

# International classification of diseases, 10th edition, clinical modification and procedure coding system: descriptive overview of the next generation HIPAA code sets

Steven J Steindel

## Correspondence to

Dr Steven J Steindel, 1021 Wadsworth Drive NW, Atlanta, GA 30318, USA; stevejs@mindspring.com

Received 6 October 2009  
Accepted 2 March 2010

## ABSTRACT

Described are the changes to ICD-10-CM and PCS and potential challenges regarding their use in the US for financial and administrative transaction coding under HIPAA in 2013. Using author constructed derivative databases for ICD-10-CM and PCS it was found that ICD-10-CM's overall term content is seven times larger than ICD-9-CM: only 3.2 times larger in those chapters describing disease or symptoms, but 14.1 times larger in injury and cause sections. A new multi-axial approach ICD-10-PCS increased size 18-fold from its prior version. New ICD-10-CM and PCS reflect a corresponding improvement in specificity and content. The forthcoming required national switch to these new administrative codes, coupled with nearly simultaneous widespread introduction of clinical systems and terminologies, requires substantial changes in US administrative systems. Through coordination of terminologies, the systems using them, and healthcare objectives, we can maximize the improvement achieved and engender beneficial data reuse for multiple purposes, with minimal transformations.

In April, 2004 President Bush directed the US healthcare system to adopt and use electronic health records (EHR) such that coverage extends to the 'entire population within a decade'.<sup>1</sup> In 2009, President Obama affirmed that announcement and Congress committed in excess of US\$17 billion via the American Recovery and Revitalization Act to ensure that it happens.<sup>2</sup> Since that time the US informatics community has focused on providing definitional infrastructure on the successful attributes those systems should have through emerging definitions for 'meaningful use' and requirements for certification of systems.<sup>3</sup>

Meanwhile, following an almost identical timeline, the financial and administrative reporting structure as defined by the Health Insurance Portability and Accountability Act of 1997 (HIPAA)<sup>4</sup> has implemented a generational update to the reporting structure and code sets used. Those providers implementing the newly prescribed EHR for their clinical use will also provide for a path in their administrative systems to accommodate the change in reporting structure from ASC X12 4010a to 5010<sup>5</sup> and move to the International Classification of Disease (ICD) version 10 revisions<sup>6 7</sup> of the reporting code sets. Wide variation exists in the total implementation cost estimates but, based on

final federal cost estimates, direct costs to providers for the administrative system updates will range between US\$9.5 and US\$16.9 billion, with ICD-10 changes amounting to between US\$2.3 and US\$2.7 billion.

The tension described above exists in our view of the code sets associated with the EHR and administrative transactions. Tang *et al*<sup>8</sup> have outlined the difficulties in using administrative classifications, primarily ICD variants, for clinical information. Chute and colleagues<sup>9 10</sup> have noted many of the inherent difficulties in using classifications to represent clinical content. Cimino and Clayton<sup>11</sup> have described how to use a reference terminology to manage some of the issues and then succinctly summarized these issues in the desiderata.<sup>12</sup> Brown *et al*<sup>13</sup> extended the work of Cimino by using SNOMED-CT as a reference terminology. This decade-plus long excursion clearly shows the difficulty the informatics community has with using classifications to express clinical content, despite the fact that the classifications are the only terminologies most likely encountered and used by most practitioners outside of research settings.

The clinical terminology community has also found difficulties in developing and deploying new products.<sup>14</sup> During the same time, from approximately 1995 to now, intense work on the development of clinical terminologies such as Read Codes Version 3,<sup>15</sup> SNOMED-RT<sup>16 17</sup> and the GALEN project<sup>18 19</sup> occurred, all trying to apply logic constructs to develop new inferences. Over this time, the National Health Service in the UK invested billions of pounds in the Read Codes Version 3 and SNOMED-CT,<sup>20</sup> now an international product of the International Health Terminology Standard Development Organization (IHTSDO) for intended future universal deployment.<sup>21</sup> Like the classifications, SNOMED-CT is not without criticism regarding completeness or logical construction.<sup>22–24</sup> The USA entered into an agreement with the intellectual property holder of SNOMED-CT, first the College of American Pathologists and now the IHTSDO, to make it freely and universally available in US healthcare.<sup>25</sup> It is still in limited use, primarily in those locations that used it before the US license, such as Kaiser Permanente<sup>26</sup> and the University of Nebraska. The only major national use now proposed is for the problem list. A SNOMED-CT subset of problems that are seen frequently in a variety of healthcare systems is available.<sup>25</sup> It remains to be seen

whether this subset will be used in ways that take advantage of SNOMED-CT's logical inference structure or—perhaps more likely—be employed simply as a large term list.

It is imperative that the informatics community understands the structure and potential impact of the new classifications so we can help guide successful introduction. They currently serve as the most used source for inferential decisions regarding health care and will continue to do so for at least the next decade, until and if a clinical code system providing internal logical support emerges and is universally accepted. Given the immediate widespread use that will occur on adoption of the new generation HIPAA classifications and the subsequent long-term effects, it is surprising that little detail on their terminological construction or changes from the previous version are available. This article will focus on these areas and suggest areas for focus and coordination.

### ICD-10-CM BACKGROUND

ICD-10 clinical modification (CM) is an evolutionary extension of ICD-10 developed and maintained by the World Health Organization (WHO) for worldwide mortality statistics.<sup>1, 27, 28</sup> Locally developed clinical modifications of the mortality codes using the previous version of ICD are common and the US version, ICD-9-CM, is our current administrative and public health statistics code set. ICD-9-CM was designed for approximately a decade of use and was scheduled for evolutionary change in approximately 2000.<sup>29</sup> The introduction of HIPAA at that time delayed that change for almost another 15 years. During that time the rigid structure of the actual codes prevented the full addition of new codes in many clinical areas needed to reflect the changes in medical care and corresponding science. It also placed the code base used in HIPAA transactions one version behind the current version of ICD-10 used for mortality reporting.

