

Entropic Measures of Complexity in a New Medical Coding System

Jerome Niyirora (✉ jerome.niyirora@gmail.com)
SUNY Polytechnic Institute

Research Article

Keywords: Medical Coding, Mapping, Complexity, Entropy

Posted Date: January 11th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-137381/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Version of Record: A version of this preprint was published at BMC Medical Informatics and Decision Making on April 9th, 2021. See the published version at <https://doi.org/10.1186/s12911-021-01485-y>.

Entropic Measures of Complexity in a New Medical Coding System

Jerome Niyirora, PhD, RHIA

SUNY Polytechnic Institute, College of Health Sciences, Utica, New York, USA

jerome.niyirora@sunypoly.edu

Abstract

Background

Transitioning from an old medical coding system to a new one can be challenging, especially when the two coding systems are significantly different. The US experienced such a transition in 2015.

Objective

This research aims to introduce entropic measures to help users prepare for the migration to a new medical coding system by identifying and focusing preparation initiatives on clinical concepts with more likelihood of adoption challenges.

Methods

Two entropic measures of coding complexity are introduced. The first measure is a function of the variation in the alphabets of new codes. The second measure is based on the possible number of valid representations of an old code.

Results

A demonstration of how to implement the proposed techniques is carried out using the 2015 mappings between ICD-9-CM and ICD-10-CM/PCS. The significance of the resulting entropic measures is discussed in the context of clinical concepts that were likely to pose challenges regarding documentation, coding errors, and longitudinal data comparisons.

Conclusion

The proposed entropic techniques are suitable to assess the complexity between any two medical coding systems where mappings or crosswalks exist. The more the entropy, the more likelihood of adoption challenges. Users can utilize the suggested techniques as a guide to prioritize training efforts to improve documentation and increase the chances of accurate coding, code validity, and longitudinal data comparisons.

Keywords: Medical Coding, Mapping, Complexity, Entropy

1. INTRODUCTION

Medical diagnoses and procedures are reported using standardized codes that are updated periodically to keep up with the latest clinical knowledge and practices. Transitioning from an old medical coding system to a new one can be challenging, especially when the two systems are significantly different. One such transition took place in the United States (US) in 2015 when the country switched from the 9th revision of the International Classification of Diseases (ICD) Clinical Modification (ICD-9-CM) to the 10th revision (ICD-10-CM). This newer revision was accompanied by a very different procedure coding system (PCS) (ICD-10-PCS), as compared to the ICD-9-CM procedure coding system (Volume 3, abbreviated here as Vol. 3). For example, each ICD-10-PCS procedure is made of 7 multi-axial characters where each axis encompasses up to 34 alphanumeric values [1]. This arrangement is a significant departure from the procedure code structure in ICD-9-CM Vol. 3, where all codes are numeric and can only be between 2 and 4 characters long. In 2015, ICD-10-PCS had about 72,000 procedure codes as compared to only about 4,000 codes in ICD-9-CM Vol. 3. The diagnosis codes between these two revisions of ICD are also quite different. For example, all diagnosis codes in ICD-10-CM are alphanumeric and can be 3 to 7 characters long, whereas ICD-9-CM diagnosis codes are mostly numeric and can only be between 3 and 5 characters long. In 2015, there were about 14,500 diagnosis codes in ICD-9-CM as compared to about 69,800 codes in ICD-10-CM [2]. Given these differences, some analysts had predicted a costly and challenging transition from ICD-9-CM to ICD-10-CM/PCS [3]. Indeed, some of the feared problems did materialize after the changeover, such as the loss in productivity [4, 5], the lack of readiness of computer systems, the inability to find some ICD-9-CM concepts in the ICD-10-CM system, and difficulties mapping ICD-10-CM to other coding systems such as SNOMED-CT [6]. Some ICD-10-CM clinical classes were also found to have more coding deficiencies than others, such as the class of external causes of morbidity (V00-Y99) [7]. In one post-ICD-10 implementation audit, it was found that one of the most significant challenges for coders was selecting the correct character in the 3rd position (Root Operation), the 4th position (Body Part), and the 5th position (Approach) of an ICD-10-PCS code [8]. While little evidence exists to suggest that reimbursement was significantly impacted by the transition, in some practices, a statistical increase in the coding-related denials was noted [9]. A few of the post-transition qualitative studies concluded that training and education were critical in overcoming many of the previously anticipated challenges [6, 10]. Besides the US, other countries have also faced challenges while transitioning to new medical coding systems. The issues ranged from coding errors to discrepancy problems when the same condition was coded in both coding systems. For example, in one analysis [11], it was found that the Swiss transition from ICD-9 to ICD-10 resulted in the initial increase of the number of coding errors for comorbidities, but, over time, the accuracy improved as the learning curve waned. In one Canadian study [12], the authors were interested in assessing the validity of ICD-10 codes after switching from ICD-9. While the authors did not find much difference in the validity of the codes from these two systems, the discrepancy was apparent for some conditions (e.g., HIV/AIDS, hypothyroidism, and dementia). The authors also observed that the quality of data had not yet improved in ICD-10 as originally expected.

Now that many countries are preparing to migrate from ICD-10 to ICD-11 [13], one can expect similar transition challenges to occur, as these two coding systems have different code structures [14], and the equivalence is at times lacking [15]. This research aims to introduce entropic measures to help users prepare for the migration to a new medical coding system by identifying and focusing preparation initiatives on clinical concepts with more likelihood of documentation deficiencies, coding errors, and longitudinal data comparison issues.

2. RELATED WORK

Not many studies have considered how to quantify the complexity of codes between two medical coding systems. In some studies, the equivalence in the number and structure of the codes between two coding systems is considered, but without accompanying measures of the dissimilarity in the codes [15]. In a few studies, an attempt is made to address the complexity between two medical coding systems. For example, in Boyd et al. [16, 17], the authors proposed using the science of networks to evaluate the difficulties of transitioning from ICD-9-CM to ICD-10-CM in the US. The authors used General Equivalence Mappings (GEMs) to create graphs where diagnoses were nodes, and the relationships in the GEMs were edges. From their analysis, the authors derived directional motifs and identified convoluted mappings, where multiple medical codes from both coding systems shared complex, entangled, and non-reciprocal mappings. The authors concluded that clinical classes with convoluted mappings were more likely to be challenging to code and costly to implement after the changeover to the new medical coding system. Besides, these authors also anticipated that clinical classes with a high ratio of ICD-10-CM to ICD-9-CM codes were more likely to affect a smooth transition. Another study that considered the complexity of transitioning between two coding systems relates to the work of Chen et al. [18], where the authors leveraged Shannon’s entropy to develop a mapping framework between ICD-10 and ICD-11 coding systems. The authors proposed three entropy-based metrics of standardizing rate (SR), uncertainty rate (UR), and information gain (IG) to validate information changes between ICD-10 and ICD-11. The authors obtained the UR measure by $\sum_{i=1}^M p_i \log 1/p_i$, where M was the number of ICD-11 candidate codes for a single ICD-10 code, and p_i was the probability of each ICD-11 code. In a special case of a uniform distribution, the authors suggested utilizing the average probability of $1/M$ to measure UR , which implied that $UR = \log M$. Among other conclusions, the authors recommended verifying ICD-10 codes with high UR measures as these codes were more likely to hinder a smooth transition to ICD-11.

3. CONTRIBUTIONS

This research complements previous studies highlighted in the Related Work section. For example, as in Chen et al. [18], this research proposes to apply Shannon’s entropy to study the complexity of the transition between two medical coding systems. Unlike in this previous study, the entropic measures in this research

account for the variation in the alphabets of candidate codes. Besides, Shannon’s entropy is also used to create a measure of coding complexity that considers not only the number of candidate codes (as in the UR measure [18]) but also the number of combinations of these codes. As shown later, failure to account for the latter information may underestimate or overestimate the related coding complexity. It should also be mentioned that the proposed methods have an advantage over convoluted measures suggested in Boyd et al. [16, 17]. Unlike in the convoluted approach, where a code is classified as either being involved in a convoluted relationship or not, the proposed methods provide non-dichotomous complexity measures of each code.

4. MATERIALS AND METHODS

4.1. Methods

4.1.1. A motivating problem

It is imagined that a manager of a given medical care facility is preparing to transition from an old medical coding system X to a new medical coding system Y . The forward ($X \rightarrow Y$) and backward ($X \leftarrow Y$) mappings between X and Y are provided. The manager is unsure about employing these mappings to identify clinical concepts that are more likely to be challenging to translate into the new medical coding system. Some of the benefits of knowing this information include being able to formulate targeted training efforts for coding and clinical documentation to foster the validity of the data in the new coding system. Besides, understanding complex translations may help take the necessary steps to ensure longitudinal data comparisons. This research aims to suggest the techniques that the manager could use to solve this dilemma.

4.1.2. Model and assumptions

Given forward mappings ($X \rightarrow Y$), the old medical coding system X is termed the source system, while the new coding system Y is termed the target system. In the backward mappings ($X \leftarrow Y$), the *source* and *target* terminologies are reversed. For model development, only forward mappings are considered here since the backward mappings would obey the same logic. From the prescribed forward mappings ($X \rightarrow Y$), it is assumed that code $x \in X$ corresponds to m number of candidate codes $y \in Y$. This relationship, referred to here as a *map*, is symbolized as $x \rightarrow \{y_1, y_2, \dots, y_m\}$ or as in the following matrix form:

$$x \rightarrow \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_m \end{bmatrix} \quad (1)$$

where each code in the map y_i , for $i : 1, \dots, m$, has n fixed number of characters (also called alphabets) a_{ij} , for $j : 1, \dots, n$. If necessary, padding may be added to a particular code to ensure that n is fixed for all codes, an approach, as shown later, simplifies calculations. Each column represents an axis or simply a

position of an alphabet in a code. The columns of a map are assumed independent. Each row of a map represents a valid code $y \in Y$. A set of more than one code in a map may be necessary to represent code $x \in X$. If $m = 0$, code x has no match in Y , which implies data loss in the new coding system. If $m = 1$, code $x \in X$ has a one-to-one relationship with code $y \in Y$. In this case, the coding complexity is expected to be zero since little surprise exists about what the new code should be. If $m > 1$, the coding complexity will be greater than zero as there is more than one candidate code in Y , thus more complexity and chances of coding or translation errors. In this research, a coding error is defined as the selection of a code where at least one alphabet is wrong or the selection of a set of codes where at least one of the codes is incorrect or missing. The expected coding complexity of a given clinical concept in X is characterized in terms of the uncertainty in the rows and columns of a map, which is measured here in bits units of Shannon's entropy [19].

