoning” involves the improper use of medication

⁷ See www.cancer.gov/ for additional information on the National Cancer Institute.

⁸ Please see www.psychiatry.org/psychiatrists/practice/dsm for additional information on the Diagnostic and Statistical Manual of Mental Disorders (DSM-5).

⁹ Please see www.cdc.gov/nchs/injury/injury_tools.htm for additional details.

¹⁰ Please see *ICD-10-CM Official Guidelines for Coding and Reporting FY 2020* (https://www.cdc.gov/nchs/data/icd/10cmguidelines-FY2020_final.pdf) for additional details.

- “Underdosing” refers to taking less of a medication than is prescribed by a provider or a manufacturer’s instruction
 - “Toxic effect” occurs when a harmful substance is ingested or comes in contact with a person.
6. ICD-10-CM medical and surgical complication codes are located in CCSR categories on the basis of the chapter in which they are found in the ICD-10-CM codebook. This means that there are complication categories in 12 body systems.¹¹ Codes specifically listed in sections with headers identifying complications (e.g., “intraoperative and postprocedural complications”, “complications in anesthesia during pregnancy”, “complications of labor and delivery”, “complication of device”, “complications predominantly related to the puerperium”, “complications of surgical and medical care”) in the ICD-10-CM codebook are assigned to a complication category.
 7. Within the Pregnancy, childbirth and the puerperium body system, codes for complicating conditions are specified as occurring during pregnancy, childbirth, or the puerperium. When possible, these conditions are grouped into separate categories related to the timing of the complication. However, there is not enough information to distinguish this timing in all cases. If the complicating condition code does not specify the timing, it is placed in a category for other specified complications of pregnancy. This category can include complications that occur during pregnancy, childbirth, or the puerperium.
 8. Codes that are described in the ICD-10-CM codebook with the keywords “infection,” “infectious,” or “infective” but are located in chapters other than Certain Infectious and Parasitic Diseases are cross-listed only to the *Certain infectious and parasitic disease* body system CCSR categories if there is a specified infection documented in the description. If no specific infection is documented, the code is not classified in a *Certain infectious and parasitic diseases* body system category. For example:
 - Code J20.0, Acute bronchitis due to mycoplasma pneumoniae, is assigned to RSP005 Acute bronchitis and INF003 Bacterial infections
 - Code J20.5, Acute bronchitis due to respiratory syncytial virus, is assigned to RSP005 Acute bronchitis and INF008 Viral infection.
 - Code J20.9, Acute bronchitis, unspecified, is assigned only to RSP005 Acute bronchitis and not to a category within the *Certain infectious and parasitic diseases* body system.
 9. Some categories in the CCSR for ICD-10-CM, such as those related to injury, include details on episodes of care. The CCSR for ICD-10-CM often separately classifies ICD-10-CM diagnoses indicating initial encounters, subsequent encounters, and sequela,

¹¹ Diseases of the blood and blood forming organs and certain disorders involving the immune system (BLD); Diseases of the circulatory system (CIR); Diseases of the digestive system (DIG); Diseases of the ear and mastoid process (EAR); Endocrine, nutritional, and metabolic diseases (END); Diseases of the eye and adnexa (EYE); Diseases of the genitourinary system (GEN); Injury, poisoning and certain other consequences of external causes (INJ); Diseases of the musculoskeletal system (MUS); Diseases of the nervous system (NVS); Diseases of the respiratory system (RSP); Diseases of the skin and subcutaneous tissue (SKN).

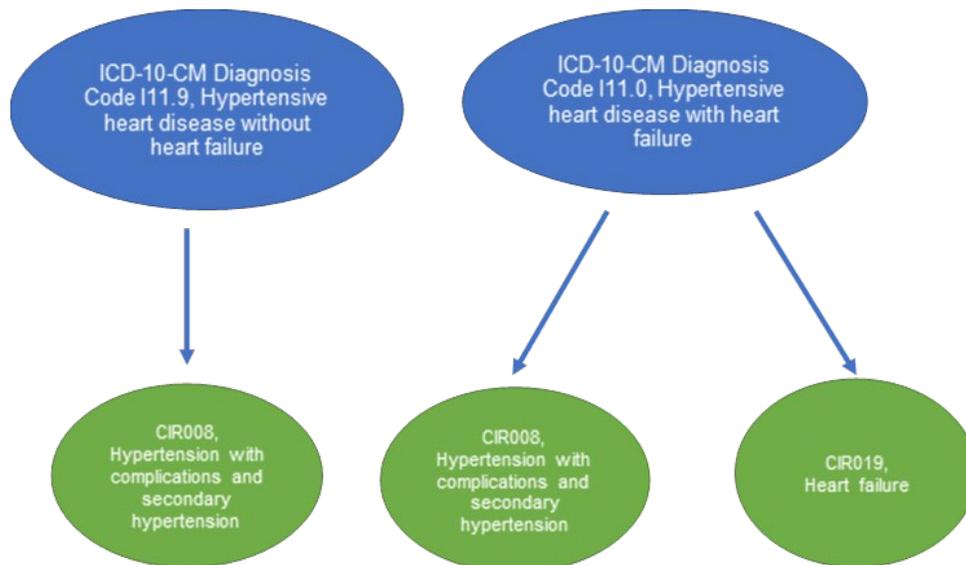
where applicable.¹² A sequela is the residual effect (condition produced) after the acute phase of an illness or injury has terminated.¹³ A sequela condition for an injury is a separate condition that arises during the treatment of the injury that needs its own treatment. A common example of a sequela condition is scar formation following a burn injury. The scar is the sequela condition and would be grouped with other scar conditions. A burn code with a seventh character of S indicates the injury responsible for the sequela condition and is assigned to the Injury; sequela encounter category. If a category contains a mix of codes with a seventh character indicating an initial encounter and codes that do not specify the episode of care, the category label does not specify an episode of care.