ICD-10-CM is maintained by the National Center for Health Statistics (NCHS).<sup>30</sup> Development was through internal resources and occurred mostly in the 1990s. Guidance on the development was through a Coordination and Maintenance Committee<sup>ii</sup>. The major change in the structure of ICD-10-CM compared with the early ICD-9-CM version was the introduction of the alphanumeric codes introduced in the international version to replace the strict numeric codes. The flexible use of letters in the base codes greatly expanded sectional (chapter) code space, allowing a more complete expression of modern medicine and making supplemental codes added to ICD-9-CM integral to the code base. A second change in code enumeration was the expansion from nine to 10 characters. As most legacy systems had a rigid database or flat-file structure for this space the change to a 10 character alphanumeric design forces updates to almost all administrative systems, part of the large provider-based implementation costs.<sup>31, 32</sup>

Codes in ICD-10-CM provide meaning that humans can understand (intelligent codes). They start with a letter designating a chapter and are followed by two numbers indicating a chapter section, with one exception noted in the present draft in the O chapter. A decimal follows these three alphanumeric characters. From zero to seven alphanumeric characters

follow the decimal. Terms assigned following the decimal are related in a hierarchical fashion to the preceding character position.

ICD-10-CM is available only from NCHS,<sup>33</sup> with private branded information now essentially just repackaging. While many files are provided with the distribution package there are two of vital importance (found in the section 'Tabular—zip'): 10tab2009.pdf—a 2369-page description of the structure, rules, headings and approximately one-half the codes; and 10-cm\_desc.txt—a file containing the 68 106 terminal-node codes of the 2009 version. Note that there is not a file with an easily computer-readable form of the rules or section-header concepts, nor is there a file with a complete collection of header, notes, descriptive information and terminal-node codes.

### ICD-10-PCS BACKGROUND

ICD-10 procedure coding system (PCS) is a revolutionary change from the current ICD-9-CM volume 3 codes for procedures.<sup>34</sup> It is maintained by the Centers for Medicare and Medicaid Services (CMS) with outside contractual support.<sup>35</sup> Like ICD-10-CM the Coordination and Maintenance Committee provides guidance regarding changes. In the late 1990s CMS quickly realized that the code space design of ICD-9-CM volume 3 would not allow expansion to include the rapidly evolving world of new medical procedures. In addition, they saw that a flexible organization of concepts was needed to accommodate unexpected changes in specifics such as mechanism and approach, as orderly placing of new developments such as those in interventional radiology and robotics was difficult.

ICD-10-PCS provides a multi-axial design to the codes, which, like ICD-10-CM, are intelligent in design. Each seven-character alphanumeric code has a section corresponding to subsections of a descriptive axis such as anatomy or surgical approach.<sup>34</sup> Only the designation of the first character is fixed to a defined section. The other six characters have their roles assigned depending on the preceding character. For example, for a code in section 0 (medical and surgical) the codes that follow are body system, root operation, body part, approach, device and qualifier with an actual set of codes for each succeeding character fixed by the preceding one. Another section could have a different defined order. The flexibility of this design allows for easy expression of the diverse world of medical procedures and, after a short learning curve, an orderly way for humans to find the codes.

ICD-10-PCS is available from CMS and the National Library of Medicine Unified Medical Language System<sup>36</sup> with private branded information now essentially just repackaging.<sup>37</sup> The distribution files are limited, with three of vital importance: UpdatedPcs2009\_tabidx.pdf—a 961-page list of the axial structure for each section and terms used by ICD-10-PCS for each axis; pcs2009\_ldsc.txt—a file containing the 72 589 terminal-node codes of the 2009 version with the descriptions as human understandable text; and pcs2009\_tdsc.txt that has the same 72 589 codes only with the descriptions in a bar delimited format of each axis. As an example of differences in the two files, the first ICD-10-PCS code listed is 0016070 and is described in the first text file as 'bypass cerebral ventricle to nasopharynx with autologous tissue substitute, open approach' and in the second as '|0 medical and surgical|0 central nervous system|1 bypass|6 cerebral ventricle|0 open|7 autologous tissue substitute|0 nasopharynx.' While ICD-10-CM actually defines and describes detail header codes for all parts, ICD-10-PCS has no defined header structure outside of the sections.

<sup>i</sup>The WHO version of ICD-10 is referred to in this paper as the 'international version' to distinguish it from the USA developed clinical modification.

<sup>ii</sup>The Coordination and Maintenance Committee consists of members appointed by NCHS, CMS, the American Hospital Association and the American Health Information Management Association.

## MATERIALS AND METHODS

A ICD-10-CM analytical file was prepared using the codes and descriptions in 10cm\_desc.txt as a base. The ICD-10-CM code found in this file is without a decimal in the fourth position, a common short form used for ICD codes, but since the content to the right and left of the decimal has different meanings the decimal was restored. As noted above, 10cm\_desc.txt only contains the terminal node codes. Header codes from 10tab2009.pdf were inserted manually for chapters A–R. The S through Z chapters, where approximately half the codes reside, used Perl scripting on text extracted from the pdf-formatted file facilitated by hand-addition to add header codes. No rules or other descriptions were placed in this file. The hand-addition of headers undoubtedly brought undetected proofreading style errors at a level equivalent to those observed in the distributed files and were not deemed significant for this analysis.

The ICD-10-PCS analytical file was prepared using the codes and descriptions in pcs2009\_ldsc.txt as a base. Header content was added by hand and consisted of the section header and designations for the next two content areas. That file was prepared using the content in UpdatedPcs2009\_tabidx.pdf also facilitated by Perl scripting.

Corresponding ICD-9-CM analytical files were prepared using the government-provided files Dtab09.rtf and required only minor augmentation with header codes. ICD-9 volume 3, used for indirect comparison with ICD-10-PCS, was found in Ptab09.rtf again with only minor augmentation of header codes.

The mortality code international version (ICD-10) analytical file was prepared from publicly available information released by WHO for the 2007 version.

Government-developed non-use-case general equivalence mappings are provided for ICD-9-CM to/from ICD-10-CM<sup>33</sup> and ICD-9 volume 3 to/from ICD-10-PCS.<sup>37 38</sup> These were converted into database form and the government interpretations of the validity of the map (exact, approximate, no map, and combination) were used to derive mapping inferences. Combination is a government-derived term used to indicate that multiple terms are need for a map and that the exactness may be situational.

Analytical files were maintained in a FileMaker Pro 10 database with each table representing a code set version and each record a term in that code set. The comparison statistics are based on the sectional counts of terms within a version derived from their respective databases. All comparison calculations were done using Microsoft Excel spreadsheets and were automatically maintained by formula relationships between cells.

## OBSERVATIONS

### Specific to ICD-10-CM

The delimited file containing the terminal codes for ICD-10-CM used for this analysis contains 68 106 codes, and released information uses this number as the term count for the classification. The printed version of ICD-10-CM has organizational header codes that convey the designers' thoughts on the structure of the classification, allows people to do organized searches from printed material and provides a 'human readable' structure to the code as noted in the enabling federal regulations.<sup>6 39</sup> These header codes were added to the released code set bringing the total codes used for this analysis to 89 576. For discussion this code set was then compared with a similar list for ICD-9-CM and the international version (table 1).