Two major sources of coding complexity are assumed here, namely source A , which captures the variation in the alphabets of a map, and source B , which relates to the combinations of the rows of a map. The entropy for source A , or $H(A)$, is calculated as:

$$H(A) = - \sum_{j=1}^n \sum_{i=1}^{k_j} p_{ij} \log_2 p_{ij} \equiv \sum_{j=1}^n H(\mathbf{a}_j) \quad (2)$$

where $k_j \leq m$ is the number of unique alphabets in column \mathbf{a}_j of matrix (1) and p_{ij} is the probability of alphabet i in position j . The more the $H(A)$ measure, the more requisite detailed documentation to express all the alphabets of a map. Likewise, the more the number of code alphabets that must be chosen separately, the more complex and time-consuming the coding.

Regarding source B , the corresponding entropy $H(B)$ is obtained by:

$$H(B) = \log_2(v) \quad (3)$$

where $v = m_0 + \sum_{i=1}^s \prod_{j=1}^{m-m_0} m_{ij}$. Here, s is the total number of possible scenarios and m_0 represents the number of stand-alone codes and, for a given scenario i , $m_{i1}, \dots, m_{i(m-m_0)}$ denote the number of candidate codes in Y that must be combined to represent code $x \in X$. As before, m is the total number of candidate codes in a map. If a map only includes stand-alone codes, where no combinations of codes are required, Equation 3 becomes comparable to the UR measure introduced in Chen et al. [18]. The more the $H(B)$ measure, the more complex the coding due to the need for more coding memory and time, since more than one candidate code in the target system is going to be required to represent a single code from the source system. See Appendix A for more details on the derivation of Equations 2 and 3.

4.1.3. Implementation

It is recommended that both $H(A)$ and $H(B)$ entropic measures be normalized into $Z(\alpha)$ and $Z(\beta)$, as exemplified in Appendix A, to allow for the comparison and ranking of complexity from different sources. If $H(A)$ and $H(B)$ measures (or their normalized counterparts) are to be utilized to prepare for the transition (e.g., documentation improvement), they should be weighed using relevant empirical distribution (e.g.,

historical frequencies of codes in a given medical facility or general practice area). Accordingly, if, say, a particular facility never performs heart transplants, it shouldn't have to spend too much training efforts on the documentation of this clinical concept. Algorithm 4.1 shows the steps that one can take to implement the suggested entropic methods.

Algorithm 4.1: Computing entropic measures

Algorithm 4.1. Step 1: Calculate $H(A)$, the entropy of the columns of a map, to estimate the coding complexity due to the variation in the alphabets of the columns of a map.

Step 2: Calculate, $H(B)$, the entropy of the rows of a map to estimate the coding complexity due to the uncertainty in the number of valid code representations in the map.

Step 3: Normalize $H(A)$ and $H(B)$ by centering these measures and then dividing them by their standard deviations. The normalized measures are symbolized here as $Z(\alpha)$ for $H(A)$ and $Z(\beta)$ for $H(B)$.

Step 4: If empirical data, based on historical visits or future forecasts, were available, one would adjust $Z(\alpha)$ and $Z(\beta)$ measures by multiplying them with the probability of a corresponding clinical concept.

Step 5: Use the adjusted or unadjusted entropic measures to prioritize transition initiatives between two medical coding systems.

4.2. Materials

Algorithm 4.1 can be applied to evaluate entropic measures between any two medical coding systems, provided mappings or crosswalks exist. For demonstration, the 2015 US transition from ICD-9-CM to ICD-10-CM/PCS medical coding systems is considered. For a brief background, when the US was preparing to migrate from ICD-9-CM to ICD-10-CM/PCS, forward and backward general equivalence mappings (GEMs) were made available to users [2, 20]. A user could determine the number of candidate codes in the target system from these mappings, given a code in the source system. These files also allowed users to apply the given supplemental five digits codes (referred to as *flags*) to determine valid combinations of candidate codes in a map. For example, a flag code of 00000 or 10000 was used to represent a one-to-one relationship. The flag code of 00000 signified the exact equivalence, whereas a flag code of 10000 represented the approximate equivalence. If the relationship were one-to-many, the third character in the flag code would be 1 (instead of 0), and the fourth and fifth characters would specify combinations of candidate codes. The fourth character enumerated the number of scenarios, while the fifth character established the order that combinations were carried out in each scenario. The data used in this paper can be obtained directly from the CMS website at <https://www.cms.gov/Medicare/Coding/ICD10/Archive-ICD-10-CM-ICD-10-PCS-GEMs>. The 2015

GEMs, instead of the newer GEMs, are utilized here since they were the most updated mappings available to users to prepare for the transition from ICD-9-CM to ICD-10-CM/PCS in 2015.

4.3. Demonstration

Appendix E demonstrates a Python code to implement Algorithm 4.1. Figure 1 exhibits the application of this algorithm on map 0052. This map relates to an ICD-9-CM Vol.3 code of 00.52 for the *implantation or replacement of transvenous electrode into left ventricular coronary venous system*.

5. RESULTS

Algorithm 4.1 was applied to both forward and backward GEMs between ICD-9-CM and ICD-10-CM/PCS. Tables 1 and 2 display the corresponding descriptive statistics for $H(A)$, $H(B)$, and UR entropic measures. Codes without match in the target system were excluded from these statistics. For comparison purposes, the normalization of the UR measure [18] is symbolized as $Z(UR)$. To implement **Step 5** of Algorithm 4.1, clinical concepts were ranked by their entropic measures. Figures 2 and 3 show ranked clinical classes from the least to the most sum of $Z(\alpha)$, $Z(\beta)$, and $Z(UR)$ measures. The classes in these figures were also ranked separately using each entropic measure. As expected, the resulting rankings based on $Z(\alpha)$, $Z(\beta)$, and $Z(UR)$ measures were not always consistent. To assess how much the rankings of these entropic measures agreed, the Kendall tau correlation coefficients were assessed, and the results are presented in Table 3. The closer to 1 the Kendall tau value (the greener the color), the more the given entropic measures agreed. An alternative approach to implementing **Step 5** of Algorithm 4.1 is performing outlier and pattern analysis and then segregate concepts that should receive more attention during the transition. An example of how such an analysis may be conducted is shown in Figure 4. To extract thematic descriptions of the outlier maps, network analysis techniques suggested in Niyirora and Aragonés [21] were applied after removing stopwords [22] and residual words (e.g., other, unspecified, etc.). Communities of words in Sub-figures 4c and 4d (distinguished by different colors) were isolated using the modularity algorithm in Gephi [23]. To gauge the frequency (or significance) of words in the outlier maps, a word cloud analysis was undertaken, where the bigger the word meant, the more significant the word (see Sub-figures 4e and 4f).

