WHO ICD Revision Information Note

## No: 7

## DATE: 01 February 2012

## **VERSION**: Draft Version 1.7

TO: RSG-SEG; RSG; TAGs; WGs;

CC:

SUBJECT: ICD-11 Code Structure, Numbering in Linearizations

KEYWORDS: Code structure, Linearization, Numbering, Code

# ISSUE:

There is a need to decide on the **numbering of ICD codes** (i.e. code structure) within the given linearizations.

## Definition:

This note describes the way ICD codes will be structured within a given linearization.

Decision and Options

1. The default numbering of all ICD codes in derivative linearizations from the ICD Foundation Layer (in particular the mortality and morbidity linearizations which is presented in the official print version) would adopt the following numbering pattern:

 **E**1**D**2**1**3**.E**4**E**5

* 1. Where E corresponds to a base 34 number (0-9 and A-Z ; excluding O,I) ;D corresponds Base 24 number (A-Z ; excluding O,I); and 1 corresponds to the base 10 integers (0-9)
	2. The rationale for this code structure is to provide maximal expressive space in a concise format, which reflects a common agreement on the human-readable length for a code. A short and standard ICD code structure is important for usability.
	3. The first three digits would represent pre-coordinated concepts, at the level of pre-coordination that presently exists in ICD10. However, 3-digit ICD10 codes that include residual categories, severity, or other multidimensional attributes would not be included. These digits can support a coding space of 7,920 rubrics (rubrics beginning with Z constitute modifiers, and cannot be used as 3-digit disease codes).
	4. By design, there is no attempt to represent chapters, blocks or other divisions within this code structure. However, to the extent practical, hierarchical relations would be retained in the 3-digit codes. All inheritance structures will be explicitly represented in the Foundation layer, and thus can be visually rendered in browsing or printed interfaces despite imperfections in apparent digit-based inheritance.
	5. The first digit after the decimal point (E4) is for expansion of the main 3-digit rubric into subtypes.
	6. The trailing decimal digit, E5, will be used to represent residual categories. At present, three are recognized: 0=no additional specification or clusters, Y=Other specified, and Z=Not specified. The values 1-X are reserved for indicating that additional information is coded, and indicating a simple grammar for linking related clusters of coded information.
	7. The inclusion of a forced number (13) is to preclude spelling possible “4-letter words” literally, in any language.
	8. The requirement that the second digit D2 be alpha, is to ensure visual distinction of ICD11 linearization codes from ICD10 codes.
1. It is desirable that various linearizations should share a common code base, if possible, (i.e. , the first three digits of the code: e.g., **ED1.** section of the code) which indicates the same group of concepts. This way of coding is referred to as **STYLE 1** (aka “the Russian doll model”) which allows telescopically nested linearizations for mortality, morbidity, primary care, verbal autopsy, and some short lists. The purpose of STYLE 1 is to facilitate the understanding of the correspondence among different linearizations and uses of ICD, within certain bounds.
2. In **STYLE 2**, different linearizations use different codes to represent similar-but-not- identical-concepts. For example, Myocardial Infarction in Verbal Autopsy is not exactly the same concept in the Morbidity Linearization. Linking similar concepts across varying linearizations will be done through the Foundation Layer via their corresponding identifiers. The equivalence of two concepts is created computationally through of Code-Code relationships across linearizations.
	1. In STYLE 2 distinguishing among various ICD linearizations would be done using **2-letter prefix codes**, (e.g. PC- Primary Care; VA- Verbal Autopsy etc.).
3. It is recognized that some linearizations would, however, have a totally different coding scheme and post-coordination, as is currently the case for External Causes and Injury chapter.
4. ICD Foundation Component – Linearization system will allow representation of alternative structures, for example emulation of the US ICD-10-CM, e.g. A11.1111 where A corresponds to the base 26 Alpha characters (A-Z). This would be an exception that would require the use of a prefix.

## Approval status:

## DEADLINE: 15 February 2012

## RSG SEG DECISION: to be Approved

## WHO DECISION: to be Approved

Note: major changes in this version include:

1. The truncation of digits from seven to five, recognizing that a multi-dimensional coding space will, by design, require a smaller coding space.
2. Forcing the second digit to be alpha
3. Clarifying that inheritance is a property of the Foundation Layer and can be visually rendered, though linearizations will attempt some reasonable inheritance where practical.
4. Introduction of a new set of residual categories (1-X), which form clustering grammars for linking multidimensional codes.
5. Acknowledging Z codes are a reserved set for modifiers