One-to-One and One-to-Many Mapping of ICD-10-CM Diagnosis Codes to CCSR

Why Cross-Classification of Codes are Necessary in the CCSR

Figure 1 provides a visual representation of how one-to-one and one-to-many mappings occur in the CCSR, using two ICD-10-CM diagnosis codes related to hypertension as examples. The ICD-10-CM diagnosis code I11.9, Hypertensive heart disease without heart failure, maps to a single CCSR category (CIR008 – Hypertension with complications and secondary hypertension), whereas ICD-10-CM diagnosis code I11.0, Hypertensive heart disease with heart failure, maps to two CCSR categories (CIR008 – Hypertension with complications and secondary hypertension and CIR019 – Heart failure).

Figure 1. Example of One-to-One and One-to-Many CCSR Mappings for Hypertension and Heart Failure



Abbreviations: CCSR, Clinical Classification Software Refined; ICD-10-CM, International Classification of Diseases, Tenth Revision, Clinical Modification.

¹² Although this approach results in a large number of injury categories, it is analytically easier for many users to combine categories than to disaggregate them.

¹³ Please see *ICD-10-CM Official Guidelines for Coding and Reporting FY 2020* (https://www.cdc.gov/nchs/data/icd/10cmguidelines-FY2020_final.pdf) for additional details.

Having code I11.0 mapped to two CCSR categories allows users to completely identify both hypertension cases and heart failure. The overarching goal of the CCSR, stated earlier, was to create clinically meaningful categories where the ICD-10-CM codes match the clinical intent of the category. The best way to accomplish that goal was to allow the coding and clinical review teams the flexibility to cross-classify individual codes into more than one CCSR category as necessary.

Prevalence of Cross-Classified Codes in the CCSR

Every ICD-10-CM code is assigned to at least one CCSR category but can be assigned to multiple categories. Version 2019.1 of the CCSR for ICD-10-CM includes more than 70,000 ICD-10-CM codes. Table 3 describes the distribution of CCSR category assignments.

Table 3. Type of CCSR Category Assignment by Number of ICD-10-CM Codes, Version 2019.1

Type of CCSR Category Assignment	Number of ICD-10-CM Codes
Codes assigned to only one category	64,791 (89.4%)
Codes assigned to two categories	5,728 (7.9%)
Codes assigned to three categories	1,535 (2.1%)
Codes assigned to four categories	361 (0.5%)
Codes assigned to five categories	21 (0.03%)
Total	72,436 (100%)

Abbreviations: CCSR, Clinical Classifications Software Refined; ICD-10-CM, International Classification of Diseases, Tenth Revision, Clinical Modification

Impact on Users

The decision to cross-classify some codes into more than one CCSR category allows researchers greater flexibility to customize the CCSR for their own analyses, but it does introduce new responsibilities for users relative to previous versions of the CCSR. The CCSR is not a medical coding grouper structured to assign a single diagnosis category to each encounter or even each ICD-10-CM diagnosis code. Users accustomed to using the CCS in this capacity will need to review cross-classified ICD-10-CM codes to make a determination on which category is most appropriate for their own research study. Users should not rely on the relative positioning of CCSR category assignments within the CCSR mapping file, as the relative ordering of categories was not part of the clinical review process. Examples of research designs that may be impacted by this are:

- Distributing all dataset records into mutually exclusive groups
- Obtaining statistics on discharge-level or record-level outcomes or measures
- Performing rank utilization by diagnoses.

These type of research questions can still use the CCSR for ICD-10-CM diagnoses, but users will be responsible for creating their own hierarchy to resolve issues introduced by the cross-classified category assignments.

Capitalizing on ICD-10-CM Diagnosis Coding Specificity

The previous section described how the specificity of ICD-10-CM diagnosis codes resulted in cross-classification of some codes in order to fully capture the clinical intent of CCSR categories. This section describes other ways the CCSR capitalized on the specificity of ICD-10-CM diagnosis codes to add value to the tool.

The CCSR tool also incorporates ICD-10-CM coding specificity by creating clinical categories not available in previous versions of the tool. The CCSR has nearly twice as many categories as the beta versions of the CCS for ICD-10-CM. All 21 body systems under the CCSR for ICD-10-CM have additional clinical categories as compared with the beta versions of the CCS, with the largest increases occurring in the *Injury, poisoning and certain other consequences of external causes* and *Neoplasms* body systems. Table 4 shows the number of CCSR categories and the number of CCS categories in the beta versions for ICD-10-CM for each body system.

Table 4. Number of CCSR and CCS (Beta Versions) Categories by Body System

Three-Character Body System Abbreviation	ICD-10-CM Body System Description	CCSR Categories	CCS Categories (Beta Versions)
INF	Certain infectious and parasitic diseases	11	10
NEO	Neoplasms	74	37
BLD	Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism	10	6
END	Endocrine, nutritional and metabolic diseases	17	11
MBD	Mental, behavioral and neurodevelopmental disorders	34	15
NVS	Diseases of the nervous system	22	11
EYE	Diseases of the eye and adnexa	12	6
EAR	Diseases of the ear and mastoid process	6	3
CIR	Diseases of the circulatory system	39	26
RSP	Diseases of the respiratory system	17	13
DIG	Diseases of the digestive system	25	20
SKN	Diseases of the skin and subcutaneous tissue	7	4
MUS	Diseases of the musculoskeletal system and connective tissue	37	12
GEN	Diseases of the genitourinary system	26	20
PRG	Pregnancy, childbirth and the puerperium	30	21
PNL	Certain conditions originating in the perinatal period	13	7
MAL	Congenital malformations, deformations and chromosomal abnormalities	10	5
SYM	Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified	17	9
INJ	Injury, poisoning and certain other consequences of external causes	76	20
EXT	External causes of morbidity	30	21
FAC	Factors influencing health status and contact with health services	25	6