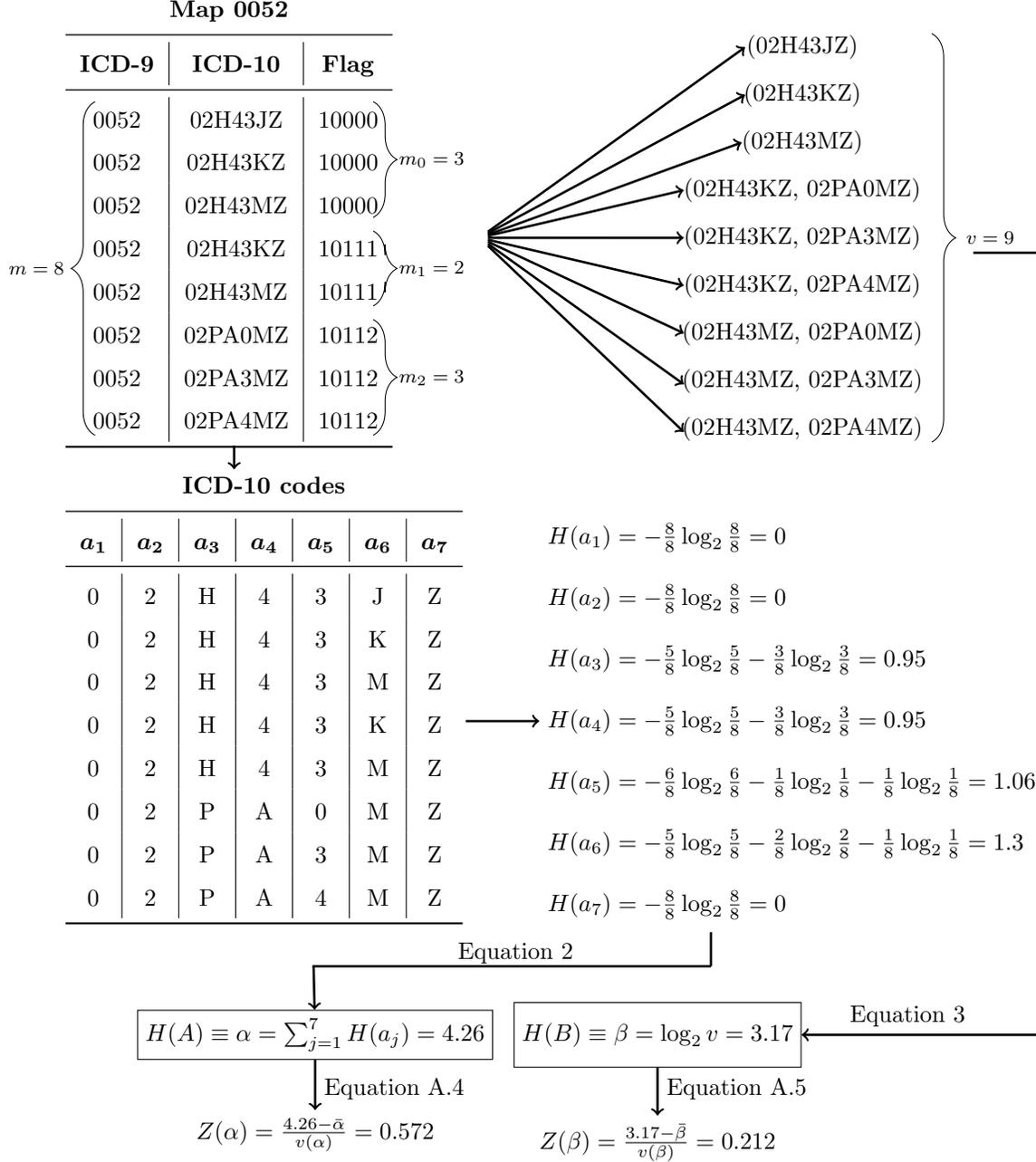


Figure 1: This figure depicts how to apply **Steps 1-3** of Algorithm 4.1 on map 0052. No empirical data were available to implement **Step 4** of this algorithm. $H(A)$ is computed per Equation 2. Equation A.3 is used to determine the number of valid representations v and $H(B)$ is calculated per Equation 3. The normalization of $H(A)$ and $H(B)$ follow Equations A.4 and A.5, respectively. The UR measure (proposed in Chen et al. [18]) of this map is obtained by $\log_2(m) = \log_2(8) = 3$. As compared to $H(B)$, the UR measure slightly underestimates the complexity of map 0052. While no actual probabilities were available for **Step 4**, it still can be speculated that if the probability of implanting or replacing any electrodes in the ventricular coronary venous system were zero for a given medical facility, both the $Z(\alpha)$ and $Z(\beta)$ measures would be $0.572 * 0 = 0.212 * 0 = 0$. The implementation of **Step 5** of this algorithm is discussed in Section 5.

Table 1: Descriptive statistics of the $H(A)$, $H(B)$, and UR entropic measures between ICD-9-CM Vol.3 and ICD-10-PCS

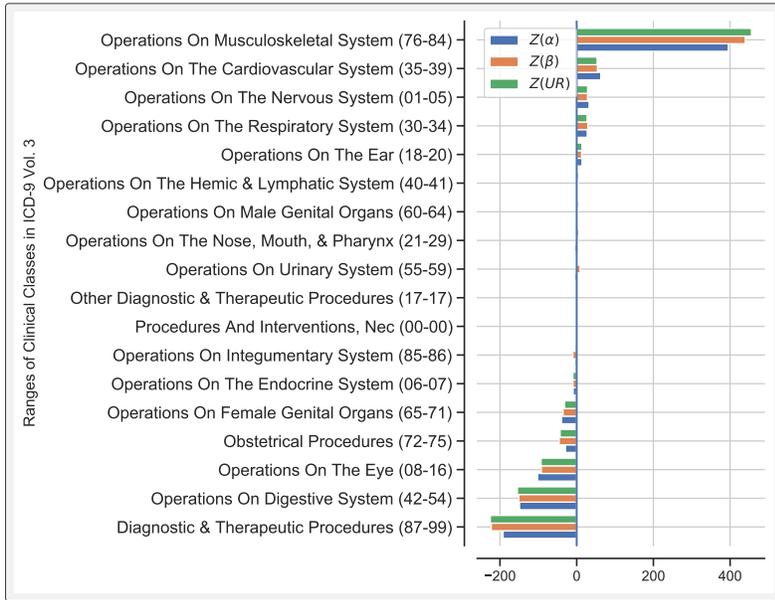
	Forward mapping			Backward mapping		
	From ICD-9-CM Vol.3 to ICD-10-PCS			From ICD-10-PCS to ICD-9-CM Vol.3		
	H(A)	H(B)	UR	H(A)	H(B)	UR
count	3672	3672	3672	71924	71924	71924
mean	2.76	3.03	2.74	0.09	0.25	0.13
std	1.92	2.16	1.85	0.32	0.84	0.40
min	0.00	0.00	0.00	0.00	0.00	0.00
25%	1.00	1.58	1.44	0.00	0.00	0.00
50%	2.58	2.58	2.58	0.00	0.00	0.00
75%	4.00	4.39	3.91	0.00	0.00	0.00
max	10.95	13.53	10.22	3.46	7.50	3.46

Table 2: Descriptive statistics of the $H(A)$, $H(B)$, and UR entropic measures between ICD-9-CM and ICD-10-CM

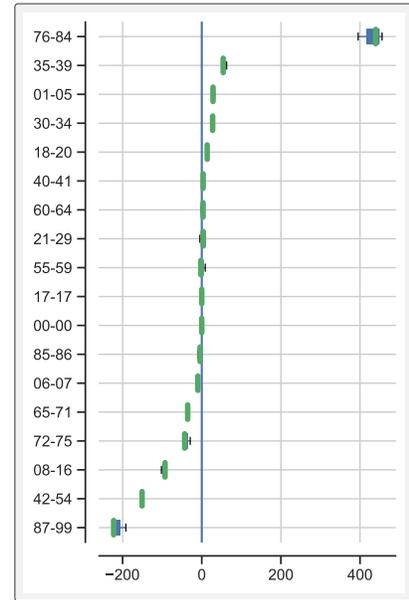
	Forward mapping			Backward mapping		
	From ICD-9-CM to ICD-10-CM			From ICD-10-CM to ICD-9-CM		
	$H(A)$	$H(B)$	UR	$H(A)$	$H(B)$	UR
count	14567	14567	14567	69823	69823	69823
mean	0.52	0.30	0.33	0.30	0.07	0.12
std	1.26	0.69	0.72	1.01	0.28	0.36
min	0.00	0.00	0.00	0.00	0.00	0.00
25%	0.00	0.00	0.00	0.00	0.00	0.00
50%	0.00	0.00	0.00	0.00	0.00	0.00
75%	0.00	0.00	0.00	0.00	0.00	0.00
max	13.10	9.06	9.06	7.31	3.58	3.58

Table 3: Kendall tau correlation among the rankings of clinical classes using the normalized entropic measures ($Z(\alpha)$, $Z(\beta)$, and $Z(UR)$). The symbol $\langle \rangle$ is used to signify mapping between the indicated medical coding systems.

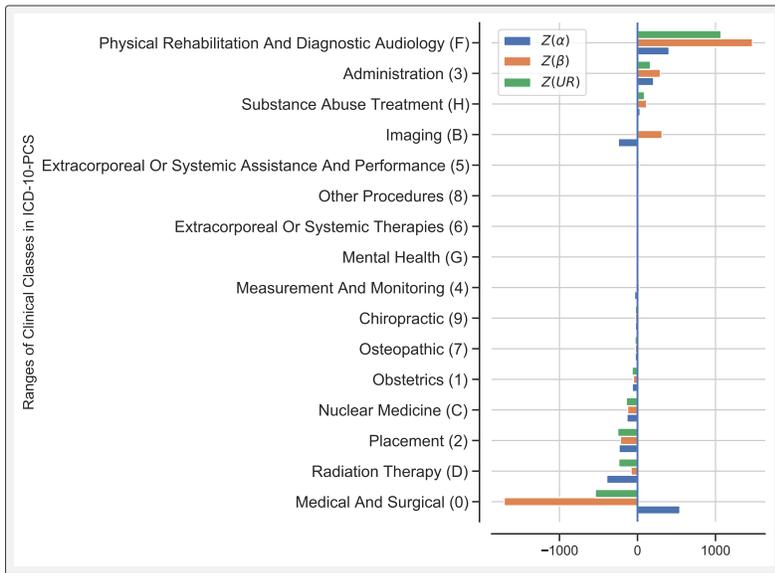
		Forward mapping			Backward mapping		
		$Z(\alpha)$	$Z(\beta)$	$Z(UR)$	$Z(\alpha)$	$Z(\beta)$	$Z(UR)$
ICD-9-CM Vol.3 \langle ICD-10-PCS	$Z(\alpha)$		0.99	0.99		0.97	0.97
	$Z(\beta)$	0.99		1.00	0.97		1.00
	$Z(UR)$	0.99	1.00		0.97	1.00	
ICD-9-CM \langle ICD-10-CM	$Z(\alpha)$		0.57	0.88		0.85	0.87
	$Z(\beta)$	0.57		0.66	0.85		0.98
	$Z(UR)$	0.88	0.66		0.87	0.98	



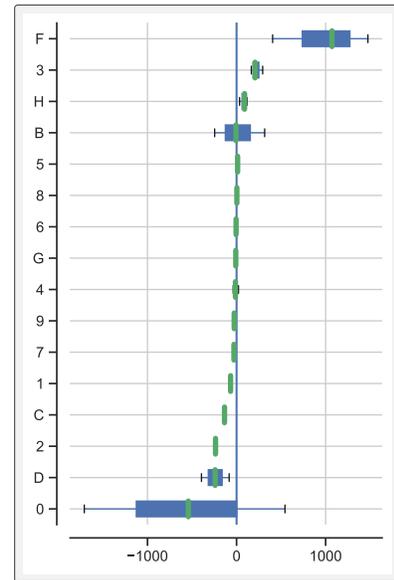
(a) Forward entropic measures from ICD-9-CM Vol.3 to ICD-10-PCS



