International Classification of Diseases (ICD)

Neoplasms Topic Advisory Group (TAG)

Face-to-Face Meeting, 7th and 8th March, 2012

The International Agency for Research on Cancer (IARC)

150 Cours Albert Thomas, 69372 Lyon CEDEX 08, France

Tel: +33 (0)4 72 73 84 85

Participants:

- Sebastien Antoni
- James Brierley
- April Fritz
- Pam Groenewald (via teleconference)
- Gunnar Henriksson (via teleconference)
- Robert Jakob
- Christian Lenzen
- Hiroshi Nishimoto
- Paola Pisani (co-chair)
- John Speakman (co-chair)
- Ulrich Vogel
- Theo Vos

Minutes

David Forman (IARC) welcomed the group. After introductions and review of the agenda, the objectives for the meeting were reviewed, as follows:

1. To draft a coding structure for a new ICD-11 Neoplasms chapter that could be presented in May 2012 as a basis for subsequent field testing;
2. To identify areas in the ICD Neoplasms chapter that need to be updated and aligned with current knowledge and practice, based on reference guides and requests from other groups, including other TAGs.

ICD is widely utilized to code morbidity and mortality for use in research (aetiology) and outcome evaluation and monitoring. Expectations for ICD-11 are long term stability of the classification (10 years), aetiology of disease and usability in low-resource settings.

Robert presented a review/recap of the current ICD revision process and the ICD-11 timeline, and the available online resources for ICD-11 revision; specifically the Redmine document repository at Mayo Clinic, the iCAT ("ICD Collaborative Authoring Tool") at Stanford, as well as the iCAT "snapshots" available on the WHO site.

The group then began addressing the objectives of the meeting. The neoplasm chapter of ICD-10 is organized by tumour "behaviour" (benign/*in situ*/malignant/uncertain) and anatomical site within "behaviour". In successive revisions of ICD a few categories defined by histopathology (called "morphology" in ICD-O) have been added to the list of 3-digit categories (mesothelioma, Kaposi sarcoma, skin melanoma). Malignancies of the haematopoietic and lymphatic system are also typically described and coded by morphology descriptors.

Several clinical working groups have suggested that:

- ICD-11 should include histopathology together with site of disease
- "Behaviour" and grade should be classified within disease groups, meaning that there should be no *a priori* separation of malignant, *in situ*, and benign into different chapters of ICD-11.

As confirmed by other TAGs and surveys of ICD stakeholders, the group recognized that, for the coding of neoplasms, virtually all major current use cases require the specification of two tumour characteristics: histology and stage/extension of the neoplasms at diagnosis as well as in follow-up care events. Detailed international coding/classification systems exist for both descriptors to guide the proposal of broad categories defined by ICD-11.

It was proposed to align ICD to current knowledge and practice, as it is contained in the WHO series on Histological and Genetic Typing of Human Tumours (WHO Blue Books, WHOBB) the ICD for Oncology (ICD-O) coding system (3rd revision), and the TNM classification system for staging (7^{th} edition). The information in the Blue Books could be used to provide definitions as well as information about the epidemiology and histopathology of individual neoplasms that could be included in the ICD. With regard to TNM the information about tumor site could be aligned with the ICD description of organ subsite (Example: C15 / Esophagus).

The TAG discussed a proposal (introduced by Paola, detailed in Annex 1) for a new five-character coding structure for ICD-11 that fulfills the above and accommodates the addition of systematic axes for stage/extension and histology within the constraints of the ICD-11 code structure as envisaged at the time of the meeting (ICD Revision Note No.7, Annex 2). The first two digits would include the chapter and organ site; the third digit behavior (0 / Benign; 1 / Borderline; 2 / In situ / Stage 0; 3 / Invasive, NOS; 4 / Invasive Stage I; 5 / Invasive Stage II; 6 / Invasive Stage III; 7 / Invasive Stage IV; 8 / Systemic; 9 / Unknown, unspecified, not applicable); separated by a dot, the fourth digit would include the organ subsite; the fifth digit morphology. See Annex 1 for details. The proposed new coding structure could be the basis for developing a draft version of the ICD-11 neoplasms chapter for presentation in May 2012.

The TAG expressed the opinion that it is not necessary that ICD-O topography codes be constrained by being the same as the C categories of ICD, as long as appropriate conversion programs are available.