Abbreviations: CCS, Clinical Classifications Software; CCSR, Clinical Classifications Software Refined; International Classification of Diseases, Tenth Revision, Clinical Modification

The following list provides a small subset of new categories available in the CCSR for each body system:¹⁴

- Certain infectious and parasitic diseases (INF) – fungal infections; foodborne intoxications
- Neoplasms (NEO) – six categories for different types of leukemia (chronic lymphocytic, acute lymphoblastic, acute myeloid, chronic myeloid, hairy cell, and all other types)
- Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism (BLD) – nutritional anemia; hemolytic anemia; aplastic anemia
- Endocrine, nutritional and metabolic diseases (END) – diabetes mellitus type 1; diabetes mellitus type 2; obesity
- Mental, behavioral, neurodevelopmental disorders (MBD) – opioid-related disorders; neonatal abstinence syndrome; fetal alcohol syndrome
- Diseases of the nervous system (NVS) – neurocognitive disorders; cerebral palsy
- Diseases of the eye and adnexa (EYE) – strabismus; refractive error
- Diseases of the ear and mastoid process (EAR) – diseases of the middle ear and mastoid (except otitis media); diseases of the inner ear and related conditions
- Diseases of the circulatory system (CIR) – chronic rheumatic heart disease; hypotension
- Diseases of the respiratory system (RSP) – sinusitis; pneumothorax
- Diseases of the digestive system (DIG) – gastrointestinal and biliary perforation; noninfective hepatitis
- Diseases of the skin and subcutaneous tissue (SKN) – pressure ulcer of skin; non-pressure ulcer of skin; contact dermatitis
- Diseases of the musculoskeletal system and connective tissue (MUS) – scoliosis and other postural dorsopathic deformities
- Diseases of the genitourinary system (GEN) – erectile dysfunction; male infertility
- Pregnancy, childbirth and the puerperium (PRG) – anesthesia complications during pregnancy; maternal intrauterine infection
- Certain conditions originating in the perinatal period (PNL) – neonatal cerebral disorders; neonatal digestive and feeding disorders
- Congenital malformations, deformations, and chromosomal abnormalities (MAL) – cleft lip or palate; chromosomal abnormalities
- Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified (SYM) – dysphagia; abnormal findings without diagnosis
- Injury, poisoning and certain other consequences of external causes (INJ) – underdosing of drugs, initial encounter; toxic effects, initial encounter
- External causes of morbidity (EXT) – intent of injury, assault; poisoning by non-drug
- Factors influencing health status and contact with health services (FAC) – personal/family history of disease; resistance to antimicrobial drugs; organ transplant status

¹⁴ See Appendix B: CCSR Categories for a complete list of categories in the CCSR. Appendix B is only available on the CCSR page of the HCUP User Support Web site - www.hcup-us.ahrq.gov/toolssoftware/ccsr/ccs_refined.jsp.

Episode of Care Coding in ICD-10-CM Diagnoses

ICD-9-CM codes did not include information on episodes of care for injuries or the trimester of specific pregnancy diagnosis codes. When the beta version of the CCS for ICD-10-CM diagnoses was developed, the categories were maintained from the CCS for ICD-9-CM.¹⁵ Therefore, previous versions of the CCS tool did not distinguish between episodes of care or trimesters of pregnancy.

In contrast, ICD-10-CM coding for injuries often uses a seventh character to distinguish episodes of care. In most injury codes with episode of care coding, the options for the seventh character are initial encounter, subsequent encounter, and sequela.¹⁶

An initial encounter, designated by a seventh character of A, is the active treatment of the injury. Coding as an initial encounter does not depend on whether the encounter represents the first time the patient is seen by a healthcare provider. It depends only on receiving active treatment for the injury. In some cases, a patient may need to be stabilized before the active treatment may begin. Therefore, multiple encounters may be coded as an initial encounter for the same injury if active treatment is ongoing.

A subsequent encounter, designated by a seventh character of D, is used when the patient has completed active treatment and is receiving follow-up care after the injury. Subsequent encounters represent routine care during healing and recovery of the injury.

A sequela, designated by a seventh character of S, is the residual effect (condition produced) after the acute phase of an illness or injury has terminated.¹⁷ A sequela condition for an injury is a separate condition that arises during the treatment of the injury that requires its own treatment.

¹⁵ See [Appendix A: Background on the Development of the CCSR](#) for additional details.

¹⁶ ICD-10-CM codes for traumatic fractures have additional seventh character values. A—initial encounter for closed fracture; B—initial encounter for open fracture type I or II; C—initial encounter for open fracture type IIIA, IIIB, or IIIC; D—subsequent encounter for closed fracture with routine healing; E—subsequent encounter for open fracture type I or II with routine healing; F—subsequent encounter for open fracture type IIIA, IIIB, or IIIC with routine healing; G—subsequent encounter for closed fracture with delayed healing; H—subsequent encounter for open fracture type I or II with delayed healing; J—subsequent encounter for open fracture type IIIA, IIIB, or IIIC with delayed healing; K—subsequent encounter for closed fracture with nonunion, M—subsequent encounter for open fracture type I or II with nonunion, N—subsequent encounter for open fracture type IIIA, IIIB, or IIIC with nonunion, P—subsequent encounter for closed fracture with malunion, Q—subsequent encounter for open fracture type I or II with malunion, R—subsequent encounter for open fracture type IIIA, IIIB, or IIIC with malunion, S—sequela. These expanded options are aggregated as initial, subsequent, or sequela for the CCSR. The additional detail of the expanded codes is not incorporated. Please see *ICD-10-CM Official Guidelines for Coding and Reporting FY 2020* (https://www.cdc.gov/nchs/data/icd/10cmguidelines-FY2020_final.pdf) for additional details.