(b) Box plots of entropic measures in (a)

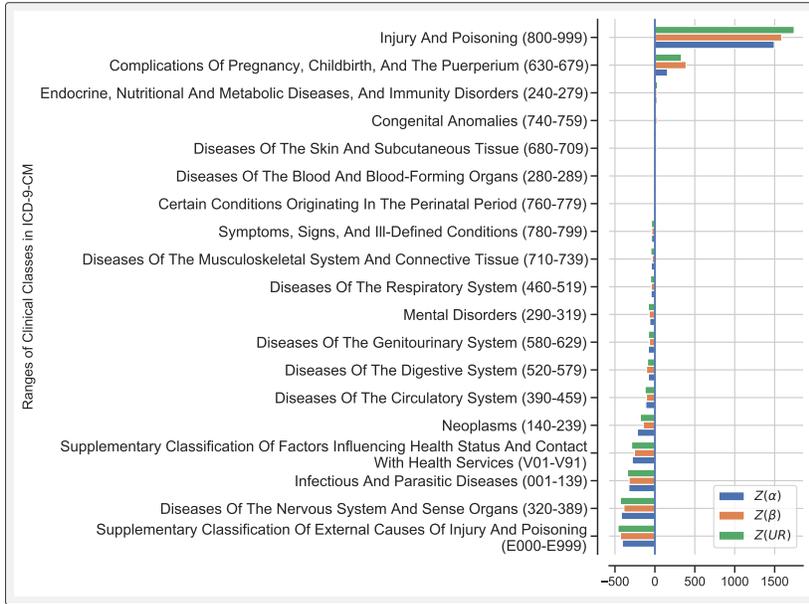


(c) Backward entropic measures from ICD-10-PCS to ICD-9-CM Vol.3

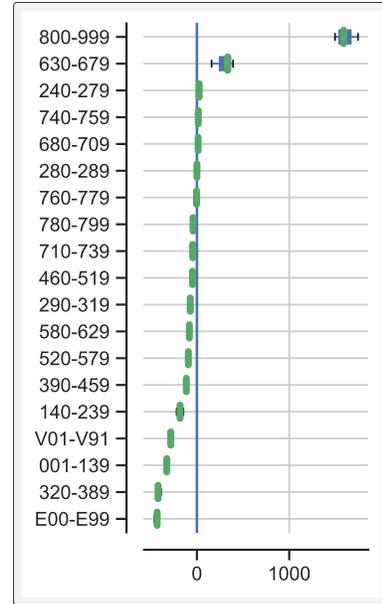


(d) Box plots of entropic measures in (c)

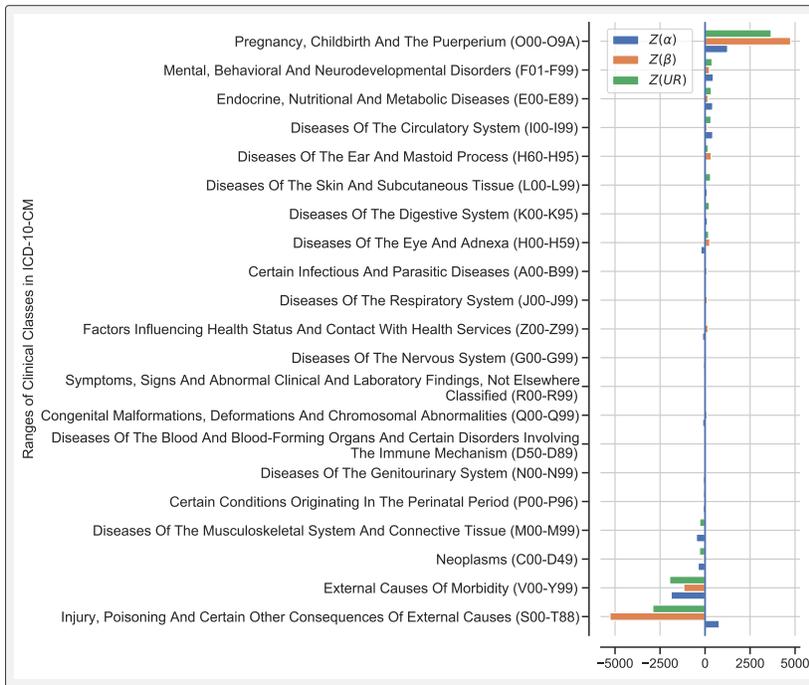
Figure 2: Forward and backward entropic measures between the procedure codes of ICD-9-CM Vol.3 and ICD-10-PCS. The x-axes represent the sum of $Z(\alpha)$, $Z(\beta)$, and $Z(UR)$ entropic measures. Sub-figures 2a and 2c show clustered bar plots of the indicated clinical classes arranged from the least to the most sum of entropic measures. Negative values signify no information gained or lost information (on average) from the source system to the target system. Positive values suggest gained information. Sub-figures 2b and 2d display related box plots that may help visually assess the variation in the entropic measures in each clinical class. The wider the box, the more the interquartile range, thus the more variability in the measures. The tighter the box and whiskers, the more the measures agree.



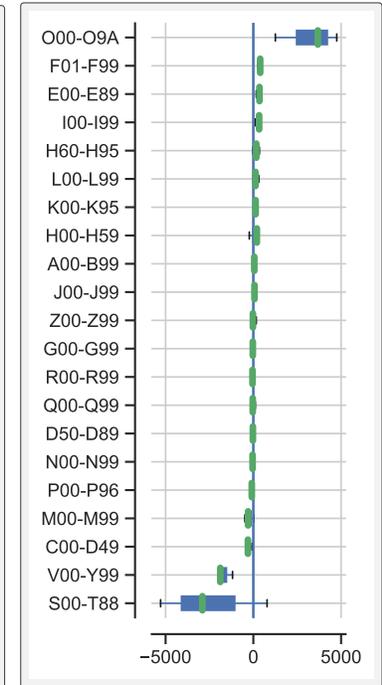
(a) Forward entropic measures from ICD-9-CM to ICD-10-CM



(b) Box plots of entropic measures in (a)

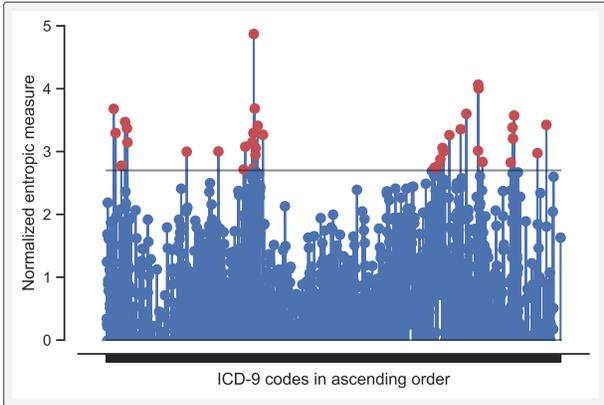


(c) Backward entropic measures from ICD-10-CM to ICD-9-CM

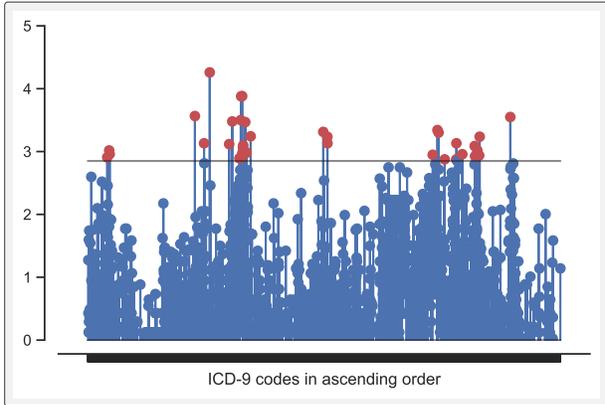


(d) Box plots of entropic measures in (c)

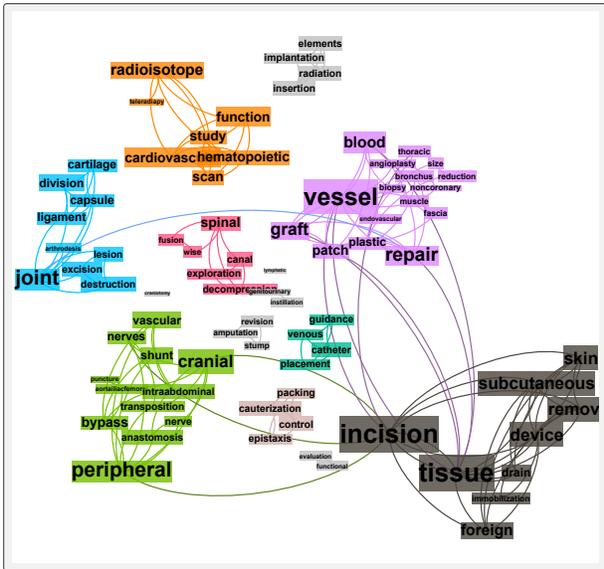
Figure 3: Forward and backward entropic measures between the diagnosis codes of ICD-9-CM and ICD-10-CM. The x-axes represent the sum of $Z(\alpha)$, $Z(\beta)$, and $Z(UR)$ entropic measures. Sub-figures 3a and 3c show clustered bar plots of the indicated clinical classes arranged from the least to the most sum of entropic measures. Negative values signify no information gained or lost information (on average) from the source system to the target system. Positive values suggest gained information. Sub-figures 3b and 3d display related box plots that may help visually assess the variation in the entropic measures in each clinical class. The wider the box, the more the interquartile range, thus the more variability in the measures. The tighter the box and whiskers, the more the measures agree.