Haematopoietic and lymphoid tissues malignancies presents a special case, as current understanding of the natural history of these neoplasms has led to a substantial overhaul of their histopathology classification (WHO BB) and diagnostic procedures, with a consequent impact on both clinical practice and aetiological research. This proposal maintains continuity

with the ICD-10 major entities for time trends analysis, incorporating at the same time the new classification. The proposal has been elaborated with the Hematology Working Group.

The TAG considered and addressed proposals submitted by other TAGs (Gastroenterology, Internal Medicine) and other stakeholders (*e.g.*, Cancer Registries, represented by IARC and TNM users, represented by UICC). Specific requests and proposals and TAG evaluations are included in the list reported below.

Open issues included how to integrate best morphology into ICD classification. Hiroshi suggested that NET, NEC, MALT, and GIST be included in the ICD classification. It was generally agreed that decisions for major histology groups should be based on the WHO Blue Book classification; however, it was debated how much level of morphological detail would be feasible in ICD and how additional coding could allow for more granularity (Ulrich). Further points of discussion, as mentioned by Hiroshi and Gunnar, included how to code overlapping lesions and with primary and secondary neoplasms, allowing a good patient follow-up from a clinical perspective. Additional topics included the classification of embryonic neoplasms and how to move classified dysplasias from different chapters within ICD into the ICD neoplasms chapter.

Questions/Pros/Cons

The TAG identified, discussed and enumerated a number of questions, and resultant pros and cons:

1. Is it possible to move the position of “behavior” (benign/malignant, *etc.*) from the current first position in ICD-10 subsequent to “anatomy” (*i.e.*, primary site), analogous to how it is done in ICD-O? Example: Breast cancer; followed by behavior (0-9).
    - Answer: yes (based on Paola’s proposal)
    - Pros: requested by multiple groups, clearly more descriptive
    - Cons: more coding space required, more characters (>4 ch), complete anatomic site requires 2 non-contiguous coding positions

2. Is it possible to incorporate a simplified one-digit axis of histology into ICD-11 as an independent qualifier, based on major histology types, and to how much detail?
    - Answer: yes (from 11 up to about 20 histology types)
    - Pros: it is already in use and follows WHO BB classification and ICD-O-3
    - Cons: not granular enough for some requests (example: GIST); requires separate code

3. Should other neoplasms such as benign or *in situ* have the same anatomical details in ICD-11as malignant neoplasms?
    - Answer: yes (consequence of new structure; see answer 1)
    - Pros: better description of benign tumors than in current ICD-10
    - Cons: many more codes

4. Should WHO BB classification of hematologic & lymphoid neoplasms be integrated into ICD11? If so, what modifications, if any, are needed to be made?

- Answer: maybe; to be checked with Hematology TAG
- Pros: update ICD to current knowledge of WHO BB and current practice
- Cons: in contrast with latest proposal from Hematology TAG incomparable with historic data due to different codes, but potentially addressable by adjusting and mapping them with historical data
- Action: Ulrich / Robert – get feedback from specialists / Hematology TAG

5. Should other WHO BB classification be integrated and how to ICD11?

- Answer: yes, if it fits into the code structure and if BB is up to date
- Pros: update ICD to current knowledge of WHO BB and current practice
- Cons: possible conflicts with historic data; retraining of coders

6. Should ICD11 topography be made consistent with TNM 7th edition?

- Answer: yes, including update
- Pros: consistent with TNM 7th edition
- Cons: none

7. Should staging be a part of ICD11 classification?

- Answer: yes
- Pros: requested by many, essential for clinical use and clinical follow-up; adds significant data that can be used for both clinical and population analysis
- Cons: it requires expanding behavior

8. How should ICD11 code multiple primary tumors? coding? Examples: C97 only for mortality coding - resolving “Dagger Asterisk” combinations

- Answer: to be resolved, pending how to handle multifocal tumors
- Pros: depends on resolution
- Cons: depends on resolution
- Action: question to be addressed at next TC

9. Should benign neoplasms and dysplasia be distinguished in ICD11?

a. Answer: to be resolved
b. Pros: depends on resolution
c. Cons: depends on resolution
d. Action: Gunnar will address question at next TC

10. Should in situ be grouped together with malignant neoplasms?

a. Answer: no
b. Pros: none
c. Cons: you would lose a lot of relevant information

11. How can we incorporate recurrences and remission into ICD-11 for any/all neoplasms?
    a. Answer: see our proposal
    b. Pros: for discussion
    c. Cons: greater granularity
    d. Action: will be addressed in Neoplasms draft