¹⁷ Please see *ICD-10-CM Official Guidelines for Coding and Reporting FY 2020* (https://www.cdc.gov/nchs/data/icd/10cmguidelines-FY2020_final.pdf) for additional details and examples.

Example of Episode of Care CCSR Categories

Table 5 illustrates how episode of care coding is incorporated into new CCSR categories for codes related to a fracture. Rows 1 and 2 represent codes indicating an initial encounter (seventh character of A and B) for treatment of a fracture, S42.022A and S42.022B. These initial encounters include the active treatment of the fractures. Rows 3–6 show the same code with seventh characters of D, G, K, and P indicating subsequent encounters (ICD-10-CM codes S42.022D, S42.022G, S42.022K, and S42.022P). The subsequent encounters could be the removal of a cast. Row 7 is again the same code but with a seventh character of S indicating a sequela (ICD-10-CM code S42.022S). The sequela condition of the fracture may be ongoing pain.¹⁸ If there is a sequela of the fracture, there would be both an ICD-10-CM code to indicate what the condition in need of treatment is (i.e., the malformation or pain) and the fracture sequela code to indicate the injury responsible for the sequela condition. In this example, there are three separate CCSR categories, one for each type of encounter. This is important because a researcher interested in the number of upper limb fractures may not want to include encounters for cast removal or ongoing pain. This separation of categories allows researchers to easily distinguish between the different types of encounters when using the CCSR to categorize injuries and other conditions with episode of care coding.

Table 5. Example of CCSR Categories for Episode of Care

Row	ICD-10-CM Code	ICD-10-CM Code Description	CCSR Category	CCSR Category Description
1	S42022A	Displaced fracture of shaft of left clavicle, initial encounter for closed fracture	INJ004	Fracture of the upper limb; initial encounter
2	S42022B	Displaced fracture of shaft of left clavicle, initial encounter for open fracture	INJ004	Fracture of the upper limb; initial encounter
3	S42022D	Displaced fracture of shaft of left clavicle, subsequent encounter for fracture with routine healing	INJ041	Fracture of the upper limb; subsequent encounter
4	S42022G	Displaced fracture of shaft of left clavicle, subsequent encounter for fracture with delayed healing	INJ041	Fracture of the upper limb; subsequent encounter
5	S42022K	Displaced fracture of shaft of left clavicle, subsequent encounter for fracture with nonunion	INJ041	Fracture of the upper limb; subsequent encounter
6	S42022P	Displaced fracture of shaft of left clavicle, subsequent encounter for fracture with malunion	INJ041	Fracture of the upper limb; subsequent encounter
7	S42022S	Displaced fracture of shaft of left clavicle, sequela	INJ073	Injury; sequela

Abbreviations: CCSR, Clinical Classifications Software Refined; International Classification of Diseases, Tenth Revision, Clinical Modification

¹⁸ Another common example of a sequela condition is scar formation following a burn injury. The scar is the sequela condition. A burn code with a seventh character of S indicates the injury responsible for the sequela condition.

Trimester Coding in ICD-10-CM Diagnoses

Many ICD-10-CM pregnancy codes (starting with characters O00–O99) have characters to indicate a specific trimester.¹⁹ This is another example of code specificity that has been incorporated into the CCSR categories. Table 6 illustrates how trimester coding is incorporated into new CCSR categories. Each row represents code O26.85, Spotting complicating pregnancy, with a different final character indicating different trimesters of pregnancy. The CCSR categories for these codes separate hemorrhages into two groups, (1) first or unspecified trimester and (2) second or third trimester. This allows users to easily separate first trimester bleeding from second or third trimester bleeding.

Table 6. Example of CCSR Categories by Pregnancy Trimester

Row	ICD-10-CM Code	ICD-10-CM Code Description	CCSR Category	CCSR Category Description
1	O26851	Spotting complicating pregnancy, first trimester	PRG009	Early, first or unspecified trimester hemorrhage
2	O26852	Spotting complicating pregnancy, second trimester	PRG010	Hemorrhage after first trimester
3	O26853	Spotting complicating pregnancy, third trimester	PRG010	Hemorrhage after first trimester
4	O26859	Spotting complicating pregnancy, unspecified trimester	PRG009	Early, first or unspecified trimester hemorrhage

Abbreviations: CCSR, Clinical Classifications Software Refined; International Classification of Diseases, Tenth Revision, Clinical Modification

Impact of New Clinical CCSR Categories on Users

The decision to create new CCSR categories unavailable in previous versions of the CCS enhances the ability for researchers to use the CCSR to identify more targeted patient populations. However, it does create a discontinuity for longitudinal trend analyses with previous versions of the CCS. Discontinuities in trend analyses are likely even for CCSR categories with similar descriptions or names to those in the CCS for ICD-9-CM, as the coding definitions of the categories may have changed.

Users wanting to look at trends should carefully analyze their data before and after the transition to ICD-10-CM to determine if additional statistical methods are necessary to address trend breaks introduced by ICD-10-CM coding and the differences between the CCS for ICD-9-CM and the CCSR for ICD-10-CM.