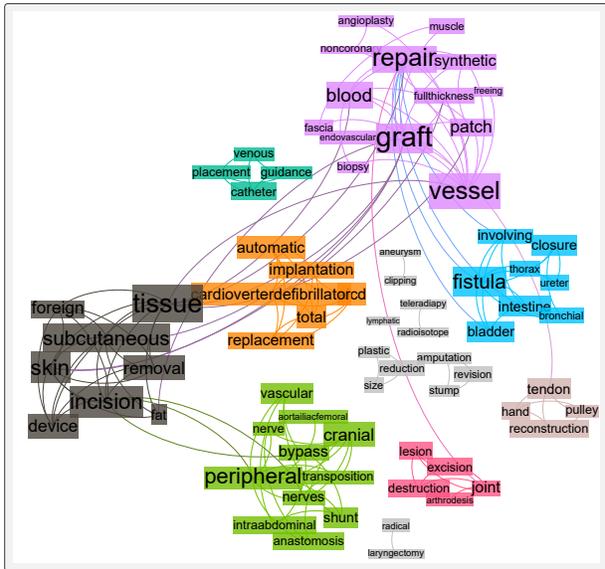
(a) $Z(\alpha)$ outlier maps are in red color



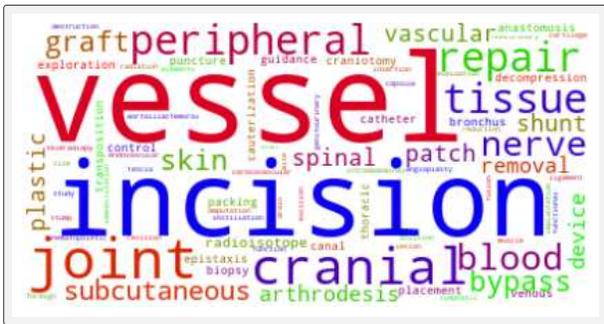
(b) $Z(\beta)$ outlier maps are in red color



(c) Network of words from the descriptions of $Z(\alpha)$ outliers



(d) Network of words from the descriptions of $Z(\beta)$ outliers



(e) Word cloud based on the descriptions of $Z(\alpha)$ outliers



(f) Word cloud based on the descriptions of $Z(\beta)$ outliers

Figure 4: An example of outlier and pattern analysis based on forward mapping from ICD-9-CM Vol.3 and ICD-10-PCS. Sub-figures 4a and 4b show red dots for outlier maps where $Z(\alpha)$ and $Z(\beta)$ scores are greater than the chosen threshold. For illustration purposes, the threshold were determined (2.7 for $Z(\alpha)$ and 2.85 for $Z(\beta)$), so that only the top 1% of the cases are isolated. The maps that were isolated are shown in Appendix C and Appendix D. Sub-figure 4c illustrates a network of words in the descriptions of $Z(\alpha)$ outlier maps while Sub-figure 4d portrays a similar network for $Z(\beta)$ outlier maps. The corresponding word clouds are respectively shown in Sub-figures 4e and 4f.

6. DISCUSSION

In 2015, ICD-10-PCS had significantly more number of mapped procedure codes ($n = 71,924$), as compared to ICD-9-CM Vol.3 ($n = 3,672$) (see Table 1). Equally, Table 2 shows more diagnosis codes for ICD-10-CM vis-à-vis ICD-9-CM. This fact alone implies that more specific information was likely to be gained by migrating from ICD-9-CM to ICD-10-CM/PCS, assuming complete clinical documentation and accurate coding. Besides, the mean statistics in these tables reveal that all the entropic measures are higher in the forward mappings, as compared to the backward mappings. This revelation further certifies that, on average, more information was gained in ICD-10-CM/PCS as compared to ICD-10-CM. The zero 25%, 50%, and 75% quartile measures suggest that at least 75% of codes in the source system had a one-to-one mapping with the target system. The implication of a one-to-one relationship is that no information is gained since $\log(1) = 0$. In other words, the codes in a one-to-one mapping may structurally look different, but if they represent the same clinical concept, then no information is gained. To a computer, a one-to-one mapping is a simple translation, but, of course, to a human coder, more complicated code structures may be more challenging to extract and translate.

The scale of the information gained (or lost) between ICD-9-CM and ICD-10-CM/PCS can be appreciated by clinical classes depicted in Figures 2 and 3. For example, Sub-figure 2a indicates that in the procedural forward mappings, the most information was gained in the class of the *Operations on Musculoskeletal System (76-84)*. The related box plot in Sub-figure 2b suggests that all three entropic measures relatively agreed on the characterization of class 76-84, given the small interquartile range. For diagnoses, Sub-figure 3a suggests that the class of the *Injury and Poisoning (800-999)* carried more information in ICD-10-CM followed by the class of *Pregnancy and Childbirth (630-679)*. Remarkably, Sub-figure 3c implies that an ICD-10-CM class related to Pregnancy and Childbirth (O00-O9A) also resulted in information gain in ICD-9-CM. These conflicting results are due to the convoluted nature of the mappings between these two medical coding systems [16].

From Figures 2 and 3, it can be appreciated that some clinical classes have negative entropic measures. The implication of this observation is that little, or no information, was going to be gained in the target system. For example, regarding the procedures, Sub-figure 2a indicates that for the class of the *Diagnostic & Therapeutic Procedures (87-99)*, little, or no information, was gained in ICD-10-PCS. Likewise, little, or no information, was going to be gained in ICD-9-CM Vol. 3 about the ICD-10-PCS class of *Medical and Surgical (0)* (see Sub-figure 2c). Though, the entropic measures relatively disagree on the latter suggestion, given a large inter-quartile range of class (0) in Sub-figure 2d. Regarding diagnoses, Sub-figure 3a suggests that little, or no information, was gained in ICD-10-CM about the ICD-9-CM class of *Supplementary Classification Of External Causes Of Injury And Poisoning (E000-E999)*. In an apparent contraction, Sub-figure 3c points to little, or no information, gained in the backward mapping about ICD-10-CM classes of *Injury, poisoning and certain other consequences of external causes (S00-T88)* and *External causes of morbidity (V00-Y99)*. This ambiguity is again the result of convoluted mapping between ICD-9-CM and ICD-10-CM/PCS coding

systems [16].

It is noteworthy that, despite a greater number of codes in ICD-9-CM/PCS, the backward max statistics in both Tables 1 and 2 are not zero. The suggestion here is that, for some clinical concepts, ICD-9-CM captured more information vis-à-vis ICD-10-CM/PCS (e.g., class (F) in Sub-figure 2c and class (O00-O9A) in Sub-figure 3c). The consequence of gaining information in the backward mapping is that some ICD-9-CM information was lost in ICD-10-CM/PCS, which created issues with longitudinal data comparisons. This dilemma also likely produced problems with verifying ICD-10-CM/PCS codes' validity, especially for classes where the information was gained in both forward and backward mappings (bidirectional), such as in the pregnancy and childbirth clinical class. Additional challenges resulting from the bidirectional information gain include conflicting documentation requirements, especially if the new coding system is trying to collect different types of information than what is commonly documented. Naturally, coding errors are likely to result if clinical documentation is lacking.

To prepare for the transition to a new medical coding system, the user can utilize the proposed entropic measures as a guide to orient training efforts. To this end, clinical classes can be ranked as a way to gauge where most information is likely to be gained or lost. Of course, the user would have more confidence if the rankings of these entropic measures agreed. Regarding the transition from ICD-9-CM to ICD-10-CM/PCS, Table 3 indicates that the most agreement existed in the forward mappings from ICD-10-CM Vol.3 to ICD-10-PCS. Another approach to prioritizing transition efforts may be based on the outlier and pattern analysis, as demonstrated in Figure 4. That is, instead of working with predefined clinical classes, the user could try to assess the impact of the transition using major themes from the descriptions of outlier maps. Many approaches to thematic analysis are available such as the Latent Dirichlet Allocation [24]. Here, a simple network was constructed and communities examined. For example, a close examination of Sub-figures 4c and 4d reveal a collection of terms that relate to various procedures for the vascular, skeletal, integumentary, and cardiac body systems. The most central words, in terms of the eigenvector centrality, were tissue, graft, subcutaneous, skin, repair, and incision. Combining these keywords, one may conclude that the procedures for the *musculoskeletal, integumentary, and cardio-vascular systems* were likely to involve more information gain in ICD-10-PCS, a conclusion that is consistent with the results in Sub-figure 2a.

7. CONCLUSION

Transitioning from an old medical coding system to a new one can be challenging, especially when the two coding systems are significantly different. The objective of this research was to propose methods that could help users prepare for the transition by identifying and focusing preparation initiatives on clinical concepts with more likelihood of transition challenges. To this end, two entropic measures of coding complexity were introduced. The first measure was a function of the variation in the alphabets of the codes in the target system, and the second measure was based on the possible number of valid representations of a code in the source system. It was recommended that the resulting entropic measures be normalized and adjusted by the

probability of a given code before isolating clinical concepts of interest. The main assumption here is that the more entropy, the more likelihood of coding errors. So, more prudent documentation is required, not only to increase the chances of accurate coding but also code validity and longitudinal data comparisons. The proposed techniques are suitable for establishing transition complexity between any two medical coding systems, provided mappings or crosswalks exist. A demonstration of how to implement these techniques was carried out using the 2015 forward and backward mappings between ICD-9-CM and ICD-10-CM/PCS.