Roadmap / Plan of Action until May 2012

1. Prepare draft of Neoplasms chapter, based on proposal
    - Solve pending issues as stated in the above proposal end of March 2012:
        - Align subsites with TNM 'T' definitions
        - Get feedback from TAG Haematology (hopefully positive)
        - Define undetermined/NOS groups for haemo-lympho malignancies
        - Map ICDO-M of haemo-lympho malignancies
        - Take decision on coding recurrences for 3rd digit
    - Deadline: informal draft in April, final draft in May 2012
    - Owner: co-chairs (first draft)
    - Comment: to be sent to co-chairs, then distributed to TAG member
2. Initiate testing of draft
3. Include additional members / experts to Neoplasms TAG
    - Comment: Include additional coding expert
4. Organize preliminary meeting?
5. Publish draft proposal on-line at WHO level
    - Deadline: May 2012

Annex 1: TAG Proposal

TAG Neoplasm proposal (v.1) March 2012

If not otherwise specified, the code structure is: $\mathbf{E_1D_21_3.E_4E_5}$ (see Annex 2)**.** The TAG expressed the opinion that there is no need to use the first digit to identify the chapter (neoplasms), which could instead be identified as a block (range of first 2 characters combinations).

Char 1-2:

Anatomy (topography), *i.e.*, primary site: Maintains continuity with ICD-10, therefore the same anatomic entities defined by ICD-10: C00 through C80, with the addition of:

- gastro-esophageal junction (see TNM instructions for definition);
- nasal cavity and middle ear upgraded as main topographic sites

A "multiple primary tumours" category is maintained (ICD-10=C97), specifying that it is to be used only for mortality. Multiple primary tumours must be coded separately for morbidity.

Codes for secondary localizations are maintained.

Pending decision: whether to delete or maintain the three ICD-10 categories defined by morphology (skin melanoma, Kaposi's sarcoma, mesothelioma), which in this proposal can be identified through the 5th digit. In order to maintain the same degree of detail as in ICD-10 at the 3-digit level (for time trends analysis) it is desirable to keep them.

The section of haemo malignancies would be reduced to the major classical groups: HD, NHL, myeloid leukaemia, lymphoid leukaemia, other (C88-C96, malignant immunoproliferative diseases, other and unspecified leukaemias, multiple myeloma, other malignant neoplasms of lymphoid, haematopoietic and related tissues), in order to maintain continuity with the past. The 5th digit provides for the new classification described by the WHOBB.

Number of $\mathbf{E_1D_2}$ combinations required: 78 3-digit ICD-10-equivalent 'solid' tumours, 5 obsolete entities for the haemo-lymphopoietic system (new entities defined by 5th digit), for a total of 83 codes, just 11 over 3*24= 72 set.

Char 3:

Tumour behaviour and stage: Two options were discussed at the meeting; the second accommodates codes specific for disease recurrence, however, recurrences may be described combining several ICD-11 distinct codes. During the preparation of the minutes, Jim raised a third option (a variant on the second option, using "r", which in TNM stands for "recurrence").

Option 1: $E_1D_21_3.E_4E_5$	**Option 2A: $E_1D_2D_3.E_4E_5$**	**Option 2B: $E_1D_2D_3.E_4E_5$**
"Behaviour"/stage:	"Behaviour"/ stage/ recurrence:	"Behaviour"/ stage/ recurrence:
0: Benign 1: Borderline 2: In situ / Stage 0 3: Invasive, stage NOS 4: Invasive Stage I 5: Invasive Stage II 6: Secondary Stage III 7: Secondary Stage IV 8: Systemic 9: Unknown, unspec, not app	0: Benign 1: Borderline 2: In situ / Stage 03: Invasive, stage NOS 4: Invasive Stage I 5: Invasive Stage II 6: Secondary Stage III 7: Secondary Stage IV 8: Systemic 9: Unknown, unspec, not app A: Remission/Free of disease B: Local disease (recurrence/persistent) C: regional disease (recurrence persistent) D: distant disease (recurrence/persistent) E: local and regional F: local and distant G: regional and distant H: local, regional, and distant	0: Benign 1: Borderline 2: In situ / Stage 0 3: Invasive, stage NOS 4: Invasive Stage I 5: Invasive Stage II 6: Secondary Stage III 7: Secondary Stage IV 8: Systemic 9: Unknown, unspec, not app A: Remission/Free of disease B: Persistent disease C: rStageI D: rStageII E: rStageIII F: rStageIV

The proposed arrangement would facilitate routine case-finding/clinical follow-up. It also moves "behaviour" within anatomy rather than the reverse as in ICD-10; all "behaviours" have the same degree of topographic detail.