The ICD-CM classification can rationally divide into two sections. Chapters I–XVIII (disease) are assigned to specific

disease sections or general clinical symptoms (chapter XVIII). Chapters XIX–XXII (non-disease) refer to injuries, their causes and general terms regarding health status and services. Codes in the disease chapters comprise 64.9% of ICD-9-CM, 82.8% of the international version, and only 29.6% of ICD-10-CM.

Formulaic expansion of concepts is a common feature in the international version and ICD-10-CM. The brief excerpt from a section of the M chapter from ICD-10-CM (figure 1) clearly illustrates the repeated use of reoccurring phrases throughout the section. Note that M90.51 through M90.57 only differ in the anatomical location at the term end and the codes under each only by laterality. Application of this type of formula phrase substitution, which can easily be done on a word-processor, occurs throughout both versions. The largest impacts in the international version are found in chapters XIII and XX. For example, in chapter XIII (diseases of the musculoskeletal system and connective tissue) 10 additional concepts (multiple sites,<sup>iii</sup> shoulder region, upper arm, forearm, hand, pelvic region and thigh, lower leg, ankle and foot, other, site unspecified) are created for all fourth level concepts found in approximately 37 major subsections, in excess of 3000 formula concepts. Chapter XXI (external causes of morbidity and mortality) and essentially all major subsections in V, W, X and Y have formulaic increases using several different blocks generally with 10 terms/block. ICD-10-CM shares many of these auto-expansion codes, although some local variation applies, and expands it in the chapters VII (eye), VIII (ear) and XIII (muscle) through the introduction of laterality for each subcode. The greatest formula increase in ICD-10-CM occurs in chapter XIX where an alpha character is added as the last code to essentially expand all S (injury) subcodes by between three and six formula terms (or more in rare cases) to further define treatment stage.

Coding for laterality was a stated goal of the expansion of ICD-9-CM to ICD-10-CM. Terms containing the term 'right' comprise 16.5% of the code set, whereas 'left' is found in 16.7%, bilateral in 0.9%, and 'unspecified,' which is also used outside of laterality concepts, occurs in 29.9% of code set.

A significant observation is the increasing raw counts of codes in the progression from ICD-9-CM to ICD-10-CM. Figure 2 compares the relative increase in the number of terms using ICD-9-CM as a base. For most chapters there is a modest and progressive increase in the number of terms. A dramatic increase is observed in chapter XIII, diseases of the musculoskeletal system and connective tissue, when comparing ICD-9-CM with the international version, due to the formulaic increase in site terms as described above. When comparing the international version to ICD-10-CM we see large increases in chapters VII, VIII and also XIII due to the introduction of laterality concepts. Only a modest increase in terms is noted in non-disease chapters from ICD-9-CM to the international version but a huge increase is observed when moving from the international version to ICD-10-CM, again due to the formulaic introduction of expansion codes. Overall the international version increased 1.3 and ICD-10-CM 7.0 times when compared with ICD-9-CM. For the disease chapters the international version increase was 1.6 times while the non-disease chapters actually decreased 0.6 times when compared with ICD-9-CM. When compared with ICD-9-CM the disease chapters increased 3.2 times while the growth in the non-disease chapters was 14.1 times.

<sup>iii</sup>Specific anatomical content for these areas is described in the ICD-10 coding instruction.

**Table 1** Percentage distribution of chapters in ICD-9-CM, ICD-10, ICD-10-CM

| Chapter | ICD-10-CM Block | Title   | % of classification |        |           |
|---------|-----------------|---|---------------------|--------|-----------|
|         |                 |   | ICD-9-CM            | ICD-10 | ICD-10-CM |
| I       | A00–B99         | Certain infectious and parasitic diseases   | 8.3                 | 5.4    | 1.4       |
| II      | C00–D48         | Neoplasms   | 6.1                 | 5.3    | 2.0       |
| III     | D50–D89         | Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism | 0.9                 | 1.2    | 0.3       |
| IV      | E00–E90         | Endocrine, nutritional and metabolic diseases   | 2.4                 | 2.6    | 1.0       |
| V       | F00–F99         | Mental and behavioural disorders  | 2.9                 | 3.0    | 1.0       |
| VI      | G00–G99         | Diseases of the nervous system  | 3.6*                | 2.4    | 0.8       |
| VII     | H00–H59         | Diseases of the eye and adnexa  | 6.9                 | 2.0    | 3.2       |
| VIII    | H60–H95         | Diseases of the ear and mastoid process   | 1.7                 | 0.8    | 1.0       |
| IX      | I00–I99         | Diseases of the circulatory system  | 3.9                 | 2.9    | 1.6       |
| X       | J00–J99         | Diseases of the respiratory system  | 2.1                 | 1.8    | 0.4       |
| XI      | K00–K93         | Diseases of the digestive system  | 4.9                 | 3.1    | 1.0       |
| XII     | L00–L99         | Diseases of the skin and subcutaneous tissue  | 1.9                 | 2.5    | 1.0       |
| XIII    | M00–M99         | Diseases of the musculoskeletal system and connective tissue  | 1.9                 | 34.0   | 8.8       |
| XIV     | N00–N99         | Diseases of the genitourinary system  | 3.6                 | 3.1    | 0.8       |
| XV      | O00–O99         | Pregnancy, childbirth and the puerperium  | 3.0                 | 3.0    | 2.8       |
| XVI     | P00–P96         | Certain conditions originating in the perinatal period  | 4.9                 | 2.4    | 0.6       |
| XVII    | Q00–Q99         | Congenital malformations, deformations and chromosomal abnormalities                                | 3.6                 | 4.4    | 1.0       |
| XVIII   | R00–R99         | Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified             | 2.8                 | 2.4    | 0.8       |
| XIX     | S00–T98         | Injury, poisoning and certain other consequences of external causes                                 | 15.4                | 4.8    | 58.3      |
| XX      | V01–Y98         | External causes of morbidity and mortality  | 7.3                 | 7.9    | 10.6      |
| XXI     | Z00–Z99         | Factors influencing health status and contact with health services                                  | 9.5                 | 4.5    | 1.5       |
| XXII    | U00–U99         | Codes for special purposes  | 0.0                 | 0.1    | 0.0       |

\*ICD-9-CM includes concepts relating to eye and ear in this chapter. Concept counts for eye and ear were manually determined for this comparison.