¹⁹ Although the majority of ICD-10-CM codes in the Pregnancy, Childbirth and the Puerperium chapter include trimester coding, the position of this character is not always the same. In other words, trimesters are not always indicated with the seventh character like episodes of care. However, trimester coding is indicated with the final character of pregnancy codes. Please see *ICD-10-CM Official Guidelines for Coding and Reporting FY 2020* (https://www.cdc.gov/nchs/data/icd/10cmguidelines-FY2020_final.pdf) for additional details and examples.

USING THE DOWNLOADABLE CCSR FILES

System Requirements

Using the CCSR for ICD-10-CM diagnoses requires a program to decompress or “unzip” files.²⁰ Approximately 2 megabytes of disk space available on one’s hard drive also will be needed to accommodate all the CCSR files. Additional space is necessary for saving CCSR output files. See below for additional details on the options for CCSR output and the disk space considerations for each option.

Downloadable Files

The CCSR for ICD-10-CM diagnosis zip file contains (1) a comma separated values (CSV) file that includes the mapping of ICD-10-CM diagnosis codes into CCSR categories with labels for the individual CCSR categories, (2) a SAS program to apply the tool to the user’s data, (3) an Excel CCSR reference file with filters applied, (4) a PDF copy of the CCSR User Guide, (5) a PDF copy of Appendix B listing all CCSR categories, and (6) an Excel version of Appendix B listing all CCSR categories with filters applied. Table 7 includes additional detail on the names and purpose of the files included in the CCSR for ICD-10-CM diagnosis zip file.

²⁰ Third-party zip utilities are available from the following reputable vendors on their official websites: ZIP Reader (Windows) (free download offered by PKWARE, Inc.), SecureZIP® for Mac or Windows (free evaluation and licensed/fee software offered by PKWARE, Inc.), WinZip (Windows) (evaluation and fee versions offered by the Corel Corporation), Stuffit Expander® (Mac) (free evaluation and licensed/fee software offered by Smith Micro Software Inc.).

Table 7. Contents of the CCSR for ICD-10-CM Diagnoses Zip File

File Name	Purpose
<p>\$DXCCSRyyyy_r.csv where yyyy represents fiscal year and r represents a release number within fiscal year^a</p>	<p>The CSV mapping file lists ICD-10-CM diagnosis codes along with a description for each ICD-10-CM diagnosis code, the CCSR categories assigned, and the full description corresponding to each CCSR category. This is a <i>horizontal</i> file that the SAS program uses as input to build the tool. The term horizontal is used because each individual ICD-10-CM diagnosis code is listed once (on a single row) with the information on the CCSR categories listed across the columns of the CSV file.</p> <p>This file can be converted to Excel where a filter can be applied to examine individual ICD-10-CM diagnosis codes or CCSR categories.</p>
<p>DXCCSR_Mapping_Program.sas</p>	<p>SAS mapping program applies the CCSR to the user's ICD-10-CM-coded data. The mapping program includes two options for the file structure of CCSR output.</p>
<p>DXCCSR-Reference-File-vyyyy-r.xlsx where yyyy represents fiscal year and r represents a release number within fiscal year</p>	<p>A Microsoft Excel reference file that includes two tabs: (1) The first tab lists all ICD-10-CM diagnosis codes with descriptions and the CCSR categories assigned with descriptions. A filter is applied to examine individual ICD-10-CM diagnosis codes or CCSR categories. (2) The second tab lists CCSR categories and numbers.</p> <p>This file differs from the CSV file in structure, but not in content. The mapping of the ICD-10-CM diagnosis codes to the CCSR categories is structured in a stacked (i.e. vertical) format, where cross-classified codes are listed in multiple rows for each different CCSR assignment.</p>
<p>DXCCSR-User-Guide-vyyyy-r.pdf where yyyy represents fiscal year and r represents a release number within fiscal year</p>	<p>A PDF User Guide for the CCSR with detailed documentation on the tool and how to use it with administrative datasets.</p>
<p>DXCCSR-AppendixB-vyyyy-r.pdf where yyyy represents fiscal year and r represents a release number within fiscal year</p>	<p>A PDF Appendix to the User Guide that lists the CCSR categories by body system.</p>

Abbreviations: CCSR, Clinical Classifications Software Refined; CSV, comma separated values; ICD-10-CM, International Classification of Diseases, Tenth Revision, Clinical Modification

^a For example, the first mapping file release to include codes valid through fiscal year 2019 is named \$DXCCSR2019_1.csv.

Flexible File Structure for Outputting the CCSR for ICD-10-CM Diagnoses

Within the SAS mapping program, users are given two options for the file structure of their output once the CCSR tool is applied to the data. The programs assume that the input file will have one record per encounter (inpatient stay, emergency department visits, etc.) with a unique record identifier and an array of ICD-10-CM diagnosis codes. There is no restriction on the

maximum length of the input diagnosis array. Users can modify the SAS code to match the structure of their data.

Output Option 1, Vertical Output File

By default, the SAS program creates a “vertical” file with one or more observations (i.e., rows) for each input record. The number of observations created from each input record depends on the number of CCSR categories triggered by the ICD-10-CM diagnoses on the input record. For example, if an input record has only one ICD-10-CM diagnosis code and that code maps to two CCSR categories, the vertical file for that input record would have two observations to display the CCSR records. If an input record has ten diagnosis ICD-10-CM diagnosis codes and they each map to just one CCSR category, the vertical file for that record would have ten observations to display the CCSR records.

The vertical file contains three data elements (i.e., columns):

- i. The record identifier (data element RECID)
- ii. A CCSR category (data element DXCCSR)
- iii. The diagnosis position (data element DX_Position) which indicates which ICD-10-CM diagnosis code on the record triggered the CCSR category assignment.