8. LIMITATIONS AND FUTURE RESEARCH

A central conjecture of this research was that clinical concepts with more entropic measures were more likely to result in a more challenging transition. The justification of this assumption emanated from the fact that more entropic measures meant more variation in the codes, thus necessitating more prudent documentation and coding. This assumption may be violated if documentation is already complete and an experienced coder knows shortcuts to circumvent the new coding complexity. Accordingly, a medical record review may be necessary to ensure that the apparent complexity from the entropic measures actually exists. Besides, the topic of correlation between code validity and related entropic measures was not explored in this research. A medical record review may also be necessary to see if any lack of validity in the codes is explained by other reasons other than the entropic measures. Other relevant topics could not be considered in this research without additional data. For example, the question of how a new medical coding system could affect reimbursement was not entrained. Also, the topic of how coding guidelines and conventions may contribute to coding errors in the new system was not discussed. Future research goals include the consideration of these other topics, especially as they relate to the upcoming (or ongoing for some countries) transition from ICD-10 to ICD-11.

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

The data used in this paper can be obtained directly from the CMS website at <https://www.cms.gov/Medicare/Coding/ICD10/Archive-ICD-10-CM-ICD-10-PCS-GEMs>. These data are also included in the supplemental materials. The Python code used here is also included in Appendix E.

Competing interests

None

Funding

No funding was received to conduct this study.

Authors' contributions

There is only one author for this manuscript

Acknowledgments

None

References

- [1] CMS, 2015 Official ICD-10-PCS Coding Guidelines, 2014. URL: <http://www.cms.gov/Medicare/Coding/ICD10/Downloads/PCS-2014-guidelines.pdf>.
- [2] NCHS, International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM), 2014. URL: <http://www.cdc.gov/nchs/icd/icd10cm.htm>.
- [3] R. Caskey, J. Zaman, H. Nam, S.-R. Chae, L. Williams, G. Mathew, M. Burton, Y. A. Lussier, A. D. Boyd, et al., The transition to icd-10-cm: challenges for pediatric practice, *Pediatrics* 134 (2014) 31–36.
- [4] S. V. Kusnoor, M. N. Blasingame, A. M. Williams, S. J. DesAutels, J. Su, N. B. Giuse, A narrative review of the impact of the transition to icd-10 and icd-10-cm/pcs, *JAMIA Open* (2019).
- [5] Z. M. Alakrawi, V. Watzlaf, S. Nemchik, P. T. Sheridan, New study illuminates the ongoing road to icd-10 productivity and optimization, *J AHIMA* 88 (2017) 40–45.
- [6] J. P. Monestime, R. W. Mayer, A. Blackwood, Analyzing the icd-10-cm transition and post-implementation stages: A public health institution case study, *Perspectives in health information management* 16 (2019).
- [7] M. Butler, Analyzing eight months of icd-10, *Journal of AHIMA* 87 (2016) 15–22.
- [8] S. Stitcher, H. Lawrence, From icd-9 to icd10: A comparative analysis of coding audit findings in year one, 2016. URL: https://hfmamd.starchapter.com/downloads/46TH_ANNUAL_INSTITUTE/3_horizon_presentation_hfma_2016_sep22.pptx.
- [9] J. B. Hellman, M. C. Lim, K. Y. Leung, C. M. Blount, G. Yiu, The impact of conversion to international classification of diseases, 10th revision (icd-10) on an academic ophthalmology practice, *Clinical ophthalmology (Auckland, NZ)* 12 (2018) 949.
- [10] J. Krive, M. Patel, L. Gehm, M. Mackey, E. Kulstad, Y. A. Lussier, A. D. Boyd, et al., The complexity and challenges of the international classification of diseases, ninth revision, clinical modification to international classification of diseases, 10th revision, clinical modification transition in eds, *The American journal of emergency medicine* 33 (2015) 713–718.
- [11] J.-M. Januel, J.-C. Luthi, H. Quan, F. Borst, P. Taffé, W. A. Ghali, B. Burnand, Improved accuracy of co-morbidity coding over time after the introduction of icd-10 administrative data, *BMC health services research* 11 (2011) 194.
- [12] H. Quan, B. Li, L. Duncan Saunders, G. A. Parsons, C. I. Nilsson, A. Alibhai, W. A. Ghali, I. investigators, Assessing validity of icd-9-cm and icd-10 administrative data in recording clinical conditions in a unique dually coded database, *Health services research* 43 (2008) 1424–1441.

- [13] WHO, Who releases new international classification of diseases (icd 11), 2020. URL: [https://www.who.int/news-room/detail/18-06-2018-who-releases-new-international-classification-of-diseases-\(icd-11\)](https://www.who.int/news-room/detail/18-06-2018-who-releases-new-international-classification-of-diseases-(icd-11)).
- [14] CDC, Update on icd-11: The who launch and implications for u.s. implementation, 2018. URL: <https://www.cdc.gov/nchs/data/icd/ICD-11-WHOV-CM-2018-V3.pdf>.
- [15] K. W. Fung, J. Xu, O. Bodenreider, The new international classification of diseases 11th edition: a comparative analysis with icd-10 and icd-10-cm, *Journal of the American Medical Informatics Association* 27 (2020) 738–746.
- [16] A. D. Boyd, J. J. Li, M. D. Burton, M. Jonen, V. Gardeux, I. Achour, R. Q. Luo, I. Zenku, N. Bahroos, S. B. Brown, et al., The discriminatory cost of icd-10-cm transition between clinical specialties: metrics, case study, and mitigating tools, *Journal of the American Medical Informatics Association* 20 (2013) 708–717.
- [17] A. D. Boyd, C. K. Jianrong‘John’Li, S. R. Zaim, J. Krive, M. Mittal, R. A. Satava, M. Burton, J. Smith, Y. A. Lussier, Icd-10 procedure codes produce transition challenges, *AMIA Summits on Translational Science Proceedings* 2018 (2018) 35.
- [18] D. Chen, R. Zhang, X. Zhu, Leveraging shannon entropy to validate the transition between icd-10 and icd-11, *Entropy* 20 (2018) 769.
- [19] D. G. Luenberger, *Information Science*, Princeton University Press, 2006.
- [20] CMS, Procedure Code Set General Equivalence Mappings ICD-10-PCS to ICD-9-CM and ICD-9-CM to ICD-10-PCS Documentation and User’s Guide, 2014. URL: <http://www.cms.gov/Medicare/Coding/ICD10/2015-ICD-10-CM-and-GEMs.html>.
- [21] J. Niyirora, O. Aragonés, Network analysis of medical care services, *Health informatics journal* 26 (2020) 1631–1658.
- [22] W. J. Wilbur, K. Sirotkin, The automatic identification of stop words, *Journal of information science* 18 (1992) 45–55.
- [23] M. Bastian, S. Heymann, M. Jacomy, Gephi: an open source software for exploring and manipulating networks, in: *Third international AAAI conference on weblogs and social media*, 2009.
- [24] D. M. Blei, A. Y. Ng, M. I. Jordan, Latent dirichlet allocation, *Journal of machine Learning research* 3 (2003) 993–1022.

Figures

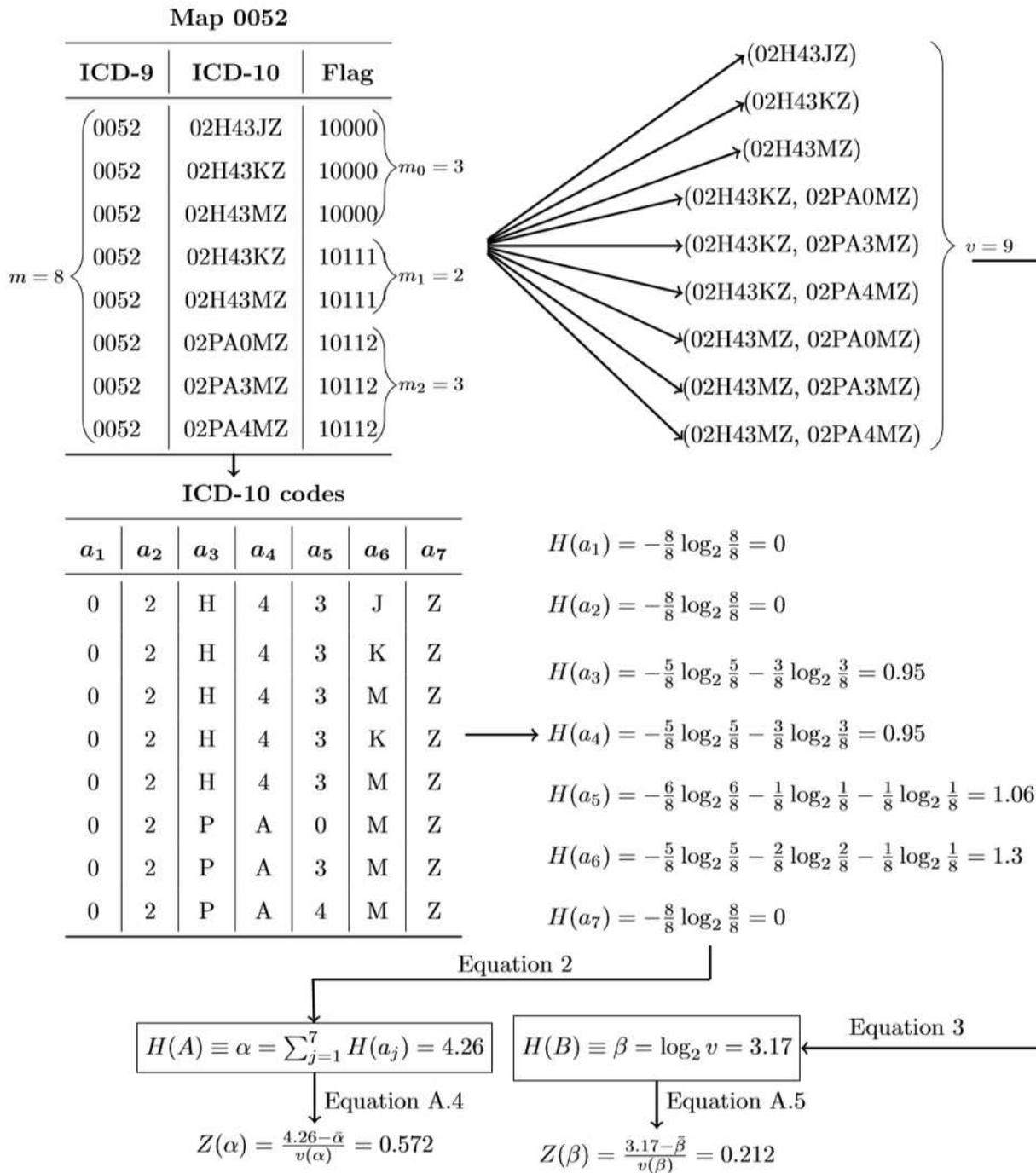
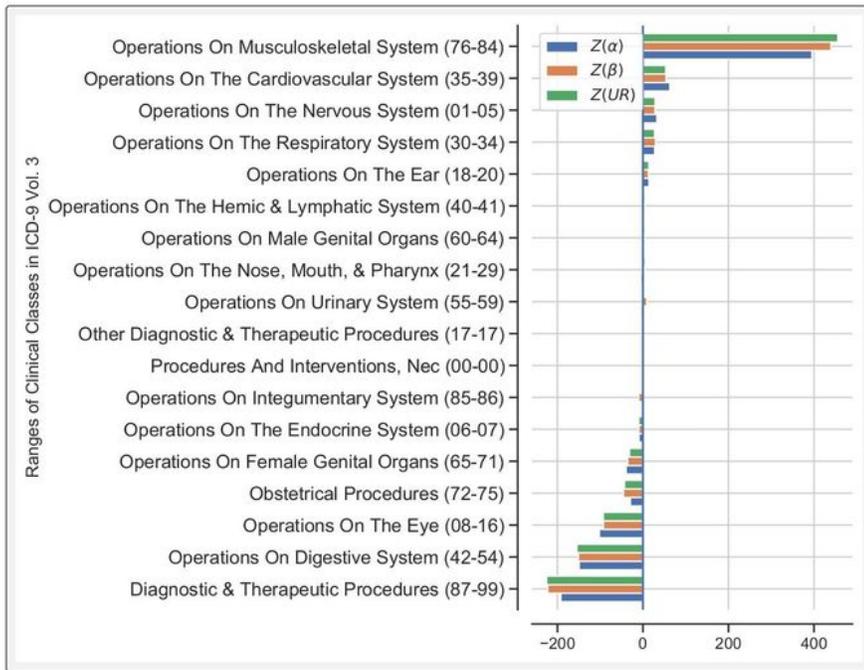