It is apparent from just the chapter numbers that the development of ICD-10-CM stems directly from the international version. To test this observation the top-level terms, those that have no digit to the right of the decimal, were compared item by item for exact matches. Of the 1397 top-level claims in the disease chapters, 83.5% had the same code and 77.1% had the same code and description, ignoring European versus American spelling. For the non-disease chapters, 80.0% had the same code but just over half, 53.8%, had identical code and description. Table 2 shows the type of differences for the disease and non-disease chapters. Many of the ICD-10-CM codes lacking an international version code represented the concept of the condition not being part of the set in this code block but classified elsewhere, a style not used greatly by the international version designers. The increased specificity of a code in one classification versus the other many times took the form of 'other (condition)' versus 'other and unspecified (condition)'. For the 'mental and behavioural disorders' and 'pregnancy, childbirth and the puerperium' chapters, however, many times these were more significant differences in concepts such as for term P05. The international version defines P05 as 'slow fetal growth and fetal malnutrition' while in the ICD-10-CM the concept is 'disorders of newborn related to slow fetal growth and fetal malnutrition'. For the non-disease chapters fewer codes were missing and, of those that were, no real pattern is observed. Most of the other differences were similar to the disease chapter except for a number of 'concept differ' terms. Many of these

were due to deprecation of the term in ICD-10-CM, replacing it with the corresponding term generally in the poison section (T). As the term remained (not retired) and indicated replacement, it was deemed different.

### Specific to ICD-10-PCS

The delimited file containing the terminal codes for ICD-10-PCS used for this analysis contains 72 589 codes, and released information uses this number as the term count for the classification. In the process of creating the analytical file, headers were created to the third level of the term ID provided in pcs2009\_tdsc.txt to convey the designers' thoughts on structure and allow people to do organized searches from printed material by providing a 'human readable' structure to the code as noted in the enabling federal regulations. These derived header codes were added to the released code set bringing the total codes used for this analysis to 73 680. Table 3 lists the ICD-10-PCS chapter designations and corresponding code distribution.

As CMS chose to re-develop a procedure classification, direct comparison with volume 3 is difficult. The major chapter, medical and surgical, comprises almost 90% of the codes and has a relationship similar to, but not exactly that of, ICD-9 volume 3 (table 4). Outside of that loose comparison one can only note that the increase in size between the two classifications is 18.0 times. One reason for the increase is the wide use of

**Figure 1** Excerpt from Chapter M Code construction.

**M90.5 Osteonecrosis in diseases classified elsewhere**  
 Code first underlying disease, such as: caisson disease (T70.3) hemoglobinopathy (D50-D64)  
**M90.50 Osteonecrosis in diseases classified elsewhere, unspecified site**  
**M90.51 Osteonecrosis in diseases classified elsewhere, shoulder**  
     **M90.511 Osteonecrosis in diseases classified elsewhere, right shoulder**  
     **M90.512 Osteonecrosis in diseases classified elsewhere, left shoulder**  
     **M90.519 Osteonecrosis in diseases classified elsewhere, unspecified shoulder**  
**M90.52 Osteonecrosis in diseases classified elsewhere, upper arm**  
     **M90.521 Osteonecrosis in diseases classified elsewhere, right upper arm**  
     **M90.522 Osteonecrosis in diseases classified elsewhere, left upper arm**  
     **M90.529 Osteonecrosis in diseases classified elsewhere, unspecified upper arm**  
**M90.53 Osteonecrosis in diseases classified elsewhere, forearm**  
     **M90.531 Osteonecrosis in diseases classified elsewhere, right forearm**  
     **M90.532 Osteonecrosis in diseases classified elsewhere, left forearm**  
     **M90.539 Osteonecrosis in diseases classified elsewhere, unspecified forearm**  
**M90.54 Osteonecrosis in diseases classified elsewhere, hand**  
     **M90.541 Osteonecrosis in diseases classified elsewhere, right hand**  
     **M90.542 Osteonecrosis in diseases classified elsewhere, left hand**  
     **M90.549 Osteonecrosis in diseases classified elsewhere, unspecified hand**  
**M90.55 Osteonecrosis in diseases classified elsewhere, thigh**  
     **M90.551 Osteonecrosis in diseases classified elsewhere, right thigh**  
     **M90.552 Osteonecrosis in diseases classified elsewhere, left thigh**  
     **M90.559 Osteonecrosis in diseases classified elsewhere, unspecified thigh**  
**M90.56 Osteonecrosis in diseases classified elsewhere, lower leg**  
     **M90.561 Osteonecrosis in diseases classified elsewhere, right lower left**  
     **M90.562 Osteonecrosis in diseases classified elsewhere, left lower leg**  
     **M90.569 Osteonecrosis in diseases classified elsewhere, unspecified lower leg**  
**M90.57 Osteonecrosis in diseases classified elsewhere, ankle and foot**  
     **M90.571 Osteonecrosis in diseases classified elsewhere, right ankle and foot**  
     **M90.572 Osteonecrosis in diseases classified elsewhere, left ankle and foot**  
     **M90.579 Osteonecrosis in diseases classified elsewhere, unspecified ankle and foot**  
**M90.58 Osteonecrosis in diseases classified elsewhere, other site**  
**M90.59 Osteonecrosis in diseases classified elsewhere, multiple sites**

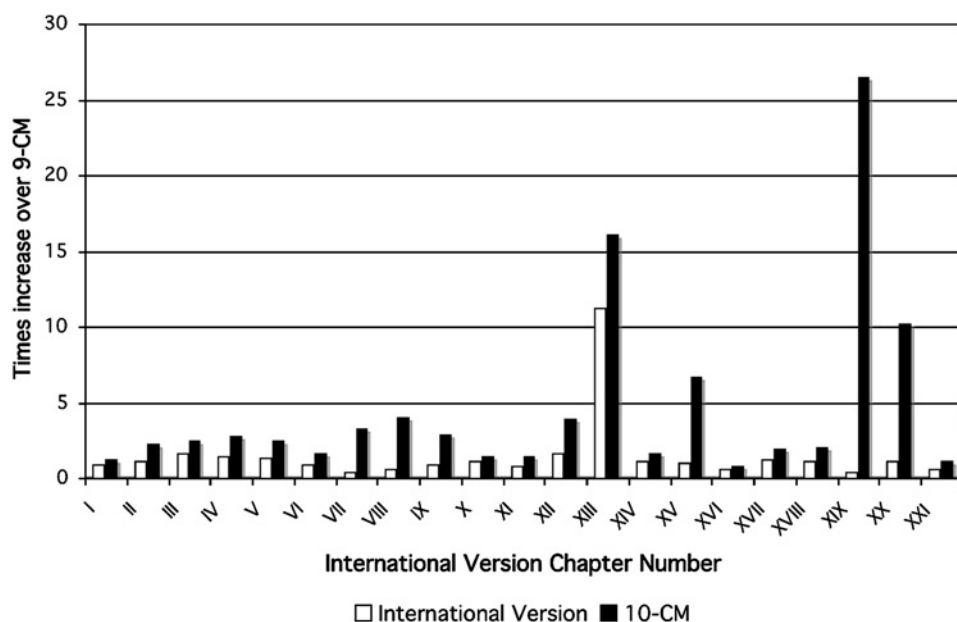
laterality with the expression 'left' or 'right' appearing in 27.1% and 27.3% of the classifications, respectively. Unlike ICD-10-CM the term 'unspecified' is not used and 'bilateral' occurs in only 2.2% of the descriptions. Another reason is the increase in chapters for administration, imaging, radiation oncology, and physical rehabilitation and diagnostic audiology, each amounting to between approximately 2% and 3% of the terms, where ICD-9 volume 3 space considerations previously prevented adequate coverage.