The vertical file option affords users more efficient storage—about 20 bytes per record. It also maintains information on the exact position of the ICD-10-CM diagnosis that triggered the CCSR category. However, because this option outputs multiple observations corresponding to a single input record and a single diagnosis code, its practical application may be challenging for users with limited statistical programming experience.

Output Option 2, Horizontal Array of all CCSR Categories

Users also can specify the SAS mapping program to create a horizontally structured CCSR output file. In this case, the SAS program creates a file with only one observation (i.e., row) for each input record that includes the record identifier and a horizontal array of 538 data elements corresponding to the CCSR categories. There is one data element for each CCSR category, with the data element name corresponding to the CCSR category along with a “DXCCSR_” prefix, where DX stands for diagnosis. For example, the data element name for the CCSR category for Essential hypertension is “DXCCSR_CIR007”.

The horizontal file option allows users to differentiate between the principal diagnosis, secondary diagnoses, or both types of diagnoses but does not retain the exact position of the ICD-10-CM diagnosis on the record in the input data file. The SAS program assigns one of four values for each DXCCSR data element:

- 0 – The CCSR was not triggered by any ICD-10-CM diagnosis code on the input record.
- 1 – The CCSR was triggered by only the principal (or first-listed diagnosis) on the input record.
- 2 – The CCSR was triggered by both the principal (or first-listed diagnosis) and any secondary diagnosis on the input record.
- 3 – The CCSR was triggered by only secondary diagnosis code(s) on the input record.

The SAS program produces a SAS data set with the following data elements:

- RECID – A record identifier

- Horizontal array of all 538 “DXCCSR_” data elements

This option is conceptually straightforward for analyses because all CCSR categories triggered by diagnoses on a record are kept at the record level without further transformation. However, the horizontal structure includes a data element for every CCSR category, and as a result, it retains CCSR categories that are not triggered by a record. Each record of horizontal output takes up approximately 1.5 KB disk space, potentially presenting computational or data storage challenges for users with large datasets or limited disk space.

Example of Output from Option 1, Vertical Output File, and Option 2, Horizontal Array of All CCSR Categories

Table 8 displays a sample input record, where four ICD-10-CM diagnoses are stored in an array of data elements I10_DX1–I10_DX4 for a record with RECID equal to 1234. Table 9 displays the vertical file output for this sample input record and Table 10 contains an example of the horizontal output file, using the same sample input record.

Table 8. Sample of Input File Record

RECID	I10_DX1	I10_DX2	I10_DX3	I10_DX4
1234	M84664A	N184	E1122	M80862A

Table 9. Example of the Vertical Output File for the Sample Record

Row	RECID	DXCCSR	DX_POSITION
1	1234	MUS014	1
2	1234	GEN003	2
3	1234	END003	3
4	1234	END005	3
5	1234	GEN003	3
6	1234	MUS013	4
7	1234	MUS014	4

The first row of Table 9 indicates that the first diagnosis code (DX_POSITION=1) for RECID 1234 triggers one CCSR, MUS014–Pathological fracture, initial encounter (DXCCSR=MUS014). The second diagnosis also maps to only one CCSR (DXCCSR=GEN003 for Chronic kidney disease and DX_POSITION=2). The third diagnosis is cross-classified into three CCSR categories listed in rows 3–5: END003–Diabetes mellitus with complication, END005–Diabetes mellitus, Type 2, and GEN003–Chronic kidney disease. The fourth diagnosis code (Rows 6–7 in Table 9) is cross classified into two CCSR including the CCSR MUS014–Pathological fracture, initial encounter that is triggered also by the principal (or first-listed) diagnosis.

In the example of the horizontal output file (Table 10) the fact that both the principal (or first-listed) diagnosis and a secondary diagnosis trigger MUS014 results in data element DXCCSR_MUS014 being assigned the value 2. The other “DXCCSR_” data elements triggered by a diagnosis on the sample record are assigned the value 3 because they are triggered by secondary diagnoses only (i.e., data elements I10_DX2–I10_DX4 in Table 8). Data element DXCCSR_BLD001 and DXCCR_SYM017 (as well as all other “DXCCSR_” data

elements not shown in Table 10) are assigned the value 0 because they are not triggered by any ICD-10-CM diagnosis codes on the input record.

Table 10. Example of Horizontal Output File

RECID	DXCSR_ BLD001	...	DXCSR_ END003	...	DXCSR_ END005	...	DXCSR_ GEN003	...	DXCSR_ MUS013	DXCSR_ MUS014	...	DXCSR_ SYM017
1234	0		3		3		3		3	2		0

Representation of ICD-10-CM Diagnosis Codes

ICD-10-CM diagnoses often are represented by 4- to 7-digit alphanumeric codes with explicit decimals. In the CSV mapping file, the ICD-10-CM diagnosis codes are enclosed in quotation marks and do not contain decimals. In addition, the SAS program that assigns the CCSR, ICD-10-CM codes are expected to be 7-digit alphanumeric character strings with implicit decimals. Table 11 provides examples for how the ICD-10-CM diagnosis codes are represented in the CSV mapping file and expected by the SAS program.