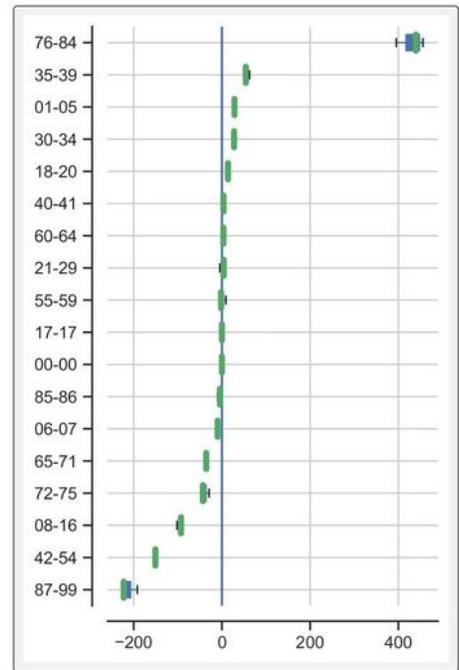
Figure 1

This figure depicts how to apply Steps 1-3 of Algorithm 4.1 on map 0052. No empirical data were available to implement Step 4 of this algorithm. $H(A)$ is computed per Equation 2. Equation A.3 is used to determine the number of valid representations v and $H(B)$ is calculated per Equation 3. The normalization

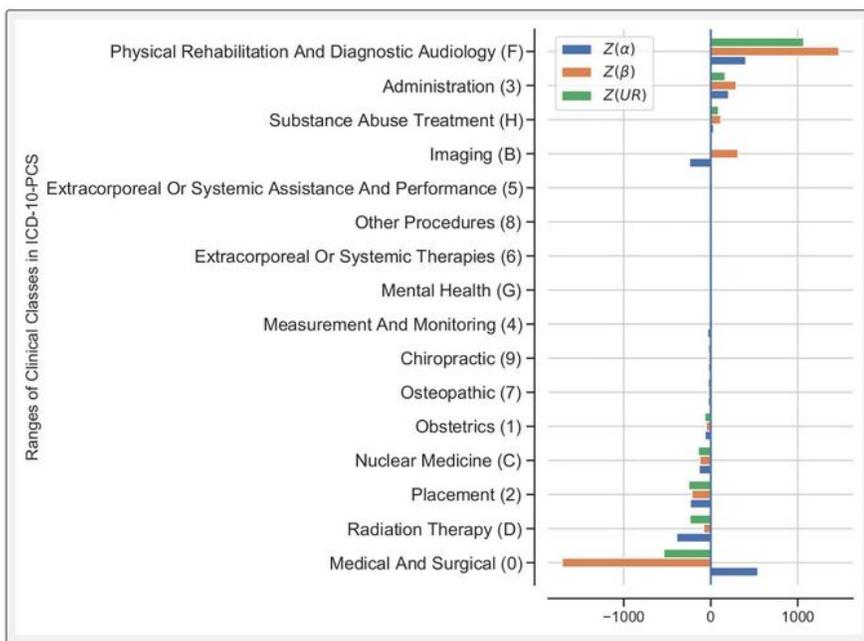
of $H(A)$ and $H(B)$ follow Equations A.4 and A.5, respectively. The UR measure (proposed in Chen et al. [18]) of this map is obtained by $\log_2(m) = \log_2(8) = 3$. As compared to $H(B)$, the UR measure slightly underestimates the complexity of map 0052. While no actual probabilities were available for Step 4, it still can be speculated that if the probability of implanting or replacing any electrodes in the ventricular coronary venous system were zero for a given medical facility, both the $Z(\alpha)$ and $Z(\beta)$ measures would be $0.572 \cdot 0 = 0.212 \cdot 0 = 0$. The implementation of Step 5 of this algorithm is discussed in Section 5.



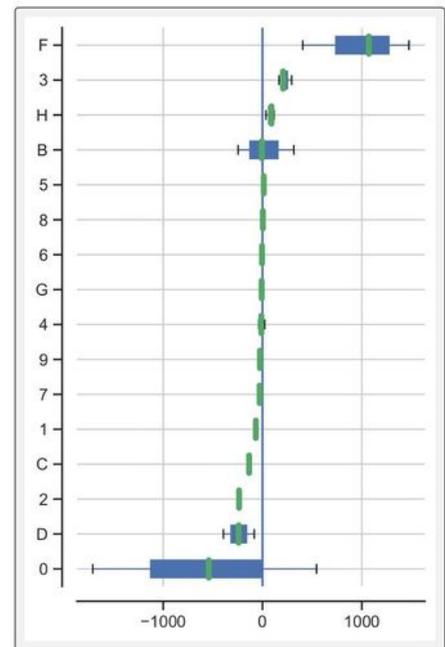
(a) Forward entropic measures from ICD-9-CM Vol.3 to ICD-10-PCS



(b) Box plots of entropic measures in (a)



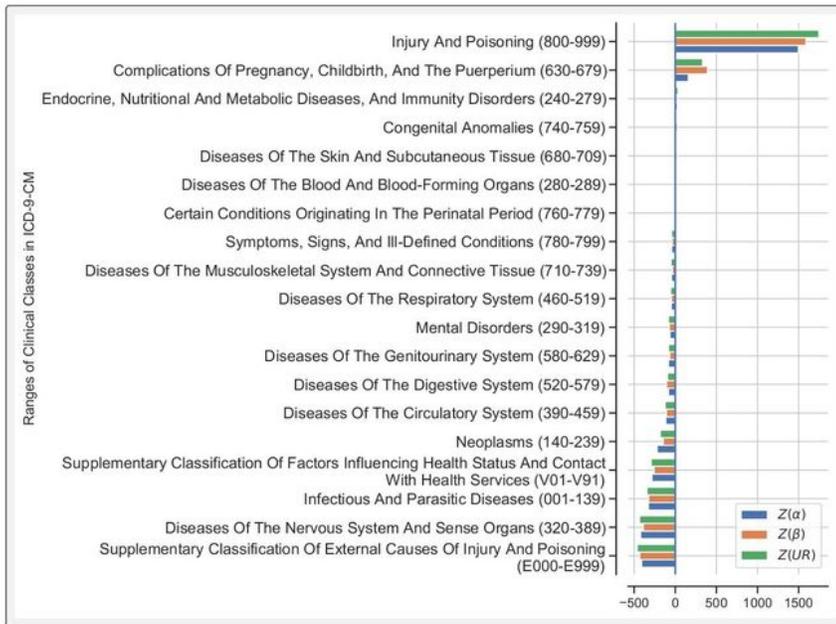
(c) Backward entropic measures from ICD-10-PCS to ICD-9-CM Vol.3



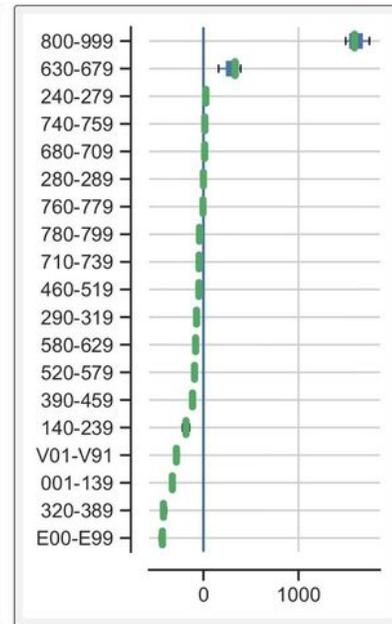
(d) Box plots of entropic measures in (c)

Figure 2

Forward and backward entropic measures between the procedure codes of ICD-9-CM Vol.3 and ICD-10-PCS. The x-axes represent the sum of $Z(\alpha)$, $Z(\beta)$, and $Z(\text{UR})$ entropic measures. Sub-figures 2a and 2c show clustered bar plots of the indicated clinical classes arranged from the least to the most sum of entropic measures. Negative values signify no information gained or lost information (on average) from the source system to the target system. Positive values suggest gained information. Sub-figures 2b and 2d display related box plots that may help visually assess the variation in the entropic measures in each clinical class. The wider the box, the more the interquartile range, thus the more variability in the measures. The tighter the box and whiskers, the more the measures agree.