**CHALLENGES**

**Intelligent versus non-intelligent codes**

ICD-10-CM and ICD-10-PCS still have intelligent codes. In ICD-10-CM they serve to limit the number of available codes due to previous choices as evidenced by the coding pattern nnn (condition); nnn.1 (specific condition 1); nnn.2 (specific condition 2);... nnn.8 (other specific conditions); nnn.81...n (other specific condition 1...n); nnn.9 (unspecified condition). The coding pattern appears in numerous places with no evidence

**Figure 2** Relative increase in terms of the international version, ICD-10-CM versus ICD-9-CM.



**Table 2** Reasons for top-level concept differences

| Error type                              | Disease |      | Non-disease |       |
|---|---------|------|-------------|-------|
|   | N       | %    | N           | %     |
| Concepts semantically same              | 22      | 1.6% | 48          | 7.2%  |
| Concepts differ—ICD-10-CM more specific | 55      | 3.9% | 44          | 6.6%  |
| Concepts differ—ICD-10 more specific    | 70      | 5.0% | 11          | 1.7%  |
| Concepts differ                         | 10      | 0.7% | 50          | 7.5%  |
| ICD-10-CM term missing                  | 48      | 3.4% | 18          | 2.7%  |
| ICD-10 term missing                     | 93      | 6.7% | 106         | 15.9% |

that those concepts appearing in the nnn.8 section are different from the ones above, except that no free code room exists. ICD-10-CM has already added a subchapter code that is non-numeric (O9A: maternal malignant neoplasms, traumatic injuries and abuse classifiable elsewhere but complicating pregnancy, childbirth and the puerperium) so flexibility does exist assuming we have a willingness to not maintain a strict relationship to the international version and to modify the rules to meet observed difficulties. Also, many sections of ICD-10-CM do not give codes for the intended 'other' conditions but have only example lists in the printed notes. It is unclear if the intelligence in the ICD-10-PCS codes will suffer from the same limitation.

### Formulaic post-coordinated terms

Although widely used for other purposes, ICD-10-CM and ICD-10-PCS are classifications intended primarily for valid statistical aggregation of the healthcare system.<sup>23</sup> ICD-10-CM has some atomic codes, generally as section headers, but few, outside of infectious organisms, in the list of terminal-node (reportable) codes. The post-coordinated nature of ICD-10-CM codes is evident from the gross construction pattern in many codes: (condition) | (comorbidity 1...n) | (modifier, like laterality, if needed). For the non-disease sections the pattern remains where condition is replaced by injury or event and comorbidity is an equivalent modifier. Use of the formulaic pattern allows for the vast addition of post-coordinated terms to the classifications and may serve as a means to improve code selection through novel

**Table 3** ICD-10-PCS chapters

| Chapter | PCS code | Title  | N     | %     |
|---------|----------|--|-------|-------|
| 1       | 0        | Medical and surgical                             | 63508 | 86.23 |
| 2       | 1        | Obstetrics                                       | 351   | 0.48  |
| 3       | 2        | Placement  | 877   | 1.19  |
| 4       | 3        | Administration                                   | 1418  | 1.93  |
| 5       | 4        | Measurement and monitoring                       | 316   | 0.43  |
| 6       | 5        | Extracorporeal assistance and performance        | 46    | 0.06  |
| 7       | 6        | Extracorporeal therapies                         | 54    | 0.07  |
| 8       | 7        | Osteopathic                                      | 103   | 0.14  |
| 9       | 8        | Other procedures                                 | 59    | 0.08  |
| 10      | 9        | Chiropractic                                     | 93    | 0.13  |
| 11      | B        | Imaging  | 2777  | 3.77  |
| 12      | C        | Nuclear medicine                                 | 584   | 0.79  |
| 13      | D        | Radiation oncology                               | 1953  | 2.65  |
| 14      | F        | Physical rehabilitation and diagnostic audiology | 1399  | 1.90  |
| 15      | G        | Mental health                                    | 44    | 0.06  |
| 16      | H        | Substance abuse treatment                        | 68    | 0.09  |

**Table 4** Correspondence between ICD-9 volume 3 and medical and surgical chapter of ICD-10-PCS

| Chapter | Title   | Vol 3% | Est PCS %* |
|---------|---|--------|------------|
| 1       | Procedures and interventions, not elsewhere classified (00)     | 2.1    |            |
| 2       | Operations on the nervous system (01–05)                        | 4.1    | 2.7        |
| 3       | Operations on the endocrine system (06–07)                      | 2.2    | 1.2        |
| 4       | Operations on the eye (08–16)                                   | 9.1    | 1.6        |
| 5       | Other miscellaneous diagnostics and therapeutic procedures (17) | 0.6    |            |
| 6       | Operations on the ear (18–20)                                   | 2.1    | 2.3        |
| 7       | Operations on the nose, mouth, and pharynx (21–29)              | 5.6    | 2.1        |
| 8       | Operations on the respiratory system (30–34)                    | 4.0    | 3.0        |
| 9       | Operations on the cardiovascular system (35–39)                 | 6.0    | 2.8        |
| 10      | Operations on the hemic and lymphatic system (40–41)            | 1.5    | 1.9        |
| 11      | Operations on the digestive system (42–54)                      | 14.3   | 5.9        |
| 12      | Operations on the urinary system (55–59)                        | 4.8    | 2.7        |
| 13      | Operations on the male genital organs (60–64)                   | 3.2    | 2.1        |
| 14      | Operations on the female genital organs (65–71)                 | 6.2    | 1.9        |
| 15      | Obstetrical procedures (72–75)                                  | 2.0    | 0.5        |
| 16      | Operations on the musculoskeletal system (76–84)                | 10.4   | 22.7       |
| 17      | Operations on the integumentary system (85–86)                  | 3.0    | 3.6        |
| 18      | Miscellaneous diagnostic and therapeutic procedures (87–99)     | 18.6   |            |

\*Correspondence of the medical and surgical chapter of PCS approximates but does not duplicate volume 3. Percentages are best on best matches between the two classification code contents.

guidance tools, either on paper or built into the next-generation clinical and practice management systems required to service HIPAA version 2. ICD-10-PCS codes are multi-axial and were never intended to be atomic.