Table 11. Example of Representation of ICD-10-CM Diagnosis Codes in the CCSR for ICD-10-CM

Condition	ICD-10-CM Diagnosis Code	Alphanumeric Code (With Implicit Decimals) in the CSV File
Single liveborn infant, delivered vaginally	Z38.00	'Z3800 '
Sepsis, unspecified organism	A41.9	'A419 '
Pneumonia, unspecified organism	J18.9	'J189 '

Abbreviations: CCSR, Clinical Classifications Software Refined; CSV, comma separated values; ICD-10-CM, International Classification of Diseases, Tenth Revision, Clinical Modification

For the accurate assignment of CCSR, the ICD-10-CM diagnosis codes in the input files must be reported as follows:

- Alphanumeric diagnosis codes must be left-justified, allowing trailing blanks to fill up the full length of 7 characters.
 - Example: Diagnosis J18.9 should be retained as 'J189 ' in the input file.
- Trailing blanks should never be zero-padded.
 - Example: Diagnosis 'J189 ' should *not* be represented as 'J189000'.
- Only fully specified ICD-10-CM codes should be included. The CCSR software does not account for substrings or parent codes, such as three-digit code blocks
 - Example: Diagnosis E11 (Type 2 diabetes mellitus) or E11x cannot be used to capture all codes that begin with the characters E11.

Running the SAS Program to Add CCSR Categories to Data

To download, modify, and run the software to apply the CCSR for ICD-10-CM diagnoses to an input dataset, follow these steps:

1. STEP 1: Users should download and extract the contents of the Zip file containing the CCSR for ICD-10-CM diagnosis tool to a saved location on their computer. Files included in the Zip file are described in Table 7 and referenced below.

2. STEP 2: Users must set up the SAS program (DXCCSR_Mapping_Program.sas) to run on their data. They must specify or modify the following where appropriate:
 - a. Change the paths in the SAS program to point to the computer location(s) of
 - i. the CSV mapping file (\$DXCCSRyyyy_r.csv)
 - ii. the input dataset
 - iii. the output dataset
 - b. Set the macro variables in the SAS program to match the data element names and file structure of the input dataset (see Table 12).

Table 12. Modifiable Macro Variable and Directory Paths

Description of Macro Variables and Directory Paths	SAS Program Syntax
Location of the CSV mapping file	FILENAME INRAW1
Specify the location of the input dataset	LIBNAME IN1
Specify the location of the output dataset(s)	LIBNAME OUT1
Choose whether to have the program to build the horizontal output file (1=yes, 0=no). The vertical output file is always built by the SAS program	%LET HORZ=1;
Specify the unique record identifier in the input dataset (KEY in most HCUP databases)	%LET RECID=KEY;
Specify the prefix used to name the ICD-10-CM diagnosis data element vector of the input dataset (I10_DX in HCUP databases)	%LET DXPREFIX=I10_DX;
Specify the data element, if available, that documents the number of diagnoses on each specific record. This will improve processing efficiency. If such a data element is not available in the input dataset, leave this macro blank and the logic will use the NUMDX value for all records	%LET NUMDX=30;
Record-specific count of diagnoses (if not available, leave it blank; I10_NDX in HCUP databases)	%LET NDXVAR=I10_NDX;
Specify the file name of the input dataset	%LET CORE=INPUT_SAS_FILE;
Specify a file name for the vertical output file	%LET VERTFILE=OUTPUT_VERT_FILE_NAME;
Specify a file name for the horizontal output file	%LET HORZFILE=OUTPUT_HORZ_FILE_NAME;

Abbreviation: CSV, comma separated values

After running of the SAS mapping program, two output datasets could be generated:

1. A vertical file with record identifier (data element RECID), CCSR category (data element DXCCSR), and the position of the diagnosis (data element DX_Position). In this file, each record identifier RECID is output multiple times to match the number of diagnosis codes present on the record.
2. A horizontal file with record identifier (RECID), categorical data elements DXCCSR_BBBnnn for the array of all CCSR. “BBBnnn” are the values of CCSR categories. See Appendix B: CCSR Categories file for specific names and descriptions of the CCSR categories.²¹

²¹ Appendix B: CCSR Categories is only available on the CCSR page of the HCUP User Support website (www.hcup-us.ahrq.gov/toolsoftware/ccsr/ccs_refined.jsp).

APPENDIX A: BACKGROUND ON THE DEVELOPMENT OF THE CCSR

In October 2015, the United States transitioned to a modified version of the World Health Organization International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM), replacing the ICD-9-CM diagnosis coding system with the ICD-10-CM diagnosis coding system for most inpatient and outpatient medical encounters. An overview of differences between ICD-10-CM and ICD-9-CM is available on the HCUP User Support website at www.hcup-us.ahrq.gov/datainnovations/icd10_resources.jsp.

Impact on the Clinical Classifications Software

In preparation for the October 2015 implementation of ICD-10-CM, the Healthcare Cost and Utilization Project (HCUP) tools were converted to the new coding system. The initial mapping was completed by linking ICD-10-CM/PCS (Procedure Coding System) codes to the current Clinical Classifications Software (CCS) Agency for Healthcare Research and Quality (AHRQ) classification assignments via the General Equivalence Mappings (GEMs) available from the Centers for Medicare & Medicaid Services (CMS) website (see <http://www.cms.gov/Medicare/Coding/ICD10/2014-ICD-10-CM-and-GEMs.html> for more information on GEMs). The GEMs files were processed to create GEMs-driven maps between ICD-9-CM and ICD-10-CM codes. Credentialed coders trained in both ICD-9-CM and ICD-10-CM reviewed the maps to ensure the validity of the CCS. The goal of the team was to create a set of clinical coding definitions that matched the categories in the CCS for ICD-9-CM and would ease the transition of users from ICD-9-CM- to ICD-10-CM-coded data. The CCS categories remained the same, although some categories received slight naming modifications to remain accurate descriptions for the codes included in ICD-10-CM. The resultant first iteration of the ICD-10-CM classification was considered a beta version.