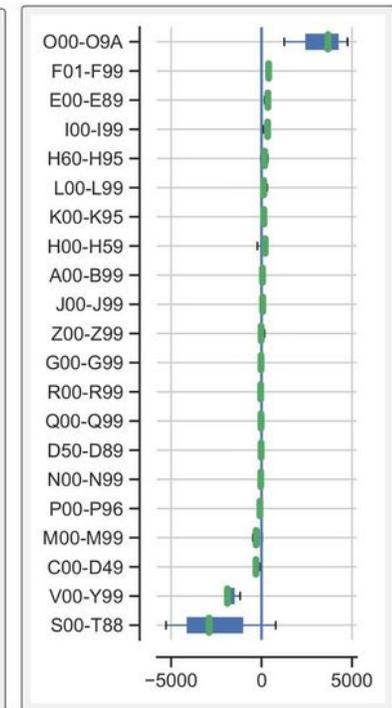
(a) Forward entropic measures from ICD-9-CM to ICD-10-CM



(b) Box plots of entropic measures in (a)



(c) Backward entropic measures from ICD-10-CM to ICD-9-CM

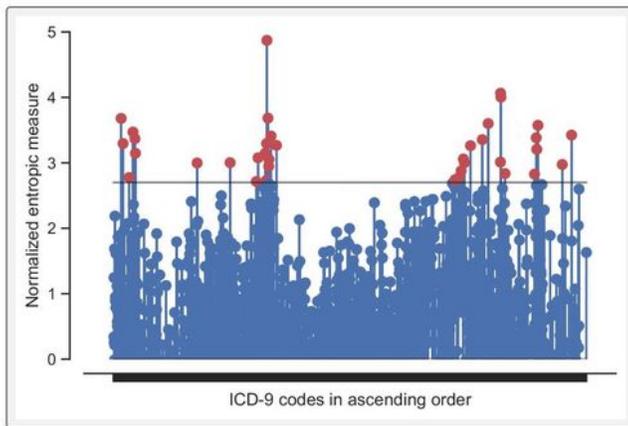


(d) Box plots of entropic measures in (c)

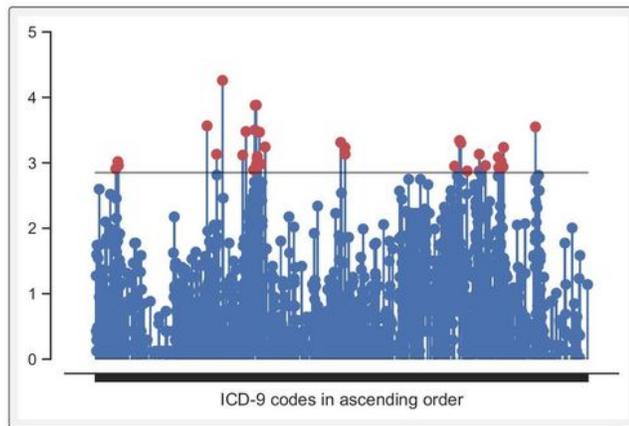
Figure 3

Forward and backward entropic measures between the diagnosis codes of ICD-9-CM and ICD-10-CM. The x-axes represent the sum of $Z(\alpha)$, $Z(\beta)$, and $Z(UR)$ entropic measures. Sub-figures 3a and 3c show clustered bar plots of the indicated clinical classes arranged from the least to the most sum of entropic measures. Negative values signify no information gained or lost information (on average) from the source system to the target system. Positive values suggest gained information. Sub-figures 3b and 3d

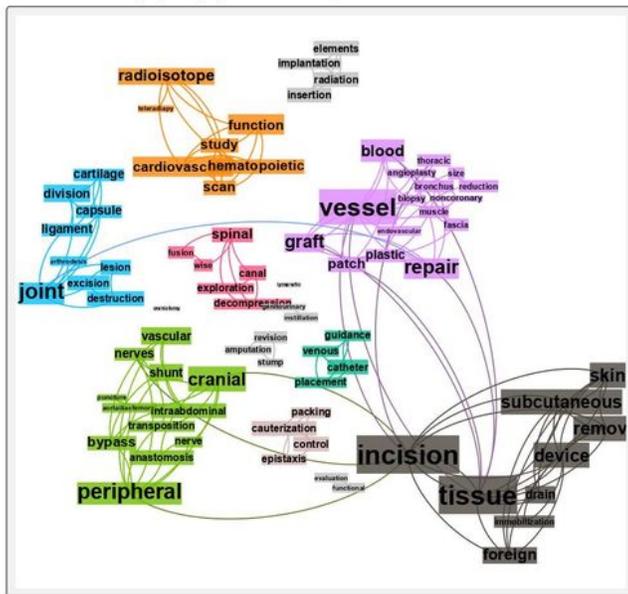
display related box plots that may help visually assess the variation in the entropic measures in each clinical class. The wider the box, the more the interquartile range, thus the more variability in the measures. The tighter the box and whiskers, the more the measures agree.



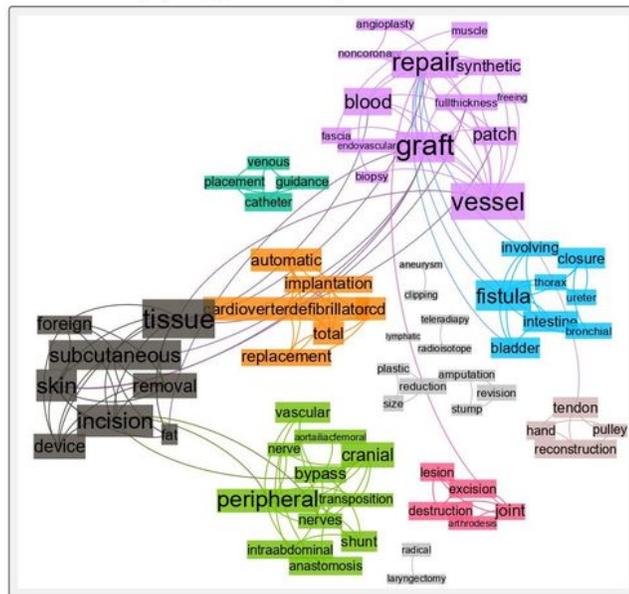
(a) $Z(\alpha)$ outlier maps are in red color



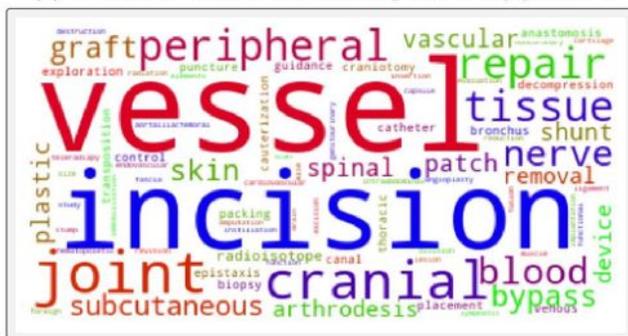
(b) $Z(\beta)$ outlier maps are in red color



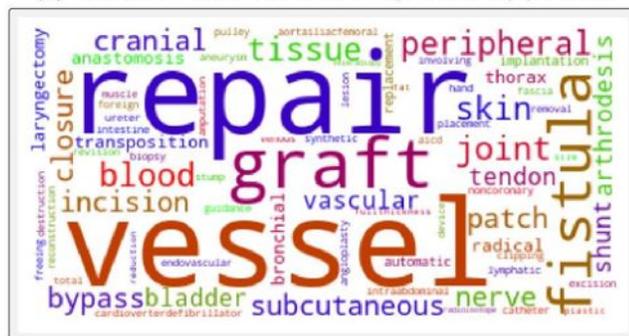
(c) Network of words from the descriptions of $Z(\alpha)$ outliers



(d) Network of words from the descriptions of $Z(\beta)$ outliers



(e) Word cloud based on the descriptions of $Z(\alpha)$ outliers



(f) Word cloud based on the descriptions of $Z(\beta)$ outliers

Figure 4

An example of outlier and pattern analysis based on forward mapping from ICD-9-CM Vol.3 and ICD-10-PCS. Sub-figures 4a and 4b show red dots for outlier maps where $Z(\alpha)$ and $Z(\beta)$ scores are greater than

the chosen threshold. For illustration purposes, the threshold were determined (2.7 for $Z(\alpha)$ and 2.85 for $Z(\beta)$), so that only the top 1% of the cases are isolated. The maps that were isolated are shown in Appendix C and Appendix D. Sub-figure 4c illustrates a network of words in the descriptions of $Z(\alpha)$ outlier maps while Sub-figure 4d portrays a similar network for $Z(\beta)$ outlier maps. The corresponding word clouds are respectively shown in Sub-figures 4e and 4f.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [PagesfromEntropicMeasuresHealthInformation12292020appendices.pdf](#)