### Member-of relationships

Classifications inherently have a hierarchy that differs from what is used in ontological construction. A classification hierarchy is needed for aggregation and not to define inheritance. Hence the hierarchical relationship is not (parent–child) but more (member-of). ICD-10-CM clearly indicates how the relationship of terms differs from ontology's with exclusion rules. The first exclusion rule specifically says that a similar term, in many cases the same condition but related to a specific cause such as 'drug induced' or in a specific state such as pregnancy or childhood, is excluded from the general classification. Subsumption now is more difficult. The second broad exclusion rule is not related to the logical structure but is a restriction on use. Basically, that rule prevents the co-reporting of codes for similar conditions that the patient has but are not related to the

present incidence, and the logical relationship between the two remains.

ICD-10-PCS has no hierarchical terms defined by the authors, although they do have an index in their released information. Lack of a formal relationship between the codes prevents logical determination of duplication in this vast assemblage of procedures. While many classifications grow as they are used, adding missing or new concepts, ICD-10-PCS has contracted, indicating that automatic detection of duplication in versions is perhaps difficult. The 1996 version of ICD-10-PCS had 197 769 codes;<sup>40</sup> in the 2008 proposed rule the number appears as approximately 87 000 codes;<sup>6</sup> and the current file has just over 72 000 codes.<sup>37</sup> While the developers of ICD-10-PCS are commended for their novel approach toward solving the complex problem of a vast number of ever-changing procedures, the success of the classification will depend on an introduction of a formalism to allow orderly growth and elimination/prevention of duplicate concepts.

### Validation of codes and testing

ICD-10-CM was essentially developed in the mid-1990s<sup>30</sup> and successfully field tested in 2003<sup>41</sup> but has not been formally retested since that time. The community now has approximately 4 years before required use, and the time can best be used to improve coverage and potential submission of concepts designed for multiple uses before legacy codes start populating databases. One example is diabetes coverage, which has expanded from 29 codes containing the phrase in ICD-9-CM to 345 in ICD-10-CM, but with few codes, outside of gestational diabetes, indicating control level or the agent used. Adding these types of concepts appears necessary and easy.

Diseases of the musculoskeletal system and connective tissue (chapter M) comprise 8.8% of the ICD-10-CM code set (table 1) and is primarily composed of a formulaic group of approximately 30 codes of introductory clinical condition, anatomical section and laterality repeats for most of the approximately 260 code groups in this chapter (figure 1). Questions arise as to the anatomical specificity of the terms used; in this example shoulder, upper arm, hand, thigh, lower leg, ankle and foot, and hip/femur, compared with the precision practitioners' use in describing similar regions and the precision used in the injury (S) chapter (eg, S32.810A: multiple fractures of pelvis with stable disruption of pelvic circle, initial encounter for closed fracture). The international version uses a similar shorthand description for anatomical sites and provides more specific definitions for the regions encompassed.<sup>42</sup>

ICD-10-PCS is totally new and has had only limited, although very successful, field testing.<sup>43</sup> More extensive validation and testing of the classification should occur before 2013.

### Code reuse

There is tension that is international in scope in coding between the provider incentives to use codes for payment versus the broader healthcare community's interest in having correct detail in the codes so they represent what the care encounter involves.<sup>44-46</sup> One study shows that ICD-10-CM on initial use does not perform significantly better than ICD-9-CM in capturing clinical detail.<sup>47</sup> In the USA providers tend to use these codes only to support payment and to help the information base of the healthcare system. Successful return on the \$30+ billion that US healthcare will invest in revamping our health information technology system is dependent on our ability to input accurate codes that have validity when reused.

The information model of both ICD-10-CM and ICD-10-PCS provides a structure that allows the isolation of detailed level codes through a stepwise selection process that mimics provider data capture. As experts in content as it applies to workflow, the informatics community should focus on helping to develop tools that better link the terminology content of information models used at various process stages (such as history to problem list to orders to billing) during the next 4 years. An example of bad practice would be the conversion of the present ICD-10-CM versions of the common superbill used in providers' offices, or their electronic system equivalence, that capture codes at the point of care as they contain ICD-9-CM codes<sup>48</sup> or ICD-10-CM codes that are not at the reporting (detail) level.<sup>49</sup>

We need to study the coding requirements and develop easily used systems, both paper and computer, that help guide one to the correct detail code, especially when dealing with the vast increase in injury codes and the complexity associated with their use. As an example, consider the coding complexity of a person with osteoporosis falling and breaking a wrist. Most likely the person will be treated in an emergency department or other urgent care facility that has no knowledge of the person's underlying condition and has the encounter coded as S62.182A: displaced fracture of trapezoid [smaller multangular], left wrist, initial encounter for closed fracture. While the S codes provide for follow-up care, that person will most likely get that care from a specialist. Current advice is for the specialist to use codes from the M chapter and most likely code subsequent encounters using M80.042D: age-related osteoporosis with current pathological fracture, left hand, subsequent encounter for fracture with routine healing. How can we coordinate both care locations to the correct sequence of codes and maintain the care model? Does this sequence of codes provide the best information for this common occurrence? To allow the successful flow of clinical observation to administrative systems to healthcare decisions, the informatics community needs to help shape this dialogue as it represents a new role of health information exchange.