Char 4:

Organ subsite: As in ICD-10, plus alignment with T definition (see oral cavity/pharynx) of TNM. Needs systematic examination of every site and comparison with TNM.

Char 5 (pending evaluation by Hematology TAG):

Morphology groups: the major ones defined in ICD-O for solid tumors and utilized as an international rule to define multiple primary tumours within a patient, with the addition of

melanoma (any site). Groups ‘lymphomas’ and ‘leukaemia’ are replaced with 12-15 entities from the new WHOBB classification of haematopoietic and lymphoid tissues (pending details of ‘NOS’ groups to be defined):

	Morphology groups	**ICD-O M codes**
1	Squamous carcinomas	M 805 - 808, M-812 - 813
2	Basal cell carcinomas	M-809 – 811
3	Adenocarcinomas	M 814, M 816, M 819 - 822, M 826 - 833, M-835 - 855, M 857, M 894
4	Other specific carcinomas	M 803 804, M 815, M 817 - 818, M-823, M 824, M-825, M-834, M 856, M 858 - 867
5	Unspecified carcinomas (NOS)	M 801, M 802
6	Sarcomas and soft tissue tumors	M 868 - 871, M 880 - 892, M 899, M 904, M 912 - 913, M 915 - 925, M 937, M 954 - 958
7	Kaposi’s sarcoma	M-914
8	Mesothelioma	M-905
9	Melanoma	M 872 – 879
A	Myeloproliferative neoplasms including abnormalities of PDGFRA, PDGFRB OR FGFR1 (first two groups in WHOBB)	Pending mapping of ICD-O M codes
C	Myelodisplastic/myeloproliferative neoplasms	Pending mapping of ICD-O M codes
D	Myelodysplastic syndromes	Pending mapping of ICD-O M codes
E	Acute myeloid leukaemia (AML)	Pending mapping of ICD-O M codes
F	Acute leukaemias of ambiguous lineage	Pending mapping of ICD-O M codes
G	Precursor lymphoid neoplasms	Pending mapping of ICD-O M codes
H	Mature B-cell neoplasms	Pending mapping of ICD-O M codes
J	Mature T-cell and NL-cell neoplasms	Pending mapping of ICD-O M codes
K	Hodgkin lymphoma	Pending mapping of ICD-O M codes
L	Histiocytic and dendritic cell neoplasms	Pending mapping of ICD-O M codes
M	Post-transplant lymphoproliferative disorders, PTLD NOS	Pending mapping of ICD-O M codes
S	Unspecified cell lineage	Pending mapping of ICD-O M

		codes
T	Unclassifiable/undetermined	Pending: define NOS entities by cell lineage/diagnostic degree of investigation
W	Other specified types of cancer	M 893, M 895 - 898, M 900 - 903, M 906 - 911, M 926 - 936, M 938 - 953, M 973 - 975, M-976
Z	Unspecified types of cancer	M 800, M-997

If there is the option of a further digit (but the TAG does not feel strongly about this):

Char 6: Histopathologic grade

Annex 2: ICD Revision Note: ICD-11 Code Structure, Numbering in Linearizations

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:

$E_1D_21_3.E_4E_5$

a. 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)

b. 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.

c. 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).

d. 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.

e. The first digit after the decimal point (E_4) is for expansion of the main 3-digit rubric into subtypes.

f. The trailing decimal digit, E_5, 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.

g. The inclusion of a forced number (1_3) is to preclude spelling possible "4-letter words" literally, in any language.

h. The requirement that the second digit D_2 be alpha, is to ensure visual distinction of ICD11 linearization codes from ICD10 codes.

2. 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.

3. 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.

a. 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.).

4. 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.

5. 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:

a) The truncation of digits from seven to five, recognizing that a multi-dimensional coding space will, by design, require a smaller coding space.

b) Forcing the second digit to be alpha

c) Clarifying that inheritance is a property of the Foundation Layer and can be visually rendered, though linearizations will attempt some reasonable inheritance where practical.

d) Introduction of a new set of residual categories (1-X), which form clustering grammars for linking multidimensional codes.

e) Acknowledging Z codes are a reserved set for modifier

World Health Organization

ANNEX 2