However, the development of the beta version of the CCS for ICD-10-CM was completed before ICD-10-CM-coded data became available. Once ICD-10-CM coded data became available, the beta version of the CCS was evaluated through preliminary analyses on HCUP data, which revealed unexpected discontinuities between the ICD-9-CM and ICD-10-CM versions of the CCS. These initial analyses encouraged AHRQ to reconsider the need for a refined ICD-10-CM version of the CCS.

Refinement of the Beta Versions of the CCS

In January 2018, AHRQ initiated planning for a refined version of the CCS, the Clinical Classifications Software Refined (CCSR) for ICD-10-CM diagnosis, which would take advantage of the specificity built into ICD-10-CM codes. The new goal was to create a CCS tool that prioritized defining the codes in each category to match the clinical intent of the category description. To adequately map the codes to clinically meaningful categories, it was important to have expert input from a coding and a clinical perspective from a variety of specialties.

The team first discussed and debated the intricacies and specificity of ICD-10-CM coding, shortcomings of the existing beta versions of the CCS tool, and potential revisions and modifications that could be implemented in a new diagnosis categorization tool. AHRQ staff reviewed the team's recommendations and made overarching decisions about the structure and design of the tool. The new CCSR tool was not restricted to include the same set of categories as the beta versions of the CCS tool, but rather the team was tasked with creating a set of categories that were identifiable in ICD-10-CM codes and valuable for health services research. Many categories were maintained from the beta versions of the CCS for ICD-10-CM with the

clinical concept retained, but the specific ICD-10-CM diagnoses included in the category may have been revised. The coding and clinical review teams were encouraged to create new categories or disaggregate existing categories when they identified opportunities to increase the value of the tool to health services researchers.

The most substantial change to the CCSR for ICD-10-CM compared with the beta versions of the CCS for ICD-10-CM diagnosis is that individual categories no longer contain a mutually exclusive set of ICD-10-CM codes. ICD-10-CM codes under the CCSR are mapped to at least one category but can be mapped to more than one category. Mutually exclusive category assignments have been a design feature of the CCS since its initial release (i.e., each ICD-9-CM diagnosis code was included in one and only one CCS category). However, individual ICD-10-CM diagnosis codes can be used to document multiple conditions or a condition and a common symptom or manifestation. Assigning only one CCS category for these codes would require prioritizing the assignment (and the preferred hierarchy of conditions might be different for different applications). An overarching goal of the CCSR was to create clinically meaningful categories where the ICD-10-CM codes match the clinical intent of the category. The best way to accomplish that goal was to allow the coding and clinical review teams the flexibility to cross-classify individual codes into more than one CCSR category as necessary.

Refinement Review Process

During the development of the CCSR, the coding definitions for each existing and proposed new category were reviewed thoroughly. The team worked through the set of ICD-10-CM codes and CCSR categories in groups organized by body system or ICD codebook chapter. This structure was retained in the final CCSR tool so that related categories would be easier to identify and so that future revisions to the CCSR would allow new categories to be inserted in a relevant section, rather than being relegated to the end of the category listing.

Within each body system, an American Health Information Management Association (AHIMA)-certified ICD-10-CM/PCS trainer suggested a set of relevant categories and assigned individual ICD-10-CM codes to each category. Each assignment was reviewed by a second AHIMA-certified ICD-10-CM/PCS trainer who checked for completeness, appropriate categorization for each code, and recognition of both specific codebook directions and published coding guidelines. Review from a third AHIMA-certified ICD-10-CM/PCS trainer was available to reconcile differences in assignment by the first two coders. Following the initial code assignments for each body system, a team of three physicians on the clinical review team also reviewed all categories and code assignments. The clinical team answered specific questions posed by the ICD-10-CM coders, such as whether to create a new category to capture a specific condition group. When necessary, specialty physicians were identified and called on to add body system specific expertise.

Following clinical review, AHRQ provided additional oversight and guidance. AHRQ staff continued to consult the expertise of the ICD-10-CM coders, the physicians, and other researchers to resolve any outstanding issues or decision points. The review process included multiple iterations of review and refinements. The team conducted various tests of the CCSR tool to identify potential improvements and modifications after its initial build with the ultimate aim of a more parsimonious set of categories. These tests resulted in additional cross-listings of codes across categories, creating new categories, renaming categories, combining or rolling up categories, and/or shifting categories from one body system to another. Such changes were the result of extensive input across multiple organizations and clinical specialties that evaluated the

CCSR category assignments for accuracy and clinical significance. The team extensively reviewed the CCSR tool at each stage of its development, using HCUP State Inpatient Databases (SID) from multiple States and years for some quality control testing. The team reviewed utilization counts of CCSR categories and individual ICD-10-CM codes, cross-tabulations by patient demographic characteristics, and correlation with Major Diagnostic Category and Medicare-Severity Diagnosis Related Group assignment. These tabulations were for informational purposes and assisted in further refinements, but there was not effort made to perfectly align the CCSR categories with MDCs and MS-DRGs, as the purpose of the classifications systems vary (i.e., CCSR is intended to assist in answering health research and policy questions, while MDCs and MS-DRGs are intended to assist with reimbursement). The ICD-10-CM coders identified all ICD-10-CM codes not categorized in body systems that match their codebook chapter location and reviewed the mapping assignments across body systems to ensure consistency with the CCSR category assignment rules.

APPENDIX B: CCSR CATEGORIES

Appendix B contains a set of tables listing the Clinical Classifications Software Refined (CCSR) for International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM) categories and numbers within each of the 21 body systems. Appendix B is only available on the CCSR page of the HCUP User Support website (www.hcup-us.ahrq.gov/toolssoftware/ccsr/ccs_refined.jsp).