### Mapping to previous versions

Bidirectional mapping of the ICD-9 and ICD-10 versions is essential for a smooth transition. The government-developed general equivalence mappings are provided for ICD-9-CM to/from ICD-10-CM<sup>33</sup> and ICD-9 volume 3 to/from ICD-10-PCS,<sup>37 38</sup> but are not particularly helpful in that more than 95% of codes are noted as having only an approximate match (table 5). Review for use-case dependent acceptance will show, as with the high-level match of the international version to ICD-10-CM discussed under observations (table 2), that many will be acceptable. Support for this observation is found in the use-case defined maps for ICD-9 to/from ICD-10 for the Medicare diagnostic-related groups,<sup>50</sup> in which the identity match is currently at 99% for clinical terms and 91% for procedures, with an overall acceptance of 95%. CMS will maintain this map for 3 years in a revenue-neutral fashion.<sup>6</sup>

### CONCLUSIONS

Clearly the US terminology situation is complex and the above-noted issues related to ICD-10-CM and ICD-10-PCS only add to this complexity. ICD-10-PCS can only be viewed as a work in progress that has great potential to solve the ever-changing world of new procedures but must solve the problem of identifying and preventing overlapping concepts. ICD-10-CM gains and suffers from its tie to the international version. The rigid structure of the international version, while much improved from ICD-9, forces an organization that may not be most

**Table 5** Percentage general equivalence mappings match-type

| Match type                  | ICD-9-CM to ICD-10-CM | ICD-10-CM to ICD-9-CM | ICD-9 volume 3 to ICD-10-PCS | ICD-10-PCS to ICD-9 volume 3 |
|-----------------------------|-----------------------|-----------------------|------------------------------|------------------------------|
| Exact                       | 10.4%                 | 4.5%                  | 0.1%                         | 0.1%                         |
| Approximate                 | 89.6%                 | 95.5%                 | 97.9%                        | 99.9%                        |
| No map                      | 1.2%                  | 1.0%                  | 0.2%                         | 0.0%                         |
| Combination*                | 6.4%                  | 12.2%                 | 2.2%                         | 6.7%                         |
| Approximate and combination | 6.4%                  | 12.0%                 | 2.2%                         | 6.7%                         |
| Approximate and no map      | 1.2%                  | 1.0%                  | 0.2%                         | 0.0%                         |

\*Combination indicates the link is to more than one code in the target system and the mapping depends on use-case and cannot be assigned outside of one.

flexible with respect to US health care. The inability to have true subsumption of like medical concepts is one indication of the problem. The way the structure forces complex coding choices is another. Despite these issues, ICD-10-CM, with some careful review regarding missing concepts, especially in the area of quality, appears close to meeting the needs of healthcare reporting for reimbursement and many secondary uses, including the one it was designed for: statistics.

The vast expansion of the number of terms in ICD-10-CM and ICD-10-PCS presents many challenges for the community. Human interface design issues need addressing in making sure the correct code is selected to represent the clinical state (which of the 10× increase in diabetes codes is most accurate?), the staging of the disease (initial, sequela, routine healing), or laterality. Fortunately, the formulaic design of ICD-10-CM and ICD-10-PCS provides much opportunity for new and creative ‘superbills’ with and without computer assistance. Central coordination is essential and is needed to better direct the introduction and changes in classification that will soon replace that used by essentially all US healthcare providers for clinical decision-making.

**Competing interests** None.

**Provenance and peer review** Not commissioned; externally peer reviewed.

## REFERENCES

1. Anon. Executive order 13335—Incentives for the use of health information technology and establishing the position of the national health information technology coordinator. *Fed Regist* 2004;**69**:24059–61.
2. US Public Law. American recovery and reinvestment act of 2009. *Public Law* 2009:111–15.
3. Office of the national coordinator H.H.S. Meaningful use 2009. <http://healthit.hhs.gov/portal/server.pt?open=512&objID=1325&parentname=CommunityPage&parentid=1&mode=2> (accessed 24 Aug 2009).
4. US Public Law. Health insurance portability and accountability act of 1996. *Public Law* 1996:104–91.
5. Office of the Secretary HHS. Health insurance reform; modifications to the Health Insurance Portability and Accountability Act (HIPAA) electronic transaction standards: 45 CFR part 162. *Fed Regist* 2009;**74**:3296–328.
6. Office of the Secretary HHS. HIPAA administrative simplification: modification to medical data code set standards to adopt ICD-10-CM and ICD-10-PCS. Proposed rule. *Fed Regist* 2008;**73**:49795–832.
7. Office of the Secretary HHS. HIPAA administrative simplification: modifications to medical data code set standards to adopt ICD-10-CM and ICD-10-PCS: 45 CFR part 162. *Fed Regist* 2009;**74**:3328–62.
8. Tang PC, Ralston M, Arrigotti MF, et al. Comparison of methodologies for calculating quality measures based on administrative data versus clinical data from an electronic health record system: implications for performance measures. *J Am Med Inform Assoc* 2007;**14**:10–15.
9. Chute CG. Clinical classification and terminology: some history and current observations. *J Am Med Inform Assoc* 2000;**7**:298–303.
10. Chute CG, Cohn SP, Campbell JR. A framework for comprehensive health terminology systems in the United States: development guidelines, criteria for selection, and public policy implications. ANSI healthcare informatics standards board vocabulary working group and the computer-based patient records institute working group on codes and structures. *J Am Med Inform Assoc* 1998;**5**:503–10.

11. Cimino JJ, Clayton PD. Coping with changing controlled vocabularies. *Proc Annu Symp Comput Appl Med Care* 1994:135–9.
12. Cimino JJ. Desiderata for controlled medical vocabularies in the twenty-first century. *Methods Inf Med* 1998;**37**:394–403.
13. Brown SH, Husser CS, Wahner-Roedler D, et al. Using SNOMED CT as a reference terminology to cross map two highly pre-coordinated classification systems. *Stud Health Technol Inform* 2007;**129**:636–9.
14. Rector AL. Clinical terminology: why is it so hard? *Methods Inf Med* 1999;**38**:239–52.
15. Schulz EB, Price C, Brown PJ. Symbolic anatomic knowledge representation in the read codes version 3: structure and application. *J Am Med Inform Assoc* 1997;**4**:38–48.
16. Spackman KA, Campbell KE, Cote RA. Snomed RT: a reference terminology for health care. *Proc AMIA Annu Fall Symp* 1997:640–4.
17. Wang AY, Barrett JW, Bentley T, et al. Mapping between SNOMED RT and clinical terms version 3: a key component of the SNOMED CT development process. *Proc AMIA Symp* 2001:741–5.
18. Rector AL, Nowlan WA. The GALEN project. *Comput Methods Programs Biomed* 1994;**45**:75–8.
19. Rogers JE, Price C, Rector AL, et al. Validating clinical terminology structures: integration and cross-validation of read thesaurus and GALEN. *Proc AMIA Symp* 1998:845–9.
20. National Health Service. SNOMED clinical terms. <http://www.connectingforhealth.nhs.uk/systemsandservices/data/snomed> (accessed 30 Aug 2009).
21. International Health Terminology Standards Development Organisation. History of IHTSDO. <http://www.ihtsdo.org/about-ihtsdo/history-of-ihtsdo/> (accessed 30 Aug 2009).
22. Heja G, Surjan G, Varga P. Ontological analysis of SNOMED CT. *BMC Med Inform Decis Mak* 2008;**8**(Suppl 1):S8.
23. Jiang G, Chute CG. Auditing the semantic completeness of SNOMED CT using formal concept analysis. *J Am Med Inform Assoc* 2009;**16**:89–102.
24. Rector AL, Brandt S. Why do it the hard way? The case for an expressive description logic for SNOMED. *J Am Med Inform Assoc* 2008;**15**:744–51.
25. National Library of Medicine. SNOMED clinical terms® (SNOMED CT®). 2009. [http://www.nlm.nih.gov/research/umls/Snomed/snomed\\_main.html](http://www.nlm.nih.gov/research/umls/Snomed/snomed_main.html) (accessed 30 Aug 2009).
26. Wiesenthal AM. Kaiser Permanente health connect: a large scale EHR deployment using SNOMED-CT. 2007. <http://www.slideshare.net/HINZ/kaiser-permanente-healthconnect-ehr-and-snomed> (accessed 30 Aug 2009).
27. Bramer GR. International statistical classification of diseases and related health problems. Tenth revision. *World Health Stat Q* 1988;**41**:32–6.
28. Anderson RN, Rosenberg HM. Disease classification: measuring the effect of the tenth revision of the international classification of diseases on cause-of-death data in the United States. *Stat Med* 2003;**22**:1551–70.
29. Health Insurance Reform. Standards for electronic transactions. *Fed Regist* 2000;**65**:50312–72.
30. Pickett D, Berglund D, Blum A, et al. A quick review of ICD-10-CM. *J AHIMA* 1999;**70**:99–100.
31. American Health Information Management Association. Destination 10: healthcare organization preparation for ICD-10-CM and ICD-10-PCS. *J AHIMA* 2004;**75**:56A–D.
32. Schwend G. Expanding the code. The methodical switch from ICD-9-CM to ICD-10-CM will bring both challenges and rewards to healthcare. *Health Manag Technol* 2007;**28**:12, 4.
33. National Center for Health Statistics. International Classification of Diseases, tenth revision, clinical modification (ICD-10-CM), 2009 update. 2009. <http://www.cdc.gov/nchs/about/otheract/icd9/icd10cm.htm> (accessed 25 Aug 2009).
34. Averill RF, Mullin RL, Steinbeck BA, et al. Development of the ICD-10 procedure coding system (ICD-10-PCS). *Top Health Inf Manage* 2001;**21**:54–88.
35. Bowman S. Brushing up on ICD-10-PCS. *J AHIMA* 2007;**78**:108–10, 12.
36. Unified Medical Language System. 2009AA ICD-10-PCS source information. 2010. <http://www.nlm.nih.gov/research/umls/sourcereleasedocs/2009AA/ICD10PCS/index.html> (accessed 25 Feb 2010).
37. Centers for Medicare and Medicaid Services. 2009 ICD-10-PCS. 2009. [http://www.cms.hhs.gov/ICD10/01m\\_2009\\_ICD10PCS.asp#TopOfPage](http://www.cms.hhs.gov/ICD10/01m_2009_ICD10PCS.asp#TopOfPage) (accessed 26 Aug 2009).
38. Centers for Medicare and Medicaid Services. ICD-10-CM/PCS to ICD-9-CM reimbursement mappings: 2009 version — documentation and user’s guide, 2009. [http://www.cms.hhs.gov/ICD10/downloads/reimb\\_map\\_guide\\_2009.pdf](http://www.cms.hhs.gov/ICD10/downloads/reimb_map_guide_2009.pdf) (accessed 26 Aug 2009).
39. Office of the Secretary HHS. HIPAA administrative simplification: modifications to medical data code set standards to adopt ICD-10-CM and ICD-10-PCS. Final rule. *Fed Regist* 2009;**74**:3328–62.
40. Averill RF, Mullin RL, Steinbeck BA, et al. Development of the ICD-10 procedure coding system (ICD-10-PCS). *J AHIMA* 1998;**69**:65–72.
41. American Hospital Association and American Health Information Management Association. ICD-10-CM field testing project, 2003. [http://www.ahima.org/icd10/documents/FinalStudy\\_000.pdf](http://www.ahima.org/icd10/documents/FinalStudy_000.pdf) (accessed 26 Aug 2009).
42. World Health Organization. International Classification of Diseases (ICD): chapter XIII — diseases of the musculoskeletal system and connective tissue (M00–M99). 2007. <http://apps.who.int/classifications/apps/icd/icd10online/> (accessed 26 Aug 2009).
43. Brooks PE. Testing ICD-10-PCS. *J AHIMA* 1998;**69**:73–4; quiz 5–6.



44. **Dudeck J.** Changes in medical documentation over the last five decades. *Stud Health Technol Inform* 2008;**134**:255–63.
45. **Lorence DP, Ibrahim IA.** Benchmarking variation in coding accuracy across the United States. *J Health Care Finance* 2003;**29**:29–42.
46. **Quan H, Sundararajan V, Halfon P, et al.** Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. *Med Care* 2005;**43**:1130–9.
47. **Quan H, Li B, Saunders LD, et al.** Assessing validity of ICD-9-CM and ICD-10 administrative data in recording clinical conditions in a unique dually coded database. *Health Serv Res* 2008;**43**:1424–41.
48. **American Academy of Family Physicians.** Superbill template. 2009. <http://www.aafp.org/fpm/icd9/fpmsuperbill.pdf> (accessed 26 Aug 2009).
49. **American Health Information Management Association.** Sample superbill. 2009. [http://www.cms.hhs.gov/ICD10/Downloads/Sample\\_superbill\\_icd10.pdf](http://www.cms.hhs.gov/ICD10/Downloads/Sample_superbill_icd10.pdf) (accessed 26 Aug 2009).
50. **Centers for Medicare and Medicaid Services.** General equivalence mappings: ICD-9-CM to and from ICD-10-CM and ICD-10-PCS. 2009. [http://www.cms.hhs.gov/MLNProducts/downloads/ICD-10\\_GEM\\_factsheet.pdf](http://www.cms.hhs.gov/MLNProducts/downloads/ICD-10_GEM_factsheet.pdf) (accessed 27 Aug 2009).